![Universiteit Leiden - The Netherlands]

Cold gas in distant galaxies

Boogaard, L.A.

Citation

Version:	Publisher's Version
License:	[Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)
Downloaded from:	https://hdl.handle.net/1887/3147175

Cover Page

Universiteit Leiden

Leiden University Repository

The handle http://hdl.handle.net/1887/3147175 holds various files of this Leiden University dissertation.

Author: Boogaard, L.A.
Title: Cold gas in distant galaxies
Issue date: 2021-02-25

2 | The MUSE Hubble Ultra Deep Field Survey: XI. Constraining the low-mass end of the stellar mass – star formation rate relation at $z < 1$

Abstract

Star-forming galaxies have been found to follow a relatively tight relation between stellar mass (M_*) and star formation rate (SFR), dubbed the 'star formation sequence'. A turnover in the sequence has been observed, where galaxies with $M_* < 10^{10}\,\mathrm{M_\odot}$ follow a steeper relation than their higher-mass counterparts, suggesting that the low-mass slope is (nearly) linear. In this paper, we characterise the properties of the low-mass end of the star formation sequence between $7 \leq \log M_*[\mathrm{M_\odot}] \leq 10.5$ at redshift $0.11 < z < 0.91$. We use the deepest MUSE observations of the *Hubble* Ultra Deep Field and the *Hubble* Deep Field South to construct a sample of 179 star-forming galaxies with high signal-to-noise emission lines. Dust-corrected SFRs are determined from Hβ λ4863 and Hα λ6565. We model the star formation sequence with a Gaussian distribution around a hyperplane between $\log M_*$, $\log$ SFR, and $\log(1 + z)$, to simultaneously constrain the slope, redshift evolution, and intrinsic scatter. We find a sub-linear slope for the low-mass regime where $\log \mathrm{SFR}[\mathrm{M_\odot\,yr^{-1}}] = 0.83^{+0.07}_{-0.06} \log M_*[\mathrm{M_\odot}] + 1.74^{+0.66}_{-0.68} \log(1 + z)$, increasing with redshift. We recover an intrinsic scatter in the relation of $\sigma_{\mathrm{intr}} = 0.44^{+0.05}_{-0.04}$ dex, larger than typically found at higher masses. As both hydrodynamical simulations and (semi-)analytical models typically favour a steeper slope in the low-mass regime, our results provide new constraints on the feedback processes which operate preferentially in low-mass halos.

L. A. Boogaard, J. Brinchmann, N. Bouché, M. Paalvast, *et al.*

2.1   Introduction

How galaxies grow is one of the fundamental questions in astronomy. The picture that has emerged is that a galaxy builds up its stellar mass mainly through star formation, which is triggered by gas accretion from the cosmic web (e.g. Dekel et al., 2009; van de Voort et al., 2012), while mergers with other galaxies play only a minor role (except for massive systems; Bundy et al., 2009).

In the past decade, star-forming galaxies have been found to form a reasonably tight quasi-linear relation between stellar mass (M_*) and star formation rate (SFR) (Brinchmann et al., 2004; Noeske et al., 2007a; Elbaz et al., 2007; Daddi et al., 2007; Salim et al., 2007) over a wide range of masses and out to high redshifts (Pannella et al., 2009; Santini et al., 2009; Oliver et al., 2010; Peng et al., 2010; Rodighiero et al., 2010; Karim et al., 2011; Bouwens et al., 2012; Whitaker et al., 2012; Stark et al., 2013; Whitaker et al., 2014; Ilbert et al., 2015; Lee et al., 2015; Renzini & Peng, 2015; Schreiber et al., 2015; Shivaei et al., 2015; Salmon et al., 2015; Tasca et al., 2015; Gavazzi et al., 2015; Kurczynski et al., 2016; Tomczak et al., 2016; Santini et al., 2017; Bisigello et al., 2018), which is often referred to as the 'main sequence of star-forming galaxies' or the 'star formation sequence'. In contrast, galaxies that are undergoing a starburst or have already quenched their star formation respectively lie above and below the relation. This main sequence is close to a similar scaling relation for halos (Birnboim et al., 2007; Neistein & Dekel, 2008; Genel et al., 2008; Fakhouri & Ma, 2008; Correa et al., 2015a,b) where the growth rate increases super-linearly[17] with halo mass, and this has been interpreted as supporting the picture where galaxy growth is driven by gas accretion from the cosmic web (e.g. Bouché et al., 2010; Lilly et al., 2013; Rodríguez-Puebla et al., 2016; Tacchella et al., 2016).

This interpretation is supported by hydrodynamical simulations of galaxy formation (Schaye et al., 2010; Haas et al., 2013a,b; Torrey et al., 2014; Hopkins et al., 2014; Crain et al., 2015; Hopkins et al., 2016), where a global equilibrium relation is found between the inflow and outflow of gas and star formation in galaxies. In this picture the star formation acts as a self-regulating process, where the inflow of gas, through cooling and accretion, is balanced by the feedback from massive stars and black holes (e.g. Schaye et al., 2010). Furthermore, semi-analytical models (e.g. Dutton et al., 2010; Mitchell et al., 2014; Cattaneo et al., 2011, 2017) and relatively simple analytic theoretical models which connect the gas supply (from the cosmological accretion) to the gas consumption can also reproduce the main features of the main sequence rather well (e.g. Bouché et al., 2010; Davé et al., 2012; Lilly et al., 2013; Dekel et al., 2013; Dekel & Mandelker, 2014; Mitra et al., 2015; Rodríguez-Puebla et al., 2016, 2017).[18]

The parameters of the M_*-SFR relation (i.e. slope, normalisation, and scatter) are thus important, as they provide us with insight into the relative contributions of different processes operating at different mass scales, in particular when comparing the values of the parameters to their counterparts in dark matter halo scaling relations. The normalisation of the star formation sequence is governed by the change in cosmological gas accretion rates and gas

---

[17]There is a tension between the shallow slope of the observed main sequence with the super-linear slope expected in models, which is set by the index of the initial dark matter power spectrum (Birnboim et al., 2007; Neistein & Dekel, 2008; Correa et al., 2015a,b).

[18]For an alternative interpretation, cf. Gladders et al. (2013); Kelson (2014); Abramson et al. (2016).

depletion timescales. The slope can be sensitive to the effect of various feedback processes acting on the accreted gas, which prevent (or enhance) star formation. The intrinsic scatter around the equilibrium relation is predominantly determined by the stochasticity in the gas accretion process (e.g. Forbes et al., 2014; Mitra et al., 2017), but can also be driven by dynamical processes that rearrange the gas inside galaxies (Tacchella et al., 2016). The M_*-SFR relation is observed to be reasonably tight, with an intrinsic scatter of only ≈ 0.3 dex (Noeske et al., 2007a; Salmi et al., 2012; Whitaker et al., 2012; Guo et al., 2013; Speagle et al., 2014; Schreiber et al., 2015; Kurczynski et al., 2016, though we caution against a blind comparison as different observables probe star formation on different timescales). Yet, it has proven to be challenging to place firm constraints on the intrinsic scatter as one needs to deconvolve the scatter due to measurement uncertainty (e.g. Speagle et al., 2014; Kurczynski et al., 2016; Santini et al., 2017).

Observationally, the slope has been difficult to measure, particularly at the low-mass end, as most studies have been sensitive to galaxies with stellar masses above $\log M_*[\mathrm{M_\odot}] \sim 10$ and often lack dynamical range in mass. In addition, while it is well known that there is significant evolution in the normalisation of the sequence with redshift, most studies have measured the slope in bins of redshift. For a flux limited sample this could introduce a bias in the slope because overlapping populations at different normalisations are not sampled equally in mass within a single redshift bin. The slope may also be mass dependent and indeed recent studies have observed that the relation turns over around a mass of $M_* \sim 10^{10}$ $\mathrm{M_\odot}$ (Whitaker et al., 2012, 2014; Lee et al., 2015; Schreiber et al., 2015; Tomczak et al., 2016) and shows a steeper slope below the turnover mass. In the low-mass regime, a (nearly) linear slope has generally been expected (e.g. Schreiber et al., 2015; Tomczak et al., 2016), motivated also by the fact that there is very little evolution in the faint-end slope of the blue stellar mass function with redshift (Peng et al., 2014). Leja et al. (2015) showed that the sequence cannot have a slope $a < 0.9$ at all masses because this would lead to a too high number density between $10 < \log M_*[\mathrm{M_\odot}] < 11$ at $z = 1$.

In addition to the observational challenges, careful modelling is required to get reliable constraints on the parameters (slope, normalisation, scatter) of the star formation sequence. It is important to properly take selection effects into account as well as the uncertainties on both the stellar masses and star formation rates (and, if spectroscopy is lacking, also on the photometric redshifts). The latter in particular, due to the fact that there is intrinsic scatter in the relation that needs to be deconvolved from the measurement errors. Common statistical techniques do not take these complications into account self-consistently, which leads to biases in the results.

Putting the existing observations in perspective, it is clear that a large dynamical range in mass is necessary to measure the slope of the star formation sequence in the low-mass regime. Deep field studies, that can blindly detect large numbers of galaxies down to masses much below 10^{10} $\mathrm{M_\odot}$, are invaluable in this regard (e.g. Kurczynski et al., 2016). Yet, such studies are challenged by having to measure all observables, distances as well as stellar masses and star formation rates, from the same photometry. This can lead to undesirable correlations between different observables. At the same time the measurements suffer from the uncertainties associated with photometric redshifts. Spectroscopic follow up is crucial in this regard, but can suffer from biases due to photometric preselection.

With the advent of the *Multi Unit Spectroscopic Explorer* (MUSE; Bacon et al. 2010) on the VLT it is now possible to address these concerns. With the deep MUSE data obtained over the *Hubble* Ultra Deep Field (HUDF; Bacon et al., 2017) and *Hubble* Deep Field South (HDFS; Bacon et al., 2015), we can 'blindly' detect star-forming galaxies in emission lines down to very low levels ($\sim 10^{-3}$ M$_\odot$ yr^{-1}) and obtain a precise spectroscopic redshift estimate at the same time (Inami et al., 2017). These data provides a unique view into the low-mass regime of the star formation sequence.

In this paper we present a characterisation of the low-mass end of the M_*-SFR relation, using deep MUSE observations of the HUDF and HDFS. We characterise the properties of the M_*-SFR relation down stellar masses of $M_* \sim 10^8$ M$_\odot$ and SFR $\sim 10^{-3}$ M$_\odot$ yr^{-1}, out to $z < 1$, and trace the SFR in individual galaxies with masses as low as $M_* \leq 10^7$ M$_\odot$ at $z \sim 0.2$. We model the relation using a self-consistent Bayesian framework and describe it with a Gaussian distribution around a plane in (log mass, log SFR, log redshift)-space. This allows us to simultaneously constrain the slope and evolution of the star formation sequence as well as the amount of intrinsic scatter, while taking into account heteroscedastic errors (i.e. a different uncertainty for each data point).

The structure of the paper is as follows. In § 2.2 we first introduce the MUSE data set and outline the selection criteria used to construct our sample of star-forming galaxies. We then go into the methods used to determine a robust stellar mass and a SFR from the observed emission lines. Before looking at the results, we discuss the consistency of our SFRs in § 2.3. We then introduce the framework of our Bayesian analysis used to characterise the M_*-SFR relation (§ 2.4) and present the results in § 2.5. We discuss the robustness of the derived parameters in § 2.A.1. Finally, we discuss our results in the context of the literature and models, and the physical implications (§ 2.6). We summarise with our conclusions in § 2.7. Throughout this paper we assume a Chabrier (2003) stellar initial mass function and a flat ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$.

2.2   Observations and methods

To study the properties of the galaxy population down to low masses and star formation rates, deep spectroscopic observations are required for a large number of sources. We exploit the unique observations taken with the MUSE instrument over the *Hubble* Ultra Deep Field (Bacon et al., 2017) and the *Hubble* Deep Field South (Bacon et al., 2015) to investigate the star formation rates in low-mass galaxies at $0.11 < z < 0.91$. We provide a brief presentation of the observations and data reduction in the next section, but refer to the corresponding papers for details.

The MUSE instrument is an integral-field spectrograph situated at UT4 of the Very Large Telescope. It has a field-of-view of $1' \times 1'$ when operating in wide-field-mode, which is fed into 24 different integral-field units. These sample the field-of-view at 0.2 '' resolution. The spectrograph covers the spectrum across 4650Å - 9300 Å with a spectral resolution of $R \equiv \lambda/\Delta\lambda \simeq 3000$. The result of a MUSE observation is a data cube of the observed field, with two spatial and one spectral axes, i.e. an image with spectroscopic information at every pixel.
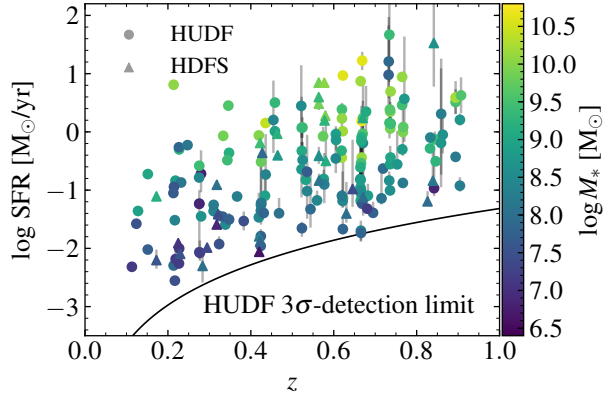
Figure 2.1: Redshift distribution of our galaxies plotted against their (dust-corrected) SFR (1σ error bars are in grey). The colour denotes the stellar mass. The solid line depicts the lowest uncorrected SFR from Hβ $\lambda4863$ we can detect in the HUDF at each redshift (which is effectively determined by the requirement that S/N(Hγ $\lambda4342$) > 3; see § 2.2.4).

2.2.1   Observations, data reduction, and spectral line fitting

The HUDF (Beckwith et al., 2006) was observed with MUSE in a layered strategy. The deepest region consists of a single $1' \times 1'$ pointing with a total integration depth of 31 hours. This deep region lies embedded in a larger $3' \times 3'$ mosaic consisting of 9 individual MUSE pointings, each of which is 10 hours deep. The average full width at half maximum (FWHM) seeing measured in the data cubes is 0.6 ʺ at 7750 Å. For the purpose of this work we use all galaxies from the mosaic region, including the deep (udf10) region, which we refer to collectively as the (MUSE) HUDF.

Because of its similar depth, we also include the MUSE observation of the HDFS (Williams et al., 2000) which was observed as part of the commissioning activities. These observations consist of a single deep field ($1' \times 1'$) with a total integration time of 27 hours and a median seeing of 0.7 ʺ.

The full data acquisition and reduction of the HUDF is detailed in Bacon et al. (2017) (for a description of the MUSE data reduction pipeline see Weilbacher et al., in prep.[19]). The data reduction of the HUDF is essentially based on the reduction of the HDFS, which is detailed in Bacon et al. (2015), with several improvements. We use HUDF version 0.42 and HDFS version 1.0, which reach a 3σ-emission line depth for a point source (1 ʺ) of 1.5 and 3.1×10^{-19} erg s^{-1} cm^{-2} (udf10 and mosaic) and 1.8×10^{-19} erg s^{-1} cm^{-2} (HDFS), measured between the OH skylines at 7000 Å.

Sources in the HUDF were identified using both a blind and a targeted approach. The latter uses the sources from the UVUDF catalogue (Rafelski et al., 2015) as prior information to extract objects from the MUSE cube. A blind search of the full cube was also conducted, using

---

[19]Weilbacher et al. (2020)

a tool specifically developed for MUSE cubes called ORIGIN (Bacon et al. 2017; Mary et al., in prep.[20]). A similar approach was already followed for the HDFS. Here sources were identified based on the Casertano et al. (2000) catalogue and blind emission line searches of the data cube were done with the automatic detection tools MUSELET[21] and LSDCat (Herenz & Wisotzki, 2017) as well as through visual inspection, and cross-correlated with the corresponding photometric catalogue, as described in Bacon et al. (2015).

The process of determining redshifts and constructing a full catalogue from the extracted sources is described in Inami et al. (2017) for the HUDF (and a similar approach was followed for the HDFS). In short, redshifts were determined semi-automatically by cross-matching template spectra with the identified sources and subsequently inspected and confirmed by at least two independent investigators. For emission line galaxies an additional constraint comes from the requirement that the emission line flux is coherent in a narrow band image around the line in the MUSE cube. The typical error on the MUSE spectroscopic redshifts is $\sigma_z = 0.00012(1 + z)$ (Inami et al., 2017), which we will take into account in the modelling (conservatively taking $\sigma_{\log(1+z)} = 0.0005$ for all galaxies; § 2.4)

For all detected sources one dimensional spectra are extracted using a straight sum extraction over an aperture around each source (based on the MUSE point spread function convolved with the Rafelski et al. (2015) segmentation map, see Bacon et al. 2017). From the extracted 1D spectra emission line fluxes are fitted in velocity space, using an updated version of the PLATEFIT code described in Tremonti et al. (2004) and Brinchmann et al. (2004, 2008). PLATEFIT assumes a Gaussian line profile for all lines, with the same intrinsic width and velocity. The result is a measurement of the flux and equivalent width of all emission lines present, with the uncertainties obtained from propagating the original pipeline errors. We define the signal-to-noise (S/N) in a particular spectral line as the line flux over the line flux error. We also determine the strength of the 4000 Å break, $D_n(4000)$, measured over $3850 - 3950$Å and $4000 - 4100$Å (Kauffmann et al., 2003). We note that the stellar absorption underlying the emission lines is taken into account by PLATEFIT.

2.2.2 Sample selection

From the HUDF and HDFS catalogues we construct our sample of star-forming galaxies using the following constraints:

1. We use Hβ λ4863 or Hα λ6565 to derive the SFR (see § 2.2.4) and in either case we always need Hβ λ4863 (to directly probe the SFR or to correct for dust extinction in Hα λ6565). As a result, we are limited to the range of redshifts where Hβ λ4863 falls within the MUSE spectral range. Subsequently, we only take objects into account that have a redshift $z < (9300/4861) - 1 = 0.913$.

2. In order to derive a robust SFR and dust correction, we only allow objects with a signal-to-noise ratio > 3 in the relevant pair of Balmer lines. This means S/N > 3 in either Hβ λ4863 and Hγ λ4342 (for Hβ λ4863 derived SFRs) or Hα λ6565 and Hβ λ4863 (for Hα λ6565 derived SFRs).

---

[20]Mary et al. (2020)
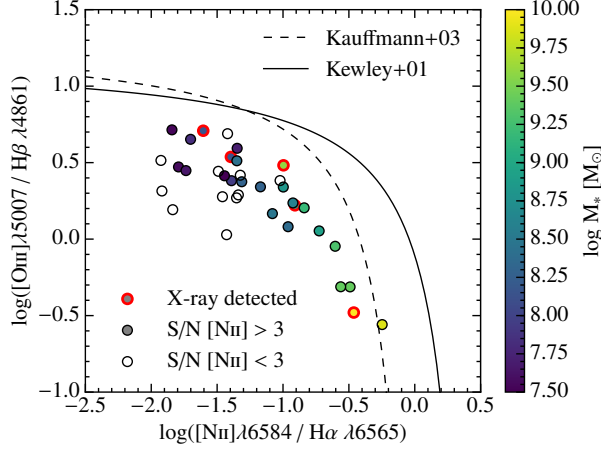[21]https://mpdaf.readthedocs.io/en/latest/muselet.html

Figure 2.2: BPT-diagram (Baldwin et al., 1981) of the sources in our Hα λ6565-subsample for which we measure [N II] λ6585. All galaxies fall in the star-forming region of the diagram. The filled and open circles have S/N([N II] λ6585) > 3 and < 3, respectively, and the 5 sources encircled in red are detected in X-rays (Luo et al., 2017). The solid and dashed curve show the AGN boundary and maximum starburst line from Kauffmann et al. (2003) and Kewley et al. (2001), respectively.

Included in the above criteria are some galaxies that are not actively star-forming and lie on the 'red-sequence'. Since these galaxies are not expected to lie on the M_*-SFR relation, we exclude them from the analysis based on their spectral features:

3. We remove 12 galaxies with a strong 4000 Å break by only allowing galaxies with a $D_n(4000) < 1.5$.

4. We omit galaxies with a rest-frame equivalent width in either Hα λ6565 or Hβ λ4863 of < 2Å.[22] This removed an additional 7 and 5 objects, respectively.

In addition, three sources were removed from the sample due to severe artefacts in their emission lines (see § 2.3). All sources selected based on the MUSE data are detected in the *HST* imaging. However, four sources were removed because there photometry was severely blended, prohibiting a mass estimate.

5. We remove potential AGN from our sample in the HUDF by cross-matching our sources with the *Chandra* Deep Field South 7Ms X-ray catalogue (Luo et al., 2017). We also confirm the location of the sources in the star-forming region of different emission line diagnostic diagrams.

A total of 16 galaxies with $z < 0.913$ from the MUSE catalogue are detected in X-rays. Five of these sources (including one AGN) show passive spectra without emission lines and did not

---

[22]Following the convention that emission-line equivalent widths (EQW) are negative, this translates to excluding EQW > −2Å.
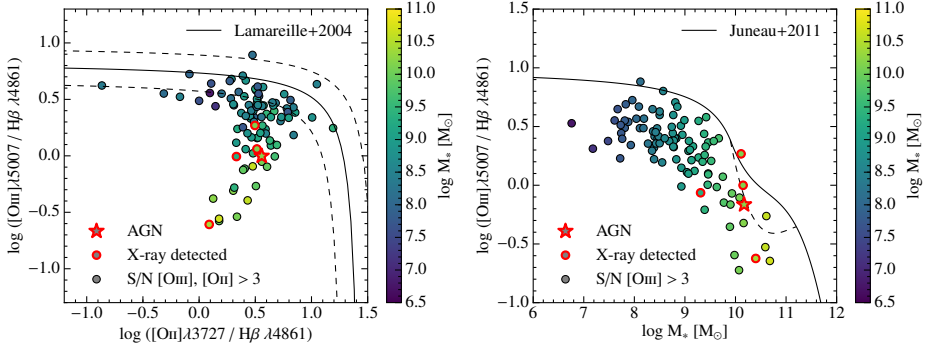
Figure 2.3: AGN diagnostics for the sources in our Hβ λ4863-subsample, including all sources which have S/N> 3 in the relevant emission lines. Overall, our sample is consistent with star-forming galaxies. We remove one X-ray detected AGN from the sample. Left: The [O II] $\lambda\lambda$3727, 3730/Hβ vs. [O III] $\lambda\lambda$4960, 5008/Hβ diagnostic from Lamareille et al. (2004) (solid line, with the uncertainty indicated by the dashed lines). Right: The mass-excitation diagram from Juneau et al. (2011). Galaxies in the region between the dashed and solid lines are on average identified as intermediate between AGN and SF.

pass the previous criteria. Cross-matching our star-forming sample (after applying criteria 1 through 4) left 11 galaxies that were detected in X-rays. Five of these sources (ID#855, 861, 863, 895, and 902) are in the Hα-subsample and six (ID#867, 869, 874, 875, 884, and 905) are in the Hβ-subsample. All of these sources were classified as 'Galaxy' in the Luo et al. (2017) catalogue (according to their 6 criteria based on X-ray luminosity, spectral index, flux-ratios and previous spectroscopic identification), except for ID# 875 which was classified as an AGN and which we subsequently removed from the sample. Luo et al. (2017) caution however that sources classified as 'Galaxy' may still host low-luminosity or heavily obscured AGN.

We plot all sources from our Hα λ6565-subsample for which we have a measurement of [N II] λ6585 in the BPT-diagram (Baldwin et al., 1981) in Figure 2.2. We include sources for which we have a low S/N (<3) measurement of [N II] λ6585 as open circles. While we can only put a subsample of our sources on this diagram, all are in the star-forming region, including the 5 galaxies which have an X-ray detection. None of the X-ray sources classified as 'Galaxy' show spectral signatures of AGN activity. In Figure 2.3 we show a similar consistency check for the Hβ λ4863-subsample. Because we lack access to the BPT diagram at these redshift, we instead use the diagnostics from both Lamareille et al. (2004) and Juneau et al. (2011). Reassuringly, our sample is overall consistent with star-forming galaxies and none of the galaxies show line-ratios clearly powered by AGN activity (including, perhaps surprisingly, the single X-ray classified AGN). There is only one source which is above the discriminating line in both plots (ID#1114), however, it is consistent within errors with being dominated by star formation and not detected in X-rays. Furthermore, its high [O III] flux can very well be driven by star formation and indeed it is part of the sample of high-[O III]/[O II] galaxies identified by Paalvast et al. (2018). Hence, except for X-ray detected AGN ID#875, we do not
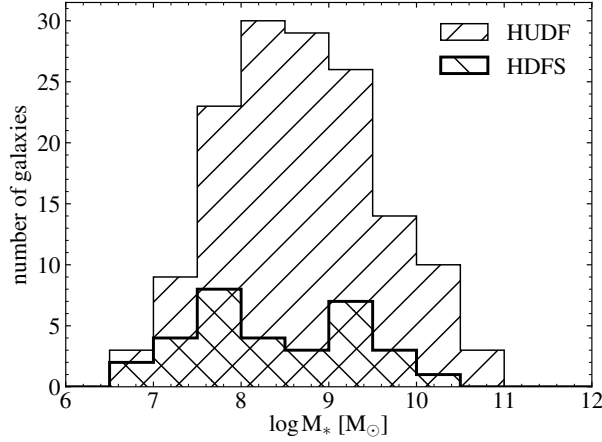
Figure 2.4: Histograms of the stellar mass distributions of the MUSE detected galaxies in the HUDF and the HDFS. The deep 30h observations allow us to detect and subsequently infer a stellar mass and SFR for galaxies down to $\sim 10^7\,M_\odot$.

remove any additional sources from the sample. Finally, we note that none of the methods to identify AGN are individually foolproof. Therefore, we check the impact of potential misclassification of AGN and confirm that excluding (1) the sources that are above the pure star-forming line in either of the diagnostic diagrams or (2) all galaxies that are detected in X-rays (even when consistent with star formation) does not significantly affect the results.

The final sample then consists of 179 star-forming galaxies, 147 from the HUDF, all with the highest redshift confidence (Inami et al., 2017), and 32 from the HDFS, between $0.11 < z < 0.91$ with a mean redshift of 0.53 (see Figure 2.1).

2.2.3 Stellar masses

The stellar masses of the galaxies were estimated using the Stellar Population Synthesis (SPS) code FAST (Kriek et al., 2009). The SPS-templates were χ^2-fitted to the broad-band photometry of the different fields for a range of parameters. For the HUDF, we rely on the deep *HST* photometry from the UVUDF catalogue (Rafelski et al., 2015) (containing WFC3/UVIS F225W, F275W and F336W; ACS/WFC F435W, F606W, F775W, and F850LP and WFC/IR F105W, F125W, F140W and F160W) while for the HDFS we take the WFPC2 photometry from Casertano et al. (2000) (F330W, F450W, F606W, and F814W). The SPS-templates were constructed from the Conroy et al. (2010) (FSPS) models using a discrete range of metallicities ($Z/Z_\odot = [0.04, 0.20, 0.50, 1.0, 1.58]$). We assumed a Chabrier (2003) initial mass function with an exponentially declining star formation history (SFR $\propto \exp(-t/\tau)$ with $8.5 < \log(\tau/\mathrm{yr}) < 10$ in steps of 0.2 dex). The redshifts were fixed to the accurate spectroscopic values determined from the MUSE spectra. Ages were allowed to vary between $8 < \log \mathrm{Age/yr} < 10.2$ in steps of 0.2 dex. We parameterised the dust attenuation curve

according to the Calzetti et al. (2000) dust law with the dust extinction in the visual taken to be within $0 < A_V < 3$ ($\Delta A_V = 0.1$ magnitudes). For all the parameters error estimates were obtained through Monte Carlo methods, by re-running the fitting 500 times while varying the input photometry within their photometric errors (see Kriek et al. 2009 for details).

Stellar masses were determined for all 179 objects in the final sample. The distribution of masses is shown in Figure 2.4. With these deep MUSE observations we are mainly probing low-mass ($<10^{9.5}\,M_\odot$) galaxies and we can still detect star formation from emission lines in galaxies with mass $\sim10^7\,M_\odot$. The mass estimates with their upper and lower confidence intervals are shown for the individual objects in Figure 2.7. The mean and standard deviation of the average errors on the mass estimates are 0.19 ± 0.06 dex for the HUDF and 0.22 ± 0.12 dex for the HDFS.

2.2.4 Star formation rates

The star formation rates are inferred from the flux in the Hα $\lambda6565$ or Hβ $\lambda4863$ recombination lines emitted by H II regions, which primarily trace recent (~10 Myr) massive star formation. Before we can infer a SFR we need to correct the measured flux in the emission lines for the attenuation by dust along the line of sight. We do this assuming a dust law according to Charlot & Fall (2000) (i.e. $\tau \propto \lambda^{-1.3}$, appropriate for birth clouds) and using the intrinsic ratio of the Balmer recombination lines ($j_{H\alpha}/j_{H\beta} = 2.86$ and $j_{H\beta}/j_{H\gamma} = 2.14$; Hummer & Storey (1987), for an electron temperature and density of $T = 10\,000$ K and $n_e = 10^3$ cm^{-3}). Hence, to derive an SFR(Hα $\lambda6565$) we also require a measurement of Hβ $\lambda4863$ and likewise for SFR(Hβ $\lambda4863$) we also require Hγ $\lambda4342$. After the dust correction we can convert the intrinsic flux to a luminosity using the measured redshift, given the assumed ΛCDM cosmology.

To determine the SFR we follow the treatment by Moustakas et al. (2006), which is essentially based on the relations from Kennicutt (1998a). Out of the SFR indicators that MUSE has access to, the Hα $\lambda6565$ line presents the least systematic uncertainties, but it is only available at low redshifts ($z \leq 0.42$ for MUSE at 9300 Å; 47 galaxies). We convert the Kennicutt (1998a) relation from a Salpeter to a Chabrier IMF ($0.1 < M[M_\odot] < 100$) by multiplying by a factor 0.62 (which is derived by computing the difference in total mass in both IMFs, while assuming the same number of massive ($>10\,M_\odot$) stars):

$$\mathrm{SFR}(\mathrm{H}\alpha\ \lambda6565) = 4.9 \times 10^{-42}\,\frac{L(\mathrm{H}\alpha\ \lambda6565)}{\mathrm{erg\,s}^{-1}}\,\mathrm{M_\odot\,yr}^{-1}, \qquad (2.1)$$

where $L(\mathrm{H}\alpha\ \lambda6565)$ is the dust-corrected luminosity. We note that this calibration assumes case B recombination and solar metallicity.

Because Hα $\lambda6565$ moves out of the optical regime at $z > 0.42$, the Hβ $\lambda4863$ luminosity is the primary tracer of SFR for the majority of our sample (132 galaxies). Given the intrinsic flux ratio between Hα $\lambda6565$ and Hβ $\lambda4863$, we can convert equation Equation 2.1 into a SFR for $L(\mathrm{H}\beta\ \lambda4863)$:

$$\mathrm{SFR}(\mathrm{H}\beta\ \lambda4863) = 1.4 \times 10^{-41}\,\frac{L(\mathrm{H}\beta\ \lambda4863)}{\mathrm{erg\,s}^{-1}}\,\mathrm{M_\odot\,yr}^{-1}, \qquad (2.2)$$

where $L(\text{H}\beta\,\lambda4863)$ is corrected for dust. We note that the Hβ $\lambda4863$ derived SFR inherits all the uncertainties from SFR(Hα $\lambda6565$), including variations in dust reddening (Moustakas et al., 2006).

We also investigate the SFR using the [O II] $\lambda\lambda3727, 3730$ nebular emission line. Here we use the calibration for the Hα $\lambda6565$ SFR (Equation 2.1), where we assume an intrinsic flux ratio between [O II] $\lambda\lambda3727, 3730$ and Hα $\lambda6565$ of unity (Moustakas et al., 2006). Since [O II] $\lambda\lambda3727, 3730$ is closest to Hβ $\lambda4863$, we use the Hβ $\lambda4863$/Hγ $\lambda4342$ ratio to determine the dust correction, scaled to the appropriate wavelength. The consequence of this is that the addition of the [O II] $\lambda\lambda3727, 3730$ line as a tracer of SFR will not add any new objects to the sample. Instead, it can be used as a useful comparison, which will be discussed in § 2.3.

To estimate the uncertainty in the SFR estimates (and dust corrections), we use Monte Carlo methods to derive a confidence interval on the SFR of every individual galaxy. We create a posterior distribution on the SFR by doing 1000 draws from a Gaussian distribution centred on the measured flux, with the variance set by the measurement error squared. The median posterior SFR can then be determined, as well as the $\pm1\sigma$ confidence intervals, by taking the 50$^{\text{th}}$, 16$^{\text{th}}$ and 84$^{\text{th}}$ percentile from the derived posterior distribution.

2.3   Consistency of SFR indicators

Before turning to the results, we first consider the consistency of the derived SFRs, by comparing the SFR estimates from different tracers for the same galaxies. In the remainder of the paper we only use the dust-corrected Balmer lines as tracers of star formation.

For a significant fraction of our galaxies ($\approx 40\%$) we find that the Balmer line ratios are below their case B values (as stated in § 2.2.4), indicative of a negative dust correction. While this might seem surprising, this is not uncommon and similar ratios have been seen in spectra from, e.g. the SDSS (Groves et al., 2012), MOSDEF (Reddy et al., 2015), KBSS (Strom et al., 2017) and ZFIRE (Nanayakkara et al., 2017). While 'unphysical', these ratios are not entirely unexpected and can have several causes.

First, these deviations can be caused by noisy spectra. Most galaxies in our sample are not very dusty and hence have a ratio close to case B. In $> 50\%$ of the cases with deviant ratios, the case B ratio is indeed within the 1σ error bars. We conservatively apply no dust correction for all these galaxies. The mean dust correction for all galaxies in our sample is $\tau(\text{H}\beta/\text{H}\gamma) \approx 0.6$ (setting galaxies with a negative dust correction to zero) or $\tau(\text{H}\beta/\text{H}\gamma) \approx 1$ (including only galaxies with a positive dust correction).

Secondly, there could be a problem with the measurement. Three objects that were significantly offset from the rest of the sample showed particular problems in their emission lines. In one object (ID#971) Hγ $\lambda4342$ was severely affected by an emission line from a nearby source ([O III] $\lambda4960$ from ID#874 at $z = 0.458$, another galaxy in our sample, coincidentally almost exactly at the observed wavelength of Hγ). For five other objects there was a clear problem with the fit to the Hβ $\lambda4863$ (ID#894, #896, #1027) or Hα $\lambda6565$ (ID#2, #1426) emission lines. We subsequently removed the first four sources from the analysis; for the latter two we disregarded the Hα $\lambda6565$ SFR and use the Hβ $\lambda4863$ SFR.

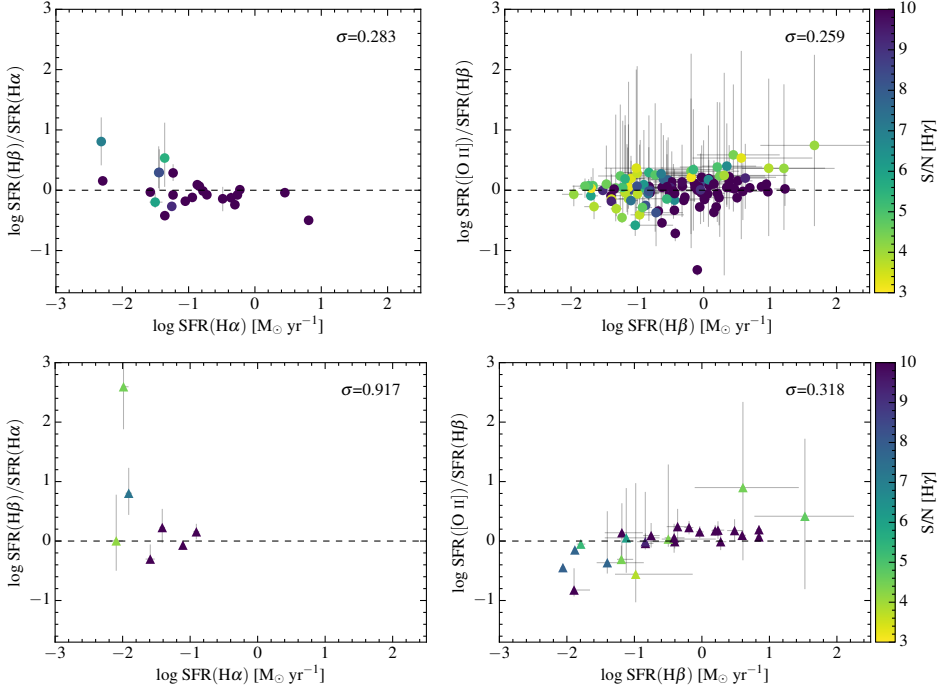A third, intriguing option is that theses objects are real. Indeed, there remains a small

Figure 2.5: A comparison of the derived star formation rate (SFR) from the Hα λ6565, Hβ λ4863 and [O II] λλ3727, 3730 luminosities for the HUDF (top panels, circles) and the HDFS (bottom panels, triangles). The left panels show the logarithm of the SFR from Hα λ6565 vs. the difference between the log Hβ λ4863 and log Hα λ6565 SFRs. The right panels show the same for Hβ λ4863 vs. [O II] λλ3727, 3730. In the top right corners σ indicates the standard deviation (in dex) around the one-to-one relation. Colour indicates the signal-to-noise ratio (S/N) in the faintest line; Hγ λ4342. Only galaxies that allowed for more than one SFR indicator are included in the plot. Overall the SFRs from Hβ λ4863 and [O II] λλ3727, 3730 agree reasonably well, considering we have not taken into account the metallicity dependence in SFR([O II] λλ3727, 3730). The scatter in Hα λ6565 vs. Hβ λ4863 SFR is largely driven by Hγ λ4342 S/N.

number of galaxies which have high-S/N spectra, but still show Balmer ratio's below their case B values.[23] Similar objects have also been observed in the other surveys already referenced, such as SDSS (Jarle Brinchmann, private communication, see also Groves et al. 2012). While these are very interesting objects on their own, a detailed analysis of these sources is beyond the scope of this paper. To be conservative and consistent, we apply no dust correction for these sources.

For some objects in the sample we measure multiple emission lines, which allows us to infer a SFR from different tracers. In any case a pair of Balmer lines (either Hα/Hβ or Hβ/Hγ)

---

[23]It is important to point out that this is not caused by stellar absorption in the continuum as this is taken into account when modelling the emission lines with PLATEFIT.

is available (§ 2.2.2), to allow for a dust correction. The majority of our sample lies at $z > 0.42$ for which Hα λ6565 is not available, but (dust-corrected) [O II] $\lambda\lambda$3727, 3730 is available as an SFR indicator. In Figure 2.5 we show a comparison for all galaxies that allowed both Hα λ6565 and Hβ λ4863 (only some galaxies at $z < 0.42$) and Hβ λ4863 and [O II] $\lambda\lambda$3727, 3730 derived SFRs (all redshifts). We note that Hβ λ4863 and [O II] $\lambda\lambda$3727, 3730 derived SFRs are corrected for dust using the same Hβ/Hγ-ratio.

In the right panels of Figure 2.5 we see that the Hβ λ4863 and [O II] $\lambda\lambda$3727, 3730 derived SFRs agree remarkably well (standard deviation $\sigma \leq 0.28$ dex), considering that we have not taken into account the metallicity dependence of the [O II] $\lambda\lambda$3727, 3730 luminosity in the SFR conversion factor (e.g. Kewley et al., 2004). A few points scatter quite a bit, most of which have large error bars. At lower SFRs we do see that [O II] $\lambda\lambda$3727, 3730 predicts a lower SFR than Hβ λ4863, which is probably because at low SFR we are also probing low-mass and low-metallicity galaxies. Stars with a lower metallicity have a higher UV flux, which causes the ionisation equilibrium for oxygen to shift from [O II] to [O III] which diminishes the observed [O II] $\lambda\lambda$3727, 3730 flux. Because of the opposite effect [O II] $\lambda\lambda$3727, 3730 occasionally predicts a higher SFR than Hβ λ4863 at the high-SFR end.

For a limited number of objects all three Balmer lines are in the spectral range of MUSE ($0.09 < z < 0.42$). We compare the Hα λ6565 and Hβ λ4863 derived SFRs in the left panel of Figure 2.5, where we find reasonable agreement (in the HUDF, where we have most sources, they have a factor of ~ 2 scatter). Most of the scatter is found at low SFR, where (on average) the S/N is also the lowest. In the HDFS one object (at low S/N) is a strong outlier, but removing this source yields a similar scatter to the HUDF. Intuitively the SFRs from Hα and Hβ should agree very well, which warrants some deeper investigation into the outliers at low SFR.

The main uncertainty in the SFR estimate is the amount of dust attenuation. In Figure 2.6 we compare the inferred optical depth from the Hβ/Hγ-ratio (τ[Hβ/Hγ]) to the optical depth determined from the Hα/Hβ ratio (τ[Hα/Hβ]). We note though that Figure 2.6 shows the measured optical depth, while we set negative τ to zero before computing the SFR. Indeed, while many sources agree well, we see that the amount of dust correction estimated from the Balmer lines is not consistent for several objects, leading to a different SFR estimate from Hα λ6565 and Hβ λ4863.

This tension is in part caused by the nature of the experiment, which requires that all three Balmer lines are in the spectral range of MUSE simultaneously. Necessarily then, Hα λ6565 will be at the long wavelength end of the spectrograph where skylines are more prevalent, occasionally adding uncertainty to its measurement. For the low-SFR sources, however, Hγ λ4342 might not be very bright, adding uncertainty to the dust correction of SFR(Hβ λ4863) for these sources (as seen at lower SFR in the left panels of Figure 2.5). Indeed, most of the outliers have a low S/N in Hγ λ4342 (as stated earlier, for the objects with a negative dust correction from Hβ/Hγ, we leave the often lower S/N measurement of Hγ λ4342 out of the analysis by setting τ(Hβ/Hγ) to zero). On the other hand, the converse is not quite true: for a large number of sources with a low S/N in Hγ λ4342 we do have a consistent SFR estimate. For all objects we use the highest S/N lines available to infer a dust-corrected SFR, i.e. for objects which have a measurement of all three Balmer line we use the Hα λ6565, Hβ λ4863 pair to infer a dust-corrected SFR, which generally has the highest S/N.
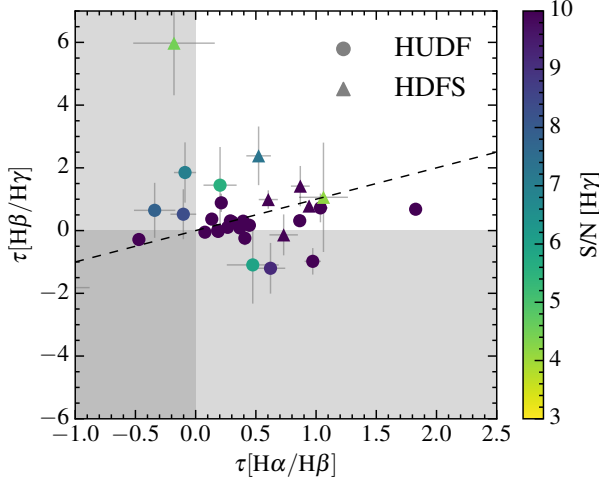
Figure 2.6: Optical depths at the wavelength of Hβ λ4863 as derived from both the Hβ λ4863/Hγ λ4342 and the Hα λ6565/Hβ λ4863-ratio, coloured by Hγ λ4342 signal-to-noise (S/N). The dashed line is the one-to-one relation. Overall the optical depths agree reasonably well, unless the Hγ S/N is low. Most galaxies actually show little dust (τ close to zero). The shaded area shows the regions of (unphysical) negative optical depth for each axis. We set the optical depth to zero for galaxies with negative τ as this is often consistent with the error bars and the offset is due to noise in the spectra. We note that some of the high-S/N outliers actually have discrepant Balmer ratios. If the inferred optical depth is very different, this will affect the comparison of the dust-corrected SFR from Hβ λ4863 and Hα λ6565 (see Figure 2.5).

   In summary, we have dust-corrected SFR measurement from the Hα λ6565 and Hβ λ4863 spectral lines for all galaxies at $z < 0.42$ and the Hβ λ4863, Hγ λ4342-pair at higher redshifts. Comparing Hα λ6565 and Hβ λ4863 SFRs, we conclude that the dust correction is the largest uncertainty on the derived SFR. We always use the highest S/N line-pair available to compute a dust-corrected SFR. Comparing the Hβ λ4863 SFRs with [O II] $\lambda\lambda$3727, 3730 at all redshifts, we see a very consistent picture (they have ≤ 0.3 dex scatter in both fields). Naturally, some variations between Hβ λ4863 and [O II] $\lambda\lambda$3727, 3730 SFRs are expected given the metallicity dependent nature of [O II] $\lambda\lambda$3727, 3730.

2.4   Bayesian model

2.4.1   Definition

The star formation sequence is commonly described by a power-law relation between stellar mass (M_*) and star formation rate (SFR), which evolves with redshift (z):

$$\text{SFR} \propto M_*^a (1 + z)^c, \tag{2.3}$$

where a and c are the power law exponents. It has been suggested that the slope (a) becomes shallower in the high-mass regime ($M_* > 10^{10}\,M_\odot$). In this work we will focus on the low-mass regime, for which we assume the slope is constant with mass. We will revisit this assumption in § 2.5.2. Given the lack of homogeneous studies with redshift it is still unclear whether the low-mass slope of the relation evolves with redshift. Here, we assume that the low-mass slope is independent of redshift over the range that we probe in this study. Likewise, given the large uncertainties in (the evolution of) the intrinsic scatter, we limit the number of free parameters in the model and assume that the intrinsic scatter does not depend on any of the other model parameters.

Following this description, we model the star formation sequence by a plane in ($\log M_*$, $\log(1 + z)$, $\log$ SFR)-space:

$$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = a \log \left(\frac{M_*}{M_0} \right) + b + c \log \left(\frac{1 + z}{1 + z_0} \right), \tag{2.4}$$

where b is now a normalisation constant. We take $M_0 = 10^{8.5}\,M_\odot$ and $z_0 = 0.55$ (close to the medians of the data) without the loss of generality. Galaxies scatter around this relation with an amount of intrinsic scatter in the vertical (i.e. $\log$ SFR) direction, which we denote by σ_{intr}. In the lack of an obvious alternative, we take the intrinsic scatter to be Gaussian in our model.

In a statistical sense we can then say that our observations ($\log M_*$, $\log(1 + z)$, $\log$ SFR) are drawn from a Gaussian distribution around the plane defined by Equation 2.4. To recover this distribution, we need to take a careful approach, taking into account the heteroscedastic errors of the measurements.

We adopt a Bayesian approach to determine the posterior distribution of the model parameters ($a, c, b, \sigma_{\mathrm{intr}}$) (see Andreon & Hurn (2010) for a lucid description of the Bayesian methodology in an astronomical context). Different approaches to construct the likelihood have been presented in the literature (see e.g. Kelly 2007 or Hogg et al. 2010). We choose to adopt a parameterisation of the likelihood following Robotham & Obreschkow (2015) (hereafter R15).

First, we state that our knowledge about galaxy i (determined by the observations) is encompassed by the probability density function of a multivariate Gaussian distribution, $\mathcal{N}(\mathbf{x}_i, C_i)$, with a mean value of:

$$\mathbf{x}_i = (\log M_{*,i}, \log(1 + z_i), \log \mathrm{SFR}_i) \tag{2.5}$$

and a diagonal covariance matrix:

$$C_i = \begin{pmatrix} \sigma^2_{\log M_*,i} & 0 & 0 \\ 0 & \sigma^2_{\log(1+z),i} & 0 \\ 0 & 0 & \sigma^2_{\log \mathrm{SFR},i} \end{pmatrix} \tag{2.6}$$

containing the variance in each parameter. This is justified as both stellar mass and star formation rate are measured independently from different data. The covariance with redshift is negligible as the error on the spectroscopic redshift is very small.

Secondly, we parameterise the model given by Equation 2.4 (which is a plane in three dimensions) in terms of its normal vector **n**, to avoid optimisation problems (R15). The galaxies scatter around this plane with an amount of intrinsic Gaussian scatter, perpendicular to the plane, which we denote by $\sigma_\perp$. We note that perpendicular scatter $\sigma_\perp$ is distinct from the (commonly reported) vertical scatter σ_{intr} which lies in the log SFR direction. After the analysis, we can simply transform the parameters $(\mathbf{n}, \sigma_\perp^2)$ back into familiar parameters $(a, c, b, \sigma_{\mathrm{intr}}^2)$ (using R15, Eq. 9).

Given the above definitions, we can express our log-likelihood[24] as the sum over N data points (see also R15):

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^{N} \left[\ln \left(\sigma_\perp^2 + \frac{\mathbf{n}^\top C_i \mathbf{n}}{\mathbf{n}^\top \mathbf{n}} \right) + \frac{(\mathbf{n}^\top [\mathbf{x}_i - \mathbf{n}])^2}{\sigma_\perp^2 \mathbf{n}^\top \mathbf{n} + \mathbf{n}^\top C_i \mathbf{n}} \right], \tag{2.7}$$

where all the parameters have been defined earlier.

Lastly, we have to define our priors on each component of **n** and on $\sigma_\perp^2$. As we want to impose limited prior knowledge, we express our priors as uniform distributions, with large bounds compared to the typical values of the parameters (we confirm that the results are robust, irrespective of the exact choice of bounds).

$$\mathbf{n} \sim \mathcal{U}^3(-1000, 1000) \tag{2.8}$$
$$\sigma_\perp^2 \sim \mathcal{U}(0, 1000),$$

where $\mathcal{U}^n$ is the n-dimensional multivariate uniform distribution and we take into account the fact that variance is always positive.

2.4.2   Execution

With the likelihood and priors in hand we determine the posterior using *Markov chain Monte Carlo* (MCMC) methods. We use the Python implementation called EMCEE (Foreman-Mackey et al., 2013), which utilises the affine-invariant ensemble sampler for MCMC from Goodman & Weare (2010). EMCEE samples the parameter space in parallel by setting off a predefined number of 'walkers', which we take to be 250.

Following Foreman-Mackey et al. (2013), we first initialise the walkers randomly in a large volume of parameter space. We then restart the walkers in a small Gaussian ball around the median of the posterior distribution (i.e. around the 'best solution'). We (generously) burn in for a quarter of the total amount of iterations for each walker which we take to be 20000 for the main run (§ 2.5.1; roughly four hundred times the autocorrelation time). We note that for all subsequent runs described below we follow the same procedure, with similar results.

We take several steps to check whether the EMCEE algorithm has properly converged. As an indication, one can look at both the mean acceptance fraction of the samples as well as the autocorrelation time (Foreman-Mackey et al., 2013). For the main run the acceptance fraction that resulted from the modelling (0.45) was well within range advocated by Foreman-Mackey

---

[24]Throughout this paper we consistently use 'log' for the base-10 logarithm and 'ln' for the base-e logarithm, with one exception: we stick to standard terminology and call $\ln \mathcal{L}$ the 'log-likelihood'.

et al. (2013) (0.2 - 0.5). The autocorrelation time was also relatively short and we let the walkers sample the posterior well over the autocorrelation time. Furthermore, we confirmed that the walkers properly explored the parameter space.

Combining the results from all walkers then gives the posterior distribution over which we can marginalise to find the posterior probability distributions for the model parameters. We will discuss the results of the modelling in § 2.5.

2.4.3   Model and data limitations

The unique aspect of the likelihood in Equation 2.7 is that it captures both the heteroscedastic errors on the observables as well as the intrinsic scatter around the plane. Furthermore, it can simultaneously describe both the slope of the sequence as well as the evolution with redshift.

It is important to determine how well we can recover the 'true' parameters with the observed data at hand. Our MUSE observations are constrained by the fact that we can only detect galaxies in a certain redshift range and cannot detect galaxies below the flux limit of the instrument (see Figure 2.1). As the flux limit varies with redshift, this could introduce a bias in our inferred parameters. The reason behind this is that the lack of low-SFR galaxies at higher redshift will bias the posterior towards shallower slopes, with a steeper redshift evolution (see Figure 2.11 for an illustration). In order to correct for such a bias, we analyse a series of simulated observations. We briefly outline the procedure here, which is described in detail in § 2.A.

In order to characterise the bias in the inferred parameters, we simulate galaxies from a mock star formation sequence for a range of values in each parameter, which we call $\mathbf{x}_{\mathrm{true,k}}$ (see Table 2.2). After applying the redshift-dependent flux limit to the mock data, we model the remaining galaxies as described in § 2.4 and recover the parameters, $\mathbf{x}_{\mathrm{out,k}}$. We then fit the transformation between the true and recovered parameters with an affine transformation ($\mathbf{x}_{\mathrm{out,k}} = A\mathbf{x}_{\mathrm{true,k}} + \mathbf{b}$) as outlined in § 2.A.2. The inverse of the best-fit transformation (Equation 2.19) can then be used to correct the posterior density distribution as measured from the MUSE data. In the following, we provide both the uncorrected (directly fitted) and the corrected values for reference.

2.5   Star formation sequence

2.5.1   Global sample

With a reliable SFR estimate in hand, we can turn to the star formation sequence between $0.11 < z < 0.91$ as observed by MUSE. Figure 2.7 shows a plot of stellar mass (M_*) versus star formation rate (SFR) for all the galaxies in the sample. The figure is based on two dust-corrected SFR indicators: the Hβ λ4863 and Hα λ6565 luminosities (Eqs. (2.2) and (2.1)). The vertical grey lines indicate the errors in (log M_*, log SFR) for each of the individual galaxies. The mean average error on the SFR is ≈ 0.2 dex in both the HUDF and the HDFS.

We are able to detect star formation in galaxies down to star formation rates as low as $0.003\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$. The galaxies appear to follow the M_*-SFR trend closely over the complete
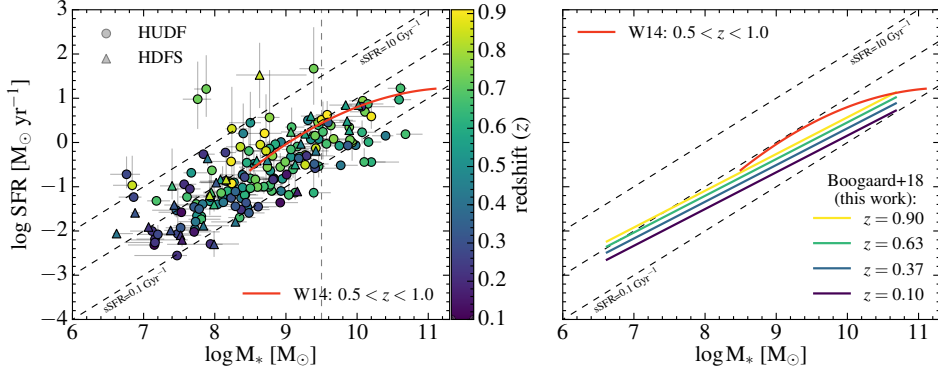
Figure 2.7: Left panel: The sample of 179 star-forming galaxies observed with MUSE, plotted on the M_*-SFR plane. The symbols indicate the field and colour indicates the redshift. The dashed lines show a constant sSFR, which is equivalent to a linear relationship: SFR $\propto M_*$. The red curve shows the model of the star formation sequence from Whitaker et al. (2014) for $0.5 < z < 1.0$. The vertical grey dashed line indicates the selection for the low-mass fit (§ 2.5.2). Right panel: Same as the left panel but with the data points removed, showing (the evolution of) the star formation sequence as seen by MUSE, according to Equation 2.11.

mass range, down to the lowest masses we can probe here $\sim 10^7$ M$_\odot$. At the high-mass end it appears we are starting to witness a flattening off of the trend, although we are primarily sensitive to the intermediate and low-mass galaxies.

We model the M_*-SFR relation with the Bayesian MCMC methodology described in detail in § 2.4. We show the resulting posterior probability density distribution for the parameters in Figure 2.8. By marginalising over the various parameters, we recover the posterior probability distributions for the individual parameters of interest ($a, c, b, \sigma_{\mathrm{intr}}$). These are plotted as histograms above the various axes in Figure 2.8. By taking the median and the 16th and 84th percentile from the posterior distributions we derive the median posterior value and a 1σ confidence interval for the parameters of interest.

The (uncorrected) best-fit (i.e. median posterior) parameters of the distribution (with their confidence intervals) that describe the star formation sequence are:

$$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = 0.79^{+0.05}_{-0.05} \log\left(\frac{M_*}{M_0}\right) - 0.77^{+0.04}_{-0.04}$$
$$+ 2.78^{+0.78}_{-0.78} \log\left(\frac{1+z}{1+z_0}\right) \pm 0.46^{+0.04}_{-0.03}, \qquad (2.9)$$

analogous to Equation 2.4. The final term represents the intrinsic scatter ($\sigma_{\mathrm{intr}} = 0.46^{+0.04}_{-0.03}$) in the vertical (log SFR) direction. We note that while it is a perfectly valid option for the parameterisation of the likelihood, the posterior distribution does not favour models with zero intrinsic scatter.

Figure 2.8 shows that some correlations exist between the different parameters of the
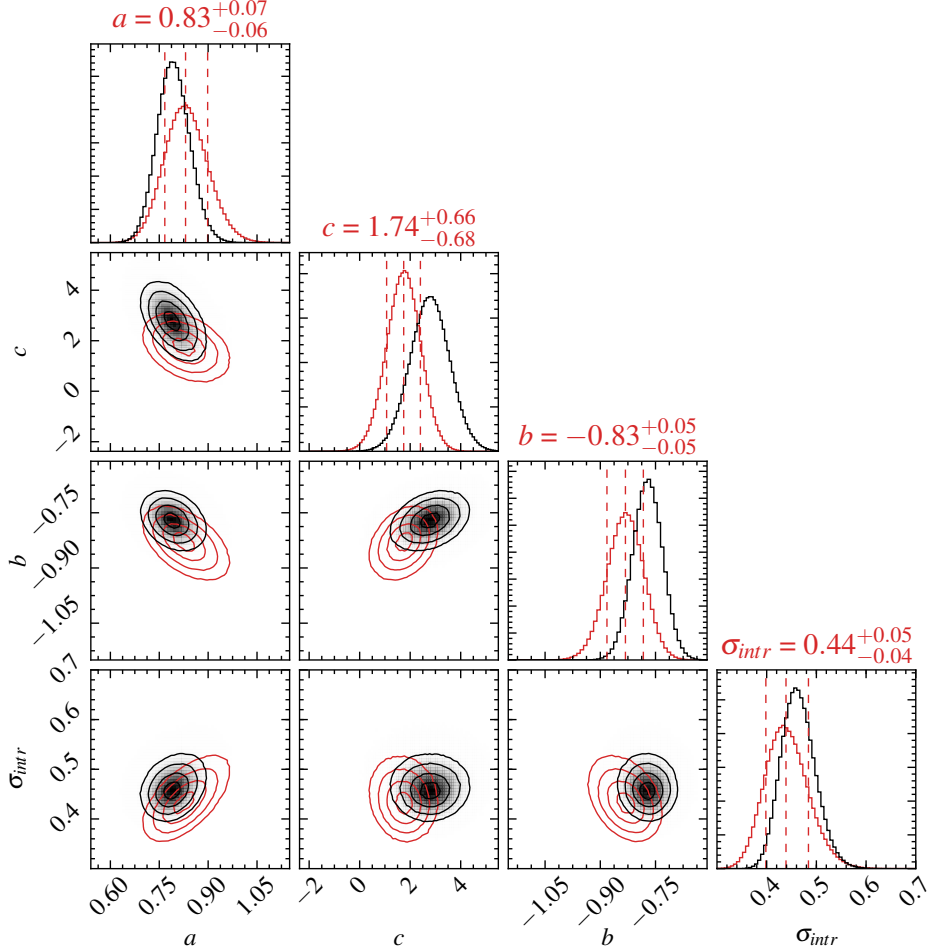
Figure 2.8: Projections of the 4D posterior distribution for the model parameters: slope (a), evolution (c), normalisation (b) and intrinsic scatter (σ_{intr}). The histograms on top show the marginalised distributions of the model parameters. The bias-corrected posterior median value and the 16th and 84th percentile are denoted by the dashed lines and by the values above the histograms. The contours show the 0.5, 1, 1.5 and 2 σ levels. The posterior directly from the modelling is shown in black, red indicates the posterior after applying the bias correction (Equation 2.19). Figure created using the corner.py module (Foreman-Mackey, 2016).
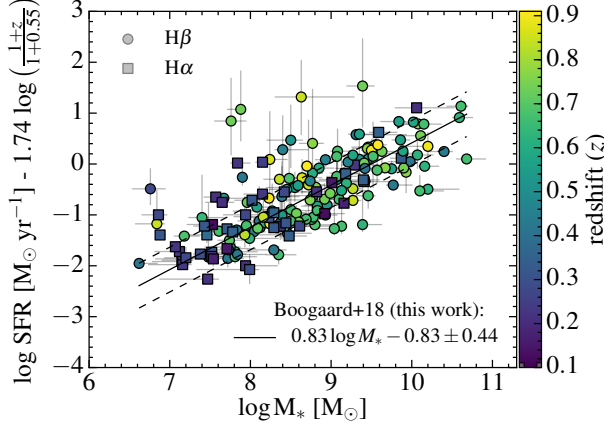
Figure 2.9: The best-fit star formation sequence for the 179 star-forming galaxies observed with MUSE. The symbols indicate the dust-corrected tracer used to infer the SFR. The solid line shows best-fit relation, as presented in Equation 2.11, and the dashed lines show the 1σ intrinsic scatter. We subtract the evolution from the y-axis and scale to the average redshift of the sample; $z = 0.55$. After accounting for evolution, the galaxies clearly follow the star formation sequence, down to the lowest masses and SFRs. The slightly larger fraction of galaxies that scatter into the high-mass, low-SFR regime may be a result of the flattening of the relation above $M_* = 10^{10}\,M_\odot$.

model, which is expected. The strongest correlation exists between slope and redshift evolution as a less steep slope requires more evolution in the normalisation to be compatible with the data. The complete covariance matrix between the different parameters is:

$$\Sigma(a, c, b, \sigma_{\mathrm{intr}}) = \begin{pmatrix} 0.003 & -0.019 & -0.001 & 0.000 \\ -0.019 & 0.620 & 0.011 & 0.000 \\ -0.001 & 0.011 & 0.002 & 0.000 \\ 0.000 & 0.000 & 0.000 & 0.001 \end{pmatrix}. \tag{2.10}$$

We correct the posterior for observational bias, by applying Equation 2.19, which is indicated by the red contours in Figure 2.8. This yields a steeper slope, with a significantly shallower redshift evolution:

$$\log \mathrm{SFR}[M_\odot\,\mathrm{yr}^{-1}] = 0.83^{+0.07}_{-0.06} \log\left(\frac{M_*}{M_0}\right) - 0.83^{+0.05}_{-0.05}$$

$$+ 1.74^{+0.66}_{-0.68} \log\left(\frac{1+z}{1+z_0}\right) \pm 0.44^{+0.05}_{-0.04}, \tag{2.11}$$

At the same time, the transformation has little effect on the intrinsic scatter. The covariance in the corrected posterior is essentially the same as the uncorrected one, with a slight increase

in covariance with intrinsic scatter.

$$\Sigma(a, c, b, \sigma_{\text{intr}}) = \begin{pmatrix} 0.004 & -0.016 & -0.002 & 0.002 \\ -0.016 & 0.459 & 0.010 & -0.003 \\ -0.002 & 0.010 & 0.002 & -0.001 \\ 0.002 & -0.003 & -0.001 & 0.002 \end{pmatrix}. \tag{2.12}$$

We compare the generative distribution (i.e. Equation 2.9) with the data in Figure 2.9. As the plane is three dimensional, we show a projection where we have subtracted the evolution with redshift from the y-axis. Overall, the distribution appears to describe the data very well and the scatter in the observations has tightened with respect to Figure 2.7. For a more familiar representation we also show the resulting star formation sequence in the right panel of Figure 2.7, for a number of different redshifts.

2.5.2 Low-mass sample ($\log M_*[M_\odot] < 9.5$)

We are primarily interested in the low-mass end of the star formation sequence. Our deep MUSE sample spans a significant mass range, between $\log M_*[M_\odot] = 6.5 - 11$. As several studies have suggested different characteristics for the star formation sequence above and below a turnover mass of $M_* \sim 10^{10} M_\odot$ (e.g. Whitaker et al., 2014; Lee et al., 2015; Schreiber et al., 2015), we repeat the above analysis excluding galaxies above a certain mass threshold. To be on the conservative side, we choose this mass threshold to lie at $M_* = 10^{9.5} M_\odot$. This excludes $31/179 \approx 17.5\%$ of the sample. We include this threshold as a dashed vertical line in Figure 2.7. We then repeat the modelling identically to what has been described in the previous sections.

The bias-corrected star formation sequence for galaxies that have a stellar mass below $M_* < 10^{9.5} M_\odot$ is:

$$\log \text{SFR}[M_\odot \, \text{yr}^{-1}] = 0.83^{+0.10}_{-0.09} \log \left(\frac{M_*}{M_0} \right) - 0.79^{+0.05}_{-0.05}$$
$$+ 2.22^{+0.75}_{-0.76} \log \left(\frac{1+z}{1+z_0} \right) \pm 0.47^{+0.06}_{-0.05}. \tag{2.13}$$

The result is essentially the same, with the main difference being a steeper redshift evolution. All parameters are within errors consistent with the relation for our complete sample (also for the uncorrected values, see Table 2.1). This reflects the fact that we are primarily sensitive to the low-mass end of the galaxy sequence. As this fit utilises only a part of the data we will refer primarily to the fit based on all the data, Equation 2.11, as the main result in the remainder of the paper. We report the (un)corrected values for all the fits in Table 2.1.

2.5.3 The effect of redshift bins (2D)

Most previous studies have not modelled the redshift evolution of the star formation sequence directly, but have instead divided the data into redshift bins and adopted a non-evolving relation: $\log \text{SFR} = a \log M_* + b$. To facilitate the comparison with the literature, we adapt our

Table 2.1: Star formation sequence parameters for different samples.

Sample	Size	a	b	c	σ_{intr}
3D		$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = a \log(M_*/M_0) + b + c \log(1+z)/(1+z_0)$			
Full	179	$0.79^{+0.05}_{-0.05}$	$-0.77^{+0.04}_{-0.04}$	$2.78^{+0.78}_{-0.78}$	$0.46^{+0.04}_{-0.03}$
$\log M_*[\mathrm{M}_\odot] < 9.5$	148	$0.79^{+0.08}_{-0.07}$	$-0.73^{+0.04}_{-0.04}$	$3.39^{+0.91}_{-0.90}$	$0.49^{+0.04}_{-0.04}$
3D – bias corrected (via Equation 2.19)					
Full	179	$0.83^{+0.07}_{-0.06}$	$-0.83^{+0.05}_{-0.05}$	$1.74^{+0.66}_{-0.68}$	$0.44^{+0.05}_{-0.04}$
$\log M_*[\mathrm{M}_\odot] < 9.5$	148	$0.83^{+0.10}_{-0.09}$	$-0.79^{+0.05}_{-0.05}$	$2.22^{+0.75}_{-0.76}$	$0.47^{+0.06}_{-0.05}$
2D		$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = a \log(M_*/M_0) + b$			
Full	179	$0.89^{+0.05}_{-0.05}$	$-0.82^{+0.04}_{-0.04}$		$0.49^{+0.04}_{-0.04}$
$0.1 < z \leq 0.5$	72	$0.86^{+0.09}_{-0.08}$	$-0.92^{+0.07}_{-0.07}$		$0.57^{+0.07}_{-0.06}$
$0.5 < z < 1.0$	107	$0.84^{+0.07}_{-0.06}$	$-0.73^{+0.06}_{-0.06}$		$0.46^{+0.05}_{-0.05}$

Notes. For a full description of the different samples, see § 2.5. $M_0 = 10^{8.5}\,\mathrm{M}_\odot$ and $z_0 = 0.55$.

model to fit the relation in the ($\log M_*$, $\log \mathrm{SFR}$)-plane, without taking the redshift evolution into account. This is easily done, by taking a two-dimensional version of our likelihood, disregarding the second, $\log(1 + z)$-component in Eq. (2.5)–(2.8) — the rest of the modelling is be identical. We note that we still take both heteroscedastic errors as well as intrinsic scatter into account (see § 2.4.1), however, we do not apply the bias correction.

We model both the entire redshift range, as well as the $0.1 < z < 0.5$ and $0.5 < z < 1.0$ range separately (similar to other studies). The results are collected in Table 2.1. For the full sample the slope is significantly steeper than when we take into account the redshift evolution, when comparing to our uncorrected fits:

$$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = 0.89^{+0.05}_{-0.05} \log\left(\frac{M_*}{M_0}\right) - 0.82^{+0.04}_{-0.04}. \tag{2.14}$$

This is also the case for the smaller samples in both redshift bins, although the results are consistent with Equation 2.9 within the error bars (which are larger due to lower number statistics). The resulting relations are, for $0.1 < z \leq 0.5$;

$$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = 0.86^{+0.09}_{-0.08} \log\left(\frac{M_*}{M_0}\right) - 0.92^{+0.07}_{-0.07}, \tag{2.15}$$

and for $0.5 < z < 1.0$;

$$\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = 0.84^{+0.07}_{-0.06} \log\left(\frac{M_*}{M_0}\right) - 0.73^{+0.06}_{-0.06}. \tag{2.16}$$

Given the significant evolution we found in the star formation sequence with redshift, this result is expected. While incidently these slopes are similar to our corrected fits, we caution that this does not imply that not modelling the redshift evolution can circumvent biases introduced by flux-limited observations.

2.6 Discussion

We have modelled the star formation sequence down to 10^8 M$_\odot$ at $0.11 < z < 0.91$ using a Bayesian framework (§ 2.4) that takes into account both the heteroscedastic errors on the observations as well as the intrinsic scatter in the relation. One major advantage of our framework is that we simultaneously model both the slope and the evolution in the M_*-SFR relation, while most previous studies have modelled these separately by dividing their sample into different redshift bins. As demonstrated in § 2.5.3, these results are not necessarily consistent, which can be attributed to evolution taking place within a single redshift bin. Another important difference is that we use the Balmer lines to trace the (dust-corrected) star formation, while most other recent studies have relied on SFRs derived from UV+IR/SED-fitting, using different dust corrections (Whitaker et al., 2014; Lee et al., 2015; Schreiber et al., 2015; Kurczynski et al., 2016).

As described in § 2.5.1, we have found that the star formation sequence (shown in Figure 2.7 and Figure 2.9) is well described by Equation 2.11 (see also Table 2.1). We now compare our results to other literature measurements and discuss each aspect of the star formation sequence separately, i.e. the redshift evolution, intrinsic scatter and the slope. We focus particularly on the slope, for which we find the strongest constraints, and continue with a discussion of the physical implications of our results.

2.6.1 Comparison with the literature

Evolution with redshift

We find that the normalisation in the star formation sequence increases with redshift as $(1 + z)^c$ with $c = 1.74^{+0.66}_{-0.68}$ ($2.22^{+0.75}_{-0.76}$ for $M_* < 10^{9.5}$ M$_\odot$). The fact that the normalisation of the star formation sequence increases with redshift is well known and attributed to the change in cosmic gas accretion rates and gas depletion timescales. Most studies have probed the higher mass regime and report values in the range of sSFR $\equiv$ SFR$/M_* \propto (1 + z)^{2.5-3.5}$ at $0 < z < 3$ (e.g. Oliver et al., 2010; Karim et al., 2011; Ilbert et al., 2015; Schreiber et al., 2015; Tasca et al., 2015). Looking specifically at the low-mass regime, Whitaker et al. (2014) reports sSFR $\propto (1 + z)^{1.9}$, similar to our result. Their more massive end indeed shows stronger evolution sSFR $\propto (1 + z)^{2.2-3.5}$. Lee et al. (2015) on the other hand, find much steeper evolution, with sSFR $\propto (1 + z)^{4.12\pm0.1}$. We note that our parameterisation assumes a power-law type of evolution of the star formation sequence with redshift. We have decided to stick to this very common first-order approximation. Still, one should keep in mind that a more complex evolution with redshift is possible, both non-linear in time as well as a different evolution in different mass regimes. We do not find strong constraints on the redshift evolution due to our relatively small redshift range from $z = 0.1$ to $z = 0.91$. Still, the results from § 2.5.3 show that it is important to take the redshift evolution into account, in order to get a robust constraint on the slope.

Intrinsic scatter

Constraining the intrinsic scatter in the star formation sequence has proven to be challenging as one has to separate the intrinsic scatter from the measurement error (e.g. Noeske et al., 2007a; Salim et al., 2007; Salmi et al., 2012; Whitaker et al., 2012; Guo et al., 2013; Speagle et al., 2014; Schreiber et al., 2015). This challenge in particular motivates our adopted model, which directly constrains the amount of intrinsic scatter in the relationship, even in the presence of measurement errors. Meanwhile, our measurements are not affected by binning, e.g. we do not boost the scatter artificially because of evolution of the star formation sequence within a single bin.

In our best fit model we find $\sigma_{intr} = 0.44^{+0.05}_{-0.04}$ dex, which is larger than the value of $\sim 0.2 - 0.4$ dex that is commonly found (e.g. Speagle et al., 2014; Schreiber et al., 2015). Kurczynski et al. (2016) determined an intrinsic scatter of $\sigma_{intr} = 0.427 \pm 0.011$ in their lowest redshift bin ($0.5 < z < 1.0$) in the HUDF, similar to our result, but found significantly smaller scatter at higher redshifts. They determined the intrinsic scatter by decomposing the total scatter ($\sigma_{Tot} = 0.525$) using the covariance matrix between M_* and SFR determined from their SED fitting.

There are several effects that could potentially affect the scatter. Measurement outliers are not a cause of concern for the intrinsic scatter as they are taken into account by the likelihood approach. However, if galaxies are included in the sample that are not on the M_*-SFR relation, such as red-sequence galaxies or starbursts, then these might artificially increase the scatter. We argue that the former is unlikely as our selection criteria based on the 4000 Å break and the Hα λ6565 or Hβ λ4863 equivalent width effectively remove all red-sequence galaxies from the sample. On the other hand, our sample does include a small number of galaxies that are offset from the relation towards high SFRs. We verified however that removing all galaxies with a sSFR > 10 Gyr^{-1} from the sample does not significantly increase or decrease the scatter.

Hypothetically, if the error bars on the SFR are underestimated, this will artificially boost the intrinsic scatter in the relationship. To determine the influence of the size of the error bars we redid the modelling while folding in an additional error on the SFR of 0.2 dex in quadrature (effectively doubling the average error bars); this decreased the scatter by 20% to ~ 0.4 dex. The sample size does not seem to affect the measurement and splitting our sample did not yield significantly larger scatter (see § 2.5.3).

Assuming our measured scatter is real, it might be that previous studies have underestimated the amount of intrinsic scatter. One potential danger might lie in the derivation of both stellar mass and SFR from the same photometry. Especially in SED modelling this might introduce correlations between M_* and SFR as both are regularised through the same star formation history in the model spectrum which could artificially decrease the scatter.

More physically, the difference could also in part be due to the fact that the Balmer lines trace the SFR on shorter timescales (stars with ages ≤ 10 Myr and masses $> 10\,M_\odot$) than the UV does (ages of ≤ 100 Myr and masses $> 5\,M_\odot$; e.g. Kennicutt 1998a; Kennicutt & Evans 2012). Simulations have indeed found that SFRs averaged over timescales decreasing from 10^8 to 10^6 yr could be significantly larger (Hopkins et al., 2014; Sparre et al., 2015), particularly if star formation histories are bursty (e.g. Dominguez et al., 2015; Sparre et al., 2017).

Furthermore, as the recent star formation histories of low-mass galaxies are more diverse, it can be expected that there is more scatter in the star formation sequence at low stellar masses. This indeed has been predicted by simulations (e.g. Hopkins et al., 2014; Sparre et al., 2017) as well as semi-analytical models (e.g. Mitra et al., 2017). Observing such a trend requires a large and highly complete sample of galaxies over an extended mass range and hence evidence has been inconclusive. Using a large sample of galaxies from the SDSS, Salim et al. (2007) reported a decrease in the scatter of -0.11 dex^{-1} from $10^8 - 10^{10.5}$ M$_\odot$, but such a trend with mass has not been confirmed by studies at higher masses (Whitaker et al., 2012; Guo et al., 2013; Schreiber et al., 2015; Kurczynski et al., 2016). Recently though, Santini et al. (2017) have found indications of decreasing scatter with mass in the Frontier Fields, albeit at higher redshifts ($z > 1.3$).

A large and complete sample of galaxies, covering the $(\log M_*, \log \text{SFR}, \log(1 + z))$-space, with independent stellar mass and SFR estimates, is required to get a firm handle on the intrinsic scatter in the star formation sequence.

Slope

We find a best-fit (median posterior) slope of the star formation sequence of $a = 0.83^{+0.07}_{-0.06}$ ($\log \text{SFR} \propto a \log M_*$). This slope is determined from galaxies that are more than an order of magnitude lower in mass than most earlier studies at $z > 0$, i.e. at 10^8 to 10^{10} M$_\odot$, whereas most previous studies (e.g. Speagle et al., 2014; Lee et al., 2015; Schreiber et al., 2015) have been primarily sensitive to a higher mass range from $10^{9.5}$ M$_\odot$ to 10^{11} M$_\odot$. For reference, we plot the polynomial fit from Whitaker et al. (2014) (down to their mass completeness limit, based on stacking) in Figure 2.7.

Recent studies have typically observed a shallower slope at the high-mass end, i.e. above 10^{10} M$_\odot$ (e.g. Whitaker et al., 2014). Gavazzi et al. (2015) find a turnover mass of $M_* \sim 10^{9.7}$ M$_\odot$ at $z = 0.55$ (after converting their result to a Chabrier IMF), increasing with redshift. As discussed in § 2.5.2, excluding galaxies above $M_* > 10^{9.5}$ M$_\odot$ has no significant effect on the slope. Only $15/179 \approx 8.5\%$ of galaxies in our sample have $M_* > 10^{10}$ M$_\odot$ and thus our result is not very sensitive to this turn-over. In light of this, we limit the following discussion to studies which specifically probe the mass range below the turnover of the star formation sequence.

Our best-fit slope of $0.83^{+0.07}_{-0.06}$ is compared to the values found by other recent studies in Figure 2.10 where we focus on studies with similar redshift ranges (i.e. $0 < z < 1$) and which extend well below $M_* < 10^{10}$ M$_\odot$. The slope in this regime is notably steeper than the consensus relation from Speagle et al. (2014) who reported $a = 0.6 - 0.7$ at our redshifts, due to the fact that this compilation is for a mass range of $\log M_*[\text{M}_\odot] = 9.7 - 11.1$, where the slope is significantly shallower. Our slope is shallower than the low-mass power-law slope from Whitaker et al. (2014) ($a = 0.94 \pm 0.03$ for $M_* < 10^{10.2}$ M$_\odot$) from the 3D-HST catalogues in CANDELS, but is consistent with the global slope of $a = 0.88 \pm 0.06$ reported by Lee et al. (2015) in a large sample of star-forming galaxies in COSMOS. Kurczynski et al. (2016) have also presented a characterisation of the star formation sequence in the HUDF, based on the CANDELS/GOODS-S (Santini et al., 2015) and UVUDF (Rafelski et al., 2015) catalogues. In their lowest redshift bin ($0.5 < z < 1.0$), which goes down to $M_* \sim 10^{7.5}$M$_\odot$ they find a
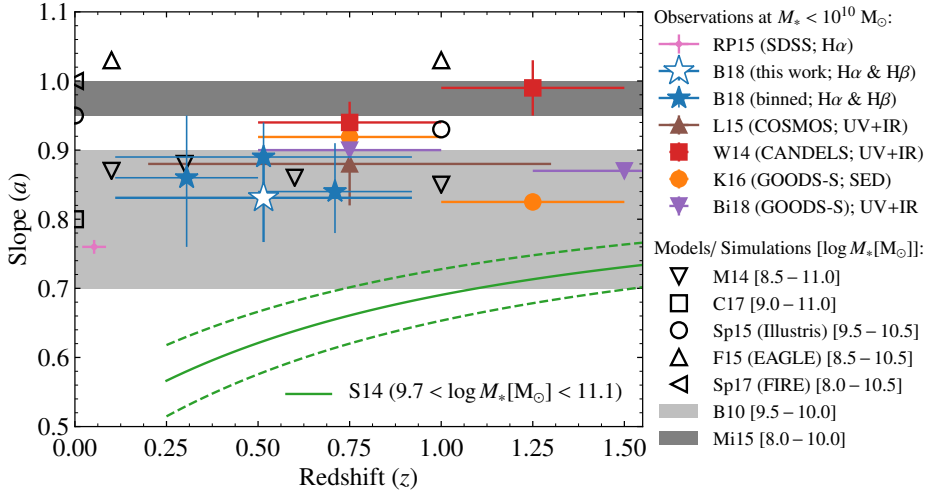
Figure 2.10: (a) A comparison of the slope (a) as a function of redshift (z) with studies from the literature that extend down to $M_* < 10^{10}$ $M_\odot$ near our redshift range. Our best-fit (bias-corrected) slope from Equation 2.11 is shown by the large star. As most studies have probed the slope in bins of redshift, we also include our results obtained using non-evolving redshift bins (Eqs. (2.14), (2.15) and (2.16); smaller blue stars). The literature results are from Renzini & Peng (2015, RP15), Kurczynski et al. (2016, K16), Bisigello et al. (2018, B18) and the low-mass ($M_* \lesssim 10^{10}$ $M_\odot$) power-law slopes from Whitaker et al. (2014, W14) and Lee et al. (2015, L15). We also add Speagle et al. (2014, SP14) for reference, though it is inferred at higher masses. We indicate the field and SFR-tracer in brackets, though note that distinct calibrations for the same tracer may be used in different studies. In addition, we add the slopes predicted by (semi-)analytical models; Bouché et al. (2010, B10), Mitchell et al. (2014, M14), Mitra et al. (2015, Mi15), Cattaneo et al. (2017, C17), and hydrodynamical simulations; Sparre et al. (2015, Sp15), Furlong et al. (2015, F15), Sparre et al. (2017, Sp17).

slope of $a = 0.919 \pm 0.017$, which is also steeper (marginally consistent) compared to what find. We note that they determined both masses and SFRs from the SED modelling, taking into account the correlations between the parameters, as their study was focused particularly on measuring the intrinsic scatter, see § 2.6.1. In the same field Bisigello et al. (2018) find a slope of 0.9 ± 0.01 ($0.5 \leq z < 1.0$), after selecting galaxies with log sSFR[Gyr^{-1}] < -9.8.

The Sloan Digital Sky Survey (SDSS; York et al. 2000; Abazajian et al. 2009) serves as a natural reference for Balmer line-derived SFRs in the local universe and since Brinchmann et al. (2004) different studies have derived the star formation sequence slope (e.g, Salim et al., 2007; Elbaz et al., 2007). The most recent of these is Renzini & Peng (2015), who measure the slope of the ridge line in the $M_* - N \times$ SFR-plane (where N is the number of galaxies in every M_*-SFR bin) and find $a = 0.76 \pm 0.01$, which is significantly flatter than our results.

Taken at face value, our slope of $a = 0.83^{+0.07}_{-0.06}$ is inconsistent with a linear slope ($a = 1$). A value (close to) unity may have been expected on the basis of simulations (see next section), which is also evident from the fact that several parameterisations of the star formation

sequence asymptote to a linear relation at low mass (e.g. Schreiber et al., 2015; Tomczak et al., 2016). An independent motivation for a near-linear value comes from the fact that there is very little evolution in the faint slope of the stellar mass function of star-forming galaxies up to $z = 2$ (see, e.g. Tomczak et al. (2014); Davidzon et al. (2017) for recent results). To first order, this may implies self-similar mass growth for low-mass galaxies (i.e. constant sSFR which implies a linear slope for the star formation sequence), unless balanced by mergers (Peng et al., 2014). Leja et al. (2015) investigated the link between the slope of the star formation sequence and the stellar mass function. While they do not provide precise constraints on the low-mass slope at low redshift (due to the challenge of disentangling growth through star formation and mergers), their results indicate that a sub-linear low-mass slope is still consistent with the stellar mass functions at $z < 1$.

Evolution of the low-mass slope

Combining results from the local universe out to redshift $z \sim 6$, Speagle et al. (2014) found evidence for an evolving slope at the high-mass end ($M_* > 10^{9.7} \, M_\odot$), where the slope gets shallower with redshift (cf. Abramson et al., 2016, Fig. 5). Given the turnover in the star formation sequence at high mass, it is important to disentangle to what extent the evolution in the slope is due to different studies being sensitive to distinct mass regimes. Our data are too sparse in redshift space to simultaneously constrain the evolution of the slope (and hence we have adopted a single power-law slope for the sequence).

In light of the potential redshift evolution of the slope, we plot the slope as a function of redshift in Figure 2.10, compared to literature results which probe the mass range $M* < 10^{10} \, M_\odot$ at $z < 1.5$. Figure 2.10 provides evidence for evolution of the low-mass slope with redshift. However, we caution against a too strong interpretation of such a trend as the literature suffers from studies probing distinct mass ranges (sometimes including the turn-over regime). What further complicates a fair comparison is that different tracers of star formation probe different timescales and additionally use varying dust corrections, which are not necessarily consistent (e.g. Davies et al., 2016). A consistent analysis of the low-mass galaxy population out to higher redshifts is important to quantify potential evolution in the low-mass slope.

2.6.2 The MS slope — a quantitative comparison to models

The galaxy main sequence (MS) is a natural outcome of hydrodynamical models (e.g. Fig. 1b in Bouche et al. 2005; Davé 2008; Genel et al. 2014; Torrey et al. 2014; Kannan et al. 2014; Hopkins et al. 2014; Sparre et al. 2015; Furlong et al. 2015) and in semi-analytical models (e.g. Somerville et al., 2008; Dutton et al., 2010; Cattaneo et al., 2011; Mitchell et al., 2014; Henriques et al., 2015; Hirschmann et al., 2016; Cattaneo et al., 2017). These models have reported a slope (and scatter) that, in general, is broadly consistent with observations, but the quantitative details regarding the slope and/or the evolution of the main sequence often do not match observations.

Since the pioneering work of Daddi et al. (2007) and Elbaz et al. (2007), it has been noted that the redshift evolution of the main sequence normalisation, in particular around $z = 2$,

is a challenge for models (e.g. Davé, 2008; Damen et al., 2009; Bouché et al., 2010; Dutton et al., 2010; Dekel & Mandelker, 2014; Torrey et al., 2014; Genel et al., 2014; Mitchell et al., 2014; Furlong et al., 2015; Sparre et al., 2015; Abramson et al., 2016; Santini et al., 2017). Here, we focus on a quantitative comparison of the slope of the main sequence (SFR $\propto M_*^a$) with various models, given that our study yields the tightest constraint on this parameter (compared to the other parameters in the model).

The Illustris simulations (Vogelsberger et al., 2014; Genel et al., 2014; Sparre et al., 2015) produce a main-sequence with a slope a that is slightly sub-linear with $a \lessgtr 1.0$. In particular, Genel et al. (2014) noted that sSFR goes as $\simeq -0.1$ with stellar mass and using the results from Sparre et al. (2015), we find that the main sequence in Illustris goes as SFR $\propto M_*^{\approx 0.95}$. The EAGLE simulations (Schaye et al., 2015; Crain et al., 2015) also allow an investigation of the main sequence and Furlong et al. (2015, their Fig. 5), showed that the sSFR is constant with M_* from 10^8 to 10^{10} $M_\odot$ at redshifts $z = 0.1, 1.0$ and 2.0, with a relatively steep decline above 10^{10} $M_\odot$. Quantitatively, below 10^{10} $M_\odot$, the slope of the main sequence a in Furlong et al. (2015) is $a \approx 1.04$. The MS slope for the Illustris and EAGLE simulations are shown in Figure 2.10 as the open circles and triangle symbols, respectively. In the FIRE simulations (Hopkins et al., 2014), Sparre et al. (2017) focused on studying the scatter in the main sequence for different tracers of SFR and shows a slope of $a \approx 0.98$ when using the FUV (their Fig. 2).

The MS slope has also been a challenge for semi-analytical models because different (regular) feedback prescriptions do not alter the MS slope as shown in Dutton et al. (2010) and discussed in Mitchell et al. (2014) (however, it can alter the slope in hydrodynamical simulations, e.g. Haas et al. 2013a,b; Crain et al. 2015). Mitchell et al. (2014) performed a detailed comparison between predictions from the GALFORM semi-analytical models with observations and their fiducial model produces a MS slope of $a \approx 0.85$ (shown in Figure 2.10 as the down-pointing triangles). Recently, the semi-analytical model of Cattaneo et al. (2017) using the GALICS2 code was set to reproduce the local luminosity function and the local MS slope simultaneously. Their MS slope is $a \approx 0.8$ (open square in Figure 2.10), but we caution their use of an extreme feedback model, where the mass loading η is $\eta \propto V^{-6}$, where V is the halo virial velocity. Such a steep scaling between galaxy mass and wind loading is not supported by the data (e.g. Schroetter et al., 2016).

Bouché et al. (2010) used a simple toy model for galaxy (self-)regulation with which they showed that variations in feedback prescriptions or in the laws of star formation have no impact on the MS slope. They argued that while ejective feedback alone is not sufficient to bring the theoretical slope of the main-sequence in agreement with observations, preventive feedback can easily do so as several studies have shown (Davé et al., 2012; Lu et al., 2015; Mitra et al., 2015, 2017). However, while the MS slope of Bouché et al. (2010) is sub-linear with $a \approx 0.8$, a quantitative analysis reveals that the slope varies rapidly with stellar mass, likely due to the limitations of the model. Indeed, the MS slope of Bouché et al. (2010) goes from 0.7 at $M_* \sim 10^{9.5}$ $M_\odot$ to 0.9 at $M_* \sim 10^{10.5}$ $M_\odot$. The range of values is indicated by the light grey box in Figure 2.10.

Mitra et al. (2015) expanded the self-regulation model of Bouché et al. (2010); Davé et al. (2012, and others) with physically motivated parameters and attempted to determine these parameters using a Bayesian MCMC approach on a set of observed scaling relations at $0 < z < 2$. Their fiducial model yields a MS with a slope that is quasi-linear with $a \sim 0.95$

in our mass regime, i.e. below 10^{10} M$_\odot$. Their MS slope is shown as the dark grey band in Figure 2.10.

Generally speaking, in the low-mass regime below 10^{10} M$_\odot$, hydrodynamical simulations have steeper MS slopes with $a \approx 1.0$ whereas our estimate ($a = 0.83^{+0.07}_{-0.06}$) at $z < 1$ and recent observations covering that mass range indicate $a < 1.0$ (see Figure 2.10). The reason that models tend to predict a steeper main sequence slope lies in the underlying feature in hydrodynamical simulations and semi-analytical models, where the growth rate for dark matter halos $\dot{M}_h$ scales with mass as $\dot{M}_h \propto M_h^{1.15}$ (Birnboim et al., 2007; Genel et al., 2008; Dekel et al., 2009; Fakhouri & Ma, 2008; Neistein & Dekel, 2008), in combination with rapid gas cooling.

2.6.3   Implications of a shallow slope

As noted originally by Noeske et al. (2007b) and discussed in Mitchell et al. (2014) and Abramson et al. (2016), a MS with a sub-linear slope, SFR $\propto M_*^a$ with $a < 1$, implies downsizing where lower-mass galaxies have longer e-folding time and a later onset of star formation. This downsizing effect would be amplified if the MS slope is substantially flatter above 10^{10} M$_\odot$ as some studies have indicated (Whitaker et al., 2014; Schreiber et al., 2015; Lee et al., 2015; Tomczak et al., 2016). This turnover has generally been attributed to either a morphological transition, such as bulge growth (Abramson et al., 2014; Lee et al., 2015; Whitaker et al., 2015), or a reduced star formation efficiency (Schreiber et al., 2016).

Our result, that the slope of the main sequence is sub-linear in the low-mass regime, implies that there are processes at work which either: (1) affect the conversion of the accreted gas into stars through increased (supernova) feedback or a decrease in the SF efficiency; or (2) prevent the accretion of gas onto low-mass galaxies. These two processes might conspire with the fact that the gravitational potential is shallower in low-mass galaxies (Mitra et al., 2015).

In hydrodynamical simulations low-mass galaxies (up to halo masses of $\sim 10^{11.5}$ M$_\odot$) obtain their gas primarily through 'cold'-accretion (Kere et al., 2005; van de Voort et al., 2011), where the gas is never heated to the virial temperature, while 'hot' accretion, where gas is first shock heated to the virial temperature and then cools and accretes, is dominant for more massive galaxies. A candidate process is feedback from gravitational heating, due to the formation of virial shocks (e.g. Faucher-Giguère et al., 2011), which becomes more effective at higher masses, however, can still play a role down to halo masses of 10^{10} M$_\odot$. The heating of gas through winds (from either supernovae or black hole feedback) can also prevent the gas from flowing into the galaxy (Oppenheimer et al., 2010; Faucher-Giguère et al., 2011; van de Voort et al., 2011), in particular in low-mass galaxies. However, Schaye et al. (2010) pointed out that this type of feedback mainly has a regulatory effect on the gas infall.

As noted by Dutton et al. (2010), Bouché et al. (2010), and Mitchell et al. (2014), in semi-analytical models, the MS slope is rather insensitive to the ejective (regular) feedback mechanisms,[25] such as the heating of gas through winds and/or the star formation efficiency (Kennicutt, 1998a) because they act primarily on the gas content. Hence, the SFR and stellar

---

[25]with mass loading $\eta \propto V^{-1}$ or $\eta \propto V^{-2}$ for momentum or energy-driven winds, respectively.

mass are affected in a similar way, leaving the slope unchanged, unless the ejective feedback prescription is strongly mass dependent with $\eta \propto V^{-6}$, as in Cattaneo et al. 2017. In addition, Mitchell et al. (2014) showed that the slope is also insensitive to the gas re-incorporation prescription (see also Mitra et al., 2015).

Preventive processes (Blanchard et al., 1992; Gnedin, 2000; Mo et al., 2005; Lu & Mo, 2007; Okamoto et al., 2008) that tend to be mass dependent can more easily impact the MS slope, the Tully-Fischer relation, and the luminosity function as argued by Bouché et al. (2010). A preventive process which can prevent the inflow of gas specifically in low-mass halos is photoionisation heating (Quinn et al., 1996). While it has been argued that this process is primarily effective in dwarf galaxies and becomes ineffective above halo masses of a few times 10^9 $M_\odot$ (e.g. Okamoto et al., 2008), Cantalupo (2010) suggest that photoionisation may still play a role for more massive halos if there is significant star formation.

2.7   Summary and conclusions

We have exploited the unique capabilities of the MUSE instrument to investigate the star formation sequence for low-mass galaxies at intermediate redshift ($0.11 < z < 0.91$). From the large number of sources detected with MUSE in the HUDF and HDFS we have constructed a sample of 179 star-forming galaxies down to $M_* \sim 10^8$ $M_\odot$, with a number of objects at even lower masses (Figure 2.4). The accurate spectroscopic redshifts from MUSE are combined with the deep photometry available over the HUDF and HDFS to determine a robust mass estimate for the galaxies in our sample through stellar population synthesis modelling.

With MUSE we can detect star-forming galaxies down to SFR $\sim 10^{-3}$ $M_\odot$ yr^{-1} (Figure 2.7). We show that we can determine robust, dust-corrected SFR estimates from Hα λ6565 and Hβ λ4863 recombination lines, by comparing the SFRs from different tracers (Figure 2.5). A dust-corrected star formation rate is inferred from the Hα λ6565 and Hβ λ4863 emission lines observed with S/N > 3 in the MUSE spectra.

We characterise the star formation sequence by a Gaussian distribution around a plane (Equation 2.4). This methodology is chosen to maximally exploit the data set taking into account heteroscedastic errors. We constrain the slope, normalisation, intrinsic scatter, and evolution with redshift from the posterior probability distribution via MCMC methods (Figure 2.8).

We analyse the robustness of our model and the influence of the MUSE detection limit on the derived properties of the star formation sequence, by determining how well we can recover the parameters from a sample of simulated relations (detailed in § 2.A). Using the results, we correct our inferred parameters for observational biases.

We report a best-fit description of the low-mass end of the galaxy star formation sequence of $\log \text{SFR} = 0.83^{+0.07}_{-0.06} \log M_* - 0.83^{+0.05}_{-0.05} + 1.74^{+0.66}_{-0.68} \log(1 + z)$ between $0.11 < z < 0.91$, shown in Figure 2.9. The full description of our parameters, including errors and normalisation, is found in Equation 2.11.

The intrinsic scatter around the sequence is found to be $\sigma_{\text{intr}} = 0.44^{+0.05}_{-0.04}$ dex (in log SFR). This is notably higher than the average value reported in literature (~ 0.3 dex), which could be attributed to a combination of the Balmer lines probing star formation on shorter

timescales and the star formation histories of low-mass galaxies being more diverse.

Excluding massive galaxies (with $M_* > 10^{9.5}$ M$_\odot$) has no significant effect on the best-fit parameters, indicating we are primarily sensitive to low-mass galaxies. Notably though, we find that the slope steepens when splitting our sample into one or multiple redshift bins, with the values going up to $\log \mathrm{SFR}[\mathrm{M}_\odot\,\mathrm{yr}^{-1}] = 0.89^{+0.05}_{-0.05} \log M_*[\mathrm{M}_\odot]$. This shows the importance of taking into account the evolution with redshift when deriving the properties of the star formation sequence.

The slope of the star formation sequence is an important observable as it provides information on the processes that regulate star formation in galaxies. Our slope is shallower than most simulations and (semi-)analytical models predict, which find a (super-)linear slope essentially due to the growth rate of dark matter halos. Feedback processes operating specifically in the low-mass regime, which affect the accretion of gas onto galaxies and/or subsequent star formation, are required to reconcile these differences. Models suggest that supernova feedback or a decreased star formation efficiency do not affect the slope of the star formation sequence. Instead, processes that prevent the accretion of gas onto low-mass galaxies are thought to play an important role in determining the slope of the star formation sequence in the low-mass regime.

Acknowledgements

Appendix 2.A   Simulations

2.A.1   Selection function and completeness

We have selected galaxies based on the signal-to-noise of their emission lines, without any photometric preselection. This means the selection function is essentially determined by the emission line sensitivity. In general, one might expect galaxies with higher S/N in their emission lines to have a higher SFR at a fixed mass, or similarly, for galaxies with the same S/N to have a higher SFR at higher redshift, which potentially introduces biases in our results. Additionally, we can only observe galaxies that have Balmer lines in the spectral range of MUSE ($z < 0.91$).

To investigate the influence of these selections, we determine how well we can recover the true parameters of the star formation sequence from a set of mock samples of galaxies, after applying the flux limit from our MUSE observations.

We determine the influence of the selection function on the inferred parameters by simulating mock data for a range of 'true' parameters. The range of values for each mock parameter is listed in Table 2.2, which combine to form a grid of $N = 1260$ points. The extent of grid is chosen such that it encompasses a wide range of possible parameters and we find that the results are consistent if we enlarge the grid even further (note that, if the grid is taken too large, non-linearities may arise at the extreme values which potentially bias the linear transformation approach of § 2.A.2). We denote each set of parameters as $\mathbf{x}_{\text{true},k} = \left(\hat{a}, \hat{c}, \hat{b}, \hat{\sigma}_{\text{intr}} \right)^T$ with $k = 1, ..., N$.

We generate realistic mock data for each set of parameters through the following procedure: We sample 100 galaxies from a uniform distribution in both mass ($7.0 < \log M_*[\text{M}_\odot] < 10.5$) and redshift ($0.1 < z < 1$). Given the mass and redshift, we compute the SFR (via Equation 2.4), i.e. assuming a mock main sequence distribution with slope $\hat{a}$ and evolution $\hat{c}$. We choose our normalisation ($\hat{b}$) such that a 10^{10} M$_\odot$ galaxy at $z = 0$ has a SFR of 1 M$_\odot$/yr, similar to our results and, e.g. the Milky Way (Chomiuk & Povich, 2011), i.e. we take a zero-point of $b_0 = -10$. We then sample up to $b_{\text{offset}} = \pm 0.4$ dex above and below this zero-point. We provide each galaxy with a random offset from the main sequence (perpendicular to the ($\log M_*$, $\log$ SFR)-relation) drawn from $\mathcal{N}(0, \hat{\sigma}_{\text{intr}})$. Finally, we apply a random measurement error for each galaxy in both log stellar mass and log SFR of 0.3 dex (i.e. drawn from $\mathcal{N}(0, 0.3)$) and in log redshift of 5×10^{-4} dex ($\sim \mathcal{N}(0, 5 \times 10^{-4})$), similar to the observations.

We then apply the same flux limit as our shallowest MUSE observations, namely in the mosaic with 3×10^{-19} erg s^{-1} cm^{-2}, and mark all 'observed' galaxies as those that fall above our detection threshold (we do not take an additional factor for dust into account as our galaxies are not very dusty on average). We then fit the observed galaxies above the flux limit. Repeating this process 30 times for each individual set of parameters $\mathbf{x}_{\text{true}}$ and marginalising over the combined posterior distribution, we determine the corresponding recovered parameters $\mathbf{x}_{\text{out},k} = (a, c, b, \sigma_{\text{intr}})^T$.

As an example, we show one the experiment for a particular set of parameters in Figure 2.11. It is clear that the recovered parameters are biased towards a shallower slope and a steeper redshift evolution. The magnitude of this bias depends on all the parameters and

Table 2.2: Grid values for our mock simulations.

	min	max	step
$\hat{a}$	0.7	1.1	0.05
$\hat{c}$	1.5	4.5	0.5
b_{offset}	-0.4	0.4	0.2
$\hat{\sigma}_{\text{intr}}$	0.3	0.6	0.1
$\hat{b} = \hat{a}\,(\log(M_0) - b_0) + \hat{c}\log(1 + z_0) + b_{\text{offset}}$			

Notes. The normalisation ($b_0 = -10$) is chosen such that a $10^{10}\,\mathrm{M}_\odot$ galaxy at $z = 0$ has a SFR of $1\,\mathrm{M}_\odot\,\mathrm{yr}^{-1}$.

becomes more severe for steeper slopes and shallower redshift evolutions.

To check our methods, we also fit all simulated galaxies (without discarding any data). Reassuringly, we recover our input parameters to within the errors, even when simulating only 100 galaxies. Since our actual sample size is 179 galaxies, we are in principle able to recover the true parameters of the relation, even in the case of intrinsic scatter and heteroscedastic errors. One feature that does draw attention is that the redshift evolution is marginally steeper than the input relation (but admittedly poorly constrained and still consistent within the error). This can be explained due to an intricacy of the model, which assumes that the intrinsic scatter about the relation is along the normal vector to the plane ($\sigma_\perp$ in § 2.4.1), i.e. also in the $\log(1 + z)$-direction. If the data are truncated and there is a non-zero slope ($|c| > 0$) in redshift space, this may introduce an artificial bias in the corresponding slope (and scatter) as the truncation boundaries are not parallel to the normal vector. Given the fact that our data (and mock sample) are limited in redshift space by the spectral range of MUSE, this means that we may have slight artificial bias towards a steeper redshift evolution. For interpreting the intrinsic scatter this is not a problem as we can project the scatter along the (physical) log SFR-axis (which is our σ_{intr}).

With our simulations in hand however, we are now in place to apply a correction for both biases identified above.

2.A.2   Transformation

The simulations show a reasonably well behaved transformation between the true and recovered slope. We therefore model the mock data with an affine transformation, to be able to transform between the measured and true parameters.

We try to find the best transformation matrix A and vector $\mathbf{b}$ between the measured and true parameters. For each set of input ($\mathbf{x}_{\text{true},k}$) and output ($\mathbf{x}_{\text{out},k}$) parameters we have:

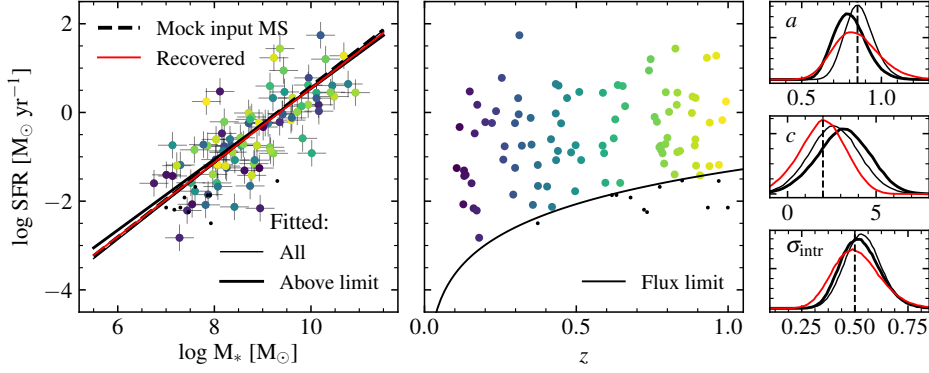$$\mathbf{x}_{\text{out},k} \approx A\mathbf{x}_{\text{true},k} + \mathbf{b} \tag{2.17}$$

Figure 2.11: Illustration of the results of the recovery experiment on mock galaxies. The points in the left and centre panels show one of the 30 realisations of 100 galaxies in $(\log M_*, \log(1+z), \log \mathrm{SFR})$-space from a mock star formation sequence: $\log \mathrm{SFR} \propto a \log M_* + c \log(1+z)$, where in this particular case $a = 0.8$ and $c = 2.0$ with $\sigma_{\mathrm{intr}} = 0.5$ dex. The colour indicates redshift, unless a mock galaxy falls below the solid line in the centre panel, indicating the flux limit of $\sim 3 \times 10^{-19}$ erg s^{-1} cm^{-2}, in which case it is a black point. The rightmost panels show the marginalised distributions (slope, redshift evolution, and intrinsic scatter) from combining all 30 realisations. The thin and thick black lines indicate the results when taking into account all mock data and only the data above the flux limit, respectively, and are compared to the input values (dashed lines). With all data points (including noise), we can recover the input parameters sequence well. When applying the flux limit a slight bias towards a shallower slope and steeper redshift evolution appears. We plot all curves in the leftmost panel at the average redshift of the sample (z_0). The red line is obtained after applying the correction to the fit of the data above the limit. These recovered curves are plotted in the leftmost panel as well and compared to the input mock relation. With our correction, we can recover the true input parameters well, even in the case of limited data.

We minimise the function

$$S(A, \mathbf{b}) = \sum_{k=1}^{N} ||\mathbf{x}_{\mathrm{out},k} - A\mathbf{x}_{\mathrm{true},k} - \mathbf{b}||_2^2 \tag{2.18}$$

with respect to each component of A and $\mathbf{b}$ in order to find the best-fit transformation A and $\mathbf{b}$ (Späth, 2004). We note that we do not take the errors on each point $\mathbf{x}_{\mathrm{out}, k}$ into account as their magnitudes are all comparable (essentially adding a constant to the equation).

With the best-fit A and $\mathbf{b}$ in hand, we can then invert the equation to obtain the relation between the observed and the recovered 'true' parameters, which denote as $\mathbf{x}'_{\mathrm{true}}$:

$$\mathbf{x}'_{\mathrm{true}} \approx A^{-1} (\mathbf{x}_{\mathrm{out}} - \mathbf{b}) \tag{2.19}$$

$$\begin{pmatrix} a' \\ c' \\ b' \\ \sigma'_{\mathrm{intr}} \end{pmatrix} = \begin{pmatrix} 1.336 & 0.014 & -0.150 & 0.171 \\ 0.638 & 0.863 & 0.574 & -2.621 \\ -0.178 & -0.008 & 1.175 & -0.185 \\ 0.285 & 0.009 & -0.044 & 1.091 \end{pmatrix} \left[\begin{pmatrix} a \\ c \\ b \\ \sigma_{\mathrm{intr}} \end{pmatrix} - \begin{pmatrix} 0.293 \\ 0.061 \\ -0.194 \\ 0.236 \end{pmatrix} \right] \tag{2.20}$$

For our simulated data, we show the distribution of the difference between the recovered parameters ($\mathbf{x}'_{\text{true}}$) and the true parameters ($\mathbf{x}_{\text{true}}$) in Figure 2.12. We recover the input parameters very well, with no mean offset between the recovered and the true parameter. This shows that the transformation (i.e. A and $\mathbf{b}$) are very well determined. Furthermore, the scatter in the differences is much smaller than the average uncertainty on each parameter obtained from the observations (of order $\sim 1\%$). As an illustration, we show the inverse transformation applied to the simulation by the red lines in Figure 2.11, which are now in good agreement with the true values (dashed lines).

In summary, the transformation obtained from the best-fit A and $\mathbf{b}$ is a very accurate description of the bias induced by the flux limit in our simulated data. We use the inverse of this transformation, Equation 2.19, in § 2.5 to correct our inferred posterior density distribution from modelling the MUSE data.
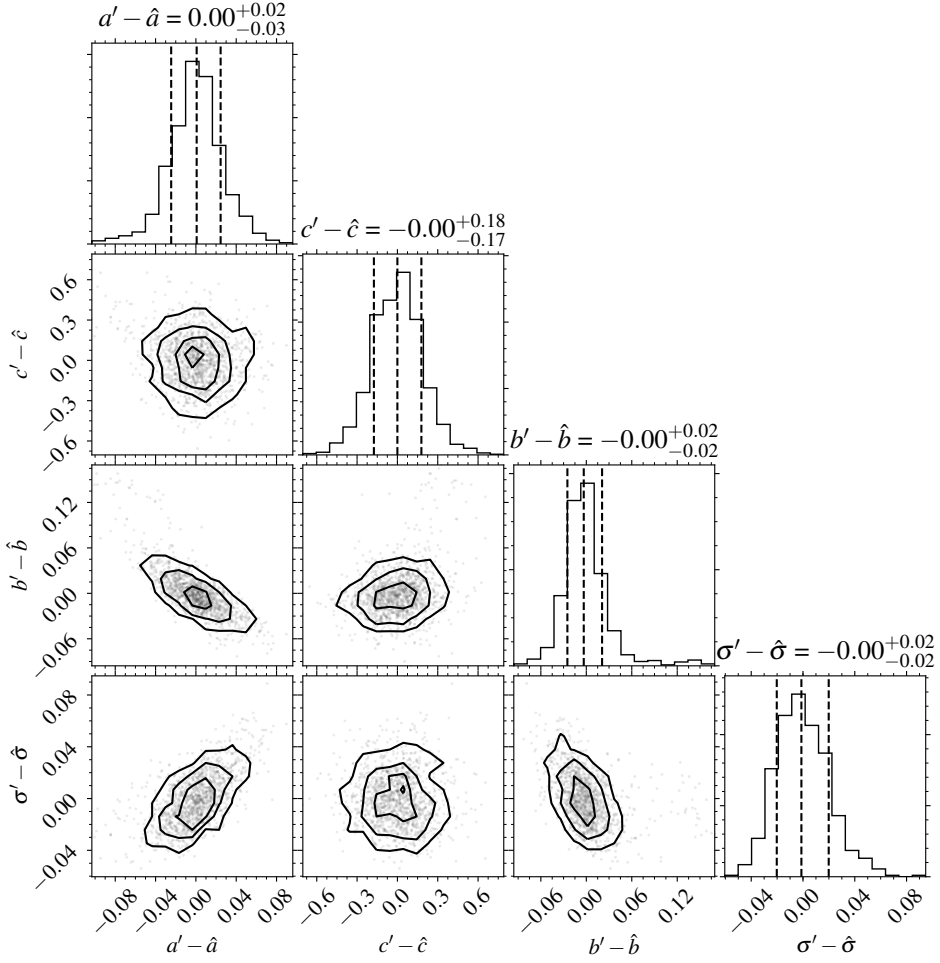
Figure 2.12: Plot of the differences between the recovered parameters, $\mathbf{x}'_{\text{true}} = \left(a', c', b', \sigma'_{\text{intr}}\right)^T$ and the true parameters, $\mathbf{x}_{\text{true}} = \left(\hat{a}, \hat{c}, \hat{b}, \hat{\sigma}_{\text{intr}}\right)^T$, for the $N = 1260$ points from our simulation; see Equation 2.19. We can recover the input parameters of our simulation very well, with no mean offset and very small scatter (compared to the uncertainty on each parameter obtained from the observations). Figure created using the `corner.py` module (Foreman-Mackey, 2016).