Early death of massive galaxies in the distant universe

Caroline M. S. Straatman

Front cover: opened dome of the Magellan Baade 6.5m Telescope at Las Campanas Observatory, Chile, just before sunset. **Back cover**: part of the cosmic field UDS on the sky, as observed in the near-IR with the FourStar instrument.

Early death of massive galaxies in the distant universe

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van de Rector Magnificus prof. mr. C. J. J. .M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op dinsdag 29 maart 2016
klokke 15:00 uur

door

Caroline Margaretha Stefanie Straatman
geboren te Leidschendam
in 1987

Promotiecommissie

Promotor	Prof. dr. M. Franx	
Co-promotor	Dr. I. F. Labbé	
Overige leden	Prof. dr. H. J. A. Röttgering	
	Prof. dr. J. Schaye	
	Prof. dr. M. T. Kriek	University of California, Berkeley
	Dr. A. van der Wel	Max-Planck-Institut für Astronomie, Heidelberg
	Prof. dr. K. Glazebrook	Swinburne University of Technology, Hawthorn, Australia
	Prof. dr. P. van der Werf	
	Dr. J. A. Hodge	

“We have a hunger of the mind which asks for knowledge of all around us, and the more we gain, the more is our desire; the more we see, the more we are capable of seeing."

Maria Mitchell (1818 – 1889)

Contents

1

Introduction

1.1 Cosmological context

The universe is often characterized by the following two words: *homogeneous* and *isotropic*. This means that the average density of matter is the same in all places in the universe (homogeneity), and at the same time, the universe looks the same in all directions as viewed by a particular observer (isotropic).

If the universe were infinite and unchanging, this would imply that wherever we look, we would always see the light of some star, and the entire sky would always be filled with light[1]. Instead we observe that the night sky is mostly dark. The most recent insights are that the universe has only existed for a finite time, and is furthermore undergoing accelerated expansion, so that light from distant sources hasn't had the time to reach us yet. The expansion also causes light from distant sources to become redshifted beyond the range of optical light that our eyes can see.

On small scales the universe is very inhomogeneous and non-isotropic. Matter is gravitationally bound together into stars, planets and galaxies. We are part of the Milky Way, a galaxy as massive as 100 billion times the mass of the Sun and itself part of the Local Group, together with its neighbour Andromeda and a number of smaller satellite galaxies. It is believed that the origins of these matter-dense regions of space lie in quantum fluctuations that occured during the very first moments of the universe.

The universe is estimated to have originated about 13.7 billion years ago from a hot, dense initial state, a phase we call the Big Bang. Shortly after the Big Bang (about 10^{-36} seconds) a period of inflation most likely took place (Guth 1981). This lasted for about 10^{-34} or more seconds and during this time the universe expanded at an astonishing rate, increasing its size by approximately 100 e-folding times, or a factor of $\sim 10^{43}$. The theory of inflation solves the so-called *horizon problem*: if the universe is homogeneous on large

[1] Olbers' paradox (Harrison 1987)

scales, this is likely because all regions in the observable universe have been in causal contact at one point in the past, even though we cannot now observe it in its entirety. Inflation also causes quantum fluctuations to be frozen into the density fluctuations that provide the initial conditions for the growth of structure in the universe.

After inflation, the universe cooled until particles were formed. During 377,000 years the universe was opaque to light, as photons could travel only short distances before interacting with an electron. At the end of this period, the universe had cooled enough for the recombination of electrons and protons into neutral hydrogen atoms, and shortly after that the photons were decoupled to travel freely through the universe. With our astronomical observations we can probe the distant universe back in time until this epoch, and we observe an imprint of the universe at the moment of decoupling, which is called the Cosmic Microwave Background (Alpher et al. 1948; Penzias & Wilson 1965). This is a faint signal redshifted to microwave wavelengths, due to the large expansion of the universe since that time. In it, we can see small fluctuations that are the beginnings of the structure we see in the universe today.

According to the Lambda Cold Dark Matter model, overdense regions are formed through the gravitational collapse of dark matter into dark matter haloes, and their subsequent hierarchical merging (White & Rees 1978). One of the greatest unresolved questions in physics and cosmology is: *What is the nature of dark matter?* To answer this question is far beyond the scope of this thesis, but it is worth noting that dark matter contributes 25.9% (Planck Collaboration et al. 2015) to the energy density of the universe. Baryonic matter, that stars, planets, humans and atoms consist of, makes up only 4.9% (Planck Collaboration et al. 2015). Important evidence of the existence of dark matter comes from studying the rotational velocities of galaxies. Based on visible mass, basic laws of motion suggest declining rotation curves towards larger radii. This is not observed and implies the presence of large quantities of dark matter mass (e.g. Freeman 1970; van Albada et al. 1985).

Baryonic matter, which consists mostly of neutral hydrogen gas, will collapse along with the dark matter, cool down, and flow to the centers of the haloes. Once the first galaxies are formed, they embark upon a complex journey of gas accretion, star-formation, feedback processes, and interactions with other galaxies. The details of the galaxy formation process are not yet well understood. In the local universe we find galaxies with a variety of morphologies, which are often linked to star-formation activity. Some galaxies are very massive and have all but stopped forming new stars. Therefore another key question in cosmology is: *How do galaxies form and evolve?* And related to

this: *What causes galaxies to stop forming new stars?* These two questions are the focus of this thesis.

1.2 Galaxies at low redshift

A logical starting point for understanding galaxies is their morphology. It is correlated to their dynamics and other galaxy properties, such as age, mass and star-formation history. The structure of galaxies is closely tied to their assembly history and the underlying dark matter distribution. Since galaxy formation models will have to be able to reproduce the great variety of galactic shapes, it is an important gauge for determining the validity of any model.

One of the earliest classification systems was that of Hubble (1926, 1936). This is the famous tuning fork, consisting of the following classes: elliptical galaxies, lenticular galaxies, disk galaxies with spiral structure, and irregular galaxies. Ellipticals and lenticulars are historically termed early-type, and spiral galaxies late-type, because originally it was thought that galaxies evolved from elliptical shapes into the seemingly more refined spiral morphologies. Early-type galaxies are, in fact, the oldest and most evolved kind of galaxies.

Some clues as to the formation history of early-type galaxies can be found by studying their physical properties. They are predominantly the most massive galaxies in the local universe and can often be found in the centers of groups and clusters of galaxies. They have old stellar populations, with distinctly red optical colours, they have little ongoing star-formation, and they are kinematically supported by random motions. Their high stellar ages indicate they assembled the bulk of their stellar mass at high redshift: redshift $(z) > 2$.

Spiral disk galaxies have blue optical colors, from light emitted by populations of young stars. They are actively forming new stars and have large rotational velocities. It is thought that the Milky Way is a spiral galaxy, with a modest rate of star-formation. It is clear that these two types of galaxies have very distinct physical properties, which, without any foreknowledge, were captured accurately by Hubble simply by studying morphology.

A well known scaling relation for star-forming galaxies is the Tully-Fisher relation, first reported by Tully & Fisher (1977). It describes a tight correlation between rotational velocity and, historically, luminosity. Since rotational velocity can be measured accurately regardless of distance, the Tully-Fisher relation was first used as a distance indicator for galaxies. In present day research, the Tully-Fisher relation is expressed in terms of stellar mass instead of luminosity and is used for kinematical studies of galaxies. At low red-

shift the Tully-Fisher relation is well established, but it remains elusive for galaxies at high redshift. Determining if and by how much the Tully-Fisher relation evolves over time is key to understanding the kinematic evolution of galaxies, which is closely tied to understanding the formation of dark matter haloes and the interplay of dark matter with baryons.

Galaxies with low star-formation rates and red optical colors are often termed *quiescent*. Several definitions exist for quiescence. These amount to either a star-formation rate maximum or a colour threshold that captures a spectral feature related to the age of the galaxy. If the progenitors of early-type galaxies were spirals, it could be that the process that changed their structure is the same as that which caused a halt to star-formation. A logical way to study the transition from the star-forming phase to the quiescent phase is to look for galaxies that are in the middle of this process. To find these, larger numbers of galaxies than those available in the local universe have to be studied.

The first large multi-wavelength galaxy survey was the Sloan Digital Sky Survey (SDSS; York et al. 2000). SDSS provided imaging in over 14,000 square degrees of sky, and spectra of more than 2 million galaxies. The survey led to several breakthroughs, such as pinpointing the bimodality in colour-mass space between star-forming and quiescent galaxies (Kauffmann et al. 2003; Blanton et al. 2005), and the determination of the relation between size and stellar mass (Shen et al. 2003). Additionally, SDSS allowed for the first time to map the three dimensional large scale structure of matter in the universe.

Despite the vast amount of galaxies observed, SDSS is a low redshift survey, covering only a limited range of time. The most effective way to study the evolution of galaxies would be to capture them as they evolve, by observing them at different epochs throughout the existence of the universe. Since SDSS a number of galaxy surveys have succesfully attempted to probe further and further into the distant universe, and assemble large samples of galaxies at high redshift. A key epoch is $1 < z < 4$, when most of the star-formation took place (Madau et al. 1996), the first galaxies ceased forming stars, and the familiar elliptical and spiral morphologies first emerged.

1.3 Galaxies at high redshift

Quiescent galaxies have been confirmed to exist out to redshifts $z \sim 2.3$ (e.g., Kriek et al. 2006). A key discovery was that at high redshift, their morphology is not the same as at low redshift. Instead of having extended elliptical shapes, they are very compact (e.g., Daddi et al. 2005; Trujillo et al. 2007; van

Dokkum et al. 2008; Damjanov et al. 2009; van der Wel et al. 2014). Their average size decreases with increasing redshift, although the general correlation between size and stellar mass, with larger sizes for more massive galaxies, remains intact. The interpretation is that elliptical galaxies grow inside-out, by first forming a dense stellar core, and later accreting more mass by star-formation and mergers with other smaller galaxies.

This does not solve the riddle of why star-formation in these galaxies ceased. Stars are formed from dense, cold gas, and a number of reasons have been suggested for the quenching of star-formation in galaxies. One such reason is feedback by active galactic nuclei, as during periods of rapid accretion onto the central black hole, a great amount of energy is released into the surrounding environment of the galaxy (e.g., Kormendy & Richstone 1995; Magorrian et al. 1998). Another example is feedback from supernovae, massive and old exploding stars that heat and dilute their surrounding gas (e.g., Mathews 1990; Ciotti et al. 1991). Less massive old stars can also influence their environment, by shedding an envelope of mass, which initially moves with the speed of its host, but interacts with surrouding gas reservoirs (e.g., Conroy et al. 2015). Finally, in time some dark matter haloes hosting galaxies become so massive that they switch from cold-mode to hot-mode accretion. This means that the cooling time of primordial gas flowing into the centers of the haloes becomes too long (e.g., Birnboim & Dekel 2003; Cattaneo et al. 2008).

A method of separating between the various proposed quenching mechanisms is charting the number density, fraction, and structural properties of quiescent galaxies to high redshift, and to link these properties to possible star-forming progenitors. Research into progenitors is now focusing on finding similarly compact star-forming galaxies, but these have proven difficult to find (e.g. Barro et al. 2014a,b; Nelson et al. 2014). Furthermore, the discovery of massive quiescent galaxies at higher and higher redshift implicates a swift formation process, with rapid star-formation at very early times. Only a subset of the star-forming population at those early epochs ($z = 4 - 10$) is known, and these tend to be UV-bright galaxies. Observations at $z > 1$ have revealed the existence of a large population of dust-obscured star-forming galaxies, with high SFRs, but similar red colours as quiescent galaxies (e.g. Reddy et al. 2005; Spitler et al. 2014). Their redness makes them difficult to find at $z > 4$ with current techniques, which means we may be missing a large fraction of the star-forming population at these redshifts. It also proves a challenge for identifying quiescent galaxies as the two kinds may easily be confused. A question that remains standing is: *When did the first galaxies become quiescent?* Pinpointing that moment in time will be essential to constrain galaxy

formation scenarios and is the topic of one of the chapters of this thesis.

Apart from a large fraction of star-forming galaxies being dust-obscured, other properties of the star-forming population are different at high redshift as well. For example, their average star-formation rate is higher, and a larger fraction of their mass is in the form of gas. Their morphology is more irregular, with clumps of star-forming matter and more visible effects from disruptions by late interactions with other galaxies. Under these circumstances, it is hardly expected that the Tully-Fisher scaling relation between stellar mass and rotational velocity holds in exactly the same way as for low-redshift galaxies. If and by how much the Tully-Fisher relation evolves is another topic of this thesis.

1.4 The FourStar Galaxy Evolution Survey

The most important challenge in astronomy is to determine accurate distances. For extragalactic observations this means precisely calculating the redshift of a source – it has been realised that redshift inaccuracies are the most dominant factor inhibiting our understanding of galaxy properties such as stellar age and mass (e.g., Chen et al. 2003; Kriek et al. 2008). The best method for this is to measure the electromagnetic spectrum of a galaxy, and use features such as emission lines from atomic transitions to determine the factor by which the spectrum was shifted towards redder wavelengths. To study galaxy evolution, this method has several drawbacks. Observing large samples, i.e., thousands, of galaxies is highly inefficient and requires a previous detection to pinpoint the location of the source. Spectroscopy is also limited to bright sources or sources with strong emission lines, and these are usually star-forming galaxies with moderate dust-obscuration, which are not necessarily representative of the full galaxy population at any redshift. Spectroscopy is therefore often used in follow-up programs of imaging surveys.

Determining redshifts for galaxy surveys that rely on imaging is done by observing sources through different filters to obtain a spectral energy distribution and fitting models to these. In a sense a spectral energy distribution is a very low resolution spectrum. Therefore this process becomes more accurate if more filters are used, with measurements at different wavelengths.

Optical light, carrying information about the age of a galaxy, is shifted into the near-IR for sources at $z > 1.5$. This is problematic for groundbased observations, because the light of typical galaxies is outshone by a factor 10^5 by the Earth's atmosphere in the IR. The FourStar instrument on the 6.5m Magellan Baade Telescope at Las Campanas Observatory in Chile provides a solution for both issues. FourStar has a set of six near-IR medium-bandwidth filters

that can capture light in small wavelength windows where the atmosphere is transparent. These six filters also provide an excellent sampling of spectral features typical for old galaxies, allowing a $1-2\%$-level redshift accuracy.

The technique of using medium-bandwidth filters was first employed for optical light by the COMBO17 survey (Wolf et al. 2004) and shown to be effective in the near-IR as well by the NEWFIRM Medium-Band Survey (Whitaker et al. 2011). The FourStar Galaxy Evolution Survey (ZFOURGE) takes this one step further by being unprecedented in depth, reaching K_s-band magnitudes (the reddest filter) of ~ 26 in AB units.

ZFOURGE is a 45 night legacy program, conducted between December 2010 and November 2012, covering a total of 400 square arcminutes in three pointings on the sky. The three pointings reduce the effect of field-to-field variance, which is caused by matter being unevenly distributed on relatively small cosmological scales. The pointings overlap with those of previous surveys, so that the near-IR data can be optimally augmented by earlier measurements ranging from the UV to the far-IR.

The aim of ZFOURGE is to shed light on how galaxies evolve by studying them at the crucial epoch between $z = 1.5$ and $z = 4.5$. It is excellently suited to identify quiescent galaxies out to $z \sim 4$, which may be the epoch in which they first appeared in the universe. With ZFOURGE we can also study scaling relations for these galaxies, such as the relation between size and stellar mass, which evolves in a different way for star-forming and quiescent galaxies.

A spectroscopic follow-up program, ZFIRE (Nanayakkara et al., 2016, submitted), was started in December 2013, employing the near-IR spectrograph MOSFIRE on the Keck I telescope on Mauna Kea in Hawai'i. The primary targets were star-forming cluster galaxies, discovered with ZFOURGE, at $z = 2.095$. At this redshift, little is known about the kinematic properties of star-forming galaxies. The spectra obtained with MOSFIRE cover the Hα emission line (rest-frame $\lambda = 6563$Å), at high spectral resolution. They are therefore an excellent tool to measure the rotational velocities of galaxies beyond $z > 2$ for the first time with single-slit spectra.

1.5 Outline and summary

In this thesis we discuss the properties of high redshift galaxies at two key epochs. We use ZFOURGE to find and study the earliest quiescent galaxies at $z \sim 4$, when the universe was only 1.6 billion years old. And we employ the ZFIRE spectroscopic data to measure the Tully-Fisher relation for star-forming galaxies at $2.0 < z < 2.5$, at the time when the cosmic star-formation rate was at its peak.

In **Chapter 2** we first present the data products from ZFOURGE. We use ultra-deep near-IR K_s–band ($2.16\mu m$) images to detect > 70000 galaxies. For each of these we derive fluxes in > 27 UV, optical and IR filters, and measure the photometric redshift. We perform an in-depth analysis of the photometric redshift accuracy, including a comparison with spectroscopically derived redshifts from literature and an analysis using galaxy pairs. Using the large sample of galaxies from ZFOURGE, we additionally investigate the efficacy of a two colours test, used to distinguish between quiescent and star-forming galaxies, at high redshift ($z > 2$).

In **Chapter 3** we present the discovery of massive quiesent galaxies at redshift $z \sim 4$. Using deep far-IR data from the MIPS instrument on the Spitzer Space Telescope and the PACS instrument on the Herschel Space Observatory, we verify that these galaxies indeed have strongly suppressed star-formation rates. From their high average stellar mass, we infer that they must have formed extremely rapidly, and quenching mechanisms were efficient even at high redshift. Lastly, we speculate that most of the star-formation in the progenitors of these galaxies was obscured by dust.

We continue our study of $z \sim 4$ quiescent galaxies in **Chapter 4**, where we investigate their sizes. We study near-IR images of both star-forming and quiescent galaxies at $z \sim 4$, which, because the light is redshifted, is a measurement of UV light emitted by the galaxies. We find that the quiescent galaxies are very compact, and much smaller than star-forming galaxies of similar stellar mass. Next, we compare with lower redshift results, to study the size evolution. We find that both quiescent and star-forming galaxies at $z \sim 4$ continue the trend of smaller average sizes towards higher redshift. We then look for compact star-forming galaxies in our sample, which could be the progenitors of similary compact quiescent galaxies at later times. We find only one, indicating these are very rare and possibly dust-obscured.

In **Chapter 5** we jump forward in time, to study the Tully-Fisher relation at redshift $2.0 < z < 2.5$. Here we make use of a sample of star-forming galaxies that were spectroscopically observed with ZFIRE. We derive rotational velocities by measuring the shear of the Hα emission line. We extensively analyse systematic effects and find that velocities measured with single-slit spectra can easily be underestimated. Taking this into account we derive a Tully-Fisher relation that is offset compared to low redshift results. We then attempt to unify previous measurements at various redshifts, and infer a gradual evolution with redshift, which is in agreement with theoretical predictions. Lastly, we find evidence of a general increase in random motions and speculate the evolution of the Tully-Fisher relation may in part reflect the conversion from gas to stars.

1.6 Future prospects

In this work we present the discovery of the furthest quiescent galaxies to date. These will be a valuable addition to the known population of quiescent galaxies through cosmic time, showing in the first place their early existence. Their number density, average stellar mass and average size will provide important constraints on galaxy formation models, dealing with the efficiency of star-formation, morphological evolution and testing of various quenching mechanisms. The question of how and why these galaxies have all but stopped forming new stars is still an open one, but we now know that galaxies can assemble most of their stellar mass rapidly and an efficient quenching mechanism is possible.

One caveat is that the existence of $z \sim 4$ quiescent galaxies has not yet been verified by other observations. The first logical step for future research is to use spectroscopy to confirm and measure more precisely their redshift. Facilities able to do this are MOSFIRE on Keck I and the spectrograph NIRSpec on the James Webb Space Telescope (JWST), which is scheduled for launch in 2018. JWST will also have an IR camera installed, that can probe the universe to further depths without hindrance by Earth's atmosphere. If quiescent galaxies exist even beyond $z > 4$, they may be found by JWST. JWST will also study the very first galaxies formed after the Big Bang and will possibly shed more light on the progenitors of early quiescent galaxies.

To better understand the evolution of the Tully-Fisher relation for star-forming galaxies at $z > 2$, the most important step is to acquire larger samples of spectroscopically observed galaxies and study these using consistent methodologies. The uncertainties on current observations – especially the discrepancies between results from different surveys – are too high to constrain by how much the relation actually evolves. Facilities for this are already in place, such as the single-slit spectrograph MOSFIRE on Keck I and the integral-field-unit KMOS on the VLT. JWST will also provide excellent quality data from its NIRSpec. Another important technical development is adaptive optics, which will provide the resolution needed to better study the complex dynamics and irregular shapes of star-forming galaxies at high redshift. Finally, both observations and models need to focus on a more detailed assessment of the interplay between gas, stars and dark matter.

As a final remark it is worth mentioning the Atacama Large Millimeter / sub-millimeter Array (ALMA). With ALMA we can study the dust and molecular gas properties of distant galaxies at high resolution. These kind of observations will yield insights into the gas content of galaxies and its conversion into stars, a highly relevant topic for galaxy evolution.

Bibliography

Alpher, R. A., Bethe, H., & Gamow, G. 1948, Phys. Rev., 73, 803

Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2014a, *ApJ*, 791, 52

Barro, G., Trump, J. R., Koo, D. C., et al. 2014b, *ApJ*, 795, 145

Birnboim, Y., & Dekel, A. 2003, *MNRAS*, 345, 349

Blanton, M. R., Eisenstein, D., Hogg, D. W., Schlegel, D. J., & Brinkmann, J. 2005, *ApJ*, 629, 143

Cattaneo, A., Dekel, A., Faber, S. M., & Guiderdoni, B. 2008, *MNRAS*, 389, 567

Chen, H.-W., Marzke, R. O., McCarthy, P. J., et al. 2003, *ApJ*, 586, 745

Ciotti, L., D'Ercole, A., Pellegrini, S., & Renzini, A. 1991, *ApJ*, 376, 380

Conroy, C., van Dokkum, P. G., & Kravtsov, A. 2015, *ApJ*, 803, 77

Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680

Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, *ApJ*, 695, 101

Freeman, K. C. 1970, *ApJ*, 160, 811

Guth, A. H. 1981, PhRvD, 23, 347

Harrison, E. R. 1987, New Scientist, 1592, 72

Hubble, E. P. 1926, *ApJ*, 64, 321

—. 1936, Realm of the Nebulae

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33

Kormendy, J., & Richstone, D. 1995, *ARA&A*, 33, 581

Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, *ApJL*, 649, L71

—. 2008, *ApJ*, 677, 219

Madau, P., Ferguson, H. C., Dickinson, M. E., et al. 1996, *MNRAS*, 283, 1388

Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, 115, 2285

Mathews, W. G. 1990, *ApJ*, 354, 468

Nelson, E., van Dokkum, P., Franx, M., et al. 2014, *Nature*, 513, 394

Penzias, A. A., & Wilson, R. W. 1965, *ApJ*, 142, 419

Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2015, ArXiv e-prints, arXiv:1502.01589

Reddy, N. A., Erb, D. K., Steidel, C. C., et al. 2005, *ApJ*, 633, 748

Shen, S., Mo, H. J., White, S. D. M., et al. 2003, *MNRAS*, 343, 978

Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJL*, 787, L36

Trujillo, I., Conselice, C. J., Bundy, K., et al. 2007, *MNRAS*, 382, 109

Tully, R. B., & Fisher, J. R. 1977, *A&A*, 54, 661

van Albada, T. S., Bahcall, J. N., Begeman, K., & Sancisi, R. 1985, *ApJ*, 295, 305

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, 788, 28

van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJL*, 677, L5

Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86

White, S. D. M., & Rees, M. J. 1978, *MNRAS*, 183, 341

Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, *A&A*, 421, 913

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579

2

The FourStar Galaxy Evolution Survey: ultraviolet to far-infrared catalogs, medium-bandwidth photometric redshifts, and stellar population properties; analysis of photometric redshift accuracy and confirmation of quiescent galaxies to z ~ 3.5

Abstract

The FourStar *galaxy evolution survey (ZFOURGE)* is a 45 night legacy program with the FourStar infrared camera on Magellan covering 400arcmin^2 in three fields, CDFS, COSMOS and UDS (overlapping CANDELS). We use 6 near-IR medium-bandwidth filters ($J_1, J_2, J_3, H_s, H_l, K_s$) ranging from 1.05$\mu$m to 2.16$\mu$m to 25 – 26 magnitude (5σ, total, AB) at a seeing of ~ 0.5″. We present K_s-band selected photometric catalogs, based on ultradeep K_s-band detection images (25.5 – 26.5 AB), including ancillary imaging in 26-40 filters per field covering wavelengths 0.03 – 8μm. The catalogs are > 80% complete to K_s < 25.3 – 25.9 AB and comprise > 70,000 galaxies. We derive photometric redshifts with EAZY and stellar population properties with FAST. Comparing photometric with spectroscopic redshifts indicates $\sigma_{z,spec}$ = 0.009, 0.008, and 0.013 in CDFS, COSMOS, and UDS. As spectroscopic samples are often biased towards bright and blue sources, we also analyse galaxy pairs finding $\sigma_{z,pairs} = 0.01 - 0.02$ at $1 < z < 2.5$ on average. We quantify how $\sigma_{z,pairs}$ depends on redshift, magnitude, SED type, and the inclusion of FourStar medium-

bands. The $\sigma_{z,pairs}$ are $\times 2$ smaller for bright and blue star forming samples, while red star forming galaxies have the worst $\sigma_{z,pairs}$, with photometric uncertainties underestimating the scatter. The inclusion of FourStar medium-bands reduces the $\sigma_{z,pairs}$ by 50% at $1.5 < z < 2.5$. We calculate SFRs based on ultraviolet-to-far-IR observations, using ultradeep Spitzer/MIPS and Herschel/PACS data. We derive rest-frame $U-V$ and $V-J$ colors, and illustrate how these colors correlate with specific SFR and dust emission to $z = 3.5$. We confirm the existence of quiescent galaxies at $z \sim 3$ and demonstrate their SFRs are suppressed by $> \times 15$.

2.1 Introduction

Over the last few decades, it has been possible to obtain new insights into the formation and evolution of galaxies in a statistically significant way, by using large samples of sources from multiwavelength photometric surveys, for example with SDSS (York et al. 2000). Improved near-IR facilities on the ground, as well as advanced space-based instruments have enabled galaxy surveys probing the universe at higher resolution, fainter magnitudes and towards higher redshifts ($z > 1.5$) (e.g. Lawrence et al. 2007; Wuyts et al. 2008; Grogin et al. 2011; Koekemoer et al. 2011; Whitaker et al. 2011; Muzzin et al. 2013a; Skelton et al. 2014). These in turn have led to great progress in tracing the structural evolution of galaxies (e.g. Daddi et al. 2005; van Dokkum et al. 2008; Franx et al. 2008; Bell et al. 2012; Wuyts et al. 2012; van der Wel et al. 2012, 2014), luminosity and stellar mass functions (e.g. Faber et al. 2007; Pérez-González et al. 2008; Marchesini et al. 2009; Muzzin et al. 2013b; Tomczak et al. 2014), the environmental effects on galaxy evolution (e.g. Postman et al. 2005; Peng et al. 2010b; Cooper et al. 2012; Papovich et al. 2010; Kawinwanichakij et al. 2014; Allen et al. 2015) and the correlation between stellar mass and star-formation rate (e.g. Noeske et al. 2007; Wuyts et al. 2011; Whitaker et al. 2012) over cosmic time.

The redshift range $1 < z < 3$, when the universe was between 2.1 and 5.6 Gyr old, is an important epoch for studies of galaxy evolution. During this period 60% of all star-formation took place (e.g. Madau et al. 1998; Sobral et al. 2013), an early population of quiescent galaxies started to appear (e.g. Daddi et al. 2005; Kriek et al. 2006; Marchesini et al. 2010) and galaxies evolved into the familiar elliptical and spiral morphologies that we see in the universe today (e.g. Bell et al. 2012). The main observational limitation to understanding the fundamentals of galaxy evolution is the availability of accurate distance estimates for mass complete samples of galaxies. These can be obtained with spectroscopy, but observations are limited to a biased population of galaxies:

bright and most often star-forming, with strong emission lines.

Instead many galaxy surveys rely exclusively on the photometric sampling of the spectral energy distributions (SEDs) of galaxies to derive redshifts. Even when deep near-infrared imaging is used to derive photometric redshifts, these surveys mostly suffer from systematic effects from the use of broadband filters. These can lead to large random errors, of the order of $\sigma_z/(1+z) = 0.1$, and may introduce biases in derived luminosities and stellar masses (Chen et al. 2003; Kriek et al. 2008). A better sampling of the SED improves the accuracy of the photometric redshifts greatly and can be obtained by the use of medium-bandwidth filters. These were first applied in the optical for the COMBO17 survey (Wolf et al. 2004). A notable feature in the SED of a galaxy is the Balmer/4000Å break at rest-frame 4000Å, which shifts into the near-IR at $z \gtrsim 1.5$. For high redshift surveys, it is therefore advantageous to split up the canonical broadband J and H filters into multiple near-IR medium-bandwith filters (van Dokkum et al. 2009), which stradle the Balmer/4000Å break at $1.5 \lesssim z \lesssim 3.5$. A set of near-IR medium-bandwidth filters was used for the NMBS, a survey using NEWFIRM on the Kitt Peak Mayall 4m Telescope, with a limiting 5σ depth in K of 23.5 AB mag for pointsources and a photometric redshift accuracy of $\sigma_z/(1+z) \sim 1-2\%$ up to $z = 3$ (Whitaker et al. 2011).

The FourStar Galaxy Evolution Survey (ZFOURGE[1]) aims to further advance the study of intermediate to high redshift galaxies by pushing to much fainter limits (25-26 AB), well beyond the typical limits of groundbased spectroscopy. This provides a unique opportunity to study the higher redshift and lower mass galaxy population in unprecedented detail, at cutting edge mass completeness limits. The power of this deep survey is demonstrated by Tomczak et al. (2014), who showed the stellar mass functions of star forming and quiescent galaxies can be accurately traced down to 10^9 $M_\odot$ at z=2, well below M^*. Furthermore Straatman et al. (2014) probed the high redshift universe and found a mass complete sample of quiescent galaxies with $M > 10^{10.6}$ $M_\odot$, already at $z \sim 4$, while Tilvi et al. (2013) used the FourStar medium-bandwidth filters to pinpoint Lyman Break galaxies at $z \sim 7$ and distinguish them from cool dwarf stars. In this paper we present the ZFOURGE data products[2], comprising 45 nights of observations with the FourStar near-infrared Camera on the 6.5m Magellan Baade Telescope at Las Campanas in Chile (Persson et al. 2013). The survey was conducted over three extragalactic fields: CDFS ($RA(J2000) = 03:32:30$ $Dec(J2000) = -27:48:30$) (Giacconi et al. 2002), COSMOS ($RA = 10:00:30$ $Dec = +02:17:30$) (Scoville et al. 2007)

[1] zfourge.tamu.edu

[2] available for download at zfourge.strw.leidenuniv.nl

and UDS ($RA = 02:17:00$ $Dec = -05:13:00$) (Lawrence et al. 2007), to reduce the effect of cosmic variance, and benefit from the large amount of public UV, optical and IR data already available. We present K_s-band selected near-IR catalogs, supplemented with public UV to IR data at $0.3-8\mu$m, far-IR data from Spitzer/MIPS at 24μm for all fields and Herschel/PACS at 100μm and 160μm for CDFS.

In Sections 2.2 and 2.3 , we discuss the survey and image processing and optimization. In Section 2.4 we discuss source detection and photometry and include a description of the ZFOURGE data products. In Section 2.5 we test the completeness limits of the survey. We derive photometric redshifts and rest-frame colors in Section 2.6 and stellar masses, stellar ages and star formation rates in Section 2.7. In Section 2.8 we show how to effectively distinguish quiescent from star forming galaxies using a UVJ diagram, validating this classification with far-IR Spitzer/MIPS and Herschel/PACS data. A summary is provided in Section 2.9. Throughout, we assume a standard ΛCDM cosmology with $\Omega_{\rm M} = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$km s^{-1}Mpc^{-1}. The adopted photometric system is AB (Oke et al. 1995).

2.2 Data

2.2.1 ZFOURGE

The FourStar Galaxy Evolution Survey (ZFOURGE, PI: I. Labbé) is a 45 night program with the FourStar instrument (Persson et al. 2013) on the 6.5 m Magellan Baade Telescope at Las Campanas, Chile. FourStar has 5 near-IR medium-bands: J_1, J_2, J_3, H_s and H_l, covering the same range as the more classical J and H broadband filters, and a K_s-band. The central wavelengths of these filters range from 1.05 μm (J_1) to 2.16 μm (K_s).

The filtercurves are shown in Figure 2.1; we have also added the filter curves of the ancillary dataset (see Section 2.2.4), showing that we cover the full UV to near-IR wavelength range. The FourStar filters overlap with broadband filters such as HST/WFC3/F125W, F140W and F160W in wavelength space, except they are narrower and sample the near-IR in more detail. The effective filter curves we use are modified to include the Lord et al. (1992) atmospheric transmission functions with a watercolumn of 2.3mm. The total integration time in each filter is shown in Table 2.1.

The sampling of the FourStar medium-bandwidth filters is illustrated in Figure 2.2, where we show the SEDs of observed galaxies in COSMOS with large Balmer/4000Å breaks at $z \gtrsim 1.5$. The FourStar near-IR filters are highlighted in red and their throughput is shown in the background. They are

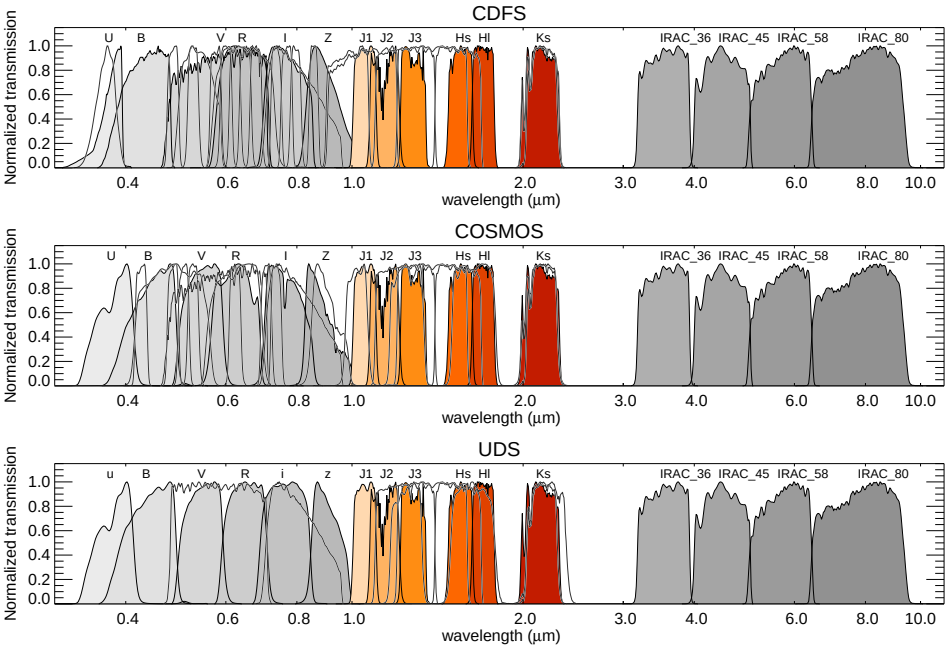

Figure 2.1: Normalized transmission corresponding to the FourStar medium-bandwidth and ancillary filters, each panel representing a different field. From top to bottom: CDFS, COSMOS and UDS. We show the FourStar J_1,J_2,J_3,H_s,H_l and K_s medium-bandwidth filters with different shades of red. The UV to optical U,B,V,R,I and Z filters and the Spitzer/IRAC filters are also highlighted (gray shaded curves). Note that these correspond to different instruments in each field. The FourStar filters overlap with other broadband near-IR filters, e.g. HST/WFC3/F125W–F160W, while providing a higher resolution sampling. Atmospheric transmission was included in all FourStar filter curves. All filters are mentioned separately in Tables 2.2 (CDFS), 2.3 (COSMOS) and 2.4 (UDS).

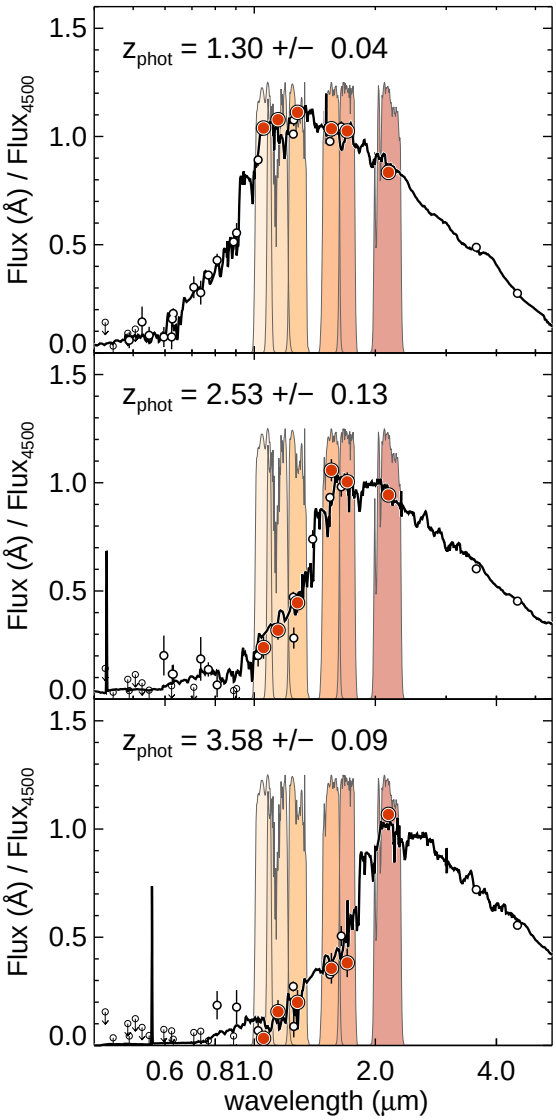

Figure 2.2: The FourStar filters provide detailed sampling of the Balmer/4000Å break of galaxies at $z \gtrsim 1.5$. Here we show the SEDs of three observed galaxies in COSMOS with large Balmer/4000Å breaks, at $z = 1.30$, $z = 2.53$ and $z = 3.58$. With increasing redshift, the Balmer/4000Å break moves through the range defined by the FourStar bands. Observed datapoints are shown as white or red dots with errorbars for ancillary and FourStar filters, respectively. Upper limits (mostly in the UV) are indicated with downwards pointing arrows. The solid curves are the EAZY best-fit SEDs (see Section 2.6). Observed and fitted SEDs are normalized at rest-frame 4500Å.

Table 2.1: FourStar observations

Cosmic field	Filter	Total integration time (hrs)	5σ depth (AB mag)
CDFS	J1	6.3	25.6
CDFS	J2	6.5	25.5
CDFS	J3	8.8	25.5
CDFS	Hs	12.2	24.9
CDFS	Hl	5.9	25.0
CDFS	Ks	5.0	24.8
COSMOS	J1	13.9	26.0
COSMOS	J2	16.0	26.0
COSMOS	J3	13.8	25.7
COSMOS	Hs	12.1	25.1
COSMOS	Hl	12.1	24.9
COSMOS	Ks	13.4	25.3
UDS	J1	7.9	25.6
UDS	J2	8.7	25.9
UDS	J3	9.3	25.6
UDS	Hs	11.0	25.1
UDS	Hl	10.4	25.2
UDS	Ks	3.9	24.7

particularly well suited to trace the Balmer/4000Å break at higher redshifts, which is crucial to derive photometric redshifts.

2.2.2 FourStar Image reduction

Pipeline

The FourStar data were reduced using a custom IDL pipeline written by one of the authors (I. Labbé) and also used in the NMBS (Whitaker et al. 2011). It employs a two-pass sky subtraction scheme based on the IRAF package xdimsum.

The pipeline processes the 4 FourStar detectors, which consist of dithered frames, separately for each $\sim 1-1.5$ hour observing block. Observed frames taken with each of the detectors are reduced and subsequently combined into a single mosaic.

Linearity corrections from the FourStar website[3] are applied to the raw data. Dark current was determined to be variable so we did not remove any dark pattern. We also found constant bias levels along columns and rows in the raw data. We therefore subtracted the median of a column/row from itself.

[3] http://instrumentation.obs.carnegiescience.edu/FourStar/calibration.html

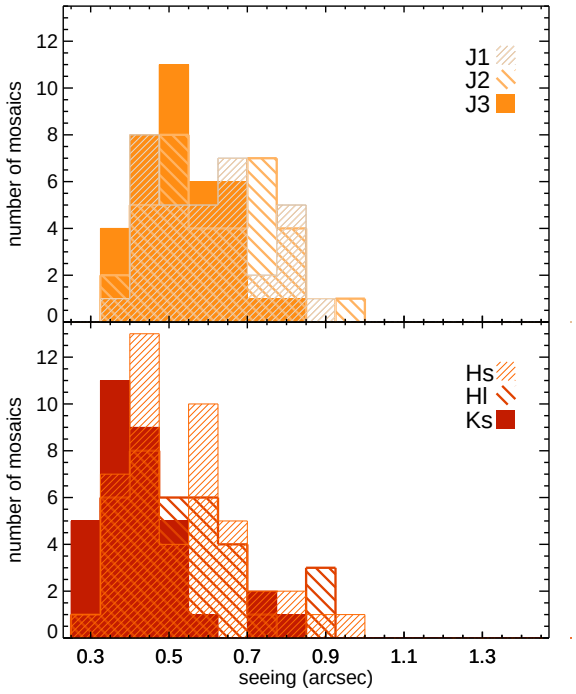

Figure 2.3: Seeing histograms of the FourStar single images, corresponding to ~ 1–1.5 hour observing blocks. Many of the images have a seeing of ~ 0.4 – 0.5″.

Master flat field data were produced using twilight observations. For the K_s-band, where thermal contributions play a role, we attempted to mitigate the impact of illumination from the warm telescope. By combining multiple dithered observations of a blank field at the end of a night when the telescope had cooled, we were able to characterise the telescope illumination pattern. Shortly afterwards we took twilight flats and subtracted the telescope illumination pattern from each exposure. The flats with the telescope contribution removed were normalised and combined into the master K_s-band flats.

Sky background models were subtracted from individual science exposures. The sky background was computed by averaging up to 8 images taken before and after that exposure. Masking routines were run to remove: (1) bad pixels via a static mask from the FourStar website (2) satellite trails (3) guider cameras entering the field of view and (4) persistence from saturated objects in previous exposures. Bad pixels only make up between 0.3 and 1.7 % of the detectors (Persson et al. 2013). In addition, the individual exposures were visually screened for any remaining tracking issues, asteroids, airplanes

and satellites.

Corrections for geometric distortion and absolute astrometric solutions are computed by crossmatching sources using astrometric reference images. In COSMOS we used the CFHT/i-band as reference (Erben et al. 2009; Hildebrandt et al. 2009), in CDFS we used ESO/MPG/WFI/I from the ESI survey (Erben et al. 2009; Hildebrandt et al. 2006) and in UDS the UKIDDS data release 8 K_s-band image (Almaini, in prep). The observations are interpolated onto a pixel grid with a resolution of 0.15 $''$/pix, which is close to the native scale of FourStar of 0.159 $''$/pix. The new grid shares the WCS tangent point (CRVAL) with the CANDELS HST images (Koekemoer et al. 2011; Grogin et al. 2011) and places CRVAL at a half-integer pixel position (CRPIX).

To optimize the signal-to-noise (SNR) of the images for each observing block (and for the final mosaics), they are weighted by their seeing, sky background levels and ellipticity of the PSF before they are combined. The seeing conditions at Las Campanas were extraordinarily good, with a median (wavelength uncorrected) seeing FWHM of $0.5''$ as shown in the histogram in Figure 2.3. Since the K_s-band cannot be observed with the HST, we only observed K_s-band when the seeing was excellent. This resulted in a median seeing in FourStar/K_s of $0.4''$.

Finally, we subtracted a background in the final mosaics using Source Extractor (SE; Bertin & Arnouts 1996) to ensure any remaining structure in the background did not impact the aperture photometry. In short, SE iteratively estimates the median of the distribution of pixel values in areas of 48×48 pixels in CDFS and COSMOS and 96×96 pixels in UDS. These estimates are smoothed on a scale of $3\times$ the background area, after which the background for the full images is calculated using a bicubic spline interpolation.

Photometric calibration

Here we describe how we derived the near-IR photometric zeropoints of the final mosaics. Since these vary significantly with changes in local precipitable water vapor and airmass, we employed a secondary standard photometric calibration scheme. First, we selected a nearby standard star. We selected relatively faint ($K_s = 14.5 - 17$ mag) spectrophotometric standard stars from the CALSPEC Calibration Database[4]. We then observed this primary standard star under photometric conditions immediately before or after a science observation in a particular filter. The science dataset was reduced and photometrically calibrated using the primary standard star observations and using an atmospheric watercolumn of 2.3mm. Secondly, we then selected bright, unsat-

[4] http://www.stsci.edu/hst/observatory/crds/calspec.html

urated stars in each of the chips of the science field for use as secondary standard stars. All other science observations of an observing block were then calibrated to the primary standard star via the secondary standard stars within each of the science fields.

In Section 2.6 we derive corrections to the zeropoints, that are typically of the order of 0.05 magnitude. We added these to the photometric zeropoints calculated here.

Image depths

We measured the depths of the FourStar images by determining the rms of the background pixels. Since pixels may be correlated on small scales, e.g. due to confusion or systematics introduced during the reduction process, we used a method in which we randomly placed 5000 apertures of $0.6''$ diameter in each background subtracted image. Due to the dither pattern the images have less coverage from individual frames at the edges. We therefore considered only regions with coverage within 80% of the maximum exposure. Sources were also masked, based on the SE segmentation maps after object detection (see Section 2.4.1).

The resulting aperture flux distributions, representing the variation in the noise, were fit with a gaussian, from which we derived the standard deviation (σ). We then applied the point-spread-functions (PSFs) derived from bright stars (further explained in Section 2.3), to determine a flux correction for missing light outside of the aperture. σ was then multiplied by 5 and converted to magnitude using the effective zeropoint (the photometrically derived zeropoints as desribed above, with a correction applied) of each FourStar mosaic, to obtain the limiting depth at 5σ confidence. The resulting depth in AB magnitude can thus be summarized as

$$\mathrm{depth}(5\sigma) = zp - 2.5\log_{10}[5\sigma apcorr] \tag{2.1}$$

with zp the zeropoint of the image and $apcorr$ the aperture flux correction (typically factors of $1.7-2.6$, depending on the seeing). The 5σ depths are summarized in Table 2.1 and have typical values of $25.5-26.0$ AB mag in J_1, J_2, J_3 and $24.9-25.2$ AB mag in H_s, H_l and $24.7-25.3$ AB mag K_s.

In Figure 2.4 we show as an example the FourStar/K_s-band image in COSMOS. We also compare with the near-IR CANDELS/HST/WFC3/F160W observations, with FWHM=$0.19''$ and a limiting 5σ depth of 26.4 AB mag. The deeper space-based F160W image has a higher resolution, but as a result of the very deep magnitude limits combined with excellent seeing conditions we can achieve almost a similar quality for our near-IR ground-based observa-

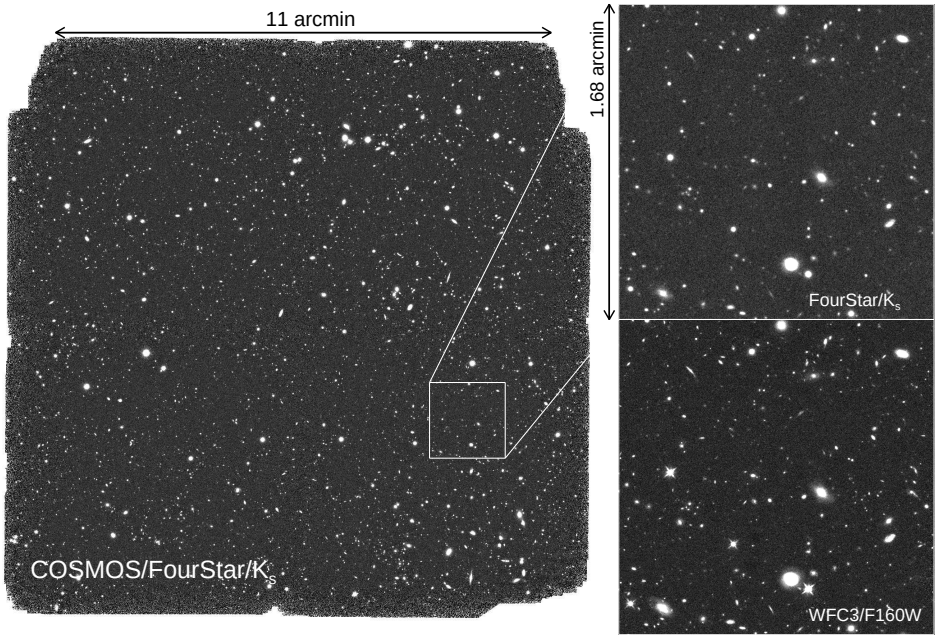

Figure 2.4: Left: The FourStar/K_s-band reduced image in COSMOS. The FourStar footprint is $13' \times 13'$. Top right: zooming in on a $1.68' \times 1.68'$ region in the COSMOS field. Bottom right: the same region with HST/WFC3/F160W.

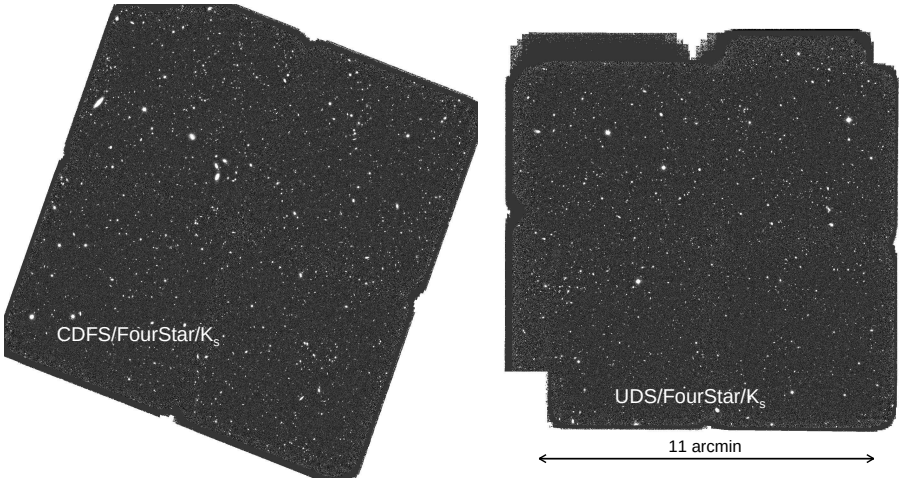

Figure 2.5: FourStar/K_s-band reduced images of CDFS and UDS

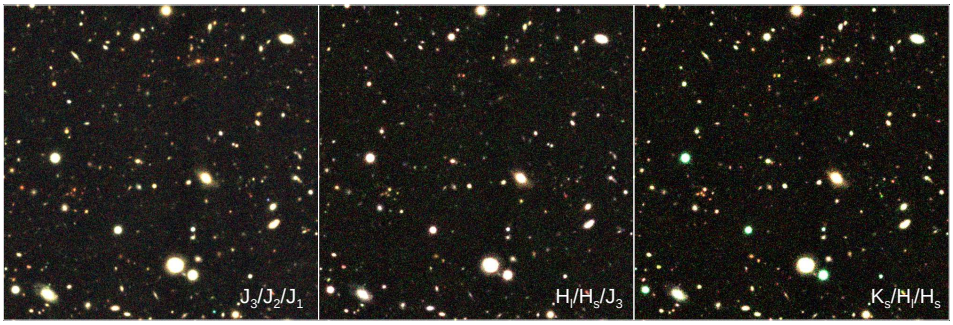

Figure 2.6: False colour images of the cutout region in Figure 2.4, demonstrating the high quality obtained with the FourStar filters, as well as the efficiency of using medium-bandwidth filters to determine the colors of galaxies within a classical J or H broadband. The filter combinations that were used in each panel are indicated at the bottom (red/green/blue).

tions. The K_s-band images in CDFS and UDS have similar depth as in COSMOS and are shown in Figure 2.5. To highlight the wealth of information provided by the fine spectral sampling of the FourStar medium-bandwidth filters we show again in Figure 2.6 the same cut-out region of Figure 2.4, using different filter combinations.

2.2.3 K_s-band detection images

We combine our FourStar/K_s-band observations with deep pre-existing K-band imaging to create super-deep detection images. In CDFS we use VLT/HAWK-I/K from HUGS (with natural seeing between 0.3 $''$ and 0.5 $''$) (Fontana et al. 2014), VLT/ISAAC/K (v2.0) from GOODS, including ultra deep data in the HUDF region (seeing= 0.5 $''$ (Retzlaff et al. 2010), CFHST/WIRCAM/K from TENIS (seeing= 0.9 $''$) (Hsieh et al. 2012), and Magellan/PANIC/K in HUDF (seeing= 0.4 $''$) (PI: I. Labbé). In COSMOS we add VISTA/K from UltraVISTA (DR2) (seeing= 0.7 $''$) (McCracken et al. 2012) and in UDS we add imaging with UKIRT/WFCAM/K from UKIDSS (DR10) (seeing= 0.7 $''$) (Almaini et al, in prep) and also natural seeing VLT/HAWK-I/K imaging from HUGS.

Using sources common to the images a distortion map was determined. Subsequent bicubic spline interpolation was used to register the images to better than 0.03$''$ across the field. We then determined the background root-mean-square flux variation (σ_{rms}) and the seeing in each image, and we used these to assign a weight using $weight = 1/(\sigma_{rms} \times seeing^2)$. The final com-

bined image stacks were obtained by a weighted average of the individual K- and K_s-band science images. Weight maps were obtained by averaging the individual exposure maps in the same way as the science images. The final K_s-band stacks have maximum limiting depths at 5σ significance of 25.5 and 25.7 AB mag in COSMOS and UDS, respectively, which are 0.2 and 1.0 magnitudes deeper than the individual FourStar/K_s-band observations. The depth in CDFS varies between 26.2 and 26.5, 1.4 to 1.7 magnitudes deeper than the FourStar/K_s-band image only. The average seeing in the three fields is $FWHM = 0.45,\ 0.58$ and $0.60''$. We use these images for source detection (Section 2.4), after calculating and subtrating the background.

2.2.4 Ancillary data

In addition to the 6 FourStar filters, we incorporate imaging in 20-34 filters into each catalog, from publicly available surveys at $0.3-8\mu$m. In CDFS we have a total of 40 bands, in COSMOS a total of 37 and in UDS a total of 26. These are summarized in Tables 2.2, 2.3 and 2.4, where we additionally show, for every image, the central wavelength, PSF FWHM (see Section 2.3), effective zeropoint, galactic extinction value and zeropoint offset derived in Section 2.6. The galactic extinction values were calculated using the $E(B-V)$ values from Schlafly & Finkbeiner (2011), interpolated between the given bandpasses and the central wavelengths of our filterset.

The CDFS UV-to-optical filters include VLT/VIMOS/U,R-imaging (Nonino et al. 2009), HST/ACS/B,V,I,Z-imaging (Giavalisco et al. 2004; Wuyts et al. 2008), ESO/MPG/WFI/U_{38},V,R_c-imaging (Erben et al. 2005; Hildebrandt et al. 2006), HST/WFC3/$F098M,F105W,F125W,F140W,F160W$ and HST/ACS/ $F606W,F814W-$ imaging (Grogin et al. 2011; Koekemoer et al. 2011; Windhorst et al. 2011; Brammer et al. 2012), 11 Subaru optical medium bands (Cardamone et al. 2010) with seeing $< 1.1''$ and CFHT/WIRCAM/K-band imaging (Hsieh et al. 2012).

In COSMOS we added CFHT/u,g,r,i,z-imaging (Erben et al. 2009; Hildebrandt et al. 2009), Subaru/$B,V,r+,z+$-imaging and 7 optical medium-bandwidth filters (Taniguchi et al. 2007) with seeing $< 1.1''$, HST/WFC3/$F125W$, $F140W,F160W$ and HST/ACS/$F606W,F814W$-imaging (Grogin et al. 2011; Koekemoer et al. 2011; Brammer et al. 2012) and UltraVISTA/Y,J,H,K_s-imaging (McCracken et al. 2012).

In UDS the additional filters are CFHT/MegaCam/U (Almaini/Foucaud, in prep), Subaru/Surpime-Cam/B,V,R,i,z (Furusawa et al. 2008), UKIRT/ WFCAM/J,H,K_s (Almaini, in prep), HST/WFC3/$F125W,F140W,F160W$ and HST/ACS/$F606W,F814W$ (Grogin et al. 2011; Koekemoer et al. 2011; Bram-

Table 2.2: CDFS passband parameters

Filter	λ_c (μm)	FWHM ($''$)	zeropoint (AB mag)	offset	galactic extinction
B	0.4318	0.73	22.219	-0.029	-0.032
I	0.7693	0.73	22.141	0.019	-0.014
R	0.6443	0.65	27.655	-0.145	-0.020
U	0.3749	0.81	26.357	-0.170	-0.037
V	0.5919	0.73	23.030	-0.010	-0.022
Z	0.9036	0.73	21.318	0.041	-0.011
Hs	1.5544	0.60	26.691	-0.033	-0.004
Hl	1.7020	0.50	26.700	-0.053	-0.004
J1	1.0540	0.59	26.372	-0.044	-0.009
J2	1.1448	0.62	26.655	-0.042	-0.006
J3	1.2802	0.56	26.671	-0.070	-0.006
Ks	2.1538	0.46	27.027	-0.086	-0.003
NB118	1.1909	0.47	24.680	0.000	-0.006
NB209	2.0990	0.45	24.792	0.000	-0.003
F098M	0.9867	0.26	25.664	0.011	-0.008
F105W	1.0545	0.24	26.278	-0.002	-0.007
F125W	1.2471	0.26	26.231	0.004	-0.005
F140W	1.3924	0.27	26.484	-0.027	-0.004
F160W	1.5396	0.27	25.950	-0.000	-0.004
F814W	0.8057	0.22	25.963	-0.004	-0.011
IA484	0.4847	0.81	25.529	-0.004	-0.024
IA527	0.5259	0.87	25.793	-0.052	-0.022
IA574	0.5763	1.01	25.871	-0.141	-0.019
IA598	0.6007	0.69	26.072	-0.034	-0.018
IA624	0.6231	0.67	25.889	0.018	-0.017
IA651	0.6498	0.67	26.223	-0.057	-0.016
IA679	0.6782	0.86	26.292	-0.077	-0.015
IA738	0.7359	0.83	26.033	-0.000	-0.013
IA767	0.7680	0.77	26.077	-0.024	-0.012
IA797	0.7966	0.74	26.051	-0.020	-0.012
IA856	0.8565	0.74	25.746	-0.005	-0.010
WFI_V	0.5376	0.96	24.187	-0.070	-0.021
WFI_Rc	0.6494	0.84	24.702	-0.035	-0.016
WFI_U38	0.3686	0.98	22.221	-0.279	-0.032
tenisK	2.1574	0.86	23.670	0.232	-0.002
IRAC_36	3.5569	1.50	20.086	-0.016	0.000
IRAC_45	4.5020	1.50	20.065	0.005	0.000
IRAC_58	5.7450	1.90	20.580	0.023	0.000
IRAC_80	7.9158	2.00	21.759	0.022	0.000

Zeropoints are the effective zeropoints. These have galactic extinction and zeropoint corrections derived in Section 2.6 incoorporated, i.e. $zp = zp_I + offset + GE$, with zp_I representing the photometrically derived zeropoint of image I, $offset$ the zeropoint correction and GE the galactic extinction value. The corrections (in units of AB magnitude) are indicated in separate columns.

Table 2.3: COSMOS passband parameters

Filter	λ_c (μm)	FWHM ($''$)	zeropoint (AB mag)	offset	galactic extinction
B	0.4448	0.61	31.615	-0.139	-0.076
G	0.4870	0.90	26.411	0.033	-0.069
I	0.7676	0.77	25.632	0.105	-0.034
IA427	0.4260	0.79	31.621	-0.143	-0.079
IA484	0.4847	0.75	31.536	-0.067	-0.069
IA505	0.5061	0.82	31.504	-0.038	-0.065
IA527	0.5259	0.74	31.477	-0.016	-0.061
IA624	0.6231	0.72	31.425	0.025	-0.050
IA709	0.7074	0.81	31.435	0.006	-0.042
IA738	0.7359	0.80	31.437	0.003	-0.039
R	0.6245	0.79	25.923	0.050	-0.047
U	0.3828	0.82	25.484	-0.166	-0.086
V	0.5470	0.80	31.345	0.115	-0.059
Rp	0.6276	0.83	31.319	0.128	-0.047
Z	0.8872	0.74	24.668	0.130	-0.030
Zp	0.9028	0.90	31.235	0.195	-0.030
Hl	1.7020	0.60	26.577	0.034	-0.010
Hs	1.5544	0.54	26.573	0.062	-0.012
J1	1.0540	0.57	26.344	0.029	-0.020
J2	1.1448	0.55	26.550	0.040	-0.018
J3	1.2802	0.53	26.583	0.011	-0.016
Ks	2.1538	0.47	26.953	-0.011	-0.006
NB118	1.1909	0.58	24.673	0.000	-0.018
NB209	2.0990	0.52	24.861	0.000	-0.006
F125W	1.2471	0.26	26.258	-0.000	-0.011
F140W	1.3924	0.26	26.475	-0.000	-0.010
F160W	1.5396	0.26	25.964	-0.000	-0.008
F606W	0.5921	0.20	26.519	0.011	-0.038
F814W	0.8057	0.21	25.925	0.042	-0.024
UVISTA_J	1.2527	0.82	29.947	0.064	-0.011
UVISTA_H	1.6433	0.81	30.004	0.004	-0.008
UVISTA_Ks	2.1503	0.79	29.970	0.036	-0.006
UVISTA_Y	1.0217	0.85	29.950	0.066	-0.016
IRAC_36	3.5569	1.70	21.620	-0.039	0.000
IRAC_45	4.5020	1.70	21.611	-0.030	0.000
IRAC_58	5.7450	1.90	21.589	-0.008	0.000
IRAC_80	7.9158	2.00	21.674	-0.093	0.000

Table 2.4: UDS passband parameters

Filter	λ_c (μm)	FWHM ($''$)	zeropoint (AB mag)	offset	galactic extinction
u	0.3828	1.06	25.558	-0.208	-0.089
B	0.4408	0.91	25.158	-0.083	-0.074
V	0.5470	0.93	25.112	-0.054	-0.058
R	0.6508	0.96	25.085	-0.036	-0.049
i	0.7656	0.98	25.025	0.010	-0.035
z	0.9060	0.99	25.037	-0.010	-0.027
J1	1.0540	0.55	26.238	-0.038	-0.022
J2	1.1448	0.53	26.506	-0.030	-0.019
J3	1.2802	0.51	26.553	-0.023	-0.015
Hs	1.5544	0.49	26.613	-0.000	-0.011
Hl	1.7020	0.51	26.542	-0.038	-0.010
Ks	2.1538	0.44	26.957	-0.073	-0.006
J	1.2502	0.91	30.999	-0.054	-0.015
H	1.6360	0.89	31.499	-0.109	-0.010
K	2.2060	0.86	31.964	-0.067	-0.006
F125W	1.2471	0.26	26.246	-0.000	-0.016
F140W	1.3924	0.26	26.465	-0.000	-0.013
F160W	1.5396	0.26	25.957	-0.000	-0.011
F606W	0.5893	0.20	26.585	-0.040	-0.054
F814W	0.8057	0.23	25.964	0.013	-0.033
Y	1.0207	0.58	27.022	0.000	-0.022
IRAC_36	3.5569	1.70	21.620	-0.039	0.000
IRAC_45	4.5020	1.70	21.600	-0.019	0.000
IRAC_58	5.7450	1.90	21.694	-0.113	0.000
IRAC_80	7.9158	2.00	21.612	-0.031	0.000

mer et al. 2012) and VLT/HAWK-I/Y (Fontana et al. 2014).

In CDFS and UDS we have additionally available FourStar narrow-bandwidth data at 1.18μm (FourStar/NB118) and 2.09μm (FourStar/NB209) (Lee et al. 2012). The narrowbands are sensitive to emission line flux. Small bandwidths in combination with high SNR for some galaxies may lead to biased photometric redshift and stellar mass estimates, because the models we use for determining redshifts and stellar population parameters do not contain well-calibrated strong emission lines. As such, they are incorporated into the catalogs, but are not used to derive photometric redshifts or stellar masses. The images have 5σ image depths of 25.2 and 24.8 AB mag in NB118 and CDFS and COSMOS, respectively and 24.4 and 24.0 AB mag in NB209.

The *Spitzer*/IRAC/3.6 and 4.5μm images used in CDFS are the ultradeep mosaics from the IUDF (PI: Labbé), using data from the cycle 7 IUDF program (PI: Labbé), IGOODS (PI: Oesch), GOODS (PI: Dickinson), ERS (PI: Fazio), S-CANDELS (PI: Fazio), SEDS (PI: Fazio) and UDF2 (PI: Bouwens). In CDFS we further use *Spitzer*/IRAC/5.8 and 8.0μm images from GOODS (Dickinson et al. 2003). In COSMOS and UDS we use the 3.6 and 4.5μm images from SEDS (Ashby et al. 2013). The 5.8 and 8.0μm data in COSMOS are from S-COSMOS (Sanders et al. 2007) and in UDS from spUDS (Dunlop et al, in prep).

The ancillary images are registered and interpolated to the same grid as the FourStar mosaics, using the program wregister in IRAF. Backgrounds were estimated with SE and manually subtracted.

We further supplement the optical/near-IR catalogs with deep far-IR imaging from *Spitzer*/MIPS at 24μm (GOODS-S: PI Dickinson, COSMOS: PI Scoville, UDS: PI Dunlop). Median 1σ flux uncertainties in 24μm for the COSMOS and UDS pointings are roughly 10μJy. The CDFS pointing is deeper with a median 1σ flux uncertainty of 3.9μJy. In CDFS we additionally make use of public *Herschel*/PACS observations from PEP (Magnelli et al. 2013) at 100μm and 160μm, with 1σ flux uncertainties of 205 and 354μJy respectively.

2.3 PSF matching

The full UV/optical to near-IR dataset contains images of varying seeing quality. The FWHMs of the PSF corresponding to each image varies between 0.2″ for the HST bands to 1.05″ for some of the UV/optical images. To measure aperture fluxes consistently over the full wavelength range, i.e. measuring the same fraction of light per object in each filter, the images have to be convolved so that the PSFs match. To achieve a consistent PSF we first characterize the PSF in all individual images, we then define a theoretical model

PSF as a reference, and finally convolve all bands to match the reference PSF.

The average PSF for each image was produced by selecting unsaturated stars with high SNR (> 150) (see Section 2.4.6, in which we describe how stars were identified from galaxies), in postage stamps of $10.65'' \times 10.65''$. For each star we measured a growthcurve, i.e the total integrated light as a function of radius, with nearby objects masked using the SE segmentation map. Outliers, such as saturated stars, were then determined based on the shape of their light profile compared with the median growthcurve, and rejected from the sample. We median averaged the remaining stars, and, after scaling to match the brightness of each individual star, used this to fill in masked regions. After renormalizing each tile by the total integrated flux at sufficiently large radius (25 pixels or $3.75''$) we again stacked the postage stamps to obtain a median star. Finally, to obtain clean sample, we again compared the light profiles of individual stars against the median light profile, and iteratively rejected stars if the average deviation from the median growthcurve squared exceeded 5%. The result was a tightly homogeneous sample of stars, from which we obtained the final median 2-dimensional PSF.

We generated as a reference PSF a model Moffat profile with full-width-at-half-maximum (FWHM) = $0.9''$ and $\beta = 2.5$. The advantage of using a model PSF rather than the average PSF from an image, is that a theoretical model is noiseless. To convolve the images to match the target model PSF, we first derive a kernel for each image individually. For this we use a deconvolution code developed by I. Labbé, which fits a series of Gaussian-weighted Hermite polynomials to the Fourier transform of the PSF. The original images were then convolved with this kernel towards the target quality. This method results in very low residuals and is optimal for images with a smaller average PSF than the target PSF or if the PSF is slightly larger (by 15%), as we show in Appendix 2.A, and which is the case for 12% of the filters included. It improves the accuracy of the final convolved PSFs, compared with e.g. maximum likelihood algorithms. For example, Skelton et al. (2014) find < 1% accuracy when convolving HST/WFC3 images, using the same technique as employed here, compared to e.g. Williams et al. (2009); Whitaker et al. (2011), who use maximum likelihood methods to match point sources to within 2−5% accuracy.

PSF growthcurves before and after convolution are shown in Figure 2.7, normalized by the model PSF. For each convolved image we obtain excellent agreement, within 1.5% at $r < 0.6''$. IRAC photometry, with FWHM> $1.5\ ''$ is treated separately in Section 2.4.5.

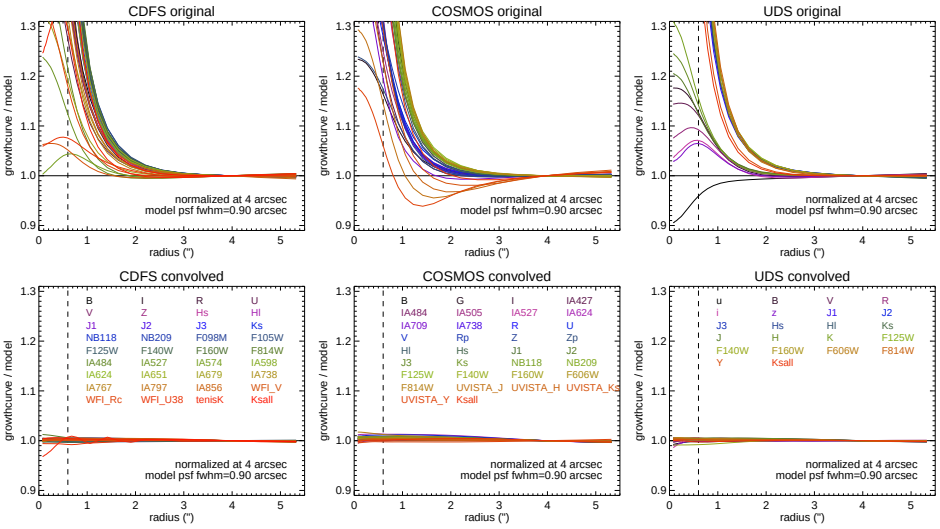

Figure 2.7: Top: Growthcurves of the median stacked PSF of stars in the unconvolved images, normalized at $4''$ radius and divided by the growthcurve of the target moffat PSF. The vertical dashed lines represent the radius at which we measure flux (Section 2.4.3). The spread in integrated flux is very large between different images, which would lead to biased color measurements. Bottom: Here we show the growthcurves of the convolved images, where each PSF is convolved to match a Moffat profile. The correspondence with the target PSF is almost one-on-one, with at most a 1.5% deviation at $r = 0.6''$.

2.4 Photometry

2.4.1 Source detection

We created detection images from the superdeep background subtracted K_s-band stacks, as described in Section 2.2.3, by noise equalizing the images, i.e. multiplying the images with the square root of the corresponding weight images. We then ran SE to create a list of sources and their locations. We optimized source detection by setting the deblending parameters of SE to `DEBLEND_THRESH`= 64 and `DEBLEND_MINCONT`= 0.0000001 and the clean parameter (`CLEAN`) to `N`. We also let SE generate a segmentation map representing the location and area of each source. The total number of sources in the catalogs is 30,911 in CDFS, 20,786 in COSMOS and 22,093 in UDS. All SE parameters used are included in the ZFOURGE data release.

2.4.2 K_s-band total flux determination

To measure the K_s-band total flux, SE was run in dual image mode, using the detection images described in the previous section and the original unconvolved K_s-band stacks to measure photometry. We used a flexible elliptical aperture (Kron 1980), to obtain SE's `FLUX_AUTO`.

This estimate is not yet the total K_s-band flux and we have to account for missing flux outside the aperture. We derived a correction factor from the stacked K_s-band PSF (separately for each field). This aperture correction varies between sources and is a function of the size of the auto-aperture that was estimated by SE.

We determined the correction by measuring the integrated PSF flux within the circularized aperture radius as well as the total integrated flux at $r < 4\,''$. Total K_s-band fluxes were then calculated using

$$F_{Ks,tot} = F_{Ks,auto} \frac{F_{PSF}(< 4\,'')}{F_{PSF}(< r_{Kron})} \tag{2.2}$$

(Labbé et al. 2003; Whitaker et al. 2011), with $F_{Ks,tot}$ the total K_s-band flux, $F_{Ks,auto}$ the flux within the auto-aperture, i.e. `FLUX_AUTO` from SE, $F_{PSF}(< 4\,'')$ the integrated PSF flux within a radius of $3''$ and $F_{PSF}(< r_{Kron})$ the flux within the circularized Kron radius.

We additionally measured the total flux using a fixed circular aperture, of $\sim 1.5\times$ the PSF FWHM of the deep K_s-band stacks. In CDFS we therefore used a $0.7''$ diameter aperture and in COSMOS and UDS a $0.9''$ diameter aperture. The measured aperture fluxes were also corrected for flux outside of the aperture.

In total we have two estimates for the total flux, one using the auto aperture flux, and one using a fixed circular aperture. We note that for small, low SNR sources, the autoscaling aperture size may be underestimated, leading to extreme aperture corrections. Therefore, we only considered the circular aperture measurements for sources if their circularised Kron radius was very small, i.e. smaller than the circular aperture radius.

2.4.3 Aperture fluxes

In addition to the total K_s-band flux, we derived flux estimates in all filters in the three ZFOURGE fields. We ran SE in dual image mode, using the combined K_s-band images for source detection and the PSF matched images to measure photometry. We use the PSF matched images to make sure the captured light within the apertures is consistent over all the images. We also included the convolved versions of the deep K_s-band stacks. We use circular apertures of $1.2''$ diameter, which are suffiently large to capture most of the light (the PSFs of the convolved images have a FWHM$= 0.9''$), but small enough to optimize SNR.

We correct all aperture fluxes to total, using the ratio between the total flux in the original deep K_s-band stacked images to the aperture flux in the PSF matched K_s-band stack, i.e.:

$$F_{F,tot} = F_{F,aper} * \frac{F_{Ks,tot}}{F_{Ks,aper}} \tag{2.3}$$

Here, $F_{F,tot}$ is the aperture flux in filter F scaled to total, $F_{F,aper}$ the unscaled aperture flux, $F_{Ks,tot}$ the total K_s-band flux described in Section 2.4.2 and $F_{Ks,aper}$ the aperture flux from the PSF-matched K_s-band image stacks.

2.4.4 Flux uncertainties

The uncertainty on the flux measured in an aperture has contributions from the background, the Poisson noise of the source, and the instrument read noise. The relative contribution from the latter two effects will be very small for the faint galaxies and medium-band filters used in this study (Persson et al. 2013). If the adjacent pixels in an image are uncorrelated, the background noise σ_{RMS} measured in an aperture containing N pixels will scale in proportion to $\sqrt{N}$. In a more realistic scenario, pixels are expected to be correlated on small scales due to interpolation or PSF smoothing and on large scale due to imperfect background subtraction, flux from extended objects, undetected sources, or systematics introduced in the reduction process, such as flat field errors. For perfectly correlated pixels, the background noise is

expected to scale as $\sigma_{RMS} \propto N$. The actual scaling of the noise in an image lies somewhere in between and can be parameterized by

$$\sigma_{NMAD} = \sigma_1 \alpha N^{\beta/2} \tag{2.4}$$

with σ_{NMAD} the normalized median absolute deviation and β taking on a value between $1 < \beta < 2$. α is a normalization parameter and σ_1 is the standard deviation of the background pixels. (Labbé et al. 2003; Quadri et al. 2007; Whitaker et al. 2011) We estimated the noise as a function of aperture size empirically by randomly dropping circular apertures of varying diameter (2000 apertures each time) in each image that was used for photometry. These are the convolved images for the aperture fluxes and the unconvolved K_s-band stacks that were used to measure total flux. We used the SE segmentation map to mask sources. We also excluded regions with low weight, such as the edges of the `FourStar` detectors.

For each aperture diameter, we fit a gaussian to the measured flux distribution and obtained the standard deviation (σ_{RMS}). We then fit Equation 2.4 to the various estimates of σ_{RMS} as a function of N pixels in each aperture, to obtain σ_1, α and β.

For circular apertures with radius r pixels, the uncertainty (e_F) on the flux measurement in filter F is

$$e_F = \sigma_{NMAD}(r)/\sqrt{w_F} = \sigma_1 \alpha (\pi r^2)^{\beta/2}/\sqrt{w} \tag{2.5}$$

with w_F the median normalized weight. Weights were obtained from the median normalized exposure images and were measured as the median in apertures with sizes corresponding to those used to measure flux.

The radius r used in Equation 2.5 was chosen to match the aperture size used for the different flux determinations. $1.2''$ diameter apertures are used for the aperture fluxes and, for the total flux, we use SE's `KRON_RADIUS`, which is based on autoscaling kron-like apertures.

The aperture flux uncertainties obtained from Equation 2.5 were scaled to total for a consistent relative error.

2.4.5 IRAC and MIPS photometry

The Spitzer/IRAC and MIPS and Herschel/PACS images have much broader PSFs than the UV, optical and near-IR images and source blending is a considerable effect. The FWHM in the IRAC images is typically $> 1.5''$ and in MIPS $> 4''$. To obtain photometry, we use a source fitting routine that models and subtracts profiles of neighbouring objects prior to measuring photometry for

a target (Labbé et al. 2006; Wuyts et al. 2008; Whitaker et al. 2011; Skelton et al. 2014, Tomczak et al, 2015, submitted).

The position and extent of each source was based on the SE segmentation maps derived from the super deep K_s-band detection images. The K_s-band images are assumed to provide a good prior for the location and extent of the unresolved far-IR flux, as sources that are bright in K are also typically bright at redder infra-red wavelengths. Each source in the K_s-band image was extracted using the segmentation map and convolved to match the PSF of the lower resolution far-IR image, assuming negligible morphological corrections. All sources were then fit simultaneously to create a model for the lower resolution image. Next, for each source in the lower resolution image, the modelled light of neighbouring sources was subtracted, after which we measured the flux on the cleaned maps within circular apertures with diameter D, using $D = 1.8''$ for IRAC and $D = 7''$ for MIPS.

To correct the far-IR aperture flux to total, the measurements were multipied by the ratio of the total K_s-band flux to the $D = 1.8''$ aperture flux on the PSF convolved K_s-band template image. Because the MIPS PSF has significant power in the wings at large radii, which are not represented in the convolution kernel, we apply an additional fixed correction of $\times 1.2$ to account for missing flux at $r > 15''$ (using values for point-sources from the MIPS instrument handbook).

Flux uncertainties were estimated from background maps that were individually generated for each source.

2.4.6 Stars

The majority of stars were identified by their observed $B - J123$ and $J123 - K_s$ colors. $J123$ here is derived as the median of the flux in the $J1$, $J2$ and $J3$ filters. Stars form a tight sequence in $J123 - K_s$ compared to galaxies. In the first two panels of figure 2.8 our selection criterion is indicated as a red line, with stars having:

$$(J123 - K_s) < 0.288(B - J123) - 0.52 \qquad [(B - J123) < 2.5]$$
$$(J123 - K_s) < 0.08(B - J123) \qquad [(B - J123) > 2.5]$$

Here we only classified sources as stars if they are below the red line at $> 2\sigma$ confidence in $J123 - K_s$. By selecting at $> 2\sigma$ confidence, we automatically reject faint sources that scatter below the red line from the star sample. This is illustrated by the histograms in the third panel of Figure 2.8, where we show the magnitude counts of stars against sources that are not now classified as stars, but would have been otherwise. These have a distribution of

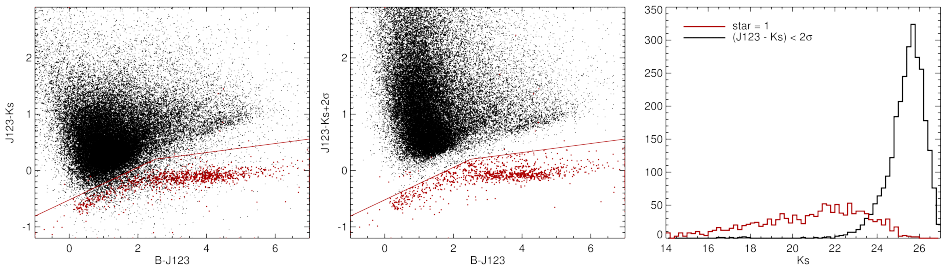

Figure 2.8: Left and Middle panel: $B-J123$ versus $J123-K_s$ for stars (red) and galaxies (black). The red line indicates the star/galaxy separation. In the first panel some sources are scattered below the red line. These are effectively removed from the star classification by selecting in $J123-K_s$ at $>2\sigma$ confidence (middle panel). Right panel: K_s-band magnitudes of stars (red) and sources that scatter below the red line (black). By only selecting at $>2\sigma$ confidence in $J123-K_s$, we exclude the faintest sources from the star sample.

magnitudes that peaks around $K_s = 25.5-26$ magnitude.

For sources that were not covered by the $J1$, $J2$ and $J3$ bands, we used broadband J or F125W where available, but only considered sources brighter than 25 mag in K_s. For sources without B-band coverage, we used $J123-K_s<0$ to classify stars, considering only sources brighter than 22 magnitude in K_s. An finally, if sources did have B-band coverage, but were saturated in B, we also used $J123-K_s<0$.

Red cool stars may not be selected in this way, as they have red $J-K$ colors. To ensure that we cover all types of stars, we fit the observed SEDs of all sources with EAZY using the speX stellar library[5]. For a few sources that were not flagged already by their $B-J123$ and $J123-K$ colors, the reduced χ^2 indicated a stellar template was a better fit to the data than the best-fit galaxy template (Section 2.6) and we flagged these sources as stars as well.

A source that is not selected by any of the methods above, is considered a saturated star if it is brighter than 16 magnitude in J_1 or K_s and at the same time could not be fit to a galaxy template, having a large reduced χ^2, which we empircally estimated by inspecting many SEDs to be $\chi^2>3000$.

In total, 1.8% of the sources in the catalogs are classified as stars.

2.4.7 A standard selection of galaxies

For convenient use of the catalogs, we have designed a `use` flag. This takes into account SNR, the star/galaxy classifications described above and the

[5] `http://pono.ucsd.edu/~adam/browndwarfs/spexprism/`

depth of the images at the respective source locations. This flag also includes sources that are well within the FourStar footprint and are observed with each of the near-IR medium-bandwidth filters. We note that the K_s-band stacks cover a somewhat larger area, especially in CDFS, which means that not all sources in the catalogs have FourStar imaging (although the majority do).

A standard selection of galaxies can be obtained by selecting on use=1 from the catalogs. The use flag is set to 1 for sources that pass the following criteria:

- Not within $r('') < 10-(m-16)$ of a bright star (< 17.5 magnitude in J_1, J_2, J_3, J or K_s), with m the apparent magnitude of the star (Table 2.6: nearstar=0).
- Detected with $SNR_{K_s} \geq 5$. The SNR is obtained from the total K_s-band detection images in seeing dependent circular apertures, $D = 0.7''$ for CDFS and $D = 0.9''$ for COSMOS and UDS.
- A minimum exposure time of at least $0.1\times$ the median exposure in all the FourStar bands (Table 2.6: wmin_fs≥ 0.1). Effectively this means that every source has at least 20 minutes exposure in each FourStar band. Because of the dither pattern, sources with lower weight that are removed by this criterion lie at the edges of the images.
- Coverage in all optical bands (Table 2.6: wmin_optical> 0).
- Not a catastrophic EAZY fit: χ^2 (reduced) ≤ 1000 (based on an empirical estimation from inspecting many fits).
- Not a catastrophic FAST fit, i.e. a finite and positive stellar mass estimate.

The use flag allows for a straightforward sample selection, representing galaxies with good photometry, i.e. high SNR sources from well exposed regions of the images. We note that for specific science goals a different selection may be optimal. We also warn that the use=1 sample may still contain problematic sources, with for example uncertain photo-z's and poorly constrained EAZY or FAST fits, and we recommend to always inspect the individual SEDs. However, the use=1 sample should be a reliable representation of the galaxy population in large statistical studies.

The total area of the Ks-band detection images is 280.9 $'^2$ for CDFS, 176.5 $'^2$ for COSMOS and 189.3 $'^2$ for UDS. Selecting only galaxies with wmin_fs> 0.1 that are not near bright stars, reduces the area to 132.2 $'^2$, 139.2 $'^2$ and 135.6 $'^2$.

2.4.8 Catalog format

We provide separate photometric catalogs for each cosmic field. These contain the coordinates, total fluxes, flux uncertainties, weight estimates, flags and

Table 2.5: Explanation of the photometric catalog header

id	ID number
x,y	pixel coordinates (scale: 0.15″ / pixel)
ra,dec	right ascension, declination (J2000)
SEflags	Source Extractor flags
iso_area	isophotal area above Source Extractor analysis treshold (pix^2)
fap_Ksall[a]	(convolved) K_s-band aperture flux within a 1.2″ diameter circular aperture
eap_Ksall	uncertainty on fap_Ksall
apcorr	aperture correction applied to fauto_Ksall to obtain f_Ksall (f_Ksall = fauto_Ksall * apcorr)
Ks_ratio	ratio between fap_Ksall and f_Ksall (Ks_ratio = f_Ksall / fap_Ksall)
fapcirc0D_Ksall[b]	aperture flux measured within a D″ diameter circular aperture
eapcirc0D_Ksall	uncertainty on fapcirc0D_Ksall
apcorr0D	aperture correction applied to fapcirc0D_Ksall to obtain fcirc0D_Ksall (fcirc0D_Ksall = fapcirc0D_Ksall * apcorr0D)
fcirc0D_Ksall[a,b]	total (aperture corrected) K_s-band flux within a D″diameter circular aperture
ecirc0D_Ksall	uncertainty on fcirc0D_Ks
fauto_Ksall	K_s-band flux within a Kron-like elliptical aperture
flux50_radius	radius (pixels) enclosing 50% of the K_s-band flux
a_vector	major axis of a Kron-like elliptical aperture
b_vector	minor axis of a Kron-like elliptical aperture
kron_radius	radius of a circularised Kron-like elliptical aperture

[a] Note that these K_s-band fluxes are derived from the superdeep combined Ks-band images. Within the catalogs only f_Ks corresponds to `FourStar`/K_s.

[b] In CDFS $D = 0.7''$, in COSMOS and UDS, $D = 0.9''$ (i.e. 1.5× the seeing FWHM).

Table 2.6: Explanation of the photometric catalog header continued

f_Ksall[a]	total (aperture corrected) K_s-band flux within a Kron-like elliptical aperture
e_Ksall	uncertainty on f_Ksall
w_Ksall	weight corresponding to f_Ksall, median normalised
f_[]	(convolved) aperture flux in filter [] within a 1.2″ diameter circular aperture, corrected to total (fap_[] = f_[] / Ks_ratio)
e_[]	uncertainty on f_[] (also scaled with Ks_ratio)
w_[]	weight corresponding to f_[], median normalised
wmin_optical	minimum w_[] of groundbased optical filters
wmin_hst_optical	minimum w_[] of HST optical filters
wmin_fs	minimum w_[] of `FourStar` filters
wmin_jhk	minimum w_[] of broadband J, H & K filters
wmin_hst	minimum w_[] of HST near-IR filters
wmin_irac	minimum w_[] of Spitzer/IRAC filters
wmin_all	minimum w_[] of all filters
star	this flag is set to 1 if the source is likely to be a star, to 0 otherwise
nearstar	this flag is set to 1 if the source is located near a bright star
use[c]	sources that pass the following criteria are set to 1: - star = 0 - nearstar = 0 - SNR ≥ 5 - wmin_fs > 0.1 - wmin_optical > 0 - not a bad EAZY fit: χ^2 (reduced) ≤ 1000 - a finite FAST stellar mass estimate
snr	signal-to-noise (=f_Ksall / e_Ksall)

[a] Note that these K_s-band fluxes are derived from the superdeep combined Ks-band images. Within the catalogs only f_Ks corresponds to `FourStar`/K_s.

[c] A standard selection of galaxies can be obtained by selecting sources with use = 1.

SNR estimates of each source. Individual sources are indicated by their ID, starting at ID = 1. A description of the columns is given in Tables 2.5 and 2.6.

The CDFS catalog contains 30,911 sources, the COSMOS catalog 20,786 and the UDS catalog contains 22,093 sources. Magnitudes for each source can be obtained by applying a zeropoint of 25 in the AB system (corresponding to a flux density of $3.631 \times 10^{-30} erg\ s^{-1}\ Hz^{-1}\ cm^{-2}$ or $0.3631 \mu Jy$). E.g. the stacked K_s-band total magnitude is $25 - 2.5 \times \log_{10}$f_Ksall.

All fluxes in the catalogs are scaled to total. They can be converted back to aperture flux (1.2″ diameter) by dividing by Ks_ratio for each source. The exceptions are fap_Ksall. fauto_Ksall and fapcirc0D_Ksall. The first is the actual (convolved) K_s-band aperture flux, and can only be converted in the other direction, towards total. The second is the auto aperture flux from SE, and we need only to apply the aperture correction, apcorr, described in Section 2.4.2, to obtain f_Ksall. The last is an alternative to fauto_Ksall, and is measured in a fixed circular aperture with diameter D, instead of the flexible elliptical Kron-like aperture from SE (using apertures of $D = 0.7''$ in CDFS and $D = 0.9''$ in COSMOS and UDS). From fapcirc0D_Ksall we can obtain the total fcirc0D_Ksall by multiplying with apcorr0D.

Each flux measurement of each source in each filter has been assigned a weight, reflecting the depth in the images at the source locations. The weigths are normalized to the median of the corresponding weight images. In the catalogs we also indicated the minimum weight for sets of filters. For example, the lowest weight of the FourStar filters is indicated by wmin_fs. If this value is greater than 0, it means a positive weight in all FourStar images.

In addition to photometric catalogs, we provide the EAZY (Section 2.6) and FAST (Section 2.7) output files, containing the photometric redshifts and stellar population properties.

2.4.9 Quality verification

Flux comparisons

Here we test whether the total fluxes derived above are reliable, by (1) testing against using a different method to measure total flux, and (2) by comparing magnitudes directly. For the latter we used the 3D-HST data set, as both surveys use many of the same images. Largely the same methods to derive photometry were used for 3D-HST. In general we find excellent correspondence between the two surveys. We show diagnostic plots in Appendix 2.B.

We tested our method of extracting total flux through SE by comparing to total flux derived with GALFIT (Peng et al. 2010a), a program which fits two-dimensional model light profiles to galaxy imaging. The fitting process bene-

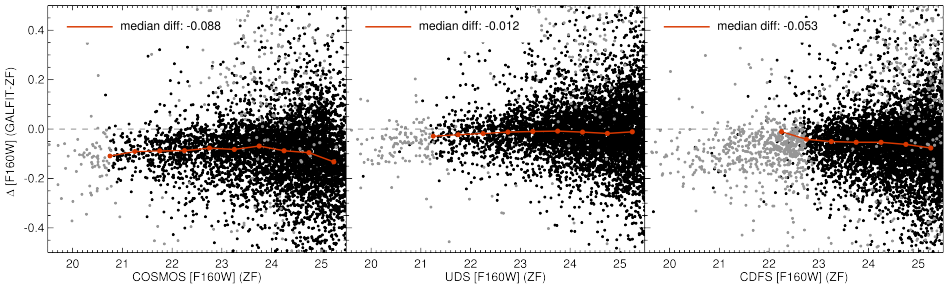

Figure 2.9: The difference between ZFOURGE and GALFIT F160W magnitudes plotted as a function of ZFOURGE magnitude, for galaxies with use=1 and SEflag< 2 (excluding blended or contaminated sources). We show sources with GALFIT flag = 1 (a suspicious fit) in gray and sources with GALFIT flag = 0 (a good fit) with black datapoints. Bad fits (GALFIT flag > 1) were ignored. The median magnitude difference for galaxies with GALFIT flag= 0 is shown by the red solid line and filled bulletpoints in bins of 0.5 mag. We also indicate the median offset in the legend. We find slightly brighter magnitudes with GALFIT, of at maximum < 0.088 magnitude.

fits from high resolution imaging, so we make use of the HST/WFC3/F160W size catalogs from van der Wel et al. (2014), based on the source catalog of 3D-HST, which contains parameters derived with GALIFT. As the F160W and K_s-band filters lie very closely together in wavelength space, we assume that the correction to total in our catalogs, which is based on the ratio between K_s-band aperture and total flux, produces a true approximation of total F160W-band flux.

The comparison with GALFIT magnitudes is shown in Figure 2.9. We use the goodness of fit flag included in the size catalogs to select source with a good (GALFIT flag = 0) or suspicous fit (GALFIT flag = 1), but not sources with bad fits (GALFIT flag > 1). We find a median offset between ZFOURGE total F160W magnitude and GALFIT magnitude of −0.012, −0.088 and −0.053 magnitude, for UDS, COSMOS and CDFS, respectively. Skelton et al. (2014) show the same comparison, with similar trends with magnitude, and find magnitude offsets for the three fields of 0.00, −0.04 and −0.03. We also compared our F160W magnitudes directly to 3DHST in Appendix 2.B, finding offsets of −0.002, 0.025 and 0.025. Taking these into account, our result corresponds well with that in Skelton et al. (2014).

The small offsets that we find between GALFIT magnitude and magnitude derived with SE, might be a result of GALFIT measuring flux beyond the auto aperture from SE, whereas we had to correct for it manually. We note that we have applied a pointsource correction to account for missing light, whereas in

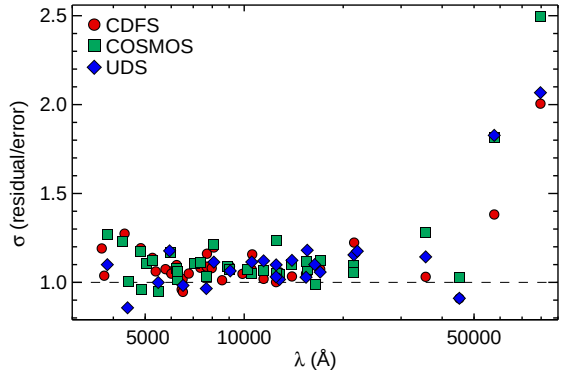

Figure 2.10: NMAD scatter in the error-normalized flux residuals as a function of wavelength, for the three ZFOURGE fields. σ is close to unity for most filters, indicating the photometric errors are accurate.

reality galaxies have different morphogies.

Flux uncertainty verification

Here we test the accurateness of the flux uncertainties derived in Section 2.4.4. We used the outcome of the SED fitting described below in Section 2.6. The residual between the best-fit template flux and the observed flux in a filter should reflect the photometric errors in the catalogs. If these are accurate, then normalizing the distribution of the residuals by the photometric error, should result in a gaussian with a width of unity. We derived the normalized median absolute deviaton (NMAD) of the distribution of the error-normalized residuals, and show the scatter (σ), for each filter in the catalog, as a function of wavelength in Figure 2.10. Overall these look very good, with σ_{NMAD} close to unity. Only for two of the IRAC bands, at 5.8 μm and 8.0 μm, the flux uncertainties could be underestimated. Some of the scatter there may result from template mismatch, instead of from the photometric errors.

Close pair contamination

It is naturally expected that some sources lie in close angular proximity of each other, and may contaminate the aperture flux of their close neighbor. This may lead to systematic errors on the aperture photometry of a source, especially if the neighbor is much brighter. The aperture diameter that we used above is 1.2 $''$. We inspected the catalogs for pairs of galaxies that lie closer than 1.2 $''$ distance away from each other. We only looked at sources that

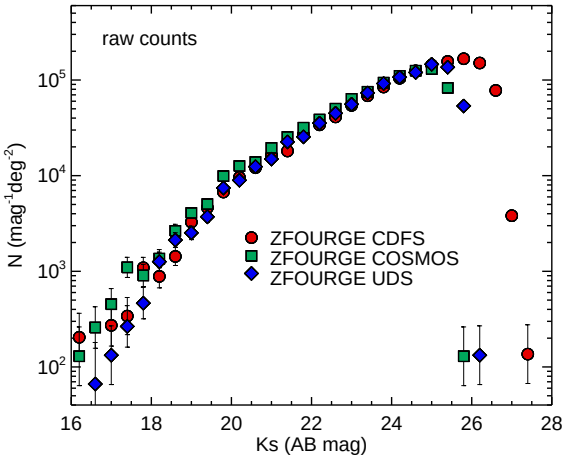

Figure 2.11: K_s-band total magnitude number counts of sources with `use=1`. We observe a turn-over in the histograms around 25 – 26 magnitude, with less detections for fainter sources.

are not already classified as stars or as being located in the neighbourhood of a bright star, as we already accounted for these sources that their flux estimate may be affected. The percentage of sources with a neighbour at $< 1.2\,''$ distance is 4.1 % in COSMOS, 3.8 % in CDFS and 4.4 % in UDS. If only the fainter part of a projected galaxy pair is affected, we estimate that ~ 2 % of the sources in each field may suffer flux effects from nearby sources.

2.5 Completeness

We counted the number of sources with `use=1` per Ks-band total magnitude bin in each catalog. This result, taking into account the effective area corresponding to the `use` flag, is shown in Figure 2.11. For the different fields, the histograms turn over at 25.5 – 26 magnitude, indicating it becomes more difficult to detect fainter sources, and we detect more scatter in magnitude.

To test how well sources are recovered from the images, we perform completeness tests, using the super-deep K_s-band detection images. We drop 10,000 mock sources, obtained from median stacking low SNR ($9 < SNR_{K_s} < 11$) sources with `use=1`, in the detection images. The stacks were scaled to a magnitude range of $18 < mag(\mathrm{AB}) < 27.5$. We used a powerlaw distribution of magnitudes, matching the slope of the number counts in Figure 2.11 between $K_s = 21$ AB and $K_s = 25$ AB. The distribution follows $dlogN/K_s = 0.24$, i.e. a fac-

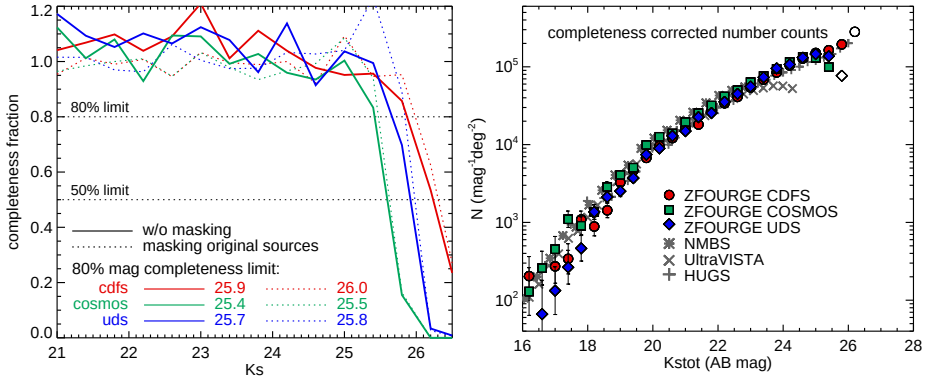

Figure 2.12: Left: We test what fraction of sources is detected by a simulation, in which mock sources are inserted into the detection images. For a realistic approach, in which sources are allowed to overlap, we recover 80% down to very deep magnitudes of $K_s = 25.4 - 25.9$. Right: Completeness corrected number counts. We correct the observed counts in Figure 2.11 using the completeness estimates in each field. Here we omit sources at magnitudes fainter than the 50% completeness limits. Sources with $< 80\%$ completness are shown with open symbols. For comparison, the K_s-band number counts of similar galaxy surveys are indicated in grayscale.

tor 1.7 more sources per unit magnitude, with N the number of sources and K_s the total K_s-band magnitude, in agreement with previously deteremined values (e.g. Fontana et al. 2014). We ran SE using the same input parameters used to generate the catalogs. We measured the observed magnitude of the input sources that were retrieved with SE. We then compared these with the input source distribution to calculate the correction as a function of observed magnitude, accounting for both completeness and scatter between input and output magnitude.

We performed the simulation in two ways. First by simply dropping mock sources randomly in the images, only excluding a few small areas around a few very bright stars. To prevent artificial crowding of simulated sources, we only dropped in 500 sources per run, and repeated the simulations a large number of times.

Next we investigated what fraction of incompleteness is due to crowding, where bright sources prevent the detection of fainter sources nearby. We masked all detected sources, using the segmentation map from SE and constrained the location of the simulated sources, such that they do not overlap. In this way, we purely tested if sources can be detected above the noise level in the images.

Table 2.7: Completeness test results

	with masking		w/o masking	
	80%	50%	80%	50%
CDFS	26.0	26.3	25.9	26.2
COSMOS	25.5	25.6	25.4	25.6
UDS	25.8	26.0	25.7	25.9

We show the results of the two tests in the left panel of Figure 2.12. Even if only stars are masked and sources are allowed to overlap (solid lines) we recover at least 80% down to very deep K_s-band magnitudes of 25.4 – 25.9 and 50% down to 25.6 – 26.2. These values correspond well with the turnover in K_s-band number counts in Figure 2.11 and the stacked K_s-band image depths (Section 2.2.3). The 50% and 80% completeness limits of both tests are tabulated in Table 2.7. The slight elevation with a higher than 100% completeness fraction at magnitudes < 24.5 for the non-masking case is due to confusion with bright sources.

We correct the number counts from Figure 2.11 using the completeness estimates as function of observed magnitude from the more conservative test (obtained w/o masking, i.e. the solid curves) in each field and show these in the right panel of Figure 2.12. We also include similar results from the NMBS (Whitaker et al. 2011), UltraVISTA (Muzzin et al. 2013a) and HUGS (Fontana et al. 2014) surveys. NMBS and UltraVISTA have shallower depths, but much larger areas than ZFOURGE. Our number counts agree with these earlier results from the literature. They also show that ZFOURGE is one of the most sensitive surveys to date, comparable to HUGS and 1 – 2 magnitudes deeper in K_s than earlier groundbased surveys. We note that, similar to NMBS, we find an excess of sources at brighter K_s-band magnitudes in COSMOS.

2.6 Photometric redshifts

2.6.1 Template fitting

Photometric redshifts were derived with EAZY (Brammer et al. 2008), by fitting linear combinations of nine spectral templates to the observed SEDs. Of these, seven are the default templates described by Brammer et al. (2008), five of which are from a library of PÉGASE stellar population synthesis models (Fioc & Rocca-Volmerange 1999), one represents a young and dusty galaxy and another is that of an old, red galaxy (see also Whitaker et al. 2011). The final two templates represent an old and dusty galaxy and a strong emission

line galaxy (Erb et al. 2010). The code has the option to include a template error function, which we use, to account for systematic wavelength-dependent uncertainties in the templates. We also make use of a luminosity prior, based on the apparent magnitude calculated from the total K_s-band flux.

Offsets in the zeropoints may systematically affect the measured flux and therefore also the derived photometric redshifts. We correct for zeropoint offsets, by iteratively fitting EAZY templates to the full optical-near-IR observed SEDs. This procedure is described in detail by Whitaker et al. (2011) and Skelton et al. (2014). Similar to Skelton et al. (2014), we use all sources in the fits, including those without a spectroscopic redshift available. We also use a two step process in which we first only vary the zeropoints of the HST-bands and then, keeping these fixed, we vary the zeropoints of the groundbased and Spitzer/IRAC data.

During this iterative fitting procedure, both the zeropoints and the templates were modified. These are separable corrections, as the templates are modified after shifting both the data and the best-fit SEDs to the rest-frame. Due to the wide range of galaxy redshifts and large number of filters in the catalogs, each part of the spectrum is sampled by a number of photometric bands. In small bins of rest-frame wavelength, we determined systematic offsets between the data and the templates and updated the templates. This allows the templates to reflect subtle features not initially included, such as the dust-absorption feature at 2175Å. After adjusting the templates, zeropoint corrections are calculated in the observed frame. The process is repeated until zeropoint corrections in all bands except U or the IRAC bands become less than 1% and this typically happens after three or four iterations.

The zeropoint offsets are listed in Tables 2.2, 2.3 and 2.4. The zeropoints in these tables are the effective zeropoints, with galactic extinction and the zeropoint offsets incorporated. The offsets are typically of the order of 0.05 magnitude. The largest offsets occur for the COSMOS and UDS U-bands, which are known to have uncertain zeropoints (Erben et al. 2009; Whitaker et al. 2011; Skelton et al. 2014). We note that template and zeropoint errors are hardest to separate from each other for the U- and IRAC 8μm bands, as these lie at the blue and red ends of the spectra, without bracketing filters.

The residuals between the best-fit templates and observed SEDs are excellent tracers of spatial variations in the zeropoint. We found small variations for all images ($< 5\%$). In particular, we were able to pinpoint small offsets between the different quadrants of the `FourStar` images in UDS. To alleviate the spatial effect, our final derivation for every filter includes two runs of the fitting process. After the first run we remove a 2 dimensional polynomial fit to the spatial residuals. This is directly incorporated into the catalogs, i.e. we

Table 2.8: Explanation of the EAZY photometric redshift catalog header

id	ID number
z_spec	spectroscopic redshift (if no redshift available, z_spec is set to -1)
z_a	photometric redshift derived without a K luminosity prior
z_m1	weighted redshift derived without a K luminosity prior
chi_a	minimum χ^2 derived without a K luminosity prior
z_p	best-fit redshift after applying the prior
chi_p	minimum χ^2 after applying the prior
z_m2	weighted redshift after applying the prior
odds	parameter indicating presence of second χ^2 minimum (1 if no minimum)
l68,u68	1 sigma confidence interval
l95,u95	2 sigma confidence interval
l99,u99	3 sigma confidence interval
nfilt	number of filters used in the fit
q_z	quality parameter
z_peak	default derived photometric redshift
peak_prob	peak probability
z_mc	randomly drawn redshift value from redshift probability distribution

apply a correction to all sources as a function of their x- and y-coordinates in the images and using the corresponding 2 dimensional offsets in each filter. Finally the fitting process is repeated in the way described above to obtain the final zeropoint offsets. In Figure 2.32 in Appendix 2.C we show the residual maps with the spatial variations.

We note that the spatial variations in zeropoint of the VLT/VIMOS/R–band image are larger than in the other images and could not be described by a polynomial function. This image is very deep, so we do not wish to discard it. We therefore impose a minimum error on the flux of 5%.

We use the output parameter `z_peak` from EAZY as indicator of the photometric redshift. `z_peak` is estimated by marginalizing over the redshift probability distribution function, $p(z)$. If $p(z)$ has more than one peak, `z_peak` only marginalizes over the peak with the largest integrated probability.

Using EAZY, we also derived various rest-frame colors, for example in the Johnson/U and V-bands (Maíz Apellániz 2006), in the J-filter from the Two Micron All Sky Survey, and at 2800Å (using a tophat shaped transmission curve). Rest-frame colors are calculated by integrating the redshifted rest-frame filter bandpasses of the best-fit template for each individual source. The process is described in more detail by Brammer et al. (2011), see also Whitaker et al. (2011). These colors are often used to make selections of quiescent versus star-forming galaxies. We discuss $U-V$ versus $V-J$ diagrams in Section 2.8. Rest-frame 2800Å luminosity can be used as a proxy of the UV

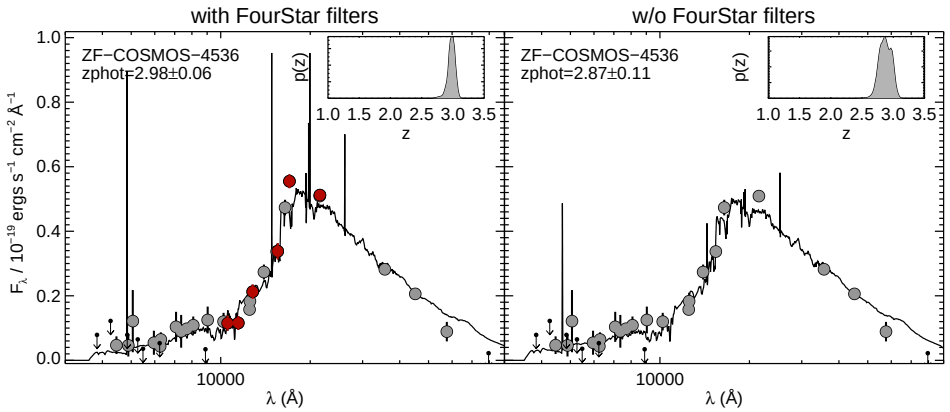

Figure 2.13: An example galaxy at $z = 2.1$ with a large Balmer/4000Å break, traced by the FourStar filters (indicated in red). Gray datapoints represent flux in ancillary filters, with downward pointing arrows representing upper limits. In the left panel we show the best-fit SED template derived using the FourStar near-IR medium-bandwidth filters. In the right panel we show the best fit, without the FourStar bands. The insets show the redshift probability functions corresponding to the fits. Including the FourStar filters leads to a factor two better constraint on the photometric redshift.

luminosity of a galaxy, which in turn can be used to derive the unobscured SFR (see Section 2.7).

We provide the full EAZY photometric redshift catalogs. See Table 2.8 for an explanation of the catalog header.

2.6.2 Photometric redshift uncertainties determined by EAZY

As a result of the use of near-IR medium-bandwidth filters, spectral features such as the Balmer/4000Å break are better sampled for galaxies at $1.5 < z < 3.5$. In Figure 2.13 we illustrate the ability of the FourStar medium-bandwidth filters to constrain galaxy SEDs and redshift probability distibutions. We can determine a photometric redshift error due to the fitting process, using the $16th - 84th$ percentiles from $p(z)$. For better constrained redshifts, $p(z)$ will be narrower and the error on z_{phot} will be smaller. We show the SED of a galaxy at a redshift of $z = 2.98 \pm 0.06$, with the uncertainty derived from the 68th percentile of the $p(z)$. This galaxy has a strong 4000Å/Balmer feature, well sampled by the FourStar medium-bandwidth filters. The photometric redshift derived without the use of medium-bandwidth filters in the near-IR, i.e. using only the available broadband groundbased

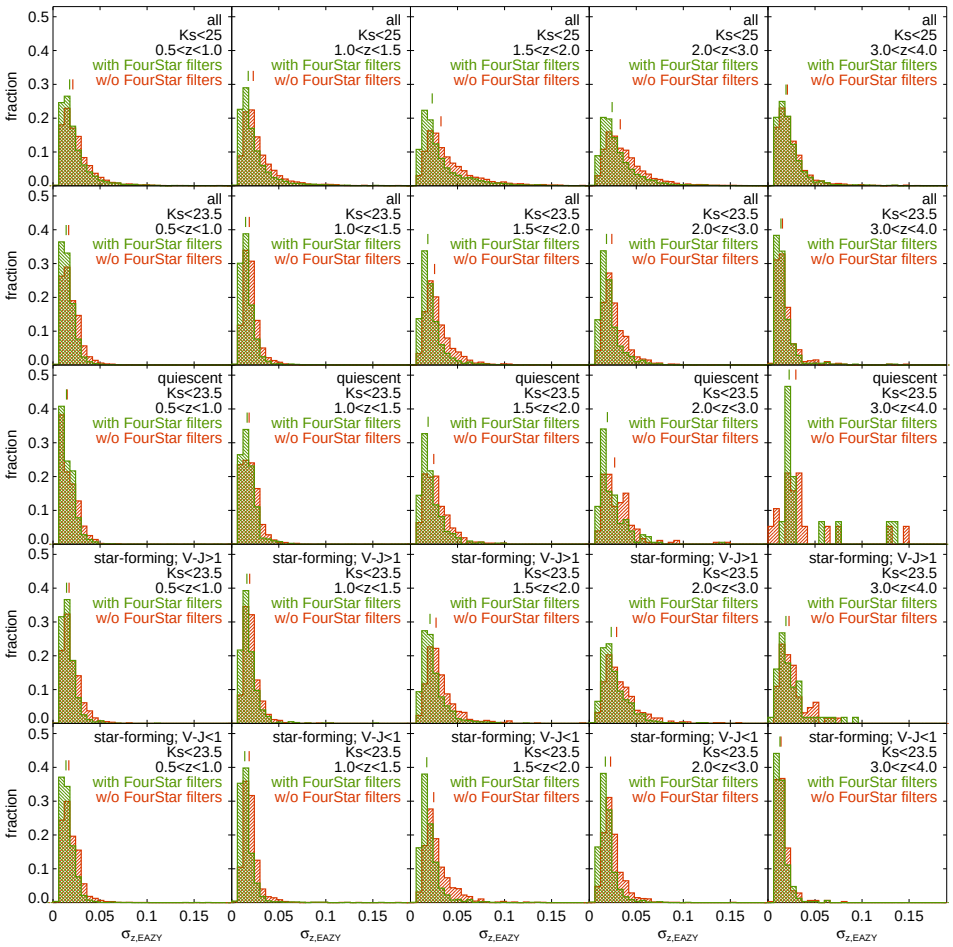

Figure 2.14: Redshift error, $\sigma_{z,EAZY}$ histograms in redshift bins (from left to right), normalized to the total number of sources in each bin. In the first and second row we inspect general magnitude-limited samples, with $K_s < 25$ AB and $K_s < 23.5$. In the third to last rows of panels we show, respectively, the error histograms of quiescent galaxies, red star-forming galaxies with $V - J \geq 1$, and blue star-forming galaxies with $V - J < 1$. The median $\sigma_{z,EAZY}$ is indicated just above the histograms in each panel, using the respective colors (green or red) of the EAZY fits with and without near-IR mediumbands. The photometric uncertaintes are systematically smaller if we include the FourStar near-IR medium-bandwidth filters when fitting SED templates (green histograms).

Y, J, H and K/K_s or spacebased F125W, F140W and F160W filters, is $z = 2.87 \pm 0.11$. The galaxy has a broader redshift probability distribution, $p(z)$, without the FourStar filters, i.e. the redshift is less tightly constrained in the fit.

In Figure 2.14 we show histograms of the errors, $\sigma_{z,EAZY} = p68(z)/(1+z)$, with $p68(z)$ the error from the 68th percentile of $p(z)$, in bins between $z = 0.5$ and $z = 4$. This is the redshift region where we expect the impact of the medium-bandwidth filters to be greatest. We show the histograms for a magnitude-limited sample, with $K_s < 25$ in the top row, and with $K_s < 23.5$ in the second row. We also show the histograms of different galaxy types, by splitting up the sample into quiescent and star-forming galaxies, using the UVJ technique (e.g. Whitaker et al. 2011). The star-forming galaxies were additionally split into blue and red by their rest-frame $U-V$ and $V-J$ colors, which we explain further in Section 2.8.

The histograms indicate that over a large range in redshift, the errors on the photometric redshifts are smaller if we include the FourStar filters. This holds for all galaxy types. The effect is especially clear around $z = 2$, and is noticable for higher redshifts as well. For example, at $1.5 < z < 2$, the median uncertainty is 40% higher without the FourStar filters, with $\sigma_{z,EAZY} = 0.036$ compared to $\sigma_{z,EAZY} = 0.025$. Whitaker et al. (2011) find a similar trend with redshift, for the mediumbands of NMBS. The peak of the histograms shifts towards higher $\sigma_{z,EAZY}$ with increasing redshift, up to $z = 3$, except for blue star-forming galaxies (with blue $U-V$ and $V-J$ colors, see Section 2.8), for which $\sigma_{z,EAZY}$ actually improves. A notable spectral feature for these glaxies is the Lyman Break at rest-frame 912Å, which is moving through the optical medium-bandwidth filters at this redshift.

2.6.3 Comparison with spectroscopic redshifts

A common comparison in the literature is to compare the photometric redshifts with spectroscopic redshifts. In Figures 2.15 and 2.16 we do this, using the compilation of publicly available spectroscopic redshifts in these fields provided by Skelton et al. (2014), with a matching radius of $1''$. We added to this first release of the MOSDEF survey (Kriek et al. 2015). The overall correspondence is excellent, as indicated by the scatter in the difference between photometric and spectroscopic redshifts. We quantify the errors in the spectroscopic redshifts, $\sigma_{z,spec}$, using the normalized median absolute deviation (NMAD) of $\Delta z/(1+z)$, i.e. 1.48× the median absolute deviation of $|z_{phot} - z_{spec}|/(1+z_{spec})$. In CDFS $\sigma_{z,spec} = 0.009$, in COSMOS $\sigma_{z,spec} = 0.008$ and in UDS $\sigma_{z,spec} = 0.013$. Only a small percentage are outliers, with $\Delta z/(1+z) > 0.15$. In CDFS 2.9% are

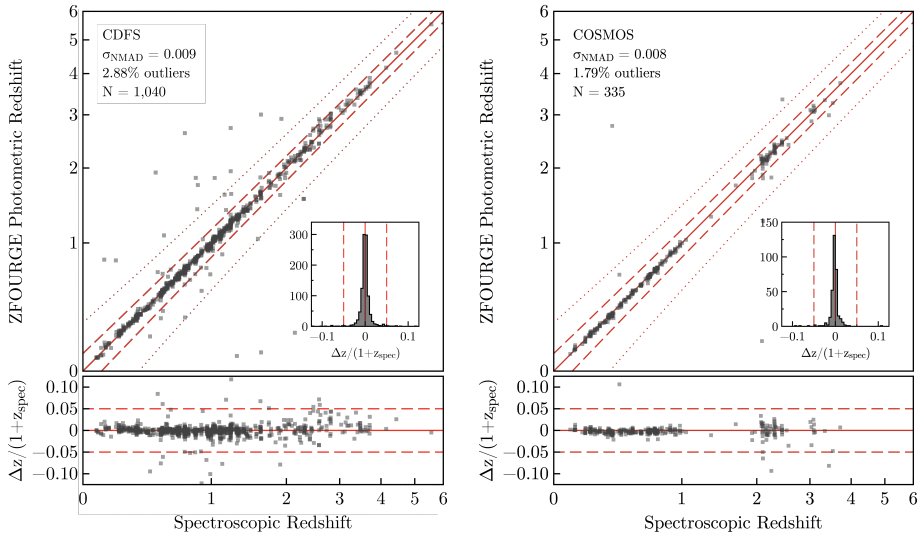

Figure 2.15: Top: Photometric redshifts from ZFOURGE versus spectroscopic redshifts in CDFS (left) and COSMOS (right). The NMAD scatter, the fraction of objects with $\Delta z/(1+z_{spec}) > 0.15$, and the number of galaxies with matches in both catalogs are shown in the upper left of the plot, while the histograms of $\Delta z/(1+z_{zspec})$ are shown as an inset in the bottom right of the plot. Bottom: the residual between the photometric and spectroscopic redshifts, divided by $1+z_{spec}$. The red solid, dashed and dotted lines indicate, respectively, $\Delta z/(1+z_{\rm spec}) = 1 \pm 0,\ \pm 0.05, \text{and}\ \pm 0.15$.

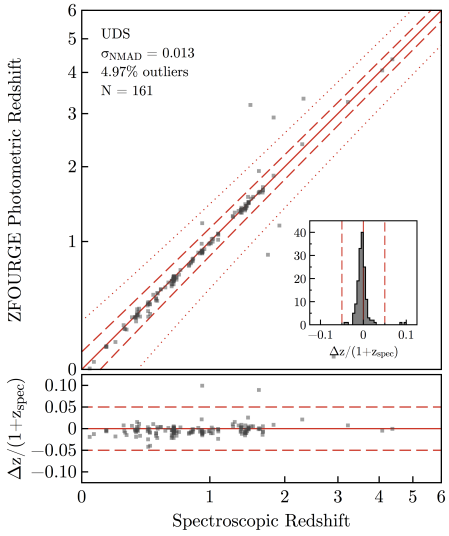

Figure 2.16: Photometric versus spectroscopic redshifts for UDS (see caption of Figure 2.15).

outliers, in COSMOS 1.8% and in UDS 4.8%. At $z > 1.5$ we find $\sigma_{z,spec} = 0.093$, $\sigma_{z,spec} = 0.031$ and $\sigma_{z,spec} = 0.044$ in CDFS, COSMOS and UDS, respectively.

We have also compared with the unpublished redshifts of the ZFIRE survey (Nanayakkara in prep, 2015), and find an NMAD of $\sigma_{z,spec} = \sim 0.02$. Full results will be shown in Nanayakkara et al, in prep.

2.6.4 Redshift pair analysis

The drawback of comparing to spectroscopic samples is that these are usually biased towards bright ($K_s < 22$) star-forming galaxies, or unusual sources, such as AGN. Therefore these comparisons are not representative of the full photometric catalog and do not allow a careful study of how photometric redshift errors depend on galaxy properties. Here we present an alternative statistical analysis by looking at galaxy pairs. This method was first described and validated by Quadri & Williams (2010). It does not rely on spectroscopic information and can be applied to the full catalogs, including faint sources. Therefore this technique provides us with a more representative photometric redshift uncertainty than possible by comparing to spectroscopic redshifts.

Due to clustering, close pairs of galaxies on the sky have a significant probability of being physically associated, and of lying at the same redshift. Other

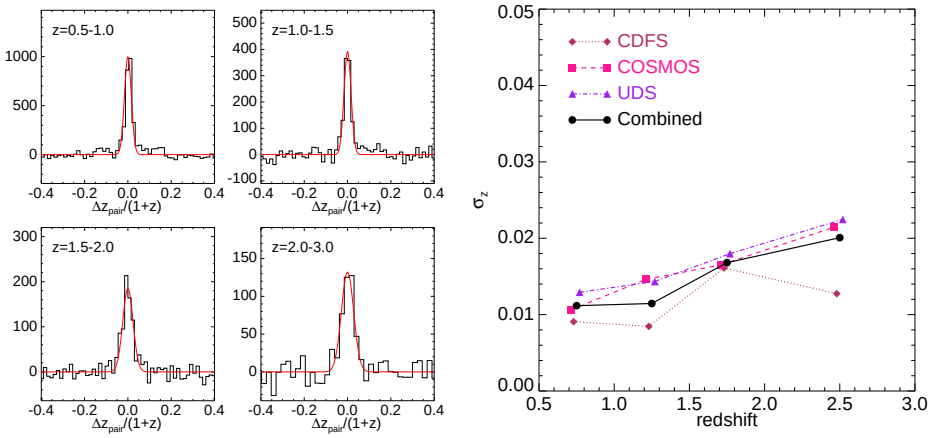

Figure 2.17: Left: distribution of $\Delta z/(1+z)$ for galaxies with $K_s < 23.5$. We fit gaussians to each histogram (red lines), from which we derive σ_z, the average uncertainty for individual galaxies, which is the standard deviation of the pair distribution divided by $\sqrt{2}(1+z_{mean})$. Right: σ_z as determined in the left panel as a function of redshift (black solid line). We also show the results for the individual fields, as indicated in the legend. σ_z increases with redshift, from $\sigma_z = 0.01$ to 0.02.

galaxy pairs will actually be chance projections along the line of sight, but this contamination by random pairs can be accounted for statistically, by randomizing the galaxy positions and repeating the analysis. Each true galaxy pair will give an independent estimate of the true redshift, and we can take the mean of the two values as our best estimate of the true redshift. The distribution of $\Delta z_{pairs}/(1+z_{mean})$ of the pairs of galaxies can then be used to estimate the average photometric redshift uncertainties. It is a narrow distribution for robustly derived redshifts, or broader due to more scatter in $\Delta z_{pairs}/(1+z_{mean})$ if the redshifts are very uncertain.

For illustration, we show the distributions of $\Delta z_{pairs}/(1+z_{mean})$ in the left panel of Figure 2.17, for pairs of galaxies with `use=1` and total K_s-band magnitude < 23.5, in four redshift bins. The pairs have angular separations between 2.5 $''$ and 15 $''$. To each distribution we fit a gaussian and determined the standard deviation. As this is the standard deviation for the redshift differences, we divide by $\sqrt{2}$ to obtain the average redshift uncertainty for individual galaxies, $\sigma_{z,pairs}$, for a particular redshift bin, i.e. $\sigma_{z,pairs}$ is obtained from $\Delta z_{pairs}/(\sqrt{2}(1+z_{mean}))$. In the right panel we show σ_z as a function of redshift. σ_z increases with redshift, but in general is excellent: varying from 1% to 2% going from $z = 0.5$ to $z = 2.5$. Calculating σ_z requires fairly large samples. This partly explains the scatter between results on individual ZFOURGE fields.

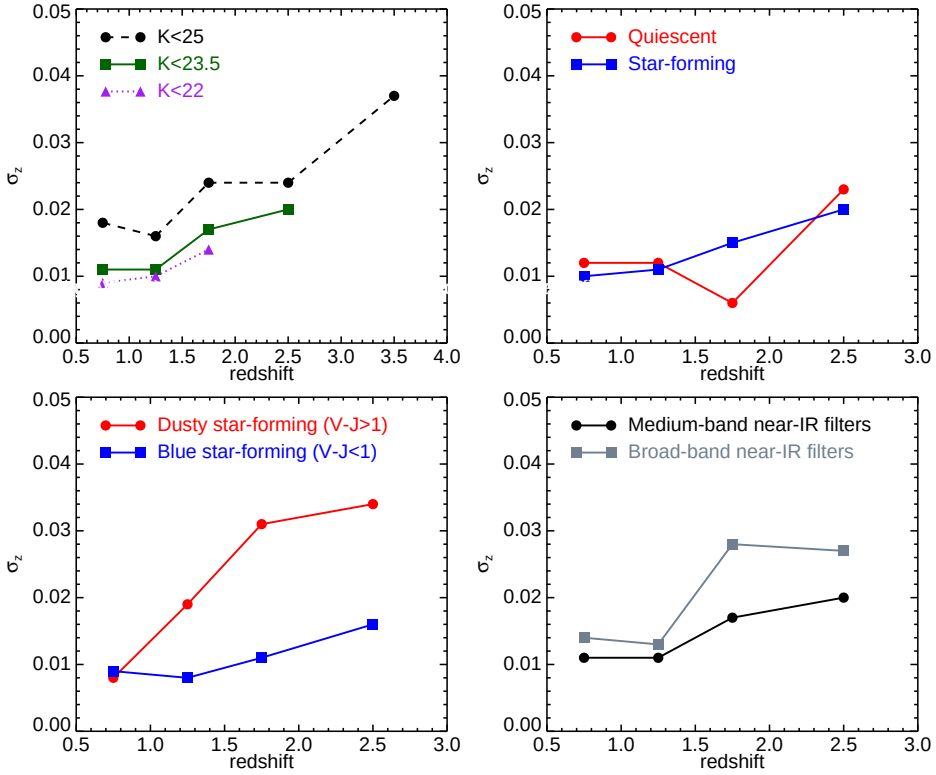

Figure 2.18: σ_z versus redshift, investigating trends with magnitude limits (top-left) or SED types (second and third panels), and investigating the effect of using the FourStar filters (top-right). σ_z tends to be smaller for brighter galaxies are considered, and for blue star-forming galaxies. σ_z is clearly smaller if the near-IR mediumband filters are used (compared to near-IR broadband), especially at $1.5 < z < 2.5$.

Other reasons for differences between the fields are different image filter sets and image depths.

Possible systematics on σ_z may include artificially narrow spikes in the $\Delta z/(1+z)$ distributions, underestimating the true redshift uncertainty. For example, because of systematic photometric errors, many sources could be fit with similiar, but wrong, redhifts. This is explained in more detail by Quadri & Williams (2010). We tested this scenario by inspecting pairs with at least one spectroscopic redshift available. We derived similar results, indicating that σ_z for photometric pairs is not systematically affected.

We expect σ_z to be sensitive to various parameters, including the type of galaxy, the magnitude and redshift. As a first exploration we will here char-

acterize how our photometric redshift uncertainty depends on these parameters. In the first panel of Figure 2.18 we show σ_z versus z for three different magnitude-limited samples, with $K_s < 22$, $K_s < 23.5$ (as above) and $K_s < 25$. For the brightest galaxies, with $K_s < 22$, the uncertainty is very small, around 1% up to $z = 2$. However, the uncertainty increases by roughly a factor towards fainter magnitudes up to $K \sim 25$, which is near our completeness limit.

We have additonally investigated the dependence of σ_z on galaxy type, using the same UVJ selected samples of quiescent, red star-forming anda blue-starforming galaxies as in Section 2.6.2. The results are shown in the second and third panels of Figure 2.18. Interestingly, the photometric redshifts of star-forming galaxies and quiescent galaxies are equally well constrained at most redshifts. The exception occurs at intermediate redshift ($1.5 < z < 2$), where instead we find much smaller redshift uncertainties for quiescent galaxies. This is the redshift range where the Balmer/4000Å break is moving through the J_1, J_2 and J_3 medium-bandwidth filters. In contrast,Quadri & Williams (2010) used shallower broadband photometry - with fewer optical filters - and found that quiescent galaxies have significantly better photometric redshifts at all redshifts. This emphasizes that the characteristics of photometric redshifts are dataset-dependent.

Comparing blue and red (dusty) star forming galaxies, we find that red galaxies have a factor $2-3$ worse $\sigma_{z,pairs}$, than do blue galaxies. The redshifts of these galaxies are difficult to constrain, even with mediumband photometry, as they have relatively featureless SEDs, and a degeneracy between redshift and the color of the reddest template allowed in the EAZY set (e.g. Marchesini et al. 2010; Spitler et al. 2014). Here we have split the sample at $V - J = 1$, but the effect will be stronger for dustier galaxies at redder $V - J$. This is a significant problem for star-forming galaxies with high mass or high SFRs, which often tend to be quite dusty.

In the last panel of Figure 2.18 we compare σ_z for the case where we have not inluded the near-IR medium-band `FourStar` filters in the EAZY fits. For the entire range considered here, the photometric redshifts are better derived if we do use the `FourStar` mediumbands. The effect is strongest at $z = 1.5-2.5$ where σ_z using the `FourStar` mediumbands is 50% smaller compared to σ_z with the `FourStar` filters removed. This confirms the efficacy of the mediumbands at intermediate to high redshift.

The pairs analysis also provides an interesting way to verify whether the redshift uncertainties that come from the EAZY template fits are reasonable. In the left panel of Figure 2.19 we compare $\sigma_{z,pairs}$ to $\sigma_{z,EAZY}$, and find that they provide heartening agreement. This figure also shows that $\sigma_{z,spec}$, the uncertainty estimated from comparing the photometric to the spectroscopic

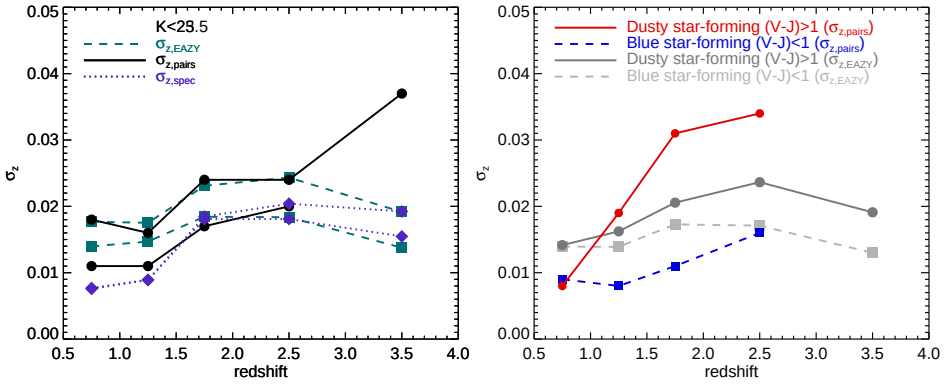

Figure 2.19: Left: comparing the three different redshift quality tests, $\sigma_{z,spec}$ from testing against spectroscopic samples, $\sigma_{z,pairs}$ from the redshift pair test, and the photometric redshift uncertainty from the $p(z)$ from EAZY, $\sigma_{z,EAZY}$. These correspond well at all redshifts. Right: comparing $\sigma_{z,pairs}$ with $\sigma_{z,EAZY}$. $\sigma_{z,EAZY}$ tends to be underestimated for red star-forming galaxies.

redshifts, provides a good estimate of the true uncertainties for the $K < 23.5$ sample.

Although the redshift uncertainties estimated by EAZY appear quite reliable for the general population of galaxies, we find that they are not trustworthy for dusty star-forming galaxies. In the right panel of Figure 2.19 we compare $\sigma_{z,pairs}$ to $\sigma_{z,EAZY}$ from the EAZY fits. The pair redshifts of blue star-forming galaxies are better than we expect from the EAZY $p(z)$. However, for red and dusty star-forming galaxies the pair redshifts are 50% worse than the EAZY $p(z)$. This effect increases with redshift and towards fainter magnitudes (not shown here). It indicates that with current methods and state-of-the-art surveys, the degeneracy between rest-frame color and redshift for dusty galaxies cannot yet be accurately resolved, and we caution that photometric redshift uncertainties for faint dusty galaxies at $z > 1.5$ are generally underestimated.

2.6.5 Redshift distributions

By improving the accuracy of the photometric redshifts, we can derive improved stellar masses and start identifying large scale structure. In Figure 2.20 we plot the K_s-band magnitudes as function of `z_peak` (or Z_{spec} where available). The histograms have independent spikes for each field, corresponding to known overdensities, e.g. at $z < 1$ in COSMOS (e.g. Kovač et al.

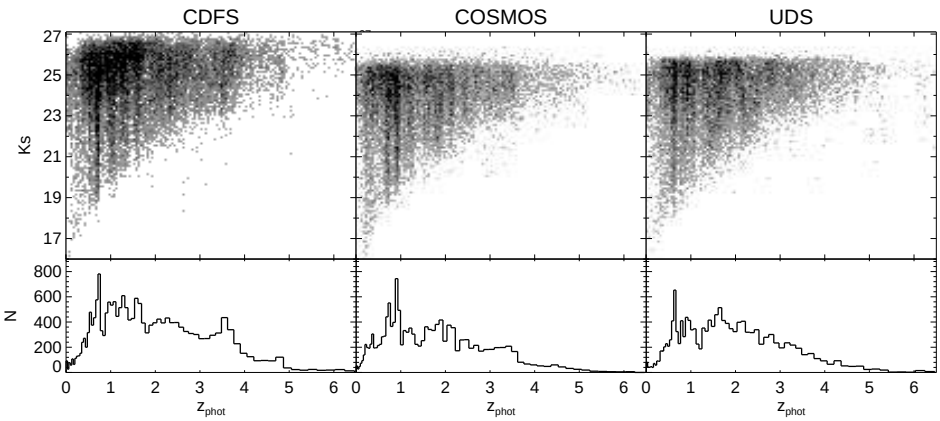

Figure 2.20: Top panels: distribution of K_s-band magnitudes as function of redshift. The grayscale indicates the density in each point, with darker colors for higher densities. Bottom panels: photometric redshift (`z_peak`) distribution. Higher density peaks are clearly visible, for example the ZFOURGE identified cluster at $z = 2.095$ in COSMOS (Spitler et al. 2012; Yuan et al. 2014).

2010; Knobel et al. 2012). One overdensity was actually discovered with ZFOURGE, and is a cluster of galaxies in COSMOS at the spectroscopically confirmed reshift of $z = 2.095$ (Spitler et al. 2012; Yuan et al. 2014).

2.7 Stellar masses and star-formation rates

Stellar population properties (stellar mass, SFR, dust extinction, and age) were derived by fitting Bruzual & Charlot (2003) models with FAST (Kriek et al. 2009), assuming a Chabrier (2003) initial mass function, exponentially declining star formation histories with timescale τ, solar metallicity and a dust law as described in Calzetti et al. (2000). For each source the redshift is fixed to the photometric redshift (`z_peak`) derived with EAZY, or the spectroscopic redshift if known. We limit dust extinction to $0 \leq A_V \leq 4$, age to $7.5 \leq log_{10}(\text{age}) \leq 10.1$ Gyr and τ to $7 \leq \tau \leq 11$ Gyr.

We estimated 80% mass completeness limits, by selecting galaxies within the range $K_s = 24.0 - 24.7$ mag and scaling their fluxes to $K_s = 25.0$ mag. Then we determined the 80th-percentile mass rank in narrow redshift bins. Galaxies above this value are the most massive objects that could plausibly fall below the K_s-band selection limit. A smooth function to these values is shown in Figure 2.21, as function of redshift. At $z = 2$ we reach a completeness limit of $\sim 10^9\ M_\odot$, and at $z = 4$ we are complete above $\sim 10^{9.5} M_\odot$. Beyond $z = 4$ the

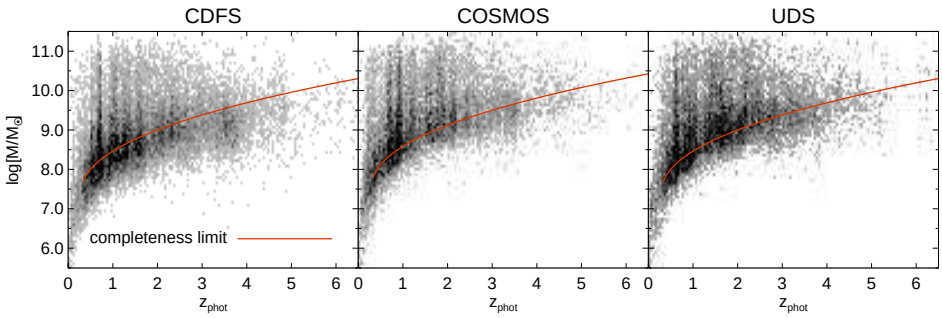

Figure 2.21: Similar to the top panels of Figure 2.20, but with stellar mass instead of K_s-band magnitude. The solid red line indicates the 80% mass completeness limit in each field.

Table 2.9: Explanation of the FAST stellar population catalog header

id	ID number
z	=z_peak (or z_spec if available)
ltau	log[tau/yr]
metal	metallicity (fixed to 0.020)
lage	log[age/yr]
Av	dust reddening
lmass	log[$M/M_\odot$]
lsfr	log[SFR/($M_\odot$/yr)]
lssfr	log[sSFR(/yr)]
la2t	log[age/τ]
chi2	minimum χ^2

completeness limit is extrapolated.

We provide the full FAST stellar population catalogs. See Table 2.9 for an explanation of the catalog header.

We note that SFRs, dust attenuations, ages and star formation histories of galaxies derived from SED fitting to UV, optical and near-IR photometry may be uncertain, especially if galaxies are highly dust-obscured. A different estimate of the SFRs can be obtained by inferring the total infrared luminosity ($L_{IR} \equiv L_{8-1000\mu m}$) of galaxies and combining this with the luminosity emitted in the UV (L_{UV} at rest-frame 2800Å). $L_{UV} + L_{IR}$ provides an estimate of the total bolometric luminosity, which can be converted to SFR under the assumption that the galaxy is continuously forming stars (Kennicutt 1998; Bell et al. 2005).

We use the conversion from Bell et al. (2005) to calculate SFRs from our

Table 2.10: Explanation of the SFR catalog header

id	ID number
z	phometric redshift (or spectroscopic redshift if available)
f_24	Spitzer/MIPS 24μm flux (mJy)
e_24	Spitzer/MIPS 24μm flux error (mJy)
f_100[a]	Herschel/PACS 100μm flux (mJy)
e_100[a]	Herschel/PACS 100μm flux error (mJy)
f_160[a]	Herschel/PACS 160μm flux (mJy)
e_160[a]	Herschel/PACS 160μm flux error (mJy)
L_IR	total integrated IR luminosity $L_\odot$
L_UV	total UV luminosity $L_\odot$
SFR	star formation rate (Equation 2.6

[a] Herschel/PACS data only available in CDFS.

data, scaled to a Chabrier (2003) IMF,

$$SFR\,[M_\odot/\mathrm{yr}] = 1.09 \times 10^{-10}\,(\mathrm{L_{IR}} + 2.2\mathrm{L_{UV}}) \tag{2.6}$$

where L_{IR} is the integrated $8-1000\mu$m IR luminosity. To derive L_{IR} we use our extracted $24-160\mu$m photometry (Section 2.4.5), to which we fit a model spectral template to calculate the total luminosity. The model template is the averaged template from Wuyts et al. (2008) (hereafter W08), generated by averaging the logarithm of the templates from the library of Dale & Helou (2002). The motivation of this approach was to introduce a simple conversion of flux to luminosity, first proposed by W08 and later validated by Wuyts et al. (2011).

Total IR luminosities are then obtained by integrating these fits between $8-1000\mu$m in the rest-frame. $L_{UV} = 1.5\nu L_{\nu,2800}$ is the estimated rest-frame 1216–3000Å UV luminosity, that we derived with EAZY. Both L_{IR} and L_{UV} are in units of $L_\odot$. This conversion assumes that the total IR luminosity reflects the amount of obscured UV light from young stellar populations. Thus by adding its contribution to that of the unobscured UV luminosity (L_{UV}) the net star-formation rate for galaxies can be measured.

In addition to the FAST catalogs, we provide catalogs with the net observed $L_{UV} + L_{IR}$ SFRs (see Table 2.10 for a description). We note that so far public Herschel/PACS data only exists or the CDFS field. For the other fields we fitted the W08 template only to the Spitzer/MIPS/24μm photometry.

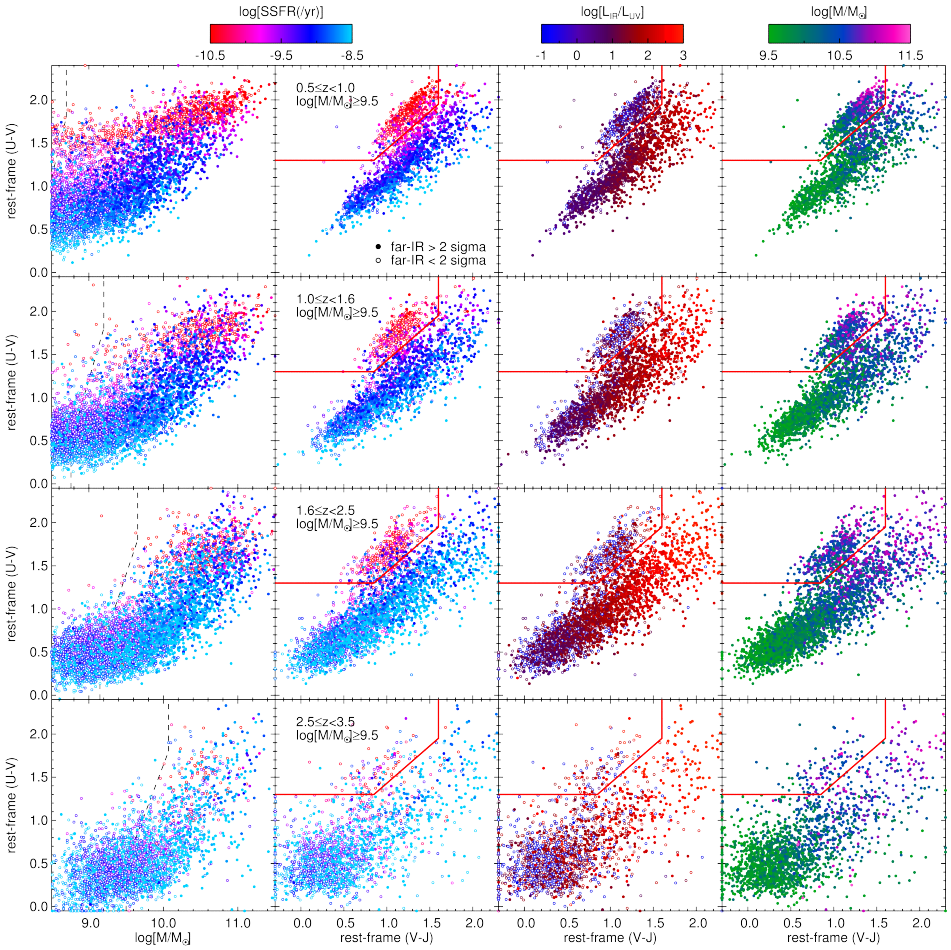

Figure 2.22: Rest-frame $U-V$ versus stellar mass (left columns) or versus $V-J$ (second, third and fourth columns) for a mass complete sample of galaxies with `use=1`, $SNR_{K_s} > 10$ and stellar mass $M > 10^{9.5} M_{\odot}$, in four redshift bins (top to bottom). The vertical dashed lines indicate our stellar mass completeness limit. The red solid line in the UVJ diagrams separates quiescent (top left) from star forming (bottom left to top right) galaxies. Galaxies that are undetected in the far-IR at $< 2\sigma$ are shown with open symbols. The sSFRs show a gradient towards blue $V-J$ and red $U-V$ colors, reflecting a gradient in age. Galaxies with the lowest sSFRs are located in the quiescent region of the diagram, with red $U-V$ and blue $V-J$ colors. Galaxies span a large range in $log[L_{IR}/L_{UV}]$, ranging from -1 for the bluest UVJ star-forming galaxies to 3 for the dustiest sources, and quiescent galaxies having low $log[L_{IR}/L_{UV}]$. A mass sequence is also visible, where massive galaxies tend to be redder.

2.8 First validation of the UVJ diagram at z = 3

Rest-frame $U-V$ is historically used to distinguish between blue, late-type galaxies with active star formation and red early-type galaxies with low star formation. At low redshift these quiescent "red and dead" galaxies dominate at high mass and form a tight "red sequence" versus luminosity or stellar mass (e.g. Tully et al. 1982; Baldry et al. 2004). In Figure 2.22 (left column) we present the rest-frame $U-V$ versus stellar mass relation of our mass complete sample, color coded by specific star formation rate (sSFR=SFR/M*, with M* stellar mass).

Clearly, at high redshift there are also large numbers of red massive galaxies with active star formation. These galaxies are red due to dust attenuation (see e.g. Brammer et al. 2009), so a single $U-V$ color cannot be used anymore to separate quiescent and star forming galaxies. We therefore use a two-color diagram, rest-frame $U-V$ versus rest-frame $V-J$ (hereafter UVJ), to efficiently separate quiescent from star forming galaxies (see e.g. Labbé et al. 2005; Wuyts et al. 2007; Williams et al. 2009). In Figure 2.22 (second column from the left), we show the UVJ diagram for the same galaxies. Red star forming galaxies are now well separated from red quiescent galaxies, and a clear red sequence in UVJ is present up to $z = 3.5$ (see also Whitaker et al. 2011).

Several trends with UVJ color can be seen in the data. First, specific star formation rates show a gradient in color space, such that redder $U-V$ and bluer $V-J$ colors correspond to lower sSFR. This can be interpreted as a stellar "age" gradient. We inspected the stellar ages derived from SED fitting for these galaxies (not shown) and indeed find that the gradient also corresponds to gradient in median stellar population age. Interestingly, this gradient can even been seen amongst SF galaxies alone, where the highest sSFR galaxies are found towards the lower right of the star-forming sequence.

Secondly, star-forming galaxies span a large range in colors due to dust attenuation. This can be seen in the right column of Figure 2.22, which shows the infrared excess IRX= $\log_{10} L_{IR}/L_{UV}$, the ratio between dust absorbed emission over unattenuated UV emission from star formation. The IRX ranges from IRX~ -1 for the bluest UVJ colors (colors typically found in dropout selected samples, (e.g. Bouwens et al. 2014) to IRX~ 3 for the dustiest sources with the reddest $U-V$ and $V-J$ colors. Quiescent galaxies are also characterized by very low L_{IR}/L_{UV} ratios, despite their red $U-V$ colors and faint UV fluxes. SED fitting shows a similar trend with best-fit dust attenuation A_V. We note that galaxies at the tip of the red star-forming sequence also appear to have lower sSFRs, so the colors of the very reddest galaxies appear to be a combination of dust and old age (see also Fumagalli in prep, Bedregal, in

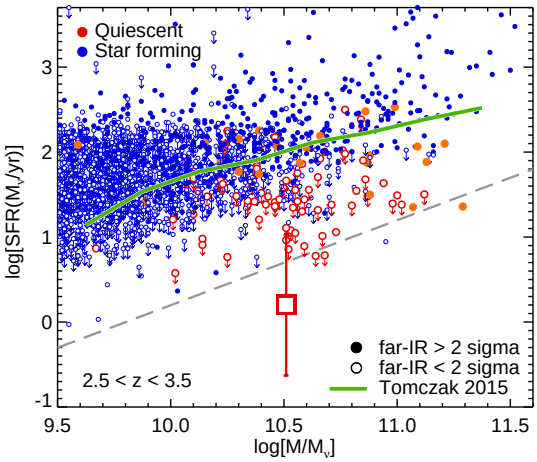

Figure 2.23: SFR versus stellar mass for galaxies at $2.5 < z < 3.5$. We show UVJ selected star-forming galaxies in blue and UVJ selected quiescent galaxies in orange/red. Galaxies with $< 2\sigma$ measurements in the far-IR are shown with open symbols. Downward pointing arrows are $< 1\sigma$ limits on SFR. The large square symbol represents the median SFR_{IR} of 21 undetected quiescent galaxies in CDFS. The thick green line is the median SFR-stellar mass relation for star-forming galaxies from Tomczak et al, in prep, 2015, and the gray dashed line represents the criterion for quiescence at $z = 3$, derived using $\mathrm{sSFR} < (3t_h)^{-1}$.

prep, Forrest, in prep).

Finally, there is a clear trend with stellar mass, such that massive galaxies tend to be redder. Up to the highest redshifts probed here ($z \sim 3.5$) massive $M > 10^{10.5} M_\odot$ galaxies are predominantly quiescent and/or dusty (see e.g. Straatman et al. 2014; Spitler et al. 2014). This is not a selection effect, as only galaxies are shown above our mass completeness limit. The $U - V$ vs stellar mass (left column) shows an absence of low mass quiescent galaxies at $z = 2 - 3$, with massive galaxies quenching first, and lower mass quiescent galaxies rapidly building up from $z = 2$ to $z = 0.5$ (see Tomczak et al. 2014).

The usage of the UVJ technique to identify quiescent galaxies has been validated to $z \lesssim 2.0$, by inspecting stacked WFC3 grism spectra (Whitaker et al. 2013) and stacked MIPS 24μm data (Fumagalli et al. 2014). However, verification at $z \gtrsim 3$ has proved difficult. Straatman et al. (2014) and Spitler et al. (2014) identified quiescent galaxies in the UVJ diagram at even at higher redshifts, to $z \sim 4$, but at these extreme redshifts the IR observations are not deep enough, and the samples are too small, to rule out significant obscured star formation. Here, we place the strongest constraints yet on the

SFRs of quiescent galaxies at $z \sim 3$ by studying the infrared fluxes from ultradeep PACS Herschel data from GOODS-Herschel and PEP. Using the net UV+IR SFR based on the ultradeep Herschel data, we show the SFR versus stellar mass diagram in Figure 2.23. The individual quiescent galaxies at $2.5 < z < 3.5$ are shown in red and orange. 80% of the quiescent galaxies is not detected in the far-IR at $< 2\sigma$ (open red symbols).

To confirm a galaxy as quiescent based on sSFR, we adopt the definition of Damen et al. (2009) which states that a galaxy is quiescent if their sSFR is $< (3t_h)^{-1}$, with t_h the Hubble time. At $z = 3$ this limit corresponds to a sSFR$< 1.6 \times 10^{-10}$/yr. The observed constraints on the sSFR of individual galaxies are not very strong, due to the limited sensitivity in the IR at such high redshift. At $z = 3$ our best estimate of the SFR comes from the combination of ultradeep Spitzer/MIPS 24 μm and Herschel/Pacs/100 and 160 μm data in the CDFS (see Section 2.7). The 1σ limiting SFR for individual sources is SFR$_{IR} < 15 M_\odot$/yr. This means that for galaxies with stellar mass $5 \times 10^{10} M_\odot$ at $2.5 < z < 3.5$, we can constrain their individual sSFRs to be $< 3.0 \times 10^{-10}$/yr. To place firmer constraints, we follow the procedure of Straatman et al. (2014) and stack all $24 - 160\mu$m fluxes, limiting our sample to $M > 10^{10}\ M_\odot$, to gain more sensitivity in the far-IR. We calculate the median of the stack of 21 IR undetected ($< 2\sigma$) galaxies in CDFS, finding a SFR$_{IR} = 1.6 \pm 3.1\ M_\odot$/yr, where the errors are derived by bootstrapping. At a mean stellar mass of $3.2 \times 10^{10} M_\odot$, this translates into a sSFR$= 0.5 \pm 1.0 \times 10^{-10}$ /yr.

This shows that UVJ selected quiescent galaxies at $z \sim 3$ indeed harbour very low levels of star formation, a factor of $> 15\times$ lower at 95% confidence than the median of star-forming galaxies at the same redshift (Tomczak et al. 2015, submitted). Quenching of star formation was thus very efficient, even at such early times.

2.9 Summary

In this paper we have presented the data products and public release of ZFOURGE, a near-IR galaxy survey aimed to increase our understanding of the evolution of galaxies at intermediate to high redshift. The near-IR data is obtained from 45 nights of observing with the `FourStar` instrument on the 6.5 m Magellan Baade Telescope at Las Campanas in Chile. `FourStar` has 5 medium-bandwidth near-IR filters (J_1, J_2, J_3, H_s, H_l) and a K_s-filter, with a wavelength range between 1.05μm and 2.16μm. These filters are particularly well suited to constrain the photometric redshifts of sources in the redshift range $1.5 < z < 3.5$. The images have excellent seeing quality, with a median of $0.5''$ and have 5σ limiting depths (for point-sources) of 24.7$-$26 AB magnitude.

We created K_s-band detection images, by co-adding the FourStar/K_s-band images with publicly available K_s-band data. The detection images have 5σ depths between 25.5 and 26.5 AB magnitude, and we detected $> 70,000$ galaxies in the three extragalactic fields CDFS, COSMOS and UDS. We derived total K_s-band fluxes for every source, and aperture fluxes in $22-36$ UV, optical and near-IR filters in each field over a wavelength range of $0.3-2.16\mu$m, using PSF matched images. We show that for point sources the PSF matched images are consistent to within $< 1\%$. We also derived IR fluxes from four Spitzer/IRAC bands at $3.6-8.0$ μm and from Spitzer/MIPS at 24 μm in all fields. In CDFS we additionally made use of two Herschel/PACS filters, at 100 and 160 μm.

We inspected the catalogs for stars, bad SED fits, sources with low exposure and sources that may be contaminated by nearby bright stars. We included in the catalogs a use flag, that can be used to select the galaxies with good photometry. The catalogs are > 80 % complete up to $25.3-25.9$ AB. We derived stellar mass completeness limits of 10^9 $M_\odot$ at $z = 2$ and $10^9.5$ $M_\odot$ at $z = 4$.

We derived photometric redshifts and stellar population properties, such as stellar mass, by fitting model spectral energy disribution templates to the data. We compared the photometric redshifts from ZFOURGE with spectroscopic redshifts from the literature and found $\sigma_{z,spec} = 0.009$ in CDFS, $\sigma_{z,spec} = 0.008$ in COSMOS and $\sigma_{z,spec} = 0.013$ in UDS. As spectroscopic samples of galaxies are often biased towards bright and blue galaxies, we performed another, independent test of the robustness of the photometric redshifts, by inspecting galaxy pairs. We found excellent results, with $\sigma_z = 0.01-0.02$ for a $K < 23.5$ magnitude limited sample, between $z = 0.5$ an $z = 2.5$. We tested how $\sigma_{z,pairs}$ behaves with redshift, magnitude, SED type and the inclusion of the FourStar mediumbands. $\sigma_{z,pairs}$ is smaller for brighter galaxies, and for blue star-forming galaxies. We compared σ_z from the pair analysis with $\sigma_{z,spec}$ from the spectroscopic test and found excellent correspondence, reinforcing the accurateness of the pair analysis test. We also compared σ_z with the uncertainties derived from the redshift probability functions after fitting, $\sigma_{z,EAZY}$. These are consistent, except for very dusty star-forming galaxies, for which $\sigma_{z,EAZY}$ is underestimated by a factor of 1.5. The redshift quality tests indicate that at $1.5 < z < 2.5$, photometric redshifts are better constrained by 50% if the near-IR FourStar medium-bandwidth filters are included, compared to SED fitting with the FourStar filters removed.

We investigated the efficacy of the UVJ diagram beyond $z = 2$, combining the ZFOURGE photometric redshifts with the far-IR Spitzer/MIPS (in all fields) and Herschel/PACS data (in CDFS). We illustrated how UVJ colors

correlate with sSFR and infrared luminosity excess (dust attenuation) and can be used to distinguish between quiescent and star-forming galaxies to $z = 3.5$. Using the UVJ diagram, we selected a sample of quiescent galaxies at $2.5 < z < 3.5$ and investigated their sSFR properties. We confirmed that these were indeed quiescent, with a on average sSFR$= 0.5 \pm 1.0 \times 10^{-10}$ and $> 15\times$ supressed SFRs,relative to the average stellar mass versus SFR relation of star forming galaxies, thereby for the first time validating the UVJ classification to $z = 3.5$.

2.10 Acknowledgements

This paper, as well as the ZFOURGE data products, could not have come about without the wonderful help of many people. We would like to thank, first of all, the staff at Las Campanas Observatory: David Osip, Jorge Araya, Herman Olivares, Gabriel Prieto, Hugo Rivera, Geraldo Valladares, Jorge Bravo, Gabriel Martin and Mauricio Navarrete. We would also like to thank the Mitchell family for their continuing support. We further thank Lisa Kewley, Tiantian Yuan, Marijn Franx, Mariska Kriek, Adam Muzzin and Jesse van de Sande for useful discussions. This work was supported by the George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, the National Science Foundation grant AST-1009707 and the NL-NWO Spinoza Grant. Australian access to the Magellan Telescopes was supported through the National Collaborative Research Infrastructure Strategy of the Australian Federal Government. GGK acknowledges the support of the Australian Research Council Future Fellowship FT140100933. KEW is supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Oak Ridge Associated Universities through a contract with NASA.

Bibliography

Allen, R. J., Kacprzak, G. G., Spitler, L. R., et al. 2015, *ApJ*, 806, 3

Ashby, M. L. N., Willner, S. P., Fazio, G. G., et al. 2013, *ApJ*, 769, 80

Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, *ApJ*, 600, 681

Bell, E. F., Papovich, C., Wolf, C., et al. 2005, *ApJ*, 625, 23

Bell, E. F., van der Wel, A., Papovich, C., et al. 2012, *ApJ*, 753, 167

Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, *ApJ*, 793, 115

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503

Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2009, *ApJL*, 706, L173

—. 2011, *ApJ*, 739, 24

Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, 200, 13

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

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682

Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, *ApJS*, 189, 270

Chabrier, G. 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Chen, H.-W., Marzke, R. O., McCarthy, P. J., et al. 2003, *ApJ*, 586, 745

Cooper, M. C., Griffith, R. L., Newman, J. A., et al. 2012, *MNRAS*, 419, 3018

Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680

Dale, D. A., & Helou, G. 2002, *ApJ*, 576, 159

Damen, M., Labbé, I., Franx, M., et al. 2009, *ApJ*, 690, 937

Dickinson, M., Giavalisco, M., & GOODS Team. 2003, in The Mass of Galaxies at Low and High Redshift, ed. R. Bender & A. Renzini, 324

Erb, D. K., Pettini, M., Shapley, A. E., et al. 2010, *ApJ*, 719, 1168

Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, Astronomische Nachrichten, 326, 432

Erben, T., Hildebrandt, H., Lerchster, M., et al. 2009, *A&A*, 493, 1197

Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2007, *ApJ*, 665, 265

Fioc, M., & Rocca-Volmerange, B. 1999, ArXiv Astrophysics e-prints, astro-ph/9912179

Fontana, A., Dunlop, J. S., Paris, D., et al. 2014, *A&A*, 570, A11

Franx, M., van Dokkum, P. G., Schreiber, N. M. F., et al. 2008, *ApJ*, 688, 770

Fumagalli, M., Labbé, I., Patel, S. G., et al. 2014, *ApJ*, 796, 35

Furusawa, H., Kosugi, G., Akiyama, M., et al. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 399, Panoramic Views of Galaxy Formation and Evolution, ed. T. Kodama, T. Yamada, & K. Aoki, 131

Giacconi, R., Zirm, A., Wang, J., et al. 2002, *ApJS*, 139, 369

Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJL*, 600, L93

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35

Hildebrandt, H., Pielorz, J., Erben, T., et al. 2009, *A&A*, 498, 725

Hildebrandt, H., Erben, T., Dietrich, J. P., et al. 2006, *A&A*, 452, 1121

Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, *ApJS*, 203, 23

Kawinwanichakij, L., Papovich, C., Quadri, R. F., et al. 2014, *ApJ*, 792, 103

Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189

Knobel, C., Lilly, S. J., Iovino, A., et al. 2012, *ApJ*, 753, 121

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36

Kovač, K., Lilly, S. J., Knobel, C., et al. 2010, *ApJ*, 718, 86

Kriek, M., van der Wel, A., van Dokkum, P. G., Franx, M., & Illingworth, G. D. 2008, *ApJ*, 682, 896

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221

Kriek, M., van Dokkum, P. G., Franx, M., et al. 2006, *ApJL*, 649, L71

Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, *ApJS*, 218, 15

Kron, R. G. 1980, *ApJS*, 43, 305

Labbé, I., Bouwens, R., Illingworth, G. D., & Franx, M. 2006, *ApJL*, 649, L67

Labbé, I., Franx, M., Rudnick, G., et al. 2003, *AJ*, 125, 1107

Labbé, I., Huang, J., Franx, M., et al. 2005, *ApJL*, 624, L81

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, 379, 1599

Lee, J. C., Ly, C., Spitler, L., et al. 2012, *PASP*, 124, 782

Lord, S. D., Hollenbach, D. J., Colgan, S. W. J., et al. 1992, in Bulletin of the American Astronomical Society, Vol. 24, American Astronomical Society Meeting Abstracts, 1182

Madau, P., Pozzetti, L., & Dickinson, M. 1998, *ApJ*, 498, 106

Magnelli, B., Popesso, P., Berta, S., et al. 2013, *A&A*, 553, A132

Maíz Apellániz, J. 2006, *AJ*, 131, 1184

Marchesini, D., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2009, *ApJ*, 701, 1765

Marchesini, D., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, 725, 1277

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, 544, A156

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, *ApJS*, 206, 8

—. 2013b, *ApJ*, 777, 18

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, 660, L43

Nonino, M., Dickinson, M., Rosati, P., et al. 2009, *ApJS*, 183, 244

Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, *PASP*, 107, 375

Papovich, C., Momcheva, I., Willmer, C. N. A., et al. 2010, *ApJ*, 716, 1503

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010a, *AJ*, 139, 2097

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010b, *ApJ*, 721, 193

Pérez-González, P. G., Rieke, G. H., Villar, V., et al. 2008, *ApJ*, 675, 234

Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, *PASP*, 125, 654

Postman, M., Franx, M., Cross, N. J. G., et al. 2005, *ApJ*, 623, 721

Quadri, R., Marchesini, D., van Dokkum, P., et al. 2007, *AJ*, 134, 1103

Quadri, R. F., & Williams, R. J. 2010, *ApJ*, 725, 794

Retzlaff, J., Rosati, P., Dickinson, M., et al. 2010, *A&A*, 511, A50

Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, *ApJS*, 172, 86

Schlafly, E. F., & Finkbeiner, D. P. 2011, *ApJ*, 737, 103

Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, 172, 1

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ArXiv e-prints, arXiv:1403.3689

Sobral, D., Smail, I., Best, P. N., et al. 2013, *MNRAS*, 428, 1128

Spitler, L. R., Labbé, I., Glazebrook, K., et al. 2012, *ApJL*, 748, L21

Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJL*, 787, L36

Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, *ApJL*, 783, L14

Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, *ApJS*, 172, 9

Tilvi, V., Papovich, C., Tran, K.-V. H., et al. 2013, *ApJ*, 768, 56

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, *ApJ*, 783, 85

Tully, R. B., Mould, J. R., & Aaronson, M. 1982, *ApJ*, 257, 527

van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, *ApJS*, 203, 24

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ArXiv e-prints, arXiv:1404.2844

van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJL*, 677, L5

van Dokkum, P. G., Labbé, I., Marchesini, D., et al. 2009, *PASP*, 121, 2

Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJL*, 754, L29

Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86

Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, ArXiv e-prints, arXiv:1305.1943

Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, 691, 1879

Windhorst, R. A., Cohen, S. H., Hathi, N. P., et al. 2011, *ApJS*, 193, 27

Wolf, C., Meisenheimer, K., Kleinheinrich, M., et al. 2004, *A&A*, 421, 913

Wuyts, S., Labbé, I., Schreiber, N. M. F., et al. 2008, ApJ, 682, 985

Wuyts, S., Labbé, I., Franx, M., et al. 2007, *ApJ*, 655, 51

Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, *ApJ*, 738, 106

Wuyts, S., Förster Schreiber, N. M., Genzel, R., et al. 2012, *ApJ*, 753, 114

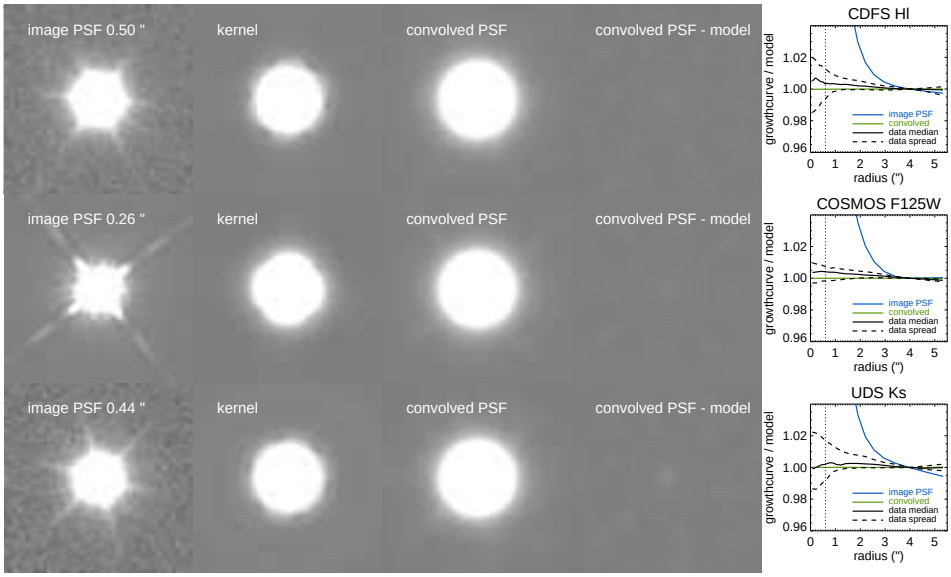

Figure 2.24: Example PSF diagnostic plots. Here we show the groundbased FourStar/H_l and K_s filters and the space based HST/WFC3/F125W filter, with various PSF widths. Postage stamps ($10.65'' \times 10.65''$) from left to right are: the median stacked PSF of the original science images, with their FWHM indicated at the top; a kernel derived using the deconvolution code developed by I. Labbé; the PSF from the leftmost panel, convolved using the corresponding kernel; the convolved PSF minus the moffat model. The rightmost panels show growthcurves for a number of cases, divided by the model growthcurve and normalized at $r = 4''$. A perfect comparison with the model means the ratio of growthcurves will be one at all radii. In blue we show the growthcurve of the orginal PSF (leftmost postage stamp) and in green the same, but after convolving. The black lines represent the median and 1σ scatter of individual pointsources in the convolved images that were used for photometry.

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579

Yuan, T., Nanayakkara, T., Kacprzak, G. G., et al. 2014, *ApJL*, 795, L20

2.A PSF convolution

In Section 2.3 we have explained how images are convolved, such that their average point source profile matches a $FWHM = 0.9''$ Moffat PSF, with the goal of obtaining consistent aperture photometry over all filters. Here were show example diagnostic plots (Figure 2.24) of the kernel derivation. The

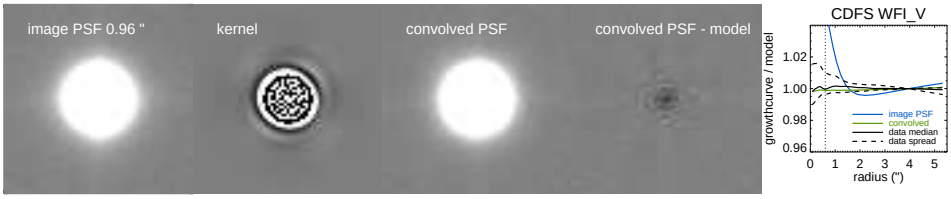

Figure 2.25: Example of a case where the orginal PSF was worse than the target PSF. See Figure 2.24 for a description of the panels. Here the image was *deconvolved*. A small residual is left after subtracting the model PSF from the deconvolved PSF, but the integrated flux at $r = 0.6''$ corresponds well, to within 1%.

four columns with postage stamps represent the original image PSF, the kernel used for convolution, the PSF after convolving with this kernel and the residual after subtracting the model moffat profile. Growthcurves representing the PSFs before and after convolution are shown in the rightmost panels. The green curve in particular shows the light profile of the convolved PSF divided by that of the model, which is close to unity. Whereas in Figure 2.7 we show the growthcurves measured on the median PSF derived from the convolved images, we inspect here the median and 1σ scatter of the growthcurves of individual stars, finding similar residuals of $< 2\%$ compared with the model PSFs.

Mathematically, we are able to produce PSF matched images in case of *deconvolution*, when the image has a worse quality, i.e. a broader PSF, than the target $FWHM = 0.9''$ moffat profile. We have applied deconvolution in some cases with care, as we cannot make any assumptions on the true underlying light profile of galaxies. We find that the strongest residuals occur once the image PSF becomes much larger than the target PSF. After inspecting the residuals by eye and taking into account the aperture radius with a diameter of $1.2''$, we include images that have up to 15% broader PSFs than the target PSF. We show an example in Figure 2.25. While the deconvolved PSF shows some residual compared to the model, the growthcurves indicate that we capture the same amount of light within 1% at $r = 0.6''$, the aperture radius that we use to derive the catalog fluxes with. In total 11 of 92 UV to near-IR images were deconvolved.

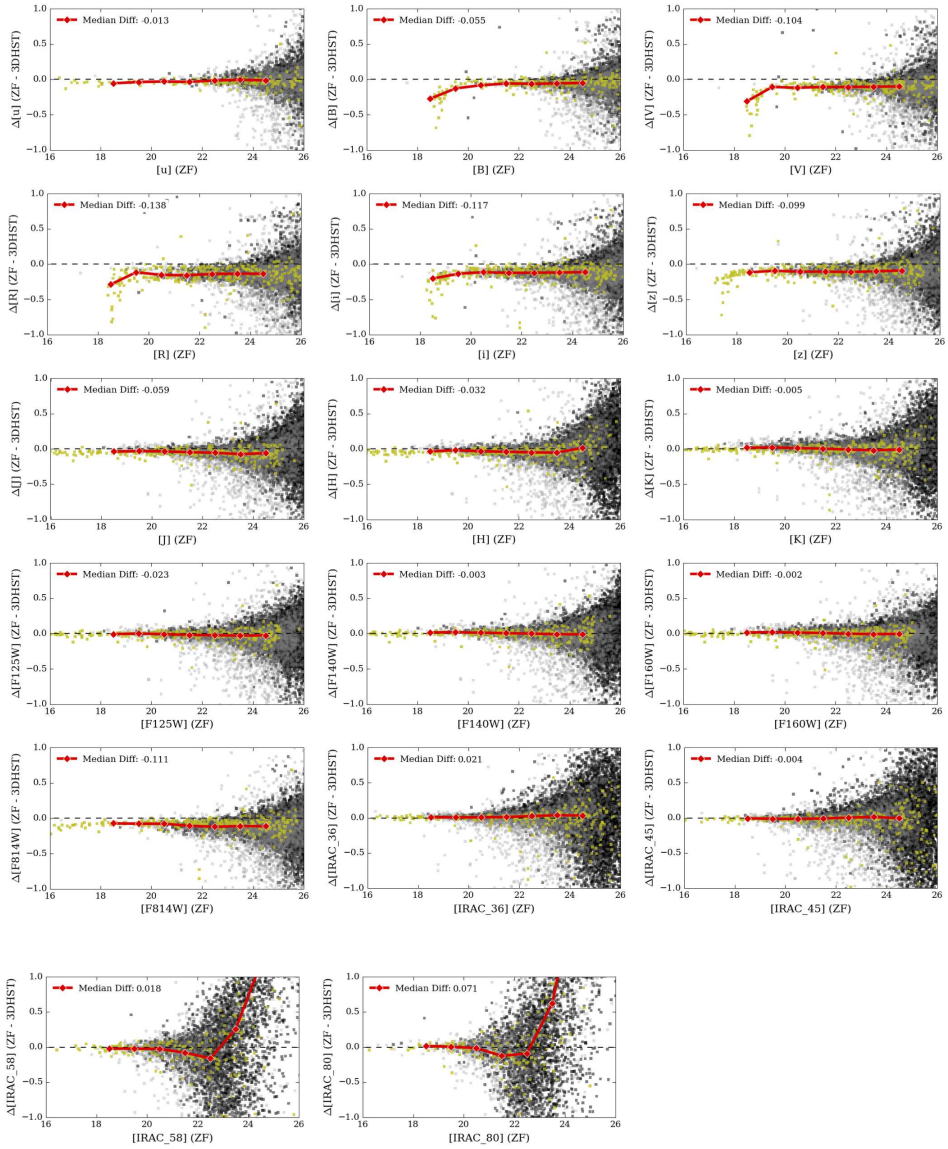

Figure 2.26: The difference between ZFOURGE and 3D-HST source magnitudes plotted as a function of ZFOURGE magnitudes in UDS for each band in common. Galaxies with `use=1` are shown as black points, point sources with `star=1` are shown as yellow points, and blended sources with SEflag= 2 (and `use=1`) are shown as grey points. The median magnitude difference for all galaxies is shown by the red solid line and large red diamond symbols in bins of 1 mag.

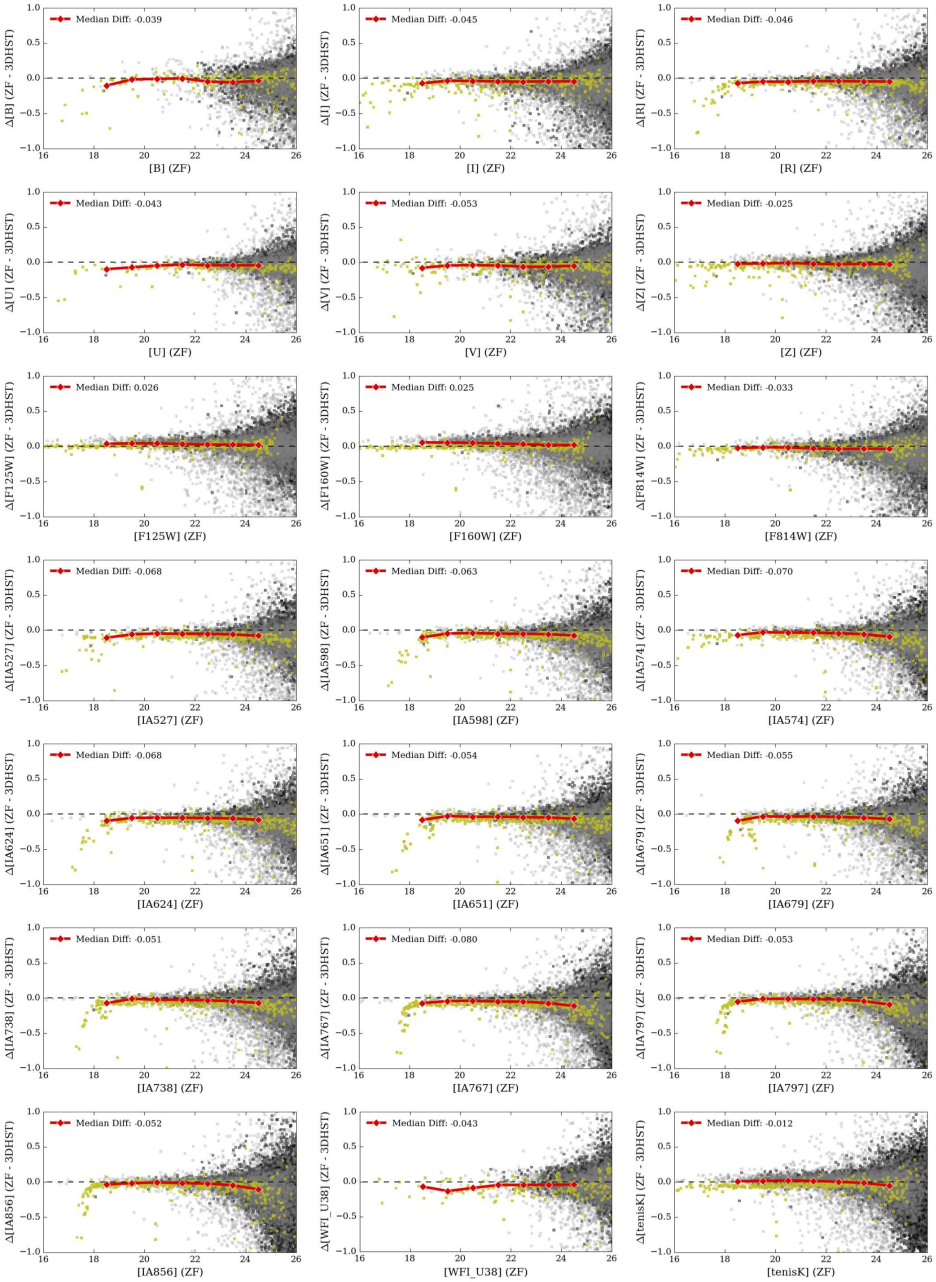

Figure 2.27: The difference between ZFOURGE and 3D-HST source magnitudes plotted as a function of ZFOURGE magnitudes in CDFS for each band in common. Symbols are the same as in Figure 2.26

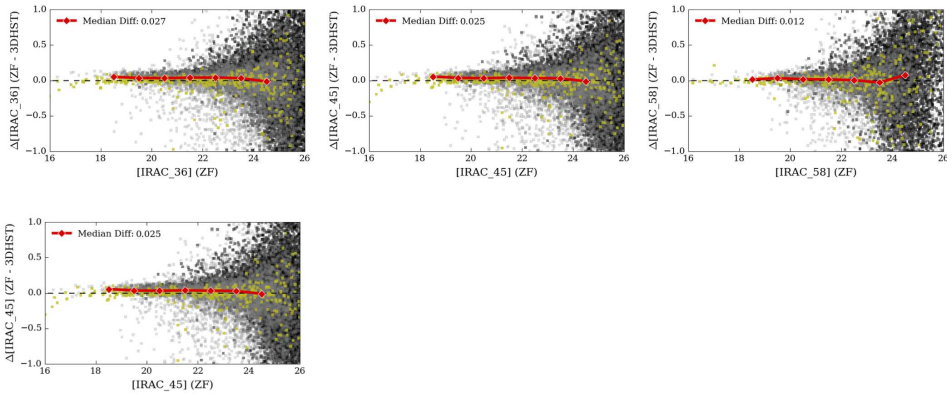

Figure 2.28: The difference between ZFOURGE and 3D-HST source magnitudes plotted as a function of ZFOURGE magnitudes in CDFS for each band in common. Symbols are the same as in Figure 2.26

2.B Comparison to the 3DHST photometric catalogs

In this section we provide a comparison of the total magnitudes measured by ZFOURGE and those measured by 3D-HST, who make use of many of the same ancillary images. They also largely use the same methods to derive photometry. For each filter in common we calculate the difference in magnitude between crossmatched sources and show this versus total magnitude as per the ZFOURGE catalogs in Figures 2.26 to 2.30. For crossmatching we used a maximum angular separation of $1''$. We separately indicated sources that were flagged as possibly blended or contaminated by neighbours by SE (`SExflags`≥ 2). Δmag has somewhat more scatter for these sources at faint magnitudes, but overall the correspondence is quite good between the surveys, with Δmag close to zero. The most notable exceptions are the V, R, i and z-bands in UDS, which tend to be ~ 0.1 magnitudes brighter in ZFOURGE.

In Figure 2.31 we show the positional offsets between source locations in ZFOURGE and the 3D-HST survey. The median offsets are $\leq 0.005''$ in RA or Dec in all fields, indicating the images are uniformly calibrated and can be reliably used for inter-survey comparisons.

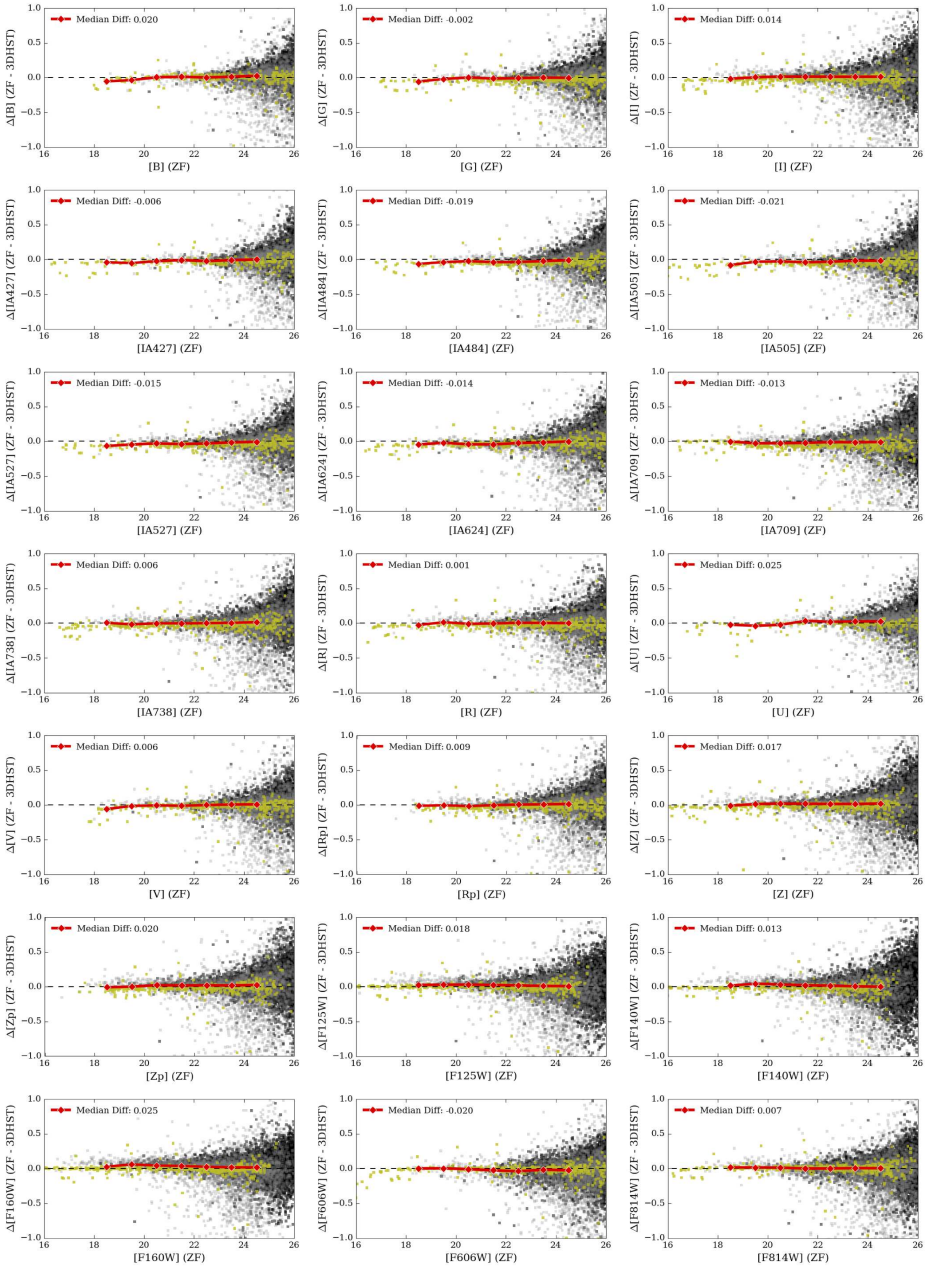

Figure 2.29: The difference between ZFOURGE and 3D-HST source magnitudes plotted as a function of ZFOURGE magnitudes in COSMOS for each band in common. Symbols are the same as in Figure 2.26

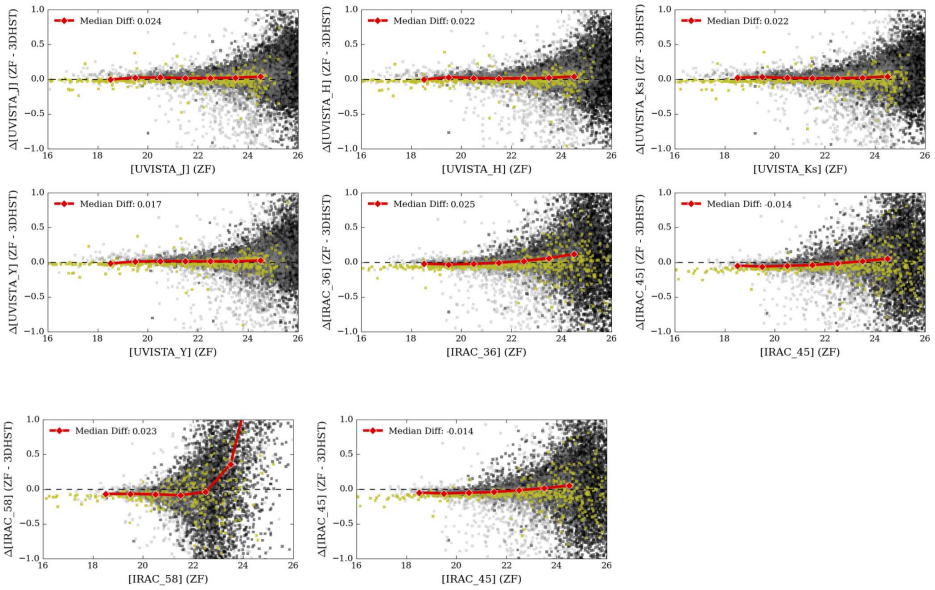

Figure 2.30: The difference between ZFOURGE and 3D-HST source magnitudes plotted as a function of ZFOURGE magnitudes in COSMOS for each band in common. Symbols are the same as in Figure 2.26

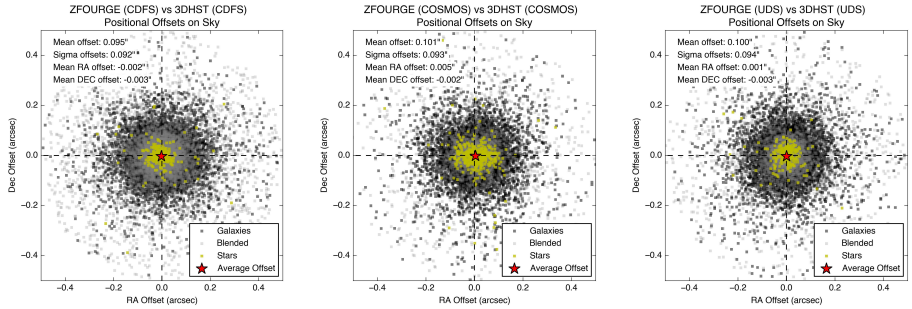

Figure 2.31: Positional offsets between source locations in ZFOURGE and 3D-HST, using the same symbols as in Figures 2.26 to 2.30. The median offsets are indicated by red stars.

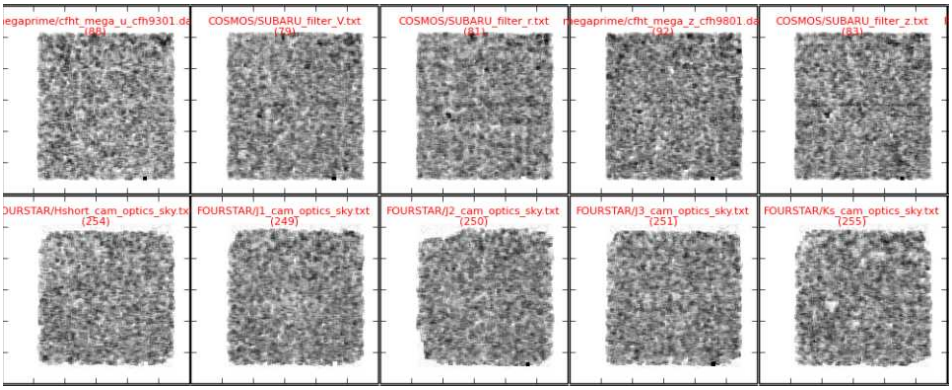

Figure 2.32: Spatial residual maps of example filters in COSMOS. The grayscale ranges from 0.95 to 1.05, i.e. a 5% flux deviation.

2.C Spatial variation in the zeropoints

In Figure 2.32 we show example spatial zeropoint residual maps derived with EAZY, by comparing the best-fit templates to the observed galaxy SEDs. The residuals are of the order of < 5 %. We fitted a two-dimensional polynomial to each offset map and used these to apply a correction to the flux of a specific filter, for all sources as a function of their x- and y-position. This was done for the full dataset. With this method we were able to trace subtle systematic offsets between the four `FourStar` detectors in J_1, J_2, H_s and H_l in UDS and fully correct for these. In one image strong, non-linear spatial effects stand out: this is in VIMOS/R in CDFS. In this image the strong spatial varation could not be removed by this first order correction. Due to its large depth, we have kept the image in our sample, but applied a minimum error floor to the V-band flux of 5% to take into account uncertainties on the zeropoint of the image.

2.D UVJ diagram field comparison

In this section we show the UVJ diagram color-coded by stellar mass (Figure 2.33), as in the third column of Figure 2.22. We show the same redshift bins, but split the diagrams into the three ZFOURGE fields. By comparing the rest-frame colors in the different fields we can look for inconsistencies, for example if the median locii of the datapoints are offset relative to each other. Here this is not the case, indicating consistent photometry.

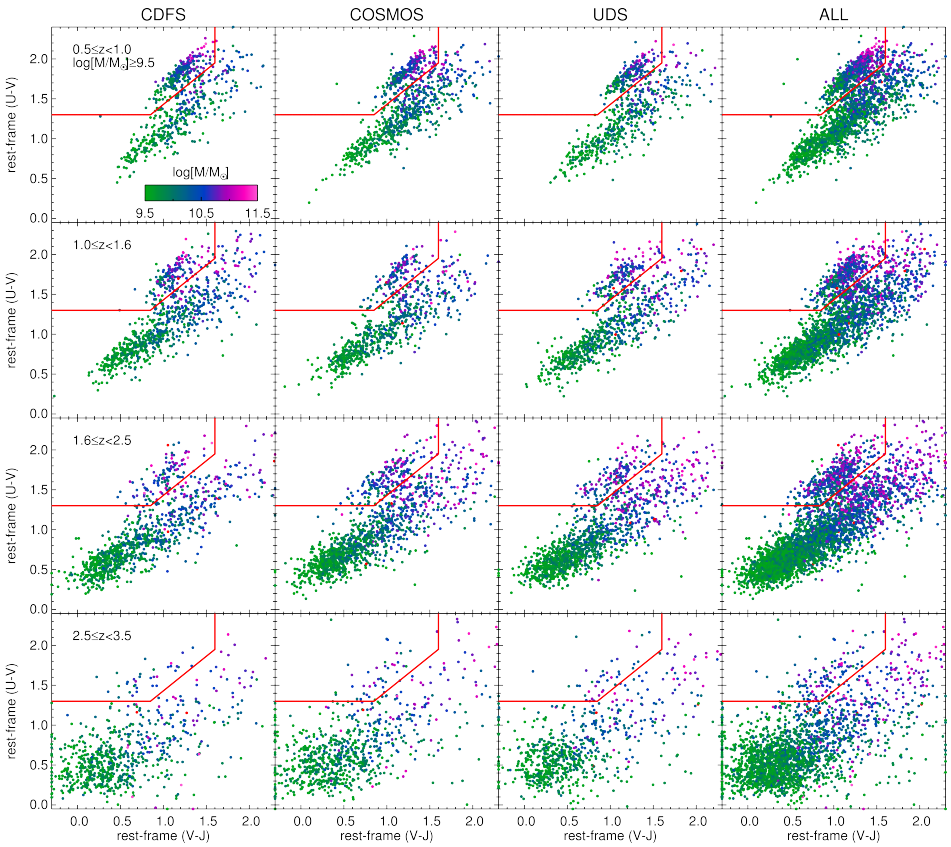

Figure 2.33: Rest-frame $U-V$ versus $V-J$ diagrams of galaxies with `use=1`, $SNR_{K_s} > 10$ and stellar mass $M > 10^{9.5} M_{\odot}$. In the first three columns we show the three ZFOURGE fields. In the last column these are combined. From top to bottom we show bins of increasing redshift. The color scaling indicates stellar mass. The rest-frame $U-V$ and $V-J$ colors in each field show the same pattern, and the same location for the red sequence, indicating consistent photometry.

3

A substantial population of massive quiescent galaxies at z ~ 4 from ZFOURGE

Abstract

We report the likely identification of a substantial population of massive $M \sim 10^{11} M_{\odot}$ galaxies at $z \sim 4$ with suppressed star formation rates (SFRs), selected on rest-frame optical to near-IR colors from the FourStar Galaxy Evolution Survey. The observed spectral energy distributions show pronounced breaks, sampled by a set of near-IR medium-bandwidth filters, resulting in tightly constrained photometric redshifts. Fitting stellar population models suggests large Balmer/4000Å breaks, relatively old stellar populations, large stellar masses and low SFRs, with a median specific SFR of $2.9 \pm 1.8 \times 10^{-11}$/yr. Ultradeep Herschel/PACS 100μm, 160μm and Spitzer/MIPS 24μm data reveal no dust-obscured SFR activity for 15/19(79%) galaxies. Two far-IR detected galaxies are obscured QSOs. Stacking the far-IR undetected galaxies yields no detection, consistent with the SED fit, indicating independently that the average specific SFR is at least 10× smaller than of typical star-forming galaxies at $z \sim 4$. Assuming all far-IR undetected galaxies are indeed quiescent, the volume density is $1.8 \pm 0.7 \times 10^{-5}$Mpc^{-3} to a limit of $\log_{10} M/M_{\odot} \geq 10.6$, which is 10× and 80× lower than at $z = 2$ and $z = 0.1$. They comprise a remarkably high fraction($\sim 35\%$) of $z \sim 4$ massive galaxies, suggesting that suppression of star formation was efficient even at very high redshift. Given the average stellar age of 0.8Gyr and stellar mass of $0.8 \times 10^{11} M_{\odot}$, the galaxies likely started forming stars before $z = 5$, with SFRs well in excess of $100 M_{\odot}$/yr, far exceeding that of similarly abundant UV-bright galaxies at $z \geqslant 4$. This suggests that most of the star formation in the progenitors of quiescent $z \sim 4$ galaxies was obscured by dust.

3.1 Introduction

The identification of a population of compact quiescent galaxies at $1 < z < 3$, characterized by suppressed star formation and very small sizes, has attracted a significant interest (e.g. Daddi et al. 2005; van Dokkum et al. 2008). It remains an open question when these galaxies first appeared. The declining number densities and fractions of quiescent galaxies with redshift at $1 < z < 3$ suggest they might be rare at $z > 3$ (e.g. Muzzin et al. 2013). Nevertheless, the high ages of some quiescent galaxies at $z \sim 2$ (Whitaker et al. 2013) suggest they could have already existed at much earlier times. If confirmed at $z > 3$, massive galaxies must have formed rapidly, early and with an effective mechanism of suppressing star formation.

Beyond $z = 3$ candidate early-type or post-starburst galaxies have been reported, despite uncertainties whether their red colors could be due to dust-reddening (e.g. Chen & Marzke 2004; Wiklind et al. 2008; Mancini et al. 2009; Fontana et al. 2009; Guo et al. 2013; Stefanon et al. 2013; Muzzin et al. 2013). In this Letter, we use the FourStar Galaxy Evolution Survey[1] (ZFOURGE; Labbé et al. in preparation) to look for the earliest examples of quiescent galaxies. The strength of ZFOURGE lies in the unique combination of depth and the medium-bandwidth filters (covering $1-2\mu$m) of the FourStar Infrared Camera (Persson et al. 2013) on the 6.5m Magellan Baade Telescope. These enable the derivation of accurate photometric redshifts and the detection of the age-sensitive Balmer/4000Å break in faint, red galaxies at $1 < z < 4.2$.

Throughout, we assume a ΛCDM cosmology with $\Omega_{\rm M} = 0.3, \Omega_\Lambda = 0.7$ and $H_0 = 70 {\rm kms}^{-1}{\rm Mpc}^{-1}$. The photometric system is AB.

3.2 Data

ZFOURGE covers three $11' \times 11'$ pointings in the fields CDFS, COSMOS and UDS, to very deep limits (~ 26 AB total mag (5σ) in J_1, J_2, J_3 and ~ 25 mag in H_s, H_l and K_s). We combine ZFOURGE with public data, including HUGS (PI:Fontana) HAWK-I and CANDELS (Grogin et al. 2011; Koekemoer et al. 2011) HST/WFC3 imaging, over a wavelength range of 0.3-8 μm. Full photometric K_s-band selected catalogs will be presented in Straatman et al. (in preparation).

We use Spitzer/MIPS 24μm data from GOODS-South (PI: Dickinson), COSMOS (PI: Scoville) an d SPUDS (PI: Dunlop) and ultradeep Herschel/PACS 100μm and 160μm imaging from the GOODS-Herschel (Elbaz et al. 2011) and

[1] http://zfourge.tamu.edu

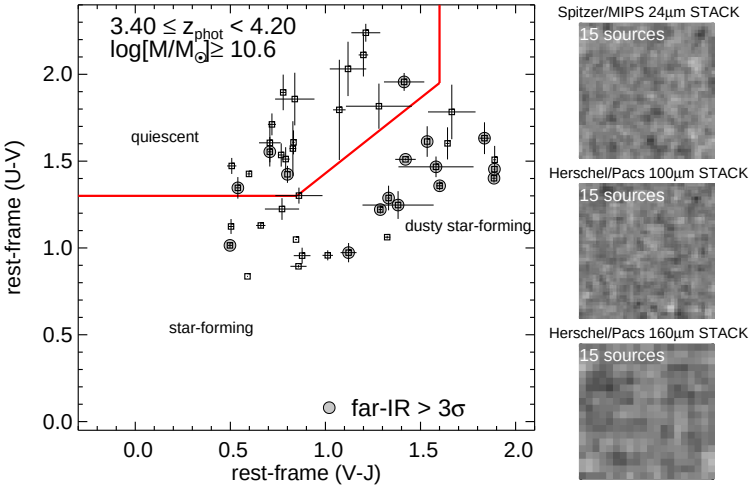

Figure 3.1: Left: Rest-frame UVJ diagram of galaxies with $\log_{10}M/M_{\odot} \geq 10.6$ at $3.4 \leq z < 4.2$. The red solid line separates quiescent galaxies (top-left region) from star-forming galaxies. Galaxies with $\geq 3\sigma$ far-IR detections are indicated with gray filled symbols and tend to be found amongst dusty star-forming galaxies. 19 objects are classified as quiescent, with 15/19 far-IR undetected. Right: 24μm, 100μm and 160μm stacks ($48'' \times 48''$) of undetected quiescent galaxies. Stacking yields no detection: $S_{24\mu m} < 0.002$mJy, $S_{100\mu m} < 0.090$mJy and $S_{160\mu m} < 0.140$mJy (1σ).

the CANDELS-Herschel program (PI:Dickinson), to independently place constraints on the on-going SFR. The ultradeep PACS 160μm imaging currently provides the best sensitivity for far-IR light from star formation at high redshift, trading off k-correction and source confusion due to increasing beam size (Elbaz et al. 2011), while 24μm data are more sensitive to the presence of hot dust associated with AGN.

Photometric redshifts and rest-frame colors were derived with EAZY (Brammer et al. 2008). Comparing ZFOURGE photometric redshifts to spectroscopic redshifts, Tomczak et al. (2013) found a scatter of $\sigma_{\delta z/(1+z)} = 0.019$. Stellar population properties were derived by fitting Bruzual & Charlot (2003) models with FAST (Kriek et al. 2009), assuming a Chabrier (2003) initial mass function, exponentially declining star formation histories with timescale τ, and solar metallicity.

Table 3.1: Quiescent galaxies at $3.4 \leq z < 4.2$

ID	ra (deg)	dec (deg)	z_{phot}	U-V	V-J
ZF-CDFS-209	53.1132774	−27.8698730	3.56±0.05	1.43±0.02	0.60±0.01
ZF-CDFS-403	53.0784111	−27.8598385	3.660[a]	1.42±0.05	0.80±0.00
ZF-CDFS-617	53.1243553	−27.8516121	3.700[a]	1.35±0.06	0.54±0.01
ZF-CDFS-4719	53.1969414	−27.7604313	3.59±0.14	1.54±0.07	0.77±0.02
ZF-CDFS-4907	53.1812820	−27.7564163	3.46±0.16	1.57±0.16	0.83±0.02
ZF-CDFS-5657	53.0106506	−27.7416019	3.56±0.07	1.61±0.07	0.83±0.01
ZF-COSMOS-13129	150.1125641	2.3765368	3.81±0.17	1.96±0.05	1.41±0.11
ZF-COSMOS-13172	150.0615082	2.3786869	3.55±0.06	1.90±0.10	0.78±0.01
ZF-COSMOS-13414	150.0667114	2.3823516	3.57±0.19	1.61±0.11	0.71±0.06
ZF-UDS-885	34.3685074	−5.2994704	3.99±0.41	1.80±0.29	1.07±0.03
ZF-UDS-1236	34.3448868	−5.2925615	3.58±0.08	1.47±0.05	0.51±0.02
ZF-UDS-2622	34.2894516	−5.2698011	3.77±0.10	1.51±0.07	0.79±0.02
ZF-UDS-3112	34.2904282	−5.2620673	3.53±0.06	1.71±0.06	0.72±0.02
ZF-UDS-5418	34.2937546	−5.2269468	3.53±0.07	1.55±0.08	0.71±0.01
ZF-UDS-6119	34.2805405	−5.2171388	4.05±0.27	1.86±0.15	0.84±0.10
ZF-UDS-9526	34.3381844	−5.1661916	3.97±0.18	2.11±0.12	1.20±0.03
ZF-UDS-10401	34.3601379	−5.1530914	3.91±0.38	1.82±0.13	1.28±0.17
ZF-UDS-10684	34.3650742	−5.1488328	3.95±0.48	2.03±0.16	1.12±0.10
ZF-UDS-11483	34.3996315	−5.1363320	3.63±0.32	2.24±0.05	1.21±0.08

ID	M ($10^{11}M_\odot$)	SFR_{SED} ($M_\odot$/yr)	$sSFR_{SED}$ (10^{-11}/yr)	τ[b] (Gyr)	age (Gyr)
ZF-CDFS-209	0.76	2.239	2.884	0.10	0.63
ZF-CDFS-403	1.15	31.623	27.542	0.25	0.79
ZF-CDFS-617	0.69	13.183	18.621	0.16	0.63
ZF-CDFS-4719	0.45	0.851	1.905	0.16	1.00
ZF-CDFS-4907	0.40	0.000	0.000	0.01	0.40
ZF-CDFS-5657	0.76	3.311	4.467	0.25	1.26
ZF-COSMOS-13129	1.78	0.000	0.000	0.01	1.58
ZF-COSMOS-13172	1.45	0.000	0.000	0.04	0.79
ZF-COSMOS-13414	0.44	0.035	0.079	0.10	1.00
ZF-UDS-885	0.60	0.001	0.001	0.03	0.40
ZF-UDS-1236	0.60	0.550	0.912	0.06	0.50
ZF-UDS-2622	0.87	16.218	18.621	0.16	0.63
ZF-UDS-3112	0.43	1.862	4.467	0.25	1.26
ZF-UDS-5418	0.44	3.020	6.761	0.16	0.79
ZF-UDS-6119	0.55	5.623	10.471	0.10	0.50
ZF-UDS-9526	0.89	16.596	18.621	0.16	0.63
ZF-UDS-10401	0.38	0.001	0.002	0.02	0.25
ZF-UDS-10684	0.85	3.802	4.467	0.25	1.26
ZF-UDS-11483	1.02	0.000	0.000	0.01	1.00

[a] Spectroscopic redshift from Szokoly et al. (2004).

[b] $SFR \sim \exp(-t/\tau)$.

Table 3.2: Quiescent galaxies at $3.4 \leq z < 4.2$ continued

ID	A(V)	Ks_tot[c] (AB)	24μm[d] (mJy)	100μm[d] (mJy)	160μm[d] (mJy)
ZF-CDFS-209	0.3	22.6	−0.001±0.004	0.163±0.163	−0.047±0.123
ZF-CDFS-403	0.8	22.4	0.100±0.005$^{*\dagger\times}$	1.272±0.199$^{*\dagger\times}$	1.686±0.201$^{*\dagger\times}$
ZF-CDFS-617	0.3	22.3	0.087±0.003*†	1.062±0.152*†	0.362±0.157†
ZF-CDFS-4719	0.3	23.4	−0.000±0.004	−0.339±0.175	−0.287±0.146
ZF-CDFS-4907	0.8	23.6	0.001±0.004	0.302±0.132	0.154±0.107
ZF-CDFS-5657	0.3	23.0	0.001±0.005†	0.193±0.259†	0.078±0.214†
ZF-COSMOS-13129	0.6	23.6	0.112±0.010*	0.895±0.356	−0.215±0.343
ZF-COSMOS-13172	0.6	22.4	0.004±0.007	−0.007±0.394	−0.323±0.420
ZF-COSMOS-13414	0.2	23.4	0.009±0.008	−0.330±0.546	0.194±0.452
ZF-UDS-885	1.3	24.0	0.012±0.008	1.009±0.537	−0.089±0.400
ZF-UDS-1236	0.4	22.6	−0.016±0.011	0.383±0.449	0.537±0.360
ZF-UDS-2622	0.9	23.0	0.013±0.010	0.761±0.372	0.152±0.442
ZF-UDS-3112	0.0	23.2	−0.010±0.010	−0.547±0.377	0.256±0.333
ZF-UDS-5418	0.5	23.3	0.049±0.010*	0.560±0.443	−0.494±0.386
ZF-UDS-6119	1.0	23.8	−0.013±0.009	0.224±0.477	0.331±0.316
ZF-UDS-9526	1.8	24.2	0.016±0.009	0.038±0.351	−0.000±0.294
ZF-UDS-10401	1.7	24.6	0.007±0.010	−0.500±0.340	−0.490±0.383
ZF-UDS-10684	1.0	24.1	0.007±0.012	−0.105±0.388	−0.503±0.544
ZF-UDS-11483	1.0	23.6	0.004±0.011	−0.149±0.409	0.715±0.325

[c] Total magnitude in the FourStar K_s−band.

[d] "*": $\geq 3\sigma$ detections. "†": X-ray detected (Szokoly et al. 2004; Xue et al. 2011). "×": radio source (Miller et al. 2013).

3.3 Selection of quiescent galaxies at z ~ 4

We use a two-color criterion (rest-frame $U-V$ versus $V-J$; Figure 3.1) to separate quiescent galaxies (red in $U-V$, but blue in $V-J$) from star-forming galaxies, (blue or red in both $U-V$ and $V-J$ colors) (e.g., Labbé et al. 2005; Williams et al. 2009). This technique has been shown to isolate the red sequence of galaxies at $z < 3$ (e.g., Whitaker et al. 2011) and was spectroscopically confirmed to identify quiescent galaxies at $z \sim 2$ (e.g. Whitaker et al. 2013).

We focus on the redshift range $3.4 \leq z < 4.2$, where the medium-bandwidth filters straddle the Balmer/4000Å break. At $z > 3.4$ the break enters the H_l filter (1.7μm), while at $z < 4.2$ the K_s–band (2.2μm) still probes light redward of the break. We limit the sample to a signal-to-noise of >7 in K_s and stellar masses of $\log_{10}M/M_{\odot} \geq 10.6$, where we are complete for passively evolving stellar populations formed at $z < 10$. This yields 44 galaxies with high quality photometry, of which 15 fall in the UVJ quiescent region and are undetected in the FIR, a significant fraction: $34 \pm 13\%$ (15/44). A summary of their properties is presented in Tables 3.1 and 3.2. Their photometric redshifts range from $z = 3.46$ to $z = 4.05$ with a mean of $z = 3.7$ and mean uncertainty $\delta z/(1+z) = 0.036$, leading to well constrained rest-frame colors. Two galaxies have spectroscopic redshifts, with a mean $(z_{phot} - z_{spec})/(1+z_{spec}) = -0.039$ (Szokoly et al. 2004).

3.4 Properties of quiescent galaxies at z ~ 4

3.4.1 Spectral energy distributions

We show representative SEDs of seven galaxies in Figure 3.2. The median SED of all 15 (far-IR undetected) quiescent galaxies, constructed by de-redshifting their photometry and normalising them to the flux density at 4500Å, is also shown.

The observed SEDs are exceedingly faint in the optical ($I \sim 27$ magnitude) and extremely red throughout the near-IR (median $I - Ks = 3.7 \pm 0.33$). The SEDs are characterized by a sharp break, with $H - Ks = 1.9 \pm 0.20$ and peaking in K_s, and a blue spectral slope in the mid-IR Spitzer/IRAC bands ($Ks - [4.5\mu\text{m}] = 0.91 \pm 0.13$). The break is reminiscent of the strong break found in quiescent galaxies at lower redshift, where it is caused by combination of the Balmer and 4000Å absorption features, indicative of a combination of relatively old stellar populations and suppressed star formation. Additionally, some galaxies exhibit a second break at bluer wavelengths, which is likely

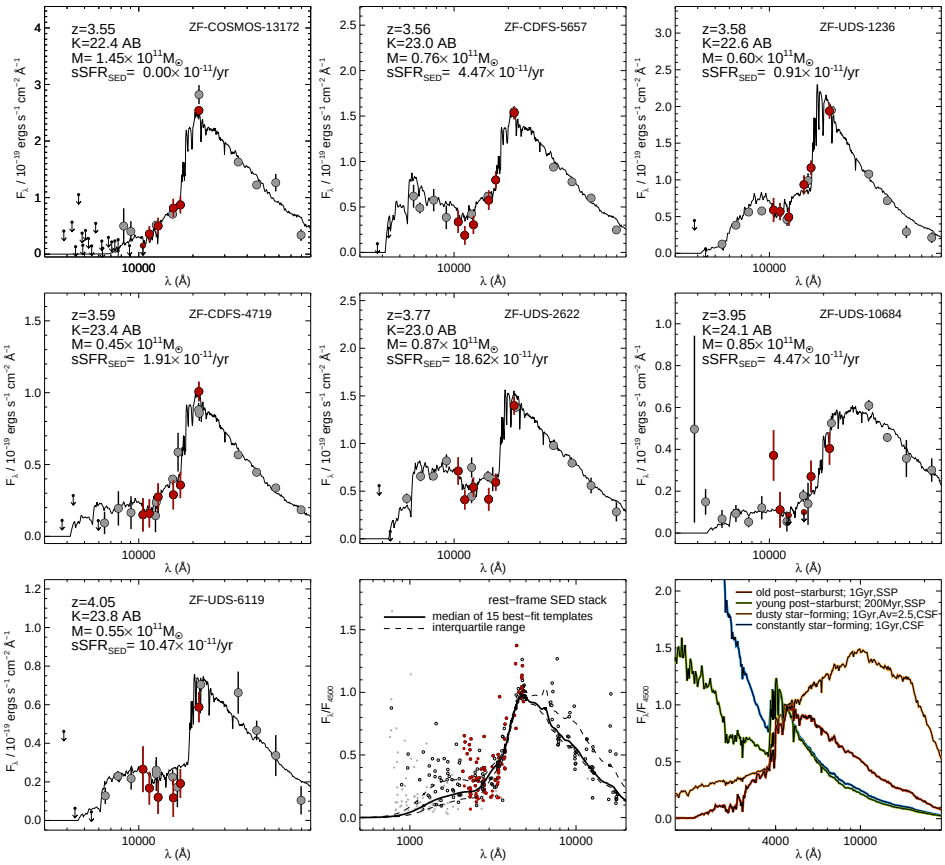

Figure 3.2: Observed SEDs of UVJ selected quiescent galaxies. Red datapoints correspond to the FourStar medium-bandwidth filters. The solid curve is the fitted model from FAST. Downward pointing arrows are 1σ upper limits. Bottom-middle: Rest-frame SED of the 15 far-IR undetected galaxies (open symbols), normalized at rest-frame 4500Å, with gray symbols corresponding to 1σ upper limits. The solid curve is the median of the best-fit template SEDs. Dashed lines mark the interquartile range. Bottom-right: Four model SEDs with constant star formation or a single stellar population (SSP) and ages from 200Myr to 1Gyr. The observed SEDs are characterized by pronounced Balmer/4000Å breaks, similar to the old post-starburst model.

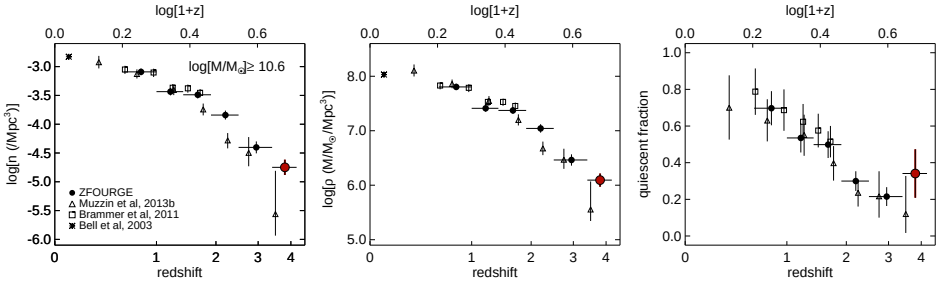

Figure 3.3: Left: Number density of quiescent galaxies in ZFOURGE. Middle: Stellar mass density. Right: Quiescent fraction. Errors are a combination of the Poisson uncertainty and cosmic variance. Horizontal errorbars indicate the width of the redshift bins. The red symbols denote the 15 UVJ selected and far-IR undetected galaxies. We compare with Bell et al. (2003) at $z = 0.1$, Brammer et al. (2011) to $z \lesssim 2$ and Muzzin et al. (2013) at $0 < z < 4$. The overall trend is a decrease in number density towards $z \sim 4$, consistent with the earlier NMBS and UltraVISTA results. However, the larger depth and sampling of ZFOURGE allows for much better constraints on the evolution at $2 < z < 4$. Surprisingly, at $z = 3.7$, $34 \pm 13\%$ of the galaxies with $\log_{10} M/M_{\odot} \geq 10.6$ could be quiescent, suggesting that the decline of the quiescent fraction could flatten at $z \gtrsim 2-3$.

the Lyman break.

As it is difficult to separate the contribution of the Balmer break and the 4000Å break from photometry alone, we quantify the size of the total combined break (D_{tot}) by estimating the flux (F_v) ratio at 4000 – 4100Å and 3500 – 3650Å on the best-fit models of each galaxy individually, extending the definition of D4000n (Balogh et al. 1999) to cover the Balmer break as well. We find the median $D_{tot} = 2.8 \pm 0.1$, which is in the range of post-starburst galaxies with suppressed star formation (e.g., a $\tau = 10$Myr model produces $D_{tot} = 3.1$ at 500Myr). In contrast, unobscured constant star-forming (CSF) models only reach $D_{tot} = 1.8$ at 1Gyr. Heavily obscured star-forming models (e.g. 1Gyr, CSF, $A_v = 2.5$) can also reach quite red $D_{tot} = 2.5$, but are ruled out as they predict very red $Ks-[4.5\mu\text{m}] = 2.0$, whereas the observed SED-slopes are bluer.

3.4.2 Stellar population fits

Models with exponentially declining SFRs fit the data well, with a median $\chi^2_{red} = 1.3$. The median best-fit age of the galaxies is 0.8Gyr, the median star formation timescale (τ) 0.1Gyr, the average stellar mass $0.8 \times 10^{11} M_{\odot}$, and the median specific SFR (sSFR) $2.9 \pm 1.8 \times 10^{-11}$/yr.

To test if models with ongoing star formation provide acceptable fits to the sample, we force τ = 250Myr, τ = 1Gyr or CSF and refitted the data, finding a median $\chi^2_{red,250Myr} = 2.1$, $\chi^2_{red,1Gyr} = 6.6$ and $\chi^2_{red,CSF} = 6.6$. This shows that models with $\tau = 250$Myr provide almost equally good fits, but longer formation timescales ($\tau > 1$Gyr) provide poor fits to the data. For all individual galaxies $\tau < 250$Myr models produced better fits than did $\tau > 1$Gyr models. We note that high redshift solutions with low sSFRs are preferred in all cases. Other solutions, e.g. at low redshift or with ongoing obscured star formation are ruled out at more than 99% confidence for 18/19 galaxies.

We refitted the data using the models of Maraston (2005), and obtained a mean stellar mass of $0.5 \times 10^{11} M_{\odot}$, with a typical offset of −0.2 dex compared to the masses in Table 3.1, and a median sSFR of 0.1×10^{-11}/yr. Hence the result is not strongly dependent on the adopted stellar population model.

Overall, the fits suggest most stars were formed at $z > 5$, followed by an epoch of suppressed star formation. As expected, the median stellar ages are lower than the typical age of 1.3 Gyr found at $z \sim 2$ by Whitaker et al. (2013). We find some galaxies with very red $U - V$ and $V - J$, pointing towards older stellar populations. However, their best-fit ages are the same as for the bluer galaxies, with larger redshift uncertainties or dust, suggesting that dust and photometric scatter are the main causes.

3.4.3 Independent constraints on SFR and AGN activity from Herschel

We derived Spitzer/MIPS 24μm, Herschel/PACS 100μm and 160μm flux intensities, measured in apertures of 7″, 8″ and 12″ diameter with aperture corrections of 2.56, 2.45 and 2.60, respectively (assuming a point source profile). Light from neighbouring sources was subtraced following Labbé et al. (2010).

We find 1, 2 and 4 $> 3\sigma$ detections at 160μm, 100μm and 24μm, respectively. These 4/19 galaxies may have obscured star formation. The total detection rate (21 ± 11%) is lower than the 50% of 24μm detections reported earlier for $z > 3$ quiescent galaxies (Stefanon et al. 2013). Two are also detected in X−ray and are type−2 QSOs (Szokoly et al. 2004; Xue et al. 2011), of which one is a radio source (Miller et al. 2013). A third X−ray detection is found amongst the far-IR undetected galaxies, for a total of 3 likely AGN.

To place tighter constraints on the average far-IR luminosity of the 15 far-IR undetected galaxies, we stack their 24μm, 100μm and 160μm images (Figure 3.1), with uncertainties derived by bootstrap resampling. The formal measurements are $S_{24\mu m} = 0.001 \pm 0.002$mJy, $S_{100\mu m} = 0.049 \pm 0.090$mJy and $S_{160\mu m} = 0.039 \pm 0.140$mJy. Hence the sources are undetected. The strongest

constraint on obscured SFR is obtained at 160μm. Using the IR templates of Wuyts et al. (2011), we find $7.1 \pm 25 M_{\odot}$/yr. Given the mean mass of the sample ($0.8 \times 10^{11} M_{\odot}$), this corresponds to a sSFR of $0.9 \pm 3 \times 10^{-10}$yr^{-1}. While these independently derived limits cannot rule out ongoing obscured star formation, they are consistent with the SED fits ($2.9 \pm 1.8 \times 10^{-11}$/yr), and are $\sim 10\times$ smaller than the sSFR= $3 - 6 \times 10^{-9}$/yr of similarly massive galaxies at $z \sim 3$ and typical UV-bright star-forming galaxies at $z \sim 4$ (e.g. Stark et al. 2013; Viero et al. 2013).

3.4.4 Contamination by emission lines

We caution that the galaxies here could in fact be vigorously star-forming, if their K_s-band fluxes were boosted dramatically by emission lines ([$OIII$] and $H\beta$), mimicking a Balmer Break (e.g. Shim et al. 2011; Stark et al. 2013). We tested this scenario by fitting CSF models to the SEDs *without fitting the* K_s-band, leading to fits with high obscuration (median $A(V) \sim 2$).

The CSF models fit the data poorly and vastly underpredict the median K_s-band: $(K_{s,obs} - K_{s,SED})_{CSF} = -0.53 \pm 0.06$ (note that standard$-\tau$ models predict the K_s-band magnitude nearly perfectly: $(K_{s,obs} - K_{s,SED})_{\tau=free} = -0.01 \pm 0.05$). Assuming this excess is due to strong ($EW_{obs} \sim 2000$Å) emission lines, the predicted median SFR is $\sim 1000 M_{\odot}$/yr, which, because of the high obscuration, should result in $7 - 18\sigma$ detections in 160μm, but is not observed.

Furthermore, we can test the hypothesis that K_s is boosted by [$OIII$] ($\lambda\lambda$4959,5007Å) and $H\beta$(4861Å) at $3.0 < z < 3.6$, by looking at existing narrowband $NB209$ (2.10μm) data, covering CDFS+COSMOS (Lee et al. 2012). Since the odds are only 20% that $NB209$ is affected by any of these lines, it effectively traces the continuum. Using a simple model, drawing uniformly random redshifts at $3.0 < z < 3.6$, the predicted median color is $(K_s - NB209) = -0.44$, nearly independent of line ratios, in strong disagreement with the observed median $(K_s - NB209) = -0.04 \pm 0.1$. We also inspected data from the 3D-HST survey (Brammer et al. 2012), with low resolution spectral coverage at $1.1 - 1.6\mu$m, as strong [$OIII$]/$H\beta$ lines in K_s would imply strong [OII] in the HST/WFC3 grism. We found only 2 detections for 13 galaxies with coverage to an emission line sensitivity of $\sim 5 \times 10^{-17} ergs/s/cm^2 (5\sigma)$: $MgII$(λ2798Å) for the QSO ZF-CDFS-617 and [OII] for the 24μm detected ZF-UDS-5418.

3.5 Implications

From hereon we adopt as operational definition of "quiescent": galaxies that satisfy the UVJ criterium and are not detected in the far-IR (e.g. Bell et al.

2012). We note however, that the current data do not allow to determine conclusively whether the galaxies have completely stopped forming stars as the sample is too faint for spectrographs on large telescopes.

3.5.1 Number densities

From the 15 quiescent galaxies we estimate the volume and stellar mass density, finding $1.8 \pm 0.7 \times 10^{-5}$Mpc^{-3} and $1.2 \pm 0.5 \times 10^{6}M_{\odot}$Mpc^{-3}, respectively. Uncertainties are the quadratic sum of the Poission uncertainty and variations due to large scale structure (Moster et al. 2011). The volume density in the $11' \times 11'$ area in the ZFOURGE–UDS field is $\sim 3\times$ higher than in the ZFOURGE–COSMOS field, underscoring the need for probing mulitple pointings to faint limits.

For comparison with other surveys, we integrated the COSMOS/ UltraVISTA mass function at $3 < z < 4$ for quiescent galaxies of Muzzin et al. (2013), based on a similar UVJ classification, to $\log_{10}M/M_{\odot} \geq 10.6$. UltraVISTA produces a number density of 2.7×10^{-6}Mpc^{-3} and a mass density of 3.1×10^{5}M$_{\odot}$ Mpc^{-3}. These are factors of ~ 7 and ~ 4, lower (albeit at only $\sim 1.6\sigma$ significance). This is likely a completeness effect, as UltraVISTA is only complete to $M \gtrsim 10^{11}M_{\odot}$. Indeed, Muzzin et al. (2013) select galaxies with $K_{s,tot,AB} < 23.4$, while 50% of the galaxies here have $K_{s,tot,AB} > 23.4$.

The number and stellar mass densities of quiescent galaxies at $0.6 \leq z < 4.2$ are shown in Figure 3.3. These were obtained from the full ZFOURGE catalogs (Straatman et al., in prep), using the same selection criteria as described in section 3.3. The number density decreases rapidly towards $z \sim 4$ ($\sim 10\times$ lower than at $z = 2$ and $\sim 80\times$ than at $z = 0.1$), suggesting that a small fraction ($10 - 15\%$) of $z \sim 2$ quiescent galaxies was already in place at $z \sim 4$. The last panel of Figure 3.3 shows the fraction of quiescent galaxies with $\log_{10}M/M_{\odot} \geq 10.6$. This strongly declines with redshift between $0.6 < z < 3$. Therefore, we would expect a value close to zero at $z \sim 4$, but we find a surprisingly high fraction of $34 \pm 13\%$. This is similar to the value at $z \sim 2.2$ ($30 \pm 8\%$), suggesting a flat quiescent fraction at $2 < z < 4$.

3.5.2 Star-forming progenitors

Given average stellar ages of 0.8Gyr and masses of $0.8 \times 10^{11}M_{\odot}$, the galaxies likely started forming their stars much earlier than $z = 5$, with SFRs well in excess of $100M_{\odot}$/yr. This raises the question what are the likely progenitors. In recent years, UV-luminous galaxies have been found in large numbers to $z \sim 10$ (e.g. Bouwens et al. 2013; Ellis et al. 2013). These are actively star-

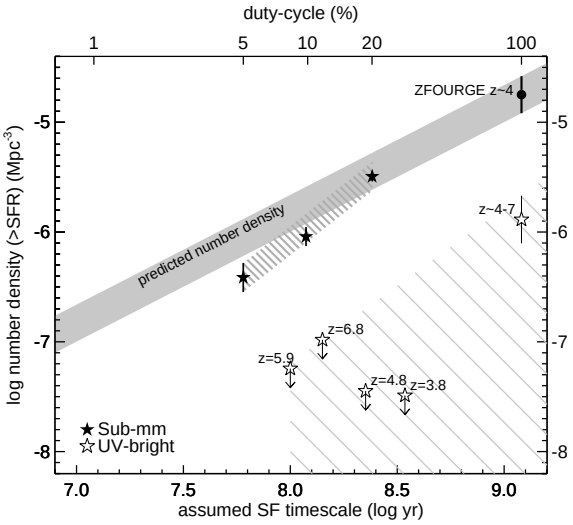

Figure 3.4: The predicted cumulative number density (n(> SFR)/Mpc^3) for the progenitors of the $z \sim 4$ quiescent galaxies as a function of the assumed formation timescale. The gray filled area shows the expected range in number density based on the $z \sim 4$ sample. If the galaxies form over a long timescale (e.g. ~ 1 Gyr), this is simply the observed number density. If they form in shorter, more intense bursts (smaller duty-cycles), the observed progenitor number densities are expected to be smaller. Open star symbols are estimates from the SFR functions of Smit et al. (2012) ($z \sim 4-7$) and the UV-luminosity functions of van der Burg et al. (2010) ($z = 3.8$ and $z = 4.8$), Willott et al. (2013) ($z = 5.9$) and Bowler et al. (2012) ($z = 6.8$). For the latter we assume that $\log_{10}\mathrm{SFR} = -0.4(M_{1600} - \langle A_{1600}\rangle) - 7.25$ (Kennicutt 1998), with M_{1600} the luminosity at 1600Å in AB mag and $\langle A_{1600}\rangle$ the dust-correction factor from Bouwens et al. (2012). Upper limits are derived from the respective survey volumes. Filled stars show sub-mm number counts (Karim et al. 2013), using a conversion of $1\mathrm{mJy} \approx 1.667 \times 10^{12} L_{\odot,IR}$ (Blain et al. 2002) and $1 M_\odot/yr = 4.5 \times 10^{-44} L_{IR}(ergs/s)$ (Kennicutt 1998) and assuming 10% of these are at $z > 4$. The number of luminous UV-bright galaxies at $z > 4$ is far too low (1-2 dex), while the sub-mm counts are much better matched, suggesting there might be sufficient numbers of heavily obscured starburst galaxies at $z > 4$ if their formation timescales are ~200Myr.

forming, although even the most luminous galaxies found so far at $z > 4$ have relatively modest UV-derived SFRs ($< 100 M_{\odot}$/yr) (Smit et al. 2012).

The expected number density at $z > 4$ of the progenitors depends on the assumed star formation timescale (gray shaded area in Figure 3.4). If the progenitors were visible at all times (i.e. a formation timescale of ~1 Gyr and number density of $1.8 \pm 0.7 \times 10^{-5}$Mpc^{-3}), then we can use the SFR functions at $z = 4 - 7$ (Smit et al. 2012) to select progenitors with sufficiently high SFRs on fixed cumulative number density (van Dokkum et al. 2010).

As shown in Figure 3.4, this number density falls ~ 1.2 dex short. If we assume shorter formation timescales (e.g. a few 100Myr), the progenitors require much higher SFRs and are predicted to be found in smaller numbers. Comparing to UV-luminosity functions from wide-field surveys, using the redshift window as the formation timescale, the number densities are > 1.5 dex too low, reflecting that sufficiently luminous UV-bright galaxies are extremely rare.

Alternatively, the main star formation episode is obscured by dust. There exists a population of high-redshift sub−mm detected galaxies, including highly obscured gas-rich mergers (e.g. Younger et al. 2007), with large SFRs ($\gtrsim$ $1000 M_{\odot}$/yr), that could be progenitors of $z \sim 4$ quiescent galaxies. Based on the 870μm source counts of Karim et al. (2013), and tentatively assuming that 10% are at $z > 4$ (e.g. Swinbank et al. 2012), we find that obscured star-bursting galaxies are sufficiently numerous. This suggests that most of the star formation in the progenitors of quiescent $z \sim 4$ galaxies could have been obscured by dust.

3.6 Summary

Using very deep imaging from ZFOURGE we find evidence for the existence of massive ($M \sim 10^{11} M_{\odot}$) galaxies with suppressed star formation at very early times ($z \sim 4$). The galaxies satisfy the UVJ criterium, which has been shown to efficiently select quiescent galaxies at $z < 3$ (Whitaker et al. 2011, 2013; Williams et al. 2009). The observed SEDs show prominent breaks, well sampled by the FourStar near-IR medium-bands, leading to accurate photometric redshifts and illustrating a key strength of the survey. The SEDs are well fit by models with strong Balmer/4000Å breaks, small $\tau (< 250$Myr), high ages (~ 0.8Gyr) and low sSFRs ($2.9 \pm 1.8 \times 10^{-11}$/yr). Consistent with this, 79% of the galaxies are undetected in deep Spitzer/MIPS and Herschel/PACS imaging. Stacking the far-IR places an independent constraint on the average sSFR of $< 3 \times 10^{-10}$yr^{-1}, a factor $> 10\times$ smaller than the average sSFR of UV-bright star-forming galaxies at these redshifts (e.g. Stark et al. 2013) and consistent

with these galaxies having strongly suppressed SFRs.

While rare (with number densities $\sim 10\times$ and $\sim 80\times$ lower than at $z = 2$ and $z = 0.1$), they make up a surprisingly high fraction of the massive galaxy population at $z \sim 4$ ($34 \pm 13\%$), higher than expected based on the declining trend over $1 < z < 3$, suggesting an effective mechanism of suppressing star formation and short formation timescales (< 1Gyr). The implied SFRs needed to form galaxies with a mean stellar mass of $0.8 \times 10^{11} M_{\odot}$ in such a short time exceeds that of similarly abundant UV-bright galaxies at $z \geqslant 4$, suggesting that most of the star formation in their progenitors was obscured by dust.

We emphasize that without spectroscopic confirmation the number of quiescent galaxies at $z \sim 4$ remains poorly constrained, but given their faint magnitudes, real progress will likely have to wait until the launch of JWST or construction of ELTs. Currently, ALMA observations can place stronger limits on the dust-obscured activity of these galaxies and help identify the progenitors at $z \gtrsim 4$.

3.7 Acknowledgements

This research was supported in part by the George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy. We would especially like to thank the Mitchell family for their continuing support. We thank Danilo Marchesini and Adam Muzzin for help with the UltraVISTA number densities. We thank Marijn Franx, Mattia Fumagalli, Shannon Patel, Jesse van de Sande, David Sobral, Renske Smit, Paul van der Werf and Claudia Maraston for useful discussions. We acknowledge support by the following grants: NSF AST-1009707, ERC HIGHZ #227749 and NL-NWO Spinoza. Australian access to the Magellan Telescopes was supported through the NCRIS of the Australian Federal Government. This work is based on observations made with Herschel, an ESA Cornerstone Mission with significant participation by NASA, through an award issued by JPL/Caltech.

Bibliography

Balogh, M. L., Morris, S. L., Yee, H. K. C., Carlberg, R. G., & Ellingson, E. 1999, *ApJ*, 527, 54

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, *ApJS*, 149, 289

Bell, E. F., van der Wel, A., Papovich, C., et al. 2012, *ApJ*, 753, 167

Blain, A. W., Smail, I., Ivison, R. J., Kneib, J.-P., & Frayer, D. T. 2002, PhR, 369, 111

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, *ApJ*, 754, 83

Bouwens, R. J., Oesch, P. A., Illingworth, G. D., et al. 2013, *ApJL*, 765, L16

Bowler, R. A. A., Dunlop, J. S., McLure, R. J., et al. 2012, *MNRAS*, 426, 2772

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503

Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, *ApJ*, 739, 24

Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, 200, 13

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

Chabrier, G. 2003, *PASP*, 115, 763

Chen, H.-W., & Marzke, R. O. 2004, *ApJ*, 615, 603

Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680

Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, *A&A*, 533, A119

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

Fontana, A., Santini, P., Grazian, A., et al. 2009, *A&A*, 501, 15

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, 207, 24

Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, *MNRAS*, 432, 2

Kennicutt, Jr., R. C. 1998, *ARA&A*, 36, 189

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221

Labbé, I., Huang, J., Franx, M., et al. 2005, *ApJL*, 624, L81

Labbé, I., González, V., Bouwens, R. J., et al. 2010, *ApJL*, 716, L103

Lee, J. C., Ly, C., Spitler, L., et al. 2012, *PASP*, 124, 782

Mancini, C., Matute, I., Cimatti, A., et al. 2009, *A&A*, 500, 705

Maraston, C. 2005, *MNRAS*, 362, 799

Miller, N. A., Bonzini, M., Fomalont, E. B., et al. 2013, *ApJS*, 205, 13

Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, *ApJ*, 731, 113

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, 777, 18

Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, *PASP*, 125, 654

Shim, H., Chary, R.-R., Dickinson, M., et al. 2011, *ApJ*, 738, 69

Smit, R., Bouwens, R. J., Franx, M., et al. 2012, *ApJ*, 756, 14

Stark, D. P., Schenker, M. A., Ellis, R., et al. 2013, *ApJ*, 763, 129

Stefanon, M., Marchesini, D., Rudnick, G. H., Brammer, G. B., & Whitaker, K. E. 2013, *ApJ*, 768, 92

Swinbank, A. M., Karim, A., Smail, I., et al. 2012, *MNRAS*, 427, 1066

Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, *ApJS*, 155, 271

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2013, ArXiv e-prints, arXiv:1309.5972

van der Burg, R. F. J., Hildebrandt, H., & Erben, T. 2010, *A&A*, 523, A74

van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJL*, 677, L5

van Dokkum, P. G., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, 709, 1018

Viero, M. P., Moncelsi, L., Quadri, R. F., et al. 2013, *ApJ*, 779, 32

Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86

Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2013, *ApJL*, 770, L39

Wiklind, T., Dickinson, M., Ferguson, H. C., et al. 2008, *ApJ*, 676, 781

Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, 691, 1879

Willott, C. J., McLure, R. J., Hibon, P., et al. 2013, *AJ*, 145, 4

Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, *ApJ*, 738, 106

Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *ApJS*, 195, 10

Younger, J. D., Fazio, G. G., Huang, J.-S., et al. 2007, *ApJ*, 671, 1531

4

The sizes of massive quiescent and star forming galaxies at z ~ 4 with ZFOURGE and CANDELS

Abstract

We study the rest-frame ultra-violet sizes of massive ($\sim 0.8\times10^{11}M_{\odot}$) galaxies at $3.4 \leq z < 4.2$, selected from the FourStar Galaxy Evolution Survey (ZFOURGE), by fitting single Sérsic profiles to HST/WFC3/F160W images from the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS). Massive quiescent galaxies are very compact, with a median circularized half-light radius $r_e = 0.63 \pm 0.18$kpc. Removing 5/16 (31%) sources with signs of AGN activity does not change the result. Star-forming galaxies have $r_e = 2.0 \pm 0.60$kpc, $3.2 \pm 1.3\times$ larger than quiescent galaxies. Quiescent galaxies at $z \sim 4$ are on average $6.0 \pm 1.7\times$ smaller than at $z \sim 0$ and $1.9 \pm 0.7\times$ smaller than at $z \sim 2$. Star-forming galaxies of the same stellar mass are $2.4 \pm 0.7\times$ smaller than at $z \sim 0$. Overall, the size evolution at $0 < z < 4$ is well described by a powerlaw, with $r_e = 5.08 \pm 0.28(1+z)^{-1.44\pm0.08}$kpc for quiescent and $r_e = 6.02 \pm 0.28(1+z)^{-0.72\pm0.05}$kpc for star-forming galaxies. Compact star-forming galaxies are rare in our sample: we find only $1/14 \Rightarrow 7\%$ with $r_e/(M/10^{11}M_{\odot})^{0.75} < 1.5$, whereas $13/16 \Rightarrow 81\%$ of the quiescent galaxies is compact. The number density of compact quiescent galaxies at $z \sim 4$ is $1.8 \pm 0.8 \times 10^{-5}$Mpc^{-3} and increases rapidly, by $> 5\times$, between $2 < z < 4$. The paucity of compact star-forming galaxies at $z \sim 4$ and their large rest-frame ultra-violet median sizes suggest that the formation phase of compact cores is very short and/or highly dust obscured.

4.1 Introduction

In recent years massive quiescent galaxies have been found beyond $z = 3$ (e.g. Chen & Marzke 2004; Wiklind et al. 2008; Mancini et al. 2009; Fontana et al. 2009; Marchesini et al. 2010; Guo et al. 2013; Stefanon et al. 2013; Muzzin et al. 2013; Spitler et al. 2014) and even at $z \sim 4$, when the universe was only 1.5 Gyrs old (Straatman et al. 2014). Quiescent galaxies at high redshift ($z > 1$) exhibit compact morphologies, with small effective radii (e.g Daddi et al. 2005; van Dokkum et al. 2008; Damjanov et al. 2009), which tend to become smaller with increasing redshift (van der Wel et al. 2014). At $z \sim 3$, they have sizes of $\sim$ 1kpc, $3-4\times$ smaller than early-type galaxies of similar stellar mass at $z \sim 0$ (Shen et al. 2003; Mosleh et al. 2013) and $2-3\times$ smaller than star-forming galaxies at the same redshift.

How compact quiescent galaxies are formed is still unclear. Simulations propose mechanisms in which gas-rich major mergers can induce central starbursts, resulting in a compact merger remnant (Hopkins et al. 2009; Wellons et al. 2015), or in which massive star-forming clumps move to the centers if galaxy disks are unstable (Dekel et al. 2009; Dekel & Burkert 2014). Alternatively they may have formed in a more protracted process at high redshift, when the universe was more dense (Mo et al. 1998).

To understand these scenarios it is necessary to identify compact quiescent galaxies and their progenitors at the highest redshifts. Compact star-forming galaxies been found in small numbers at $z = 2-3$ (Barro et al. 2014a,b; Nelson et al. 2014), but many host AGN, complicating the interpretation of the observations. At the same time, rest-frame ultra-violet (UV) or optically measured sizes of star-forming galaxies may be affected by dust-obscured central regions, thereby increasing their effective radii.

In this work we investigate the sizes of a stellar-mass complete sample of star-forming and quiescent galaxies at $z \sim 4$. Throughout, we assume a standard ΛCDM cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70$km s^{-1}Mpc^{-1}. The adopted photometric system is AB.

4.2 Sample selection

The galaxies were selected using deep K_s-band images from the FourStar Galaxy Evolution Survey (ZFOURGE; Labbé et al. in prep.), a near-IR survey with the FourStar Infrared Camera (Persson et al. 2013), covering three $11' \times 11'$ pointings, located in the fields CDFS (Giacconi et al. 2002), COSMOS (Scoville et al. 2007) and UDS (Lawrence et al. 2007). The ZFOURGE K_s-band selected catalogs are at least 80% complete down to $K_s = 24.53$, 24.74 and

25.07 mag in each field, respectively (Papovich et al. 2015). Photometric redshifts and stellar masses were derived using 5 near-IR medium-bandwidth filters on FourStar (J_1, J_2, J_3, H_s, H_l), which provide fine sampling of the age-sensitive Balmer/4000Å break at $1.5 < z < 4$, in combination with public data over a wavelength range $0.3-8\mu$m (Straatman et al. 2014). Here we make additional use of HST/WFC3/F160W data from CANDELS (Grogin et al. 2011; Koekemoer et al. 2011; Skelton et al. 2014), to examine galaxy sizes and Spitzer/MIPS 24μm data from GOODS-South (PI: Dickinson), COSMOS (PI: Scoville) and SPUDS (PI: Dunlop) to measure infrared flux.

The galaxies in this work have photometric redshifts $3.4 \leq z < 4.2$, stellar masses of $\log_{10}(M/M_{\odot}) \geq 10.55$ and K_s-band signal-to-noise (SNR) of SNR> 7. They are separated into quiescent and star-forming according to their rest-frame $U-V$ versus $V-J$ colours (Labbé et al. 2005; Williams et al. 2009; Spitler et al. 2014), yielding 19 quiescent and 25 star-forming galaxies (Straatman et al. 2014). Of these, 34 have HST/WFC3/F160W coverage. One quiescent galaxy has SNR< 3 in F160W and is not included. Another star-forming galaxy with a highly uncertain redshift solution was also rejected from the sample, along with two star-forming galaxies that appear to consist of two sources each in the higher resolution HST images. In total we study 16 quiescent and 14 star-forming galaxies. We include a control sample at $2 \leq z < 3.4$ (326 sources) at similar mass and SNR.

4.3 Galaxy sizes from HST/WFC3 imaging

4.3.1 Sérsic fits

Sizes and structural parameters were measured by fitting Sérsic (Sersic 1968) profiles on $6'' \times 6''$ HST/WFC3/F160W image stamps using GALFIT (Peng et al. 2010). In particular, we measure the half-light radius, encapsulating half the sources' integrated light. The corresponding parameter in GALFIT is the half-light radius along the semi-major axis ($r_{1/2,maj}$), which can be converted to circularized effective radius ($r_e = r_{1/2,maj}\sqrt{(b/a)}$), with b/a the axis ratio.

We manually subtracted the background in each image stamp, masking sources and using the mode of the pixel flux distribution. Sky estimation in GALFIT was turned off. Neighbouring objects at $r > 1.1''$ from the source were effectively masked by setting their corresponding pixels in the image to zero flux and increasing those in the noise image by $\times 10^6$. Close neighbouring objects were fitted simultaneously.

We created mean PSFs for each field by stacking image stamps of bright stars (masking all neighbouring sources). As many of the galaxies are small

Table 4.1: Properties of 16 quiescent and 14 star forming galaxies

ID	R.A. (deg)	Decl (deg)	z	$K_{s,tot}$ (mag)	H_{tot}[a] (mag)
QUIESCENT					
ZF-CDFS-209	53.1132774	-27.8698730	3.56	22.6	24.1
ZF-CDFS-403	53.0784111	-27.8598385	3.660[d]	22.4	23.7
ZF-CDFS-4719	53.1969414	-27.7604313	3.59	23.4	25.2
ZF-CDFS-4907	53.1812820	-27.7564163	3.46	23.6	25.0
ZF-CDFS-5657	53.0106506	-27.7416019	3.56	23.0	24.6
ZF-CDFS-617	53.1243553	-27.8516121	3.700[d]	22.3	23.5
ZF-COSMOS-13129	150.1125641	2.3765368	3.81	23.6	25.2
ZF-COSMOS-13172	150.0615082	2.3786869	3.55	22.4	24.4
ZF-COSMOS-13414	150.0667114	2.3823516	3.57	23.4	25.4
ZF-UDS-10684	34.3650742	-5.1488328	3.95	24.1	25.9
ZF-UDS-11483	34.3996315	-5.1363320	3.63	23.6	26.0
ZF-UDS-2622	34.2894516	-5.2698011	3.77	23.0	24.6
ZF-UDS-3112	34.2904282	-5.2620673	3.53	23.2	24.9
ZF-UDS-5418	34.2937546	-5.2269468	3.53	23.3	24.9
ZF-UDS-6119	34.2805405	-5.2171388	4.05	23.8	25.5
ZF-UDS-9526	34.3381844	-5.1661916	3.97	24.2	25.9
STACK	-	-	3.66	-	-
STAR FORMING					
ZF-CDFS-261	53.0826530	-27.8664989	3.40	23.2	24.2
ZF-CDFS-400	53.1025696	-27.8606110	4.10	24.3	25.1
ZF-CDFS-509	53.1167717	-27.8559704	3.95	24.2	25.1
ZF-COSMOS-12141	150.0815277	2.3637166	4.00	24.0	24.7
ZF-COSMOS-3784	150.1817627	2.2390490	3.58	22.9	23.9
ZF-UDS-11279	34.3843269	-5.1402941	3.72	25.0	26.6
ZF-UDS-4432	34.3581772	-5.2409291	3.76	23.8	24.5
ZF-UDS-4449	34.3409157	-5.2405076	3.84	23.1	24.4
ZF-UDS-4462	34.3408661	-5.2402906	3.92	23.0	24.0
ZF-UDS-5617	34.3407745	-5.2240300	4.17	24.5	26.0
ZF-UDS-8379	34.4104004	-5.1821156	3.77	23.8	25.2
ZF-UDS-8399	34.4105759	-5.1825032	3.44	24.4	25.3
ZF-UDS-8580	34.3544159	-5.1797152	4.07	23.7	24.6
ZF-UDS-9165	34.3225441	-5.1713767	4.06	23.4	24.2
STACK	-	-	3.84	-	-

[a] F160W, SNR and circularized KRON radius (r_{KRON}) crossmatched from 3D-HST (Skelton et al. 2014; van der Wel et al. 2014)

[d] *zspec* (Szokoly et al. 2004)

Table 4.2: Properties of 16 quiescent and 14 star forming galaxies continued

ID	H_{Galfit} [b] (mag)	SNR_{F160W} [a]	$M/10^{11}$ ($M_\odot$)	r_{KRON} [a] (")	$r_{1/2,maj}$ (")
QUIESCENT					
ZF-CDFS-209	24.3 ± 0.0	64.6	0.76	0.23	0.06 ± 0.01
ZF-CDFS-403	23.5 ± 0.0	118.0	1.15	0.22	0.12 ± 0.03
ZF-CDFS-4719	25.2 ± 0.1	33.5	0.45	0.23	0.12 ± 0.03
ZF-CDFS-4907	25.1 ± 0.1	38.2	0.40	0.28	0.08 ± 0.02
ZF-CDFS-5657	24.2 ± 0.1	26.7	0.76	0.33	0.52 ± 0.16
ZF-CDFS-617	23.5 ± 0.0	135.1	0.69	0.22	0.10 ± 0.02
ZF-COSMOS-13129	24.9 ± 0.1	10.8	1.78	0.46	0.52 ± 0.13
ZF-COSMOS-13172	24.4 ± 0.1	37.2	1.45	0.27	0.08 ± 0.02
ZF-COSMOS-13414	25.4 ± 0.1	14.0	0.44	0.32	0.20 ± 0.06
ZF-UDS-10684	25.2 ± 0.2	8.5	0.85	0.32	0.50 ± 0.17
ZF-UDS-11483	25.9 ± 0.2	8.9	1.02	0.35	0.11 ± 0.05
ZF-UDS-2622	24.5 ± 0.1	29.9	0.87	0.30	0.13 ± 0.03
ZF-UDS-3112	24.9 ± 0.1	25.7	0.43	0.30	0.07 ± 0.02
ZF-UDS-5418	24.9 ± 0.1	20.7	0.44	0.30	0.07 ± 0.02
ZF-UDS-6119	25.4 ± 0.2	10.6	0.55	0.32	0.26 ± 0.15
ZF-UDS-9526	25.8 ± 0.3	11.5	0.89	0.21	0.10 ± 0.05
STACK	-	-	0.81	-	-
STAR FORMING					
ZF-CDFS-261	24.5 ± 0.1	27.1	1.07	0.40	0.61 ± 0.14
ZF-CDFS-400	25.1 ± 0.2	23.9	0.52	0.33	0.24 ± 0.13
ZF-CDFS-509	25.0 ± 0.0	29.1	0.41	0.25	0.31 ± 0.06
ZF-COSMOS-12141	24.1 ± 0.2	18.8	0.45	0.34	0.81 ± 0.27
ZF-COSMOS-3784	23.8 ± 0.1	26.6	0.36	0.38	0.53 ± 0.13
ZF-UDS-11279	26.4 ± 0.3	4.5	0.46	0.32	0.15 ± 0.10
ZF-UDS-4432	24.2 ± 0.2	17.5	0.83	0.37	0.75 ± 0.39
ZF-UDS-4449	24.9 ± 0.1	17.2	0.41	0.35	0.44 ± 0.10
ZF-UDS-4462	24.0 ± 0.1	27.9	0.39	0.26	0.39 ± 0.09
ZF-UDS-5617	24.5 ± 0.3	6.3	0.42	0.37	2.33 ± 0.72
ZF-UDS-8379	25.2 ± 0.1	14.0	0.65	0.25	0.30 ± 0.07
ZF-UDS-8399	25.0 ± 0.1	11.9	0.43	0.23	0.69 ± 0.16
ZF-UDS-8580	24.7 ± 0.1	19.8	0.66	0.26	0.36 ± 0.08
ZF-UDS-9165	24.6 ± 0.1	33.8	0.68	0.31	0.11 ± 0.03
STACK	-	-	0.55	-	-

[a] F160W, SNR and circularized KRON radius (r_{KRON}) crossmatched from 3D-HST (Skelton et al. 2014; van der Wel et al. 2014)

[b] GALFIT and 3DHST magnitudes are consistent within 0.05 ± 0.03 mag on average, with dispersion 0.24.

Table 4.3: Properties of 16 quiescent and 14 star forming galaxies continued

ID	r_e (kpc)	b/a	n_{sercic}	A_v	24μm[c] (μJy)
QUIESCENT					
ZF-CDFS-209	0.27 ± 0.07	0.37 ± 0.08	4.00	0.3	-0.9 ± 3.5
ZF-CDFS-403	0.82 ± 0.18	0.85 ± 0.05	7.78 ± 0.94	0.8	99.8 ± 148.5†
ZF-CDFS-4719	0.60 ± 0.14	0.48 ± 0.08	1.88 ± 0.84	0.3	1.9 ± 3.4
ZF-CDFS-4907	0.56 ± 0.13	0.86 ± 0.12	3.28 ± 0.90	0.8	1.4 ± 3.6
ZF-CDFS-5657	3.22 ± 0.93	0.72 ± 0.11	4.45 ± 0.98	0.3	1.7 ± 3.8†
ZF-CDFS-617	0.55 ± 0.11	0.59 ± 0.03	4.00	0.3	86.3 ± 3.4†*
ZF-COSMOS-13129	2.15 ± 0.48	0.34 ± 0.08	0.56 ± 0.24	0.6	110.1 ± 10.2*
ZF-COSMOS-13172	0.49 ± 0.12	0.64 ± 0.13	3.94 ± 1.11	0.6	2.7 ± 7.6
ZF-COSMOS-13414	0.83 ± 0.29	0.34 ± 0.14	1.51 ± 1.00	0.2	7.1 ± 8.7
ZF-UDS-10684	2.42 ± 0.77	0.47 ± 0.18	4.63 ± 1.68	1.0	8.8 ± 12.8
ZF-UDS-11483	0.52 ± 0.25	0.43 ± 0.24	4.59 ± 2.01	1.0	1.8 ± 10.2
ZF-UDS-2622	0.76 ± 0.19	0.66 ± 0.10	4.00	0.9	12.2 ± 10.6
ZF-UDS-3112	0.39 ± 0.13	0.66 ± 0.19	4.00	0.0	-10.9 ± 10.6
ZF-UDS-5418	0.50 ± 0.14	0.83 ± 0.17	4.00	0.5	48.4 ± 10.6
ZF-UDS-6119	1.26 ± 0.75	0.49 ± 0.20	4.00	1.0	-12.5 ± 8.7
ZF-UDS-9526	0.39 ± 0.35	0.34 ± 0.24	2.03 ± 2.28	1.8	38.7 ± 8.7†*
STACK	0.85 ± 0.35	-	4.14 ± 0.71	-	-
STAR FORMING					
ZF-CDFS-261	3.54 ± 0.80	0.62 ± 0.06	1.21 ± 0.25	1.9	12.1 ± 4.4†
ZF-CDFS-400	1.45 ± 0.78	0.78 ± 0.11	3.40 ± 1.40	0.9	31.3 ± 3.6†*
ZF-CDFS-509	1.55 ± 0.32	0.52 ± 0.05	0.51 ± 0.17	1.0	-4.5 ± 4.1
ZF-COSMOS-12141	3.58 ± 1.09	0.40 ± 0.10	4.92 ± 1.35	1.1	0.9 ± 8.0
ZF-COSMOS-3784	3.40 ± 0.78	0.77 ± 0.10	1.88 ± 0.33	0.5	-2.4 ± 10.2
ZF-UDS-11279	0.96 ± 0.54	0.81 ± 0.23	1.00	2.2	29.3 ± 12.5
ZF-UDS-4432	3.61 ± 1.74	0.46 ± 0.11	4.27 ± 1.65	1.5	669.0 ± 10.7*
ZF-UDS-4449	1.90 ± 0.41	0.38 ± 0.07	0.23 ± 0.14	1.0	-
ZF-UDS-4462	2.09 ± 0.45	0.60 ± 0.08	1.69 ± 0.27	0.8	22.6 ± 9.4
ZF-UDS-5617	10.74 ± 3.30	0.45 ± 0.18	4.92 ± 1.51	1.3	9.5 ± 9.7
ZF-UDS-8379	1.50 ± 0.34	0.50 ± 0.09	0.52 ± 0.28	2.6	355.8 ± 25.0*
ZF-UDS-8399	2.28 ± 0.49	0.20 ± 0.05	0.14 ± 0.17	2.5	106.6 ± 25.1*
ZF-UDS-8580	1.82 ± 0.37	0.54 ± 0.05	0.18 ± 0.09	1.1	7.1 ± 8.4
ZF-UDS-9165	0.66 ± 0.14	0.72 ± 0.09	1.00	0.3	43.3 ± 10.1*
STACK	2.62 ± 1.15	-	2.17 ± 2.41	-	-

[c] †: X−ray detection (Xue et al. 2011); *: $L_{IR} > 7 \times 10^{12} L_{\odot}$.

we investigate the impact of PSF choice. We repeated the fitting using the hybrid PSF models of van der Wel et al. (2012) and find marginally larger (< 5%) sizes. In particular, for the smallest galaxies ($r_e < 0.20''$), we find a median $r_e/r_{e,PSFvdW} = 0.93 \pm 0.05$.

Errors on the individual measurements were calculated using a Monte Carlo procedure. After subtracting the best-fit GALFIT models from the sources, we shifted the residuals by a random number of pixels, added back the model and used this as input for GALFIT. Repeating this $> 200\times$ for each galaxy, errors were calculated as the 1σ variation on these measurements. We report our results in Tables 4.1, 4.2 and 4.3.

In the fits, the Sérsic index (n_{sersic}) was restricted to $0.1 < n_{sersic} < 8.0$. If n_{sersic} reached the extreme value 0.1 or 8.0, GALFIT was rerun while forcing $n_{sersic} = 1$ for star-forming and $n_{sersic} = 4$ for quiescent galaxies. These values correspond to the median n_{sersic} of galaxies with well-constrained fits and $\mathrm{SNR}_{F160W} > 15$.

At $z \sim 4$, this happens for $6/16 \Rightarrow 38\%$ quiescent and $2/14 \Rightarrow 14\%$ star-forming galaxies. To explore systematic effects introduced by the choice of profile, we set $n_{sersic} = 1.0$ or $n_{sersic} = 4.0$ for bright ($mag_{F160W}(AB) < 24.5$) and compact sources ($r_e < 0.20''$) and find on average $r_{e,n=1}/r_{e,n=4} = 0.80 \pm 0.13$, corresponding to a systematic uncertainty of 20%. We add this in quadrature to the uncertainties from the Monte Carlo procedure for each galaxy. Systematic biases of this level do not affect the main results. For comparison, van der Wel et al. (2012) derived typical systematic uncertainties on size of ~ 12% for faint F160W= $24-26$ and small $r_e < 0.3''$ galaxies.

As many galaxies have modest SNR, we tested the reliability of our measurements by a simulation, in which we inserted source models, convolved with the instrument PSF, in the F160W images. These have adopted magnitudes of $25 < mag_{F160W}(AB) < 26$ and size of $0.06 < r_e('') < 0.3$. We find $r_{e,out}/r_{e,in} = 0.97 \pm 0.05$, with $r_{e,in}$ and $r_{e,out}$ the input and output effective radii, showing that we can recover the sizes of faint compact sources without bias. As an additional test we determine the size distribution of point sources, by inserting PSFs in the images and measuring their size. We can constrain the size of bright objects to $0.01''$ at 95% confidence, which we adopt as a minimum uncertainty on the sizes.

We crossmatched our sample at $2 \leq z < 4.2$ with the size catalogs of van der Wel et al. (2014), based on the 3D−HST photometric catalogs (Skelton et al. 2014). We find that the sizes and Sérsic indices agree well, with a median $r_{e,ZFOURGE}/r_{e,3DHST} = 1.004 \pm 0.01$ and $n_{ZFOURGE} - n_{3DHST} = -0.012 \pm 0.058$.

We test for color gradients between rest-frame UV sizes and rest-frame optical sizes, using a rest-frame color and stellar-mass matched control sample

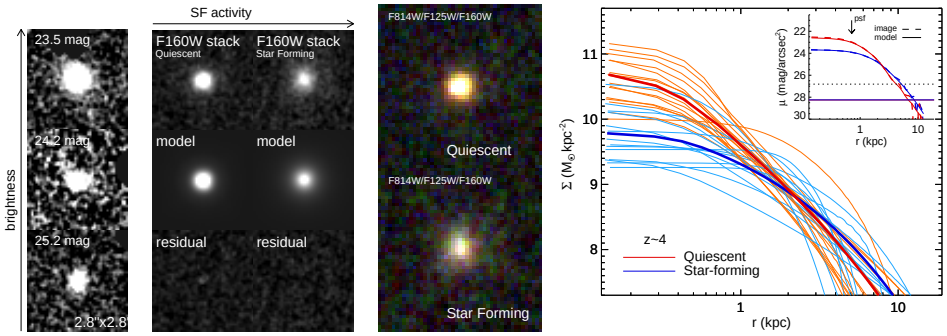

Figure 4.1: Left: Example galaxies at $z \sim 4$ of varying magnitude. Second: stacks of the quiescent and star forming subsamples, with the corresponding best-fit models and residuals after subtracting the models. Third: F814W/F125W/F160W stack color composites. Right: Stellar mass surface density profiles. Thin orange and blue lines represent individual measurements of quiescent and star-forming galaxies, respectively. Thick lines represent the stacks. The inset shows the surface brightness profiles of the stacks, with horizontal lines indicating 3σ brightness limits of 28.3 mag/arcsec^2, measured in annuli of $0.06''$ (0.43kpc) width at $r > 1''$. The background limit for individual galaxies (dotted line) is 26.8 mag/arcsec^2.

at $z \sim 3$. We find F160W (rest-frame 4000Å) sizes are $0 \pm 6\%$ and $6 \pm 11\%$ smaller than F125W (rest-frame 3000Å) sizes for star-forming and quiescent galaxies, respectively.

4.3.2 Stacking

We also measure the average sizes by stacking the background subtracted image stamps of the two subsamples, normalizing each by mean stellar mass. Neighbouring sources were masked. The final stacks were obtained by calculating the mean value at each pixel location of the image stamps.

We ran GALFIT using the same input parameters as for the individual galaxies. Errors were estimated by bootstrapping, i.e. randomly selecting galaxies, recreating the image stacks and rerunning GALFIT.

In Figure 4.1 we show the stacks and examples of individual galaxies. The stack of quiescent galaxies is redder than the stack of star-forming galaxies and has a more compact morphology. We also show stellar mass surface density profiles ($\Sigma(\mathrm{M}_\odot/\mathrm{kpc}^2) = M(< r)/(\pi r^2)$), obtained from the light profile measured in concentric apertures of radius r and assuming a constant mass-to-light ratio. For the stacked profiles we used the mean mass of the galaxies in each stack. They are consistent with the individual profiles within the un-

certainties, suggesting that the stack does not reveal an extended low surface brightness component, down to a surface brightness limit of 28.3 mag/arcsec^2.

4.3.3 Contamination by AGN

A substantial fraction of sources show signs of AGN activity either from X−ray detections or strong 24μm (rest-frame 5μm, tracing hot dust). As WFC3/F160W ($\lambda = 1.5396\mu$m) corresponds to rest-frame $2960 - 3500$Å (UV) at $3.4 \leq z < 4.2$, it could be that an AGN is dominating their central light, leading to small sizes of the single Sérsic fits.

In the quiescent sample we find four X−ray detected galaxies, two of which are spectroscopically confirmed type-II QSOs (Szokoly et al. 2004). Another has strong 24μm, which could either point towards dust-obscured star-formation or AGN activity. Several have small positive residuals after subtracting the best fit, suggesting the presence of a central point source. These 5/16 (31%) galaxies were re-fit with two components, a Sérsic model and a pointsource-like model (represented by a Gaussian with FWHM= 0.1pixels) to trace possible AGN light. In these models, the point source accounts for $4.3 - 68\%$ of the total light (with 57% and 68% for the type-II QSOs, but on average 6.2% for the remaining 3 AGN candidates). The average size of the Sérsic component increases by 1.5× (from a median $r_e = 0.13 \pm 0.12''$ to $r_e = 0.20 \pm 0.03''$).

Amongst the star-forming galaxies two are X−ray detected, and four are very bright at 24μm (L$> 7 \times 10^{12} L_\odot$ or $SFR > 1200 M_\odot/yr$). Re-fitting with a two-component model attributes $0.9 - 39.4\%$ of the light to a point source, while the extended component changes in size by 0.65× (from $r_e = 0.31 \pm 0.15''$ to $r_e = 0.19 \pm 0.02''$). We note that for the most extended sources, adding central light reduces the Sérsic index n_{Sersic} of the extended component, and can result in a smaller r_e.

We additionally estimated the possible AGN contribution from the galaxy SEDs. We first determine the best fitting powerlaw bluewards of rest-frame 0.35μm and at observed 8μm (Kriek et al. 2009). Then we fit the sum of the powerlaw and the original best-fit EAZY template to the data. The contribution of the AGN powerlaw template to F160W is $1.1 - 7.4\%$ for the 5 quiescent galaxies and $0.9 - 2.9\%$ for the 6 star-forming galaxies.

While the two-component fits and SEDs indicate that a point source contribution is probably small, the true contribution and its effect on the sizes remain unclear.

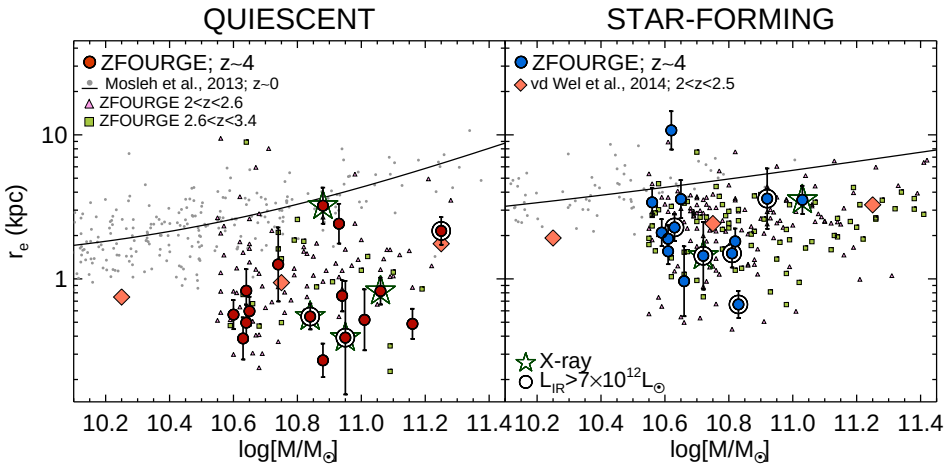

Figure 4.2: Circularized effective radius for galaxies at $z \sim 4$. In purple and green we show our control sample at $2 \leq z < 2.6$ and $2.6 \leq z < 3$ and in orange the median of van der Wel et al. (2014) at $2 < z < 2.5$. The black solid line is the $z \sim 0$ relation of Mosleh et al. (2013). X−ray detections and bright 24μm sources are indicated with stars and open circles. The median sizes are $r_e = 0.63 \pm 0.18$kpc (quiescent galaxies) and $r_e = 2.0 \pm 0.60$kpc (star-forming galaxies).

4.4 Results

We show the effective radius as a function of stellar mass in Figure 4.2. Quiescent galaxies at $z \sim 4$ are very compact, with a bootstrapped median size $r_e = 0.63 \pm 0.18$kpc. When we remove AGN we find a similar result: $r_e = 0.57 \pm 0.18$kpc.

Star-forming galaxies have $r_e = 2.0 \pm 0.60$kpc. They are $3.2 \pm 1.3\times$ larger than quiescent galaxies. Both samples have a large spread in size, with some almost as large as at $z \sim 0$, showing that at $z \sim 4$ the population is already very diverse. On average the sizes lie well below the $z \sim 0$ relation (Mosleh et al. 2013), by $6.0 \pm 1.7\times$ for quiescent and $2.4 \pm 0.7\times$for star-forming galaxies. Quiescent galaxies are also $1.9 \pm 0.7\times$smaller than at $2 \leq z < 2.2$.

In Figure 4.3 we show Sérsic index versus size for the $z \sim 4$ galaxies, and a sample at similar mass at $2 \leq z < 2.2$. Star-forming galaxies have smaller Sérsic index, with on average $n_{sersic} = 1.3 \pm 0.7$, compared to $n_{sersic} = 3.2 \pm 1.2$ for quiescent galaxies. The difference between the two populations is also clear from the stellar mass density profiles in Figure 4.1, with quiescent galaxies having steeper profiles and more centralized flux. In Figure 4.3 we also plot $\langle\Sigma\rangle_{max}$, defined as the average stellar mass density inside the radius where

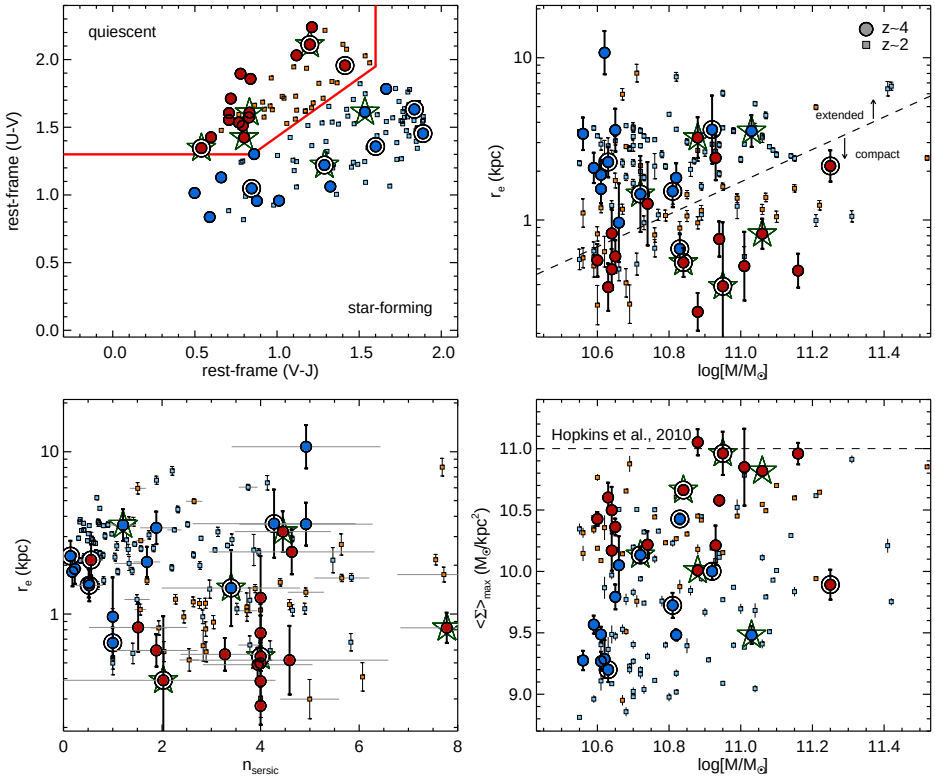

Figure 4.3: Top-left: UVJ diagram of $z \sim 4$ galaxies (symbols as in Figure 4.2). Small squares represent galaxies at $2.0 \leq z < 2.2$. Top-right: Stellar mass versus size. Bottom-left: Sérsic index versus size. Bottom-right: stellar mass versus maximum stellar mass density. The horizontal dashed line is the empirical limit of Hopkins et al. (2010). Only one $z \sim 4$ star-forming galaxy is compact. On average quiescent galaxies have smaller sizes, higher Sérsic indices and higher central densities than star-forming galaxies.

$\Sigma(M_\odot/kpc^2)$ falls of by a factor of two (Hopkins et al. 2010), with uncertainties from the Monte Carlo procedure described in section 4.3.1.

Quiescent galaxies at $z \sim 4$ have a median $\langle\Sigma\rangle_{max} = 3.3 \pm 1.1 \times 10^{10} M_\odot kpc^{-2}$, much higher ($\sim 10\times$) than for star-forming galaxies: $\langle\Sigma\rangle_{max} = 0.3 \pm 0.1 \times 10^{10} M_\odot$ kpc^{-2}, and more similar to $2 \leq z < 2.2$ quiescent galaxies: $\langle\Sigma\rangle_{max} = 1.7 \pm 0.3 \times 10^{10} M_\odot kpc^{-2}$.

When stacking we find $r_e = 0.85 \pm 0.35$kpc (quiescent) and $r_e = 2.6 \pm 1.2$kpc (star-forming), and Sérsic indices $n_{sersic} = 4.17 \pm 0.90$ and $n_{sersic} = 2.18 \pm 2.03$, respectively. The effective radius of the quiescent stack is slightly larger than the median of the individual galaxies, by $1.3 \pm 0.3\times$ at $< 1\sigma$ significance, but overall the results are consistent.

In Figure 4.4 we show the median sizes at the respective mean redshifts of the two subsamples. Comparing with lower redshift, they continue to follow a trend of decreasing size with increasing redshift. Our control sample of galaxies at $2 \leq z < 3.4$ with $10.5 \leq \log_{10}(M/M_\odot) < 11$ corresponds well with the results of van der Wel et al. (2014), which suggest the same trend.

We fit a relation of the form $r_e = A(1+z)^B$kpc at $0 < z < 4$, using the measurements of van der Wel et al. (2014) at $z < 2$. We find $r_e = 5.08 \pm 0.28(1+z)^{-1.44\pm0.08}$kpc for quiescent and $r_e = 6.02 \pm 0.28(1+z)^{-0.72\pm0.05}$kpc for star-forming galaxies. We note that our sample at $z \sim 4$ includes higher mass ($\log_{10}(M/M_\odot) \geq 11$) galaxies. If we remove the most massive galaxies, we find the same evolutionary relation.

To test for incompleteness for diffuse galaxies, we redshift a stellar-mass matched sample with $r > 2$kpc and $n_{sersic} < 2.5$ at $z \sim 2.5$ to $z = 3.7$ and find 70% completeness.

4.5 Discussion

Our results show that the galaxies at $z \sim 4$ in this study obey similar relations between size and star-forming activity as galaxies at lower redshift. Quiescent galaxies are compact, while star-forming galaxies are more extended and diffuse. The difference is also clear when selecting purely on size: if we define compactness as $r_e/(M/10^{11}M_\odot)^{0.75} < 1.5$ (van der Wel et al. 2014), 13/14 (93%) of massive compact galaxies would be classified as quiescent, and 13/16 (81%) of larger galaxies as star-forming (Figure 4.3).

The number density of compact, $\log_{10}(M/M_\odot) \geq 10.55$, quiescent galaxies at $z \sim 4$ is $1.8 \pm 0.8 \times 10^{-5} Mpc^{-3}$, increasing by $> 5\times$ between $3.4 \leq z < 4.2$ and $2 \leq z < 2.2$, towards $1.0 \pm 0.3 \times 10^{-4} Mpc^{-3}$. This suggests we are probing a key era of their formation, and we would expect to see their star-forming progenitors in abundance.

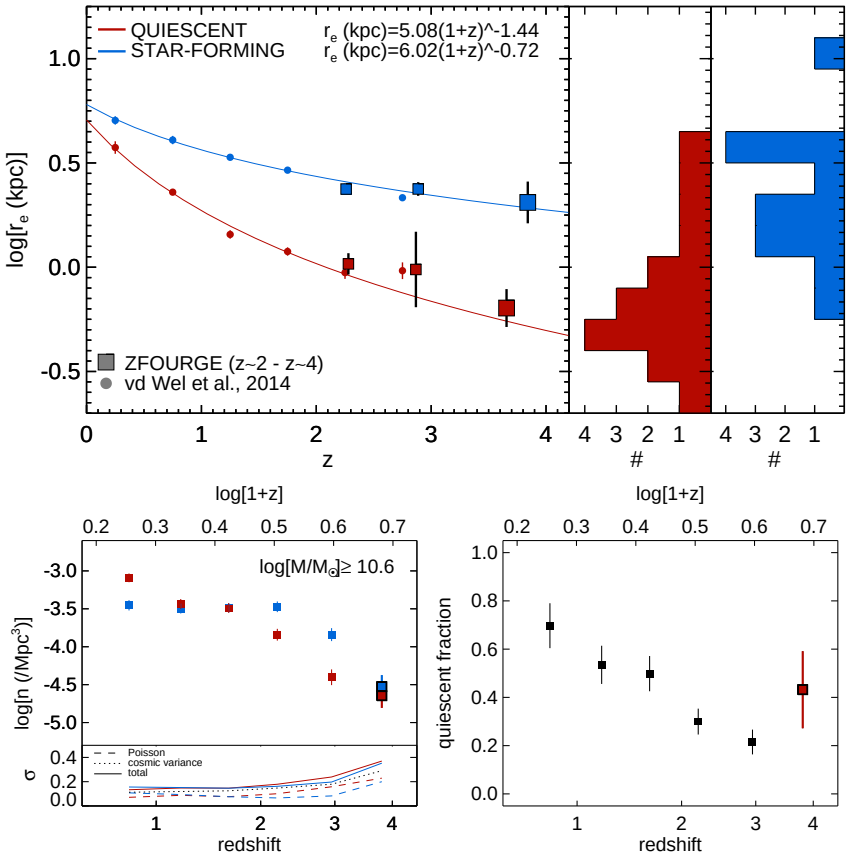

Figure 4.4: Top: Effective radius versus redshift for galaxies with $10.5 < \log_{10}(M/M_{\odot}) < 11.0$ at $2 \le z < 3.4$ (van der Wel et al. 2014) and $\log_{10}(M/M_{\odot}) \ge 10.55$ at $3.4 \le z < 4.2$ (filled squares). Quiescent galaxies follow $r_e = 5.08 \pm 0.28(1+z)^{-1.44\pm0.08}$kpc and star-forming galaxies $r_e = 6.02 \pm 0.28(1+z)^{-0.72\pm0.05}$kpc (solid curves). The histograms show the size distribution at $z \sim 4$. Bottom: Number density (left) and quiescent fraction (right), including galaxies without HST coverage. In the left panel we include the relative Poissonian uncertainties and the effect of cosmic variance. The total uncertainty on number density increases to 40% at $z \sim 4$.

Small effective radii for star-forming galaxies have been reported at $z = 2-3$ (Barro et al. 2014a,b; Nelson et al. 2014). They are rare in our sample: we find 1/14 with $r_e/(M/10^{11}M_\odot)^{0.75} < 1.5$. On average, star-forming galaxies at $z \sim 4$ are twice as large as quiescent galaxies at $z \sim 2$. If they are the direct progenitors of $z < 4$ compact quiescent galaxies, we expect them to be similar, not only in size, but also in Sérsic index and central surface density (Nelson et al. 2014). However, we find smaller n_{Sersic} for star-forming galaxies, while the central densities indicate that they must increase in $\langle\Sigma\rangle_{max}$ by $5-10\times$, to match the more cuspy profiles of $z = 2-4$ quiescent galaxies.

In a recent simulation, Wellons et al. (Ilustris; 2015) trace the evolution of galaxies to $z = 2$. They indeed identified two theoretical formation tracks: one in which a brief and intense central starburst prompted by a gas-rich major merger causes the galaxies' half-mass radius to decrease dramatically. The second is that of a more gradual but early formation, with small galaxy sizes due to the higher density of the universe. In the second case, nearly all of the stellar mass is in place at $z > 4$.

Comparing with the observations, we find that 19/44 of massive $z \sim 4$ galaxies are classified as quiescent, whereas all similarly massive galaxies in Illustris are still actively star-forming, with a typical $SFR = 100-200M_\odot/yr$. This level of star-formation is ruled out at $> 3\sigma$ by Herschel observations of the $z \sim 4$ quiescent galaxies (Straatman et al. 2014). At the same time, the fraction of compact galaxies in our sample is 47% ,versus ~ 20% in Illustris. Hence massive galaxies appear to quench their star-formation earlier and to be more compact than in simulations.

The paucity of compact star-forming galaxies at $z \sim 4$ and their large median rest-frame UV size is puzzling. At face value it suggests that the rapid increase in number density of compact quiescent galaxies cannot be explained by simple shutdown of star-formation in typical star-forming galaxies of similar stellar mass. A possible solution is a rapidly forming dense core, i.e. a central starburst. Then the chance to observe the progenitors in our sample is small, as it is proportional to the duration of the main star-forming episode. For example, if compact cores of $2 \leq z < 2.2$ quiescent galaxies formed at random times between $2.5 < z < 6$, with a typical 100Myr central starburst duration, their predicted number density at $z \sim 4$ would be $\sim 6 \times 10^{-6}\mathrm{Mpc}^{-3}$. The observed number density of compact star-forming galaxies is $1.4 \pm 1.4 \times 10^{-6}\mathrm{Mpc}^{-3}$: smaller, but in a similar range given the large uncertainties.

We note that the remarkably high fraction of quiescent galaxies at $z \sim 4$ (Figure 4.4) is still uncertain. Current limits on the average dust-obscured SFR are weak ($< 75M_\odot/yr(3\sigma)$, Straatman et al. 2014), hence some of the quiescent galaxies could be star-forming. Cosmic variance is significant (~ 30%).

Highly obscured massive star-forming galaxies might also be missed by near-IR surveys (e.g. Daddi et al. 2009; Caputi et al. 2012), although the abundance and redshift distribution of such galaxies is still very uncertain. Finally, extended ($r > 3$kpc) galaxies with small n_{sersic} and low surface brightness are more difficult to detect than compact galaxies (e.g. Trujillo et al. 2006).

We caution that the light profiles measured here may not be representative of the stellar mass distribution due to color gradients, with rest-frame UV sizes larger than rest-frame optical sizes. This would imply that the size evolution is stronger. However, using control sample at $z \sim 3$, we find no difference between UV and optical, consistent with van der Wel et al. (2014), who show this effect is $\lesssim 10\%$ at $z \sim 2$ and decreasing with redshift.

Galaxy sizes may also be overestimated if dust is obscuring a central starburst. Submm sizes of obscured starbursting galaxies could be small: $<$ 1kpc (e.g. Ikarashi et al. 2014; Simpson et al. 2015). A direct comparison of ALMA submm and rest-frame optical/UV morphologies for the same objects with measured stellar mass will reveal the effect of dust obscuration on UV/optically measured galaxy sizes.

4.6 Acknowledgements

This research was supported by the George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, the National Science Foundation grant AST-1009707 and the NL-NWO Spinoza Grant. Australian access to the Magellan Telescopes was supported through the National Collaborative Research Infrastructure Strategy of the Australian Federal Government. GGK is supported by an Australian Research Council Future Fellowship FT140100933. KEW is supported by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by Oak Ridge Associated Universities through a contract with NASA. We thank Arjen van der Wel, Darren Croton, Duncan Forbes and Alister Graham for useful discussions.

Bibliography

Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2014a, *ApJ*, 791, 52

Barro, G., Trump, J. R., Koo, D. C., et al. 2014b, *ApJ*, 795, 145

Caputi, K. I., Dunlop, J. S., McLure, R. J., et al. 2012, *ApJL*, 750, L20

Chen, H.-W., & Marzke, R. O. 2004, *ApJ*, 615, 603

Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, 626, 680

Daddi, E., Dannerbauer, H., Stern, D., et al. 2009, *ApJ*, 694, 1517

Damjanov, I., McCarthy, P. J., Abraham, R. G., et al. 2009, *ApJ*, 695, 101

Dekel, A., & Burkert, A. 2014, *MNRAS*, 438, 1870

Dekel, A., Birnboim, Y., Engel, G., et al. 2009, *Nature*, 457, 451

Fontana, A., Santini, P., Grazian, A., et al. 2009, *A&A*, 501, 15

Giacconi, R., Zirm, A., Wang, J., et al. 2002, *ApJS*, 139, 369

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, 197, 35

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, *ApJS*, 207, 24

Hopkins, P. F., Hernquist, L., Cox, T. J., Keres, D., & Wuyts, S. 2009, *ApJ*, 691, 1424

Hopkins, P. F., Murray, N., Quataert, E., & Thompson, T. A. 2010, *MNRAS*, 401, L19

Ikarashi, S., Ivison, R. J., Caputi, K. I., et al. 2014, ArXiv e-prints, arXiv:1411.5038

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, 197, 36

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221

Labbé, I., Huang, J., Franx, M., et al. 2005, *ApJL*, 624, L81

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, 379, 1599

Mancini, C., Matute, I., Cimatti, A., et al. 2009, *A&A*, 500, 705

Marchesini, D., Whitaker, K. E., Brammer, G., et al. 2010, *ApJ*, 725, 1277

Mo, H. J., Mao, S., & White, S. D. M. 1998, *MNRAS*, 295, 319

Mosleh, M., Williams, R. J., & Franx, M. 2013, *ApJ*, 777, 117

Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, 777, 18

Nelson, E., van Dokkum, P., Franx, M., et al. 2014, *Nature*, 513, 394

Papovich, C., Labbé, I., Quadri, R., et al. 2015, *ApJ*, 803, 26

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097

Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, *PASP*, 125, 654

Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, 172, 1

Sersic, J. L. 1968, Atlas de galaxias australes

Shen, S., Mo, H. J., White, S. D. M., et al. 2003, *MNRAS*, 343, 978

Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2015, *ApJ*, 799, 81

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, 214, 24

Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJL*, 787, L36

Stefanon, M., Marchesini, D., Rudnick, G. H., Brammer, G. B., & Whitaker, K. E. 2013, *ApJ*, 768, 92

Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, *ApJL*, 783, L14

Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, *ApJS*, 155, 271

Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, *ApJ*, 650, 18

van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, *ApJS*, 203, 24

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, *ApJ*, 788, 28

van Dokkum, P. G., Franx, M., Kriek, M., et al. 2008, *ApJL*, 677, L5

Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, *MNRAS*, 449, 361

Wiklind, T., Dickinson, M., Ferguson, H. C., et al. 2008, *ApJ*, 676, 781

Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, 691, 1879

Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2011, *ApJS*, 195, 10

5

ZFIRE: the evolution of the stellar mass Tully-Fisher relation to redshift $2.0 < z < 2.5$ with MOSFIRE _

Abstract

Using observations made with MOSFIRE on Keck I as part of the ZFIRE survey, we present the stellar mass Tully-Fisher relation at $2.0 < z < 2.5$. The sample was drawn from a stellar mass limited, K_s–band selected catalog from ZFOURGE over the CANDELS area in the COSMOS field. We model the shear of the Hα emission line (rest-frame 6564.614Å) to derive rotational velocities at 2.2× the scale radius of an exponential disk ($V_{2.2}$). We correct for the blurring effects of a two-dimensional PSF and the fact that the MOSFIRE PSF is better approximated by a Moffat than a Gaussian, which is typically assumed. We find for the Tully-Fisher relation at $2.0 < z < 2.5$ that $\log V_{2.2} = (2.19 \pm 0.049) + (0.247 \pm 0.094)(\log M/M_{\odot} - 10)$ and infer an evolution of the zeropoint of $\Delta \mathrm{M}/M_{\odot} = -0.26 \pm 0.14$ dex compared to $z = 0$ when adopting a fixed slope. We also derive the alternative kinematic estimater S_{05}, and find $\log S_{05} = (2.10 \pm 0.033) + (0.228 \pm 0.059)(\log M/M_{\odot} - 10)$, with an evolution of $\Delta \mathrm{M}/M_{\odot} = -0.54 \pm 0.13$ dex compared to $0.1 < z < 1.2$. We investigate and review various systematics, ranging from PSF effects, projection effects, and systematics related to stellar mass derivation and selection biases. We find that discrepancies between literature values are reduced when taking these into account. After correction of the observations, we find a gradual evolution in the Tully-Fisher stellar mass zeropoint from $z = 0$ to $z = 2.5$. This corresponds reasonably well with the predictions from semi-analytic models.

5.1 Introduction

A major goal for galaxy evolution models is to understand the interplay between dark matter and baryons. In the current ΛCDM paradigm, galaxies are formed as gas cools and accretes into the centers of dark matter haloes. The gas maintains its angular momentum, settling in a disk at the center of the gravitational potential well (Fall & Efstathiou 1980) where it forms stars. This process can be disrupted by galaxy mergers, gas inflows, AGN and star formation feedback, which can affect the shape, star-formation history and kinematics of galaxies (e.g. Hammer et al. 2005).

From studies at $z = 0$ of the kinematic properties of disk galaxies a correlation has emerged between disk rotational velocity and, initially, luminosity. This relation is now named the Tully-Fisher relation, first reported by Tully & Fisher (1977), and originally used as a distance indicator. At $z = 0$ the Tully-Fisher relation is especially tight if expressed in terms of stellar mass instead of luminosity (Bell & de Jong 2001). If studied at high redshift, it can be an important test of the mass assembly of galaxies over time, as it describes the relation between angular momentum and stellar mass, and the conversion of gas into stars versus the growth of dark matter haloes by accretion (e.g. Fall & Efstathiou 1980; Mo et al. 1998). With the increasing succes of multi-wavelength photometric surveys to study galaxy evolution, much insight has already been obtained into the structural evolution of galaxies to high redshift (e.g. Franx et al. 2008; van der Wel et al. 2014; Straatman et al. 2015), and their stellar mass growth and star-formation rate histories (e.g. Whitaker et al. 2012; Tomczak et al. 2014, 2015). The study of galaxy kinematics at $z > 1$ has been lagging behind, because of the faint magnitudes of high redshift galaxies and the on-going development of sensitive near-IR multiobject spectrographs needed for efficient follow-up observations.

In the past few years, studies of the Tully-Fisher relation at $0 < z < 1$ were performed with the multiplexing optical spectrographs DEIMOS on Keck I (Kassin et al. 2007; Miller et al. 2011) and LRIS on Keck II (Miller et al. 2012), and optical Integral Field Unit (IFU) spectrographs such as VLT/GIRAFFE (Puech et al. 2008), but beyond $z > 1$ progress has been comparitively slow because of the reliance on mostly single-object integral field spectrographs, such as SINFONI (Cresci et al. 2009; Gnerucci et al. 2011; Vergani et al. 2012) on the VLT. These studies resulted in contradictive estimates of a potential evolution of the stellar mass zeropoint of the Tully-Fisher relation with redshift. For example, studies by Puech et al. (2008); Vergani et al. (2012); Cresci et al. (2009) and Gnerucci et al. (2011) indicate evolution already beyond $z > 0.6$. At $z = 0.6$ this amounts to $\Delta M/M_{\odot} \sim 0.3$ dex (Puech et al. 2008). At $z = 2.2$

$\Delta M/M_{\odot} \sim 0.4$ dex (Cresci et al. 2009) and at $z = 3$ $\Delta M/M_{\odot} \sim 1.3$ dex (Gnerucci et al. 2011). At the same time Miller et al. (2011, 2012) find no significant evolution up to $z = 1.7$. The latter result, in combination with the strong evolution observed at $z = 2.2$ and $z = 3$, would suggest that star-forming galaxies very rapidly establish the dynamical state typical at low redshift in the period just before $z = 1.7$.

Part of the inferred evolution however, or lack thereof, could be explained by selection bias, for example by preferentially selecting the most dynamically evolved galaxies at each redshift. This acts as a progenitor bias, (van Dokkum & Franx 2001), where the high redshift sample is an increasingly biased subset of the true distribution, leading to an underestimate of the evolution. Dynamically evolved galaxies could make up only a small fraction of the total population at high redshift, as irregular, dusty and dispersion dominated galaxies become more common towards higher redshifts (e.g. Abraham & van den Bergh 2001; Kassin et al. 2012; Spitler et al. 2014). Similarly, previous surveys at the highest redshift at $z > 2$ tend to be biased towards the less dust-obscured or blue star-forming galaxies, such as Lyman Break galaxies, and often required previous rest-frame UV selection or a spectroscopic redshift from optical spectroscopy (e.g. Förster Schreiber et al. 2009; Gnerucci et al. 2011). As a consequence these samples may not be representative of massive galaxies at high redshift, which are more often reddened by dust-obscuration (e.g. Reddy et al. 2005; Spitler et al. 2014).

The different results between these studies could also be due to systematics arising from the different methodologies used to derive stellar mass, rotational velocity, and the different types of spectral data (one-dimensional long-slit spectra versus two-dimensional IFU data). As Miller et al. (2012) note, a striking discrepancy exists between their long-slit results (no evolution) and IFU studies by Puech et al. (2008); Vergani et al. (2012) and Cresci et al. (2009) ($\Delta M/M_{\odot} = 0.3-0.4$ dex). Sample size may also play a role: the highest redshift studies are based on small samples of only 14 galaxies at $z = 2.2$ (Cresci et al. 2009) and 11 galaxies at $z = 3$ (Gnerucci et al. 2011).

A non-evolving Tully-Fisher relation would be a puzzling result, as the average properties of galaxies evolve strongly with redshift. For example, the average star-formation rate of star-forming galaxies at fixed stellar mass tends to increase with redshift (e.g. Tomczak et al. 2015), as does their gas fraction (e.g. Papovich et al. 2015). At the same time their average size tends to be smaller (e.g. van der Wel et al. 2014), which would by itself imply higher velocities at fixed stellar mass.

It is clear that more studies with larger numbers of galaxies are needed to shed light on the observationally key epoch at $z \sim 2$. In this study we use new

spectra of galaxies at $2.0 < z < 2.5$ from the ZFIRE survey (Nanayakkara; in prep). These were obtained from the newly installed MOSFIRE instrument on Keck I, a sensitive near-IR spectograph that allows batch-observations of large numbers of galaxies at the same time. The primary aim of ZFIRE is to spectroscopically confirm and study galaxies in two high redshift cluster clusters, one in the UDS field (Lawrence et al. 2007) at $z = 1.62$ (Papovich et al. 2010) and one in the COSMOS field (Scoville et al. 2007) at $z = 2.095$ (Spitler et al. 2012; Yuan et al. 2014). However, ZFIRE also targets many foreground and background galaxies at redshifts $1.5 < z < 4.0$. In this paper, we use the rich data set over the COSMOS field to study the Tully-Fisher relation at $2.0 < z < 2.5$. Our aim is to provide improved constraints on the evolution of the stellar mass Tully-Fisher relation with redshift.

In Section 5.2 we describe our data and sample of galaxies, in Section 5.3 we describe our analysis, in Section 5.4.2 we derive the Tully-Fisher relation at $2 < z < 2.5$ and in Section 5.5 we discuss our results in an evolutionary context. Throughout, we use a standard cosmology with $\Omega_\Lambda = 0.7$, $\Omega_\mathrm{m} = 0.3$ and $H_0 = 70$ km/s/Mpc. At $z = 2.2$ one arcsecond corresponds to 8.3 kpc.

5.2 Observations and selections

5.2.1 Observations

Spectroscopic data

This study makes use of data obtained with the Multi-Object Spectrometer for InfraRed Exploration (MOSFIRE; McLean et al. 2010) on Keck-I on Mauna Kea in Hawaii. The observations over COSMOS were carried out in 6 pointings with a $6.1' \times 6.1'$ field of view. The observations were conducted on December 24-25, 2013 and February 10-13, 2014. Galaxies were observed in 8 masks in the $K-$band, which covers $1.93 - 2.45\mu$m, and can be used to measure Hα and [NII] emission lines for galaxies at $z \sim 2$. Two $H-$band masks were also included in the observations. The $H-$band coverage is $1.46 - 1.81\mu$m, overlapping with $H\beta$ and [OIII]. For this work, we limit the analysis to the Hα (rest-frame 6564.614Å) emission line data in the $K-$band. Further details on the $H-$band masks can be found in Nanayakkara et al (in prep).

The total exposure time was 2 hours for each $K-$band mask. A $0.7''$ slit width was used, yielding spectral resolutions $R \sim 3500$. At $z = 2.2$, the median redshift of the sample of galaxies in this study, this corresponds to 27 km/s per pixel. The seeing conditions were $0.65 - 1.10''$, with a median of $0.7''$. We used a standard two-position dither pattern (ABBA). Before and after science

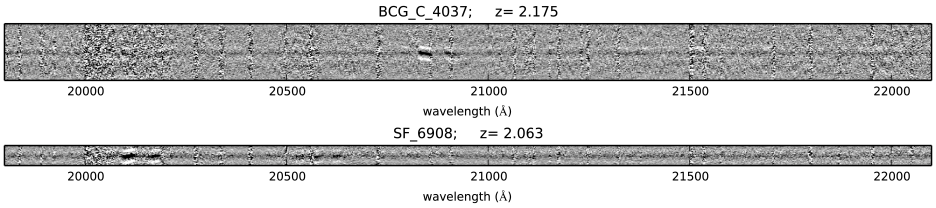

Figure 5.1: Two example Keck MOSFIRE spectra (inverted grayscale) at $z = 2.175$ and $z = 2.063$, with H$\alpha\lambda$6565 clearly visible at $\lambda = 20843.2$Å (top) and $\lambda = 20109.6$Å (bottom). Other lines are visible as well, most notably [NII]$\lambda\lambda$6550,6585 and [SII]$\lambda\lambda$6718,6733.

target exposures, we measured the spectrum of an A0V type standard star in 0.7″ slits to be used for telluric corrections and standard stars to be used for flux calibration in a slit of width 3″ to minimize slit loss. Each individual mask also contained a star for monitoring purposes, such as measuring the seeing conditions.

The raw data were reduced using the publicly available data reduction pipeline[1] developed by the MOSFIRE instrument team, resulting in two-dimensional spectra that were background subtracted, rectified and wavelength calibrated to vacuum wavelengths, with a typcial residual of < 0.1Å (Nanayakkara et al, in prep). To make up for the lack of skylines at the red end of the K–band, we used both night sky lines and and a Neon arc lamp for wavelength calibration.

Based on the standard star, we applied a telluric correction and flux calibration to the two-dimensional spectra, similar to the procedure used by (Steidel et al. 2014), and using our own custom IDL routines. The uncertainty on the absolute flux calibration is 8% and the uncertainty on the absolute wavelength calibration is 50 km/s (Nanayakkara; in prep).

In Figure 5.1 we show two example spectra at $z = 2.175$ and $z = 2.063$, with strong Hα emission at observed frame $\lambda = 20843.2$Å and $\lambda = 20109.6$Å, respectively. For both spectra the line profile exhibits the characteristic shape associated with rotation along the line of sight, where light from one half of the galaxy is relatively blueshifted due to motion towards the observer and the other half is relatively redshifted, with a turnover in the middle. Other lines are visible in the spectrum as well, most notably [NII]$\lambda\lambda$6550,6585 and [SII]$\lambda\lambda$6718,6733.

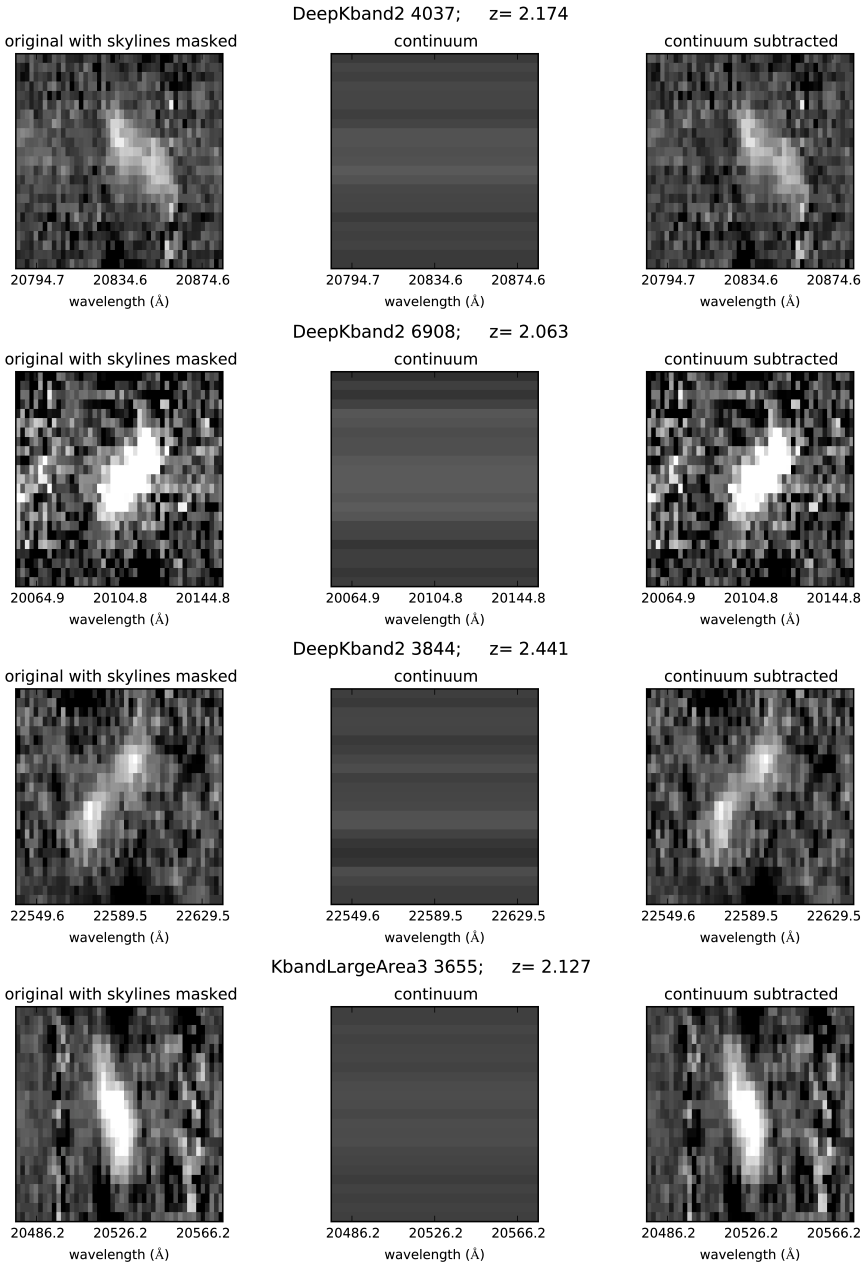

Figure 5.2: Continuum subtraction. Shown here are four examples from the sample. Left panels: the original spectral image stamps with Hα emission lines. Middle panels: the estimated continuum. Right: the spectral image stamps with the contiuum subtracted.

Continuum subtraction

From each two-dimensional spectrum we extracted spectral image stamps of 300Å wide centered on the Hα emission lines. Night sky emission was masked using the publicly available night sky spectra taken during May 2012 engineering, at wavelengths where the sky spectrum exceeds 10^{-24} ergs/s/cm^2/arcsec^2. We also masked 40Å wide boxes centered on the Hα line and the [NII] doublet. We subtracted the continuum using the following method: for each pixel row (one row corresponding to a one-dimensional spectrum with a length of 300Å) we determined the median flux and the standard deviation. Next we iteratively rejected pixels at $> 2.5\sigma$ from the median and recalculated both values. We repeated this a total of three times. The final median flux was our estimate of the continuum in that particular pixel row, which was then subtracted accordingly. In Figure 5.2 we show examples of the spectral image stamps, the estimated continuum and the continuum subtracted stamps, within a smaller wavelength range (100Å) and without the Hα line masked, for clarity. We will use the 100Å spectral image stamps in the remainder of our analysis.

PSF determination

At $z < 2$ galaxies are generally small ($R_e < 1.0''$; van der Wel et al. 2014), so the PSF needs to be properly characterized. Not only the FWHM of the PSF needs to be tracked, but even the detailed shape of the PSF can have a noticeable effect on the smoothing of the Hα line and its rotation profile. A simple Gaussian is often assumed, but this leads to underestimating the shear of the emission line – and hence the velocity – if the true PSF has stronger wings. Because the Tully-Fisher relation is very steep (e.g. Bell & de Jong 2001; Reyes et al. 2011), a small change in velocity could lead to significant offsets.

We first attempted to derive the PSF from the collapsed spectra of the monitor stars, which received the same exposure as the galaxies in the masks. The collapsed spectra were obtained by averaging over the flux in the wavelength direction, after masking skylines. The intensity profile had a very steep profile, which was well fit by a Gaussian profile. Although adopting a Gaussian profile is common (e.g. Kriek et al. 2015), this was unexpected, because the MOSFIRE PSF in deep K_s–band imaging (Marchesini; private communication) clearly has strong wings, which are better fit with a Moffat profile (see Figure 5.3). Even small wings are important, because the effect of the PSF on convolution does not scale with the amount of flux in the wings, but with the

[1] http://www2.keck.hawaii.edu/inst/mosfire/drp.html

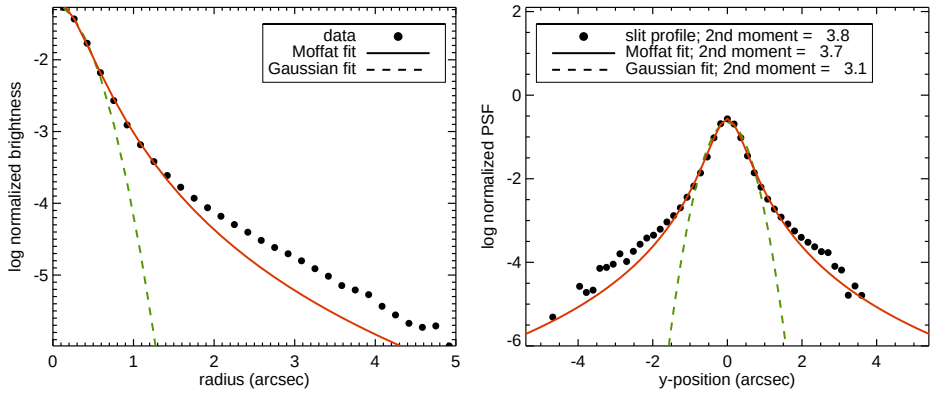

Figure 5.3: Left: surface brightness profile of the two-dimensional K_s–band image PSF (dots) as a function of radius, with the best-fit Moffat (solid red line) and Gaussian (dashed green line). The Gaussian is quite steep, whereas the Moffat gives a better approximation of the flux at large radii. Right: a simulated one-dimensional spectral PSF, obtained from integrating the two-dimensional K_s–band PSF and the best-fits in a $0.7''$ virtual slit. The second order moment of the Moffat is close to that of the actual PSF, but that of the Gaussian is much smaller.

second order moment of the PSF (Franx et al. 1989), which is given by

$$F_2 = \int (r-\mu)^2 f(r) dr \tag{5.1}$$

for a general function $f(r)$ centered on $r = \mu$. Even a few percent flux in the wings can have a significant effect, due to the r^2 weighting. For illustration, we calculate the second moment for a simulated spectral PSF derived from a deep MOSFIRE image at FWHM= $0.6''$ seeing. The image PSF was created by median stacking 5 unsaturated bright stars, after background subtraction and normalization. We then measured the brightness profile of the PSF as a function of radius and fitted a Moffat and a Gaussian function, as shown in the left panel of Figure 5.3. To reproduce the one-dimensional spectral PSFs, we integrated the two-dimensional image PSF and its two model fits within a $0.7''$ virtual slit. Finally, we calculated the second order moments. As shown in the right panel of Figure 5.3, the true PSF ($F_2 = 3.8$) is severely underestimated by a Gaussian approximation ($G_2 = 3.1$), whereas a Moffat fit produces good correspondence ($M_2 = 3.7$).[2]

[2]Note, to avoid noise amplifaction at large radii due to the r^2 weighting, we evaluate the second order moment at $r < 2.6''$. The Gaussian is scaled up by 12% for a consistent comparison to a Moffat in one dimension.

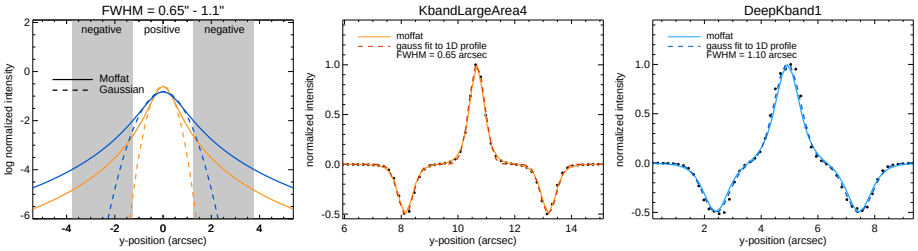

Figure 5.4: Examples of spatial profiles of MOSFIRE PSFs. The solid and dashed curves are theoretically derived Moffat and Gaussian intensity profiles, respectively. They are shown at logarithmic scale in the left panel. A Moffat is a good representation of the original MOSFIRE PSF, but sky subtraction in the reduction process leaves negative imprints on each sides, which will subtract the strong wings. This makes the reduced PSF appear Gaussian. This is illustrated by the two examples of spatial profiles of monitor stars in the middle and right panels, with best and worst seeing, respectively. The black datapoints represent the star spectra collapsed in the wavelength direction. Overplotted are the simulated theoretical PSFs, showing that Moffat PSFs are now nearly indistinguishable from Gaussians.

Clearly it is important to account for the flux in the wings of the PSF. However, it turns out to be rather difficult to reconstruct the true shape of the PSF accurately from the spatial profile of a monitor star spectrum. The reason is that standard reduction of the ABBA dither pattern results in one positive and two negative imprints each 2.52″ apart, meaning the PSF wings are largely subtracted out and the resulting profile is too steep. The problem is seeing dependent and becomes worse if the seeing is larger. We therefore proceeded to reconstruct the true PSF separately for every mask (with seeing varying from 0.65 to 1.1″). The proces is illustrated in Figure 5.4.

As the central region of the PSFs are still well approximated by a Gaussian, we used Gaussian fits to the collapsed spectra of the monitor stars to characterize the seeing FWHM for each of the 8 K–band masks. We then reconstructed the approximate true PSF by first integrating a two-dimensional Moffat ($\beta = 2.5$) PSF over the width of a 0.7″ wide virtual slit and subtracting 1/2 times the intensity offset by 2.52″ on either side to simulate the reduction process. Because the FWHM of a Gaussian fit to the resulting spectral PSF is 12% broader than the original Moffat FWHM, we scaled the FWHM of the two-dimensional Moffat to match the simulated spectral PSF to the observations. Figure 5.4 illustrates two extreme cases of best and worst seeing.

We verified the effect of using a Gaussian or Moffat profile in our modelling by calculating rotational velocities using either the Moffat PSFs derived

above, or Gaussian fits to the collapsed star spectra. The mean velocity is 4% smaller if a Gaussian is assumed, with up to 15% effects for some individual cases.

5.2.2 Target sample selection

The primary ZFIRE sample was designed to spectroscopically confirm a large cluster of galaxies at $z = 2.095$ (Spitler et al. 2012; Yuan et al. 2014) within the COSMOS field (Scoville et al. 2007). The sample was optimized by focusing mostly on near-IR bright star-forming galaxies, with strong expected signatures such as Hα emission. Star-forming galaxies as part of the cluster were selected based on their rest-frame $U-V$ and $V-J$ colors, with photometric redshifts between $2.0 < z < 2.2$. K-band magnitudes of $K < 24$ were priority sources, but fainter sources could be included as well. Non star-forming galaxies were prioritized next and lastly, field galaxies (not necessarily at the cluster redshift) could be used as fillers for the mask. In total 187 unique sources were listed for $K-$band observations. 36 of these were observed in two different masks and 2 in three different masks, leading to a total of 227 spectra.

Spectroscopic targets were originally obtained from the photometric redshift catalogs of the FourStar (Persson et al. 2013) Galaxy Evolution Survey (ZFOURGE; Straatman et al., submitted). The ZFOURGE catalogs were derived from ultra-deep near-IR K_s-band imaging (~ 25.5 mag). FourStar has a total of 6 near-IR medium bandwidth filters (J_1, J_2, J_3, H_s, H_l), that accurately sample the rest-frame 4000Å/Balmer break at redshifts $1.5 < z < 4$. We combined these with a wealth of already public multiwavelength data at $0.3-24\mu$m to derive photometric redshifts, using the EAZY software (Brammer et al. 2008). These redshifts were used as a prior for the MOSFIRE masks. The typical redshift uncertainty is $1-2\%$ for galaxies at $1.0 < z < 2.5$ (Straatman et al., submitted).

For this work we make use of the ZFOURGE stellar masses. These were calculated by fitting Bruzual & Charlot (2003) stellar template models, using the software FAST (Kriek et al. 2009), assuming a (Chabrier 2003) initial mass function, exponentially declining star formation histories, solar metallicities and a Calzetti et al. (2000) dust law. Galaxy sizes, axis-ratios and position angles are obtained from the size catalog of galaxies from the 3D-HST/CANDELS survey (van der Wel et al. 2014; Skelton et al. 2014). These were crossmatched to ZFOURGE by looking for matches within $< 0.7''$. The sizes were derived by fitting two-dimensional Sérsic (Sersic 1968) surface brightness profiles to HST/WFC3/F160W images, using the software GALFIT

(Peng et al. 2010).

From the original $N = 227$ ZFIRE K_s-band sample, we first selected 150 unique galaxies with $2.0 < z_{spec} < 2.5$, where we used spectroscopic redshifts derived from one-dimensional collapsed spectra (Nanayakkara in prep). Using the F160W position angles, we determined offsets with respect to the MOSFIRE masks: $\Delta\alpha = PA - \alpha_{mask}$, with PA the position angle of the major axis of the galaxy and α_{mask} the slit angle from the mask. If the slit was rotated by 180°, we first subtracted 180° and then determined $\Delta\alpha$. We refined the sample by selecting only galaxies with $|\Delta\alpha| < 40°$, resulting in a sample of 65 galaxies. Some were included in more than one mask, and we have 87 spectra in total that follow these criteria. The Hα emission was inspected by eye for contamination from night sky emission, and we only kept those instances that were largely free from skylines, removing 19. Out of the remaining 68 spectra, 24 have very low SNR and were also omitted. We also looked for signs of AGN, by crossmatching with radio and X-ray catalogs. This revealed one AGN, which we removed. The final high quality sample contains 43 spectra of 38 galaxies, and these form the basis for the kinematic analysis which we discuss next.

5.3 Analysis

5.3.1 Hα rotation model

We modeled the rotation curves by fitting two-dimensional (λ, r) intensity models to the spectral image stamps containing the Hα emission. We used the empirically motivated arctan function to model the velocity curve (Courteau 1997; Willick 1999; Miller et al. 2011):

$$v(r) = V_0 + \frac{2}{\pi} V_a \arctan\left(\frac{r - r_0}{r_t}\right) \tag{5.2}$$

with $v(r)$ the velocity at radius r, V_0 the central velocity, V_a the asymptotic velocity, r_0 the dynamic center and r_t the turnover, or kinematic, scale radius. r_t is a transitional point between the rising and flattening of the arctan curve.

For relatively small proper motion if viewed on a cosmological scale, we can express the velocity as function of the wavelength difference with respect to the central wavelength λ_0 as:

$$\frac{v}{c} = \frac{\Delta\lambda}{\lambda_0} = \frac{\lambda - \lambda_0}{\lambda_0} \tag{5.3}$$

Therefore we initially fit our model in wavelength space, and then afterwards convert the offset in λ to velocity. In terms of wavelength, Equation 5.2

becomes:

$$\lambda(r) = \lambda_0 + \frac{2}{\pi}\lambda_a \arctan\left(\frac{r - r_0}{r_t}\right) \tag{5.4}$$

We also model the spatial intensity of the emission, assuming an exponential disk:

$$I(r) = I_0 \exp\left[\frac{(r - r_0)}{R_s}\right] \tag{5.5}$$

with $I(r)$ the intensity at radius r and I_0 the central intensity. r_0 is the same in Equations 5.2, 5.4 and 5.5, and the coordinates (λ_0, r_0) represent the velocity centroid of the galaxy in Hα. R_s is the scalelength of an exponential disk. At a given r, the intensity as a function of wavelength is modelled by a Gaussian profile, centered on $\lambda(r)$:

$$I(\lambda, r) = I(r) \exp\left[-\frac{(\lambda - \lambda(r))^2}{2(\sigma^2 + \sigma_{instr}^2)}\right] \tag{5.6}$$

with σ the velocity dispersion and σ_{instr} the instrumental broadening. The latter was obtained from a Gaussian fit to a skyline. We allowed σ to vary in the fit, but assumed it to be independent of radius. With Equations 5.4 to 5.6 we built a two-dimensional model of the Hα emission line, which was then smoothed with the PSF derived in Section 5.2.1. To avoid undersampling effects, we built the initial model on a grid with 3× the spatial and wavelength resolution of the spectra. We also used a 3× refined PSF. After convolving we rebinned the model by a factor 1/3. We also subtracted half the intensity of the model at ±14 pixels to reproduce the dithering pattern. Parameters that can vary in the model are $\lambda_0, \lambda_a, r_0, r_t, I_0, R_s$ and σ.

5.3.2 Fitting procedure

We fit the intensity model to 100Å wide spectral image stamps, centered on the Hα emission line. We used the Python scipy `optimize.curve_fit` algorithm, which is based on the Levenberg-Marquardt algorithm. This algorithm can be used to solve non-linear least squares minimization problems. The Levenberg-Marquardt algorithm can find local minima, but these are not necessarily the global minima, i.e. the best fits, that we are looking for. Therefore, we assessed each galaxy's spectral image stamp individually and we chose initial parameters for the model to be a reasonable match to the observed Hα emission.

In addition to the Hα stamps, we extracted corresponding images from the error spectra that are available for each observation. The error spectra

represent standard errors on the flux in each pixel. The error stamps were matched by wavelength location to the Hα spectral image stamps, and we included these as weight arrays in the fit. We did not mask skylines or pixels with low SNR, but simply used the (much) smaller weights from the error images at those locations.

In Figures 5.5 - 5.8 we show the initial guesses and best-fit models for the four example galaxies shown earlier. The best-fit models are good representations of the Hα emission, with small residuals.

We estimated uncertainties on the parameters $\lambda_0, \lambda_a, r_0, r_t, I_0, R_s$ and σ, by applying a Monte Carlo procedure. For every source, we subtracted the best-fit two-dimensional model from the spectral image stamp, obtaining the residual images shown in the right-hand panels of Figures 5.5-5.8. We then shifted the residual pixels by a random number of rows and columns, preserving local pixel-to-pixel correlations. The magnitude of the shift was drawn from a Gaussian distribution centered on zero, allowing negative values, i.e., shifting in the opposite direction, and with a standard deviation of two pixels. The number of rows and columns to be shifted were generated independently from each other. We then added the best-fit model back to the shifted residual and re-ran our fitting procedure. We repeated this process 100 times, obtaining for each parameter a distribution of values. We calculated the standard deviations for each parameter and used these as the uncertainties.

5.3.3 Velocities

We measured the velocities from Equation 5.2 at 2.2 times the scale radius (R_s) of the exponential brightness profile. We chose $r = 2.2R_s$ as this is the radius where the rotation curve of a self-gravitating ideal exponential disc peaks (Freeman 1970). It is also a commonly adopted parameter in literature (e.g. Miller et al. 2011). Its main advantage is that it gives a consistent approximation of the rotational velocity accross the sample, while avoiding extrapolations towards large radii and low SNR regions of the spectrum.

We corrected the velocities for inclination and misalignment of the slit angle ($\Delta\alpha$) with respect to the PA. For the inclination correction we used:

$$v'_{2.2} = \frac{v_{2.2}}{\sin(i)} \tag{5.7}$$

with

$$i = \cos^{-1}\sqrt{\frac{(b/a)^2 - q_0^2}{1 - q_0^2}} \tag{5.8}$$

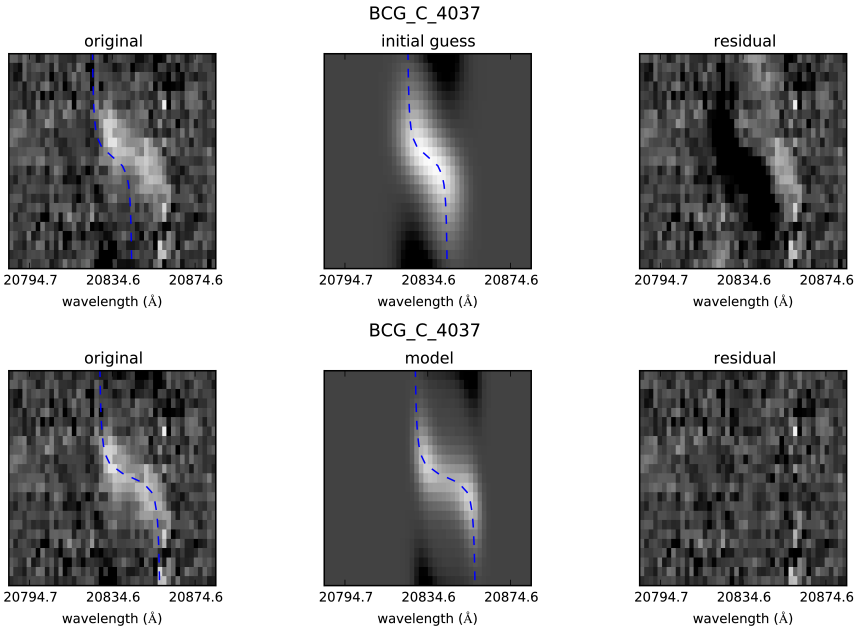

Figure 5.5: Initial guess (top) and best-fit model (bottom) for galaxy BCG_C_4037. From left to right: the spectral image stamps; the initial guess/best-fit model; the residual after subtracting the initial guess/best-fit model. The blue dashed curves are the model arctan functions.

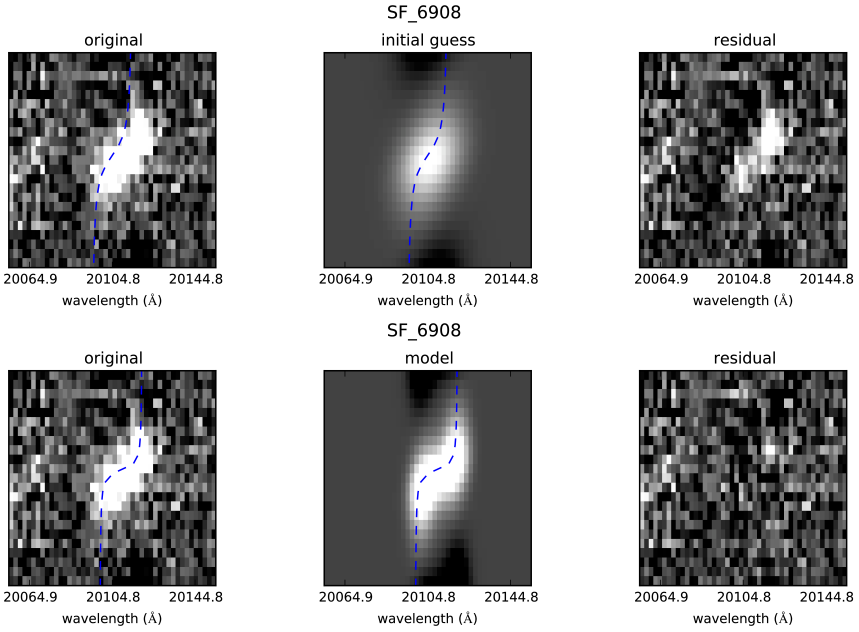

Figure 5.6: Same as Figure 5.5, here for galaxy SF_6908.

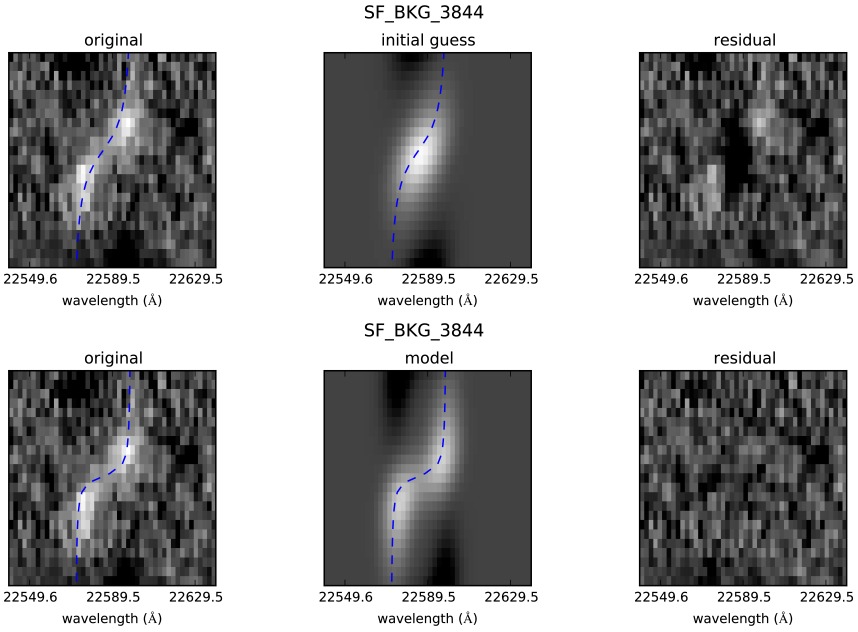

Figure 5.7: Same as Figure 5.5, here for galaxy SF_3844.

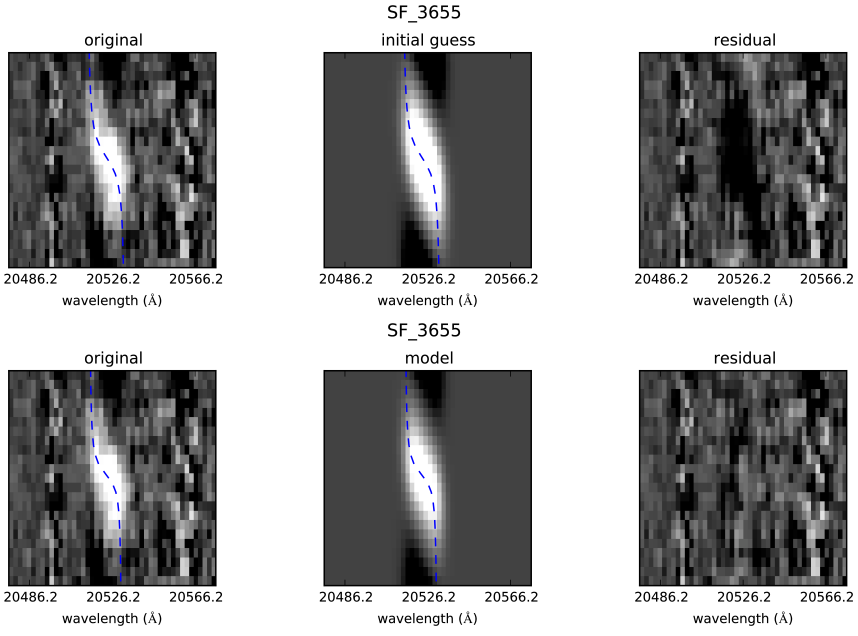

Figure 5.8: Same as Figure 5.5, here for galaxy SF_3655.

We adopt here the convention that $i = 0°$ for galaxies viewed face-on and $i = 90°$ for edge-on galaxies. b/a is the axis ratio. We used the axis-ratio's derived with GALFIT from van der Wel et al. (2014). $q_0 \simeq 0.1 - 0.2$ represents the intrinsic flattening ratio of an edge-on galaxy. Following convention we adopt $q_0 = 0.19$ (Pizagno et al. 2007; Haynes & Giovanelli 1984)

The correction to account for slit mismatch is derived from simple trigonometry:

$$V_{2.2} = \frac{v'_{2.2}}{|\cos(\Delta\alpha)|} \tag{5.9}$$

From hereon we use capital $V_{2.2}$ to indicate the corrected velocity measurements. Any uncertainties on the axis-ratio and PA were propagated and added to the velocity uncertainty from the Monte Carlo procedure. Example velocity curves (corrected and uncorrected) are shown in Figures 5.9 and 5.10.

5.3.4 Two-dimensional PSF and projection effects

When considering slit spectra, with one spatial dimension, we need to account for systematic effects due to the two-dimensional nature of the PSF smoothing. The main effect is that two-dimensional smoothing will effectively lead to an underestimation of the line-of-sight motion captured in one-dimensional spectra, as a flux component from lower velocity regions is mixed in. The effect depends on the apparent size of the galaxy relative to size of the PSF and the size of the slit, i.e., mixing occurs even for an infinitely thin slit if the seeing is significant, and vice versa.

To assess this effect, we first investigated the limiting maximal case of an edge-on, flat, optically thin and circularly symmetric galaxy. Because the galaxy is symmetric and flat, when viewed edge-on all flux will fall within the slit. Since the galaxy is optically thin, the integrated line-of-sight velocity at a projected distance from the center has the maximum contribution from lower velocities sampled at larger radii.

To quantify this, we simulated a two-dimensional exponential disk of one million particles. The disk has scale radius $R_s = 0.36''$, representative of our $2.0 < z < 2.5$ sample, and particles uniformly random distributed over angles α. We assigned velocities v, perpendicular to the radial direction to each particle, using Equation 5.2 with a turn-over radius $r_t = 0.1''$ and $2V_a/\pi = 200$ km/s.

The one-dimensional projection of the line-of-sight velocities of this model are shown in the leftmost panel of Figure 5.11. We show the true arctan velocity curve at radius r, and the line-of-sight velocities versus projected distance along the slit. In the middle panel we show the line-of-sight velocities of particles in a $0.18''$ bin (corresponding to a single MOSFIRE pixel) at $r = 2.2R_s$.

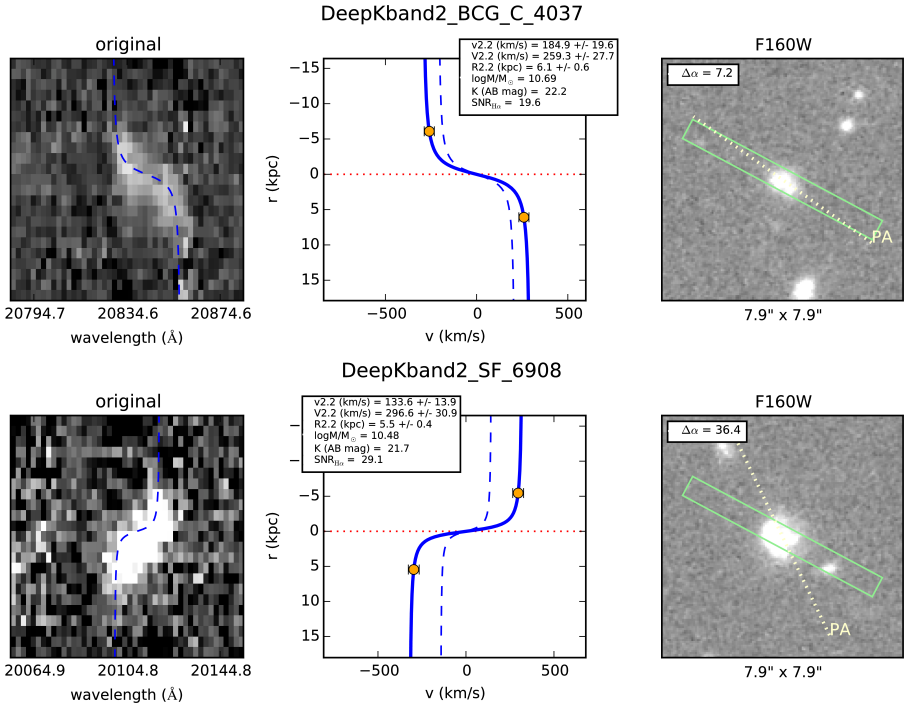

Figure 5.9: Velocity curves for the galaxies in Figures 5.5 and 5.6. The lefthand panels are the spectral image stamps with the best-fit model overplotted. The middle panels show the best-fit arctan functions without corrections ($v(r)$; Equation 5.2; dashed lines) and with inclinations and slit corrections applied (solid lines). The orange datapoints indicate $V_{2.2}$ at $r = \pm 2.2 R_s$. The spatial center (r_0) is indicated by the dotted line. In the righthand panels we show the corresponding HST/WFC3/F160W images. The green box shows the dimensions and orientation of the slit compared to the galaxies. The dotted line indicates the PA of the major axis.

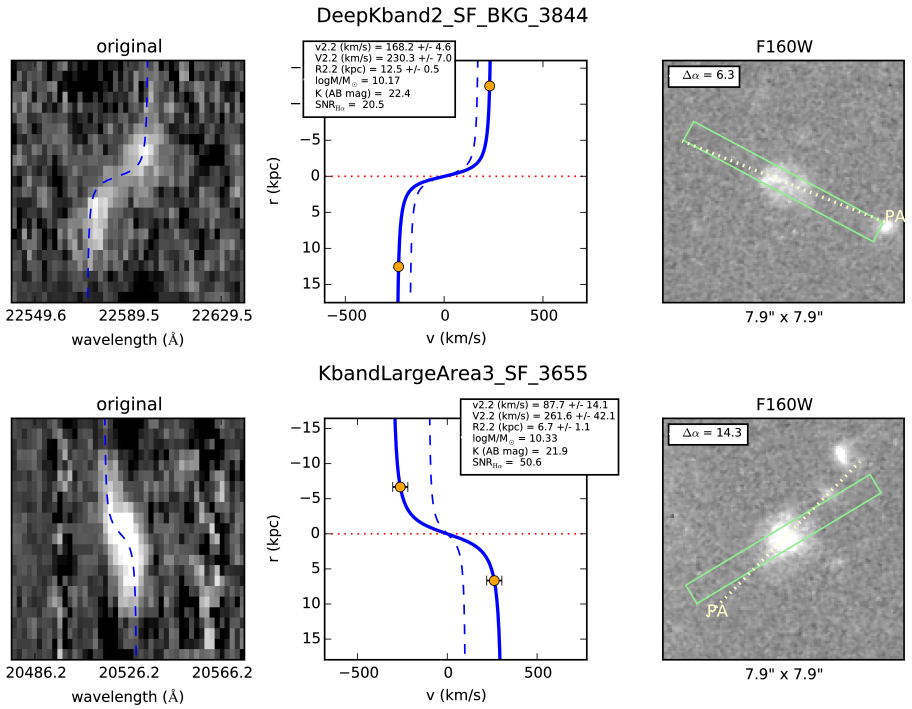

Figure 5.10: Same as Figure 5.9, here for the galaxies in Figures 5.7 and 5.8

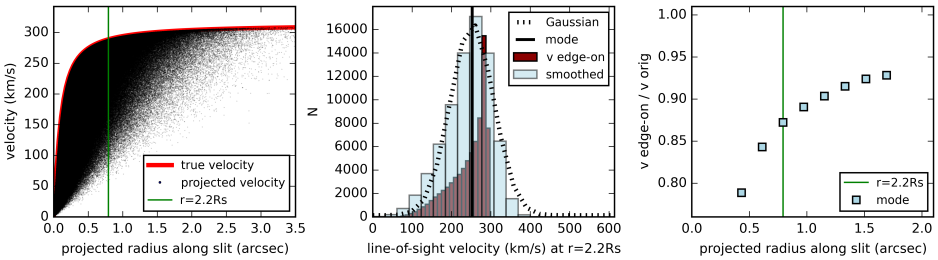

Figure 5.11: Left: the radial velocity distribution of particles in a simulated galaxy, with $z = 2.2$, $R_s = 0.36''$, $r_t = 0.1''$ and $2V_a/\pi = 200$ km/s. The red line represents the input velocities of the particles, which follows an arctan curve. The dots are the projected velocities along the line of sight for an inclination of $i = 90°$. The green line indicates $r = 2.2R_s$. Middle: distribution of projected velocities at $r = 2.2R_s$ (darkred). The blue histogram is the distribution smoothed with σ_{instr}, to which we fitted a Gaussian (dashed curve), with a center at 252 km/s. Right: The the best-fit Gaussian centroids of smoothed distributions at various radii. The datapoint intersecting with the green line (indicating $r = 2.2R_s$) corresponds to the distribution in the middel panel. This result is 13% less than the actual rotation velocity at $r = 2.2R_s$ of 289 km/s.

The average line-of-sight velocity underestimates the true rotational velocity at $2.2R_s$. To mimick the instrumental broadening, we smoothed projected velocities with a Gaussian with $\sigma = \sigma_{instr} = 34$ km/s, and then rebinned to spectral pixels of 31 km/s, corresponding to the MOSFIRE resolution and dispersion at $z = 2.2$. We then fitted a Gaussian to the histogram to obtain the center of the distribution. In a spectrum this would be equivalent to fitting Equation 5.6, with the velocity determined by the location of the brightness peak. We show the best-fit velocities in the rightmost panel, along with the results at different radii. At $2.2R_s$, the difference between the fitted velocity and the actual rotation velocity is 13% for the maximal edge-on case.

To calculate the expected systematic effect for typical galaxies in our sample, we considered the average galaxy, under typical inclination (45 degrees) and seeing conditions (Moffat PSF with FWHM= $0.80''$ and $\beta = 2.5$). To simulate the seeing, we uniformly sampled radial offsets from the growthcurve of the Moffat PSF, and added these to the particles, along with a random angular phase. We also gave each particle a random velocity kick sampled from a Gaussian with $\sigma = 50$ km/s, to simulate a typical velocity dispersion. As above, we simulated a two-dimensional spectrum. First we overlaid a virtual slit and assigned the particles within the slit to $0.18''$ wide spatial bins and 31 km/s wide velocity bins, using their line-of-sight velocities and a shift corresponding to the position of the particle in the slit. The shift adds a broadening to the spectrum that corresponds to the instrument broadening. As a final step, we subtracted half the brightess at ± 14 pixels to mimick the dithering pattern.

We modelled the rotational velocity in the same way as for the observed spectra. After correcting for the inclination of 45 degrees, we find $V_{2.2} = 272$ km/s, a factor 1.06 smaller than the input velocity (289 km/s). To assess the impact of noise on these systematic offsets, we repeat the measurements by adding representative amounts of noise to the simulations. The typical integrated SNR of the spectra in our sample is 23 and we scaled the brightness of the simulated galaxy to obtain a similar SNR. The noise in each pixel was randomly drawn from a normal distribution using the information in one of the error spectra. Adding noise we find $V_{2.2} = 272 \pm 9$ km/s, where the uncertainty was derived from repeating our measurement 100 times, drawing new values for the noise each time. In Figure 5.12 we show the simulated spectrum without noise and with a SNR of 23, along with the best-fit models and residuals.

We found similar results for different seeing values, for example $V_{2.2} = 274$ km/s (a factor 1.05) for a seeing of $0.65''$ or $V_{2.2} = 270$ km/s (a factor 1.07) for a seeing of $0.65''$. Given that the effect is reasonably small, and not very dependent on details of the seeing, from hereon we will apply a fixed correction

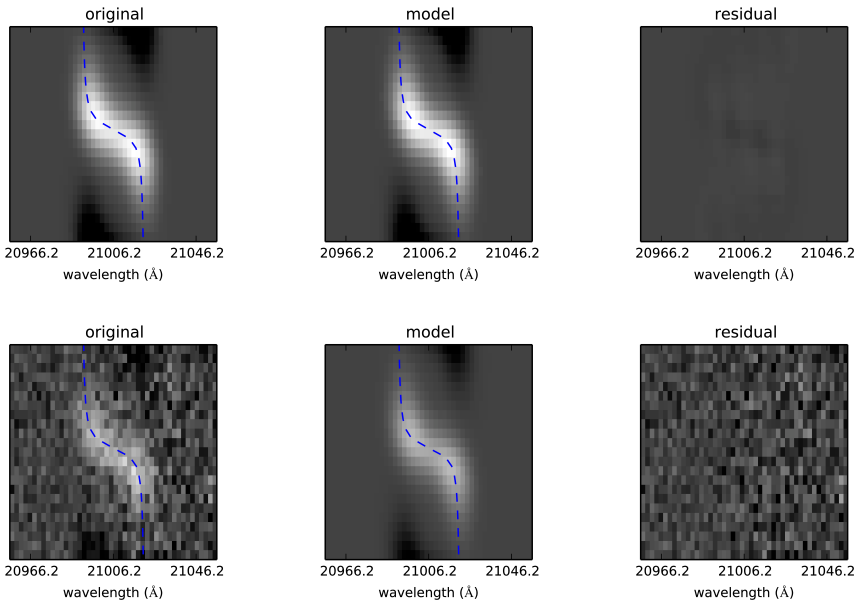

Figure 5.12: A simulated spectrum of a typical galaxy with $z = 2.2$, $R_s = 0.36''$, $r_t = 0.1''$, $\sigma = 50$ km/s, a seeing of $0.8''$, $i = 45°$ and $V_{2.2} = 289$ km/s. The top panels show the pure spectrum with the best-fit model and residual. The bottom panels show the case with noise added, and the brightness scaled to result in an integrated SNR= 23. For the case without noise we measure $V_{2.2} = 272$ km/s, and with noise $V_{2.2} = 272 \pm 9$ km/s. This shows that an approximating the two-dimensional PSF by a one-dimensional PSF reduces velocities by 6%.

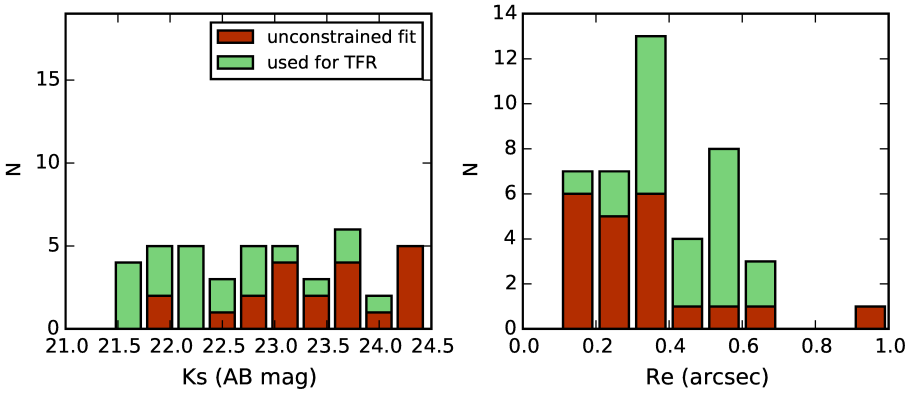

Figure 5.13: K_s-band magnitude and effective F160W radius R_e stacked histograms of 38 galaxies in the high quality sample. For 21 spectra of 19 galaxies the fits were poorly constrained. The remaining 19 galaxies were used to derive the Tully-Fisher relation. The galaxies with unconstrained fits have fainter K_s-band magnitudes and smaller sizes on average, which resulted in a selection bias towards larger and/or brighter galaxies.

of 1.06 to our velocities.

5.3.5 Results

Of the 43 spectra of in the high quality sample, we obtained good fits for 22, while for 21 spectra we obtained poorly constrained fits. The best-fit parameters of the rotation model and their uncertainties, along with $v_{2.2}$ and $V_{2.2}$, are shown in Tables 5.1 and 5.2.

The poor fits were caused by various reasons. In some cases the spectra were just noisy, with very large random uncertainties (> 50%) on the velocities. In other cases, galaxies were unresolved and/or showed no clear rotation or turnover in the rotation. In these cases, r_t was poorly constrained leading to unrealistic and poorly constrained solutions (e.g. $r_t >> R_s$).

We therefore removed these 21 spectra (of 19 galaxies) from the sample. To evaluate if removing the failed fits introduces biases relative to the target sample we show the distribution of the K_s−band magnitudes and sizes in Figure 5.13. The K_s−band magnitudes for the good fits are brighter than those of the full target sample (median K_s = 22.2 versus median K_s = 22.9) and the galaxies are slightly larger (median $R_e = 0.43''$ versus $R_e = 0.35''$). So removing these galaxies does bias the sample to somewhat brighter and larger galaxies.

Table 5.1: Results

ID	mask	seeing ($''$)	$z_{centroid}$	V_a (km/s)	r_t ($''$)
BCG_C_4037	DeepKband2	0.80	2.1750±0.00004	212.2±173.2	1.5e-01±1.2e-01
–	mask2	0.67	2.1747±0.00005	190.8±16.1	8.9e-02±2.7e-02
DU_3598	mask2	0.67	2.2279±0.00007	145.3±38.0	1.3e-01±8.2e-02
DU_6553	mask1	0.71	2.1245±0.00012	128.5±36.6	4.5e-08±1.3e-02
SF_2715	mask2	0.67	2.0824±0.00004	46.2±122.7	2.2e-06±4.5e-01
SF_2723	mask2	0.67	2.0851±0.00004	97.4±8.2	3.7e-07±5.3e-03
SF_2765	mask1	0.71	2.2279±0.00005	282.9±331.4	4.5e-01±7.2e-01
SF_3074	mask1	0.71	2.2266±0.00005	116.9±20.5	2.7e-02±3.1e-02
SF_3527	KbandLargeArea4	0.65	2.1890±0.00005	88.5±7.5	1.4e-03±4.7e-03
SF_3633	DeepKband2	0.80	2.0991±0.00007	170.7±19.7	7.9e-02±3.2e-02
–	mask1	0.71	2.0982±0.00006	325.0±360.3	3.5e-01±5.0e-01
SF_3655	KbandLargeArea3	1.09	2.1263±0.00003	110.0±60.6	2.2e-01±3.8e-01
SF_3680	mask3	0.68	2.1753±0.00009	132.7±165.3	1.2e-01±4.4e-01
SF_3714	mask3	0.68	2.1761±0.00005	81.1±57.2	2.1e-02±3.2e-02
SF_4099	mask3	0.68	2.4391±0.00004	50.6±11.9	5.6e-03±3.1e-02
SF_4645	DeepKband1	1.10	2.1011±0.00008	184.7±39.6	1.7e-01±5.0e-02
SF_4930	DeepKband2	0.80	2.0974±0.00002	69.8±155.7	1.5e-01±1.2e+00
SF_5630	KbandLargeArea4	0.65	2.2427±0.00006	152.5±126.1	2.2e-01±2.9e-01
SF_6908	mask1	0.71	2.0631±0.00006	152.5±202.1	5.3e-02±4.6e-01
–	DeepKband2	0.80	2.0633±0.00003	145.6±111.6	8.2e-02±6.2e-02
SF_8108	mask2	0.67	2.1622±0.00006	224.4±59.6	1.4e-01±9.1e-02
SF_BKG_3844	DeepKband2	0.80	2.4404±0.00001	177.0±10.3	1.2e-01±1.7e-01

ID	mask	$SNR_{H\alpha}$	R_s ($''$)	σ (km/s)	$v_{2.2}$ (km/s)
BCG_C_4037	DeepKband2	16	0.33±0.03	21.9±9.8	185.1±8.1
–	mask2	20	0.34±0.04	61.0±9.2	176.5±12.0
DU_3598	mask2	11	0.34±0.06	75.8±11.9	128.9±27.6
DU_6553	mask1	22	0.02±0.02	84.0±10.1	128.5±17.1
SF_2715	mask2	21	0.18±0.02	64.7±9.3	46.2±9.7
SF_2723	mask2	13	0.13±0.01	60.8±4.0	97.4±6.8
SF_2765	mask1	38	0.22±0.02	75.7±10.1	147.3±22.4
SF_3074	mask1	16	0.37±0.05	78.1±11.0	114.4±15.9
SF_3527	KbandLargeArea4	71	0.23±0.04	67.1±11.1	88.3±7.2
SF_3633	DeepKband2	20	0.44±0.09	52.2±5.7	161.9±17.1
–	mask1	19	0.36±0.02	83.1±8.5	238.8±10.7
SF_3655	KbandLargeArea3	41	0.35±0.06	27.3±12.9	90.5±11.9
SF_3680	mask3	10	0.19±0.03	0.0±17.7	108.9±20.8
SF_3714	mask3	33	0.22±0.01	70.1±3.9	78.9±7.6
SF_4099	mask3	15	0.27±0.03	25.0±6.7	50.3±8.0
SF_4645	DeepKband1	3	0.26±0.04	0.0±16.1	151.1±22.3
SF_4930	DeepKband2	14	0.24±0.04	49.5±7.2	57.7±15.4
SF_5630	KbandLargeArea4	26	0.27±0.02	67.2±9.2	117.8±21.7
SF_6908	mask1	24	0.34±0.05	89.3±15.3	145.7±24.5
–	DeepKband2	19	0.30±0.02	46.8±7.6	134.0±11.7
SF_8108	mask2	17	0.15±0.03	49.6±19.3	165.5±27.4
SF_BKG_3844	DeepKband2	16	0.70±0.13	53.0±6.2	168.2±18.2

Table 5.2: Results continued

ID	mask	$V_{2.2}$ (km/s)	$sin(i)$	α_{mask} (deg)	$\|cos(\Delta\alpha)\|$
BCG_C_4037	DeepKband2	259.7±11.8	0.72	-62.0	0.99
–	mask2	247.7±17.1	–	-47.3	0.99
DU_3598	mask2	167.0±37.5	0.97	-47.3	0.80
DU_6553	mask1	252.3±33.9	0.61	134.0	0.83
SF_2715	mask2	58.2±12.2	0.81	-47.3	0.99
SF_2723	mask2	239.4±18.0	0.49	-47.3	0.83
SF_2765	mask1	212.7±32.5	0.78	134.0	0.88
SF_3074	mask1	134.5±18.9	0.87	134.0	0.97
SF_3527	KbandLargeArea4	100.8±8.3	0.89	2.0	0.98
SF_3633	DeepKband2	185.0±20.1	0.95	-62.0	0.92
–	mask1	322.3±17.9	–	134.0	0.78
SF_3655	KbandLargeArea3	269.7±35.5	0.35	59.0	0.97
SF_3680	mask3	137.5±27.0	0.87	14.8	0.91
SF_3714	mask3	117.3±11.6	0.72	14.8	0.93
SF_4099	mask3	77.4±12.5	0.72	14.8	0.90
SF_4645	DeepKband1	156.5±23.9	0.97	2.0	1.00
SF_4930	DeepKband2	66.9±20.3	1.00	-62.0	0.86
SF_5630	KbandLargeArea4	148.9±28.8	0.99	2.0	0.80
SF_6908	mask1	277.7±46.7	0.56	134.0	0.94
–	DeepKband2	297.5±26.2	–	-62.0	0.80
SF_8108	mask2	171.1±30.3	0.98	-47.3	0.98
SF_BKG_3844	DeepKband2	230.4±25.0	0.73	-62.0	0.99

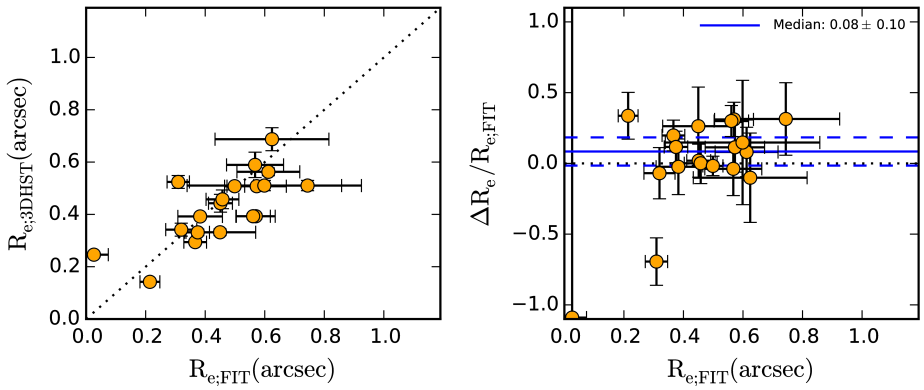

Figure 5.14: Left: $R_e = 1.678 * R_s$ in the $K-$band from our fits versus R_e in the HST/WFC3/F160W-band measured by van der Wel et al. (2014), for 18 spectra (removing two flagged as bad fits by van der Wel et al. (2014)). The dotted line indicates the one-to-one relation. Right: $\Delta R_e/R_{e;FIT} = (R_{e;FIT} - R_{e;3DHST})/R_{e;FIT}$ as a function of $R_{e,FIT}$. The bootstrapped median and 1σ error on the median are shown as the solid and dashed lines, respectively.

The 19 galaxies for which we will derive the Tully-Fisher relation have high velocities and velocity dispersion, with a median $V_{2.2} = 178$ km/s, $\sigma = 61$ km/s and a typical $V/\sigma = 3$. We note that these dispersions could be slightly overestimated, e.g., the dispersion reflects mixing of velocity gradients on scales smaller than the seeing.

At high redshift measuring the kinematic profile of a galaxy is more difficult due to the smaller angular scales for distant galaxies, and seeing effects and SNR play a larger role. We therefore verified our size measurements. We converted the best-fit $R_s$derived from the $K-$band spectra to effective radius (R_e), using $R_e = 1.678R_s$ (valid for exponential disks), and compared this with the effective radii from HST/WFC3/F160W image reported by van der Wel et al. (2014). On average we find good agreement, with some scatter, and we derived a bootstrapped median $\Delta R_e/R_e = 0.08 \pm 0.10$. The most prominent outliers, with $|\Delta R_e/R_e| > 0.5$, occur for two small galaxies. The most extreme outlier is an approximately face-on galaxy, with a skyline partly overlapping the Hα emission. The other has a very irregular morphology, and was flagged by van der Wel et al. (2014) as a suspicous GALFIT result. We note that van der Wel et al. (2012) explored systematics for their GALFIT measurements in F160W and found that for galaxies with magnitudes up to $F160W \sim 23$, systematic uncertainties were of the order of $1-2\%$, and $1-10\%$ for magnitudes up to $F160W \sim 24$.

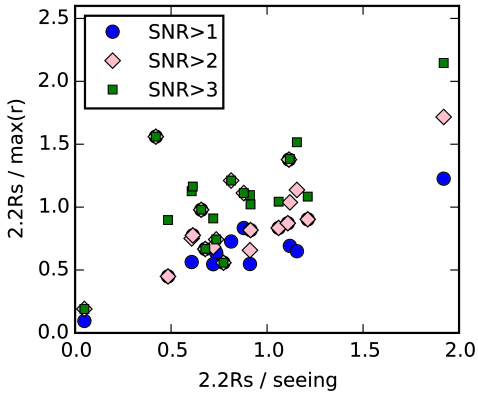

Figure 5.15: SNR, seeing and spatial extent of the signal for all objects in our sample. On the x-axis we show how $r = 2.2R_s$ scales with the FWHM of the seeing. On the y-axis we show how it relates to the extent of the signal. 82% of the sample has $2.2R_s < \max(r)$, with $\max(r)$ the radius within which the SNR is > 1.

We also investigated how $2.2R_s$ related to the seeing and the SNR. We measured the SNR within $5R_s$ above and below the center of the line, but never beyond $1.26''$ to avoid the negative imprints of the emission line in the spectrum. We also defined a wavelength region within which to measure SNR, defined by the maximum shear of the line, plus a buffer of $3FWHM_\lambda = 3(2\sqrt{2\ln 2})\sqrt{\sigma^2 + \sigma^2_{instr}}$Å. The SNR within these limits was calculated by summing the flux and summing the squares of the equivalent pixels in the noise spectrum, and dividing the first by the square root of the latter. The total SNR is included in Table 5.1. We also calculated the SNR for each row of pixels individually and determined the spatial extent of emission with SNR> 1, SNR> 2 and SNR> 3. The results are shown in Figure 5.15, relative to $2.2R_s$. For 82% of the spectra, the emission extends beyond $2.2Rs$ at SNR> 1. The majority of the sample has $2.2R_s$ smaller than the FHWM of the seeing. The determination of R_s itself seems to be correct, as verified above.

We note that 5/22 fits resulted in very small r_t, with $r_t < 0.02''$. This is clearly much less than the resolution of a pixel: $0.18''$. To investigate the potential impact of small r_t on the velocities, we re-fit the spectra limiting r_t to $r_t > 0.02''$. Except for one bad fit, the average velocities for the whole sample go up by 17%. If we had simply ignored the galaxies with very small r_t, then the average velocity would have gone up by 6%. This may indicate that the velocities are underestimated for sources with small r_t, but without knowing the true r_t, the effect is difficult to quantify. We note that a value of zero for r_t is possible in the presence of non-circular motion, for example if

the galaxy has a bar (Franx & de Zeeuw 1992). For the three galaxies with $r_t < 0.001''$ we inspected the F160W images, but found no indications of a bar-like morphology. One galaxy has a very irregular morphology, and the other two appear to be morphologically compact or face-on galaxies. One indeed has $n_{sersic} = 4.5$ and both have inclinations < 40 deg, so that the shear is likely unresolved.

Three galaxies were included in two masks (BCG_C_4037, SF_3633 and SF_6908 in Tables 5.1 and 5.2). As they were observed under different seeing conditions and have different SNR and slit-angle, they provide a useful check on consistency. Encouragingly, we find that these three galaxies have velocities, redshifts and scale parameters that agree between masks within their uncertainties. We averaged their velocities to derive the Tully-Fisher relation in the next Section.

5.4 The Tully-Fisher relation at 2.0 < z < 2.5

5.4.1 Tully-Fisher sample

We show F160W images of the remaining 19 galaxies in the Tully-Fisher sample in Figure 5.16, and illustrate the orientation of their major axis and the MOSFIRE slits. Physical properties of the sample are shown in Table 5.3, Figure 5.17 and Figure 5.18. We also compare with the primary 187 ZFIRE targets as well as with the general population of galaxies at this redshift obtained from ZFOURGE. For the ZFOURGE sample we selected galaxies with stellar mass $M/M_{\odot} > 10^9$. The 19 galaxies in our sample cover the full range of the star-forming region of the UVJ diagram (below the red line), but they have higher SFRs compared to the SFR-stellar mass relation for star-forming galaxies at $2.0 < z < 2.5$ (Tomczak et al. 2015). They lie at the bright, high-mass end of the general galaxy population. They have a large spread in size (Figure 5.18), including even a massive compact galaxy with effective size $R_e = 0.14''$, but on average they are larger than predicted by the size-mass relation at $2.0 < z < 2.5$ (van der Wel et al. 2014).

Of the 19 galaxies, 6 are spectroscopically confirmed to be part of the $z = 2.095$ galaxy cluster. However, due to the small number of cluster galaxies, a study of the effects of environment on the evolution of the Tully-Fisher relation is not feasible.

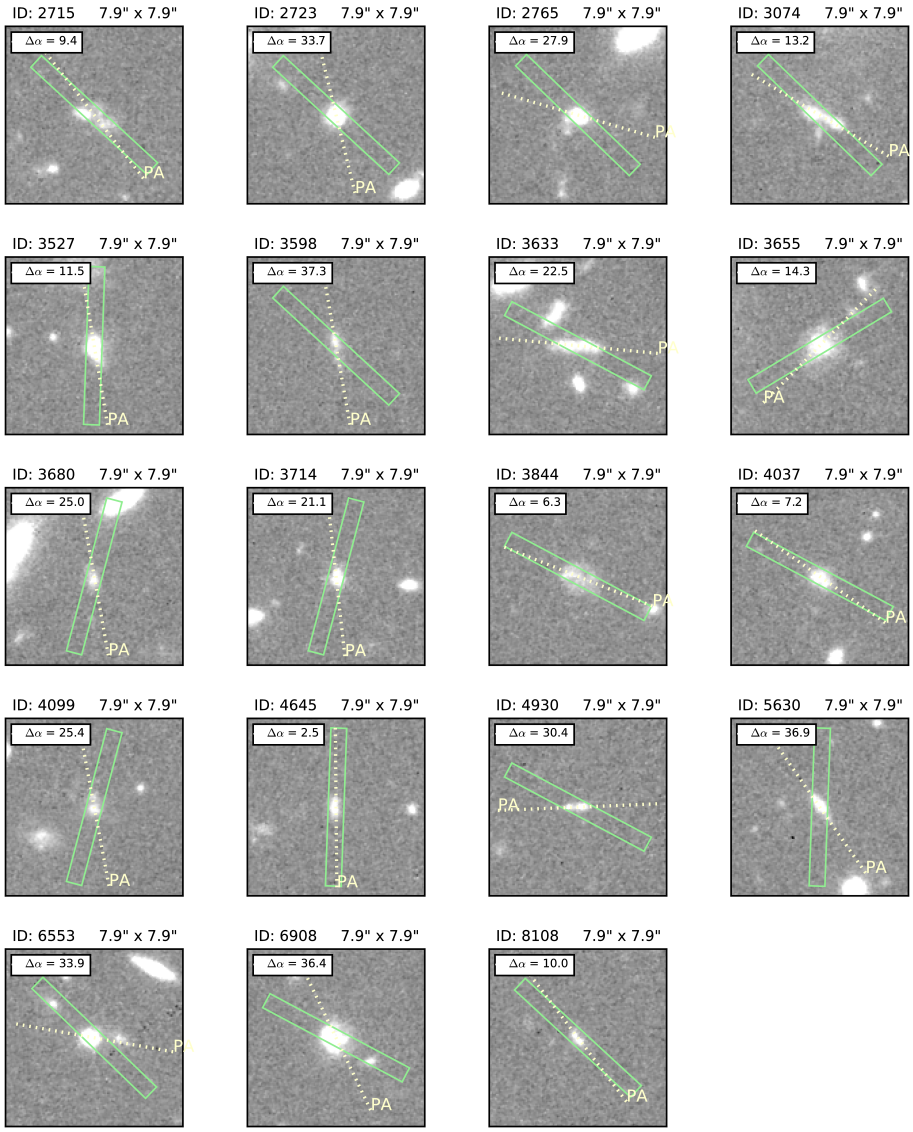

Figure 5.16: F160W images of the galaxies in our Tully-Fisher sample, shown in the same way as in Figure 5.9.

Table 5.3: Full sample

ID	R.A. (deg)	Decl (deg)	z_{spec}	K_s (AB mag)	F160W (AB mag)	$M/10^{10}$ ($M_\odot$)
BCG_C_4037	150.0981293	2.2428052	2.1738	22.2	23.1	4.9
DU_3598	150.1120911	2.2368469	2.2281	22.9	23.7	2.3
DU_6553	150.0686035	2.2797525	2.1246	21.7	22.4	5.4
SF_2715	150.0895386	2.2235634	2.0827	22.6	23.0	0.8
SF_2723	150.1172638	2.2238791	2.0857	21.5	22.0	7.8
SF_2765	150.119339	2.2241209	2.2282	21.9	22.5	2.4
SF_3074	150.1209106	2.2288201	2.2269	22.2	22.6	1.4
SF_3527	150.1825714	2.2358665	2.1892	21.9	22.5	2.1
SF_3633	150.1249237	2.236979	2.1008	22.1	22.7	2.5
SF_3655	150.1691284	2.2383816	2.1267	21.9	22.4	2.1
SF_3680	150.063446	2.237031	2.1757	24.0	24.1	0.2
SF_3714	150.0707703	2.2381561	2.1764	22.7	23.3	1.5
SF_4099	150.0718231	2.243396	2.4395	23.1	23.7	2.2
SF_4645	150.0743256	2.2516196	2.1014	23.6	23.8	0.3
SF_4930	150.0559387	2.2557058	2.0975	23.5	23.8	0.3
SF_5630	150.2009735	2.2665324	2.2429	22.9	23.3	0.9
SF_6908	150.0834198	2.2857671	2.0634	21.7	22.2	3.0
SF_8108	150.0622711	2.3044007	2.1625	23.8	24.1	0.5
SF_BKG_3844	150.1094666	2.2400432	2.4410	22.4	23.0	1.5

ID	SFR ($M_\odot$/yr)	R_e ($''$)	GALFIT flag[a]	b/a	n_{sersic}	P.A. (deg)
BCG_C_4037	95.7	0.39±0.01	0	0.71±0.02	0.6	-54.8
DU_3598	69.9	0.59±0.05	0	0.30±0.03	1.1	-10.0
DU_6553	114.5	0.25±0.00	0	0.80±0.01	1.2	-79.9
SF_2715	34.4	0.52±0.02	1	0.61±0.03	0.6	-37.9
SF_2723	66.4	0.14±0.00	0	0.88±0.02	4.5	-13.6
SF_2765	101.1	0.29±0.01	0	0.64±0.01	1.8	-73.9
SF_3074	30.9	0.69±0.04	0	0.51±0.02	2.3	-59.2
SF_3527	168.8	0.39±0.01	0	0.48±0.01	1.0	-9.5
SF_3633	150.5	0.51±0.02	1	0.37±0.02	0.9	-84.5
SF_3655	127.3	0.56±0.01	0	0.94±0.01	1.1	44.7
SF_3680	16.9	0.34±0.02	0	0.51±0.04	0.9	-10.2
SF_3714	23.1	0.33±0.01	0	0.71±0.02	0.7	-6.3
SF_4099	57.5	0.46±0.04	0	0.71±0.03	1.7	-10.6
SF_4645	14.1	0.33±0.01	0	0.32±0.03	0.5	-0.5
SF_4930	15.3	0.41±0.02	2	0.09±0.02	0.2	87.6
SF_5630	111.3	0.44±0.03	0	0.24±0.02	2.4	-34.9
SF_6908	138.9	0.51±0.01	0	0.84±0.01	0.4	-25.6
SF_8108	16.3	0.31±0.02	2	0.27±0.03	0.2	-37.3
SF_BKG_3844	65.2	0.64±0.02	2	0.69±0.02	0.2	-68.3

[a] 0: good fit; 1: suspicious fit; 2: bad fit (van der Wel et al. 2012).

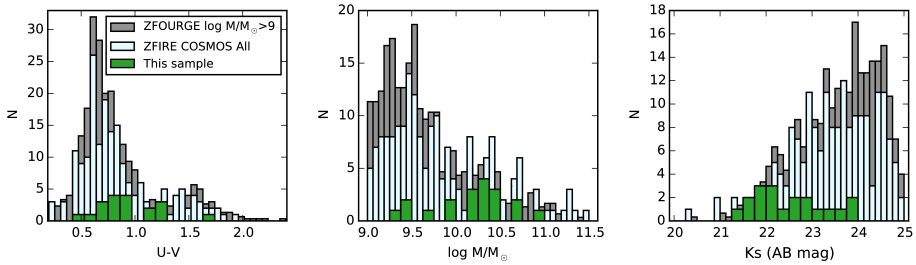

Figure 5.17: Rest-frame $U-V$ colours, stellar masses and ZFOURGE K_s-band magnitudes for the 19 galaxies used here to derive the Tully-Fisher relation (green), the ZFIRE target sample (lightblue), and a parent sample drawn from ZFOURGE with $2 < z < 2.5$ and $M/M_{\odot} > 10^9$ (gray). The gray histograms were reduced by a factor of three for reasons of visibility. The 19 galaxies of this study have a large range in $U-V$, stellar mass and brightness.

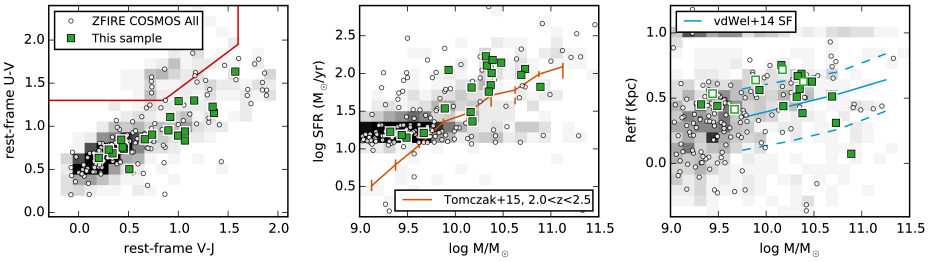

Figure 5.18: Left: UVJ diagram of the ZFIRE sample (open symbols) and the 19 galaxies studied in this work (squares). The underlying histogram is the full distribution of $2.0 < z < 2.5$ galaxies with $M > 10^9$ $M_{\odot}$from ZFOURGE. This diagram separates quiescent galaxies from star-forming galaxies based on their rest-frame $U-V$ and $V-J$ colors, obtained from ZFOURGE photometry. The 19 galaxies in the sample span the full range in color typical of star-forming galaxies (region below the red line). Middle: stellar mass versus the logarithm of SFR. The orange line shows the median SFR as function of stellar mass of star-forming galaxies at $2.0 < z < 2.5$ (Tomczak et al. 2015). Most of the galaxies in the sample are above the SFR-stellar mass relation at that redshift (Tomczak et al. 2015). Right: stellar mass versus effective radius, with the size-mass relation at $2.0 < z < 2.5$ for star-forming galaxies shown as a blue line. The dashed lines are the corresponding 16th and 84th percentiles. Open square datapoints are flagged as suspicious or bad fits in the catalogs of van der Wel et al. (2014). Our sample has a large spread in size, with mostly larger sizes than the size-mass relation, but also a very massive, compact galaxy.

Table 5.4: Tully-Fisher variables

ID	$V_{2.2;TF}$ (km/s)	$S_{05;TF}$ (km/s)	σ_{TF} (km/s)
BCG_C_4037	268.9±22.0	194.6±15.5	41.5±13.4
DU_3598	177.1±39.8	146.4±24.8	75.8±11.9
DU_6553	267.5±35.9	206.9±23.6	84.0±10.1
SF_2715	61.7±13.0	78.0±9.3	64.7±9.3
SF_2723	253.8±19.1	189.5±12.8	60.8±4.0
SF_2765	225.5±34.4	176.5±22.4	75.7±10.1
SF_3074	142.6±20.0	127.5±13.1	78.1±11.0
SF_3527	106.8±8.8	101.0±8.7	67.1±11.1
SF_3633	268.9±28.6	201.8±19.3	67.6±10.2
SF_3655	285.9±37.7	204.0±26.5	27.3±12.9
SF_3680	145.7±28.7	103.1±20.3	0.0±17.7
SF_3714	124.4±12.2	112.5±7.2	70.1±3.9
SF_4099	82.1±13.3	63.2±9.0	25.0±6.7
SF_4645	165.9±25.4	117.3±17.9	0.0±16.1
SF_4930	70.9±21.5	70.4±11.9	49.5±7.2
SF_5630	157.8±30.5	130.3±19.1	67.2±9.2
SF_6908	304.8±56.7	226.0±38.6	68.0±17.1
SF_8108	181.4±32.1	137.5±22.3	49.6±19.3
SF_BKG_3844	244.3±26.5	180.7±18.0	53.0±6.2

5.4.2 The Tully-Fisher relation

The Tully-Fisher relation is the relation between rotational velocity and stellar mass. We show our rotation measurements (also shown in Table 5.4) versus stellar mass in the left panel of Figure 5.19, using the stellar masses taken from the ZFOURGE catalogs. Our sample has a large scatter, with the root-mean-square $\sigma_{rms} = 0.21$ dex. We added σ_{rms} in quadrature to the velocity uncertainties and performed a linear regression to the data following:

$$\log V_{2.2} = B + A(\log M/M_{\odot} - 10) \tag{5.10}$$

The Tully-Fisher relation is by convention shown in diagrams with stellar mass on the y-axis. However, the dominant uncertainty here is that in velocity and therefore we performed regression with $V_{2.2}$ as the dependent variable. This is also a method very commonly used in literature which acts against Malmquist bias (Bamford et al. 2006; Weiner et al. 2006b; Kelly 2007).

We obtain from the fit B =(2.19 ± 0.049) and A =(0.247 ± 0.094). We derived the uncertainties by bootstrapping the sample 1000 times, and taking the standard deviation from the bootstrapped distributions of B and A. The slope of the Tully-Fisher relation, (0.247 ± 0.094), is consistent with previous results at $z = 0$. For example, Reyes et al. (2011) find A = 1/3.5 = 0.29 and Bell & de

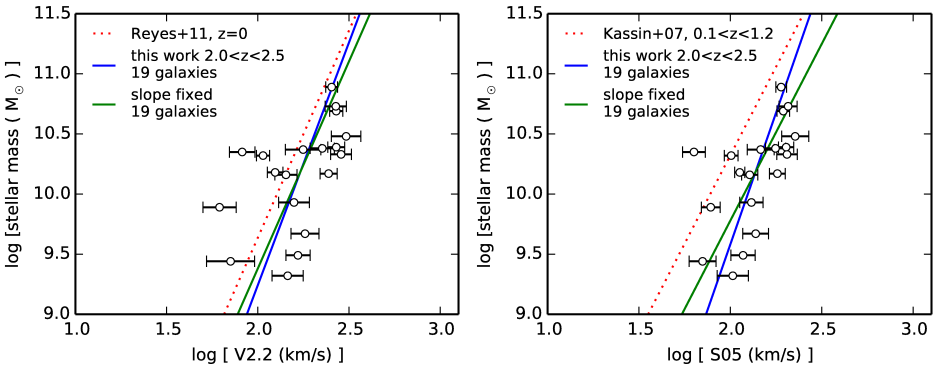

Figure 5.19: Left: Stellar mass versus velocity and the best-fit Tully-Fisher relation for galaxies at $2.0 < z < 2.5$: $\log V_{2.2} = (2.19 \pm 0.049) + (0.247 \pm 0.094)(\log M/M_{\odot} - 10)$ (blue line). The dotted red line is the $z = 0$ result from (Reyes et al. 2011). The green line is the best-fit result with the slope fixed to that at $z = 0$. The scatter in velocity is 0.21 dex. Right: Stellar mass versus $S_{05} = \sqrt{0.5V_{2.2}^2 + \sigma^2}$. The scatter in S_{05} is smaller than in velocity: 0.16 dex. We derive a best-fit relation $\log S_{05} = (2.07 \pm 0.03) + (0.224 \pm 0.060)(\log M/M_{\odot} - 10)$ (blue line), with a steeper slope than at $0.1 < z < 1.2$ (dotted red line; Kassin et al. 2007). The green line is the best-fit with the slope fixed to match the relation of Kassin et al. (2007).

Jong (2001) find $A = 1/4.5 = 0.22$. Our study has too few numbers to significantly constrain evolution in the slope between $z = 0$ and $z = 2.2$. If we fix the slope to that at lower redshift we can study the evolution of the zeropoint. Setting $A = 0.29$, we find $B = (2.18 \pm 0.040)$. Compared to $z = 0$ (Reyes et al. 2011), this implies an evolution of the zeropoint (in stellar mass) of $\Delta M/M_{\odot} = -0.26 \pm 0.14$ dex. We included here a small correction of -0.05 dex in stellar mass to account for the Kroupa (2001) IMF used by Reyes et al. (2011) instead of the Chabrier (2003) IMF used here. As an additional consistency check, we fitted the Tully-Fisher relation only to the galaxies with highest SNR, fixing the slope to $A = 0.29$. We obtained a consistent result, with $B = (2.16 \pm 0.052)$ for the 10 galaxies with spectra with SNR> 20.

The scatter of the residual velocities with respect to the Tully-Fisher relation is significant, with $\sigma = 0.18$ dex. This is more than at $z = 0$, and may be related to star-forming galaxies at high redshift showing more variety in kinematics, and the increase of non-rotationally supported galaxies (e.g. Kassin et al. 2007). An alternative to the stellar mass-velocity relation is the stellar mass-S_{05} relation, with $S_{05} = \sqrt{0.5V_{2.2}^2 + \sigma^2}$. This relation was first coined by Weiner et al. (2006a), and Kassin et al. (2007) showed that the scatter decreases significantly if S_{05} is used. They also found that it does not evolve significantly between $z = 0.1$ and $z = 1.2$.

We calculated S_{05} for the galaxies in our sample (right panel in Figure 5.19) and find the scatter is indeed smaller: 0.16 dex. This is similar to what Kassin et al. (2007) derived at $0.1 < z < 1.2$ (also 0.16 dex) and to a recent study by Price et al. (2015), using MOSFIRE at $1.4 < z < 2.6$ (0.17 dex). We derived the best-fit relation to the data with A and B free in the fit and found $\log S_{05} = (2.10 \pm 0.033) + (0.228 \pm 0.059)(\log M/M_{\odot} - 10)$. Here the slope is steeper than at $0.1 < z < 1.2$, where Kassin et al. (2007) found that $A = 0.34$. This is in agreement with the previous study by Cresci et al. (2009), who did not derive a best-fit to their data, but they do find higher S_{05} values towards smaller stellar mass compared to the $0.2 < z < 1.2$ relation. Keeping the slope fixed at $A = 0.34$, we found $B = (2.07 \pm 0.031)$. This implies a zeropoint evolution of $\Delta M/M_{\odot} = -0.54 \pm 0.13$ dex compared to $0.1 < z < 1.2$. Price et al. (2015) also find an offset for galaxies at high redshift, implying $\Delta M/M_{\odot} \sim -0.3$ dex, and their data does not indicate a steeper slope. Their offset from $0.1 < z < 1.2$ is inconsistent with and less than what we find (by 0.14 dex), which could be due to the inclusion of galaxies at $z < 2$, their assumption of a Gaussian PSF, and our correction for two-dimensional PSF effects on velocity of 6%. Both their result and ours point towards evolution of the zeropoint of the stellar mass-S_{05} relation between $z < 1.2$ and $z \gtrsim 2$, but no evolution for the scatter in S_{05}.

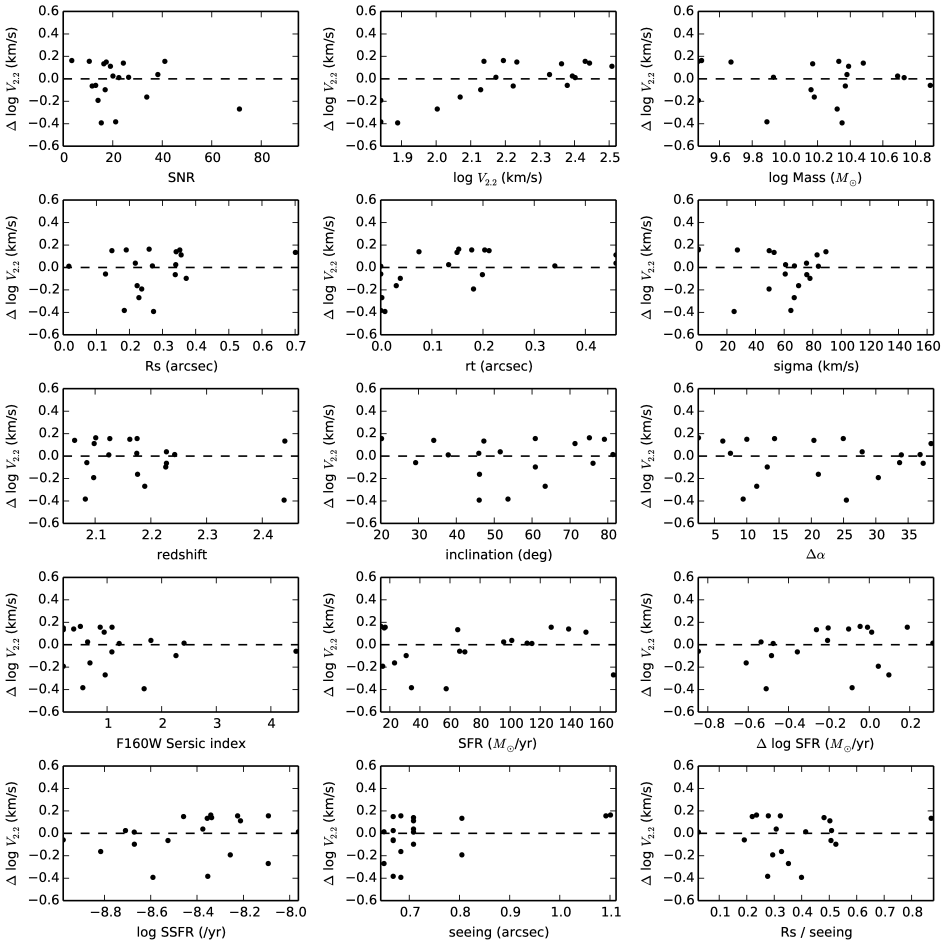

Figure 5.20: Difference between the observed velocity and the velocity predicted by the best-fit Tully-Fisher relation. We plot against model and observational parameters. From left to right, top to bottom: SNR, log $V_{2.2}$, log stellar mass, best-fit R_s, best-fit r_t, best-fit σ, redshift, inclination, $\Delta\alpha$, Sersic index, SFR, SFR minus predicted SFR at $2.0 < z < 2.5$ (Tomczak et al. 2015), SSFR, seeing and R_s relative to the seeing. We find almost no correlations, except for best-fit r_t, with a negative offset in $\Delta\log V_{2.2}$ for fits with $r_t \approx 0$.

To investigate if there are remaining systematic trends we show in Figure 5.20 the velocity residuals of the best-fit Tully-Fisher relation with respect to various parameters and properties of the galaxies (such as SFR). We define the residual as $\Delta \log V_{2.2} = \log V_{2.2} - \log V_{TFR}$ with V_{TFR} the rotational velocity predicted from the fit for a specific stellar mass. There are no systematic effects related to Sersic Index, R_s, stellar mass, SFR, SSFR, or offset from the SFR-stellar mass relation at $2.0 < z < 2.5$ (Tomczak et al. 2015). In addition there is no clear relation with inclination, PA, or seeing. A few prominent outliers have a negative $\Delta \log V_{2.2}$, i.e., they are located to the left of the Tully-Fisher relation in Figure 5.19. These have average values very small r_t, the kinematic scale radius. As we have shown in Section 5.3.3, resolution effects may play a role in determining r_t, and we derive somewhat higher velocities if we assume a higher value for r_t. If we exclude the five galaxies with very small r_t ($< 0.02''$) from the sample, we find a velocity offset of 6%, or $\Delta \log M/M_\odot = 0.24$ dex, implying a total zeropoint evolution of $\Delta \log M/M_\odot = -0.50 \pm 0.11$ dex for the remaining galaxies.

5.5 Discussion

5.5.1 Comparison to literature

To put our result into context, we show the evolution of the stellar mass zeropoint in Figure 5.21, and include previous results from literature. These were all derived from the stellar mass-velocity relation, with quite strong discrepancies between different studies[3]. If the results of Miller et al. (2011, 2012) are adopted, it appears that there is little evolution to $z \sim 1.7$, and strong evolution at $z > 2$. However, before comparing with other studies at different redshifts, several major caveats have to be taken into account: studies use different galaxy selections, different methodologies to derive velocity and stellar mass, and different types of spectroscopic observations. We will discuss these first and then review and compare the studies.

The first is selection bias. At $z > 2$ star-forming galaxies have different properties on average than at $z = 0$. For example, they have higher SFRs, higher gas masses and smaller sizes (e.g. Papovich et al. 2015; van der Wel et al. 2014). At $z > 2$ dust-obscured galaxies are more common, and for these galaxies the Hα luminosity is attenuated (e.g. Reddy et al. 2005; Spitler et al.

[3]We show the stellar mass zeropoint offsets as quoted by the authors, which were carefully derived and corrected for the different IMFs used in $z = 0$ studies. We verified the corrections applied to each datapoint, but could not confirm the IMF-correction by Conselice et al. (2005). The correction from Vergani et al. (2012) was unclear.

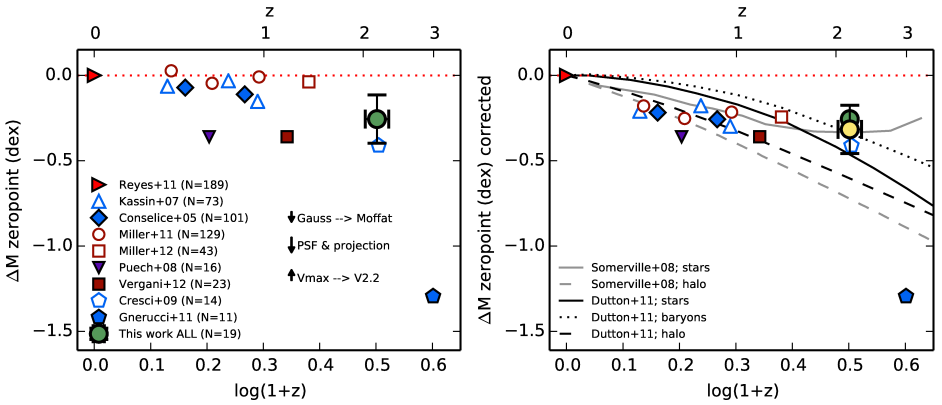

Figure 5.21: Left: the evolution of the stellar mass zeropoint with redshift. The green datapoint with errorbars is for $V_{2.2}$ from our survey at $2.0 < z < 2.5$. The horizontal errorbar indicates the standard deviation of redshift in our sample. Results from other surveys (as quoted in the corresponding papers) are shown with symbols as indicated in the legend. The magnitude of the systematic effects that we have corrected for are indicated by arrows in the middle of the panel, and we have also indicated the effect of choice of velocity estimator ($V_{2.2}$ or V_{max}). Right: we corrected several results from literature for these effects, and found a more consistent picture. We also show the predictions from semi-analytical models (solid lines), and we shifted our datapoint down by 0.06 dex (yellow datapoint) to match V_{max} used in the models. The dashed lines show the model results at the virial radius and the dotted line indicates the evolution for the total baryonic mass.

2014). Samples that are UV or Hα selected may therefore not be a complete distribution of star-forming galaxies at high redshift and changes in incompleteness may mimick evolution with redshift. Mergers and galaxies with irregular morphologies are also more common than at $z = 0$ (e.g. Abraham & van den Bergh 2001; Mortlock et al. 2013). These galaxies have less ordered velocity fields (e.g. Kassin et al. 2007) higher velocity dispersions relative to circular velocities, and are often excluded from Tully-Fisher samples because it is difficult to describe these galaxies with smooth rotating models (Cresci et al. 2009; Gnerucci et al. 2011). At high redshift the angular extent of galaxies is often small compared to the seeing, which may give the appearance that the galaxy is dispersion dominated if the velocity gradient is unresolved. If the selection requires ordered rotation, this leads to biases towards larger galaxies (e.g. Miller et al. 2012).

Here we have attempted to introduce as little selection bias as possible, but it could not be entirely avoided. As we have shown in Section 5.3.5, we have excluded galaxies with poor fits, which tended to be galaxies with smaller sizes and fainter magnitudes than the overall photometric sample. Our sample is therefore also somewhat biased towards the more extended, brighter star-forming galaxy population.

Another caveat when comparing different results from literature is methodology. In many studies the PSF is assumed to be Gaussian, but for our MOSFIRE data a Moffat profile is a better approximation. The difference between using a Gaussian and a Moffat in our modelling leads to a 0.06 dex shift in the stellar mass zeropoint of the Tully-Fisher relation. In addition, several different possibilities exist to model the velocity field, e.g. the one-dimensional arctan model we used here (and also used by e.g. Miller et al. 2011, 2012) or a two-dimensional integrated mass model (Cresci et al. 2009; Gnerucci et al. 2011). Different choices for the radius at which to evaluate velocity exist as well. In some cases R_{80} is used, encapsulating 80% of the optical light (Reyes et al. 2011). In other cases V_{max} is used, or the asymptotic velocity V_a in the arctan model, which is often extrapolated at a radius beyond the optically observed extent of the galaxy (e.g. Weiner et al. 2006b). Most studies in Figure 21 employ V_{max}. We prefer $V_{2.2}$, because it is more robust, and it is used in several other studies (e.g. Miller et al. 2011, 2012). The relation Reyes et al. (2011) derived for V_{80}, which is close to V_{max}, implies a 4% increase in velocities relative to$V_{2.2}$, or a ~ 0.06 dex effect on the inferred stellar mass zeropoint. Lastly, uncertainty on the stellar mass has to be taken into account. We derived our stellar mass from fitting to SEDs obtained from photometry, which depends on several assumptions of the stellar population models. Differences between a Salpeter (1955), Kroupa (2001), Diet Salpeter (Bell et al. 2003) and

Chabrier (2003) IMF are $0.05-0.3$ dex. In addition, different stellar populations models can produce stellar masses different by a factor of 2 (e.g. the review of Conroy 2013). Also, fitting models to SEDs versus applying M/L ratios based on $(g-r)$ colors (Bell et al. 2003) can amount to up to a factor of 2 differences (Reyes et al. 2011). In some cases, full photometry is even missing and less reliable conversions from light to stellar mass have to be made (e.g. Puech et al. 2008).

Another important issue is simply that the datasets between surveys are of a different kind, such as single-slit data versus integral field spectroscopy. For example, Cresci et al. (2009), who use IFS, employ a three-dimensional method, by modelling a datacube with x,y and λ dimensions. This kind of modelling already includes effects from the two-dimensional PSF and projection, whereas (y,λ) modelling of single-slit data using a one-dimensional PSF (as performed in this study and by Conselice et al. (2005); Kassin et al. (2007); Miller et al. (2011, 2012) at high redshift) results in systematically underestimating the velocity.

In summary, methodology and data sets can introduce significant offsets. Expressed in stellar mass zeropoint offsets, these corrections are: 0.06 dex for the choice of PSF (Gaussian versuss Moffat), 0.09 dex for two-dimensional PSF and projection effects (compared to a simple, one-dimensional approach often used in single-slit data), 0.06 dex for the choice of velocity indicator V_{max} versus $V_{2.2}$. Reviewing the studies at different redshifts with this in mind (shown in the left panel of Figure 5.21) we can try to understand these discrepancies. For example, there exist clear differences between Puech et al. (2008) and other studies (e.g. Miller et al. 2011) at $z\sim0.5$ of $0.3-0.4$ dex. The studies of Miller et al. (2011, 2012) are based on one-dimensional modeling of single-slit data, and probably require corrections for two-dimensional PSF smoothing, whereas Puech et al. (2008) have uncertain stellar masses derived on incomplete photometric data. Similarly, there is a large discrepancy between the results of Vergani et al. (2012) and Miller et al. (2012), where Vergani et al. (2012) use two-dimensional IFS data, and evaluate at V_{max} instead of $V_{2.2}$. At $z=2.2$ there is some difference between our study (using $V_{2.2}$) and Cresci et al. (2009) (reporting on V_{max}), but correcting for the velocity indicator the results are more consistent. The apparent strong evolution between $z=1.7$ (Miller et al. 2012) and our results at $z=2.2$ could be explained by the PSF corrections (two-dimensional PSF and Gaussian vs Moffat, amounting to 0.15 dex total). The differences between $0<z<2$ studies and the $z=3$ result of Gnerucci et al. (2011) could be related to their rest-frame UV selection and requirement of optical (rest-UV) spectroscopic redshifts, which tends to be much bluer than near-IR selected samples at $1<z<2$.

Different studies also compare to different $z = 0$ relations. The most common references are the results from Bell & de Jong (2001), Pizagno et al. (2005), and Reyes et al. (2011). These are based on different IMF, choice of velocity indicator, and method to derive stellar mass. In most high redshift studies, estimates of the evolution of the Tully-Fisher relation are derived very carefully, but a major factor of uncertainty is the derived slope of the relation at $z = 0$. For example, Bell & de Jong (2001) derive a much steeper slope, with $A = 1/4.5$, than Pizagno et al. (2005) and Reyes et al. (2011). This is illustrated by Vergani et al. (2012), who find a -0.36 ± 0.11 dex evolution compared to Pizagno et al. (2005) and only -0.05 ± 0.16 dex compared to Bell & de Jong (2001). We show the first result in Figure 5.21, because it is more consistent with our comparison to $z = 0$ based on the study by Reyes et al. (2011). Other studies that compare to Bell & de Jong (2001) are those by Conselice et al. (2005), Kassin et al. (2007), Cresci et al. (2009) and Gnerucci et al. (2011). Miller et al. (2011) and Miller et al. (2012) use the $z = 0$ relation from Reyes et al. (2011) and Puech et al. (2008) derive the stellar mass Tully-Fisher relation both at $z = 0$ and $z = 0.6$ and compare internally.

5.5.2 Interpretation of the evolution of the Tully-Fisher relation

Taking into account the various systematic differences between studies at high redshift, we have endeavoured to revise earlier results and interpret them in a framework of semi-analytical models. As a first step, encouraged by the fact that reasonably expected corrections (related to PSF and velocity indicator) improve agreement between various earlier work at $0 < z < 2.0$, we went ahead and corrected previous results as described in Section 5.5.1.

Single-slit observations are corrected for two-dimensional PSF effects and choice of PSF effects (Conselice et al. 2005; Kassin et al. 2007; Miller et al. 2011, 2012)[4]. We also applied a shift to our result and that of Miller et al. (2011, 2012) to correct for the use of $V_{2.2}$ instead of V_{max}, as the latter was used by the models. The results are shown in the right panel of Figure 5.21. With the various offsets from the $z = 0$ relation corrected, we find a much more consistent picture. The datapoints from Puech et al. (2008) at $z = 0.6$ and Gnerucci et al. (2011) at $z = 3$ are still outliers. As mentioned, Puech et al. (2008) probably have uncertain stellar masses, whereas the offset of

[4] Please note that these corrections are rather coarse. For example the precise correction for two-dimensional PSF and projection effects depends on the slitwidth as well as the angular size of the galaxy. The effect may thus be smaller below $z < 1$. The shifts applied here are not meant to be exact, nor to find fault with individual studies. Rather our aim is to discuss a possible explanation for the inconsistencies between different surveys.

Gnerucci et al. (2011) could be related to their rest-frame UV selection and requirement of optical (rest-UV) spectroscopic redshifts.

We now discuss the observations in a framework of semi-analytical models. In the right panel of Figure 5.21 we show the predictions by Somerville et al. (2008) and Dutton et al. (2011). Dutton et al. (2011) who previously reported to be able to reproduce the evolution of the stellar mass Tully-Fisher relation observed by Conselice et al. (2005), Kassin et al. (2007) and Cresci et al. (2009). The models of Somerville et al. (2008) and Dutton et al. (2011) are roughly similar, with the most important difference that in the model of Dutton et al. (2011) the baryonic component of galaxies contains both gas and stars, whereas Somerville et al. (2008) assumed disks without gas.

Our result at $z = 2.2$ is consistent with the predictions from the models for the stellar mass zeropoint and the baryonic zeropoint. The predicted evolution from $z = 0$ to $z = 1$ is smaller in the models than in the observations (see footnote 4). At $1 < z < 2.5$ there is good agreement between the models and the observations, with a gradual zeropoint evolution. With the various offsets from the $z = 0$ relation corrected, we find a much more consistent picture, which is in general agreement with the semi-analytic models, in which individual galaxies grow in a self-similar way, i.e., along the scaling relations. The datapoints from Gnerucci et al. (2011) is still an outlier, based on a small sample of 11 galaxies, with possible selection biases. If representative however, it may point to non-self similar evolution at high redshift ($z > 3$).

Our finding of large velocity dispersions ($V/\sigma \sim 3$) confirms the emerging picture that at high redshift disk galaxies are more often pressure supported. The observed evolution is in agreement with the predictions from semi-analytic models that disks form roughly along the Tully-Fisher relation by the accretion of gas, but the offset in stellar mass could also reflect the conversion from gas to stars over time. The evolution in disk rotation also depends on feedback, with more modest evolution if feedback is taken into account (Sales et al. 2010). An important assumption in the model of Dutton et al. (2011) is that galaxy disks are entirely supported by rotation at all redshifts. This is likely not true, as indicated by the general increase of velocity dispersion in high redshift galaxies (e.g. Kassin et al. 2007, 2012; Wisnioski et al. 2015; Price et al. 2015). At high redshift, galaxies more often have disordered kinematics, which causes a larger velocity dispersion, hence the alternative estimator S_{05}. On average the ratio V/σ decreases towards higher redshift, and also depends on the Hα SSFR and stellar mass, in the sense that V/σ decreases with increasing SSFR and towards lower stellar mass (Price et al. 2015). This effect is related to the gas fraction, which is smaller in higher mass galaxies with lower SFR, and decreases with time.

A more detailed assessment of the interplay between gas, stars and dark matter at high redshift is needed in both observations and models to explain the observed evolution of the Tully-Fisher relation. It is clear that the issue is complex, because apart from model uncertainties producing observational results that are consistent across redshifts remains challenging.

5.6 Summary

In this work we have derived the stellar mass-velocity and S_{05}-velocity scaling relations at $2.0 < z < 2.5$, making use of 22 MOSFIRE single-slit spectra of 19 star-forming galaxies, as part of the ZFIRE survey. The diagnostic used was the Hα emission line, and we fitted model spectral image stamps to the data to trace the rotational velocities and dispersions of the Hα gas in the galaxies.

We conducted a careful check of systematics and corrected our results where necessary, and subsequently fitted and interpreted the stellar mass Tully-Fisher evolution to $0 < z < 2.5$. We found the following main results:

- The MOSFIRE PSF can be best approached by a Moffat function with $\beta = 2.5$, instead of a Gaussian. Assuming a Gaussian PSF instead leads to 4% underestimated velocities on average, implying a 0.06 dex effect on the stellar mass zeropoint of the Tully-Fisher relation.

- Two-dimensional PSF and slit projection effects cause flux from lower velocity regions of a galaxy to be mixed within the slit. For a slitwidth of $0.7''$ and a median scenario with a seeing of $0.8''$ and inclination of 45 deg, this results in 6% smaller velocities, or a 0.09 dex effect on the Tully-Fisher stellar mass zeropoint.

- Taking this into account, we derived the stellar mass Tully-Fisher relation $\log V_{2.2} = (2.19 \pm 0.049) + (0.247 \pm 0.094)(\log M/M_{\odot} - 10)$ and inferred an evolution of $\Delta \mathrm{M}/M_{\odot} = -0.26 \pm 0.14$ dex compared to $z = 0$.

- The best-fit modified Tully-Fisher relation, the S_{05}-velocity relation, is $\log S_{05} = (2.10 \pm 0.033) + (0.228 \pm 0.059)(\log M/M_{\odot} - 10)$, with an inferred zeropoint evolution of $\Delta \mathrm{M}/M_{\odot} = -0.54 \pm 0.13$ dex compared to $0.1 < z < 1.2$.

- We reviewed previous results in literature, which have strong discrepancies between IFS and single-slit studies over a large redshift range. We give as an explanation for these discrepancies that single-slit results may suffer from PSF and projection effects.

- After correcting literature values for the various systematic effects, the overall evolution of the stellar mass zeropoint at $0 < z < 2.5$ is reasonably well matched by the semi-analytic models of Dutton et al. (2011). However, in detail some discrepancies with the models remain, especially at $0.5 < z < 1$ and at $z > 3$. Futhermore, our data confirm previous observations of increased contributions from non-rotationally supported galaxies, which are not included in the models. The increase of the average velocity dispersion towards higher redshift is related to the higher gas fractions in galaxies. It is therefore possible that the evolution in $\Delta M/M_{\odot}$ (partly) reflects the conversion from gas to stars.

To explain the observed evolution of the Tully-Fisher relation, a more detailed assessment of the interplay between gas, stars and dark matter at high redshift is needed in both observations and models. With current facilities such as MOSFIRE on Keck, that allow for efficient observations of large numbers of galaxies in the near-IR, this has become a real possibility. However, future studies that use consistent methodologies are needed to derive a better picture of the evolution of the Tully-Fisher relation.

5.7 Acknowledgements

We wish to thank the W. M. Keck Observatory support staff for their enthusiastic support. We recognize and acknowledge the very significant cultural role and reverence that the summit of Mauna Kea has always had within the indigenous Hawaiian community. We thank Gabe Brammer and Danilo Marchesini for kindly sharing a MOSFIRE K_s-band image, and George Bekiaros for help with software. We are grateful to Susan Kassin for in-depth discussions. CMSS gratefully acknowledges the support of the Australian Government through an Endeavour Research Fellowship.

Bibliography

Abraham, R. G., & van den Bergh, S. 2001, Science, 293, 1273

Bamford, S. P., Aragón-Salamanca, A., & Milvang-Jensen, B. 2006, *MNRAS*, 366, 308

Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, *ApJS*, 149, 289

Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503

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

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682

Chabrier, G. 2003, Publications of the Astronomical Society of the Pacific, 115, 763

Conroy, C. 2013, *ARA&A*, 51, 393

Conselice, C. J., Bundy, K., Ellis, R. S., et al. 2005, *ApJ*, 628, 160

Courteau, S. 1997, *AJ*, 114, 2402

Cresci, G., Hicks, E. K. S., Genzel, R., et al. 2009, *ApJ*, 697, 115

Dutton, A. A., van den Bosch, F. C., Faber, S. M., et al. 2011, *MNRAS*, 410, 1660

Fall, S. M., & Efstathiou, G. 1980, *MNRAS*, 193, 189

Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, *ApJ*, 706, 1364

Franx, M., & de Zeeuw, T. 1992, *ApJL*, 392, L47

Franx, M., Illingworth, G., & Heckman, T. 1989, *AJ*, 98, 538

Franx, M., van Dokkum, P. G., Schreiber, N. M. F., et al. 2008, *ApJ*, 688, 770

Freeman, K. C. 1970, *ApJ*, 160, 811

Gnerucci, A., Marconi, A., Cresci, G., et al. 2011, *A&A*, 528, A88

Hammer, F., Flores, H., Elbaz, D., et al. 2005, *A&A*, 430, 115

Haynes, M. P., & Giovanelli, R. 1984, *AJ*, 89, 758

Kassin, S. A., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, 660, L35

—. 2012, *ApJ*, 758, 106

Kelly, B. C. 2007, *ApJ*, 665, 1489

Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221

Kriek, M., Shapley, A. E., Reddy, N. A., et al. 2015, *ApJS*, 218, 15

Kroupa, P. 2001, *MNRAS*, 322, 231

Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, 379, 1599

McLean, I. S., Steidel, C. C., Epps, H., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 1

Miller, S. H., Bundy, K., Sullivan, M., Ellis, R. S., & Treu, T. 2011, *ApJ*, 741, 115

Miller, S. H., Ellis, R. S., Sullivan, M., et al. 2012, *ApJ*, 753, 74

Mo, H. J., Mao, S., & White, S. D. M. 1998, *MNRAS*, 295, 319

Mortlock, A., Conselice, C. J., Hartley, W. G., et al. 2013, *MNRAS*, 433, 1185

Papovich, C., Momcheva, I., Willmer, C. N. A., et al. 2010, *ApJ*, 716, 1503

Papovich, C., Labbé, I., Quadri, R., et al. 2015, *ApJ*, 803, 26

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, 139, 2097

Persson, S. E., Murphy, D. C., Smee, S., et al. 2013, *PASP*, 125, 654

Pizagno, J., Prada, F., Weinberg, D. H., et al. 2005, *ApJ*, 633, 844

—. 2007, *AJ*, 134, 945

Price, S. H., Kriek, M., Shapley, A. E., et al. 2015, ArXiv e-prints, arXiv:1511.03272

Puech, M., Flores, H., Hammer, F., et al. 2008, *A&A*, 484, 173

Reddy, N. A., Erb, D. K., Steidel, C. C., et al. 2005, *ApJ*, 633, 748

Reyes, R., Mandelbaum, R., Gunn, J. E., Pizagno, J., & Lackner, C. N. 2011, *MNRAS*, 417, 2347

Sales, L. V., Navarro, J. F., Schaye, J., et al. 2010, *MNRAS*, 409, 1541

Salpeter, E. E. 1955, *ApJ*, 121, 161

Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, 172, 1

Sersic, J. L. 1968, Atlas de galaxias australes

Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, ArXiv e-prints, arXiv:1403.3689

Somerville, R. S., Barden, M., Rix, H.-W., et al. 2008, *ApJ*, 672, 776

Spitler, L. R., Labbé, I., Glazebrook, K., et al. 2012, *ApJL*, 748, L21

Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJL*, 787, L36

Steidel, C. C., Rudie, G. C., Strom, A. L., et al. 2014, *ApJ*, 795, 165

Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2015, *ApJL*, 808, L29

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, *ApJ*, 783, 85

—. 2015, ArXiv e-prints, arXiv:1510.06072

Tully, R. B., & Fisher, J. R. 1977, *A&A*, 54, 661

van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, *ApJS*, 203, 24

van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ArXiv e-prints, arXiv:1404.2844

van Dokkum, P. G., & Franx, M. 2001, *ApJ*, 553, 90

Vergani, D., Epinat, B., Contini, T., et al. 2012, *A&A*, 546, A118

Weiner, B. J., Willmer, C. N. A., Faber, S. M., et al. 2006a, *ApJ*, 653, 1027

—. 2006b, *ApJ*, 653, 1049

Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJL*, 754, L29

Willick, J. A. 1999, *ApJ*, 516, 47

Wisnioski, E., Förster Schreiber, N. M., Wuyts, S., et al. 2015, *ApJ*, 799, 209

Yuan, T., Nanayakkara, T., Kacprzak, G. G., et al. 2014, *ApJL*, 795, L20

6

Samenvatting van dit proefschrift in het Nederlands

6.1 Inleiding

Een van de belangrijkste onopgeloste vraagstukken in de hedendaagse sterrenkunde is: *Hoe worden sterrenstelsels gevormd en hoe evolueren ze?* Een vraag die niet zo gemakkelijk valt te beanwoorden. Dit komt omdat de ontwikkeling van een stelsel plaats heeft gedurende een paar miljard jaar en de afstand tot de te bestuderen stelsels ontzettend groot is. De evolutie van sterrenstelsels hangt nauw samen met de verdeling van donkere materie in het heelal en een zo mogelijk nog groter mysterie betreft de eigenschappen van die donkere materie.

De oorsprong van sterrenstelsels gaat volgens sommige theorieën terug tot een korte periode van inflatie vlak na de oerknal, waarin het heelal zeer snel uitdijdde. Door die snelle uitdijing bleven zogenoemde quantumfluctuaties in stand en die vormden de basis van de latere dichtheidsfluctuaties waaruit sterrenstelsels zijn ontstaan. In het huidige universum zijn sterrenstelsels alom tegenwoordig. We leven zelf in een sterrenstelsel, namelijk de Melkweg, zichtbaar als een lichtgevende band aan de hemel.

Zover in de tijd als tot de oerknal kunnen we niet terugkijken, omdat het heelal pas na 377 000 jaar doorschijnend werd. Voor dat moment was de dichtheid van het heelal zo groot, dat lichtdeeltjes altijd vrijwel meteen een interactie ondergingen met andere deeltjes. De allereerste lichtdeeltjes die zich vrijelijk over lange afstanden door de ruimte konden bewegen nemen we waar als de kosmische achtergrondstraling. Studies van de kosmische achtergrondstraling hebben al geleid tot veel inzichten, bijvoorbeeld over de samenstelling van het heelal. De materie waaruit sterren, planeten, de flora en fauna op aarde en de waarneembare gedeeltes van sterrenstelsels bestaan, oftewel baryonische materie, beslaat maar 4.9% van de totale energiedicht-

heid van het universum. Dat is vijf keer zo weinig als de totale hoeveelheid donkere materie.

In de kosmische achtergrondstraling zien we ook temperatuurfluctuaties: het begin van de latere structuur van donkere materie en sterrenstelsels. De eerste sterrenstelsels zijn ontstaan na het door de zwaartekracht ineenstorten van materie in wat we ook wel halo's van donkere materie noemen. Baryonische materie, meestal waterstofgas, beweegt zich in dat proces helemaal tot in het centrum van zo'n halo en hoopt zich opeen tot sterren in een sterrenstelsel.

Een sterrenstelsel kan verschillende vormen aannemen. Zo is de Melkweg, voor zover wij dat vanuit ons perspectief kunnen bepalen, een spiraalvorming stelsel. Een eigenschap van spiraalstelsels is dat ze over het algemeen stervormend zijn. Dat betekent dat uit wolken gas in het stelsel nog altijd relatief veel sterren worden geboren. Een spiraalvormig stelsel kenmerkt zich daarnaast door een wat blauwere kleur, vanwege het sterke ultraviolette spectrum van jonge sterren. Zo'n stelsel kan al wel heel massief zijn. Zo heeft de Melkweg ongeveer 100 miljard keer zoveel massa als één keer onze zon.

Een ander type stelsel is elliptisch van vorm, niet blauw van kleur maar rood, en heeft gemiddeld een oude populatie van sterren, terwijl er weinig nieuwe sterren meer bijkomen. Vanwege deze eigenschappen wordt een dergelijk stelsel ook wel passief of "rood en dood" genoemd. Passieve stelsels behoren voornamelijk tot de meest massieve. Een van de vragen die astronomen bezighouden met betrekking tot passieve stelsels is: *Wat is de reden dat er in sterrenstelsels na verloop van tijd geen nieuwe sterren meer worden gevormd?* Er zijn verschillende verklaringen voor het "uitdoven" van een stelsel. Deze hebben er allemaal mee te maken dat het gas waaruit sterren worden gevormd niet genoeg afkoelt – een voorbeeld is het invallende gas in de meest massieve halo's – of wordt verstoord door factoren in de omgeving. Zo'n factor is bijvoorbeeld de grote hoeveelheid energie die vrijkomt bij de aanwas van materie rond het centrale massieve zwarte gat in een stelsel of de energie die vrijkomt door exploderende, massieve sterren. Minder massieve sterren kunnen materie uit hun buitenste lagen verliezen en ook op die manier invloed hebben op hun omgeving.

In hoeverre een van deze processen de evolutie van een sterrenstelsel bepaalt, hebben we met waarnemingen nog niet kunnen aantonen. We weten zelfs nog niet wanneer stelsels voor het eerst passief werden. Er zijn wel aanwijzingen dat massieve, passieve stelsels al 11 miljard jaar geleden bestaan hebben – relatief kort na de oerknal – en dat zou wijzen op de mogelijkheid van een snelle evolutie met een efficient uitdovingsmechanisme.

Onderzoek naar de evolutie van sterrenstelsels gebeurt de laatste jaren

op grote schaal, door data te verzamelen van tienduizenden tot honderdduizenden sterrenstelsels, en daarbij de grenzen te verleggen tot steeds dieper in het heelal, ofwel steeds verder terug in de tijd. Dit wordt meestal in internationaal verband gedaan en gebruikmakend van de meest geavanceerde telescopen en instrumenten. Door de uitdijing van het heelal nemen we het licht van verre bronnen waar op een langere golflengte dan oorspronkelijk, een effect dat we roodverschuiving noemen. Dit heeft tot gevolg dat observaties zich nu richten op infrarood licht, een uitdaging vanwege de optische dichtheid van de atmosfeer in het infrarood.

In dit proefschrift presenteren we een survey van verre sterrenstelsels: The FourStar Galaxy Evolution Survey (ZFOURGE) en maken we vervolgens van de data gebruik om sterrenstelsels diep in het heelal te bestuderen. We beschrijven de ontdekking van de verste passieve sterrenstelsels tot nu toe en analyseren hun eigenschappen. We onderzoeken ook stervormende stelsels in het verre universum, hoewel op iets kleinere afstand. Van deze stelsels meten we met behulp van spectroscopische waarnemingen de rotatiesnelheid en relateren deze aan de totale stermassa van de stelsels. Deze relatie wordt ook wel de Tully-Fisher relatie genoemd (uitgevonden door de astronomen R. Tully en J. Fisher) en is al tot in detail bestudeerd voor sterrenstelsels die zich dicht bij ons bevinden. Door deze ook op hoge roodverschuiving te bestuderen, hopen we iets te leren over de aanwas van stermassa en de relatie met donkere materie gedurende de actieve periode van sterrenstelsels.

6.2 Dit proefschrift

Hoofdstuk 1 van dit proefschrift geeft een algemene introductie over sterrenstelsels, wat we al weten over hun evolutie en hoe we ze, gebruikmakend van surveys, in het verre heelal bestuderen.

In **hoofdstuk 2** presenteren we de data van ZFOURGE. ZFOURGE beslaat 45 nachten aan waarnemingen in het infrarood met het FourStar instrument op de Magellan Baade Telescoop in Chili. We analyseren de diepte van de opnames en beschrijven hoe we meer dan 70 000 lichtbronnen hebben gedetecteerd, tot zeer zwakke lichtsterktes. De opnames zijn gemaakt op zes verschillende golflengtes en vervolgens gecombineerd met data van andere instrumenten met een groot spectraal bereik, van het ultraviolet tot het verre infrarood. Hiermee stellen we voor iedere bron een spectraal distributie samen waarmee we zeer accuraat de roodverschuiving (een afstandsindicator) van de sterrenstelsels kunnen bepalen, evenals andere eigenschappen zoals leeftijd en de totale massa in sterren.

Passieve en stervormende stelsels hebben verschillende spectraal distri-

buties, die leiden tot de typische rode en blauwe kleuren. Om de stelsels van elkaar te onderscheiden, wordt daarom vaak een test met twee kleuren gebruikt, gebaseerd op de intensiteit van het licht in drie verschillende filters. De nauwkeurigheid van deze test is nog niet bewezen voor stelsels op hoge roodverschuiving. In **hoofdstuk 2** onderzoeken we met data in het verre infrarood, afkomstig van de Spitzer en Herschel ruimtetelescopen, de sterformatie van verre sterrenstelsels in ZFOURGE en bevestigen we het bestaan van massieve, passieve stelsels in de afgelopen 11.5 miljard jaar.

In **hoofdstuk 3** gaan we nog een stap verder en onderzoeken we of we op nog grotere afstand passieve stelsels kunnen vinden. Een moeilijkheid daarbij is dat stervormende stelsels met veel stofdeeltjes, die het ultraviolette licht van jonge sterren absorberen, ook een rode kleur hebben en sterk lijken op passieve stelsels. De test met twee kleuren helpt bij het onderscheiden van passieve stelsels en rode en blauwe stelsels met sterformatie. We kijken echter ook naar de data in het verre infrarood van de Spitzer en Herschel ruimtetelescopen, waarmee we de sterformatie in stelsels met veel stof kunnen bepalen.

Aan de hand van de kleurtest kunnen we met de diepe data van ZFOURGE 19 masssieve, passieve stelsels onderscheiden, op een gemiddelde afstand van 12 miljard jaar. Hiervan hebben er vier sterke detecties in het verre infrarood. Dat duidt erop dat er in deze stelsels mogelijk juist wel veel sterformatie aan de gang is. Voor de overige 15 stelsels vinden we geen duidelijke aanwijzing voor actieve sterformatie, maar de data is niet diep genoeg om het voor de stelsels individueel uit te sluiten. We kijken daarom naar het gemiddelde van de verre infrarood data en leiden af dat de sterformatie in de passieve stelsels 10 keer lager is dan gemiddeld voor stervormende stelsels op dezelfde afstand en met dezelfde massa.

De 15 passieve stelsels zijn zeer massief, met een gemiddelde totale stermassa van ongeveer 80 miljard keer de massa van de zon. Gezien de jonge leeftijd van het heelal (ongeveer 1.6 miljaar jaar), betekent dit dat de sterformatie in deze stelsels ooit extreem hoog moet zijn geweest. We vergelijken onze bevindingen met huidige observaties van blauwe stervormende stelsels uit de periode waarin deze 15 massieve stelsels zich hebben gevormd, maar vinden niet genoeg voorbeelden van zulke hoge sterformatie. We speculeren dat de jonge versies van de 15 passieve stelsels wel eens zeer actieve stelsels met veel stofdeeltjes kunnen zijn geweest, die we zouden kunnen waarnemen door te kijken op submillimeter golflengtes.

Eerdere waarnemingen van passieve stelsels op hoge roodverschuiving hebben laten zien dat deze toen veel compacter waren dan nu en ook veel kleiner dan stervormende stelsels. De huidige interpretatie is dat deze stelsels

van binnen naar buiten groeien. Eerst wordt er een compacte kern gevormd, waarna er nog aanwas plaatsvindt door interactie met andere sterrenstelsels. Een interessante vraag daarbij is of het mechanisme dat leidt tot een compacte kern er ook voor zorgt dat het sterrenstelsel uitdooft. In **hoofdstuk 4** onderzoeken we de groottes van de 15 passieve stelsels in ZFOURGE en plaatsen die in een evolutionaire context. We maken gebruik van data afkomstig van de Hubble ruimtetelescoop, met een zeer hoge resolutie, en concluderen dat de passieve sterrenstelsels inderdaad zeer compact zijn. Ook passen ze binnen de evolutionaire trend van steeds kleinere stelsels op steeds grotere afstand. Wat betreft stervormende stelsels 12 miljard jaar geleden vinden we juist dat die gemiddeld veel groter zijn. Stervormende compacte stelsels, die mogelijk de voorlopers zijn van compacte passieve stelsels, zijn zeer zeldzaam. Dit duidt erop dat de formatie van een compacte kern op een zeer korte tijdschaal gebeurt en mogelijk verborgen is door grote hoeveelheden stof in de sterrenstelsels.

In **hoofdstuk 5** onderzoeken we de Tully-Fisher relatie (stermassa versus rotatiesnelheid) voor verre stervormende stelsels, zo'n 10 tot 11 miljard jaar geleden. We onderzoeken of en hoe sterk deze relatie verandert met de tijd, wat een belangrijke aanwijzing zou zijn over hoe efficient stelsels gas omzetten in sterren en hoe de stermassa zich ontwikkelt ten opzichte van de donkere materie. Er zijn al een aantal studies verricht voor verre sterrenstelsels. Deze studies maken gebruik van verschillende methodes en geven verschillende resultaten. Een duidelijk beeld van de evolutie van de Tully-Fisher relatie is er dus nog niet.

We meten de rotatiesnelheiden van 19 massieve sterrenstelsels aan de hand van infrarode spectra, gemeten met het MOSFIRE instrument op de Keck-I telescoop in Hawaii. De metingen zijn onderdeel van de spectroscopische survey ZFIRE, gebaseerd op lichtbronnen in ZFOURGE. We bestuderen een optische emissielijn die voortkomt uit waterstofgas, maar vanwege de grote afstand is verschoven naar het infrarood. Vanwege de rotatie binnen een stelsel is de lijn uitgerekt, met een gedeelte dat meer en een gedeelte dat minder roodverschoven is. Hieruit leiden we de rotatiesnelheid af. We onderzoeken daarbij uitgebreid mogelijke systematische effecten, om er zeker van te zijn dat we de rotatiesnelheid niet onderschatten.

We vinden dat stervormende stelsels op hoge roodverschuiving bij een bepaalde rotatiesnelheid gemiddeld een lagere totale stermassa hebben, wat duidt op evolutie van de Tully-Fisher relatie. We vergelijken ons resultaat met eerdere studies, maar corrigeren die eerst voor de door ons onderzochte systematische effecten. Dit is een belangrijke stap, omdat de eerdere resultaten zeer sterk van elkaar verschilden, en we nu voor het eerst kunnen aan-

tonen wat daarvan de oorzaak zou kunnen zijn, namelijk het gebruik van verschillende methodes om de rotatiesnelheid te meten. Het uiteindelijke resultaat wijst op een geleidelijke evolutie over de laatste 10 tot 11 miljard jaar,

We vergelijken de waarnemingen ook met modellen en vinden dat de voorspelde evolutie redelijk overeenkomt met onze waarnemingen op hoge roodverschuiving (10 tot 11 miljard jaar geleden), maar niet met de gecorrigeerde resultaten van eerdere studies in de periode daarna. Een verklaring zou kunnen zijn dat sterren in stelsels op hoge roodverschuiving vaak mindere geordende banen volgen, terwijl dat niet is meegenomen in de modellen waarmee we vergelijken. Waar het verschil precies vandaan komt blijft gissen, totdat we met betere data een duidelijker beeld van de Tully-Fisher relatie kunnen vormen.

6.3 Blik op de toekomst

We hebben in dit proefschrift aangetoond dat passieve sterrenstelsels al zeer vroeg, slechts 1.6 miljard jaar na de oerknal, in het universum voorkwamen. De volgende stap is nu om het vroege bestaan en de gemiddelde eigenschappen, zoals massa en grootte, van deze stelsels toe te voegen aan modellen. Er wacht nog wel een andere taak, namelijk het bevestigen van de vondst door het meten van de precieze roodverschuiving met behulp van spectroscopie. De mogelijkheid tot het meten van infrarode spectra van sterrenstelsels is er al, bijvoorbeeld met MOSFIRE. In 2018 staat ook de lancering van de James Webb op het programma, die gaat zorgen voor nog diepere infrarood data, waarmee we mogelijk de allereerste passieve stelsels, en wellicht ook hun voorlopers, kunnen gaan vinden.

Met MOSFIRE, de James Webb en nog een aantal andere instrumenten kunnen we ook voor grote hoeveelheden stervormende sterrenstelsels spectroscopische waarnemingen gaan doen. Het is van belang om voor een breed roodverschuivingsbereik en met een consistente methodologie de rotatiesnelheid van de stelsels te bepalen, om de evolutie van de Tully-Fisher relatie nog beter te kwantificeren voor we proberen die te begrijpen.

Tenslotte is het nog de moeite waard om de ALMA telescoop in de Atacama woestijn in Chili te noemen. Met deze telescoop kunnen we submillimeter opnames maken van zeer actieve, maar door stof omgeven sterrenstelsels en beter inzicht krijgen in alle fases van de evolutie van sterrenstelsels.

7

Curriculum vitae

I was born on the 20th of September in the year 1987, in Leidschendam in the Netherlands. I lived my first 1.5 years in the town of Voorburg, before I moved to the village of Leiderdorp, where I grew up. There I attended PCBS De Driemaster from 1991 to 1999. In that period I also developed my interest in space, fascinated as I was by the beautiful pictures of planets and galaxies that were of high quality already in the nineties, and initially even wanting to become an astronaut.

In 1999 I started attending Stedelijk Gymnasium Leiden, and graduated in 2005 with a focus on natural and life sciences. The curriculum also included Latin and Greek, and a strong emphasis on the English language. In my final year I participated in the NOVA waarneemwedstrijd INT (a "write-the-best-observing-proposal" competition), and was awarded a trip to the Isaac Newton Telescope at the Roque de los Muchachos Observatory on La Palma, together with three other students from different schools. This was my first observing trip and despite extremely bad weather, I was able to obtain optical imaging of the merging Antennae Galaxies in enough filters to make a false colour image and observe young blue and old red populations of stars, which was the aim of my project.

In 2005 I started my Bachelor's degree studies at Leiden Observatory, the astronomy department of Leiden University, and graduated in 2009. I then continued my studies to obtain my Master's degree, and graduated in 2011. The curriculum included courses in mathematics, physics, astrophysics and programming languages, with special attention paid to computing, data analysis and presentation techniques. The Bachelor's also included an observing trip to the same INT on La Palma, this time to study a binary star system.

An important part of the curriculum was research. For my Bachelor's thesis I studied the effects of pixel binning on mass estimates of distant globular star clusters, together with Marinus Israel and dr. Bernhard Brandl. During my Master's degree studies I first studied photon dominated regions in the

Magellanic Clouds, using far-infrared data from the Infrared Space Observatory, under supervision of prof. dr. Frank Israel. I then moved on to the field of galaxy formation, to write my Master's thesis under supervision of prof. dr. Joop Schaye and dr. Olivera Rakic. I analysed Quasi-Stellar absorption spectra obtained by instruments on the Keck-I telescope in Hawaii and the Very Large Telescope in Chile, to study the enrichment of intergalactic gas.

While I was an undergraduate in Leiden I was involved in various extracurricular activities. In 2005 I joined the choir of LSKO Collegium Musicum (CM), and remained a member of this student association until 2013. In 2007 I joined the board of CM for a year as treasurer, and was involved in organising concerts and managing the financials. From 2009 to 2011 I held the same position in their Anniversary Committee. Within the astronomy department, I was a member of the Education Advisory Committee from 2006 to 2008, assessing the quality of the curriculum through feedback from students and meeting regularly with faculty members. During this entire period I was a tutor in natural and life science courses for secondary school students.

I commenced my PhD studies at Leiden University in 2011, with dr. Ivo Labbé as daily supervisor and co-promotor and prof. dr. Marijn Franx as promotor, on the topic of galaxy evolution. I was part of an international collaboration of astronomers from the United States (US), Australia, and The Netherlands, and attended meetings at Swinburne Institute of Technology in Melbourne (Australia), Macquarie University in Sydney (Australia), Texas A&M University in College Station (US), and Carnegie Institute of Technology in Pasadena (US). The collaboration was originally established to share observing time at the 6.5m Magellan Baade Telescope at Las Campanas Observatory in Chile, and use the then newly commissioned instrument FourStar for an ambitious survey of galaxies at near-infrared wavelengths: the FourStar Galaxy Evolution Survey, or ZFOURGE, with dr. Ivo Labbé as principal investigator. I was actively involved in acquiring the data and the subsequent analysis of the images and travelled twice to Chile for observations, once in December 2011 and once in February 2012.

In 2014 I was awarded an Endeavour Research Fellowship from the Australian government to spend four months at Swinburne Institute of Technology in Melbourne, where I worked on a follow-up program: ZFIRE, with close collaborators and ZFOURGE members prof. dr. Karl Glazebrook and dr. Glenn Kacprzak, who are experts on spectral analysis. To obtain additional data for ZFIRE I travelled to the 10m Keck I telescope on Mauna Kea in Hawaii (US) in March 2015.

In the first year of my PhD studies I was teaching assistent for "Inleiding Astrofysica", an introductory course for the Bachelor's. In the second and

third years I assisted at “Modern Onderzoek”, a course introducing Bachelor students to modern research in astronomy. I attended two schools, the NOVA Fall School in Dwingeloo in 2011 and the 30th Jerusalem Winter School in Theoretical Physics in Jerusalem (Israel) in 2013.

Finally, I presented my work at the European Week of Astronomy and Space Science in Turku (Finland) in 2013, the Winter Conference in Aspen (US) in 2014, and the Annual Scientific Meeting of the Astronomical Society of Australia in Sydney (Australia) in 2014.

During my PhD studies I developed an interest in data analysis and statistics. I am now a data scientist at Stedin, an electric utility company in Rotterdam.

8

List of publications

In preparation

Straatman, C. M. S., & et al. 2016, ZFIRE: the evolution of the stellar mass Tully-Fisher relation to redshift $2.0 < z < 2.5$ with MOSFIRE, *ApJ*

Submitted to journal

Straatman, C. M. S., Spitler, L. R., Quadri, R., et al. 2016, The FourStar Galaxy Evolution Survey: ultraviolet to far-infrared catalogs, medium-bandwidth photometric redshifts, and stellar population properties; analysis of photometric redshift accuracy and confirmation of quiescent galaxies to $z \sim 3.5$, *ApJ*

Published, 1st author

Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, A Substantial Population of Massive Quiescent Galaxies at z ~ 4 from ZFOURGE, *ApJL*, 783, L14

Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2015, The Sizes of Massive Quiescent and Star-forming Galaxies at z ~ 4 with ZFOURGE and CANDELS, *ApJL*, 808, L29

Published, 2nd author

Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, Exploring the z = 3-4 Massive Galaxy Population with ZFOURGE: The Prevalence of Dusty and Quiescent Galaxies, *ApJL*, 787, L36

Published, contributing author

Allen, R. J., Kacprzak, G. G., Spitler, L. R., et al. 2015, The Differential Size Growth of Field and Cluster Galaxies at z = 2.1 Using the ZFOURGE Survey, *ApJ*, 806, 3

Bouwens, R. J., Oesch, P. A., Labbe, I., et al. 2015, Most Luminous z=9-10 Galaxies: A First Determination of the Bright End of the z~9 and z~10 UV Luminosity Functions using all five CANDELS Fields, ArXiv e-prints, arXiv:1506.01035

Cowley, M. J., Spitler, L. R., Tran, K.-V. H., et al. 2016, ZFOURGE catalogue of AGN candidates: an enhancement of 160-μm-derived star formation rates in active galaxies to z = 3.2, *MNRAS*, 457, 629

Forrest, B., Tran, K.-V. H., Tomczak, A. R., et al. 2016, UV to IR Luminosities and Dust Attenuation Determined from ~4000 K-Selected Galaxies at 1<z<3 in the ZFOURGE Survey, ArXiv e-prints, arXiv:1602.01096

Kacprzak, G. G., Yuan, T., Nanayakkara, T., et al. 2015, The Absence of an Environmental Dependence in the Mass-Metallicity Relation at z = 2, *ApJL*, 802, L26

Kawinwanichakij, L., Papovich, C., Quadri, R. F., et al. 2014, The Distribution of Satellites around Massive Galaxies at 1 < z < 3 in ZFOURGE/CANDELS: Dependence on Star Formation Activity, *ApJ*, 792, 103

Kawinwanichakij, L., Quadri, R. F., Papovich, C., et al. 2016, Satellite Quenching and Galactic Conformity at 0.3 < z < 2.5, *ApJ*, 817, 9

Kewley, L. J., Yuan, T., Nanayakkara, T., et al. 2015, Z-FIRE: ISM properties of the z = 2.095 COSMOS Cluster, ArXiv e-prints, arXiv:1506.07525

Papovich, C., Labbé, I., Quadri, R., et al. 2015, ZFOURGE/CANDELS: On the Evolution of M* Galaxy Progenitors from z = 3 to 0.5, *ApJ*, 803, 26

Rees, G. A., Spitler, L. R., Norris, R. P., et al. 2016, Radio galaxies in ZFOURGE/NMBS: no difference in the properties of massive galaxies with and without radio-AGN out to z = 2.25, *MNRAS*, 455, 2731

Spitler, L. R., Labbé, I., Glazebrook, K., et al. 2012, First Results from Z - FOURGE: Discovery of a Candidate Cluster at z = 2.2 in COSMOS, *ApJL*, 748, L21

Tilvi, V., Papovich, C., Tran, K.-V. H., et al. 2013, Discovery of Lyman Break Galaxies at z ~ 7 from the zFourGE Survey, *ApJ*, 768, 56

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, Galaxy Stellar Mass Functions from ZFOURGE/CANDELS: An Excess of Low-mass Galaxies since z = 2 and the Rapid Buildup of Quiescent Galaxies, *ApJ*, 783, 85

Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, The SFR-M* Relation and Empirical Star-Formation Histories from ZFOURGE* at 0.5 < z < 4, *ApJ*, 817, 118

Tran, K.-V. H., Nanayakkara, T., Yuan, T., et al. 2015, ZFIRE: Galaxy Cluster Kinematics, H alpha Star Formation Rates, and Gas Phase Metallicities of XMM-LSS J02182-05102 at zcl = 1.6232, *ApJ*, 811, 28

Yuan, T., Nanayakkara, T., Kacprzak, G. G., et al. 2014, Keck/MOSFIRE Spectroscopic Confirmation of a Virgo-like Cluster Ancestor at z = 2.095, *ApJL*, 795, L20

9

Acknowledgements

This final chapter is dedicated to the many people who have been there for me during the years I have worked on this thesis. I could never have guessed the impact on my live the last years have been, and I would like to thank everyone who has supported me throughout.

First, I must thank my supervisors at Leiden University for creating this wonderful and interesting research opportunity, and letting me rack my brains about it. Ivo, thank you big time for letting me be involved in all aspects of astronomy, and the almost unlimited travel opportunities to observing sites, meetings and conferences. I've had countless questions and you've always made time to answer them. I am honoured that you were my mentor. Marijn, thank you for the positive interest in my work, for inviting me to groupmeetings and for the great feedback you have given me especially for the final chapter.

In Leiden I was fortunate to meet many great minds. Monica, it was great to meet you and to share a house with you for a while. Best of luck to you and Bernard in the United States and on your respective career paths. Thanja, it was so much better grading assignments together and I am grateful for the friendship that followed. I have the fondest memories of our dinners together. Goodluck to you and Daan the coming years in Germany. Berenice, thank you so much for all your great advice on careers. It was great to meet you again in London and see how you found your place. Bram, it's just been great knowing you. Congratulations to you and Angelica on getting married and having your wonderful daughter Jule. I wish you great happiness together. Marco, thanks for reassuring me that I would make it through. Xander, your support and encouragement were invaluable, thank you. Thanks also to Tiffany, Ainil, Alex Rimoldi, Matt, Jens, Willem, David Huijser, Andra, Sara, Matteo, Maryam, David Carton, Irène, Ali, Koen, Marijke, Remco, Axel and Jaya. I was also part of a stimulating galaxy research group and would like to thank Jesse, Mattia, Adam, Shannon, Renske, Allison, Daniel, Kenneth, Corentin,

Michael, Rychard, Benne, Mauro, Bart, David Sobral, Simone, Moein, Saskia and Joanna for a great research environment and interesting scientific discussions.

It was my pleasure for four years to share office HL404 with Carl, Sebastiaan and Paula. Carl, thanks for brightening up the room with your perpetual happy appearance. I loved our chessgames and philosophical conversations. I was glad to see you achieve your goals and I wish you all the best for your new position. Sebastiaan, thanks for being a great friend who, even if you were busy, was always ready for coffee and conversation. Make sure you decorate your new office with lots of Star Wars posters! Thank you also for introducing me to Florie, I wish you both a happy future together and a successful time in Chicago. Paula, without you the office for certain would have been a lot less colourful. Thanks for your understanding, your openness and the addition of basic salsa moves to the curriculum. I hope you salsa your way into happiness and success.

I also owe thanks to the staff at Leiden Observatory, especially the computergroup: Aart, Erik and David Jansen. Not only was it fun to be neighbours, but without you solving the numerous computerproblems I encountered my thesis would have taken much longer to finish. In particular I'd also like to thank Jeanne, Anita and Liesbeth van der Veld. Evelijn, thank you for great advice and always having solutions ready at hand.

I was lucky to spend some time abroad in Melbourne, and making great friends at Swinburne Institute of Technology. Karl, first of all a big thank you for pointing out this opportunity to me, giving me access to the data and making time free to be my supervisor. Thanks for sharing your knowledge and for your invaluable support. Glenn, you have been a great friend and daily supervisor. I owe you for your contagious cheerfulness. Themiya, you are a most valuable friend. It was fun exploring together, from Tasmania to Hawaii and from Adelaide to Texas. George, thanks for your great sense of humor and for coding advice. Paolo, in an office without windows the only light has to come from the people in it, and I was lucky to have your company. Rebecca and Weston, it was great knowing you and I will never forget our trip to Tasmania. Martina, it was great to meet you and thanks for the lovely weekend at Lakes Entrance. Tyler, thanks for being so awesome. Dany, it was fun to meet you, and then meeting you again on the other side of the world. Congratulations on your marriage, and lots of happiness to you and Margreet. Angela, thanks for your cheerfulness. Manisha, it was lovely to meet you. Alex, I have the best memories of our friendship. You made Melbourne feel like home and I'll never forget that. Emily, Shirin, David Fisher, Rob, Stephi, Sarah, Chris, and David Lagattuta, it was great meeting you all. Liz, thanks

for making my time at Swinburne a pleasant one.

Most valuable to me as well was being part of an inspiring international team, without whose great support I would not have been able to finish this work. Even more importantly, in addition to science there was always the joy of hanging out together, going out for dinners and beers, and meeting each other at places around the world. Eric, it was a pleasure to meet you at the extraordinary site at Las Campanas and to see your passion for the FourStar instrument. Thank you for telling me your stories about observing and always taking a kind interest in my work. Lee, you've been both a friend and a colleague, and working together with you has been a very valuable experience. I'm thrilled to now be able to congratulate you on your permanent position at Macquarie, which will hopefully lead to many more great research opportunities. Michael, thank you for all your quick feedback and the interesting conversations that we've had. Adam, it was a pleasure to work with you, with many thanks for the fun trip to the Blue Mountains. Vy, thanks for your organisational skills and for always looking after our wellbeing. I will sorely miss your ready restaurant and foodie advice. Ryan, it was a pleasure to be on the team with you, and to listen to your wonderful stories. Casey, I'm glad to have met such an enthusiastic and intelligent person. Andy, Nicola, Gabe, Leo, Pieter, Ben, Josha, Nancy, Glen, Tilvi, Kate, Tiantian and Lisa, it was an honour to meet such talented astronomers. I wish you all the very best.

I am deeply grateful to all my friends and my entire family. In sometimes unexpected situations you have supported me and cheered me up. In particular I owe thanks to a number of people. Everdien, thank you for our longstanding friendship, for your kind understanding and for your positive outlook. I am glad that you will once more stand at my side as paranimf. Annika, I owe you for helping me to put things in perspective. It was fantastic to see you at work in Beregszász. Thanks as well for agreeing to be my paranimf. Paula, you've been a wonderful friend. Thanks a million for your listening ear and great advice. Margreet, I am glad I got to know such a warm and caring person. Thanks for always being there for me and for your everlasting support. A special thanks to Marloes and to all my musical friends, in particular to Marieke, Irene, Marie-Astrid, Adriane and Maarten.

Finally, I would like to express my deepest gratitude to my parents, who have always encouraged me to open my mind and to think beyond the ordinary. In loving memory of my father. Thank you for having supported my choice of career and for having been proud of what I do. I wish you could be here. Lieve mama, your unconditional love and dedication have led as much as anything to the completion of this thesis. For that, and for your patience, support, encouragement and many things more, I thank you.