UNIVERSITE D'AIX-MARSEILLE

ECOLE DOCTORALE PHYSIQUE ET SCIENCES DE LA MATIERE

Pour obtenir le titre de

Docteur en Sciences

Discipline : Astrophysique

Pin-Wei WANG

Star formation rate and the assembly of galaxies in the early universe

PhD supervised by Olivier LE FÈVRE and prepared in the Laboratoire d'Astrophysique de Marseille

08/04/2015

Composition of the Jury:

Philippe Amram	LAM, Marseille	Jury President
Sébastien Foucaud	Shanghai Jiao Tong University	Rapporteur
Giovanni Zamorani	INAF, Bologna	Rapporteur
Paolo Cassata	Universidad de Valparaiso	Examinateur
Olivier Le Fèvre	LAM, Marseille	Directeur de thèse

ABSTRACT

The main purpose of this study is to identify and study the population of high redshift galaxies in the redshift range ($4.5 < z < 6.5$) in the COSMOS field. I use the near infrared data from the UltraVista survey conducted with the Vista telescope in combination with multi-wavelength data available in the COSMOS field. The VIMOS Ultra Deep spectroscopic redshift survey conducted with VIMOS spectrograph (VUDS) is used as a control sample for the selection of high redshift candidates. The DR2 release of the UltraVista survey provides very deep near infrared data (24.8 ABmag in K_s, 5σ) for a large area of 1.39 deg^2, and enables a large volume sampling to identify rare objects as well as a reduction in cosmic variance compared to smaller fields probed until now. I perform the multi-band photometry and produce a photometric catalog ranging from the B band to the 4.5μm Spitzer band. I define and use efficient search techniques and quantify the associated selection function in terms of completeness and purity. Here I demonstrate that traditional colour-colour techniques used to select drop-out galaxy sample at these high redshifts are subject to contamination and completeness issues leading to significant limitations in producing a complete census of the galaxy population. I show that adding multi-band criteria that include the use of full information derived from both optical photometry from the Subaru telescope (B, V, r, i, z') and near-infrared photometry (Y, J, H, K_s) from the Vista telescope and 3.6 and 4.5 μm from Spitzer can help in cleaning-up samples. This analysis leads me to select galaxies at $z \geq 4.5$ using photometric redshifts computed from the full spectral energy distribution (SED) combined with well tuned magnitude limits based on the depth of the data in each band. I use the photometric redshift measurement code Le Phare to perform spectral fitting on each object to obtain the photometric redshifts directly. Several photometric redshift degeneracies with galactic cool dwarf stars and lower redshift red dusty galaxies force a visual inspection to further clean the data from interlopers. At the end of this process I produce a unique catalogue of 2 036 galaxies with $4.5 \leq z \leq 5.5$ and 330 galaxies with $5.5 \leq z \leq 6.5$, the largest and most complete catalogue of bright sources at these redshifts existing today. I compute the Luminosity Function using the code ALF with Olivier Ilbert. Thanks to the large volume covered by UltraVista I am able for the first time to place strong constraints on the bright end of the LF. I find that the LF at $z \sim 5$ is well fit by a Schechter function. At $z \sim 6$ I find that the bright end might be more populated than expected from a Schechter function, in line with results from other authors, an indication that the mass assembly processes have evolved quickly in a short 0.5-1 Gyr timescale. Finally I integrate the luminosity functions to compute the luminosity density and derive the star formation rate density (SFRD) in $4.5 \leq z \leq 6.5$.

My results show a high SFRD comparable to the latest results derived from the HST data, with an improved accuracy linked to the better constraints at the bright end of the LF. The main remaining uncertainties are the slope α of the luminosity function and the dust correction, which will require much deeper observations in large volumes, using next generation observatories like the James Webb Space Telescope.

Résumé

L'objectif de cette thèse est d'identifier et d'étudier la population à haut décalage spectral ($4.5 < z < 6.5$) dans le champ COSMOS. Pour cette étude j'ai utilisé des données dans le proche infrarouge venant du sondage UltraVista, observé avec le télescope Vista, associé à des données multi-longueur d'onde disponibles dans le champ COSMOS, ainsi que le sondage ultra profond de VIMOS (VUDS) utilisé comme un échantillon de contrôle pour la sélection des candidats à grand décalage spectral. La publication des données DR2 du sondage UltraVista donne accès à des données infrarouges très profondes (K_{AB}=24.8, 5σ) sur une portion de 1.39 deg^2. Elle permet un échantillonage important en volume pour identifier des objets rares ainsi qu'une réduction de la variance cosmique, chose que les champs sondés jusqu'à maintenant n'ont pas permis car étant trop petits.

Dans le cadre de ce sondage, j'ai extrait la photométrie et produit le catalogue photométrique de la bande u à la bande IRAC à 4.5 μm. A partir de ce catalogue j'ai défini et utilisé des techniques efficaces de recherche de galaxies et quantifié la fonction de sélection associée en terme de complétude et de pureté. Je démontre dans ce travail, que les techniques de sélection couleur-couleur traditionnellement utilisées pour sélectionner les échantillons de galaxies à ces décalages spectraux élevés sont sujets à des problèmes de contamination et de complétude entrainant des limitations importantes dans la production d'un comptage complet de ce type de population de galaxies.

Je montre également qu'utiliser des critères multi-longueur d'onde qui incluent l'utilisation de l'information complète tirée de la photométrie optique (avec le télescope Subaru) et infrarouge (avec le sondage UltraVista) ainsi que les données Spitzer à 3.5 et 4.5 μm peut apporter une aide pour nettoyer les échantillons. Cette analyse m'a amené à sélectionner des galaxies à z$\geq$4.5 en utilisant les décalages spectraux photométriques estimés à partir de la distribution spectrale d'énergie complète (basée sur le code Le Phare) ainsi que des limites en magnitudes basées sur la profondeur des données dans chaque bande. Cette sélection m'a permi de construire un catalogue unique de 2036 galaxies dans l'intervalle $4.5 \leq z \leq 5.5$ et de 330 galaxies dans l'intervalle $5.5 \leq z \leq 6.5$ faisant de ce catalogue le plus grand et le plus complet à ce jour. En utilisant ces catalogues, j'ai calculé la fonction de luminosité en utilisant le code ALF. Grâce à l'importante couverture en volume du sondage UltraVista, il a été possible, pour la première fois, de placer d'importantes contraintes sur les faibles luminosités de la fonction de luminosité. J'ai trouvé que la fonction de luminosité à $z \sim 5$ est bien reproduite par une fonction de Schechter. A $z \sim 6$, j'ai observé qu'aux faibles luminosités la fonction de luminosité semble être plus peuplée qu'une fonction de Schechter le laisse présager, en accord avec les résultats d'autres études. Ceci est une indication que les processus d'assemblage de la masse ont évolué rapidement, sur une période de 0.5 à 1 milliard d'années.

Finalement, j'ai intégré la fonction de luminosité pour en déduire la densité de luminosité et dériver la densité de formation stellaire entre $z = 4.5$ et $z = 6.5$. Mes résultats montrent une densité de formation stellaire importante, en comparaison des derniers résultats avec les données du télescope Hubble, ainsi qu'une précision plus grande liée aux meilleures contraintes sur la fin lumineuse de la fonction de luminosité. Les incertitudes principales restantes sont sur la pente α de la fonction de luminosité et sur la correction de l'absorption par la poussière. Des observations encore plus profondes et sur de grands volumes seront nécessaires, ce qui deviendra possible avec les observatoires de future génération, tel que le James Webb Space Telescope.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The main goal of my thesis work is to identify and study galaxies in the very early phases of assembly at redshifts $4.5 \leq z \leq 6.5$, corresponding to an epoch more than 12.5 Gyr in the past, just after reionisation. I aim to measure the luminosity function and derive the star formation rate density to understand how the star formation has evolved since these early epochs.

My manuscript is organised as follows:

In Chapter 2, I recall the modern cosmology background and the current understanding on galaxy formation and evolution. I identify important physical properties, such as the luminosity function and star formation rate which help us investigate the early universe.

In Chapter 3, I introduce the observational methods deployed to search for these distant objects. The now classical Lyman Break photometric technique is complemented by narrow band filter searches for Lyman-α. I introduce the photometric redshift technique which uses all the multi-band photometry to fit the spectral energy distribution with model spectra to find redshifts and also measure physical parameters like stellar mass and star formation rate. I then review the main existing surveys and samples used to study galaxies at $z > 4.5$, and I identify the current lack of deep near infrared photometry over wide fields, the main motivation for the UltraVista photometric survey which is the core dataset used in my work. To end Chapter 3 I describe how the luminosity function can be computed, with the approximation of the Schechter function, and I then explain how the LF is integrated to compute the luminosity density and how the star formation rate density (SFRD) is finally estimated.

The next step of my work is to construct a multi-wavelength photometric catalogue combining the new near infrared survey UltraVista with the deep existing data in the COSMOS field. UltraVista is the deepest survey conducted with the VISTA telescope

at the Paranal Observatory in Chile. It covers the YJH and K_s photometric bands over an unprecedented 1.5 square degree area. I used images from the data release number 2 of UltraVista, and I perform the photometry using SExtractor reaching down to $K_{AB} = 24.8, 5\sigma$, as described in Chapter 4. I also present the VIMOS Ultra Deep Survey (VUDS), a spectroscopic survey of 10 000 galaxies reaching up to $z \sim 6$, which I use as a control sample in the next chapters.

After the catalogue construction, I proceed in describing two important aspects of my work in chapter 5: simulating galaxies at $z > 4.5$ and the selection of my galaxy sample at $4.5 \leq z \leq 6.5$. The first part of chapter 5 presents the simulation I created to answer several questions: (1) What are the optimum selection criteria for selecting high redshift galaxies? (2) How many galaxies are lost when constructing the multi-wavelength catalog after applying the standard photometric process on images? (3) What are the completeness rate and the contamination of selecting high redshift galaxies when relying on colour criteria only. These questions are crucial to figure out when computing the luminosity function and then convert into SFRD. The last and important part in this chapter is how I produce the final high redshift galaxy sample. After a careful evaluation of the selection methods I decided to use photometric redshifts directly computed from spectral energy distribution (SED) fitting instead of using traditional colour-colour drop-out techniques. I introduce the concept of SED fitting and the general photometric redshift codes widely used in astronomical community. I describe the code Le Phare, the initial parameters setting and the galaxy template set that I use. I emphasise the need to visually examine the SED of all candidates to remove the difficult cases of degeneracies between red obscured galaxies and cool galactic dwarf stars with high redshift galaxies. I compare the photometric redshift selection to colour selection and demonstrate that this method is much less affected by uncertainties than colour selection methods. The final part of chapter 5 presents the final sample generated by this method. I obtain a sample of 2 036 galaxies with $4.5 \leq z \leq 5.5$ and 330 galaxies with $5.5 \leq z \leq 6.5$, the largest sample at these magnitudes and redshifts today.

In chapter 6, I present the global properties of my sample including the redshift distributions as well as magnitude, size, mass and star formation rate distributions. The sample includes a large number of bright galaxies thanks to the large area / volume probed by the UltraVista survey. I then compare my photometric sample and the VUDS spectroscopic sample, and find that my sample is comparable over the same magnitude range but goes down more than one magnitude fainter and provides a larger sample for more robust statistics.

The core results of my PhD dissertation are presented in chapter 7. I compute the rest-frame UV luminosity functions, luminosity densities and dust-corrected star formation

rate densities from my deep UltraVista sample. Luminosity functions are significantly improving the accuracy of the bright end measurements compared to previous studies. I then derive luminosity densities which are consistent to previous studies but better constrained. The luminosity function at redshift $z \sim 6$ shows a significant excess at the bright-end compared to a simple Schechter function. It brings evidence that massive galaxies in this redshift range are more efficient to assemble mass than we expected before. I explain the importance of dust attenuation correction when transforming UV luminosity densities to star formation rate densities. Finally I present the star formation rate densities at $z \sim 5$ and $z \sim 6$, and show the mild evolution of the SFRD between redshift ~ 4 to redshift ~ 6 based on my results.

The last part of my work is in chapter 8, where I contributed to the work on the galaxy merger rate at a redshift $z \sim 3$ an important property to understand the evolution of the SFRD and mass assembly.

I give some conclusions and perspectives in Chapter 9

Chapter 2

Introduction on Galaxy Evolution and Star Formation

Galaxy evolution is one of the most active research areas in astrophysics with a lot of dedicated effort worldwide. People are trying to address some of the fundamental questions related to the formation and evolution of galaxies as follows:

- When and how did galaxies form?
- How and why did they evolve along cosmic time?
- What is the link between the formation of a galaxy and the formation of stars?

Even if we already have some general ideas based on previous studies, these questions are still a major focus of investigations. There are many different ways to study galaxy formation and evolution. Throughout my thesis I study galaxies in their early evolution stages at redshifts $4.5 < z < 6.5$, and derive their luminosity function when the universe was $\sim$1Gyr old.

In this chapter, I will give a brief introduction to galaxy formation and evolution in the context of the current cosmological model, with a summary of the fundamentals in galaxy formation in dark matter halos. I will describe the current knowledge obtained from deep galaxy surveys on the history of star formation and discuss the many questions that remain open.

2.1 Cosmological Background

Any study of galaxy evolution must consider the cosmological framework. In 1929, Edwin Hubble revealed one of the most important scientific discoveries in history. It suggests that distant extragalactic objects are found to have a Doppler shift interpretable as relative velocity away from the Earth and these recession velocities of various galaxies are proportional to their distance from the observer. This famous Hubble Law shows that our universe is expanding. Consequently, the universe was extremely dense and hot in the past. Moreover, this model suggests that at some moment all of space was contained in a single point, which is considered the beginning of the universe: $TheBigBang$. This discovery is the cornerstone of modern cosmology. The Hubble law in the local universe is expressed by the equation:

$$V = H_0 D \tag{2.1}$$

where "H_0" is the constant of proportionality (Hubble constant) between the distance "D" to a galaxy and its velocity "V". This is an direct interpretation of the expansion of the observable universe and the light of distant objects is redshifted owing to the expansion of the Universe. The light of farther objects is redshifted. The ratio of the observed to emitted wavelength is given by:

$$\frac{\lambda_o}{\lambda_e} = (1 + z) \tag{2.2}$$

where the quantity z is termed the redshift. The size of the Universe now compared to its size when the light was emitted is also $(1+z)$. Redshift is a measure of lookback time, owing to the finite speed of light. With the current measurements of the cosmological model parameters, the relationship between redshift and epoch is well established as shown in Figure 2.1.

In 1964 Arno Penzias and Robert Wilson accidentally discovered the cosmic background microwave radiation (CMB) as a radiation detected in the microwave bands at any sky location, now mapped over the whole sky (as shown in Figure 2.2). The CMB has an almost perfect black body spectrum and is assumed to be the remnant radiant heat from the "Big Bang". This substantial discovery provided a snapshot of the distribution of the very early seeds of the large scale structure and the distribution of galaxies. During the 1970s and 80s, several observations showed that there is not sufficient visible matter in the universe to account for the apparent strength of gravitational forces

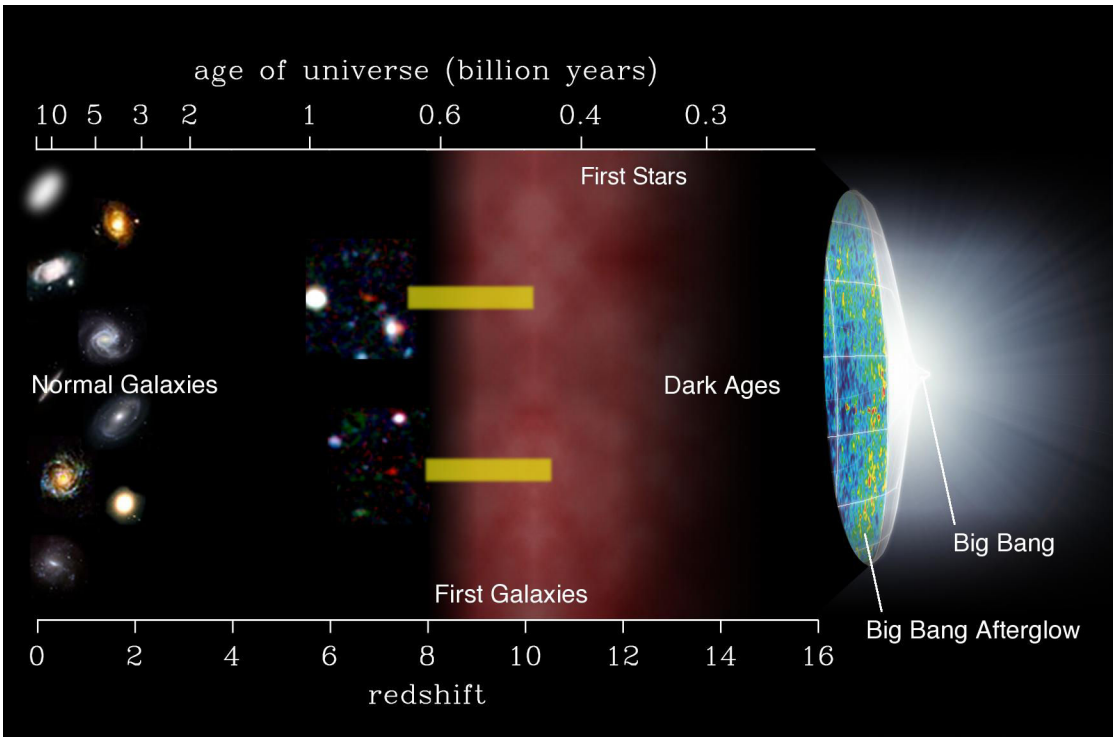


FIGURE 2.1: This figure shows the relation between redshift and the age of the universe. The timeline shows several major events in the history of the universe, beginning with the Big Bang and advancing through the haze of the Dark Ages, the formation of the first stars and galaxies, and eventually arriving at the present day, with normal galaxies like our Milky Way. The yellow bars next to the two high-redshift galaxies show the time period during which most of their stars formed. These galaxies might be the first in the universe to have formed, shortly after the first stars. (Image courtesy Ivo Labbé)

within and between galaxies. This led to the idea that up to 90% of the matter in the universe is dark matter that does not radiate or interact with normal baryonic matter. While dark matter has always been controversial, its presence is inferred by various observations: the anisotropies in the CMB, galaxy rotation curve, galaxy cluster velocity dispersions, large-scale structure distributions, gravitational lensing studies, and X-ray measurements of galaxy clusters. In the last decades of 20th century, measurements of high redshift supernovae provide a compelling evidence that our Universe is not only expanding, but it also has been accelerating since the universe was about half of its present age. To explain this acceleration, traditional cosmology requires that much of the energy in the universe consists of a component with large negative pressure: "Dark Energy". Negative pressure is believed to be a property of vacuum energy, but the exact nature and existence of dark energy remains one of the great mysteries of the current cosmological model. Possible candidates for dark energy include a cosmological constant and quintessence. Figure 2.3 shows the results from the Planck team in 2013, in accordance with a universe that consists of 68.3% dark energy, 26.8% dark matter, 4.9%. According to theory, the energy density of matter decreases with the expansion of the universe and the dark energy density remains constant as the universe expands.

Therefore matter made up a larger fraction of the total energy density of the universe in the past than it does today, but its fractional contribution will fall in the far future as dark energy becomes the dominant factor.

The model which currently is able to explain the various sets of observations and cosmological probes is called ΛCDM (Lambda Cold Dark Matter). It also can be parameterised by the different density parameters in each component, at the present day - radiation (Ω_r), matter(Ω_m) and cosmological constant (Ω_Λ) expressed as:

$$\Omega_\Lambda + \Omega_m + \Omega_r = \Omega_{total} \tag{2.3}$$

Here we account for negligible contributions from radiation in the matter-dominated epoch and present.

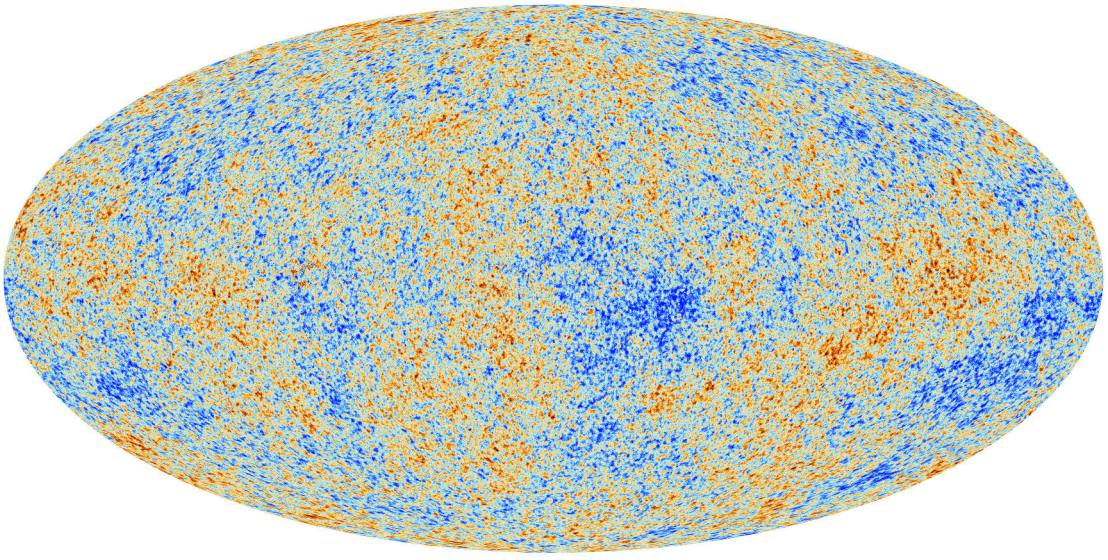


FIGURE 2.2: The detailed, all-sky picture of the infant universe created from four years of Planck data. The image reveals 13.77 billion year old temperature fluctuations (shown as colour differences) that correspond to the seeds that grew to become the galaxies. Credit: ESA and Planck Collaboration

2.2 Hierarchical Picture and Galaxy Formation in Dark Matter Halos

Galaxy formation is described in the context of structure formation theories as a result of tiny fluctuations produced shortly after the Big Bang. The simplest model for galaxy formation which is in general agreement with observed phenomena is the Λ CDM

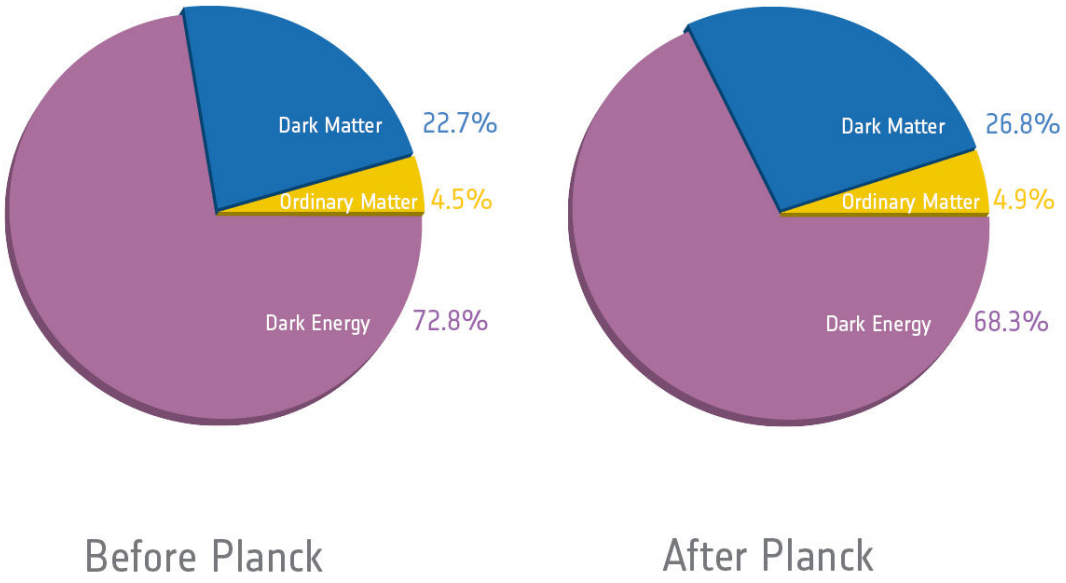


FIGURE 2.3: Planck's high-precision cosmic microwave background map has allowed scientists to extract the most refined values yet of the Universe's ingredients. Normal matter that makes up stars and galaxies contributes just 4.9% of the Universe's mass/energy inventory. Dark matter, which is detected indirectly by its gravitational influence on nearby matter, occupies 26.8%, while dark energy, a mysterious force thought to be responsible for accelerating the expansion of the Universe, accounts for 68.3%. The 'before Planck' figure is based on the WMAP 9-year data release presented by Hinshaw et al., (2012). Credit: ESA and Planck Collaboration

cosmology as described in the previous section. According to the ΛCDM paradigm, structure formation started from small overdensity scales and can be observed as primordial fluctuations in the CMB (as shown in Figure 2.2). After this epoch the small scale perturbations started to grow with time under gravity − over-dense regions pull matter from under-dense regions. When the overdensity of perturbation exceeds the average critical density then it starts to collapse into a dark matter halo under its own gravity. Within the currently favoured cosmology structure forms hierarchically, from the bottom-up. Dark matter halos (and possibly the galaxies they host) are built by a series of discrete merging events and typical dark matter halo mass expected to host the first galaxies at the redshifts 4.5−6.5 as shown in Figure 2.4, which is the crucial redshift range related to our work. The history of dark halo mergers can be illustrated in the form of the 'merger tree'. The merger process occurs in many different types, in general the two most important types can be defined as:

- **Major merger** − Two dark haloes and their associated galaxies merge with approximate masses and sizes, then produces the more massive central galaxy inside the halo.
- **Minor merger** − One of the haloes is significantly larger than the other(s) in size and mass. The smaller dark matter components merge into the massive dark matter halo.

Discussing galaxy formation, we are concerned with the interior baryonic matter of the dark matter halo, which is more complex. The galaxies inside the dark matter haloes continue to grow in mass (and size) either by accreting material from their neighbourhood or by merging with other galaxies, but also drives away much of the dust and gas (mass loss) through a variety of stellar feedback mechanisms or central feedback mechanisms that often include a stage in which there are active galactic nuclei (AGN). This is thought to be the driving force behind many quasars. Star formation quenching is an important aspect of galaxy evolution following the transformation of late-type star-forming galaxies into early-type passive galaxies. Quenching can be produced by a number of processes including stellar or AGN feedback or environmental processes like gas stripping.

This merging picture is issued from theoretical prediction, but observations also show many examples of merging galaxies (as shown in Figure 2.5). One study found that large galaxies merged with each other on average once over the past 9 billion years (z$\sim$ 2−3). Note that the Milky Way and the Andromeda Galaxy are believed to be in a collision path in about 4.5 billion years. The merging of these galaxies would classify as major as they have similar sizes. The result would likely be an elliptical galaxy is we follow detailed simulations. The larger galaxy will often "eat" the smaller, absorbing most of its gas and stars with little other major effect on the larger galaxy. Our home galaxy, the Milky Way, is thought to be currently absorbing smaller galaxies in this fashion, such as the Canis Major Dwarf Galaxy, and possibly the Magellanic Clouds. The Virgo Stellar Stream is thought to be the remains of a dwarf galaxy that has mostly merged with the Milky Way. In chapter 8, we present a detailed discussion about the importance of merging events in galaxy evolution, from a measurement of the merger rate at high redshifts ($2 < z < 4$).

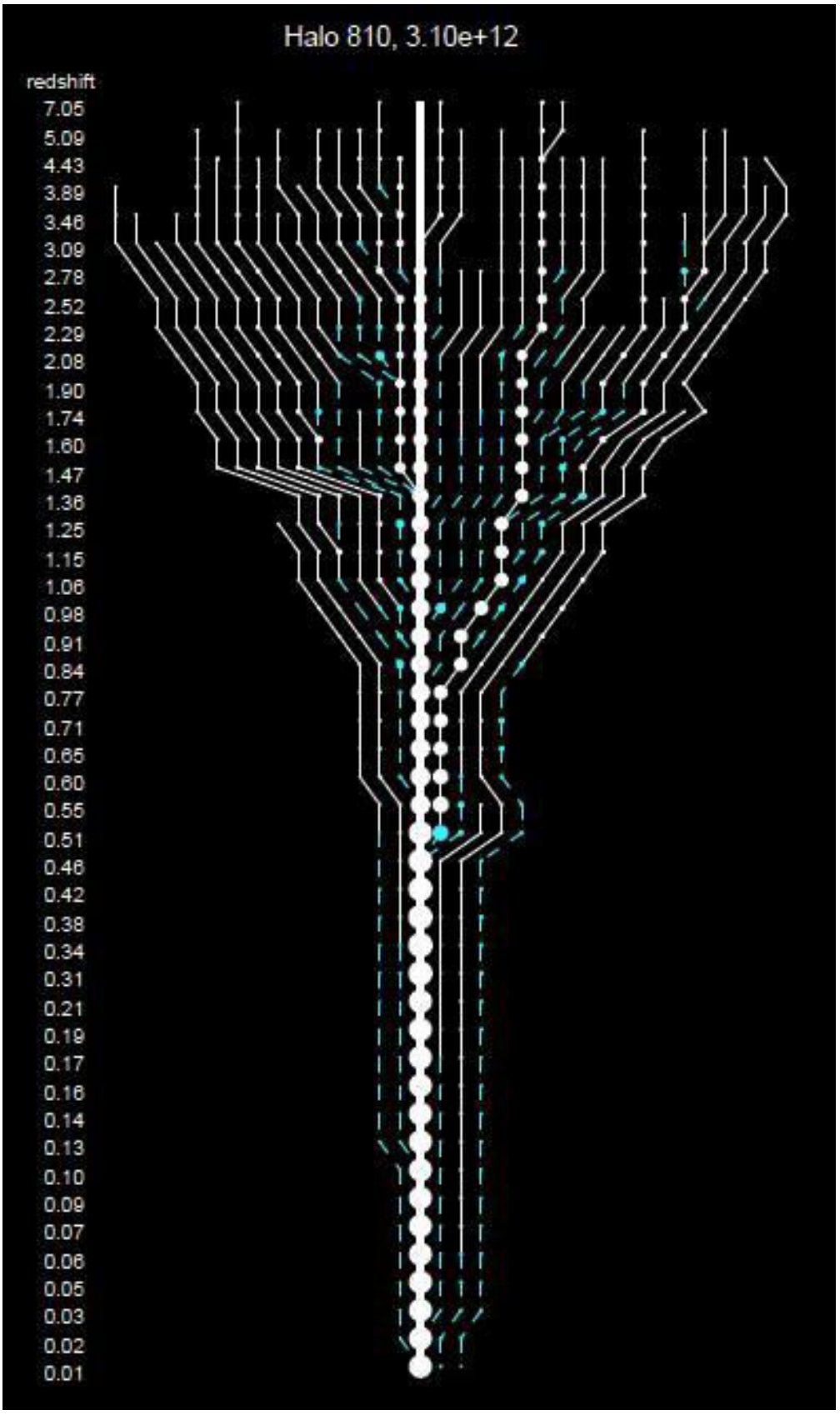


FIGURE 2.4: Example of a "merger tree" showing the merger history of a single, Milky Way-size dark matter halo over cosmic time, extracted from a cosmological N-body simulation. The redshift of each snapshot is printed on the left-hand side. White lines represent isolated halos that will eventually merge into the main progenitor halo (the centre line). White lines become cyan when these infalling halos cross the main progenitor's virial radius. This image (colours modified) is taken from Stewart et al. (2008).

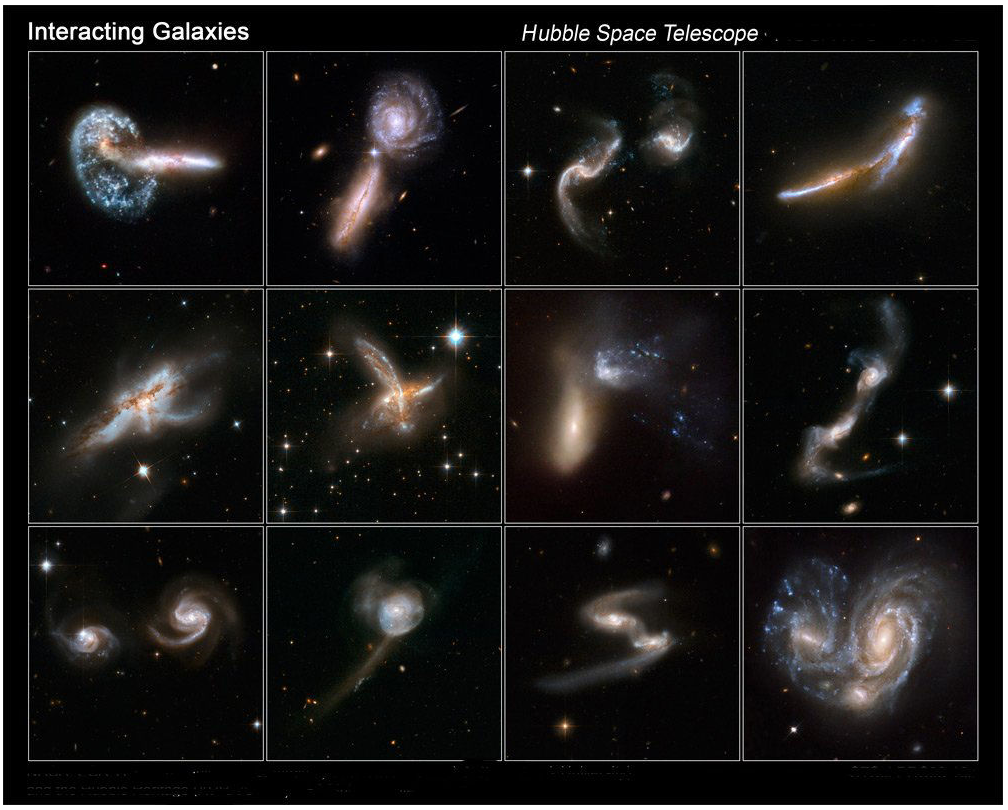


FIGURE 2.5: Collection of Hubble Space Telescope (HST) images of interacting galaxies and merging galaxies. Credit: NASA, ESA, the Hubble Heritage (STScI/AURA)-ESA/Hubble Collaboration, and A. Evans (University of Virginia, Charlottesville/NRAO/Stony Brook University

2.3 Historical Perspective on High Redshift Galaxies Surveys

A redshift survey is a survey of a field of the sky to measure the luminosity, redshift, and other properties of astronomical objects : usually galaxies and also galaxy clusters and quasars. These surveys aim to provide fundamental data on galaxies and their distribution. Basically surveys can be divided into two types : **photometric surveys** and **spectroscopic surveys**.

- **Photometric surveys**– Photometric surveys aim at getting magnitudes of objects in a selected field of sky imaged by a wide-field instrument; Multi-wavelength photometry can be used to get a photometric redshift and the spectral energy distribution (SED) can give some estimates of physical parameters like stellar mass, star formation rate. Additionally, colour selection may also be applied to assist discrimination between stars and galaxies. The advantages of photometric survey are wide field cameras which allow us to observe a large number of objects, and the depth of the photometry, and the main disadvantage of photometric redshifts is

that they are not precise enough to trace large scale structures or identify galaxy pairs. Several large photometric surveys have been conducted recently. In the optical bands this includes the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) obtained with the Canada France Hawaii Telescope (CFHT), the Cosmic Evolution Survey (COSMOS) done with the Subaru telescope and the Hubble Space Telescope (HST). In the near-infrared one can cite the Cosmic Assembly Near-infrared Deep Extragalactic Survey (CANDELS) with HST ACS and WFC3 observations, and the Ultra-Deep Survey (UltraVista) with the VISTA telescope. With the deepest photometric surveys like those done with HST, searches for very-high-redshift galaxies have been pushed in the past decade to identify some candidates galaxies over $z \sim 10$.

- **Spectroscopic surveys**– Spectroscopic surveys aim to get redshifts and spectra. The photometrically selected galaxies are observed by spectroscopy, most commonly at visible wavelengths, to measure the wavelengths of prominent spectral lines; comparing observed and theoretical wavelengths then gives the redshift for each galaxy. Advantages of spectroscopic survey are the precision of redshift and spectral line indicators. In 1977, the first systematic redshift survey was the Center for Astrophysics (CfA) Redshift Survey of around 2 200 galaxies which was the first attempt to map the large scale structure of the universe. After 90s, astronomers obtained spectra for several hundred galaxies simultaneously due to the invention of fibre-optic spectrographs and multi-slit spectrographs, and much larger redshift surveys became feasible. Famous examples of survey in the local universe ($0 < z < 0.5$) are the 2dF Galaxy Redshift Survey (221 000 redshifts, completed 2002 as shown in Figure 2.6); the Sloan Digital Sky Survey (SDSS, approximately 2 million redshifts by 2013). After the initial success of the Canada-France-Redshift Survey (CFRS, Lilly et al., 1995, ApJ, 455, 108), the most notable large high redshift surveys are the zCOSMOS Survey ($1 < z < 3$), the VIMOS-VLT Deep Survey (VVDS) ($0 < z < 5$), the DEEP2 survey and the VIPERS survey at $z \sim 1$ and the VIMOS Ultra Deep Survey (VUDS) ($2 < z < 6$) ; these have around ∼30 000, ∼50 000, ∼30 000, ∼100 000 and ∼10 000 redshifts respectively, and are mainly focused on galaxy evolution. There are also several attempts at $z > 6$ (Vanzella et al. 2009 ; Ouchi et al. 2010; Stark et al. 2010, 2011; Jones et al. 2012; Schenker et al. 2012). Table 3.1 also lists several notable redshift surveys in the past twenty years.

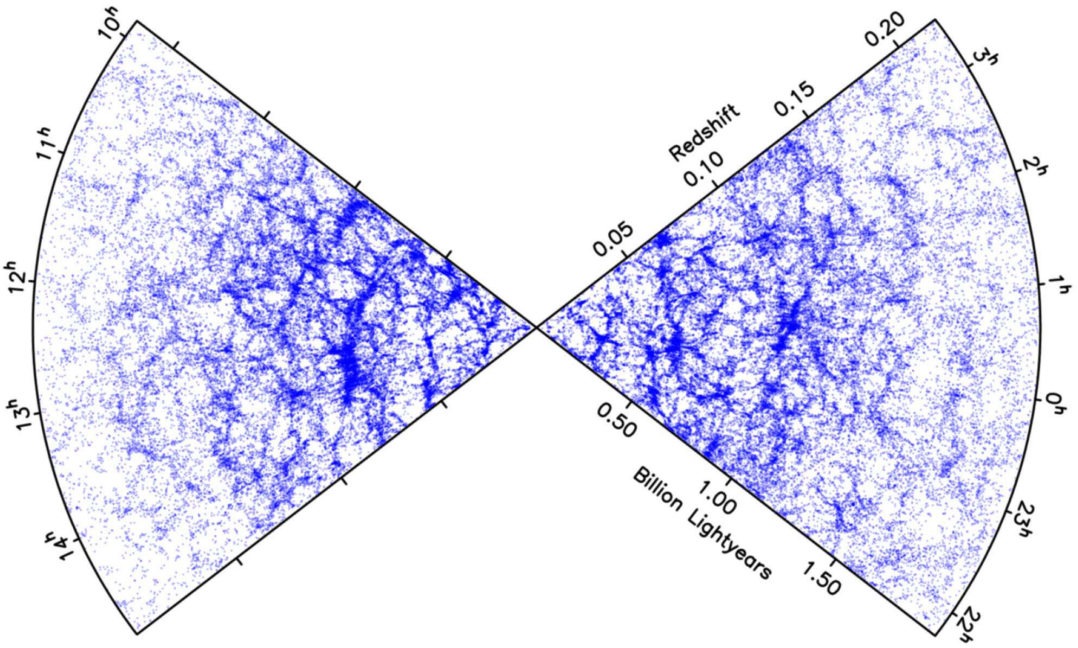


FIGURE 2.6: The projected distribution of galaxies in 2dF galaxy redshift survey, as a function of redshift and RA. Credit (http://www.roe.ac.uk/ jap/2df)

2.4 Current Picture on Star Formation Evolution

The evolution of the star formation rate (SFR) density is considered one of the most fundamental observational descriptions of the galaxy population as a whole. Matching and explaining this observed evolution is often used as a benchmark of success for theoretical models of galaxy formation. The evolution of the SFR density is constructed by integrating the luminosity function at a specific wavelength sensitive to star formation (e.g., rest-frame UV, Hα, far-IR, or radio) in order to obtain the associated luminosity density. Then, a conversion between luminosity and SFR (as described in chapter 3.4) is used to obtain the associated SFR density. The first attempts to chart the cosmic star-formation history were presented by Madau et al. (1996) SFR using the evolution of the luminosity density from the Canada-France Redshift Survey (CFRS, Lilly et al. 1996) up to $z \sim 1$ and samples at $z \sim 3-4$ identified with HST. Since then, extensive new Star Formation Rate Density (SFRD) measurements have been obtained, with a summary from $z \sim 0$ to $z \sim 10$ recently presented in the review by Madau and Dickinson (2014). Different SFR indicators are used including rest-frame UV, far-IR, radio, and X-ray continuum luminosities, and both Balmer and forbidden [OII] emission lines. The recent result of dust-corrected SFRDs are shown in Figure 2.7 as obtained by Cucciati et al. 2012 based on VVDS-Deep and VVDS-UltraDeep surveys. It shows the peak in the SFRD evolution at $z \sim 2$. This peak is preceded at earlier cosmic epochs by a rapid increase of a factor 6 from $z \sim 4.5$, then followed by a decrease by a factor of 12 to

the local universe. In several versions of the cosmic star formation history, the global SFRD remains roughly constant between $z \sim 2$ and $z \sim 4$, and then declines toward higher redshift. At $z > 4.5$, there is a large uncertainty in measurements and lack of observations compared to the results at $z < 4.5$. Bouwens et al. (2014) compiled all the data from the deep HST imaging programs (GOODS, COSMOS, CANDELS,...) to produce an average of the SFRD up to redshifts $z \sim 10$ (shown in Figure 2.8). From this compilation Bouwens et al. claim that the SFRD continuously decrease from $z \sim 2$ to $z \sim 10$, with even a faster decline above $z \sim 8$. However data above redshift 4 associate large uncertainty because these very deep surveys came from the small area. As we mentioned in the previous section, $z > 4.5$ is a crucial epoch for studying galaxy formation which requires improved measurements of the SFRD at $z > 4.5$. This is the main goal of this thesis..

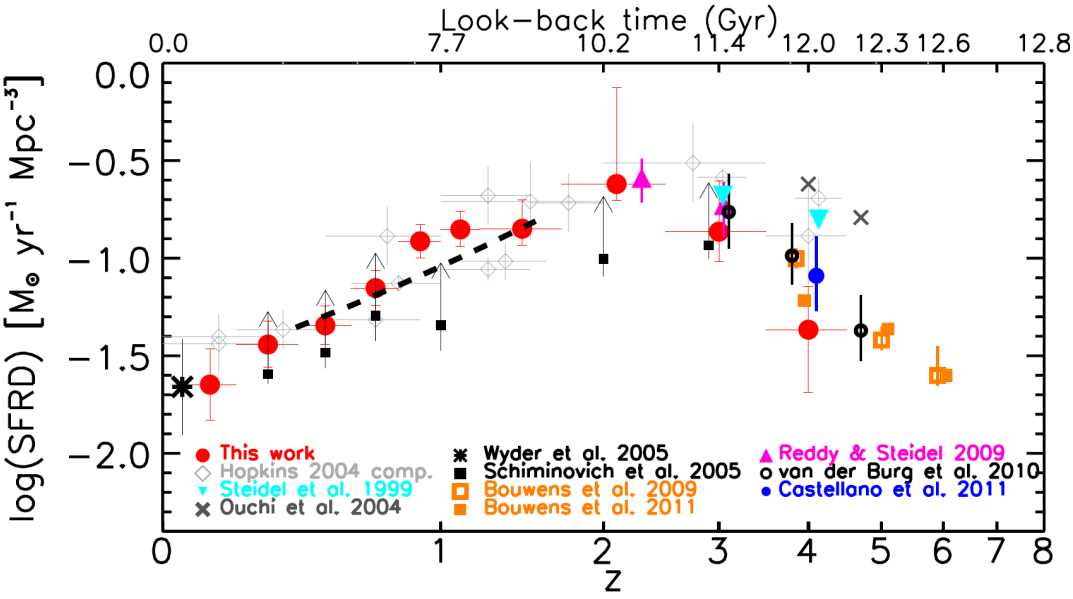


FIGURE 2.7: Total dust-corrected UV-derived SFRDs as a function of redshift from the VVDS Deep+Ultra-Deep sample (red filled circles) by Cucciati et al. 2012. The black dashed line is the SFRD(z) implied from the stellar mass density in Ilbert et al. (2010). Their SFRDs, as well as those from the literature, are derived using the FUV-band Luminosity Densities converted into SFRD with the scaling relation from Madau et al. (1998). All data have been homogenised with the same IMF (Salpeter 1955). Compilation in Hopkins (2004) taking only the FUV and NUV determinations (grey open diamonds); Steidel et al. (1999) (LBG sample, cyan solid upside-down triangles); Ouchi et al. (2004) (SDF and SXDF LBGs sample, thick grey crosses); Wyder et al. (2005) (GALEX-2dFGRS, bold asterisk); Schiminovich et al. (2005) (GALEX-VVDS at $z < 1.2$ and HDF above, small solid squares with arrows); Bouwens et al. (2009) and Bouwens et al. (2011) (LBGs, orange open and filled squares, respectively); Reddy & Steidel (2009) (LBG, solid magenta triangles; van der Burg et al. (2010) (CFHTLS, black open circles); Castellano et al. (2012) (LBG, blue open circles).

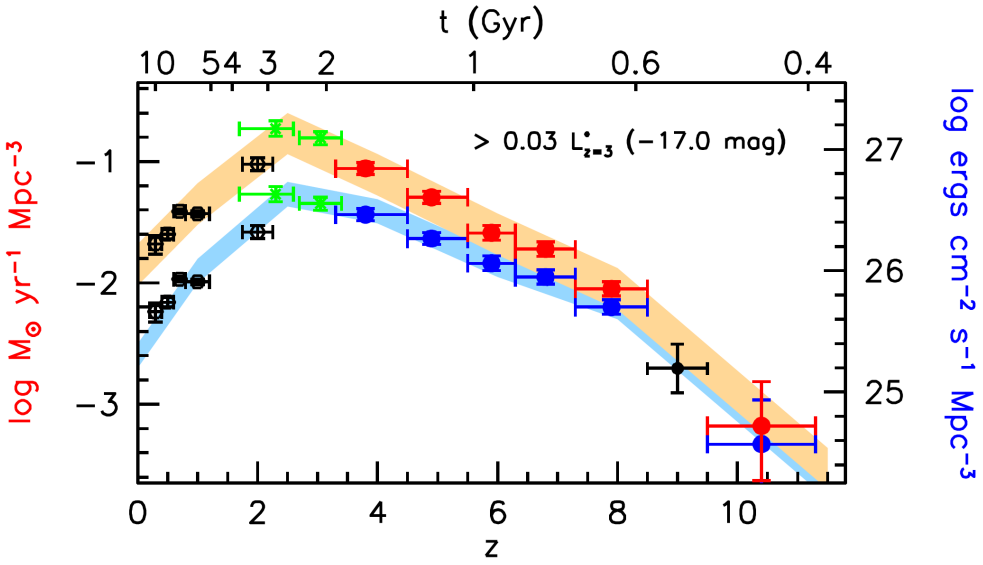


FIGURE 2.8: Updated Determinations of the derived SFR (left axis) and UV luminosity (right axis) densities versus redshift. The left axis gives the SFR densities would infer from the measured luminosity densities, assuming the Madau et al. (1998) conversion factor relevant for star-forming galaxies with ages of $> 10^8$ yr (see also Kennicutt 1998). The right axis gives the UV luminosities we infer integrating the present and published LFs to a faint-end limit of ~ 17 mag (0.03 L_* z =3) which is the approximate limit to $z \sim 8$ data set. (Credit: Bouwens et al. 2014)

2.5 Goal of this Thesis Work

In this thesis I aim to measure the luminosity function at $4.5 < z < 6.5$ using the largest and deepest infrared imaging survey available today, UltraVista. I perform ultra-deep multi-band photometry in $YJHK_s$ bands, matched to the $BVriz'$ photometry in the COSMOS field. I identify galaxies in two redshift ranges, $4.5 < z < 5.5$ and $5.5 < z < 6.5$ based on photometric selection criteria. I compare my samples with existing spectroscopic samples at the same redshifts. I then derive galaxy counts and produce LF to finally compute the SFRD. In the remainder of my thesis is organised as follows: Chapter 3 provides an overview of the many different techniques that have recently been employed for identifying high redshift ($z > 4.5$) galaxies. The existing high redshift samples from the literature available in the last years and current measurements of physical properties of galaxies will then be described. Chapter 4 describes the construction of my photometric catalogue from UltraVista and multi-wavelength data and the basic data assessment. Chapter 5 describes how I extract a robust sample of galaxies at $4.5 < z < 6.5$, mainly using photometric redshifts. I produce photometric simulations to identify the selection function of my photometric samples, and I compute ideal galaxy colour-tracks from SED templates. I describe the procedure used for image simulation and how I retrieve the important statistical factors for the estimation of LF, such as detection rate, completeness rate and contamination rate. Chapter 6 lists the properties of the selected high redshift samples at $4.5 < z < 6.5$. The LFs and the evolution of the LFs is presented and discussed in chapter 7 where I then derive the SFRD. In chapter 8, I focus on the merging of galaxies, one of the main physical mechanisms going along with the evolution of the LF, SFRD and stellar mass, and present our work on the galaxy merger rate at $z \sim 3$. Finally, the conclusions are summarised in chapter 9. Unless otherwise stated, all results presented here assume a concordance cosmology with parameters $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, $\mathrm{H}_0 = 70 km/s/Mpc$. Unless specified, all magnitudes are in AB system (Oke, J.B. 1974).

Chapter 3

Identification of Galaxies at $z > 4.5$ and Measuring Their Properties

As discussed in chapter 2, observations of galaxies with very high redshifts play a central role in cosmology because they provide insight into the star formation history, mechanisms of galaxy formation, and mass assembly processes. In this chapter, I will discuss the strategies for identifying galaxies at very high redshift and present the main galaxy surveys that have been conducted, focusing on $z > 4.5$. I then present the luminosity function, which is a fundamental tool used to uncover the evolutionary processes, and its application to our infant universe.

3.1 Methods to Identify High Redshift Galaxies

Because of their extreme distances, galaxies at high redshift naturally appear to an observer as faint objects, so it is very challenging to find them even if their intrinsic luminosities are already large. Furthermore, distant galaxies are rare on the plane of the sky compared to the very high number of foreground galaxies, making the search even more difficult. There are two main techniques to locate galaxies at very high redshifts $z > 4.5$, both based on the spectral imprint of hydrogen. In the first case, broad-band photometry aims at identifying absorption breaks due to neutral hydrogen surrounding the source and hydrogen clouds present on the line of sight between us and the source (inter-galactic medium - IGM). In the second case, narrow-band techniques target the strong emission in the Lymanα line. In this section, we review all of these most common and complementary techniques for identifying distant objects.

3.1.1 Lyman Break Technique

The Lyman-break technique is a technique to search for star-forming galaxies at high redshift proposed in the early 1990s. This technique exploits a major feature of the UV spectra of galaxies with ongoing star formation, namely the hydrogen ionisation edge or Lyman limit. This method works very well to select high redshift objects because light with wavelength shorter than 912Å is entirely absorbed by neutral hydrogen gas, which is very abundant within and around star-forming galaxies. As a result of this absorption of ionising photons, the spectrum instantly dims by more than an order of magnitude at wavelengths shorter than the Lyman limit (Lyman-break). It is a very efficient method because the wavelength of the filter bandpass is what determines the redshift range of the candidates. The filter suite shown in Figure 3.1 is very sensitive in the redshift range $2 < z < 3.5$. However, by excluding the U band and using the B band to probe shortward of the Lyman limit, together with the V and I bands to probe longward of it, one can target higher redshift intervals, which in this case $3.5 < z < 4.5$. Other redder filters can be used for even higher redshift intervals. For searching even higher redshift objects ($z > 5$), additional absorption is introduced by the IGM, in which the light from the galaxy travels through in its journey towards the observer. In such cases, the continuum break limit changes into 1216Å from 912Å producing what is commonly referred to as a drop-out galaxy.

In history, this technique has been used to identify thousands of galaxies (so-called Lyman break galaxies, or LBGs) at $z \sim 3$ (Steidel et al. 1996, 2003), and even push to higher redshift by using different filter sets, at $z \sim 4, 5$ (Ouchi et al. 2004) $z \sim 6$ (Bouwens et al. 2007) and beyond. Figure 3.2, from Steidel et al. (2004), provides an illustration of the rest-frame UV-selection criteria in UGR colour space, tuned to find galaxies at $2.7 < z < 3.5$ at $dz = 0.1$ redshift interval. The advantages of this simple colour selection is that it can not only be applied using imaging with broad-band filters, which allows potentially large samples of high-redshift galaxies, but also yields redshifts accurate to $\delta z = 0.1 - 0.2$. However, for a proper assessment of completeness/contamination and the determination of redshifts with sufficient accuracy and precision to allow robust clustering measurements, spectroscopic follow-up is still essential. Given sufficiently good signal to noise ratio and appropriate broad-band filters, the selection of Lyman break galaxies at $z > 4.5$ is a reasonable option because the strong break produced by the combination of the IGM extinction and Lyman limit is relatively easy to identify. Although spectroscopic measurements are necessary to obtain a high-precision redshift, spectroscopy is typically much more time consuming than imaging, so

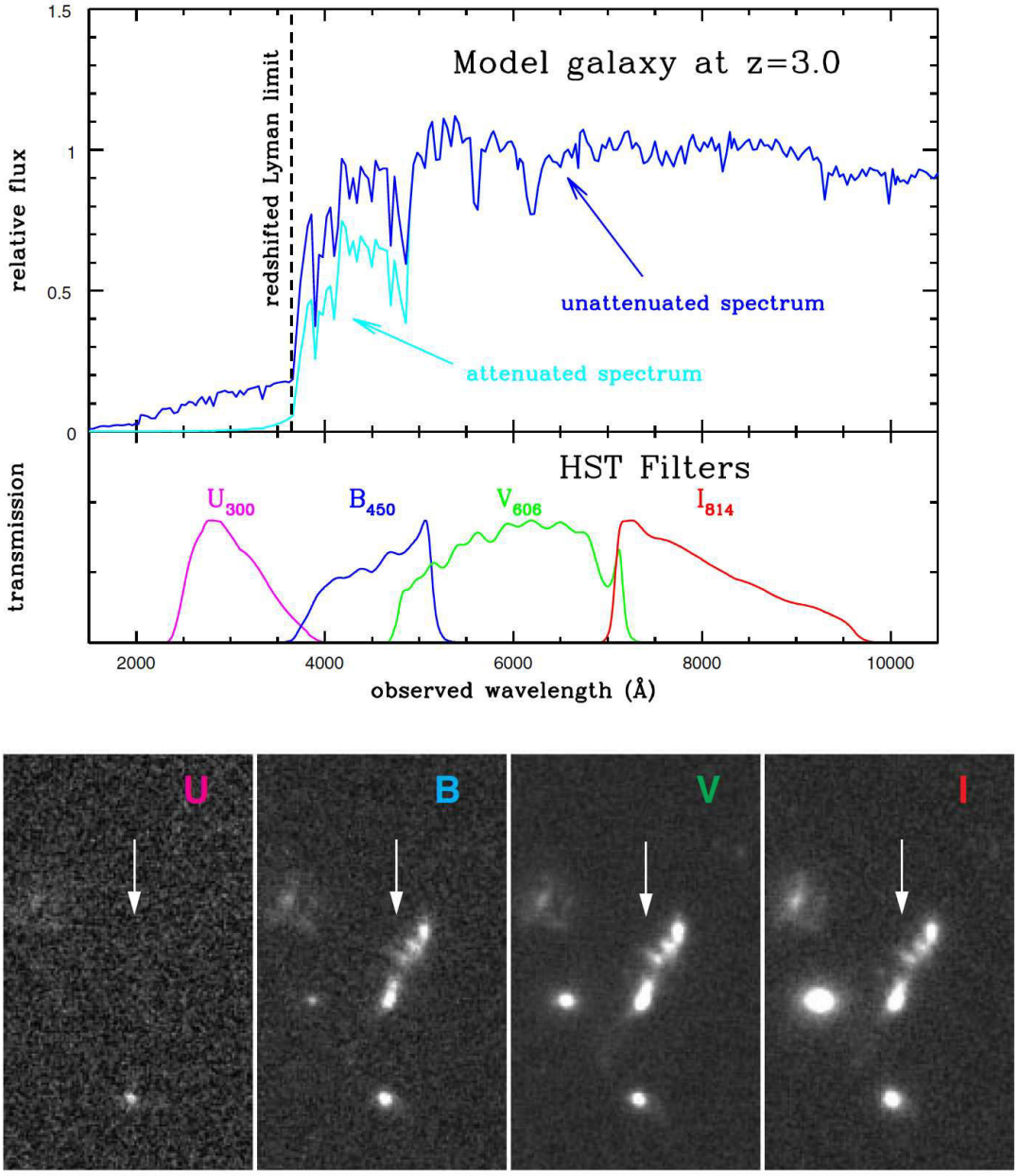


FIGURE 3.1: The UV spectrum of a simulated star-forming galaxy observed at redshift z =3, and the idea behind the Lyman-break technique. The top panel shows that the spectrum redshifted at z =3 shows a sharp decrease of the intensity of the light at wavelengths shorter than the Lyman limit. The attenuated spectrum illustrates the absorption of UV light by intervening neutral hydrogen. The middle panel shows the transmission curves of the four filters used to observe the Hubble Deep Field (HDF) with HST. When observing through a given filter, only radiation with wavelengths within the filter curve can reach the detectors. The bottom panels reproduce the images of a real Lyman-break galaxy at z =2.8 (the multiple object with a sausage-like morphology indicated by the arrow) through each filter. Notice how the galaxy drops out of the U filter, since this samples the spectrum to the left of the Lyman limit. Three other galaxies can be observed which do not drop out from the U filter (they are actually detected in the filter), but rather progressively dim, indicating that they are at relatively modest redshifts. Credit: (R.Ellis 1998)

the pre-selection of candidate galaxies via the Lyman break technique greatly improves the efficiency of high redshift galaxy spectroscopic surveys.

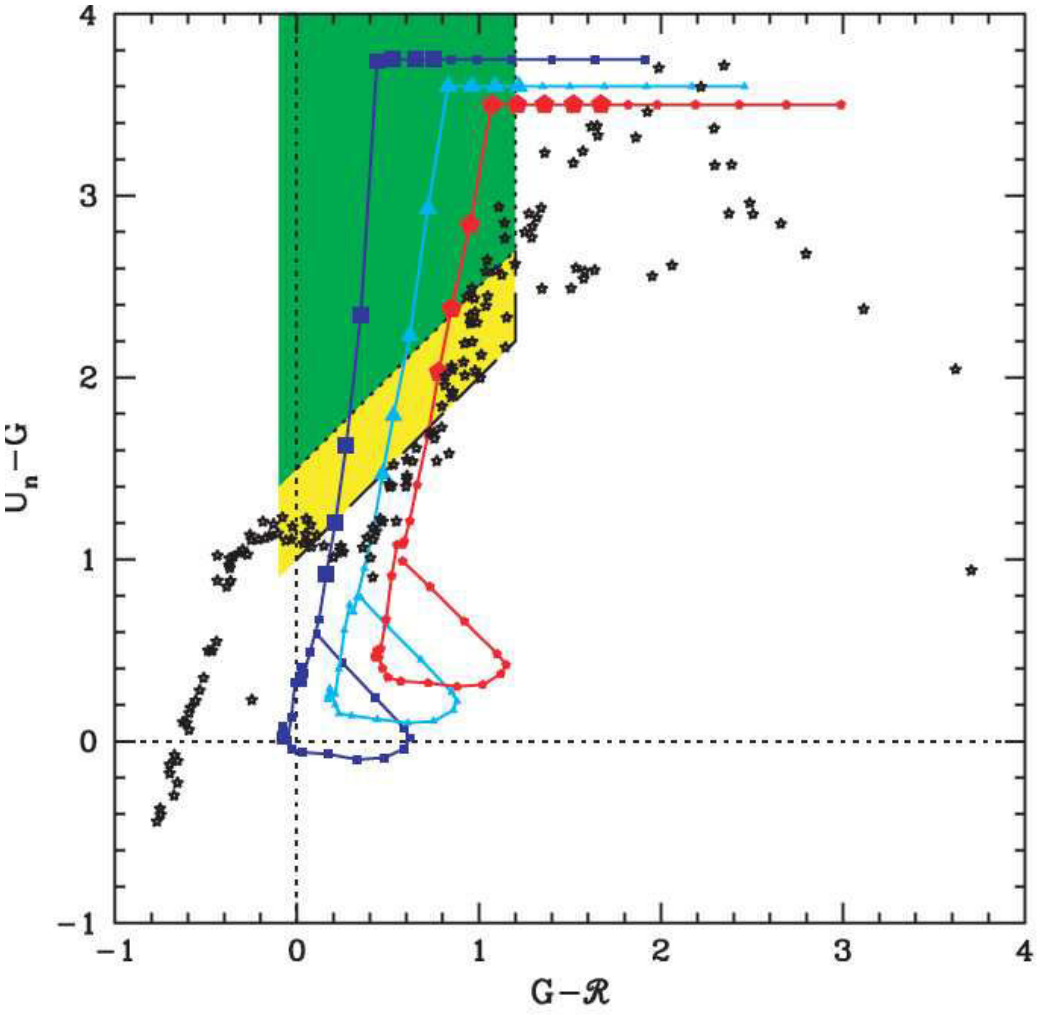


FIGURE 3.2: Expected colours for model star-forming galaxies in the targeted redshift range, for three assumed values of internal extinction $E(B\text{-}V) = 0$, 0.15, 0.30 using the Calzetti et al. 2000 prescription, for squares, triangles, and pentagons, respectively. The large points on each curve correspond to galaxies in the redshift interval $2.7 \leq z \leq 3.4$. The $U_n - G$ colours have been truncated for clarity, in practice, limited dynamic range will prevent $U_n - G$ limits from exceeding 4.0. These points are discussed in some detail in Steidel et al. (1999). The expected location of the stellar locus, based on the Gunn & Stryker (1983) atlas, is shown with black stars.

3.1.2 Narrow Band Selection

In contrast to the identification of high-redshift galaxies on the basis of broadband photometric filters, the use of a narrowband filter is efficient at selecting high redshift star-forming galaxies with strong emission lines. The most common emission line for which narrowband filters are designed for is hydrogen Lyman$-\alpha$ at the rest wavelength λ=1216Å, the transition between energy levels 1 and 2 of the hydrogen atom. The

young massive hot OB stars produced in galaxies with undergoing star-formation emit intense UV radiation and ionise Hydrogen, going back to the fundamental state with the emission of a Lyman$-\alpha$ photon. The Lyman$-\alpha$ line emission is redshifted as expressed in equation 3.1, then astronomers can use a narrow band filter tuned to the redshift of interest to find an excess flux compared to adjacent filters.

$$1 + z = \frac{\lambda_{obs}}{1216\mathring{A}} \tag{3.1}$$

Figure 3.3 (from Gronwall et al. 2007), shows the filter set used to identify Lyman$-\alpha$ emission at $z = 3.11$. The narrow-band filter is centred at $\lambda = 5000\mathring{A}$. The comparison of the flux in the narrowband filter and in broadband filters close in wavelength, allows to identify excess emission in the narrow band, and the identification of Lyman$-\alpha$ emitters (LAEs). There are many impressive results to date particularly obtained from the Suprime$-$cam imaging camera on the Subaru 8m telescope (Hu et al. 2004, Taniguchi et al. 2005, Ouchi et al. 2005, Iye et al 2006, Ota et al. 2010) and more than 2,000 LAEs have been identified with ground-based telescopes at $z > 2$. In order to get theses emitters with highest signal-to-noise ratio (S/N), narrow band filters are normally designed to avoid strong sky emission features. The OH molecules in the Earth atmosphere produce bands of emission which leave only a few wavelength windows free of sky emission, corresponding to $z = 4.7, 5.7, 6.6, 6.9$. Compared to the lyman break technique, the vast majority of objects with a narrowband excess when selecting LAE candidates at $z > 5$ are contaminants. Many of these are emission-line objects (galaxies or AGN) at lower redshifts, with the narrow-band excess being produced by, for example, CIV emission at 1549Å, MgII at 2798Å, [OII] at 3727Å, [OIII] at 5007Å, or Hα at 6563Å. The star-forming LAEs tend to be significantly fainter on average than the UV-continuum-selected objects (LBGs), and therefore offer a probe of the faint end of the luminosity function. However, the faint nature of these objects leads to a challenge in assembling high S/N multiwavelength imaging and spectra for individual LAEs.

3.1.3 Photometric Redshift Selection

Using photometric redshifts to infer the high redshift nature of galaxies has been widely used in recent years because this can provide redshift information for large number of faint galaxies for which spectroscopic follow-up is very costly except for a relatively small and bright fraction of the galaxy population. Despite being less accurate and precise than spectroscopy, photometric redshifts still provide a way to estimate distances for galaxies too faint for spectroscopy or samples too large to be practical for complete spectroscopic coverage. Recently, the number of studies which utilise this method for a

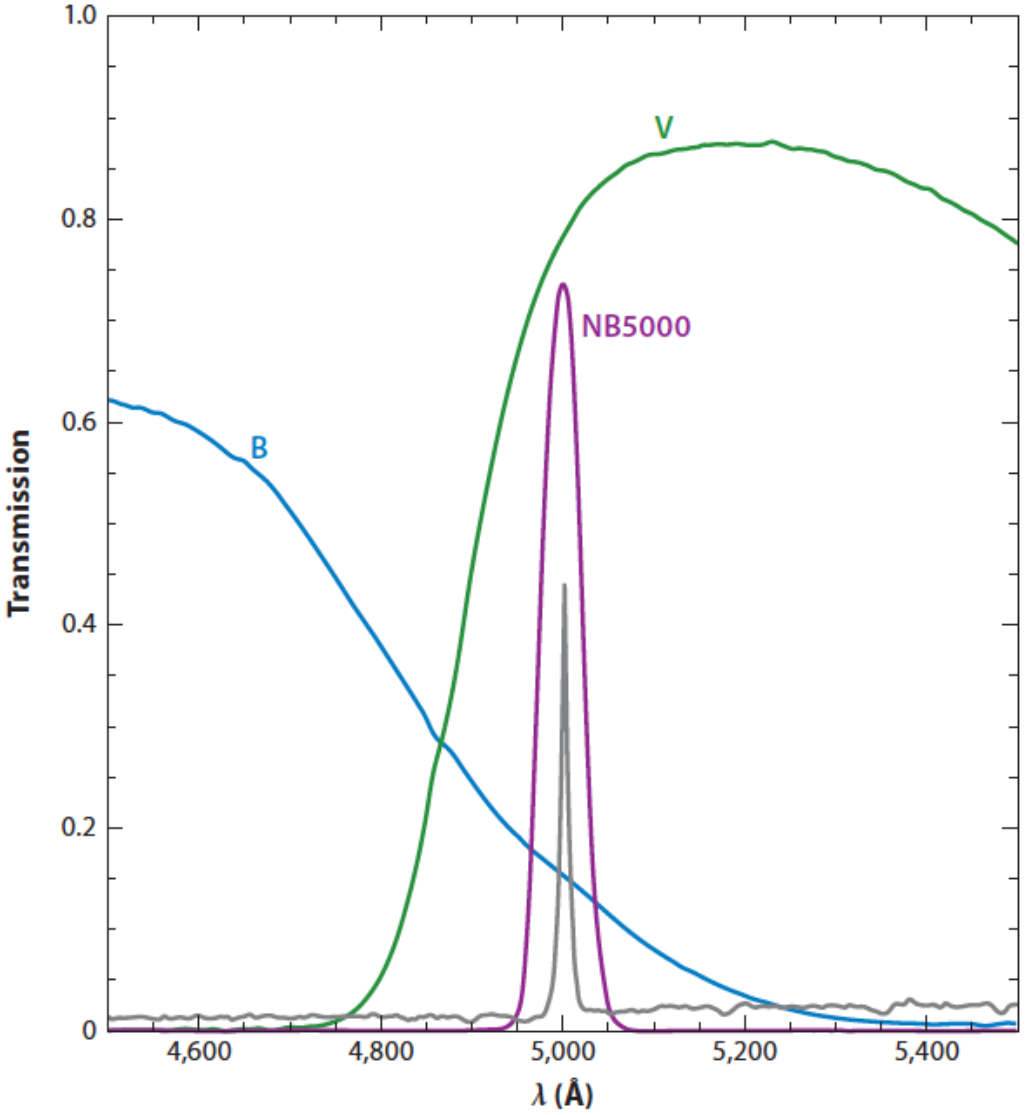


FIGURE 3.3: Filter bandpasses used for narrowband selection of Lyα emitters (LAEs) at z = 3.1, where a spectrum of a typical Lyα-emitting galaxy (overlaid in grey for comparison) would be detected as having a red broadband minus narrowband colour. This technique preferentially selects objects with bright line emission (and, often, faint continua). (From Gronwall et al. 2007.)

number of applications have grown rapidly.

Photometric filters are sensitive to a range of wavelengths and the technique relies on making many assumptions about the nature of galaxy spectra. This technique is usually divided into two groups, template fitting and empirical fitting. The template fitting technique derives the photometric redshift by minimising the χ^2 value when comparing an observed SED with the SED computed from a template library that includes spectral-energy distributions for a variety of galaxy types (representing different redshifts, star formation histories, chemical abundance, and mixtures of dust and stars). The empirical technique uses a training set of galaxies with known spectroscopic redshifts to derive a relation between observed photometry and redshifts. There is a large number of codes for both techniques, many of which are publicly available today. Figures 3.4 shows how this technique works using the template spectrum of an elliptical galaxy. As the galaxy gets closer to us (decreasing in redshift), the spectrum we observe shifts through the various bands, and different amounts of light will be measured in each. Using a template library of various galaxy types, one can then compare the observed colours in each band to the colours that would be produced by each type of galaxy at various redshifts. When the method works, the closest match will correspond to the template that is the correct galaxy type at the correct distance. Between the two methods most largely used, here we choose the SED-fitting because the lack of spectroscopic data in our research redshift range ($4.5 < z < 6.5$) and the flexibility when different sets of data are used. The distribution of χ^2 as a function of redshift is used to compute the redshift probability distribution function (PDF). In the PDF a good photometric redshift measurement has a single strong and narrow peak, while degenerate solutions would show two or more peaks. In this study I use the code Le Phare developed by S. Arnouts and O. Ilbert. This is one of the most commonly used codes at the international level today. It will be extensively described in the following chapters.

3.2 Existing Surveys / Samples

As mentioned in the first chapter, over the past 20 years galaxy surveys have become a powerful tool for studying galaxy evolution, so the number of surveys has truly exploded (as shown in Table 3.1 and 3.2). Large telescopes, in space and on the ground, are capable of taking images of distant galaxies or obtaining spectroscopic data, reaching all the way to the currently highest redshift galaxy with a confirmed spectroscopic redshift at $z = 7.51$ (Finkelstein et al. 2013) and dropout candidates to beyond $z = 10$ (Bouwens et al. 2011, shown in Figure 3.5). But it is increasingly difficult to obtain optical and

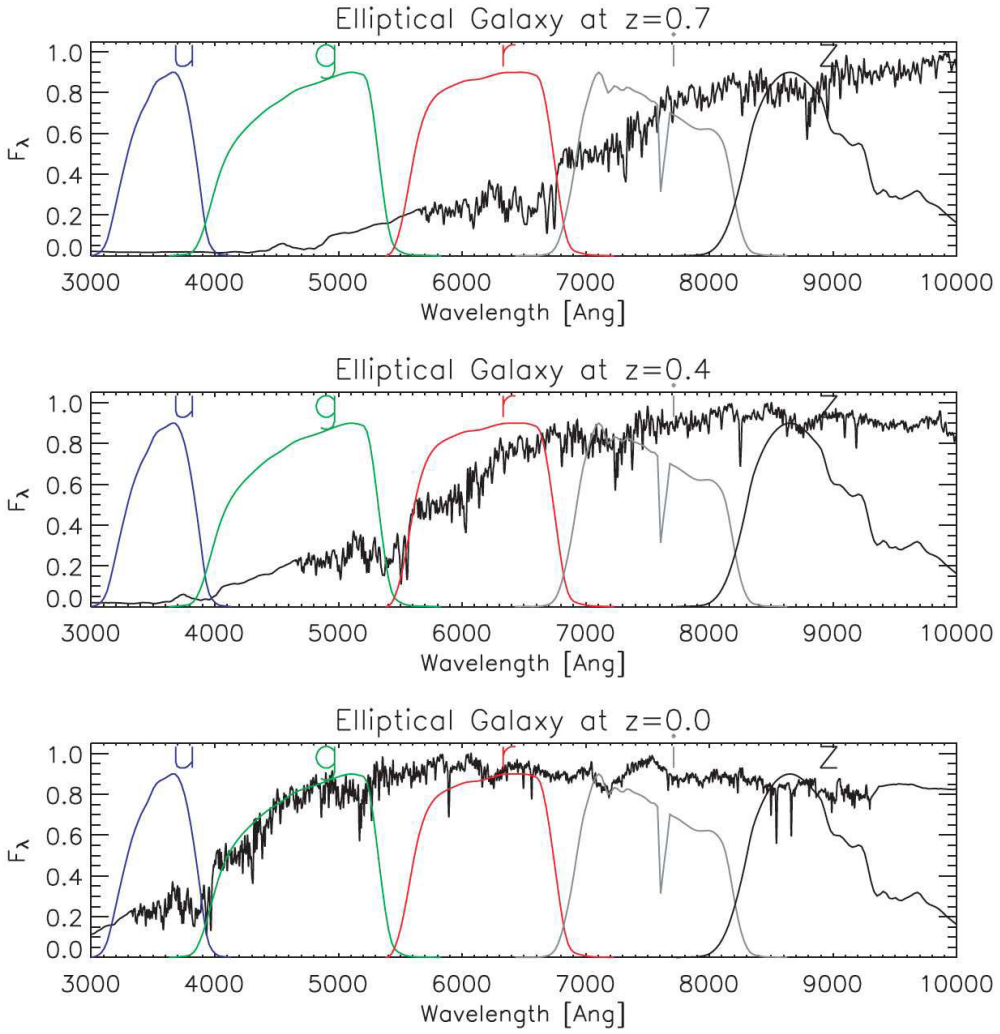


FIGURE 3.4: A model spectrum of an elliptical galaxy, taken from Bruzual & Charlot (2003), shown at three redshifts. The model assumes a single burst of star formation 11 Gyr ago and solar metallicity; the effect of evolution is not shown for simplicity. Also over plotted are the response functions (including atmospheric absorption) for the five SDSS filters. The break in the spectrum at 4000 Å, and its migration through the SDSS filters is clearly seen. Credit:(Padmanabhan et al. 2007)

near-IR spectra of faint objects at these highest redshifts.

There are two steps in the identification of galaxies at high$-z$: first photometric selection, followed by spectroscopic measurement to confirm the candidate. At $z > 4.5$ most galaxies used in the literature to compute luminosity functions are photometric candidates without spectroscopic confirmation. There are only a few spectroscopic surveys at $z > 4.5$. The VIMOS Ultra Deep Survey (Le Fèvre et al. 2014), targets galaxies at $2 < z < 6$ for spectroscopy down to a given optical magnitude ($i' < 25$). Compared

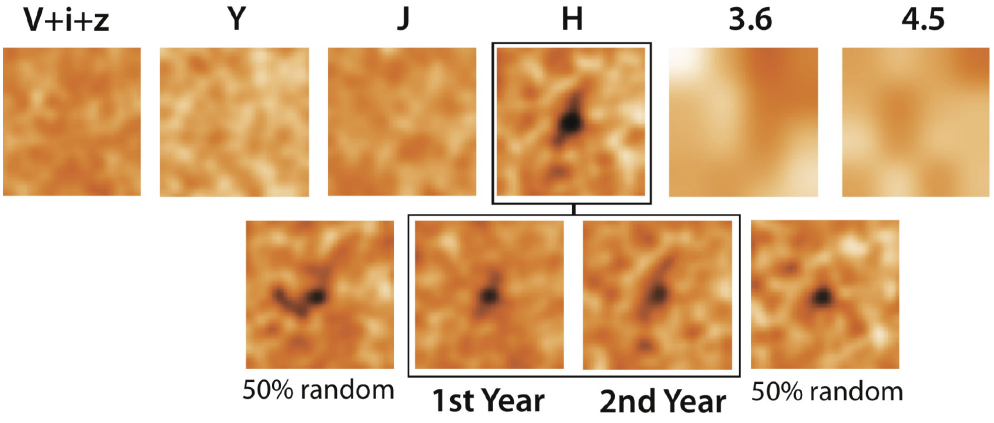


FIGURE 3.5: Optical and near−infrared images of the candidate $z \sim 10$ galaxy, UDFj−39546284, from the HUDF. Top row: the leftmost panel shows the HUDF ACS ($V_{606}, i_{775}, z_{850}$) data; the next three panels show the similarly deep HUDF09, near−infrared WFC3/IR ($Y_{105}, J_{125}, H_{160}$) data (reaching to 5σ depths of 29.8 AB mag); and the last two panels show the longer wavelength Spitzer IRAC 3.6 and 4.5 μm observations. Bottom row: the two middle panels show images of the $z \sim 10$ candidate in the first and second year of H_{160}−band observations (each representing 50% of the total); the two outer panels show two random 50% subsets of the data. Credit: (Bouwens et al. 2011).

to the spectroscopic survey like VUDS, there are also some high redshift surveys conducted by Keck-DEIMOS or VLT-FORS2. Stark et al. 2013 used KECK spectroscopy obtaining $\sim$ 100 spectroscopic confirmed faint lyman break galaxies, Ouchi et al .2010 also used the same instrument to target 200 $z \sim 7$ Lymanα emitter candidates and confirmed 24 of them spectroscopically. Vanzella et al . 2014 obtained $\sim$120 LBGs with spectro$-z > 4$ with the FORS2 spectrograph at VLT.

As a matter of fact, photometric redshifts have played an increasingly common role in describing the properties of distant galaxies. At $z > 4.5$, the majority of photometric surveys are UV restframe photometric surveys, mainly obtained from HST data. HST is equipped with the very sensitive ACS camera at optical wavelengths and with the third version of the Wide Field Camera (WFC3) with excellent near-IR sensitivity. Deep images of HUDF and other fields open an unprecedented window into the distant Universe (see Figure 3.6). HST provides photometric samples of several thousand Lyman break candidates at $z \sim 4$ and of order 1,000 at $z \sim 5$ and $z \sim 6$, reaching significantly fainter magnitude that allow astronomers characterise the luminosity function. The HUDF (Beckwith et al. 2006) observed a single ACS pointing ($\sim$11 arcmin^2) located within the GOODS South region with very long exposure times and reaching fainter than the original HDF with better sensitivity at higher redshifts. Both GOODS and the HUDF have been repeatedly revisited with new observations from HST over the years to add

deeper optical imaging as well as NIR data, first from Near Infrared Camera and Multi-Object Spectrometer (NICMOS) (Thompson et al. 2006, Conselice et al. 2011) and later with WFC3 in the HUDF09 and HUDF12 programs (Bouwens et al. 2011b, Ellis et al. 2013) and the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS) of several deep survey fields including GOODS and COSMOS (Grogin et al. 2011, Koekemoer et al. 2011). These IR observations make Lyman break selection at $z \sim 6$ far more robust and extend the method out to $z \sim 8$.HST successfully constructs the largest photometric samples in numbers at $z > 4.5$ but the area surveyed from HST surveys is small (~ 600 arcmin^2 for CANDELS the largest near-IR survey with WFC3) and one has to worry about cosmic variance. To be statistically meaningful, measurements of the current LD and SFRD require surveys covering a larger fraction of the sky.

Here we also need to briefly introduce the COSMOS survey and UltraVista surveys which play the crucial role for our study. The COSMOS survey investigates the evolution of galaxies and dark matter with large-scale structure (LSS) over the redshift range $z > 0.5 - 6$. The survey includes multi wavelength imaging and spectroscopy from X-ray to radio wavelengths covering a 2 deg^2 field of view, including HST imaging which is the largest filed covered with HST ever. COSMOS also provides deep optical imaging obtained with the prime-focus camera (Suprime-Cam) on the 8.2 m Subaru Telescope. These observations produce a reference catalogue of about one million galaxies at various redshifts, and the wavelength coverage of the photometry allows to study the stellar content and estimate photometric redshifts with reasonable accuracy. Besides, in order to extend our study over $z > 4.5$ near-IR observations is the key element that the UltraVista can provide us. The UltraVista survey has the largest field covered (~ 1.5 deg^2) with four near-IR broad bands (Y, J, H, K_s) and one narrow band (NB118) at extreme depth. To maximise the value of the new deep near-infrared data, the UltraVISTA survey is centred on the COSMOS field. This feature is essential to us to obtain better photometric redshift measurement by constructing a multiwavelength catalogue combining UltraVista with the Subaru optical photometry. All technical details of these two surveys are described in chapter 4.

3.3 Measuring the Luminosity Function

The luminosity function (LF) offers constraints on the overall abundance of objects per volume observed at a given wavelength. The LF is defined as the number of stars or galaxies per luminosity interval and per unit volume. LFs are used to provide a statistical description of a given population. As galaxies of different luminosities are formed and

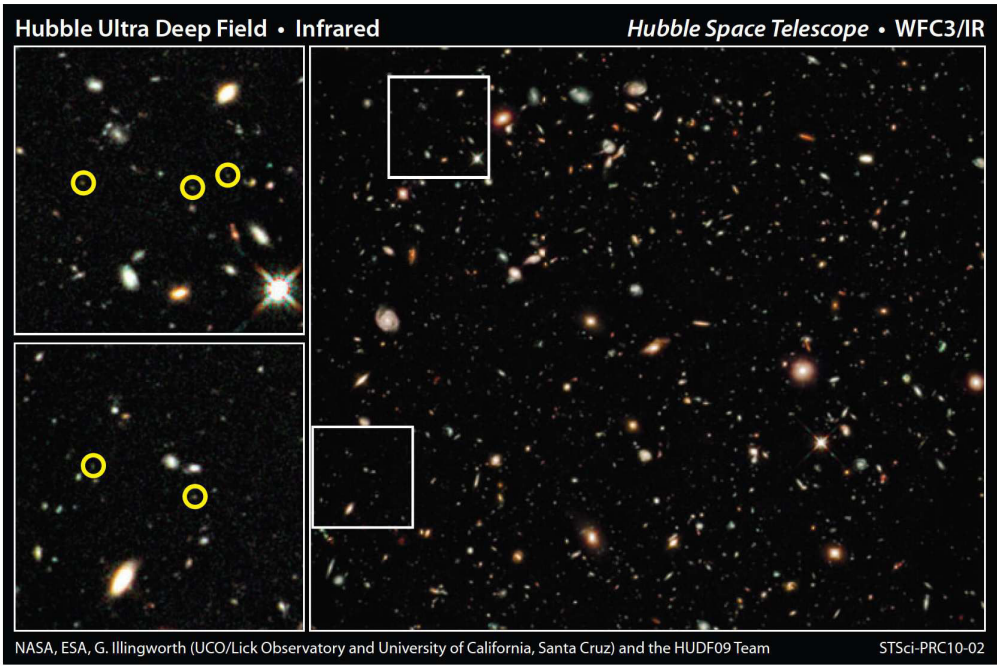


FIGURE 3.6: Early galaxies in the Hubble Space Telescopes deepest view of the Universe. The image was taken with the newly installed WFC3/IR camera; the positions of newly discovered galaxies at $z \sim 7$ are indicated by the circles in the zooms on the lefthand side. Figure courtesy of NASA.

TABLE 3.1: Deep spectroscopic redshift surveys in past 20 years (Not exhaustive).

Suvey	Instrument	Redshift range	Galaxies
2dFGRS	2dF/AAT	$0 < z < 0.5$	~220,000
SDSS	SDSS/Apache Point	$0 < z < 0.5$	1,800,000
CFRS	CFHT-MOS	$0 < z < 1.2$	600
LBG	KECK-LRIS	$2.5 < z < 4.5$	1000
DEEP2	KECK-DEIMOS	$0.7 < z < 1.4$	~50,000
VVDS	VLT-VIMOS	$0 < z < 5$	~50,000
VIPERS	VLT-VIMOS	$0.5 < z < 1.2$	100,000
zCOSMOS	VLT-VIMOS	$0 < z < 3$	~30,000
VUDS	VLT-VIMOS	$2 < z < 6$	~10,000
Vanzella et al . 2014	VLT-FORS2	$4 < z < 7$	~ 120

TABLE 3.2: Deep photometric redshift surveys in past 20 years (Not exhaustive).

Suvey	Instrument	Area
Ukidss-UDS	UKIRT	0.77deg^2
SXDS	Subaru	1.3deg^2
CFHTLS	CFHT	4deg^2
UltraVISTA	VISTA	0.73deg^2
VIDEO	VISTA	12deg^2
HUDF	HST	$5.76\ \text{arcmin}^2$
XDF	HST	$4.6\ \text{arcmin}^2$
CANDELS(wide)	Multiwavelength	720arcmin^2
CANDELS(deep)	Multiwavelength	130arcmin^2

evolve in dark matter halos the LF constantly evolves at all luminosities, and comparing LFs at different redshifts therefore allows to trace the evolution along time of galaxies. The shape of the LF can tell us about the evolution of galaxies, for instance we can see what is the effect of AGNs or supernovae on the bright-end slope and faint-end slope of UV LF, as both AGNs and supernovae quench the star formation. For the analysis of the evolution of LFs, different methods can be used to fit and describe the properties of galaxies. The Schechter function (Schechter 1976) is the most common form used to characterise the LF. In this parametrisation, the LF is described by the Schechter parameters α, M^* and ϕ^*. The M^* and ϕ^* are the characteristic magnitude and number density of a galaxy population at the knee of the function, and α is the faint-end slope of the function (as shown in Figure 3.7). The Schechter function is written as follows :

$$\phi(M)dM = 0.4ln(10)\phi^* 10^{0.4(M^*-M)(\alpha+1)} exp[-10^{0.4(M^*-M)}]dM \tag{3.2}$$

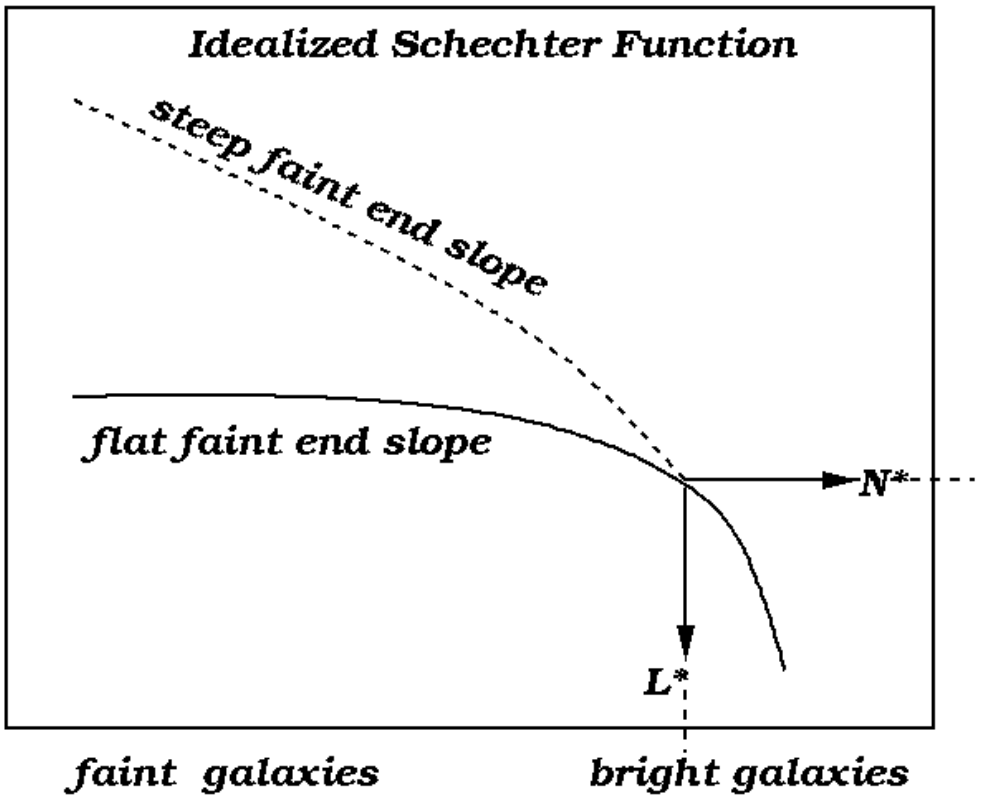


FIGURE 3.7: Schematic representation of the Schechter formulation that describes the galaxy luminosity function. This function is defined by a characteristic luminosity, L^*, and normalised space density at that luminosity, N^* (more generally called $\Phi_\star$), along with the slope of the faint end. The solid line indicates a relatively flat faint end slope, that represents what is usually obtained in samples that are not corrected for surface brightness selection effects. The dashed line shows the case of a steep faint end slope in which most of the galaxies in the Universe are faint/low mass objects. Credit: (http://zebu.uoregon.edu/tandf/timages/toc6.html).

To compute the LF requires obviously the knowledge of the number of galaxies per luminosity bin, but also of the cosmological volume in which this population is accounted for. In addition, it is important to describe the absolute luminosity output of galaxies, and therefore the knowledge of absolute magnitudes is also necessary. A standard way to compute LFs is using the the V_{max} method (Schmidt 1968) to account for the systematic scattering of galaxies in parameter space. The method does not require any assumptions on the shape of the LF. Owing to its simplicity, this method is the most used in high redshift surveys. The following equations describe the algorithm of the V_{max} method. The maximum observable comoving volume $V_{obs,i}$ in which galaxy i can be detected, is given by

$$V_{obs,i} = \int_w \int_{z_{min,i}}^{z_{max,i}} \frac{d^2V}{dwdz} dwdz \tag{3.3}$$

The number densities ϕ are derived in each absolute magnitude bin k as follows:

$$\phi_k^{Ref} dM = \frac{1}{V_{total}} \sum_{i=1}^{N_g} \frac{V_{total}}{V_{obs,i}} W(M_k^{Ref} - M_i^{Ref}) \tag{3.4}$$

where the window function W is defined as,

$$W(M_k^{Ref} - M_i^{Ref}) = \begin{cases} 1 & \text{if } -dM/2 \leq M_k^{Ref} - M \leq dM/2 \\ 0 & \text{otherwise,} \end{cases} \tag{3.5}$$

where V_{total} is the comoving volume between z_{low} and z_{high}. This method is able to recover the total number of galaxies per unit volume from the galaxy distribution in the redshift bin. We apply the Poisson counting statistics to estimate the errors of the LF (Marshall 1985), as follows:

$$\sigma_\phi = \sqrt{\sum_{i=1}^{N_g} \frac{1}{(V_{obs,i})^2} W(M_k^{Ref} - M_i^{Ref})} \tag{3.6}$$

While the V_{max} method is a simple approach in computing the LF well adapted to very high redshift samples, we have to keep in mind that according to Ilbert et al. (2004), there might be some biases induced using this method for UV LF. Other estimators are used to compute LFs, like the C+ and STY methods, as described in Ilbert et al. (2004). Finally it must be remembered that all luminosity functions are affected to some extent by cosmic variance (Moster et al. 2011), and extremely high-redshift

surveys need to cover sufficient area (helped by covering independent lines of sight) to offer a representative picture of the galaxy population at each epoch.

3.4 Measuring Luminosity Density and Star Formation Rate Density

As we mentioned in chapter 2, the star formation history (SFH) is a crucial element to understand galaxy evolution (see the review by Madau and Dickinson 2014). Many physical processes affect star formation activity, such as the gas accretion, mergers and AGN feedback during galaxy evolution. Although the luminosities and colours of galaxies represent basic and fundamental observables, we seek to translate these measurements into physical quantities such as the star formation rate density (SFRD). Before measuring the SFRD, we should start from the luminosity density (LD) which is usually used to compute the SFRD. The mean comoving LD in each redshift bin is derived as follows:

$$LD = \int_{faint}^{bright} \phi(L) dL \tag{3.7}$$

where $\phi(L)$ is the luminosity function assuming a Schechter (1976) functional form as described in the previous section. The integration of the LF is intuitive, but in practice the effect of different limits in the integral need to be considered. Given the typical shape (shown in Figure 3.8) of the LF for bright galaxies, setting L_{bright} as above or $L_{bright} = \infty$ does not make a big difference. The problematic part is how to set α (the faint-end slope) which will alter LD calculations dramatically, because faint galaxies are more populated than bright galaxies and they are difficult to be detected. This uncertainty will propagate to the next step when people attempt to transform LD to SFRD. Figure 2.7 and Figure 3.9 clearly show the that current estimates of the SFRD with $z > 4.5$ still suffer from shortage of data and associated significant uncertainty. In this work we aim to add new reliable data to better measure the SFRD at the highest redshifts observable from ground-baed infrared telescopes like VISTA.

In the second part of this section, we summarise several simple SFR tracers from specific luminosities, as well as the techniques used to model the stellar populations of high-redshift galaxies based on multiwavelength SEDs. Different SFR indicators are measured at different wavelengths with different methods. These include: LD estimated from the infrared (IR) produced by the re-radiated of ultraviolet (UV) photons by dust (e.g., Caputi et al. 2007; Pérez-González et al. 2005; Reddy et al. 2008); nebular emission lines such as Hα, Hβ and [OII], produced by ionised gas surrounding hot young

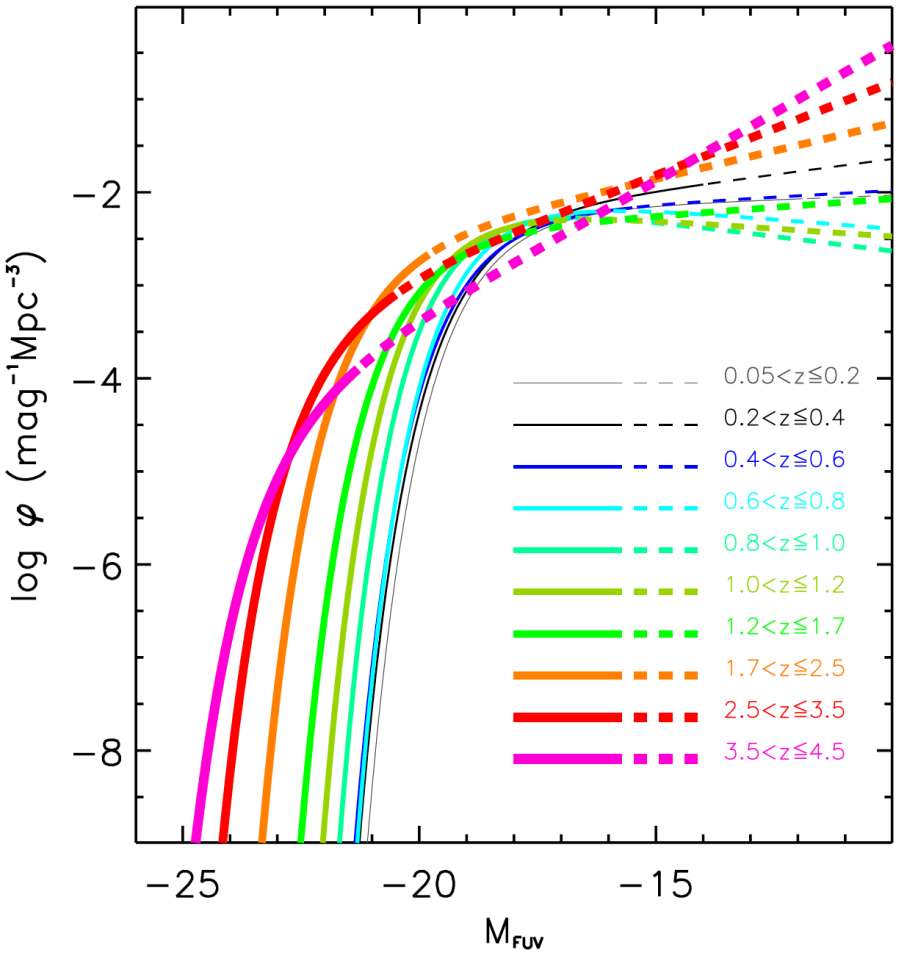


FIGURE 3.8: The VVDS rest-frame FUV-band luminosity functions from $z \sim 0.1$ up to $z \sim 4$ (from the thinnest to the thickest curve), fitted with a Schechter functional form parametrised with the Schechter parameters (α, ϕ^*, M_*). Lines are dashed for magnitudes fainter than the LF magnitude bias in each redshift bin Credit: (Cucciati et al. 2012).

OB stars (e.g., Tresse et al. 2002; Sawicki et al. 2005; Reddy et al. 2008); X-ray emission produced by X-ray binary systems in late-type galaxies (Norman et al. 2004; Lehmer et al. 2008); and SFR from radio emission produced by relativistic electrons in supernova remnants (Condon 1992; Seymour et al. 2008). But more directly, LD have been measured from the UV luminosities of galaxies produced from OB stars (Reddy et al. 2008; Oesch et al. 2010; Cucciati et al. 2012). Except potential contamination by other objects such as stars and AGNs, UV is one of the most straightforward tracer of SFR as purely relying on the luminosity of new stars. In practice, the main indicators of SFR, namely Hα, IR and UV each has their own drawbacks. They have been compared to each other by Hirashita et al. (2003): Hα is a very reliable indicator, if the Balmer decrement is measured precisely enough to correct for the extinction of Hα photons; the IR luminosity traces the SFR quite well but there is a risk that the SFR is

underestimated for SFR $< 1 M_{\odot} \mathrm{y} r^{-1}$ and UV luminosites should be corrected for dust extinction. Therefore, modern versions of SFR estimates rely both on UV and IR to explore the SFRH simultaneously. Figure 3.9 presents the newest SFRH results based on UV, IR and UV+IR measurements (Madau & Dickinson 2014).

In my work, I use UV luminosity to convert to SFR which is useful for high redshift galaxies where UV emission is redshifted into optical or Near-IR. To transform UV fluxes into star formation rates, I use the luminosity-SFR relation from Madau et al. (1998). It yields:

$$SFRD(z) = 1.4 \times 10^{-28} LD_{UV}(z) 10^{0.4 A_{UV}(z)} \tag{3.8}$$

where the SFRD is in $M_{\odot} yr^{-1} Mpc^{-3}$ units and the LD in ergs$^{-1} Hz^{-1} Mpc^{-3}$. This formula includes the dust attenuation A_{UV}(z), and assumes a Salpeter (1955) initial mass function (IMF) which includes stars from 0.1 to 125 solar masses. We primarily apply the recipe from Calzetti et al. (2000) to derive dust attenuations at UV band and A(λ) denoted as follows:

$$A(\lambda) = E(B-V) k(\lambda) \tag{3.9}$$

where $E(B-V)$ is the intrinsic colour excess of the stellar continuum of a galaxy, and k(λ) is the starburst reddening curve.

3.5 Summary

I presented the methods to select galaxies at the high redshifts ($z > 4.5$), focusing on the continuum properties producing the drop-out in the continuum that I will use to select my sample as described in Chapter 5, both from colour-colour and photometric selection. From the count of galaxies I explained how to derive the luminosity function, luminosity density and ultimately the evolution of the SFRD.

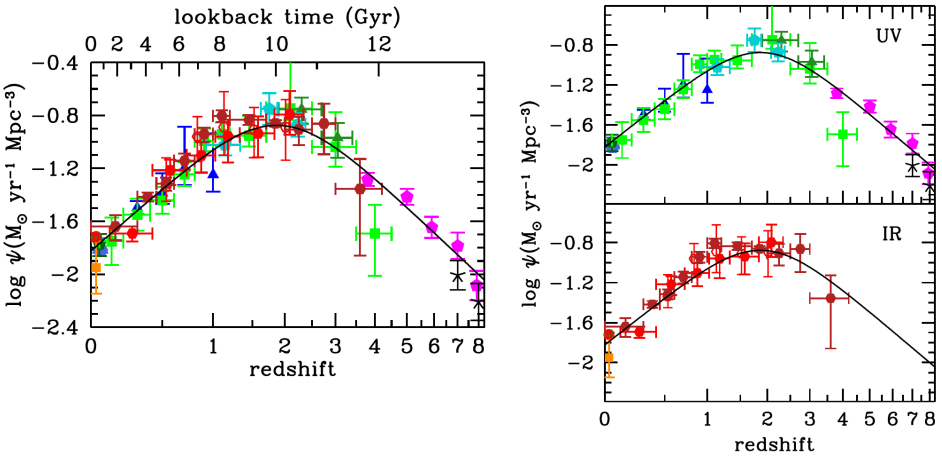


FIGURE 3.9: The history of cosmic star formation from (top right panel) FUV, (bottom right panel) IR and (left panel) FUV+IR rest-frame measurements. All data points are compiled from resent literature from UV and IR during 2003 to 2013. Lists of UV data: Wyder et al. 2005 (Blue-grey hexagon), Schiminovich et al. 2005 (Blue triangles), Robotham & Driver 2011 (Dark green pentagon), Cucciati et al. 2012 (Green squares), Dahlen et al. 2007 (Turquoise pentagons), Reddy & Steidel 2009 (Dark green triangles), Bouwens et al. 2012 (Magenta pentagons) and Schenker et al. 2013 (Black Cross). Lists of IR data: Sanders et al. 2003 (Brown circle), Takeuchi et al. 2003 (Dark orange square), Magnelli et al. 2011 (Red open hexagons), Magnelli et al. 2013 (Red filled hexgons) and Gruppioni et al. 2013 (Dark red filled hexagons). Credit: (Madau & Dickinson 2014).

Chapter 4

Photometric and Spectroscopic Data

In the field of astrophysics, all of the useful information obtained about the universe and the objects within it comes from gathering radiation and analysing it. Images provide us the crucial information about the structure of objects in space, but to get a more detailed view of what is going on in the Universe, light has to be studied in much more detail. This is why photometry and spectroscopy are the two most common techniques used by observers to study the Universe. Photometry is a way to measure how much light we receive from objects in space. It is a measure of the amount of light received in a specified wavelength range. Different filters are used to let different wavelengths of light through to the detectors, and measure the amount of light entering the detector in each wavelength range. Spectroscopy is the detailed study of an object by spectrometers which spread light out into its wavelengths, creating a spectrum. Within this spectrum, astronomers can study emission and absorption lines which are the fingerprints of atoms and molecules. Astronomers can learn a great deal about an object in space by studying its spectrum. By identifying the atomic and molecular fingerprints in a spectrum and the shape of its continuum, we can learn about an object's composition. The intensity and width of spectral lines tell us about an object's temperature and density. Spectra can provide a precise estimation of the redshift of objects, a very important property used throughout this work.

These two types of data associate different advantages and drawbacks. For our work, one can ask why we still need to utilise a photometric dataset if spectroscopic data can provide more accurate redshift measurements? Photometric redshifts are less accurate than spectroscopic redshifts, but spectroscopy is very time consuming and photometric

data can provide information on larger datasets. In a given amount of time we can get photometric redshifts for many galaxies, or an accurate spectroscopic redshift for only a few galaxies (for instance the difference is a factor 50 to 60 for luminous Red Galaxies in the Sloan Digital Sky Survey). Most importantly, for the same integration time, photometry is deeper than spectroscopy which disperses photons on the detector. Using photometry one can then get large samples at the most extreme depths reachable with a telescope. At redshifts larger than z∼5 existing samples of galaxies are predominantly based on photometry. At these redshifts, obtaining spectra of extremely faint sources is getting more problematic and only few galaxies have spectroscopic redshifts confirmed over redshift 5. But there is also an inevitable disadvantage of photometric data, as photometric redshifts from SED fitting depend on the accuracy of the photometry as well as on the reference SED models used in the fitting. Fortunately, very deep multi-wavelength photometric datasets are now available covering from the optical to the near-infrared and the far infrared, and allow to identify robust galaxy samples at $z > 5$. Very deep photometry in the YJH and K_s bands is now available in the COSMOS field from the UltraVista survey, adding an exquisite dataset to the available optical and Spitzer data. A key element is the large field of view of ∼1.5 deg^2 covered by these observations compared to much smaller, albeit deeper, HST-WFC3 data. This allows to probe the bright end of the Luminosity Function much more robustly. Moreover, the VUDS spectroscopic survey provides a sample of 10,000 galaxies with spectroscopic redshifts up to $z \sim 6$, which can be used to calibrate the selection of the photometric redshift candidates at $z > 5$. I specify the details of the UltraVista andVUDS samples in the following sections.

4.1 The COSMOS Field

All the data used in this thesis are located in the Cosmic Evolution Survey field (COSMOS, P.I. Nicholas Scoville). The COSMOS survey started as an HST (Hubble Space Telescope) Treasury Program covering ∼ 2 square degree (1.4° × 1.4°) of an equatorial field, centred on R.A.(J2000)=10:00:28.6 and DEC.(J2000)=+02:12:21.0 with the Advanced Camera for Surveys (ACS) in the I band (shown in Figure 4.1). Parallel observations with the Wide Field/Planetary Camera (WFPC) and Near Infrared Camera and Multi-Object Spectrometer (NICMOS), which are installed on HST, were also obtained. This survey has also assembled several public ancillary datasets including optical spectra, deep X-ray imaging by the X-ray Multi-Mirror Mission (XMM) and radio imaging by the Very Large Array (VLA), ground-based optical/IR imaging by the Subaru telescope, ultraviolet (UV) imaging from the Galaxy Evolution Explorer (GALEX) and mid infrared data from the Spitzer telescope. The COSMOS field is also the target of

numerous spectroscopic surveys, including zCOSMOS (Lilly et al. 2007), Keck surveys (Capak, Mobasher et al. 2011 and Ouchi, Mobasher et al. 2009), and most importantly for this work the VIMOS Ultra Deep Survey (VUDS, Le Fèvre et al. 2014), which together provide several tens of thousand of spectra and redshifts.

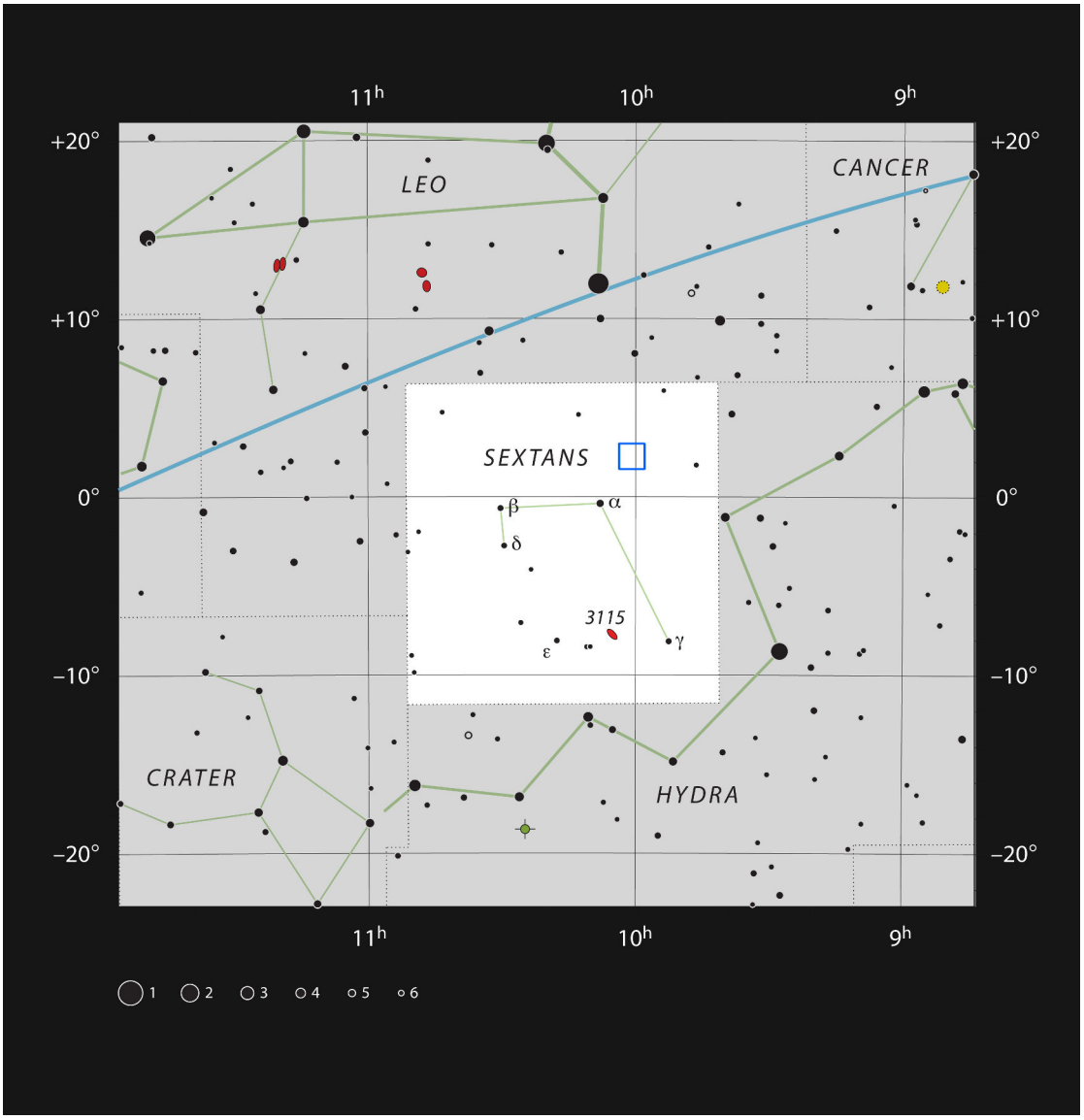


FIGURE 4.1: This chart shows the location of the COSMOS field in the constellation of Sextans (The Sextant). This map shows most of the stars visible to the unaided eye under good conditions, and the COSMOS field is marked as a blue square. Through a small telescope nothing can be seen here apart from a few faint stars, but this small patch of sky has been studied in great detail by telescopes on the ground and in space (Credit by ESO, IAU and *Sky & Telescope*).

4.2 Multi-band Photometry in COSMOS

The COSMOS data were collected on a variety of telescopes and instruments, but only the most important and deepest optical and mid-infrared data are used in this paper. The location of UltraVista is superimposed on the part of Cosmological Evolution Survey (COSMOS) field and covers 1.5 deg^2. All the optical data for our analysis is from Subaru Suprime-Cam. It also provides deep $z'-$band data which is a crucial band (shown in Figure 4.2) to extend the selection of high redshift galaxies to $z > 5.5$. In addition, data from mid-infrared plays an important role to help separate contaminating lower redshift red objects and galactic stars from the high redshift photometric sample of interest for our study and collect more physical information based on SED fitting. Mid-infrared imaging within the COSMOS field was obtained with the *Spitzer*/infrared array camera (IRAC) including the original survey from Sanders et al. (2007) followed by the recently completed the *Spitzer* Large Area Survey with Hyper-Suprime-Cam (SPLASH) survey.

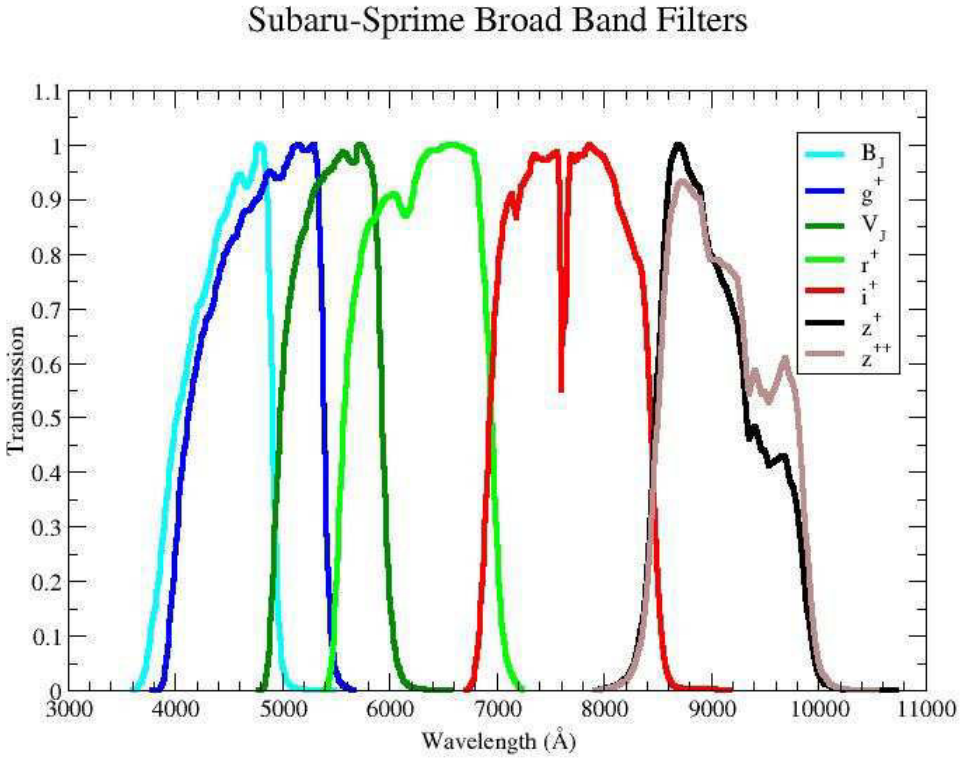


FIGURE 4.2: The broad band filter sets of Suprime-Cam deployed on 8.3m Subaru telescope and their normalised filter response curves including the atmospheric transmission. These values are estimated at airmass of 1.2. (Credit: http://www.astro.caltech.edu/ capak/filters/index.html)

TABLE 4.1: Coordinates of four ultra-deep (DR2) stripes.

Stripe ID	Coordinates(R.A.(J2000),Dec(J2000))
stripe1	149.3300, 2.76, 149.4600, 2.76, 149.4600, 1.66, 149.3300, 1.66
stripe2	149.7000, 2.76, 149.8300, 2.76, 149.8300, 1.66, 149.7000, 1.66
stripe3	150.0600, 2.76, 150.2000, 2.76, 150.2000, 1.66, 150.0600, 1.66
stripe4	150.4300, 2.76, 150.5700, 2.76 ,150.5700, 1.66, 150.4300, 1.66

4.3 Ultra Vista Ultra Deep Survey

The VISTA (Visible and Infrared Survey Telescope for Astronomy) is a reflecting telescope with a 4.1 metre mirror, located at the Paranal Observatory in Chile with only one instrument: VIRCAM (Vista InfraRed CAMera) deployed. VISTA is the largest survey telescope in the world dedicated to surveying the sky at near-infrared wavelengths. Observing at wavelengths longer than optical allows VISTA to study objects that may be almost impossible to see in visible light because they are cool, obscured by dust clouds or because their light has been redshifted tremendously by the cosmological expansion, especially when these objects are in the very early Universe. There are six public surveys conducted by VISTA in the first five years of operations. They are VIKING (VISTA Kilo-Degree Infrared Galaxy Survey), VMC (VISTA Magellanic Survey), VVV (VISTA Variables in the Via Lactea), VHS (VISTA Hemisphere Survey), VIDEO (VISTA Deep Extragalactic Observations Survey) and the UltraVista, the deepest survey carried out with VISTA telescope which is a key element of our study.. These surveys cover different areas of the sky (shown in Figure 4.3) to different image depths to deal with different astronomical topics.

UltraVista is allocated a total of 1800h of observations. The main goals of UltraVista include understanding mass assembly at very high redshift, searching for the first galaxies and dust-obscured star formation. There are two broad band photometric surveys in the Y, J, H, and K_s bands (shown in Figure 4.4), with a deep survey covering the full 1.5 deg^2 field and an ultra-deep survey covering 0.73 deg^2 into 4 strips. There is also one narrow band (NB118) survey for searching Lα emitters around redshift 8.8. The UltraVISTA field coverage is shown in Figure 4.5, including the ultra-deep strips in green, and taking into account the 1 arcmin trimming due to incomplete coverage from dithering. The coordinates of the corners of the ultra-deep field (green) are listed in Table 4.1.

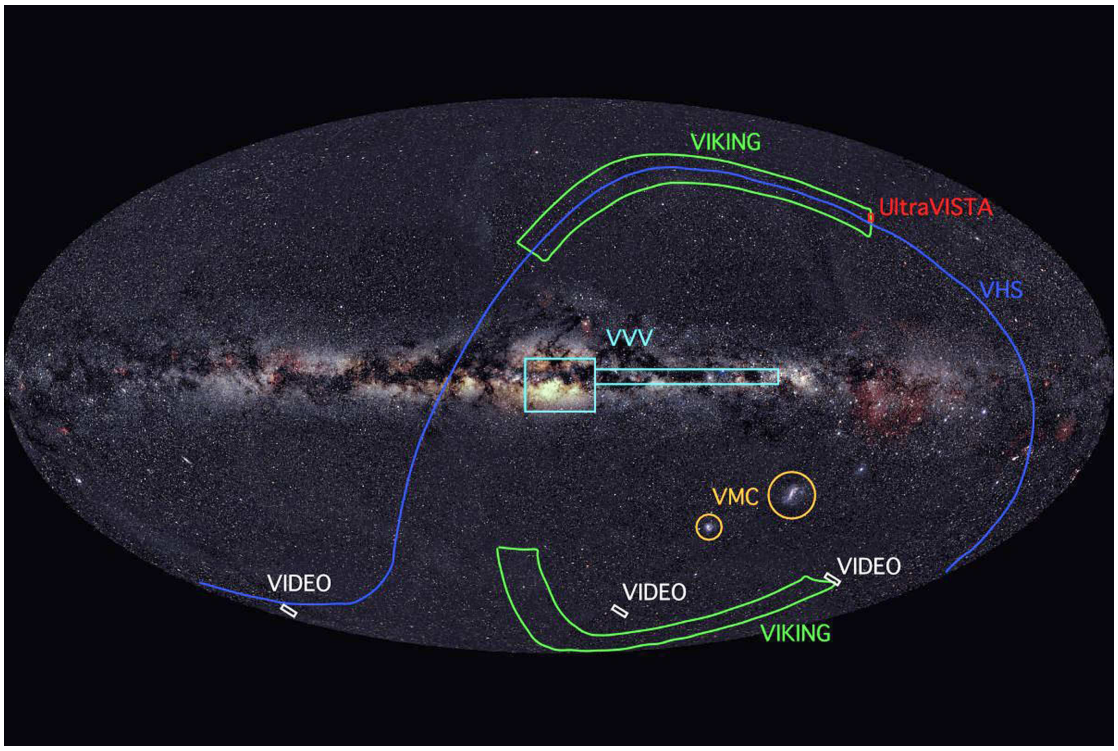


FIGURE 4.3: A schematic representation of the sky coverage of the various VISTA surveys. The image is a projection of the entire sky with the Milky Way across the centre. (Credit: VISTA ESO)

4.3.1 Photometric Depth

The determination of imaging depths across all bands is crucial to obtain accurate errors for use in the SED fitting analysis, and also in making appropriate colour cuts in the selection process. However, defining global limiting depths for data over degree-scale fields in the optical and near-infrared becomes problematic, as the combined pointings may have different integration times and seeing. Therefore, we obtained local depths over each image from the clipped median absolute deviation of the 200 closest apertures to each point. Apertures were placed randomly on the blank regions of the images that had been background subtracted using SExtractor, where blank regions were defined using the segmentation map of each image. The medians of the local depths for the imaging utilised here are presented in Table 4.2, within the 2.0 arcsec diameter circular aperture used for the photometry in our work.

4.3.2 Photometric Catalogue

I created the primary catalogues for the UltraVISTA fields using SExtractor v2.8.6 with a detection threshold of 2.2σ and a minimum object size of 5 pixels in dual-image mode to create multi-wavelength catalogues. The UltraVISTA catalogue was selected in

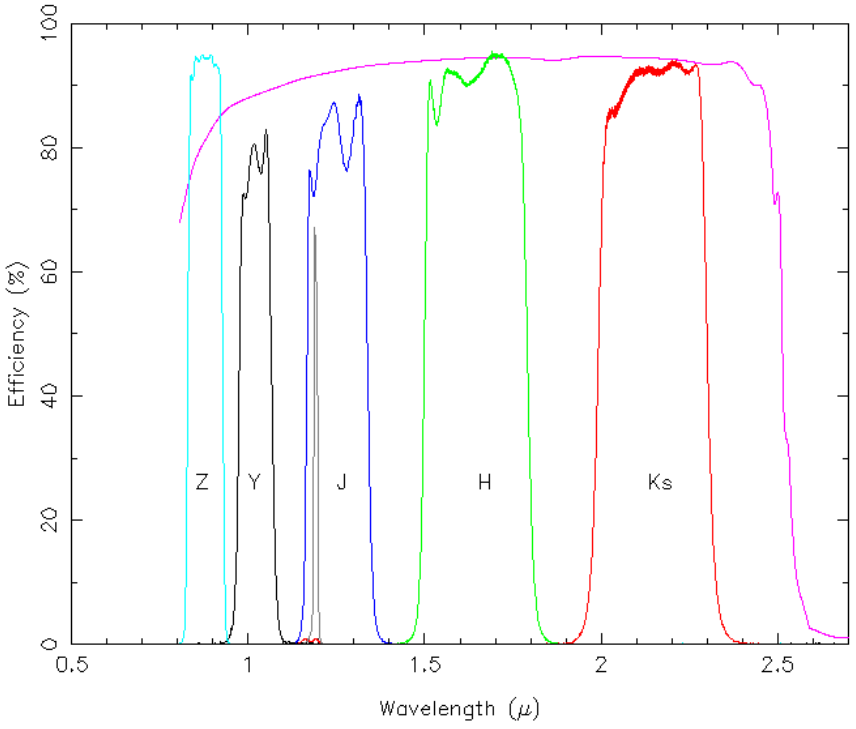


FIGURE 4.4: The filter set of VIRCAM deployed on 4.1m VISTA telescope and their filter transmission and detector Quantum efficiency (QE). A sample detector QE curve is shown superposed in purple. The narrow grey curve in the J-band is the NB118 narrow band filter.(Credit: http://casu.ast.cam.ac.uk/surveys-projects/vista/technical/filter-set)

TABLE 4.2: The 5σ limiting magnitudes for the relevant optical and near-infrared data used in this study.

COSMOS/UltraVISTA Filter	Deep (5σ)	Ultra Deep (5σ)	Source
B	28.2	28.2	Subaru/SuprimeCam
V	27.5	27.5	Subaru/SuprimeCam
r	27.7	27.7	Subaru/SuprimeCam
i	27.1	27.1	Subaru/SuprimeCam
z'	26.1	26.1	Subaru/SuprimeCam
Y	24.6	25.7	UltraVISTA
J	24.4	25.1	UltraVISTA
H	23.9	24.7	UltraVISTA
K_s	23.7	24.8	UltraVISTA
3.6 μm	25.3	25.3	Spitzer/SPLASH
4.5 μm	25.1	25.1	Spitzer/SPLASH

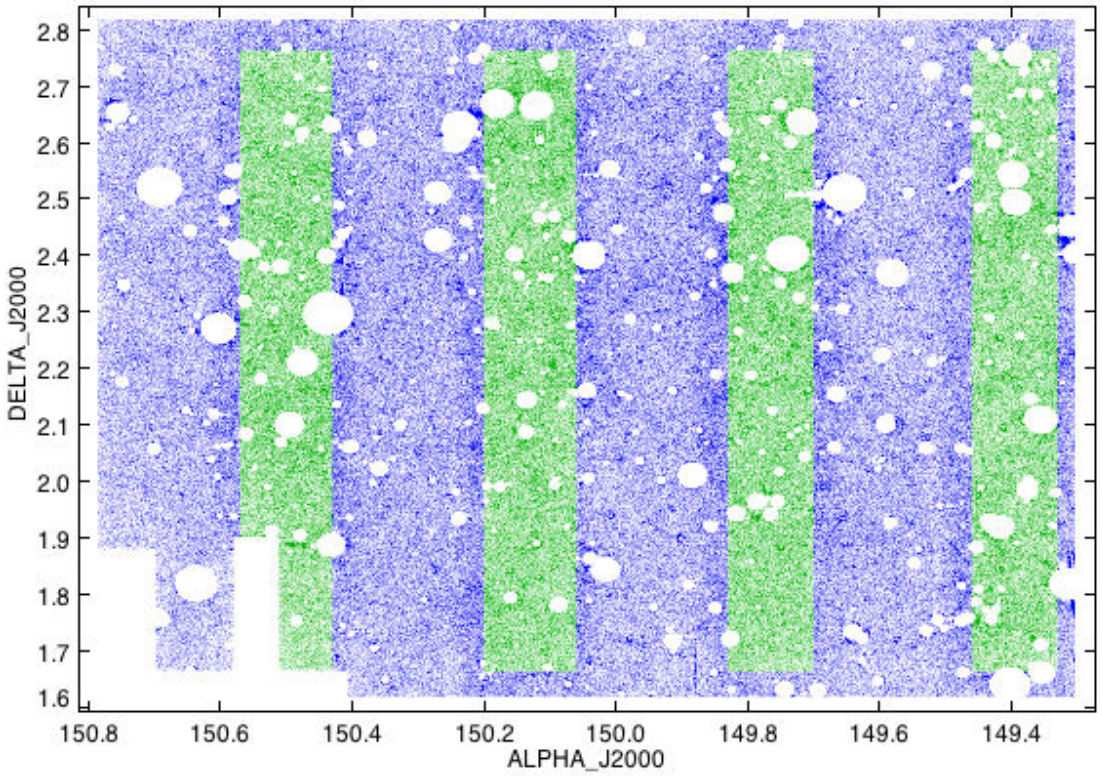


FIGURE 4.5: The footprint of the UltraVista field showing the ultra-deep data as green regions and the deep data as the blue region. All the blank regions are areas where bright stars, spikes, and bad pixels are masked.

$Y+J+H+K_s$ 4 images stacks which is coadded using the swarp software (Bertin et al. 2002) as χ^2 weighted to ensure I remain sensitive to faint object detection in any band for redshifts reaching over 7. The initial catalogue for the UltraVISTA dataset contains 586,382 objects, 229,029 objects within the ultra-deep part and the rest 357,353 objects locate in the deep stripes.

4.3.3 Photometric Counts

Figure 4.6 shows the K_s-band number counts extracted from the UltraVista catalogues that I obtained from the data release 2 (DR2) in comparison with the first data release of UltraVista (McCracken et al. 2012). Our photometric counts in two regions agree with the UltraVista Data Release 1 (DR1), and reach 1 mag deeper in the ultra-deep stripes. After this basic data assessment, we can further investigate the distribution of objects in optical and near infrared colour-colour space. The BzK diagram (Daddi et al. 2004) allows us to separate galaxies and stars cleanly as shown in Figure 4.7. This two dimensional density plot shows all objects detected to 5σ in K_s band aperture magnitude. The feature seen as a long ridge relatively blue in $(z' - K_s)$ shows the locus of galactic stars.

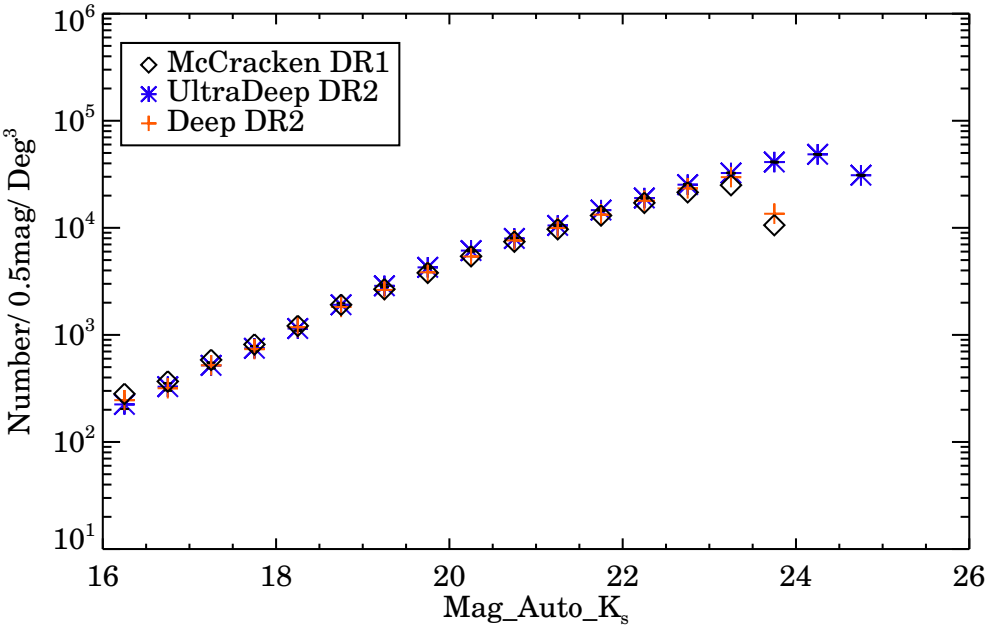


FIGURE 4.6: Galaxy number counts for our UltraVISTA DR2 catalog, in addition to previous data release (DR1) made by McCracken et al. (2012).

4.3.4 Stellar Objects and Spurious Sources

A sturdy star-galaxy classification is essential and it has important implications for our analysis because stars must be correctly identified in order to apply our automatic masking algorithm of diffraction spikes. Although SExtractor produces a continuous stellar classification index parameter ranging from 0 (extended sources) to 1 (point sources), the main drawback of this index is that it is not sufficient to identify faint stars from galaxies and miss a certain number of bright stars (shown in Figure 4.8). This figure also shows that star selection by CLASS_STAR > 0.8 for example, lost a number of bright stars ($K_s < 19$) and significantly mix fainter galaxies and stars at $K_s > 20.5$.

There is a substitute method to overcome this problem which is based on another SExtrator parameter MU_MAX (peak surface brightness above the background level). As illustrated in Figure 4.9, point sources occupy a well-defined locus in this plane. This alternative approach has the advantage of a tighter correlation of the stellar locus and a clear break indicating the magnitude at which the stars saturate up to $K_s \sim 21$. Using the MU_MAX parameter also can help us to pick out spurious detections where the objects are too small to be consistent with the point spread function (PSF). It means objects with very faint magnitude and high peak surface brightness are considered as

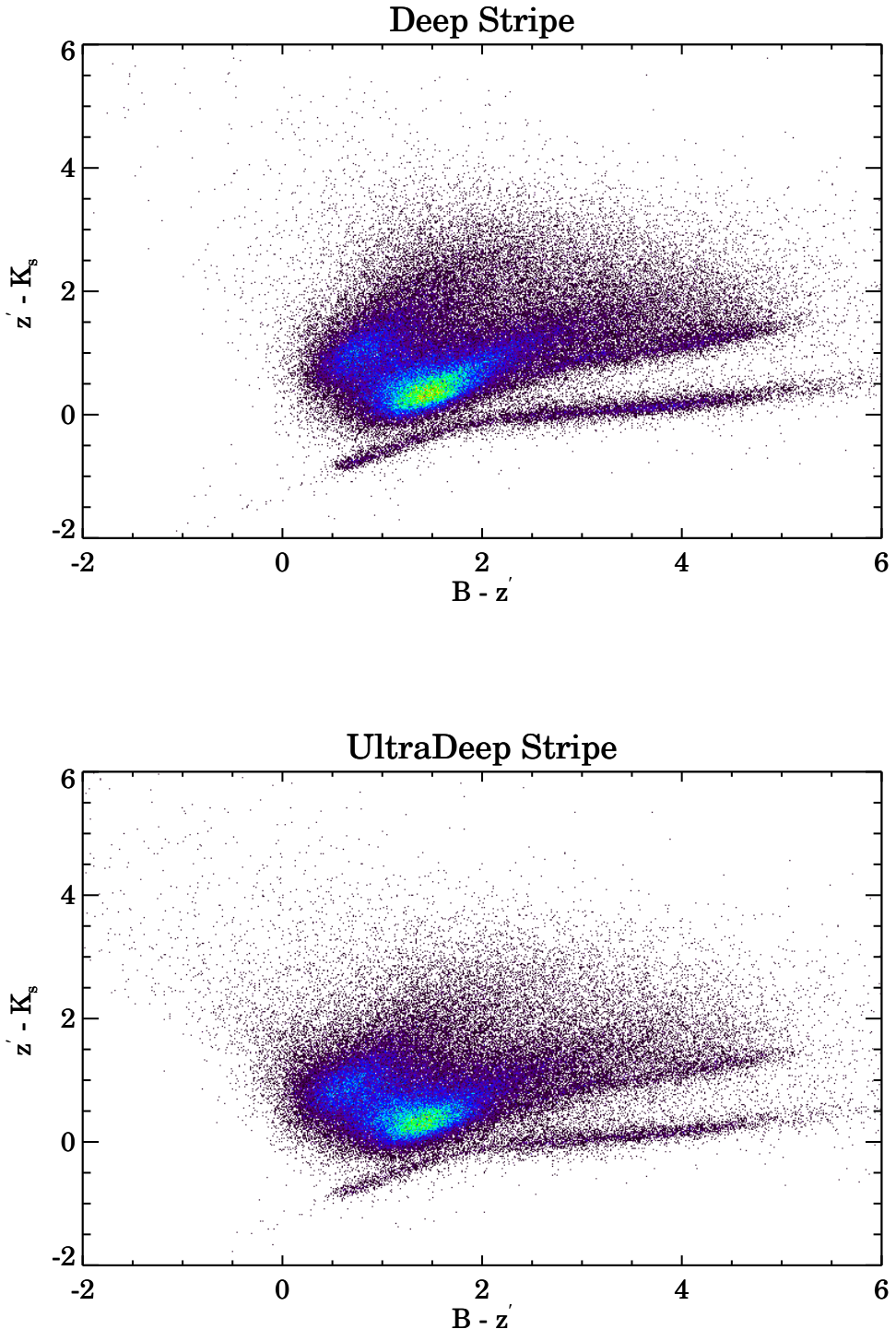


FIGURE 4.7: Comparing colour-colour plot of $(B - z')$ vs $(z' - K_s)$ corrected aperture colour for our UltraVISTA DR2 catalog. All sources detected to a 5σ limit in K_s auto magnitudes are shown. The continuous stripe on the lower part of the diagram represents the locus of galactic stars.

false detections which is obviously nonphysical. Normally these objects are mainly artifacts, hot pixels, and residual cosmic rays, and we remove them for further analysis. The criteria for our selection of bright stars and spurious objects are presented as :

- Bright Star : $-0.25 <$ Mu_MAX_$K_s - K_s$_MAG_Auto< 0.25 & $K_s < 20.5$
- Spurious Sources : Mu_MAX_$K_s < 26.2$ & K_s_MAG_Auto> 29.3

In conclusion, we have used photometric star indicators to identify brights stars that affect the sky background level and the areas that need masking around them and are not used for our analysis. ”

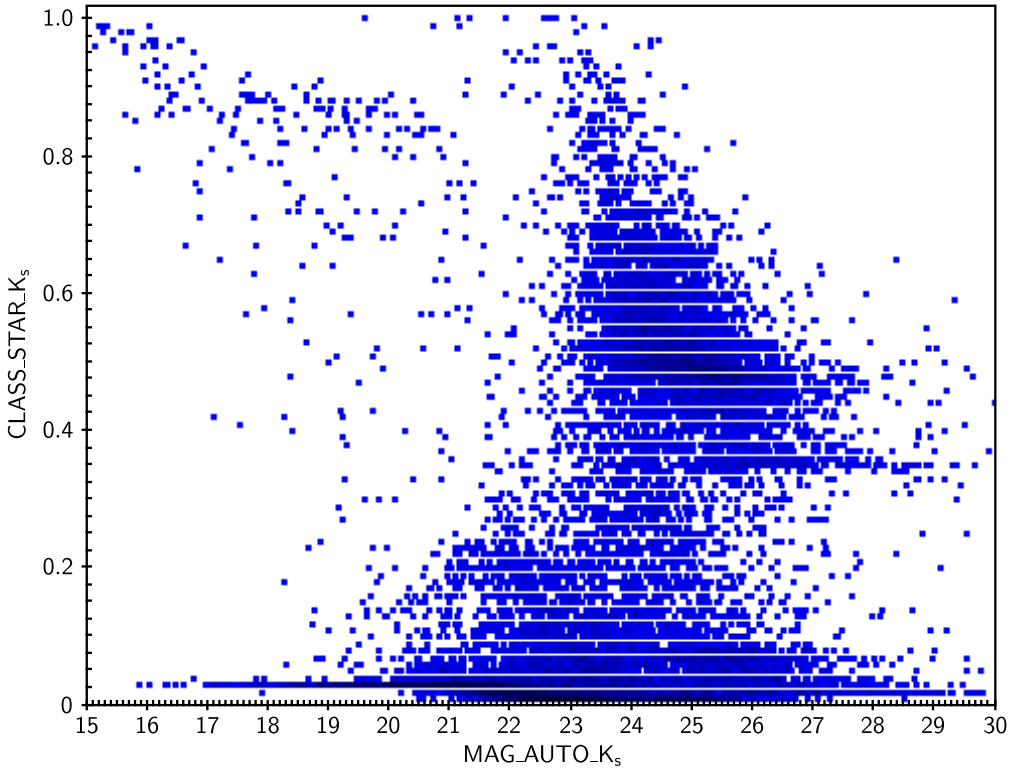


FIGURE 4.8: SExtractor stellar index (CLASS_STAR) for our point-source selection based on the peak surface brightness of objects

4.4 VIMOS Ultra Deep Survey

As I mentioned in the above sections, spectroscopic follow-up and redshift confirmation of photometric candidates are crucial to our high redshift studies. Here I am introducing the VIMOS Ultra Deep Survey (VUDS) which we use in this work in particular as a

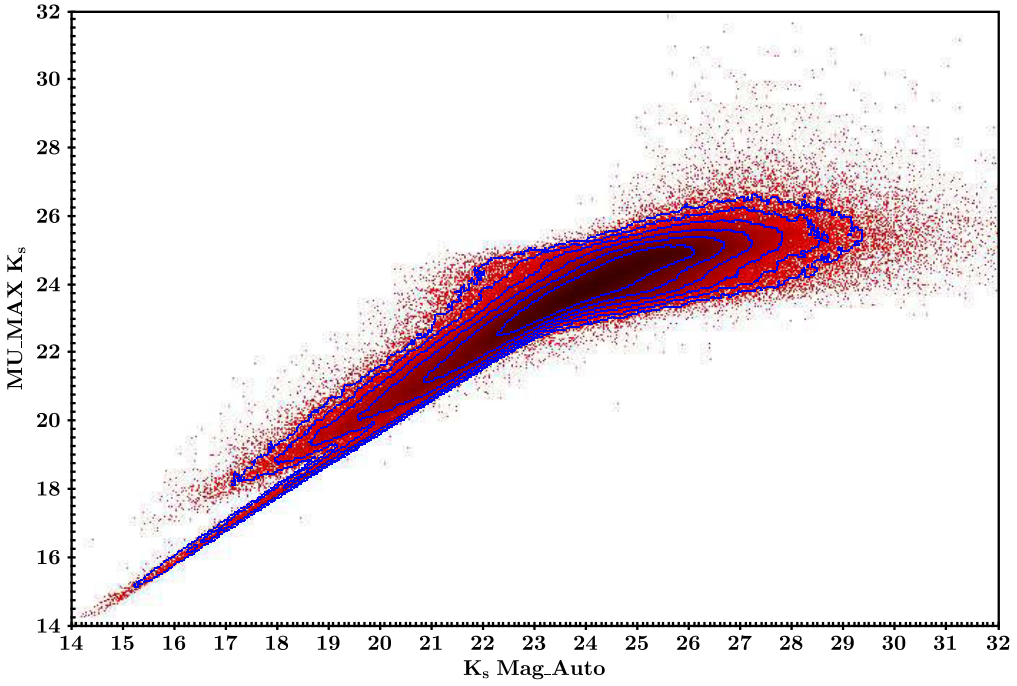


FIGURE 4.9: Classification of point sources, galaxies, and artifacts within the MU-MAX/MAGAUTO plane with blue contours showing the density of sources..

control sample for our photometric candidates selection. VUDS is a large spectroscopic survey of $\sim 10,000$ galaxies with the VIMOS multi-object spectrograph at the European Southern Observatory Very Large Telescope (Le Fèvre et al. 2003). The main goal of this survey is to study galaxy formation and evolution at $2 < z < 6$. Figure 4.10 shows some successful spectroscopic confirmed sample in $5 < z < 6$ which strongly relate to our study and Table 4.3 lists all high spectroscopic confirmed objects in all VUDS fields: the COSMOS fields, the VIMOS VLT Deep Survey 2 Hours (VVDS2H) field and the Extended Chandra Deep Field-South Survey (ECDFS) field. VUDS covers a total area of 1 deg^2 in three independent fields, which includes part of COSMOS field. This sample offers an important reference to test the robustness of our methodology for selecting high redshift candidates. Details about the survey strategy, target selection, as well as data processing and redshift measurements are presented in (Le Fèvre et al. 2014).

In this work I am using only the objects with spectra z probability more than 75% (flag =2, 3, 4, 9) based on the reliability flag (shown in Table 4.4). Additionally, the paper describing in details the VUDS survey is attached below. I participated to the VUDS survey in contributing photometric selection and photometric analysis of the spectroscopic sample. The level of my contribution entitles me to be listed as a VUDS "team member" in all VUDS publications with three orders of priority each listed by

TABLE 4.3: High redshift VUDS spectroscopic confirmed galaxies in 3 regions (COSMOS, VVDS2H, ECDFS).

Fields	Redshift Bin	Amount
VVDS2H	$5.5 > z > 4.5$	97
	$z > 5.5$	12
COSMOS	$5.5 > z > 4.5$	149
	$z > 5.5$	15
ECDFS	$5.5 > z > 4.5$	16
	$z > 5.5$	6

TABLE 4.4: Robustness of the measurement of VUDS Flags in the following conception.

Flag (VUDS)	Description
0	No redshift
1	$50 - 75\%$ probability
2	$75 - 85\%$ probability
3	$95 - 100\%$ probability
4	100% probability
9	$\sim 80\%$ probability and each spectrum has a single emission line

alphabetical order: "core members - builders of the survey", "team members", and "associate members"

4.4.1 Paper: The VIMOS Ultra-Deep Survey:$\sim$ 10 000 galaxies with spectroscopic redshifts to study galaxy assembly at early epochs $2 < z \simeq 6$

The VIMOS Ultra-Deep Survey: ~10 000 galaxies with spectroscopic redshifts to study galaxy assembly at early epochs $2 < z \simeq 6$ ⋆

O. Le Fèvre[1], L.A.M. Tasca[1], P. Cassata[1], B. Garilli[3], V. Le Brun[1], D. Maccagni[3], L. Pentericci[4], R. Thomas[1], E. Vanzella[2], G. Zamorani[2], E. Zucca[2], R. Amorin[4], S. Bardelli[2], P. Capak[12], L. Cassarà[3], M. Castellano[4], A. Cimatti[5], J.G. Cuby[1], O. Cucciati[5,2], S. de la Torre[1], A. Durkalec[1], A. Fontana[4], M. Giavalisco[13], A. Grazian[4], N. P. Hathi[1], O. Ilbert[1], B. C. Lemaux[1], C. Moreau[1], S. Paltani[9], B. Ribeiro[1], M. Salvato[14], D. Schaerer[10,8], M. Scodeggio[3], V. Sommariva[5,4], M. Talia[5], Y. Taniguchi[15], L. Tresse[1], D. Vergani[6,2], P.W. Wang[1], S. Charlot[7], T. Contini[8], S. Fotopoulou[9], C. López-Sanjuan[11], Y. Mellier[7], and N. Scoville[12]

[1] Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France
[2] INAF–Osservatorio Astronomico di Bologna, via Ranzani,1, I-40127, Bologna, Italy
[3] INAF–IASF, via Bassini 15, I-20133, Milano, Italy
[4] INAF–Osservatorio Astronomico di Roma, via di Frascati 33, I-00040, Monte Porzio Catone, Italy
[5] University of Bologna, Department of Physics and Astronomy (DIFA), V.le Berti Pichat, 6/2 - 40127, Bologna
[6] INAF–IASF Bologna, via Gobetti 101, I–40129, Bologna, Italy
[7] Institut d'Astrophysique de Paris, UMR7095 CNRS, Université Pierre et Marie Curie, 98 bis Boulevard Arago, 75014 Paris, France
[8] Institut de Recherche en Astrophysique et Planétologie - IRAP, CNRS, Universit de Toulouse, UPS-OMP, 14, avenue E. Belin, F31400 Toulouse, France
[9] Department of Astronomy, University of Geneva ch. d'cogia 16, CH-1290 Versoix
[10] Geneva Observatory, University of Geneva, ch. des Maillettes 51, CH-1290 Versoix, Switzerland
[11] Centro de Estudios de Física del Cosmos de Aragón, Teruel, Spain
[12] Department of Astronomy, California Institute of Technology, 1200 E. California Blvd., MC 249–17, Pasadena, CA 91125, USA
[13] Astronomy Department, University of Massachusetts, Amherst, MA 01003, USA
[14] Max-Planck-Institut für Extraterrestrische Physik, Postfach 1312, D-85741, Garching bei München, Germany
[15] Research Center for Space and Cosmic Evolution, Ehime University, Bunkyo-cho 2-5, Matsuyama 790-8577, Japan

e-mail: olivier.lefevre@lam.fr

Received...; accepted...

ABSTRACT

We present the VIMOS Ultra Deep Survey (VUDS), a spectroscopic redshift survey of ~10 000 very faint galaxies to study the major phase of galaxy assembly in $2 < z \simeq 6$. The survey covers 1 deg^2 in 3 separate fields: COSMOS, ECDFS and VVDS-02h, with the selection of targets based on an *inclusive* combination of photometric redshifts and colour properties. Spectra covering $3650 < \lambda < 9350$Å are obtained with VIMOS on the ESO-VLT with integration times of 14h. Here we present the survey strategy, the target selection, the data processing, as well as the redshift measurement process, emphasizing the specific methods adapted to this high redshift range. The spectra quality and redshift reliability are discussed, and we derive a success rate in redshift measurement of 91%, or 74% limiting the dataset to the most reliable measurements, down to a limiting magnitude $i_{AB} = 25$, and measurements are performed all the way down to $i_{AB} = 27$. The mean redshift of the main sample is z~3 and extends over a large redshift range mainly in $2 < z < 6$. At $3 < z < 5$, the galaxies cover a large range of luminosities $-23 < M_{NUV} < -20.5$, stellar mass $10^9 M_\odot < M_* < 10^{11} M_\odot$, and star formation rates 1 $M_\odot$/yr$< SFR < 10^3 M_\odot$/yr. We discuss the spectral properties of galaxies using individual as well as stacked spectra. The comparison between spectroscopic and photometric redshifts as well as colour selection demonstrate the effectiveness of our selection scheme. From about ~ 90% of the data analysed so far, we expect to assemble >6 000 galaxies with reliable spectroscopic redshifts in $2 < z < 6$ when complete. This makes VUDS the largest at these redshifts and offers the opportunity for unprecedented studies of the star-forming galaxy population and its distribution in large scale structures during the major phase of galaxy assembly. Staged releases of the data will appear on http://cesam.lam.fr/vuds/, starting in the summer of 2015.

Key words. Galaxies: evolution – Galaxies: formation – Galaxies: high redshift – Cosmology: observations – Cosmology: large-scale structure of Universe – Astronomical Databases: surveys

1. Introduction

The study of the first billion years of galaxy evolution is one of the key frontiers of modern cosmology. The current theoretical paradigm rests on the hierarchical build-up of dark matter halos in a ΛCDM cosmology (see e.g. Mo, van den Bosch & White 2010). Galaxies formed in these deep potential wells are expected to transform primordial gas into stars from the initial reservoir and to be fed from new accreted gas. As dark matter halos merge, galaxies in them are also expected to merge, these events deeply transforming the dynamics and overall star, gas, and dark matter content of the merging galaxies into the newly formed one. Complex processes are invoked to regulate the growth of galaxies, including supernovae or AGN feedback possibly quenching star formation, or the role of different environments impacting the way galaxies are nurtured. Increasingly sophisticated galaxy simulations in representative volumes (e.g. Springel et al. 2008) coupled to semi-analytic description of galaxy evolution (e.g. Guo et al. 2011) need better constraints from observations to be thoroughly tested.

Impressive progress has been made on the observational front over the past two decades in an attempt to test and detail a galaxy formation and evolution scenario from robust measurements. A key element driving observational progress is the need to cover all major phases of galaxy evolution from the early formation and galaxy assembly to today, a formidable endeavour. Deep galaxy surveys have flourished to conduct this exploration. The latest few billion years have been extensively mapped by surveys like the 2dFGRS (Colless et al. 2001) and from the various stages of the SDSS (Abazajian et al. 2009), setting a firm observational reference after more than 13 billion years of evolution. At larger redshifts deep surveys are providing a complex picture with strong evolutionary features like the build-up along cosmic time of stellar mass in galaxies of different types or the star formation history. Imaging and spectroscopic surveys are playing a complementary role, the deepest studies being performed in photometry and augmented with photometric redshifts (e.g. Ilbert et al. 2006, Rix et al. 2004, Koekemoer et al. 2007, Koekemoer et al. 2011), and with spectroscopic surveys bringing accurate redshifts and spectro-photometry, spectral features properties, as well as internal velocity information. Fundamental quantities have been measured from these surveys including the cosmic star formation history (e.g. Lilly et al. 1996, Madau et al. 1996, Tresse et al. 2007, Bouwens et al. 2009, Cucciati et al. 2012, Madau & Dickinson 2014), as well as the history of stellar mass assembly (e.g. Arnouts et al. 2007, Ilbert et al. 2013).

At redshifts $z \sim 1$ the pioneering CFRS survey (Lilly et al. 1995, Le Fèvre et al. 1995) was followed by more extensive galaxy redshift surveys covering larger volumes like the DEEP2 (Davis et al. 2003), the VVDS (Le Fèvre et al. 2005a, Le Fèvre et al. 2013a), zCOSMOS (Lilly et al., 2007), now reaching the 100 000 redshift mark at $z \sim 1$ with VIPERS (Guzzo et al. 2014). These surveys have brought a wealth of quantitative and accurate measurements now in large enough areas of a few tens of square degrees and volumes of $\sim 5 \times 10^7 h^{-3}\mathrm{Mpc}^3$ so that the most fundamental statistical quantities describing the galaxy population like the luminosity function (LF), the mass function (MF) or the correlation function (CF) are becoming very accurate and less affected by the cosmic variance related to the small fields of earlier studies. At higher redshifts ($z >\sim 2$), the rapid progress has been driven by the effectiveness in selecting high redshift galaxies and, most importantly, by the impressive gains in sensitivity and efficiency provided by high multiplex multi-slit spectrographs like DEIMOS (Faber et al. 2003) on the Keck telescope, and VIMOS (Le Fèvre et al. 2003) on the VLT. The effectiveness of the Lyman-break galaxies (LBG) selection has provided the capability of finding large numbers of galaxies at $z > 2.5$ (e.g. Steidel et al. 2003, Bouwens et al. 2009, Ellis et al. 2013). It is supplemented by narrow band imaging techniques isolating Lyman-α emitters (LAE; Taniguchi et al. 2005, Shimasaku et al. 2006; Ouchi et al. 2008). In addition to these pre-selection techniques, deep purely magnitude selected surveys are conducted to probe a large population mostly free of pre-selection biases. The largest to date probing redshifts $z > 1.5$ is the $i - band$ magnitude selected VVDS survey covering up to $z \sim 6.5$ (Le Fèvre et al. 2005a, 2013a). Other magnitude selected surveys have attempted using redder bands to alleviate selecting only galaxies with strong rest-frame UV continuum. The K20 survey used K-band magnitude selection down to $K = 20$ to identify extremely red objects (EROs) at $z \sim 1.5 - 2$ (Cimatti et al. 2002), and the GMASS survey selected objects on the basis of Spitzer NIR photometry with $m_{4.5\mu m} < 23.0$ (AB) to identify a few hundred galaxies with $1.5 < z < 3$ (Cimatti et al. 2008). However, and despite these successful studies, the approach using pure magnitude selection is costly in observing time when going much beyond redshift $z \sim 2$ or so.

Performing a complete galaxy census is a basic *Astronomy* input necessary for any subsequent *astrophysical* analysis. While at redshifts ~ 1 this census is mostly complete down to stellar mass $10^8\mathrm{M}_\odot$, it is not yet the case at redshifts $z > 1$ for several reasons. First, the colour selection schemes applied to photometric samples to extract the high redshift populations, while efficient to identify galaxies, are affected by significant incompleteness, loosing some fraction of the population at the selected redshift, or by contamination from galaxies at other redshifts. While the latter can be dealt with by obtaining spectroscopic redshifts, the former remains a serious difficulty especially at faint magnitudes and at the highest redshifts. Unfortunately the level of incompleteness strongly depends on the photometric filters used for imaging, the depth of the observations, as well as the image quality, which requires a case by case study involving source simulations complicated by the need to make a priori hypotheses on the properties of galaxies one is trying to establish. It was realised that colour-colour selection like the LBG technique at $z > 2.5$ or like the BzK working at $z \simeq 2$ (Daddi et al. 2004) would miss a part of the general galaxy population in their selection process (Le Fèvre et al. 2005b), therefore making the galaxy census incomplete (Le Fèvre et al. 2013b). The consequences of this maybe far-reaching, as incompleteness in counts leads to underestimates in luminosity density, star formation rates, or mass density, just to cite these few. An important aspect of on-going and future studies is to revisit galaxy counts as a function of redshift making sure that no significant population is missing and that no significant bias is introduced in deriving astrophysical quantities.

⋆ Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, under Large Program 185.A-0791.

A key element is then the availability of large samples of galaxies with a well defined and well controlled selection function. Spectroscopic redshift surveys play a key role as they provide samples with confirmed redshifts. Photometric redshift surveys are widely used and have now reached an impressive accuracy. However the level of 'catastrophic failures' when photometric redshifts disagree with their training set of spectroscopic redshifts, even if low at a few percent (Ilbert et al. 2013), could still produce large unknowns because of the shape of the redshift distribution N(z) of flux limited samples. An error of 1% at the peak, $z \simeq 1$, of the N(z) of a $i_{AB} = 25$ sample could spread galaxies with wrong photometric redshifts to higher redshifts, e.g. at $z \sim 3$ where the projected galaxy density is less than 10 times the N(z) at peak, representing several tens of percent of uncertainty. One recent example is the difficulty to distinguish $z \sim 5$ very massive objects from lower redshifts lower mass $z \sim 2$ galaxies (Wiklind et al. 2008, Caputi et al. 2012). Obtaining a spectroscopic redshift therefore remains a fundamental measurement.

The total number of galaxies spectroscopically confirmed today at $z > 2$ is still limited. Published LBG samples reach ∼2000 redshifts at $z \sim 3$, only a few hundred at $z \sim 4$, and several tens of galaxies beyond that (e.g. Steidel et al., 2003, Vanzella et al., 2009, Stark et al. 2010, Bielby et al. 2013). Samples of LAE galaxies selected with narrow band techniques and confirmed in spectroscopy reach a few hundred objects beyond $z = 3$ (Ouchi et al. 2008, Kashikawa, 2011). The high redshift tail of the VVDS at $z > 2$ contains about 1000 galaxies (Le Fèvre et al. 2013b). To overcome the uncertainties linked to small existing spectroscopic samples, and to understand the biases and limitations of photometry-based studies in their ability to provide a complete census of the galaxy population, extremely deep spectroscopic surveys over large volumes are needed.

Here we present VUDS, the VIMOS Ultra Deep Survey specifically designed to study the early phases of galaxy evolution in $2 < z < 6+$ from ∼10 000 targeted galaxies. The VUDS sample contains an unprecedented number of galaxies with secure spectroscopic redshifts at this epoch, obtained in three different fields: COSMOS, ECDFS and VVDS-02h. The survey design including the target selection is based mainly on photometric redshifts, as presented in Section 2. The VIMOS multi-slit spectroscopic observations, the data reduction and redshift measurement scheme are discussed in Section 3. Properties of the VUDS sample are presented in Section 4, including the redshift distribution of the sample, the distribution of intrinsic properties like stellar mass and star formation rate, and the average spectra properties based on high signal-to-noise stacked spectra. After comparing to other surveys in Section 5, we summarize our results in Section 6 and conclude on the usefulness of the VUDS to study the early phases of galaxy assembly with unprecedented accuracy.

All magnitudes are given in the AB system unless specified, and we use a Cosmology with $\Omega_M = 0.3, \Omega_\Lambda = 0.7$ and $h = 0.7$.

2. Survey design

2.1. Science drivers

When the VUDS survey observing time proposal was approved at the end of 2009, spectroscopic redshift surveys were still of limited scope at the epoch of galaxy assembly significantly beyond redshifts $z \sim 2 - 3$. The VUDS survey therefore aimed to address several core science goals at the epoch $2 < z < 6+$, including: (i) the early history of the global star formation rate, (ii) the build up of the mass function for different galaxy populations, and the contribution to mass assembly and star formation of merging, feedback and accretion, (iii) the study of very young galaxies in their early stages of assembly, including the earliest stellar populations like Pop-III. (iv) the identification of proto-structures and the effects of environment at early stages of galaxy evolution, (v) the study of the clustering of galaxies to infer the mass growth of underlying dark matter halos,

These broad science goals still remain hot science topics today and are the main focus of VUDS.

2.2. Survey strategy overview

The VUDS survey is designed to study the galaxy population beyond redshift $z \simeq 2$ and up to the highest redshifts possible in the wavelength range up to $\lambda = 9350$Å accessible with the VIMOS spectrograph on the VLT (see below), corresponding to Lyα at a redshift up to $z \simeq 6.6$. A total of one square degree is observed in three well separated fields, the COSMOS field, the extended Chandra Deep Field South (ECDFS) and the VVDS-02h field, in order to mitigate cosmic variance effects. This program was awarded Large Program status with 640h of VLT-VIMOS observing time.

A key feature of the survey is the target selection. It is primarily based on photometric redshifts and their probability distribution function (PDF), complemented by colour-colour selections and analysis of the SED shape when not already in the photo-z selection. When the geometry of the observed slit-masks allows it after placing the previous priority targets, a purely magnitude-selected random sample of galaxies with $23 \leq i_{AB} \leq 25$ has been added to the target list. This is further described in Section 2.5.

The second key feature is the extended wavelength coverage of the VIMOS spectroscopic observations covering $3650 \leq \lambda \leq 9350$Å, which minimizes any instrument-induced redshift desert and strongly reduces the degeneracies in redshift measurements (Le Fèvre et al. 2013a).

The last key point is the integration time of $\simeq$ 14h per target, which allows reaching a S/N on the continuum at 8500Å of $S/N = 5$ for $i_{AB} = 25$, and $S/N = 5$ for an emission line with a flux $F = 1.5 \times 10^{-18}$ erg/s/cm^2/Å.

2.3. Survey fields

VUDS covers three of the most observed extragalactic fields: the COSMOS field, the ECDFS and the VVDS-02h (also known as the CFHTLS-D1/XMM-LSS field). A total of 0.5 square degrees are observed in the COSMOS field, 0.31 square degrees in the VVDS-02h field, and 0.14 square degrees in the ECDFS. The fields location and covered area for each are summarized in Table 1, and identified in Figures 1 to 3.

2.4. Multi-wavelength data and photometric catalogues

By design the three survey fields accumulate a wealth of deep multi-wavelength data and spectroscopy. This is an important part of VUDS as multi-wavelength coverage of the spectral energy distribution (SED) is essential in combination with accurate spectroscopic redshifts, in particular to derive key physical quantities like absolute magnitudes, stellar masses, star formation rates or ages. We describe for each field the most relevant existing data. We note that when the spectroscopic targets were selected over the 2010-2012 period we used the datasets available at the time as an input to the selection function. The physical parameters derived from multi-wavelength photometry reported here (e.g. stellar mass $M_\star$ and star formation rate SFR) and in other VUDS studies make use of the latest datasets available.

The COSMOS field (Scoville et al. 2007) is extensively observed at all wavelengths. Its location was setup close to one of the original VVDS fields, with the center moved to a region with less galactic extinction when better extinction maps became available. The field was observed with HST-ACS in the F814W filter (Koekemoer et al. 2007) providing high spatial resolution imaging down to a depth $AB(F814W) = 27.2$ (5σ). Extensive imaging observations were obtained in g, r, i, and z bands from the Subaru SuprimeCam (Taniguchi et al. 2007), as well as from the CFHT Megacam from the CFHT-Legacy Survey including the u-band, and photometry is available from 12 medium bands (references available in Ilbert et al. 2009). The UltraVista survey is acquiring very deep near-infrared imaging in the Y, J, H and K bands using the VIRCAM camera on the VISTA telescope (McCracken et al. 2012). The UltraVista DR2 release reaches a depth $K_{AB} = 24.8$ at 5σ (in 2 arcsec apertures), and it is planned to increase the depth down to $K_{AB} = 25.5$ in half the area by the end of the survey. The VUDS VIMOS pointings are placed in a way to optimize the overlap with the deeper UltraVista stripes (Figure 1). Following on the initial Spitzer observations (Sanders et al. 2007), deeper Spitzer-warm observations are obtained by the Splash program, reaching $AB = 25.4$ (5σ at 3.6μm) (Capak et al. in prep.) with a smaller area on one UltraVista stripe observed from the SEDS survey pushing down to $AB = 26$ at 3σ (Ashby et al. 2013). The COSMOS field has deep Herschel imaging reaching 8 mJy at 250 μm (Oliver et al. 2012). The CANDELS survey (Grogin et al. 2011) in this field includes WFC3 imaging in the F125W (J) and F160W (H) filters down to $H_{AB} = 27.1$ (5σ); the CANDELS area is almost entirely covered by the VUDS pointings (Figure 1).

Extensive spectroscopy was obtained in the COSMOS field prior to VUDS. The zCOSMOS survey (Lilly et al. 2007) contains 20 000 galaxies with $0 < z < 1.2$ selected with $i_{AB} \leq 22.5$ and $\sim$ 6000 galaxies selected with $g_{AB} \simeq 25 - 25.5$ with $1.4 < z < 3$ and a median redshift $z = 2.15$ (Lilly et al. in preparation). Deep spectroscopy from Keck has identified several hundred galaxies in the same redshift range as VUDS (e.g. Capak et al. 2011). VUDS targeting avoided to re-observe those galaxies with existing reliable redshifts, when known. A wealth of observations at other wavelengths are available in this field, we refer the reader to the COSMOS web site (`http://cosmos.astro.caltech.edu/`) for the detailed list and properties.

The ECDFS field is the focus of a number of deep multi-wavelength surveys. Built around the field with deep 1Ms Chandra observations (Giacconi et al. 2002), the field was extended and is covered with deep UBVRI imaging down to $R_{AB} = 25.3$ (5σ, Cardamone et al. 2010 and references therein). Several spectroscopic surveys have been conducted (Le Fèvre et al. 2004, Cardamone et al. 2010, Cooper et al. 2012), with more than 5000 galaxies with redshifts $0 < z < 2$, but the number of spectroscopically confirmed galaxies at $z > 2$ remains small. The central part of this field accumulates a number of deep HST imaging with the HDF-South (Williams et al. 2000), GEMS (Rix et al. 2004), and ERS (Windhorst et al. 2011) surveys, and more recently with the CANDELS survey (Grogin et al. 2012) adding WFC3 near-IR imaging reaching as deep as $H_{AB} = 27.3 - 27.6$. The SERVS Spitzer-warm obtained 3.6 and 4.5 μm imaging data down to $AB = 23.1$ (Mauduit et al. 2012).

The VVDS-02h field was originally defined as the 'deep field' of the VVDS survey. It was observed in BVRI at CFHT (Le Fèvre et al. 2004), followed by deeper u', g, r, i observations as the D1 deep field in the CFHTLS survey reaching $i_{AB} = 25.44$ at 50% completeness in the latest DR7 (Cuillandre et al. 2012). Deep infrared imaging was obtained with the WIRCAM at CFHT in YJHK bands down to $Ks_{AB} = 24.8$ also at 50% completeness (Bielby et al. 2012). Extensive multi-slit spectroscopy was obtained with VIMOS as the 'Deep' and 'Ultra-Deep' surveys of the VVDS, with magnitude-selected samples down to $I_{AB} = 24$ and $i_{AB} = 24.75$ respectively. A total of 11 139 objects have spectroscopic redshift measurements from the VVDS, including 774 with $z > 2$. This field was observed in all Spitzer bands as part of the SWIRE survey (Lonsdale et al. 2003), reaching a magnitude in the 3.6 microns band of $AB = 21.5$. More recently the SERVS survey obtained deeper data with Spitzer in the 3.6 μm and 4.5 μm bands down to $AB = 23.1$ at 3.6 μm (Mauduit et al. 2012). It is one of the fields of the HERMES survey (Oliver et al. 2012), matched to CFHTLS and VVDS data (Lemaux et al. 2014a). A number of other deep data are available including X-ray (Pierre et al. 2004) and radio (Bondi et al. 2003) observations.

In each field photometric catalogues including at least from the u-band to the Spitzer 4.5 μm band have been assembled, matching the different imaging sources, and extracting photometry in 2 arcsecond apertures in dual mode using SExtractor (Bertin & Arnouts 1996). On the COSMOS field we used the photometric catalogues published by the COSMOS collaboration (see e.g. a full description of the u-band to 4.5μm catalogues in the context of the UltraVista survey, Ilbert et al. 2013). In the VVDS-02h field we used the CFHTLS photometry for u-band to z-band (Cuillandre et al. 2012) and the WIRDS photometry for YJHK (Bielby et al. 2012), PSF-matched. In these two fields, the source detection was performed using χ^2 images resulting from a combination of images from u-band to z-band. In the ECDFS field we used the catalogues assembled by Cardamone et al. (2010), and the reader is referred to this paper for the details on its construction.

2.5. Spectroscopic target selection

As shown from the VVDS, the redshift distribution of faint galaxies, e.g. down to $i_{AB} \simeq 25$ is peaked at $z \sim 1.3 - 1.5$, with a rapidly decreasing high redshift tail (Le Fèvre et al.

2013b). Magnitude selection is then inefficient in picking-up $z > 2$ galaxies among the far more numerous foreground. This prompted the use of the Lyman-break selection technique, as pioneered by Steidel and collaborators (Steidel et al., 1996). This technique makes simple use of the colour properties of galaxies as a function of redshift, identifying a locus in colour-colour space associated to specific redshift ranges.

A more sophisticated technique to pre-select high redshift galaxies follows the development of the photometric redshift technique and its demonstrated success in terms of the redshift measurement accuracy with typical errors of less 5% on $\Delta z/(1+z)$ and a low rate of catastrophic errors when compared to spectroscopic redshifts (e.g. Ilbert et al. 2010, Cardamone et al. 2010). In essence, photometric redshift techniques are a generalisation of colour selection techniques such as LBG or BzK to the complete spectral energy distribution (SED) which must satisfy a colour distribution across wavelengths compatible with reference galaxy templates. We elected to use the best possible photometric redshifts as our main selection criterion to define spectroscopic targets. Photometric redshifts have been computed with the code Le Phare (Arnouts et al. 1999, Ilbert et al. 2006) using the multi-band photometric catalogues covering from the u-band to the Spitzer 4.5μm band. The techniques are described in Ilbert et al. (2009), and more recently in Ilbert et al. (2013). The observed photometric data are correlated against 33 templates covering a range from early to late type galaxies of different ages, star formation histories and metallicities, leaving the E(B-V) extinction as a free parameter. Emission lines with a flux empirically computed from the UV continuum flux are added to the templates. The 'best' redshift is assigned from the median of the marginalized redshift probability distribution function (PDF).

To select the spectroscopic targets in VUDS we followed an *inclusive* rather than exclusive strategy, adding samples pre-selected from several different criteria, as described below:

- The spectroscopic targets must satisfy $z_{phot} + 1\sigma \geq 2.4$, and $i_{AB} \leq 25$. As degeneracies are known to occur, we selected sources for which either the first or the second peak in the photometric redshift probability distribution function satisfy this criterion. This corresponds to 87.7% of our primary sample.
- Sources which were not selected from the primary z_{phot} criterion but satisfy one of the ugr, gri or riz LBG colour-colour selection criteria have been added to the target list, representing 8.4% of the primary sample.
- The full SED over all filters available in ugrizYJHK is used to identify galaxies with a break in the continuum and not identified by any of the z_{phot} or LBG criteria, extending the *dropout* technique. These targest are allowed to have $K_{AB} \leq 24$ and hence may not necessarily be brighter than $i_{AB} = 25$. This adds another 3.9% to the primary sample.
- Finally, when space on the slit mask was still available after the above selections, a random 'secondary' sample of galaxies with $23 \leq i_{AB} \leq 25$ has been selected, representing 11.5% of the total sample.

For the last half of the survey observations, priority for slit placement was given to the targets with $z_{estimate} \geq 4$ to increase the number of objects at these redshifts in the final sample.

In adding these different selection criteria we aim at maximizing the pre-selection of objects above $z = 2.4$. Given the dispersion of photometric redshift errors, we expect the redshift distribution of VUDS sources to start rising at $z \simeq 2$. While the selection strategy described above could lead to an increase in the contamination from objects at lower redshifts, this contamination remains quite small as described in Section 4.1. For each of the targeted samples, slits are placed at random using the `vmmps` slit mask design software (Bottini et al. 2005), providing a fair and representative sample of the general population.

These various selection criteria are summarized in Table 2. Besides the magnitude-selected sample used as a filler in the slit-mask design, the targeted high-z sample has been selected for 87.7% on the basis of photometric redshifts, with 65.9% from the first peak in the photometric redshift PDF, and ∼ 21.8% from the second peak. The rest ∼ 12.3% is made up of the galaxies satisfying standard LBG and dropout colour-colour selection. The selection of VUDS galaxies is well defined and well controlled as both the parent catalogues and selection criteria are well documented and can be easily reproduced; this will be extensively described in a forthcoming paper. We note here that the selection technique providing the largest number of sources is the photometric redshift technique. Only about 50% of our sample satisfy LBG criteria over the full redshift range, therefore using LBG selection alone we would have missed ∼50% of the star-forming galaxy population. This can be easily understood as the photometric redshift selection produces a continuous redshift distribution while in each redshift range defined by any two colour filter set the LBG selection produces a bell-shaped redshift distribution, missing an increasingly large fraction of sources at the redshift extremes when the continuum features produced by the Lyman limit and IGM absorption move through the 3 filters. When using several colour-colour selections to cover a large redshift range (e.g. ugr, gri, riz) the redshift distribution from LBG selection is the sum of several of these bell-shaped functions (as seen e.g. in Bouwens et al. 2014, fig.1). While correcting for this incompleteness in photometric LBG samples is usually done using simulations, LBG source selection as an input to spectroscopy over a large redshift range $2 < z <\sim 6$ like in VUDS would then miss ∼50% of the possible targets from the onset. Our strategy using photometric redshifts rather than LBG selection was designed to avoid this severe difficulty. The effect of our selection function and a comparison between the different selection criteria will be analysed in details in a forthcoming paper.

3. Observations, data processing, and redshift measurement

3.1. VIMOS on the VLT

The VIsible Multi-Object Spectrograph (VIMOS) is installed on the European Southern Observatory Very Large Telescope unit 3 Melipal. VIMOS is a wide field imaging multi-slit spectrograph (Le Fèvre et al. 2003), offering broad band imaging capabilities, as well as multi-slit spectroscopy. VIMOS is a high performance Multi-Object Spectrograph (MOS) with 4 parallel channels, each a complete imaging-spectrograph with a field of view 8×7 arcmin^2, or a total field of 224 arcmin^2. The key features of VIMOS are a high multiplex (number of slits) and the excellent sky subtraction accuracy reaching $\sigma_{sky\ residual} \simeq 0.1\%$ of the sky signal (Le Fèvre et al. 2013a).

For VUDS we use the low resolution multi-slit mode of VIMOS; with the 4 channels, this offers the largest multiplex for multi-slit spectroscopy. Following Scodeggio et al. (2009) we have optimized the slit length allowing lengths as small as 6 arcseconds, maximizing the number of observed slits.

3.2. VIMOS observations

We observe a total of 16 VIMOS pointings, at the coordinates of the pointings identified in Table 3. Fifteen pointings are observed with both the LRBLUE grism covering $3650 \leq \lambda \leq 6800$Å and the LRRED grism covering $5500 \leq \lambda \leq 9350$Å, leading to a full wavelength coverage of $3650 \leq \lambda \leq 9350$Å. With slits one arc-second wide, these grisms provide a spectral resolution $R = 230$ quite uniform over the wavelength range. At this resolution, each of the four 2048×4096 pixels detectors can accommodate 3–4 full length spectra along the dispersion direction, and given the projected space density of VUDS targets we therefore observe on average $N_{slits} \simeq 600$ individual slits simultaneously. Each slit may contain not only the VUDS pre-selected target but also serendipitous objects falling in the slit by chance.

One pointing on the ECDFS (#3) has been observed with the MR grating. This setup covers $5000 \leq \lambda \leq 9500$Å with a spectral resolution $R = 580$. With this resolution about 2 full length spectra can be placed along the dispersion direction, and a total of ∼ 220 objects have been observed in this pointing. The targets for this pointing have been optimized towards the highest redshifts $z > 4.5$; the improved resolution compared to the LR grisms making it somewhat easier, in principle, to identify emission lines like Lyα between the OH sky emission features (but possibly at the expense of sensitivity on continuum measurements).

Most VUDS observations were obtained after the CCD detectors were upgraded in summer 2010. At that time, the original blue–sensitive thinned E2V CCDs were changed to red–optimized thick E2V CCDs in 2010 (Hammersley et al. 2010). The global VIMOS sensitivity at 9000Å increased by ∼ ×2, making it comparable to the FORS2 sensitivity in the red with a field of view 4.8× larger, and significantly reducing the fringing above 8000Å thanks to the thicker substrate.

To reach a total integration of 14h, each of the LRBLUE or LRRED grism observations consists on average of 13 observing blocks (OBs) executed at the telescope. Each OB includes three spectroscopic exposures of 1250 to 1350 seconds obtained dithering $-0.75, 0, +0.75$ arcseconds along the slit. The OBs specify the observing conditions that must be met, including a seeing better than 1 arcsecond, sky transparency set to 'clear', airmass less than 1.5, dark time with constraints on the moon phase and distance from the field (lunar illumination 0.3 to 0.5, and distance to the observed field of more than 60-90 degrees). Arc lamp and flat field calibrations are obtained after each set of 3 OBs, and flux calibration on standard stars are performed in the standard ESO procedure. When ready, OBs are sent to the VLT-VIMOS service observing queue via the P2PP tool and get executed when the atmospheric and moon conditions are met.

3.3. Spectroscopic data processing, redshift measurement, and reliability flags

The spectroscopic data processing followed the same general principles as defined for the VVDS (Le Fèvre et al. 2005a, 2013a). We summarize this process below and emphasize the specific data processing steps that we follow for VUDS.

The general outline of the VUDS data processing follows what is now a standard for multi-slit spectroscopy. We use the VIPGI environment to process the spectra (Scodeggio et al. 2005). First the 40 individual 2D spectrograms coming from the 13 OBs for one of the LRBLUE or LRRED observations are extracted finding the location of the slit projection on the detector using the expectation from the slit mask design. Sky subtraction is performed with a low order spline fit along the slit for each wavelength sampled. The sky subtracted 2D spectrograms are combined with sigma clipping to produce a single stacked 2D spectrogram calibrated in wavelength and flux. The 2D spectrogram is collapsed along the dispersion direction to produce a slit profile in which objects are identified. The spectral trace of the target and other detected objects in a given slit are linked to the astrometric frame to identify the corresponding target in the parent photometric catalogue.

This processing is performed separately for each of the LRBLUE and LRRED observations. It is to be noted that, in the wavelength range of overlap $5500 \leq \lambda \leq 6800$Å, the end spectra are the average of the LRBLUE and LRRED, accumulating about 28h of observation. The 1D sky-corrected stacked spectra are extracted following the observed slit profile. They are fully calibrated both in wavelength and in flux using spectrophotometric standard stars. The 2D and 1D fully calibrated spectra are cross matched per observed slit, and joined to form spectra with full $3650 \leq \lambda \leq 9350$Å wavelength coverage. In this process additional objects can be detected in the slit and have subsequently their 1D spectra extracted. These objects are called "secondary detections", and a photometric counterpart is searched for in the photometric catalogue. If an object is found within less than one arcsecond from the spectral trace, the secondary spectrum identification is assigned the identifier of the object in the photometric catalogue. If no object is found in the photometric catalogue, the spectrum of the object is extracted and a new entry is produced in the catalogue with coordinates at the location corresponding to the trace of the object. It might happen that a single emission line is identified upon visual examination of the 2D spectrogram of a given slit, but not extracted by the automated procedure: these objects are flagged and then manually extracted. This procedure is particularly important for objects with single emission lines and no or very little detected continuum, which often turn out to be Lyα emitters at high redshifts (e.g. Cassata et al. 2011).

Upon comparison between the photometric magnitudes and the magnitudes derived from the calibrated spectra, we realized that the $u-$band part of the spectra (and to a lesser extent the $g-$band) was lacking photons at the ∼ 40% (∼ 15%) level. We proceeded to add three well-defined corrections to the spectra: (1) atmospheric absorption, (2) atmospheric refraction, and (3) Galactic extinction. The atmosphere absorbs photons depending on the airmass along the light path; this is corrected using the prescription defined for the Paranal observatory in Patat et al. (2011). In addition atmospheric refraction acts as a small prism before entering the telescope, spreading the incoming

light into a spectrum with length depending on the airmass and paralactic angle, the angle of the slit to the zenithal angle. With slit-masks placed on sky objects using an $r-$band filter prior to spectroscopic observations, and slits one arcsecond wide, this may introduce a significant loss of uv-blue photons falling out of the slit. The Galactic extinction on distant sources also produces a chromatic correction, which has been applied using the E(B-V) maps of Schlegel et al. (1998). In this process we also proceed to produce spectra with continuum flux which are calibrated on the $i-$band photometric flux, therefore correcting for any slit losses occurring in that band if the object extension is larger than the slit width.

Adding these corrections we are able to correct the uv-blue spectroscopic flux in such a way that there is excellent agreement between the spectroscopic flux and the photometric flux measurements at better than the 5% level in all wavelengths $3650 \leq \lambda \leq 9350$Å (Thomas et al. 2014). This is of particular importance when fitting the spectra to reference templates in order to derive internal galaxy properties (Thomas et al., in prep.). One of the last observations obtained on this program includes a number of repeated observations on a sub-sample of VUDS galaxies. When processed it will allow to estimate redshift (velocity) measurements accuracy, as well as to have an independent check of the reliability of each of the redshift reliability flags (see below). Based on the previous VVDS, zCOSMOS and VIPERS surveys with the same VIMOS instrument, we expect the redshift accuracy to be in the range $dz/(1+z) = 0.0005 - 0.0007$, or an absolute velocity accuracy $150 - 200$ km/s (Le Fèvre et al. 2013a). Relative velocities in the same slit (along the slit profile) can be measured to a better precision using e.g. accurate spectral line fitting.

The 2D and 1D spectra are then available for spectroscopic redshift measurements using the EZ environment (Garilli et al. 2010). The core algorithm to find a redshift is the cross-correlation with templates, confronted to a separate estimate of an emission line redshift when applicable. A key element for the cross-correlation engine to deliver a robust measurement is the availability of reference templates covering a large range of galaxy and star types, as well as a large range of rest wavelengths. This last point has to be carefully dealt with when measuring the highest redshift galaxies, as it is necessary that the templates go far enough in the UV, bluer than the Lyman-912Å limit, for the wavelength overlap between the observed galaxy and the redshifted template to be large enough to provide a robust correlation signal. We have used templates built over the years from VIMOS observations for the VVDS (Le Fèvre et al. 2005a, 2013a) and the zCOSMOS survey (Lilly et al. 2007). As is relevant for $z > 2$ we have used templates with and without Lyman$-\alpha$ emission.

The redshift measurements are first obtained from an automated run with EZ. This serves as a basis for a visual examination of each spectrum, with an iteration on the redshift measurement using EZ in manual mode, if necessary. All spectra are visually examined. For about half of the spectra the redshift solution found from the automated EZ run is retained as the final one without further iteration and the human intervention consists only in assigning a reliability flag based on the spectral features identified. However, we find that at these high redshifts more than half of the spectra need manual intervention to properly measure a redshift. This is mainly due to residual defects in the spectra like sky features residuals after sky subtraction or second order spectra superimposition. The manual intervention consists first in cleaning the spectrum from obvious noise residuals at the location of strong sky lines or zero order projection. The EZ algorithm is then run again and the main peaks of the correlation are examined within the graphical interface allowing to overplot the best template fit at each peak in the correlation signal. The best redshift is in general the second or third strongest correlation peak, or, in case of single line emission, the redshift is derived from a Gaussian fit of the line.

As the manual intervention remains an important feature in this process, we have implemented a method to minimize measurement biases linked to one single person. One spectrum is measured by two team members separately, and these measurements are then confronted to produce a single measurement agreed on by the two measurers. With this scheme, we have implemented the same redshift reliability estimator as developed for the CFRS (Le Fèvre et al. 1995) and refined for the VVDS (Le Fèvre et al. 2005, 2013), zCOSMOS (Lilly et al. 2007) or the VIPERS (Guzzo et al. 2013) surveys. The reliability of a redshift measurement is expressed with a flag giving the range of probability for a redshift to be right. The reliability flag may take the following values:

- 0: No redshift could be assigned (the redshifts are then set to 9.9999).
- 1: 50–75% probability to be correct
- 2: 75–85% probability to be correct
- 3: 95–100% probability to be correct
- 4: 100% probability to be correct
- 9: spectrum with a single emission line. The redshift given is the most probable given the observed continuum, it has a ∼ 80% probability to be correct.

The probabilities associated to these reliability flags are remarkably stable because of the process involving several independent people, smoothing out individual biases (Le Fèvre et al. 2013a).

VUDS enters a redshift domain which has never been probed by spectroscopic redshift surveys on this scale. The expertise of the VUDS team members in measuring redshifts grew steadily as more and more of the data was being processed. Upon examination of the first measurements we realised that redshifts were probably wrongly assigned for a small but sizeable (∼ 10%) fraction of the objects. Several standard cases for erroneous measurements were identified, including the possible confusion between early M stars (M0-M3) and $z \sim 5$ absorption-line only galaxies (going both ways), the assignment of Lyα-1215Å instead of OII-3727Å (or vice-versa), the setting of a continuum break to the Balmer-D4000 break rather than the Lyα break (or vice-versa). In addition, the reliability flag as defined above was sometimes either too cautious when a measurer found a spectrum he/she was not yet familiar with, or too optimistic.

In view of this, we opted to create a 'Tiger Team' (OLF, PC, EV, BG, DM, VLB, OLF, LP, LTa) in charge to conduct an additional redshift check, provided with a set of well defined reference cases and their treatment. This redshift check was done again by two independent people going through all the measured spectra, separately identifying which ones needed to have their redshift and/or reliability flag modified, and agreeing on the modifications. The check of all pointings was done

by four pairs of 2 people, each couple proposing a list of modifications to be examined and agreed upon by the other 'Tiger team' members. We compare in Figure 4 the *old redshifts* with the *Tiger team redshifts* which summarizes this process. At the end of this process, about 10% of the objects had either a redshift or a flag change. While this does not guarantee that there are no more obvious 'catastrophic failures' in spectroscopic redshift measurements, this process ensures a homogeneous treatment of all spectra and the reduction of the main degeneracies present in measuring the redshifts of high redshift galaxies. Obviously, for the fainter objects where the information content of the spectra is not sufficient to solve a possible degeneracy, the reliability flag is assigned to the 'flag 1' category.

One key element of the selection function of the VUDS sample is the target sampling rate (TSR) defined as the ratio of the observed galaxies (all reliability flags) to the underlying parent photometric populations from which the spectroscopic targets have been selected. We find a global TSR of ∼ 30% for the VUDS survey, similar for the three observed fields. With respect to the total population with $i_{AB} \leq 25$, the parent sample of galaxies satisfying the VUDS selection criteria represents 10%, and hence the observed VUDS sample represents 3.3% of all galaxies with $i_{AB} \leq 25$. The most reliable redshifts including flag 2, 3, 4 and 9 account for 70.2% of the sample, and we advocate the use of these objects for most science analysis, properly weighting galaxies if appropriate. Objects with flag 1 represent 21.4% and 8.4% of the targeted sample could not be measured (flag 0). The spectroscopic success rate (SSR) as a function of magnitude is shown in Figure 5. Down to $i_{AB} = 25$ 91% of targets have a redshift measurement (flags $1, 2, 3, 4, 9$), and 74.3% have a reliable measurement (flags $2, 3, 4, 9$). This fraction is decreasing to 58% in the last 0.25 magnitude bin before $i_{AB} = 25$. The spectroscopic success rate then goes down with magnitude with a reliable redshift obtained for $\simeq$ 30% of the galaxies targeted at $i_{AB} = 26$. A complete description of the survey selection function including the analysis of the spectroscopic success rate at different redshifts will be provided elsewhere (Tasca et al. in prep.).

The experience gained in this process is invaluable for future massive high redshift surveys. As in Le Fèvre et al. (2013a), we emphasize that redshift measurements *at these high redshifts* is a complex process which deserves dedicated and expert care beyond a simplistic 'good' vs. 'bad' redshift measurement scheme to fully exploit the information content of faint object spectroscopy at the instrumental limit. This is further discussed in Section 3.5 in view of the aposteriori comparison between spectroscopic and photometric redshifts.

3.4. VUDS spectra

We present sample spectra over the redshift range of the survey in Figures 6 to Figures 9. The signal-to-noise per sampling element (∼ 7Å) of the spectra at 1500Å rest-wavelength has a mean of $S/N = 4.5$, and $\sigma_{S/N} = 2.1$, the S/N per spectral resolution element being ∼ 2× higher. This gives access to a range of spectral features and properties for each individual galaxy. The main spectral lines identified in individual spectra are Lyman−α 1215Å (in emission or in absorption), OIλ1303, CIIλ1334, the SiIV-OIV doublet at λ1394-1403Å, SiIIλ1527, CIVλ1549, FeIIλ1608, HeIIλ1640, AlIIλ1671, AlIIIλ1856, CIII]λ1909 (see the list in Table 4). Below Lyman−α, and depending on the IGM absorption, Lyman−β, Lyman−γ and the Lyman limit at 912Å can be identified. The average spectral properties in different redshift ranges are described in section 4.3.

An overview of the spectra of the VUDS galaxy population over $2 < z <\sim 6$ is shown in Figure 10. This figure is built from all spectra with reliability flags 3 and 4 in $2 < z < 4$ and all spectra with flags 2, 3 and 4 for $z > 4$ assembled in one single image, one spectrum per image line. The display is quite striking as the eye is able to follow up some of the weakest spectral features up to the highest redshifts. The Lyα line is readily visible in absorption or emission all along the redshift range, with the blue wing of the broad damped Lyα clearly visible. Going to the higher redshifts one can note the Lyβ, Lyγ, and Lyman-break (912Å).

3.5. Comparison between spectroscopic and photometric redshifts

With VUDS targets selected in large part from their photometric redshifts (Table 2), we make here a comparison between the photometric redshifts z_{phot} computed from the multi-wavelength data set and used as an input to the target selection, and the spectroscopic redshifts z_{spec}. We compare z_{phot} and z_{spec} for the VUDS sample with reliable flags 3 and 4 in Figure 11 (bottom-left panel). As close to 100% of the z_{spec}(flags 3,4) are the "truth", we can directly test the accuracy and degeneracies of the z_{phot}. We present in Figure 12 the distribution of $\delta z = (z_{spec} - z_{phot})/(1 + z_{spec})$ for flags 3+4 and 2+9 for $z < 1.5$ and z> 2 separations arbitrarily chosen to distinguish a low redshift regime where the z_{phot} computation heavily relies on rest-frame visible domain features like the D4000 spectral break, and a high redshift domain where the computation rests on UV rest-frame features like the Lyα break produced by the intervening IGM and the Lyman-continuum limit. The width of the distribution for flags 3+4 below redshift $z = 1.5$ is an excellent $\sigma(\delta z/(1 + z)) = 0.02$, particularly in the COSMOS and ECDFS fields which benefit from medium band photometry. Above redshift $z = 2$ it is about double which is still excellent but signals the increasing difficulty to assign accurate photometric redshifts from broad band photometry only.

It is immediately visible in comparing the photometric redshifts and spectroscopic redshifts for the best reliability flags (Figure 11, bottom-left panel) that there is a secondary tight relation with a low z_{phot} corresponding to a high z_{spec}, which is easily explained by the degeneracy between the D4000 continuum break and the Lyman−α break when computing z_{phot}, and our use of the secondary peak of the z_{phot} PDF to select spectroscopic targets. We draw in Figure 11 the relation $z_{phot} = (1215/4000)\times(1+z_{spec})-1$ and ±15% around this mean. As seen in the bottom-left panel, a number of galaxies are well within these limits, representing 9% of the flag 3 and 4 sample at $z > 2$. In all, 95.2% of our flag 3+4 sample at $z_{spec} > 2$ satisfy either $z_{spec} = z_{phot}$ or $z_{phot} = (1215/4000) \times (1 + z_{spec}) - 1$ within 15%, meaning that our selection function is particularly effective in picking–up galaxies at these redshifts, with low 'catastrophic failure' rate, as further discussed in Tasca et al. (in prep.).

We can use the distribution of (z_{spec}, z_{phot}) for flags 3 and 4 to evaluate the level of agreement between z_{phot} and z_{spec} for other spectroscopic reliability flags 2, 1 and 9. The comparison between z_{phot} and z_{spec} for reliability flag 2, 1 and 9 is presented in Figure 11. The distribution of galaxies for flags 1, 2 and 9 is qualitatively similar to the distribution for flags 3+4, with the 1:1 and the 4000:1215 relationships well populated. For flags 2+9 and $z > 2$ we find $\sigma(\delta z) = 0.044$ hence not much different from the flags 3+4 distribution, and with about 75% to 80% of the objects in this category within 15% of the 1:1 and 1215:4000 relations, which fully supports that these redshifts have a high level of reliability. As noted above, the secondary peak in these distributions for $z > 2$ is produced by the degeneracy betweeen the D4000 and 1215Å continuum breaks. This peak is more pronounced for flag 2+9, with about 23% of the objects in this category within 15% of the 1215:4000 relation, meaning that the objects with these reliability flags are more prone to this degeneracy. A possible reason is that the magnitudes of flag 2+9 objects are on average fainter than the flag 3+4 counterparts. Defining the catastrophic failure rate in the selection of $z > 2.4$ VUDS galaxies as the fraction of galaxies which are outside either $\mid z_{spec} - z_{phot} \mid \leq 0.15 \times (1 + z_{spec})$ or $z_{phot} - (1215/4000) \times (1 + z_{spec}) - 1 \leq 0.15$, it is 20% for flag 2, 37% for flag 1 and 24% for flag 9. Taking into account the intrinsic catastrophic failure rate of 5% for photometric redshifts as observed for flag 3+4 this gives a qualitative estimate on the reliability level of the different spectroscopy flags.

Based on the excellent match between spectroscopic and photometric redshifts obtained for flag 3 and 4, we have added a decimal point to the reliability flag as defined in Section 3.3 translating the level of agreement between the spectroscopic redshift and the photometric redshift for each galaxy. This decimal point may take five different values from 1 for a poor agreement, to 5 for an excellent agreement; more specifically: *.1 means that the spectroscopic and photometric redshifts have a difference $dz = \mid z_{spec} - z_{phot} \mid /(1 + z_{spec}) \geq 0.5$, *.2 is for $0.3 \leq dz < 0.5$, *.3 for $0.2 \leq dz < 0.3$, *.4 for $0.1 \leq dz < 0.2$, and *.5 for $dz < 0.1$. Adding a *.5 decimal therefore further increases the reliability of the spectroscopic redshift measurement, while a *.1 decimal would rather lower it. However, for the higher spectroscopic reliability flag 3 and 4, a low photometric decimal *.1 or .2 would rather indicate that the photometric redshift is likely to be incorrect. This scheme allows to define a sample depending on the level of robustness required by a particular analysis.

This analysis brings two general comments for studies based on photometric redshifts. First, if VUDS targets had been selected based only on the primary peak of the z_{phot} PDF, ∼ 17.5% of the sample with $z \geq 2$ would have been lost compared to our selection. Second, for studies based only on a z_{phot} sample, which necessarily assign a redshift using the primary peak of the redshift PDF, 14.2% of high redshift galaxies would be missed because they would be wrongly placed at low redshift instead of $z > 2.3$. This average value shows a variation with redshift, which will be discussed in a forthcoming paper.

4. General properties of the VUDS sample

In this section we report on the main properties of the VUDS sample, giving an overview of the parameter space probed by the survey. As of this writing, the VUDS sample contains 6 250 objects with a measured redshift, including 6 003 galaxies, 20 AGNs and 227 stars, and no redshift measurement could be obtained for 750 objects. These numbers will increase by 15-20% when data processing will be completed. The projected sky distribution of the VUDS sample follows the layout of the VIMOS pointings, as identified in Figures 1 to 3. With 8 pointings on the COSMOS field, ∼ 4150 objects previously without measurements have been observed and ∼ 3700 have now spectroscopic redshift measurements (these numbers are being revised following the last data processing), covering a total area of 1800 arcmin^2. Over 1125 arcmin^2 in the VVDS-02h field the data for ∼ 2300 targets have been processed and ∼ 2100 objects have now spectroscopic redshifts, and we expect 20% more when data processing will be completed. In the ECDFS, one pointing has been processed so far with ∼ 550 objects observed and ∼ 500 with a redshift measurement, and with 2 more pointings to process (one with LR and one with MR grisms) we expect this number to increase by about 60%. We use the current sample to discus general properties of the sample below.

4.1. Sample properties

The efficiency of the survey target selection can be estimated from the redshift distribution of the observed sources.

The redshift distribution of the VUDS sample is shown in Figure 13. The N(z) is bimodal, with a high redshift component from $z \sim 2$ to $z \sim 6.5$, and a low redshift component mainly at $z < 1.5$. The sample above $z = 2$ is the sample of interest for the main science goals of VUDS; it represents about 80% of the total sample, with currently more than 4500 objects with redshift measured with $z \geq 2$, and, extrapolating for the remaining data to be processed the VUDS sample, it will contain ∼ 6000 objects with $2 < z < 6.5$ in the end. VUDS is today the largest sample of galaxies with spectroscopic redshifts in any of the redshift bins $2 < z < 3, 3 < z < 4, 4 < z < 4.7$ or $4.7 < z < 5.3$, as further discussed in Section 5.

The sample below $z \simeq 2$ is made of several sub-samples: (i) galaxies for which the second peak of the z_{phot} PDF is at $z > 2.4$ but which turned out to be at the lower redshift indicated by the first peak, (ii) galaxies for which the selection criteria for high redshift failed and which are rather at low redshifts, (iii) galaxies with $i_{AB} \leq 25$ which have been used as mask fillers and for which the N(z) is expected to peak at $z \sim 1.5$ (Le Fèvre et al. 2013b), and (iv) some serendipitous sources falling in the slits by chance. The mean redshift for the sample above $z = 2$ is $z = 3.0$. About 10% of the sample is above $z = 4$, and the high redshift tail goes up beyond redshift $z = 6$, the highest reliable redshift so far being $z = 6.5363$. The current redshift distribution for $z \geq 2$ in the COSMOS, ECDFS and VVDS-02h fields is shown in Figure 14. The redshift bin of $dz = 0.01$ enables to show the strong clustering present at all redshifts probed by the survey. Some of the densest peaks are remarkable examples of clustering in the early universe, as discussed in Cucciati et al. (2014) and Lemaux et al. (2014b).

As our sample is z_{phot} selected, it is interesting to check were the VUDS galaxies are distributed in several standard colour-colour diagrams. Following other studies (e.g. Kurk et al. 2013, Le Fèvre et al. 2013b), this *a posteriori* analysis gives those studies using colour-colour selection an indication of both the efficiency of the selection and the contamination by galaxies at other redshifts than the redshift of interest in the

colour space selected. We examine here the *gzK* diagram used to select galaxies with $1.4 < z < 2.5$, and the *ugr* diagram for galaxies with $2.5 < z < 3.5$. Other colour-colour diagrams like the *gri* or *riz* for $3.5 < z < 4.5$ and $4.5 < z < 5.5$, will be investigated in a forthcoming paper.

The BzK selection is based on the identification of the Balmer and D4000 break crossing the *z* band for $1.4 < z < 2.5$ (Daddi et al. 2004). We present the *gzK* diagram in Figure 15, with the colour-colour area adjusted to take into account the different wavelength coverage of the *g* filter compared to the B filter (Bielby et al. 2012, Le Fèvre et al. 2013b). The *gzK* criteria are efficient to select 92% of the VUDS galaxies with $K \leq 24$. However, we note that the level of contamination of a *gzK* selected sample down to $K_{AB} = 24$ would be quite high as 58% of galaxies in the selection area of the *gzK* diagram selecting the $1.4 < z < 2.5$ redshift range would be outside this range at $z < 1.4$ or $z > 2.5$ (right panel of Figure 15); at $K_{AB} = 22$ we find this contamination to be lower, at the ∼ 30% level. This trend is similar to that found by Le Fèvre et al. (2013b), who identified a magnitude-dependent contamination level. As the VUDS data is selected with $z_{phot} > 2.4$ the contamination level is further enhanced compared to the pure $i - band$ magnitude selection of the VVDS.

Going to higher redshifts, we show the distribution of VUDS galaxies with $2.5 < z < 3.5$ in the *ugr* colour-colour diagram in Figure 16. Here again a large fraction ∼ 80% of galaxies appear in the expected locus of the colour-colour diagram. The contamination by galaxies at $z < 2.5$ or $z > 3.5$ is quite high at about 40%. We note that we did not attempt to optimize the redshift range on the basis of the exact shape of the photometric bands used for the CFHTLS photometry used in this plot, but we rather elected to show the distribution of a large population.

This a posteriori colour-colour analysis shows that the VUDS sample is behaving as generally expected for galaxies at these redshifts. The photometric redshift selection allows to identify galaxies beyond the classical colour-colour locus of the *gzK* and *ugr* diagrams. We also point out the strong contamination present in colour-colour selected samples from galaxies at other redshifts outside the targeted redshift range.

4.2. Absolute magnitudes, stellar masses and star formation rates

The distribution of apparent magnitudes with redshift is shown in Figure 17. In all ∼ 90% of the distribution is within $23 \leq i_{AB} \leq 25$, while ∼ 10% have $25 \leq i_{AB} \leq 27$.

The VUDS sample covers a large range of galaxy physical properties. Using the VUDS spectroscopic redshifts we perform SED fitting on the multi-wavelength photometry using the code Le Phare (Ilbert et al. 2006), as described in Ilbert et al. (2013). We summarize here a few key points, but we refer to the recent description of the fitting process in Ilbert et al. (2013) for a detailed account and associated limitations. Galaxy luminosities are transformed into stellar mass using the best fit synthetic template, from a list of templates built from Bruzual & Charlot (2003) stellar population synthesis models with 3 metallicities ($Z = 0.004$, $Z = 0.008$, and solar $Z = 0.02$), and exponentially declining and delayed SFR with 9 different τ values from 0.1 to 30 Gyr. We use a Calzetti (2000) extinction law, and emission lines are added to the synthetic spectra as described in Ilbert et al. (2009). The output of this fitting process includes among other parameters: absolute magnitudes integrated into standard bands, stellar masses, star formation rates, and extinction. The distributions in absolute NUV (2300Å rest) magnitude, stellar masses, and star formation rate of the sample are presented in Figure 18. At redshifts $z \sim 3 - 4$ the NUV absolute magnitude ranges from $NUV = -20.5$ to $NUV = -23$, the stellar mass from $10^9 M_\odot$ to $10^{11} M_\odot$, and the star formation rate from below $1 M_\odot$/yr up to several hundred $M_\odot$/yr.

The VUDS survey therefore covers a large parameter space both in the observed properties and the physical properties of the sample galaxies.

4.3. Average spectral properties: stacked spectra

Stacked spectra have been produced from previous surveys, like Shapley et al. (2003) for $z \sim 3$, or for B, V and R dropout galaxies (Vanzella et al. 2009). The large statistics and wavelength coverage of VUDS offers the opportunity to produce composite spectra over a large redshift coverage. The average spectral properties of galaxies over the redshift range $2 \leq z_{spec} \leq 6.5$ are derived from stacking VUDS spectra in different redshift bins. For each redshift bin, the average spectra are produced using the `odcombine` task in IRAF, averaging spectra after scaling to the same median continuum value, i.e. luminosity weighted, and weighting spectra to their mean flux in the same rest-frame wavelength range. Average spectra using sigma clipping, removing at each wavelength data with a value 1.5 – 3 times the 1σ value, were compared to the straight weighted average described above, and very little difference was observed when the number of spectra is large (> 50). For smaller samples, sigma clipping helps improving the S/N of the stacks by removing e.g. the left-over signatures of the sky subtraction process.

The average spectra of all galaxies for several redshift bins are shown in Figures 19 to 23, together with the average spectra of those with and without Ly-α emission. We discuss some of the key features below.

We find that the fraction of galaxies with any trace of Lyα emission is strongly changing with redshift with $30.6 \pm 1.8\%$ in $2 \leq z \leq 3$, $38.3 \pm 2.9\%$ in $3 \leq z \leq 4$, $61.6 \pm 7.0\%$ in $4 \leq z \leq 4.7$, and $66.6 \pm 14.2\%$ in $4.7 \leq z \leq 5.3$. The detailed properties of the Lyα emitting fraction and implications are examined in details in Cassata et al. (2015).

Besides Lyα a number of spectral features are noteworthy, as indicated in Table 4. The main absorption features redder than Lyα include the SiIIλ1260, OIλ1303, CIIλ1334, SiIVλ1394/1403, SiIIλ1527, CIVλ1549 lines. In emission, weak CIII]λ1909 is quite common, and one can identify S-shape absorption-emission features for e.g. SiIIλ1260, OIλ1303 or CIVλ1549, indicative of strong outflows. The HeIIλ1640 emission is ubiquitous, as seen on all spectra where this line is in the wavelength range (Figures 19 to 21). This line, when several hundreds of km/s in width, indicates the presence of strong winds around Wolf-Rayet stars. If narrow it may indicate some other processes like the presence of a population of low-metallicity stars with properties akin to Population III (Cassata et al. 2013).

Below the Lyα line, the main features identified in galaxies at these redshifts are SiIIλ1192, and then the Lyβ and Lyγ lines, followed by the Lyman-limit producing a continuum break at

912Å. It is interesting to note that the flux below 912Å in our stacked spectra is not zero as would be expected if the lyman continuum escape fraction was 0%. We observe a significant detection of flux below the Lyman-limit in the stacked spectra of galaxies at $3 < z < 4$, $4 < z < 4.7$ and $4.7 < z < 5.3$. The ratio of continuum flux density in the [1400,1500]Å range over the flux in [800,900]Å range below the 912Å Lyman-limit is $f_{1500}/f_{900} = 32 \pm 3, 39 \pm 5$, and 33 ± 10, in these three redshift bins respectively. Although these measurements are based on a much larger sample covering a larger redshift range, these values are comparable to the observed value of 58 ± 18 reported by Shapley et al. (2006) at $z \sim 3$. While it is tempting to interpret this in terms of the Lyman-continuum escape fraction, we note that at the faint magnitudes we are observing, the flux observed below 912Å in stacked spectra could also be coming in part from objects contaminated by lower redshift interlopers producing observed flux below that of the Lyman-limit at the rest-frame of the distant source (e.g. Vanzella et al. 2010). The *observed* non-zero flux below 912Å has consequences on the selection of high redshift galaxies based on the Lyman-break technique, as the expected colours of high redshift galaxies are affected by a priori hypotheses, e.g. on the f_{1500}/f_{900} ratio (see e.g. Cooke et al. 2014). This and the corrected Lyman escape fraction will be analysed in forthcoming papers.

Our stacked spectra beautifully show the evolution of the 'stair-case' pattern of IGM absorption as a function of redshift produced by continuum blanketing from the Lyman series of galaxies along the line of sight in the volume probed (Madau 1995). The comparison of the observed mean and distribution of IGM absorption properties as a function of redshift in VUDS to the models of Madau (1995) and Meiksin (2010) is extensively discussed in Thomas et al. (2014).

A detailed analysis of the spectral properties of stacked spectra and individual galaxies in the VUDS sample will be presented in forthcoming papers.

5. Comparison with other spectroscopic surveys at high redshifts

Large spectroscopic surveys at $z > 2$ published in the literature are understandably relatively limited compared to surveys at $z < 1.5$ because of the faintness of the sources and requirements on deep imaging to perform pre-selection.

We compile a list, likely not exhaustive, of the main surveys in the redshift range $2 < z < 7$ in Table 5, and we show the corresponding redshift distribution of these surveys in Figure 24. The largest numbers of galaxies can be found in the range $2 < z < 3$ with the LBG-selected surveys of Steidel et al. (1999, 2003, 2004), the zCOSMOS-Deep survey (Lilly et al. 2007, Lilly et al. in prep.), the VLRS (Bielby et al. 2013), and the VVDS-Deep/Ultra-Deep (Le Fèvre et al. 2013a). In the redshift range $2.5 < z < 3.5$, these surveys accumulate ∼ 2700 galaxies with spectroscopic redshifts. The VUDS survey contributes ∼ 2800 spectroscopic redshifts, therefore as much as all these other surveys combined, but in a single well-controled survey.

Beyond redshift $z \sim 3.5$ both the number of spectroscopic surveys and the number of galaxies with measured spectroscopic redshifts dramatically fall down. Spectroscopic campains on the GOODS-South area with VIMOS and FORS2 on the VLT have produced about 114 galaxies with redshifts $z \sim 3-6$. The Keck-DEIMOS surveys of Stark et al. 2010 contribute more than 300 galaxies in the range $3 < z < 6.5$. VUDS brings more than 800 new galaxies with spectroscopic redshifts in the range $3.5 < z < 5$, or about twice the number of all other surveys combined as can be seen in Figure 24.

At the highest redshifts $z > 5$, we enter the game of single redshift confirmations following-up on galaxies identified as *dropout*, i.e. with a sharp change in flux between two adjacent photometric bands (e.g. Stark et al. 2010, Curtis-Lake et al. 2012), following narrow-band LAE candidates (e.g. Ouchi et al. 2010), or serendipitous identification of LAE (Cassata et al. 2011). These different surveys are bringing several tens of spectroscopic identifications beyond $z = 5$. VUDS contributes so far about 60 new sources with spectroscopic redshifts $z > 5$. This sample is being consolidated, particularly at $z > 6$, and will be the subject of future studies.

An important element to keep in perspective is the area sampled by a survey. Fluctuations in number density of objects observed in deep surveys result from the cosmic variance in the distribution of galaxies in large scale structure. The expected cosmic variance for a given area, redshift, and depth, can be computed e.g. following the prescription of Moster et al. (2011). Surveys with one square degree at $z \sim 3$ are subject to an uncertainty of about 10% from cosmic variance, in addition to the uncertainties related to the number of objects in a survey. For fields of the GOODS size (∼ 150arcmin^2), the cosmic variance at $z \sim 3$ is expected to be in excess of 50% for galaxies with stellar masses $M_* > 10^{10.5} M_\odot$.

We compare in Figure 25 the number of spectra per square degree vs. the area for different surveys, a high redshift version of that presented in Baldry et al. (2010) and updated by Le Fèvre et al. (2013a). While at $z \sim 3$, several surveys have surveyed about 1 square degree (Steidel et al. 2003, Lilly et al. 2007, Bieby et al. 2013, or Le Fèvre et al. 2013a), the highest redshift surveys at $z > 3$ have surveyed only about 0.1 deg^2 (e.g. Vanzella et al. 2009, Stark et al. 2010). In this context, VUDS is a unique survey with selected galaxies at any redshift $z > 2$ in three different fields for a total of ∼ 1deg^2.

6. Summary

The VIMOS Ultra Deep Survey (VUDS) is a deep spectroscopic redshift survey aiming to study the early phases of galaxy assembly at $2 < z < 6.5$ from a sample of ∼ 10 000 galaxies observed with the VIMOS multi-slit spectrograph at the ESO-VLT. The survey target selection is based on photometric redshifts derived from extensive multi-wavelength data for 87.7% of the sample, supplemented by colour-colour and SED selection for 12.3%. Most of the sample is limited down to $i_{AB} = 25$, but galaxies are observed as faint as $i_{AB} = 27$.

The combination of a wide wavelength coverage from 3650Å to 9350Å, and exposure times of ≃ 14h, lead to a spectroscopic success rate in redshift measurement of about 91% (74% for spectroscopic reliability flags 2 to 9) down to $i_{AB} = 25$. The comparison of photometric redshifts to the VUDS spectroscopic redshifts shows that the VUDS strategy minimizes the loss of galaxy populations compared to more restrictive selection criteria.

We report on the general properties of the sample based on ∼ 80% of the data which has already been processed. The redshift distribution of the current sample at $z \geq 2$ peaks at a

mean $z = 3$, and extends beyond $z = 6$. A secondary sample at $z < 2$ is the result of the selection function and provides interesting very low intrinsic luminosity galaxies. The average spectral properties of galaxies with $z > 2$ are discussed based on high S/N stacks of VUDS spectra in several increasing redshift bins. Galaxies with and without Lyα in emission are found at any redshift, but it is found that the fraction of galaxies with Lyα in emission increases with redshift, as is further quantified in Cassata et al. (2015). Using stacked spectra we find that there is observed flux below the 912Å Lyman limit in all the redshift ranges explored, the origin of this is being investigated and will be the subject of future papers.

Following an early measurement of the merger rate at $z \sim 3$ (Tasca et al. 2013), several papers presenting results from the VUDS are published (Cucciati et al. 2014, Cassata et al. 2015, Lemaux et al. 2014b, Amorin et al. 2014) or submitted (Durkalec et al. 2014, Tasca et al., 2014, Thomas et al., 2014). A number of other analyses are in progress.

VUDS is the first deep spectroscopic survey covering such a large redshift range with such a large sample of galaxies with confirmed spectroscopic redshifts. It is ideally suited for detailed studies of the galaxy population at early times $2 < z < 6$. When the full data set will be completed, it is foreseen to make VUDS data releases publicly available.

Acknowledgements. We thank ESO staff for their continuous support for the VUDS survey, particularly the Paranal staff conducting the observations and Marina Rejkuba and the ESO user support group in Garching. This work is supported by funding from the European Research Council Advanced Grant ERC-2010-AdG-268107-EARLY and by INAF Grants PRIN 2010, PRIN 2012 and PICS 2013. AC, OC, MT and VS acknowledge the grant MIUR PRIN 2010–2011. DM gratefully acknowledges LAM hospitality during the initial phases of the project. This work is based on data products made available at the CESAM data center, Laboratoire d'Astrophysique de Marseille. This work partly uses observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

References

Abazajian et al., 2009, ApJS, 182, 543
Amorin, R., et al., 2014, A&A, 568, 8
Arnouts, S., et al., 2007, A&A, 476, 137
Ashby, M.L.N., et al., 2013, ApJ, 769, 80
Baldry, I., et al., 2010, MNRAS, 404, 86
Bertin, E., Arnouts, S., 1996, A&AS, 117, 393
Bielby, R., et al., 2012, A&A, 545, 23
Bielby, R., et al., 2013, MNRAS, 430, 425
Bondi, M., et al., 2003, A&A, 403, 857
Bouwens, R. et al., 2009, ApJ, 705, 936
Bouwens, R. et al., 2014, arXiv:1403.4295
Bruzual, G., & Charlot, S., 2003, MNRAS, 344, 1000
Calzetti, D., et al., 2000, ApJ, 533, 682
Capak, P., et al., 2011, Nature, 470, 233
Caputi, K., et al., 2012, ApJ, 750, 20
Cardamone, C., et al., 2010, ApJS, 189, 270
Cassata, P., Le Fèvre, O., et al., 2011, A&A, 525, 143
Cassata, P., Le Fèvre, O., et al., 2013, A&A, 556, 68
Cassata, P., Le Fèvre, O., et al., 2015, A&A, 573, in press
Cimatti, A., et al., 2002, A&A, 381, 68
Cimatti, A., et al., 2008, A&A, 482, 21
Colless, M., et al., 2001, MNRAS, 328, 1039
Cooke, J., Ryan-Weber, E. V., Garel, T., Daz, C. G., 2014, MNRAS, 411, 837
Cooper, M.C., et al., 2012, MNRAS, 425, 2116
Cucciati, O., et al., 2012, A&A, 539, 31
Cucciati, O., et al., 2014, A&A, 570, 16
Cuillandre, J.C., et al., 2012, Proc. SPIE 8448, Observatory Operations: Strategies, Processes, and Systems IV, 84480
Curtis-Lake, E., et al., 2012, MNRAS, 422, 1425
Daddi, E., et al., 2004, ApJ, 617, 746
Davis, M., et al., 2003, SPIE, 4834, 161
Durkalec, A., et al., submitted, arXiv:1411.5688
Ellis, R.S., et al. 2013, ApJ, 763, L7
Faber, S. M., et al. 2003, Proc. SPIE, 4841, 1657
Garilli, B., et al., 2010, PASP, 122, 827
Giacconi, R., et al., 2002, ApJS, 139, 369
Grogin, N., et al., 2011, ApJS, 197, 35
Guo, Q., et al., 2011, MNRAS, 413, 101
Guzzo, L., et al., 2014, arXiv:1303.2623
Hammersley, P. et al., 2010, The Messenger, 142, 8
Ilbert, O., et al., 2006, A&A, 457, 841
Ilbert, O., et al., 2009, ApJ, 690, 1236
Ilbert, O., et al., 2013, A&A, 556, 55
Kashikawa, N., et al., 2011, ApJ, 734, 119
Koekemoer, A., et al., 2007, ApJS, 172, 196
Koekemoer, A., et al., 2011, ApJS, 197, 36
Le Fèvre, O., Crampton, D., Lilly, S.J., Hammer, F., Tresse, L., 1995, ApJ, 455, 60
Le Fèvre, O., et al., 2003, SPIE, 4841, 1670
Le Fèvre, O., et al., 2004, A&A, 417, 839
Le Fèvre, O., et al., 2005a, A&A, 439, 845
Le Fèvre, O., et al., 2005b, Nature, 437, 519
Le Fèvre, O., et al., 2013a, A&A, 559, 14
Le Fèvre, O., et al., 2013b, A&A, arXiv:1307.6518
Lemaux, B., et al., 2014a, A&A, 572, 90
Lemaux, B., et al., 2014b, A&A, 572, 41
Lilly, S.J., Le Fèvre, O., Crampton, D., Hammer, F., 1995a, ApJ, 455, 50
Lilly, S. J., Le Fèvre, O., Hammer, F., Crampton, D., 1996, ApJ, 460, 1
Lilly, S.J., Le Fèvre, O., et al., 2007, ApJS, 172, 70
Lonsdale, C.C., et al., 2003, PASP, 115, 897
Madau, P., 1995, ApJ, 441, 18
Madau, P., Ferguson, H.C., Dickinson, M., Giavalisco, M., Steidel, C.C., Fruchter, A., 1996, MNRAS, 283, 1388
Madau, P., & Dickinson, M., 2014, ARAA, in press (arXiv:1403.0007)
Mauduit et al., 2012, PASP, 124, 714
McCracken, H.J., et al., 2012, A&A,
Mo, H., van den Bosch, F., White, S., Galaxy Formation and Evolution, Cambridge University Press, 2010
Moster, B.P., Somerville, R.S., Newman, J.A., Rix, H-W., 2011, ApJ, 731, 113
Oliver, S., et al., 2012, MNRAS, 424, 1614
Ouchi, M., et al., 2008, ApJS, 176, 301
Patat, N., et al., 2011, A&A, 527, 91
Pierre, M., et al., 2004, JCAP, 09, 011
Popesso, et al., 2009, A&A, 494, 443
Rix, H.W., et al., 2004, 2004, ApJS, 152, 163
Sanders, D., et al., 2007, ApJS, 172, 86
Schlegel, D., et al., 1998, ApJ, 500, 525
Scodeggio, M., et al., 2005, PASP, 117, 1284
Scodeggio, M., et al., 2009, The Messenger, 135, 13
Scoville, N., et al., 2007, ApJS, 172, 1
Shapley, A., et al., 2006, ApJ, 651, 688
Shimasaku, K., 2006, PASJ, 5, 313
Springel, V. et al., 2008, MNRAS, 391, 1685
Stark, D., et al., 2010, MNRAS, 408, 1628
Steidel, C.C., Giavalisco, M., Pettini, M., Dickinson, M., Adelberger, K.L., 1996, ApJ, 462, 17
Steidel, C.C., Adelberger, K.L., Giavalisco, M., Dickinson, M., Pettini, M., 1999, ApJ, 519, 1
Steidel, C.C., et al., 2003, ApJ, 592, 728
Steidel, C.C., et al., 2004, ApJ, 604, 534
Taniguchi, Y., et al., 2005, PASJ, 57, 165
Taniguchi, Y., et al., 2007, ApJS, 172, 9
Tasca, L., et al., 2014, A&A, 565, 10
Tasca, L., et al., 2014, submitted, arXiv:1411.5687
Thomas, R., et al., 2014, submitted, arXiv:1411.5692
Tresse, L., et al., 2007, A&A, 472, 403
Vanzella, E., et al., 2008, A&A, 478, 83

Vanzella, E., et al., 2009, ApJ, 695, 1163
Vanzella, E., et al., 2010, MNRAS, 404, 1672
Wiklind, T. et al., 2008, ApJ, 676, 781
Williams, R., et al., 2000, AJ, 120, 2735
Windhorst, R., et al., 2011, ApJS, 193, 27

Table 1. VUDS fields

Field	α_{2000}	δ_{2000}	b	l	Area	*Depth*
COSMOS	$10h00m04.0s$	+02 deg 12′40″	42.1	236.8	1800 arcmin^2	$i_{AB} \simeq 25$
ECDFS	$03h32m28.0s$	−27 deg 48′30″	−54.0	223.5	675 arcmin^2	$i_{AB} \simeq 25$
VVDS-02h	$02h26m00.0s$	−04 deg 30′00″	−57.5	172.0	1125 arcmin^2	$i_{AB} \simeq 25$

Table 2. VUDS spectroscopic target selection criteria

Criterion	Value	Limiting magnitude	Fraction of targets
Photometric redshift, 1st PDF peak	$z_{phot} + 1\sigma \geq 2.4$	$22.5 \leq i_{AB} \leq 25$	58.3%
Photometric redshift, 2nd PDF peak	$z_{phot} \geq 2.4$	$22.5 \leq i_{AB} \leq 25$	19.3%
LBG ugr	$2.7 < z < 3.5$	$i_{AB} \leq 25$	1.9%
LBG gri	$3.5 < z < 4.5$	$i_{AB} \leq 25$	1.9%
LBG riz	$4.5 < z < 5.5$	$i_{AB} \leq 25$	3.6%
Break SED-based	–	$K_{AB} \leq 24$	3.5%
Magnitude-selected	–	$23 \leq i_{AB} \leq 25$	11.5%

Table 3. VUDS: observed VIMOS fields

Pointing	α_{2000}	δ_{2000}	Grism
COSMOS-P01	$09h59m02.39s$	+01 deg 54′35.9″	LRBLUE & LRRED
COSMOS-P02	$10h00m04.08s$	+01 deg 54′35.9″	LRBLUE & LRRED
COSMOS-P03	$10h01m05.76s$	+01 deg 54′35.9″	LRBLUE & LRRED
COSMOS-P04	$09h59m02.39s$	+02 deg 12′41.4″	LRBLUE & LRRED
COSMOS-P05	$10h00m04.08s$	+02 deg 12′41.4″	LRBLUE & LRRED
COSMOS-P06	$10h01m05.76s$	+02 deg 12′41.4″	LRBLUE & LRRED
COSMOS-P07	$10h00m04.08s$	+02 deg 30′46.7″	LRBLUE & LRRED
COSMOS-P08	$10h01m05.76s$	+02 deg 30′46.7″	LRBLUE & LRRED
ECDFS-P01	$03h32m25.99s$	−27 deg 41′59.9″	LRBLUE & LRRED
ECDFS-P02	$03h32m34.00s$	−27 deg 53′59.9″	LRBLUE & LRRED
ECDFS-P03[a]	$03h32m15.00s$	−27 deg 49′59.9″	MR
VVDS02-P01	$02h26m44.51s$	−04 deg 16′42.8″	LRBLUE & LRRED
VVDS02-P02	$02h25m40.34s$	−04 deg 16′42.8″	LRBLUE & LRRED
VVDS02-P03	$02h26m44.51s$	−04 deg 34′50.3″	LRBLUE & LRRED
VVDS02-P04	$02h25m40.34s$	−04 deg 34′50.3″	LRBLUE & LRRED
VVDS02-P05	$02h24m36.14s$	−04 deg 44′57.9″	LRBLUE & LRRED

[a] The ECDFS-P03 has been observed with a VIMOS PA=70deg on the sky.

Table 4. Main spectral features observed in VUDS spectra

Spectral line	λ_{rest} (Å)	Line type
Lyman−limit	912.0	Continuum break
Lyman$-\gamma$	972.0	HI absorption
Lyman$-\beta$	1025.2	HI absorption
SiIIλ1192	1192.0	ISM, blend 1190+1193
Lyman$-\alpha$	1215.7	HI emission & absorption
SiIIλ1260	1260.4	ISM
OI+SiII-1303	1303.2	ISM, blend
CIIλ1334	1334.5	ISM
SiIVλ1394	1393.8	ISM
SiIVλ1403	1402.8	ISM
SiIIλ1527	1526.7	ISM
CIVλ1549	1549.1	ISM, blend 1548.2+1550.8
FeIIλ1608	1608.5	ISM
HeIIλ1640	1640.0	Nebular
AlIIλ1671	1670.8	ISM
FeIIλ1855	1854.7	ISM
FeIIλ1863	1862.8	ISM
CIII]λ1909	1908.7	Nebular, blend 1907+1909
FeIIλ2344	2343.5	ISM
FeIIλ2371	2370.5	ISM
FeIIλ2402	2402.6	ISM
FeIIλ2594	2593.7	ISM
MgIIλ2796	2796	ISM

Table 5. Comparison of the VUDS survey with other spectroscopic redshift surveys at $z > 2$ in the literature, by order of increasing mean redshift

Survey	*Area* deg^2	*Depth* i_{AB} *eq.*	N_{obj} *in Survey*	N_{obj} *at* $z \geq 2$	z_{range}	z_{mean}	*Selection*	*Reference*
VVDS-Deep	0.74	24.00	11 601	634	0 – 5	0.92	$17.5 \leq I_{AB} \leq 24.0$	Le Fèvre et al. 2013
VVDS-UDeep	0.14	24.75	941	341	0 – 4.5	1.38	$23.0 \leq i_{AB} \leq 24.75$	Le Fèvre et al. 2014
Steidel-z2	0.48	$R = 25.5$	851	588	1.4 – 2.5	2.0	$BM - BX$	Steidel et al. 2004
zCosmos-Deep	1	23.75	∼ 7 500	∼ 4 100	1.5 – 3	2.1	$B_{AB} \leq 25$+colour	Lilly et al. 2007, Lilly et al. in prep.
VLRS	1.62	24.7	2 135	2 135	2 – 3.5	2.8	$23 < R < 25$	Bielby et al. 2013
LBG-z3	0.38	24.8	1 000		2.7 – 3.5	3.2	$R_{AB} < 25.5$+colour	Steidel et al. 2003
GOODS	0.09	$z = 26$	887	114	3	3.5	$(i_{775} - z_{850}) > 0.6, z_{850} < 26$	Vanzella et al. 2009 and ref. therein
VVDS-LAE	0.74	–	217	217	0 – 6.7	3.5	$23.0 \leq i_{AB} \leq 24.75$	Cassata et al. 2011
LBG-z4	0.38	25.0	300	300	3.5 – 4.5	4.0	$I_{AB} < 25$colour	Steidel et al. 1999
Dropout-z456	0.09	–	310	310	3 – 6.5	4.5	Dropout, Lyα break	Stark et al. 2010
LAE-z6	1	∼ 27	16	16	∼ 6.5	6.5	LAE narrow band	Ouchi et al. 2010
VUDS	1	25 – 27	∼ 10 000	∼ 6 000[a]	2 – 6.7	3.0	$i_{AB} < 25$+photo-z	This paper

[a] Adjusted to take into account new data not yet processed

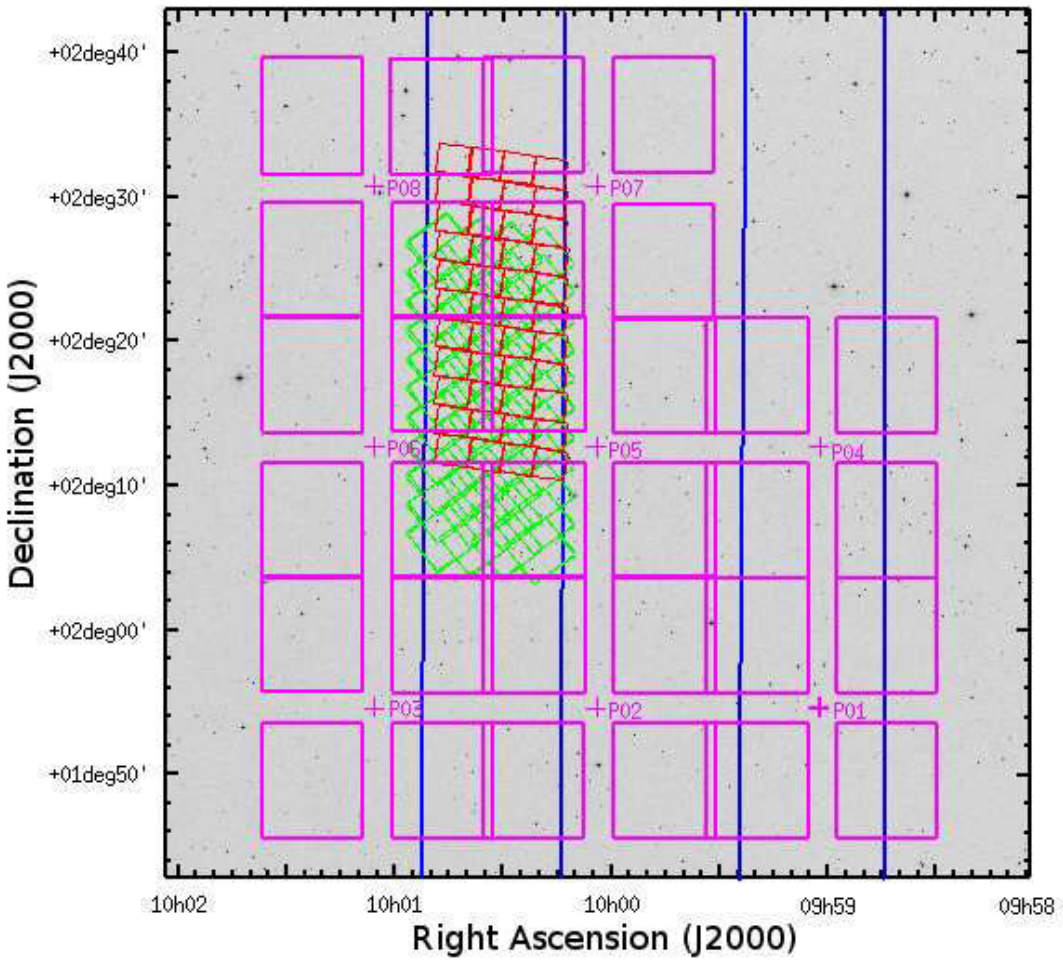


Fig. 1. Layout of the observed 8 VUDS VIMOS pointings in the COSMOS field. The center of each VIMOS pointing is identified by a cross with the pointing number (see Table 3), while the imprint of the 4 quadrants is in magenta. The positions of the UltraVista Deep stripes overlapping with the VUDS area are identified by the blue regions. The CANDELS ACS-F814W (in green) and WFC3-F160W (in red) areas are indicated. The size of the image is 1×1 deg^2.

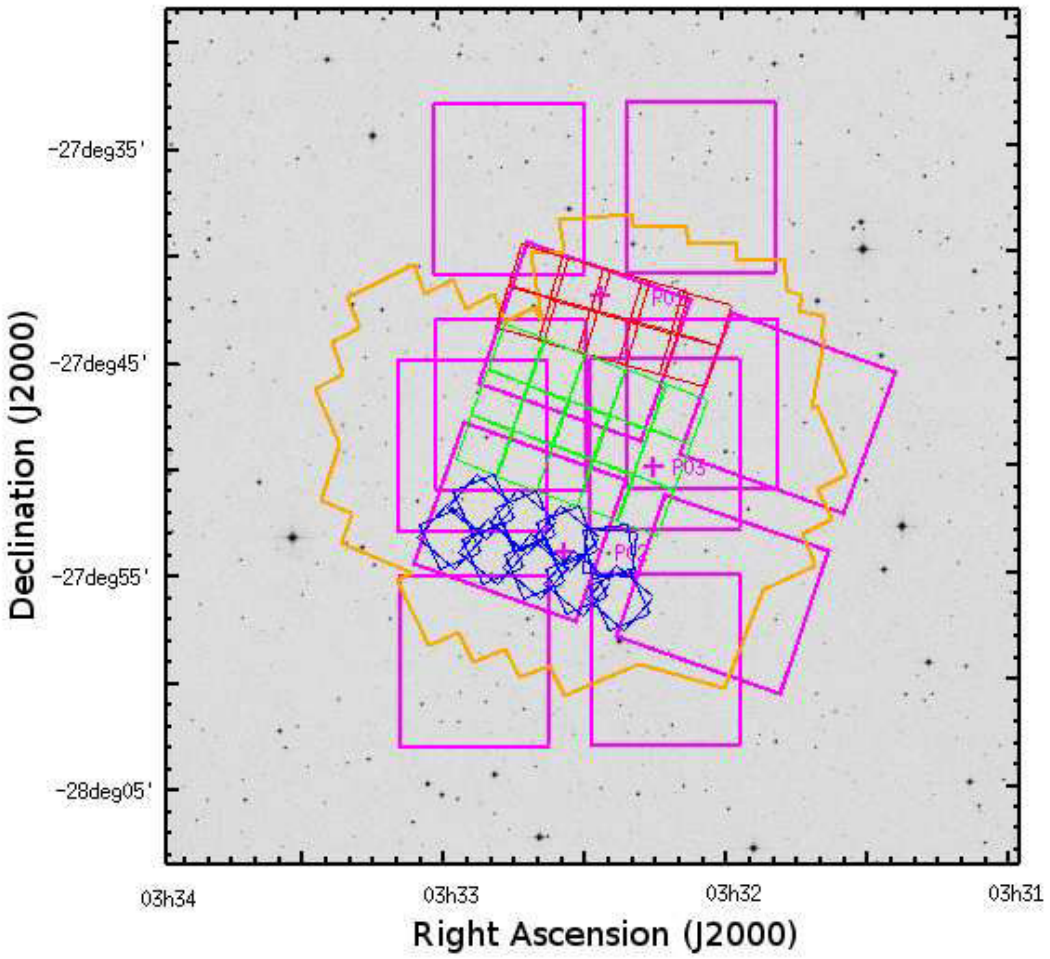


Fig. 2. Layout of the observed 3 VUDS VIMOS pointings in the ECDFS field. The center of each VIMOS pointing is identified by a cross with the pointing number (see Table 3), while the imprint of the 4 quadrants is in magenta. The positions of the CANDELS WFC3 deep area is indicated in green, and the CANDELS wide area in blue. The red region indicates the WFC3 coverage of the ERS. The outline of the existing ACS-F814W imaging is identified in orange. The size of the image is 0.67×0.67 deg^2.

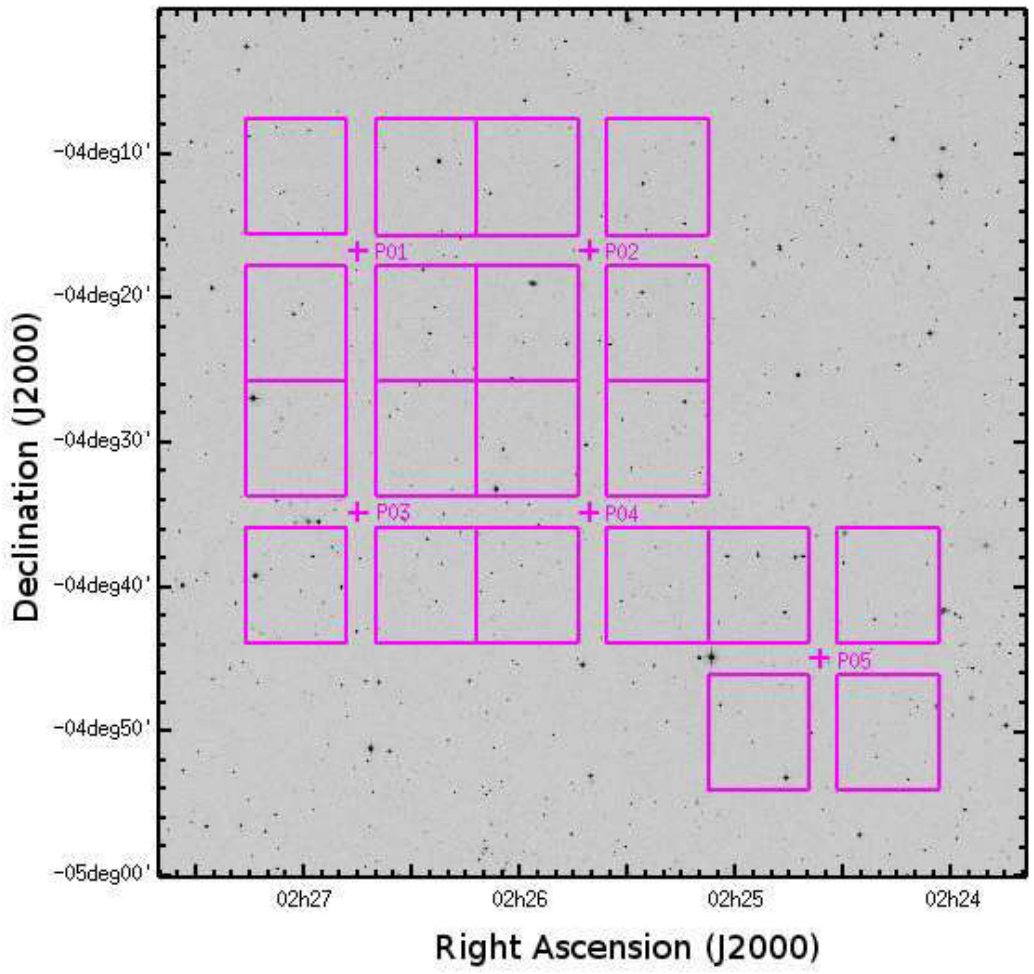


Fig. 3. Layout of the observed 5 VUDS VIMOS pointings in the VVDS-02h field. The center of each VIMOS pointing is identified by a cross with the pointing number (see Table 3), while the imprint of the 4 quadrants is in magenta. The size of the image is 1×1 deg^2. All of the VIMOS pointings are covered by CFHTLS visible photometry and WIRDS near infrared photometry.

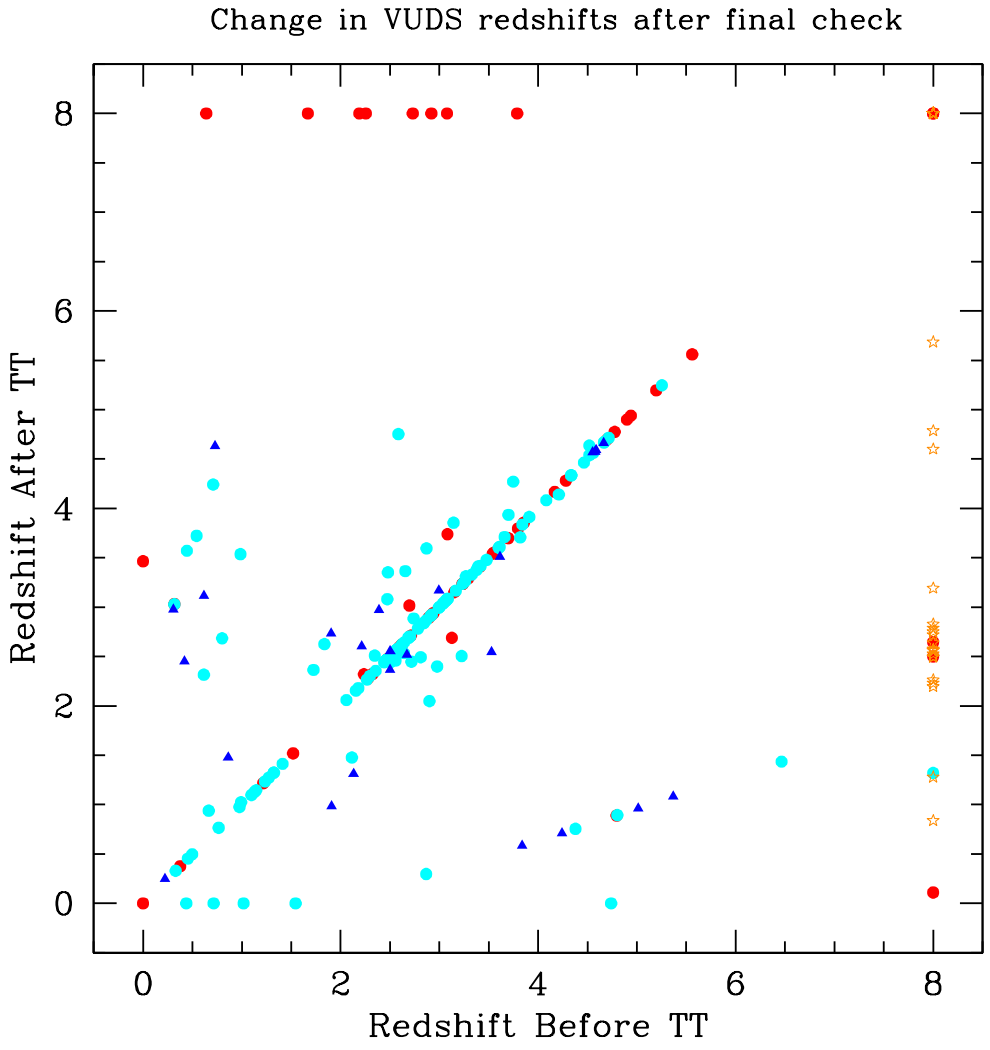


Fig. 4. Comparison between VUDS spectroscopic redshifts before and after the *Tiger Team* (TT) check of redshifts measured after a first measurement pass (see text). Only those objects with a change are considered here. Represented in this figure are about 10% of all objects which had a redshift change (6%) or a flag change (4%). Red circles represent objects for which the reliability flag was downgraded after the Tiger Team work; filled cyan circles are those objects with upgraded flags; and blue triangles are objects keeping the same flag. Objects which had undetermined redshifts (flag=0) before the TT work are placed at $z_{before} = 8$ and represented with orange starred symbols. Objects which have undetermined redshifts after the TT work (either keeping their original undetermined status or the TT work decided to downgrade them) are the red circles at $z_{after} = 8$.

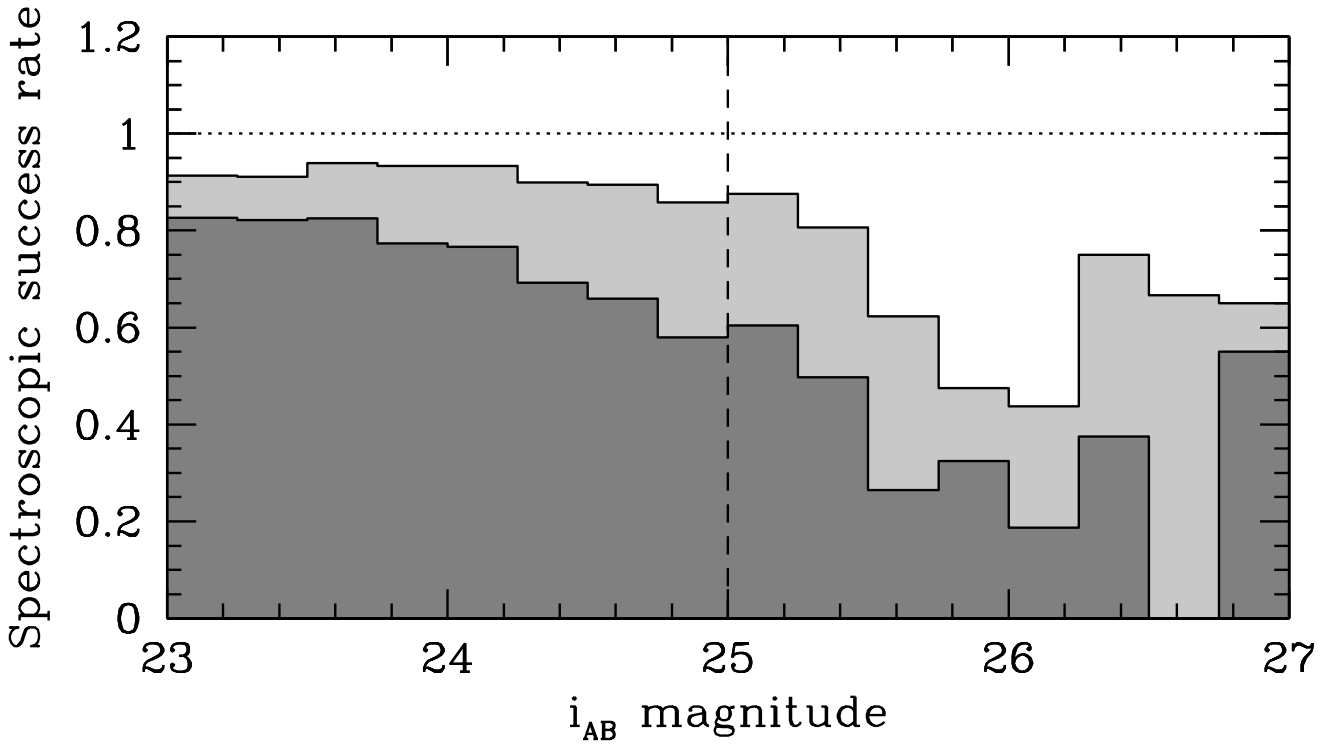


Fig. 5. Spectroscopic success rate versus i_{AB} magnitude, for all objects with a redshift measurement (light grey) and all objects with a > 75% reliable redshift measurement (flags 2, 3, 4, and 9; dark grey).

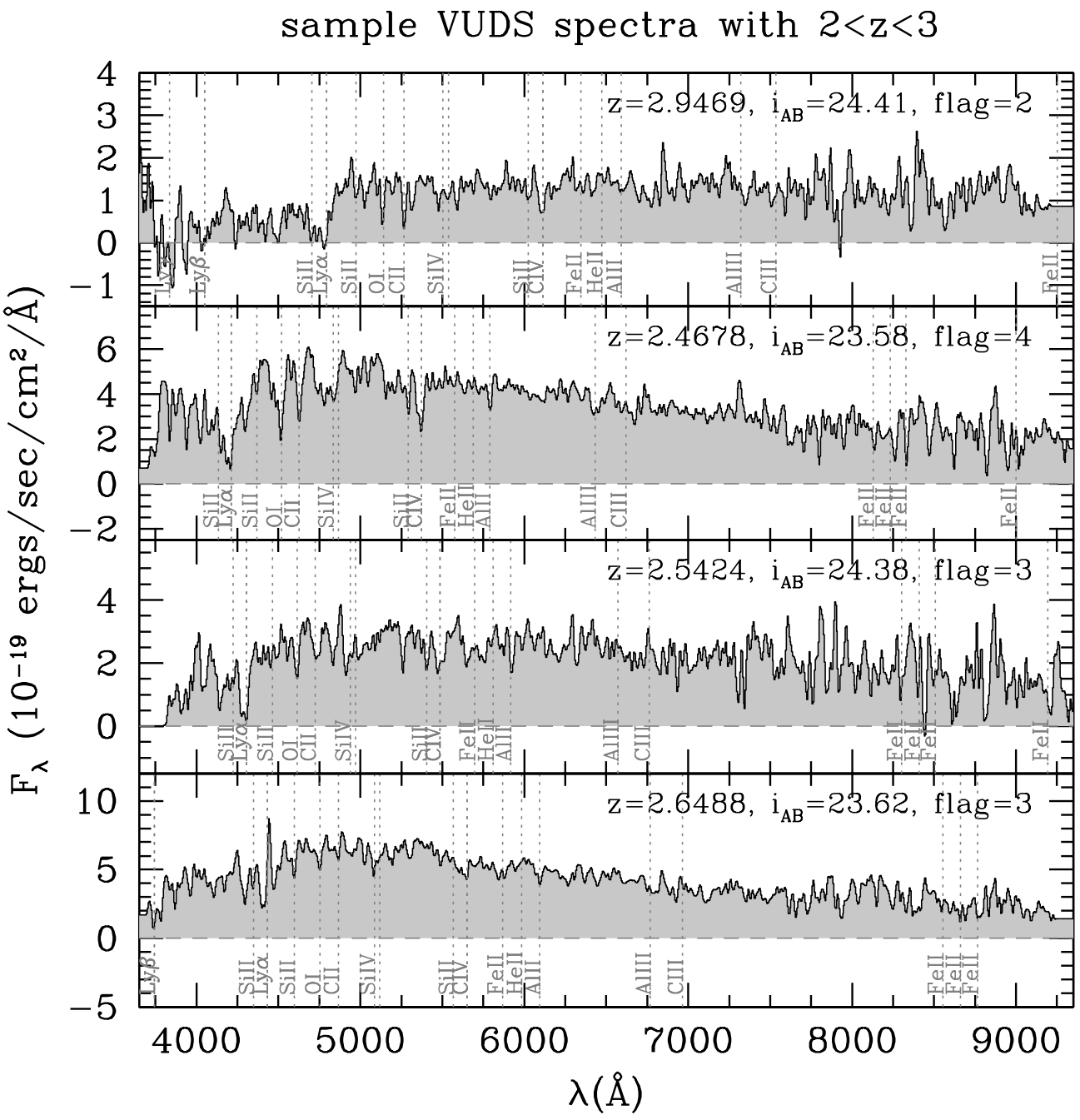


Fig. 6. Sample VUDS spectra in the range 2 < z < 3. In this range Lyα just enters the VIMOS bandpass, and spectral features can be identified up to Iron lines at ~2370Å. Most of the population does not have Lyα emission.

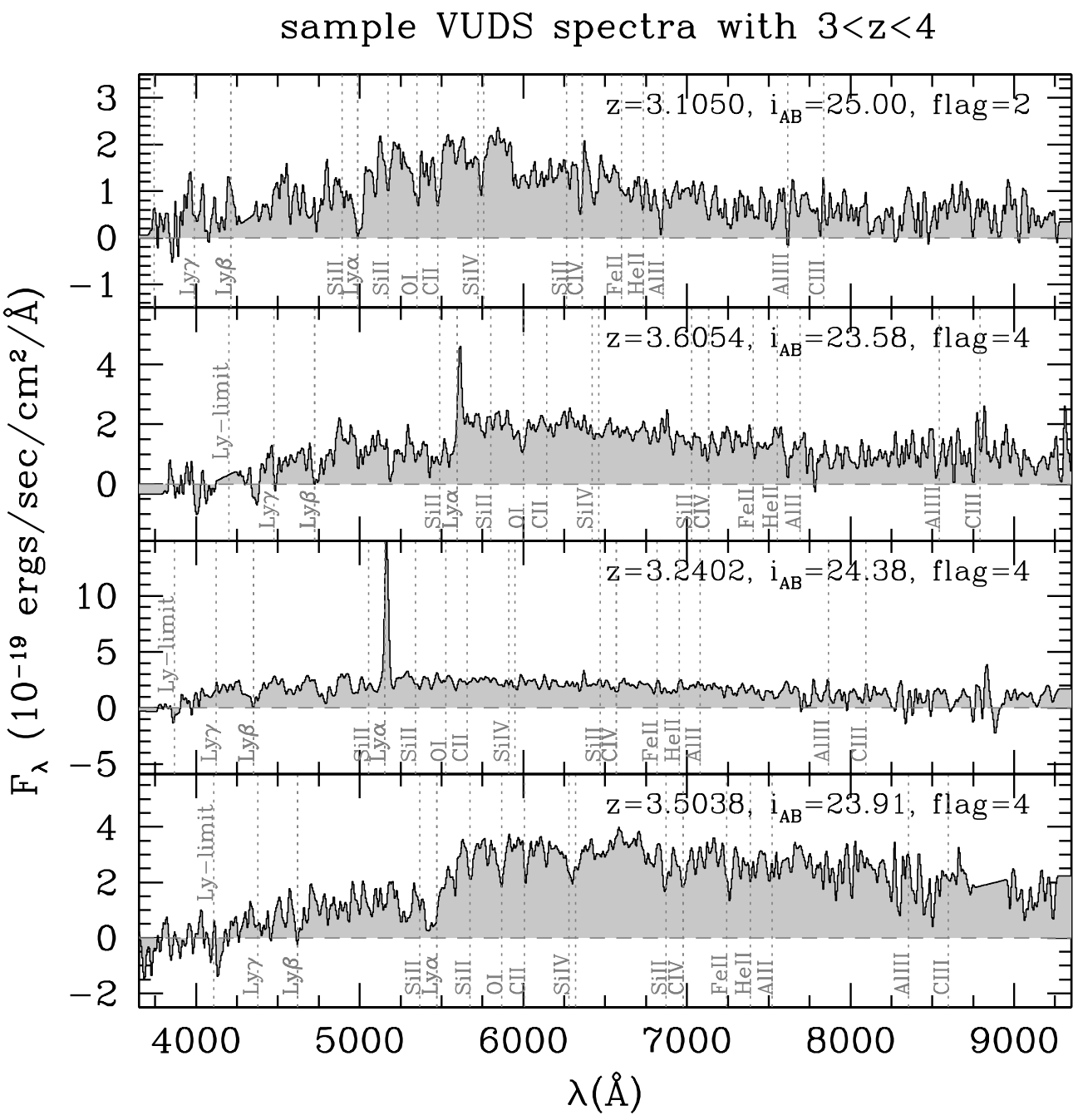


Fig. 7. Sample VUDS spectra in the range $3 < z < 4$. In the UV the Lyman limit at 912Å enters the VIMOS bandpass at z=3, and the VIMOS spectra are well sensitive to the IGM transmission. Further in the red spectral features can be identified up to CIII]. In this redshift range the fraction of Lyα emitters starts to rise but the dominant population is without Lyα emission.

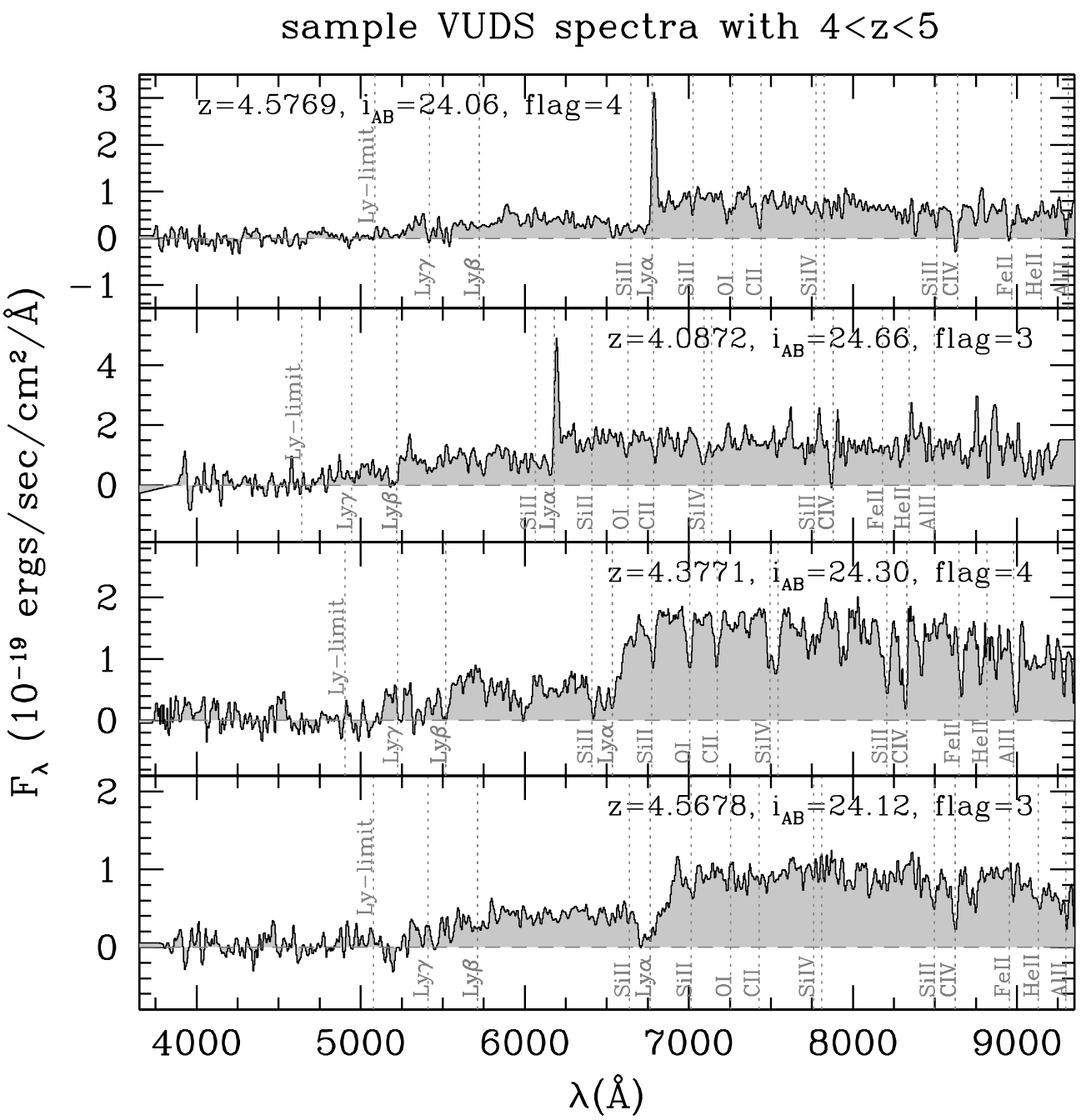


Fig. 8. Sample VUDS spectra in the range 4 < z < 5. The observed UV rest-frame covers from below the Lyman limit up to CIV-1549Å.

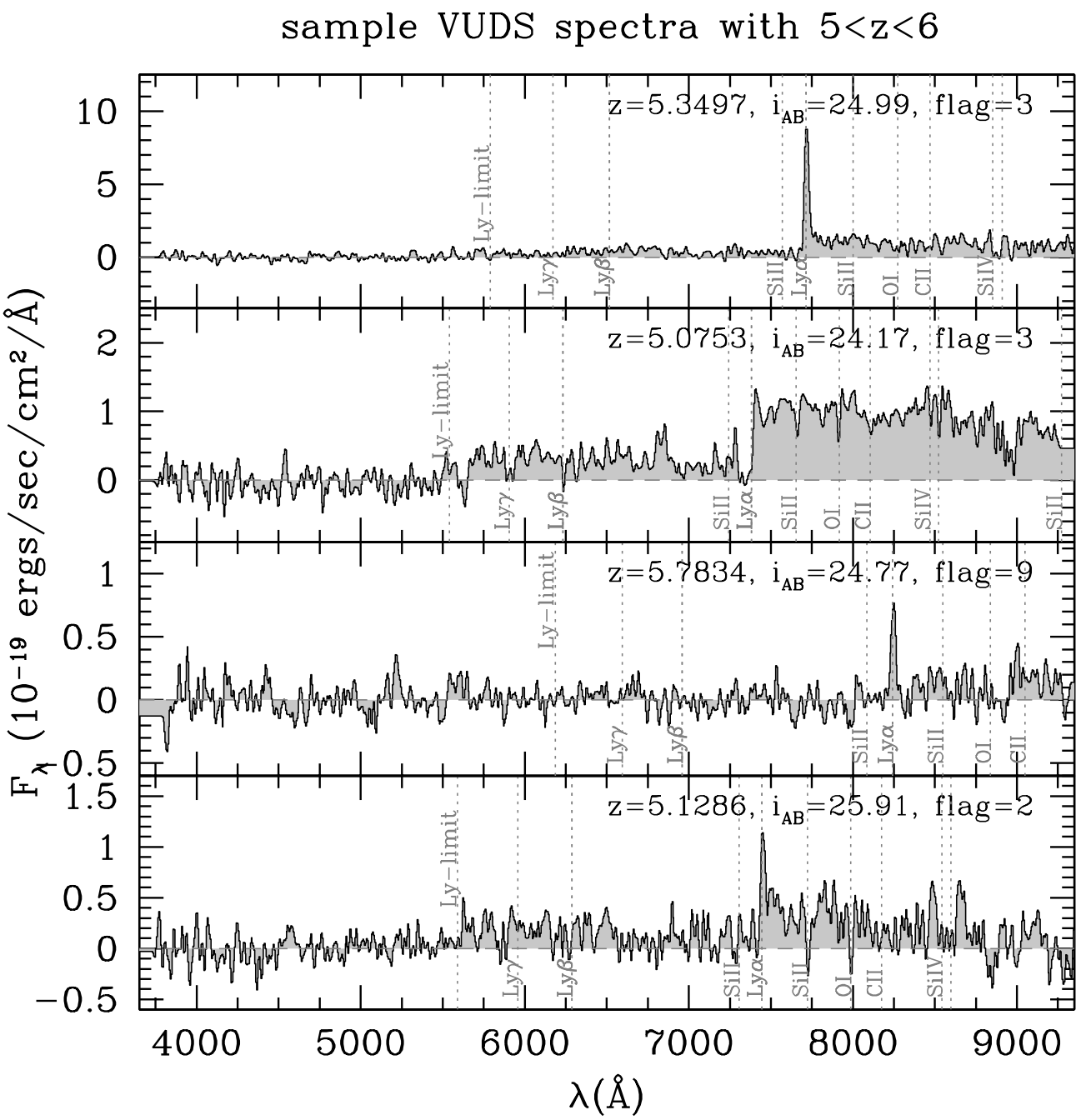


Fig. 9. Sample VUDS spectra in the range $5 < z < 6$. The IGM transmission becoming high the main spectral features used for redshift determination are the spectral break produced by the IGM below Lyα and the Lyα line itself. The number of galaxies with Lyα in emission is about equivalent to galaxies with Lyα absorption (see Cassata et al., 2014, for a detailed discussion).

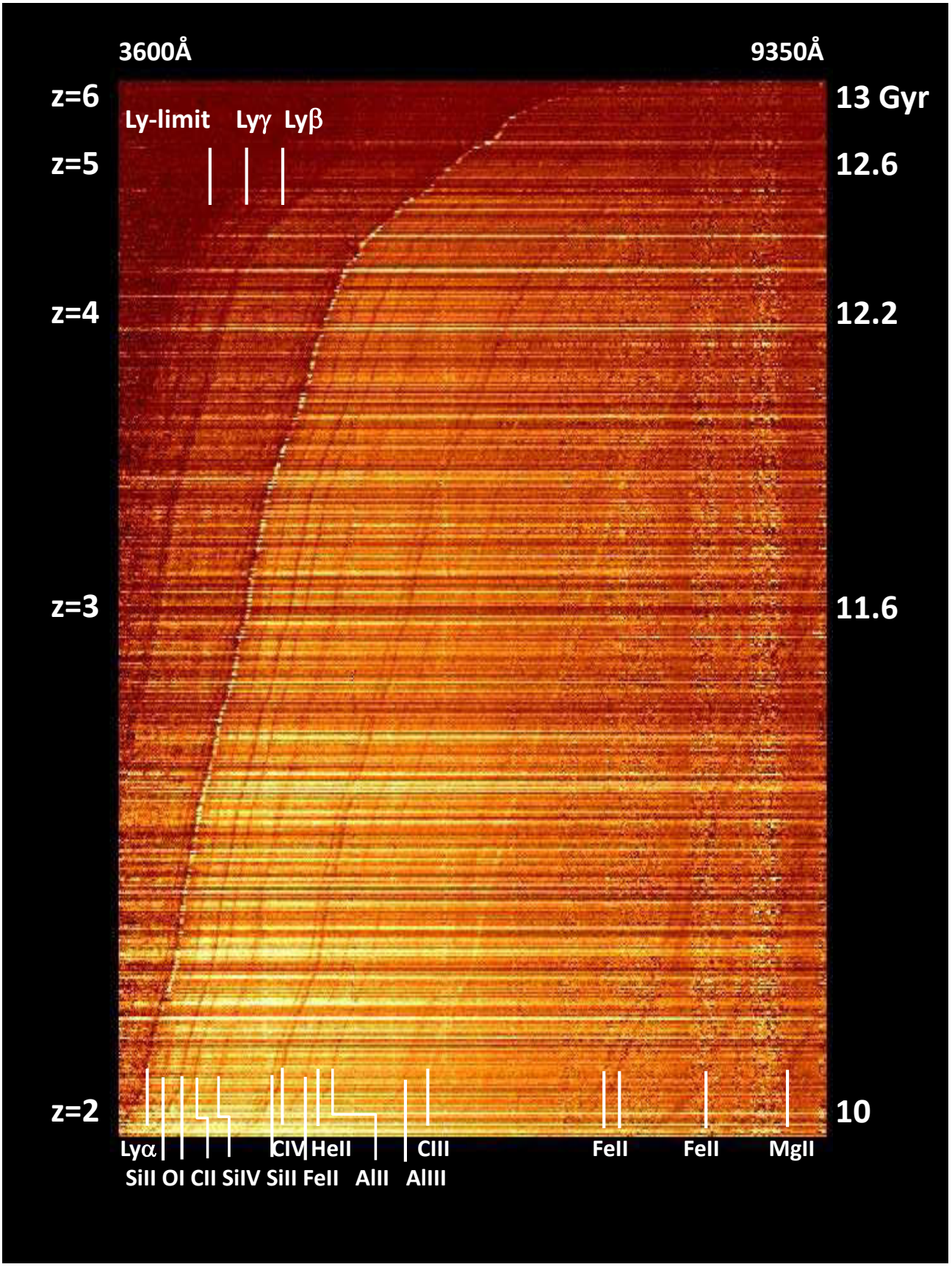


Fig. 10. Overview of the spectra obtained in the VUDS survey, with the wavelength going from 3650 to 9350Å on the X-axis. The redshift is increasing (in a non linear way) as indicated along the left Y-axis, with the corresponding look-back time indicated along the right Y-axis. The image is build with the spectra of all $2 < z < 4$ VUDS galaxies with flags 3 and 4, and all spectra with flags 2, 3 and 4 for $z > 4$, ordered one per image line by increasing redshift. All the main emission and absorption lines can be readily identified on this image, even faint ones, thanks to the increased contrast produced by the continuous display of spectra. The main spectral lines are identified on the top left (below Lyα) and at the bottom of the plot (above Lyα), as listed in Table 4. The vertical bands appearing at fixed wavelength in the red correspond to increased residual noise after subtraction of the strong atmospheric OH-bands.

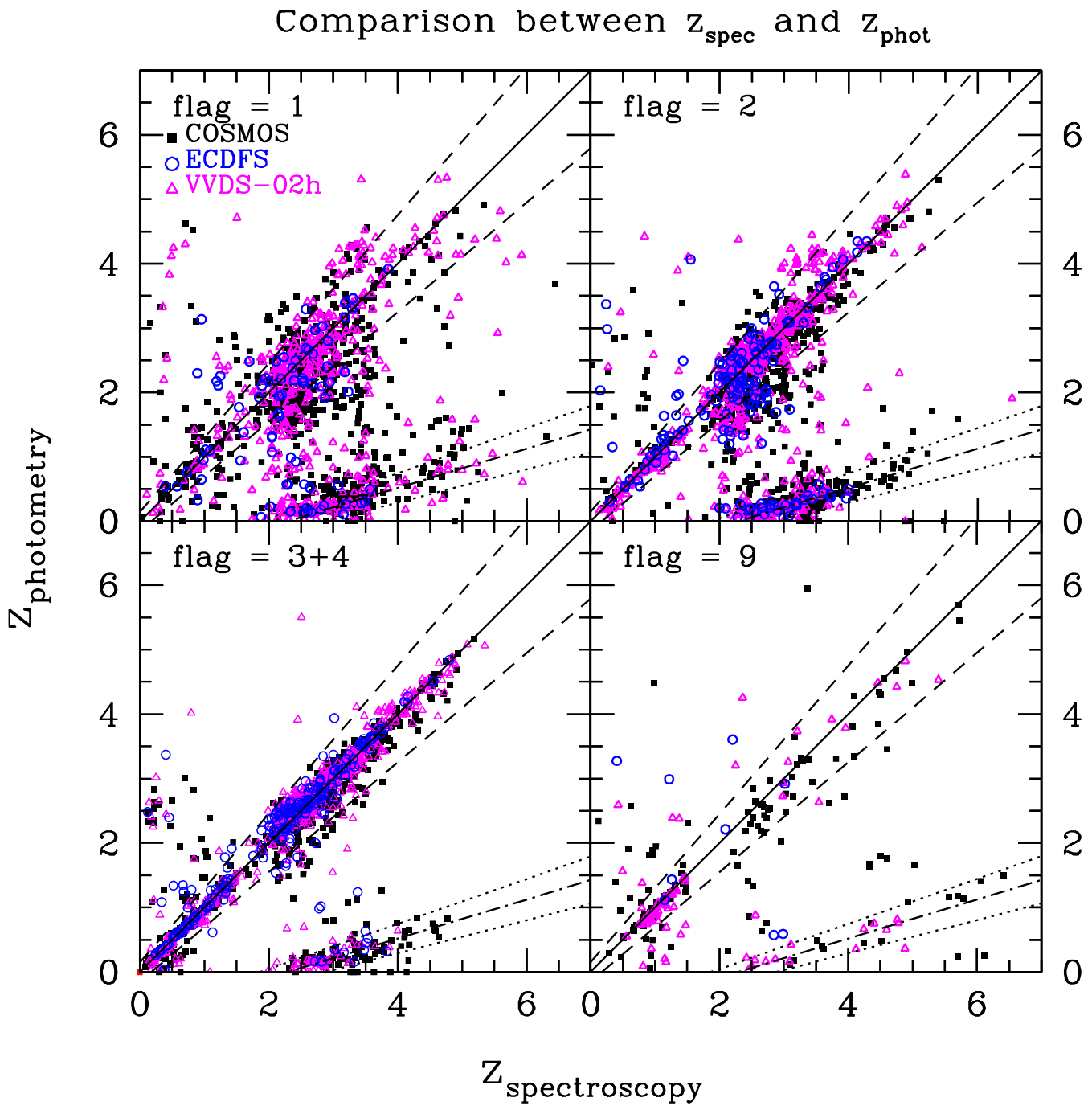


Fig. 11. A posteriori comparison between VUDS spectroscopic redshifts and the SED-derived photometric redshifts using the Le Phare code. (top-left): VUDS galaxies with spectroscopic flags 1 ($\simeq 50-70\%$ reliable), (top-right): VUDS galaxies with spectroscopic flags 2 ($\simeq 75-85\%$ reliable), (bottom-left): VUDS galaxies with spectroscopic flag 3 and 4 (95-100% reliable), (bottom-right): VUDS galaxies with spectroscopic flags 9 ($\simeq 80\%$ reliable). Galaxies in the three different fields are identified as filled squares (COSMOS field), open triangles (VVDS-02h), and open circles (ECDFS). The 1:1 equality relation is drawn as a continuous line, and 15% errors expressed as $0.15 \times (1 + z)$ are drawn as dashed lines. The known degeneracy of photometric redshifts between a Balmer–4000Å break and a Lyα–1215Å break is identified by the dot-dash line, with 15% errors drawn with dotted lines.

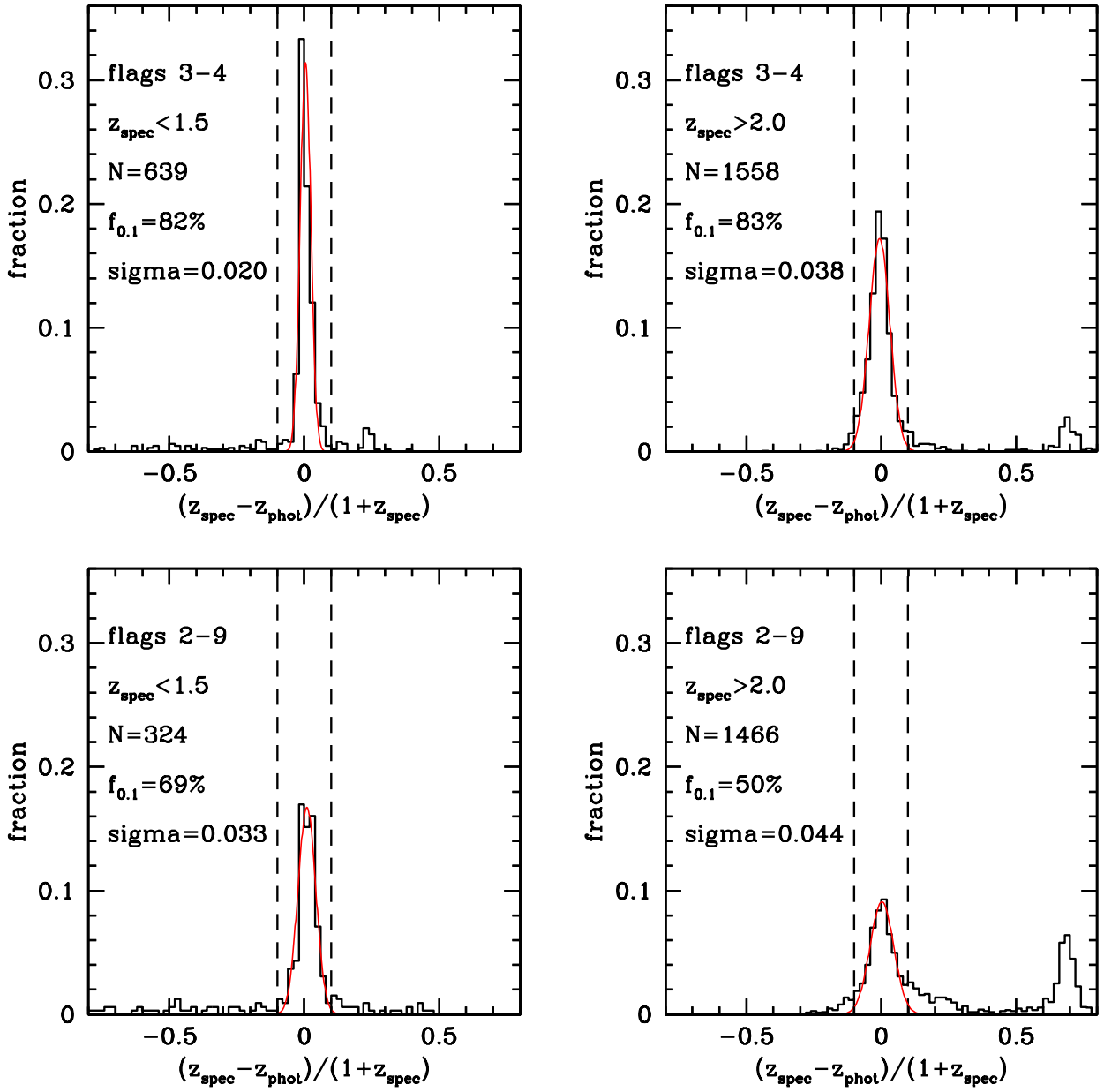


Fig. 12. Distribution of the difference between photometric redshifts and VUDS spectroscopic redshifts for flags 3+4 (top panels) and flags 2+9 (bottom panels), for redshifts $z < 1.5$ (left panels) and redshifts $z > 2$ (right panels). The number of galaxies is indicated in each panel together with the σ of the distribution, and the fraction of galaxies satisfying $\delta z = (z_{spec} - z_{phot})/(1 + z_{spec}) \leq 0.1$. The distribution peaked at 0 is for the main sample, while a secondary peak appearing beyond $\delta z = 0.5$ is produced by the degeneracy between the D4000 and the 1215Å continuum breaks (see text).

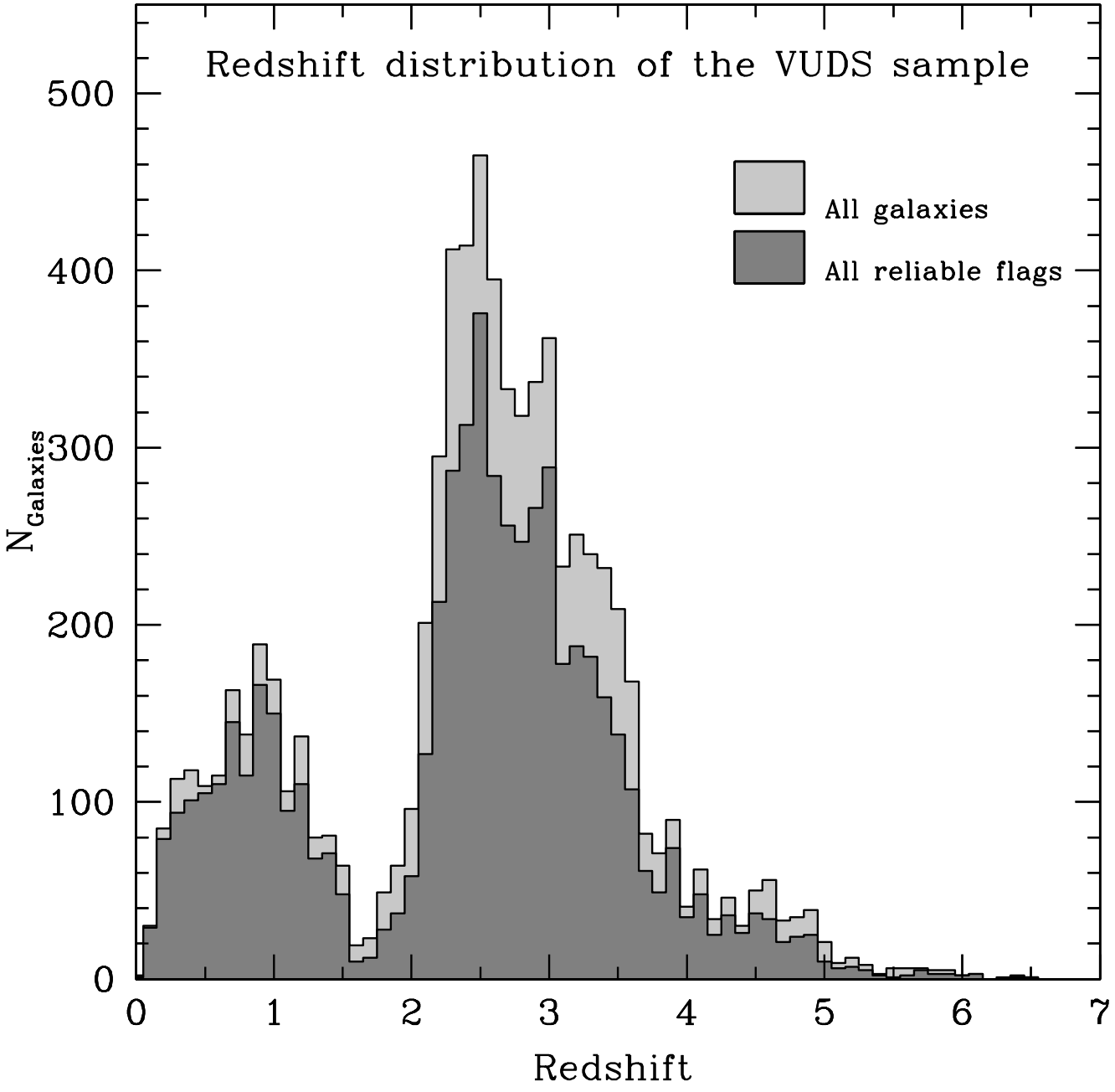


Fig. 13. The current redshift distribution from 7825 galaxies already processed in the VUDS survey, for all objects with a redshift measurement (light grey) and all objects with a > 80% reliable redshift measurement (flags 2, 3, 4, and 9; dark grey).

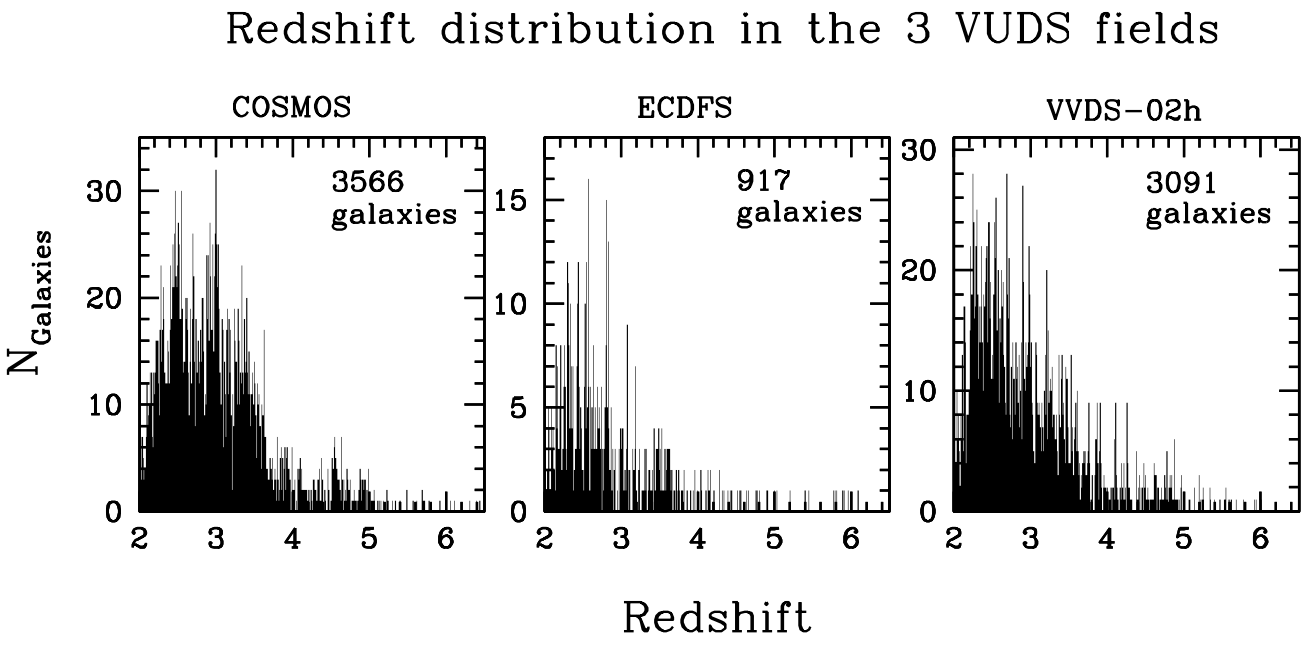


Fig. 14. The distribution of redshifts measured by VUDS in the 3 fields: 3566 galaxies in COSMOS, 917 in ECDFS, and 3091 in VVDS-02h (all objects with a redshift measurement are used). Note that the number of galaxies in each field will be higher than in these figures when the sample will be complete.

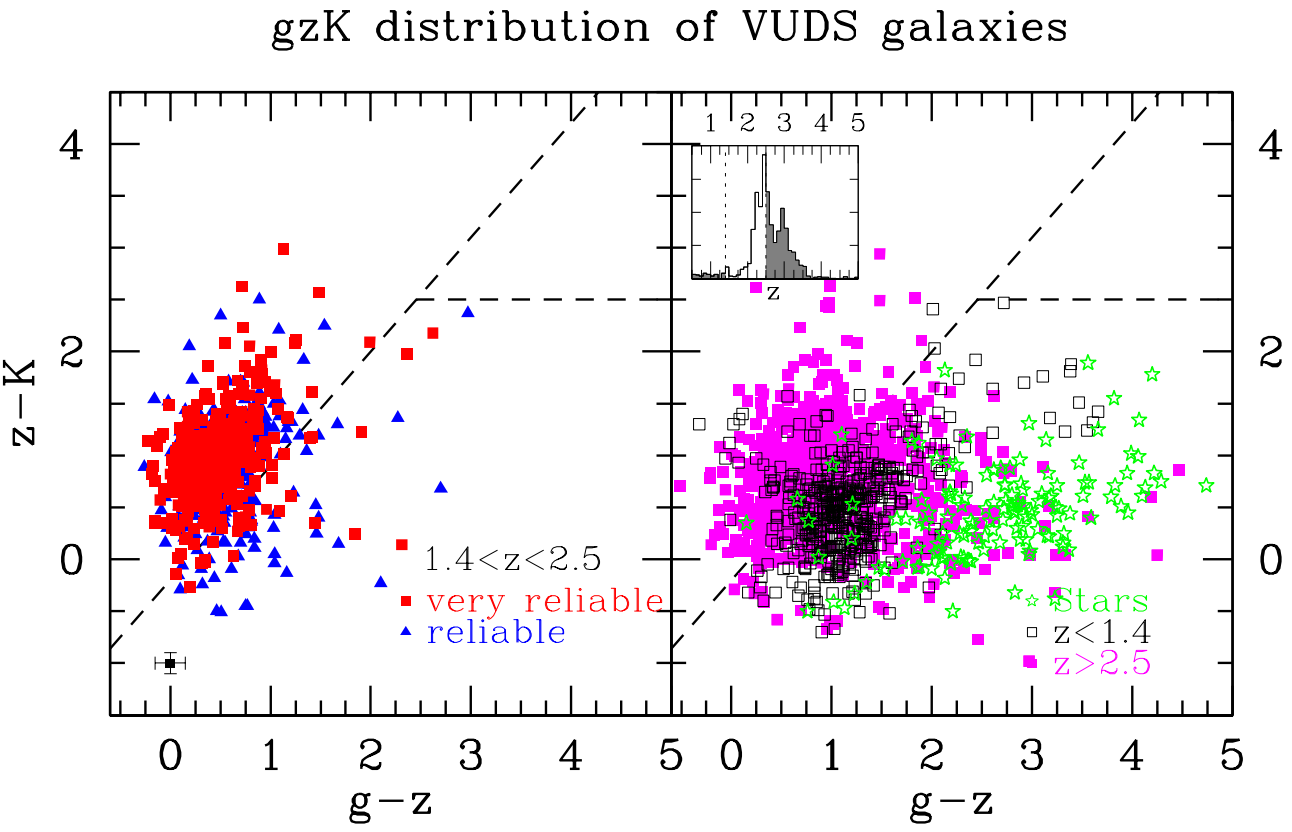


Fig. 15. (left panel) (g-z) vs. (z-K) colour-colour diagram for VUDS galaxies with $1.4 < z < 2.5$ (flag 3+4: red squares, flag 2: blue triangles). The average colour errors are indicated on the lower left corner. Galaxies in this plot are detected in all three bands at more than 3σ (95% of the sample). The *gzK* selection would have selected 91% of the VUDS galaxies with $1.4 < z < 2.5$. (right panel) same for galaxies either with $z < 1.4$ (open squares) or $z > 2.5$ (magenta filled squares). Stars are represented by green starred symbols. There is a high level of contamination as 58% of galaxies in the selection area of the *gzK* are outside the redshift range $1.4 < z < 2.5$, coming mostly from galaxies at $z > 2.5$. (inset) redshift distribution of galaxies in the gzK selection area, the grey shaded histogram represents galaxies outside the redshift range $1.4 < z < 2.5$.

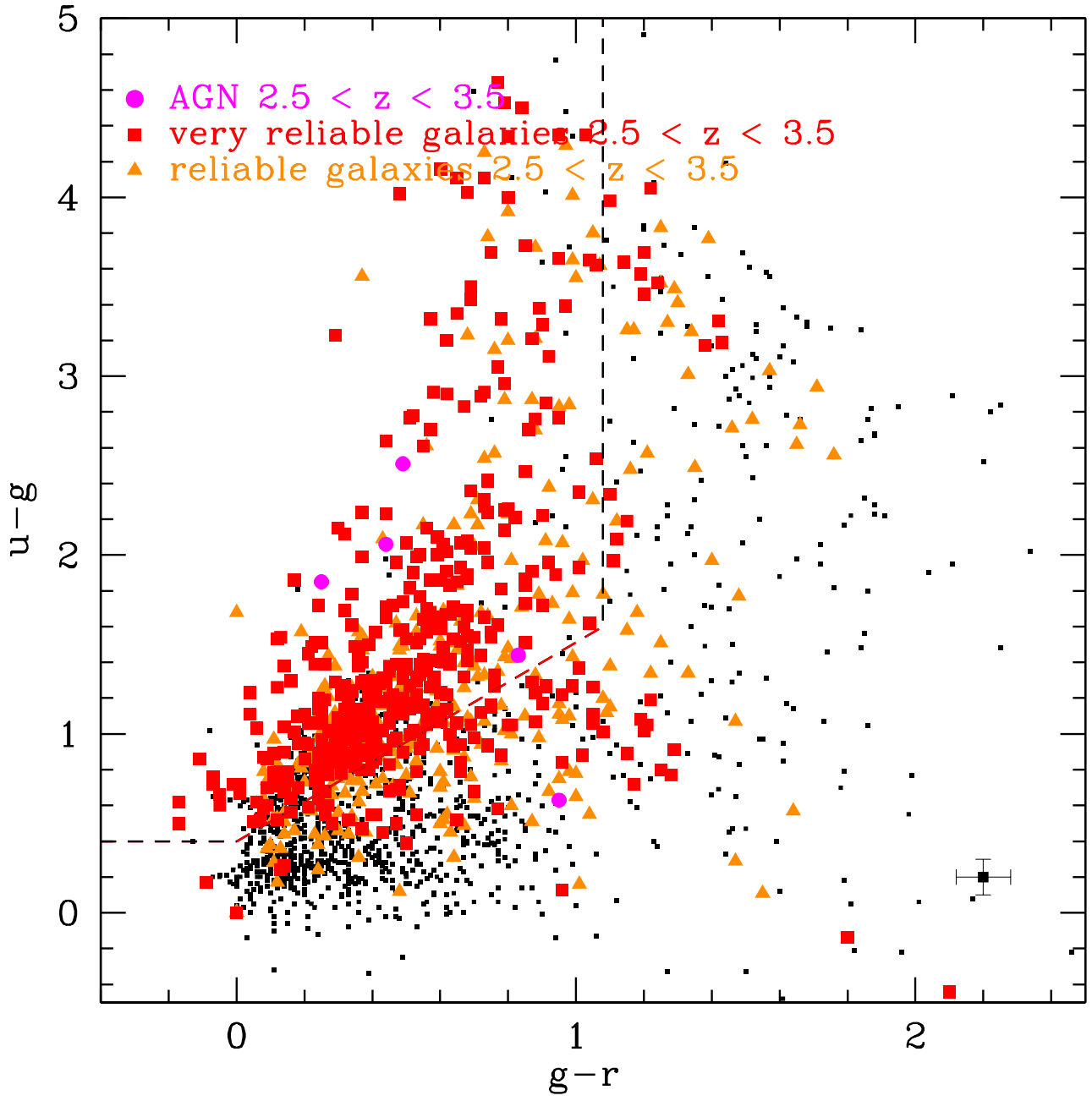


Fig. 16. (u-g) vs. (g-r) colour-colour diagram for VUDS galaxies with $2.5 < z < 3.5$. Galaxies with flag 3+4 are represented by red squares, galaxies with flag 2 by orange triangles while type I AGN are represented by filled magenta circles. The *ugr* selection would have selected 80% of the VUDS galaxies with $2.5 < z < 3.5$, but with a high level of contamination as 40% of galaxies in the selection area of the *ugr* are outside the redshift range $2.5 < z < 3.5$.

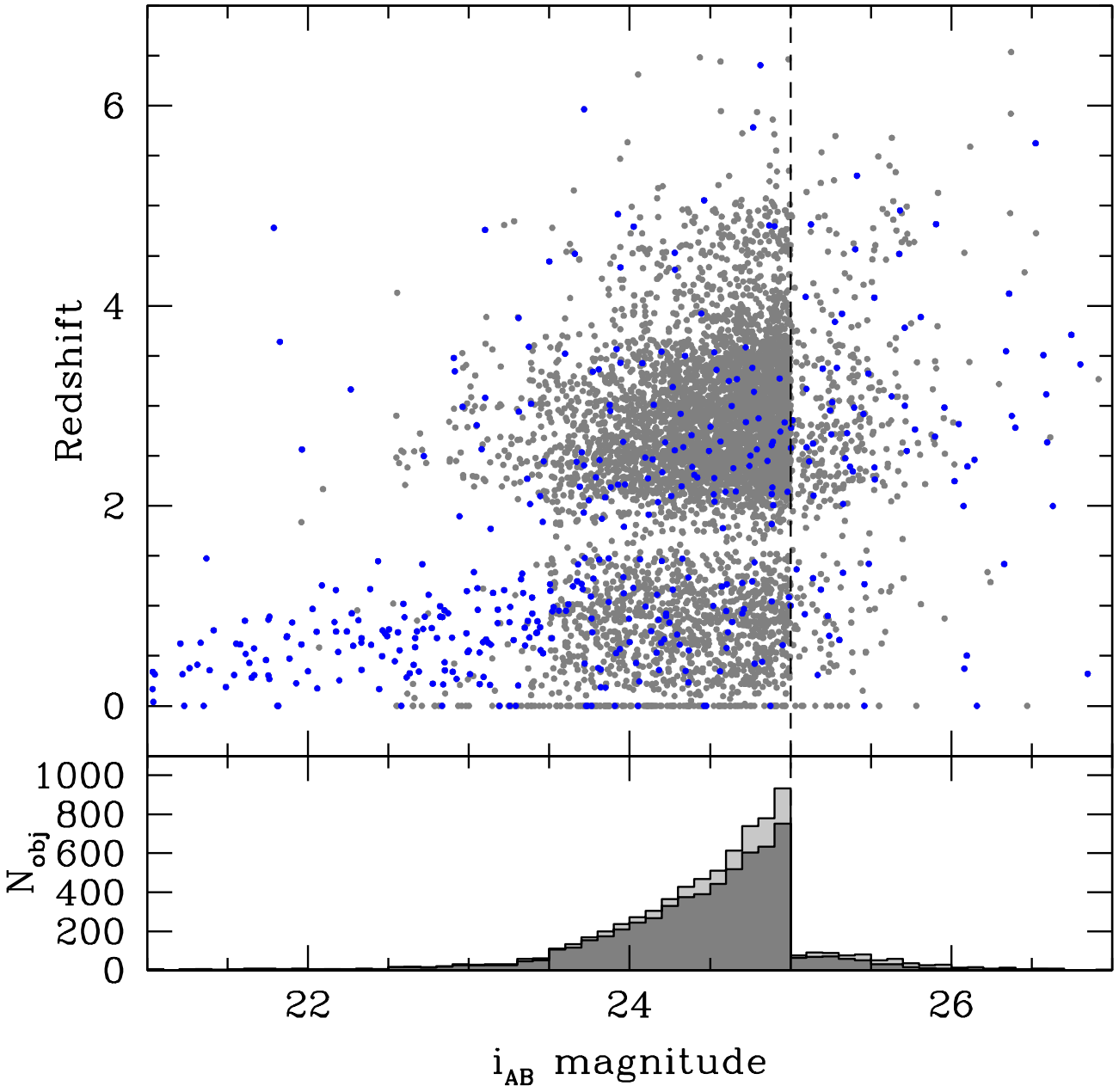


Fig. 17. (top panel) i_{AB} magnitude - redshift distribution for the full VUDS sample (grey points) and for serendipitous objects in the slits (blue points). (bottom panel) Distribution of i_{AB} magnitudes in the VUDS survey for all objects observed (light grey), and for those with a redshift measurement (any non 0 flag, dark grey). The $i_{AB} = 25$ imposed on the sample selected by photometric redshifts is indicated. The fainter objects are pre-selected from one of the other colour or SED criteria.

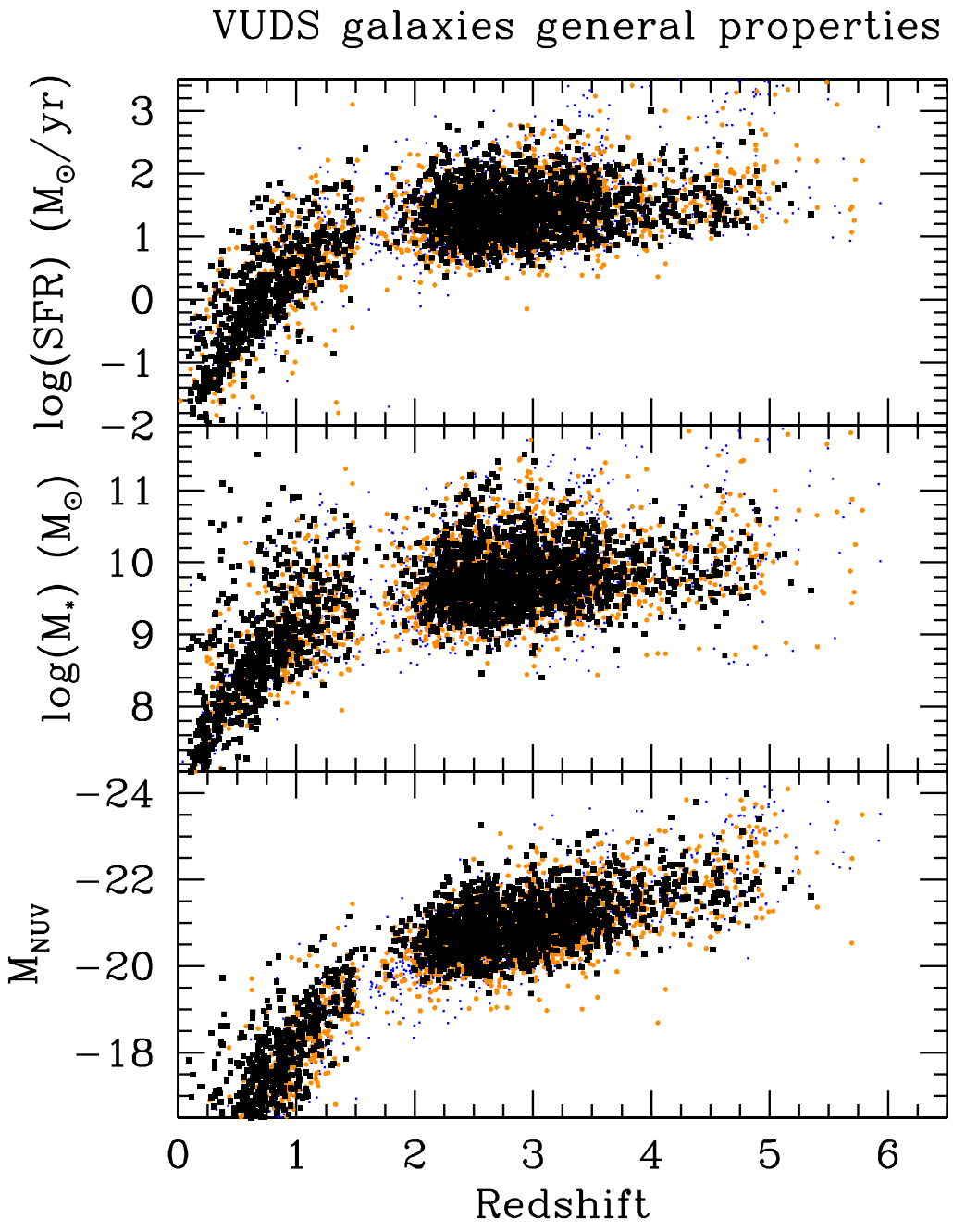


Fig. 18. The distribution of absolute U-band magnitudes (bottom), stellar masses (center), and star formation rate (top). Stellar masses and star formation rates are derived from template SED fitting at the spectroscopic redshift (see text). Black squares are for galaxies with reliability flags 3 and 4, orange filled circles are for flag 2, and blue dots are for flag 1.

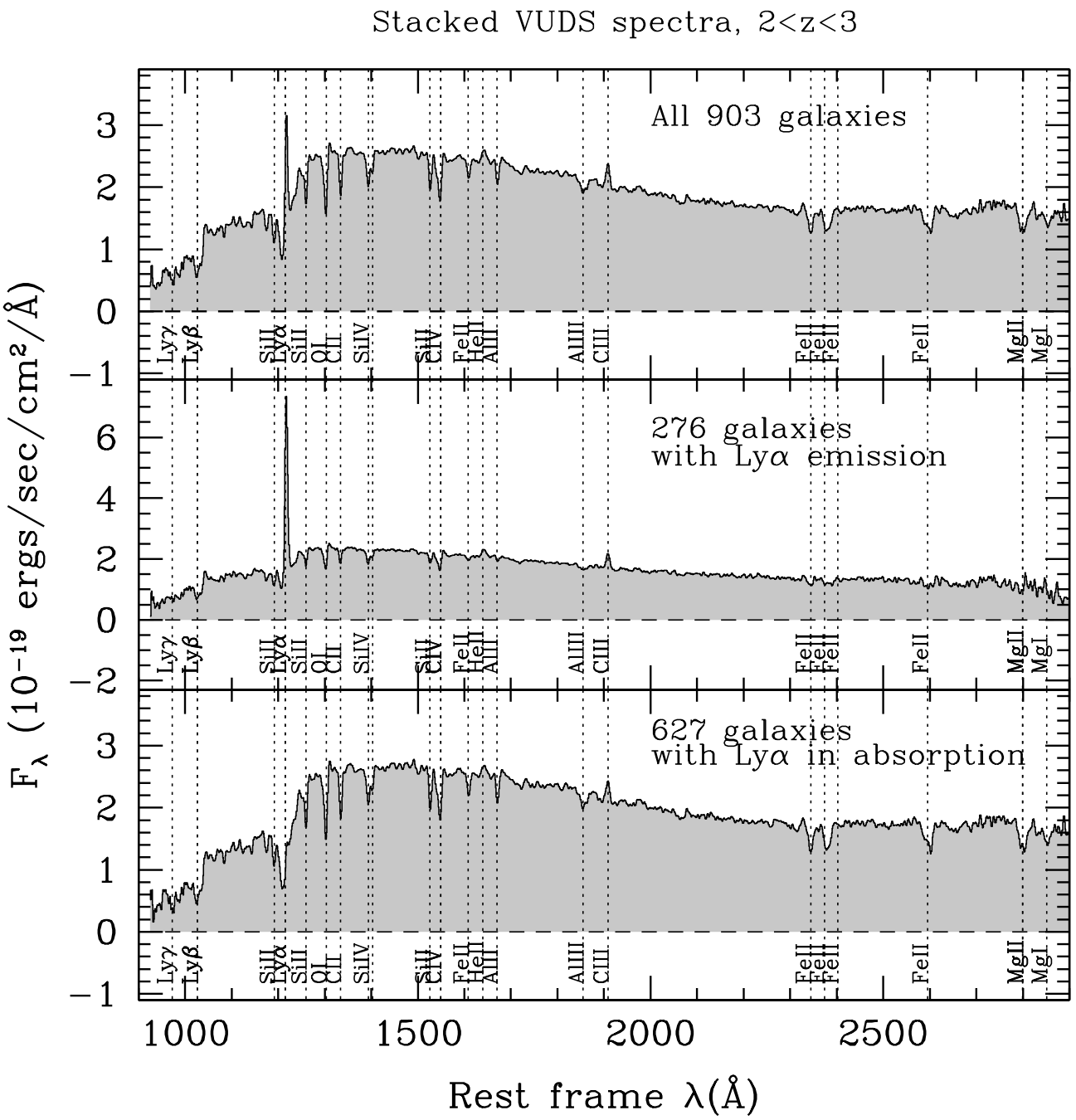


Fig. 19. Average rest-frame spectra (F_λ) of galaxies with flags 3 and 4 in VUDS with $2 \leq z \leq 3$: *(Top)*: stack of all galaxy spectra; *(Middle)*: stack of galaxies with Ly-α in emission; *(Bottom)*: stack of galaxies with Ly-α in absorption. (this stack is from a subset of the final sample available at the time of this writing).

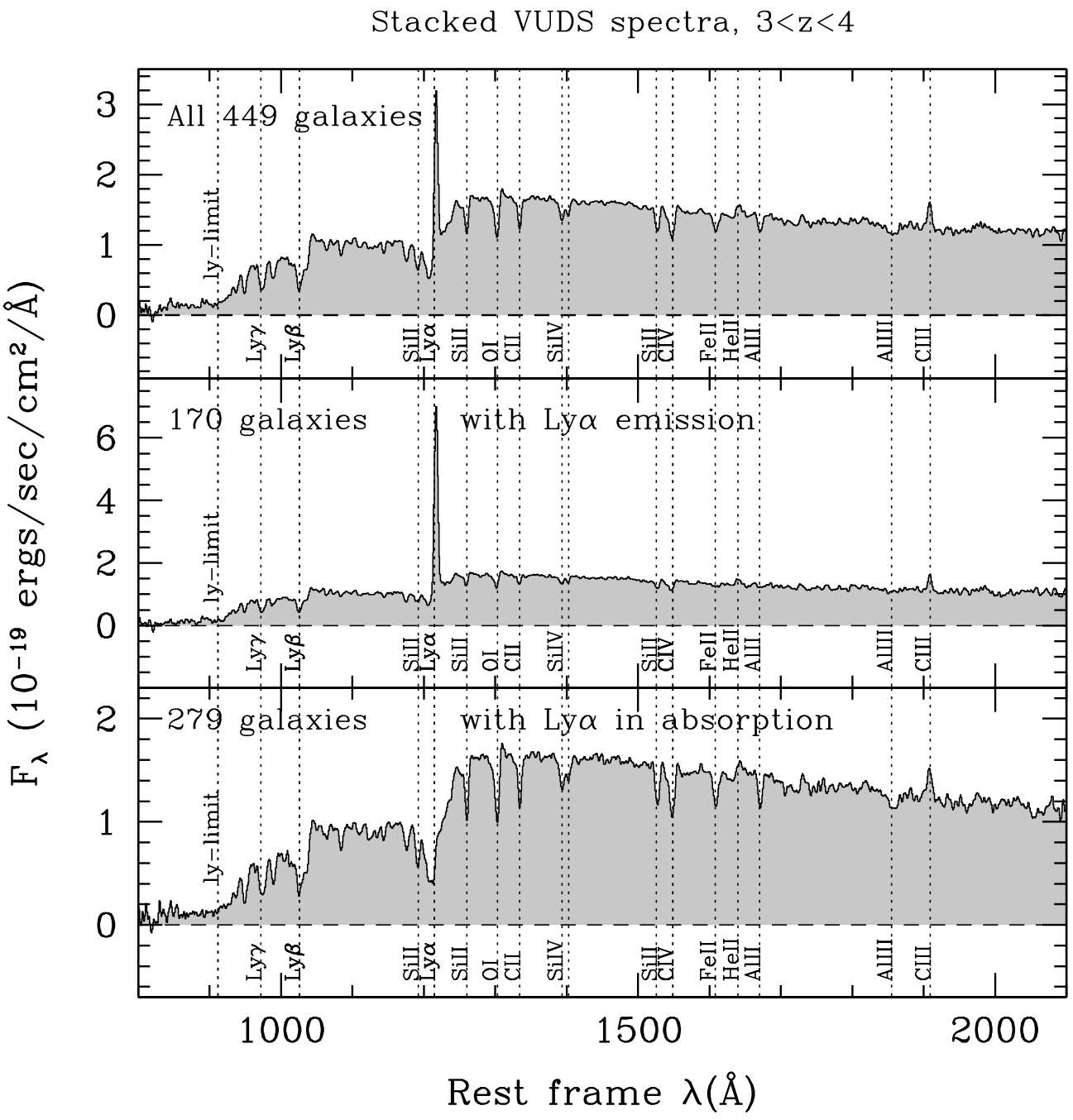


Fig. 20. Average rest-frame spectra (F_λ) of galaxies with flags 3 and 4 in VUDS with $3 \leq z \leq 4$: *(Top)*: stack of all galaxy spectra; *(Middle)*: stack of galaxies with Ly-α in emission; *(Bottom)*: stack of galaxies with Ly-α in absorption. (this stack is from a subset of the final sample available at the time of this writing).

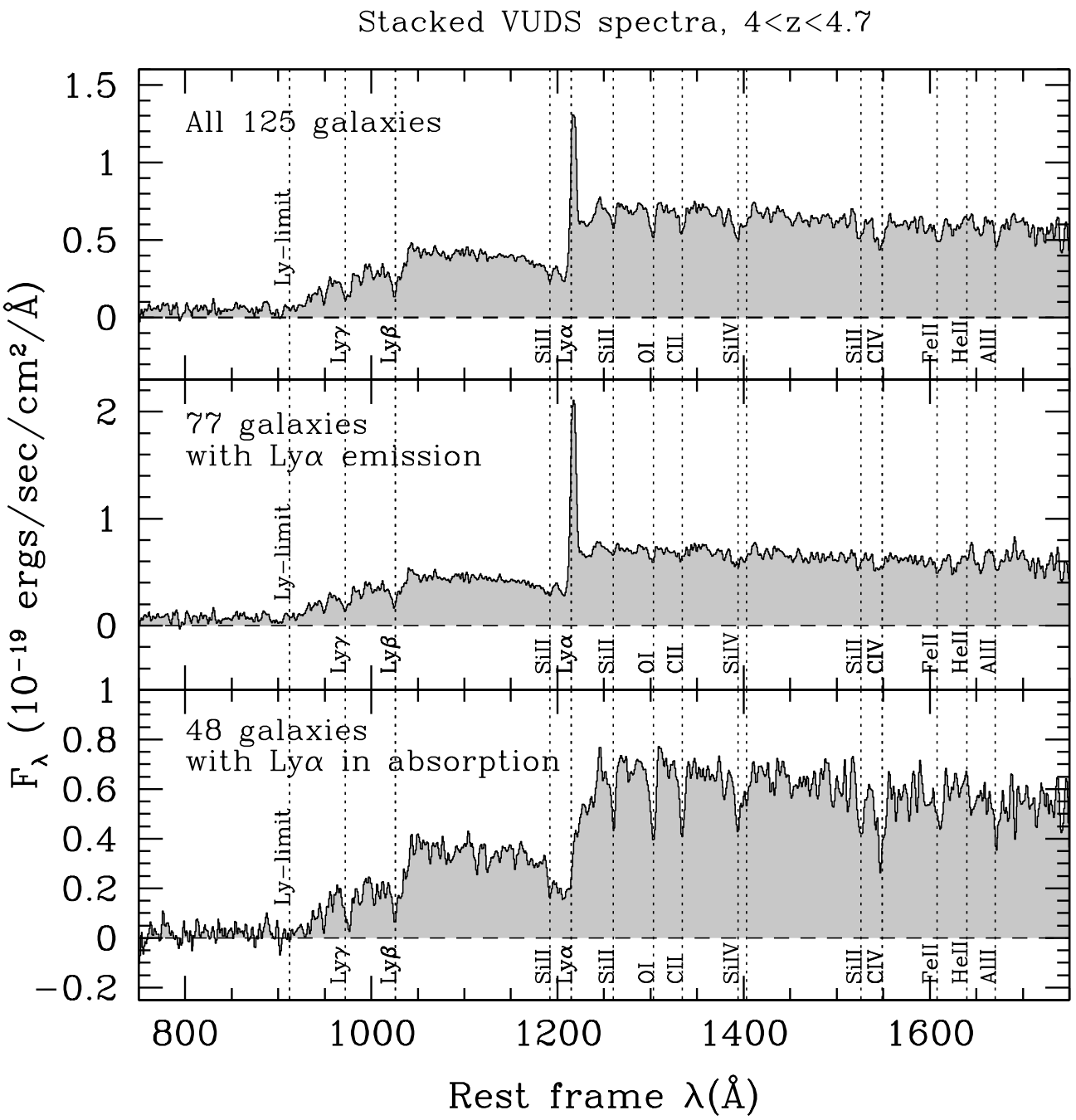


Fig. 21. Average rest-frame spectra (F_λ) of galaxies with flags 2, 3 and 4 in VUDS with $4 \leq z \leq 4.7$: *(Top)*: stack of all galaxy spectra; *(Middle)*: stack of galaxies with Ly-α in emission; *(Bottom)*: stack of galaxies with Ly-α in absorption. (this stack is from a subset of the final sample available at the time of this writing).

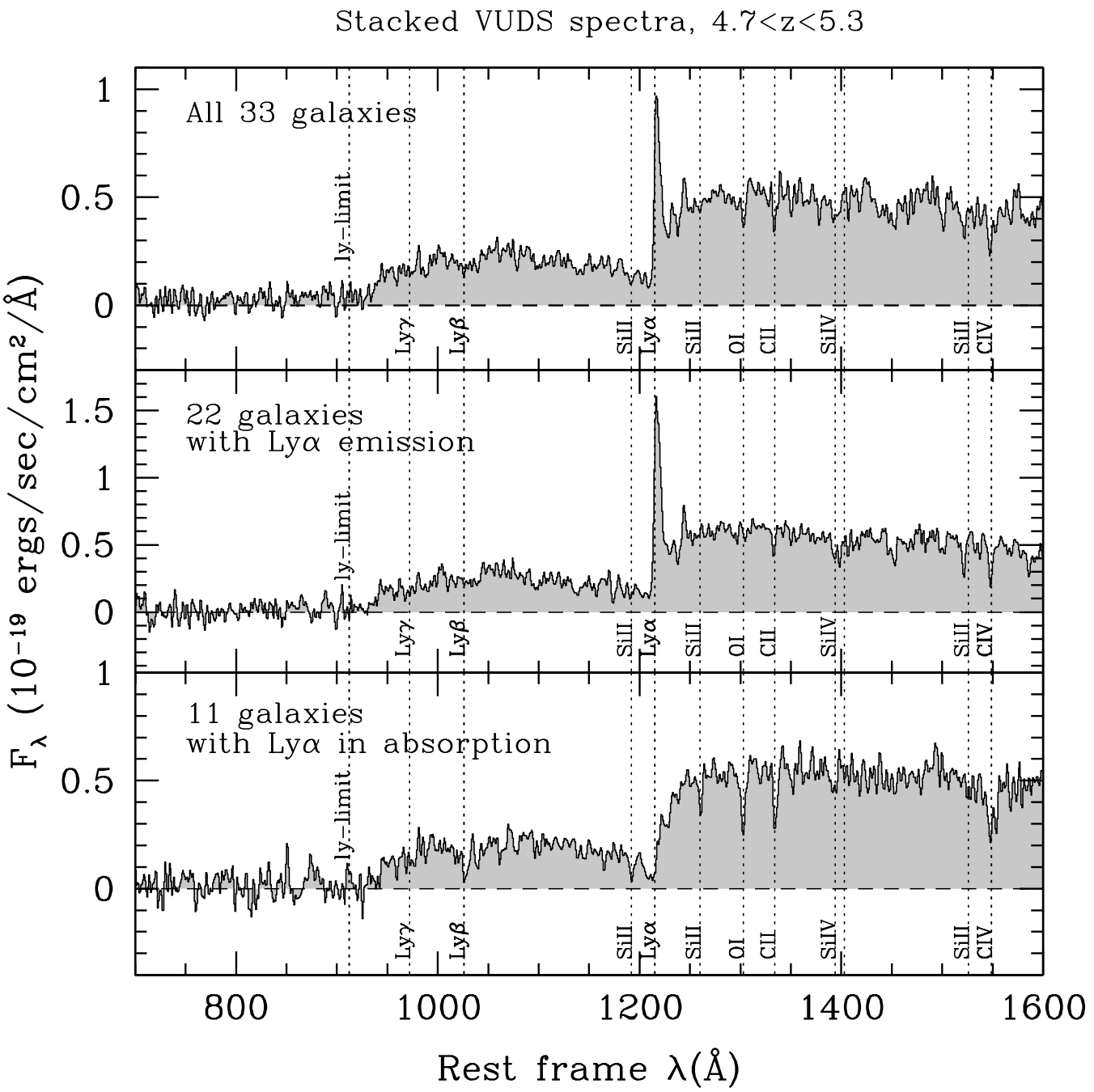


Fig. 22. Average rest-frame spectra (F_λ) of galaxies with flags 2, 3 and 4 in VUDS with $4.7 \leq z \leq 5.3$: *(Top)*: stack of all galaxy spectra; *(Middle)*: stack of galaxies with Ly-α in emission; *(Bottom)*: stack of galaxies with Ly-α in absorption. (this stack is from a subset of the final sample available at the time of this writing).

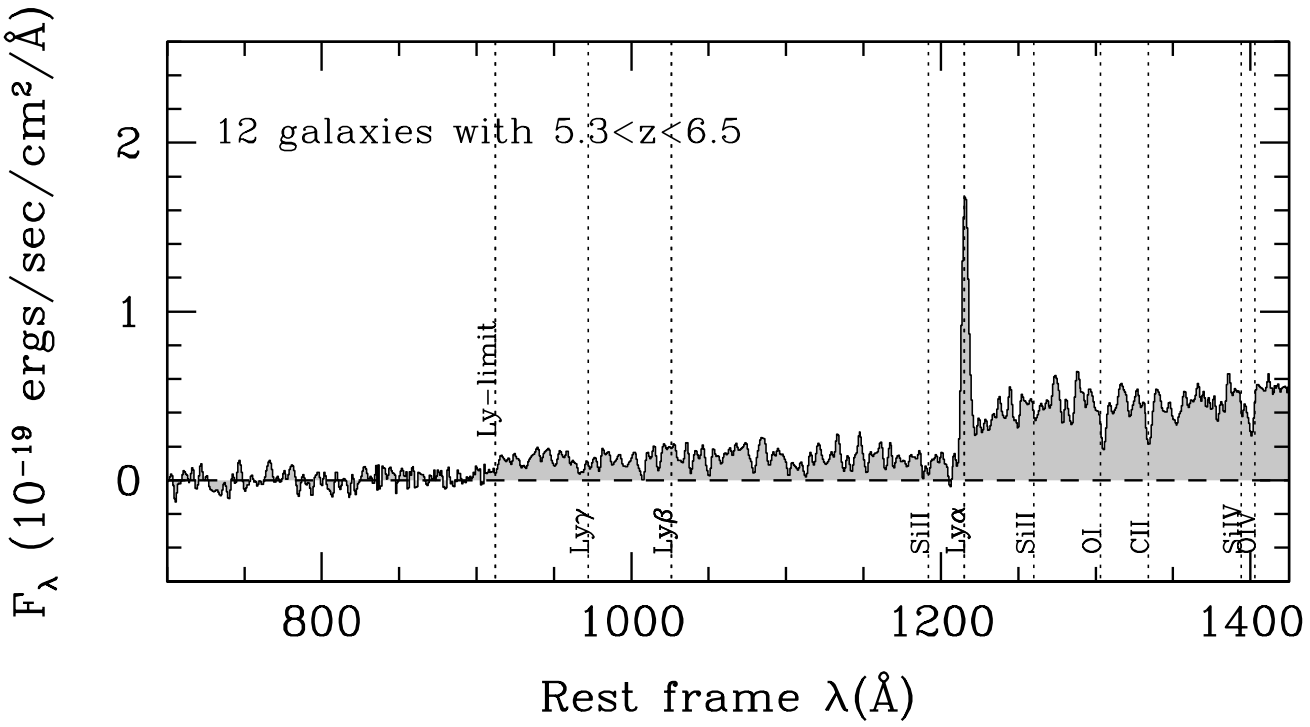


Fig. 23. Average rest-frame spectra (F_λ) of galaxies with flags 2, 3 and 4 in VUDS with $5.3 \leq z \leq 6.5$ (this stack is from a subset of the final sample available at the time of this writing).

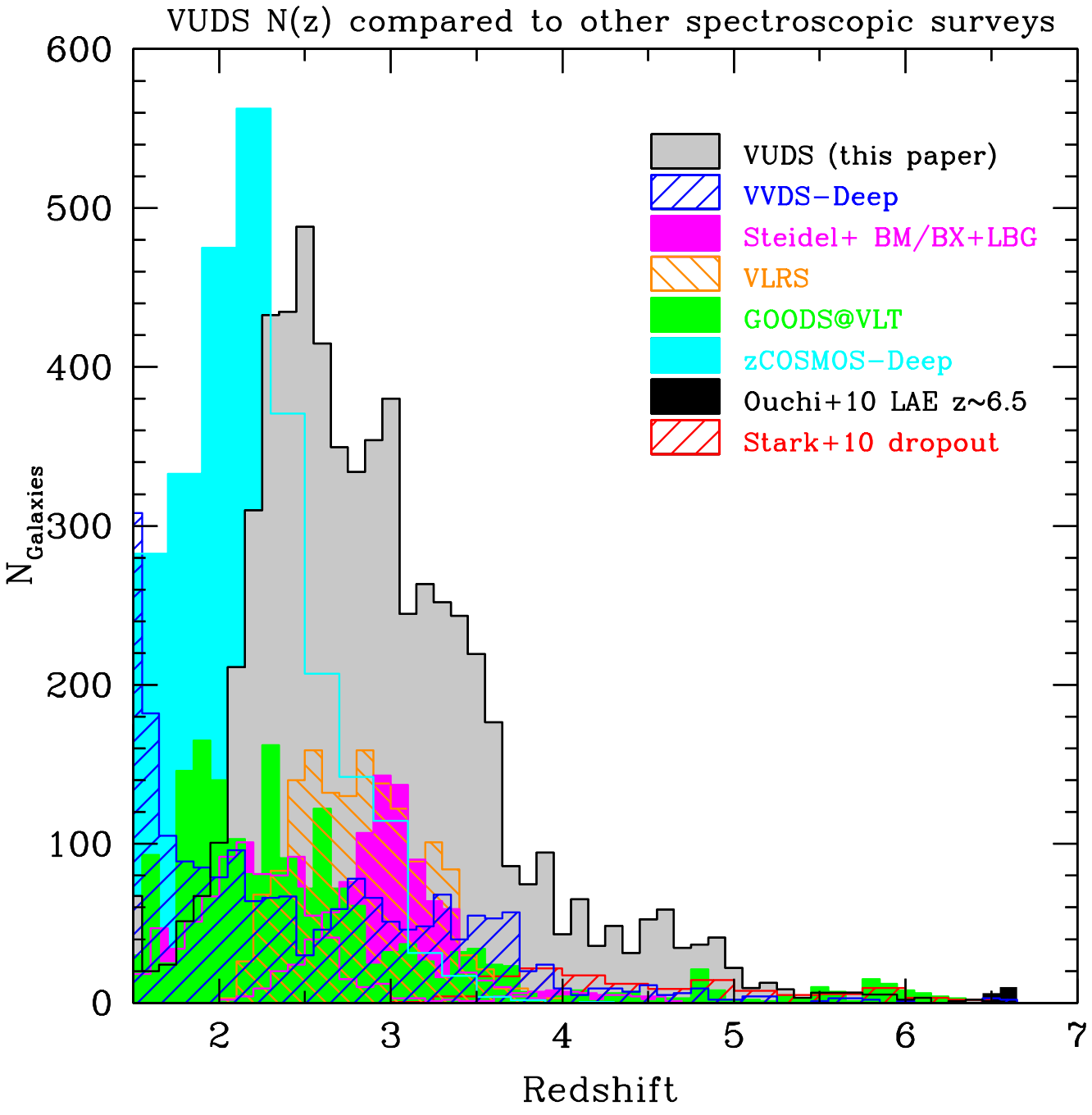


Fig. 24. Comparison of the redshift distribution of the VUDS survey with other published spectroscopic surveys at $z > 2$ (as listed in Table 5). The VUDS counts from existing measurements have been scaled by 1.05 to account for the data not yet processed at the time of this writing.

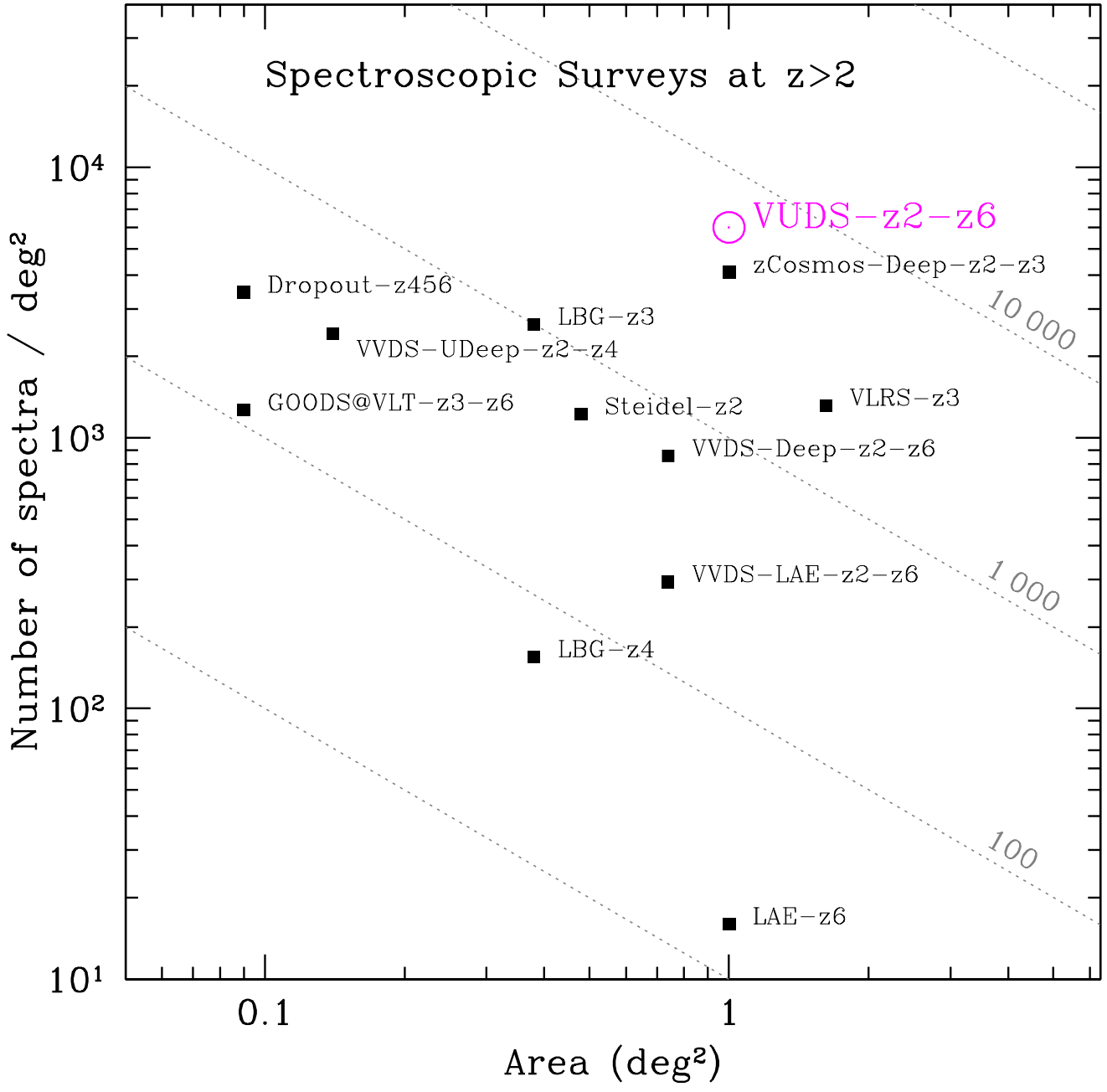


Fig. 25. Comparison of the number density of spectra in the VUDS survey with other published spectroscopic surveys at $z > 2$ (as listed in Table 5). The VUDS counts have been scaled by 1.2 to account for the data not yet processed at the time of this writing. Dashed lines are for surveys with a total number of spectra as indicated.

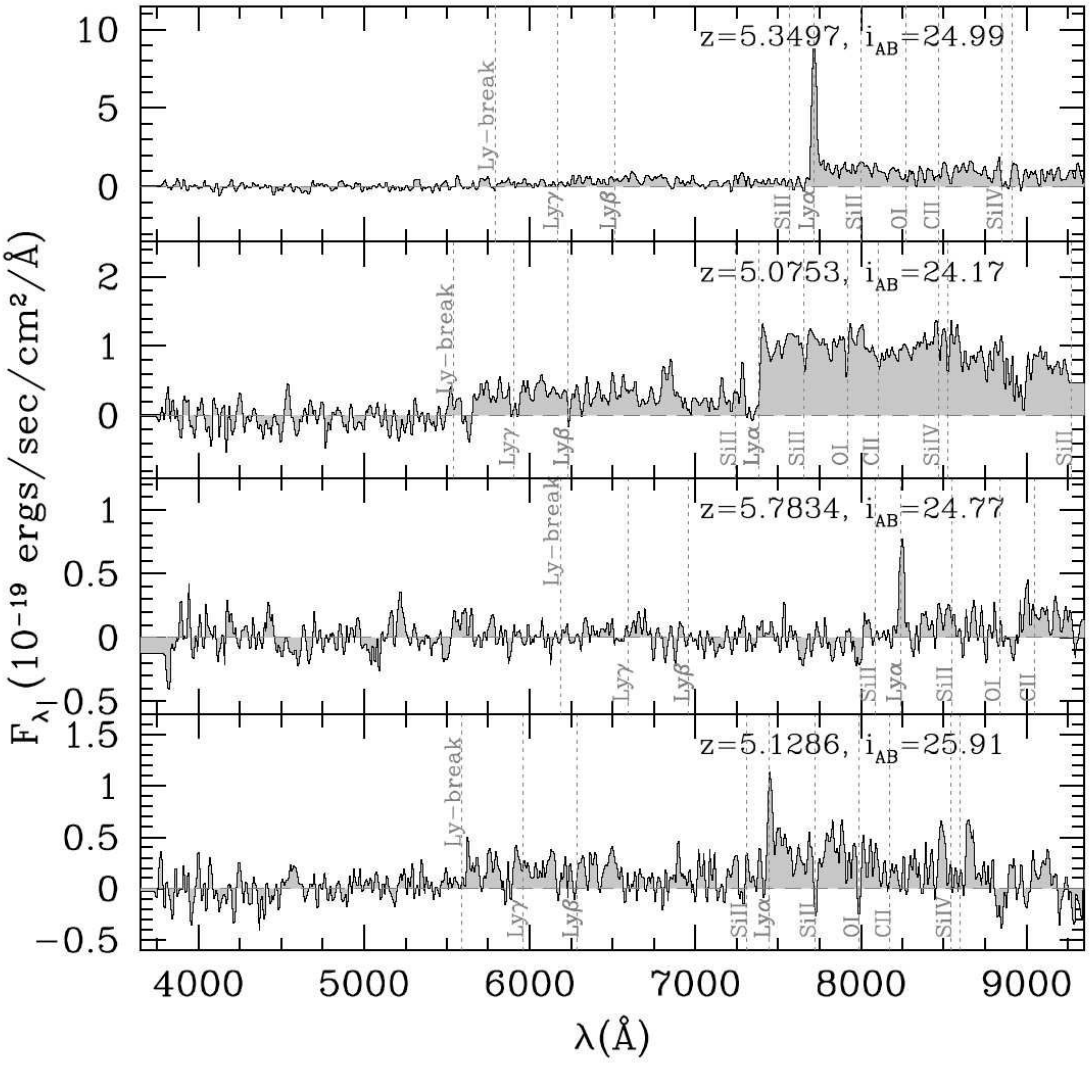


FIGURE 4.10: The $5 < z < 6$ spectroscopic confirmed sample from VUDS (Le Fèvre et al. 2014)

Chapter 5

Candidates Selection of Galaxies $z > 4.5$ and Simulations

In this chapter I describe the selection of a sample of galaxies at $z \geq 4.5$ and associated uncertainties. I produce a set of simulations including galaxy tracks in colour-colour space, and image simulations produced from creating and injecting simulated galaxies into the images, and explain how their detection rate compare to the real galaxy sample. Distant galaxies at $z > 4.5$ are faint and their projected density is low compared to the numerous foreground galaxy population. I aim to understand what is the completeness and contamination rate for galaxy samples which are selected. To quantify completeness and contamination I have run a simulation placing 20 000 mock galaxies with known redshifts and magnitudes in the different UltraVista images, and run the same photometric detection algorithm with SExtractor as I did on catalogue construction.

I then describe the colour-colour selection methods and identify severe limitations. I have explored the robustness of the colour-selection criteria and found that several contaminants are difficult to remove from colour selection only, mostly stars and lower redshift $z \sim 1 - 3$ reddened galaxies. This motivates the definition of a more general method based on photometric redshifts. I describe how I have applied this method on the UltraVista data to select a robust sample of galaxies at $z \geq 4.5$, including a classification scheme based on visual examination of the SED.

This chapter describes this whole process: Section 5.1 explains how simulations are performed, in Section 5.2 I describe sample selection methods focussing on the photometric redshift selection, in Section 5.3 I apply this selection method to the UltraVista data to produce the sample of galaxies at $z \geq 4.5$ used in the following chapters of this manuscript.

5.1 Simulating Galaxies at $z > 4.5$

5.1.1 Goals

With image simulations I aim to measure several quantities to better understand the selection function associated to the selection of a distant galaxy sample. First I want to measure the recovery rate of galaxies as a function of magnitude. Then by applying the same selection criteria as used on the observed data I want to compute the completeness and purity of the samples. There are three main parameters that I aim to measure using simulations:

- **Detection Rate**: For detection rate, I simulate galaxy magnitudes based on model SED, these galaxies are added into the real UltraVista $YJHK_s$ images so that they have the proper S/N, and SExtractor is run on the image to perform the photometry of these simulated objects. From the photometry I derive the detection rate.
- **Galaxy Tracks**: For galaxy tracks I use templates that are redshifted, and at each redshift their location in colour-colour diagrams is computed. These tracks are used to identify the locus of galaxies in colour-colour space. I use the photometric redshift code Le Phare to produce all the SED templates.
- **Completeness and Purity**: Another important goal of my simulations is to evaluate completeness and purity of traditional colour selection of high redshift galaxies. Briefly speaking, when applying the colour criteria to discriminate the candidates in the specific redshift bins, I need to understand how many galaxies are in reality in this redshift bin but don't satisfy the colour criteria (completeness) and check how many interlopers satisfy colour criteria but are not at the right redshift (purity). Especially photometric errors affect the magnitude and may drive galaxies in or out of the selection criteria.

More precisely, completeness and purity are defined as follows when using colour-colour selection:

- **Completeness**: High redshift objects in the colour box / All the high redshift objects (red asterisks in the black dashed line box / all the red asterisks)
- **Purity**: High redshift objects in the colour box / All the objects in the colour box (red asterisks in the black dashed line box / all the asterisks in black dashed line box)

TABLE 5.1: Parameters of LePhare used for simulations

Parameter	Value
Redshift range	$0 - 7.5$
Metallicity	0.004,0.008,0.02 ($Z/Z_{\odot}$)
E(b-v)	0, 0.5
SFRH	0.03,0.1,0.5,1,3 (Gyr)
Emission Line	Yes

These important calculations are performed in the next section after the definition of selection criteria for our sample.

5.1.2 Using Le Phare to simulate galaxy magnitudes and colours

In these simulations one wants to produce mock galaxies from the local universe to $z \sim 7.5$ based on all types of ideal galaxies templates. Le Phare is the tool that I use to make each model SED derived from the BC03 (Bruzual & Charlot 2003) stellar population synthesis model with 3 metallicities, a declining star formation history, 2 dust-reddening E(B−V) settings and including emission lines (all the details are listed in Table 5.1).

5.1.3 Galaxy Tracks

In the dropout technique used to select high-z galaxies, the goal is to identify the Lyman break based on a strong colour change when the break is detected in a three-colour set. Several colour−colour diagrams are used to select galaxies at different redshifts based on the position of the break. The locus of galaxies in colour-colour plots as a function of redshift are shown in Figure 5.1, using extreme SED models from the local universe to redshift 7.5 with E(B−V) = 0 and E(B−V) = 0.5, respectively.

These tracks can be used to identify the area in colour-colour space where to expect galaxies in a given redshift range.

5.1.4 Image simulations

All photometric-based high-z galaxies selection techniques rely on the ability to properly measure the magnitudes and colours of galaxy samples as complete as possible compared to the underlying population. The computation of magnitudes using photometry algorithms like implemented in the popular SExtractor code (Bertin and Arnouts 1996) is sensitive to noise variations across an image as well as to the contamination from close

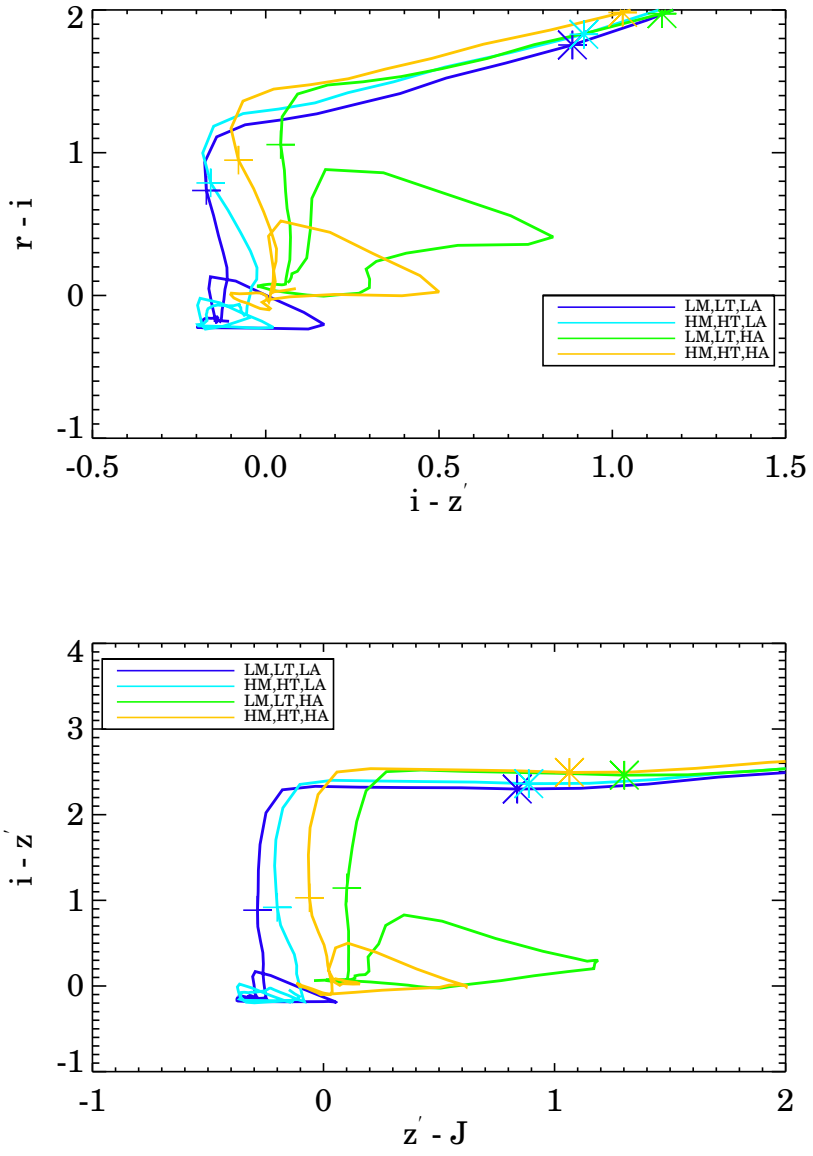


FIGURE 5.1: Galaxy track plot in two typical colour space (riz' and $iz'J$ for selecting galaxies with $4.5 < z < 5.5$ and $5.5 < z < 6.5$ respectively. Each track shows a galaxy followed from redshift 0 to redshift 7.5 in each colour-colour space, plus signs and asterisks show the lower limit and upper limit in each redshift bins. Different colours of lines mean different combination of extreme SED templates. (LM, HM denote lowest metallicity and highest metallicity; LT, HT denote lowest τ and highest τ; LA, HA denote youngest age and oldest age in this redshift bin.)

neighbours. As objects get fainter, the density of sources increases in such a way that the probability to have a close neighbour within the aperture (or isophote) used to compute the magnitude is increasing to more than 25% (for a 3 arcsec aperture), and 'decontamination' issues are not straightforward and depend on the distribution of light for a given object (extended or compact, irregular or symmetric, ...).

The best and pragmatic way to evaluate limitations in the photometry process is to produce simulations where objects of known magnitudes and colours are injected randomly into a real image. Object detection and photometric measurements are then performed with the same code and parameters as used for the real non simulated image. One then gets a catalogue of objects extracted from the simulated image that can be compared to the known properties of simulated objects.

I have adopted this logic and I have produced simulated images for each of the photometric bands used in this study ($Bvriz'$ and $YJHK_s$). One has to be careful to produce simulations that are not affecting the counts statistics too notably or otherwise face biases in measurements. I choose to add 20 000 simulated galaxies in each image, corresponding to only a few percent of the total count of galaxies in these fields. I simulated galaxies with a flat magnitude distribution selected randomly in the K_s range from 23.5 to 25.5. I choose to simulate galaxies as point-sources, convolved with the observed seeing of each image, under the assumption of the high-z galaxies I'm interested in are unresolved or barely resolved under the typical FWHM$= 0.8 - 1$ arcsecond seeing of the observations. All mock galaxies were inserted into all images avoiding masked regions resulting from the photometric catalog process. I then ran SExtractor with the exact same parameter set used to construct the photometric catalogues of observed sources. This produced catalogues of simulated sources. From the 20 000 mock galaxies 13 312 were recovered in this process.

The second goal of simulations is to compute the completeness and purity of these candidates selected by my photometric criteria. These factors are very important to estimate the correct (true) number of galaxies at the redshifts of interest. It is absolutely needed to compute these corrections because there is no perfect colour criteria allowing us to select all high redshift candidates without any low redshift interlopers. The only way is to pre-select candidates avoiding interlopers as much as possible, and then use the completeness and purity estimates to obtain a corrected estimate of the number counts. The more accurate the completeness and purity estimates are, the closer to the true number counts one can be.

I show in Figure 5.2 a small extract from a real image compared to the same image but including the simulated galaxies.

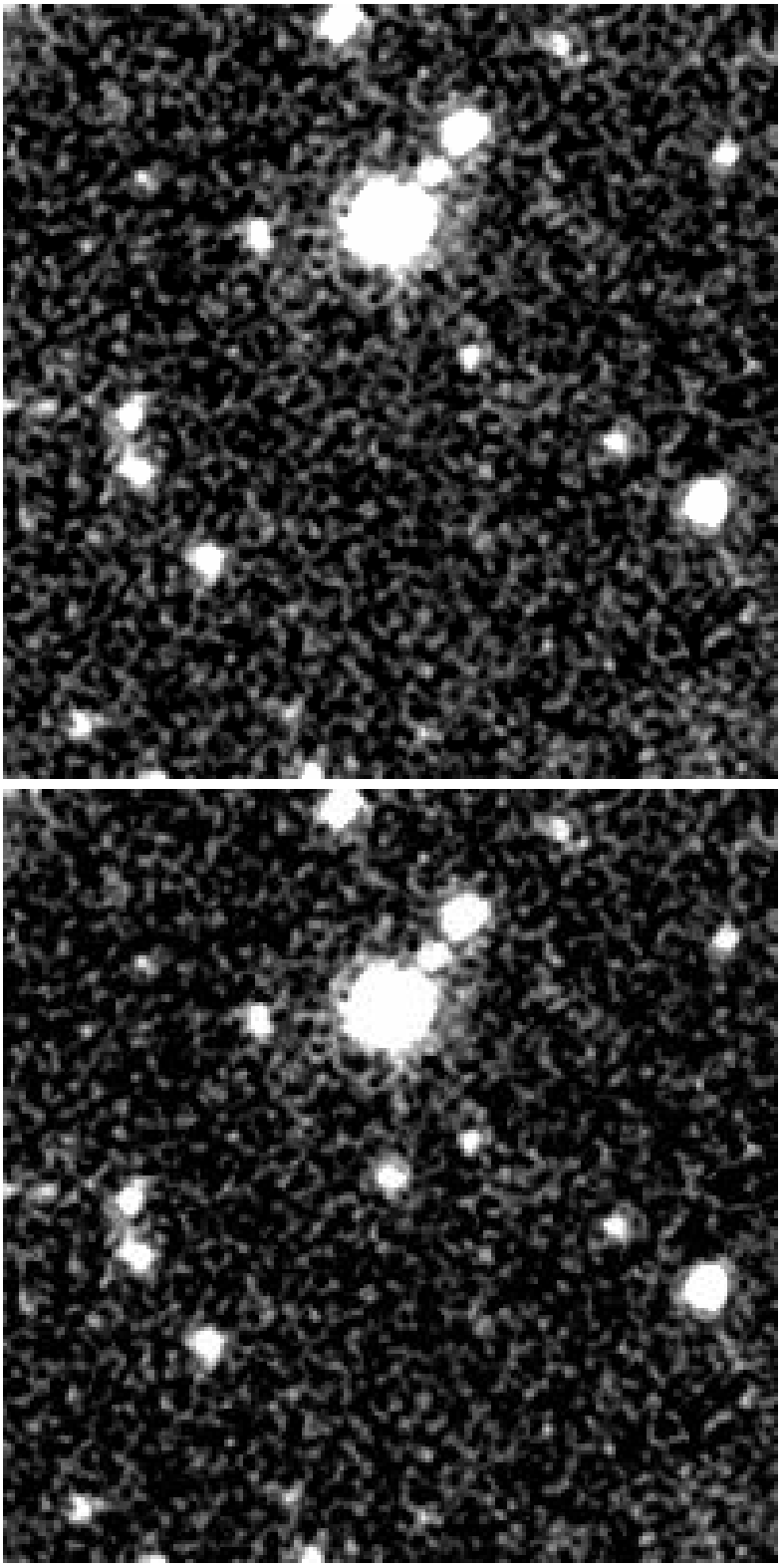


FIGURE 5.2: Two poststamps of original image (upper panel) and simulated image (lower panel) in K_s band. In the center of simulated image shows the mock galaxy with $K_s = 23.5$. The sizes of two poststamps are $30'' \times 30''$.

TABLE 5.2: Detection rate from image simulation

magnitude bins	Detection Rate (Ultra-Deep field)	Detection Rate (Deep field)
	z' band	
23.0 – 23.5	0.936782	0.933526
23.5 – 24.0	0.918580	0.907921
24.0 – 24.5	0.904130	0.889371
24.5 – 25.0	0.875000	0.836839
25.0 – 25.5	0.848985	0.691382
25.5 – 26.0	0.826975	0.665122
26.0 – 26.5	0.785216	0.579805
26.5 – 27.0	0.697015	0.352073
	Y band	
23.0 – 23.5	0.932735	0.925581
23.5 – 24.0	0.918060	0.914191
24.0 – 24.5	0.892169	0.888889
24.5 – 25.0	0.878505	0.838421
25.0 – 25.5	0.844530	0.712367
25.5 – 26.0	0.806554	0.603824
26.0 – 26.5	0.735661	0.387839
26.5 – 27.0	0.492038	0.201663

5.1.5 Magnitude completeness

The first goal of my simulation is estimating the detection rate of galaxies with different fluxes. I computed the detection rate over the magnitude range of interest as the ratio between the simulated galaxies recovered from the photometry and the input list of simulated galaxies. The result of this calculation is shown in Figure 5.3 for z' and Y bands, and tabulated in Table 5.2 for two different UltraVista areas: the Deep and the UltraDeep image areas. At the bright end, I find that the recovery rates of all three bands are over ∼90%. The ∼ 10% loss is essentially due to the blending effects of simulated sources with brighter sources in the image. Going to fainter magnitudes the detection rate drops, and the drop happens about one magnitude brighter for the deep compared to the ultra-deep, as expected. This detection rate as a function of magnitude will be used to correct the galaxy counts in each redshift bin using the appropriate band: the z' band for z∼5 and Y band for z∼6.

5.2 Sample Selection Methods at $z \geq 4.5$: overview

Several methods are used to select galaxies beyond $z \sim 4.5$. I recall why magnitude selection is not appropriate, then I describe the colour selection method, classically used to select *drop-out galaxies*, and show its limitations in terms of contamination and completeness. I then describe the more general photometric redshift method based on

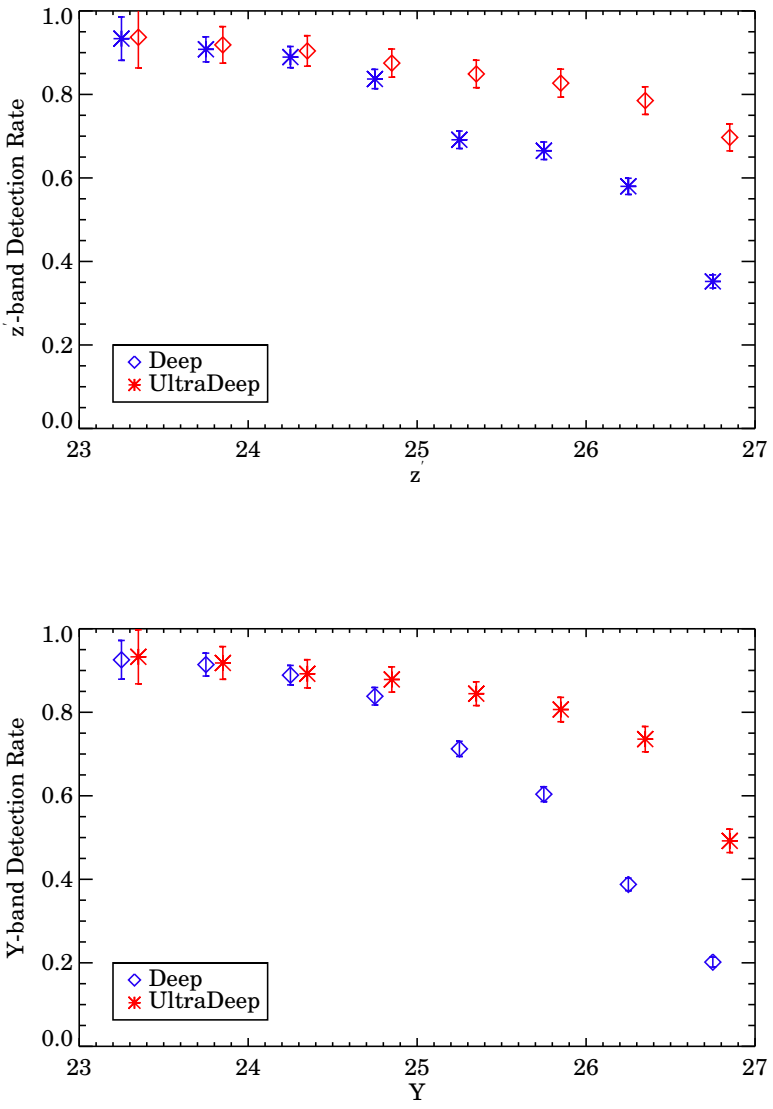


FIGURE 5.3: Comparison of detection rates in deep and ultra-deep UltraVista areas. Similar detection rates around 90% are found for bright galaxies in both areas. The ultra-deep area shows a deeper limiting magnitude as expected.

SED fitting, which I select as the main selection technique to identify a high redshifts galaxy sample at $z \geq 4.5$ in the UltraVista data.

5.2.1 Magnitude selection

Magnitude selection is a simple method which limits the possibility of complex biases when selecting large samples. However, the projected density of galaxies at $z \geq 4.5$ on the plane of the sky is expected to be a small fraction of the total fraction of galaxies

selected this way. For a sample selected with magnitude AB≤25, the redshift distribution of the resulting sample would peak at a redshift $z \sim 1 - 1.5$, and show a high redshift tail going to $z \sim 6$ and beyond (Le Fèvre et al. 2013b). The projected number density of $z < 4.5$ foreground galaxies is about a factor 100 larger than the number of galaxies with $z \geq 4.5$, so looking for galaxies at these high redshifts from magnitude selected samples is very inefficient and additional selection criteria are therefore needed.

5.2.2 Colour selection

As described in Chapter 4 the main technique used to select high redshift $z > 2.5$ galaxies is the Lyman break selection. When redshift increases, the Lyman break to select Lyman break galaxy samples (LBG) is replaced by the Lyman-α drop produced by the almost total extinction of flux produced by the IGM along the line of sight. Using this drop produces *dropout* galaxy samples. This is the main technique used to select galaxies at $z > 4$ as summarised recently by Bouwens et al. (2014). I summarise below what are the colour-colour criteria that can be used to select dropout galaxies in several redshift ranges beyond $z = 4.5$.

5.2.2.1 Redshift ~ 5

From galaxy tracks I define the colour criteria ($r-$ dropout galaxies) in the $(r-i)-(i-z')$ colour-colour plane as follow:

- **Colour box**: $r - i > 0.75, -0.4 < i - z' < 1.2, r - i > 1.6(i - z') + 0.7$
- **Magnitude cut**: B fainter than 27.7, z' brighter than 26.08 (5σ value) and 3 of 4 near-infrared bands (Y, J, H, K_s) should be detected (brighter than 1σ value)

5.2.2.2 Redshift ~ 6

For redshifts over 5.5, pre-selection criteria are defined from galaxy tracks based on simulations (Figure 5.1). Compared to redshift $\sim$ 5 galaxies, the colour cut seems to be working better. However there is still a significant number of interlopers overlapping with redshift ∼6 galaxies in the corresponding colour-colour space.

For this redshift range, I define the colour criteria ($i-$ dropout galaxies) as follows:

- **Colour box**: $i - z' > 0.8, z' - J < 1.2$

TABLE 5.3: Completeness and Purity rate of colour selection from image simulation and VUDS

dropout filter	completeness Rate	Purity Rate
$r-dropout(simulation)$	0.83(786/947)	0.73(786/1080)
$i-dropout(simulation)$	0.82(679/832)	0.76(679/895)
$r-dropout(VUDS)$	0.532	0.453
$i-dropout(VUDS)$	0.25	0.016

- **Magnitude cut**: Y brighter than 25.9 in deep strips and 27 in ultra-deep strips (1.5σ in deep or ultra-deep strips), B fainter than 29.18 (2σ value), V fainter than 28.04 (3σ value) and 2 of 3 near-infrared bands (J, H, K_s) should be detected (brighter than 1σ value)

5.2.2.3 Colour completeness

The Table 5.3 gives the Completeness and Purity of the colour selection scheme defined above, based on the photometric simulations alone. One can see that the completeness decreases when going to higher redshifts.

We can check the efficiency of this colour-colour selection by comparing to galaxies with the most reliable redshifts in the VUDS spectroscopic survey which is selected with photometric redshifts. We show the $(r-i)-(i-z')$ plot of all VUDS galaxies with $4.5 \leq z \leq 5.5$ compared to the selection criteria listed above in Figure 5.4. I used only VUDS galaxies with the most reliable redshifts: flags 2,3,4 and 9 for $z \sim 5$ and all flags for $z \sim 6$ to maximise statistics (keeping in mind that some of the flag 2 may be incorrect). I find that 46.8% (72/154) of the VUDS galaxies do not satisfy the selection criteria although they are known to be at the right redshifts; this gives a completeness rate of 53.2% (82/154). For galaxies which satisfy the selection criteria, I find that this includes not only galaxies in the expected redshift range but also galaxies at other (lower) redshifts. The number of interlopers is quite high in part due to the range of IGM transmission not taken into account in the simulated tracks (see Thomas et al. 2015) which in some cases is higher producing a higher continuum between 912Å and 1216Å (UV continuum to Lymanα emission). From this I compute a purity of 45.3% (82/182) (Table 5.3). I do the same using the $(i-z')-(z'-J)$ colour-colour plot, I find that 25% (2/8) completeness rate for picking up redshift 6 galaxies and the purity rate is only 2% (2/119) by colour criteria. But there is only 8 spectroscopically confirmed redshift 6 galaxy which is deficient on statistically analysis. So the computation of completeness and purity over redshift 5.5 are still based on my simulation.

Figure 5.7 and Figure 5.9 clearly demonstrate the difficulties in separating high-redshift galaxies from and extreme red galaxy located in $z \sim 2-3$ and dwarf-star contaminants without further SED analysis. Theses are typical two cases that colour of some low-z interlopers could also meet the pre-selection criteria, and contamination of galactic cool star can still be present because of the photometric error.

5.2.3 Photometric redshift selection

I conclude from the above section that colour-colour selection of galaxies with $z > 4.5$ can be severely affected by contamination and by a low completeness rate. I therefore explored another more robust method trying to avoid these problems as much as possible.

The colour-colour selection makes use of only three photometric bands, a limited information compared to the full set of magnitude measurements available in the COSMOS-UltraVista area. The full Spectral energy distributions (SEDs) provides information about the properties of these faint galaxies over a broad wavelength range, and one can hope that by adding information compared to only three bands a more accurate redshift information will be gained.

Different physical processes occurring in galaxies all leave their signatures on the spectrum, each dominating at different wavelengths. Detailed analysis of the SED of a galaxy should allow us to fully understand the properties of that galaxy. SED fitting is thus the attempt to analyse a galaxy SED and to derive one or several physical properties simultaneously from fitting models to an observed SED. This is in contrast to searching a single feature that could constrain a single parameter (a prominent example would be the use of recombination line, such as Hα to derive the star formation rate). Now SED fitting is an almost universally used technique that has matured significantly in recent years. Photometric redshift are the primary measurements obtained from SED fitting (e.g. Ilbert 2006, 2009, 2013). With the multiple bands of COSMOS-UltraVista, photometric redshift accuracy and catastrophic errors in measurements are considerably reduced as shown in Figure 5.5.

Selecting high redshift samples of galaxies with photometric redshifts is now more and more used as the depth and number of bands in several well studied surveys fields are improving. Using photometric redshifts greatly helps to minimise contamination of a sample by low redshifts galaxies, and hence maximises purity. Because a broad SED is observed, the photometric redshift of faint galaxies is more stable than a crude estimate based on the detection of a continuum break with three bands. For these reasons I have selected this method to select the galaxies of my final sample.

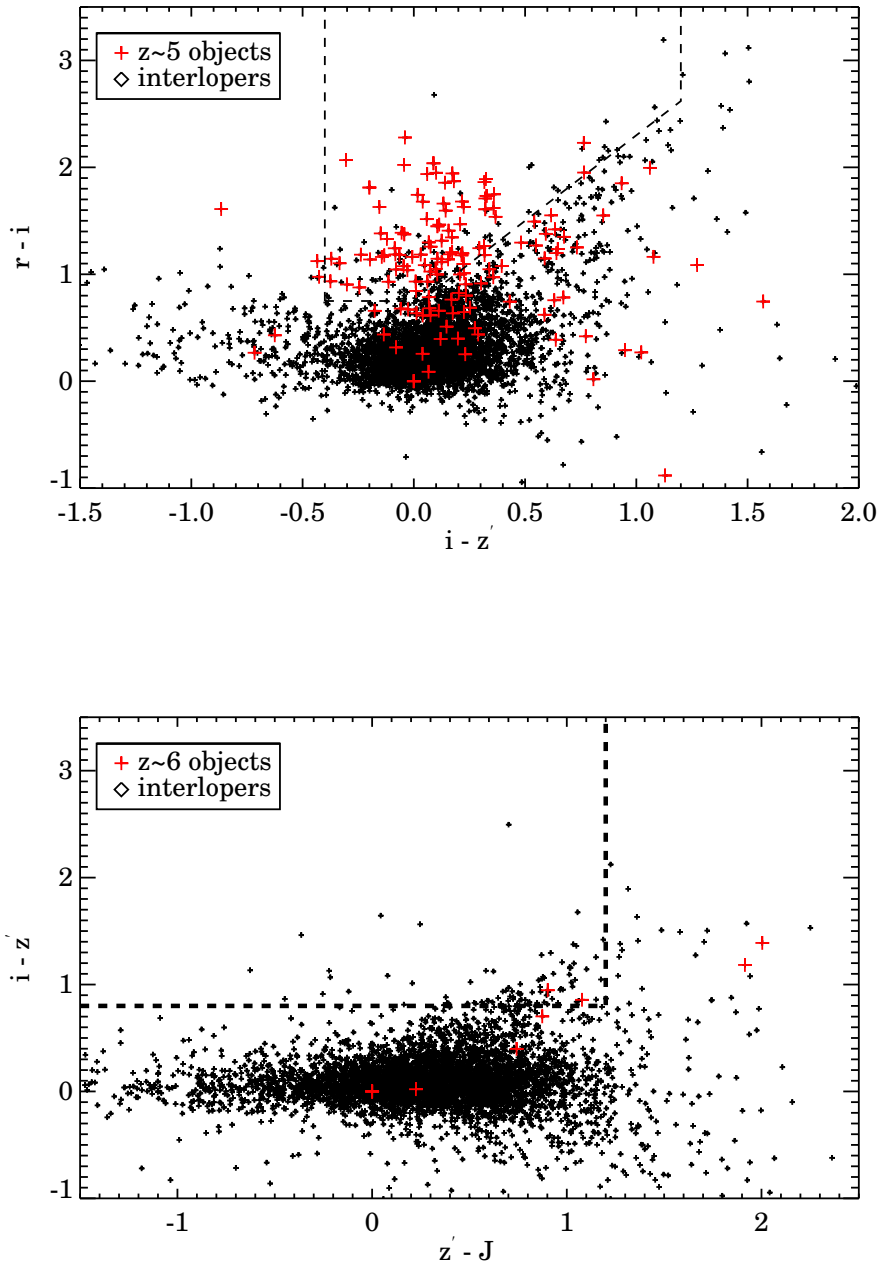


FIGURE 5.4: $(r-i)$ vs. $(i-z')$ colour-colour diagram used to select r-dropout galaxies at $4.5 \leq z \leq 5.5$. The black dashed line box shows the selection box defined using galaxy tracks as defined in Figure 5.1. The red-starred points are galaxies securely identified in this redshift range from the VUDS spectroscopic sample. We can immediately see that the colours of galaxies in this range cover a larger range than identified from tracks. I use the VUDS sample to compute the completeness rate making the ratio of VUDS galaxies with $4.5 \leq z \leq 5.5$ inside the box over the total number of VUDS galaxies in this range. I can also identify what fraction of galaxies at redshifts outside $4.5 \leq z \leq 5.5$ are contaminating the selection (black points in the box). (bottom) $(i-z')$ vs. $(z'-J)$ colour-colour diagram used to select i-dropout galaxies at $5.5 \leq z \leq 6.5$. The black dashed line box shows the selection box defined using galaxy tracks as defined in Figure 5.1. The red-starred points are galaxies securely identified in this redshift range from the VUDS spectroscopic sample. I can also identify what fraction of galaxies at redshifts outside $5.5 \leq z \leq 6.5$ are contaminating the selection (black points in the box).

There are several photometric redshift codes, such as EAZY (Brammer, van Dokkum & Coppi, 2008), Hyperz (M. Bolzonella, J.−M. Miralles & R. Pelló, 2000), Le Phare, ZEBRA (Feldmann, Carollo et al. 2006) which are freely distributed to the astronomical community. In my thesis, I use Le PHARE which was designed by Stéphane Arnouts & Olivier Ilbert. LePhare (PHotometric Analysis for Redshift Estimation) is a set of Fortran programs to compute photometric redshifts using the SED fitting technique. The package is composed of three parts:

- A preliminary phase to select the SED models, the set of filters and to compute the template magnitudes, using stand-alone programs. They allow to extract basic informations relative to the filters(λmean, AB-corrections, attenuation, ...) and SEDs (k-correction vs z, colour-colour diagrams and etc.).
- The photometric redshift code, based on a simple χ^2 fitting method.
- A generator of realistic multicolour catalogues taking into account observational effects.

5.3 Selection of galaxies at $z \geq 4.5$ in the UltraVista COSMOS survey

5.3.1 Overview of the adopted method

The SED-fitting selection method I have devised can be summarised in the following points:

- Perform the multi-band matched photometry, as described in Chapter 4
- Define SED-fitting parameters and templates, using Le Phare and BC03 models
- Perform the SED fitting with Le Phare, including photometric redshifts measurements as well as stellar mass, star formation rates and dust extinction
- Apply same magnitude cut as I apply in colour-selection as described in section 5.2.2
- Visually examine ALL candidate galaxies in the selected redshift range and classify them in terms of their probability to be at the redshift of interest. This is done by two observers separately, and their values are reconciled

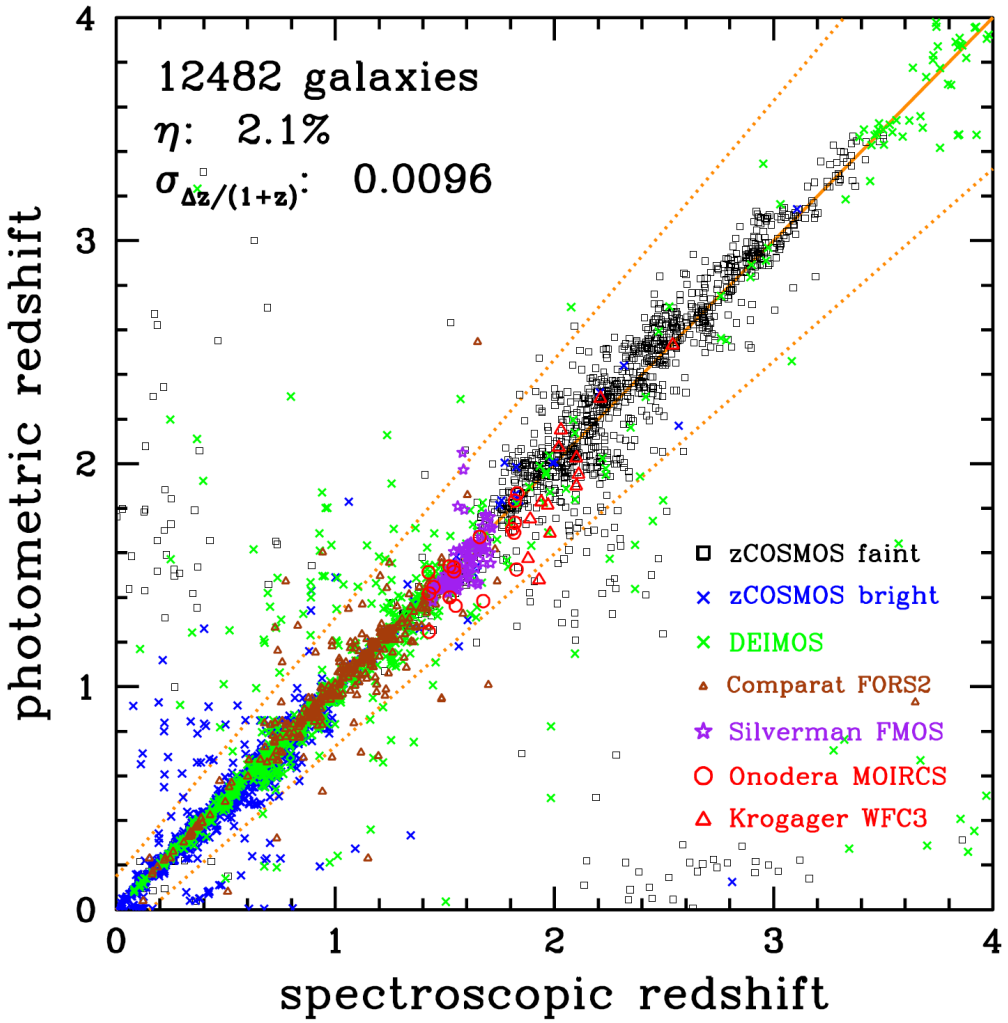


FIGURE 5.5: Current state of the art in photometric redshift measurements in the COSMOS field (from Ilbert et al. 2013)

- Produce a final catalogue with all observables.

These steps are detailed below.

5.3.2 Computing photometric redshifts

I used "Le Phare" to complete full SED-fitting analysis to derive the redshift-probability distribution for each galaxy candidate. In this study, the Bruzual & Charlot (2003) stellar evolution models (BC03) are employed in my analysis, considering models with metallicities ranging from 1 solar metallicity to 0.02 solar metallicity. The ages of the stellar populations models were allowed to range from 10 Myr to 14 Gyr, but were required to be less than the age of the Universe at each redshift. The star formation histories considered were exponentially declining with characteristic time-scales in the

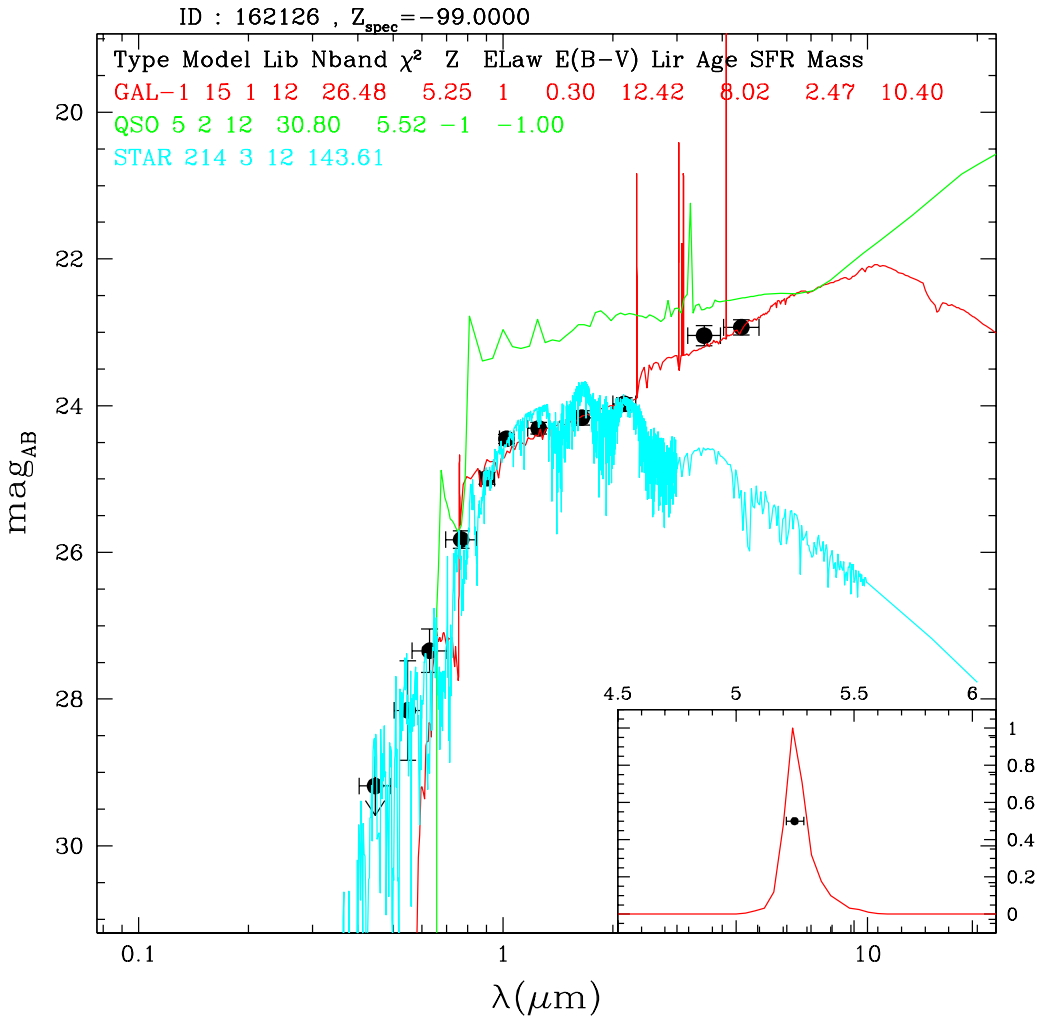


FIGURE 5.6: SED of a high redshift galaxy candidate with $z_{phot} = 5.25$. It shows how difficult it is to discriminate dwarf stars and high redshift galaxy candidates: only the mid-infrared Spitzer data is able to confirm the high redshift nature of the galaxy.

range 30 Myr to 30 Gyr. Dust reddening was described by the Calzetti et al. (2000) attenuation law, and allowed to vary within the range 0 < E(B-V) < 0.6. To discriminate between galactic cool star and real high redshift objects it is important to include the fitting of the stellar spectral type of L, M and T dwarf stars. Templates also include emission lines as computed by Le Phare from the UV-continuum and standard Case-B photo-ionisation. I emphasise the importance to include 2 mid-infrared (3.6μm and 4.5μm) band from Spizter which can strongly assist us in the SED fitting and classification process. We show in Figure 5.6 that these two bands are the key information able to discriminate between cool stars and high redshift galaxies.

5.3.3 Contamination: Lower Redshift Galaxies and Stars

In this section I discuss the possible contamination producing degeneracies between high redshift objects in the redshift range of interest and other categories of astronomical objects. There are three main contaminants with SED properties close to high redshift candidates in my redshift range ($4.5 < z < 6.5$):

- Foreground extreme red galaxy
- Dwarf stars
- Spurious Sources in the photometric process

Foreground dusty galaxies in the redshift range between $1 < z < 3$ can have a red continuum mimicking a break in the continuum that the SED-fitting would interpret as the Lyman-break dropout. The SED of such a typical low redshift interloper is shown in Figure 5.7. Combined with photometric errors this degeneracy is sometimes hard to solve, and colours of some low-z galaxies could also meet the selection criteria for high redshift due to photometric errors. Theses cases may impact dramatically the high redshift galaxy counts if simply taking the SED photometric fitting by Le Phare at face value. Indeed as low redshift galaxies largely dominate in numbers even a small number of incorrect photometric redshifts placing low redshift galaxies incorrectly at high redshifts can strongly affect the galaxy counts in the redshift range of interest. This pushes for a careful visual examination of the SED of each single candidate.

The other main contaminants are galactic cool stars with M−, T− and L− stellar types mostly. Using photometric stellarity criteria MU_MAX/MAG_AUTO plot (Figure 4.9) can help us to remove the bright stellar objects, but it does not work well on removing most of the fainter galactic cool stars. Galactic stars can then satisfy pre-selection criteria and photometric redshift estimation to be compatible with a high redshift galaxy candidate, especially getting worse in the high-end of my redshift bins ($z > 6$). It does not mean SED fitting is not crucial anymore, it only means that the selection can not just depend on photometric redshift without visual checking individually. Figure 5.8 shows a typical dwarf star SED fitting plot which shows the great similarity with the SEDs of high redshift objects. Compare to SEDs of low redshift interlopers, the degeneracy can not be broke by photometric redshift easily without visual inspection. As mentioned in section 5.3.2, two things needs to be included before starting SED fitting. First thing is the well-complete dwarf star library which helps us recognise dwarf stellar SEDs ,and second is two mid-infrared band which helps us constrain my SED fitting better (3.6μm and 4.5μm of IRAC). Finally I do visually classification for those candidates after

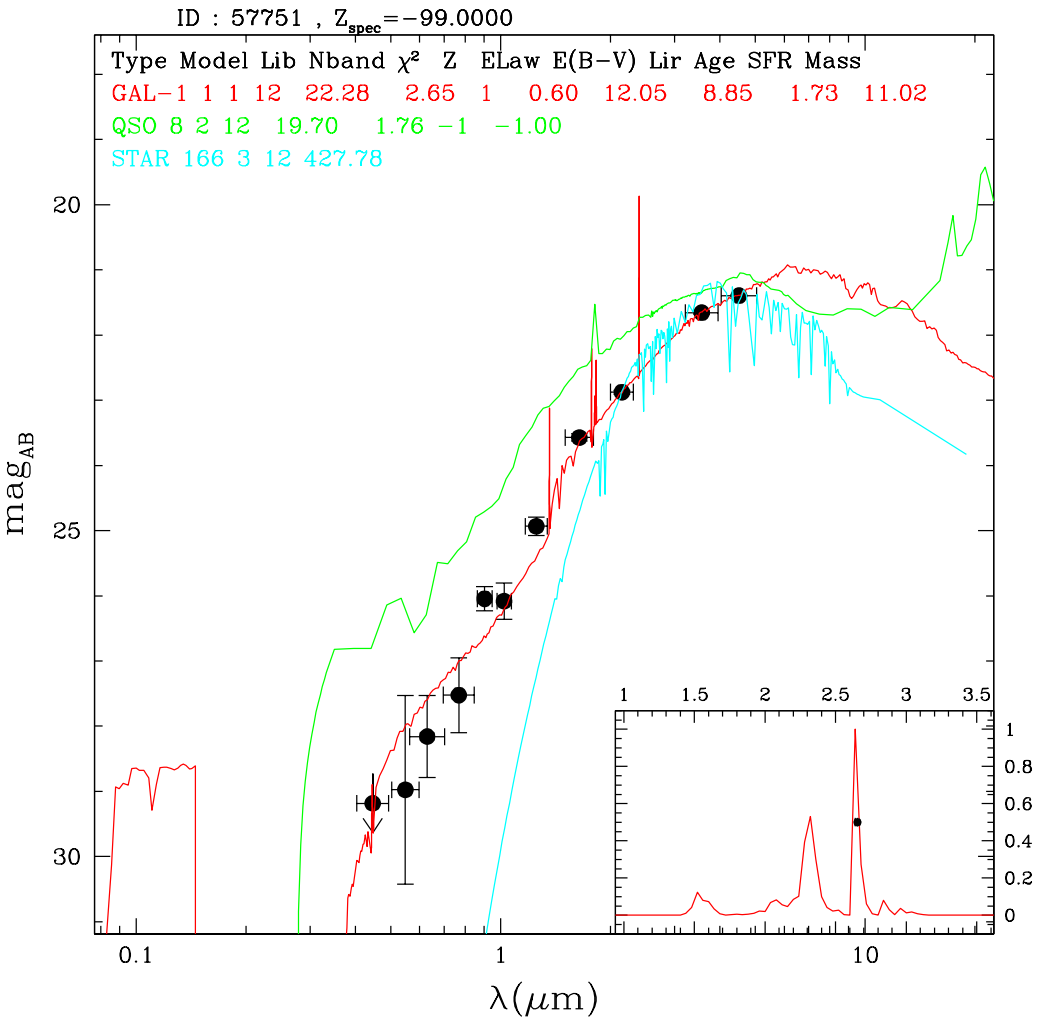


FIGURE 5.7: SED plot of a massive galaxy with extreme red colours. The 12 COSMOS-UltraVista photometric broad bands show a strong break as high redshift LBG.

photometric redshift fitting, human eye still perform efficiently precise for distinguishing faint galactic dwarf and high redshift candidates as shown in Figure 5.8 and Figure 5.9.

Compare to two main contamination mentioned before, the rest of factors is the minority. From the start when I built the parent catalogue, I mask out the bright objects and its spikes, cosmic rays, boundary region with very low signal-to-noise level and artefacts on the images. The further MU_MAX/MAG_AUTO plane (Figure 4.9) is also help us to remove part spurious sources. Then, after pre-selection criteria with σ detection level on more than 3 near-infrared bands and visual inspection after SED fitting, all the procedure suggest that my candidates are not spurious sources.

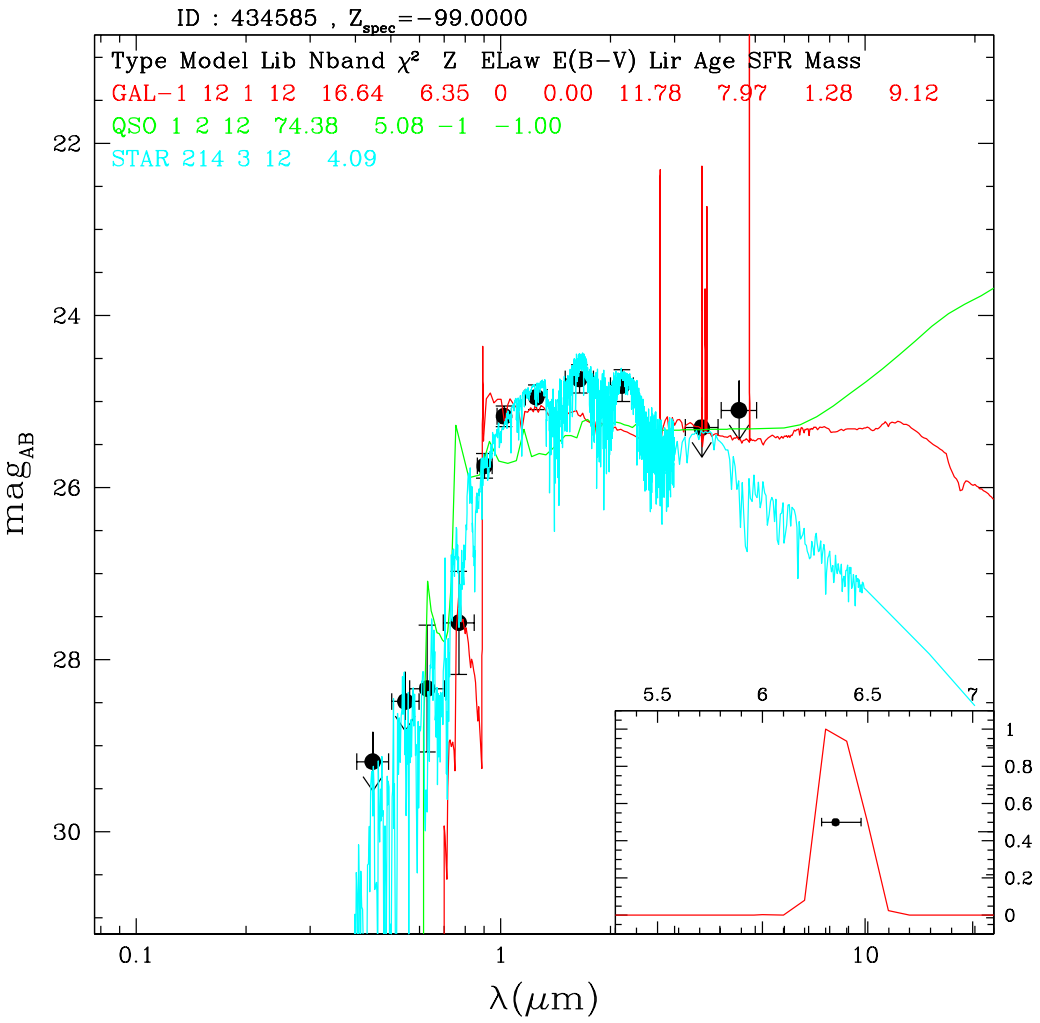


FIGURE 5.8: One of the typical SED plot of galactic cool star. It can still be present just rely on simple colour selection without SED analysis.

5.3.4 Visual Inspection

There may be many sophisticated algorithmic ways to help solve the degeneracy problem, but each of them would add further complications in understanding the ability to eliminate contamination and hence know the selection function. Here I am advocating the importance of visual inspection, especially when we need to check some tiny features on SED and images which can not be pre-targeted by any automatic procedure. Given the manageable number of <10 000 objects pre-selected on the basis of the SED-based photometric redshifts, I selected to visually examine and classify each single object candidate.

To minimise biases linked to one single person, two independent persons (me & Olivier Le Fèvre) classified all candidates separately and their classification was then combined

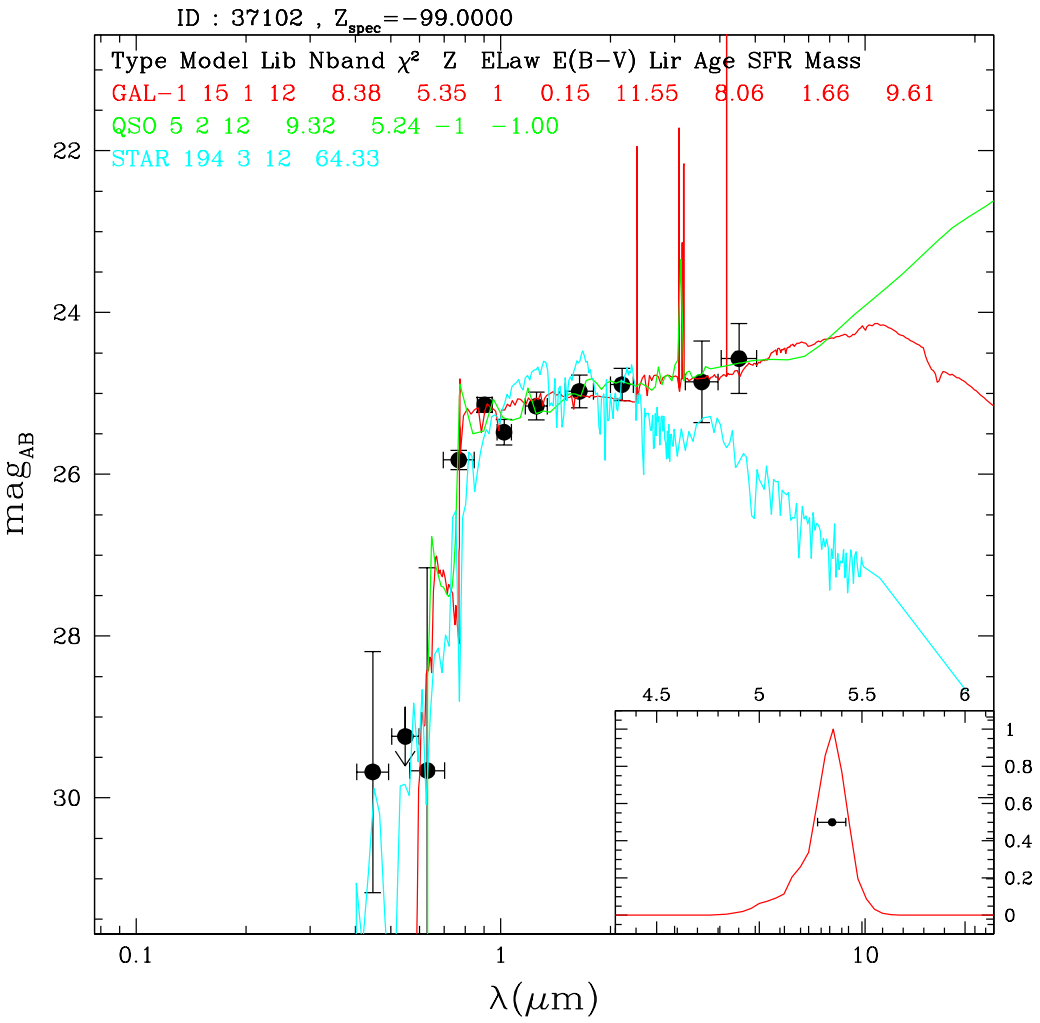


FIGURE 5.9: SED plot of a reliable candidate at $z = 5.35$. There is a well defined continuum break, and a continuum slowly rising to the red. This would be visually classified as a flag 1.

into one single value per object. Table 5.4 presents the list of flags used to classify the different types of objects expected in our sample. Table 5.5 provides the numbers of candidates changes after SED fitting compare to the numbers just depends on colour-colour box and sigma threshold which have been described in section 5.2.2. As can be seen the SED fitting process combined with visual inspection helped us to reduce by at least $\sim 50\%$ the number of interlopers in this high-z study. For the computation of the luminosity function in next chapters, only flag 1, 1.5, 1.7 and 2 are used. In Appendix A, I also provide the images and plots of SED-fittings of several credible candidates (Flag=1).

TABLE 5.4: Classification of candidates after sed-fitting by LePhare .

Flag (After SED-Fitting)	Description
1	Highly-credible candidate
2	Credible candidate
3	Degenerate object
4	Highly-credible low-redshift interloper
5	Credible low-redshift interloper
6	Star
7	QSO-like object
8	Unclassified
1.5	One inspector assigned flag 1 and another assigned flag 2
1.7	One inspector assigned a flag 1

TABLE 5.5: Candidates at each step of selection.

Redshift bin	Colour-selection	SED-fitting	Visual inspection
$4.5 < z < 5.5$	5,416	3,506	2,036
$5.5 < z < 6.5$	2,510	1,031	330

5.3.5 Comparison of colour-colour and Photometric redshift selection

I show in Figures 5.10 and 5.11 colour-colour plots for the sample selected with photometric redshifts, and compared to the VUDS sample. It is clear that the sample selected with photometric redshifts spans a larger colour-colour space than the dropout box defined from simulations.

5.4 The UltraVista $z \geq 4.5$ galaxy sample

After following the selection method outlined above, I obtain a sample of robust $z \geq 4.5$ galaxy candidates. The best sample used for the computation of the luminosity function in next chapters is made of flags 1, 1.5, 1.7 and 2. The final number of galaxies in each redshift bin is indicated in Table 5.5. After my selection, I assembled a sample of 2036 galaxies with $4.5 \leq z \leq 5.5$, and 330 galaxies with $5.5 \leq z \leq 6.5$. This is an unprecedented large sample of bright galaxies at these redshifts. The properties of galaxies in these samples are described in the next chapter.

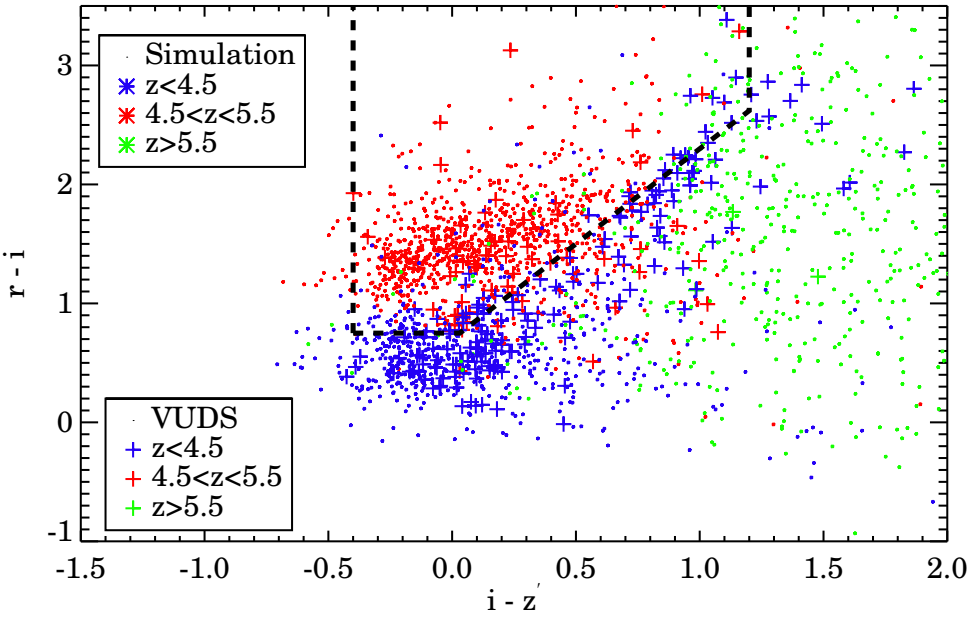


FIGURE 5.10: riz' colour−colour diagram for the simulated galaxies with $4.5 \leq z \leq 5.5$ represented by dots. The black dashed-line represents the colour−colour selection box as defined from simulations. VUDS galaxies in the same redshift range are shown as plus. Lower redshift objects (z < 4.5) are shown as blue points, and higher redshift objects (z > 5.5) are shown as green points.

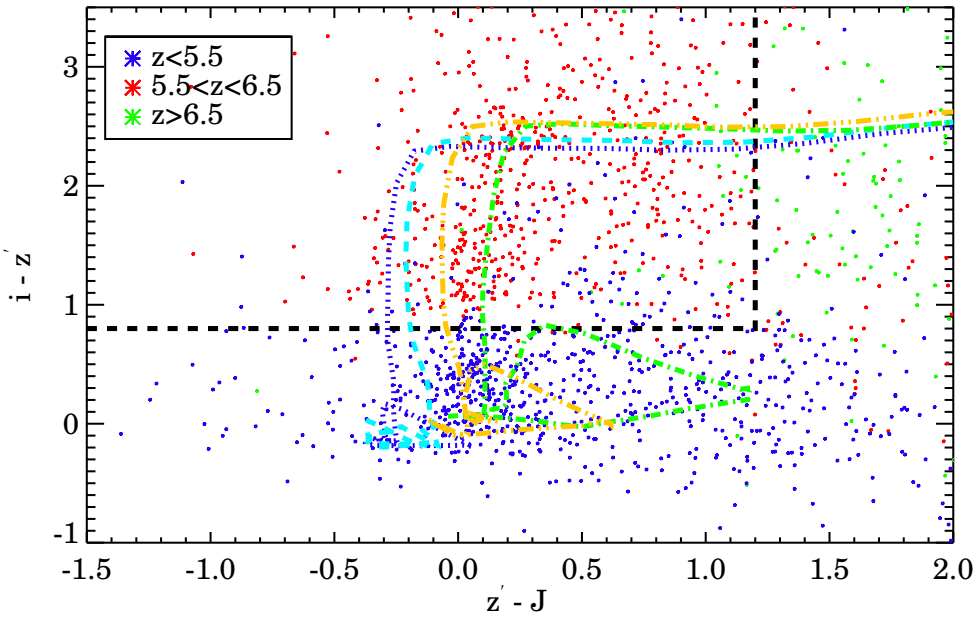


FIGURE 5.11: $iz'J$ colour-colour diagram for the simulated galaxies with $5.5 \leq z \leq 6.5$ represented by asterisk. The black dashed-line represents the colour−colour selection box as defined from simulations. Our coloured dashed lines shows 4 galaxy tracks based on different combinations by stellar age and metallicity from BC03 SED template.

Chapter 6

Sample Properties

In this chapter I present the properties of the high redshift galaxy sample identified in Chapter 5. To characterise the sample I present the redshift distribution of the high redshift candidates ($4.5 < z < 6.5$), the distribution of intrinsic properties including colour, size, absolute magnitude, stellar mass and star formation rate. These properties are derived respectively in two redshift bins: $4.5 < z < 5.5$ and $5.5 < z < 6.5$. These properties are giving an overview of the parameter space probed by this high redshift study.

6.1 Redshift and Apparent Magnitude Distribution

Each of the different methods used for the selection of high redshift galaxy candidates leads to a specific redshift distribution N(z). As our selection is mostly based on photometric selection and flux limit, we expect that the redshift distribution follows a shape as identified in e.g. Le Fèvre et al. (2013). Without redshift limits, a magnitude limited sample down to magnitudes AB ~ 25 is expected to show a redshift distribution peaking at $z \sim 1 - 1.5$, and a high redshift tail with a continuously decreasing number of galaxies going to higher redshifts (see Figure 6.1). Indeed the luminosity of distant sources decreases following the cosmology luminosity distance, hence sources of constant intrinsic luminosity will increasingly drop-out from a magnitude limited sample as redshift is increasing. Lyman-break selected samples have a distinct redshift distribution (Figure 6.2), as produced by the colour-colour selection when the Lyman-break goes through the filters. As our data is basically selected by limiting magnitude in each of two redshift bins, we rather expect a redshift distribution of our sample following Figure 6.1, i.e. a decreasing number of galaxies with increasing redshift in each of our bins. This section

presents the observed redshift distribution and apparent magnitude distribution of our sample in each specific redshift bins.

6.1.1 Redshift∼5

The redshift distribution of the 2036 candidates in this redshift range is presented in Figure 6.3. We observe that the number of galaxies continuously decreases with redshift in the range $4.5 < z < 5.3$, as expected. This is an indication that our sample is representative of the galaxy population in this redshift range.

For redshifts $5.3 < z < 5.5$ while we still expect the number of galaxies to fall down, we note a slight increase in the number of sources in our candidate sample. To check whether this could be the result of an observational bias we compare the redshift distribution of the most reliable candidates (those with flags 1 and 1.5) to our less reliable ones (with flag 2) in the lower panel of Figure 6.3. We see that the excess counts in this redshift bin is mostly produced by the flag 2 candidates, and it is therefore likely that a fraction of these has an incorrect redshift as a result of photometric redshift degeneracy. However, this concerns a small fraction, about 5%, of galaxies with $4.5 < z < 5.5$, and hence we do not expect this bias to affect the results on the luminosity function as presented in Chapter 7. We also point out that the except counts in $5.3 < z < 5.5$ might be the result of cosmic variance produced by overdensities in this still relatively small field.

Here we use the $z'-$band magnitude which is the critical band for selecting $z \sim 5$ candidates to investigate the magnitude distribution in this redshift range ($4.5 < z < 5.5$). The distribution of apparent magnitudes with redshift is shown in Figure 6.4. In all $\sim 98\%$ of the distribution is within $24 < z' < 27$, while $\sim 2\%$ have $z' > 27$. Figure 6.5 shows the K_s magnitude distribution versus redshift of the same sample. Similarly to the z' band distribution, the majority of $z \sim 5$ candidates are within $24 < K_s < 27$. We note that compared to the z' band distribution, some candidates are much fainter in K_s. This indicates that one should not apply a too restrictive detection level on near-infrared detection for bands redder than the band just redder than Lyman-α: for a star-forming galaxy with little or no dust, the slope of the UV rest-frame continuum is very blue and while a galaxy can be detected just redder than Lyman-α, its flux may drop sufficiently that it may not be detected at redder wavelengths.

6.1.2 Redshift∼6

The redshift distribution of the 330 candidates in this redshift range is presented in Figure 6.6. We observe that the number of galaxies continuously decreases with redshift

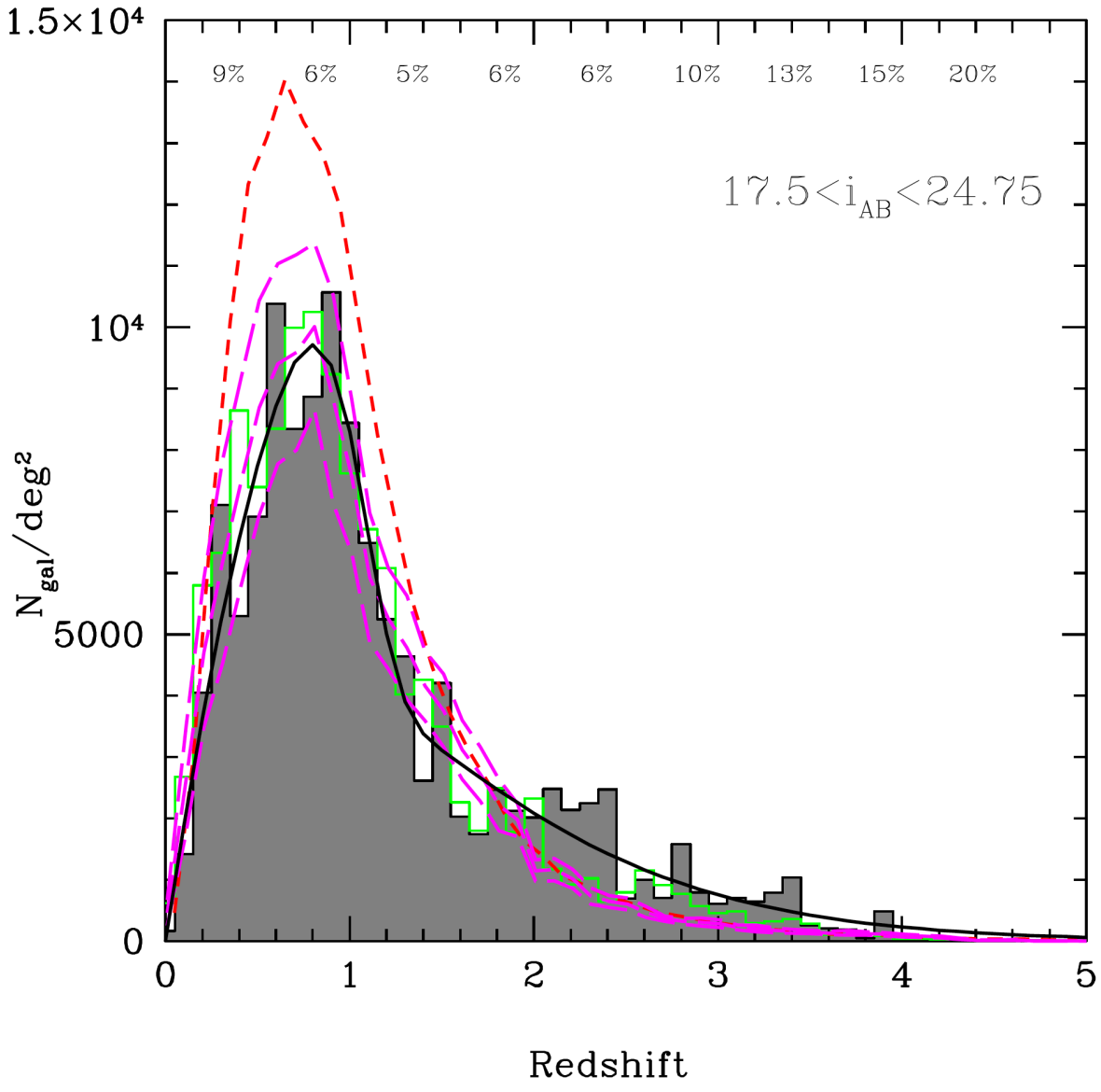


FIGURE 6.1: Spectroscopic redshift distribution N(z) (number of galaxies per square degree) of galaxies with $17.5 \leq i_{AB} \leq 24.75$, using the VVDS-Deep (10 044 galaxies) and VVDS-UltraDeep samples (721 galaxies), in 0.61 deg^2. The N(z) from the De Lucia and Blaizot (2007) SAM based on the Millennium simulation using the WMAP1 cosmology is shown as the dotted red line and the SAM based on the latest Millennium-WMAP1 as the dashed magenta lines (Henriques et al., 2012) representing the mean and $\pm 1\sigma$ values from 24 mocks. The open green histogram is the N(z) derived from the updated v2.0 photometric redshift sample from Ilbert et al. (2009), including UltraVista data, on 1.73 deg^2 in the COSMOS field. Estimates of cosmic variance are listed on the top of the plot. (Credit: Le Fèvre et al. (2013)

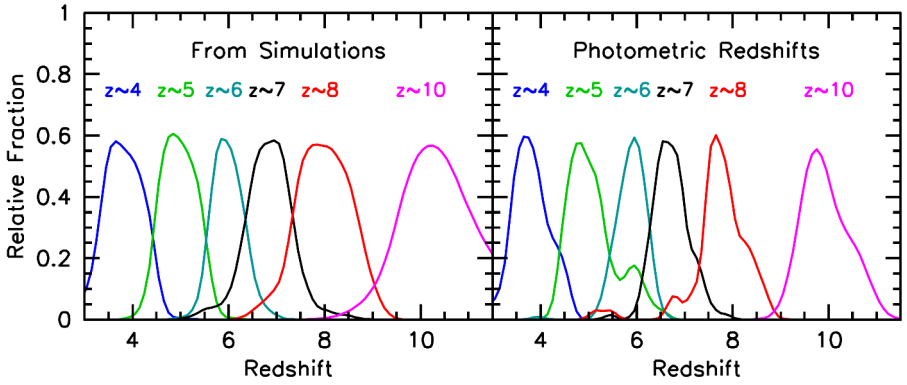


FIGURE 6.2: (left) The expected redshift distributions resulting from the drop-out selection for $z \sim 4, z \sim 5, z \sim 6, z \sim 7, z \sim 8$, and $z \sim 10$ samples. This is produced using Monte-Carlo simulations with the filter set as for the XDF HST observations. The mean redshifts for these samples are 3.8, 4.9, 5.9, 6.8, 7.9, and 10.4, respectively. Each selection window is smoothed by a normal distribution with scatter $\sigma_z \sim 0.2$. (right) Redshift distribution for sources at $z \sim 4, z \sim 5, z \sim 6, z \sim 7, z \sim 8$, and $z \sim 10$ samples using the EAZY photometric redshift code (with similar smoothing as in the left panel). These simulations demonstrate that these filters can effectively isolate galaxies within fixed redshift ranges. It is clear, however, that the bell-shaped redshift distribution is not the real physical distribution of galaxies, but the result of the convolution of the true redshift distribution and the selection function produced by the filters used. (Credit: Bouwens et al. 2014)

in the range $5.5 < z < 6.2$, as expected. This is again an indication that our sample is representative of the galaxy population in this redshift range.

For redshifts $6.2 < z < 6.5$ while we still expect the number of galaxies to fall down, we note a slight increase in the number of sources. To check whether this could be the result of an observational bias we compare the redshift distribution of the most reliable candidates (those with flags 1 and 1.5) to our less reliable ones (with flag 2) in the lower panel of Figure 6.6. We see that the excess counts is produced by the flag 2 candidates, and it is therefore likely that a fraction of these has an incorrect redshift as a result of photometric redshift degeneracy. However, this concerns a small fraction of the galaxies with $6.2 < z < 6.5$, and hence we do not expect this bias to affect the results on the luminosity function as presented in Chapter 7.

The distribution of apparent Y-band magnitudes is shown in Figure 6.7. Figure 6.8 shows the distribution of K_s magnitude with redshift of same $z \sim 6$ samples. Some candidates are much fainter in K_s, and similarly to the $z \sim 5$ sample we note that one should not apply a too restrictive detection level on near-infrared detection for bands redder than the band just redder than Lyman-α.

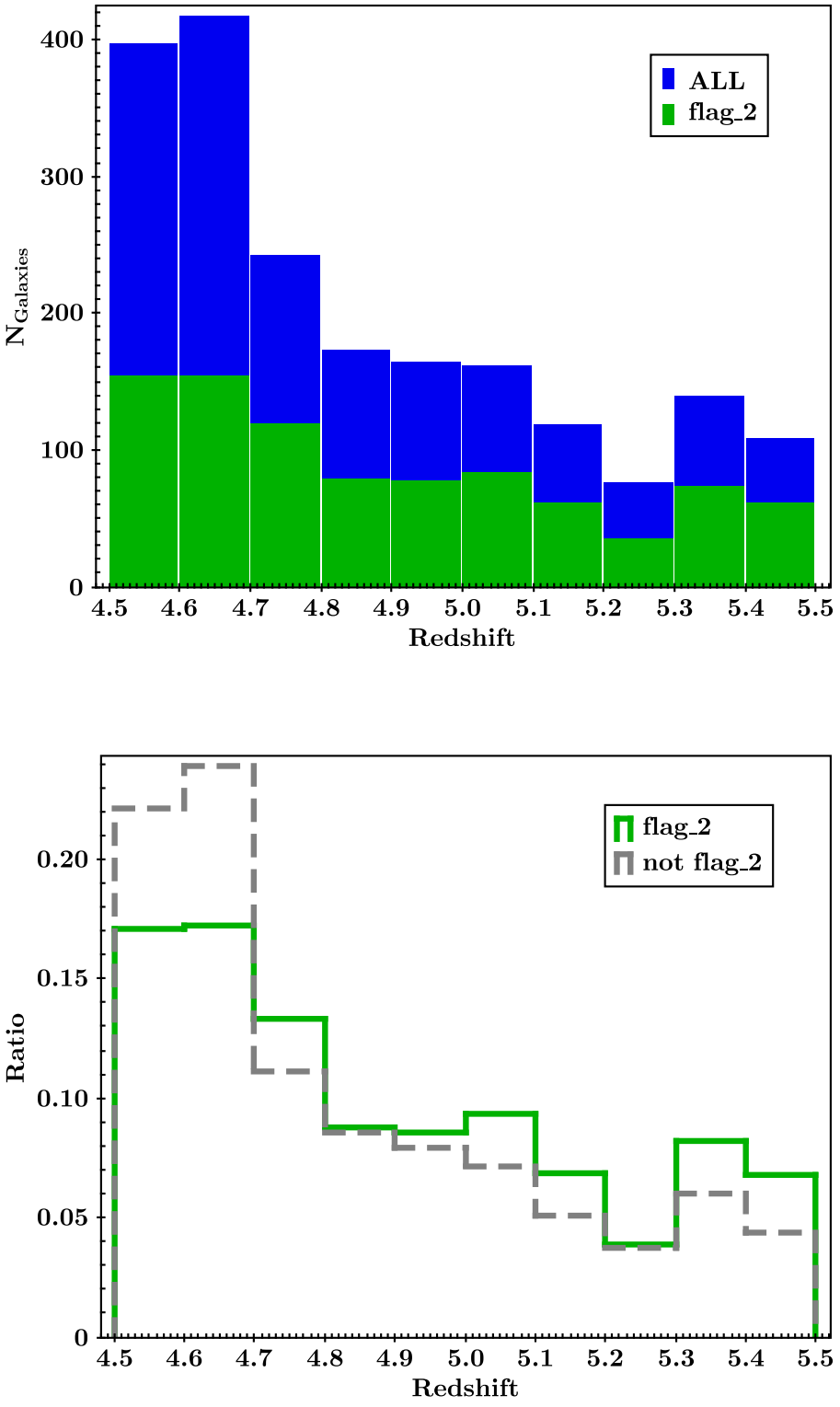


FIGURE 6.3: (Upper panel) The redshift distribution of $z \sim 5$ UltraVista galaxies satisfying all our criteria. (Lower Panel) The redshift distribution of $z \sim 5$ UltraVista galaxies divided into 2 subsamples (flag=2 and flag=1 or 1.5). The grey line shows the distribution of galaxies with flag=1 and flag=1.5, while the green line shows the distribution of galaxies with flag=2.

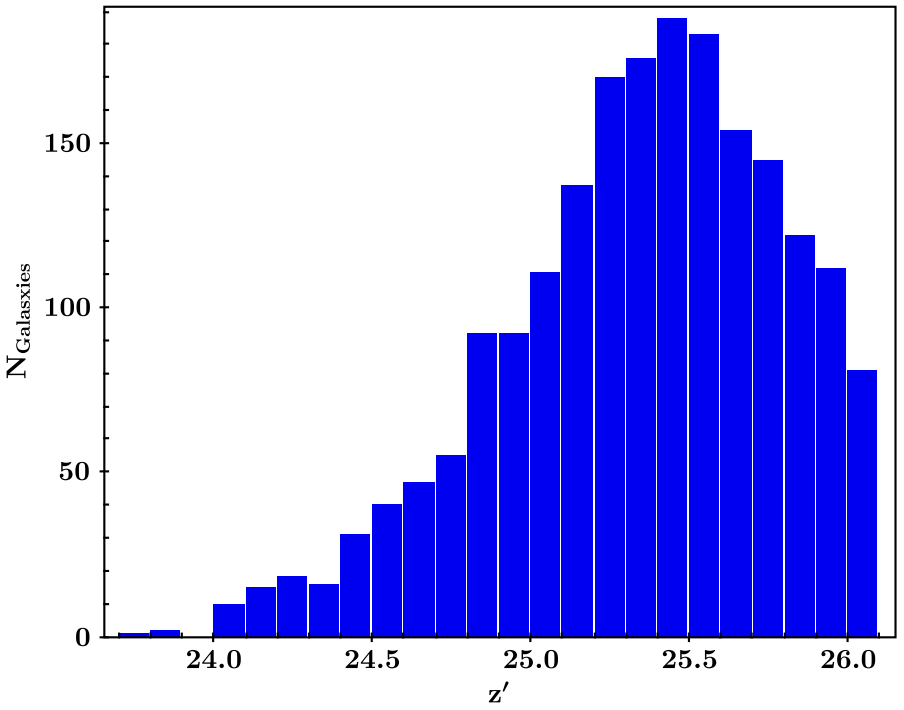


FIGURE 6.4: The $z' - band$ distribution of all the $z \sim 5$ UltraVista galaxies.

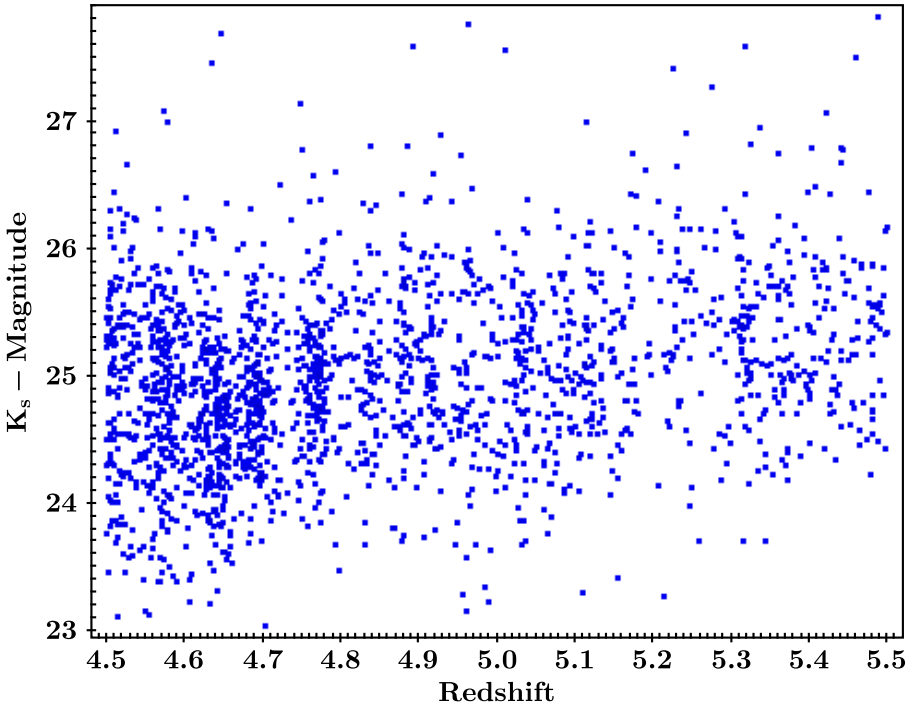


FIGURE 6.5: The distribution of $K_s - band$ magnitude vs. redshift for all the $z \sim 5$ UltraVista galaxies.

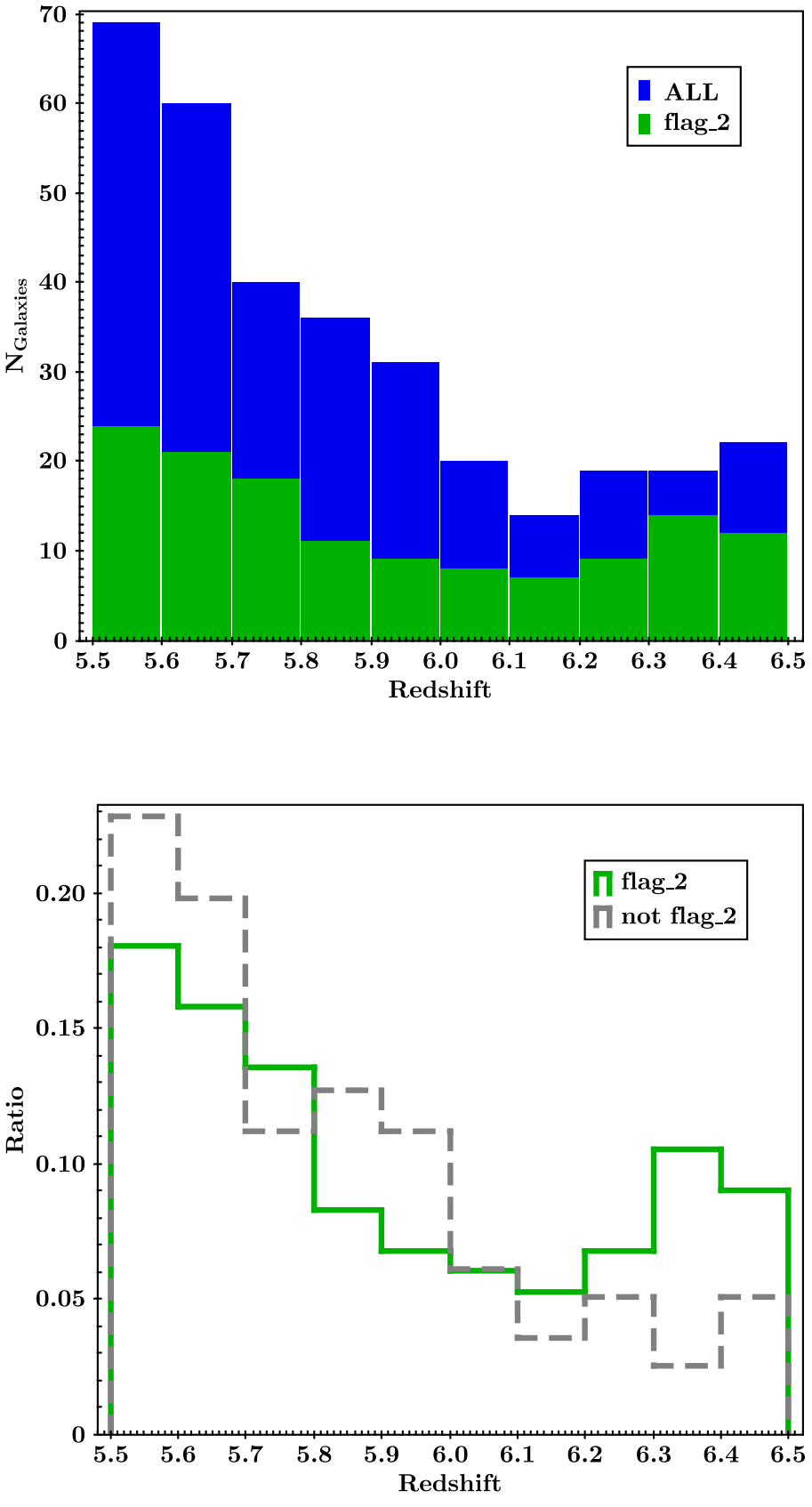


FIGURE 6.6: (Upper panel) The redshift distribution of $z \sim 6$ UltraVista galaxies satisfying all our criteria. (Lower Panel) The redshift distribution of $z \sim 6$ UltraVista galaxies divided into 2 subsamples (flag=2 and flag=1 or 1.5). The grey lines shows the distribution of galaxies with flag=1 and flag=1.5, while the green line shows the distribution of galaxies with flag=2.

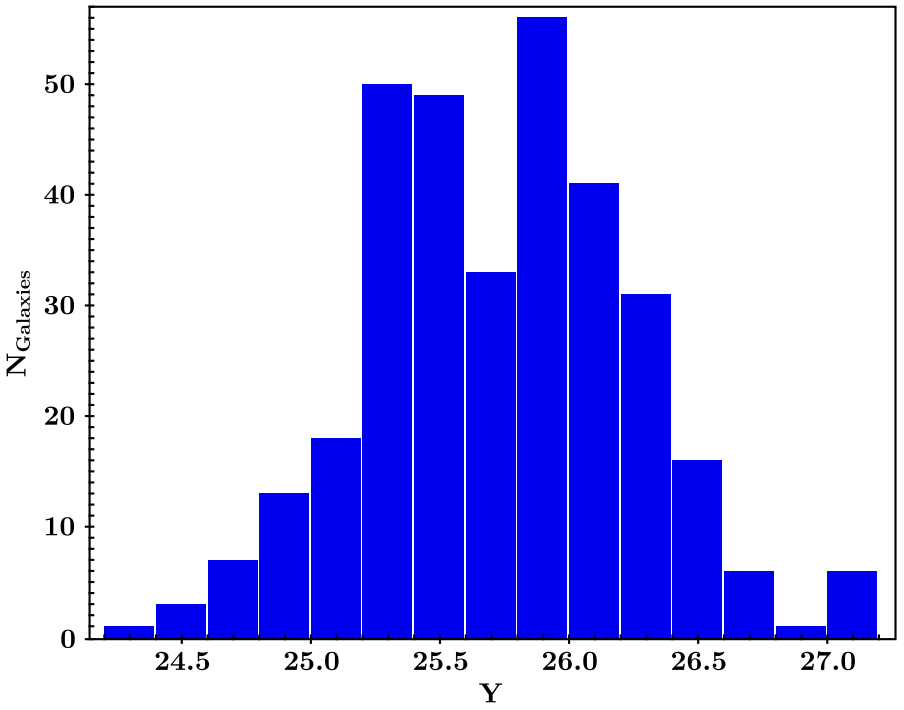


FIGURE 6.7: The $Y-band$ distribution of all the $z \sim 6$ UltraVista galaxies.

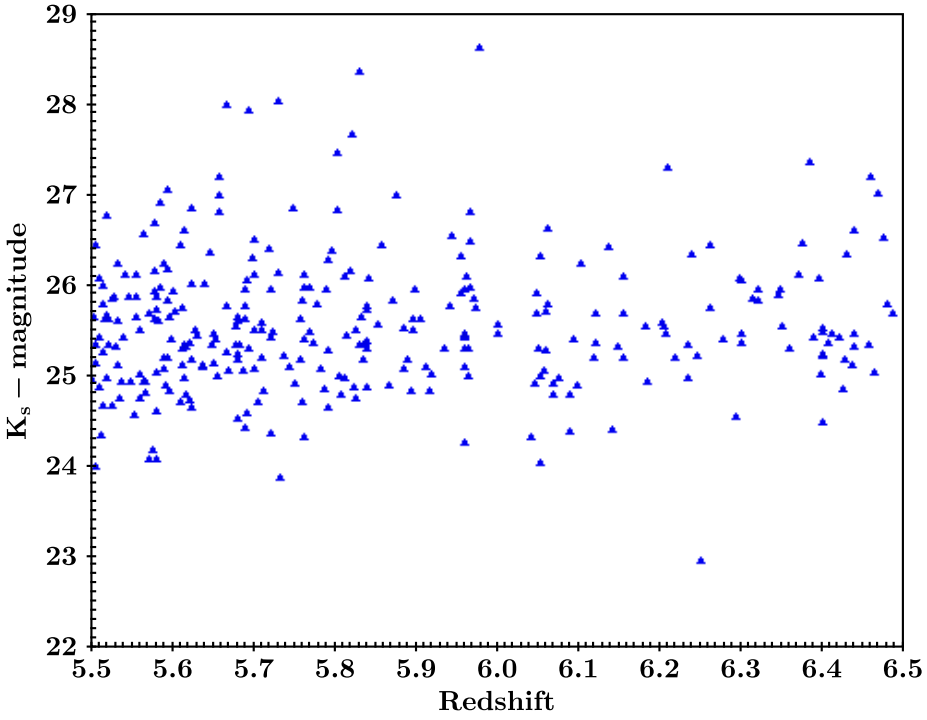


FIGURE 6.8: The distribution of K_s-band vs. redshift of all the $z \sim 6$ UltraVista galaxies.

6.2 Colour Distribution

As our high redshift sample is selected using photometric redshifts, it is interesting to check a posteriori where theses galaxies are distributed in several standard colour-colour diagrams used in the literature to select high$-z$ galaxies. We examine here the riz' diagram traditionally used to select drop-out galaxies with $4.5 < z < 5.5$ and the $iz'J$ diagram for galaxies with $5.5 < z < 6.5$. Each of these diagrams are discussed in forthcoming sections.

6.2.1 Redshift∼5

Figure 6.9 and Figure 6.10 present the colour distribution and colour-colour plot of our UltraVista redshift ~ 5 candidates, and the dashed lines in these 2 plots represent the traditional colour selection threshold adjusted using galaxy track (described in chapter 5). In our redshift $z \sim 5$ sample, typical colour selection works pretty well as 78% (1581/2036 objects) of our photometric redshift selected sources are selected by the colour-colour selection criteria, so loosing a fraction of the population at the selected redshift with an incompleteness of 22%. The contamination from galaxies at redshifts outside $4.5 < z < 5.5$ in the colour-colour selection area is a much more severe problem as identified in Figure 6.10. 3835 sources are in the selection area but with a redshift outside the target range, which is 2.43 (3835/1581) times higher than the number of galaxies at the right redshift. This shows that the drop-out colour-colour technique used to select galaxies in this redshift range in most studies may be seriously affected by contamination, and that this contamination must therefore be properly estimated and corrected for to obtain reliable galaxy counts. Our selection technique using photometric redshifts is immune to this problem and no specific contamination correction need to be applied to our data.

6.2.2 Redshift∼6

Figure 6.11 and Figure 6.12 present the colour distribution and colour-colour plot of our UltraVista redshift ~ 6 candidates, and the dashed lines in these 2 plots represent the traditional colour selection limits. In our redshift $z \sim 6$ sample, typical colour selection works pretty well as 96% (318/330 objects) of our photometric redshift selected sources are selected by the colour-colour selection criteria, so with an incompleteness of 4%, loosing a small fraction of the population at the selected redshift. The contamination from galaxies at redshifts outside $5.5 < z < 6.5$ in the colour-colour selection area is a much more severe problem as identified in Figure 6.12, 2192 sources are in the

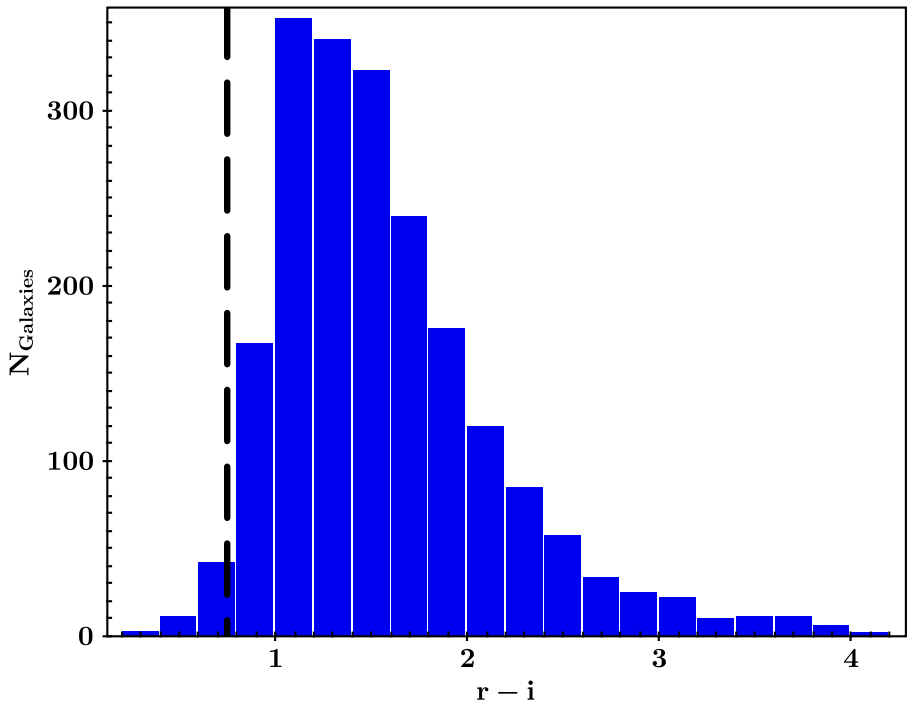


FIGURE 6.9: The typical $r-i$ colour distribution of all redshift$\sim$ 5 UltraVista galaxies, the black dashed line shows the typical $r-i$ threshold of redshift$\sim$ 5 selection.

selection area but with a redshift outside the target range, which is 6.89 (2 192/318) times higher than the number of galaxies at the right redshift. This shows that the drop-out colour-colour technique used to select galaxies in this redshift range in most studies may be seriously affected by contamination, and that this contamination must therefore be properly estimated and corrected for to obtain reliable galaxy counts. Our selection technique using photometric redshifts is immune to this problem and no specific contamination correction need to be applied to our data.

6.3 Size, Absolute Magnitude, Mass and Star Formation Rate

In this section, we summarise several galaxy properties in each redshift bins. To characterise the size of galaxies we use the ISOAREA_IMAGE, an output of SExtractor photometry. We analyse the estimation of physical properties from the SED fitting on the multi-wavelength photometry as the output of the code Le Phare. Galaxy luminosities are transformed into stellar mass using the best fit synthetic template which

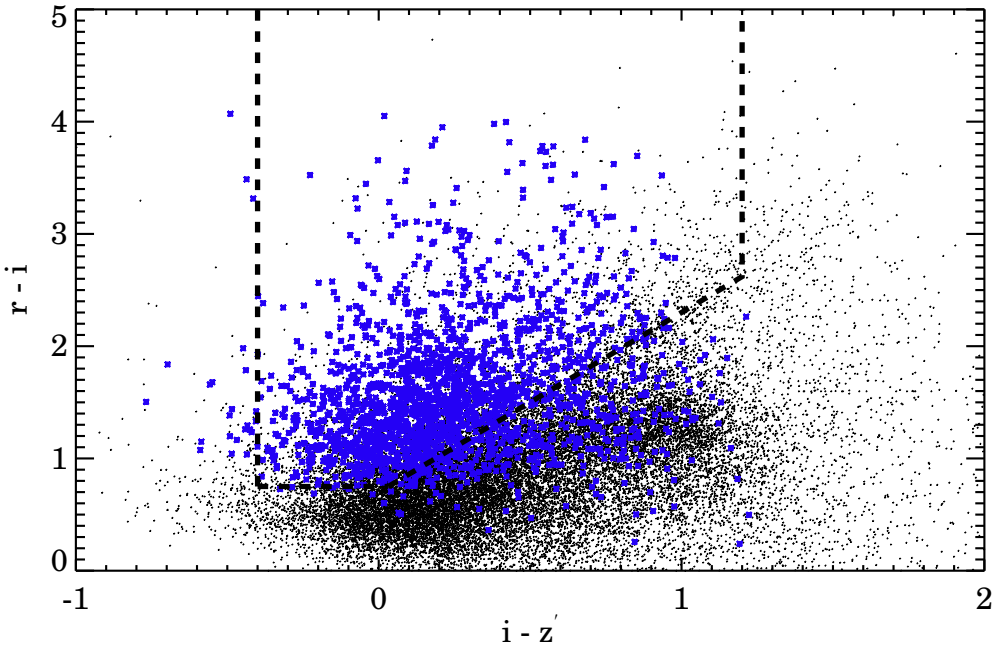


FIGURE 6.10: riz' colour-colour diagram of 2036 UltraVista galaxies (blue dots) with redshift ~ 5 ($4.5 < z < 5.5$), two dashed lines are the typical colour threshold selecting $z \sim 5$ objects. The small black dots are objects outside the redshift range ($4.5 < z < 5.5$). There are 3 835 galaxies in the colour box but with a redshift outside the range of interest. All the galaxies in the plot are limiting in $z' = 26.08$, (5σ value).

is generated using the Stellar Population Synthesis (SPS) model of Bruzual & Charlot(2003). We considered exponentially declining and delayed SFR with 9 possible τ values ranging from 0.1 to 30 Gyr and 3 metallicities (Z = 0.004, Z = 0.008, and solar Z = 0.02). We also assume the Calzetti (2000) extinction law; emission line contributions are included using an empirical relation between the UV light and the emission line fluxes. The output of SED fitting process includes several physical parameters: absolute magnitude in UV, stellar masses and star formation rates. All the distributions of galaxy properties are presented in following sub-sections in each redshift bins, respectively.

6.3.1 Redshift∼5

The distributions in absolute UV (1500Å rest-frame) magnitude, stellar masses, and star formation rate of the redshift ~ 5 samples are presented in Figure 6.14, 6.15 and 6.16. At redshifts $z \sim 5$ the UV absolute magnitude ranges from $M_{1500} = -22.4$ to $M_{1500} = -19.8$, the stellar mass from $10^{7.7} M_{\odot}$ to $10^{11.5} M_{\odot}$ and the majority of star formation rate ranges from below $4 M_{\odot}/yr$ up to several hundred $M_{\odot}/yr$. We find one

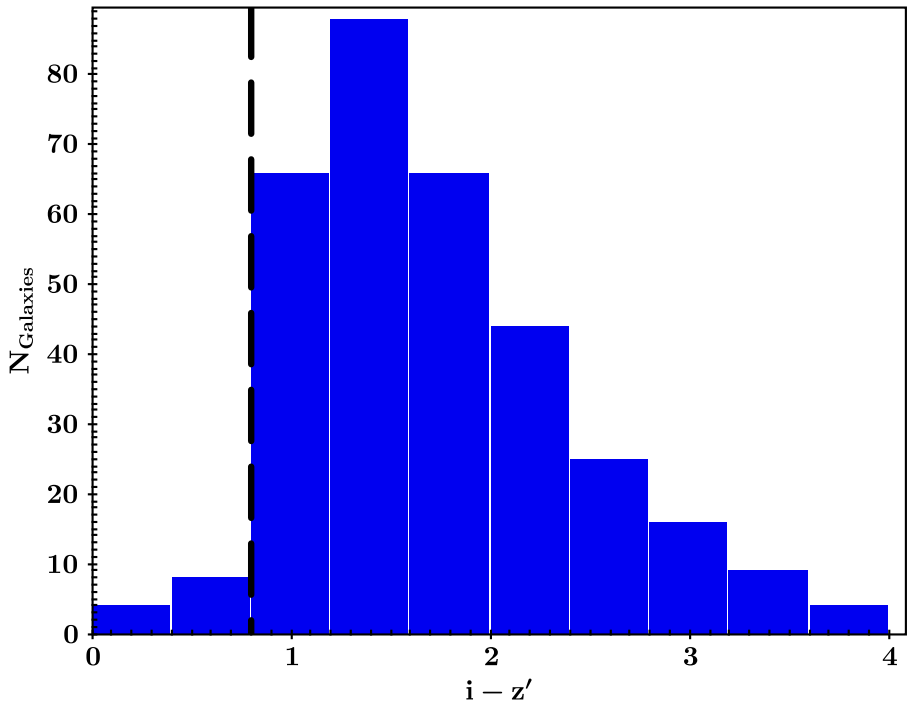


FIGURE 6.11: The typical $i - z'$ colour distribution of all redshift$\sim$ 6 UltraVista galaxies, the black dashed line shows the typical $i-z'$ threshold of redshift$\sim$ 6 selection.

extremely active starburst galaxy with a star formation rate reaching a few thousand solar masses produced in a year. The SFR-M relation for all $z \sim 5$ UltraVista galaxies is shown in Figure 6.17. It is immediately visible that the distribution of redshift ~ 5 galaxies does not follow a single main sequence relation. On average our sample galaxies are more than 1 dex above the local main sequence (MS) of star forming galaxies in the SDSS (Peng et al. 2010) at any redshift considered in this study. A significant fraction of our data also lie above the Daddi et al. (2007) MS at $z \sim 2$.

6.3.2 Redshift$\sim$6

The distributions in absolute UV (1500Å rest-frame) magnitude, stellar masses, and star formation rate of the redshift $\sim$ 6 samples are presented in Figure 6.19, 6.20 and 6.21. At redshifts $z \sim 6$ the UV absolute magnitude ranges from $M_{1500} = -22.4$ to $M_{1500} = -19.8$, the stellar mass from $10^{7.7} M_{\odot}$ to $10^{10.5} M_{\odot}$ and the majority of star formation rate ranges from below $3 M_{\odot}/yr$ up to several hundred $M_{\odot}/yr$. We find one extremely active starburst galaxy with a star formation rate reaching a few thousand solar masses produced in a year. Here we also check the SFR-M relation in redshift

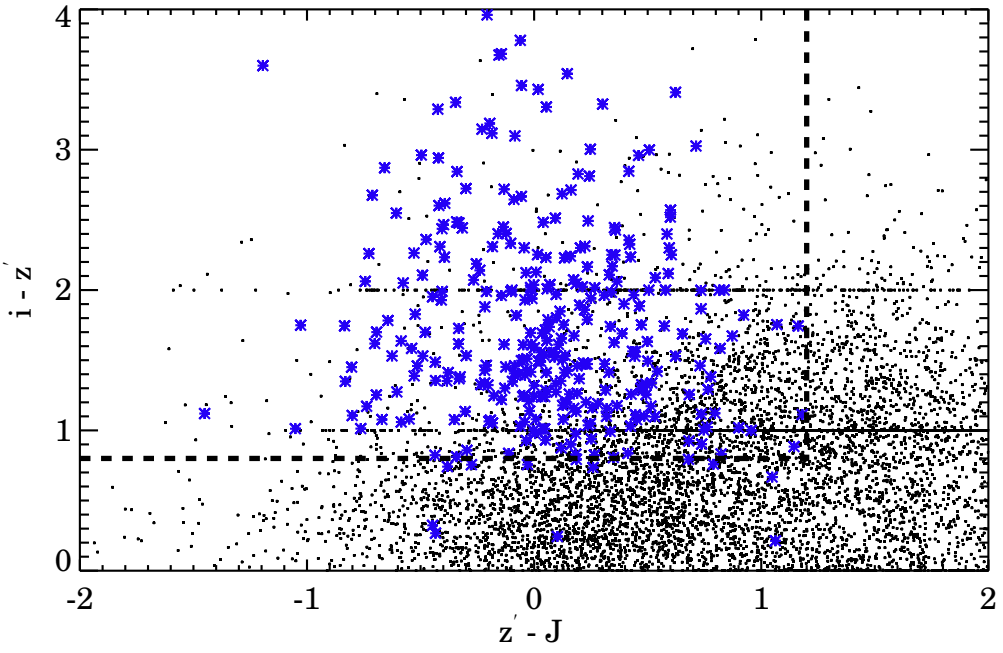


FIGURE 6.12: $iz'J$ colour-colour diagram of 330 UltraVista galaxies (blue asterisks) with redshift ~ 6 ($5.5 < z < 6.5$), the dashed lines are the typical colour threshold selecting $z \sim 6$ objects. The small black dots are objects outside the redshift range ($5.5 < z < 6.5$). There are 2 192 galaxies with redshifts outside the redshift of interest in the selection box. All the galaxies in the plot are limiting in $Y = 26$, 27 (1.5σ value in Deep strips and Ultra Deep strips)

~ 6 (Figure 6.22). Galaxies in this redshift range still more active than MS galaxies in redshift $\sim$1 and $\sim$2 as we found the similar trend in redshift ~ 5.

6.4 Comparison with the VUDS spectroscopic sample $4.5 < z < 6.5$

As our sample is selected solely from photometry, it is important to compare this sample with existing spectroscopic samples. As a member of the VIMOS Ultra Deep Survey (VUDS) team, I have access to the VUDS data. VUDS is the largest spectroscopic survey available today in $2 < z < 6.5$ (see Chapter 4). The VUDS median redshift is $z = 3$, but the high redshift tail goes up to $z \sim 6.5$. I am therefore able to compare the properties of my photometric candidates to the VUDS galaxies in two redshift domains $4.5 < z < 5.5$ and $5.5 < z < 6.5$. VUDS is selected based on photometric redshifts

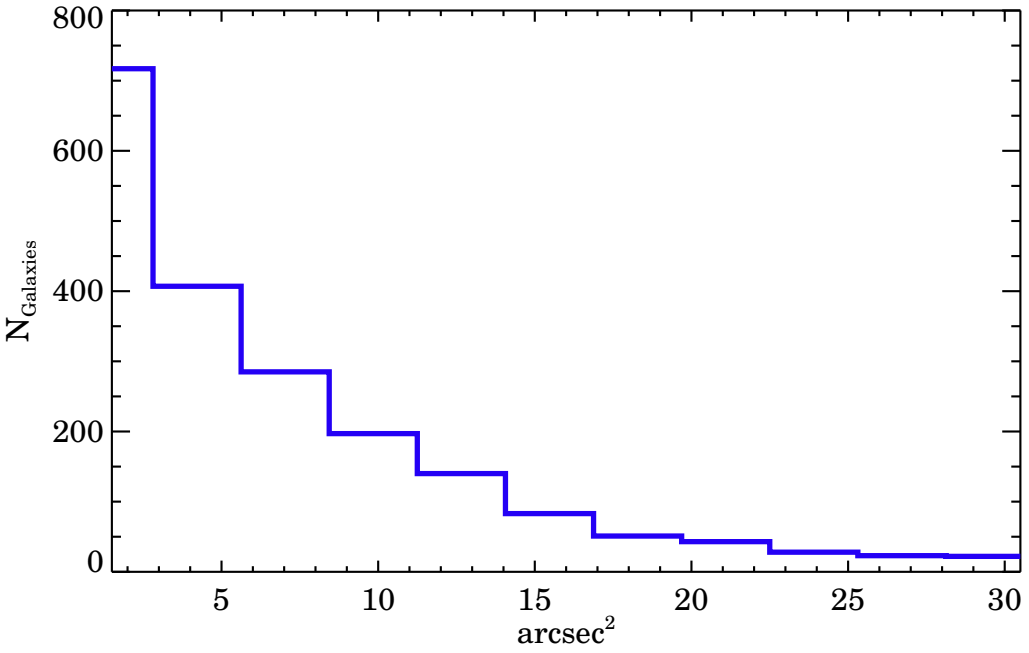


FIGURE 6.13: The area distribution of all redshift ~ 5 UltraVista galaxies.

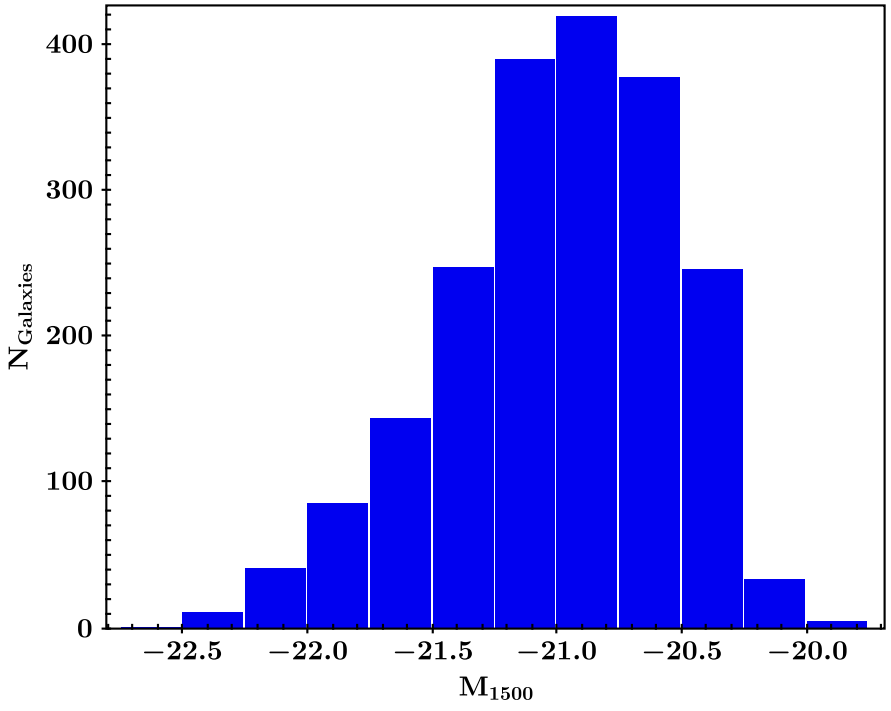


FIGURE 6.14: The distribution in absolute M_{1500} (1500Å rest-frame) magnitude of all redshift ~ 5 UltraVista candidate galaxies.

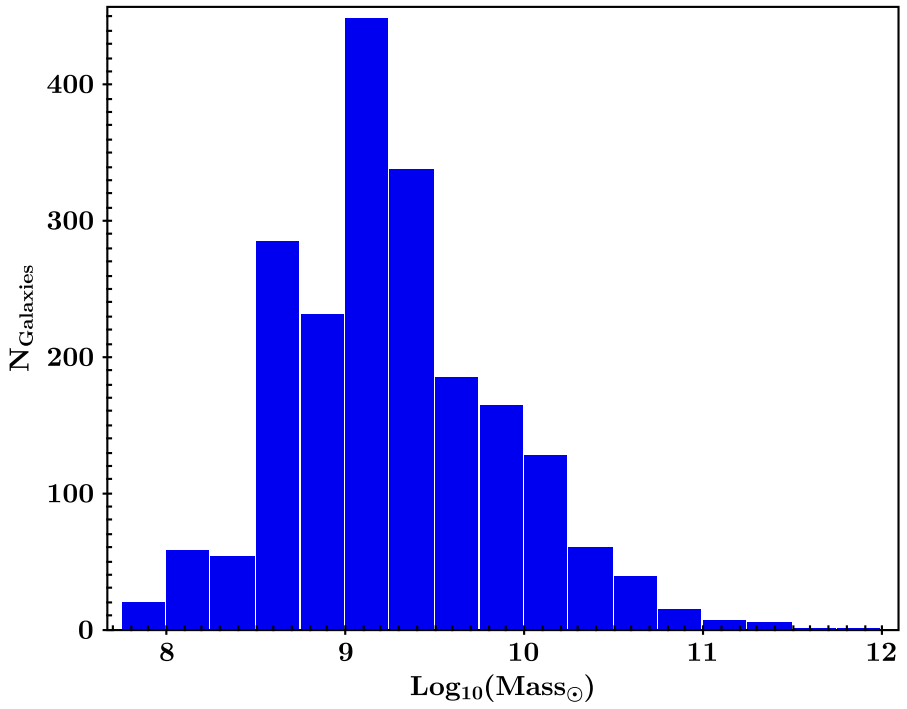


FIGURE 6.15: The stellar mass distribution of all redshift ~ 5 UltraVista galaxies.

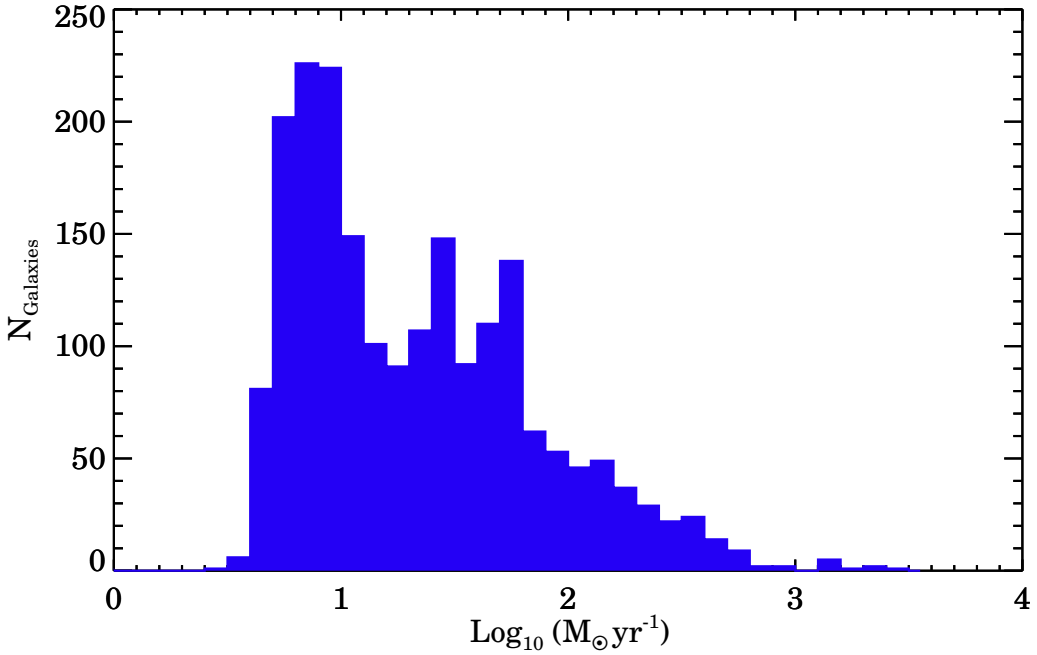


FIGURE 6.16: The star formation rate distribution of all redshift $\sim$ 5 UltraVista galaxies.

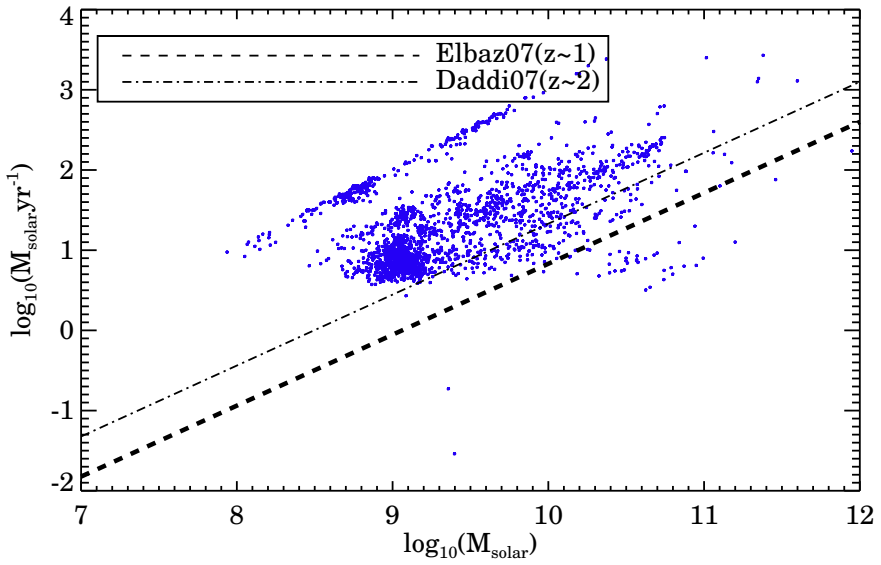


FIGURE 6.17: The SFR$-$M$_\star$ relation for UltraVista $z \sim 5$ galaxies. This is compared to the main sequence measured from the MS at $z \sim 1$ by Elbaz et al. (2007; dashed line), and the MS of Daddi et al. (2007; dash-dotted line) at $z \sim 2$. Our result appears to lie significantly above the Daddi et al. (2007) main sequence.

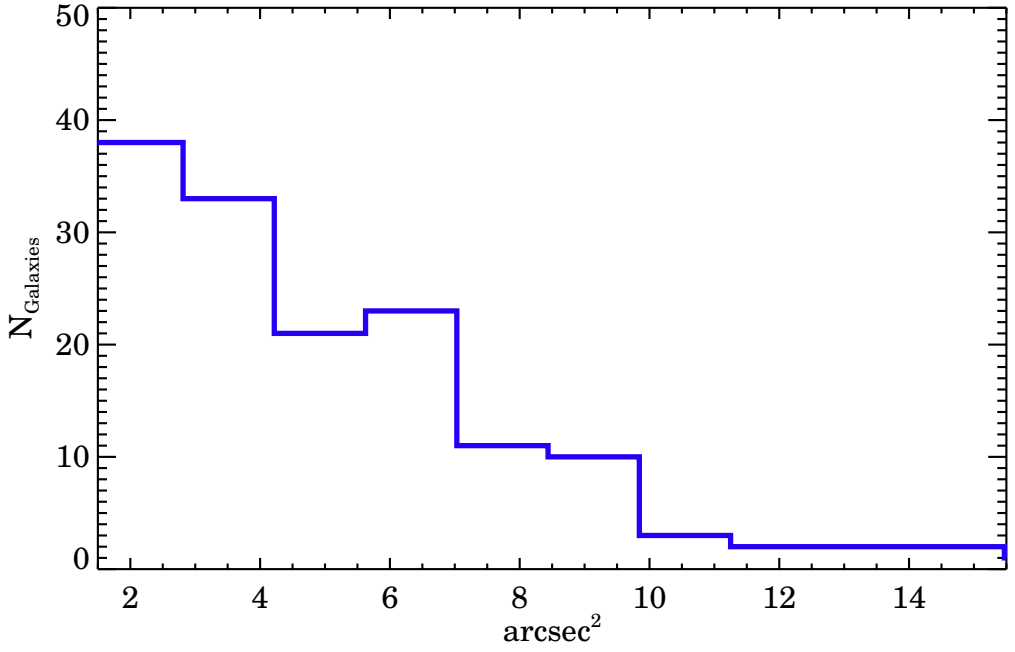


FIGURE 6.18: The area distribution of all redshift ~ 6 UltraVista galaxies.

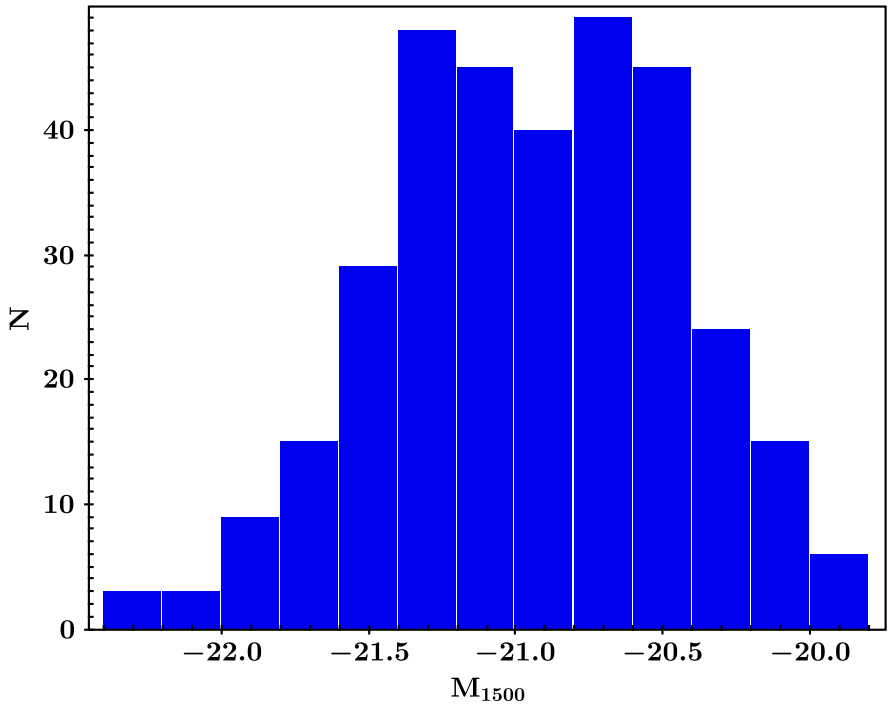


FIGURE 6.19: The distribution in absolute M_{1500} (1500Å rest-frame) magnitude of all redshift ~ 6 UltraVista galaxies.

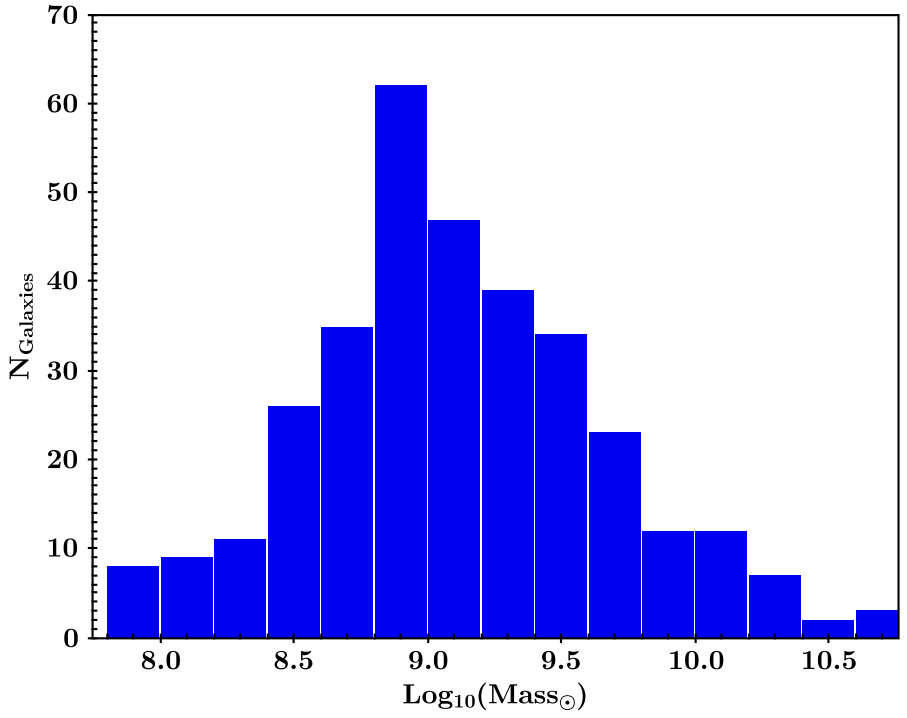


FIGURE 6.20: The stellar mass distribution of all redshift ~ 6 UltraVista galaxies.

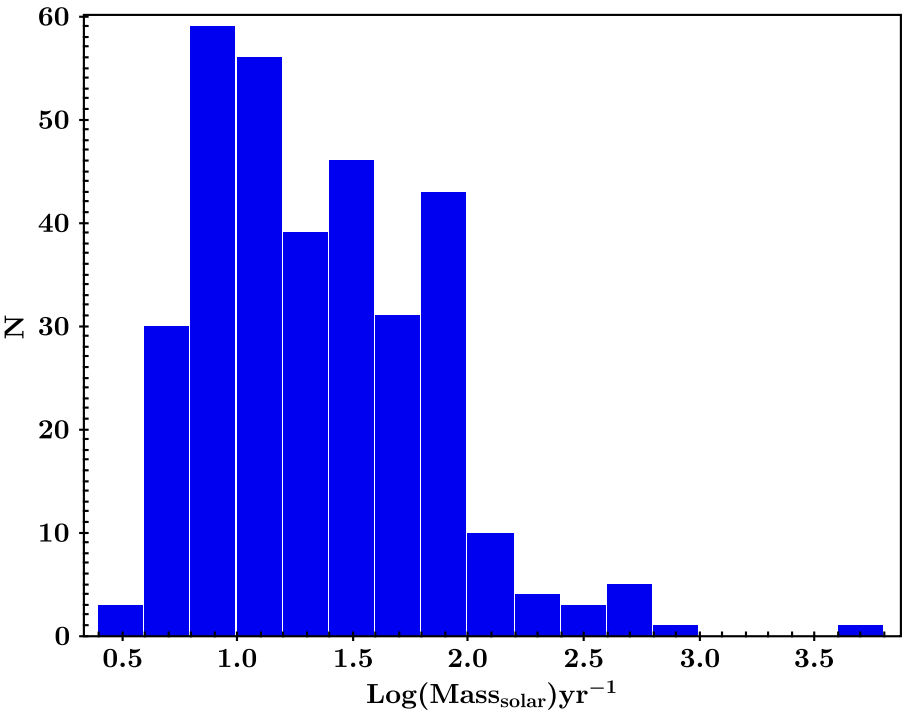


FIGURE 6.21: The star formation rate distribution of all redshift $\sim$ 6 UltraVista galaxies.

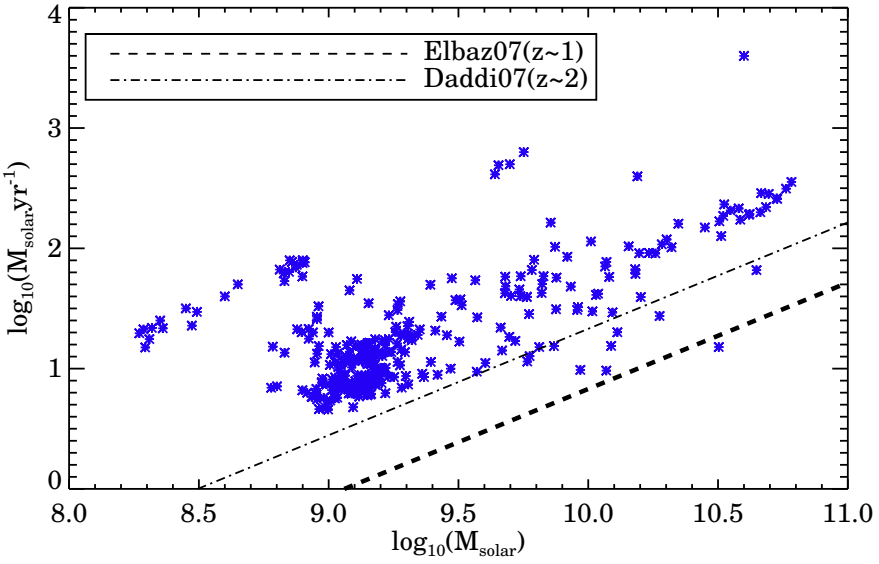


FIGURE 6.22: The SFR−M relation for UltraVista $z \sim 6$ galaxies. This is compared to the main sequence measured from the MS at $z \sim 1$ by Elbaz et al. (2007; dashed line), and the MS of Daddi et al. (2007; dash-dotted line) at $z \sim 2$. Our result appears to lie significantly above the Daddi et al. (2007) main sequence.

$z_{phot} \geq 2.4 \pm 1\sigma$. The VUDS sample can therefore be used as a control sample, even if the statistics is much lower than my photometric candidates. This is described below.

6.4.1 Comparison to the $z \sim 5$ VUDS sample

Figure 6.23 and Figure 6.24 compare the redshift distribution and i magnitude distribution of all $z \sim 5$ VUDS galaxies in all three fields (COSMOS, ECDFS and VVDS2H) to our UltraVista candidates. There are 262 galaxies in VUDS in this redshift range and 154 of them are galaxies with reliable redshift measurements (zflag 2, 3, 4 and 9, see Le Fèvre et al. 2014). Among the 94 VUDS galaxies in the COSMOS field area in common to our UltraVista area 78 are in our UltraVista photometric catalog, and therefore most of the VUDS galaxies are detected (∼83%) in the UltraVista photometry. We compare below the colour, mass and star formation rate properties of the UltraVista sample to the VUDS sample.

- **Colour-colour plot** – Figure 6.25 shows the $(r-i)-(i-z')$ distribution of all the 154 redshift ~ 5 VUDS galaxies with reliable flags compared to the UltraVista candidates. The dashed lines define the standard colour-colour region for selecting $z \sim 5$ candidates. It is clear that a fraction (∼47%) of good $z \sim 5$ VUDS galaxies would be excluded if a colour-colour selection was applied, showing that colour-colour selection is too restrictive. This supports our approach to use photometric redshifts. The UltraVista sample has a colour distribution with candidates outside the colour-colour area, a property supported by galaxies with a similar colour-colour distribution in the VUDS sample of galaxies with confirmed spectroscopic redshifts.

- **Stellar Mass & SFR distribution** – Figure 6.26 compares the stellar mass $M_\star$ and SFR distribution of the UltraVista sample to the VUDS sample. At redshifts $z \sim 5$ the stellar mass ranges from $10^8 M_\odot$ to $10^{12} M_\odot$, and the distribution of VUDS and UltraVista shows more less massive galaxies in UltraVista compared to VUDS. The plot of SFR distribution shown in Figure 6.27 also shows that there are more star forming galaxies found in VUDS compared to the UltraVista and the majority of star formation rate is from $10 M_\odot/yr$ up to several hundred $M_\odot/yr$ in both samples.

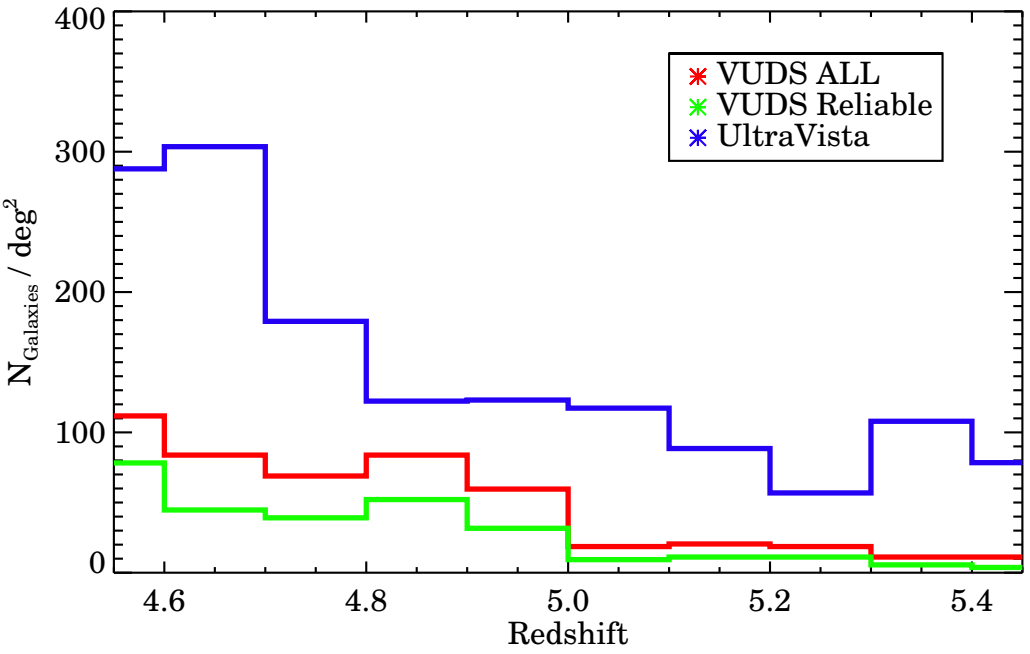


FIGURE 6.23: The redshift distribution of VUDS galaxies compared to our UltraVista candidates with $4.5 \leq z \leq 5.5$. The red histogram shows total $z \sim 5$ VUDS samples from all 3 fields(COSMOS,ECDFS,VVDS2H), the green histogram shows only the VUDS galaxies with the most reliable redshifts (zflag=2, 3, 4, 9). The blue-colour histogram shows the redshift distribution of the UltraVista sample. The number of galaxies have been normalised to one square degree area. The difference in counts is due to the different limiting magnitudes of VUDS (i_{AB}=25 and UltraVista (equivalent to i_{AB}=26, see Figure 6.24.

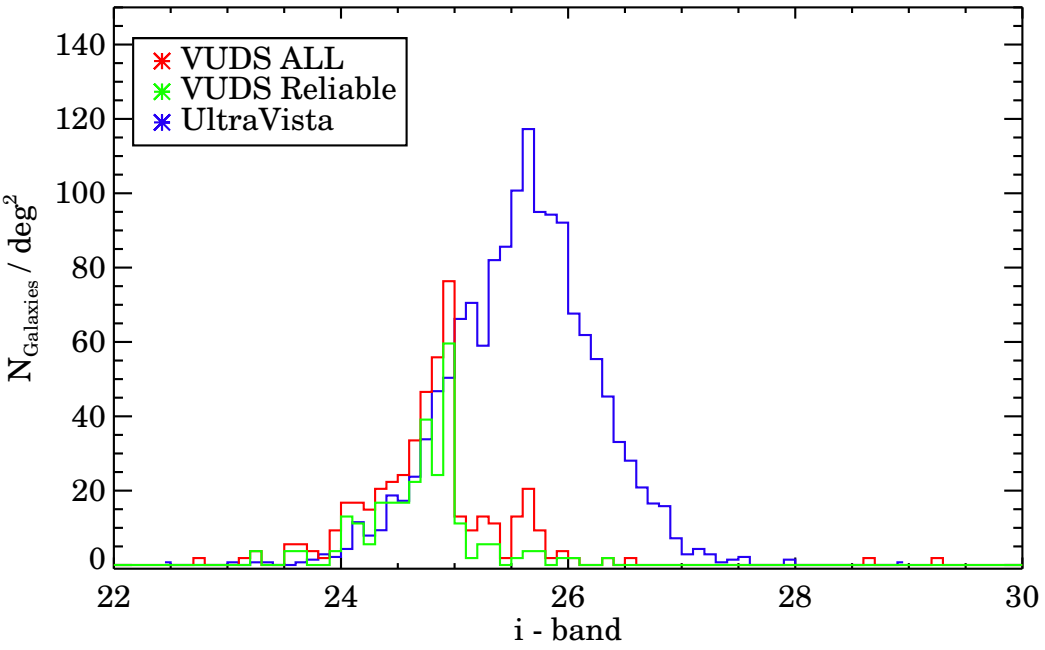


FIGURE 6.24: The i magnitude distribution with redshift of VUDS galaxies compared to our UltraVista candidates with $4.5 \leq z \leq 5.5$. The red histogram shows total $z \sim 5$ VUDS samples from all 3 fields(COSMOS,ECDFS,VVDS2H), the green histogram shows only the VUDS galaxies with the most reliable redshifts (zflag=2, 3, 4, 9). The blue-colour histogram shows the redshift distribution of the UltraVista sample. The number of galaxies have been normalised to one square degree area. One can see that down to i_{AB}=25 the two distributions agree very well, and that UltraVista is going about one magnitude deeper than VUDS beyond i_{AB}=26.

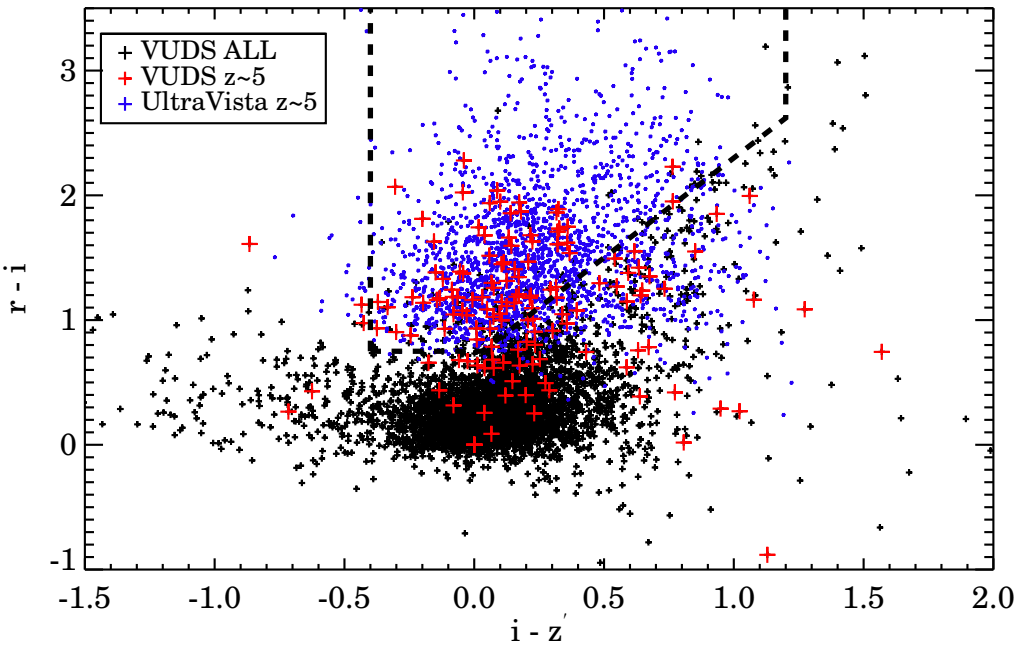


FIGURE 6.25: $(r-i)-(i-z')$ colour-colour diagram of all reliable VUDS galaxies with redshift ~ 5 ($4.5 < z < 5.5$) (red crosses) compared to the Ultra-Vista candidates (blue crosses) in the same redshift range. The VUDS galaxies with redshifts outside $4.5 \leq z \leq 5.5$ are shown as black dots. . The dashed lines identify the colour selection used when selecting $z \sim 5$ drop-out objects from a *riz* colour plot.

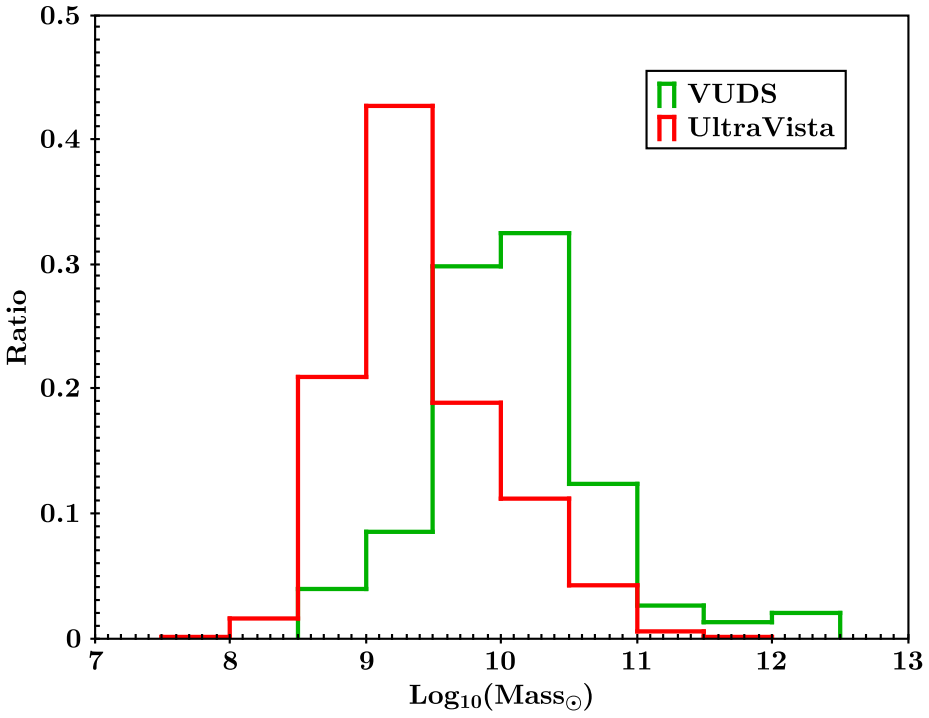


FIGURE 6.26: The distribution of stellar masses $M_\star$ of all reliable VUDS galaxies with redshift ~ 5 ($4.5 < z < 5.5$).

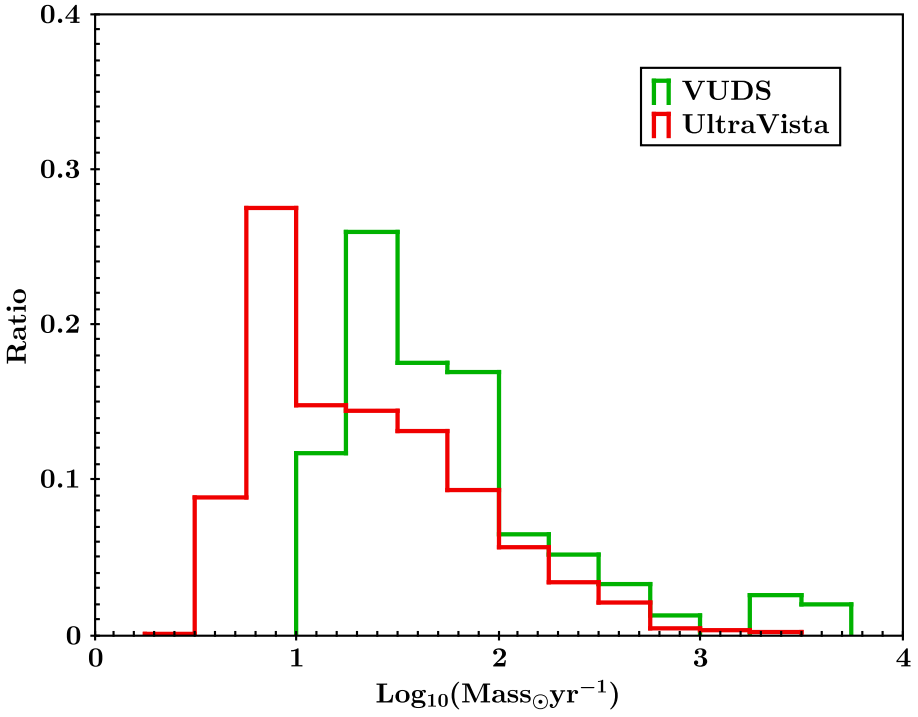


FIGURE 6.27: The star formation rate distribution of of all reliable VUDS galaxies with redshift ~ 5 ($4.5 < z < 5.5$).

Chapter 7

High Redshift Galaxy Luminosity Function

As discussed in chapter 3, the luminosity function (LF) is a count of the number of galaxies as a function of luminosity (or absolute magnitude), per unit volume. LFs can help us understanding galaxy formation because the luminosity of a galaxy is linked to the number of stars present either recently formed (UV luminosity function) or older (IR LF), and therefore is connected to the stellar mass assembly along cosmic time. Comparing LFs at different redshifts allows to follow the evolution of luminosity in the Universe. LFs can also help us to be compared to galaxy evolution models and simulations to try to better understand the galaxy evolution scenario. In this chapter, I compute LFs in $4.5 < z < 6.5$ using the UltraVista data described in Chapters 5 and 6, and I compare to other results in the literature. I then derive the Star Formation Rate Density and discuss it in the framework of the galaxy formation scenario.

7.1 Area and Volume

The definition of the LF is a number of galaxies per unit volume within a specific range of luminosity, so the first step to estimate the LF is to estimate the real area of our observations and its correspond volume in each redshift bins. Figure 4.5 shows the whole field of the UltraVista survey, with a total area covered 1.58 deg^2. Due to the subtraction of 0.19 deg^2 of masked areas corresponding to halos and spikes from bright objects and artefacts, the final area for LFs calculation is 1.39 deg^2 (1.58 deg^2−0.19 deg^2).

TABLE 7.1: Comoving Volume of each redshift bins

Redshift range	Volume (Mpc^3)
$4.5 < z < 5.5$	1.36×10^7
$5.5 < z < 6.5$	1.21×10^7

From the area calculation, I need one step forward to estimate the volume sampled in each redshift bin. The volume I use here is the comoving volume which can be obtained from the assumption of cosmological parameters of the cosmological model. The definition of comoving volume is a region of space which expands with cosmological expansion. Therefore, the number of galaxies in a comoving volume would not change with time, if galaxies were neither destroyed nor created. The comoving volume in a solid angle dΩ and redshift interval dz is :

$$dV_c(z) = \frac{c}{H_0} \frac{(1+z)^2 d_A}{\sqrt{\Omega_M (1+z)^3 + \Omega_k (1+z)^2 + \Omega_\Lambda}} d\Omega dz \tag{7.1}$$

where c is the speed of light, H_0 is the Hubble constant and the three density parameters (Ω_M, Ω_k, Ω_Λ) are as defined in chapter 2. Table 7.1 presents the comoving volume of 2 redshift bins within 1.39 deg^2 solid angle.

7.2 Computing the Luminosity Function at $4.5 \leq z \leq 6.5$

I use my UltraVista catalogues to compute the LF in two redshift bins $4.5 \leq z \leq 5.5$ and $5.5 \leq z \leq 6.5$. I use the tool ALF with Olivier Ilbert to compute the LF using 4 indicators: the V_{max}, SWML, C^+ and STY. The values are used to compute a fit using a Schechter function (Chapter 3). I report the results of these computations in the following sections.

7.2.1 The Luminosity Function at z $\sim$ 5

The luminosity function computed in $4.5 \leq z \leq 5.5$ is presented in Figure 7.1, with the results from the four different indicators. The V_{max} values listed in Table 7.2. I find that the characteristic Luminosity is M*=-20.618$\pm^{0.175}_{0.199}$, the volume density ϕ^*=0.764$\pm^{0.084}_{0.132}$ ($10^{-3} Mpc^{-3}$), and the slope α=-1.127$\pm^{0.398}_{0.388}$ as listed in Table 7.4. The uncertainty on the slope remains quite high because the UltraVista data are not sufficiently deep to fully constrain the faint-end slope. When fixing the slope to $\alpha = -1.78$ (from Bouwens et al. 2014) we find M*=-20.965$\pm^{0.056}_{-0.058}$ and $\phi^* = 0.520\pm^{0.015}_{0.015}$.

TABLE 7.2: LF with redshift 5

Magnitude ($z' - band$)	log
-22.651	-6.210 $\pm^{0.244}_{0.612}$
-22.385	-5.604 $\pm^{0.139}_{0.206}$
-22.158	-4.947 $\pm^{0.0728}_{0.0875}$
-21.900	-4.644 $\pm^{0.0555}_{0.0637}$
-21.675	-4.383 $\pm^{0.0456}_{0.0510}$
-21.436	-4.120 $\pm^{0.0392}_{0.0430}$
-21.192	-3.923 $\pm^{0.0363}_{0.0396}$
-20.976	-3.795 $\pm^{0.0357}_{0.0389}$
-20.781	-3.637 $\pm^{0.0479}_{0.0538}$
-20.510	-3.480 $\pm^{0.133}_{0.193}$

I compare the results of my LF calculation to other LFs in the literature in Figure 7.2. My LF is in excellent agreement with other previous studies (Bouwens et al 2014, Finkelstein et al. 2014 and Van der Burg et al. 2010), over the whole magnitude range. Compared to other studies we bring a significant improvement on the measurement of the bright end of the LF: the advantage of the large area and corresponding volume of the UltraVista sample is fully evident with significantly reduced error bars below M_{1500}=-21.

7.3 The Luminosity Function at z $\sim$ 6

The luminosity function computed in $5.5 \leq z \leq 6.5$ is presented in Figure 7.3, with the results from the three different indicators. The V_{max} values listed in Table 7.3. I find that the characteristic Luminosity is M^*=-20.195$\pm^{0.138}_{0.149}$ and the volume density $\phi_\star$=1.527$\pm^{0.137}_{0.137}$ ($10^{-3} Mpc^{-3}$), obtained fixing the slope α=-1.78 as listed in Table 7.4.

The UltraVista LF is compared to other LFs in the literature in Figure 7.4. At the faint end M_{1500}=-21.5 probed by UltraVista our LF is good agreement with Bouwens et al. (2014) and Bowler et al. (2014). At brighter magnitudes, the UltraVista redshift $\sim$ 6 luminosity function seems to be above Willott (2013) and Bowler (2014), in agreement with Finkelstein et al. (2014) and somewhat lower than Bouwens et al. (2014), but our errors are significantly lower than these studies (as shown in Figure 7.4 and Table 7.3). It was possible to obtain a robust measurement of the bright end because the UltraVista survey area is $\sim$10 times larger than other surveys at the same depth, for a volume surveyed large enough to find the relatively rare bright galaxies. At the bright end of the $z \sim 6$ LF $M_{1500} < -22$ we find that the number density exceeds what is expected from a pure Schechter function. This is comparable to the results recently obtained by Bowler et al. (Bowler et al. 2014b,a) from a different analysis of the same dataset.

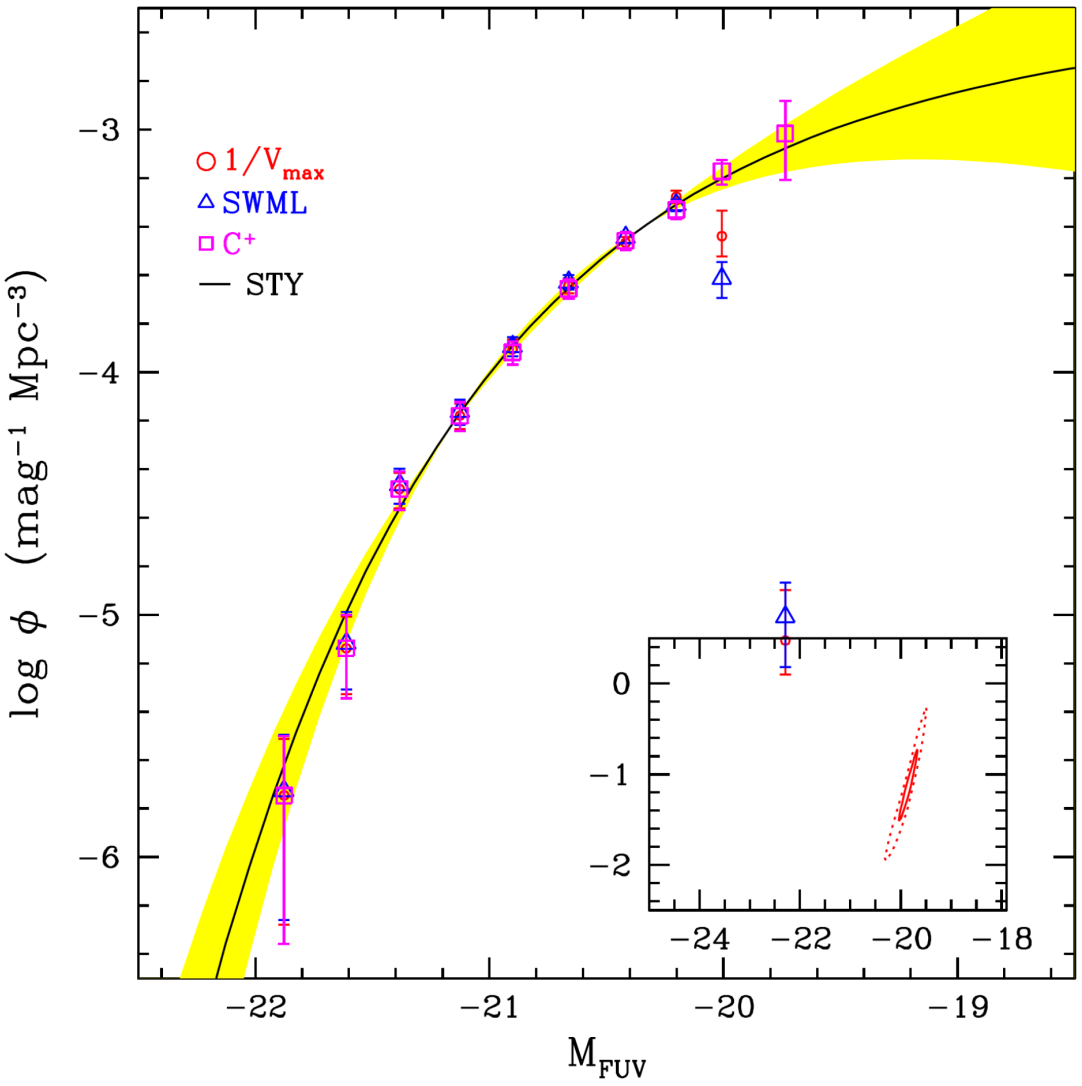


FIGURE 7.1: The rest-frame UV LF at $4.5 \leq z \leq 5.5$ computed from our sample of 2036 candidates selected from the UltraVista photometry. Four different indicators are used to compute the LF (V_{max}, SWML, C+ and STY), as indicated in the upper left panel. The error ellipse in the (α, $M_\star$) plane is indicated in the lower right panel.

TABLE 7.3: LF with redshift 6

Magnitude ($Y-band$)	log
-22.311	$-5.859\ ^{+0.232}_{-0.534}$
-22.216	$-5.893\ ^{+0.301}_{-15.8}$
-22.082	$-5.636\ ^{+0.199}_{-0.377}$
-21.921	$-5.328\ ^{+0.166}_{-0.273}$
-21.800	$-5.001\ ^{+0.119}_{-0.165}$
-21.663	$-5.370\ ^{+0.165}_{-0.270}$
-21.562	$-4.710\ ^{+0.0936}_{-0.119}$
-21.432	$-4.772\ ^{+0.0950}_{-0.122}$
-21.308	$-4.435\ ^{+0.0710}_{-0.0849}$
-21.185	$-4.310\ ^{+0.100}_{-0.130}$

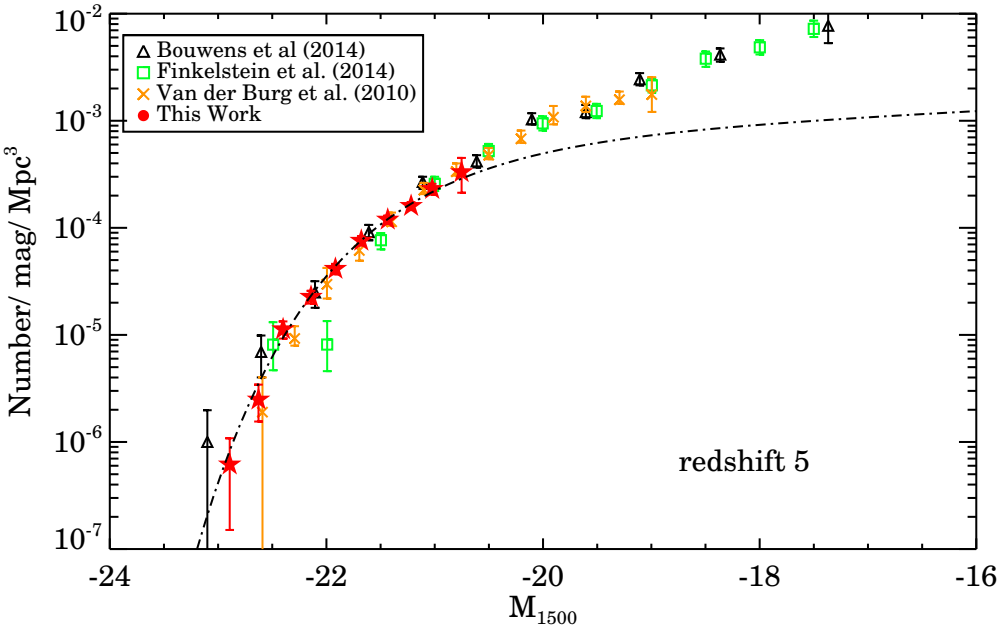


FIGURE 7.2: The rest-frame UV LF at $z \sim 5$ from this work (in red), compared to other results from the literature.

TABLE 7.4: Determination of Best-Fit Scechter Parameter

Redshift range	M*	$\phi(10^{-3}Mpc^{-3})$	α
$4.5 < z < 5.5$	$-20.618\pm^{0.175}_{0.199}$	$0.764\pm^{0.084}_{0.132}$	$-1.127\pm^{0.398}_{0.388}$
$4.5 < z < 5.5$	$-20.965\pm^{0.056}_{0.058}$	$0.520\pm^{0.015}_{0.015}$	-1.78 (fixed)
$5.5 < z < 6.5$	$-20.195\pm^{0.138}_{0.149}$	$1.527\pm^{0.137}_{0.137}$	-1.78 (fixed)

7.4 Evolution of the Luminosity Function and Star Formation rate between z ∼ 5 and z ∼ 6

As reported above and in Table 7.4 I find a significant evolution of M* between redshift 5 to redshift 6 from 0.42 to 0.77 mag. The comparison between our two LFs is shown in Figure 7.5.

I then compute the luminosity densities at $z \sim 5$ and $z \sim 6$ integrating the LF down to 0.03L* at each redshift following Bouwens et al. (2014). This is then transformed into Star Formation Rate Densities (SFRD) based on the luminosity-SFR relation from Madau et al. (1998) shown in equation 2.8. The result of dust-corrected UV-derived SFRDs from our UltraVista sample is shown in Figure 7.6.

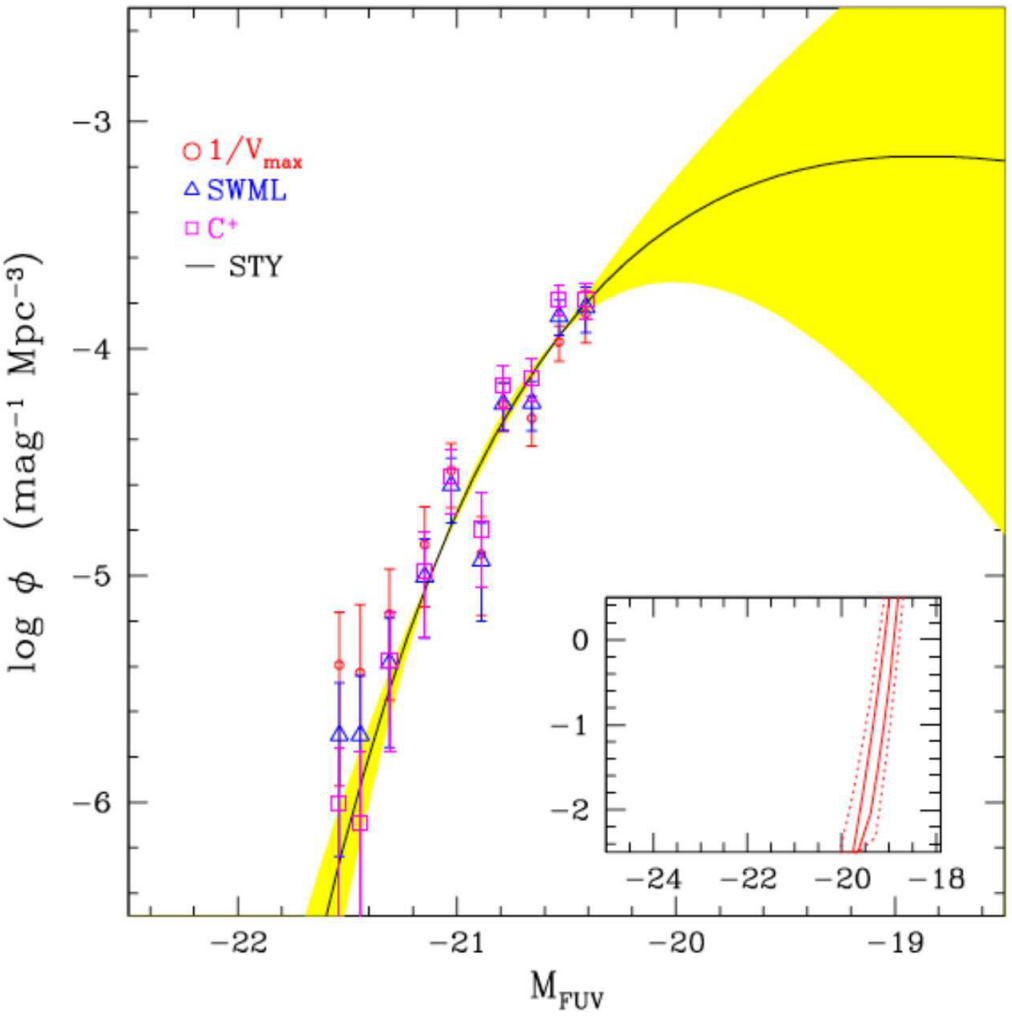


FIGURE 7.3: The rest-frame UV LF at $5.5 \leq z \leq 6.5$ computed from our sample of 330 candidates selected from the UltraVista photometry. Four different indicators are used to compute the LF (V_{max}, SWML, C+ and STY), as indicated in the upper left panel. The error ellipse in the (α, $M_\star$) plane is indicated in the lower right panel.

There are two main uncertainties related to SFRD measurements: uncertainties on the dust attenuation correction and uncertainties on the LF slope α. The dust attenuation (A_{UV}) correction may affect the final result dramatically. Estimating the dust attenuation properties in a galaxy is complicated and there are various methods to derive the dust attenuation using a multiwavelength dataset. Unfortunately, compared to the dust studies in the local universe, there are only a few studies (Bouwens et al. 2009, 2011 and 2014) over redshift 5 which estimated the dust extinction (Figure 7.6). Here I choose to apply the factors suggested from Bouwens et al. 2011 & 2014, which $A_{uv} = 0.85$ and $A_{uv} = 0.2$ in redshifts 5 and redshift 6 respectively. Using $A_{uv} = 0.5$ and $A_{uv} = 0.15$ changes the SFRD as shown in Figure 7.6.

Compared to the literature our SFRD measurements are comparable to Bouwens et al. (2014), which is higher than other measurements. I think this shows that our results are

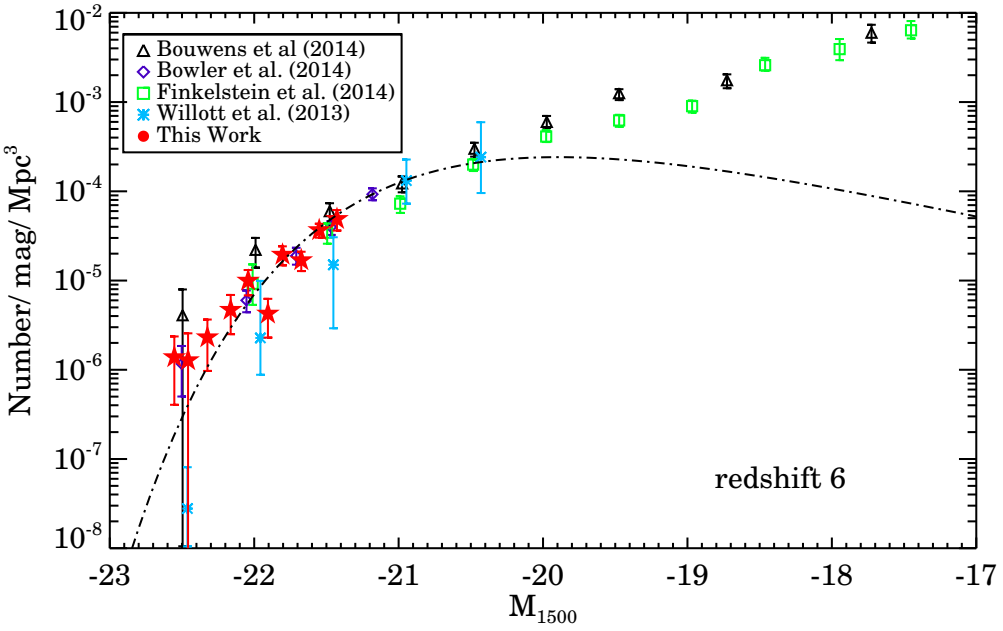


FIGURE 7.4: The rest-frame UV LF at $z \sim 6$ from this work (in red), compared to other results from the literature. The dot-dashed line is the Schechter function fit to our data.

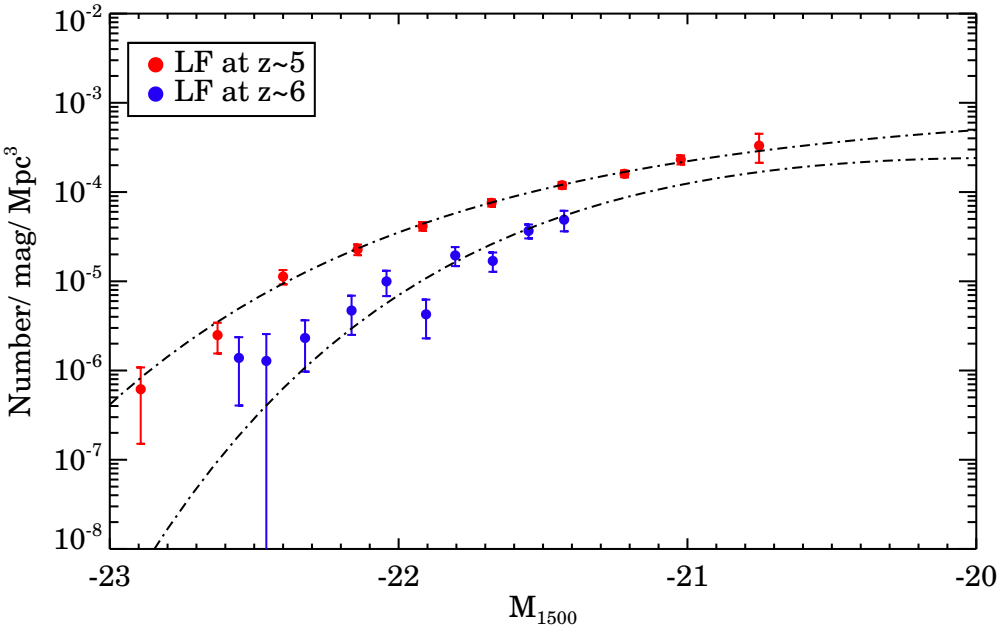


FIGURE 7.5: The best-fitting of Schechter function fits to the rest-frame UV LF at $z = 5, 6$.

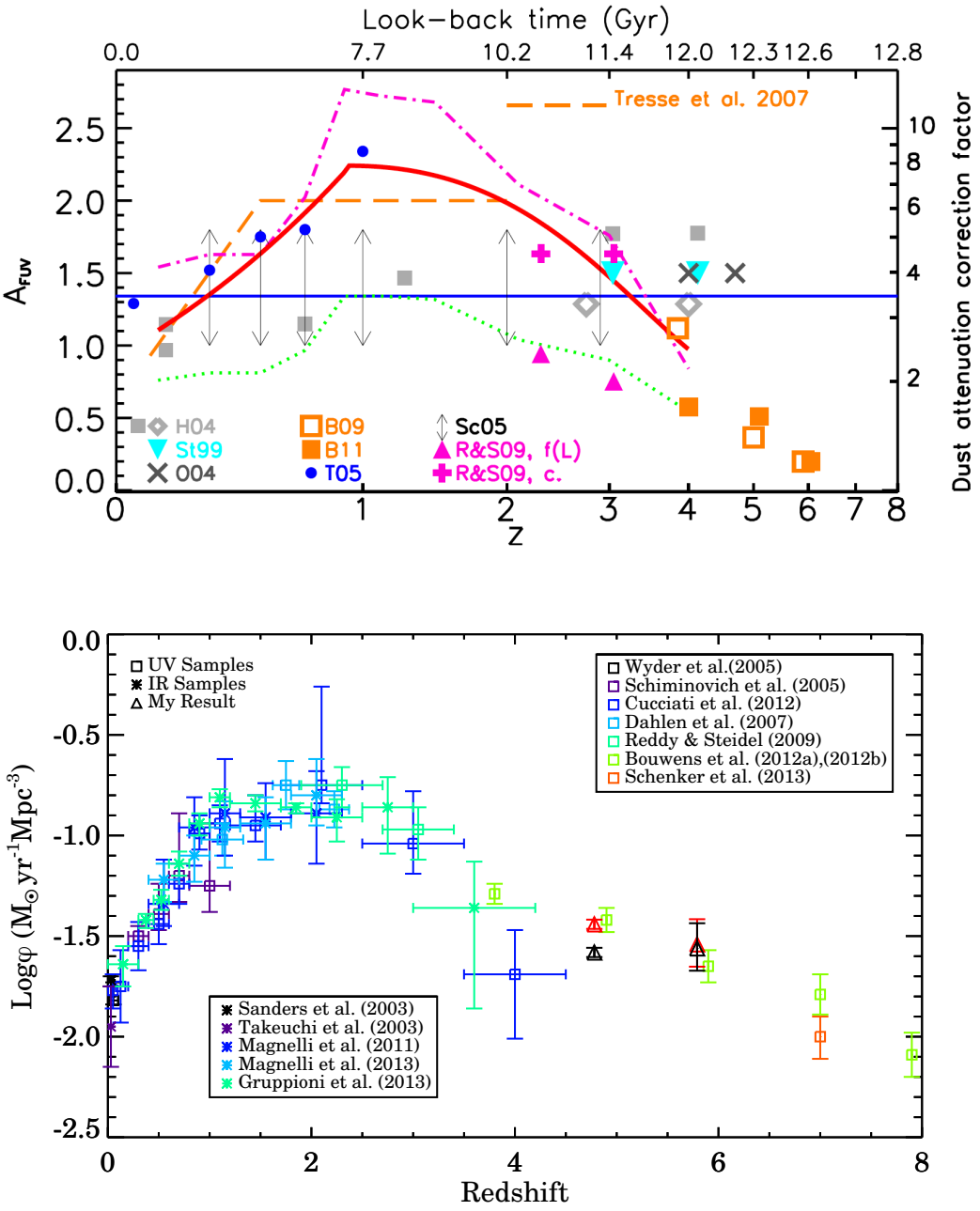


FIGURE 7.6: (Top) Dust attenuation A_{FUV} in magnitudes as a function of redshift including different results and fitting from literatures. (Credit: Cucciati et al. 2012)(bottom)Total dust-corrected UV derived SFRD, the two red triangle points are SFRD measurements at z = 5 to z = 6 based on my work. At z=5 and 6 the points with black triangles are computed from my sample with a lower extinction 0.5 and 0.15 instead of 0.85 and 0.2. All the other points are compiled from other literature results mainly derived from UV and IR samples (see in the Figure).

more complete in counts than other datasets, except Bouwens et al. (2014) who benefit from a compilation of all existing datasets obtained with HST. There is a continuous increase of the SFRD from z∼6 to z∼ 2, showing the dramatic rise in star formation activity over this important period of galaxy evolution.

My results are presented in the following paper (currently in a draft form), to be submitted to A&A.

7.4.1 Paper: The bright end of the UV luminosity function at $4.5 < z < 6.5$ from the UltraVista survey

Astronomy & Astrophysics manuscript no. uvista˙LF˙z567
June 11, 2015

The bright end of the UV luminosity function at $4.5 < z < 6.5$ from the UltraVista survey. ⋆

P.W. Wang[1], O. Le Fèvre[1], P. Cassata[1], O. Ilbert[1], L.A.M. Tasca[1], and et al.[1]

[1] Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France e-mail: bin-wei.wang@lam.fr
[2] et al.

Received ...; accepted ...

ABSTRACT

Context. The distribution of galaxy luminosities in the early phases of galaxy evolution and the derived luminosity density and star formation rate density at z>4.5 are still poorly constrained by observations, particularly for the bright, relatively rare, galaxies. Here we aim to find and characterise high redshift galaxy candidates in the redshift range $4.5<z<6.5$ in the well studied COSMOS field. We use the UltraVista survey digital release 2 (DR2) together with all multi-wavelength data available in this field to search for galaxies in this redshift range using photometric redshifts derived from the spectral energy distribution (SED), compared to colour-selected samples following the Lyman-break and drop-out technique. The UltraVista survey provides deep near infrared data down to a limiting magnitude K_{sAB}=24.8 (5σ) for a large 1.39 deg^2 area minimizing cosmic variance. We evaluate the completeness and contamination from low redshift dusty interlopers and galactic cool dwarf stars of colour-colour selection and we quantify the fraction of galaxies lost in our selection process using image simulations. After visual examination of images and SED of all candidates satisfying our criteria, we obtain a sample of 2036 and 330 galaxies in redshift ranges $4.5<z<5.5$ and $5.5<z<6.5$, respectively. We compare this photometric sample to the spectroscopic sample obtained from the VIMOS Ultra Deep Survey (VUDS) in the overlapping redshift range to further evaluate our completeness and contamination. From these data we build UV-rest luminosity functions over the redshift range 4.5 - 6.5 and derive integrated luminosity and star formation rate densities. We find that the bright end of the luminosity function $M_{UV} < -21$ is well represented by a standard Schechter function at $z \sim 5$, but is significantly more populated than would be expected from such a function at the bright end of the $z \sim 6$ LF. We infer that these numerous bright galaxies have significantly contributed to the re-ionisation of the universe.
Aims.
Methods. Results. Galaxies: evolution – Galaxies: formation – Galaxies: high redshift – Galaxies: star formation –
Conclusions.

Key words.

1. Introduction

The luminosity distribution of galaxies is one of the main indicators used to follow galaxy evolution. The galaxy luminosity function, i.e. the number of galaxies of a given luminosity per unit volume, is intimately linked to the star formation process and to galaxy mass assembly. In the hierarchical mass assembly picture galaxies form from the collapse of matter in the deep potential wells of dark matter haloes with gas cooling and fragmentation on galaxy scales (White & Rees 1978; Bromm et al. 2009). The luminosity output of a galaxy is then directly related to its stellar mass, modulated by stellar evolution processes. The UV emission, when corrected from dust extinction, is a measure of the star formation rate and thus the UV luminosity function is commonly used to estimate the global star formation rate in the universe (Madau & Dickinson 2014). The UV luminosity function (LF) is integrated into the UV luminosity density (LD), in turn transformed into a star formation rate using standard calibrators. This approach was first successfully implemented up to redshifts z~4 using the LF from the Canada-France Redshift Survey (Lilly et al. 1996) and HST Hubble Deeep Field data (Madau et al. 1996), and since then has been constantly improved to reach the highest possible redshifts (e.g. Cucciati et al. 2012, Bouwens et al. 2014).

The census of galaxies as a function of their luminosity is therefore the backbone of one of our most important view on galaxy evolution. Computing the LF and deriving the LD requires a sample of galaxies representative at all luminosities above the characteristic luminosity $L_\star$, and an estimate of the faint-end slope α in the classical representation of the LF using a Schechter function (Schechter 1976). While the LD and star formation rate history are reasonably well defined up to z≃4 (e.g. Madau & Dickinson 2014), at higher redshifts the shape of the LF (a Schechter form or a double power law), as well as characteristic quantities like the density $\Phi_\star$, luminosity $L_\star$ and faint-end slope α, are still the subject of considerable debate and uncertainties. At redshifts z>4, the most constrained LFs today are build from galaxy samples identified from deep multi-band photometry assembled using HST ACS visible and WFC3 infrared camera data (see e.g. the compilation from Bouwens et al. 2014). When compared to the deepest ground-based data, differences of

⋆ Based on data obtained with the European Southern Observatory VISTA telescope, Paranal, Chile, under Large Program 179.A–2005. Based in part on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, under Large Program 185.A–0791.

up to a factor ∼ 2 in number counts have been reported at z∼6 (Bowler et al. 2014a). It is unclear where these differences may come from, but two main elements are pointed out: the methods to identify galaxies in a particular redshift range, and field to field cosmic variance. Cosmic variance is mitigated by sampling volumes as large as possible in several unrelated fields, and it is estimated that cosmic variance in a single ∼1.5 deg^2 field like the COSMOS-UltraVista is equivalent to the variance over the 5 fields of the CANDELS survey covering 0.2 deg^2, each with a variance ∼10% (Bouwens et al. 2014). Identifying galaxies in a particular redshift range from deep multi-band photometric data is generally presented as a straightforward task, but the assumptions and limitations of the many steps required to reach representative number counts are often, and unfortunately, not fully documented. Two main techniques, quite close in essence, are used to identify distant galaxies at $z > 4$, either the *drop-out* technique, the high redshift analogue of the Lyman-break technique used at redshifts z∼3 based on the continuum drop produced by the strong extinction below Lyman-α 1215Å produced by the inter-galactic medium (IGM), or photometric redshifts derived from the fit of the spectral energy distribution (SED) using reference galaxy templates. These techniques present a number of limitations that are poorly explored, and each study uses its own particular tuning of the search parameters which makes it difficult to reproduce. The extreme example being the search for redshift z∼8-10 candidates from exactly the same dataset in two different studies and finding two completely different sets of candidates (refs), or the claim for very massive galaxies at $z \sim 5$ which turned out to be at lower redshifts $z \sim 1.5 - 2$ (refs). So far there is no systematic validation of these search techniques confirming candidates using e.g. spectroscopic redshift samples, so that the completeness and contamination of samples assembled so far is largely unchecked.

HST-WFC3 data provide exquisite sensitivity reaching down to H_{AB}=26-27 but the field area covered is small and the largest field covered so far is a total of ∼750 arcmin^2 in 5 fields from the CANDELS survey (Grogin et al. 2011). This offers a good handle on the faint end of the luminosity function, reported to be quite steep in the range α=-1.5 to -2.0 for redshifts z∼ $4 - 10$ (Bouwens et al. 2014), but these data do not cover enough volume to place constraints on the bright end of the LF. This is one of the driving arguments to conduct surveys over large areas, only possible with dedicated ground-based instrumentation. The largest fields observed to date are the UDS (Warren et al. 2007) and UltraVista (McCracken et al. 2012) covering 0.8 and 1.5 deg^2, at depths H_{AB}=25 and 25.5, respectively. LFs at z∼6 and z∼7 have been recently derived from the UltraVista DR2 release (Bowler et al. 2014b,a), concluding that the bright end is more populated than previously thought, and finding significant differences with the study of Bouwens et al. (2014).

In this paper we present a new analysis of the UltraVista DR2 to derive LFs in 2 redshift domains $4.5 < z < 5.5$ and $5.5 < z < 6.5$. We summarize the properties of the UltraVista data and all multi-wavelength data from the COSMOS field in Section 2.1. The search technique applied to identify galaxy candidates in these redshift ranges is discussed in Section 3. Simulations are conducted to estimate the likelihood to retrieve a galaxy with a given redshift given the distribution of sources, and the probability of contamination by galaxies outside the redshift range of interest, as presented in Section 4. The properties of the sample are discussed in Section 5. Comparison with LBG-selected samples and with the VIMOS Ultra Deep Survey spectroscopic redshift sample is discussed in Section 6. We present our galaxy catalogues and derive the LFs, LDs and SFRD in Section 7. We discuss our results and conclude in Section 8.

We use a cosmology with $H_0 = 70h\ km\ s^{-1}\ Mpc^{-1}$, $\Omega_{0,\Lambda} = 0.70$ and $\Omega_{0,m} = 0.30$. All magnitudes are given in the AB system.

2. Data

2.1. The UltraVista deep photometric sample in DR2

UltraVista is an ongoing survey with the Visible and Infrared Camera (VIRCAM) on the ESO-Vista telescope which provides the deepest wide field ground-based near infrared imaging available today (Dalton et al. 2006). The broad-band (Y, J, H, K_s) photometric survey includes a deep survey covering the full 1.5 deg^2 field, and an ultra-deep survey covering 0.73 deg^2 into 4 stripes, as well as a narrow-band NB-118 photometric survey. All the work from this paper is based on the deep (DR1) and ultra-deep (DR2) broad-band surveys (McCracken et al. 2012). UltraVISTA ultra-deep stripes are shown in Figure 1, and the coordinates of the corners of the ultra-deep field are listed in table 1.

Besides the near-infrared data, there are two main other datasets used in this paper covering the optical and mid-infrared wavelength domains. The UltraVista survey is superimposed on the 2 deg^2 multi-wavelength Cosmological Evolution Survey (COSMOS) field (Scoville et al. 2007). All the *griz* optical data for our analysis is from Subaru Suprime-Cam (Taniguchi et al. 2007). The deep z'-band data is a particularly important band to work at the high redshifts of interest here. In addition, we use Mid-infrared imaging within the COSMOS field obtained with the Spitzer/infrared array camera (IRAC), including the sCOSMOS data (Sanders et al. 2007) and the more recent data from the Spitzer Large Area Survey with Hyper-Suprime-Cam (SPLASH) (Steinhardt et al. 2014), keeping the deep channel 1 (3.6μm) and channel 2 (4.5μm) photometry. Mid-infrared plays an important role to help identify contamination from lower redshift objects and obtain better physical information based on the better constrained SED fitting.

2.2. Spectroscopic sample from the VIMOS Ultra Deep Survey

The VIMOS Ultra Deep Survey – VUDS is a large spectroscopic survey observing ∼10 000 galaxies with the VIMOS multi-object spectrograph at the European Southern Observatory Very Large Telescope (Le Fèvre et al. 2003). The main goal of this project is to study galaxy formation and evolution at $2 < z < 6$ using reliably determined spectroscopic redshifts (Le Fèvre et al. 2014). VUDS covers a total area of 1 deg^2 in three independent fields, including 0.5 deg^2 in the COSMOS field. The VUDS sample was selected based on photometric redshift $z_{phot} \geq 2.4 \pm 1\sigma_{zphot}$, and $i_{AB} \leq 25$, with z_{phot} the first *or* the second peak in the photometric redshift probability distribution function. Details about the survey strategy, target selection, as well as data processing and redshift measurements are presented in Le Fèvre et al. (2014).

In the VUDS catalogue we find 178 galaxies in the redshift range $4.5 \leq z \leq 6.5$ and reliable spectroscopic redshifts, including flag 2 and 9 (∼ 75 – 80% reliable) and flag 3 and 4 (∼ 100% reliable) as described in Le Fèvre et al. (2014)). This includes 107 galaxies in the COSMOS field. Another 104 VUDS galaxies, 52 in the COSMOS field, have a low reliability redshift (flag 1) in that same redshift range. This is one of the largest sample a

galaxies with spectroscopic redshifts in this redshift range available today. Even though this sample is rather small compared to photometric samples, the set of galaxies with robust spectroscopic redshifts forms an important reference to compare to the galaxies we select based on photometry alone, and we will use this sample to verify the completeness and contamination of our selection function in the following.

2.3. Building the catalogue

We build the primary UltraVISTA multi-wavelength catalogue using SExtractor v2.8.6 (Bertin & Arnouts 1996) with a detection threshold of 2.2σ and a minimum object size of 5 pixels in dual-image mode. The UltraVISTA sources were identified from a χ^2 weighted $Y+J+H+K_s$ stacked produced from each individual images in each band to ensure that faint objects with redshifts higher than $z = 4.5$ are not missed. The initial UltraVISTA catalogue for the dataset consisted of 229,029 objects within the ultra-deep part and 357,353 objects in the deep UltraVISTA area imaging that overlaps with the multi-wavelength auxiliary data in optical bands (COSMOS). Figure 2 shows the galaxy number counts extracted from our catalogue in comparison with the first UltraVista data release DR1 (McCracken et al. 2012), our counts agree well with the existing K_s counts, and demonstrate that the dataset corresponding to the DR2 reaches 1 magnitude deeper than the DR1. We determined imaging depths across all bands using 1σ measurements in multiple isolated 2.0 arcsec diameter circular apertures as used for the optical and near-infrared photometry and 2.8 arcsec aperture for the SPLASH data. These limiting magnitudes are reported in Table 2.

3. Selection of candidate galaxies at $4.5 \leq z \leq 6.5$

3.1. Pre-selection

The now standard way to identify galaxies at $z > 4.5$ is to use the drop-out technique, an extension to high redshifts of the Lyman Break Galaxy (LBG) technique. As discussed in Bouwens et al. (2014) the LBG selection produces a redshift distribution which is modulated by the sensitivity of a particular three colour photometry. As the break produced below Lyman-α by the combined Lyman-limit and IGM transmission is moving through three adjacent photometric redshifts, the resulting redshift distribution N(z) is bell-shaped (see e.g. Figure 1 of Bouwens et al. (2014)) because the difference in photometry for redshifts at the edges of the redshift domain targeted by the filters is getting small and a larger fraction of galaxies is lost. This is usually recovered by the means of extensive galaxy simulations whereby galaxies of different spectral types are moved in redshift through the filter set, their photometry is computed and the corresponding model images are added to deep images, and the fraction of galaxies recovered at the right redshift after this procedure is used as a corrective factor to compute volume quantities like the luminosity of mass functions. Simulation of galaxies outside the redshift range of interest are also performed to identify the best locus of a selection box in colour-colour space maximizing the selection of galaxies in the right redshift range, while minimizing the contamination from galaxies at other redshifts. This process relies on the hypothesis that one knows the distribution of key physical properties like the colours of galaxies resulting from a combination of stellar and gas emission, dust and IGM extinction.

In this study we elected to follow a different route. We use the same input data as for colour-colour selection, namely the multi-band photometry, but rather than using only three bands for a particular redshift range, we use all the photometric bands to assess the photometric redshift of a galaxy. In this way we expect a uniform redshift sensitivity over the redshift range probed. In addition, the identification of high redshift galaxies with $z \geq 4.5$ is subject to degeneracies in colour space which can confuse high redshift $z > 4.5$ galaxies with lower redshift $z \sim 1.5 - 2.5$ galaxies and with red dwarf stars in our Galaxy. Using three bands - two colours - is only partially able to discriminate these populations resulting in a loss of galaxies in the redshift range of

Table 1. Coordinates of four ultra-deep (DR2) strips in decimal.

Strip ID	Coordinates(RA,Dec)
strip1	149.33, 2.76, 149.46, 2.76, 149.46, 1.66, 149.33, 1.66
strip2	149.70, 2.76, 149.83, 2.76, 149.83, 1.66, 149.70, 1.66
strip3	150.06, 2.76, 150.20, 2.76, 150.20, 1.66, 150.06, 1.66
strip4	150.43, 2.76, 150.57, 2.76 ,150.57, 1.66, 150.43, 1.66

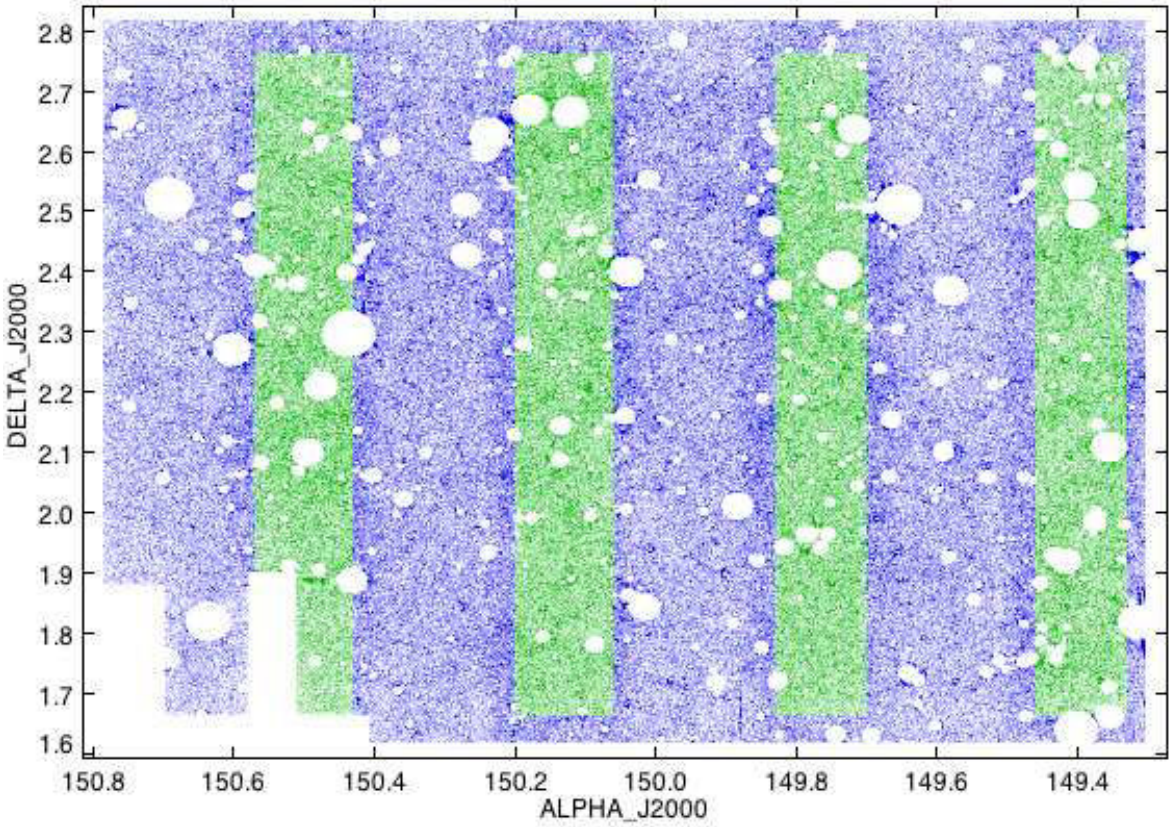


Fig. 1. The footprint of UltraVista field which is showing the ultra-deep (DR2) data populate inside the green region and the deep data inside the blue region (DR1).

Table 2. The 5σ limiting magnitudes for the relevant optical and near-infrared data used in this study.

COSMOS/UltraVISTA Filter	Deep (5σ)	Ultra Deep (5σ)	Source
B	28.2		Subaru/SuprimeCam
V	27.5		Subaru/SuprimeCam
r	27.7		Subaru/SuprimeCam
i	27.1		Subaru/SuprimeCam
z	26.1		Subaru/SuprimeCam
Y	24.6	25.7	UltraVISTA
J	24.4	25.1	UltraVISTA
H	23.9	24.7	UltraVISTA
K_s	23.7	24.8	UltraVISTA
3.6 μm	25.3		Spitzer/SPLASH
4.5 μm	25.1		Spitzer/SPLASH

interest, and a high level of contamination (see Sect. 6). Figures 6 and 8 clearly demonstrate the difficulties in separating high-redshift galaxies from and extreme red galaxy located in z~ 2-3 and dwarf-star contaminants without further SED analysis. Theses are two typical cases for which the colour of some low-z interlopers could also meet the pre-selection criteria, and contamination of galactic cool star can still be present because of the photometric error. Because they use more information than the basic three photometric bands used in colour-colour LBG selection, accurate photometric redshifts are able to better discriminate between galaxies at different redshifts and galactic stars.

The SED fitting is performed using the publicly available code Le Phare (Ilbert et al. 2006). We are primarily interested in analysing the probability distribution function (PDF) of the photometric redshift for each galaxy in the UltraVista catalogue to identify the candidates in the redshift range $4.5 \leq z_{phot} \leq 6.5$.

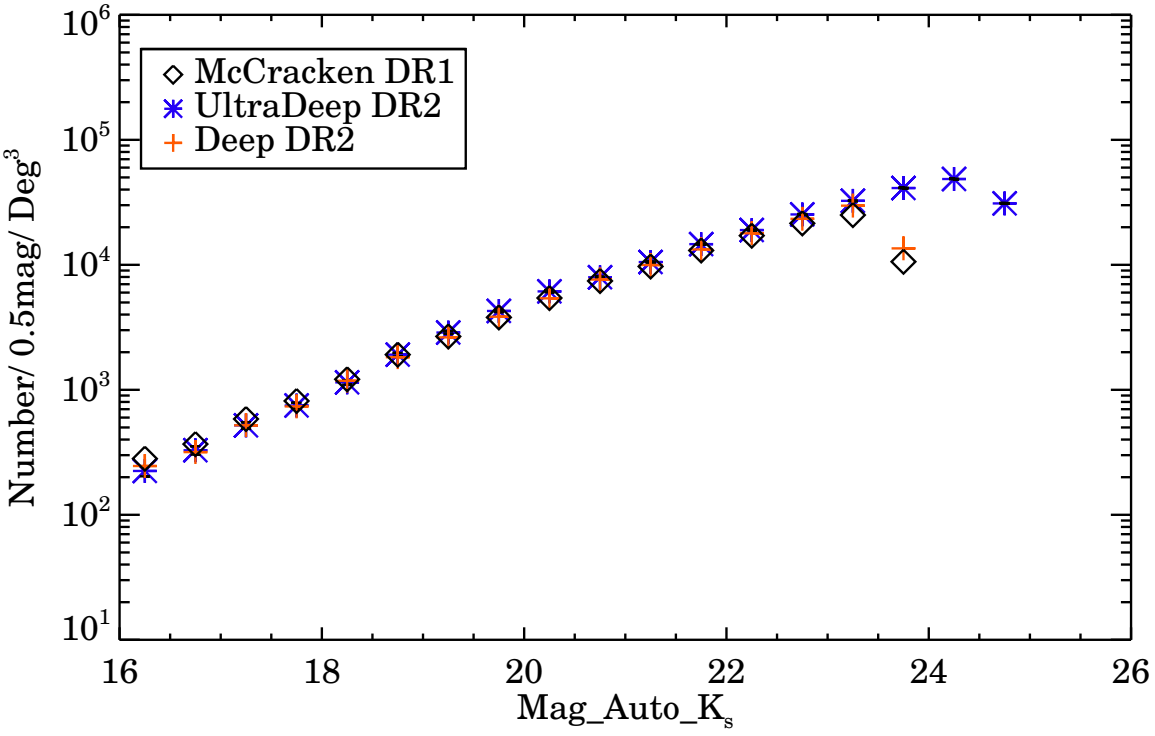


Fig. 2. Galaxy number counts for UltraVISTA, in addition to previous data release (DR1) made by McCracken et al. (2012).

We make also full use of the physical parameter derived from the SED fitting along with the photometric redshift measurement, primarily the stellar mass, star formation rate and dust extinction. In this paper we use the Bruzual & Charlot (2003) stellar evolution models considering models with metallicities ranging from 1 to 0.02 solar metallicity. The ages of the stellar populations models are allowed to range from 10 Myr to 14 Gyr, and are required to be less than the age of the Universe at each redshift. The star formation histories considered here are exponentially declining with characteristic timescales in the range 30 Myr to 30 Gyr. Dust reddening is described by the Calzetti et al. (2000) attenuation law, and allowed to vary within the range 0 < E(B-V) < 0.6. To discriminate galactic cool star and real high redshift objects, it is essential to include galactic star templates including the reddest known stars and we include in our template set L, M and T dwarf stars spectral templates. We also include emission lines in our SED fitting, as these can significantly modify broad band magnitude fluxes (Ilbert et al. 2009; de Barros et al. 2014). The most prominent emission lines like Ly-α, [OII]3727, H-β, [OIII]4959/5007, and Hα are added based on the available UV photons computed using the UV continuum of the model and photoionisation models based on case–B recombination.

In addition to the photometric redshift selection, the sample produced in this study is continuum magnitude limited down to z_{AB} = 26.5 for 4.5 ≤ z ≤ 5.5 and Y_{AB} = 25.5 for 5.5 ≤ z ≤ 6.5. In addition we impose detection above the 1σ limit for all other bands redder than Lyman-α. These magnitude limits are expected to produce a N(z) decreasing with redshift over the range probed following the equivalent limiting magnitude of the survey, as identified for example in the VVDS i-band limited spectroscopic survey (Le Fevre et al. 2013).

3.2. Visual classification: final catalog

At the high redshift probed here, and because of the degeneracies still present even in photometric redshifts, we took a careful approach to select a final list of candidate galaxies. Two team members (PWW and OLF) examined and classified by eye the SED of all candidate galaxies for which Le Phare gave 4.5 ≤ z_{phot} ≤ 6.5. We used the following classification scheme:

- flag 1: excellent candidate at the z_{phot} computed by Le Phare. The SED shows a well defined break below Lyα, the photometric points above Lyα are well behaved following either a power-law for young star-forming galaxies, or a more evolved, redder, SED with often a significant detection in the Spitzer-warm bands. The χ^2 from the SED fit with Le Phare is the lowest between galaxy, AGN, or star templates. The z_{phot} PDF shows a single narrow peak (give typical dz width) around the best redshift.
- flag 2: good candidate at the z_{phot} computed by Le Phare. The SED shows a well defined break below Lyα, the photometric points above Lyα show some scatter but are compatible with a power-law for young star-forming galaxies, or a more evolved, redder, SED with often a significant detection in the Spitzer-warm bands, and the χ^2 from Le Phare is the lowest between galaxy, AGN, or star templates. The z_{phot} PDF

shows a single peak, which may be broader than for flag 1 (give typical dz width), especially for the faintest sources.
- flag 3: degenerate z_{phot}. The χ^2 of the Le Phare fit does not allow to discriminate between galaxies at high or low redshifts, or between high redshift galaxies and galactic stars: the z_{phot} PDF shows two or more peaks of equivalent probability.
- flag 4: good low redshift candidate: The SED shape, usually red, is not compatible with a typical drop-out SED supposed to show a sharp transition at Ly-α at these redshifts, but rather shows a more continuous SED. The z_{phot} PDF shows two peaks (give typical dz width) one at high-z retained by Le Phare and one at low-z with nearly equivalent probability. The lower redshift case shows a high dust extinction.
- flag 5: possible low redshift candidate.
- flag 6: galactic star
- flag 7: good AGN/QSO candidate. The SED shows a well defined break below Lyα, the photometric points above Lyα are well behaved following a power-law as expected from QSOs. The z_{phot} PDF shows a single narrow peak (give typical dz width) around the best redshift. The main classification criterion used here is the maximum flux reach by the SED, when going significantly brighter than a magnitude $AB = 24$ it is more likely that this is a QSO than a galaxy. At the redshifts probed here the SED of a QSO and a star-forming galaxy are rather similar and hard to discriminate.
- flag 8: bad candidate. This corresponds to several failures cases in computing photometric magnitudes: either the SED shows a high dispersion of points, one or more points in the SED have unrealistic values indicating a catastrophic measurement, one or more of the photometric points below the putative break below Ly-α show a strong detection.
- flag 9: candidate that cannot be classified. This is for objects for which the photometry shows a high dispersion of points, and the Le Phare fit has not converged on a meaningful z_{phot}.

In Table 3 we give the numbers of galaxies for each flag category compared to the total number of galaxies as originally selected after the blind run of Le Phare fitting. The excellent and good candidates (flags 1 and 2) represent 58% and 32% of the original Le Phare lists at $4.5 \leq z_{phot} \leq 5.5$ and $5.5 \leq z_{phot} \leq 6.5$, respectively. In total we obtain a sample of 2036 excellent/good candidate galaxies with $4.5 \leq z_{phot} \leq 5.5$ and 330 with $5.5 \leq z_{phot} \leq 6.5$. This forms our best sample analysed in this paper.

Our classification scheme is based on an initial sample preselected from the Le Phare photometric redshifts in $4.5 \leq z_{phot} \leq 6.5$, therefore we cannot recover those galaxies with $4.5 \leq z_{phot} \leq 6.5$ but for which the Le Phare z_{phot} is wrong. In particular there is a well known degeneracy between the Lyman break and the Balmer break, occasionally confusing the SED fitting and assigning a low redshift corresponding to the Balmer break rather than a high redshift associated to the Lyman break. This effect is quantified in the VUDS spectroscopic survey as both the first and second peaks of the z_{phot} PDFs have been used to pre-select targets for spectroscopy. As shown in Le Fèvre et al. (2014) about 10% of the galaxies with high redshifts $z > 4$ are wrongly placed at $z < 2$ by photometric redshift codes like Le Phare. This effect will be taken into account when deriving number counts and luminosity functions in the following.

4. Selection function: photometric simulations

4.1. Simulations

In this section we describe the creation and injection of simulated galaxies into the images to determine the fraction of lost and undetected sources resulting from the blending effects and photometric errors especially at the faint-end of the data. These simulations are also used to predict the tracks of galaxies at these redshifts in colour-colour diagrams compared to observed galaxy samples.

We produce model SEDs from Le Phare using a BC03 (Bruzual & Charlot 2003) stellar population synthesis model with metallicity 0.02, 0.004 and 1 (solar metallicity), a declining star formation history with tau is equal to 0.03, 1, 5, 10 and 30Gyr and a dust-reddening e(B-V) set to 0 and 0.5. We assume that galaxies are point-like sources, which is a reasonable hypothesis for ground-based 0.8 − 1.0 arcsecond FWHM imaging given the observed sizes of galaxies at these redshifts (van der Wel et al. 2014). We generate 20 000 galaxies with a broad range of SED defined by the above model parameters and a K_s band in AB magnitude randomly distributed between 23.5 and 25.5. The magnitudes of these model galaxies are computed in each band of the COSMOS photometry including the UltraVista photometry presented here, and the corresponding images are added to each real image of the field (avoiding masked regions).

4.2. Detection rate

We run SExtractor in dual-image mode on the images including the simulated galaxies using the same parameters set used for the observed images. We then retrieve all real objects plus the mock galaxies. From a total of 20 000 mock galaxies, we recover 13 312 galaxies. We use this resulting catalog of simulated sources to compute the detection rate as a function of magnitude. We find that it is 21% to 87% in the deep region, and 43% to 88% in ultra-deep region, as shown in Figure 5.

5. Properties of the UltraVista sample at $4.5 \leq z \leq 6.5$

5.1. Magnitude and Redshift Distribution

The properties of our sample galaxies are summarized in the following. Figure 8 shows the magnitude-redshift distribution for the two redshift ranges. Our sample is limited to $z_{AB} = 26$ for $4.5 \leq z \leq 5.5$, and to $Y_{AB} = 25.5$ for $5.5 \leq z \leq 6.5$. The redshift distribution of the full sample identified by redshift range is presented in Figure 9. The number of galaxies decreases with redshift as expected, with a dispersion around the mean around 10% at 1σ, corresponding to the expectation of cosmic variance in this field (Moster et al. 2011).

5.2. Luminosity, stellar mass and dust

The distribution of galaxy luminosities, stellar mass $M_\star$ and dust extinction as a function of redshift are shown in Figure 10

6. Comparison to LBG selection and VUDS spectroscopic samples

6.1. Comparison with the VUDS spectroscopic sample

We apply the same selection criteria to the VUDS spectroscopic sample as applied to the UltraVista data, and produce a refer-

Table 3. Result of the visual SED classification of galaxies at $4.5 \leq z_{phot} \leq 6.5$ on the Le Phare pre-selection list.

Redshift bin	Pre-selection	flag 1	flag 2	flag 3	flag 4	flag 5	flag 6	flag 7	flag 8	flag 9
$4.5 < z < 5.5$	3 506	434	663	38	901	894	375	4	171	26
$5.5 < z < 6.5$	1 031	76	103	18	133	375	223	4	99	0

ence catalogue of sources with spectroscopic redshifts. VUDS contains a total of 293 galaxies with spectroscopic redshifts in $4.5 < z < 6.5$, including 161 galaxies in the COSMOS field, 109 in the VVDS-02h field, and 23 in the ECDFS field. There are 150 sources in the UltraVista photometry matching with VUDS sources in the area in common between the two surveys in the COSMOS field.

The redshift distribution of the VUDS sample is compared to the photometric sample obtained in Section 3 as shown in Figure 15. Down to the $i_{AB} = 25$ limit of VUDS, the two distributions are in excellent agreement. The UltraVista photometry reaches deeper than the VUDS spectroscopy by more than one magnitude, an important asset when computing the luminosity function as described in the next sections.

6.2. Comparison with LBG and drop-out selection

We analyse here what are the main differences between the sample selection based on photometric redshifts used in this paper, and the LBG dropout selection used in the literature. We compare the colour-colour locus of galaxies in our sample to a classical colour-colour selection box. For drop-out galaxies in three bands photometry we use the following colour criteria defined using tracks of galaxies as modelled in Sect. 4, and comparable to what is generally practised in the literature for LBG selected samples (Bouwens et al. 2014):

r-dropout at $z \simeq 5$:

- colour box: $r\text{-}i > 0.75$, $-0.4 < i\text{-}z < 1.2$, $r\text{-}i > 1.6(i\text{-}z)+0.7$
- i and z brighter than 5σ

i-dropout at $z \simeq 6$:

- colour box: $i\text{-}z > 0.8$, $z\text{-}j < 1.2$
- B fainter than 2σ and V fainter than 3σ

Galaxies selected by photo-z distribute differently in the (r-i,i-z) diagram for $4.5 \leq z \leq 5.5$ (Figure 3) and in the (i-z,z-j) diagram for $5.5 \leq z \leq 6.5$ (Figure 4) compared to the locus imposed by LBG selection. It is immediately evident that the photometric redshift selection picks-up additional galaxies than the LBG selection does. The photometric redshift selection includes 28% and 4% more galaxies at $z \sim 5$ and $z \sim 6$, respectively.

We can compare LBG-selected samples at these redshifts to the VUDS spectroscopic sample to assess the completeness and contamination rate of LBG selection. The VUDS sample is primarily selected using photometric redshifts computed using Le Phare in the same way as we have selected the UltraVista candidates, and VUDS then obtains spectroscopy to confirm these candidates. We plot in Fig. 4 the VUDS galaxies with very reliable spectroscopic redshifts; these follow the general distribution of our UltraVista sample as expected from the photometric redshift selection. It is again clear that galaxies with secure spectroscopic redshifts are indeed identified outside of a colour-colour box one would naturally draw from simulated galaxy tracks. Using the colour-colour LBG selection box, we compute that at least 47% of the galaxies in the redshift range $4.5 \leq z \leq 5.5$ and with $i_{AB} \leq 25$ are lost by LBG selection. Using the VUDS sample we also compute that LBG selected samples at these redshifts include 70% of false candidates of galaxies at redshifts outside the same redshift range.

7. Luminosity functions and density in 4.5<z<6.5

We use the galaxy counts in each redshift range to compute luminosity functions. LFs are computed using the ALF code (Ilbert et al. 2005) with four different indicators: the V_{max}, C^+, SMWL, and STY. We parametrize luminosity functions following the Schechter function (Schechter 1976) with three main parameters: the characteristic luminosity M^*, density ϕ^* and faint end slope α:

$$\phi(M)dM = 0.4ln(10)\phi^* 10^{0.4(M^*-M)(\alpha+1)} exp[-10^{0.4(M^*-M)}]dM \quad (1)$$

We present luminosity functions in Figures 16 and 17.

From the LF we compute the luminosity density as:

$$LD = \int_{faint}^{bright} \phi(L)LdL \quad (2)$$

and transform it to star formation rate density using:

$$SFRD(z) = 1.4 \times 10^{-28} LD_{UV}(z) 10^{0.4A_{UV}(z)} \quad (3)$$

The main uncertainty in computing the SFRD is to estimate the dust extinction A_{UV}. Discussion TBD.

8. Discussion and conclusion

We use the deep near-infrared data from UltraVista to identify galaxies in the redshift range $4.5 < z < 6.5$. We use photometric redshift to select candidates, and find that this is a very effective selection procedure requiring less approximations and fewer uncertainties than LBG drop-out selection. We assemble a sample of more than 2 000 galaxies in $4.5 < z < 5.5$, and more than 300 in $5.5 < z < 6.5$. The sample is compared to other existing photometric and spectroscopic samples.

Because of the large area sampled the bright end of the LF is determined more robustly than any previous study, giving a unique reference. It was possible to obtain an improved measurement of the bright end because the UltraVista survey area is ∼ 10 times larger than other surveys at the same depth, for a volume surveyed large enough to find the relatively rare bright galaxies. The luminosity function at $z \sim 5$ is well represented by a Schechter function and compares well with the most recent estimates in the literature (Bouwens et al. 2014). At $z \sim 6$, the LF shows an excess of bright galaxies compared to a standard Schechter function. At the bright end of the z∼ 6 LF $M_U < -22$ we find 50% more galaxies than previous surveys. This is comparable to the results recently obtained by Bowler et al. (Bowler et al. 2014b,a) from a different analysis of the same dataset.

This result has important consequences on the luminosity density (LD) at these high redshifts: integrating the LF we find that the LD is as high as the highest measurements to date. This directly translates to a higher star formation rate density: we plot

our SFRD measurements in Figure 20, compared to other results from the literature. We find a level of star formation in the first phases of galaxy evolution as high as reported by Bouwens et al. (2014). This SFRD level compared to lower redshift measurements indicates that the star formation has increased relatively mildly by a factor of ∼3 from $z \sim 6$ to $z \sim 2$. The high SFRD at $z \sim 6$ indicates a significant UV background just after reionisation. More TBD

Acknowledgements. This work is supported by funding from the European Research Council Advanced Grant ERC-2010-AdG-268107-EARLY. This work is based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatories under ESO programme ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy survey Unit on behalf of the UltraVISTA consortium. This work is based on COSMOS survey data products made available at the CESAM data center, Laboratoire d'Astrophysique de Marseille. This work partly uses observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS. This work is based in part on observations made with the Spitzer Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under NASA contract 1407.

References

Bertin, E. & Arnouts, S. 1996, A&AS, 117, 393
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, ArXiv e-prints
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2014a, ArXiv e-prints
Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2014b, MNRAS, 440, 2810
Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F. 2009, Nature, 459, 49
Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31
Dalton, G. B., Caldwell, M., Ward, A. K., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 0
de Barros, S., Schaerer, D., & Stark, D. P. 2014, A&A, 563, A81
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Ilbert, O., Tresse, L., Zucca, E., et al. 2005, A&A, 439, 863
Le Fevre, O., Cassata, P., Cucciati, O., et al. 2013, ArXiv e-prints
Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, ed. M. Iye & A. F. M. Moorwood, 1670–1681
Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2014, ArXiv e-prints
Lilly, S. J., Le Fevre, O., Hammer, F., & Crampton, D. 1996, ApJ, 460, L1
Madau, P. & Dickinson, M. 2014, ARA&A, 52, 415
Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, MNRAS, 283, 1388
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, ApJ, 731, 113
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Schechter, P. 1976, ApJ, 203, 297
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Steinhardt, C. L., Speagle, J. S., Capak, P., et al. 2014, ApJ, 791, L25
Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Warren, S. J., Hambly, N. C., Dye, S., et al. 2007, MNRAS, 375, 213
White, S. D. M. & Rees, M. J. 1978, MNRAS, 183, 341

Magnitude distribution of the UltraVista candidate galaxies with $4.5 < z < 6.5$ after applying our photometric selection. (top) z-band distribution for $4.5 < z < 5.5$, (bottom) $Y-$band distribution for $5.5 < z < 6.5$.

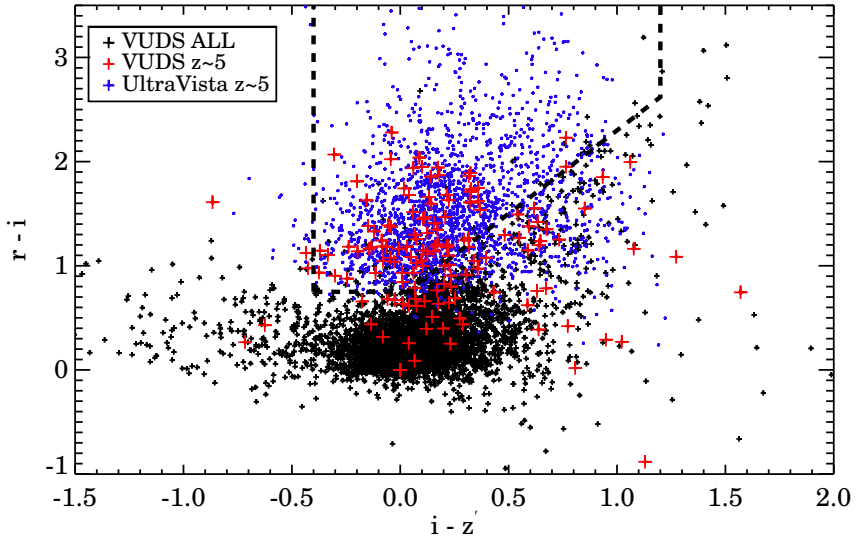


Fig. 3. Colour-colour diagram used to select *r*-dropout galaxies over VUDS fields. The black solid-line box shows how we differentiate z~ 5 candidates and low redshift interlopers from the spectroscopic sample (VUDS).

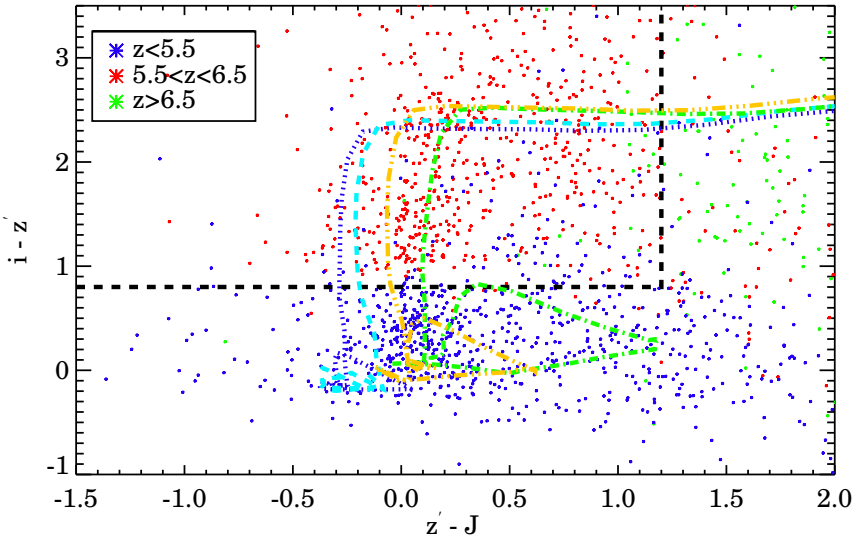


Fig. 4. *izJ* colour-colour diagram used to select *i*-dropout galaxies over UltraVista fields. The black dashed-line box shows how we differentiate z~ 6 candidates and low redshift interlopers. Four coloured dashed lines shows 4 galaxy tracks based on different combinations by stellar age and metallicity from BC03 SED template. Lower redshift objects (z < 5.5) shown in blue points, intermediate redshift objects (5.5 < z < 6.5) shown in red points and high redshift (z > 6.5) shown in green points.

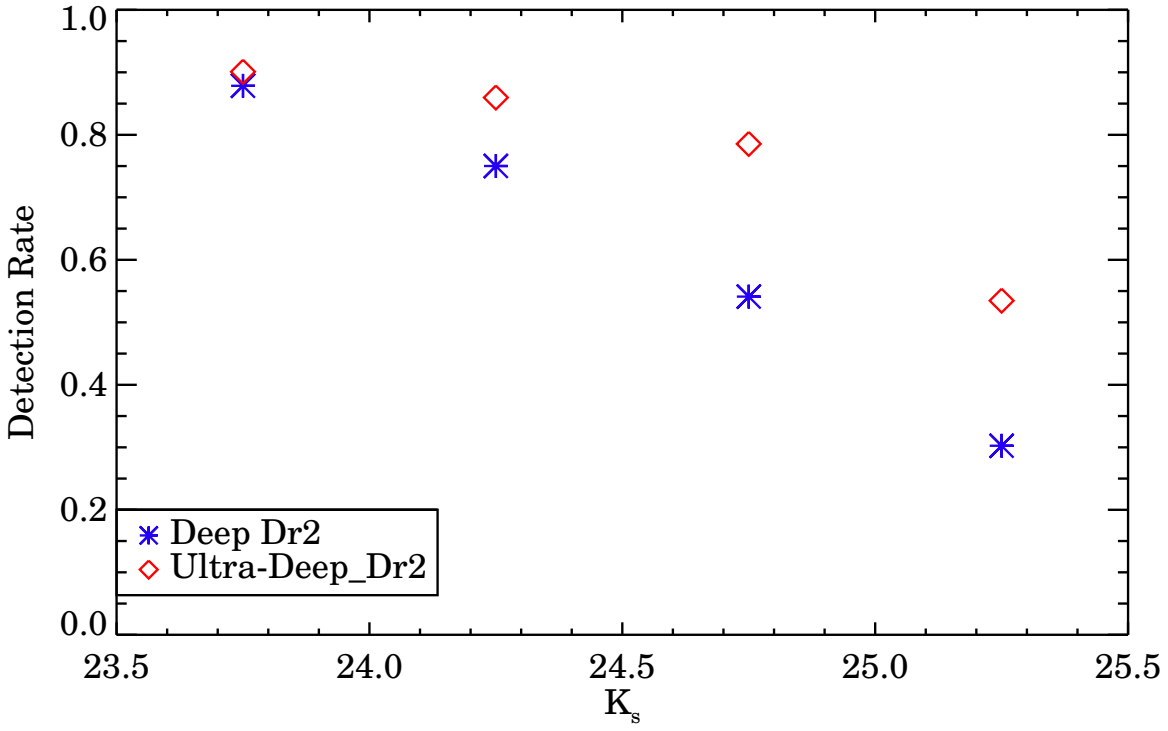


Fig. 5. Detection rate of K_s band magnitude from our image simulation in the deep and ultra-deep regions.

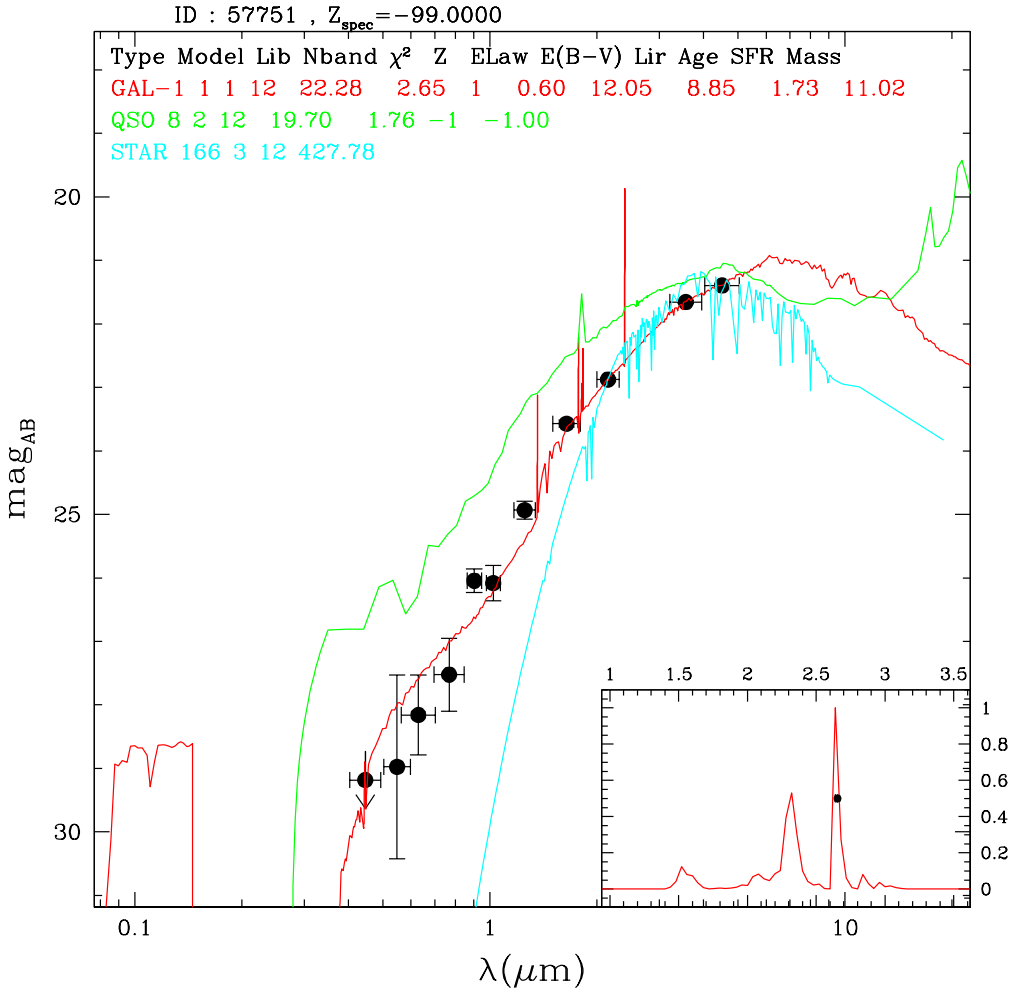


Fig. 6. SED plot of a massive extreme red galaxy in 12 photometric broad bands showing the strong colour break as high redshift LBG.

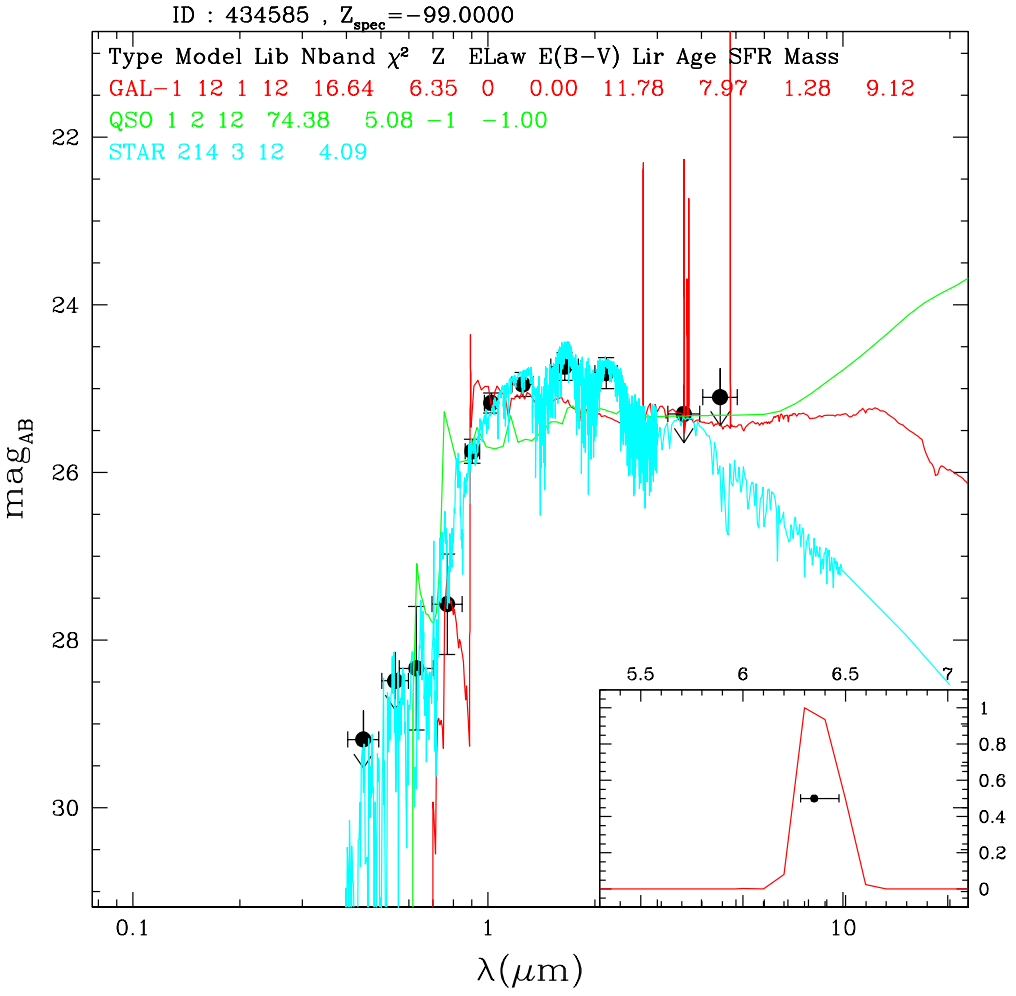


Fig. 7. One of the typical SED plot of galactic cool star. It can still be present just rely on simple colour selection without SED analysis.

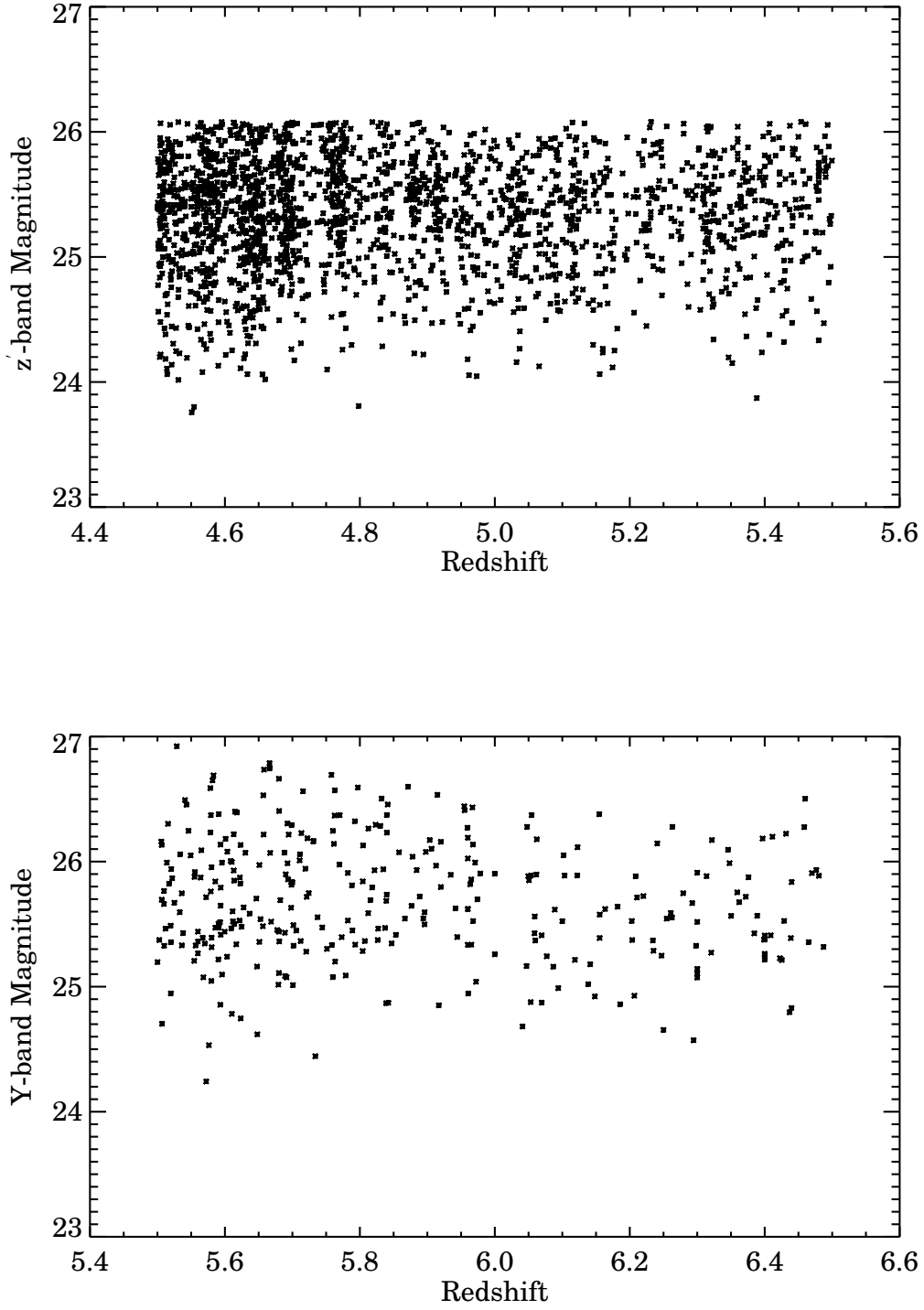


Fig. 8. Magnitude distribution of the UltraVista candidate galaxies with 4.5 < z < 6.5 after applying our photometric selection. (top) z-band distribution for 4.5 < z < 5.5, (bottom) $Y-$band distribution for 5.5 < z < 6.5.

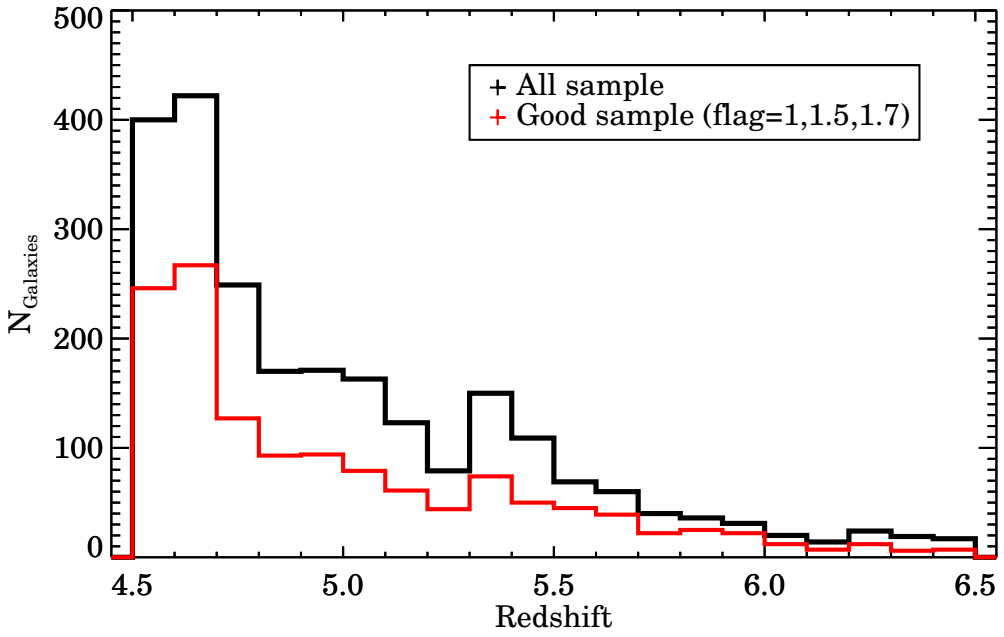


Fig. 9. Redshift distribution of the UltraVista candidate galaxies with 4.5 < z < 6.5 after applying our photometric selection.

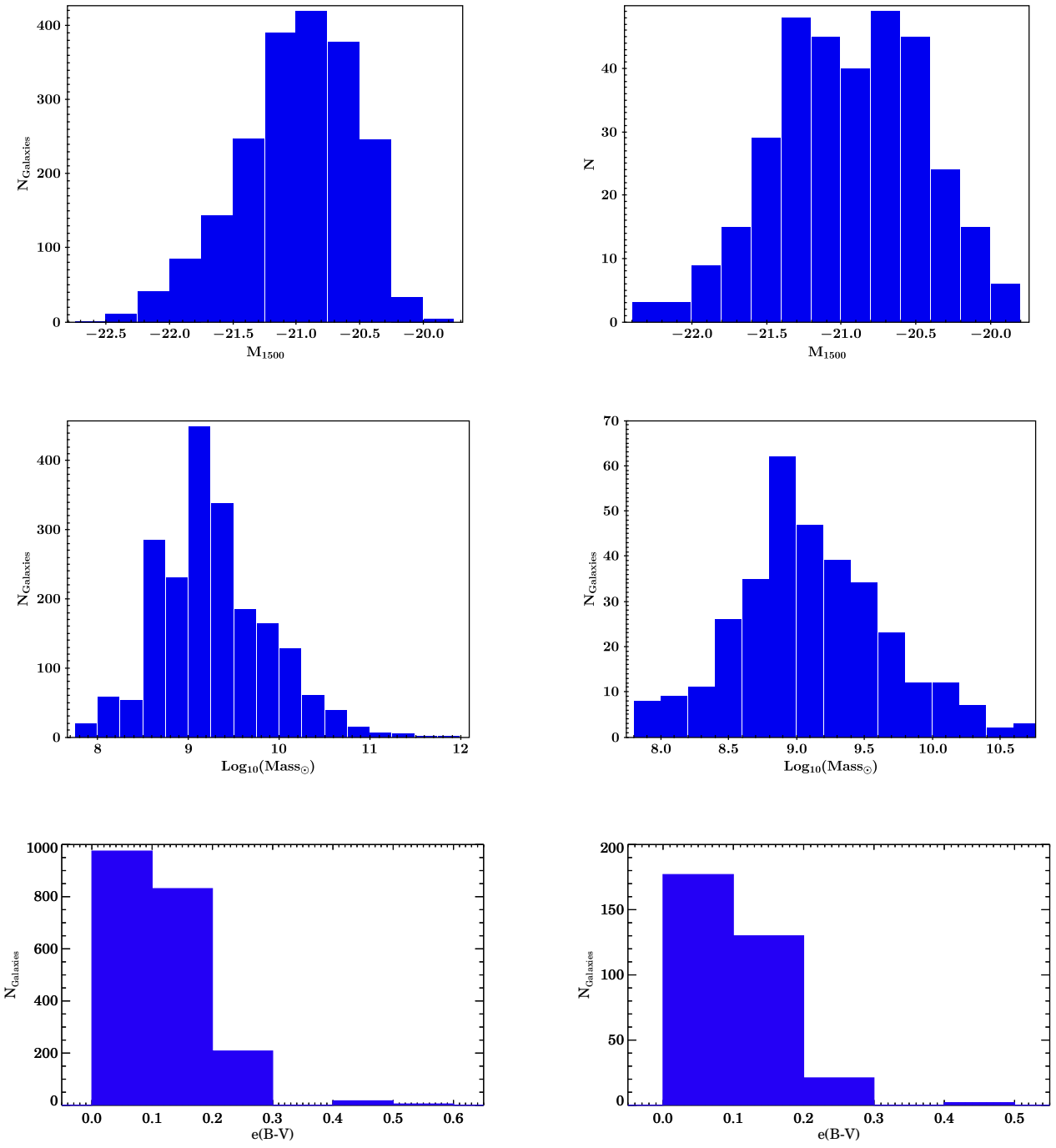


Fig. 10. Distribution in Luminosity (top panel), stellar mass $M_\star$ (middle panel) and dust extinction (lower panel) as a function of redshift for the $4.5 < z < 6.5$ UltraVista sample. Plots located in left side with $4.5 < z < 5.5$ and plots located in right side with $5.5 < z < 6.5$

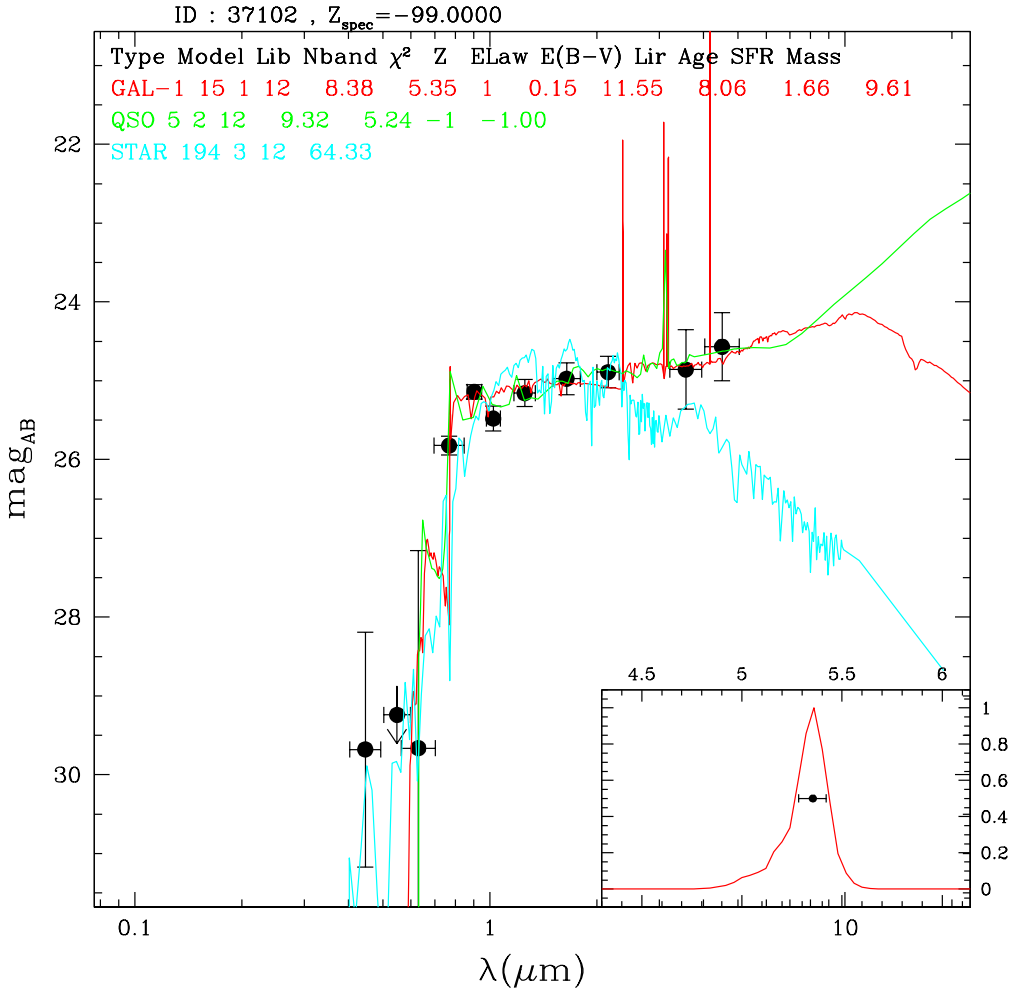


Fig. 11. Example SED of a galaxy at z_{phot} = 5.35 with a well-defined photometric redshift PDF. The photometric redshift is calculated by fitting to all available photometric bands including the IRAC channel 1 and channel 2 filters.

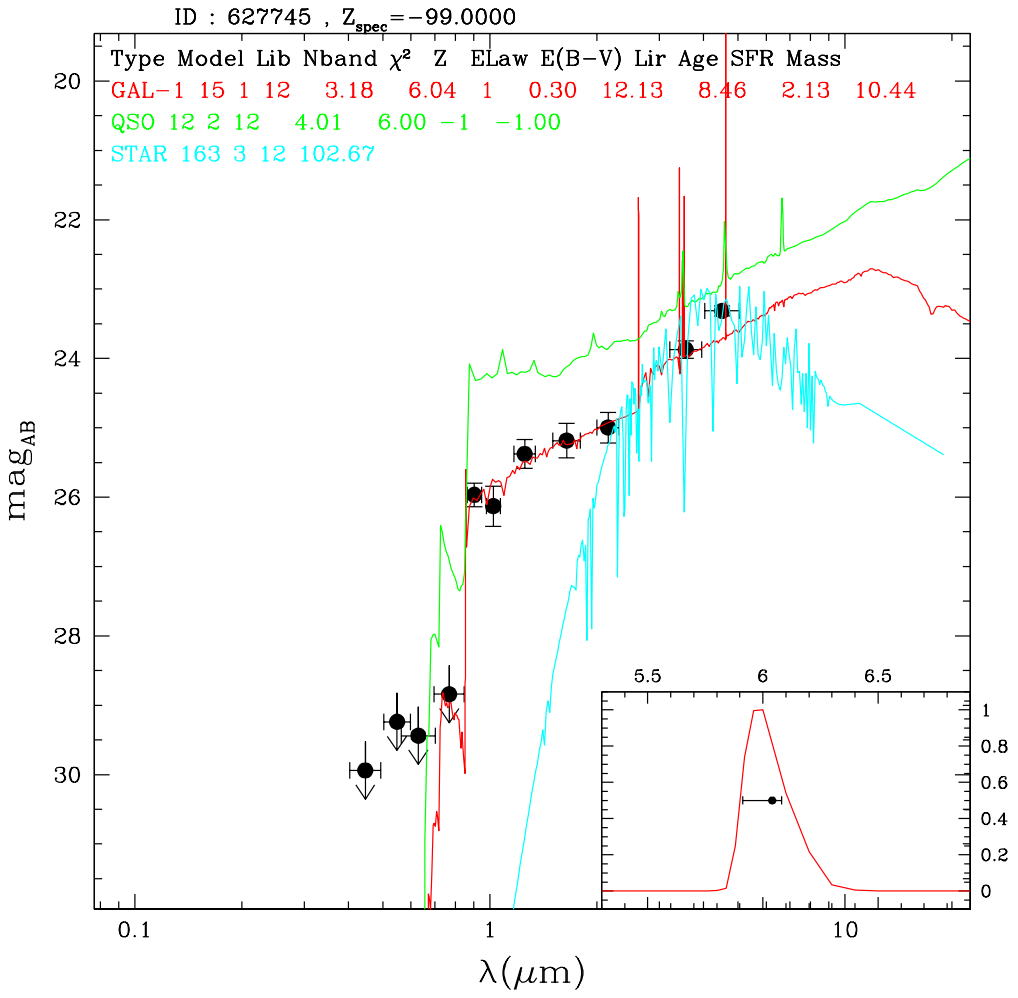


Fig. 12. Example SED of a galaxy at z_{phot} = 6.04 with a well-defined photometric redshift PDF.

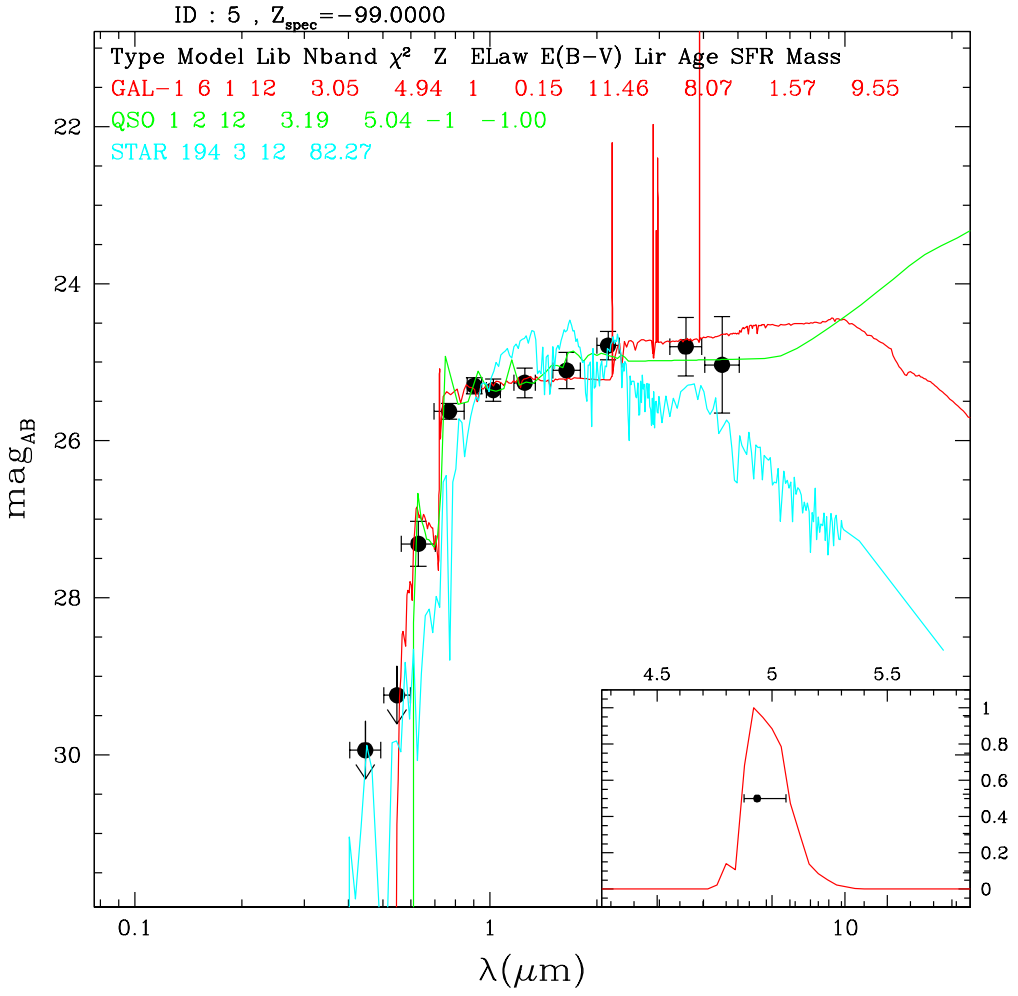


Fig. 13. Average SED produced from stacking from all credible SEDs of z~5 candidates (TBD).

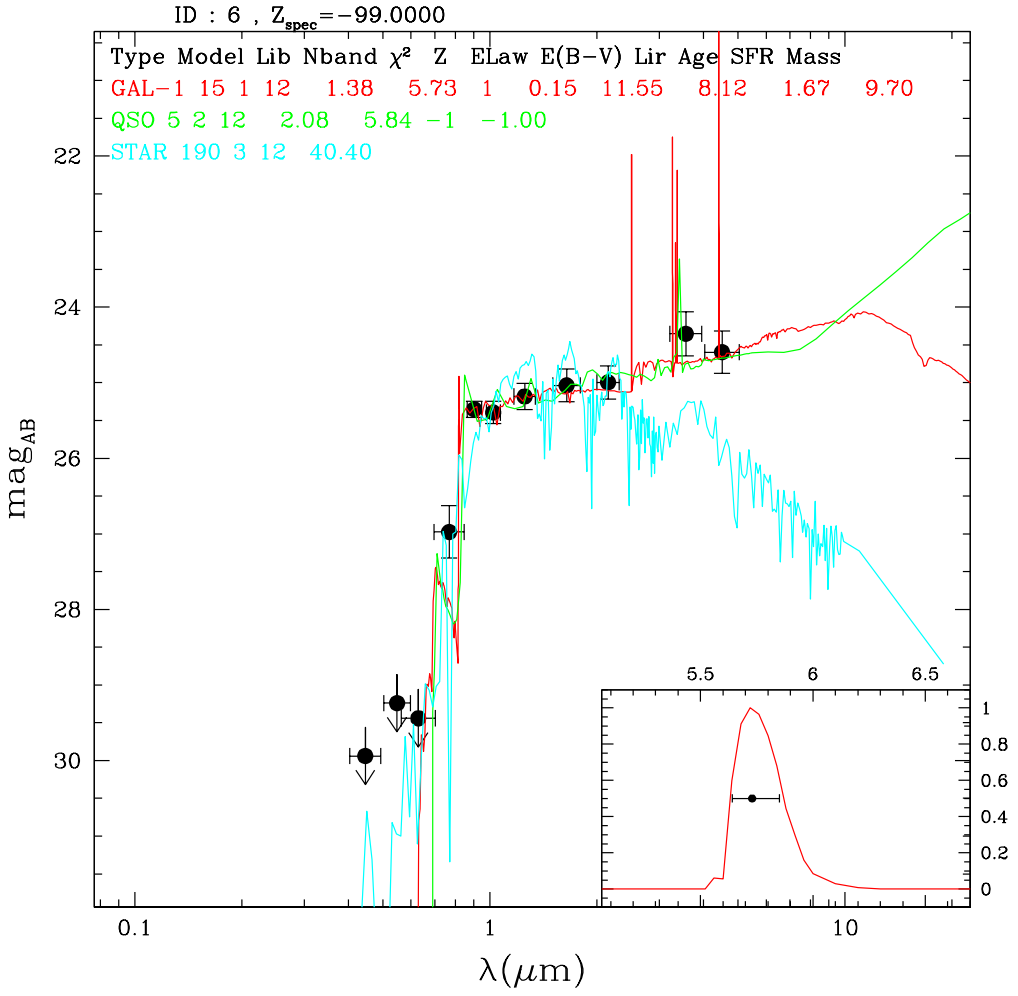


Fig. 14. Average SED produced from stacking from all credible SEDs of z~6 candidates (TBD).

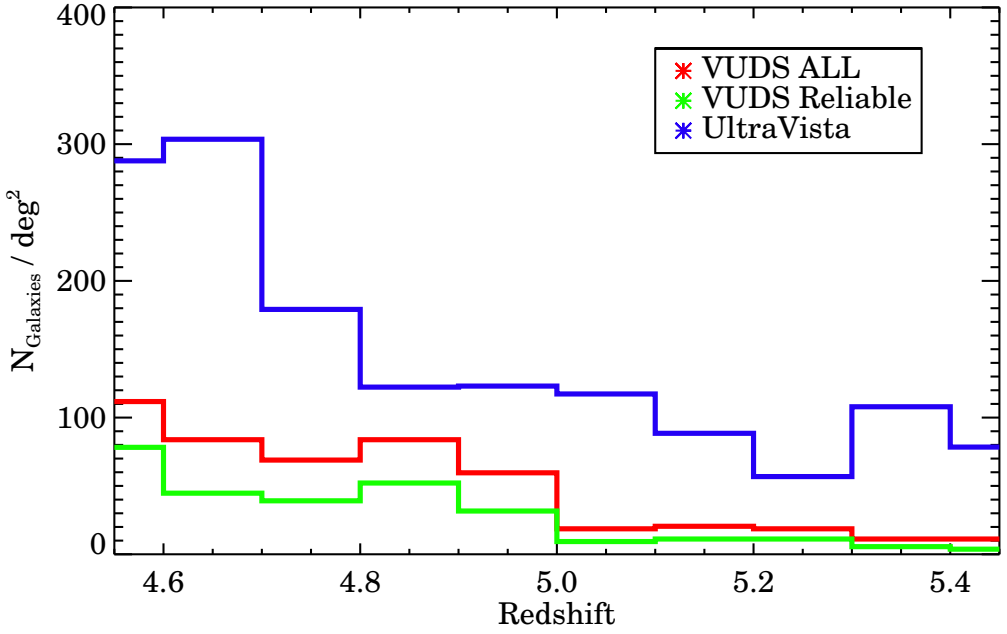


Fig. 15. Redshift distribution of galaxies with spectroscopic redshifts $4.5 < z_{spec} < 6.5$ in the VUDS survey compared to the redshift distribution to the photometric sample obtained in this paper from the UltraVista survey. The galaxy counts for each survey are converted to number of galaxies per square degree using the sampling rates and corresponding areas covered by each survey.

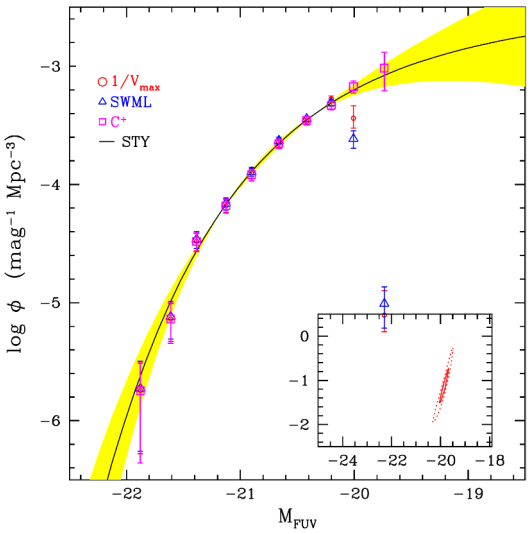


Fig. 16. The rest-frame UV LF at $z \sim 5$ computed from our sample of 2036 candidates selected from the UltraVista photometry. Four different indicators are used to compute the LF (V_{max}, SWML, C^+ and STY), as indicated in the upper left panel. The error ellipse in the (α, $M_\star$) plane is indicated in the lower right panel.

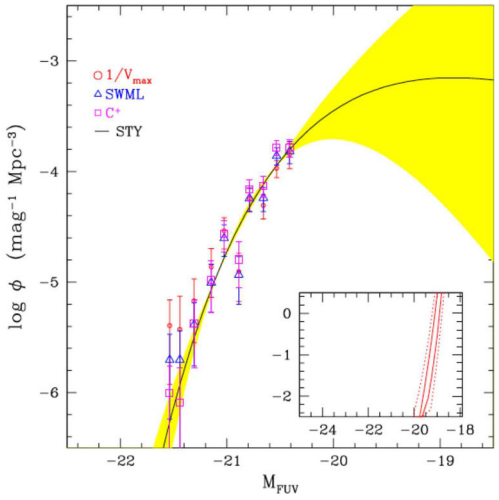


Fig. 17. The rest-frame UV LF at $z \sim 6$ computed from our sample of 330 candidates selected from the UltraVista photometry. Four different indicators are used to compute the LF (V_{max}, SWML, C^+ and STY), as indicated in the upper left panel. The error ellipse in the (α, $M_\star$) plane is indicated in the lower right panel.

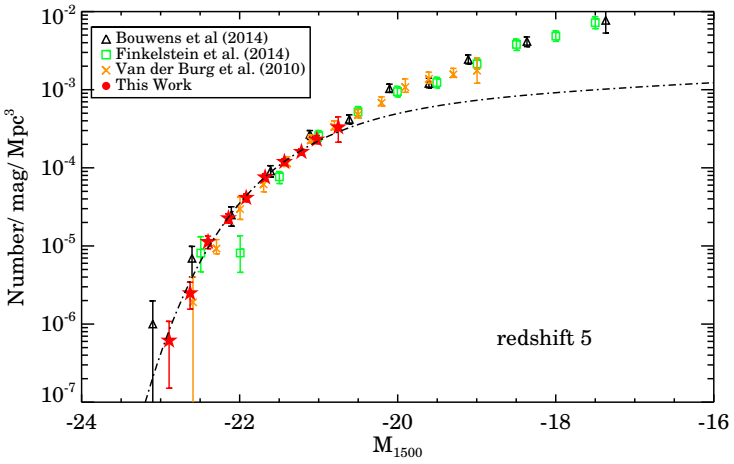


Fig. 18. The rest-frame UV LF at $z \sim 5$ from this work (in red), compared to other results from the literature.

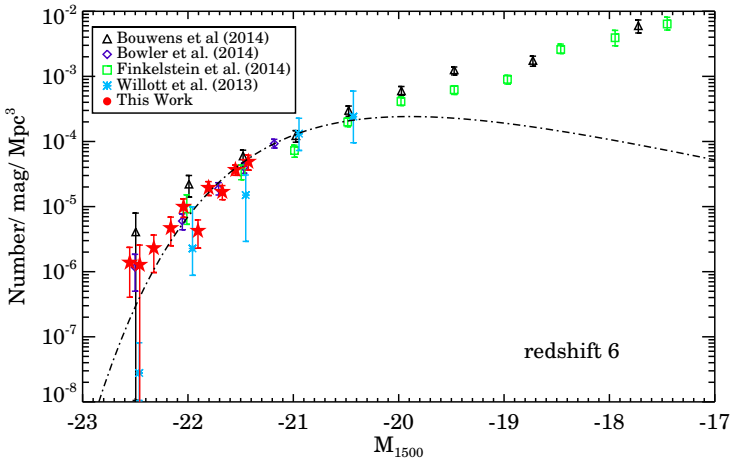


Fig. 19. The rest-frame UV LF at $z \sim 6$ from this work (in red), compared to other results from the literature.

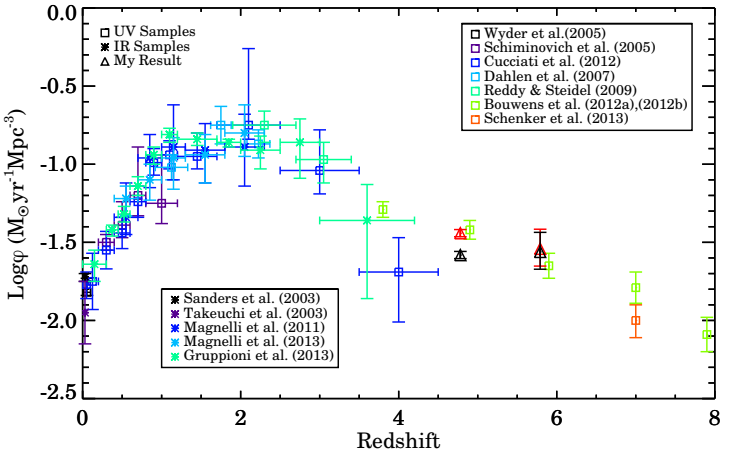


Fig. 20. Total dust-corrected UV derived SFRD, the two red triangle points are SFRD measurements at z = 5 to z = 6 based on my work. At z=5 and 6, the points with black triangles are computed from my sample with a lower extinction 0.5 and 0.15 instead of 0.85 and 0.2. All the other points are compiled from other literature results mainly derived from UV and IR samples (see in the Figure).

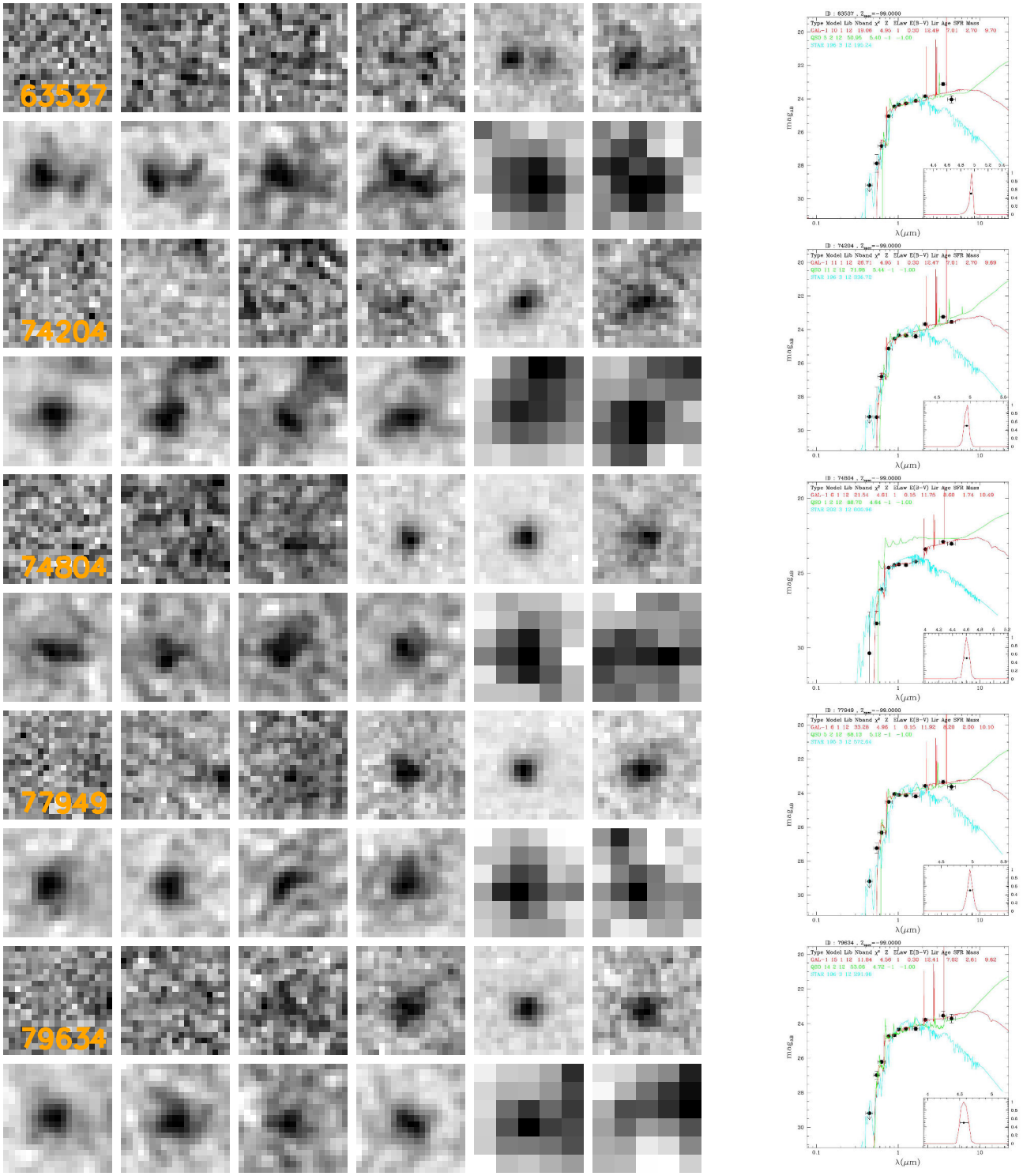


Fig. 21. poststamps and SED plots from Lephare photometric redshift code.

Chapter 8

Galaxy Merger Rate in the High Redshift Universe

The galactic merger rate is one of the fundamental measures of galaxy evolution which can tell us how galaxies grow over time through encounters with other galaxies. Galaxy mergers can transform star-forming disks (blue cloud) to spheroids (red sequence), trigger starbursts, feed super-massive black holes and ignite AGN (DiMatteo & Springel 2005) and even enrich the IGM. Past measurements of this number have yielded a wide range of results (as shown in Figure 8.1). For better understanding the SFRD through cosmic time, especially at high redshift, the galaxy merger rate is a crucial property we need to investigate.

Observationally, there are two method to identify galaxy mergers, one is checking morphological transformation and another is counting close galaxy pairs. Each of them associates drawback and cautionary part. At first, accurately identifying mergers by morphological indicator is extremely difficult at high redshift. Galaxies of various morphologies and sizes interact very differently compare to the local universe, and it is hard to characterise the morphological pattern or significancy based due to the surface brightness dimming effect and limit of resolution. Counting pairs from imaging alone is affected by projection effects along the line of sight, and one cannot get information on the velocity difference between the two galaxies to assess the reality of the pair. Simply speaking, either the pair observed in imaging has a physical connection and the two galaxies will eventually merge, or the pair results from a projection effect and should not be counted as a merger. It is therefore necessary to obtain relative velocity information from spectroscopic observation of the two components of the pair. Typically, a pair with transverse separation of about 25 kpc and velocity separation of 500 km/s is expected

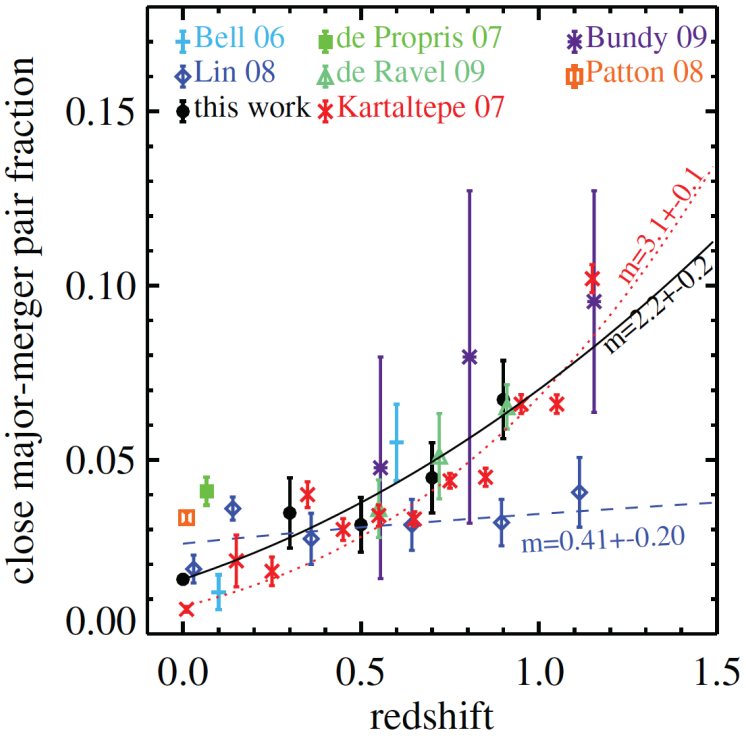


FIGURE 8.1: Comparisons of observed pair fractions in the literature. When it is appropriate, results of different authors were corrected so they are consistent with a common definition of close major-merger pairs with the maximum projected separation of r=20 kpc and the maximum primary-to-secondary mass ratio is 3. (Credit: Xu et al. 2012)

to merge in 0.5 to 1 Gyr.

I participated to the spectroscopic identification and study of close physical pairs of galaxies at redshifts $2 < z < 4$ based on VUDS and VVDS with the discussion on the impact of major mergers in shaping galaxies at these early cosmological times. I provide a summary of this paper below. I have contributed the infrared photometry and images that we used in this study.

8.1 Scope of the study

Along their evolution, the mass of galaxies increases. A physical framework that describes, in a coherent way, the mass growth of galaxies is still missing. In the current paradigm of galaxy formation and evolution, the hierarchical model, the mass of a galaxy is assembled from a relative contribution of different physical processes. The two main

processes that are thought to be responsible of the mass growth are the cold gas accretion and galaxy mergers. In the first, the cold gas is accreted onto the galaxies from the filaments of the cosmic web and sustains a high star formation rate. However, direct evidence for this accretion process is difficult to identify observationally. Galaxy mergers are a key process in the galaxy evolution scenario. Predicted in the hierarchical growth of structure in the Λ-CDM cosmological framework, mergers of dark matter halos not only increase the dark matter mass but also leads to the merging of the galaxies that lies into these dark matter halos. The typical dynamical timescale for the merger of two galaxies is about 0.5-1.0 Gyr. Consequently, the present day massive halos are expected to be the results of several mergers since its formation, following the so-called merger trees.

While cold gas accretion is hard to observe, mergers are easier to identify and very well documented at low to high redshifts. The identification of mergers can be done from the use of two methods.

- An aposteriori method that relies on the morphology of the resulting merged galaxy through the identification of post-merger morphological signatures.
- An apriori method that is based in the identification of pairs of physically bound galaxies about to merge.

While at low redshifts the first method can be easily used because post-merger low surface brightness signatures can be detected, at very high redshifts the identification of major merging pairs is the only practical alternative (Le Fèvre et al, 2000). From the pair fraction, one can compute a merger rate per volume which gives an access to the amount of stellar mass that has been assembled from the merger process (eg, Lopez Sanjuan et al 2011 & 2013). At z∼1, pairs have been identified from different surveys. With the VVDS (see below) it has been observed that the merging fraction is three times higher than in the local universe. At higher redshifts, pair identification is more sparse and the number of confirmed pairs at $z > 2$ is very limited.

8.2 Data

In this framework we used the VIMOS VLT Deep survey (VVDS) and VIMOS Ultra Deep Survey (VUDS) to identify pairs at $z > 1.8$.

The VVDS survey (Le Fèvre et al, 2013a) is a spectroscopic redshift survey composed of 35 000 galaxies with redshifts between 0 and 6.7. It is organised in three magnitude-selected galaxy sample:

- The 'Wide' sample, selected with $17.5 \leq i_{AB} \leq 22.5$ on an area of 8.6 deg^2. It is composed of 25 805 galaxies.
- The 'Deep' sample, selected with $17.5 \leq i_{AB} \leq 24$ in 0.74 deg^2 and composed of 11 486 galaxies.
- The 'Ultra-Deep' sample, with 938 galaxies, with $23 \leq i_{AB} \leq 24.75$ selected in an area of 512 arcmin^2.

Additionally, 422 AGNs and 217 Lyman Alpha emitters were detected.

VUDS is described in Chapter 4. We used a preliminary VUDS catalog of pairs available in early 2014 to compute the merger fraction and merger rate at $z \sim 3$.

8.3 The identification of pairs and their properties

Pairs in the VVDS and the VUDS surveys are identified with the projected transverse separation r_p and the velocity difference Δv defined as

$$\Delta v = c\frac{v_1 - v_2}{1 - v_1 v_2}. \tag{8.1}$$

v_x represents the normalised velocity defined as $v_x = [(1+z_x)^2 - 1]/[(1+z_x)^2 + 1]$. Our galaxy sample was then scanned with $r_p \leq 25h^{-1}$kpc and $\Delta v \leq 500$km.s^{-1}. Galaxies with quality flags from 2 to 9 were considered for the primary target and from 1 to 9 for the companions (primary and secondary objects). From this criteria, 12 pairs have been identified from $z = 1.82$ to $z = 3.65$.
The properties of those pairs are summarised in Table 8.1.

It shows that our pairs span a large mass range, from $10^{9.7}$M$_\odot$ to 10^{11}M$_\odot$. The galaxies in the twelve pairs show an important on-going star formation and one of them shows a sign of AGN actvity. The mass ratio of the two galaxies for each pairs is in the range 1<M1/M2<6 and with 11 out of 12 satisfies a major merger pair criterion 1<M1/M2<4. Under merger hypothesis, the merging timescale of our pairs goes from 0.6 to 2.3 Gyr. Meaning that these pairs would merge before the peak of star formation at $z \simeq 1.5$ (Cucciati et al, 2012)
From those properties, we derived a first measurement for the major merger fraction, $f_{merge} \sim 15 - 20\%$ at $1.8 < z < 4.0$. This measurement is high, and of the same order that the one at lower redshift (1<z<1.8) in the MASSIV survey (Lopez Sanjuan et al, 2013).

Field	Pair Galaxies ID numbers	Redshift z_1	Stellar Masses[(1)] $10^{10} M_\odot$	Mass ratio[(2)]	Flux ratio[(3)]	$r_p^{(4)}$ h^{-1}kpc	$\Delta v^{(5)}$ km s^{-1}	$T_{merg}^{(6)}$ Gyr	$z_{assembly}^{(7)}$
COSMOS	511001467a&b	2.0970	0.3/0.15	2.0	1.9 ± 0.2	18.4 ± 0.3	87	2.3	1.1
	510788270a&b	2.9629	0.5/0.7	1.4	1.7 ± 0.3	15.4 ± 0.2	38	1.5	1.8
	510175610/510778438	3.0939	0.8/5.0	6.3	4.2 ± 0.2	9.4 ± 0.2	161	0.7	2.3
VVDS-02h	520452183/520450423	1.8370	5.6/3.0	1.9	2.2 ± 0.1	22.3 ± 1.2	390	1.1	1.3
	520478238/520478087	2.2460	2.0/1.0	2.0	1.8 ± 0.4	25.3 ± 1.2	28	1.8	1.3
	910260902/910261083	2.3594	3.4/7.9	2.3	2.0 ± 0.1	12.3 ± 1.1	54	0.6	1.9
	910302317/910302929	1.8171	1.1/1.8	1.6	1.4 ± 0.1	8.7 ± 1.2	181	0.6	1.5
ECDFS	530034527a&b	3.6500	1.2/0.9	1.3	1.5 ± 0.1	6.6 ± 0.2	0	0.5	2.9
	530050663a&b	2.3250	0.4/0.2	2.0	1.9 ± 0.2	5.5 ± 0.3	54	0.6	1.9
	530042814/530042840	2.9903	0.12/0.1	1.2	1.3 ± 0.1	6.0 ± 0.2	278	1.0	2.1
	530046916a&b	2.8684	1.4/0.6	2.3	2.5 ± 0.1	6.5 ± 0.2	279	0.5	2.4
	530036900a&b	3.3300	1.3/1.1	1.2	1.5 ± 0.1	7.0 ± 0.2	76	0.6	2.6

TABLE 8.1: Properties of the twelve identified pairs in the VVDS and VUDS survey. (1):Stellar mass estimates from SED fitting; (2) Stellar mass ratio between the two galaxies in the pair; (3) Flux ratio as measured on H or K-band; (4) r_p : transverse separation between the two galaxies in the pair, errors in r_p have been conservatively estimated taking a one pixel error on the difference in the centroid measurement of each galaxy in the pair; (5) $\Delta\nu$: velocity separation along the line of sight, errors in $\Delta\nu$ are estimated to be <100 km.s^{-1} when the two redshifts are from the same slit (10 pairs), and $\sim$ 200 km.s^{-1} when redshifts are from two different slits (2 pairs); (6) T_{merg}: timescale for the pair to merge, using the Kitzbichler & White (2008) prescription; from errors in mass ratio, r_p and $\Delta\nu$ errors on T_{merg} are $\sim$10-20% ; (7) $z_{assembly}$: redshift by which the two galaxies will have merged and the Δz corresponding to the merger timescale

8.4 Conclusions

These study showed evidence for major merging at very high redshifts $1.8 < z < 4.0$. This means that the hierarchical assembly which produce massive galaxies from the merger of smaller galaxies, is already at play at these redshifts. We compute that a galaxy at $z \sim 3$ will grow more than half of its mass through mergers along its life to z=0.

While it is shown here that the merging is at play at very early times, it is important to emphasise that it does not exclude any other mass growth processes (Tacconi et al. 2010; Daddi et al. 2010) like cold gas accretion. Thus, the assembly of the total mass of a galaxy comes from the relative contribution of these different processes. The ultimate goal would be to place an upper limit to the total mass growth from other physical processes.

8.4.1 Paper: Evidence for major mergers of galaxies at $2 \leq z < 4$ in the VVDS and VUDS surveys

A&A 565, A10 (2014)
DOI: 10.1051/0004-6361/201321507

Astronomy & Astrophysics

Evidence for major mergers of galaxies at $2 \lesssim z < 4$ in the VVDS and VUDS surveys⋆

L. A. M. Tasca[1], O. Le Fèvre[1], C. López-Sanjuan[2], P.-W. Wang[1], P. Cassata[1], B. Garilli[3], O. Ilbert[1], V. Le Brun[1], B. C. Lemaux[1], D. Maccagni[3], L. Tresse[1], S. Bardelli[4], T. Contini[5], S. Charlot[6], O. Cucciati[4], A. Fontana[7], M. Giavalisco[8], J.-P. Kneib[1,9], M. Salvato[10], Y. Taniguchi[11], D. Vergani[12], G. Zamorani[4], and E. Zucca[4]

[1] Aix-Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388 Marseille, France
e-mail: lidia.tasca@oamp.fr
[2] Centro de Estudios de Física del Cosmos de Aragón, Plaza San Juan 1, planta 2, 44001 Teruel, Spain
[3] INAF – IASF, via Bassini 15, 20133 Milano, Italy
[4] INAF – Osservatorio Astronomico di Bologna, via Ranzani, 1, 40127 Bologna, Italy
[5] Institut de Recherche en Astrophysique et Planétologie – IRAP, CNRS, Université de Toulouse, UPS-OMP, 14 avenue E. Belin, 31400 Toulouse, France
[6] Institut d'Astrophysique de Paris, UMR 7095 CNRS, Université Pierre et Marie Curie, 98bis boulevard Arago, 75014 Paris, France
[7] INAF – Osservatorio Astronomico di Roma, via di Frascati 33, 00040, Monte Porzio Catone, Italy
[8] Astronomy Department, University of Massachusetts, Amherst, MA 01003, USA
[9] LASTRO, École polytechnique fédérale de Lausanne, Suisse
[10] Max-Planck-Institut fur extraterrestrische Physik, Giessenbachstrasse, 85748 Garching bei Munchen, Germany
[11] Research Center for Space and Cosmic Evolution, Ehime University, Bunkyo–cho, 790–8577 Matsuyama, Japan
[12] INAF–IASF Bologna, via P. Gobetti 101, 40129 Bologna, Italy

Received 19 March 2013 / Accepted 25 February 2014

ABSTRACT

Context. The mass assembly of galaxies can proceed through different physical processes. Here we report on the spectroscopic identification of close physical pairs of galaxies at redshifts $2 \lesssim z < 4$ and discuss the impact of major mergers in building galaxies at these early cosmological times.

Aims. We aim to identify and characterize close physical pairs of galaxies destined to merge and use their properties to infer the contribution of merging processes to the early mass assembly of galaxies.

Methods. We searched for galaxy pairs with a transverse separation $r_p \leq 25\ h^{-1}$ kpc and a velocity difference $\Delta v \leq 500$ km s^{-1} using early data from the VIMOS Ultra Deep Survey (VUDS) that comprise a sample of 1111 galaxies with spectroscopic redshifts measurements at redshifts $1.8 \leq z \leq 4$ in the COSMOS, ECDFS, and VVDS–02h fields, combined with VVDS data. We analysed their spectra and associated visible and near-infrared photometry to assess the main properties of merging galaxies that have an average stellar mass $M_\star = 2.3 \times 10^{10}\ M_\odot$ at these redshifts.

Results. Using the 12 physical pairs found in our sample we obtain a first robust measurement of the major merger fraction at these redshifts, $f_{\rm MM} = 19.4^{+9}_{-6}\%$. These pairs are expected to merge within 1 Gyr on average each producing a more massive galaxy by the time the cosmic star formation peaks at $z \sim 1$–2. Using the pairs' merging time scales, we derive a merging rate of $R_{\rm MM} = 0.17^{+0.08}_{-0.05}$ Gyr^{-1}. From the average mass ratio between galaxies in the pairs, the stellar mass of the resulting galaxy after merging will be ∼60% higher than the most massive galaxy in the pair before merging. We conclude that major merging of galaxy pairs is on-going at $2 \lesssim z < 4$ and is significantly contributing to the major mass assembly phase of galaxies at this early epoch.

Key words. galaxies: formation – galaxies: high-redshift

1. Introduction

The contribution of different physical processes to galaxy mass assembly along cosmic time is still unknown, and a clear picture describing how galaxies assemble, supported by observational evidence, has yet to emerge. Looking back at the average history of a galaxy observed today, we are still unable to identify how and when its stellar mass has been acquired and which physical processes are possible contributors.

The mass build-up of galaxies is expected to proceed from a relatively small number of processes (for a summary, see e.g. Springel et al. 2005b; Silk & Mamon 2012). New stars can form from the gas reservoir of a galaxy, either acquired at birth or replenished from a more continuous accretion process along the galaxy lifespan since formation. Major and minor merging between galaxies is identified in numerous spectacular examples in the local universe (e.g. in the RNGC catalogue, Sulentic & Tifft 1973; Barnes & Hernquist 1992). Merging is efficient at assembling mass, because it produces a significant increase in mass of up to a factor of two for equal mass mergers and for each merging event. Other processes are expected to modulate

⋆ Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, under Large Programmes 070.A–9007, 177.A–0837, and 185.A–0791. Based on observations obtained with MegaPrime/MegaCam, a joint project of the CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii.

the total mass gains. AGN and SNe feedback have been proposed as mechanisms capable of quenching star formation, as well as of supporting winds that drive some mass fraction into the inter-galactic medium (IGM), hence reducing the increase in stellar mass from in-situ star formation (Silk 1997; Murray et al. 2005; Cattaneo et al. 2009). The environment of galaxies is also expected to affect mass growth, with interactions between galaxies and the dense intra-cluster medium, such as harassment or stripping, thereby rapidly removing a significant part of the gas content of a galaxy (e.g. Moore et al. 1996). These processes are expected to ultimately combine along cosmic time to produce the mass distribution observed in the well-defined Hubble sequence of galaxy types in the nearby universe.

In recent years, cold gas accretion fuelling star formation has received focussed attention, following numerical simulations (Kereš et al. 2005). In this picture, cold gas flows along the filaments of the cosmic web into the main body of a galaxy to support vigorous star formation. This mechanism has been proposed as the main mode of galaxy assembly (Dekel et al. 2009) and is often cited in recent literature as the preferred scenario for galaxy assembly (e.g. Kereš et al. 2009; Dijkstra & Loeb 2009; Bouché et al. 2010; Di Matteo et al. 2012). However, as of today, only limited and indirect observational evidence exists in support of this picture (Cresci et al. 2010; Kacprzak et al. 2012), with detailed observational investigations failing to identify direct supporting evidence (Steidel et al. 2010) until the recent claim for accretion detected at $z \sim 2$ (Bouché et al. 2013), which demonstrates the difficulty directly identifying the accretion process at work.

The merging of galaxies is another key process that contributes to galaxy assembly. The hierarchical growth of dark matter haloes is a key prediction of the ΛCDM model for galaxy formation (Davis et al. 1985; Springel et al. 2005a; Hopkins et al. 2006). In this picture, the merging of dark matter (DM) haloes would not only lead to an increase in the DM halo masses, but also naturally lead to the merging of the galaxies associated with each of the haloes (Kauffmann et al. 1993). While galaxy assembly seems to produce the more massive galaxies early in a seemingly anti-hierarchical downsizing pattern (De Lucia et al. 2006), it is nonetheless expected that merging of galaxies would continue to occur as DM haloes continue to merge along cosmic time. Since the dynamical time scale for haloes to merge is of the order of 0.5−1 Gyr (Kitzbichler & White 2008; Lotz et al. 2010), it is expected that a massive halo today, as identified in DM halo merger trees from numerical N-body simulations, will have experienced several mergers since its formation. Observational evidence of merging activity at different epochs can therefore shed light on the contribution of this process in assembling mass in galaxies.

Evidence of merging is direct and well documented. Mergers have been identified since early days of photographic galaxy atlases and classified alongside the Hubble sequence of morphological types (e.g. the RC3 catalogue, de Vaucouleurs et al. 1991). The merging process of two disc galaxies has been proposed as one of the mechanisms for producing early-type bulge-dominated galaxies, supported from simulations (e.g. Mihos & Hernquist 1996; Bournaud et al. 2005). Although this picture may be too simplistic (Bournaud et al. 2011), it is however clear that mergers do occur in the low redshift universe and that major mergers may lead to large modifications of the physical properties of the galaxies involved. Mergers are identified either a posteriori from morphological signatures like wisps, tails, or irregular shapes produced by on-going or post-merger dynamics, or a priori from the identification of pairs of physically bound galaxies destined to coalesce. Going to high redshifts and using pairs in early merging stages rather than post-merger remnants remains the most robust way to derive a merger fraction (e.g. Le Fèvre et al. 2000). This is because major merging pairs are easier to identify at these redshifts than the post-merger morphological signatures, which are of low surface brightness and may escape detection. The pair fraction can be transformed into a merger rate per volume or per galaxy (Patton et al. 2000; Kitzbichler & White 2008; Lin et al. 2008; de Ravel et al. 2009; López-Sanjuan et al. 2011, 2013), using the dynamical time scale for a pair of galaxies with a given mass ratio and projected physical separation (Kitzbichler & White 2008; Lotz et al. 2010). The integrated merging rate over the lifespan of a galaxy since formation then gives the total amount of stellar mass assembled from the merger process.

The secure identification of pairs and measurement of pair fractions at different redshifts is therefore an important observational measurement to perform. Spectroscopic redshift measurements of both members of the pair are required to eliminate the risk of background or foreground contamination along the line of sight and confirm that the pair is physically bound. At $z \sim 1$, pairs have been observed from deep galaxy spectroscopic surveys with the confirmation of the pair nature with spectroscopic redshifts (e.g. Lin et al. 2008; de Ravel et al. 2009). From the VIMOS VLT Deep Survey (VVDS), de Ravel et al. (2009) finds that the merger fraction is higher by a factor ∼3 at $z = 1$ than in the local Universe and further shows that the merger rate and its evolution depend significantly on the stellar mass (luminosity) of the galaxy population. At redshifts beyond $z \sim 1$, only a few direct identifications of pairs and measurements of the merger fraction and merger rate exist. López-Sanjuan et al. (2013) report a high pair fraction of ∼20−22% at $1 < z < 1.8$ from 3D spectroscopy measurements in MASSIV (Mass Assembly Survey with Sinfoni in VVDS, Contini et al. 2012). At higher, redshifts Conselice et al. (2003) has used the CAS (concentration, asymmetry clumpiness) methodology, which relies on image shapes and the expected signature of on-going or past mergers, to perform a measurement of the merger fraction up to $z \sim 3$. Cooke et al. (2010) provides spectroscopic identification of five pairs of galaxies in their LBG sample at $z \sim 3$, claiming that merging is triggering a significant part of the Lyα emission. The number of confirmed pairs is therefore small beyond $z \sim 2$, and larger samples have to be assembled to enable measurements of the pair fraction and merger rate, which is accurate to a few percent.

Here we present a sample of galaxy pairs identified at $z > 1.8$ in the VIMOS Ultra Deep Survey (VUDS) and VVDS. The VUDS is an on-going survey (Le Fèvre et al. 2014) with ultra-deep spectroscopy obtained with VIMOS on the VLT targeting galaxies with $z > 2$ in three well studied fields: the COSMOS, ECDFS, and VVDS-02h (XMM-LSS/CFHTLS-D1). The VVDS has been extensively discussed elsewhere (Le Fèvre et al. 2005). We are using the "Final Data Release" of this survey as described in Le Fèvre et al. (2013b). These spectroscopic redshift survey samples are searched to identify a sample of real physical pairs with redshifts $1.8 < z < 4$, based on the observed projected separation r_p and velocity difference Δv. We discuss the derived pair properties using all available spectroscopy, as well as visible and near-IR imaging and photometric data. The VUDS and VVDS spectroscopic redshift surveys are described in Sect. 2. The methodology for identifying pairs as well as confirming that they are at close physical separation rather than a random projection along the line of sight is presented in Sect. 3. We then examine the pair properties in Sect. 4. The pair fraction

and merging rate at $1.8 < z < 4$ are discussed in Sect. 5 and we conclude in Sect. 6.

Throughout this work, we adopt a cosmology with $H_0 = 100\,h$ km s^{-1} Mpc^{-1}, $h = 0.7$, $\Omega_{0,\Lambda} = 0.73$, and $\Omega_{0,\mathrm{m}} = 0.27$. All magnitudes are given in the AB system.

2. Spectroscopic observations and parent sample

To find pairs, we have explored the VUDS, VVDS-Deep, and VVDS-UltraDeep surveys, providing a sample of galaxies with spectroscopic redshifts measured with VIMOS on the ESO-VLT (Le Fèvre et al. 2003). The VIMOS spectra were obtained with 4.5 h of integration for the VVDS-Deep survey, covering $5500 \leq \lambda \leq 9350$ Å, and 16 h and 14 h integrations in each of the LRBLUE and LRRED grism settings for the VVDS-UltraDeep and VUDS surveys, respectively, covering a combined wavelength range $3600 \leq \lambda \leq 9350$ Å.

The VVDS-deep (Le Fèvre et al. 2005, 2013b) and VVDS-UltraDeep (Le Fèvre et al. 2013b) surveys are based on i-band magnitude selection with $17.5 \leq i_{AB} \leq 24$ and $23 \leq i_{AB} \leq 24.75$, each covering up to $z \sim 6$ (Le Fèvre et al. 2013a). These two VVDS surveys are located in the VVDS-02h field centred at $\alpha_{2000} = 02\mathrm{h}26\mathrm{m}00\mathrm{s}$ and $\delta_{2000} = -04\,\mathrm{deg}\,30'00''$.

The VUDS is an on-going spectroscopic survey also using VIMOS, targeting $z > 2$ galaxies in three fields: COSMOS, ECDFS, and VVDS-02h. The baseline target selection for spectroscopy is using photometric redshifts $z_{\mathrm{phot}} > 2.4$ measured using all the photometry available in the survey fields that use the Le Phare code (Ilbert et al. 2006). The photometric redshift accuracy obtained from the multi-wavelength data and calibrated on existing spectroscopic redshifts is $\sigma_{\delta z/(1+z)} \simeq 0.01$ for magnitudes $i_{AB} < 25$ in the COSMOS field (see e.g. Ilbert et al. 2013). There is a slight degradation by a factor <2 on the photometric redshift accuracy in the other two fields because of there are fewer photometric bands observed. In addition, we supplemented the z_{phot} selection by several colour-selection criteria, adding those galaxies likely to be in this redshift range, but not selected from the primary z_{phot} selection. Allowing for errors in z_{phot}, this selection provides a sample with $2 \lesssim z \lesssim 6$, as described in Le Fèvre et al. (2014).

All VIMOS data are processed with the VIPGI package (Scodeggio et al. 2005). Following automated measurements, each galaxy is examined visually and independently by two people, each assigning a spectroscopic redshift. These measurements are compared before assigning the final redshift measurement. A reliability flag is assigned to each redshift measurement representing the probability for the redshift to be right. As consistently shown from the VVDS (Le Fèvre et al. 2005), zCOSMOS (Lilly et al. 2007), and VIPERS (Guzzo et al. 2014) surveys, the reliability of flags reflect the statistical process of redshift assignment between independent observers and does not depend on the survey type or its intrinsic quality, with flags 1, 2, 3, 4, and 9 having a probability of being right of ~50, 87, 98, 100, 90%, respectively. (Flags 1x with x = 1, 2, 3, 4, 9 indicate a broad line AGN; flag 2x are objects falling serendipitously in a slit next to a main target; and both have probability distributions similar to the main flag categories.)

Redshifts of each galaxy are measured using the EZ engine based on cross-correlating observed spectra to reference templates (Garilli et al. 2010). It is noteworthy that the redshift accuracy in spectra velocity measurements using cross-correlation is better than the instrument resolution, because, simply put, the centroid of a line can be measured to better accuracy than the size of the resolution element, and combining a number of lines as done in the cross-correlation can improve the velocity measurement further. The accuracy of redshift measurements and associated errors have been extensively described in Le Fèvre et al. (2013b) using about 1000 independent measurements from independent observations of the same galaxies, while the measurement error for each galaxy is ~200 km s^{-1} and the absolute velocity accuracy of the whole sample is about 40 km s^{-1}.

In addition to the VIMOS spectroscopic data, a large set of imaging data is available in the three fields covered by our pair search. The COSMOS field (Scoville et al. 2007) has a full coverage with the HST/ACS F814W filter (Koekemoer et al. 2007) and includes, among other data, ugriz photometry from Subaru (Taniguchi et al. 2007) and, more recently, *YJHK* photometry from the UltraVista survey (McCracken et al. 2012). Spectroscopic redshifts from the zCOSMOS survey are also available (Lilly et al. 2007). The ECDFS is covered by the MUSYC survey in *UBVRIz'* (Gawiser et al. 2006) and partly by the CANDELS survey with the ACS and WFC3 on HST (Koekemoer et al. 2011). The VVDS-02h field has deep CFH12K *BVRI* photometry (Le Fèvre et al. 2004), and even deeper CFHTLS ugriz photometry (e.g. Ilbert et al. 2006), as well as *JHK* photometry from the deep survey with CFHT-WIRCAM (WIRDS Bielby et al. 2012).

At redshifts $z > 2$, most of the galaxies in VVDS and VUDS are star-forming galaxies with star-formation rates from ~2 to 100 $M_\odot$, and stellar masses from $\sim 5 \times 10^9$ to 10^{11} $M_\odot$. As discussed in Le Fèvre et al. (2013b), only the reddest galaxies are missing from the i-band selection of the VVDS, corresponding to heavily obscured or passive early-type galaxies. Heavily obscured objects have star formation rates comparable to the UV-selected population (Lemaux et al. 2013) so that their merging properties are expected to be similar to our UV-based selection. Since passive galaxies represent less than a few percent of the global galaxy population at $z \sim 3$ as identified from the mass function (e.g. Ilbert et al. 2013), our sample is representative of the moderately star-forming galaxy population, which is largely dominant at $z \sim 3$.

3. Pair identification

Pairs have been identified using the projected transverse separation r_{p} and the velocity difference $\Delta v = c(v_1 - v_2)/(1 - v_1 v_2)$ where v_x, the normalized velocity of the xth galaxy, is given as $v_x = [(1 + z_x)^2 - 1]/[(1 + z_x)^2 + 1]$. The survey samples have first been scanned for separation $r_{\mathrm{p}} \leq 25\ h^{-1}$ kpc and $\Delta v \leq 500$ km s^{-1}. We chose these separations because it is expected that pairs would merge in about 1 Gyr (e.g. Kitzbichler & White 2008), meaning that a pair observed at $z \sim 3$ would have merged by the peak in star formation activity at $z \sim 1.5$–2 (e.g. Cucciati et al. 2012).

Given the limitation of our ground-based seeing observations, we are not able to identify pairs superimposed along the line of sight or separated spatially by less than 1 arcsec, the average image quality (FWHM) of the imaging data, corresponding to $5\ h^{-1}$ kpc at $z \sim 3$. In addition, pairs with larger r_{p} or Δv could also merge, albeit on a longer time scale, and with a lower probability (Kitzbichler & White 2008).

We used all objects with reliability flags 2 to 9 (galaxies) and 12 to 19 (broad-line AGN) for the primary targets, including a total of 1111 galaxies, and an additional 811 galaxies with flags 1 to 9 and 11 to 19, as well as 21 to 29 (objects falling serendipitously in the slit), for the companions.

Table 1. Pair properties.

Field	Pair galaxies ID numbers	Redshift z_1	Stellar masses[1] $10^{10}\ M_\odot$	Mass ratio[2]	Flux ratio[3]	r_p[4] h^{-1} kpc	Δv[5] km s^{-1}	T_{merg}[6] Gyr	$z_{assembly}$[7]
COSMOS	511001467 a&b	2.0970	0.3/0.15	2.0	1.9 ± 0.2	18.4 ± 0.3	87	2.3	1.1
	510788270 a&b	2.9629	0.5/0.7	1.4	1.7 ± 0.3	15.4 ± 0.2	38	1.5	1.8
	510175610/510778438	3.0939	0.8/5.0	6.3	4.2 ± 0.2	9.4 ± 0.2	161	0.7	2.3
VVDS-02h	520452183/520450423	1.8370	5.6/3.0	1.9	2.2 ± 0.1	22.3 ± 1.2	390	1.1	1.3
	520478238/520478087	2.2460	2.0/1.0	2.0	1.8 ± 0.4	25.3 ± 1.2	28	1.8	1.3
	910260902/910261083	2.3594	3.4/7.9	2.3	2.0 ± 0.1	12.3 ± 1.1	54	0.6	1.9
	910302317/910302929	1.8171	1.1/1.8	1.6	1.4 ± 0.1	8.7 ± 1.2	181	0.6	1.5
ECDFS	530034527 a&b	3.6500	1.2/0.9	1.3	1.5 ± 0.1	6.6 ± 0.2	0	0.5	2.9
	530050663 a&b	2.3250	0.4/0.2	2.0	1.9 ± 0.2	5.5 ± 0.3	54	0.6	1.9
	530042814/530042840	2.9903	0.12/0.1	1.2	1.3 ± 0.1	6.0 ± 0.2	278	1.0	2.1
	530046916 a&b	2.8684	1.4/0.6	2.3	2.5 ± 0.1	6.5 ± 0.2	279	0.5	2.4
	530036900 a&b	3.3300	1.3/1.1	1.2	1.5 ± 0.1	7.0 ± 0.2	76	0.6	2.6

Notes. [(1)] Stellar mass estimates from SED fitting; [(2)] stellar mass ratio between the two galaxies in the pair, errors on the mass ratio have been computed from the error on the ratio of the H or K band flux measurements (see text); [(3)] flux ratio as measured on H or K-band; [(4)] r_p: transverse separation between the two galaxies in the pair, errors in r_p have been conservatively estimated taking a one pixel error on the difference in the centroid measurement of each galaxy in the pair; [(5)] Δv: velocity separation along the line of sight; errors in Δv are estimated to be <100 km s^{-1} when the two redshifts are from the same slit (10 pairs), and ∼200 km s^{-1} when redshifts are from two different slits (2 pairs); [(6)] T_{merg}: time scale for the pair to merge, using the Kitzbichler & White (2008) prescription; from errors in mass ratio, r_p and Δv errors on T_{merg} are ∼10−20%; [(7)] $z_{assembly}$: redshift by which the two galaxies will have merged, obtained combining the observed redshift and the Δz corresponding to the merger time scale.

In looking for pairs, there is a possibility that the two identified objects could be two giant HII regions of the same galaxy, which would then be prominent in the UV rest frame probed by the i-band, but which would appear as a single galaxy in the H or K band, which probes red-wards of the D4000 or Balmer breaks at wavelengths that are less sensitive to contamination by younger stellar populations. In contrast, the persistence of separate morphological components in the redder bands is an indication of the existence of two galaxies even if they are embedded in a diffuse light background. We therefore examined the i-band and H/K-band images of the pair candidates selected from their separation. As described in the next section, all of our pairs are well separated, and both galaxies are seen from bluer to redder bands. We are therefore confident that we are dealing with true physical pairs rather than double HII regions of the same galaxy.

In this process, we have identified 12 pairs with redshifts from $z = 1.82$ to $z = 3.65$. All primary and companion galaxies have flags ≥ 2 except one companion galaxy (ID: 520478087) with a lower reliability flag 1. Our pair sample therefore benefits from reliable spectroscopic identification. The pair properties are discussed in the next section.

4. Pair properties

From the multi-wavelength dataset, we have derived the main properties of the galaxies in the identified pairs. An important parameter for this study is the stellar mass ratio used to separate "major-merging pairs", with a mass ratio between the two galaxies of $1 \leq M_1/M_2 \leq 4$, from "minor-merging pairs", for which $M_1/M_2 > 4$. To measure this ratio, we used two different approaches and verified that they provide the same sample of major-merging pairs. First, the stellar mass of each galaxy of a pair has been derived from SED fitting of the available photometric data at the measured redshift (see e.g. Ilbert et al. 2010) and range from 10^9 to $10^{11}\ M_\odot$ (Table 1). While this method is known to lead to absolute uncertainties of up to a factor 2 (e.g. Bolzonella et al. 2010) depending on the number of bands used or on the IMF, the relative comparisons of masses derived from this method are robust when derived from the same set of SED models, since it is mainly the mass scaling that changes (e.g. Ilbert et al. 2013).

Since our major-merger sample is selected based on a mass ratio, we investigated how our sample would be affected by systematic errors on the mass ratio. To verify the robustness of our major pair sample to the SED-derived mass ratio, we computed the flux ratio of each pair using the K-band (or H-band when not available) flux of each component of a pair. The luminosity of a galaxy above $\lambda_{rest} \gtrsim 4000$ Å can be considered as a rough proxy for stellar mass because it is the most sensitive to the older stellar populations accumulated over the life of a galaxy (Bruzual & Charlot 2003), and therefore the flux ratios should be close to a mass ratio. The flux ratio are reported in Table 1 and are in excellent agreement with the SED-derived mass ratio. Given the agreement between these two methods, we used the error on the K-band (or H-band) flux ratio as a proxy for mass ratio errors, which is easier to control than the SED-derived mass ratio error, with typical errors ranging from 0.1 to 0.4 for the faintest pairs. The distribution of mass ratio reported in Table 1 indicates that all pairs have a mass ratio $M_1/M_2 \leq 2.5$, which is consistent with their flux ratio and well above our $M_1/M_2 \leq 4$ limit. Therefore given the 1σ uncertainties on flux ratio estimates reported in Table 1 it is unlikely that any of the major merger pairs are misclassified minor mergers. We conclude from this analysis that our major-merger pair sample is not affected by stellar mass ratio uncertainties. Because our primary galaxy sample is to first order magnitude-selected and galaxies are star-forming with similar SEDs, the magnitude limit ranging from $i_{AB} = 24$ to 25 corresponds to a stellar-mass limit in the range $M_* = 10^9$ to $5 \times 10^9\ M_\odot$ depending on redshift over the redshift range considered.

Taking the mass ratio between the two galaxies in a pair into account, along with the separations Δv and r_p, we used the formalism of Kitzbichler & White (2008, formula 10) to compute the merging time scale for each pair T_{merg}, as described in

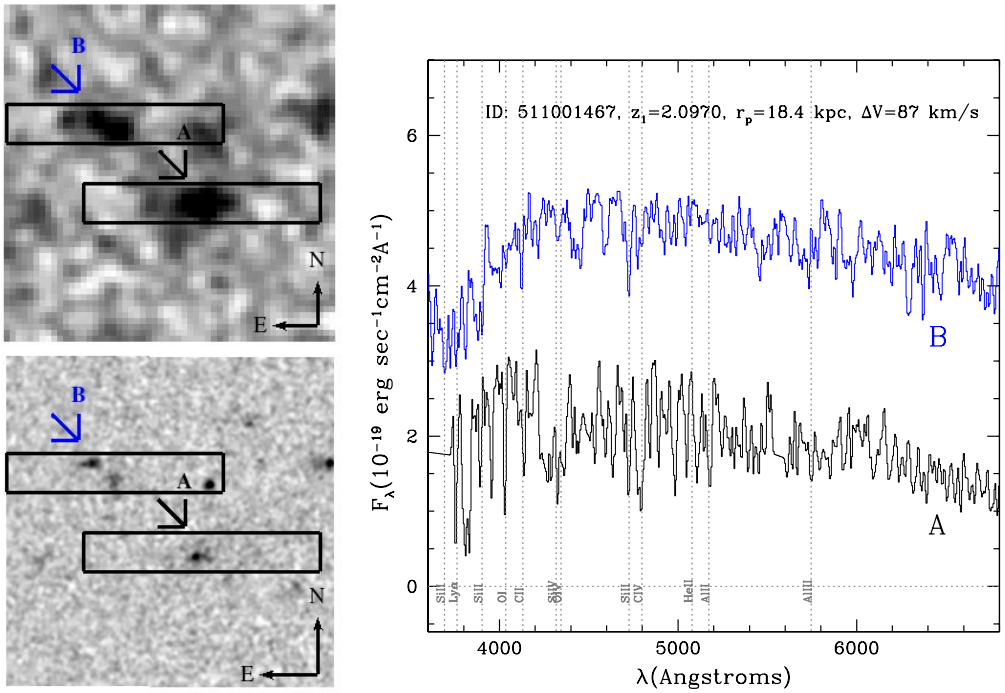


Fig. 1. Pair COSMOS-511001467 A/B: 5″ × 5″ sum of UltraVista YJHK images (*top-left*) and HST F814W image (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slits is shown by the rectangles. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

de Ravel et al. (2009). These time scales are compatible with the Lotz et al. (2010) estimates, as discussed in López-Sanjuan et al. (2011). We list the main pair properties in Table 1, including $T_{\rm merg}$ and the redshift $z_{\rm assembly}$ by which the pair would have merged into a single galaxy, taking the redshift of the pair and $T_{\rm merg}$ into account. Images and spectra of each pair are presented in Figs. 1 to 12. We describe the properties of each pair below.

COSMOS-511001467 A/B (Fig. 1): redshifts of the two galaxies in this pair have been measured from two different VUDS observations at $z_1 = 2.0970$ and $z_2 = 2.0961$, for a velocity difference of $\Delta v = 87$ km s^{-1}. The two galaxies are separated by 18.4 h^{-1} kpc, and are easily visible both in the HST/ACS F814W image and in the UltraVista near-infrared images. The most massive galaxy has a stellar mass of 0.3×10^{10} $M_\odot$, and the stellar mass ratio between the two galaxies is estimated to be $M_1/M_2 = 2$, hence a major merger. The south-west galaxy shows a compact nucleus surrounded by a faint diffuse component, while the north-eastern galaxy is seen as two components in both HST/ACS F814W and UltraVista images. The spectra of the two galaxies have a UV flux rising to the blue up to Lyα, typical of star-forming galaxies at these redshifts. Given the physical separation and mass difference, and following the prescription of Kitzbichler & White (2008), this pair is expected to merge within 2.3 Gyr.

COSMOS-510788270 A/B (Fig. 2): two galaxies separated by 15.4 h^{-1} kpc are at a redshift $z = 2.9629$ with a velocity difference $\Delta v = 38$ km s^{-1}, measured from two different VUDS observations. The most massive galaxy has a mass 0.7×10^{10} $M_\odot$, and the mass ratio between the two galaxies is estimated to be $M_1/M_2 = 1.5$, so a major merger. One of the spectra shows Lyα in emission with $EW(\mathrm{Ly}\alpha)_{\rm rest} = 15$ Å, while Lyα is observed in absorption for the other galaxy. Images are barely resolved at the HST/ACS resolution, showing a slightly extended and elliptical light distribution. This pair is expected to merge within 1.5 Gyr.

COSMOS-510175610/510778438 (Fig. 3): this pair at a redshift $z_1 = 3.0939$ is separated by 9.4 h^{-1} kpc and $\Delta v = 161$ km s^{-1}, measured in the same slit from VUDS observations. The brightest and most massive galaxy has a compact, although irregular, morphology, while the companion to the west, which is confirmed at the same redshift, is more diffuse and has a low surface brightness. Another companion is visible about 3 h^{-1}kpc to the north-east, and although its photometry is compatible with the redshift of the pair, no spectroscopic redshift information is available to confirm that it is physically linked to this system. Both spectra have been obtained from the VUDS survey. One of the VIMOS spectra shows a weak Lyα in emission, while the other is purely in absorption. The mass of the brightest galaxy is 5×10^{10} $M_\odot$, and the mass ratio is $M_1/M_2 = 6.3$, considered to be a minor merger. This pair is expected to merge within 0.7 Gyr.

VVDS-02h-520452183/520450423 (Fig. 4): the spectra of these two galaxies obtained from the same VVDS slit show a broad-line AGN with $z = 1.8370$ and an absorption-line galaxy with $z = 1.8333$ for a velocity difference $\Delta v = 391$ km s^{-1}. This pair is separated by 22.3 h^{-1} kpc, and the estimated mass ratio is $M_1/M_2 = 1.9$, with the most massive galaxy, the AGN host, having a stellar mass 5.6×10^{10} $M_\odot$. This estimate is, however, quite uncertain given the presence of the AGN. Another object is observed in between the two main galaxies, but no redshift information is available. The AGN host is compact, barely resolved at the seeing of the best CFHTLS image ($FWHM = 0.6$ arcsec), while the companion galaxy is slightly elongated and irregular in shape. Given the physical separation and mass difference, this pair is expected to merge within 1.1 Gyr, although this estimate is affected by the uncertain mass estimate of the AGN host.

VVDS-02h-520478238/520478087 (Fig. 5): the two galaxies observed spectroscopically in the same VUDS slit have $z_1 = 2.2460$, $z_2 = 2.2463$ for a velocity difference $\Delta v = 28$ km s^{-1}, and are separated by $r_{\rm p} = 25$ h^{-1} kpc. They are part of a group

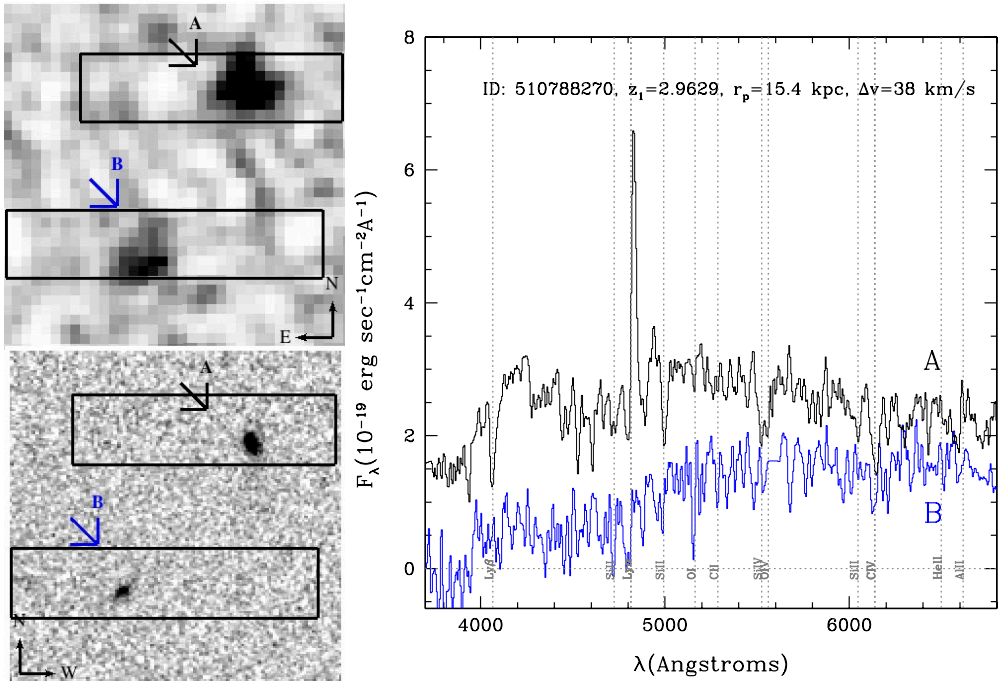


Fig. 2. Pair COSMOS-510788270 A/B: 5″ × 5″ sum of UltraVista YJHK images (*top-left*) and HST F814W image (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slits is shown by the rectangles. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

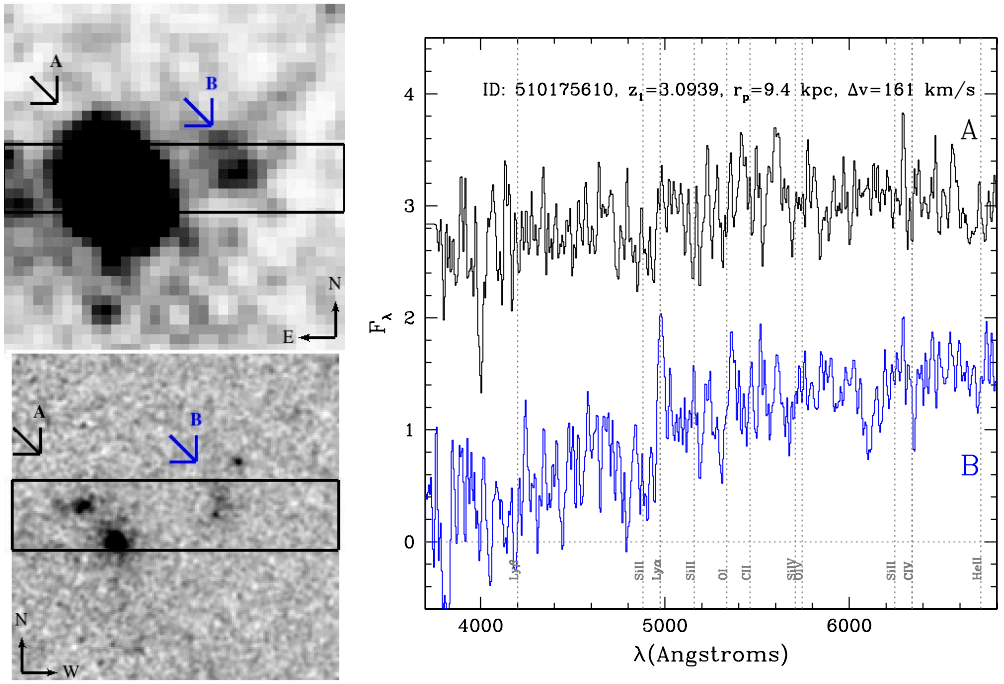


Fig. 3. Pair COSMOS-510175610/510778438: 5″ × 5″sum of UltraVista YJHK images (*top-left*) and the HST F814W image (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

of four galaxies within 5 arcsec (30 h^{-1} kpc), with photometric redshifts compatible with the redshift of the pair. The eastern component is the most massive with $M_\star = 2.0 \times 10^{10}\,M_\odot$. The mass ratio, $M_1/M_2 = 2$, indicates a major merger and results in an expected time to merge within 1.8 Gyr.

VVDS-02h-910260902/910261083 (Fig. 6): the two galaxies, spectroscopically observed in the same VVDS slit with $z_1 = 2.3594$ and $z_2 = 2.3588$, are identified both in the CFHTLS *i*-band image and the *K*-band WIRDS image, with an $r_p = 12.3\ h^{-1}$ kpc and $\Delta v = 54$ km s^{-1} separation. The slit was placed on the centroid of the blended CFHTLS image of the two galaxies, but still included a significant fraction of the flux of both galaxies to yield two well separated spectra. The south-east galaxy is the most massive with $7.9 \times 10^{10}\,M_\odot$ and the mass ratio is $M_1/M_2 = 2.4$, the later indicating a major merger. While the massive galaxy shows a symmetric elongated shape,

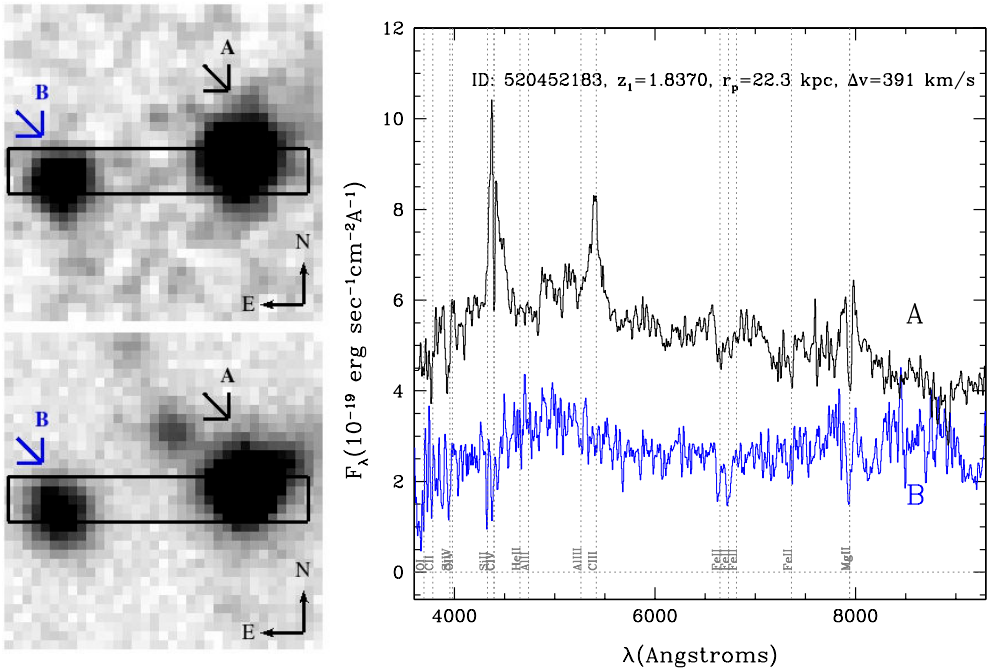


Fig. 4. Pair VVDS-2h-520452183/520450423: $6.5'' \times 6.5''$ composite WIRDS JHKs image (*top-left*) and CFHTLS *i*-band image (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

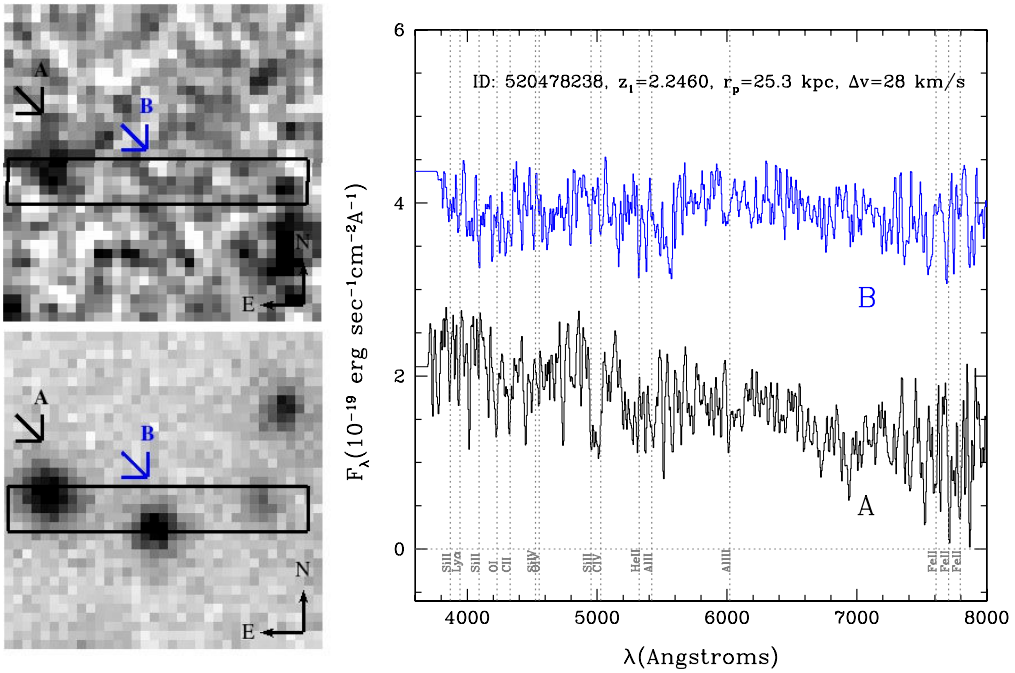


Fig. 5. Pair VVDS-2h-520478238/520478087: $6.5'' \times 6.5''$ composite WIRDS JHKs image (*top-left*) and CFHTLS *i*-band image (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

the north-west galaxy shows an irregular morphology, and the two are connected by a faint bridge. Both galaxies show Lyα in emission, with $EW(\mathrm{Ly}\alpha)_{\mathrm{rest}} = 50$ Å and 60 Å, and integrated *Ly*α line flux $L_{Ly\alpha} = 10^{42}$ and 3×10^{42} erg s^{-1}, respectively, indicating strong star formation at the level of 1–2 $M_\odot$/yr (see e.g. Cassata et al. 2011). Given the observed separation of this pair, it is expected that it will merge within 0.6 Gyr.

VVDS-02h-910302317/910302929 (Fig. 7): the two main galaxies observed in spectroscopy in the same VVDS slit are at $z_1 = 1.8171$ and $z_2 = 1.8154$, respectively, separated by 8.7 h^{-1} kpc along the east-west direction, and by $\Delta v = 181$ km s^{-1}. The brightest/most massive galaxy has a mass 1.8×10^{10} $M_\odot$ and is made of two main components, visible in both the *i*-band and *K*-band images. The mass ratio between

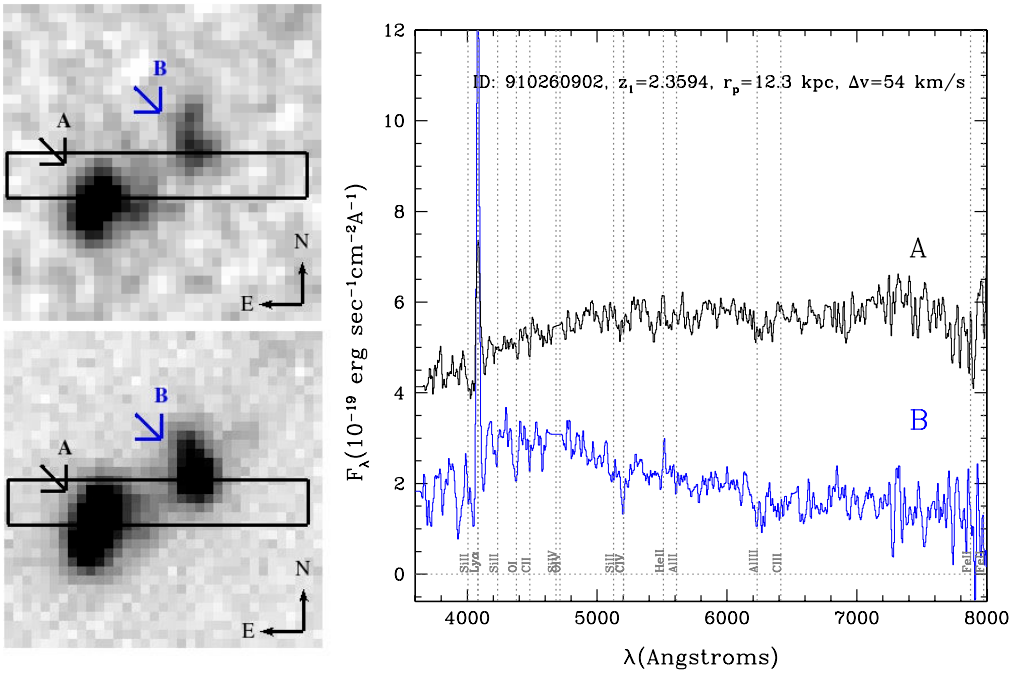


Fig. 6. Pair VVDS-2h-910260902/910261083: 6.5″ × 6.5″ composite WIRDS *JHK*s image (*top-left*) and CFHTLS *i*-band image (*bottom-left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

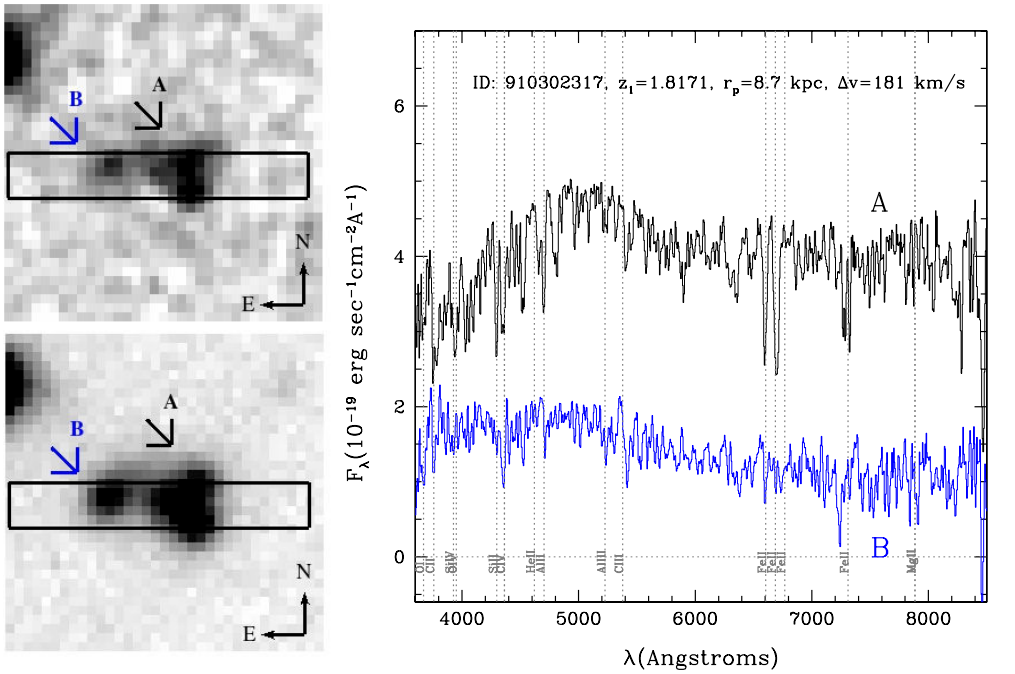


Fig. 7. Pair VVDS-2h-910302317/910302929: 6.5″ × 6.5″ composite WIRDS *JHK*s image (*top-left*) and CFHTLS *i*-band image (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

the two galaxies in the pair is $M_1/M_2 = 1.7$, which indicates a major merger. This pair is expected to merge within 0.6 Gyr.

ECDFS-530034527 (Fig. 8): this is the highest redshift for which we have obtained a spectroscopic confirmation of a physical pair, with both galaxies at the same redshift $z = 3.6500$, separated by 6.8 h^{-1} kpc, as measured from the same VUDS slit. In Fig. 8 another companion is observed to the north-east, but there is no confirmation of its redshift. The stellar mass of the brightest and most massive galaxy is 1.2×10^{10} $M_\odot$, and the mass ratio with the companion is $M_1/M_2 = 1.3$, which indicates a major merger. The two spectra obtained in the VUDS survey show absorption line spectra, with only a weak Lyα emission for the brighter galaxy. This pair is expected to merge within 0.5 Gyr.

ECDFS-530050663 (Fig. 9): this pair is made of two components identified in the *H*-band CANDELS image, with redshifts

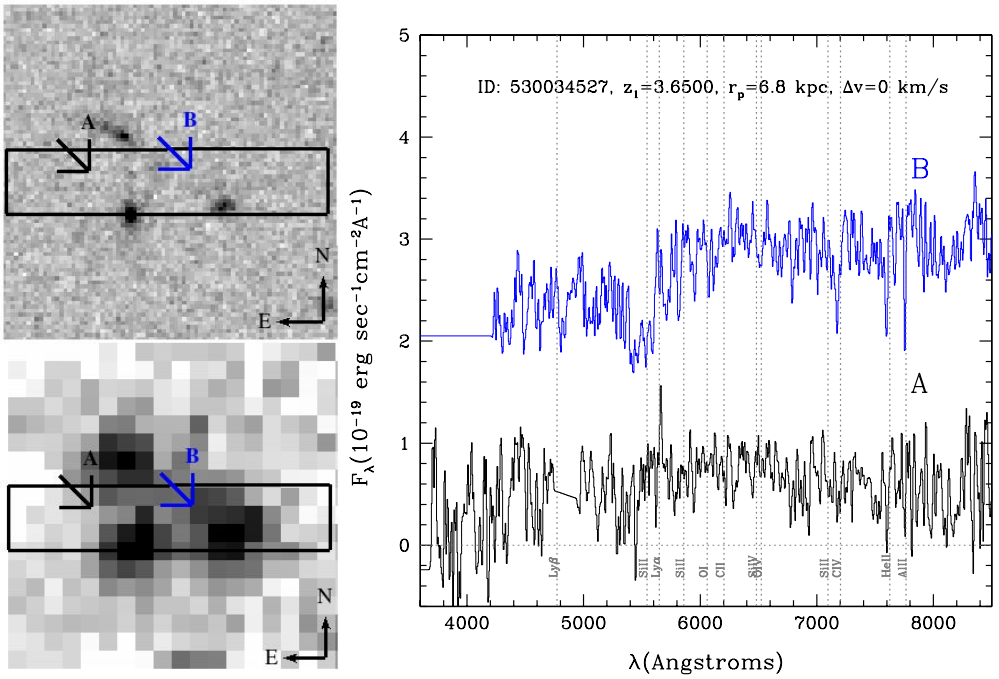


Fig. 8. Pair ECDFS-530034527 A/B: $5'' \times 5''$ HST/WFC3 F160W CANDELS image (*top-left*) and composite *BVR* image from the MUSYC survey MUSYC (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

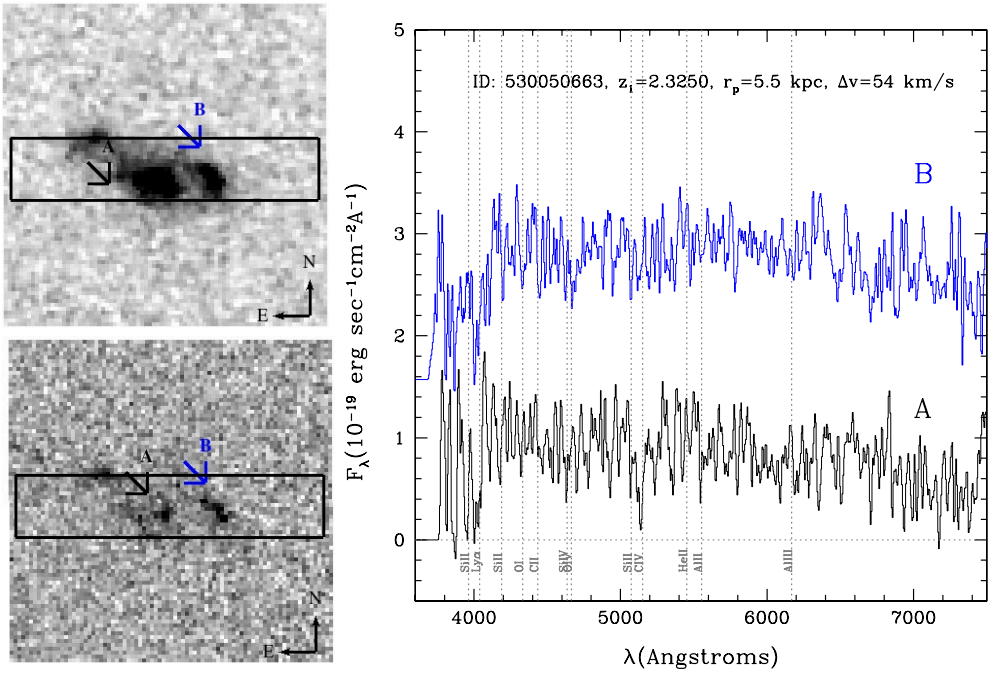


Fig. 9. Pair ECDFS-530050663 A/B: $5'' \times 5''$ HST/WFC3 F160W image (*top-left*) and HST/ACS F850W image from the CANDELS survey (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

$z_1 = 2.3250$ and $z_2 = 2.3244$ measured in the same VUDS slit. The stellar mass of the main component is $0.4 \times 10^{10}\ M_\odot$, and the mass ratio between components is $M_1/M_2 = 2$: a major merger. A third component is observed to the north-east, all three components being embedded in a low surface brightness emission. The pair identification is therefore ambiguous, since this configuration might be indeed a merger at an advanced stage surrounded by tidal debris, or the result of three giant star-forming regions in a single forming galaxy. Under the merger hypothesis, the two main components are expected to merge within 0.6 Gyr.

ECDFS-530042814/530042840 (Fig. 10): the two galaxies are seen very well in the composite *BVR* MUSYC images and the *K*-band image. These galaxies are at a mean redshift $z = 2.9903$ with a velocity difference $\Delta v = 278$ km s^{-1} as measured in the same VUDS slit. These two main galaxies are separated by 6 h^{-1} kpc, and a third component is observed 6.5 h^{-1} kpc to

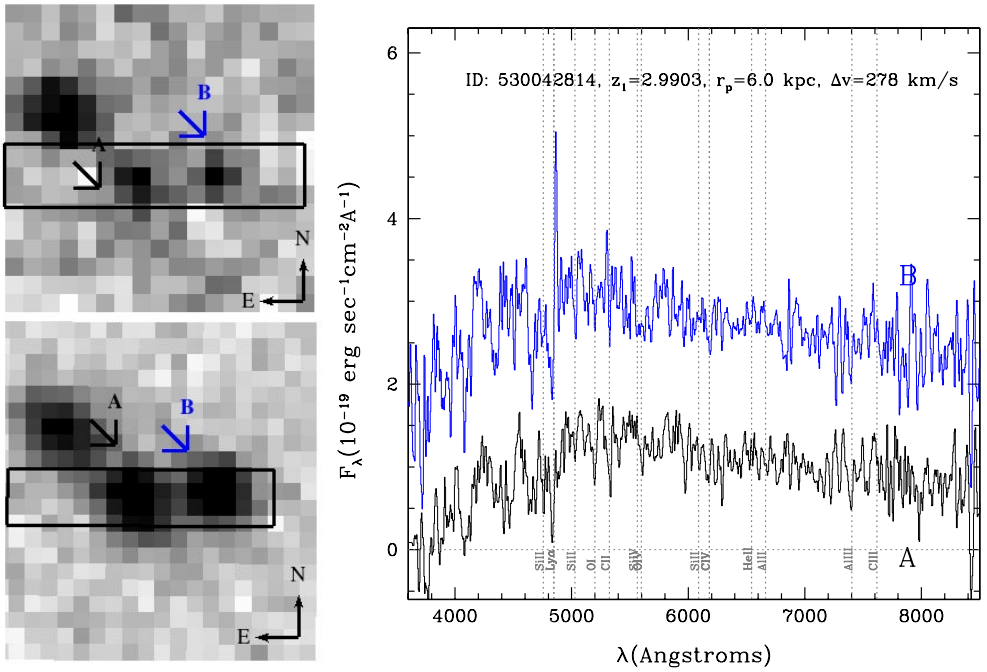


Fig. 10. Pair ECDFS-530042814/530042840: $5'' \times 5''$ composite *JK* image (*top-left*) and composite *BVR* image from the MUSYC survey (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

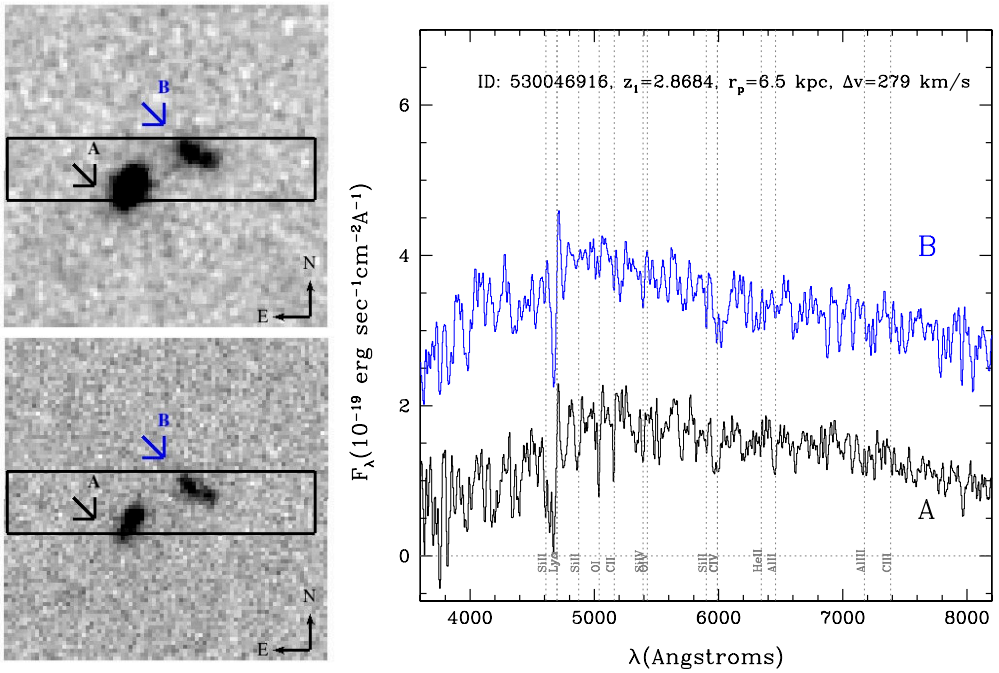


Fig. 11. Pair ECDFS-530046916 A/B: $5'' \times 5''$ HST/WFC3 F160W image (*top-left*) and HST/ACS F850W image from the CANDELS survey (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS spectra for both components in the pair. The spectra have been arbitrarily shifted in flux to avoid overlap.

the north-east. While the main galaxy shows Lyα in absorption, the other galaxy measured with VIMOS shows Lyα in emission, both spectra having a UV slope indicating strong star formation. The mass of the main galaxy is estimated to be $0.12 \times 10^{10}\ M_{\odot}$, and the mass ratio is $M_1/M_2 = 1.2$ almost an equal mass major merger. This pair is expected to merge within 1.0 Gyr.

ECDFS-530046916 (Fig. 11): two close but well separated images of two galaxies are observed in both HST z-band (F850W) and H-band (F160W) images, with a redshift around $z = 2.868$ measured in the same VUDS slit. The brightest/most massive galaxy has a regular morphology compatible with a disc, while the companion galaxy has an irregular shape. The mass of the main galaxy is estimated to be $1.4 \times 10^{10}\ M_{\odot}$, and the mass

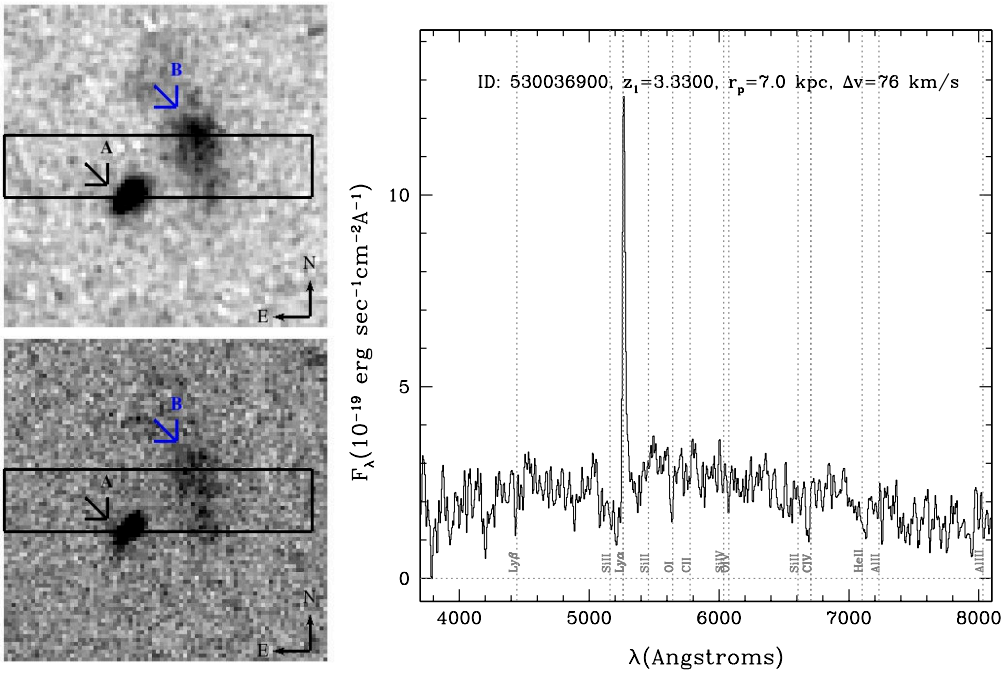


Fig. 12. Pair ECDFS-530036900 A/B: 5″ × 5″ HST/WFC3 F160W image (*top-left*) and HST/ACS F850W image from the CANDELS survey (*bottom left*). The location of the 1 arcsec width VLT/VIMOS slit is shown by the rectangle. The more massive object is labelled as A. *Right panel*: VIMOS VUDS spectrum for component A. The redshift of component B has been measured by Popesso et al. (2009) and not shown in the figure.

ratio is $M_1/M_2 = 2.3$, indicating a major merger. This pair will probably merge within 0.5 Gyr.

ECDFS-530036900/26931 (Fig. 12): two well separated galaxies are observed in the HST z-band (F850W) and H-band (F160W) images from CANDELS, with 7.0 h^{-1} kpc and $\Delta v = 78$ km s^{-1} separation at a redshift of $z_1 = 3.3300$. One galaxy is quite compact, while the companion shows a sharp point-like component surrounded by a nebulous extension. The compact galaxy has been observed by VUDS and shows strong Lyα emission, while the fainter galaxy has been observed as part of the VIMOS-GOODS survey (Popesso et al. 2009). The mass of the main galaxy is estimated to be 1.3×10^{10} $M_\odot$, and the mass ratio is $M_1/M_2 = 1.2$, an almost equal mass major merger. This pair will probably merge in less than 0.6 Gyr.

5. The pair fraction and merging rate at $2 \lesssim z < 4$

From the identified pairs, we derive a first robust measurement of the major merger pair fraction at these redshifts. This requires a complete understanding of the survey selection function, including the target selection, spatial sampling, and spectroscopic success rate. We follow the same method as described in de Ravel et al. (2009, 2011) and López-Sanjuan et al. (2011), as the instrumental set-up using VIMOS is identical. The pairs have been identified from a parent photometric sample used for selecting targets in the VUDS and VVDS. The probability of selecting a galaxy from this parent sample is the target sampling rate (TSR), defined as the ratio TSR = $N_{\rm target}/N_{\rm phot}$ of $N_{\rm target}$ objects targeted in the spectroscopic observations over $N_{\rm phot}$ the number of objects in the parent photometric catalogue. The TSR of the VVDS sample has been described in Le Fèvre et al. (2013b) and varies from 6.5% for the Ultra-Deep sample and up to 29% for the Deep sample. For the VUDS sample, we have computed the TSR using the parent photometric catalogue used in selecting targets for the first set of existing observations. In the redshift range probed here, we find a TSR$_{\rm VUDS}$ = 10%, independent of magnitude. Once targeted for VIMOS observations, an object has an associated probability of having a measured spectroscopic redshift, defined as the spectroscopic success rate (SSR). The SSR is computed as the ratio SSR = $N_{\rm spec}/N_{\rm target}$ where $N_{\rm spec}$ is the number of objects with a spectroscopic redshift measurement. In both VVDS and VUDS, we have been using a flag system related to the probability of the redshift being correct, as described extensively in Le Fèvre et al. (2013b). As the VUDS observational set-up is nearly identical to that of the VVDS, we used the SSR of the VVDS-Ultra Deep sample with an average SSR = 80% (Le Fèvre et al. 2013b). In addition, we need to apply a correction factor accounting for the ground-based seeing-limited observations that prevent us from observing pairs separated by less than one arcsecond on the plane of the sky, the average image quality of the spectroscopic observations. This is computed as the ratio of the number of spectroscopic pairs over the number of photometric pairs in the parent photometric catalogue.

Following de Ravel et al. (2009) and López-Sanjuan et al. (2011), we obtain the pair fraction as $f_{\rm MM} = N_{\rm p}/N_1$, where $N_1 = 1111$ is the number of principal galaxies, and $N_{\rm p}$ is the corrected number of pairs

$$N_{\rm p}/N_1 = \frac{\sum_{k=1}^{N_{\rm p,obs}} w^k_{\rm spec,1} \times w^k_{\rm spec,2} \times w^k_{\rm comp} \times w^k_\theta}{\sum_{i=1}^{N_{\rm p,obs}} w^i_{\rm spec}}$$

assuming $w_{\rm spec}$ approximately constant, this leads to

$$N_{\rm p}/N_1 = \frac{w^2_{\rm spec} \sum_k w^k_\theta \times w^k_{\rm comp}}{w_{\rm spec} \sum_i 1} = \frac{w_{\rm spec} \sum_k w^k_\theta \times w^k_{\rm comp}}{1111}$$

where $w_{\rm spec}$ is computed for each field as $w_{\rm spec} = 1/{\rm TSR} \times 1/{\rm SSR}$, with $w_{\rm spec}$(COSMOS) = 0.124, $w_{\rm spec}$(VVDS) = 0.089 and $w_{\rm spec}$(ECDFS) = 0.081.

Here, $w_{\rm comp}$ is estimated for each pair as $\rho(M_2)/\rho(M_{\rm lim})$ where $\rho(M_2)$ is the number density of galaxies with stellar mass higher than $M_2 = M_1 \times (1/4)$, M_1 is the stellar mass of the more massive galaxy in the pair, 1/4 is the major merger ratio, $\rho(M_{\rm lim})$ is the number density of galaxies with stellar mass higher than the mass limit of the survey at each redshift, and number densities are computed using the mass function from UltraVista (Ilbert et al. 2013). If $M_2 > M_{\rm lim}$, obviously $w_{\rm comp} = 1$.

Then, w_θ is computed for each pair as $a/(N_{zz}/N_{\rm pp})$, where N_{zz} is the number of projected pairs in the spectroscopic catalogue for a given angular separation, $N_{\rm pp}$ is the number of projected pairs in the initial photometric catalogue for the same separation, and a the value of $N_{zz}/N_{\rm pp}$ at large separations. For our survey $w_\theta \sim 1$ for $\theta > 5$ arcsec, and it is lower than 1 at smaller separations when the companion galaxy lies by chance in the slit.

After correcting for the selection function as described above we find $N_{\rm p} = 216$ and so we derive a measurement for the major merger fraction of $f_{\rm MM} = 19.4^{+9}_{-6}\%$ for a mean redshift $z = 2.6$ over $1.8 < z < 4.0$. The pair fraction error is estimated by combining the Poisson error on the pair number in quadrature with the errors in $w_{\rm comp}$ estimated using an uncertainty on the mass limit of a factor 1.5.

Because of the number of confirmed spectroscopic pairs and the well controlled selection function, this is one of the best measurements to date at these high redshifts. This value is comparable to values observed from pair identification at lower redshifts $1 < z < 1.8$ in the MASSIV survey (López-Sanjuan et al. 2013), where they found $f_{\rm MM} = 20.8, 20.1, 22.0\%$ at $z = 1, 1.3, 1.6$, respectively, i.e. similar to our higher redshift value. Measurements have also been obtained from CAS morphological analysis up to $z = 3$ Conselice et al. (2003), who find merger fractions ∼10–40% in the redshift range 2−3 depending on absolute magnitude, with large uncertainties related to photometric redshift estimates and detection of low surface brightness features.

The merger rate can only be derived from the knowledge of the merger time scale of each pair, but this cannot be obtained by direct observation. To derive the merger time scale we use the observed projected spatial and velocity separations, and a 0.2 dex error in mass estimation along with the prescription from Kitzbichler & White (2008) derived from numerical simulations. The merging time scale for each pair is listed in Table 1; with separations from 6 to 25 h^{-1} kpc and mass ratio 1/6 to 1, the average merging time scale of our sample is $\langle T_{\rm merg} \rangle = 1$ Gyr, and the median is 0.7 Gyr.

These merger time scales are used to infer the merger rate, which is the ratio of the merger fraction to the volume probed by the VUDS and VVDS surveys in the redshift range where pairs are identified. We computed the major merger rate as $R_{\rm MM} = f_{\rm MM} \times T^{-1}_{\rm merg}$, where $f_{\rm MM}$ is as computed above and $T_{\rm merg} = 1.12$ Gyr from the average $T_{\rm merg}$ in Table 1. After applying the same weighting scheme as for the rest of the analysis, we find a merger rate $R_{\rm MM} = 0.174^{+0.08}_{-0.05}$ Gyr^{-1}. We further point out that since our pair sample is not strictly mass-selected, the merger fraction and merger rates derived here are averaged over the mass range considered; investigating variations of the merger fraction with stellar mass will require larger samples than discussed in this paper.

6. Discussion and conclusions: the role of major merging at $2 \lesssim z < 4$

We have identified 12 pairs of galaxies with redshifts $1.81 \leq z \leq 3.65$ from the on-going VUDS combined with the VVDS. Both components of the pairs have a confirmed spectroscopic redshift obtained with VIMOS on the VLT and, therefore, comprise a unique sample of true physical pairs at these redshifts. The galaxies in our sample span a wide mass range from $10^{9.1}$ $M_\odot$ to 10^{11} $M_\odot$. The majority of galaxies in our pairs show signs of strong star formation, a common property at these redshifts, as proved by either strong Lyα emission, strong UV continuum, or both, and with one of them showing AGN activity. The mass ratio of the merger is in the range $1 < M_1/M_2 < 6$, with 11 of 12 pairs satisfying a major merger pair criterion $1 < M_1/M_2 < 4$.

We find a pair fraction $f_{\rm MM} = 19.4^{+9}_{-6}\%$ and a merger rate $R_{\rm MM} = 0.174^{+0.08}_{-0.05}$ Gyr^{-1}. This estimate is the first to use a robust sample of pairs with confirmed spectroscopic redshifts at $2 < z < 4$. Previous work used either morphological indicators (Conselice et al. 2003) or photometric pairs (Bluck et al. 2009), which accumulate several sources of uncertainties different from those associated to spectroscopic pairs, but when taking these into account, our results are in broad quantitative agreement with these studies in identifying a high pair fraction or merger rate at these epochs.

With the merger time scales as derived in Sect. 5, it is interesting to note that most of these pairs will have merged before the peak of star formation at $z \simeq 1.5$ (see e.g. Cucciati et al. 2012). The contribution of these major mergers to the mass growth of individual galaxies is substantial. With an average mass ratio of 1.75 for the 11 major merger pairs, the most massive galaxies involved in these mergers will have increased their stellar mass by ∼60% from $z \sim 3$ to $z \sim 1.5$ from the merging process alone.

Our observations therefore provide unambiguous evidence of major merging occurring at $2 \lesssim z < 4$. It is clear that hierarchical assembly, with massive galaxies being built from the merging of less massive ones, is at work at these redshifts and that major merging is contributing to the assembly of mass in galaxies at early times. This mass assembly simply results from the sum of the masses in each galaxy in a merging pair, a simple and effective way to increase mass at each merging event. Minor mergers, with a mass ratio greater than the factor four probed here, are also expected to contribute to this mass growth, but remains unconstrained at these redshifts. An additional increase in stellar mass from star formation triggered by the merging process is also possible, with a range of mass production identified in the literature, ranging from relatively large (e.g. Kocevski et al. 2011) to more limited star bursts (e.g. Mullaney et al. 2012), depending on the duration and strength of the merger induced burst. Merging is therefore a clear path to move low mass galaxies towards the higher end of the mass function, which contributes to the evolution of the stellar mass function (e.g. Ilbert et al. 2013).

These results are to be placed in the context of the currently favoured picture of galaxy assembly, with cold accretion playing a key role in building-up mass in galaxies (e.g. Dekel et al. 2009). As expected in the hierarchical picture of DM halo growth by merging, our results indeed show that galaxies are merging at a high rate and that therefore merging substantially contributes to the build-up of galaxies with high stellar masses before the peak in cosmic star formation activity. A galaxy at $z \sim 1.5$ therefore includes different populations of stars that each may have different ages and chemical enrichment, and may have formed in different environments before integrating their post-merger galaxy. Merging in combination with other mass assembly processes may account for the overall increase with redshift of the stellar mass density of the galaxy

population (Ilbert et al. 2013). The large gas reservoirs identified around high redshift galaxies (Tacconi et al. 2010; Daddi et al. 2010) imply the continuous formation of stars, which could potentially be sustained by new gas brought in from a cold accretion process, although direct observational evidence of accretion remains scarce (Bouché et al. 2013). The total mass growth of a galaxy must therefore come from all these different processes at work in parallel at these epochs. By precisely knowing the contribution of merging to global mass assembly measured e.g. by the growth of the stellar mass density, one would ultimately be able to place upper limits to the total mass growth from other processes, including cold gas accretion.

The integrated contribution of merging processes to the complete history of mass assembly requires knowledge of the evolution of the merger rate since early times (i.e. significantly beyond $z \sim 2$). Building on the sample presented in this paper, the VUDS survey, when complete, will enable a robust measurement of the merging rate out to $z \sim 4$ and an estimate of the total amount of mass assembled by the merging process since the early universe.

Acknowledgements. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of the NRC and CNRS. This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. We thank the ESO staff for their support and efficient execution of the observation programme. This work is supported by funding from the European Research Council Advanced Grant ERC–2010–AdG–268107–EARLY.

References

Barnes, J. E., & Hernquist, L. 1992, ARA&A, 30, 705
Bielby, R., Hudelot, P., McCracken, H. J., et al. 2012, A&A, 545, A23
Bluck, A. F. L., Conselice, C. J., Bouwens, R. J., et al. 2009, MNRAS, 394, L51
Bolzonella, M., Kovač, K., Pozzetti, L., et al. 2010, A&A, 524, A76
Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001
Bouché, N., Murphy, M. T., Kacprzak, G. G., et al. 2013, Science, 341, 50
Bournaud, F., Jog, C. J., & Combes, F. 2005, A&A, 437, 69
Bournaud, F., Chapon, D., Teyssier, R., et al. 2011, ApJ, 730, 4
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Cattaneo, A., Faber, S. M., Binney, J., et al. 2009, Nature, 460, 213
Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183
Contini, T., Garilli, B., Le Fèvre, O., et al. 2012, A&A, 539, A91
Cooke, J., Berrier, J. C., Barton, E. J., Bullock, J. S., & Wolfe, A. M. 2010, MNRAS, 403, 1020
Cresci, G., Mannucci, F., Maiolino, R., et al. 2010, Nature, 467, 811
Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31
Daddi, E., Bournaud, F., Walter, F., et al. 2010, ApJ, 713, 686
Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. 1985, ApJ, 292, 371
De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499
de Ravel, L., Le Fèvre, O., Tresse, L., et al. 2009, A&A, 498, 379
de Ravel, L., Kampczyk, P., Le Fèvre, O., et al. 2011, unpublished [arXiv:1104.5470]
de Vaucouleurs, G., de Vaucouleurs, A., Corwin, H. G., et al. 1991, Third Reference Catalogue of Bright Galaxies, Volume 1–3, XII, 2069, 7 figs (New York: Berlin Heidelberg Springer-Verlag)
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451
Di Matteo, T., Khandai, N., DeGraf, C., et al. 2012, ApJ, 745, L29
Dijkstra, M., & Loeb, A. 2009, MNRAS, 400, 1109
Garilli, B., Fumana, M., Franzetti, P., et al. 2010, PASP, 122, 827
Gawiser, E., van Dokkum, P. G., Herrera, D., et al. 2006, ApJS, 162, 1
Guzzo, L., Scodeggio, M., Garilli, B., et al. 2014, A&A, in press, DOI 10.1051/0004-6361/201321489
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2006, ApJS, 163, 1
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, ApJ, 709, 644
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Kacprzak, G. G., Churchill, C. W., Steidel, C. C., Spitler, L. R., & Holtzman, J. A. 2012, MNRAS, 427, 3029
Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160
Kitzbichler, M. G., & White, S. D. M. 2008, MNRAS, 391, 1489
Kocevski, D. D., Lemaux, B. C., Lubin, L. M., et al. 2011, ApJ, 736, 38
Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Le Fèvre, O., Abraham, R., Lilly, S. J., et al. 2000, MNRAS, 311, 565
Le Fèvre, O., Saisse, M., & Mancini, D. 2003, in SPIE Conf. Ser. 4841, eds. M. Iye, & A. F. M. Moorwood, 1670
Le Fèvre, O., Mellier, Y., McCracken, H. J., et al. 2004, A&A, 417, 839
Le Fèvre, O., Vettolani, G., & Garilli, B. 2005, A&A, 439, 845
Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013a, A&A, submitted [arXiv:1307.6518]
Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013b, A&A, 559, A14
Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2014, A&A, submitted [arXiv:1403.3938]
Lemaux, B. C., Le Floc'h, E., Le Fèvre, O., et al. 2013, A&A, submitted [arXiv:1311.5228]
Lilly, S. J., Le Fèvre, O., Renzini, A., & Zamorani, G. 2007, ApJS, 172, 70
Lin, L., Patton, D. R., Koo, D. C., et al. 2008, ApJ, 681, 232
López-Sanjuan, C., Le Fèvre, O., de Ravel, L., et al. 2011, A&A, 530, A20
López-Sanjuan, C., Le Fèvre, O., Tasca, L. A. M., et al. 2013, A&A, 553, A78
Lotz, J. M., Jonsson, P., Cox, T. J., & Primack, J. R. 2010, MNRAS, 404, 575
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
Mihos, J. C., & Hernquist, L. 1996, ApJ, 464, 641
Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Nature, 379, 613
Mullaney, J. R., Pannella, M., Daddi, E., et al. 2012, MNRAS, 419, 95
Murray, N., Quataert, E., & Thompson, T. A. 2005, ApJ, 618, 569
Patton, D. R., Carlberg, R. G., Marzke, R. O., et al. 2000, ApJ, 536, 153
Popesso, P., Dickinson, M., Nonino, M., et al. 2009, A&A, 494, 443
Scodeggio, M., Franzetti, P., Garilli, B., et al. 2005, PASP, 117, 1284
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Silk, J. 1997, ApJ, 481, 703
Silk, J., & Mamon, G. A. 2012, Res. Astron. Astrophys., 12, 917
Springel, V., Di Matteo, T., & Hernquist, L. 2005a, MNRAS, 361, 776
Springel, V., White, S. D. M., Jenkins, A., et al. 2005b, Nature, 435, 629
Steidel, C. C., Erb, D. K., Shapley, A. E., et al. 2010, ApJ, 717, 289
Sulentic, J. W., & Tifft, W. G. 1973, The revised new general catalogue of non-stellar astronomical objects (Tucson: University of Arizona Press)
Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, Nature, 463, 781
Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9

Chapter 9

Conclusion and Future Perspective

Throughout my thesis, my UltraVista/COSMOS catalogue has provided me with an exceptional resource of ultradeep, wide-area, multiwavelength optical (B, V, r, i, z') and near Infrared data (Y, J, H, K_s) for studying star-forming galaxies at high redshift. Using SED fitting on the multi-wavelength data followed by visual inspection, I was able to study a large sample of galaxies with $4.5 \leq z \leq 6.5$. Relative to previous observations, the UltraVista area is 10 times larger than HST near infrared surveys, and is the deepest ground-based survey. This allows me to place improved constraints on the bright end of the luminosity function, and identify a possible excess of bright galaxies at these redshifts. This allowed me to show the mild evolution of the star formation rate density from $z \sim 6$ to the peak of star formation at $z \sim 2$.

In this thesis, I have contributed new results on the properties of galaxies in the first billion years of cosmic time, joining a number of investigations carried out over the past 20 years. Certainly the next 20 years should be even more exciting. First, the next generation of giant ($>$ 20m) ground-based near-infrared telescopes (shown in Figure 9.1) equipped with sophisticated adaptive-optics systems (TMT, E-ELT, GMT) will also enable detailed near-infrared high-resolution spectroscopy of the most distant galaxies which can take over all the studies done by 8-10m telescope so far. In addition, for telescopes in space, there are also several ambitious projects in progress, such as EUCLID satellite and James Webb Space Telescope (JWST). EUCLID deep field covers 20 deg^2 down to $J \sim 26$ mag. (Laureijs et al. 2011) which can measure bright-end (brighter than 3L*) of LF with $z > 7$ and obtain meaningful clustering measurements of galaxies at such high redshifts. The JWST is equipped with a 6.5m diameter mirror

and next generation NIR cameras and spectrographs which can probe the first galaxies ($z > 10$) and investigate their physical properties directly. All of these improvement of instrumental resources will be completed in next 10 years. In the very near future, there are some new wide-area optical imaging surveys such as Hyper-Suprime CAM (HSC) on Subaru (Takada 2010) which can yield enormous scientific benefits over a broad range of areas and in the infrared, the UltraVista survey will be completed in the next couple of years at a depth one magnitude deeper than the data I used in this thesis. New generation ground-based multi-object near infrared spectrographs including Prime Focus Spectrograph (PFS) on Subaru (Takada et al., 2010), KMOS on the VLT (Sharples et al. 2006), and MOSFIRE on Keck (McLean et al. 2008) are also devoted on follow-up studies of brighter $z > 7$ galaxies. Optimistically speaking, some unsolved fundamental questions about galaxies will be tackled or improved by observational constraints by these next generation equipments. The study presented in this manuscript can be considered as a precursor of the next generation studies which will be conducted with these new powerful facilities.

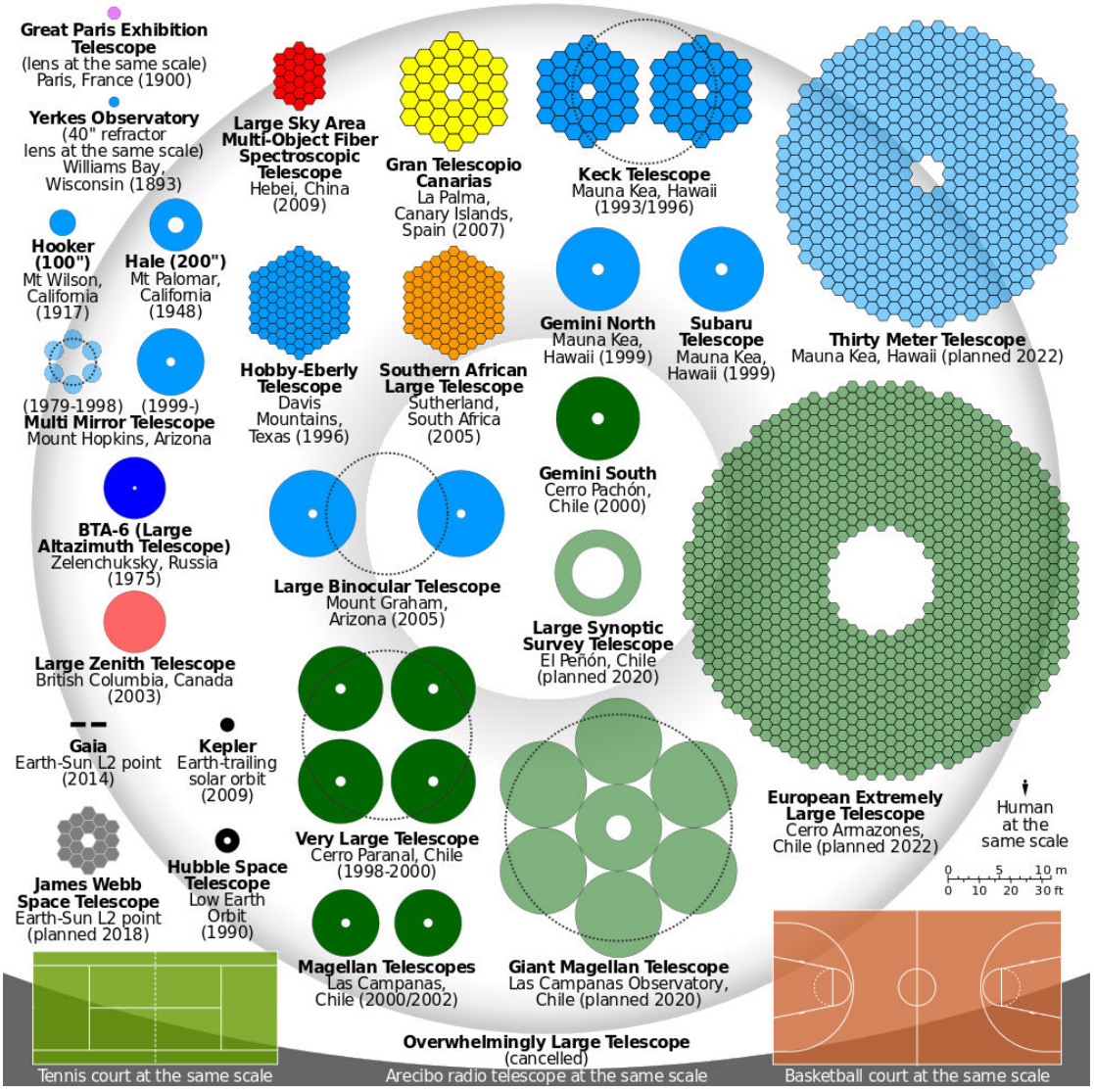


FIGURE 9.1: A comparison of primary mirror sizes. Note how the planned TMT and E-ELT are much larger than anything else in existence. (Credit:http://www.webastro.net/)

Appendix A

Images and SED-fittings of several reliable candidates

Here I randomly select 10 images associate their plots of SED-fitting of my several credible high redshift ($4.5 < z < 6.5$) candidates (Flag=1) in this appendix.

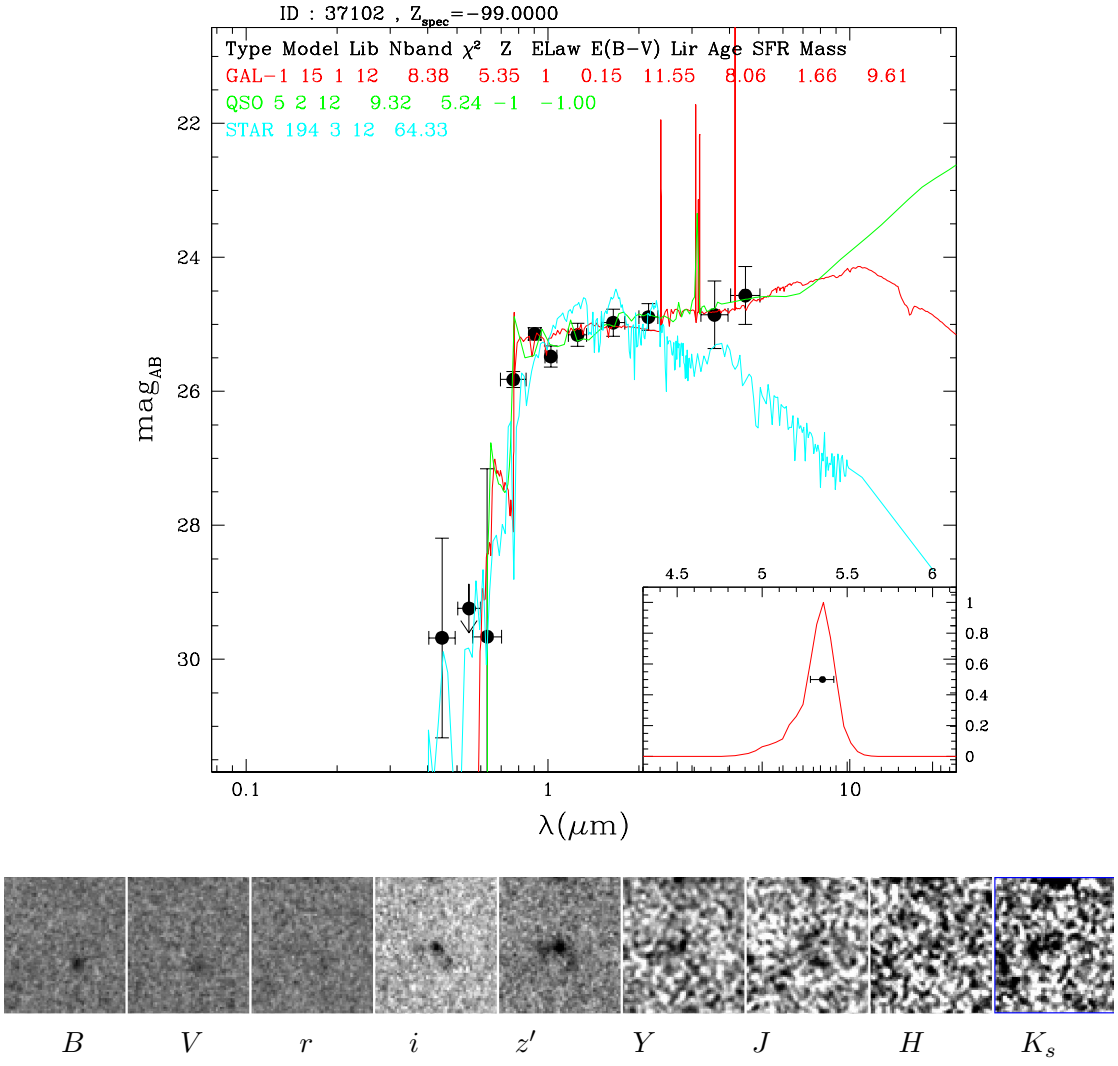


FIGURE A.1: SED-fitting and multiwavelength image of my credible candidates with $z = 5.35$, size of each poststamp is 20arcsec × 20arcsec.

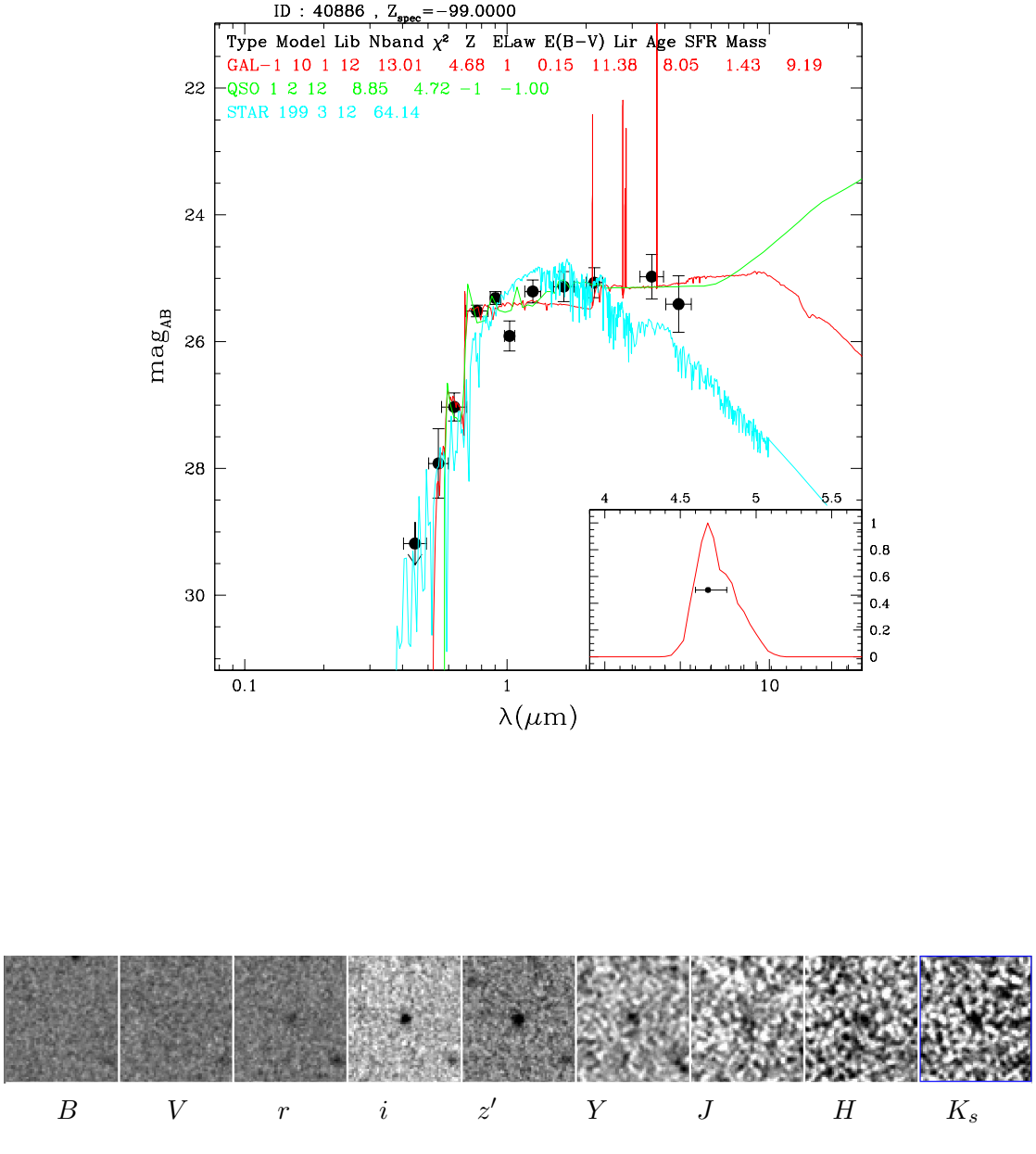
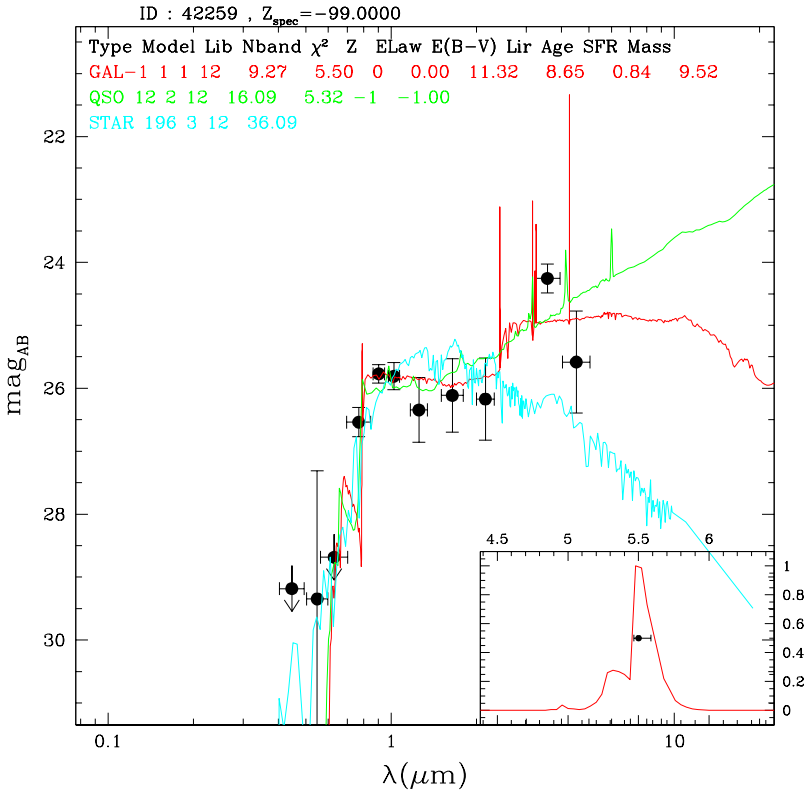


FIGURE A.2: SED-fitting and multiwavelength image of my credible candidates with $z = 4.68$, size of each poststamp is 20arcsec × 20arcsec.



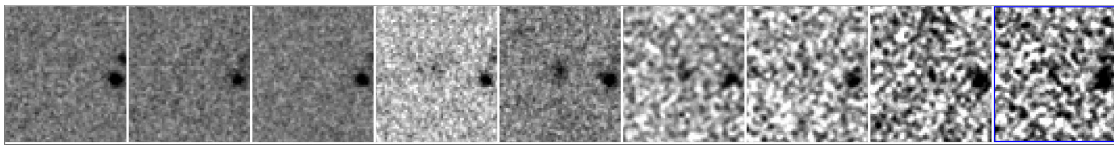


FIGURE A.3: SED-fitting and multiwavelength image of my credible candidates with $z = 5.5$, size of each poststamp is 20arcsec × 20arcsec.

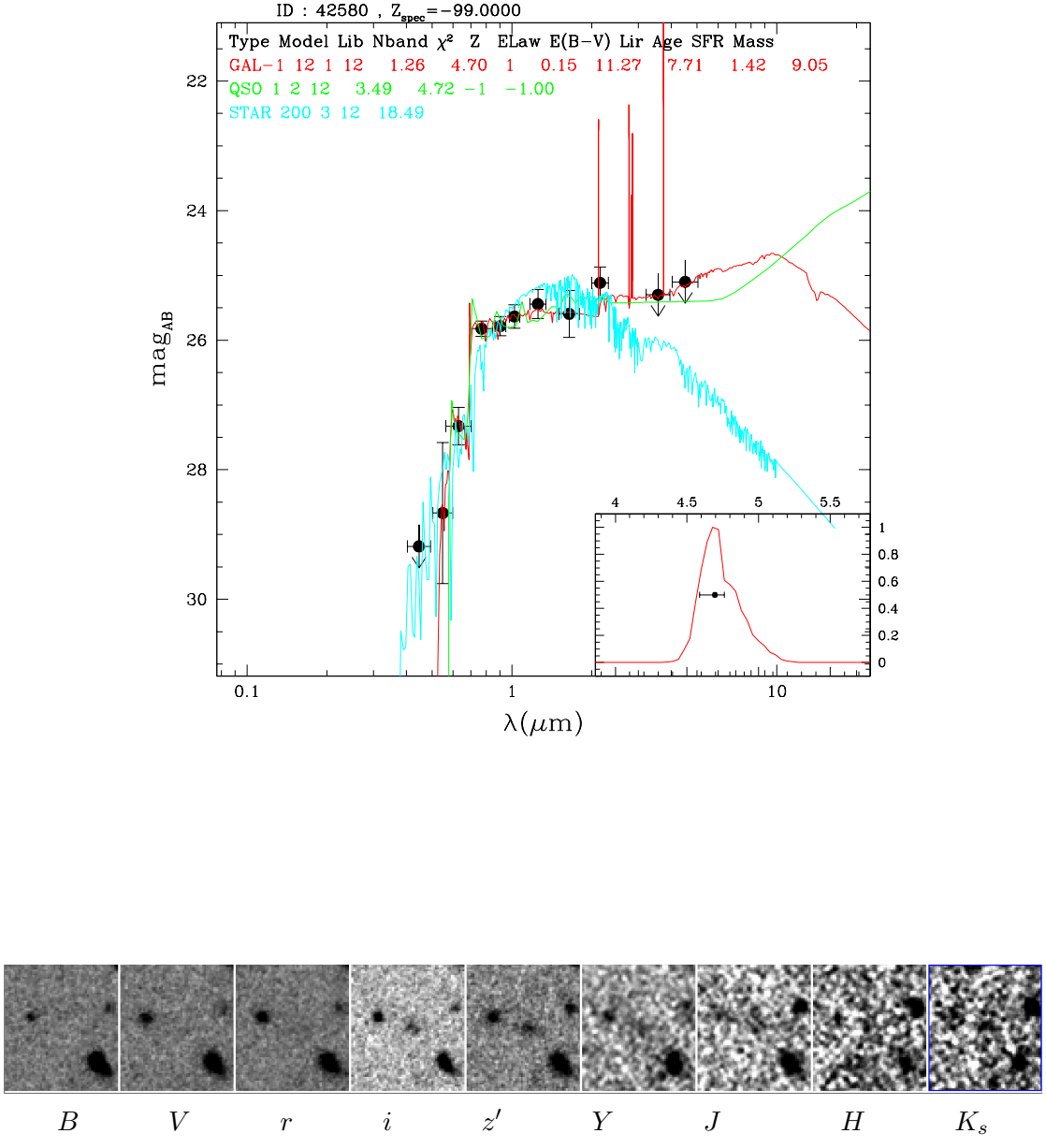


FIGURE A.4: SED-fitting and multiwavelength image of my credible candidates with $z = 4.7$, size of each poststamp is 20arcsec × 20arcsec.

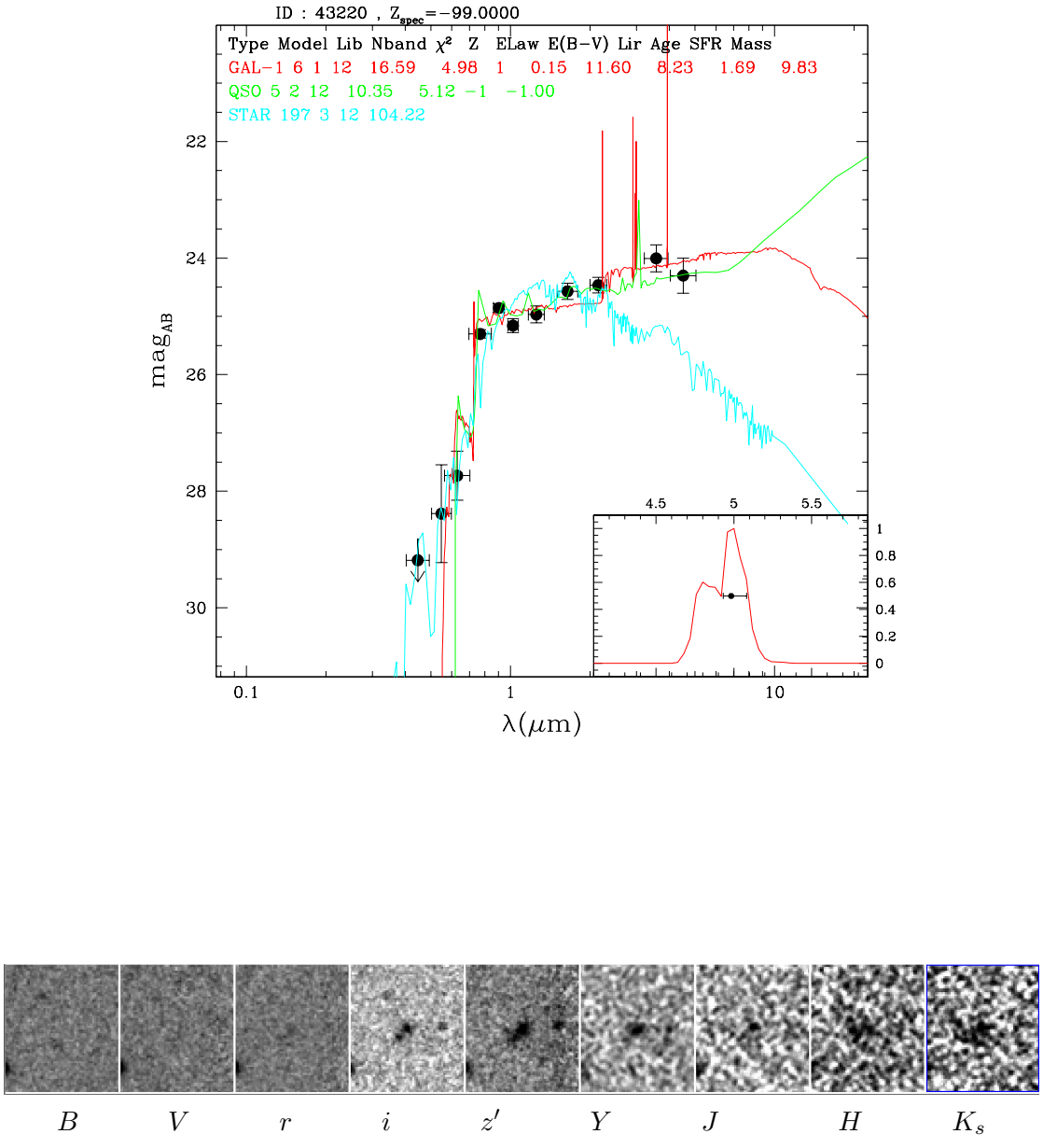


FIGURE A.5: SED-fitting and multiwavelength image of my credible candidates with $z = 4.98$, size of each poststamp is 20arcsec × 20arcsec.

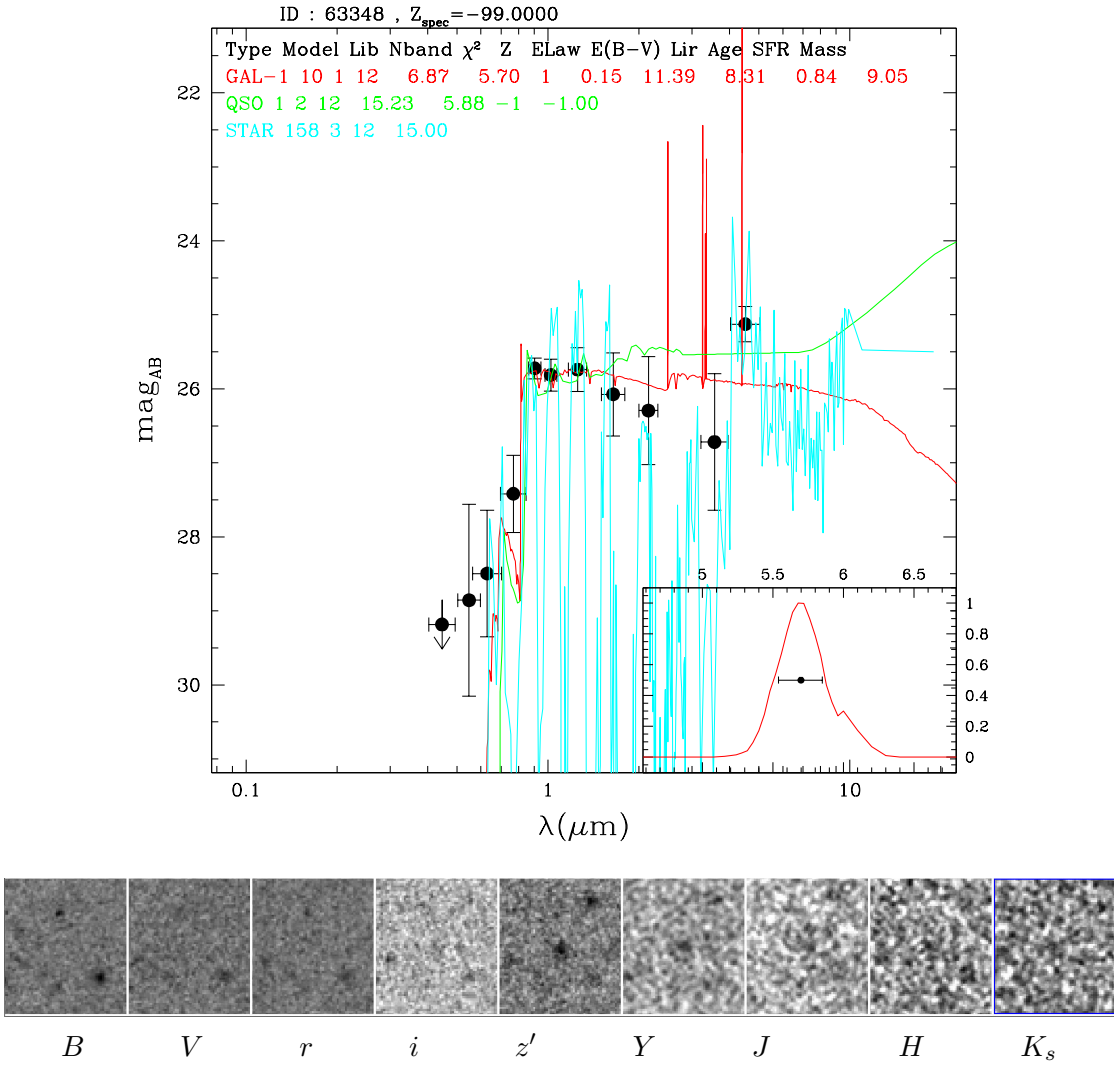


FIGURE A.6: SED-fitting and multiwavelength image of my credible candidates with $z = 5.7$, size of each poststamp is 20arcsec × 20arcsec.

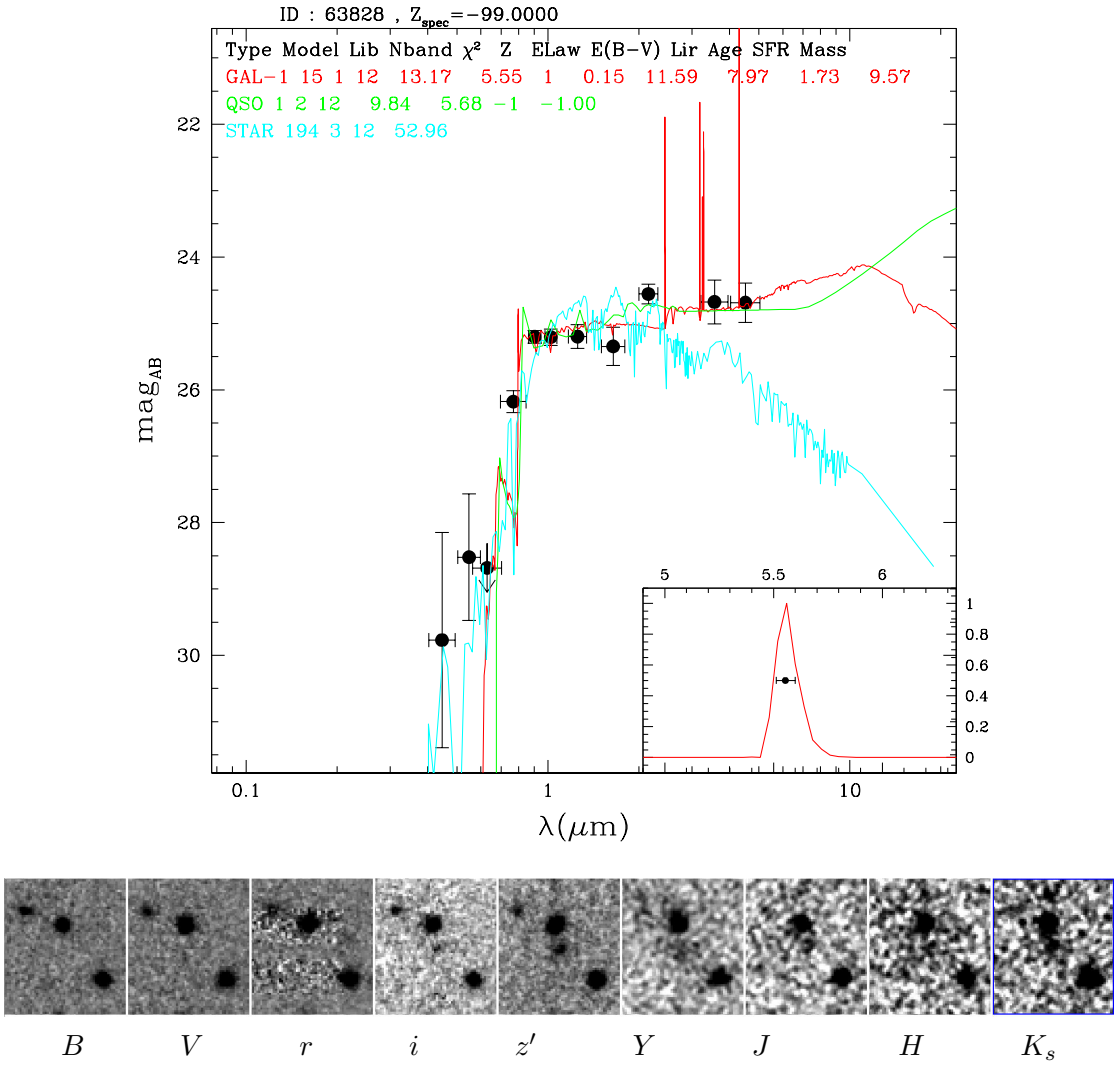


FIGURE A.7: SED-fitting and multiwavelength image of my credible candidates with $z = 5.55$, size of each poststamp is 20arcsec × 20arcsec.

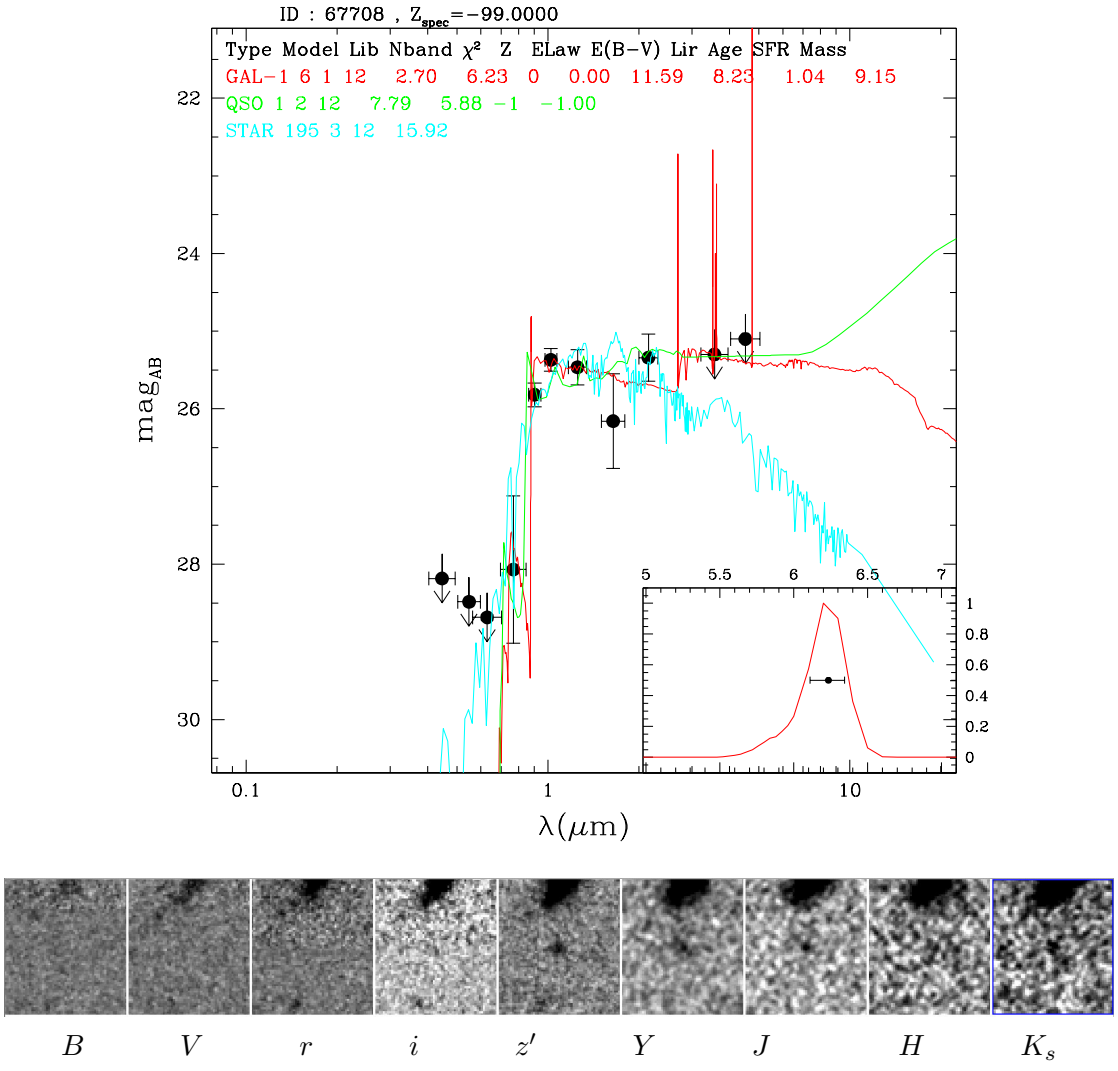


FIGURE A.8: SED-fitting and multiwavelength image of my credible candidates with $z = 6.23$, size of each poststamp is 20arcsec × 20arcsec.

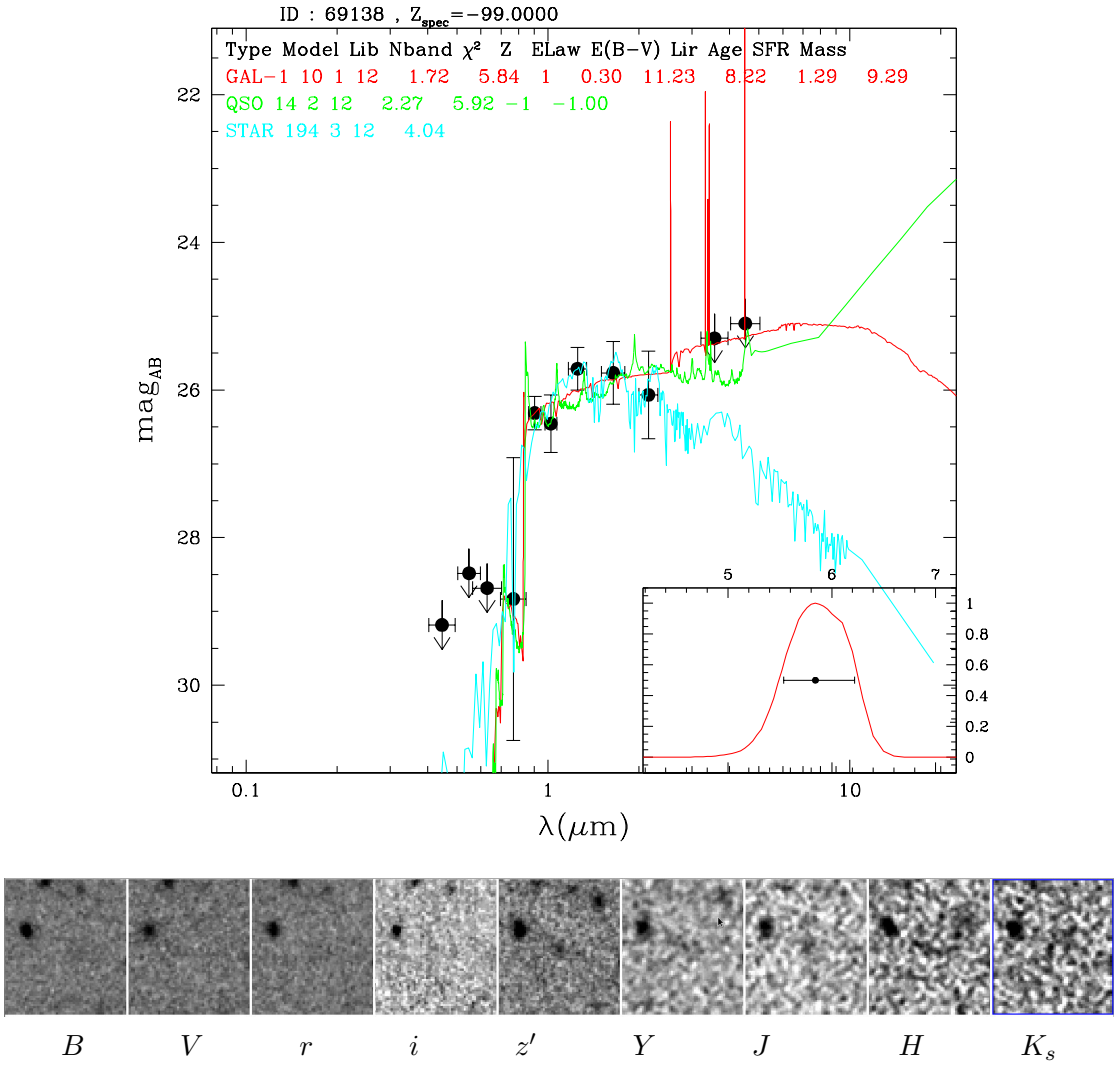


FIGURE A.9: SED-fitting and multiwavelength image of my credible candidates with $z = 5.84$, size of each poststamp is 20arcsec × 20arcsec.

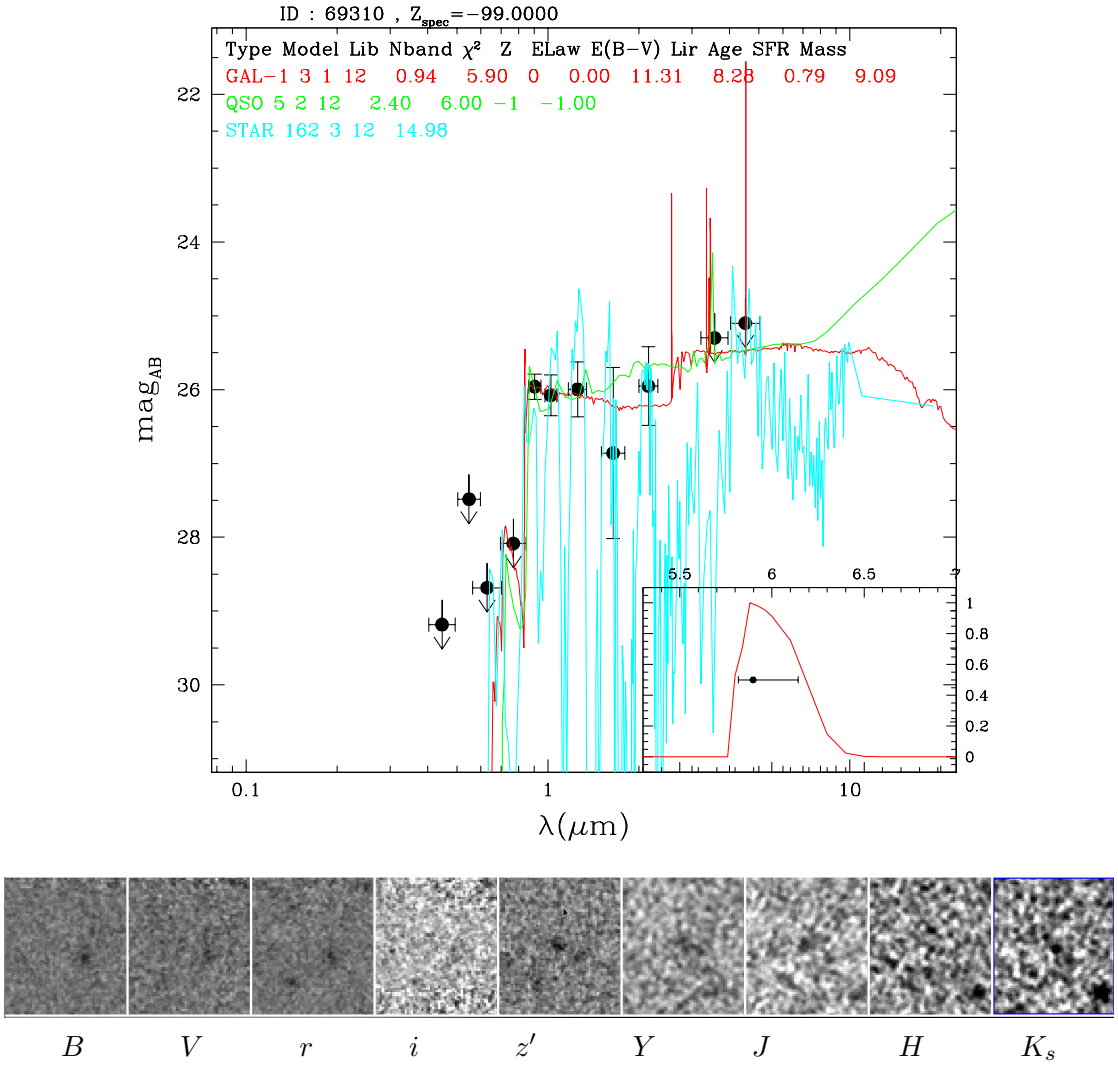


FIGURE A.10: SED-fitting and multiwavelength image of my credible candidates with $z = 5.9$, size of each poststamp is 20arcsec × 20arcsec.

Bibliography

[1] S De Barros, D Schaerer, and D P Stark. Z 3 6. 3(2013), 2014.

[2] Anton M. Koekemoer, S. M. Faber, Henry C. Ferguson, Norman a. Grogin, Dale D. Kocevski, David C. Koo, Kamson Lai, Jennifer M. Lotz, Ray a. Lucas, Elizabeth J. McGrath, Sara Ogaz, Abhijith Rajan, Adam G. Riess, Steve a. Rodney, Louis Strolger, Stefano Casertano, Tomas Dahlen, Mark Dickinson, Timothy Dolch, Adriano Fontana, Mauro Giavalisco, Andrea Grazian, Yicheng Guo, Nimish P. Hathi, Kuang-Han Huang, Arjen van der Wel, Hao-Jing Yan, Viviana Acquaviva, David M. Alexander Omar Almaini, Matthew L. N. Ashby, Marco Barden, Eric F. Bell, Frédéric Bournaud, Thomas M. Brown, Karina I. Caputi, Paolo Cassata, Peter Challis, Ranga-Ram Chary, Edmond Cheung, Michele Cirasuolo, Christopher J. Conselice, Asantha Roshan Cooray, Darren J. Croton, Emanuele Daddi, Romeel Davé, Duilia F. de Mello, Loic de Ravel, Avishai Dekel, Jennifer L. Donley, James S. Dunlop, Aaron a. Dutton, David Elbaz, Giovanni G. Fazio, Alex V. Filippenko, Steven L. Finkelstein, Chris Frazer, Jonathan P. Gardner, Peter M. Garnavich, Eric Gawiser, Ruth Gruetzbauch, Will G. Hartley, Boris Häussler, Jessica Herrington, Philip F. Hopkins, Jia-Sheng Huang, Saurabh Jha, Andrew Johnson, Jeyhan S. Kartaltepe, Ali Ahmad Khostovan, Robert P. Kirshner, Caterina Lani, Kyoung-Soo Lee, Weidong Li, Piero Madau, Patrick J. McCarthy, Daniel H. McIntosh, Ross J. McLure, Conor McPartland, Bahram Mobasher, Heidi Moreira, Alice Mortlock, Leonidas a. Moustakas, Mark Mozena, Kirpal Nandra, Jeffrey a. Newman, Jennifer L. Nielsen, Sami Niemi, Kai G. Noeske, Casey J. Papovich, Laura Pentericci, Alexandra Pope, Joel R. Primack, Swara Ravindranath, Naveen a. Reddy, Alvio Renzini, Hans-Walter Rix, Aday R. Robaina, David J. Rosario, Piero Rosati, Sara Salimbeni, Claudia Scarlata, Brian Siana, Luc Simard, Joseph Smidt, Diana Snyder, Rachel S. Somerville, Hyron Spinrad, Amber N. Straughn, Olivia Telford, Harry I. Teplitz, Jonathan R. Trump, Carlos Vargas, Carolin Villforth, Cory R. Wagner, Pat Wandro, Risa H. Wechsler, Benjamin J. Weiner, Tommy Wiklind, Vivienne Wild, Grant Wilson, Stijn Wuyts, and Min S. Yun. CANDELS: The Cosmic Assembly

Near-infrared Deep Extragalactic Legacy Survey - The Hubble Space Telescope Observations, Imaging Data Products and Mosaics. 35:36, 2011. ISSN 0067-0049. doi: 10.1088/0067-0049/197/2/35. URL http://arxiv.org/abs/1105.3754.

[3] H. J. McCracken, B. Milvang-Jensen, J. Dunlop, M. Franx, J. P. U. Fynbo, O. Le Fèvre, J. Holt, K. I. Caputi, Y. Goranova, F. Buitrago, J. P. Emerson, W. Freudling, P. Hudelot, C. López-Sanjuan, F. Magnard, Y. Mellier, P. Mø ller, K. K. Nilsson, W. Sutherland, L. Tasca, and J. Zabl. UltraVISTA: a new ultra-deep near-infrared survey in COSMOS. *Astronomy & Astrophysics*, 544:A156, 2012. ISSN 0004-6361. doi: 10.1051/0004-6361/201219507. URL http://arxiv.org/abs/1204.6586.

[4] K. I. Caputi, M. J. Michaowski, M. Krips, J. E. Geach, M. L. N. Ashby, J.-S. Huang, G. G. Fazio, a. M. Koekemoer, G. Popping, M. Spaans, M. Castellano, J. S. Dunlop, a. Fontana, and P. Santini. PdBI COLD DUST IMAGING OF TWO EXTREMELY RED H [4.5] ¿ 4 GALAXIES DISCOVERED WITH SEDS AND CANDELS. *The Astrophysical Journal*, 788(2):126, June 2014. ISSN 0004-637X. doi: 10.1088/0004-637X/788/2/126. URL http://stacks.iop.org/0004-637X/788/i=2/a=126?key=crossref.7f6e3ebe7e9586f4cf7fa9a958781087.

[5] Charles L. Steinhardt, Josh S. Speagle, Peter Capak, John D. Silverman, Marcella Carollo, James Dunlop, Yasuhiro Hashimoto, Bau-Ching Hsieh, Olivier Ilbert, Olivier Le Fevre, Emeric Le Floc'h, Nicholas Lee, Lihwai Lin, Yen-Ting Lin, Dan Masters, Henry J. McCracken, Tohru Nagao, Andreea Petric, Mara Salvato, Dave Sanders, Nick Scoville, Kartik Sheth, Michael a. Strauss, and Yoshiaki Taniguchi. STAR FORMATION AT 4 ¡ z ¡ 6 FROM THE SPITZER LARGE AREA SURVEY WITH HYPER-SUPRIME-CAM (SPLASH). *The Astrophysical Journal*, 791:L25, 2014. ISSN 2041-8213. doi: 10.1088/2041-8205/791/2/L25. URL http://adsabs.harvard.edu/abs/2014ApJ...791L..25S.

[6] P. Capak, H. Aussel, M. Ajiki, H. J. McCracken, B. Mobasher, N. Scoville, P. Shopbell, Y. Taniguchi, D. Thompson, S. Tribiano, S. Sasaki, a. W. Blain, M. Brusa, C. Carilli, a. Comastri, C. M. Carollo, P. Cassata, J. Colbert, R. S. Ellis, M. Elvis, M. Giavalisco, W. Green, L. Guzzo, G. Hasinger, O. Ilbert, C. Impey, K. Jahnke, J. Kartaltepe, J-P. Kneib, J. Koda, a. Koekemoer, Y. Komiyama, a. Leauthaud, O. Lefevre, S. Lilly, R. Massey, S. Miyazaki, T. Murayama, T. Nagao, J. a. Peacock, a. Pickles, C. Porciani, a. Renzini, J. Rhodes, M. Rich, M. Salvato, D. B. Sanders, C. Scarlata, D. Schiminovich, E. Schinnerer, M. Scodeggio, K. Sheth, Y. Shioya, L. a. M. Tasca, J. E. Taylor, L. Yan, and G. Zamorani. The First Release COSMOS Optical and Near-IR Data and Catalog. *The Astrophysical Journal Supplement Series*, 172(1):19, September 2007. ISSN 0067-0049. doi: 10.1086/519081. URL http://arxiv.org/abs/0704.2430.

[7] Ol Fevre, P. Cassata, and O. Cucciati. The VIMOS VLT Deep Survey: the redshift distribution N (z) of magnitude-limited samples down to iAB= 24.75 and KsAB= 22. *arXiv preprint arXiv: ...*, July 2013. URL http://arxiv.org/abs/1307.6518.

[8] Masami Ouchi, Bahram Mobasher, Kazuhiro Shimasaku, Henry C. Ferguson, S. Michael Fall, Yoshiaki Ono, Nobunari Kashikawa, Tomoki Morokuma, Kimihiko Nakajima, Sadanori Okamura, Mark Dickinson, Mauro Giavalisco, and Kouji Ohta. Large Area Survey for z=7 Galaxies in SDF and GOODS-N: Implications for Galaxy Formation and Cosmic Reionization. *The Astrophysical Journal*, 706(2): 20, December 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/706/2/1136. URL http://arxiv.org/abs/0908.3191.

[9] D. Calzetti, L. Armus, R. C. Bohlin, a. L. Kinney, J. Koornneef, and T. Storchi-Bergmann. The Dust Content and Opacity of Actively Star-Forming Galaxies. 20: 27, 1999. ISSN 0004-637X. doi: 10.1086/308692. URL http://arxiv.org/abs/astro-ph/9911459.

[10] Y. Taniguchi, N. Scoville, T. Murayama, D. B. Sanders, B. Mobasher, H. Aussel, P. Capak, M. Ajiki, S. Miyazaki, Y. Komiyama, Y. Shioya, T. Nagao, S. S. Sasaki, J. Koda, C. Carilli, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, O. LeFevre, S. Lilly, a. Renzini, M. Rich, E. Schinnerer, P. Shopbell, N. Kaifu, H. Karoji, N. Arimoto, S. Okamura, and K. Ohta. The Cosmic Evolution Survey (COSMOS): Subaru Observations of the HST COSMOS Field. 2006. ISSN 0067-0049. doi: 10.1086/516596.

[11] Andrew M. Hopkins and John F. Beacom. On the normalisation of the cosmic star formation history. page 15, 2006. ISSN 0004-637X. doi: 10.1086/506610. URL http://arxiv.org/abs/astro-ph/0601463.

[12] a. van der Wel, M. Franx, P. G. van Dokkum, R. E. Skelton, I. G. Momcheva, K. E. Whitaker, G. B. Brammer, E. F. Bell, H.-W. Rix, S. Wuyts, H. C. Ferguson, B. P. Holden, G. Barro, a. M. Koekemoer, Yu-Yen Chang, E. J. McGrath, B. Häussler, a. Dekel, P. Behroozi, M. Fumagalli, J. Leja, B. F. Lundgren, M. V. Maseda, E. J. Nelson, D. a. Wake, S. G. Patel, I. Labbé, S. M. Faber, N. a. Grogin, and D. D. Kocevski. 3D-HST+CANDELS: THE EVOLUTION OF THE GALAXY SIZE-MASS DISTRIBUTION SINCE z = 3. *The Astrophysical Journal*, 788:28, 2014. ISSN 0004-637X. doi: 10.1088/0004-637X/788/1/28. URL http://adsabs.harvard.edu/abs/2014ApJ...788...28V$\delimiter"026E30F$nhttp://stacks.iop.org/0004-637X/788/i=1/a=28?key=crossref.281bee0dcbf71fccb5ab5e017e376adc.

[13] Robert C. Kennicutt and Jr. The Global Schmidt Law in Star Forming Galaxies, 1997. ISSN 0004-637X. URL http://arxiv.org/abs/astro-ph/9712213.

[14] R. a a Bowler, J. S. Dunlop, R. J. McLure, H. J. McCracken, B. Milvang-Jensen, H. Furusawa, J. P U Fynbo, O. Le Fèvre, J. Holt, Y. Ideue, Y. Ihara, a. B. Rogers, and Y. Taniguchi. Discovery of bright z 7 galaxies in the UltraVISTA survey. *Monthly Notices of the Royal Astronomical Society*, 426(September):2772–2788, May 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21904.x. URL http://arxiv.org/abs/1205.4270v3.

[15] Benjamin P. Moster, Rachel S. Somerville, Jeffrey a. Newman, and Hans-Walter Rix. A Cosmic Variance Cookbook. 113(2002):8, 2010. ISSN 0004-637X. doi: 10.1088/0004-637X/731/2/113. URL http://arxiv.org/abs/1001.1737.

[16] Matthew Colless, P Gavin Dalton, Steve Maddox, Will Sutherland, Peder Norberg, Shaun Cole, Joss Bland-hawthorn, Terry Bridges, Russell Cannon, Chris Collins, Warrick Couch, Nicholas Cross, Kathryn Deeley, Roberto De Propris, Simon P Driver, George Efstathiou, Richard S Ellis, Carlos S Frenk, Karl Glazebrook, Carole Jackson, Ofer Lahav, Ian Lewis, Stuart Lumsden, Darren Madgwick, John a Peacock, Bruce a Peterson, Ian Price, Mark Seaborne, and Keith Taylor. The 2dF Galaxy Redshift Sur vey : spectr a and r edshifts. 1063:1039–1063, 2001.

[17] R. a a Bowler, J. S. Dunlop, R. J. McLure, a. B. Rogers, H. J. McCracken, B. Milvang-Jensen, H. Furusawa, J. P U Fynbo, Y. Taniguchi, J. Afonso, M. N. Bremer, and O. Le Fèvre. The bright end of the galaxy luminosity function at z 7: Before the onset of mass quenching? *Monthly Notices of the Royal Astronomical Society*, 440(3):2810–2842, April 2014. ISSN 13652966. doi: 10.1093/mnras/stu449. URL http://mnras.oxfordjournals.org/cgi/doi/10.1093/mnras/stu449.

[18] R. J. Bouwens, G. D. Illingworth, P. a. Oesch, M. Trenti, I. Labbe', L. Bradley, M. Carollo, P. G. van Dokkum, V. Gonzalez, B. Holwerda, M. Franx, L. Spitler, R. Smit, and D. Magee. UV Luminosity Functions at redshifts z˜4 to z˜10: 11000 Galaxies from HST Legacy Fields. *eprint arXiv:1403.4295*, 3, 2014. URL http://adsabs.harvard.edu/abs/2014arXiv1403.4295B.

[19] Charles C. Steidel, Kurt L. Adelberger, Alice E. Shapley, Max Pettini, Mark Dickinson, and Mauro Giavalisco. 3: Survey Description and Full Data Set. *The Astrophysical Journal*, 592:728–754, 2003. ISSN 0004-637X. doi: 10.1086/375772.

[20] E. Vanzella, M. Giavalisco, M. Dickinson, S. Cristiani, M. Nonino, H. Kuntschner, P. Popesso, P. Rosati, a. Renzini, D. Stern, C. Cesarsky, H. C. Ferguson, R. a. E.

Fosbury, and the Goods Team. Spectroscopic Observations of Lyman-Break Galaxies at Redshift ˜ 4, 5 and 6 in the GOODS-South Field. page 58, 2009. doi: 10.1088/0004-637X/695/2/1163. URL http://arxiv.org/abs/0901.4364.

[21] C. Kevin Xu, Yinghe Zhao, N. Scoville, P. Capak, N. Drory, and Y. Gao. = 1. *The Astrophysical Journal*, 747(2):85, March 2012. ISSN 0004-637X. doi: 10.1088/0004-637X/747/2/85. URL http://stacks.iop.org/0004-637X/747/i=2/a=85?key=crossref.65065cf6f91b0da007a517c100488f35.

[22] S. J. Lilly, L. Tresse, F. Hammer, David Crampton, and O. Le Fevre. The CANADA-FRANCE REDSHIFT SURVEY VI: Evolution of the galaxy luminosity function to z ˜ 1. 1995. ISSN 0004-637X. doi: 10.1086/176560. URL http://arxiv.org/abs/astro-ph/9507079.

[23] O. Le Fevre, L. a. M. Tasca, P. Cassata, B. Garilli, V. Le Brun, D. Maccagni, L. Pentericci, R. Thomas, E. Vanzella, G. Zamorani, E. Zucca, R. Amorin, S. Bardelli, P. Capak, L. Cassara, M. Castellano, a. Cimatti, J. G. Cuby, O. Cucciati, S. de la Torre, a. Durkalec, a. Fontana, M. Giavalisco, a. Grazian, N. P. Hathi, O. Ilbert, B. C. Lemaux, C. Moreau, S. Paltani, B. Ribeiro, M. Salvato, D. Schaerer, M. Scodeggio, V. Sommariva, M. Talia, Y. Taniguchi, L. Tresse, D. Vergani, P. W. Wang, S. Charlot, T. Contini, S. Fotopoulo, C. Lopez-Sanjuan, Y. Mellier, and N. Scoville. The VIMOS Ultra-Deep Survey: ˜10,000 galaxies with spectroscopic redshifts to study galaxy assembly at early epochs 2¡z¡˜6. *arXiv preprint arXiv:1403.3938*, 2014.

[24] Chris J. Willott, Ross J. McLure, Pascale Hibon, Richard Bielby, Henry J. McCracken, Jean-Paul Kneib, Olivier Ilbert, David G. Bonfield, Victoria a. Bruce, and Matt J. Jarvis. AN EXPONENTIAL DECLINE AT THE BRIGHT END OF THE z = 6 GALAXY LUMINOSITY FUNCTION. *The Astronomical Journal*, 2013. ISSN 0004-6256. doi: 10.1088/0004-6256/145/1/4.

[25] Volker Bromm and Naoki Yoshida. The First Galaxies. *Annual Review of Astronomy and Astrophysics*, 49(1):75, September 2011. ISSN 0066-4146. doi: 10.1146/annurev-astro-081710-102608. URL http://arxiv.org/abs/1102.4638.

[26] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, a. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. Nine-Year Wilkinson Microwave Anisotropy Probe (Wmap) Observations: Cosmological Parameter Results. *The Astrophysical Journal Supplement Series*, 208:19, 2013. ISSN 0067-0049. doi: 10.1088/0067-0049/208/2/19. URL http://adsabs.harvard.edu/abs/2013ApJS..208...19H.

[27] R. J. Bouwens, G. D. Illingworth, P. a. Oesch, I. Labbe, P. G. van Dokkum, M. Trenti, M. Franx, R. Smit, V. Gonzalez, and D. Magee. UV-Continuum Slopes of ¿4000 z˜4-8 Galaxies from the HUDF/XDF, HUDF09, ERS, CANDELS-South, and CANDELS-North Fields. *The Astrophysical Journal*, 793(2):40, September 2013. ISSN 15384357. doi: 10.1088/0004-637X/793/2/115. URL http://arxiv.org/abs/1306.2950.

[28] Masahiro Takada, Nobuyuki Kawai, and Shigehiro Nagataki. Subaru Hyper Suprime-Cam Project. 120:120–127, 2010. doi: 10.1063/1.3509247. URL http://scitation.aip.org/content/aip/proceeding/aipcp/10.1063/1.3509247.

[29] L. a. M. Tasca, O. Le Fevre, C. Lopez-Sanjuan, P. W. Wang, P. Cassata, B. Garilli, O. Ilbert, V. Le Brun, B. C. Lemaux, D. Maccagni, L. Tresse, S. Bardelli, T. Contini, O. Cucciati, a. Fontana, M. Giavalisco, J. P. Kneib, M. Salvato, Y. Taniguchi, D. Vergani, G. Zamorani, and E. Zucca. The role of major mergers in shaping galaxies at 2¡z¡4 in the VUDS and VVDS surveys. page 12, March 2013. URL http://arxiv.org/abs/1303.4400.

[30] E. Daddi, M. Dickinson, G. Morrison, R. Chary, a. Cimatti, D. Elbaz, D. Frayer, a. Renzini, a. Pope, D. M. Alexander, F. E. Bauer, M. Giavalisco, M. Huynh, J. Kurk, and M. Mignoli. Multiwavelength study of massive galaxies at z˜2. I. Star formation and galaxy growth. page 19, 2007. ISSN 0004-637X. doi: 10.1086/521818. URL http://arxiv.org/abs/0705.2831.

[31] R Laureijs, J Amiaux, S Arduini, J L Augueres, J Brinchmann, R Cole, M Cropper, C Dabin, L Duvet, a Ealet, B Garilli, P Gondoin, L Guzzo, J Hoar, H Hoekstra, R Holmes, T Kitching, T Maciaszek, Y Mellier, F Pasian, W Percival, J Rhodes, G Saavedra Criado, M Sauvage, R Scaramella, L Valenziano, S Warren, R Bender, F Castander, a Cimatti, O Le Fevre, H Kurki-Suonio, M Levi, P Lilje, G Meylan, R Nichol, K Pedersen, V Popa, R Rebolo Lopez, H W Rix, H Rottgering, W Zeilinger, F Grupp, P Hudelot, R Massey, M Meneghetti, L Miller, S Paltani, S Paulin-Henriksson, S Pires, C Saxton, T Schrabback, G Seidel, J Walsh, N Aghanim, L Amendola, J Bartlett, C Baccigalupi, J P Beaulieu, K Benabed, J-G Cuby, D Elbaz, P Fosalba, G Gavazzi, a Helmi, I Hook, M Irwin, J P Kneib, M Kunz, F Mannucci, L Moscardini, C Tao, R Teyssier, J Weller, G Zamorani, M R Zapatero Osorio, O Boulade, J J Foumond, a Di Giorgio, P Guttridge, a James, M Kemp, J Martignac, a Spencer, D Walton, T Blümchen, C Bonoli, F Bortoletto, C Cerna, L Corcione, C Fabron, K Jahnke, S Ligori, F Madrid, L Martin, G Morgante, T Pamplona, E Prieto, M Riva, R Toledo, M Trifoglio, F Zerbi, F Abdalla, M Douspis, C Grenet, S Borgani, R Bouwens, F Courbin, J M Delouis, P Dubath, a Fontana, M Frailis, a Grazian, J Koppenhöfer, O Mansutti,

M Melchior, M Mignoli, J Mohr, C Neissner, K Noddle, M Poncet, M Scodeggio, S Serrano, N Shane, J L Starck, C Surace, a Taylor, G Verdoes-Kleijn, C Vuerli, O R Williams, a Zacchei, B Altieri, I Escudero Sanz, R Kohley, T Oosterbroek, P Astier, D Bacon, S Bardelli, C Baugh, F Bellagamba, C Benoist, D Bianchi, a Biviano, E Branchini, C Carbone, V Cardone, D Clements, S Colombi, C Conselice, G Cresci, N Deacon, J Dunlop, C Fedeli, F Fontanot, P Franzetti, C Giocoli, J Garcia-Bellido, J Gow, a Heavens, P Hewett, C Heymans, a Holland, Z Huang, O Ilbert, B Joachimi, E Jennins, E Kerins, a Kiessling, D Kirk, R Kotak, O Krause, O Lahav, F van Leeuwen, J Lesgourgues, M Lombardi, M Magliocchetti, K Maguire, E Majerotto, R Maoli, F Marulli, S Maurogordato, H McCracken, R McLure, a Melchiorri, a Merson, M Moresco, M Nonino, P Norberg, J Peacock, R Pello, M Penny, V Pettorino, C Di Porto, L Pozzetti, C Quercellini, M Radovich, a Rassat, N Roche, S Ronayette, E Rossetti, B Sartoris, P Schneider, E Semboloni, S Serjeant, F Simpson, C Skordis, G Smadja, S Smartt, P Spano, S Spiro, M Sullivan, a Tilquin, R Trotta, L Verde, Y Wang, G Williger, G Zhao, J Zoubian, and E Zucca. Euclid Definition Study Report. *arXiv.org*, 1110(July):3193, 2011. URL http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2011arXiv1110.3193L\&link_type=ABSTRACT$\delimiter"026E30F$npapers2://publication/uuid/F9BA842F-467A-4AFE-85EF-33EC4BDB0563.

[32] Db Sanders, H Aussel, O Ilbert, Js Kartaltepe, Je Barnes, Mc Liu, Jp Williams, M Salvato, N Scoville, G Helou, T Brooke, Aw Blain, P Capak, Rs Ellis, Pl Shopbell, H Aussel, N Scoville, Ja Surace, Dt Frayer, K Sheth, G Helou, B Bhattacharya, L Yan, D Calzetti, Sm Fall, M Giavalisco, a Koekemoer, B Mobasher, C Carilli, Cm Carollo, S Lilly, a Comastri, E Daddi, E Daddi, M Elvis, a Franceschini, a Renzini, G Hasinger, C Impey, Le Fèvre O, Hj McCracken, Hj McCracken, a Renzini, M Rich, E Schinnerer, Y Taniguchi, Dj Thompson, Dj Thompson, and Cm Urry. S-COSMOS: The Spitzer legacy survey of the Hubble Space Telescope ACS 2 deg COSMOS field I: Survey strategy and first analysis. pages 86–98, 2007. ISSN 0067-0049. doi: 10.1086/517885. URL http://hdl.handle.net/2381/11535.

[33] O. Ilbert, L. Tresse, E. Zucca, S. Bardelli, S. Arnouts, G. Zamorani, L. Pozzetti, D. Bottini, B. Garilli, V. LeBrun, O. Le Fèvre, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, G. Vettolani, a. Zanichelli, C. Adami, M. Arnaboldi, M. Bolzonella, a. Cappi, S. Charlot, T. Contini, S. Foucaud, P. Franzetti, I. Gavignaud, L. Guzzo, a. Iovino, H. J. McCracken, B. Marano, C. Marinoni, G. Mathez, a. Mazure, B. Meneux, R. Merighi, S. Paltani, R. Pello, a. Pollo, M. Radovich, M. Bondi, a. Bongiorno, G. Busarello, P. Ciliegi, Y. Mellier, P. Merluzzi, V. Ripepi, and D. Rizzo. The VIMOS-VLT Deep Survey: Evolution of the galaxy luminosity function up to z=2 in first epoch data. *Astronomy and*

..., 876:17, 2004. ISSN 00046361. doi: 10.1051/0004-6361:20041961. URL http://arxiv.org/abs/astro-ph/0409134.

[34] S. J. Warren, N. C. Hambly, S. Dye, O. Almaini, N. J G Cross, a. C. Edge, S. Foucaud, P. C. Hewett, S. T. Hodgkin, M. J. Irwin, R. F. Jameson, a. Lawrence, P. W. Lucas, a. J. Adamson, R. M. Bandyopadhyay, J. Bryant, R. S. Collins, C. J. Davis, J. S. Dunlop, J. P. Emerson, D. W. Evans, E. a. Gonzales-Solares, P. Hirst, M. J. Jarvis, T. R. Kendall, T. H. Kerr, S. K. Leggett, J. R. Lewis, R. G. Mann, R. J. McLure, R. G. McMahon, D. J. Mortlock, M. G. Rawlings, M. a. Read, M. Riello, C. Simpson, D. J B Smith, E. T W Sutorius, T. a. Targett, and W. P. Varricatt. The United Kingdom infrared telescope infrared deep sky survey first data release. *Monthly Notices of the Royal Astronomical Society*, 375:213–226, 2007. ISSN 00358711. doi: 10.1111/j.1365-2966.2006.11284.x.

[35] Piero Madau and Mark Dickinson. Cosmic Star-Formation History. *Annual Review of Astronomy and Astrophysics*, 52(1):415–486, August 2014. ISSN 0066-4146. doi: 10.1146/annurev-astro-081811-125615. URL http://www.annualreviews.org/doi/abs/10.1146/annurev-astro-081811-125615.

[36] Steven L. Finkelstein, Russell E. Ryan Jr., Casey Papovich, Mark Dickinson, Mimi Song, Rachel Somerville, Henry C. Ferguson, Brett Salmon, Mauro Giavalisco, Anton M. Koekemoer, Matthew L. N. Ashby, Peter Behroozi, Marco Castellano, James S. Dunlop, Sandy M. Faber, Giovanni G. Fazio, Adriano Fontana, Norman a. Grogin, Nimish Hathi, Jason Jaacks, Dale D. Kocevski, Rachael Livermore, Ross J. McLure, Emiliano Merlin, Bahram Mobasher, Jeffrey a. Newman, Marc Rafelski, Vithal Tilvi, and S. P. Willner. The Evolution of the Galaxy Rest-Frame Ultraviolet Luminosity Function Over the First Two Billion Years. October 2014. URL http://arxiv.org/abs/1410.5439v1.

[37] Rychard J. Bouwens, Garth D. Illingworth, Marijn Franx, and Holland Ford. UV Luminosity Functions at z˜4, 5, and 6 from the HUDF and other Deep HST ACS Fields: Evolution and Star Formation History. *The Astrophysical* ..., 4:33, 2007. doi: 10.1086/521811. URL http://arxiv.org/abs/0707.2080.

[38] Piero Madau, Lucia Pozzetti, and Mark Dickinson. The Star Formation History of Field Galaxies. 1:33, 1997. ISSN 0004-637X. doi: 10.1086/305523. URL http://arxiv.org/abs/astro-ph/9708220.

[39] O. Cucciati, L. Tresse, O. Ilbert, O. Le Fevre, B. Garilli, V. Le Brun, P. Cassata, P. Franzetti, D. Maccagni, M. Scodeggio, E. Zucca, G. Zamorani, S. Bardelli, M. Bolzonella, R. M. Bielby, H. J. McCracken, a. Zanichelli, D. Vergani, and a. Pollo. The Star Formation Rate Density and Dust Attenuation Evolution over

12 Gyr with the VVDS Surveys. *Astronomy & ...*, 31:22, 2011. ISSN 0004-6361. doi: 10.1051/0004-6361/201118010. URL http://arxiv.org/abs/1109.1005.

[40] R J Bouwens, G D Illingworth, I Labbe, P a Oesch, M Trenti, C M Carollo, P G van Dokkum, M Franx, M Stiavelli, V González, D Magee, and L Bradley. A candidate redshift z10 galaxy and rapid changes in that population at an age of 500Myr. *Nature*, 469:504–507, 2011. ISSN 0028-0836. doi: 10.1038/nature09717. URL http://adsabs.harvard.edu/cgi-bin/nph-data_query?bibcode=2011Natur.469..504B\&link_type=ABSTRACT$\delimiter"026E30F$npapers://939e88c8-a723-4abe-bb43-ccd202439e84/Paper/p14846.

[41] R S Ellis, M Takada, H Aihara, and Et Al. Extragalactic science and cosmology with the subaru prime focus spectrograph (pfs). *ArXiv:1206.0737*, (May 2009): 1–42, 2012.

[42] O. Ilbert, M. Salvato, E. Le Floc'h, H. Aussel, P. Capak, H. J. McCracken, B. Mobasher, J. Kartaltepe, N. Scoville, D. B. Sanders, S. Arnouts, K. Bundy, P. Cassata, J. P. Kneib, a. Koekemoer, O. Le Fevre, S. Lilly, J. Surace, Y. Taniguchi, L. Tasca, D. Thompson, L. Tresse, M. Zamojski, G. Zamorani, and E. Zucca. Galaxy Stellar Mass Assembly between 0.2¡z¡2 from the S-COSMOS survey. *The Astrophysical Journal*, 709(2):37, February 2009. ISSN 0004-637X. doi: 10.1088/0004-637X/709/2/644. URL http://arxiv.org/abs/0903.0102.

[43] Bau-Ching Hsieh, Wei-Hao Wang, Haojing Yan, Lihwai Lin, Hiroshi Karoji, Jeremy Lim, Paul T. P. Ho, and Chao-Wei Tsai. ¿ 7. *The Astrophysical Journal*, 749(1):88, April 2012. ISSN 0004-637X. doi: 10.1088/0004-637X/749/1/88. URL http://stacks.iop.org/0004-637X/749/i=1/a=88?key=crossref.cda83f64d92d170c5d13263a8ffcbc92.

[44] E. Bertin and S. Arnouts. SExtractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117(2):393–404, 1996. ISSN 0365-0138. doi: 10.1051/aas:1996164.

[45] Nikhil Padmanabhan, David J. Schlegel, Uroš Seljak, Alexey Makarov, Neta a. Bahcall, Michael R. Blanton, Jonathan Brinkmann, Daniel J. Eisenstein, Douglas P. Finkbeiner, James E. Gunn, David W. Hogg, Željko Ivezić, Gillian R. Knapp, Jon Loveday, Robert H. Lupton, Robert C. Nichol, Donald P. Schneider, Michael a. Strauss, Max Tegmark, and Donald G. York. The clustering of luminous red galaxies in the Sloan Digital Sky Survey imaging data. *Monthly Notices of the Royal Astronomical Society*, 378:852–872, 2007. ISSN 00358711. doi: 10.1111/j.1365-2966.2007.11593.x.

[46] Peter Capak, C L Carilli, N Lee, T Aldcroft, H Aussel, E Schinnerer, G W Wilson, M S Yun, a Blain, M Giavalisco, O Ilbert, J Kartaltepe, K Lee, H Mccracken, B Mobasher, M Salvato, S Sasaki, K S Scott, K Sheth, Y Shioya, D Thompson, M Elvis, D B Sanders, N Z Scoville, and Y Tanaguchi. SPECTROSCOPIC CONFIRMATION OF AN EXTREME STARBURST AT REDSHIFT 4 . 547 1 We report the spectroscopic confirmation of a submillimeter galaxy (SMG) at z p 4 . 547 with an estimated L IR p (0 . 5 2 . 0) # 10 13 L , . The spectra , mid-IR , and X-ray pr. *Flux*, (2007):53–56, 2008.

[47] S. G. Djorgovski, S. M. Castro, D. Stern, and a. Mahabal. On the Threshold of the Reionization Epoch. page 13, August 2001. ISSN 0004637X. doi: 10.1086/324175. URL http://arxiv.org/abs/astro-ph/0108069.

[48] O. Ilbert, H. J. McCracken, O. Le Fèvre, P. Capak, J. Dunlop, a. Karim, M. a. Renzini, K. Caputi, S. Boissier, S. Arnouts, H. Aussel, J. Comparat, Q. Guo, P. Hudelot, J. Kartaltepe, J. P. Kneib, J. K. Krogager, E. Le Floch, S. Lilly, Y. Mellier, B. Milvang-Jensen, T. Moutard, M. Onodera, J. Richard, M. Salvato, D. B. Sanders, N. Scoville, J. D. Silverman, Y. Taniguchi, L. Tasca, R. Thomas, S. Toft, L. Tresse, D. Vergani, M. Wolk, and a. Zirm. Mass assembly in quiescent and star-forming galaxies since z 4 from UltraVISTA. *Astronomy & Astrophysics*, 556:A55, 2013. ISSN 0004-6361. doi: 10.1051/0004-6361/201321100. URL http://adsabs.harvard.edu/abs/2013A\&A...556A..55I.

[49] Remco F. J. van der Burg, Hendrik Hildebrandt, and Thomas Erben. The UV galaxy luminosity function at z=3-5 from the CFHT Legacy Survey Deep fields. 74:14, 2010. ISSN 0004-6361. doi: 10.1051/0004-6361/200913812. URL http://arxiv.org/abs/1009.0758.