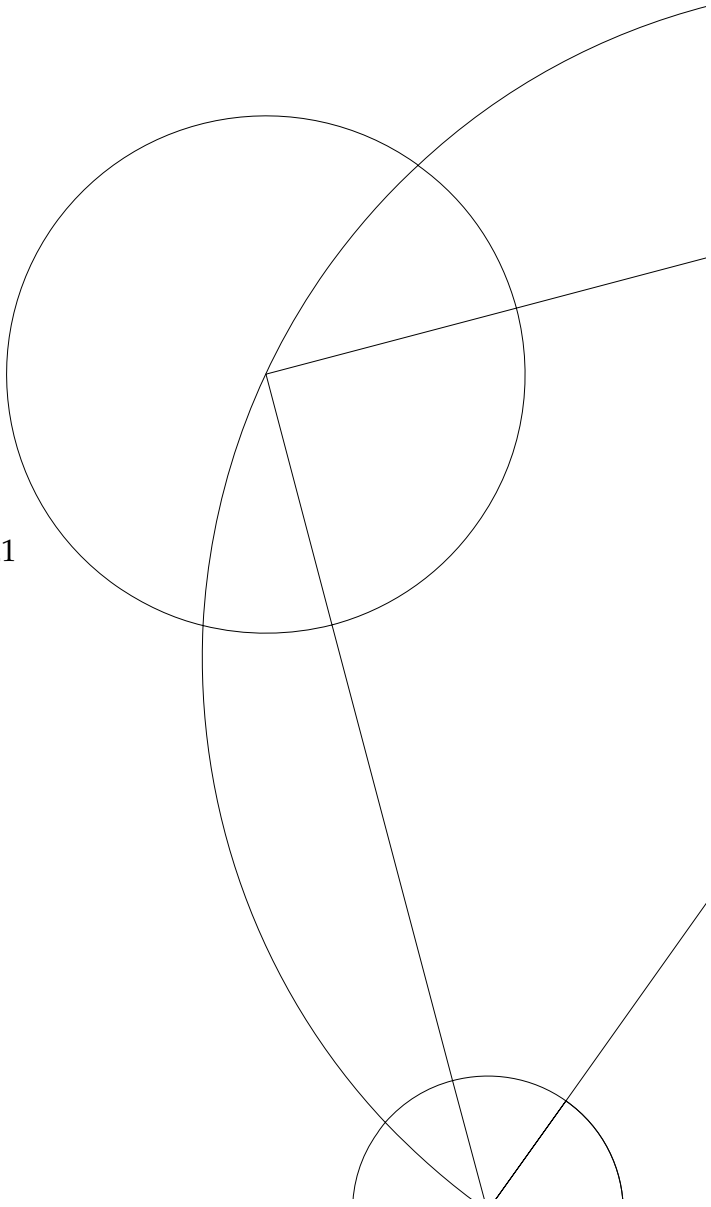
UNIVERSITY OF COPENHAGEN

UNIVERSITY OF COPENHAGEN

MSc THESIS

Galaxy evolution studies with the new, state of the art, multi-wavelength catalog in the COSMOS field

Author: ***Athanasios*** **ANASTASIOU**
Supervisors: *Georgios* MAGDIS, *Iary* DAVIDZON



January 15, 2021

UNIVERSITY OF COPENHAGEN

Abstract

Faculty of Sciences
Department of Physics

Master of Science

Galaxy evolution studies with the new, state of the art, multi-wavelength catalog in the COSMOS field

by Athanasios ANASTASIOU

Galaxy formation and evolution is a field driven forward by large multiwavelength surveys. Through the acquired data, we managed to improve our models of galaxy evolution. However, there are pieces of observational evidence still in tension with theoretical models. One of these is the observation of massive quenched galaxies being at place already at $2 < z < 3$, which is in contrast with the hierarchical paradigm of dark halo formation. In this thesis I will present the new COSMOS 2020 catalog, a deep panchromatic survey with updated imaging and spectroscopic data over $2\,\mathrm{deg}^2$ of sky.

New deeper data in UV, optical and near-infrared are added from Canada-France-Hawaii Telescope, Subaru/Hyper Supreme Cam DR2, UVISTA DR4 alongside all legacy surveys conducted on COSMOS field with *Spitzer Space Telescope*. A novelty introduced in the data release by the COSMOS collaboration is the implementation of two different photometric extraction methods to make two versions of the same catalog: one with traditional aperture photometry (`SExtractor`) and other with forced model photometry model (`Farmer`). The new catalog is reaching a limiting depth of $K_s = 25.2\,\mathrm{mag}$ (3", 3σ) which is $0.5\,\mathrm{mag}$ deeper than the previous version of the catalog (COSMOS 2015). Multiple number counts comparisons are presented here, to show the agreement with the literature as well as the progress in terms of depth (although I found that Farmer is missing *i*-bright sources). The derived photometric redshifts are compared with a large spectroscopic sample from different surveys. Catastrophic outliers constitute only 3.1% of the sample while the photometric precision is $\sigma = 0.011$. A separation of stars and galaxies is done based on the basis of the gzK_s diagram and the χ^2 comparison between galactic and stellar template fit. With the help of $NUVrJ$ diagram and the SFR-M_* correlation formula I selected quiescent galaxy candidates. The final sample is cross-matched with confirmed quiescent galaxies from the literature. A spectroscopic follow-up of the candidates in the future can help investigate the open questions of mass assembly and star formation quenching.

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Motivation

1.1 Galaxy formation and evolution

According to the Planck Collaboration et al. (2018) dark matter constitutes more than 80% of the Universe's total matter content. In the beginning gas and dark matter were mixed. But as the Universe expanded the gas was able to cool down and move to the center of dark matter halos. Halos are for dark matter the equivalent arrangement of what galaxies are for luminous matter. After the gas collapses the seeds for galaxies and stars are created.

Galaxy formation and evolution is a topic studied by astronomers for approximately a century. Great efforts have been made in the past to understand the physical processes that dominate the life of galaxies. A huge step that revolutionized the field was the introduction of large multiwavelength surveys. Ambitious surveys like Two-Micron All Sky Survey (2MASS)[1] mapped the whole sky in the infrared (IR) and created a catalog of 500 million objects. The 2dF Galaxy Redshift Survey (2dFGRS)[2] studied the large-scale structure of the Universe and was the largest redshift survey of its time. That was until the launch of the Sloan Digital Sky Survey (SDSS)[3], a multi-spectral and imaging redshift survey that conducted an even wider survey in the local Universe with multiple interlocking science goals over the years. Later surveys such as the Cosmic Evolution Survey (COSMOS)[4], the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS)[5] and the Great Observatories Origins Deep Survey (GOODS)[6] have opened up great potential for research. By acquiring photometric data in a plethora of different bands spanning from X-rays to sub millimeters for a collection of sources, researchers are able to build catalogs of astronomical sources thus creating statistical samples of galaxies.

These statistical samples need to have specific characteristics in order to be reliable and useful for scientific analysis. First of all, it is crucial to have a large number of galaxies in the catalog. The reason for this is to reduce the statistical uncertainties in our calculations induced by the sample size. Furthermore, a large sample ensures that all different types of galaxies are included in order for astronomers to study them. The best-known classification for galaxies is the morphology classification done by Hubble (1926). In this scheme galaxies are categorized based on their shape and structure through optical observations. The classes are elliptical, spiral, irregular and lenticular. An elliptical galaxy approximates an ellipsoid with smooth light

[1] https://irsa.ipac.caltech.edu/Missions/2mass.html
[2] http://www.2dfgrs.net/
[3] https://www.sdss.org/
[4] http://cosmos.astro.caltech.edu/
[5] http://arcoiris.ucolick.org/candels/
[6] https://www.stsci.edu/science/goods/

distribution, while a spiral galaxy have a disk structure with spiral arms and a central bulge. Lenticulars are considered a transition between ellipticals and spirals and lastly, irregulars are galaxies whose shape present weak or no structure (Schneider, 2006). But as any statistical sample, a catalog must not be biased. It has been found that the Universe is not homogeneous in scales < 1 Gpc (Davis et al., 1985). As a result any number count or density measurement are subject to uncertainty (cosmic variance) due to large-scale density fluctuations. To minimize the effects of cosmic variance, a survey should probe a volume much larger than the typical clustering scale of the observed sources (Moster et al., 2011). Having said that, an average galaxy evolves with timescales from a few million to billion years. This makes impossible, considering the human lifespan, to observe and study a single galaxy from the first moments of creation until its death. Thankfully the light from galaxies travels towards us with a finite speed so when it reaches our telescope we see the galaxy as it was when the light was emitted. This time between the emission and the detection of the light is called lookback time. So capturing light from galaxies from various lookback times allows researchers to link different "cosmic snapshots" of the Universe to gain insight on the evolutionary paths of galaxies. However, the further a galaxy is the dimmer it's light is when it reaches Earth. That is why we need our surveys to be "deep" in order to observe distant faint galaxies. This is achieved with a combination of technical aspects (e.g. high-sensitivity cameras) and survey strategy (e.g. long exposure times).

The introduction of multiwavelength surveys helped the astronomical community reach many decisive points in the path of understanding galaxy formation and evolution. By introducing color selection criteria it was discovered that galaxies follow a bimodal distribution. Apart from the morphological classification mentioned earlier galaxies can be categorized in two classes based on their colors, one including red quiescent galaxies and the other blue star-forming galaxies (Baldry et al., 2004). Quiescent galaxies are objects that appear red in color, have little to no gas left and stopped forming stars. On the other hand star-forming galaxies appear in bluer colors and have significant star formation activity. Since galaxies can be thought as conglomerates of billions of stars we can study their evolution by looking at their star formation history (SFH) and the build-up of the stellar mass (M_*). In addition the stellar mass has a direct connection with the dark matter halo that the galaxy resides in so it can also be used to examine dark matter properties and clustering (Wechsler and Tinker, 2018). We managed to study the evolution of stellar mass by focusing on the rest-frame optical light that originates from the low-mass stars that dominate the mass among the stars (Dickinson et al., 2003). Observing the rest-frame UV-continuum light that is dominated by short-lived massive stars enabled us to study star formation over a broad redshift range only to find that the star formation rate (SFR) -the total mass of stars formed per year- was higher in the past as shown in Fig. 1.1. Both star formation rate and stellar mass are rest-frame quantities, which would be impossible to estimate without knowing the distance, thus the redshift, of the given galaxy. Similarly to Doppler's effect, redshift causes an increase in the wavelength of a signal emitted by a galaxy with a receding velocity with respect to the observer. According to Hubble's law (Hubble, 1929) the receding velocity of an object due to the expansion of the universe is proportional to its distance. Redshift can be measured with two different methods. Spectroscopic redshift (spec-z) is calculated by measuring the shift of a known spectral feature in the spectrum of a galaxy. This reveals the receding velocity and in turn the distance. Photometric redshift (photo-z) is calculated using special software that generates the best spectrum fit to the inputted photometric fluxes of given bands. Even though the uncertainty of

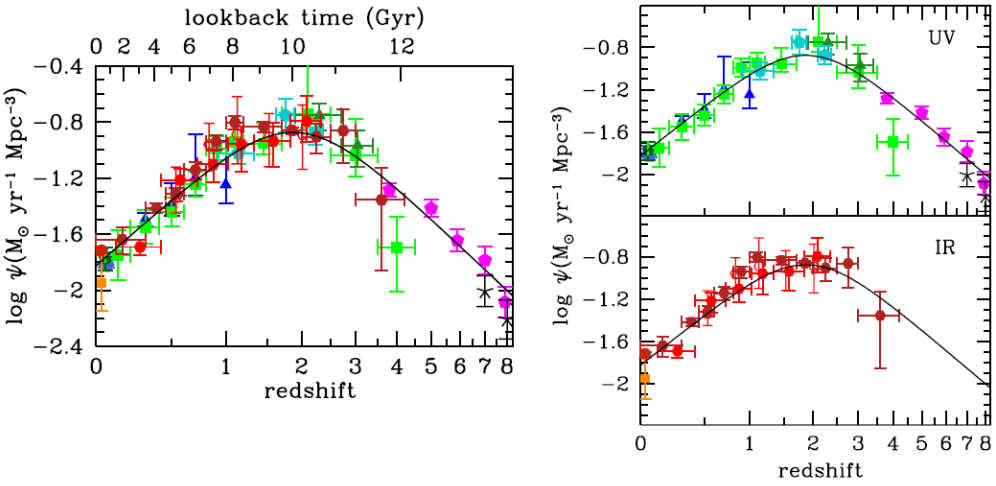


FIGURE 1.1: The cosmic star formation history inferred from FUV, (top right panel) IR, (bottom right panel) and FUV+IR rest-frame measurements. It is clear from the diagram that the star formation rate density (SFRD) peaked at around $z \approx 2$. Image taken from (Madau and Dickinson, 2014)

the photometric method can be as high as $\delta z = 0.5$ (Bolzonella, Miralles, and Pello', 2000) its advantage lies in the ability to estimate the redshifts of millions of sources in a quick and computationally inexpensive way. On the other hand spectrographs give an absolute value of the redshift with negligible systematic uncertainties but cannot target more than 50-100 targets per pointing. Many different surveys using either methods found numerous galaxies at high redshifts with the most distant one reaching $z = 11.1$ discovered by Oesch et al. (2016) which is just 400 Myr after the Big Bang.

Despite all these achievements, our knowledge of galaxy formation and evolution is still incomplete. The current most-accepted cosmological model is the Λ-Cold Dark Matter model (Λ-CDM). Simulations based on this model had relatively great success in reproducing the observable Universe (Springel et al., 2005; Springel, 2005; Schaye et al., 2015). However these simulations also revealed a tension between theory and observations. Why do we see massive galaxies with old stars while Λ-CDM models predict a younger star population in those galaxies (Thomas et al., 2005)? Moreover, why do we observe high-z massive galaxies that contradict our theoretical paradigm about hierarchical formation (Daddi et al., 2005; Toft et al., 2007)?

1.2 Stellar mass assembly and Downsizing

Numerous works have studied the evolution of stellar mass assembly in the local Universe (Bolzonella et al., 2010; Bundy et al., 2006; Pozzetti et al., 2010) as well as at high redshifts $z \geq 2$ (Daddi et al., 2007; Pérez-González et al., 2008; Ilbert, O. et al., 2013; Laigle et al., 2016).All the aforementioned studies found an interesting feature manifested and confirmed in different ways. This feature is the decline of the formation of massive galaxies ($M_* > 10^{10} M_\odot$) with cosmic time, a trend called "cosmic downsizing" (Cowie et al., 1996). In other words the most massive galaxies have already been formed by $z \approx 3$, while those at lower masses formed later.

This is in disagreement with the current Λ-CDM model, which predicts that galaxies and dark matter halos grow through collision and aggregation of different galaxies- a process called merging- in a hierarchical formation scenario (Primack, 2015). But recent works have shown otherwise. Schreiber, C. et al. (2015) run an extended census on the SFRs on all masses between $0 < z < 4$ and showed that star-forming galaxies follow a universal scaling law termed as main sequence of galaxies (Fig 1.2). The most interesting finding of the work though is the tight correlation between SFR-M_* with no clear redshift dependence. This feature reveals that more than two-thirds of the stars in the Universe were created in main-sequence galaxies and that mergers are not a dominant event of star formation. This is also one of the features that cosmological simulations based on the Λ-CDM model cannot reproduce in early cosmic time, despite their overall success, without the incorporation of feedback mechanisms. These mechanisms are the prevention of cooling flows due to accretion in Active Galactic Nuclei (AGN) (Bower et al., 2006) and the reheating of the disk gas from stellar winds and supernovae (SNe) (Stringer et al., 2009). A further description of the mechanisms follows later on.

There are different kinds of observational evidence that seem to confirm this downsizing trend. Thomas et al. (2005) and Gallazzi et al. (2005) showed that in massive galaxies reside older stellar populations than in galaxies with lower stellar mass. Bundy et al., 2006; Brinchmann et al., 2004 confirmed that star-forming galaxies have lower masses in low redshift compared to the star-forming in high redshift. Finally studying the evolution of the galaxy stellar mass function Bundy et al. (2006),Pozzetti, L. et al. (2007) and Pérez-González et al. (2008) have discovered that the high-mass end evolves very slowly compared to the low-mass end meaning that massive galaxies formed much earlier than the low-mass galaxies. Despite that, there are disagreements suggesting that the discrepancy between models and observations on cosmic downsizing is based on mistreatment of physical processes and misaccounts of errors. More specifically, Fontanot et al. (2009) points out that when stellar mass errors are accounted for then the discrepancies in the number densities of massive galaxies between predictions and observations weaken or disappear. Furthermore, he suggests that the implementation of feedback mechanisms such as AGN,SNe and stellar winds are still crude.

However galaxies are not isolated systems unaffected by external interactions. Observational evidence imply that there is connection between galactic internal properties (SFR, M_*) and the surrounding environment. A striking connection appeared when it was found that most of the elliptical galaxies reside in the cores of clusters while the disk star-forming galaxies are most frequently located in the outskirts or in the cosmological field ((Postman and Geller, 1984; Dressler, 1980)). After this "morphology-density" relation, later works found other types of correlations. Balogh et al. (2004) showed that there is connection between the observed colors and the environment. Galaxies sitting in dense environments are red in their majority. Kauffmann et al. (2004) found that there are more properties and features that correlate with the environment in the local Universe ($z \sim 0.1$). Stellar mass distribution grows by a factor of two towards higher masses when shifting from a low to high density region. At fixed stellar mass SFR is lower near the center of clusters while double the number of galaxies host AGN in the general plane than in clusters. On the other hand Elbaz et al. (2007) shows that the connection is reversed at $z \sim 1$; galaxies in more dense environments present higher values of SFR than those in less dense environments.

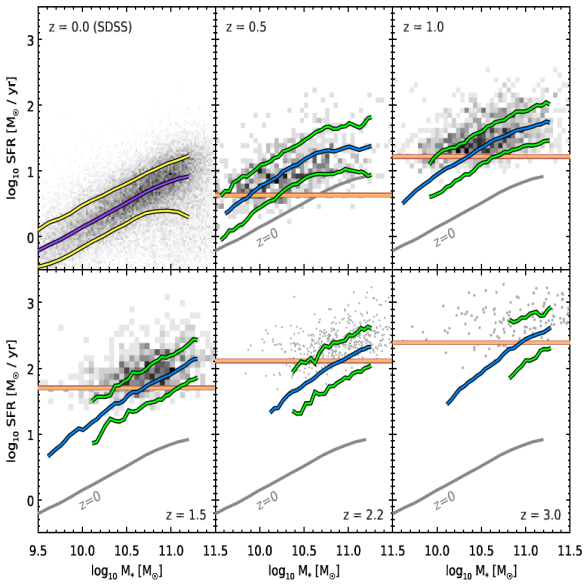


FIGURE 1.2: The correlation of SFR-M_* in different redshift bins. The blue line is the averaged SFR of the sample while the green line is the 1σ dispersion. Orange line marks the SFR detection limit. Image taken from (Schreiber, C. et al., 2015)

1.3 Star formation and quenching

Another important point where there is tension between theory and observations is the so-called "quenching" mechanism. The term quenching is used to describe either the termination of any star formation activity or the conservation of the quiescent state for the rest of the galaxy's lifetime (Man and Belli, 2018). A detailed description of this mechanism, which may involve multiple physical processes, has not been achieved yet. There is evidence that the suppression of star formation is more efficient in massive galaxies with a quenching mass threshold of $M_Q \propto (1+z)^{3.5}$ above which galaxies are quenched (Bundy et al., 2006). Possible mechanisms responsible for quenching are described below. A quick summary is given in Fig1.3.

- A reduction in the dark matter accretion onto the galaxy/halo system from cosmological filaments may lead to "cosmological starvation". Feldmann and Mayer (2015) shows that dark matter growth for a set of galaxy radii stops at about $z \sim 3.5$ affecting the accretion rate of cool gas as well. After that point star formation slows down significantly. However stellar winds and SNe can still replenish the gas reservoir of a galaxy therefore an additional quenching mechanism (see below) is needed to conserve the quenching.

- Inefficiency of gas cooling can also cause the termination of star formation leading to quenching. According to theory, gas collapses into the dark matter halo and is heated from virial shocks (Rees and Ostriker, 1977); after that the gas should cool radiatively. Bower et al. (2006) and Croton et al. (2006) showed that for massive galaxies ($\geq 10^{12} M_\odot$) the presence of a slow accreting supermassive black hole can effectively cut the cooling flows of gas and prevent it from collapsing and forming stars. Their simulations based on the hierarchical Λ-CDM model incorporated this feedback and reproduced a very well match with observations especially on the distribution of massive galaxies at high

redshifts. Other potential gas heating mechanisms are gas clump or satellite galaxies falling into massive halos ($(\geq 10^{13} M_{\odot})$. In dense environments when satellite galaxies are infalling to the center of the cluster they experience shock heating. If the energy deposited on the gas is enough to bring the temperature over the virial value then the gas is efficiently removed (Metzler and Evrard, 1994). Moreover the satellite galaxies are subjects to ram pressure stripping. As they move within the clusters they experience a "wind" caused by the intracluster medium that can effectively heat or even remove the gas. Khochfar and Ostriker (2008) took into account these scenarios and suggested that the gravitational potential energy gained by the infalling galaxies can be transferred as heat to the gas thus halting star formation. Finally, energy release via stellar winds, type-Ia supernovae and thermally pulsating asymptotic giant branch stars can also significantly heat the gas (Ciotti et al., 1991).

- Star formation efficiency may also decrease as a result of various reasons. In order for gas to form stars it must lose it's kinetic energy. Several mechanisms can inject turbulence to the gas and prevent it from further collapse. The stellar bulge can stop the fragmentation of the gas disk through shear. This process is called morphological quenching (Martig et al., 2009). Other possible sources of turbulence can be; the formation and presence of stellar bar (Khoperskov et al., 2018) which collects the gas in the bar region and through shear injects turbulence to gas; low-power AGN can induce random gas motions through their molecular outflow (Alatalo et al., 2015) and lastly magnetic fields can prevent the gravitational collapse of molecular clouds (Tabatabaei et al. (2018)).

- An intense burst of star formation may cause rapid consumption of gas. In the case where the gas is depleted faster than it gets replenished then the galaxy will be quenched after a short period of intense star formation. This can be caused by abrupt compression of gas suggested to be triggered by disk instabilities (Zolotov et al., 2015) or positive AGN feedback(Ishibashi and Fabian, 2012).

- Quenching may be also caused by AGN feedback triggered by merging. Simulations suggest that galaxy mergers trigger starburst activity due to the gas inflows (Mihos and Hernquist, 1996). These inflows feed the central black hole through accretion and provide feedback energy heating the circumnuclear gas. The feedback is strong enough to expel all the gas from the center of the galaxy shutting down both black hole accretion and star formation (Di Matteo, Springel, and Hernquist, 2005).

1.3.1 Summary and next steps

The connection between AGN and galaxies is critical to the understanding of quenching. Furthermore looking at the gas properties of recently quenched galaxies is crucial. Estimating the cold gas reservoir or the hot gas properties in high-z quenched galaxies will give valuable answers. Also evaluating the importance of gas outflows from observations will be enlightening. Lastly, the study of the stellar population and its history through spectroscopic observations can put constraints on the parameters involved in quenching.

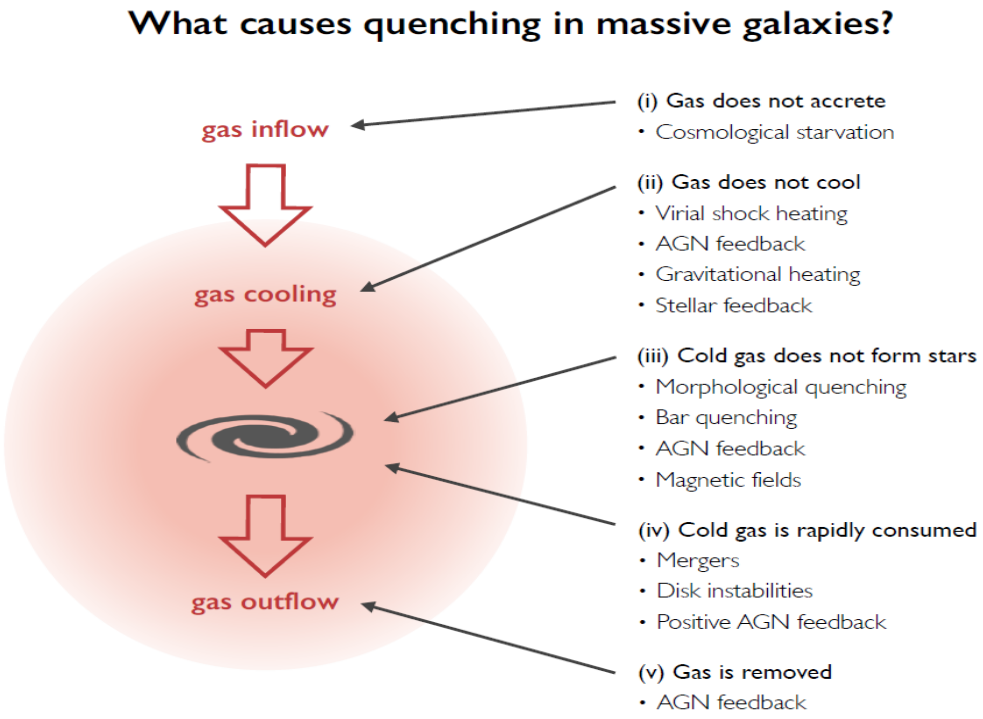


FIGURE 1.3: Schematic representation of the possible mechanisms responsible for quenching. Image taken from (Man and Belli, 2018).

1.4 Motivation

In this work I use the new COSMOS 2020 catalog and try to answer the open questions of galaxy formation and evolution. As mentioned above, large deep astronomical surveys are one of the most important tools to probe the Universe. The COSMOS project has been one of the pioneering project in the study of galaxies over a broad range of redshift. The new catalog offers images from UV to mid-IR with deeper depths in the majority of the bands than the other previous versions. Its wide angle coverage ensures the compilation of a large, deep and non-biased source sample. With this sample we aim at selecting high-z massive quiescent galaxy candidates. The spectroscopic follow-up observations of the best candidates, will enable the thorough study of their properties and gain insight on the cosmic downsizing as well as the early quenching of star formation. The structure of this thesis is described in the following points.

- In Chapter 2 I give an overview of the COSMOS 2020 design and present the set of filters and the new data collected from all the different telescopes of the survey. A full description of the photometric extraction software is given.
- In Chapter 3 I describe the methods and their theoretical background that are used to validate the catalog as well as compare with other works as a verification tool. Various comparisons are made with other surveys to ensure the advanced possibilities of the new catalog.
- The analysis of the data is described in Chapter 4. The strategy design of source classification is presented and explained. Along with that the tools of successfully selecting high-z quiescent galaxy candidates are presented. The final sample of candidates is cross-matched with spectroscopically confirmed galaxies

in the field.

- The results of the data analysis are presented and discussed in Chapter 5. I conclude this work with a summary of the findings as well as the strategy for the future follow-up research plans.

A standard Λ-CDM cosmology is used with $H_0 = 70\ km^{-1}Mpc^{-1}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$. All magnitudes are in the AB system (Oke, 1974).

Chapter 2

Data

2.1 Overview

The COSMOS field is a 2 deg^2 area centered at RA = 10:00:28.6 DEC = +02:12:21.0 (Scoville et al., 2007, J2000). It is an area with relatively low number of bright stars on the field compared with the rest of the celestial sphere. Furthermore it is devoid of bright sources in X-ray, UV, and radio wavelengths along with a low and uniform Galactic extinction of $< E_{(B-V)} > \simeq 0.02$ mag. In this way the sources of interest (in our case, galaxies at $z > 3$) have a very low level of contamination from nearby objects. Another important characteristic is the location of the field: COSMOS is strategically selected – near the celestial equator – in order to be accessible by astronomical facilities from both hemispheres.

Throughout the years many different multiwavelength surveys were conducted in the field (e.g., Capak et al., 2007; Ilbert et al., 2008; McCracken et al., 2010; McCracken et al., 2012; Laigle et al., 2016). Furthermore, the COSMOS field covers a large area that drastically reduces cosmic variance uncertainties (see Fig. 2.1) especially at high redshift. For comparison, the areas of the fields presented in the Figure are 0.003 deg^2, 0.08 deg^2, 0.22 deg^2 and 0.19 deg^2 for UDF, GOODS, GEMS, and EGS respectively. Out of all the fields presented in the graph, COSMOS is the one with the least cosmic variance error.

The previous version of the COSMOS catalog was COSMOS 2015, presented in Laigle et al. (2016). The catalog offered deep (25–26 mag) multiwavelength data in the range from UV to mid-IR (i.e., 0.25μm–8μm). In the K_s band the limiting depth was 24.7 mag at 3σ in the deepest regions. In other words for a source that has $K_s = 24.7$ mag the signal is expected to be 3 times the value of the background noise. The method used to calculate the depth in COSMOS2015 was by measuring the sky background in 3" empty apertures. Note that multiwavelength surveys before COSMOS 2015 had significantly fainter limiting depths. For example the VISTA Deep Extragalactic Observations (VIDEO) was also a survey in optical and near-IR (NIR) data and designed to study the galaxy structure and evolution as a function of epoch and environment. The limiting depth of VIDEO was $K_s = 23.8$ mag at 5σ over the whole 12 deg^2 covered by the survey (Jarvis et al., 2012). Another example is the VIMOS-VLT Deep Survey (VVDS), which was a purely-magnitude limited spectroscopic survey studying galaxy evolution throughout cosmic time and had a depth of $i = 24.7$ mag at 5σ limit (Le Fèvre et al., 2013); this is 1.5 magnitudes shallower than COSMOS 2015, where the i-band depth was 26.2 mag. One can also compare to even shallower surveys probing the local Universe, e.g. the maximum depth of SDSS was $r = 22.7$ mag[1] at $z \simeq 0.1$ while in Laigle et al. it is $r = 26.5$ mag. It is also interesting to mention the characteristics of the first generation of "deep" extragalactic

[1] https://www.sdss.org/dr14/imaging/other_info/

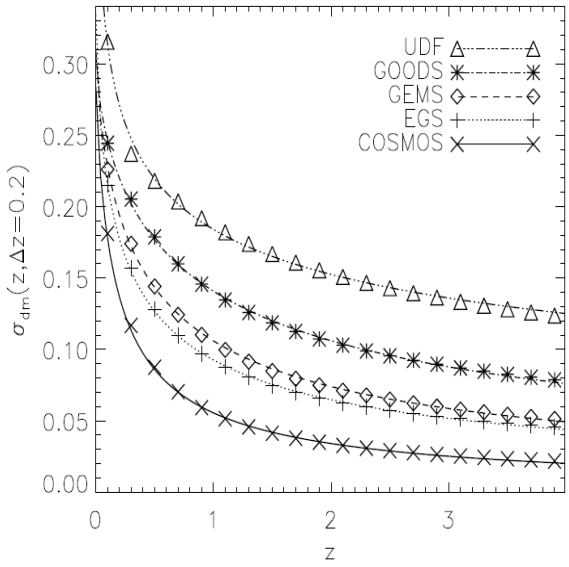


FIGURE 2.1: The dependence of cosmic variance on mean redshift with reference bin sizes of $\Delta z_{ref} = 0.2$. The fields presented here are Hubble Ultra Deep Field (UDF), GOODS, GEMS (Galaxy Evolution from Morphology and SEDs), EGS (Extended Groth Strip) and COSMOS. Figure taken from (Moster et al., 2011)

surveys, like Classifying Objects by Medium-Band Observations in 17 Filters (Wolf et al., 2003, COMBO-17), which were able to probe the "distant" universe up to $z \sim 2$. COMBO-17 studied the luminosity distribution at $0.2 < z < 1.2$ of a galaxy sample limited at $R \leq 24$ mag. This comparison emphasizes the dramatic improvement made by Laigle et al. (2016).

Nevertheless, there were also surveys that reached deeper limiting magnitudes than COSMOS 2015. Fontana et al. (2014) carried out Hawk-I UDS and GOODS Survey (HUGS), a NIR imaging survey in the UKIRT Infrared Deep Sky Survey- Ultra Deep Survey (UKIDSS-UDS) field and the Great Origins Observatories Deep Survey-South (GOODS-South) field. The data consist of deep high-quality imaging in the K and Y bands, executed with the High Acuity Wide field K-band Imager camera (HAWK-I) at the European South Observatory-Very Large Telescope (ESO-VLT). At 1σ limit per arcsec^2 the survey results into a K-band limiting depth of 28.0 mag and 27.3 mag in GOODS-S and UDS respectively.

Another survey that went deeper than COSMOS 2015 was CANDELS. The observations executed with the *Hubble* Space Telescope/Wide Field Camera 3 (HST/WFC3) in the Ultra Deep Survey (UDS) field , reached a limiting depth (as a function of position) of 28.26 mag at 1σ for the majority of the area (Galametz et al., 2013). CANDELS also covered the COSMOS field in Nayyeri et al. (2017). Although released shortly after Laigle et al. (2016), it would be instructive to compare with Nayyeri et al. since that study is still state of the art in terms of HST observations, with F160W band (15369 Å) reaching a depth of 27.56 mag at 5σ. In their catalog, Nayyeri et al. included also the ground-based observations from Laigle et al. (2016): this shows the importance of ground-based data, which cannot be disregarded even when space-based data are available.

Given the unprecedented depths, one can argue that CANDELS is more efficient in searching for high-z faint galaxies than COSMOS 2015. However the area covered

by those HST observations ($\sim 216\,\mathrm{arcmin}^2$ in the COSMOS field, $0.2\,\mathrm{deg}^2$ in total) is small enough to be susceptible to cosmic variance uncertainties. CANDELS covers only 3% of the whole COSMOS field, thus cosmic variance dominates the uncertainties. Therefore, COSMOS 2015 may be shallower but it is statistically more robust.

One of the main features of COSMOS 2015 was the inclusion of new data sets, namely the IR data from UVISTA-Data Release 2 (UVISTA-DR2, McCracken et al., 2012), the Spitzer Large Area Survey with Hyper-Suprime-Cam (SPLASH, Steinhardt et al., 2014), and the Y-band from HSC. In UVISTA-DR2 there were two regions of the COSMOS field with different exposure: a "deep" region covering $1.5\,\mathrm{deg}^2$ of the field (limiting depth of $K_s = 24.0$ mag at 3σ in 3" aperture) and four "ultra-deep" stripes with significantly longer exposure ($K_s = 24.7$ mag at 3σ in 3" aperture). The two UVISTA regions remained even after DR2 and they can be seen in Fig. 2.2. Approximately 6×10^5 objects were included in the $1.5\,\mathrm{deg}^2$ area covered by UVISTA deep exposure while 1.5×10^5 of them reside in the ultra-deep stripes ($0.62\,\mathrm{deg}^2$). A more detailed comparison between COSMOS 2020, COSMOS 2015, and the rest of the literature will be provided in Chapter 3 and 4.

The present catalog contains photometric data from: the Hyper Supreme Cam (HSC), installed on the Subaru 8.3m telescope[2] on Mauna Kea (Hawaii); the UltraVISTA survey (UVISTA) from the Visible and Infrared Survey Telescope for Astronomy 4.1m telescope[3] (VISTA) on Paranal Observatory(Chile), the Canada-France-Hawaii 3.58m telescope (CFHT)[4] also on Mauna Kea; and from the Infrared Camera Array (IRAC) on board the *Spitzer* Space Telescope[5].

Compared to the previous version of the catalog, COSMOS 2020 includes deeper data in all bands except IRAC channel 3 and 4 because those channels are not operational any longer. In fact, in May 2009 the cryogen that was on board *Spitzer* to keep instruments cool ran out[6]. Without the cryogen the thermal noise produced by the telescope itself made it impossible for certain instruments to acquire reliable data. Among them there were channel 3 and 4, while channel 1 and 2 were less affected and continued to acquire data (although with higher level of noise).

The footprint of the various telescopes mentioned here is shown in Fig. 2.2. Note that UVISTA (now at the fourth data release) still has two different exposures. However, the difference in depth between the two areas is smaller at least in the K_s band , thanks to the observations carried out in the "deep" area during the last few years. The footprint of the different telescopes is not the same: there are patches that are not covered in some telescopes (see Fig. 2.2) but these are mostly at the edges of the field. Therefore, most of the scientific analysis is focused on the central area of COSMOS to avoid inhomogeneity in the final catalog. Moreover, there can be unreliable pixels in the image because of bad photometry (e.g. saturated stars, tracks of artificial satellites, technical failures on the camera's sensors).

To visualize the wavelength coverage and the way it is sampled, I plot in Fig. 2.3 the transmission curves of the various filters. The filters are the spectral "windows" that determine the wavelength band that will pass to the detector. These curves include the different effects of atmospheric transmission, instrument transmission, filter transmission, and detector's quantum efficiency. As a reference, I have included in the Figure the spectrum of a quiescent galaxy at $z = 2$ which shows the

[2] https://subarutelescope.org/en/
[3] http://www.vista.ac.uk/
[4] https://www.cfht.hawaii.edu/
[5] http://www.spitzer.caltech.edu/
[6] https://www.nasa.gov/mission_pages/spitzer/news/spitzer-20090506.html

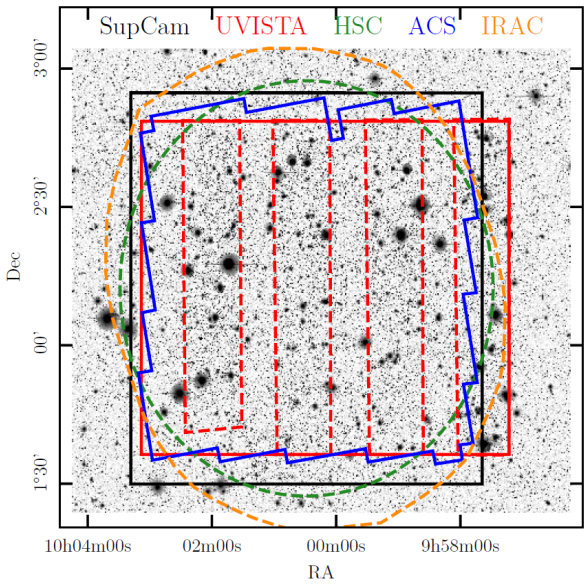


FIGURE 2.2: Schematic of the coverage on the COSMOS field. The background is the stacked detection image from $izYJHK_s$ bands. The solid lines mark the survey limits while the dashed lines represent the regions with the deeper exposures. On the top on the image, the instruments/surveys are presented. The Advanced Camera for Surveys (ACS) from Hubble Space Telescope (HST) is presented here as a reference. Image taken from Weaver et. al., (in prep)

emission of the galaxy as a function of wavelength. The spectrum is synthetic and is generated by a software (Bruzual and Charlot, 2003) that incorporates our current theoretical knowledge about galaxy and stellar evolution. This is called stellar population synthesis model and will be described in Chapter 3. The Figure shows that the set of filters of COSMOS 2020 samples the galaxy spectrum with relatively small gaps (with the exception of the gap between 2.5 and 3 μm). The COSMOS team used also narrow and medium band filters, which sample galaxy spectra more accurately. However, they need a longer exposure time and also cover a narrower wavelength range. Thus, it is of pivotal importance to include as many different broad-band filters as possible in a photometric survey. This is also the reason to keep filters at overlapping wavelengths: there is no drawback in having more information since it translates to more data points. This is true of course only if the additional data points do not introduce serious systematic uncertainties.

In this work I will not use data from the narrow and medium band filters, therefore they are not described here. In Chapter 3 I will show that the data from these filters can be used to estimate important information such as redshift and galaxy stellar mass but other than that, they are not needed in any of the analysis done in this thesis. In Fig 2.4 the depths of the COSMOS 2020 filters are shown along the full spectrum. Only the filters of the updated data sets are presented, plus IRAC channel 3 and 4. In Table 2.1 all the details of the filters are presented. Apart from the new data, a set of archival data from older surveys is also used for the derivation of photo-z and are presented in Table 2.1. A full description of the data sets follows below.

2.2 Data sets

2.2.1 UV and Optical

The UV regime is covered by the two u bands from the CFHT Large Area U-band Deep Survey (CLAUDS, Sawicki et al., 2019). One is the old u^* band that was also included in the previous COSMOS catalog (Laigle et al., 2016) while the newly-installed u filter is used for the first time here. They are centered at 3858 Å and 3709 Å reaching a 27.2 mag and 27.1 mag at 3σ limit respectively. Along with the new data, the previous data from the Galaxy Evolution Explorer filters are included in the analysis (FUV and NUV, Zamojski et al., 2007). Capturing UV light is critical as it originates from young high-mass stars that dominate the emission of the galaxy and reveals star formation activity. Unfortunately in reality the dust content of the galaxies is enough to greatly redden UV light since dust grains preferably absorb ultraviolet photons and re-emit them in IR. Moreover UV bands capture galaxy emission only in the low-z Universe since at high-z the spectrum gets shifted out of the filter's band-pass. However, if a source is not detected in UV then this makes it a candidate for a $z > 2$ object. Consequently, this component is a good first indicator for selecting high-z sources.

From HSC we include the data from the latest public data release (UVISTA-DR4) (Aihara et al., 2019). The g, r, i, z, y optical bands range between 4200–10000Å and have limiting depths that are remarkably deeper compared with previous studies. It is also noteworthy that while in COSMOS 2015 the data were from the Suprime-Cam instrument, in COSMOS 2020 the data comes from the superior Hyper Suprime-Cam. This fairly new instrument provides advanced capabilities to the user compared with the old Suprime-Cam. The 104 main charged coupled device cameras (CCD) cover a 1.5 deg field of view with upgraded sensitivity on all the filters. Furthermore, the quantum efficiency reaches an impressive 94% at 650nm (Miyazaki et al., 2017). The quantum efficiency of a CCD is the measure of the ability of a detector to absorb and record incoming photons. In comparison with Laigle et al. (2016) and the similar bands from Suprime-Cam used there, the limiting magnitudes are $\sim$0.5–1.0 mag deeper in this work.

2.2.2 Near-Infrared

The data of Y, J, H, K_s bands from the VISTA telescope are utilized here to cover the NIR part of the spectrum (McCracken et al., 2012). The images were taken from December 2009 to June 2016 with the VIRCAM instrument and were released in the Ultra-VISTA DR4. There is a notable increase in the exposure times of all ultra-deep exposures, as well as the deep exposure of K_s, compared with DR2 which was used in Laigle et al. (2016) as seen in Table 2.2. This results into $\sim$ 0.5 mag deeper images compared to DR2.

The K_s band is one of the most important in the survey. The rest-frame optical and NIR light is a very good proxy for galaxy stellar mass. That is because it is emitted by the low-mass stars, which dominate the stellar mass of a galaxy (Gavazzi, Pierini, and Boselli, 1996). Up to $z \simeq 2$ the K_s band captures this rest-frame emission and greatly contributes to the calculation of the stellar mass. Indeed, Fontana et al. (2006) shows that at $z < 2$ mid-IR data from *Spitzer* are not required to derive the stellar masses and just K_s data are enough.

On top of that the use of K_s band bridges a very large spectral range between NIR and mid-IR which is crucial for the estimation of physical quantities at high redshift.

In fact at $z \leq 3.5$ Balmer break is shifted in the coverage of K_s which is critical for a reliable estimation of photo-z (Fontana et al., 2014).

2.2.3 Mid-Infrared

For this part of the spectrum data from the Spitzer/IRAC channels 1 to 4 is used (Moneti et al. in prep.). The novelty is that images from all the surveys ever done in COSMOS field by IRAC are co-added in a single image. By stacking images from the same band together one can achieve a great enhancement of the signal to noise ratio (S/N). According to Howell (2006) when the noise is dominated by the Poisson term then S/N grows as $\sqrt{N_*}$ where N_* is the total number of photons collected from the source. So by stacking images together astronomers can "smooth" the background noise and boost the signal. The surveys were the images were used for stacking are, *Spitzer*-COSMOS (Sanders et al., 2007, S-COSMOS), SEDS (Ashby et al., 2013), SPLASH (Steinhardt et al., 2014), the Spitzer-Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (Ashby et al., 2015, S-CANDELS) and the Spitzer Matching Survey of the inUltra-Vista ultra-deep stripes Survey (Ashby et al., 2018, SMUVS).

Spitzer's mid-IR data is a cornerstone for probing the high-z Universe. Through stellar mass estimates and number density measurements astronomers can put important constraints on current models. As mentioned above, Spitzer mid-IR data do not contribute significantly to the stellar mass estimation at $z < 2$, but at $z > 4$ they are essential. At these redshifts the rest-frame NIR light shifts towards mid-IR so the K_s band no longer captures this emission. Instead it captures the optical light which is dust sensitive and thus does not trace the stellar mass properly (Ashby et al., 2018). As a consequence it is crucial to include IRAC data in our analysis in order to derive accurate stellar mass estimations.

2.3 Photometry: Two different methods to measure it

All the data acquired have to follow multi-step processes in order to extract magnitudes from the sources. However, it is not in the purposes of this work to describe these standard processes (data reduction, frame combination and astrometric solution). For a detailed description i refer the reader to (Weaver et al., in prep.) and (Moneti et al., in prep.).

As surveys progress in reaching deeper and deeper depths, an issue is becoming more and more and pressing. The phenomenon of blended sources are a common problem. Even if a pair of sources is millions of light years apart, they would appear to overlap on the image if their line of sights with respect to us are very close. Boucaud et al. (2020) point out that as the survey depth increases, the number of sources that contaminate or overlap goes up by 50%. The majority of astronomical tools cannot account for the flux of overlapping sources in a statistically justified method nor diagnose if a source is blended with another. In the case of S/N measurements by a software, a blended source can be very problematic.

Furthermore, traditional methods fail to extract certain pieces of information available in the data. The most common method–aperture photometry–uses fixed aperture sizes to measure the flux of the objects. This done by extracting a point spread function (PSF) which is the 3-dimensional response function of an imaging system to a point source. However the PSF differs significantly with each band due to the characteristics of each wavelength. For example the same object in u can have

a 0.5" full width at half maximum (FWHM) of the PSF while in K_s as large 1". In order to smooth these variations so that there is no band-dependency on the flux, researchers homogenize all the different PSFs to a reference PSF. This step degrades information from high-resolution images and introduces systematic errors related with astrometric offsets and variable source size.

After painting this picture it is apparent that new methods and approaches are required to overcome these obstacles. Therefore in this work two different methods of photometric extraction are used on the same data in an effort to test new tools, upgrade our methods and acquire more robust data products.

2.3.1 `SExtractor`

The first catalog is compiled with the traditional method of aperture photometry using `SExtractor` (SE Bertin and Arnouts, 1996). This software starts by estimating the sky background. The background histogram of the image is clipped until it converges at $\pm 3\sigma$ around its median. Next is the detection of the sources on the image. In older literature this was performed on either a single-band image or on each image independently. The method used here is the one described in Szalay, Connolly, and Szokoly (1998). Szalay, Connolly, and Szokoly showed that the method of image stacking bypasses the problems of other approaches such as single bandpass detection or merging of independent bandpass detections. The single bandpass detection can miss objects that have unusual colors or objects close to the detection limit in other bands. On the other hand, independent detection in each band requires careful handling of objects that are detected in some of the bands but not in all of them. Therefore, a composite image of a selection of bands (in our case $izYJHK_s$) is built where each pixel value is a representation of a probability that this pixel is just sky background and not a real source. By utilizing this probability distribution (a χ^2 distribution) along with the distribution of pixels representing object flux one can create an image with only object-probable pixels. The formula used is:

$$\chi^2_{det} = \frac{1}{N} \sum_i \frac{(F_{im,i} - F_{back,i})^2}{\sigma^2_{rms,i}}$$

where F_{im} is the flux value of pixels, F_{back} is an estimate of the sky background flux, σ_{rms} is the RMS uncertainty at a given pixel, i is the index for the different filters used to create the detection image and N is the total number of co-added images.

The software uses a threshold and convolves the original image with a template frame. To sufficiently detect faint galactic sources the best template frame is the PSF. After detecting the sources, a routine that deblends the sources is executed. All the sources are re-thresholded to a 30 level exponentially-spaced profile with the bottom being the background threshold and the top the peak value. An algorithm starts downwards and checks whether it should extract one or two sources. Once finished, in the case of extracting two sources the "blended" pixels lying lower than the separation threshold have to be reallocated to the sources. This is performed by calculating the probability of each pixel belonging to either source. This probability is based on the contribution of flux made by the different sources. In the event of low threshold being adopted, a problem arises with spurious detections near the edges of the objects. Due to the fact that around sources the background noise is locally higher, it can cause the unwanted detection of noise peaks. To solve this situation the software incorporates a filter method that checks whether or not a source would have been detected if there were no close-by objects. Moving on, the user is now

ready to extract magnitudes from the sources. In this work we use 3 arcsec apertures to estimate the enclosed flux. Lastly, the software performs an estimation on the star/galaxy classification of each object based on various estimators (Sebok classifier (Sebok, 1979), Q classifier (Le Fevre et al., 1986),the "$-r^2$" moment (Kron, 1980) and ψ parameter (Maddox, Efstathiou, and Sutherland, 1990)) but all of them lack reliability when it comes to classification of blended objects (Bertin and Arnouts, 1996).

2.3.2 `Farmer`

The other catalog is compiled with the `The Tractor` software (Lang, Hogg, and Mykytyn, 2016). `Tractor` treats the extraction process as an optimization problem. It uses physically derived prescriptions to model each source. However the software only contains the suite of models and the way to optimize them. For the rest of the steps (source detection, deblending, selection etc) the COSMOS Team has developed `The Farmer`, a `python` routine to fill in the necessary requirements to build a large astronomical catalog.

In `Farmer` source detection is performed with the same method (Szalay, Connolly, and Szokoly, 1998) as described above in the `Sextractor`. After the sources are detected, a deblending process follows. A major difference with `Sextractor` is how the overlapping sources are handled. The software defines blobs that contain sources and models them simultaneously via a joint likelihood. Taking into account the number of sources in the blob gives the advantage of assessing the effect of neighboring sources on each other. Next the user is ready to "force" a model on each source selected from a set of five models similar to the ones available in SDSS (Stoughton et al., 2002). These are:

1. **Point Source**: A simple Delta function for stars and unresolved sources with a three-parameter PSF (two positional and one of the total integrated flux).
2. **Simple Galaxy**: A radially symmetric exponential profile of a fixed size is used for nearly resolved galaxies. The profile is described by $I(r) = I_0 exp(-1.68 r/r_e)$ where $I(r)$ is the flux surface density, r the distance from the center, I_0 the flux surface density at the center and r_e the effective radius enclosing half of the total flux. Same parameters as the Point Source model.
3. **Exponential Galaxy**: an identical profile as the Simple Galaxy but with a variable size and arbitrary position angle and axis ratio. Used for resolved galaxies. Has three extra parameters compared to the previous: the effective radius, the axis ratio and the position angle.
4. **de Vaucoulers Galaxy**: a similar to the Exponential Galaxy model but more suited for elliptical galaxies. The profile is $I(r) = I_0 exp(-7.67(r/r_e)^{(1/4)})$ (de Vaucouleurs, 1948). Uses the same parameters as the Exponential Galaxy.
5. **Fixed Composite Galaxy**: a combination of the two previous profiles each with it's own set of parameters. In total there nine parameters. The effective radius, total integrated flux and positional parameters are "shared" between the combination of the two models. However, there is a separate axis ratio and position angle parameter contributed by the two components. The last parameter added is f_{dev} which is the fraction of flux contributed by the de Vaucoulers model .

For each source the software is running another χ^2 check:

$$\chi^2 = \sum_i \frac{(F_{model,i} - F_{image,i})^2}{\sigma_i^2}$$

where i is a given pixel in a source segment, $F_{model,i}$ is the flux of the model in pixel i, $F_{image,i}$ is the background subtracted raw flux in pixel i, and σ_i is the pixel-to-pixel measure of the uncertainty. By attempting every model choice the software can find the minimum χ^2 and determine the best-fit model.

The user now has to convolve the parametrized model of each source to an appropriate PSF. After this step, one can perform "forced" photometry. Since the model and the structural parameters are fixed, the only parameter left to optimize and extract is the flux. For additional information about Farmer see (Weaver et al., 2019).

2.4 Spectroscopic Data

A variety of spectroscopic samples from different surveys are going to be used for our photometric validation in Chapter 3. Since we are only interested in validation, only the most secure targets are included.

The zCOSMOS survey (Lilly et al., 2007) is divided into a bright sample ($i \leq$ 22.5 mag) and a faint sample ($B_J < 25$ mag) (Kashino et al., in prep). We include 8280 galaxies from the bright sample and 739 from the faint sample, covering an important magnitude range. From the VIMOS Ultra Deep Survey (VUDS Le Fèvre et al., 2015), with a range of $i < 25$ mag, we include 944 galaxies.

From the Complete Calibration of the Color-Redshift Relation Survey (Masters et al., 2019, C3R2) 2056 galaxies are included. This sample is picked to fill the color space , thus it is not representative in magnitudes.

We also include 4353 galaxies from the DEep Imaging Multi-Object Spectrograph (DEIMOS) from the Keck Telescope. This sample has galaxies selected from X-rays to radio (Hasinger et al., 2018). The importance of this sample resides exactly on this wide span of selection . Since there are reports of unreliable results for certain selection population (Casey et al., 2012), it is an excellent opportunity to test photometric redshifts accuracy.

A sample of 832 bright star-forming galaxies from (Kashino et al., 2019) taken with NIR Fiber Multi-Object Spectrograph (FMOS), is included. The sample is at $z \sim 1.6$ which has been identified as a very difficult regime for recovering redshifts.

Finally, we add 447 sources from (Rosani et al., 2020) that are observed with Multi Unit Spectroscopic Explorer (MUSE) at the VLT. It includes faint star-forming galaxies ($i > 26$ mag) and Lyα emitters making it extremely useful for comparison at these faint magnitudes.

In Chapter 4 I will also use samples of confirmed high-z quiescent galaxies. They will be used as assessment of our final selected sample. The first sample is taken from Stockmann et al. (2020) and consists of 14 ultra-massive ($log(M_*/M_\odot) \sim 11.5$) at $z \sim 2$ confirmed with X-SHOOTER spectrograph from VLT. The other sample is composed of 8 galaxies from Schreiber et al. (2018) observed with Multi-Object Spectrometer for Infra-Red Exploration (MOSFIRE, Keck) and 1 galaxy from Valentino et al. (2020) observed with X-SHOOTER. This sample at $z \geq 3.5$ is extremely important for the validation of our results.

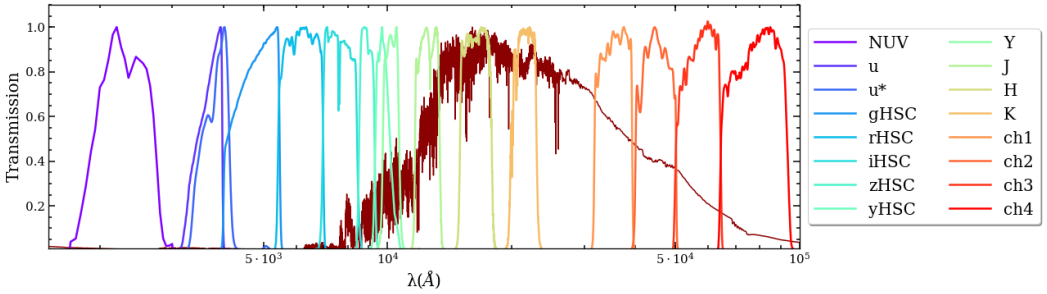


FIGURE 2.3: All the transmission profiles of the filters used in this work. The curves are normalized to a throughput of one and include the effects of atmospheric transmission,camera optics, telescope, filters and detector. Due to normalization we cannot see each filter's and telescope's relative efficiency but this way it gives a more clear view of the profile shapes. The dark red line is an artificial SED of a quiescent galaxy at $z = 2$.

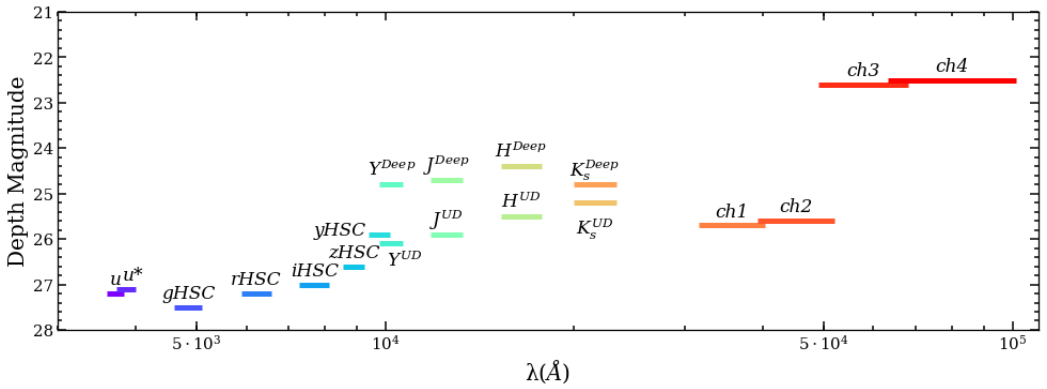


FIGURE 2.4: Depth of each filter at 3σ measured in empty 3" diameter aperture. The length of each line corresponds to the full width half maximum of the transmission curve.

TABLE 2.1: Overview of the characteristics of the all the filters used.

Instrument /Telescope (Survey)	Band[a]	Central[b] λ [Å]	Width[c] [Å]	Depth 3σ[c] (2"/3") ± 0.1
GALEX	FUV	1526	224	...
	NUV	2307	791	25.5[d]
MegaCam	u	3709	518	27.8/27.2
/CFHT	u^*	3858	598	27.1/27.1
HSC	g	4847	1383	28.1/27.5
/Subaru	r	6219	1547	27.8/27.2
(HSC-SSP DR2)	i	7699	1471	27.6/27.0
	z	8894	766	27.2/26.6
	y	9761	786	26.5/25.9
Suprime-Cam	B	4485	887	27.8/27.1
/Subaru	g^+	4805	1255	26.1/25.6
	V	5485	936	26.8/26.2
	r	6310	1372	27.2/26.5
	i^+	7615	1484	26.7/26.1
	z^+	8964	820	25.7/25.1
	z^{++}	9049	1313	25.2/24.7
VIRCAM	$Y^{\rm UD}$	10216	923	26.6/26.1
/VISTA	$Y^{\rm Deep}$			25.3/24.8
(UltraVISTA DR4)	$J^{\rm UD}$	12525	1718	26.4/25.9
	$J^{\rm Deep}$			25.2/24.7
	$H^{\rm UD}$	16466	2905	26.1/25.5
	$H^{\rm Deep}$			24.9/24.4
	$K_s^{\rm UD}$	21557	3074	25.7/25.2
	$K_s^{\rm Deep}$			25.3/24.8
IRAC	ch1	35686	7443	26.4/25.7
/*Spitzer*	ch2	45067	10119	26.3/25.6
	ch3	57788	14082	23.2/22.6
	ch4	79958	28796	23.1/22.5

[a] Complete set of the filters used in this work. The A_{Deep} and A_{UD} notation is to denote the deep and ultra-deep area respectively.
[b] Central wavelength of the transmission curve, defined as the median wavelength weighted by the transmission.
[c] Width of the transmission curve, defined using the half-maximum transmission points and corresponding to the wavelength range the filter covers.
[d] Depths [ABmag] at 3σ computed on PSF-homogenized images in empty apertures.
[e] Value given in Zamojski et al. (2007) corresponding to a 3σ depth.

TABLE 2.2: Exposure times in DR2 (COSMOS 2015) and DR4 (COSMOS 2020)

Filter	Exposure time per pixel in DR2 (Ultra-Deep/Deep)[hrs]	Exposure time per pixel in DR4 (Ultra-Deep/Deep)[hrs]
Y	53.2/11.1	150.2/11.0
J	34.9/12.8	164.7/12.8
H	29.4/13.3	181.8/13.2
K_s	81.5/10.6	157.4/84.6

Chapter 3

Methods

In this section I will describe the different methods used to validate the catalog such as number counts plot and the gzK_s diagram as well as the methods used to derive and validate the photometric redshifts.

3.1 Galaxy number counts

A useful tool for both data validation and theoretical studies is to plot the distribution of number of galaxies in magnitude bins per unit area. The first theoretical models presented to interpret the observed number counts were derived from the assumptions of a static zero-curvature (Euclidean) universe and no luminosity evolution (Sanders, 2005). In this case the slope of the distribution was expected to be 0.6 (Sandage, 1988). But after comparison with data it was understood that this kind of models did not reproduce the observed Universe (Koo and Kron, 1992). Nevertheless, the number counts plot reveal information on the space curvature, the galaxy luminosity evolution, and the diversity of the galaxy population.

However, the data sets to derive this distribution are sensitive to several factors. First of all, photometric errors caused by bad photometric calibration can introduce systematics, which can be more important than photon noise. This can lead to a serious miscalculation of the magnitudes, hence the number counts. Secondly, in the case of a low-S/N detection threshold, spurious detections can be introduced into the data set. This leads to an overestimation of the number of sources in the field. There is also the effect of 'flux-boosting'; sources affected by positive noise fluctuation appear brighter than they are. Lastly, the reliability of number counts depends on the "completeness" of the sample. Completeness is defined as the fraction of extracted sources; it describes how efficiently astronomical objects (at a given magnitude) are recovered from the detection image. The actual completeness of a survey depends on data quality and software configuration. A common method to estimate the completeness is to insert mock sources in the detection image within a wide range of magnitudes. The goal is to test how many of these sources can the software recover as well as how many of them are photometrically reliable (within the detection limit of the survey). Moreover, comparing number counts from other studies is a powerful method to trace potential errors in the data and photometry.

First of all, I present the number counts in the K_s band from the SE version of the present catalog (C20 SE) in Fig 3.1. They are compared with the previous version of the COSMOS catalog (Laigle et al., 2016) (C15) and for both K_s^{Deep} and K_s^{UD}. In this way the comparison is a direct diagnostic of the quality of the data, since the photometric extraction method is the same in the two catalogs (i.e., SE aperture photometry with 3" diameter). The selection of the sources for the number counts was

done based on the column FLAG_COMBINED, which represents the photometric validity of the source. The effective area used here is the effective area after removing regions with bad photometry (e.g., near large saturated stars) and it will be used also in the following figures. The C15 data are taken from Table 3 of Laigle et al. (2016). There is good agreement between the two curves. From the "turnover point", where counts start to decline, we can clearly see that C20 SE is $\sim$ 0.5 mag deeper than C15. In terms of completeness at 25 mag C20 SE has 28% more objects than C15 while in 25.5 mag 43% more objects. Moreover, in Laigle et al. (2016) there was a significant difference in the number counts between the Deep and UDeep area, since the 3σ limit was very different. It is not the case anymore: in 2020 the two regions have very similar exposure time and the difference at the turnover point is small. Fig 3.1 also shows a different slope in the bright end: this is probably due to a more efficient removal of stellar interlopers that are usually bright in the current work. The method use to classify sources is discussed in Chapter 4.

Next, in Fig 3.2 we see a comparison between the two different versions of the catalog (SE and Farmer). Since the two versions originate from the same data, the comparison between them is a test for the different methods applied. For the selection of the sources I also used FLAG_COMBINED to remove bad photometry sources. The Farmer version seems significantly deeper than the SE version ($\geq$ 0.5 mag). At 25 mag Farmer is recovering 5% more objects than SE and even more at 25.5 mag where Farmer recovers 25% more objects. This is mainly due to the fact that the Farmer pipeline applies a more aggressive deblending than the classical SE configuration. The flux of these faint sources, when they are extracted by SE, is more likely to be overestimated because of flux contamination within the aperture by nearby objects. It might also be that Farmer extracts a large number of fake detections, but from a visual inspection it does not seem the case. Therefore, this figure reveals that model photometry provides better results than traditional aperture photometry methods.

I also compare the C20 SE catalog with HUGS (Fontana et al., 2014). The selection is based on the K_s magnitudes so the comparison is straightforward. However, in their work two different number counts are presented. One set of raw "uncorrected" data and another set with "corrected" counts through a simulation. This simulation tested the completeness of the catalog by inserting fake objects within a range of magnitudes. The ability of the software to recover these fake objects quantified the level of completeness. Consequently this lead to calculated assumptions on the number of objects that are present in the survey but are below the permitted signal to noise ratio. In Fig 3.3 the number counts are shown. Even though the corrected simulated counts are still the deepest among K_s-selected samples with the limiting depth exceeding 27 mag, the raw uncorrected counts flatten at $K_s \sim 25$ mag. The Farmer version clearly reaches deeper while the SE version seems to match the same limiting depth. More specifically, at 25.5 mag C20 Farmer is 19% more complete than the uncorrected counts.

Another instructive comparison would be with surveys that include *HST* data. It is important to evaluate the quality of ground-based observations while approaching the era of the next generation of space telescopes (e.g., *James Webb* Space Telescope). In Fig 3.4 there is a comparison with the CANDELS survey in the COSMOS field (Nayyeri et al., 2017). Note here that the selection in the CANDELS survey was done based on the *F*160*W* filter, at a wavelength similar to the UVISTA *H* band (16466 Å). It is evident that CANDELS is deeper than C20 by $\sim$ 0.5mag. Having said that I have to note that this might be due to a horizontal offset caused by systematics of flux calibration. The catalog from Nayyeri et al. (2017) did not include

magnitudes and the conversion formula was taken from the literature using a zero point that was not calibrated for the specific survey. A vertical offset can be seen here and it is probably accounted to underestimation of the area. However we can still conclude that NIR data from *HST* are unparalleled even though they are always combined with complementary ground-based data.

3.1.1 Missing objects

Another informative test is to compare the number of detections in our catalog with the first data release of the HSC-Subaru Strategic Program (Aihara et al., 2017, HSC-SSP). The HSC-SSP is an imaging survey relying on HSC at the Subaru Telescope. This survey among other fields, targeted COSMOS with a limiting depth of $i \sim 27$ mag at 5σ. We want to compare the multi-band detection strategy of our catalog with the i-band image detection of HSC-SSP, in order to see if (and what kind of) objects are missing from the former. This is a concern because the two wavelengths are more sensitive to different SEDs: the i band is more effective for detecting blue, star-forming galaxies, while a K_s band detects better red passive galaxies (which are fainter in i). However in COSMOS 2020 the inclusion of the i, z band in the detection image is expected to increase the completeness of the blue objects (Weaver et al., in prep.).

To test this hypothesis I present in Fig 3.5 the cumulative distribution of the missing objects. In order to locate the missing objects i cross-matched the sources of the two catalogs based on their coordinates (right ascension and declination). Every source that is in the HSC catalog and not present in the C20 catalog within 1" radius of the HSC coordinates, is marked as a missing object. The Farmer version is missing 10.6% of the HSC-SSP objects (within its detection limit) while SE version is missing 7.37%. The difference between the two versions is small but noteworthy. In order to check if there is a preference in the galaxy magnitude the two catalogs are missing, a number counts distribution of the missing objects is presented in Fig 3.6. This figure shows the number counts of the missing objects divided by the total number of galaxies per bin. For $i > 23$ mag the two catalogs have similar fractions of missing objects per bin. On the contrary, the Farmer version is missing predominantly galaxies with bright magnitudes ($i \leq 23$ mag) while the SE version has a more stable trend, which starts to increase only when moving towards the faintest magnitudes.

A similar comparison was presented in Laigle et al. (2016) with the Cosmos 2007 catalog Capak et al. (2007) and it was shown that 16.1% was missing at the detection limits ($i < 26.1$ mag at 5σ in 3" aperture) of the objects present in Capak et al. (2007). The majority of the missing objects were blue, faint star-forming galaxies. I made the same comparison with COSMOS 2020 to see if there is any improvement. Still, the majority of the bright galaxies are missed by Farmer version. A noteworthy behavior is the exponential growth of missing objects for both versions close to the detection limit, higher than in the comparison with HSC-SSP. The explanation for this trend lies in the configuration of the extraction process in Capak et al. (2007): a peculiar choice of parameters lead to a significant number of artifacts caused by the deblending process. As a result many objects appearing close to the detection limit are actually single astronomical objects that were split and recorded as two or more sources in the 2007 catalog.

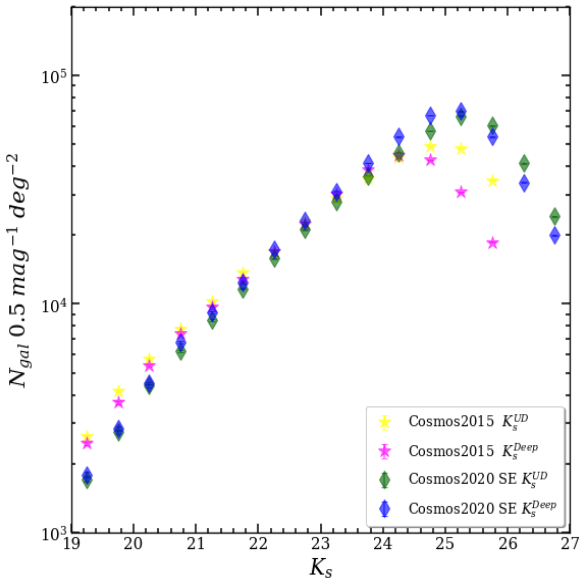


FIGURE 3.1: Number counts comparison in K_s mag between Laigle et al. (2016) and C20 SE. The K_s^{Deep} and K_s^{UD} notation is used to describe the Deep and UltraDeep exposure respectively for both catalogs. The stars marker represent C15 (UltraDeep = yellow , Deep = magenta) while the diamonds are for C20 SE (UltraDeep = green , Deep = blue). All the errors associated to data points are Poisson errors.

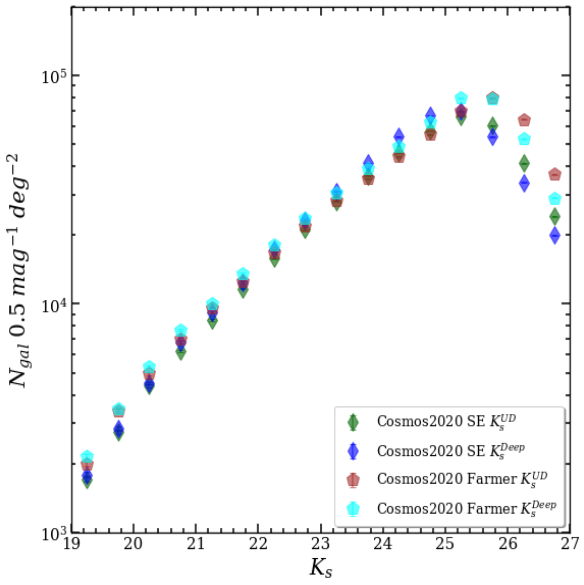


FIGURE 3.2: Comparison of number counts between the two different versions of this work (Farmer and SE). The diamonds represent C20 SE (UltraDeep = green , Deep = blue) and the pentagons C20 Farmer (UltraDeep = brown , Deep = cyan).

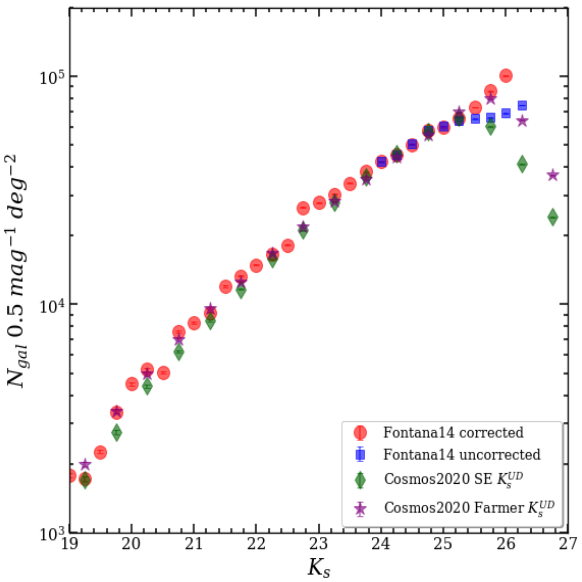


FIGURE 3.3: Comparison of C20 SE (green diamonds), C20 Farmer (purple stars) with corrected (red dots) and uncorrected number counts (blue squares) from Fontana et al. (2014) in the K_s band.

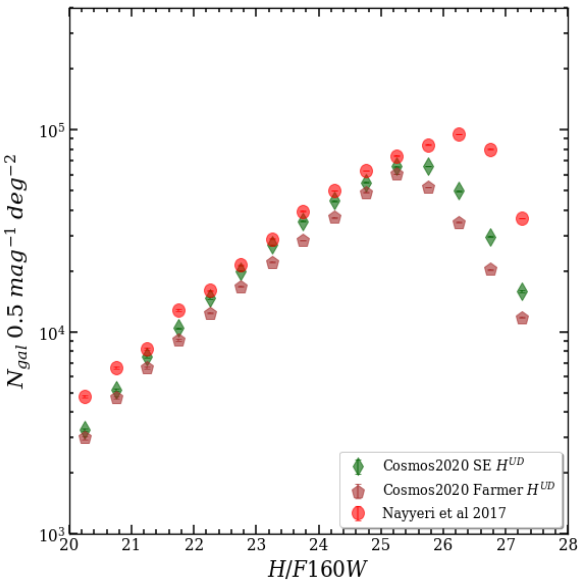


FIGURE 3.4: Number counts comparison between C20 and CANDELS (Nayyeri et al., 2017). The comparison is between the *HST* $F160W$ and UVISTA H-band as seen on x axis.

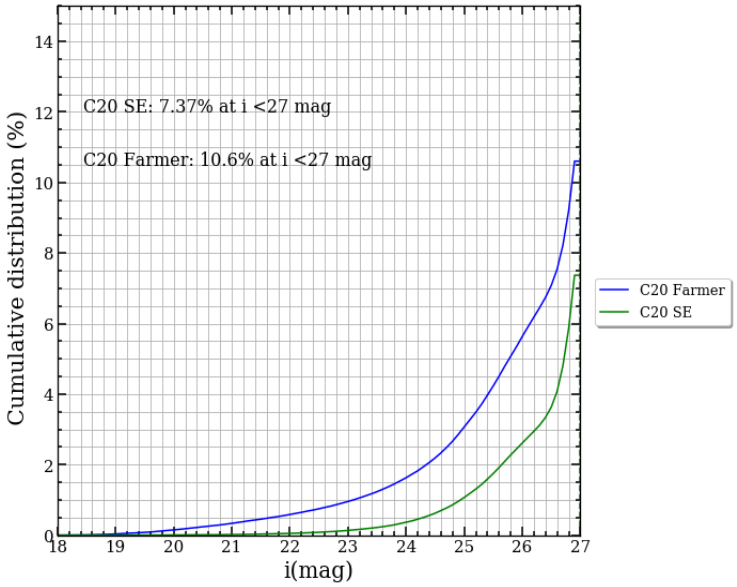


FIGURE 3.5: Cumulative distribution of the missing objects. The curves represent the percentage of the missing objects over the total number of galaxies in Aihara et al. (2017) with $i < 27$ mag (detection limit of the survey) as a function of magnitude.

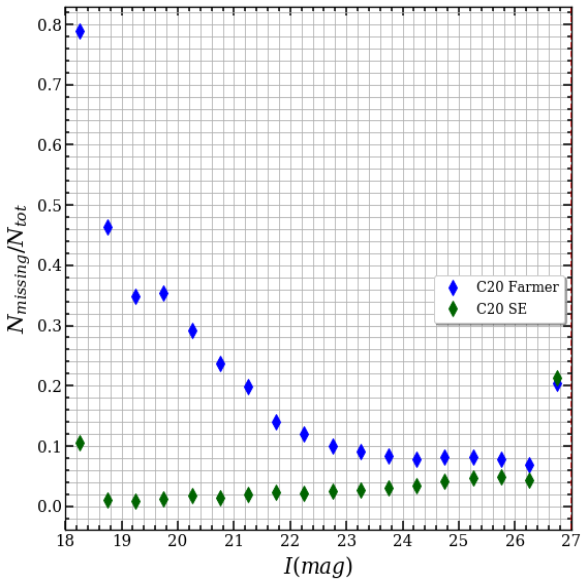


FIGURE 3.6: Number counts of the missing objects of HSC-SSP divided by the total number of galaxies per bin. The y axis shows the fraction of missing galaxies per bin of magnitude.

3.2 Photometric redshift

To measure physical properties of a galaxy one first needs to assess its distance, i.e. redshift. The idea of measuring redshifts from photometric observations dates back to Baum (1962). The author plotted different wavelength observations in order to reconstruct the spectrum of elliptical galaxies from the Virgo cluster and measure the redshift from the horizontal offset.

The main observable in photometry remains the same from then and is the flux of the sources in given bands, which together constitute their SED. Depending on the redshift of each source, their SED gets shifted towards longer wavelength by a factor of $(1 + z)$. The shift can be quantified from spectral features with known rest-frame wavelength. Two important features are the Balmer break and the Lyman break. The Balmer break ($\sim$ 3646 Å) is the difference in flux intensity of both sides of the limit of the Balmer series. It is caused by the absorption of ionizing photons from the second energy level of Hydrogen (see Mihalas, 1967) along with the absorption lines by ionised metals in stellar atmospheres. The combination of these two creates a continuum absorption below the Balmer limit. The Lyman break ($\simeq$ 912 Å) is also a difference in flux intensity, in this case caused by the absorption of the ionizing photons at the limit of the Lyman series and the absorption of the intergalactic medium (e.g., Salvato, Ilbert, and Hoyle, 2019). It was first implemented by Steidel et al. (1996) for the selection of high-z star-forming galaxies (z = 3 - 4) using the Lyman break. Galaxies at these redshift move their rest-frame UV emission to $\sim$ 3600Å. This wavelength does not get absorbed by Earth's atmosphere and can reach our ground-based telescope detectors unobscured. If the galaxy appears in optical (> 3600Å) but not in UV (< 3600Å) then it is a "Lyman-break galaxy" thus a high-z candidate (Steidel et al., 1998).

Strong emission/absorption lines are also spectral characteristics that can contribute to the calculation of photo-z. When a known line is successfully detected and recognized, the derivation of redshift is much easier and more reliable. Some of the strongest and most contributing lines are Lyα, FeII, MgII, [OII], [OIII], Hα and Hβ.

When there is sufficient sampling from broadband filters then it is possible to put constraints on the spectral continuum and derive the redshift based on the detection of the aforementioned features. As mentioned briefly in Chapter 1 the main advantage of the photo-z technique is the ability to estimate redshifts for a whole imaging survey with relatively short observation time and resources. On the opposite, the precision of the typical photo-z estimate is worse compared to a spectroscopic estimation by 1–2 orders of magnitude (since spectroscopic redshift are determined by high-resolution features like absorption and emission lines). It is therefore of crucial importance to assess the quality of photo-z results with a representative spectroscopic sample.

3.2.1 SED fitting

The method used in this work to acquire photo-z is template fitting. Special software (e.g. Bruzual and Charlot, 2003; Maraston, 2005), based on stellar population synthesis models, generate a set of synthetic spectra based on the current theoretical knowledge of star and galaxy evolution. These galaxy "templates" are based on various assumptions. One of the most important is the initial stellar mass function. It is an empirical function that describes the mass distribution of a stellar population that would emerge from a collapsing gas cloud of a given mass (Salpeter, 1955; Chabrier,

2003). Along with that a set of tunable parameters like e.g. age, dust content, metallicity are included in the templates. In addition, empirical templates can be extracted from observed spectra and expanded by model extrapolation to a larger wavelength range (Polletta et al., 2007).

Those templates are compared to the observed SED of an astronomical source through the following χ^2 formula:

$$\chi^2 = \sum_{f=1}^{N_f} \left(\frac{(F^f_{obs} - A \times F^f_{pred}(z,T)}{\sigma^f_{obs}} \right)^2 \tag{3.1}$$

where F^f_{pred} is the predicted flux in the band f for a galaxy model (template) T at redshift z. F^f_{obs} is the observed flux and σ^f_{obs} the corresponding error in the same band. The f index iterates over the set of N_f filters used in the survey; A is the normalization factor that takes into account the intrinsic luminosity of the source (Ilbert et al., 2009). The most likely z value is calculated from the χ^2 minimization while z, T, A are free parameters. The final value of z can be determined as the Eq. (3.1) solution with the smallest χ^2, or derived by studying the whole redshift probability distribution $P(z)$, represented by:

$$P(z) \propto exp\left(-\frac{\chi^2(z) - \chi^2_{min}}{2} \right) \tag{3.2}$$

3.2.2 Factors that affect photo-z estimates

Here I will present a brief description of the effects the main parameters in the stellar population synthesis models (age, metallicity, dust) have on the SED fitting procedure.

The age of the bulk of the stellar population in a galaxy is an important factor that affects the photo-z accuracy. In the case of a young stellar population, the Lyman break emerges in the spectrum at $\sim 10^7$ years (Bruzual A. and Charlot, 1993). If the observed galaxy has a stellar population younger than 10^7 years than the accuracy drops by a factor of 100 for $z \leq 3$ (Bolzonella, Miralles, and Pello', 2000). With the absence of a strong spectral feature such as the Lyman break the fitting software does not have enough information to constraint the result. An older stellar population can be recognized from a strong Balmer discontinuity, but such a feature can be mimicked by a model with a much younger age and a conspicuous amount of dust. Indeed, the UV part is extremely sensitive to dust and can change drastically depending on the dust content and grain properties (Salvato, Ilbert, and Hoyle, 2019). As a result a young blue star-forming galaxy may appear as an old red quiescent galaxy .

Another important quantity that can result in degenerate z solutions is metallicity. Metallicity is a term used to describe the abundance of elements heavier than H or He. The presence of "metals" absorbs the UV radiation making the galaxy appear "redder". As seen in Fig 3.7, for a given SED observed in optical-NIR, the model of a metal-rich young galaxy can fit the data as well as a model made by an old metal-poor stellar population.

These kinds of degeneracy (age-dust, age-metallicity) cause severe problems in constraining not only the redshift but also the physical parameters that can be extracted from the SED fitting (see below).

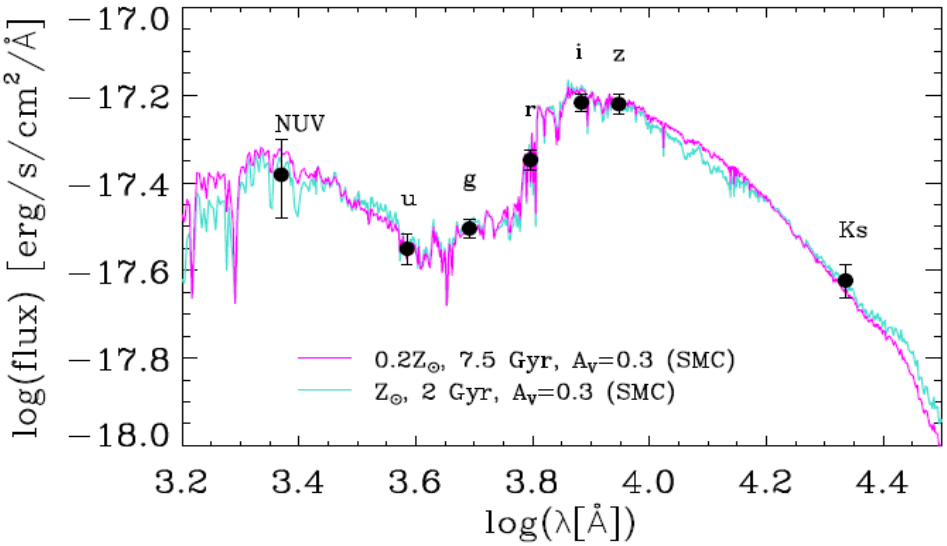


FIGURE 3.7: SED template fitting of photometric observations from a galaxy taken from the VIMOS Public Extragalactic Survey (VIPERS). The two templates plotted have the same dust content but different metallicities (Z). They represent a metal-rich young galaxy (cyan line) (2 Gyrs, $Z = 0.2Z_{\odot}$) and an old galaxy (7.5 Gyr,$Z = Z_{\odot}$) (magenta line). However, the shape of the spectrum is the same and both solutions have a 98% probability. Image credit: I.Davidzon

Modeling of emission lines in the templates plays an important role in the photo-z derivation. Emission lines can enhance a broad band magnitude by up to 0.4 mag (Ilbert et al., 2009). Therefore including in the templates emission from [OII], [OIII], Hβ, Hα ,and Lyα is critical to obtain accurate photo-z. Neglecting the emission lines -especially those with high equivalent widths- can lead to overestimation of derived physical properties such as stellar mass and SFR by a factor of 3 (Labbé et al., 2013) It has been shown that when templates include emission lines they can decrease the dispersion of photo-z (compared to spectroscopic redshifts) by a factor of 2.5 (Ilbert et al., 2009).

The quality of photo-z also depends on the wavelength range sampled by the survey. As already discussed in Chapter 2, the inclusion of NIR filters is crucial to constrain the rest-frame optical light for galaxies at $2 \leq z \leq 5$. Without NIR, especially at faint magnitudes the SED fitting software can misidentify the Lyman and Balmer break features. In Fig. 3.8 an example of this degeneracy is presented. While spectroscopic redshift z_{spec}=0.311, the photo-z estimate is z_{phot}=2.97. In this case it is purely due to the lack of NIR data that the fitting gives a wrong result. Since the low and high-z template have both the same shape in UV and optical, the only way to discriminate between the two solutions would be the inclusion of data at $\lambda > 1\,\mu$m. It is also interesting to note that in Fig. 3.8 the $P(z)$ distribution has the main peak around the wrong redshift; nonetheless, there is also a secondary peak around the correct value.

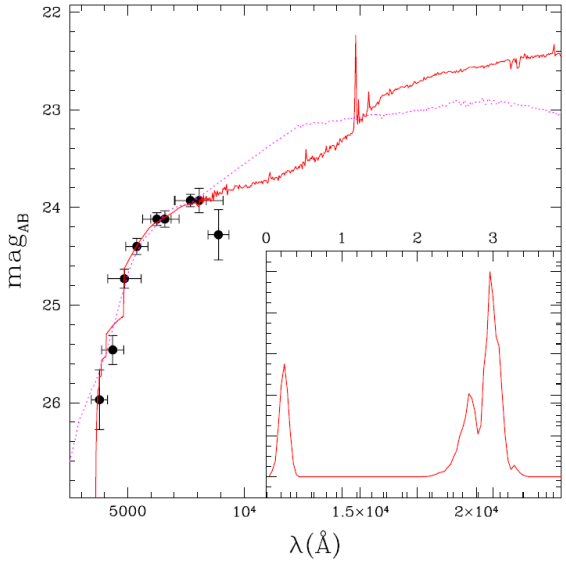


FIGURE 3.8: An example showing a case of degeneracy caused by the lack of NIR filters. The black points are the photometric data from the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS Ilbert et al., 2006). The solid line represents a template with z_{phot} = 2.97 while the dotted line z_{phot} = 0.24. The spectroscopic redshift is 0.311 for this galaxy. The bottom right panel shows the PDFz with the possible solutions. Image taken from (Ilbert et al., 2006).

3.2.3 Assessment of photometric redshifts through comparison with spectroscopic samples

One of the most common and important method to assess the quality of the photo-z is by comparing the results with a sample of spectroscopic redshifts. There are two different ways often used to make the comparison. One of them is examining the population of the outliers. An outlier is considered a galaxy with $\mid z_{phot} - z_{spec} \mid /(1 + z_{spec}) > 0.15$ (Hildebrandt et al., 2010). The percentage with respect to the full sample is denoted by η. A direct measurement of the accuracy σ is the normalized median absolute deviation $1.48 \times median(\mid z_{phot} - z_{spec} \mid /(1 + z_{spec}))$ (NMAD) (Hoaglin, Mosteller, and Tukey, 1983). This dispersion estimate describes the scatter between the photo-z (prediction) and the spec-z (truth). However it is more robust than regular dispersion calculation since it's less sensitive to the outliers.

In Fig 3.9 I present the comparison of the two versions of the catalog with our spectroscopic sample. The selection of the sources is based upon the validity of the photometry and the redshift values assigned to these sources. The two versions have almost identical values for catastrophic outliers and accuracy. For SE version is $\eta = 3.1\%, \sigma = 0.011$ in a sample of 13463 galaxies , while in Farmer version is $\eta = 3.2\%, \sigma = 0.012$ in a sample of 13228 galaxies.

Even though the difference in the total number of galaxies between the two samples is small, it is instructive to investigate the reason behind this difference. There are two possible explanations for that. The first one is connected to the way Farmer detects the objects on the image and extracts the flux. In the Farmer catalog there are 2 different pairs of coordinate columns. One with the "true" values (`ALPHA_J2000, DELTA_J2000`) and one with the coordinates assigned to the source by Farmer (`ALPHA_DETECTION, DELTA_DETECTION`). As the software is trying to detect

the source, it fails to locate the center and tries to fit the model on wrong coordinates. As a consequence, this results in a failed fit and thus the source is never recovered. On the other hand, determining coordinates with SE is more straightforward and robust.

Trying to test this hypothesis, I cross-matched the two column pairs of coordinates in Farmer. From a total of 889,639 objects in the Farmer catalog, 5,157 do have a separation larger than 1 arcsec between the two coordinate pairs. The matched pairs have a distribution of separation that peaks at $0.12 \sim 0.14$. Moreover, the number of objects at this peak are $\sim$ 30% of the total objects of the catalog. Of the unmatched objects 2,483 of them are marked as photometrically valid and almost half of them are undetected in the *i*-band. Both the valid and non-valid unmatched objects peak their *i*-mag distribution at $\sim$ 23mag. All in all, in connection with Fig 3.6 we can see that there is a systematic offset probably in the *i*-band, which is the main reason for the difference in the total number of galaxies in the cross-matched samples. A secondary reason and less effective is a difference in the `FLAG COMBINED` values of some objects between the two catalogs that could be due to different masked regions.

In comparison with Laigle et al. (2016) the improvement is huge. The C15 catalog was cross-matched with a variety of spectroscopic samples. However most of them had less than 1000 objects or did not cover the high-z regime. The comparison with *zCOSMOS* Lilly et al., 2007, which is the largest among the samples compared (8608 galaxies), yields very good results ($\eta = 0.51\%, \sigma = 0.007$). Despite that, the sample is limited to z=1.19 so it does not contribute to the assessment of high-z sources. The two most relevant comparisons are with (Kartaltepe et al., 2010) ($0.02 < z < 5.87$, $z_{med} = 0.93$, 2022 galaxies) and (Le Fèvre et al., 2015, VUDS) ($0.1 < z < 4.93$, $z_{med} = 2.7$, 998 galaxies). The characteristics of the comparison are $\eta = 7.96\%, \sigma = 0.014$ and $\eta = 13.13\%, \sigma = 0.028$ for the two samples respectively. This shows that there is a 60-70 % improvement in the reduction of catastrophic outliers with an increase of 20-50 % accuracy.

In Fig 3.10 I am showing a comparison between the computed z_{phot} from both versions of the catalog. The percentage of outliers is $\sim$ 0.9% with a $\sigma = 0.009$ either selecting SE of Farmer version as "truth". As a result, both aforementioned plots denote that there is no significant difference in the derived redshifts between the two photometric extraction methods.

3.2.4 Physical properties derived from SED fitting

The SED fitting is a versatile tool. After determining the redshift of galaxies, we can calculate several important rest-frame physical properties. I will mention and describe a few that are commonly derived in large surveys, namely luminosity, stellar mass and star formation rate.

Galaxy luminosity in a given rest-frame band is the most direct quantity that can be derived. The most simple way is to use the following formula;

$$L_n = 4\pi d^2 F_n \tag{3.3}$$

where d is the distance to the source (can be derived from Hubble's law since we know the redshift) and F_n is the brightness of the source (flux over surface area) in the n band. It is more sensitive to photometric accuracy and less to redshift accuracy so it can provide quite robust results. Note that we can only infer luminosities for the available bands of the survey and can't make any direct measurements of the bolometric luminosity.

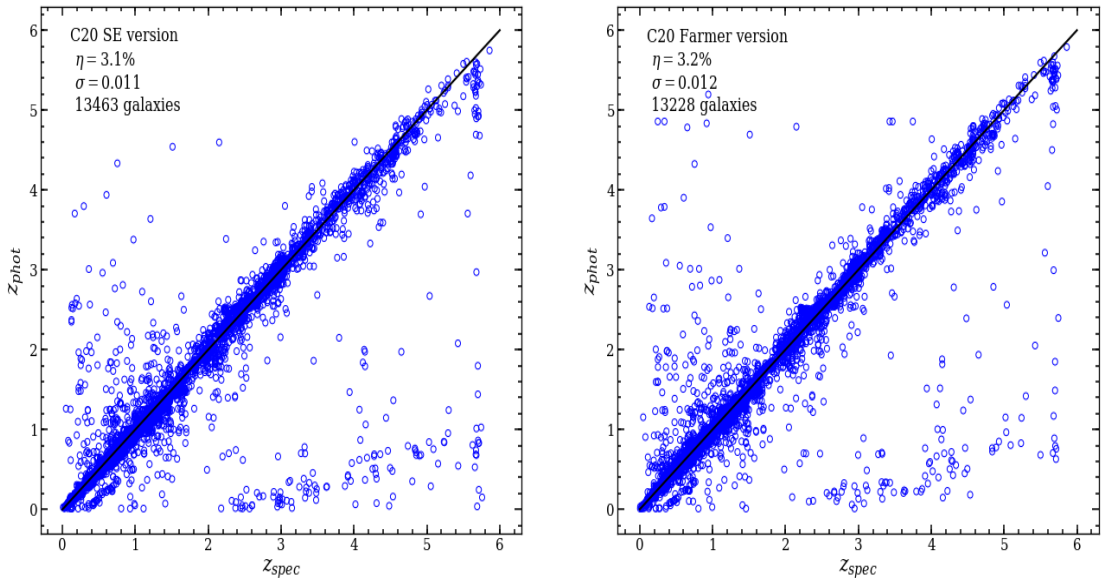


FIGURE 3.9: Comparison between photometric and spectroscopic redshift of the two versions of the catalog with our spectroscopic sample . The C20 SE version is on the left while the C20 Farmer version on the right. On the top left of both plots you can see the catastrophic outliers percentage , the accuracy based on the NMAD and the total number of galaxies plotted.

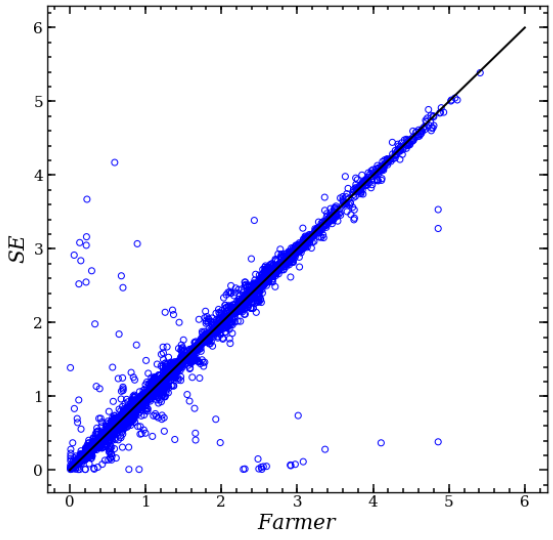


FIGURE 3.10: Comparison of the z_{phot} derived from the two versions of the catalog. On y axis is the SE redshift and on x axis the Farmer redshift.

The galaxy stellar mass can be computed from the ratio between mass and luminosity (M/L). In general there is a linear correlation between mass and luminosity (Girardi et al., 2002). However, using an extrapolated bolometric luminosity from the SED fit would introduce many uncertainties in the light-to-mass conversion Therefore the best choice is to calculate the mass based on the M/L in a reference band, possibly one that is a good proxy of the bulk of the stellar population and not affected by degeneracies between age, dust and chemical abundance (Walcher et al., 2010). The best choice for that is the rest-frame NIR ($\sim$2 μm) which means e.g. telescope filters like K_s at $z < 3$, or *Spitzer*/IRAC channels at higher redshifts. In our analysis, instead of a simple M/L conversion, we will consider the full SED and its best-fit template, because from the stellar population synthesis model we know also the mass of the latter (this approach is not feasible with empirical templates like Polletta et al., 2007).

A key quantity to extract from an SED fit is SFR. Kennicutt (1998) made a huge contribution to the galaxy evolution field when he showed the relation between the IR luminosity ($8 - 1000\mu$m) and the SFR of galaxies. As mentioned before, young stellar populations emit predominantly in the UV which in turn gets absorbed preferably by the dust content. After the dust is heated it emits back in IR. Later works like Kennicutt et al. (2009) have taken into account other parameters that contribute (range of ages of the stellar populations), but the core is the same; IR emission in an excellent SFR tracer. In this sense the extracted IR emission from the SED fit can be used to calculate SFR with the use of following formula (Kennicutt, 1998);

$$\frac{SFR}{1 M_\odot\, yr^{-1}} = \frac{L_{FIR}}{2.2 \times 10^{43} \mathrm{erg}\, s^{-1}} \tag{3.4}$$

In our case again the SFR will be calculated straight from the best-fit SED model. However, this approach has a serious of caveats as discussed in (Davidzon et al., 2018). Many works studied the evolution of specific SFR (sSFR) (Stark et al., 2007; Behroozi, Wechsler, and Conroy, 2013). The quantity sSFR is defined as SFR/M_*. It is used to disentagle the mass dependency of SFR. For example it shows the difference between the formation of 10 $M_\odot$/yr of a $10^7 M_\odot$ galaxy with a $10^{11} M_\odot$ galaxy. If we want to describe it in more physical terms we would say it is the timescale which a galaxy doubles it's M_*. All the past works extract SFR and M_* straight from the SED fit and ended up in disagreement with theory (Weinmann, Neistein, and Dekel, 2011). Serious systematics are introduced to the final values from varying factors that affect SED such as nebular emmision contamination, assumptions about star formation history, age, dust and metallicity. For further information see Davidzon et al. (2018).

3.3 The gzK_s color-color diagram

The gzK_s diagram is a physically motivated validation method that I will use to assess the the photometry and the redshift derivation of the survey. This diagram is an empirical color selection introduced by Daddi et al. (2005) as a novel photometric technique to select and classify galaxies mostly in the range $(1 < z < 2)$. Variations of the diagram (VJL, iHM) can extend the redshift range and identify star-forming and quiescent galaxies at $z \sim 2.5$ and 3.5 respectively (Daddi et al., 2005; Guo et al., 2012).

The original diagram had two colors based on B, z, and K_s band. The color index on the x axis is $B - z$ while on y axis is $z - K_s$. Using this selection on samples of

the K20 survey (Cimatti et al., 2002), Daddi et al. (2005) that galaxies occupy certain regions in the diagram depending on their redshift and star formation activity. The samples with photometric redshifts were cross-matched with their spectroscopic redshifts to enhance the validity of the method. In Fig 3.11 we can see an example of a BzK_s diagram taken from Fang et al., 2015. The top left region selects the star-forming population according to the criterion:

$$(z - K_s) - (B - z) \geq -0.2 \tag{3.5}$$

The reason behind this is the little effect on $B - z$ color of the duration of star formation, in connection with the increase of $z - K_s$ with age due to the development of Balmer breaks that fall further than the z band for $z \geq 2$ (Daddi et al., 2005). On the other hand, the old passive high-z red galaxies are concentrated on the region defined by the criterion:

$$(z - K_s) - (B - z) \geq -0.2 \cup (z - K_s) > 2.5 \tag{3.6}$$

Just below these two regions lie the lower redshift galaxies with no particular separation between star-forming and passive. The region defined below as seen in Fig 3.11 with the dashed line is the region occupied by stars. Note that all these criteria were quite robust since the interlopers were quite low (Daddi et al., 2005, e.g in the GOODS area of the K20 survey were 13%).

This technique was later utilized by other works as well (Lane et al., 2007; Blanc et al., 2008; McCracken et al., 2010; Laigle et al., 2016) establishing it as a solid validation tool. In this work I am going to use it here as a validation tool and later in Chapter 4 for galaxy-star classification. Since this survey does not include B band (4488Å) I will use the g band (4847Å) that has a negligible central wavelength difference. In Fig 3.12 I present the gzK_s diagram of both versions of the catalog. The sources represented here were selected based on their photometry. Only sources with a detection on all three bands (gzK_s) are included and only K_s-bright objects ($K_s < 24$). The latter serves to displaying purposes (otherwise the different regions would be hazy and not distinguished).

We can see that all different loci are clear,well-defined and in agreement with the literature. The central locus represents the bulk of the sample with low-z galaxies ($z \approx 1$). The top left region has another well-defined locus with star-forming galaxies ($z \geq 2$). As we go to the top right we see the spread of the passive high-z galaxies ($z \geq 2.5$). To separate stars I use an empirical broken line based on the shape of the "sequence". The following are the x and y coordinates of the initial, middle, and end point of the broken line for the SE catalogue:

$$(-3.07, -1.99), (1.46, 0.1), (8.5, 2.08); \tag{3.7}$$

while for Farmer version I used

$$(-3.07, -1.99), (1.35, 0.15), (8.5, 2.08) \tag{3.8}$$

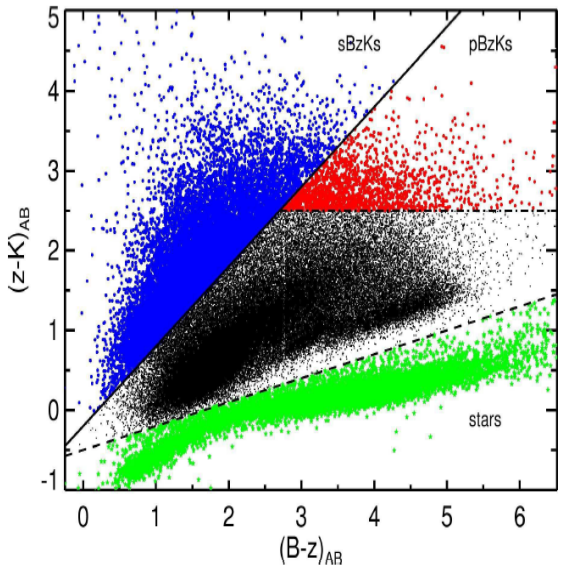


FIGURE 3.11: The BzK_s diagram of sources taken from the CANDELS survey on the COSMOS field from the *HST* F160W band. The blue and red dots are star-forming galaxies (sBZks) and passive galaxies (pBZks) respectively both at $(z \sim 2)$ (pBZks) galaxies. Black dots are lower redshift galaxies and green dots are stars. Image taken from (Fang et al., 2015)

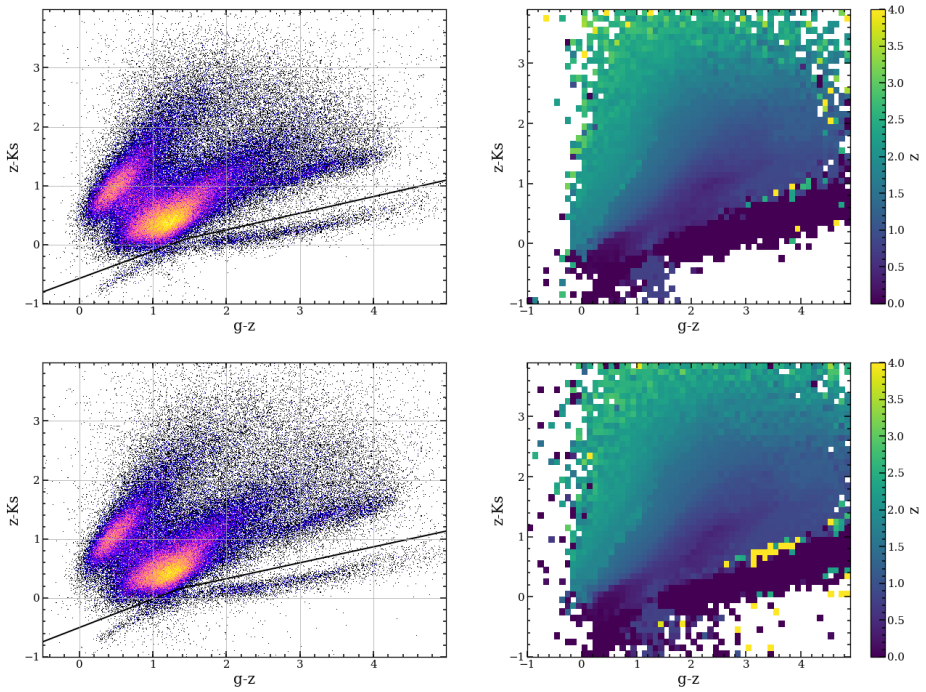


FIGURE 3.12: The gzK_s diagram of the two versions of the catalog.The top plots are derived from the SE version while the two bottom from the Farmer version. The two left plots are density plots of the sources with the brighter colors indicating larger density of the area. The two right plots are also density plots but the color coding this time represents the median redshift value of the bin.

Chapter 4

Data Analysis and Results

In this Chapter I will present the analysis mostly performed on the SE catalog. Due to the results of the validation methods in the previous Chapter, I will not extensively continue the analysis on the Farmer catalog. Since the Farmer catalog is based on a novel method that has not been tested sufficiently through other works and is still a pipeline in development, I decided not to risk presenting results with a high level of uncertainty. The results from Farmer will be mentioned briefly and only be used as a cross-match for those of SE.

4.1 Star/Galaxy classification

Since we are interested only in galactic sources we have to efficiently remove the stars from our sample. For this task I will use two different methods in order to achieve the highest level of purity in our sample.

As I showed in the previous Chapter, the gzK_s diagram can be used to classify sources based on their type, their activity and their redshift. Thus we can use the diagram to mark the objects occupying the stellar sequence as stars. More specifically, I take all the objects that are below the broken line in Fig. 3.12 and flag them as stars. The total number of objects removed this way is 177,517 out of total 1,720,700.

However, as mentioned before, there is a 10-15% fraction of stellar interlopers in the regions occupied by galaxies in gzK_s. This is mainly due to the fact that stars with red $z - K_s$ colors are not selected in the gzK_s (Straatman et al., 2016). Thus I will also create another flag to remove the stars that is based on the SED best-fit. Having said that i cannot simply take the difference between the χ^2 of the galaxy template fit and the stellar template fit (Andrae, Schulze-Hartung, and Melchior, 2010). The galaxy templates have 3 degrees of freedom while the stellar templates only 2. Therefore I will use the reduced chi-squared ($\chi^2_{red} = \frac{\chi^2}{N_{filt}-K}$, where N_{filt} is the number of filters that recovered this source and K is the degrees of freedom of the model) (Davidzon et al., 2017). For convenience I will note it as χ^2 here.

Among the most popular conditions to mark stars that can be found in the literature is $\chi^2_{gal} < \chi^2_{star}$ (Laigle et al., 2016) and $\chi^2_{gal} - \chi^2_{star} > 1$ (Davidzon et al., 2017). Since we want to acquire the best level of purity and eliminate any possibilities of stellar interlopers, I will adopt a more conservative condition. I will use the condition $\chi^2_{gal} - \chi^2_{star} > -1$. Every source that fulfills this condition is flagged as a star. In Fig. 4.1 the distribution of the χ^2 difference for each source can be seen. With this method the total number of sources marked as stars are 543,656. Between the two methods a total of 54.040 stars overlap. As a conclusion, the two methods are both necessary because each of them "detected" and removed different stellar populations. Without both of them a serious fraction of stellar interlopers would contaminate the galaxy sample.

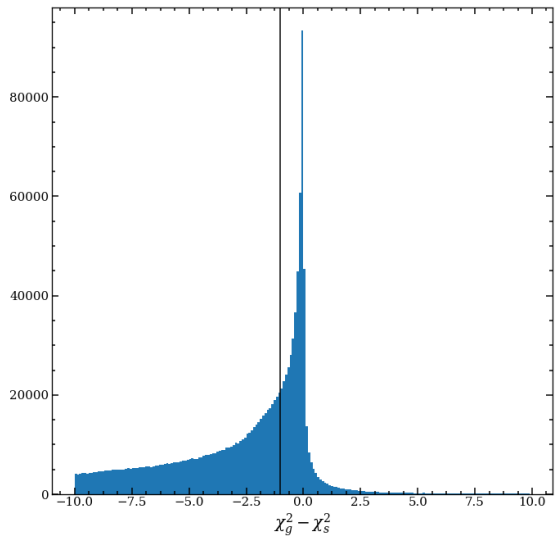


FIGURE 4.1: The distribution of the χ^2 difference between the galaxy template and the stellar template. The black vertical line at -1 marks the threshold which above that sources are marked as stars.

4.2 Galaxy classification with rest-frame color-color diagram

After successfully removing the stellar population from the sample, I can now move on to classify the galaxies based on their star formation rate. This way i will be able to isolate the quiescent sample and study it further.

One very efficient method to do so was presented by Williams et al. (2009). In that work the observed magnitudes of J, K and IRAC Channel 2 (4.5μm) for objects at $z \sim 2.5$ were converted into rest-frame magnitudes $U\,V\,J$ respectively. The choice of this rest-frame magnitudes was not arbitrary. The index $U - V$ is a useful indicator of SFR since U band is a good tracer. The index $V - J$ traces successfully the dust content since dust scatters visual light. Combining these two indices in a diagram a very interesting bimodality emerged; a diagonal sequence that extended from blue to red $V - J$ and locus just above that is red in $U - V$ but blue in $V - J$.

These features in the diagram classify the galaxies in the same sense as the gzK_s. The bottom left region marks the unobscured star-forming galaxies with little to no dust content. As the sequence moves to the top right , dust content increases thus red dusty galaxies with obscured star formation lie. On the top left of this sequence the quiescent galaxies concentrate. An example can be see in Fig 4.2. The breakthrough of this rest-frame diagram is that it successfully breaks the dust degeneracy; the red dusty galaxies occupy different region than the quiescent galaxies. This separation would not be possible with a single color (e.g. $U - V$) since the quiescent sample would be contaminated by red dusty star-forming galaxies.

However, later works pointed some weaknesses that were able to bypass by adopting a variation of the initial UVJ diagram (Ilbert et al., 2010; Ilbert, O. et al., 2013; Laigle et al., 2016; Davidzon et al., 2017). This variation is the $NUVrJ$ diagram. It was shown that $NUV - r$ is a more reliable indicator to distinguish between fully quiescent galaxies and ones with a low residual star formation activity (Martin et al., 2007; Arnouts et al., 2007). In the original diagram these two populations occupy the

same region thus cannot be distinguished (Davidzon et al., 2017). Moreover, the dynamical range of the $NUV - r$ index is approximately twice as large as the $U - V$. Therefore, the index is less sensitive to uncertainties introduced by observations and photometry (Ilbert, O. et al., 2013).

Taking into consideration the above reasons, I will use the $NUVrJ$ diagram to select high-z quiescent galaxy candidates. In Fig 4.3 I present the $NUVrJ$ diagram of the galaxy sample in 4 redshift bins spanning from $z = 2$ to $z = 5$ with density color coding. The criterion to classify a galaxy as quiescent is:

$$(NUV - r) > 3(r - J) + 1 \; and \; (NUV - r) > 3.1 \tag{4.1}$$

Although the criterion was defined empirically (Ilbert, O. et al., 2013), it is physically motivated. The slash line moves perpendicular with increasing sSFR. In addition, the $V - J$ (dust content) moves the galaxies parallel to the border separating effectively the "true" quiescent galaxies on the left side from the dusty star-forming on the right.

Looking at the diagram we can see important aspects of galaxy evolution. We can see that a large population of quiescent galaxies are already in place at $2 < z < 3.5$ as observed by previous surveys. At $z > 3.5$ it is much more difficult to observe quiescent galaxies because the age of the universe ($t_{cosmic} < 1.8\,\mathrm{Gyr}$) is shorter than the timescale of most of the quenching mechanisms. At $2 < z < 3.5$ we also see that there are more galaxies with $2 < NUV - r < 3$, which are in transition towards the quiescent locus. Moreover it is noteworthy that dusty star-forming galaxies (upper-right corner of the $NUVrJ$ diagram) exist from at least $z \geq 5$. Unfortunately some of these dusty galaxies can be scattered inside the quiescent locus because of the large uncertainty in their color estimates.

One of the most important achievements of this survey is the extend of the redshift range that the $NUVrJ$ can cover. The original diagram from Williams et al. (2009) studied galaxies up to $z \sim 2.5$. This limited redshift range was because of the filters used in UDS. The 'reddest" filter used was IRAC Channel 2 (4.5μm). As a result the maximum redshift of which they could interpolate the intrinsic UVJ fluxes from the observed data was $z \sim 2.5$. On the other hand in our survey we extended our wavelength coverage by including all of IRAC channels (coverage reaching 7.9μm) thus making it possible to recover rest-frame magnitudes from galaxies at $z < 6$.

Implementing this method I find 35 high-z quiescent galaxy candidates (QG). From a cross-match check with the Farmer version I find that 29 of them are present in the Farmer version. However, two of them have non-valid magnitudes and redshift and one of them is classified as a star. On the other hand, doing the same analysis with Farmer yields 558 QG of which 7 appear in the SE sample.

Nonetheless, recent works have shown that the UVJ diagram (and the $NUVrJ$ in addition) do not necessarily constrain age and metallicity (Leja, Tacchella, and Conroy, 2019; Schreiber et al., 2018). It was demonstrated that observed UVJ features are correlated with galaxy scaling relationships and will eventually evolve with cosmological time. Furthermore, sub-solar metallicity galaxies need 3 Gyr to move into the quiescent region (Tacchella et al., 2018), therefore at $z \geq 3$ it is unlikely to find them in the quiescent UVJ colors. On top of that, Leja, Tacchella, and Conroy (2019) investigate the correlation between sSFR and UVJ colors to show that at sSFR < $10^{-10.5} yr^{-1}$ the correlation starts to saturate until sSFR=$10^{-11} yr^{-1}$ which is fully saturated. Having said that, it is also mentioned that $NUV - r$ color index correlates better at low sSFRs can be used to select QG more efficiently. Last but not least, a

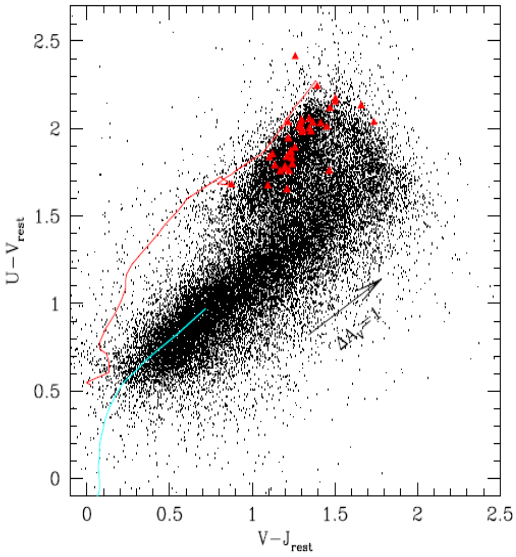


FIGURE 4.2: The UVJ diagram from a $z_{phot} < 2.5$ sample from the UDS (black points). The red triangles are spectroscopically confirmed old passive galaxies. The blue and red lines are evolutionary tracks of constantly star-forming galaxies and passively evolving galaxies (Bruzual and Charlot, 2003). Image taken from (Williams et al., 2009).

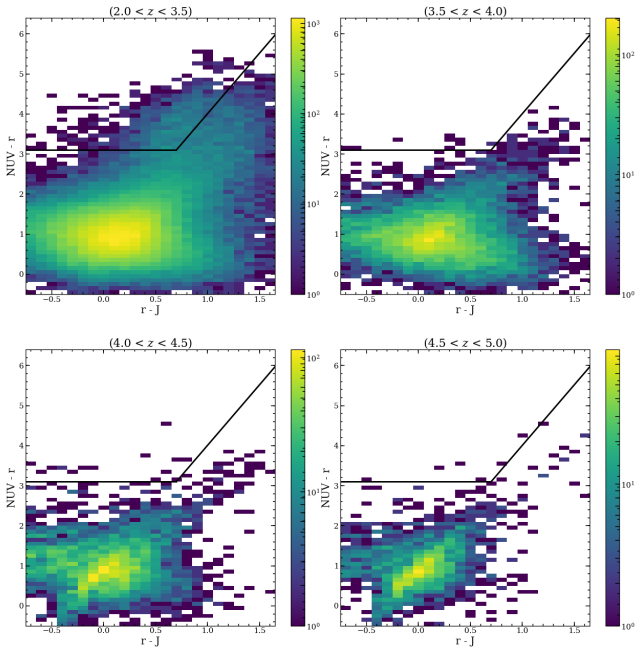


FIGURE 4.3: The $NUVrJ$ density plot in 4 different redshift bins shown on the top of each plot. The color coding is the density and the values are shown in the colorbar on the right of each plot. The broken line is the criterion to classify quiescent galaxies.

series of works report that $\sim 10\% - 30\%$ of the UVJ-quiescent galaxy sample preserve non-negligible star-forming activity (Belli et al., 2017; Schreiber et al., 2018; Díaz-García et al., 2019).

4.3 SFR and Stellar mass

It has become clear from the aforementioned studies that the UVJ diagram and it's variations are not complete and substantially contaminated. To counter this I will also use another method of selecting quiescent galaxies.

This method is based on the evaluation of SFR and sSFR of the galaxies. Previous studies discovered a strong correlation between the SFR and M_*. This "main sequence" of star-forming galaxies was found to extend from $z \simeq 0$ (Noeske et al., 2007) to $z = 7$ (Stark et al., 2013; Steinhardt et al., 2014). The shape of the sequence is directly dependent on z as seen in Fig 4.4 and is based on observations. At low-z and $log(M_*) < 10.5$ the shape is a unity slope. With increasing redshift, the normalization factor of the sequence grows. An interesting feature is the "bending" of the curve at low-z and high M_*, that has no solid explanation.

The fitting of the data points produces a parametric formula of the SFR of main-sequence galaxies:

$$log_{10}(SFR_{MS}[M_\odot/yr]) = m - m_0 + a_0 r - a_1[max(0, m - m_1 - a_2 r)]^2 \tag{4.2}$$

with $r = log_{10}(1+z)$, $m = log_{10}(M_*/10^9 M_\odot)$, $m_0 = 0.5$, $a_0 = 1.5$, $a_1 = 0.3$, $m_1 = 0.36$ and $a_2 = 2.5$ (Schreiber, C. et al., 2015). All these terms are defined empirically for best-fit and also are assigned errors. However, the accuracy of the SFR-M_* correlation formula is beyond the purpose of this work so I will use these terms without assigned errors. In short this formula calculates SFR_{MS} which is the SFR that the galaxy would have if it was sitting in the main sequence. Using this formula I will select quiescent galaxy candidates at $3.5 < z < 5$. I will consider as quiescent any galaxy that has SFR 10 times less than it's main-sequence star-forming counterpart of same mass and redshift. In mathematical terms, any galaxy is quiescent if:

$$\frac{SFR_{MS}}{SFR} > 10 \tag{4.3}$$

After implementing this method to the SE version i find 9 quiescent galaxy candidates. Cross-matching with the quiescent sample from the $NUVrJ$ sample I get 1 match. Since our selection criteria are strict, I ran some additional checks on the SFR-M_* sample. After plotting the $NUVrJ$ of the SFR-M_* sample i noticed that 2 galaxies sit very close to the quiescent region borderline [$(NUV - r, r - J)$ =(3.63,1.07) and (3.03,0.37)]. However, the positions of the galaxies on the $NUVrJ$ diagram are subjects to rest-frame color uncertainties which mainly emerge from photometric uncertainties and k-correction (Davidzon et al., 2017). Having that in mind, one can safely assume that the positions of these two galaxies is inside the quiescent locus if the rest-frame color uncertainties were taken into account. As a result I will consider that the overlapping galaxies of the two methods are 3.

As an additional check to the selection I look at the sSFR values of the selected galaxies in the catalog. It is commonly used in the literature to consider a galaxy as quiescent when $log_{10}(sSFR) < -11$ and additionally to consider a galaxy in the process of quenching ("post-starbust") when $log_{10}(sSFR) < -10$ (Ilbert et al., 2010;

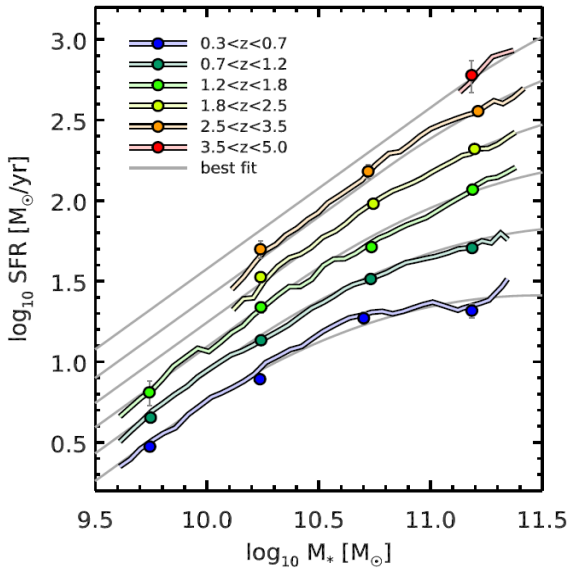


FIGURE 4.4: The evolution of SFR-M_* as a function of redshift. The colored curves represent the positions of sliding bins of mass (filled circles). The colors of each line correspond to different redshift as can be seen from the legend. The faint gray lines are the best fit to the data. Image taken from (Schreiber, C. et al., 2015)

TABLE 4.1: Comparison of the final QG sample selected from SE with Farmer. The first value in each cell is from Farmer and the second is from SE. The third galaxy is flagged as star in Farmer.

RA	DEC	z_{phot}	$log_{10}(M_*)$	$log_{10}(SFR)$	$log_{10}(sSFR)$
150.45457/150.45457	2.456/2.45599	3.764/3.578	10.716/10.766	-0.188/-0.449	-10.888/-11.249
150.5509/150.5509	2.17304/2.17304	3.611/3.590	10.6001/10.787	0.299/0.274	-10.300/-10.525
150.57693/150.57693	1.6711/1.67111	8.143/4.01	8.301/9.7	0.734/-4.3	-7.573/-14.0

Ilbert, O. et al., 2013; Valentino et al., 2020). In our case we have $log_{10}(sSFR)$ = (-14, -10.52, -11..25). This enhances the robustness and reliability of our selection.

Applying the same method in Farmer, I find that 227 galaxies qualify the check. The intersection between the $NUVrJ$ and this method is 49 galaxies. Looking at sSFR, 16 galaxies have $log_{10}(sSFR) < -10$ (only 2 of them have $log_{10}(sSFR) < -11$). When matching these 49 Farmer candidates with the 3 SE candidates I get no matches. Taking a closer look at the Farmer catalog, I find that one of them is indeed marked as star according to the χ^2-check. The other two though are very close to the borderline in $NUVrJ$.As in fact one of them is sitting exactly on the borderline. Keeping in mind the absence of errors, we can conclude that the results between the two catalogs (taking SE as a prior) are matching. In Table 4.1 the properties of the sources are presented. The fact that the majority of the properties of the two valid sources have similar values means probably that the photometric extraction process has less impact than the SED fitting technique which is responsible for deriving redshift and rest-frame physical properties.

4.4 The sample of quiescent galaxies

In the previous subsections i described how I executed the methods to successfully select quiescent galaxy candidates from the catalogs. Even though the methods

themselves have been tested and verified over the years in various works, the results of our analysis has to be cross-matched with the literature. In this section I will cross-match the QG sample and our catalog with other works that spectroscopically confirmed quiescent galaxy candidates in the COSMOS field.

Regardless the outcome of the cross-match it will give us valuable information. If the confirmed galaxies appear in the final sample, then we can be sure that the catalog was compiled correctly and at least reproduces the results of confirmed observations. In that case any additional galaxies that are in the sample and do not appear in the literature, have higher possibility of being an actual non-detected-before quiescent galaxy. In the case that none of the confirmed quiescent ones appear in the sample, then we have to check if they do exist in the catalog and what is the reason they do not make it into the final sample. If there is a satisfactory explanation for that (e.g. masked region) then we can also be hopeful that the final sample may contain useful results. In the unfortunate case that the confirmed galaxies do not appear at all in the catalog, then probably our results lack reliability.

High-z quiescent galaxies are faint objects and very difficult to find them. Even in the literature of the COSMOS field their numbers are very few. Thus, any mismatch between the results may be due to observational reasons (low S/N, bad photometry etc.). Therefore, before I cross-match directly with the high-z confirmed ones, I will compare with an intermediate-z sample ($z \sim 2$) from VLT/X-SHOOTER spectrograph (Stockmann et al., 2020). This test will reveal if the robustness of C20 photometry and redshift calculation robustness.

There are 14 galaxies in the COSMOS field that are spectroscopically confirmed in Stockmann et al. (2020). All of them appear in the SE version. In Table 4.2 a detailed presentation of all the sources are shown. There are 3 objects with invalid data due to masking. Two more are flagged as stars according the gzK_s diagram. The rest of the galaxies with valid data and flags are plotted in a z_{spec} vs z_{phot} in Fig 4.7. There is significant scatter between the two methods which could be a signal for possible systematic errors. However, all of them appear in the $NUVrJ$ quiescent locus but only 3 of them have $log_{10}(sSFR) < -10$.

In the same comparison with Farmer, 13 out 14 objects match in the catalog. It could be that the missing source is connected with Farmer's issue of source detection and coordinate assignment. There also 3 objects with invalid data (2 in a masked region, 1 recovery failure). All 10 valid sources sit in the $NUVrJ$ quiescent locus and 7 of them have $log_{10}(sSFR) < -11$.

The next step is to test our QG sample with confirmed high-z quiescent galaxies from the literature. The sample from the literature with the comparison is presented in Table 4.3. It is a total of 9 galaxies from Schreiber et al. (2018) and Valentino et al. (2020). Cross-matching our QG sample with those galaxies gives back 0 objects. However, 8 out of 9 of these galaxies appear in the catalog. Two of them have invalid redshift. A third one is marked with a redshift flag of an X-ray source (z=99.9). Also one of them is marked as a star according to the χ^2-check.

As I already mentioned, there is significant meaning on checking the reasons why the galaxies do not make it into the final sample. In the case of ZF-COS-20133, even though there is great difference between z_{spec} and z_{phot}, its confirmed z_{spec} is lower than our criteria. ZF-COS-14907 has also big difference between the two redshifts as it can be seen in the Table. However, this source has a very low confidence level (Schreiber et al., 2018). Same is true for ZF-COS- 10559 in an opposite way. It is marked as rejected in Schreiber et al. (2018) and assigned z_{spec}=2.637 but in SE has z_{phot}=3.993. An extra check on its data shows that it's very close to the borderline in $NUVrJ$ so if errors were considered it would probably be inside the

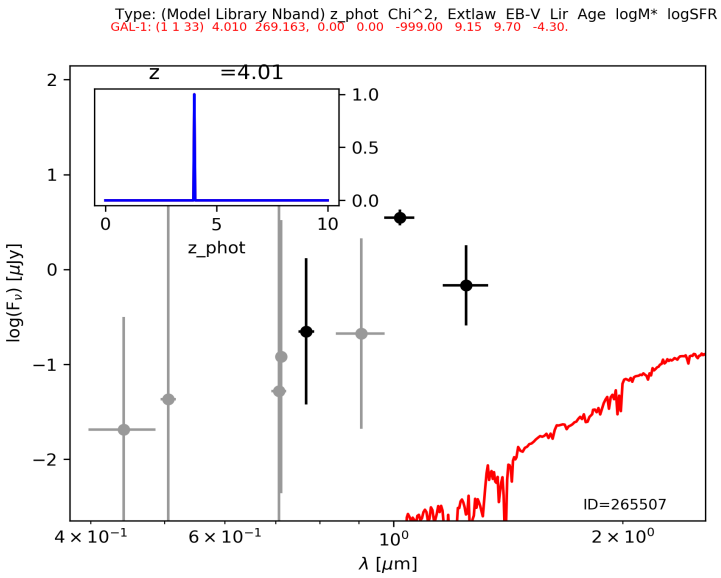


FIGURE 4.5: The failed SED fit of one of the sample sources.

quiescent locus. The most peculiar case is galaxy 466654 which is reported as a secure estimate (Valentino et al., 2020). Even though the two redshift values are very close (z_{phot}=3.642 & z_{spec}=3.775), in C20 SE catalog it appears as a blue dust-free star-forming galaxy. Not only it sits in the corresponding region in the $NUVrJ$ plane but it is also assigned an SFR higher than its main sequence counterpart. In the same manner, no confirmed source appears in the final QG sample of Farmer. In the full catalog, 5 out of 9 galaxies are recovered. Again we see that Farmer is missing sources which is probably connected to the source recovery problem.

Looking at the SEDs of the 3 candidates we see that one of them is a catastrophic error (see Fig 4.5). So we the final sample consists of two sources. Their SEDs can be seen in Fig 4.6. Their fit looks very reliable with both presenting a Balmer break that is being captured. The final QG sample of C20 SE is presented in Table 4.4. Even though the comparison with other samples was not the most promising one and can be translated as a signal for reconsidering certain configurations and parameters of the catalog, these 2 new candidates have been selected with strict criteria. Therefore, we can be positive of their potential as interesting spectroscopy targets that can give valuable information on mass assembly and quenching.

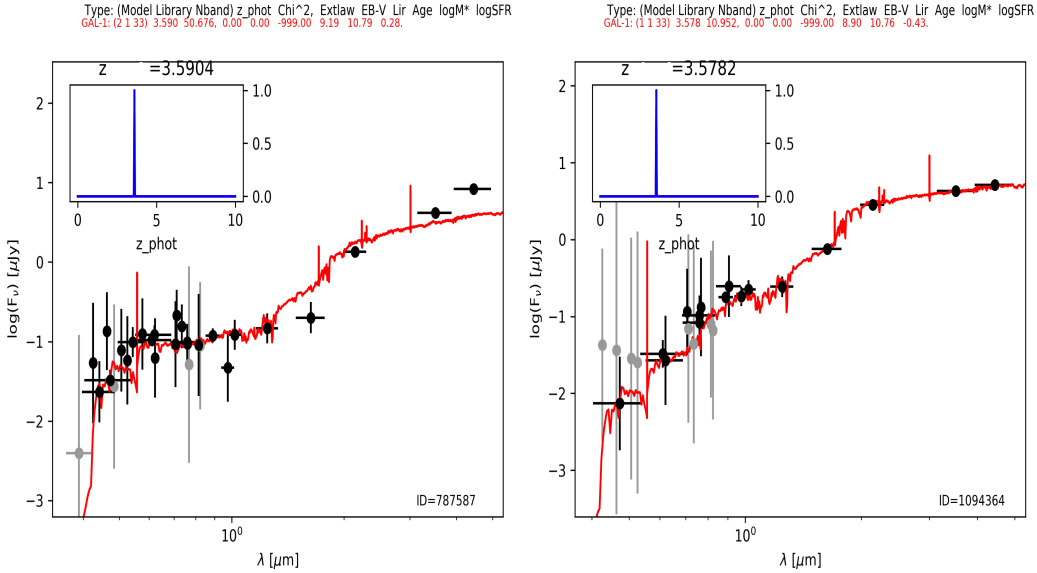


FIGURE 4.6: The SEDs of the two valid sources of the final QG sample.The top left panel represents the redshift probability distribution function.

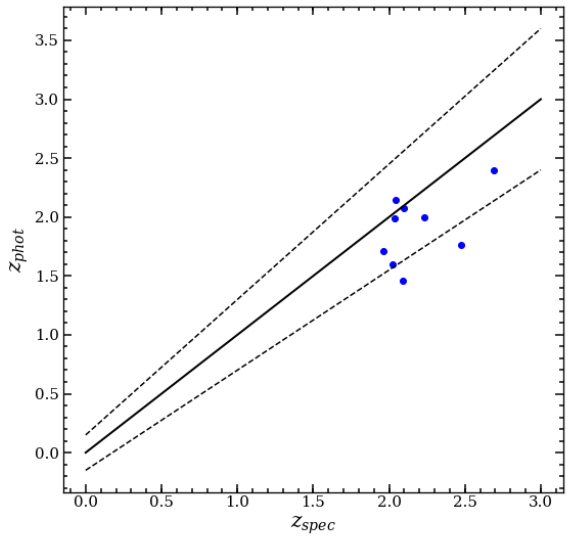


FIGURE 4.7: A comparison between the spectroscopic sample from (Stockmann et al., 2020) with the photo-z assigned by C20 SE. The dashed lines are $z_{phot} = z_{spec} \pm 0.15(1 + z_{spec}$ and represent the limits of acceptable scatter.

TABLE 4.2: Comparison of C20 SE with intermediate-z quiescent galaxies from (Stockmann et al., 2020)

Target ID	RA	DEC	z_{spec}	z_{phot}	$log_{10}(sSFR)$	Star
UV-108899	150.17661	2.0608871	2.2312	2.1436	-11.6	0
UV-250513	149.82227	2.6531196	2.0814	2.3996	-9.15357	0
CP-561356	150.20888	1.8502616	2.6963	1.9854	-13.29986	0
UV-105842	150.26265	2.0177791	2.0195	1.5989	-9.69816	0
UV-171687	149.88702	2.3506956	2.1020	2.0013	-9.34688	0
UV-90676	150.48750	2.2700379	2.4781	1.9943	-13.31517	1
CP-1291751	149.86954	2.3167057	2.0253	1.7626	-9.43427	0
UV-155853	149.55630	2.1672480	1.9816	1.4574	-9.80036	0
UV-171060	149.78951	2.3413286	2.0995	-99.0	-13.31974	0
UV-230929	150.20842	2.7721019	2.1679	2.076	-12.60532	0
UV-239220	149.43275	2.5106428	2.0057	-99.0	-8.3451	1
UV-773654	150.74574	2.0104926	2.0328	1.7076	-9.6117	0
CP-1243752	150.07394	2.2979755	2.0903	1.6407	-9.49676	1
CP-540713	150.32512	1.8185385	2.0409	-99.0	-8.93562	0

TABLE 4.3: Comparison of C20 SE with high-z quiescent galaxies

Target ID[a]	RA	DEC	z_{spec}	z_{phot}	$log_{10}(sSFR)$	Star
466654	149.41958	2.00755	3.775	3.6423	-8.0957	0
ZF-COS-10559	150.07147	2.2911844	2.637	3.9937	-8.59957	0
ZF-COS-14907	150.12422	2.3374486	4.193	2.0914	-8.14288	0
ZF-COS-17779	150.04651	2.3673911	3.415	99.9	-14.7287	1
ZF-COS-18842	150.08728	2.3960431	3.782	-	-	-
ZF-COS-19589	150.06671	2.3823645	3.715	-99.0	-8.10251	0
ZF-COS-20032	150.11256	2.3765266	2.473	4.4119	-8.94564	1
ZF-COS-20115	150.06149	2.3787093	3.714	-99.0	-11.2646	0
ZF-COS-20133	150.12173	2.374594	3.48	3.1044	-10.0769	0

[a] The first galaxy is from (Valentino et al., 2020). All the rest are from (Schreiber et al., 2018)

TABLE 4.4: Final sample of C20 SE high-z quiescent galaxy candidates

Target ID	RA	DEC	z_{phot}	$log_{10}(M_*)$	$log_{10}(SFR)$	$log_{10}(sSFR)$
787587	150.5509	2.17304	3.5904	10.78702	0.27457	-10.52523
1094364	150.45457	2.45599	3.5782	10.76675	-0.44999	-11.24999

Chapter 5

Discussion and Conclusion

5.1 Discussion

In this section I will discuss the results of our analysis along with the caveats and future perspectives.

5.1.1 Scientific significance of the quiescent sample.

The results of the COSMOS 2020 catalog analysis are a valuable resource for further scientific research. The 3 quiescent galaxy candidates can be excellent targets for follow-up spectroscopy. Studying their properties can reveal crucial information on galaxy mass assembly and star formation quenching. It was only recently that the first QG at $z > 3$ was spectroscopically confirmed (Glazebrook et al., 2017). After that, more confirmations followed collecting about 30 high-z massive quiescent galaxies (most of them at $z < 3.5$) which were extensively used for several studies. Schreiber et al. (2018) compared the star formation history and number density of quiescent galaxies at $3 < z < 4$ with different models. In Valentino et al. (2020), 3 quiescent galaxies at $3 < z < 4$ were targeted with spectroscopy to study the connection between quenching and potential sub-millimeter-bright progenitors at $4 < z < 5$. Finally, the largest sample of massive quiescent galaxies at $z > 3$ (Forrest et al., 2020) was used to confirm the extremely short star formation events as well as further study the connection of AGN hosts and early quenching.

Likewise, the new candidates selected here can be observed spectroscopically (or in X-ray) to confirm the existence of an accreting massive black hole in their centers. And since AGNs are not the only possible factor that can halt star formation, a scan on the local environment of the candidates (e.g., identifies nearby galaxies at similat redshift) can also reveal useful information. A cluster of galaxies indeed can produce interactions like gas heating or ram pressure that can prevent gas from cooling down.

It is true that our sample is small in size, including only 2 galaxies. However, there are only 7 known QGs at $z > 3.5$ in the literature. Our selection was done favoring purity over completeness. Although there are known issues in our analysis that will be discussed below, the implemented strategy is expected to result in a very small yet robust sample.

5.1.2 The Farmer results

The Farmer pipeline is an ambitious new method of photometric extraction that as I described in Chapter 2 is designed to solve known problems of traditional aperture extraction methods. Its approach is driven by physically motivated models and tools. Indeed, the results of the number counts showed that Farmer is able to recover

more sources per magnitude and in even larger depth than SE. However, the comparison with the HSC-SSP survey revealed that Farmer is prone in missing i-bright sources. Later in Chapter 4, the cross-match with the QGs from the literature showed that Farmer missed almost half of the confirmed sample. From this we can conclude that the missing sources problem of Farmer is not limited in i-bright sources (which are not the target of this catalog) but extends to the population we are interested in (high-z quiescent galaxies). As a result we cannot trust Farmer results in this state.

Investigating this open issue of Farmer would be complicated. First of all because the Farmer pipeline is still an ongoing-project. It is frequently updated with new configurations and parameters trying to fix known issues and bugs of previous versions. So it is still too early for a final stable version. In a few words, even if we investigated and eventually solved our current issue, there is no guarantee that there is not another problem that we missed with our tests and validation. Secondly, investigating the issues of Farmer pipeline is beyond the goals of this thesis.

Having said that, our analysis with Farmer was not unfruitful. The fact that we were able to discover this issue here, grants the COSMOS team this valuable knowledge. This will help them investigate the problem further, look for similar possible issues and eventually reform their strategy of the catalog.

5.1.3 `LePhare` configuration

Looking at the comparison to the literature it is clear that results are sub-optimal. Apart from our conservative selection (resulting in the rejection of a few targets) there are also confirmed sources that had a large difference between assigned z_{phot} and z_{spec}. On top of that, in Table 4.1 we can see that for the two valid candidates the derived physical properties have similar values. Therefore, we can safely assume that since there is this similarity, there is little dependence on the photometric extraction method (SE vs Farmer) on the derivation of the physical properties. On the contrary, the responsibility for these results falls on the SED fitting software.

The software used in this work to convert observed to rest-frame magnitudes and fit the SEDs to the data points is `LePhare` (Arnouts et al., 1999; Ilbert et al., 2006). From the results, it seems that `LePhare` is misconfigurated for our purpose. The configuration used was a more generic one with a balanced orientation between low-z and high-z solutions. However, since our catalog is aiming at selecting high-z sources, this configuration does not benefit our results. More specifically, the software does not include templates of passive and dusty galaxies. As a result galaxies with these features are fitted with other templates, ending up in either wrong redshifts or failed fittings. Even so, it is fair to say that adding these templates has its risk, since this will contaminate the quiescent sample with interlopers. Apart from that, there is also an intrinsic problem with the SED fitting at high-z. Since there are so few spectroscopically confirmed galaxies at $z > 3.5$, we have limited knowledge on the SEDs of such objects (Stefanon et al., 2015).

As we probe to higher redshifts, the objects become fainter and fainter with many of them pushing to the noise limit. In cases like that, the photometric uncertainties are quite impactful. Ignoring them lowers the reliability of our results. When these uncertainties are taken into account, the mean errors σ can reach up to 0.3 mag (Davidzon et al., 2017). Such values of errors, mean that the position of a source in the $NUVrJ$ diagram is precarious. This way, the selection is affected drastically. Consequently, incorporating rest-frame color errors would greatly benefit the selection and its robustness.

Nevertheless, the comparisons with the literature gave quite useful information. It showed us that even with state-of-the-art data, it is no trivial matter to successfully recover high-z sources. All parameters must be taken into account and tuned very carefully. The results of our literature comparison are a useful guide to the direction of optimal configuration.

5.1.4 The redshift distribution of the full sample.

The redshift distribution of the selected candidates from both methods can be instructive. Depending on the shape of the distribution we can comment on the contribution of C20 towards recovering quiescent candidates at $z \simeq 4$ which are extremely rare and faint. We look at both samples separately ($NUVrJ$ & SFR-M_*). Looking at the distribution in Fig 5.2 we see that the majority of the $NUVrj$ sample sits at $z < 4.3$. On the other hand, the SFR-M_* sample in Fig 5.1 shows a similar trend but since it is less than 10 objects, we cannot infer any strong conclusion. Still, this is an indication that even with significantly deeper data than before there is a barrier that makes the number of recovered galaxies at $z \geq 4$ extremely low.

5.1.5 The role of star/galaxy separation

We already mentioned the benefits of our conservative strategy. Purity over completeness gives us a reliable robust sample. The price we pay for that though is a very small sample.

In retrospect, the star/galaxy classification criteria I used is responsible for misclassifying possible high-z candidates as stars. For example in Fig 5.3 we can see the SED fitting of source ID1011306. This source is confirmed in the literature (see Table 4.3) but here is classified as a star. Had we adopted a more lenient criteria (e.g. $\chi^2_g - \chi^2_s > 0 \Rightarrow star$) then we would have avoided that.

The conservative strategy that we followed is only applicable here because we are interested in single targets. In case we wanted to perform statistical analysis (e.g. stellar mass function), then this strategy would have been inefficient. Restricting our sample would have led to high uncertainties. In order to perform a statistical analysis, completeness (with an acceptable limit of interlopers) is what we should aim for.

As we see both extremes have their pros and cons. A more "balanced" approach that could help mitigate those problems is the next step towards the future. Since we are dealing with data that are often very close to the detection limit, assigning classification probabilities might be more useful and realistic than decisive classifications (Henrion et al., 2011). Therefore, for future projects, working with Bayesian probabilistic approach for star/galaxy classification may help us overcome issues like that.

5.1.6 Follow up with future facilities

The scientific importance described in 5.1.1 justifies the call for spectroscopic follow-up of our targets. Facilities like the Multi-Object Spectrometer For Infra-Red Exploration (MOSFIRE McLean et al., 2010; McLean et al., 2012) at the Keck Observatory and X-SHOOTER spectrograph at the VLT (Vernet et al., 2011) have been traditionally excellent instruments for spectroscopy. All confirmed QGs at $z > 3$ mentioned or used for comparison purposes here were observed with either of them (Glazebrook et al., 2017; Schreiber et al., 2018; Stockmann et al., 2020; Valentino et al., 2020).

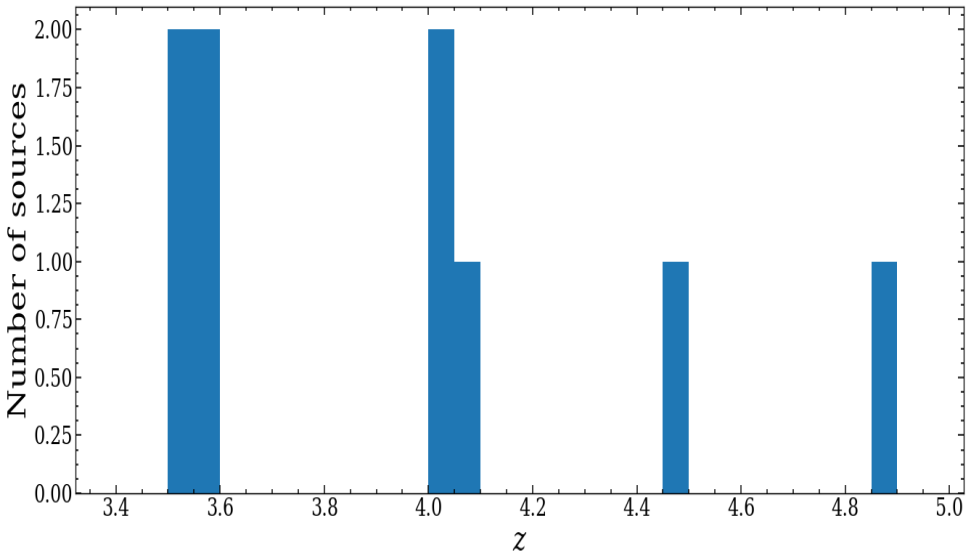


FIGURE 5.1: Redshift distribution in 0.05 bins of the SFR selected sample.

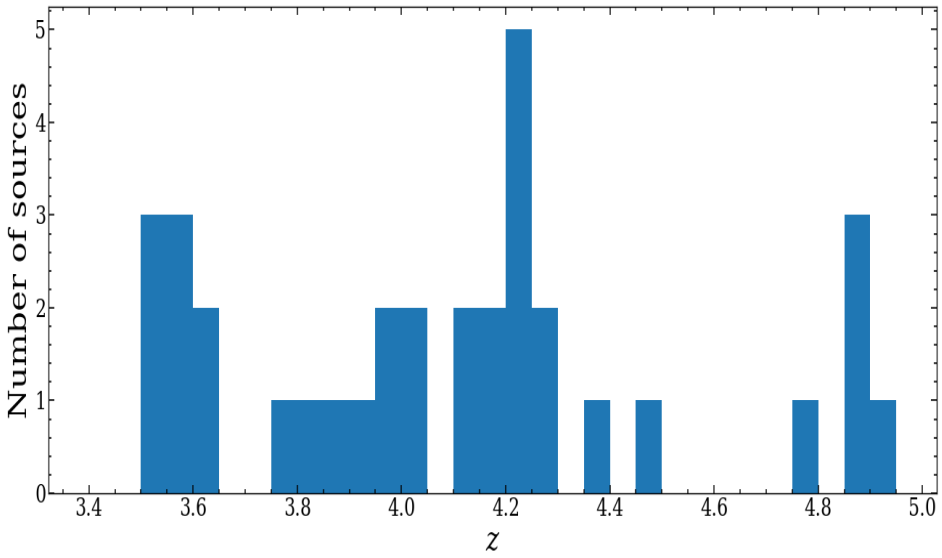


FIGURE 5.2: Redshift distribution in 0.05 bins of the *NUVrJ* selected sample.

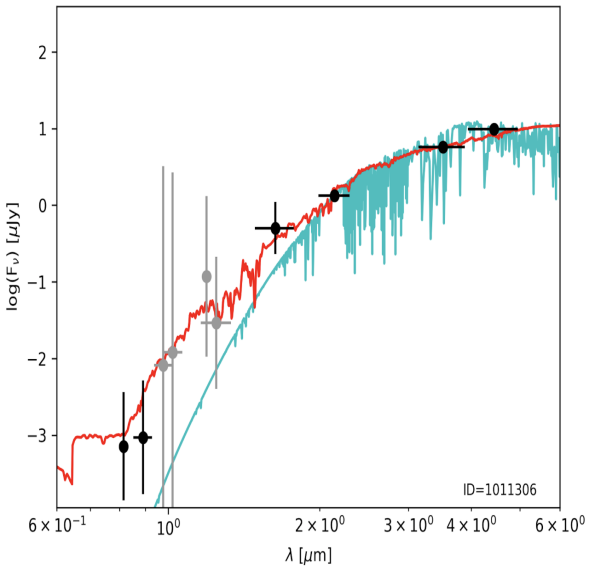


FIGURE 5.3: The SED fitting of the data points of source ID=1011306. The red spectrum is a galaxy template, while the cyan spectrum a stellar template. This source is misclassified as a star due to the strict criteria in the χ^2-check.

Indeed their wide wavelength coverage combined with the high resolution makes them unparalleled in their field.

However we are entering a new era of space observations with the upcoming launch of *James Webb Space Telescope* (JWST). This next generation telescope will open up great potential with its state-of-the-art instruments. In fact, a simple comparison shows that utilizing the capabilities of JWST Near-Infrared Spectrograph (NIRSpec) enhances accuracy and saves resources.

For example, the confirmed quiescent galaxy COS-466654 with z_{spec}=3.775 needed an exposure time of 8.6 hours with VLT/X-SHOOTER (Valentino et al., 2020). The same observation with NIRSpec would take $\sim 1h20m$ to achieve a similar S/N as with X-SHOOTER (F. Valentino, private communication). On top of that, COS-466654 has a magnitude of $K_s = 22.26$ mag which is much brighter than the majority of the candidates in our full sample. Comparing the median magnitude of our full sample ($K_s \sim 24.95$ mag) with the corresponding of the COSMOS galaxies in Schreiber et al. (2018) ($K_s \sim 24.69$ mag), it is obvious that our sample is more resource-demanding.

In terms of technical capabilities the NIRSpec instrument is remarkable at meeting the increased demand of high-z targets. Its wavelength coverage (0.6-5.3μm) is three times wider than MOSFIRE (0.97-2.14μm) and more than double of X-SHOOTER (0.3-2.5μm). Therefore, at $z > 3.5$ the NIRSpec is able to capture emission and absorption lines of elements that with MOSFIRE and X-SHOOTER fall outside of the coverage. Strong lines like CaI, FeI, [NII] and $H\alpha$ are an example and as we move to higher z more lines are missed from the ground. Nevertheless, NIRSpec resolving power ($R \sim 2700$) is less than the maximum capacity of MOSFIRE ($R \sim 3500$) and X-SHOOTER ($R \sim 4000$) but the fact that as a space-based facility bypasses atmospheric noise, absorption and turbulence, makes up for that.

In addition, NIRSpec is equipped with an integral field unit (IFU) that slices the

image into smaller cells and extracts the spectrum out of each cell. On single targets this feature is useful since it can trace gas motions inside the galaxy through the Doppler effect. Furthermore, this technology is ideal for multiple-targeting of sources as well. Sources extended over a few arcseconds in the field of view, can be targeted with a single exposure. It is extremely useful in cases of catalogs extracted over a small field of view like ours. This way all the candidates can be observed with just one exposure. Ground-based IFUs like K-band Multi Object Spectrograph (KMOS), Visual and IR Multi-Object Spectrograph (VIMOS) and Fibre Large Array Multi Element Spectrograph (FLAMES) have produced excellent results over the years. But as mentioned before the wavelength coverage of these instruments is limited compared to NIRSpec. Looking forward to the future, we want to be able to break the limit of $z \sim 4$. At $z > 4.5$ the Balmer break is missed by the best ground-based instruments limited at 2.5μm. We do not know if quenched galaxies exist at $z > 4.5$, but NIRSpec will help us answer that question.

5.2 Conclusion

In this thesis I have described and presented COSMOS 2020, a new deep multi-wavelength large survey including photometric data from 24 broadband filters and precise physical properties. The current version of the catalog is compiled with remarkably deeper data in almost all bands. In comparison with COSMOS 2015, the limiting depth is increased by 0.5 mag. A novelty introduced for the first time here is `Farmer` a new photometric extraction method that was used alongside the traditional software `SExtractor`. I described the principles of operation of both methods. Farmer is an extraction method using model fitting on the sources contrary to SE that uses apertures to extract the flux.

Using the number counts tool I confirmed that the catalog is in agreement with previous works from the literature and in fact is reaching deeper limiting depth than most of them. Farmer version seemed to recover more and fainter sources than SE. However, in a cross-match comparison with HSC-SSP, an *i*-selected catalog, it was seen that Farmer is failing at recovering *i*-bright sources while SE is more consistent. Later I described the process of SED fitting for the derivation of z_{phot} and presented a comparison with a confirmed spectroscopic sample from the literature. The results seem to be in very good agreement with the outliers fraction being $\eta = 3.1\%$ for SE and $\eta = 3.2\%$ for Farmer. As an additional method to validate the catalog, I plotted the results in the gzK_s color-color diagram. This diagram breaks the metallicity degeneracy and concentrates the galaxies in certain regions. The diagram agrees excellent with the literature so we are convinced about the validity of the photometric extraction.

Moving on with the analysis of the sample, I started with the star/galaxy separation. Using the gzK_s as well as a chi^2-check from the template fitting, I classified the sources accordingly to our criteria. Later, I used the $NUVrJ$ diagram and the SFR-M_* correlation formula to select the quiescent galaxies from the sample. Because of the conservative strategy that we followed on our selections and criteria that favored purity over completeness, the final intersection between the two selection methods contained 2 candidates. To validate the results, I cross-matched them with spectroscopically confirmed galaxies from the literature. The comparison revealed that due to a combination of selection criteria and parameter configuration, the catalog is missing a non-negligible number of confirmed sources. Nevertheless, all the weak

spots of the catalog revealed are valuable for future corrections. I hope that the final sample provided by this work can help ,through future follow-up observations, shed light on the fundamental questions of galaxy formation and evolution.

Bibliography

Aihara, Hiroaki et al. (Oct. 2017). "First data release of the Hyper Suprime-Cam Subaru Strategic Program". In: *Publications of the Astronomical Society of Japan* 70.SP1. ISSN: 2053-051X. DOI: 10.1093/pasj/psx081. URL: http://dx.doi.org/10.1093/pasj/psx081.

Aihara, Hiroaki et al. (Oct. 2019). "Second data release of the Hyper Suprime-Cam Subaru Strategic Program". In: *Publications of the Astronomical Society of Japan*. ISSN: 2053-051X. DOI: 10.1093/pasj/psz103. URL: http://dx.doi.org/10.1093/pasj/psz103.

Alatalo, Katherine et al. (Jan. 2015). "Suppression of Star Formation in NGC 1266". In: ApJ 798.1, 31, p. 31. DOI: 10.1088/0004-637X/798/1/31. arXiv: 1410.4556 [astro-ph.GA].

Andrae, Rene, Tim Schulze-Hartung, and Peter Melchior (2010). *Dos and don'ts of reduced chi-squared*. arXiv: 1012.3754 [astro-ph.IM].

Arnouts, S. et al. (Dec. 1999). "Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North". In: MNRAS 310.2, pp. 540–556. DOI: 10.1046/j.1365-8711.1999.02978.x. arXiv: astro-ph/9902290 [astro-ph].

Arnouts, S. et al. (Dec. 2007). "The SWIRE-VVDS-CFHTLS surveys: stellar mass assembly over the last 10 Gyr. Evidence for a major build up of the red sequence between z = 2 and z = 1". In: A&A 476.1, pp. 137–150. DOI: 10.1051/0004-6361:20077632. arXiv: 0705.2438 [astro-ph].

Ashby, M. L. N. et al. (May 2013). "SEDS: The Spitzer Extended Deep Survey. Survey Design, Photometry, and Deep IRAC Source Counts". In: ApJ 769.1, 80, p. 80. DOI: 10.1088/0004-637X/769/1/80.

Ashby, M. L. N. et al. (June 2015). "S-CANDELS: The Spitzer-Cosmic Assembly Near-Infrared Deep Extragalactic Survey. Survey Design, Photometry, and Deep IRAC Source Counts". In: ApJS 218.2, 33, p. 33. DOI: 10.1088/0067-0049/218/2/33. arXiv: 1506.01323 [astro-ph.GA].

Ashby, M. L. N. et al. (Aug. 2018). "Spitzer Matching Survey of the UltraVISTA Ultradeep Stripes (SMUVS): Full-mission IRAC Mosaics and Catalogs". In: ApJS 237.2, 39, p. 39. DOI: 10.3847/1538-4365/aad4fb. arXiv: 1801.02660 [astro-ph.GA].

Baldry, I. K. et al. (Dec. 2004). "Color bimodality: Implications for galaxy evolution". In: *The New Cosmology: Conference on Strings and Cosmology*. Ed. by Roland E. Allen, Dimitri V. Nanopoulos, and Christopher N. Pope. Vol. 743. American Institute of Physics Conference Series, pp. 106–119. DOI: 10.1063/1.1848322. arXiv: astro-ph/0410603 [astro-ph].

Balogh, Michael L. et al. (Nov. 2004). "The Bimodal Galaxy Color Distribution: Dependence on Luminosity and Environment". In: ApJ 615.2, pp. L101–L104. DOI: 10.1086/426079. arXiv: astro-ph/0406266 [astro-ph].

Baum, W. A. (Jan. 1962). "Photoelectric Magnitudes and Red-Shifts". In: *Problems of Extra-Galactic Research*. Ed. by George Cunliffe McVittie. Vol. 15. IAU Symposium, p. 390.

Behroozi, Peter S., Risa H. Wechsler, and Charlie Conroy (June 2013). "The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8". In:

ApJ 770.1, 57, p. 57. DOI: 10.1088/0004-637X/770/1/57. arXiv: 1207.6105 [astro-ph.CO].

Belli, Sirio et al. (May 2017). "KMOS^{3D} Reveals Low-level Star Formation Activity in Massive Quiescent Galaxies at 0.7 < z < 2.7". In: ApJ 841.1, L6, p. L6. DOI: 10.3847/2041-8213/aa70e5. arXiv: 1703.07778 [astro-ph.GA].

Bertin, E. and S. Arnouts (June 1996). "SExtractor: Software for source extraction." In: A&AS 117, pp. 393–404. DOI: 10.1051/aas:1996164.

Blanc, Guillermo A. et al. (July 2008). "The Multiwavelength Survey by Yale-Chile (MUSYC): Wide K-Band Imaging, Photometric Catalogs, Clustering, and Physical Properties of Galaxies at z ∼2". In: ApJ 681.2, pp. 1099–1115. DOI: 10.1086/588018. arXiv: 0803.0763 [astro-ph].

Bolzonella, M. et al. (Nov. 2010). "Tracking the impact of environment on the galaxy stellar mass function up to $z \sim 1$ in the 10k zCOSMOS sample". In: *Astronomy and Astrophysics* 524, A76. ISSN: 1432-0746. DOI: 10.1051/0004-6361/200912801. URL: http://dx.doi.org/10.1051/0004-6361/200912801.

Bolzonella, Micol, Joan-Marc Miralles, and Roser Pello' (2000). *Photometric Redshifts based on standard SED fitting procedures*. arXiv: astro-ph/0003380 [astro-ph].

Boucaud, Alexandre et al. (Jan. 2020). "Photometry of high-redshift blended galaxies using deep learning". In: MNRAS 491.2, pp. 2481–2495. DOI: 10.1093/mnras/stz3056. arXiv: 1905.01324 [astro-ph.GA].

Bower, R. G. et al. (Aug. 2006). "Breaking the hierarchy of galaxy formation". In: *Monthly Notices of the Royal Astronomical Society* 370.2, 645–655. ISSN: 1365-2966. DOI: 10.1111/j.1365-2966.2006.10519.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2006.10519.x.

Brinchmann, J. et al. (July 2004). "The physical properties of star-forming galaxies in the low-redshift Universe". In: MNRAS 351.4, pp. 1151–1179. DOI: 10.1111/j.1365-2966.2004.07881.x. arXiv: astro-ph/0311060 [astro-ph].

Bruzual, G. and S. Charlot (Oct. 2003). "Stellar population synthesis at the resolution of 2003". In: MNRAS 344.4, pp. 1000–1028. DOI: 10.1046/j.1365-8711.2003.06897.x. arXiv: astro-ph/0309134 [astro-ph].

Bruzual A., Gustavo and Stephane Charlot (Mar. 1993). "Spectral Evolution of Stellar Populations Using Isochrone Synthesis". In: ApJ 405, p. 538. DOI: 10.1086/172385.

Bundy, Kevin et al. (Nov. 2006). "The Mass Assembly History of Field Galaxies: Detection of an Evolving Mass Limit for Star-Forming Galaxies". In: *The Astrophysical Journal* 651.1, 120–141. ISSN: 1538-4357. DOI: 10.1086/507456. URL: http://dx.doi.org/10.1086/507456.

Capak, P. et al. (Sept. 2007). "The First Release COSMOS Optical and Near-IR Data and Catalog". In: *The Astrophysical Journal Supplement Series* 172.1, 99–116. ISSN: 1538-4365. DOI: 10.1086/519081. URL: http://dx.doi.org/10.1086/519081.

Casey, C. M. et al. (Dec. 2012). "A Redshift Survey of Herschel Far-infrared Selected Starbursts and Implications for Obscured Star Formation". In: ApJ 761.2, 140, p. 140. DOI: 10.1088/0004-637X/761/2/140. arXiv: 1210.4928 [astro-ph.CO].

Chabrier, Gilles (July 2003). "Galactic Stellar and Substellar Initial Mass Function". In: *Publications of the Astronomical Society of the Pacific* 115.809, 763–795. ISSN: 1538-3873. DOI: 10.1086/376392. URL: http://dx.doi.org/10.1086/376392.

Cimatti, A. et al. (Aug. 2002). "The K20 survey". In: *Astronomy and Astrophysics* 392.2, 395–406. ISSN: 1432-0746. DOI: 10.1051/0004-6361:20020861. URL: http://dx.doi.org/10.1051/0004-6361:20020861.

Ciotti, Luca et al. (Aug. 1991). "Winds, Outflows, and Inflows in X-Ray Elliptical Galaxies. I." In: ApJ 376, p. 380. DOI: 10.1086/170289.

Collaboration, Planck et al. (2018). *Planck 2018 results. VI. Cosmological parameters*. arXiv: 1807.06209 [astro-ph.CO].

Cowie, Lennox L. et al. (Sept. 1996). "New Insight on Galaxy Formation and Evolution From Keck Spectroscopy of the Hawaii Deep Fields". In: AJ 112, p. 839. DOI: 10.1086/118058. arXiv: astro-ph/9606079 [astro-ph].

Croton, Darren J. et al. (Jan. 2006). "The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies". In: MNRAS 365.1, pp. 11–28. DOI: 10.1111/j.1365-2966.2005.09675.x. arXiv: astro-ph/0508046 [astro-ph].

Daddi, E. et al. (June 2005). "Passively Evolving Early-Type Galaxies at 1.4 $\leq$z $\leq$ 2.5 in the Hubble Ultra Deep Field". In: *The Astrophysical Journal* 626.2, pp. 680–697. DOI: 10.1086/430104. URL: https://doi.org/10.1086%2F430104.

Daddi, E. et al. (Nov. 2007). "Multiwavelength Study of Massive Galaxies at z$\sim$2. I. Star Formation and Galaxy Growth". In: ApJ 670.1, pp. 156–172. DOI: 10.1086/521818. arXiv: 0705.2831 [astro-ph].

Davidzon, I. et al. (Sept. 2017). "The COSMOS2015 galaxy stellar mass function . Thirteen billion years of stellar mass assembly in ten snapshots". In: A&A 605, A70, A70. DOI: 10.1051/0004-6361/201730419. arXiv: 1701.02734 [astro-ph.GA].

Davidzon, Iary et al. (Jan. 2018). "An Aernate Approach to Measure Specific Star Formation Rates at 2\z\7". In: ApJ 852.2, 107, p. 107. DOI: 10.3847/1538-4357/aaa19e. arXiv: 1712.03959 [astro-ph.GA].

Davis, M. et al. (May 1985). "The evolution of large-scale structure in a universe dominated by cold dark matter". In: ApJ 292, pp. 371–394. DOI: 10.1086/163168.

de Vaucouleurs, Gerard (Jan. 1948). "Recherches sur les Nebuleuses Extragalactiques". In: *Annales d'Astrophysique* 11, p. 247.

Di Matteo, Tiziana, Volker Springel, and Lars Hernquist (Feb. 2005). "Energy input from quasars regulates the growth and activity of black holes and their host galaxies". In: Nature 433.7026, pp. 604–607. DOI: 10.1038/nature03335. arXiv: astro-ph/0502199 [astro-ph].

Dickinson, Mark et al. (Apr. 2003). "The Evolution of the Global Stellar Mass Density at 0<z<3". In: ApJ 587.1, pp. 25–40. DOI: 10.1086/368111. arXiv: astro-ph/0212242 [astro-ph].

Dressler, A. (Apr. 1980). "A catalog of morphological types in 55 rich clusters of galaxies". In: ApJS 42, pp. 565–609. DOI: 10.1086/190663.

Díaz-García, L. A. et al. (Nov. 2019). "Stellar populations of galaxies in the ALHAMBRA survey up to z 1". In: *Astronomy Astrophysics* 631, A156. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201832788. URL: http://dx.doi.org/10.1051/0004-6361/201832788.

Elbaz, D. et al. (June 2007). "The reversal of the star formation-density relation in the distant universe". In: A&A 468.1, pp. 33–48. DOI: 10.1051/0004-6361:20077525. arXiv: astro-ph/0703653 [astro-ph].

Fang, Guan-Wen et al. (May 2015). "Morphology and structure of BzK-selected galaxies atz 2 in the CANDELS-COSMOS field". In: *Research in Astronomy and Astrophysics* 15.6, 819–827. ISSN: 1674-4527. DOI: 10.1088/1674-4527/15/6/005. URL: http://dx.doi.org/10.1088/1674-4527/15/6/005.

Feldmann, Robert and Lucio Mayer (Jan. 2015). "The Argo simulation - I. Quenching of massive galaxies at high redshift as a result of cosmological starvation". In: MNRAS 446.2, pp. 1939–1956. DOI: 10.1093/mnras/stu2207. arXiv: 1404.3212 [astro-ph.GA].

Fontana, A. et al. (Dec. 2006). "The Galaxy mass function up to z =4 in the GOODS-MUSIC sample: into the epoch of formation of massive galaxies". In: A&A 459.3,

pp. 745–757. DOI: 10.1051/0004-6361:20065475. arXiv: astro-ph/0609068 [astro-ph].

Fontana, A. et al. (Sept. 2014). "The Hawk-I UDS and GOODS survey (HUGS): Survey design and deep K-band number counts". In: *Astronomy and Astrophysics* 570. DOI: 10.1051/0004-6361/201423543.

Fontanot, Fabio et al. (Aug. 2009). "The many manifestations of downsizing: hierarchical galaxy formation models confront observations". In: MNRAS 397.4, pp. 1776–1790. DOI: 10.1111/j.1365-2966.2009.15058.x. arXiv: 0901.1130 [astro-ph.CO].

Forrest, Ben et al. (Sept. 2020). "The Massive Ancient Galaxies At $z > 3$ NEar-infrared (MAGAZ3NE) Survey: Confirmation of Extremely Rapid Star-Formation and Quenching Timescales for Massive Galaxies in the Early Universe". In: *arXiv e-prints*, arXiv:2009.07281, arXiv:2009.07281. arXiv: 2009.07281 [astro-ph.GA].

Galametz, Audrey et al. (June 2013). "CANDELS Multiwavelength Catalogs: Source Identification and Photometry in the CANDELS UKIDSS Ultra-deep Survey Field". In: ApJS 206.2, 10, p. 10. DOI: 10.1088/0067-0049/206/2/10. arXiv: 1305.1823 [astro-ph.CO].

Gallazzi, Anna et al. (Sept. 2005). "The ages and metallicities of galaxies in the local universe". In: MNRAS 362.1, pp. 41–58. DOI: 10.1111/j.1365-2966.2005.09321.x. arXiv: astro-ph/0506539 [astro-ph].

Gavazzi, G., D. Pierini, and A. Boselli (Aug. 1996). "The phenomenology of disk galaxies." In: A&A 312, pp. 397–408.

Girardi, Marisa et al. (Apr. 2002). "Observational Mass-to-Light Ratio of Galaxy Systems from Poor Groups to Rich Clusters". In: ApJ 569.2, pp. 720–741. DOI: 10.1086/339360. arXiv: astro-ph/0112534 [astro-ph].

Glazebrook, Karl et al. (Apr. 2017). "A massive, quiescent galaxy at a redshift of 3.717". In: Nature 544.7648, pp. 71–74. DOI: 10.1038/nature21680. arXiv: 1702.01751 [astro-ph.GA].

Guo, Yicheng et al. (Apr. 2012). "Rest-frame UV-Optically Selected Galaxies at 2.3 <∼z <∼3.5: Searching for Dusty Star-forming and Passively Evolving Galaxies". In: ApJ 749.2, 149, p. 149. DOI: 10.1088/0004-637X/749/2/149. arXiv: 1110.3801 [astro-ph.CO].

Hasinger, G. et al. (May 2018). "The DEIMOS 10K Spectroscopic Survey Catalog of the COSMOS Field". In: ApJ 858.2, 77, p. 77. DOI: 10.3847/1538-4357/aabacf. arXiv: 1803.09251 [astro-ph.GA].

Henrion, Marc et al. (2011). "A Bayesian approach to star-galaxy classification". In: *Monthly Notices of the Royal Astronomical Society* 412.4, 2286–2302. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2010.18055.x. URL: http://dx.doi.org/10.1111/j.1365-2966.2010.18055.x.

Hildebrandt, H. et al. (Nov. 2010). "PHAT: PHoto-z Accuracy Testing". In: A&A 523, A31, A31. DOI: 10.1051/0004-6361/201014885. arXiv: 1008.0658 [astro-ph.CO].

Hoaglin, David C., Frederick Mosteller, and John W. Tukey (1983). *Understanding robust and exploratory data anlysis*.

Howell, Steve B. (2006). *Handbook of CCD Astronomy*. 2nd ed. Cambridge Observing Handbooks for Research Astronomers. Cambridge University Press. DOI: 10.1017/CBO9780511807909.

Hubble, Edwin. (Jan. 1926). "No. 324. Extra-galactic nebulae." In: *Contributions from the Mount Wilson Observatory / Carnegie Institution of Washington* 324, pp. 1–49.

Hubble, Edwin (1929). "A relation between distance and radial velocity among extra-galactic nebulae". In: *Proceedings of the National Academy of Sciences* 15.3, pp. 168–173. ISSN: 0027-8424. DOI: 10.1073/pnas.15.3.168. eprint: https://www.pnas.

org/content/15/3/168.full.pdf. URL: https://www.pnas.org/content/15/3/168.

Ilbert, O. et al. (Oct. 2006). "Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey". In: A&A 457.3, pp. 841–856. DOI: 10.1051/0004-6361:20065138. arXiv: astro-ph/0603217 [astro-ph].

Ilbert, O. et al. (Dec. 2008). "COSMOS PHOTOMETRIC REDSHIFTS WITH 30-BANDS FOR 2-deg2". In: *The Astrophysical Journal* 690.2, 1236–1249. ISSN: 1538-4357. DOI: 10.1088/0004-637x/690/2/1236. URL: http://dx.doi.org/10.1088/0004-637X/690/2/1236.

Ilbert, O. et al. (Jan. 2009). "Cosmos Photometric Redshifts with 30-Bands for 2-deg^2". In: ApJ 690.2, pp. 1236–1249. DOI: 10.1088/0004-637X/690/2/1236. arXiv: 0809.2101 [astro-ph].

Ilbert, O. et al. (Feb. 2010). "Galaxy Stellar Mass Assembly Between 0.2 < z < 2 from the S-COSMOS Survey". In: ApJ 709.2, pp. 644–663. DOI: 10.1088/0004-637X/709/2/644. arXiv: 0903.0102 [astro-ph.CO].

Ilbert, O. et al. (2013). "Mass assembly in quiescent and star-forming galaxies since z from UltraVISTA". In: *A&A* 556, A55. DOI: 10.1051/0004-6361/201321100. URL: https://doi.org/10.1051/0004-6361/201321100.

Ishibashi, W. and A. C. Fabian (Dec. 2012). "Active galactic nucleus feedback and triggering of star formation in galaxies". In: MNRAS 427.4, pp. 2998–3005. DOI: 10.1111/j.1365-2966.2012.22074.x. arXiv: 1209.1480 [astro-ph.GA].

Jarvis, Matt J. et al. (Oct. 2012). "The VISTA Deep Extragalactic Observations (VIDEO) survey". In: *Monthly Notices of the Royal Astronomical Society* 428.2, 1281–1295. ISSN: 0035-8711. DOI: 10.1093/mnras/sts118. URL: http://dx.doi.org/10.1093/mnras/sts118.

Kartaltepe, Jeyhan S. et al. (Sept. 2010). "A Multiwavelength Study of a Sample of 70 μm Selected Galaxies in the COSMOS Field. II. The Role of Mergers in Galaxy Evolution". In: ApJ 721.1, pp. 98–123. DOI: 10.1088/0004-637X/721/1/98. arXiv: 1006.4956 [astro-ph.CO].

Kashino, Daichi et al. (2019). "The FMOS-COSMOS Survey of Star-forming Galaxies at z 1.6. VI. Redshift and Emission-line Catalog and Basic Properties of Star-forming Galaxies". In: *The Astrophysical Journal Supplement Series* 241.1, p. 10. ISSN: 1538-4365. DOI: 10.3847/1538-4365/ab06c4. URL: http://dx.doi.org/10.3847/1538-4365/ab06c4.

Kauffmann, Guinevere et al. (Sept. 2004). "The environmental dependence of the relations between stellar mass, structure, star formation and nuclear activity in galaxies". In: MNRAS 353.3, pp. 713–731. DOI: 10.1111/j.1365-2966.2004.08117.x. arXiv: astro-ph/0402030 [astro-ph].

Kennicutt, Robert C. (Sept. 1998). "STAR FORMATION IN GALAXIES ALONG THE HUBBLE SEQUENCE". In: *Annual Review of Astronomy and Astrophysics* 36.1, 189–231. ISSN: 1545-4282. DOI: 10.1146/annurev.astro.36.1.189. URL: http://dx.doi.org/10.1146/annurev.astro.36.1.189.

Kennicutt, Robert C. et al. (Sept. 2009). "DUST-CORRECTED STAR FORMATION RATES OF GALAXIES. I. COMBINATIONS OF H AND INFRARED TRACERS". In: *The Astrophysical Journal* 703.2, 1672–1695. ISSN: 1538-4357. DOI: 10.1088/0004-637x/703/2/1672. URL: http://dx.doi.org/10.1088/0004-637X/703/2/1672.

Khochfar, Sadegh and Jeremiah P. Ostriker (June 2008). "Adding Environmental Gas Physics to the Semianalytic Method for Galaxy Formation: Gravitational Heating". In: ApJ 680.1, pp. 54–69. DOI: 10.1086/587470. arXiv: 0704.2418 [astro-ph].

Khoperskov, S. et al. (Jan. 2018). "Bar quenching in gas-rich galaxies". In: A&A 609, A60, A60. DOI: 10.1051/0004-6361/201731211. arXiv: 1709.03604 [astro-ph.GA].

Koo, David C. and Richard G. Kron (1992). "Evidence for Evolution in Faint Field Galaxy Samples". In: *Annual Review of Astronomy and Astrophysics* 30.1, pp. 613–652. DOI: 10.1146/annurev.aa.30.090192.003145. eprint: https://doi.org/10.1146/annurev.aa.30.090192.003145. URL: https://doi.org/10.1146/annurev.aa.30.090192.003145.

Kron, R. G. (June 1980). "Photometry of a complete sample of faint galaxies." In: ApJS 43, pp. 305–325. DOI: 10.1086/190669.

Labbé, I. et al. (Nov. 2013). "The Spectral Energy Distributions of z ~8 Galaxies from the IRAC Ultra Deep Fields: Emission Lines, Stellar Masses, and Specific Star Formation Rates at 650 Myr". In: ApJ 777.2, L19, p. L19. DOI: 10.1088/2041-8205/777/2/L19. arXiv: 1209.3037 [astro-ph.CO].

Laigle, C. et al. (June 2016). "THE COSMOS2015 CATALOG: EXPLORING THE 1>z<6 UNIVERSE WITH HALF A MILLION GALAXIES". In: *The Astrophysical Journal Supplement Series* 224.2, p. 24. DOI: 10.3847/0067-0049/224/2/24. URL: https://doi.org/10.3847%2F0067-0049%2F224%2F2%2F24.

Lane, K. P. et al. (July 2007). "The colour selection of distant galaxies in the UKIDSS Ultra Deep Survey Early Data Release". In: MNRAS 379.1, pp. L25–L29. DOI: 10.1111/j.1745-3933.2007.00327.x. arXiv: 0704.2136 [astro-ph].

Lang, Dustin, David W. Hogg, and David Mykytyn (Apr. 2016). *The Tractor: Probabilistic astronomical source detection and measurement*. ascl: 1604.008.

Le Fevre, O. et al. (Jan. 1986). "Electronographic BV photometry of three distant clusters of galaxies." In: A&A 154, pp. 92–99.

Le Fèvre, O. et al. (Apr. 2015). "The VIMOS Ultra-Deep Survey: ~10 000 galaxies with spectroscopic redshifts to study galaxy assembly at early epochs 2 < z 6". In: A&A 576, A79, A79. DOI: 10.1051/0004-6361/201423829. arXiv: 1403.3938 [astro-ph.CO].

Le Fèvre, O. et al. (Oct. 2013). "The VIMOS VLT Deep Survey final data release: a spectroscopic sample of 35016 galaxies and AGN out toz 6.7 selected with 17.5 iAB 24.75". In: *Astronomy and Astrophysics* 559, A14. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201322179. URL: http://dx.doi.org/10.1051/0004-6361/201322179.

Leja, Joel, Sandro Tacchella, and Charlie Conroy (July 2019). "Beyond UVJ: More Efficient Selection of Quiescent Galaxies with Ultraviolet/Mid-infrared Fluxes". In: ApJ 880.1, L9, p. L9. DOI: 10.3847/2041-8213/ab2f8c. arXiv: 1907.02970 [astro-ph.GA].

Lilly, S. J. et al. (Sept. 2007). "zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering 0 < z < 3 in the COSMOS Field". In: ApJS 172.1, pp. 70–85. DOI: 10.1086/516589. arXiv: astro-ph/0612291 [astro-ph].

Madau, Piero and Mark Dickinson (Aug. 2014). "Cosmic Star-Formation History". In: ARA&A 52, pp. 415–486. DOI: 10.1146/annurev-astro-081811-125615. arXiv: 1403.0007 [astro-ph.CO].

Maddox, S. J., G. Efstathiou, and W. J. Sutherland (Oct. 1990). "The APM galaxy survey. II Photometric corrections." In: MNRAS 246, p. 433.

Man, Allison and Sirio Belli (Sept. 2018). "Star formation quenching in massive galaxies". In: *Nature Astronomy* 2.9, pp. 695–697. ISSN: 2397-3366. DOI: 10.1038/s41550-018-0558-1. URL: https://doi.org/10.1038/s41550-018-0558-1.

Maraston, Claudia (Sept. 2005). "Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies". In: MNRAS 362.3,

pp. 799–825. DOI: 10.1111/j.1365-2966.2005.09270.x. arXiv: astro-ph/0410207 [astro-ph].

Martig, Marie et al. (Dec. 2009). "Morphological Quenching of Star Formation: Making Early-Type Galaxies Red". In: ApJ 707.1, pp. 250–267. DOI: 10.1088/0004-637X/707/1/250. arXiv: 0905.4669 [astro-ph.CO].

Martin, D. Christopher et al. (Dec. 2007). "The UV-Optical Galaxy Color-Magnitude Diagram. III. Constraints on Evolution from the Blue to the Red Sequence". In: ApJS 173.2, pp. 342–356. DOI: 10.1086/516639. arXiv: astro-ph/0703281 [astro-ph].

Masters, Daniel C. et al. (2019). "The Complete Calibration of the Color–Redshift Relation (C3R2) Survey: Analysis and Data Release 2". In: *The Astrophysical Journal* 877.2, p. 81. ISSN: 1538-4357. DOI: 10.3847/1538-4357/ab184d. URL: http://dx.doi.org/10.3847/1538-4357/ab184d.

McCracken, H. J. et al. (Jan. 2010). "The COSMOS-WIRCam Near-Infrared Imaging Survey. I. BzK-Selected Passive and Star-Forming Galaxy Candidates at z gsim 1.4". In: ApJ 708.1, pp. 202–217. DOI: 10.1088/0004-637X/708/1/202. arXiv: 0910.2705 [astro-ph.CO].

McCracken, H. J. et al. (Aug. 2012). "UltraVISTA: a new ultra-deep near-infrared survey in COSMOS". In: A&A 544, A156, A156. DOI: 10.1051/0004-6361/201219507. arXiv: 1204.6586 [astro-ph.CO].

McLean, Ian S. et al. (July 2010). "Design and development of MOSFIRE: the multi-object spectrometer for infrared exploration at the Keck Observatory". In: *Ground-based and Airborne Instrumentation for Astronomy III*. Ed. by Ian S. McLean, Suzanne K. Ramsay, and Hideki Takami. Vol. 7735. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 77351E. DOI: 10.1117/12.856715.

McLean, Ian S. et al. (Sept. 2012). "MOSFIRE, the multi-object spectrometer for infrared exploration at the Keck Observatory". In: *Ground-based and Airborne Instrumentation for Astronomy IV*. Ed. by Ian S. McLean, Suzanne K. Ramsay, and Hideki Takami. Vol. 8446. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 84460J. DOI: 10.1117/12.924794.

Metzler, C A and A E Evrard (1994). "A Simulation of the Intracluster Medium with Feedback from Cluster Galaxies". In: *Astrophys. J.* 437.astro-ph/9309050, pp. 564–583. URL: https://cds.cern.ch/record/543013.

Mihalas, Dimitri (July 1967). "Statistical-Equilibrium Model Atmospheres for Early-Type Stars. I. Hydrogen Continua". In: ApJ 149, p. 169. DOI: 10.1086/149239.

Mihos, J. Christopher and Lars Hernquist (June 1996). "Gasdynamics and Starbursts in Major Mergers". In: ApJ 464, p. 641. DOI: 10.1086/177353. arXiv: astro-ph/9512099 [astro-ph].

Miyazaki, Satoshi et al. (Sept. 2017). "Hyper Suprime-Cam: System design and verification of image quality". In: *Publications of the Astronomical Society of Japan* 70.SP1. S1. ISSN: 0004-6264. DOI: 10.1093/pasj/psx063. eprint: https://academic.oup.com/pasj/article-pdf/70/SP1/S1/23692157/psx063.pdf. URL: https://doi.org/10.1093/pasj/psx063.

Moster, Benjamin P. et al. (Mar. 2011). "A COSMIC VARIANCE COOKBOOK". In: *The Astrophysical Journal* 731.2, p. 113. ISSN: 1538-4357. DOI: 10.1088/0004-637x/731/2/113. URL: http://dx.doi.org/10.1088/0004-637X/731/2/113.

Nayyeri, H. et al. (Jan. 2017). "CANDELS MULTI-WAVELENGTH CATALOGS: SOURCE IDENTIFICATION AND PHOTOMETRY IN THE CANDELS COSMOS SURVEY FIELD". In: *The Astrophysical Journal Supplement Series* 228.1, p. 7. ISSN: 1538-4365. DOI: 10.3847/1538-4365/228/1/7. URL: http://dx.doi.org/10.3847/1538-4365/228/1/7.

Noeske, K. G. et al. (May 2007). "Star Formation in AEGIS Field Galaxies since z=1.1: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-forming Galaxies". In: ApJ 660.1, pp. L43–L46. DOI: 10.1086/517926. arXiv: astro-ph/0701924 [astro-ph].

Oesch, P. A. et al. (Mar. 2016). "A REMARKABLY LUMINOUS GALAXY ATZ= 11.1 MEASURED WITHHUBBLE SPACE TELESCOPEGRISM SPECTROSCOPY". In: *The Astrophysical Journal* 819.2, p. 129. ISSN: 1538-4357. DOI: 10.3847/0004-637x/819/2/129. URL: http://dx.doi.org/10.3847/0004-637X/819/2/129.

Oke, J. B. (Feb. 1974). "Absolute Spectral Energy Distributions for White Dwarfs". In: ApJS 27, p. 21. DOI: 10.1086/190287.

Pérez-González, Pablo G. et al. (Mar. 2008). "The Stellar Mass Assembly of Galaxies fromz= 0 toz= 4: Analysis of a Sample Selected in the Rest-Frame Near-Infrared withSpitzer". In: *The Astrophysical Journal* 675.1, pp. 234–261. DOI: 10.1086/523690. URL: https://doi.org/10.1086%2F523690.

Polletta, M. et al. (July 2007). "Spectral Energy Distributions of Hard X-Ray Selected Active Galactic Nuclei in the XMM-Newton Medium Deep Survey". In: ApJ 663.1, pp. 81–102. DOI: 10.1086/518113. arXiv: astro-ph/0703255 [astro-ph].

Postman, M. and M. J. Geller (June 1984). "The morphology-density relation - The group connection." In: ApJ 281, pp. 95–99. DOI: 10.1086/162078.

Pozzetti, L. et al. (Nov. 2010). "zCOSMOS – 10k-bright spectroscopic sample". In: *Astronomy and Astrophysics* 523, A13. ISSN: 1432-0746. DOI: 10.1051/0004-6361/200913020. URL: http://dx.doi.org/10.1051/0004-6361/200913020.

Pozzetti, L. et al. (2007). "The VIMOS VLT Deep Survey* - The assembly history of the stellar mass in galaxies: from the young to the old universe". In: *A&A* 474.2, pp. 443–459. DOI: 10.1051/0004-6361:20077609. URL: https://doi.org/10.1051/0004-6361:20077609.

Primack, Joel R. (2015). *Cosmological Structure Formation*. arXiv: 1505.02821 [astro-ph.GA].

Rees, M. J. and J. P. Ostriker (June 1977). "Cooling, dynamics and fragmentation of massive gas clouds: clues to the masses and radii of galaxies and clusters." In: MNRAS 179, pp. 541–559. DOI: 10.1093/mnras/179.4.541.

Rosani, G. et al. (2020). "Bright Lyman- emitters among Spitzer SMUVS galaxies in the MUSE/COSMOS field". In: *Astronomy Astrophysics* 633, A159. ISSN: 1432-0746. DOI: 10.1051/0004-6361/201935782. URL: http://dx.doi.org/10.1051/0004-6361/201935782.

Salpeter, Edwin E. (Jan. 1955). "The Luminosity Function and Stellar Evolution." In: ApJ 121, p. 161. DOI: 10.1086/145971.

Salvato, Mara, Olivier Ilbert, and Ben Hoyle (June 2019). "The many flavours of photometric redshifts". In: *Nature Astronomy* 3, pp. 212–222. DOI: 10.1038/s41550-018-0478-0. arXiv: 1805.12574 [astro-ph.GA].

Sandage, Allan (Jan. 1988). "Observational tests of world models." In: ARA&A 26, pp. 561–630. DOI: 10.1146/annurev.aa.26.090188.003021.

Sanders, D. B. et al. (Sept. 2007). "S-COSMOS: The Spitzer Legacy Survey of the Hubble Space Telescope ACS 2 deg^2 COSMOS Field I: Survey Strategy and First Analysis". In: ApJS 172.1, pp. 86–98. DOI: 10.1086/517885. arXiv: astro-ph/0701318 [astro-ph].

Sanders, Robert H. (2005). "Observational Cosmology". In: *The Physics of the Early Universe*. Ed. by E. Papantonopoulos. Vol. 653, p. 105. DOI: 10.1007/978-3-540-31535-3_4.

Sawicki, Marcin et al. (Sept. 2019). "The CFHT Large Area U-band Deep Survey (CLAUDS)". In: *Monthly Notices of the Royal Astronomical Society*. ISSN: 1365-2966.

DOI: 10.1093/mnras/stz2522. URL: http://dx.doi.org/10.1093/mnras/stz2522.

Schaye, Joop et al. (Jan. 2015). "The EAGLE project: simulating the evolution and assembly of galaxies and their environments". In: MNRAS 446.1, pp. 521–554. DOI: 10.1093/mnras/stu2058. arXiv: 1407.7040 [astro-ph.GA].

Schneider, Peter (2006). *Extragalactic Astronomy and Cosmology*.

Schreiber, C. et al. (Oct. 2018). "Near infrared spectroscopy and star-formation histories of $3 \leq z \leq 4$ quiescent galaxies". In: *aap* 618, A85, A85. DOI: 10.1051/0004-6361/201833070. arXiv: 1807.02523 [astro-ph.GA].

Schreiber, C. et al. (2015). "The Herschel view of the dominant mode of galaxy growth from z = 4 to the present day". In: *A&A* 575, A74. DOI: 10.1051/0004-6361/201425017. URL: https://doi.org/10.1051/0004-6361/201425017.

Scoville, N. et al. (Sept. 2007). "The Cosmic Evolution Survey (COSMOS): Overview". In: *The Astrophysical Journal Supplement Series* 172.1, 1–8. ISSN: 1538-4365. DOI: 10.1086/516585. URL: http://dx.doi.org/10.1086/516585.

Sebok, W. L. (Oct. 1979). "Optimal classification of images into stars or galaxies - a Bayesian approach." In: AJ 84, pp. 1526–1536. DOI: 10.1086/112570.

Springel, Volker (Dec. 2005). "The cosmological simulation code GADGET-2". In: MNRAS 364.4, pp. 1105–1134. DOI: 10.1111/j.1365-2966.2005.09655.x. arXiv: astro-ph/0505010 [astro-ph].

Springel, Volker et al. (June 2005). "Simulations of the formation, evolution and clustering of galaxies and quasars". In: Nature 435.7042, pp. 629–636. DOI: 10.1038/nature03597. arXiv: astro-ph/0504097 [astro-ph].

Stark, D. P. et al. (Apr. 2007). "A New Measurement of the Stellar Mass Density at $z\sim5$: Implications for the Sources of Cosmic Reionization". In: ApJ 659.1, pp. 84–97. DOI: 10.1086/511325. arXiv: astro-ph/0604250 [astro-ph].

Stark, Daniel P. et al. (Jan. 2013). "KECK SPECTROSCOPY OF 3 <z< 7 FAINT LYMAN BREAK GALAXIES: THE IMPORTANCE OF NEBULAR EMISSION IN UNDERSTANDING THE SPECIFIC STAR FORMATION RATE AND STELLAR MASS DENSITY". In: *The Astrophysical Journal* 763.2, p. 129. ISSN: 1538-4357. DOI: 10.1088/0004-637x/763/2/129. URL: http://dx.doi.org/10.1088/0004-637X/763/2/129.

Stefanon, Mauro et al. (Apr. 2015). "Stellar Mass Functions of Galaxies At $4 < z < 7$ from an IRAC-selected Sample in Cosmos/Ultravista: Limits on the Abundance of Very Massive Galaxies". In: ApJ 803.1, 11, p. 11. DOI: 10.1088/0004-637X/803/1/11. arXiv: 1408.3416 [astro-ph.GA].

Steidel, C. C. et al. (Dec. 1998). "Lyman Break Galaxies at $z\sim3$ and Beyond". In: *arXiv e-prints*, astro-ph/9812167, astro–ph/9812167. arXiv: astro-ph/9812167 [astro-ph].

Steidel, Charles C. et al. (May 1996). "Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts z > 3". In: *The Astrophysical Journal* 462.1, L17–L21. ISSN: 1538-4357. DOI: 10.1088/1538-4357/462/1/l17. URL: http://dx.doi.org/10.1088/1538-4357/462/1/L17.

Steinhardt, Charles L. et al. (Aug. 2014). "Star Formation at 4 < z < 6 from the Spitzer Large Area Survey with Hyper-Suprime-Cam (SPLASH)". In: ApJ 791.2, L25, p. L25. DOI: 10.1088/2041-8205/791/2/L25. arXiv: 1407.7030 [astro-ph.GA].

Stockmann, Mikkel et al. (Jan. 2020). "X-shooter Spectroscopy and HST Imaging of 15 Massive Quiescent Galaxies at z 2". In: ApJ 888.1, 4, p. 4. DOI: 10.3847/1538-4357/ab5af4. arXiv: 1912.01619 [astro-ph.GA].

Stoughton, Chris et al. (Jan. 2002). "Sloan Digital Sky Survey: Early Data Release". In: AJ 123.1, pp. 485–548. DOI: 10.1086/324741.

Straatman, Caroline M. S. et al. (Oct. 2016). "The FourStar Galaxy Evolution Survey (ZFOURGE): Ultraviolet to Far-infrared Catalogs, Medium-bandwidth Photometric Redshifts with Improved Accuracy, Stellar Masses, and Confirmation of Quiescent Galaxies to z $\sim$ 3.5". In: ApJ 830.1, 51, p. 51. DOI: 10.3847/0004-637X/830/1/51. arXiv: 1608.07579 [astro-ph.GA].

Stringer, M.J. et al. (2009). "Mock observations with the Millennium simulation: Cosmological downsizing and intermediate redshift observations". In: *Mon. Not. Roy. Astron. Soc.* 393, p. 1127. DOI: 10.1111/j.1365-2966.2008.14186.x. arXiv: 0806.2232 [astro-ph].

Szalay, Alexander, Andrew Connolly, and Gyula Szokoly (Nov. 1998). "Simultaneous Multicolor Detection of Faint Galaxies in the Hubble Deep Field". In: *The Astronomical Journal* 117. DOI: 10.1086/300689.

Tabatabaei, F. S. et al. (Nov. 2018). "Discovery of massive star formation quenching by non-thermal effects in the centre of NGC 1097". In: *Nature Astronomy* 2, pp. 83–89. DOI: 10.1038/s41550-017-0298-7. arXiv: 1710.05695 [astro-ph.GA].

Tacchella, Sandro et al. (Dec. 2018). "A Redshift-independent Efficiency Model: Star Formation and Stellar Masses in Dark Matter Halos at z 4". In: ApJ 868.2, 92, p. 92. DOI: 10.3847/1538-4357/aae8e0. arXiv: 1806.03299 [astro-ph.GA].

Thomas, Daniel et al. (Mar. 2005). "The Epochs of Early-Type Galaxy Formation as a Function of Environment". In: *The Astrophysical Journal* 621.2, 673–694. ISSN: 1538-4357. DOI: 10.1086/426932. URL: http://dx.doi.org/10.1086/426932.

Toft, S. et al. (Dec. 2007). "Hubble Space TelescopeandSpitzerImaging of Red and Blue Galaxies atz$\sim$ 2.5: A Correlation between Size and Star Formation Activity from Compact Quiescent Galaxies to Extended Star-forming Galaxies". In: *The Astrophysical Journal* 671.1, pp. 285–302. DOI: 10.1086/521810. URL: https://doi.org/10.1086%2F521810.

Valentino, Francesco et al. (Feb. 2020). "Quiescent Galaxies 1.5 Billion Years after the Big Bang and Their Progenitors". In: ApJ 889.2, 93, p. 93. DOI: 10.3847/1538-4357/ab64dc. arXiv: 1909.10540 [astro-ph.GA].

Vernet, J. et al. (Dec. 2011). "X-shooter, the new wide band intermediate resolution spectrograph at the ESO Very Large Telescope". In: A&A 536, A105, A105. DOI: 10.1051/0004-6361/201117752. arXiv: 1110.1944 [astro-ph.IM].

Walcher, Jakob et al. (Aug. 2010). "Fitting the integrated spectral energy distributions of galaxies". In: *Astrophysics and Space Science* 331.1, 1–51. ISSN: 1572-946X. DOI: 10.1007/s10509-010-0458-z. URL: http://dx.doi.org/10.1007/s10509-010-0458-z.

Weaver, John et al. (Nov. 2019). "The Farmer: Improved model-based photometry for the next generation of galaxy surveys". In: *The Art of Measuring Galaxy Physical Properties*, p. 9. DOI: 10.5281/zenodo.3554205.

Wechsler, Risa H. and Jeremy L. Tinker (Sept. 2018). "The Connection Between Galaxies and Their Dark Matter Halos". In: *Annual Review of Astronomy and Astrophysics* 56.1, 435–487. ISSN: 1545-4282. DOI: 10.1146/annurev-astro-081817-051756. URL: http://dx.doi.org/10.1146/annurev-astro-081817-051756.

Weinmann, Simone M., Eyal Neistein, and Avishai Dekel (Nov. 2011). "On the puzzling plateau in the specific star formation rate at z= 2-7". In: MNRAS 417.4, pp. 2737–2751. DOI: 10.1111/j.1365-2966.2011.19440.x. arXiv: 1103.3011 [astro-ph.CO].

Williams, Rik J. et al. (2009). "DETECTION OF QUIESCENT GALAXIES IN A BI-COLOR SEQUENCE FROMZ= 0-2". In: *The Astrophysical Journal* 691.2, pp. 1879–

1895. DOI: 10.1088/0004-637x/691/2/1879. URL: https://doi.org/10.1088%2F0004-637x%2F691%2F2%2F1879.

Wolf, C. et al. (Mar. 2003). "The COMBO-17 survey: Evolution of the galaxy luminosity function from 25000 galaxies with $0.2 < z < 1.2$". In: *Astronomy and Astrophysics* 401.1, 73–98. ISSN: 1432-0746. DOI: 10.1051/0004-6361:20021513. URL: http://dx.doi.org/10.1051/0004-6361:20021513.

Zamojski, M. A. et al. (Sept. 2007). "Deep GALEX Imaging of the COSMOS HST Field: A First Look at the Morphology of $z \sim 0.7$ Star-forming Galaxies". In: ApJS 172.1, pp. 468–493. DOI: 10.1086/516593. arXiv: astro-ph/0701478 [astro-ph].

Zolotov, Adi et al. (July 2015). "Compaction and quenching of high-z galaxies in cosmological simulations: blue and red nuggets". In: MNRAS 450.3, pp. 2327–2353. DOI: 10.1093/mnras/stv740. arXiv: 1412.4783 [astro-ph.GA].