Rijksuniversiteit Groningen

Bachelor Thesis

The star formation rate in mergers and galaxy pairs in the KiDS-N-W2 and COSMOS fields

Author:
P.B. Folkertsma

Supervisor:
L. Wang

Abstract

We present an analysis of the relation between the star formation rates of merging galaxies. We do this by selecting mergers from close pairs and identification with a neural network. We select a functional form of the galaxy main sequence in two catalogues built upon the KiDS-N-W2 and COSMOS2020 fields. The catalogues contain 519403 and 26075 galaxies respectively. We compare how the merger fraction changes as a function of distance to the galaxy main sequence. We also compare the star formation rate of the merging galaxies to a control sample. We cover a redshift range of $0.1 < z < 3.0$. We find and compare pair fractions as a function of redshift to existing literature. We find that pairs have a decrease in star formation rate of $32 - 84\%$ compared to their control samples, our mergers have a decrease in star formation rate of $3.2 \pm 0.3\%$. We discuss potential sources of error in this analysis and compare with existing literature.

Contents

1 Introduction

In order to understand the properties of galaxies and how these affect each other it is important to look at galaxies at many different stages of their evolution. One process that has a major impact on the evolutionary track of galaxies is a galaxy merger (Garay-Solis et al., 2023). In this thesis the main focus is on star formation activity within merging galaxies. We compare the Star Formation Rate (SFR) in $M_\odot\,yr^{-1}$ of merging and non-merging galaxies in two large datasets.

1.1 Galaxy Mergers

The gravitational interactions between merging galaxies is understood to cause increased star formation via two pathways. Firstly, molecular hydrogen, which is the main fuel of star formation, can be formed from atomic hydrogen at a faster rate due to the gravitational interactions. The increased abundance of molecular hydrogen drives an increase in star formation. The second cause of increased star formation is the increase in density of the pre-existing molecular gas. This increased density is driven by the gravitational interactions, which decrease the angular momentum of gas and cause it to flow towards the galactic centre. This increases the star formation activity in the galactic centre (Violino et al., 2018).

We use two methods to select merging galaxies. The first is visual classification, which is increasingly being replaced in favour of machine learning algorithms, due to the rapidly increasing sizes of datasets and reproducibility of these models. Examples of machine learning models that are used to identify mergers are found in La Marca et al. (2024); Ferreira et al. (2020). The other method selects close pairs of galaxies from their separation in 3D space. The separation is derived from the the right ascension, declination, and redshift of a source. Examples of this method can be found in Mundy et al. (2017); Duncan et al. (2019); Ventou et al. (2019); Desmons et al. (2023), some also include limits on the mass ratio of galaxy pairs.

An important detail to note on the difference between close pairs and visual detection is that both methods probe different moments in the full timeline of a galaxy merger. Close pairs probe the time before the galaxies merge. Machine learning classification probes the full timeline of a galaxy merger, as long as morphological disturbances are detectable. For this reason these two methods should not be directly compared when discussing the impact of mergers on galaxy evolution, but instead they complement each other and allow for an increased understanding of the conditions in galaxies during a merger.

In order to classify the galaxies as a merger or non-merger a Convolutional Neural Network (CNN, Lecun et al. 1998) was trained by La Marca et al. (2024). The model was trained on data from the hydrodynamical simulations IllustrisTNG (Pillepich et al., 2018) and Horizon-AGN (Dubois et al., 2014). Galaxies were selected from the IllustrisTNG simulation as a merger when they had a merger event less than 300 Myr ago or would merge within 800 Myr. They only consider mergers within a redshift range of $0.1 < z < 1.0$ and with mass ratios of $1/4 < M_{*,1}/M_{*,2} < 4$. For training they randomly selected an equal number of non-mergers. For the Horizon-AGN simulation they use the same criteria to select mergers. They exclude any galaxy with a stellar mass of $M_* < 10^9 M_\odot$. For more details on the selection of the data and model training see La Marca et al. (2024) and the references therein.

There are varied results on the relation between (pre-)mergers and their SFR. For example, Garay-Solis et al. (2023) find no significant difference in SFR between interacting galaxies and their control sample, but they do note that this is in contrast to previous results. They use a sample of 236 star-forming galaxies of which 110 are pairs, merging or post-merger galaxies and the other 126 are non-mergers. These galaxies are from the Calar Alto Legacy Integral Field Area (CALIFA) survey (Sánchez et al., 2012) with additional infrared data from the Atacama

Pathfinder EXperiment (APEX) telescope (Güsten et al., 2006) and the Combined Array for Research in Millimeter-wave Astronomy (CARMA, Bock et al. 2006). Saintonge et al. (2012) show that not all mergers are necessarily associated with an increase in star formation, and that this is dependent on the molecular gas mass fraction (f_{H_2}) and depletion time of said gas. The depletion time is how long it would take to use up all molecular hydrogen at the current SFR. They use a sample of 365 galaxies with CO measurements to trace molecular hydrogen. The sample is constructed using observations from the Sloan Digital Sky Survey (SDSS, York et al. 2000) and the CO Legacy Database for the Galex-Arecibo-SDSS survey (COLD GASS, Saintonge et al. 2011). Meanwhile, Ellison et al. (2022) find that star formation can rapidly decrease in mergers, but only after the merging galaxies have coalesced. No such quenching of star formation was present in close pairs. They use a neural network to detect mergers in galaxies in SDSS (York et al., 2000) using imaging from CFIS Bickley et al. (2022). They further visually check the 2000 mergers identified by the neural networks, after which they are left with 699 post-mergers. Wilkinson et al. (2022) get a similar result, finding that mergers play a significant role in quenching star formation. They identify mergers via morphology metrics, a neural network, and visual classification. Scudder et al. (2012) find a significant increase in star formation activity in close pairs of 30-60% compared to their control sample. They select galaxy pairs from SDSS imagery (York et al., 2000). Ultimately they have a pair sample of 1899 galaxies.

1.2 CIGALE

To study the relation between mergers and starbursts we use several quantities: the SFR, stellar mass (M_*) in $M_\odot$, and redshift of each galaxy, the functional form of the galaxy Main Sequence (MS), and merger classifications for the galaxies. The SFR and M_* were derived using the Code Investigating GALaxy Emission (CIGALE, Boquien et al. 2019) for Spectral Energy Distribution (SED) fitting (La Marca et al., 2024). A SED is a plot of the spectral energy density received from a source as a function of wavelength or frequency. The exact inner workings of the code are better described in Boquien et al. (2019), but its purpose is to fit a user defined model of a galactic SED to several points obtained from the integrated fluxes through several wavelength bands. It is important to note that having more wavelength bands leads to a higher accuracy of the results, because the code has to distinguish between degenerate parameters, e.g. the age, metallicity and dust attenuation (Boquien et al., 2019). For this thesis our primary interest lies with the SFR and M_* derived from CIGALE.

1.3 Galaxy Main Sequence

The galaxy MS is a relation between the SFR of star-forming galaxies and M_* which has been well established at redshifts of $0 < z < 6$, most recently by Popesso et al. (2023). Much more work on it has been done by others, this includes Leslie et al. (2020); Speagle et al. (2014), and the many references within these and Popesso et al. (2023). The MS is mainly dependent on M_* and the redshift, it usually exhibits a turn-over in SFR at higher M_* (around $10^{9.5}\,M_\odot$) (Leslie et al., 2020; Popesso et al., 2023). The SFR-M_* diagrams for our data are shown in Sec. 2 and the MS is shown in Sec. 3.2.

1.4 Description of this Thesis

In this thesis we use two catalogues, one is the KiDS-N-W2/GAMA09 catalogue from de Jong et al. (2013); Kuijken et al. (2019); Driver et al. (2011), which was cross-matched and expanded upon by La Marca et al. (2024). The other is the COSMOS2020 catalogue from Weaver et al.

(2022), which was expanded upon in Wang et al. (2024). Wang et al. (2024) added infrared data from the PACS and SPIRE instruments of Herschel. We refer to the data La Marca et al. (2024) as the KiDS sample and from Wang et al. (2024) as the COSMOS sample. More information on these samples is available in Sec. 2 and, of course, La Marca et al. (2024); Wang et al. (2024). We investigate the relation between the SFR of merging galaxies and non-merging galaxies. We do this by comparing the merger fraction as a function of redshift to the distance to the MS. We also compare the SFR of mergers and a control sample of non-mergers. Mergers are identified using the classifications by La Marca et al. (2024) and via the close pair method. This thesis is structured as follows: in Sec. 2 we describe the data used for our analysis, in Sec. 3 we describe the methods used to determine the galaxy MS and close pairs, in Sec. 4 we describe the results and discuss their implications and limitations, and in Sec. 5 we summarise our findings. This thesis also includes additional figures in Appendix A. We assume $\Omega_M = 0.2865$, $\Omega_\Lambda = 0.7135$ and $H_0 = 69.32$ km s^{-1} Mpc^{-1} (La Marca et al., 2024), a flat universe in the ΛCDM model.

2 Data

2.1 KiDS Sample

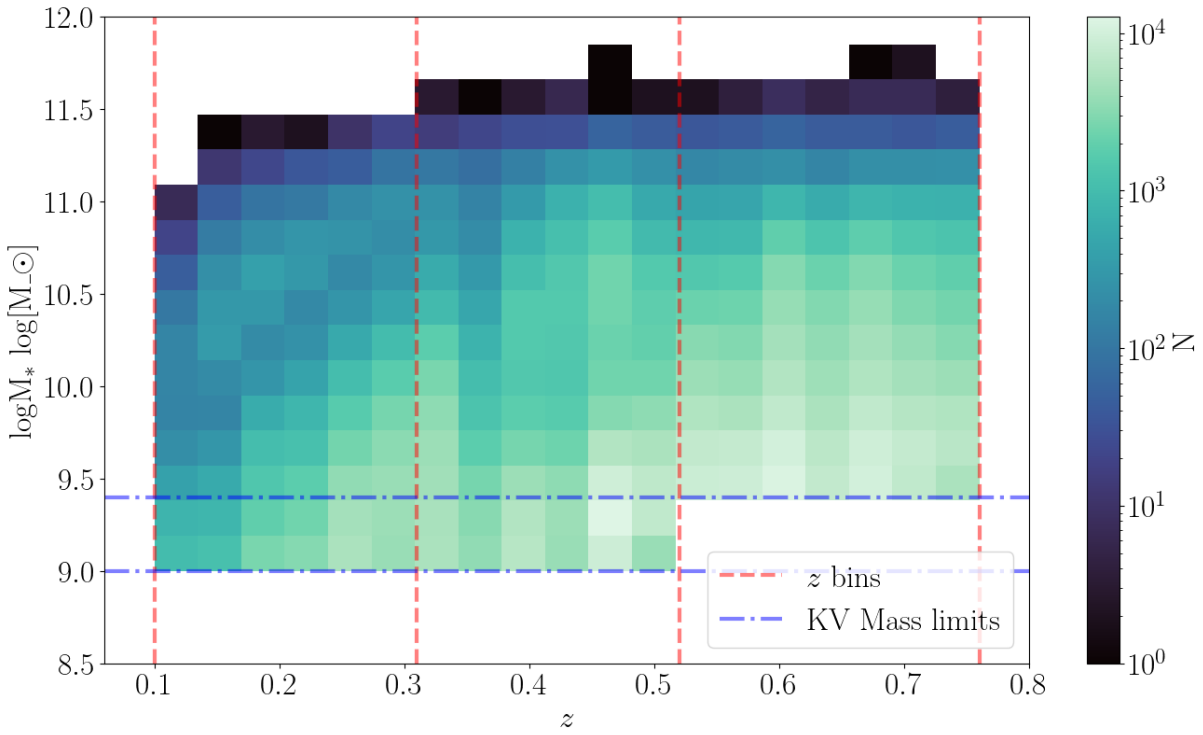

Figure 1: Plot of the redshift against $\log M_*$ in $\log[M_\odot]$ for the KiDS sample. The blue dashdotted line indicates the KiDS-VIKING mass limits applied by La Marca et al. (2024), the red dashed lines indicate the redshift bins: [0.1, 0.31), [0.31, 0.52), and [0.52, 0.76). The brightness of each bin indicates the number of galaxies in the bin, as is indicated on the colourbar.

To analyse the connection between merging galaxies and starburst galaxies we use the catalogue compiled by La Marca et al. (2024), referred to as the KiDS sample. It has a redshift range of $0.1 \leq z \leq 0.76$. The galaxies in the sample are selected from the Kilo-Degree Survey (KiDS, de Jong et al. 2013), specifically the KiDS-N-W2 field (Kuijken et al., 2019). This is supplemented with data from the Galaxy And Mass Assembly survey (GAMA, Driver et al. 2011), the VISTA Kilo-degree INfrared Galaxy survey (KiDS-VIKING, de Jong et al. 2013, Edge et al. 2013), Hyper-Suprime-Cam Subaru Strategic Program (HSC-SSP, Aihara et al. 2018), the latest data release of the NASA Wide-field Infrared Survey Explorer (WISE, Wright et al. 2010), AllWISE (Cutri et al., 2013), the Herschel Astrophysical Terahertz Large Area Survey (H-ATLAS, Valiante et al. 2016), the extended ROentgen Survey with an Imaging Telescope Array (eROSITA, Predehl et al. 2021), and from the eROSITA Final Equatorial Depth Survey (eFEDS, Brunner et al. 2022). The KiDS-VIKING and HSC-SSP are matched and combined, excluding any sources that are not part of both datasets (La Marca et al., 2024). The HSC-SSP images were used by La Marca et al. (2024) to detect mergers. The KiDS-VIKING photometric redshifts (photo-z) were also used. The photo-z was replaced by the spectroscopic redshift if that was available. Spectroscopic redshifts are redshifts determined from the spectrum of a galaxy, which leads to a low uncertainty in the calculated redshift (Newman and Gruen, 2022). Photo-z are determined from the integrated fluxes of a galaxy in a set of filters, a greater number of and more narrow filters increases the accuracy of the measurement (Newman and Gruen, 2022). In the best cases

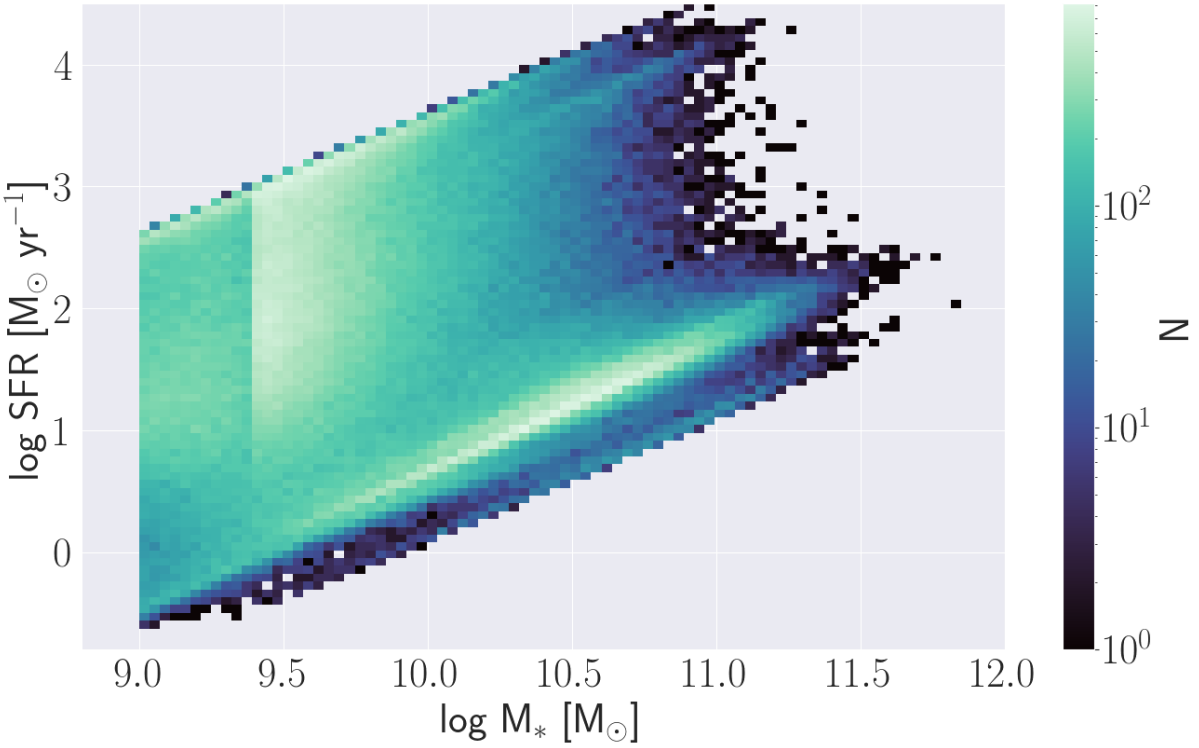

Figure 2: SFR-M_* diagram of the stellar mass limited KiDS sample. The break in the distribution at $M_* = 2.5 \cdot 10^9\,M_\odot$ is the stellar mass limit of z-bin 3 ([0.52, 0.76)), seen in Fig. 1. The brightness of each bin indicates the number of galaxies in the bin, as is seen on the colourbar.

the wavelength resolution of spectroscopic surveys is around twice as high as photometric surveys. In the worst case, when the bands used for photo-z's are few and broad, the resolution ($\lambda/\Delta\lambda$) can be as high as 30,000 (Newman and Gruen, 2022). So, even in the best case spectroscopic redshifts provide a more accurate measurement of the redshift than photo-z. For large datasets photo-z have to be used, because measuring spectroscopic redshifts is much more time consuming (Newman and Gruen, 2022). The photo-z in the KiDS sample are from eFEDS, HSC, and KiDS-VIKING. The final KiDS sample contains 24 wavebands covering most of the wavelength range of 300 nm to 800 μm, excluding the eROSITA x-ray waveband (La Marca et al., 2024). The flux in these wavebands was used in CIGALE SED fitting (Boquien et al., 2019) to determine properties of the galaxies in the survey, most importantly for this thesis the SFR and M_*. In order to identify the galaxies in HSC-SSP (Aihara et al., 2018) images as mergers La Marca et al. (2024) uses a convolutional neural network (Lecun et al., 1998) on the subsample of galaxies that have a fitted AGN component. The model is trained on the IllustrisTNG (Pillepich et al., 2018) and Horizon-AGN (Dubois et al., 2014) hydro-dynamical simulations (La Marca et al., 2024), which were processed to have similar visual artefacts and noise to real data. For more information on the training see La Marca et al. (2024).

The full KiDS sample is shown in Fig. 2 as a SFR-M_* diagram. This diagram compares M_* in $\log M_\odot$ to the SFR in $\log M_\odot \mathrm{yr}^{-1}$. The galaxy MS is visible as a bright streak across the plot. The KiDS sample covers an area in the sky of 65 deg^2, the sample includes 519403 galaxies in the stellar mass limited sample and 42881 in the subsample for which a merger classification was given.

The KiDS sample was divided into three redshift bins defined as follows: [0.1, 0.31), [0.31, 0.52) and [0.52, 0.76), following the same convention as in La Marca et al. (2024). These redshift

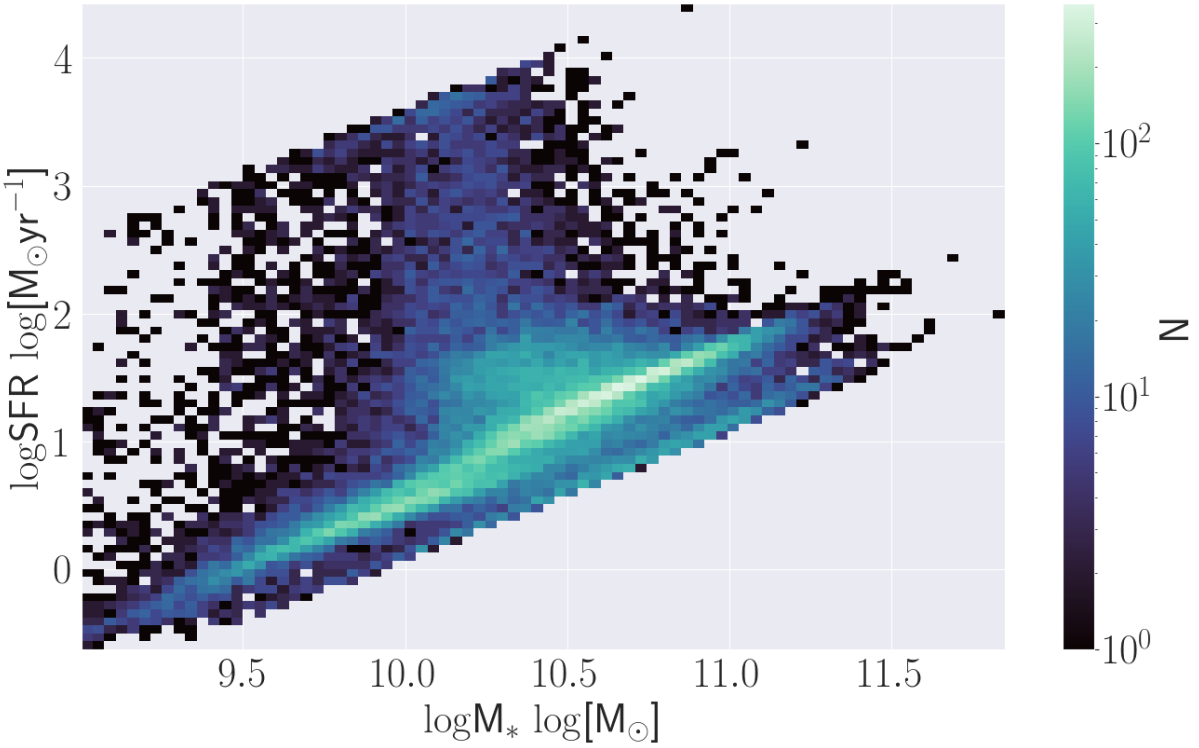

Figure 3: SFR-M_* diagram limited to the sample of galaxies that have a fitted AGN component (La Marca et al., 2024). These galaxies were identified as merger, non-merger or indeterminate. The galaxy MS is clearly visible as a bright streak across the plot.

bins contain 69140, 183554 and 266709 galaxies respectively in the stellar mass limited KiDS sample and 9159, 10614 and 23108 galaxies in the subsample with merger classifications.

We preferentially use the Bayes results over the best fit results of CIGALE, because the former are more reliable (La Marca et al., 2024).

2.2 COSMOS Sample

The COSMOS sample is constructed from observations in 40 (Weaver et al., 2022) + 5 (Wang et al., 2024) bands. The 40 bands are part of the COSMOS2020 catalogue (Weaver et al., 2022). The COSMOS2020 bands covered a wavelength range of $\approx 0.2\,\mu$m to $10\,\mu$m and were also used to determine the SFR and M_* of galaxies in the survey using CIGALE SED fitting (Boquien et al., 2019). The 6 bands from Wang et al. (2024) cover the $24-850\mu$m regime. The first 40 bands are from the COSMOS GALaxy Evolution eXplorer (GALEX) catalogue compiled by Zamojski et al. (2007), the MegaCam instrument of the Canada-France-Hawaii Telescope (CFHT, Sawicki et al. 2019), the Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS, Leauthaud et al. 2007), HSC-SSP (Aihara et al., 2018), the older Suprime-Cam instrument from SSP (Taniguchi et al., 2007, 2015), DR4 of the UltraVISTA survey (McCracken et al., 2012), and the Cosmic Dawn Survey using Spitzer (Euclid Collaboration et al., 2022). The bands added by Wang et al. (2024) are from the PACS and SPIRE instruments of the Herschel space telescope. They used the COSMOS2020 catalogue as a list of possible sources that could have detectable far-IR emission. They trained a machine learning model on CIGALE SED fits to predict the far-IR fluxes of the sources in the aforementioned list. This is used to select sources from this list to measure the flux of at five infrared wavelengths, these fluxes are added to the COSMOS sample (Wang et al.,

2024).

The area of the sky covered by the COSMOS sample from Wang et al. (2024) is $1.278\,\mathrm{deg}^2$.

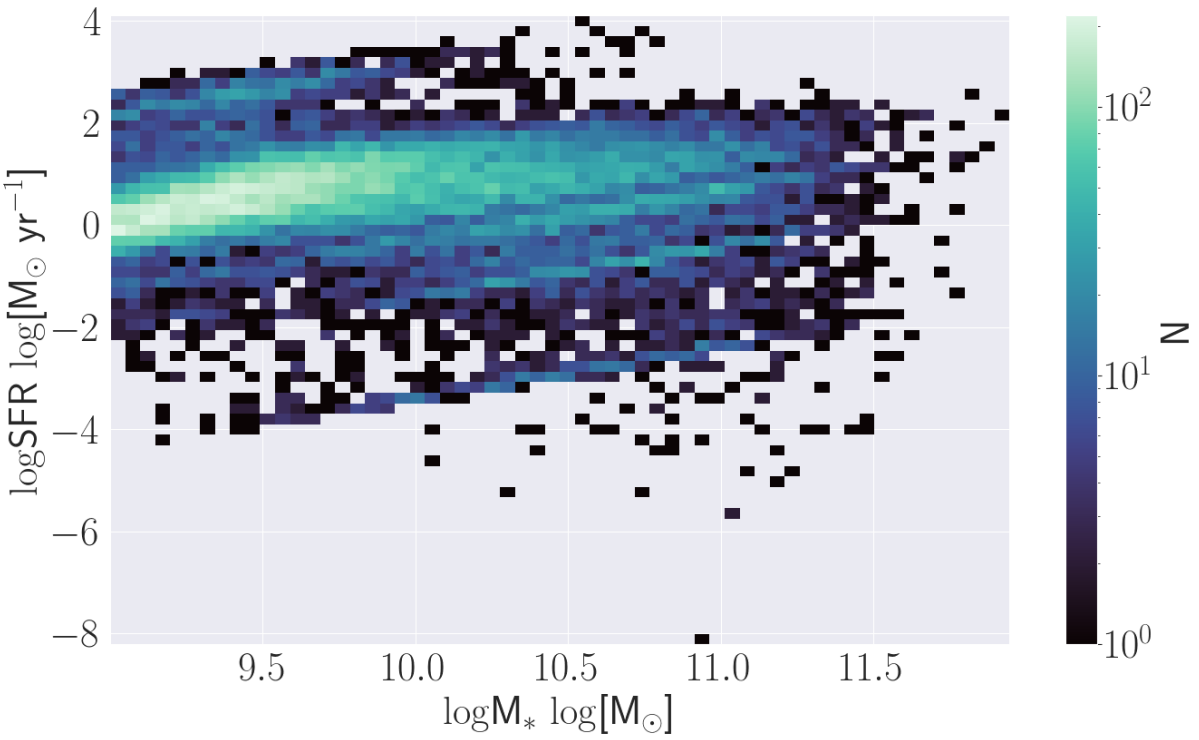

Figure 4: SFR-M_* diagram for the COSMOS sample.

The COSMOS sample contains 26075 galaxies. It was divided into four redshift bins as follows: [0.5, 1.0), [1.0, 1.5), [1.5, 2.0) and [2.0, 3.0). These redshift bins contain 6868, 7909, 5774 and 5524 galaxies respectively.

We preferentially use the Bayes results over the best fit results of CIGALE, because the former are more reliable (La Marca et al., 2024).

3 Methods

3.1 Selection of Close Pairs

In order to provide a different merger classification we selected close pairs of galaxies based on two quantities. Firstly, the projected separation, r_p, between the galaxies, which is calculated as the distance between galaxies that are projected onto the same sphere. And the difference in radial velocity, derived from their redshift. These two quantities combined provide an estimate of how close two galaxies are in 3d space. Two galaxies are identified as a close pair if these quantities satisfy the following conditions (Ventou et al., 2019; Mundy et al., 2017; Desmons et al., 2023; Duncan et al., 2019):

1. $r_p < 50\,h^{-1}$ kpc
2. $\Delta v < 500\,\mathrm{km\,s^{-1}}$

The threshold for r_p is 72.13 kpc using the value of h given in Sec. 1, Δv is calculated from the redshift of the two galaxies using

$$v = \frac{cz}{1+z} \tag{1}$$

(Ventou et al., 2019), and r_p is the projected separation given by

$$r_p = d_A(z)\Delta\theta \tag{2}$$

(Mundy et al., 2017), where $\Delta\theta$ is the angular separation between two galaxies, given by

$$\Delta\theta = \sqrt{(\Delta\alpha \cdot \cos(\delta))^2 + \Delta\delta^2}, \tag{3}$$

and $d_A(z)$ is the angular diameter distance, given by

$$d_A(z) = \frac{v}{H_0(1+z)}. \tag{4}$$

We will also compare how our results change when we apply a mass limit for each galaxy pair of $1/4 < M_{*,1}/M_{*,2} < 4$, similar to mass ratio limits in Violino et al. (2018); Scudder et al. (2012).

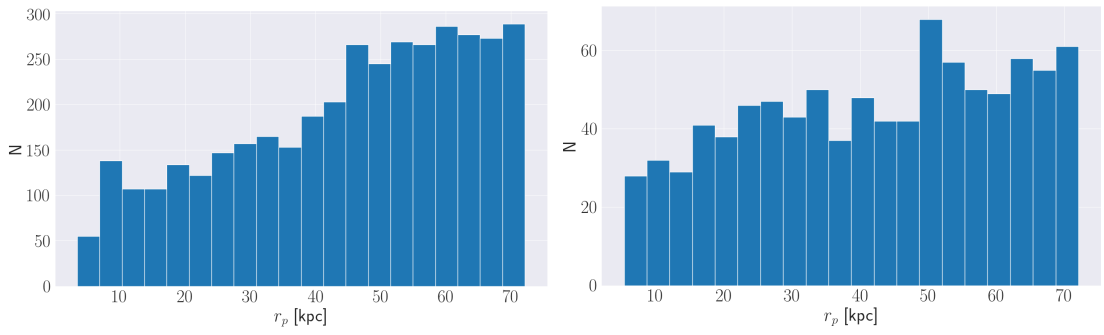

Figure 5: Histogram of the projected separation, r_p, of paired galaxies in the KiDS sample on the left and COSMOS sample on the right. It shows an increase in number of paired galaxies as the separation increases.

The selection criteria led to 11023 galaxies in close pairs or triplets in the KiDS sample. This gives a close pair fraction of $f_{pair} = 0.021$, which is lower than the fraction of 0.09 ± 0.01 found by Desmons et al. (2023) in the GAMA G02 field. It should be noted that they found their

fraction in a lower redshift range of $0.04 \leq z \leq 0.20$, which includes only a small part of our redshift range of $0.1 \leq z \leq 0.76$. Due to our higher redshift range it is possible that similar galaxies that are detected by Desmons et al. (2023) fall below the magnitude limits in the KiDS sample, due to their higher redshift. This decreases the number density of galaxies in our field and in turn decreases the expected number of galaxy pairs. Fig. 5 shows a histogram of the projected separation, r_p, given the two criteria. It shows a clear increase in the number of paired galaxies at larger pair separations. It also shows that at very low separations the number of pairs decreases drastically. We use 20 bins for each histogram, because we use the same amount to analyse the relationship between the distance to the MS and r_p in Sec. 4.1.

Desmons et al. (2023) use a sample of 2327 galaxies to find close pairs and find 250 paired galaxies. The size of their field is $2.38\,\mathrm{deg}^2$, which is much smaller than the $65\,\mathrm{deg}^2$ of the KiDS-N-W2 field used in this thesis. Due to their lower redshift range of $0.04 \leq z \leq 0.20$ it is possible that their higher pair fraction is caused by a more complete sample, because low mass galaxies may fall within the magnitude limits of GAMA and HSC-SSP (Aihara et al., 2018; Desmons et al., 2023) at their lower redshifts.

In the COSMOS sample we found 2889 paired galaxies, leading to a close pair fraction of 0.11. Ferreira et al. (2020) find a 'before merger' fraction, which is comparable to our close pair fraction, of $0.041 - 0.181$ in the same redshift range as our data. They use a lower distance criterion of $20\,h^{-1}\,\mathrm{kpc}$. They also use a mass limit of $M_* > 10^{10}\,\mathrm{M}_\odot$, compared to the mass limit of $M_* > 10^9\,\mathrm{M}_\odot$ as seen in Fig. 4. Man et al. (2012) find a pair fraction of $0.08 - 0.15$ in a redshift range of $0.0 - 3.0$ with a distance criterion of 30 kpc and the additional requirement that the mass ratio of the pair is $1/4 \leq M_1/M_2 \leq 4$. Overall the pair fractions found in the literature are comparable to ours. Fig. 5 shows the histogram of the pair separations. The number of galaxies at a lower separation decreases as expected.

When we include the mass ratio requirement of $1/4 < M_{*,1}/M_{*,2} < 4$ we find 7587 paired galaxies and a pair fraction of 0.015 in the KiDS sample. In the COSMOS sample we find 1706 paired galaxies and a pair fraction of 0.065.

Fig. 6 shows the pair fractions (f_{pair}) in both fields, with or without the mass ratio requirement, as a function of redshift. We see a peak in the second z-bin for both fields, but this is deceptive as the second z-bin of the KiDS field is $0.31 \leq z < 0.52$, while the second z-bin of the COSMOS field is $1.0 \leq z < 1.5$. The last z-bin of the COSMOS field has a relatively higher pair fraction when we apply the mass ratio criterion compared to previous bins. This can be explained by the detection of lower mass galaxies being less likely in this bin. This makes it more likely for a random pair of galaxies to satisfy the mass ratio criterion, due to the lower mass range in the sample. The pair fractions from Ferreira et al. (2020), who also use the same mass ratio criterion, deviate in the last two z-bins. The last z-bin shows a particularly large difference in f_{pair} between our results and theirs.

3.2 Selection of the Galaxy Main Sequence

There are several functional forms of the galaxy main sequence. Some include a turn-over around $\log M_* \approx 9.5$, e.g. in Popesso et al. (2023); Leslie et al. (2020), while others are linear with a time dependent slope and intercept, e.g. in Speagle et al. (2014). We compare

$$\log SFR(M_*, t) = (a_1 - a_2 t)\log M_* - (b_1 - b_2 t)\,, \tag{5}$$

$$\log SFR(M_*, t) = a_1 - a_2 t - \log\left(1 + \frac{10^{b_1 - b_2 t}}{10^{\log M_*}}\right), \tag{6}$$

and

$$\log SFR(M_*, t) = (a_1 - a_2 t)\log M_* - (b_1 - b_2 t) - c_1 \log^2 M_*, \tag{7}$$

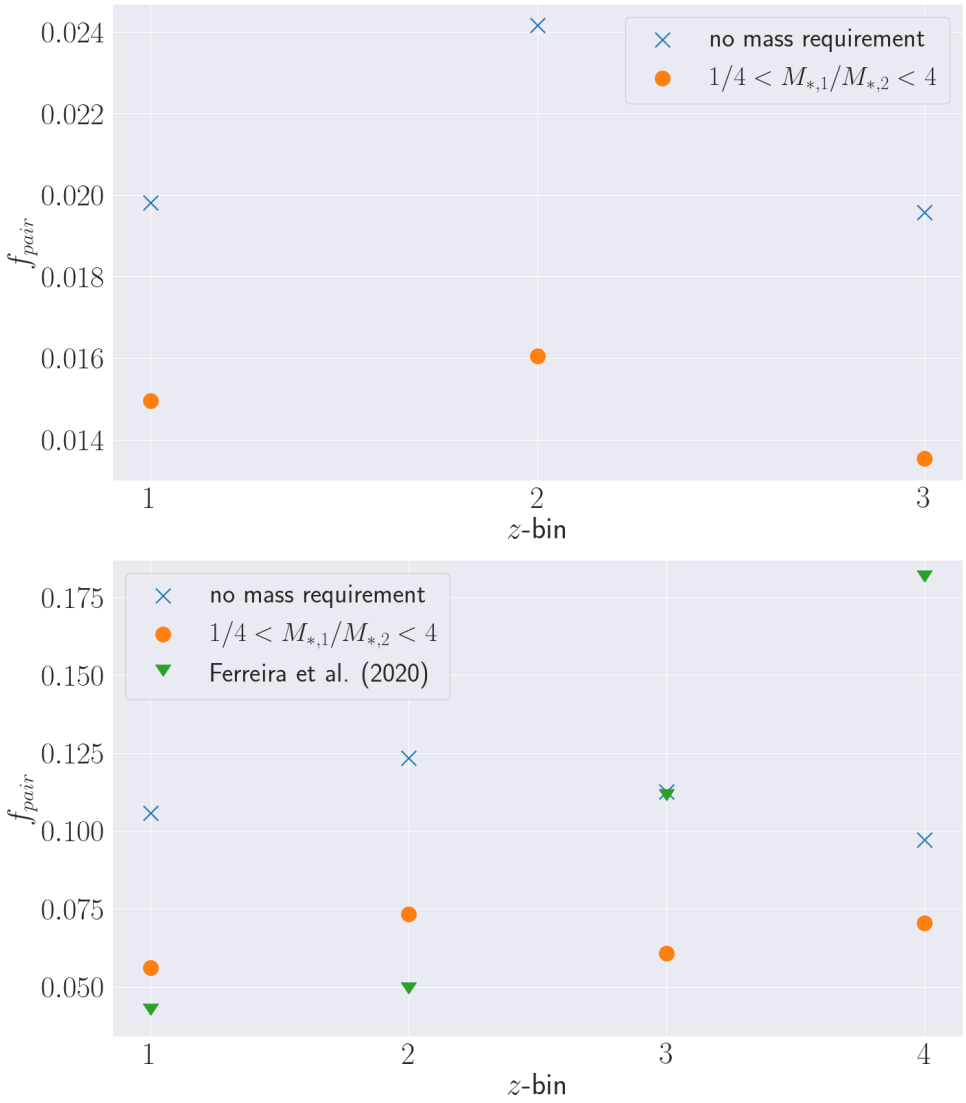

Figure 6: Comparison of the pair fractions as a function of redshift. The z-bins mentioned on the x-axis are defined in Sec. 2. The blue crosses are the results without the mass ratio requirement of $1/4 < M_{*,1}/M_{*,2} < 4$ outlined in Sec. 3.1, and the orange dots are with this requirement. *Top:* KiDS sample results. *Bottom:* COSMOS sample results, including the results from Ferreira et al. (2020) in green triangles.

	Eq. 5	Eq. 6	Eq. 7
Source	Speagle et al. (2014)	Leslie et al. (2020)	Popesso et al. (2023)
a_1	0.84 ± 0.02	$2.97^{+0.08}_{-0.09}$	4.722 ± 0.012
a_2	0.026 ± 0.003	0.22 ± 0.01	0.034 ± 0.002
b_1	6.51 ± 0.24	$11.06^{+0.15}_{-0.16}$	26.134 ± 0.0015
b_2	0.11 ± 0.03	$0.12^{+0.03}_{-0.02}$	0.20 ± 0.02
c_1			0.1925 ± 0.0011

Table 1: Table of the values of the parameters in the galaxy MS functions which were found in their respective sources. Asymmetric errors are included when given.

to see how each fits to the data. Eq 5 is from Speagle et al. (2014), Eq 6 from Leslie et al. (2020), and Eq 7 from Popesso et al. (2023). t is the age of the universe in Gyr, which is calculated as

$$t(z) = \frac{1}{H_0} \int_0^{1/(1+z)} \frac{\mathrm{d}a}{a\sqrt{\Omega_\Lambda + \Omega_M a^{-3}}}. \tag{8}$$

The other parameters, a_1, a_2, b_1, b_2, and c_1, are given in their respective papers, which we summarise in Table 1. Appendix C in Leslie et al. (2020) contains several other functional forms of the galaxy MS from the literature, including fitted parameters.

In addition to these functional forms with given parameters we also fit the MS ourselves with the functional form

$$\log SFR(M_*) = a_1 \log M_* - b_1 \tag{9}$$

for each of the redshift bins defined in Section 2. The minus sign is added for consistency with functional forms from the literature. We compare the values of a_1 and b_1 when fitting Eq. 9 in each redshift bin to evaluate whether fitting Eq. 5 would give a similar result for the quantities $a_1 - a_2 t$ and $b_1 - b_2 t$. This will tell us whether the MS as a function of redshift displays a similar trend in the KiDS sample and COSMOS sample as it does in the literature.

For the COSMOS sample we chose Eq. 7 from Popesso et al. (2023) as the functional form of the MS. Fig. 7 shows the location of every galaxy in the COSMOS sample on the SFR-M_* diagram if it were a MS galaxy using Eq. 7. Fig. 8 shows Eqs. 5, 6, and 7 on the SFR-M_* diagram. When compared to the SFR-M_* diagram on the top left it is clear that Eq. 5 does not fit well, because it does not have the turnover present in the data. Eq. 6 lies slightly above the MS when compared to Eq. 7, which is why we use the latter as the functional form of the MS in the COSMOS sample. Fig. 9 shows the SFR-M_* diagrams with Eq. 7 for each of the four redshift bins outlined in Sec. 2.

For the KiDS sample we compared Eqs. 5, 6, and 7 to the visually identified main sequence, this is shown in Fig. 10. As can be clearly seen, all MS functions deviate strongly from the MS, we will discuss potential causes of this mismatch and difference with the COSMOS sample later. For now we fit the power law in Eq. 9 through the MS in each of the three z-bins. Unlike Eqs. 5, 6, and 7, this is not dependent on the redshift of each galaxy. After fitting the equations we will compare their parameters to see if there is a redshift or time dependence.

In order to fit the galaxy MS we divided M_* into 14 equally sized bins with a width of 0.20 dex, this width was chosen based on the median error $\sigma_{\log M_*}$ of 0.10 dex. As can be seen in Fig. 2, there are many galaxies that lie above the main sequence and to ensure a good fit we use a cut to the specific SFR (sSFR). The sSFR is defined as $sSFR = SFR/M_*$, due to the linear MS in our data using the sSFR allows us to make the MS a horizontal line. This in turn allows for a cut to be applied to the data to remove galaxies that are clearly not on the MS. We chose to remove

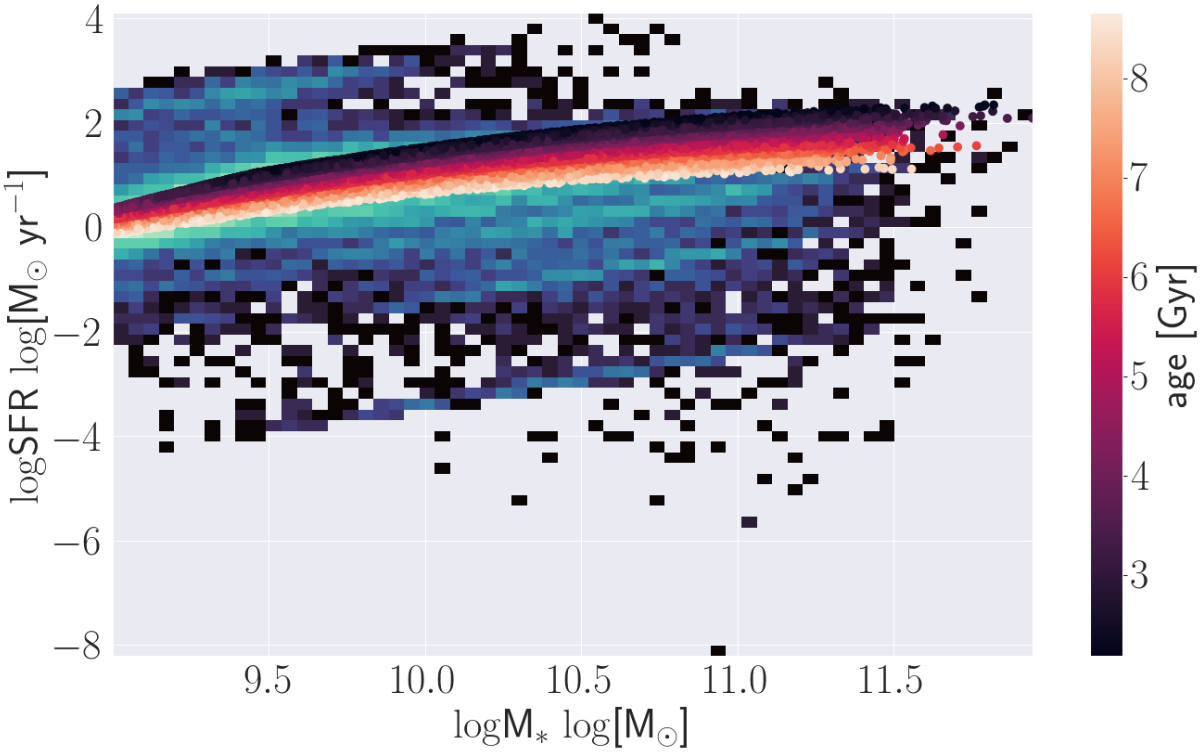

Figure 7: SFR-M$_*$ diagram with Eq. 7 plotted using the parameters from Table 1. The plot gives the location of every galaxy in the COSMOS sample if it were on the main sequence. The colour of the MS indicates the age of the universe at the redshift of each galaxy.

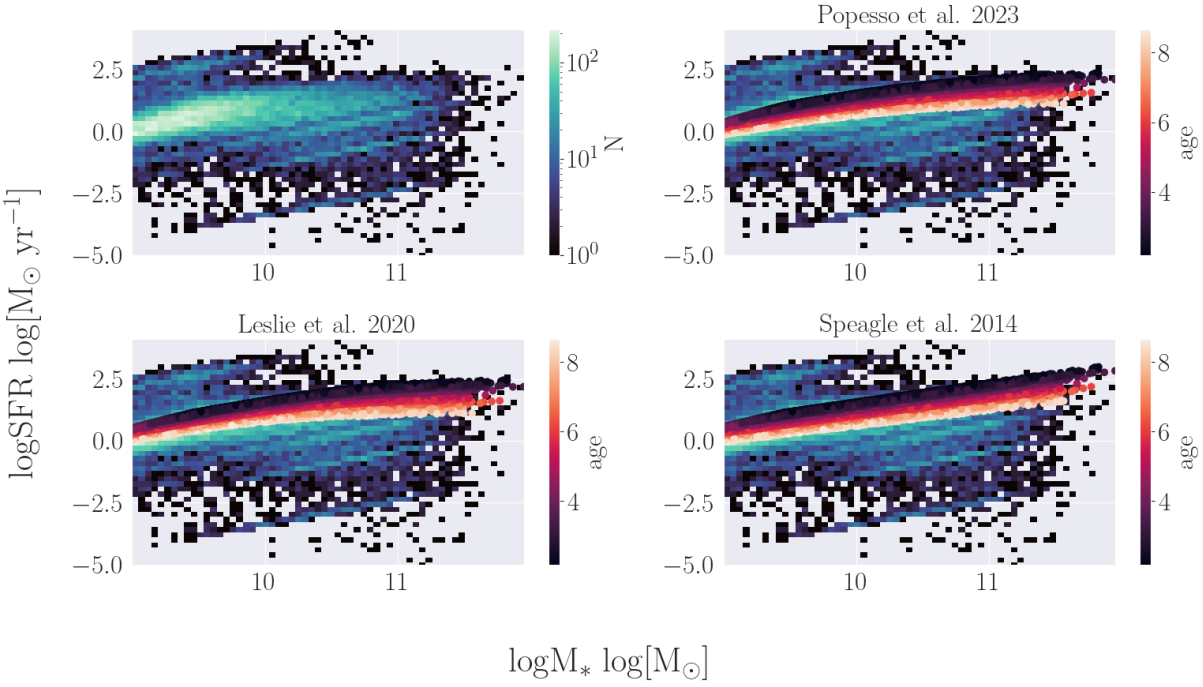

Figure 8: *Top-left:* SFR-M$_*$ diagram of the COSMOS sample. *Top-right:* Eq. 7 from Popesso et al. (2023). *Bottom-left:* Eq. 6 from Leslie et al. (2020). *Bottom-right:* Eq. 5 from Speagle et al. (2014).

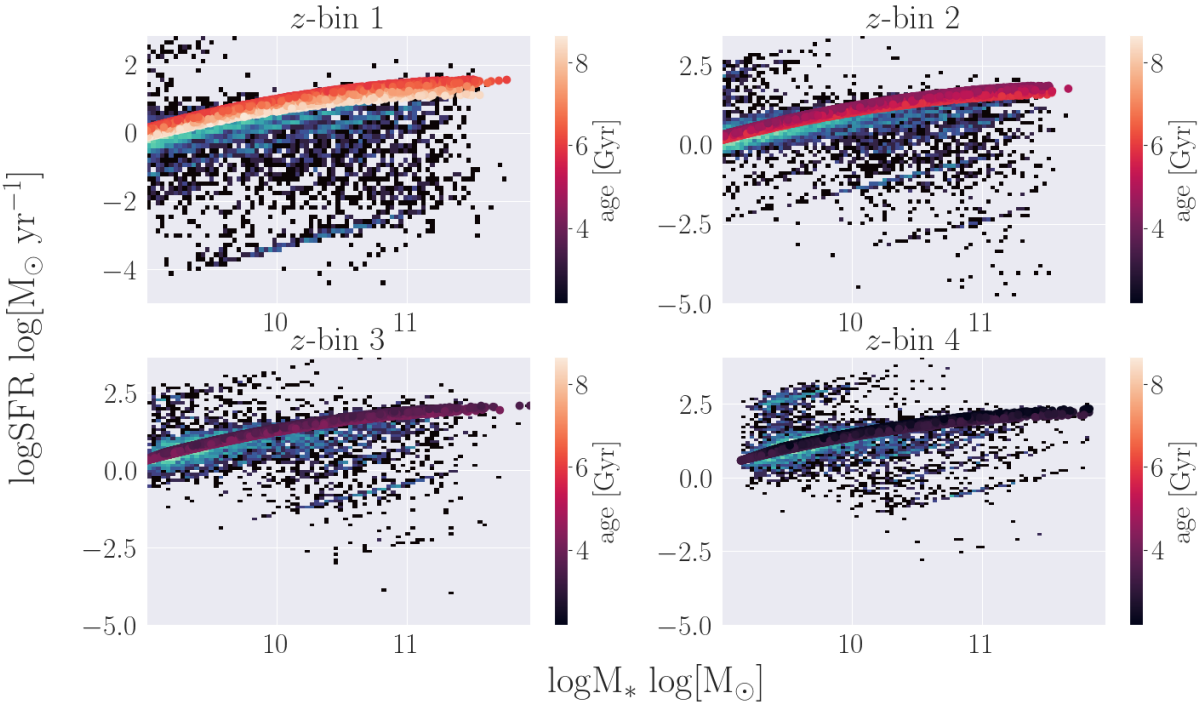

Figure 9: SFR-M$_*$ diagrams of the COSMOS sample throughout the four z-bins defined in Sec. 2 with the MS function from Popesso et al. (2023) (Eq. 7).

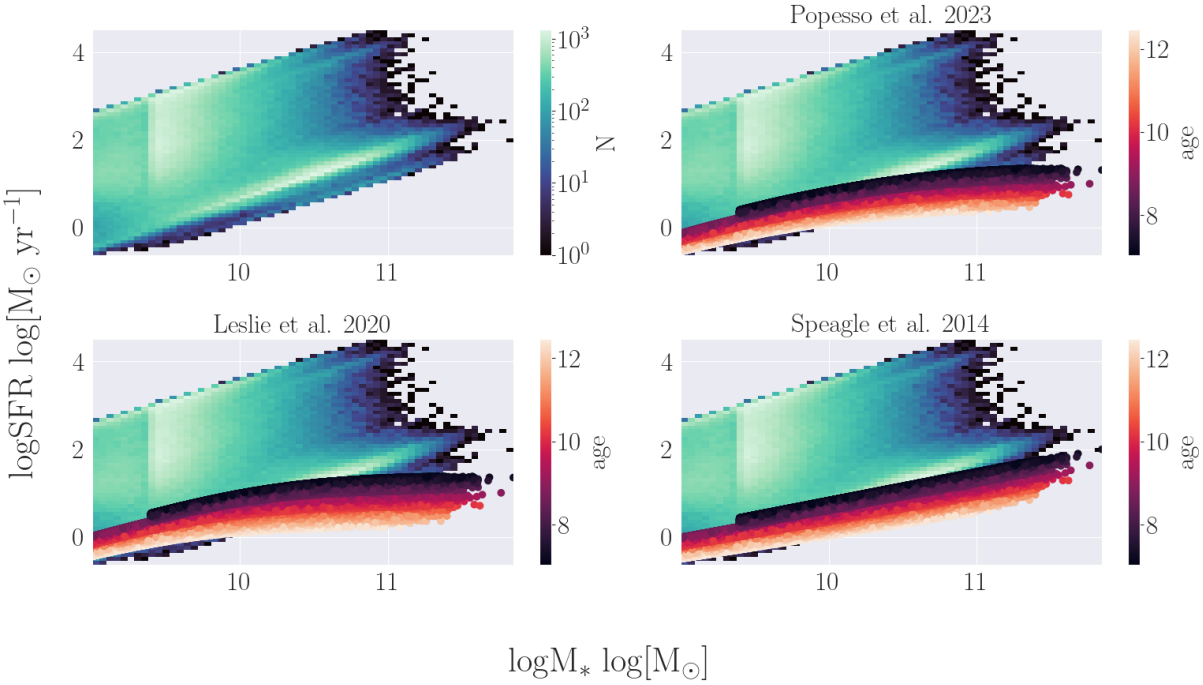

Figure 10: SFR-M$_*$ diagram showing the KiDS sample and Eqs. 5, 6, and 7, in the bottom left, bottom right, and top right respectively, with the parameters given in table 1. The colourbar for the first plot indicates the number of galaxies in each bin on a logarithmic scale. The other colourbars give the age of the universe in Gyr for each galaxy.

all galaxies with sSFR > -8.5, this value was chosen because it compromised between removing many high SFR galaxies, while keeping the galaxies that lie on the main sequence. Fig. 11 shows the number density of galaxies in a sSFR-M_* plot and the cut that was applied. The last two M_* bins were excluded from the fit, because the position of the MS could not be accurately determined with to the low number of galaxies in these bins. In each of the 12 remaining M_* bins the mode of the SFR distribution was determined, these distributions can be seen in Fig. 28, 29 and 30 in Appendix A. We use the standard deviation of these SFR distributions as the error in the SFR of the points used to fit the MS.

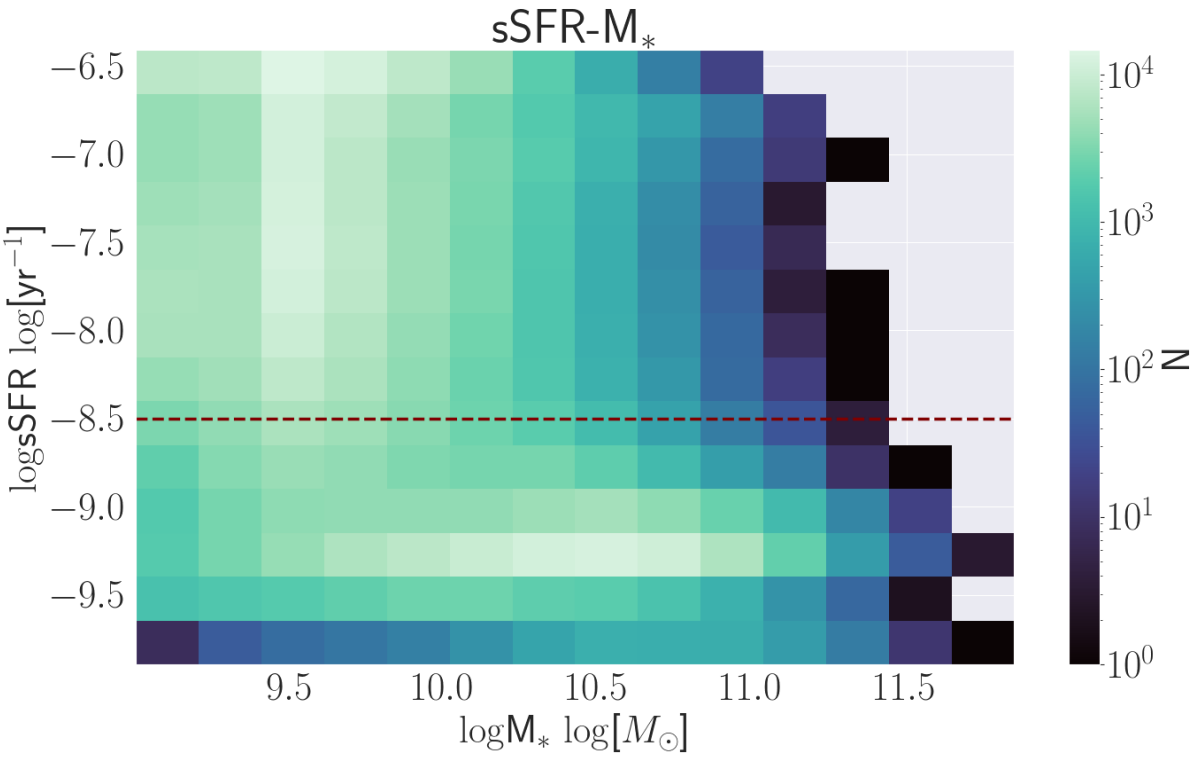

Figure 11: Plot of logsSFR against $\log M_*$ in $\log \mathrm{yr}^{-1}$ and $\log M_\odot$ respectively. The red dotted line indicates the cut applied to the sSFR. The brightness of the colour indicates the number of galaxies in each bin on a logarithmic scale, as can be seen on the colourbar. The size of the M_* bins was chosen based on the median size of the error in $\log M_*$

Fig. 12 shows the values of the parameters $a(t)$ and $b(t)$ in $\log SFR = a(t) \log M_* - b(t)$. We expect the parameters to have a linear relation with time from Speagle et al. (2014); Leslie et al. (2020); Popesso et al. (2023), see Eqs. 5, 6, and 7. However, Fig. 12 shows no linear correlation between the parameters and the age of the universe. We also find that $a(t)$ and $b(t)$ are systematically higher in our fit compared to the parameters of Eq. 5 from Speagle et al. (2014); Leslie et al. (2020).

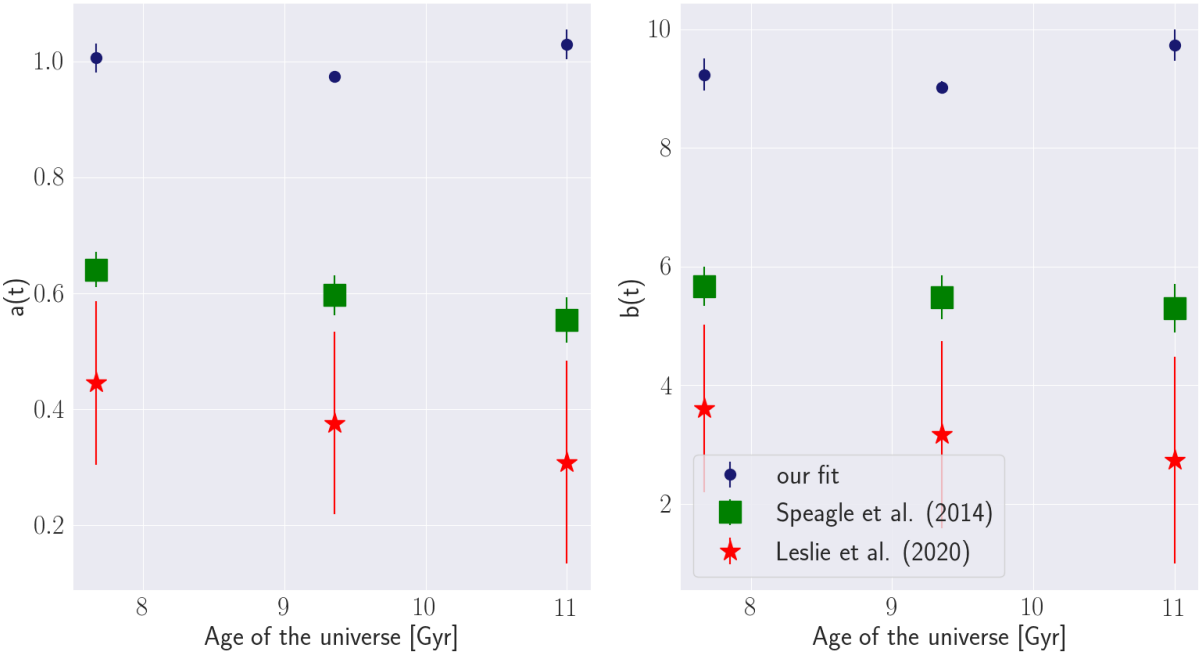

Figure 12: Plot of the values of the parameters $a(t)$ and $b(t)$ in $\log SFR = a(t) \log M_* - b(t)$. We also include the values of $a(t) = a_1 - a_2 t$ and $b(t) = b_1 - b_2 t$ from Speagle et al. (2014); Leslie et al. (2020). The parameters of Eq. 5 from Speagle et al. (2014) are shown as green squares. This equation was also fit by Leslie et al. (2020), the result of that fit is shown as red stars. The result of our fit is shown in the purple dots. While they both seem to follow a similar curve, there is no linear correlation and decrease with time as is found by Speagle et al. (2014); Leslie et al. (2020); Popesso et al. (2023).

The first clue to the difference in how well the MS functions from Speagle et al. (2014); Leslie et al. (2020); Popesso et al. (2023) fit to the data is the differing shapes of the SFR-M_* diagrams. We see two main reasons for this. The first is the difference in number of wavelength bands used in the COSMOS sample and KiDS sample. The latter does cover a larger wavelength range, as described in Sec. 2, but has about half the number of bands and the COSMOS sample includes many narrow bands. The other effect is the setup of CIGALE. Wang et al. (2024) use more parameters for the Star-Formation History (SFH) than La Marca et al. (2024), because La Marca et al. (2024) were more focused on AGN. For example, Wang et al. (2024) use an age of the main population of 500, $1000 - 13000$ (step 1000) Myrs, but La Marca et al. (2024) use $4000 - 13000$ (step 1000) Myrs. With CIGALE in Wang et al. (2024) using more parameters for the star-formation history and having more data available per galaxy this leads to a more physically reasonable SFR-M_* diagram in the COSMOS sample.

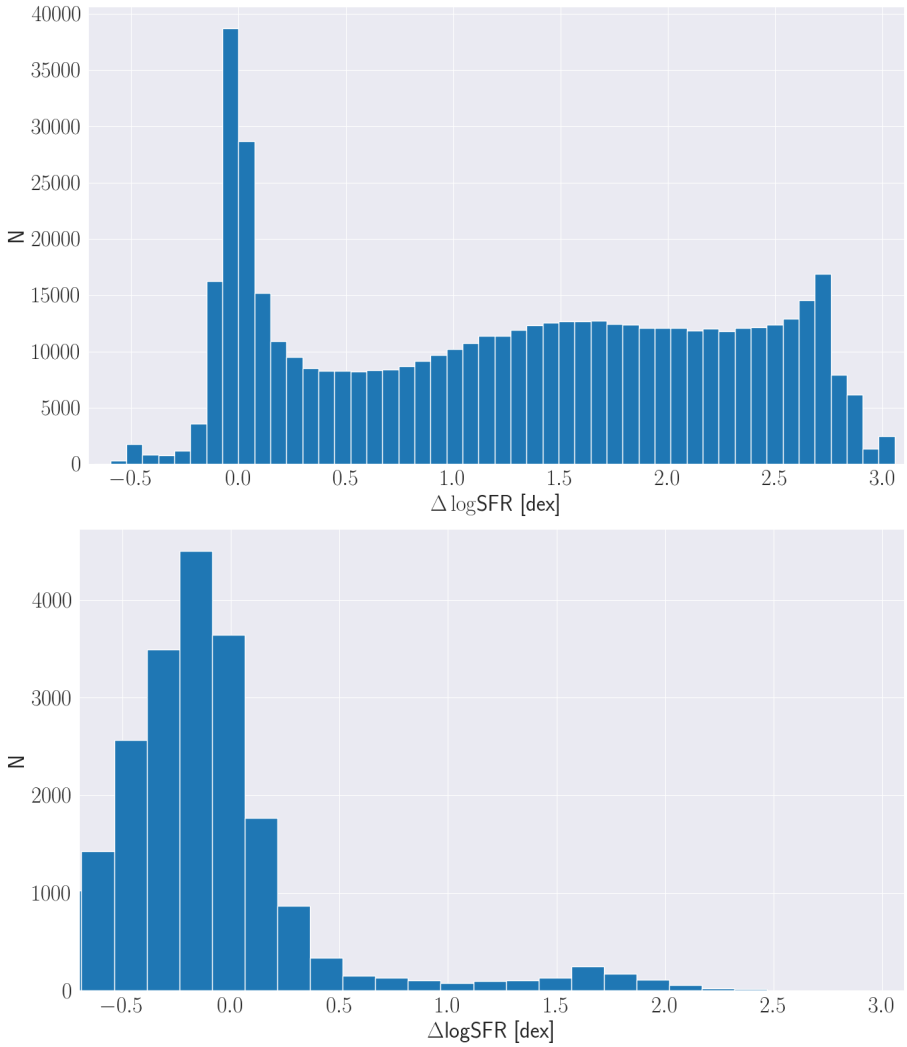

Figure 13: Histogram of the distance to the MS of paired galaxies in the KiDS sample on the top and COSMOS sample on the bottom. We restrict both plots to the range $-0.7 < \Delta \log SFR < 3.1$. We can see that the COSMOS sample has a relatively larger population of galaxies that lie below the MS, while the KiDS sample has a relatively larger population of galaxies that lie above the MS.

4 Results

4.1 Distance to Galaxy MS in Close Pairs

With the functional form of the MS in both datasets we calculate the distance from the main sequence as

$$\Delta \log SFR = \log SFR - \log SFR_{MS} = \log \frac{SFR}{SFR_{MS}}. \tag{10}$$

$\Delta \log SFR$ is the logarithm of a fraction and hence has the unit dex. The distribution of $\Delta \log SFR$ for both catalogues is shown in Fig. 13. By comparing the two it is clear that there are relatively more galaxies which lie above the main sequence in the KiDS sample. This is also clear from looking at their SFR-M$_*$ diagrams (Fig. 2 & 4 respectively). We restrict both plots to the range $-0.7 < \Delta \log SFR < 3.1$ in order to compare them. Fig 14 shows two histograms of the errors in $\Delta \log SFR$ for the KiDS and COSMOS samples. When including errors in later analysis we exclude any galaxy whose error is > 1 dex. For the COSMOS sample

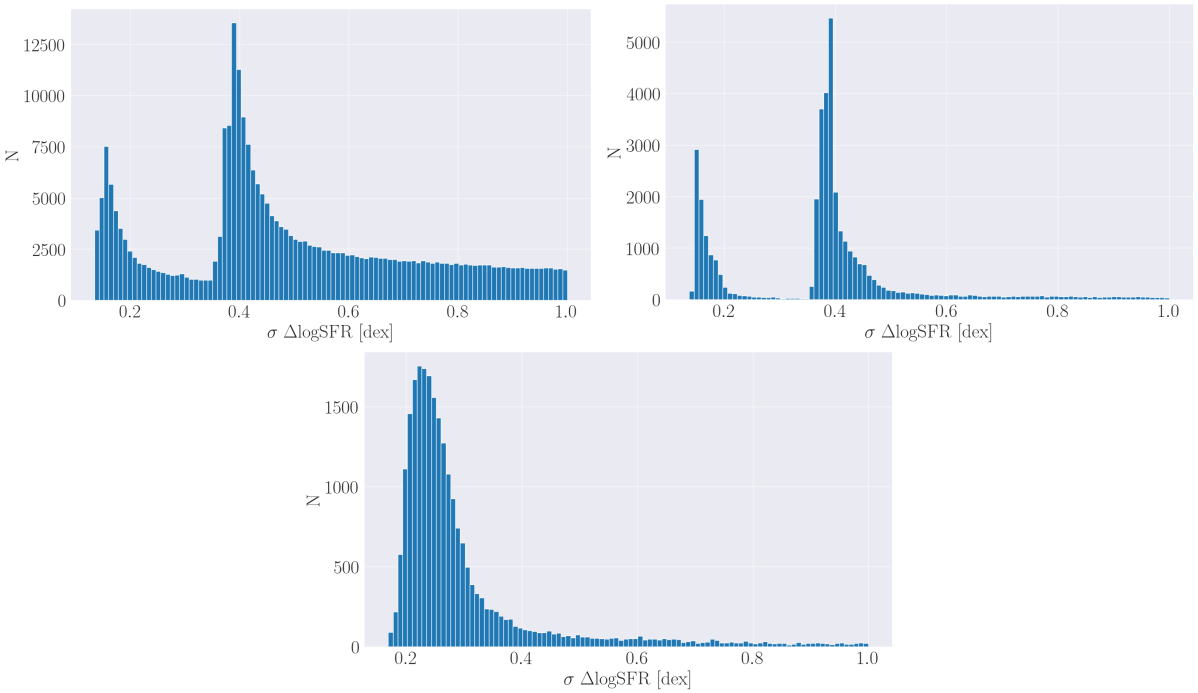

Figure 14: Histograms of the errors in the distance to the MS, $\Delta \log SFR$. We exclude any error above 1 dex. *Top:* KiDS sample on the left, with the merger subsample on the right. The two peaks are caused by the small errors in $a(t)$ and $b(t)$ in the second z-bin, as seen in Fig. 12. These smaller errors lead to a smaller error in the distance to the MS, so the first peak contains mostly galaxies from that z-bin and the other peak contains mostly galaxies from the other two z-bins. *Bottom:* COSMOS sample.

this only decreases the number of galaxies by $\approx 8\%$, but for the KiDS sample the number of galaxies is almost reduced by half. This shows that the COSMOS sample is more reliable than the KiDS sample. The errors in the KiDS sample and subsample show two peaks. The first peak is from the data in the second z-bin, where the errors on the parameters $a(t)$ and $b(t)$ are much smaller than in the first and third z-bin. This can be seen in Fig. 12. When these small errors are propagated we find a smaller error in $\Delta \log SFR$. The sample also shows a greater error in M_*, which can be seen from the taper. This taper is not present in the subsample, which shows that these have a lower error in M_* and also informs us that the subsample is not entirely representative of the whole KiDS sample.

Fig. 15 compares the distance to the MS to the pair fraction in each bin for both catalogues. It shows a decrease in the pair fraction as the distance to the MS increases in the KiDS sample. The COSMOS sample shows no correlation between the pair fraction and distance to the MS over the same $\Delta \log SFR$ range. Fig. 15 also shows the relation between the pair separation, r_p, in kpc and the distance to the MS in $\log M_* \, yr^{-1}$. The increase is positive for the KiDS sample, but negative for the COSMOS sample. As can be seen in Fig. 5, there are enough paired galaxies in each r_p bin. The erratic behaviour of both samples is not caused by the bin size, but instead shows the variation present in the data.

Fig. 16 is the result of adding the errors from Fig. 14 and removing galaxies for which $\sigma_{\Delta \log \mathrm{SFR}} > 1$. We include results with error separately, because we also exclude galaxies with $\sigma_{\Delta \log \mathrm{SFR}} > 1$. This has a significant enough effect on the results to discuss these separately. We find significant outliers when comparing the Δ logSFR of KiDS sample galaxies to the pair fraction in the Δ logSFR bins. These outliers cannot be explained by a small sample size, because each bin has plenty of galaxies as is seen in Fig. 13. The COSMOS sample has small sample

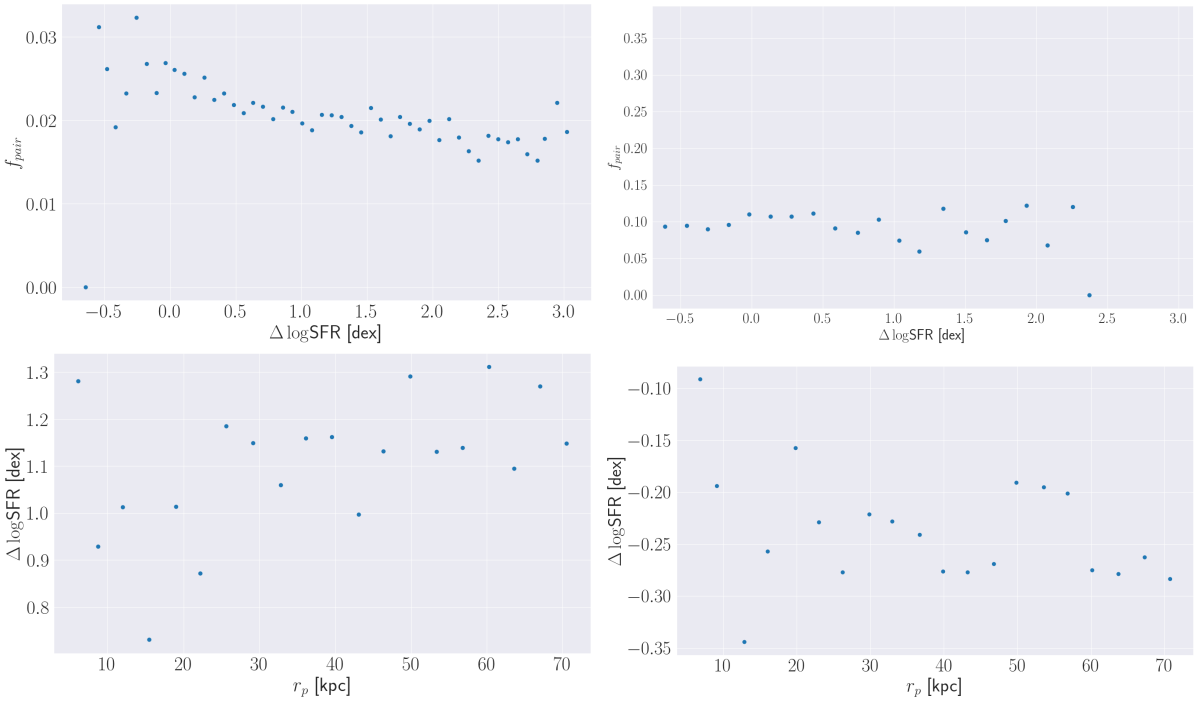

Figure 15: *Left:* Results for the KiDS sample. *Right:* Results for the COSMOS sample. *Top:* Comparison of the distance to the MS and the fraction of paired galaxies in each distance bin. Both datasets show a decrease in pair fraction as the distance to the MS increases. *Bottom:* Comparison of the pair separation and distance to the MS, which is determined as the median of $\Delta \log \mathrm{SFR}$ in each r_p bin.

size for Δ logSFR > 0.5 dex, it shows a slight increase at lower Δ logSFR, but no correlation for Δ logSFR > 0.5 dex. We also compare Δ logSFR as a function of projected separation, r_p. We find that there is an increase in SFR for high projected separation in the KiDS sample. Two bins have a Δ logSFR which is ≈ 1.2 dex lower than the bins to either side. The sample sizes of both bins are ≈ 200, so this likely doesn't cause the lower Δ logSFR. The COSMOS sample shows a very constant Δ logSFR of ≈ -0.2 dex. All points lie beneath the MS, which tells us that the star formation is slightly quenched in close pairs in the COSMOS sample.

To constrain our results we also applied a third criteria for our pairs, outlined in Sec. 3.1. To reiterate, we constrain our pairs to a mass ratio of $1/4 < M_{*,1}/M_{*,2} < 4$. The results with this constraint are shown in Fig. 17, in contrast to our previous results in Fig. 15. We find that the trend in the pair fraction is less pronounced here, the pair fractions at the lowest distance to MS can be explained by the low number of galaxies in these bins. We find that the trends in distance to the MS as a function of r_p are also less pronounced compared to the results without the mass ratio constraint.

Fig. 18 shows the same plots as Fig. 15, 16, and 17. Here we include both the mass ratio requirement and the error requirement of $\sigma < 1$ dex. We find that the distance to MS is constant as a function of r_p for both the KiDS and COSMOS sample. For the COSMOS sample the average distance to the MS varies around -0.2 dex, meaning that pairs in this sample mostly lie below the main sequence. We find no significant difference with Fig. 16. For the KiDS sample we find that the distance to MS is almost constant around 2 dex, but there are three significant outliers. Compared to Fig. 16 we find that the mass ratio requirement results in a more constant $\Delta \log \mathrm{SFR}$. In the top plots we compare the distance to the MS with the pair fraction in each $\Delta \log \mathrm{SFR}$ bin. We find that the pair fraction for both samples is slightly lower compared to

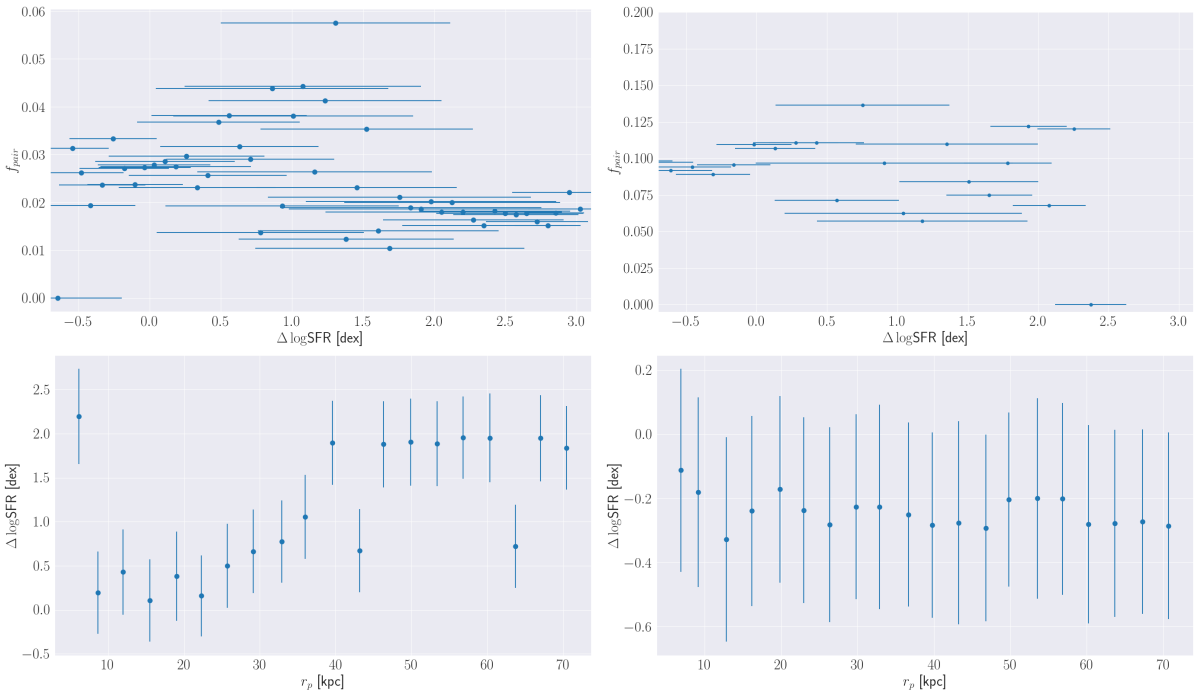

Figure 16: *Left:* Results for the KiDS sample. *Right:* Results for the COSMOS sample. *Top:* Comparison of the distance to the MS and the fraction of paired galaxies in each distance bin. The outliers in the KiDS sample cannot be explained by a small sample size, because those bins contain > 5000 galaxies each. The COSMOS results are most reliable up to a Δ logSFR of 0.5, after which the number of galaxies in each bin decreases drastically. The sample sizes per bin can be read from Fig. 13 for both the KiDS and COSMOS sample. *Bottom:* Comparison of the pair separation and distance to the MS, which is determined as the median of $\Delta \log \mathrm{SFR}$ in each r_p bin.

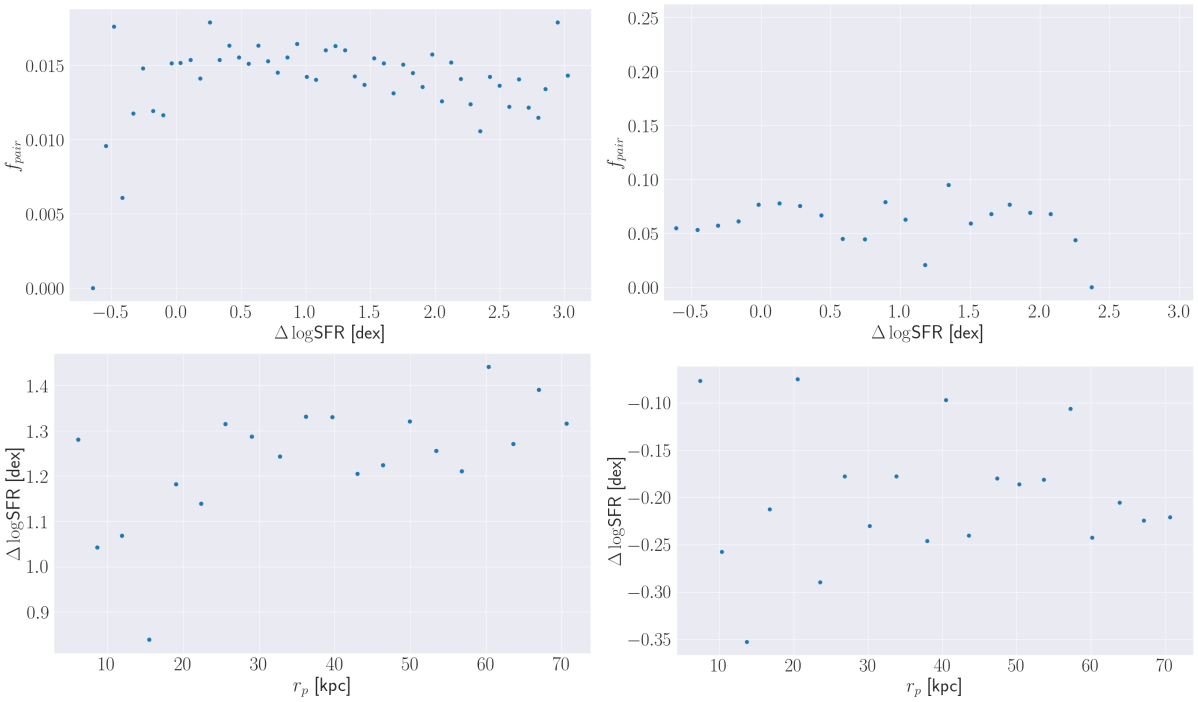

Figure 17: Result for mass ratio of $1/4 < M_{*,1}/M_{*,2} < 4$. *Left:* Results for the KiDS sample. *Right:* Results for the COSMOS sample. *Top:* Comparison of the distance to the MS expressed in $\log \mathrm{M}_* \,\mathrm{yr}^{-1}$ and the fraction of paired galaxies in each distance bin. *Bottom:* Comparison of the pair separation and distance to the MS, which is determined as the median of $\Delta \log \mathrm{SFR}$ in each r_p bin. The COSMOS sample shows no correlation. The KiDS sample shows a steeper increase compared to the results without a mass ratio constraint (Fig. 15).

Fig. 16. The same behaviour is seen as in Fig. 16, with a spread in the KiDS result around $\Delta \log \mathrm{SFR} \approx 0.5$ dex and a spread in the COSMOS result for $\Delta \log \mathrm{SFR} > 0.5$ dex. The latter is likely caused by a small sample size.

Overall we find that the distance to the MS is constant with pair separation when we apply both the mass ratio constraint ($1/4 < M_{*,1}/M_{*,2} < 4$) and $\Delta \log \mathrm{SFR}$ error constraint ($\sigma < 1$ dex). For the KiDS sample it is constant around $\Delta \log \mathrm{SFR} \approx 2.0$ dex, this sample does have three significant outliers at low pair separation. For the COSMOS sample this is $\Delta \log \mathrm{SFR} \approx 0.2$ dex. Compared to the literature we would expect the distance to the MS to decrease when r_p increases, because the gravitational perturbations that increase star formation decrease in strength (Violino et al., 2018).

When we compare the pair fraction to the distance to the MS we find that it remains roughly constant when we include the mass ratio and error constraints. For the KiDS sample we have significant outliers around $\Delta \log \mathrm{SFR} \approx 1.0$ dex. For the COSMOS sample there is significant scatter for $\Delta \log \mathrm{SFR} > 0.5$ dex, but this can be caused by a small sample size, as can be seen in Fig. 5.

4.2 Comparison of the SFR of Close Pairs and Control Galaxies

To compare the SFR of our close pairs to the entire sample we also created a control group. We create a control group for each paired galaxy that had to satisfy the following criteria:

$$|z_{control} - z_{sample}| \leq 0.05, \tag{11}$$

$$|\log M_{*,control} - M_{*,sample}| \leq 0.1 \,\mathrm{dex}. \tag{12}$$

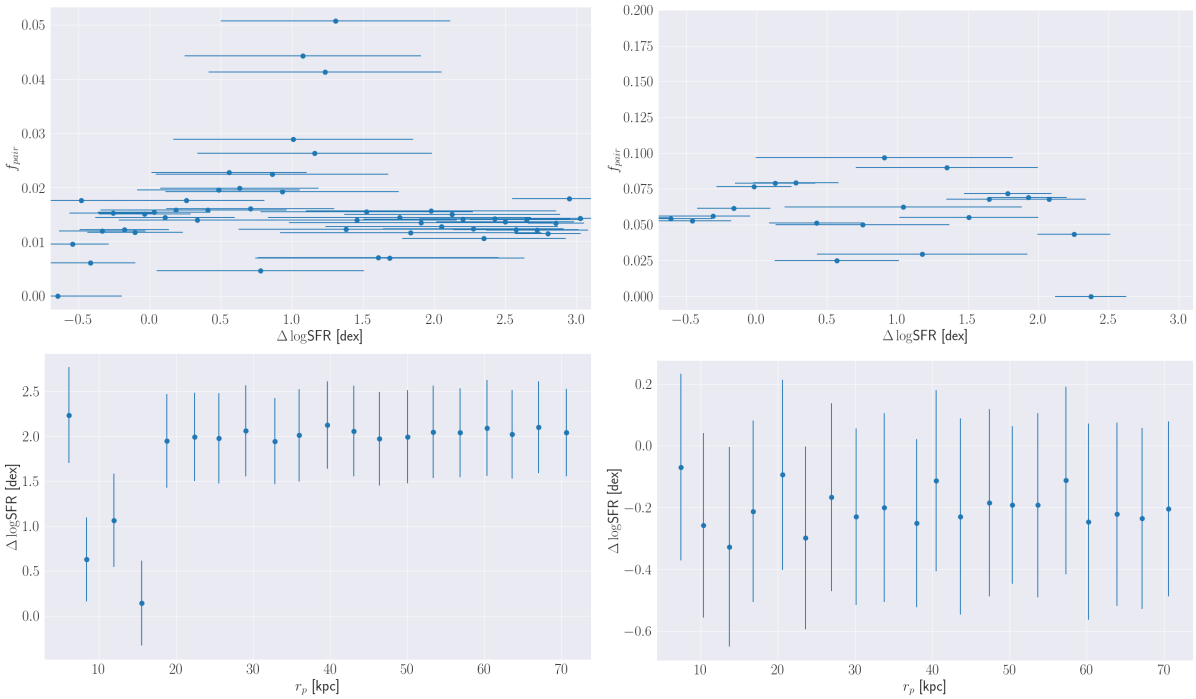

Figure 18: Results when we apply both the mass ratio requirement of $1/4 < M_{*,1}/M_{*,2} < 4$. *Left:* Results for the KiDS sample. *Right:* Results for the COSMOS sample. *Top:* Comparison of the distance to the MS expressed in $\log \mathrm{M}_* \, \mathrm{yr}^{-1}$ and the fraction of paired galaxies in each distance bin. The COSMOS sample shows a roughly constant pair fraction, with the noise above $\Delta \log \mathrm{SFR} \approx 0.5$ likely caused by a lower sample size. *Bottom:* Comparison of the pair separation and distance to the MS, which is determined as the median of $\Delta \log \mathrm{SFR}$ in each r_p bin.

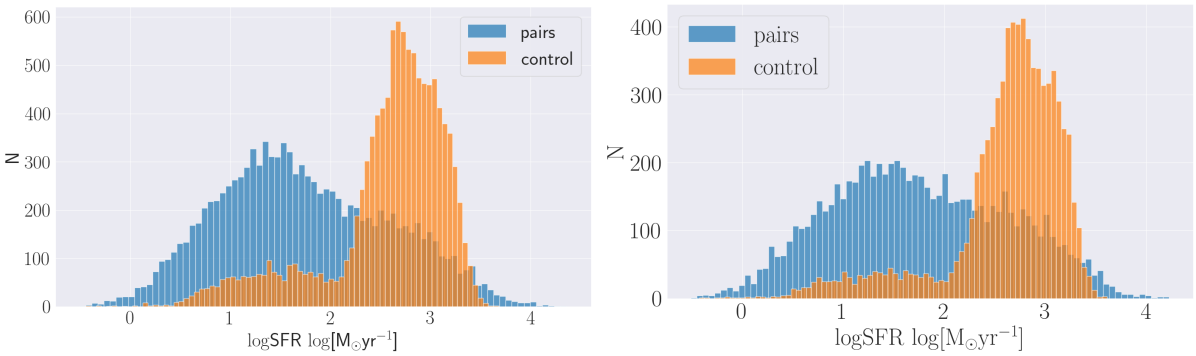

Figure 19: Histograms of logSFR in $\log \mathrm{M}_\odot \mathrm{yr}^{-1}$ for paired galaxies (blue) and their control sample (orange) in the KiDS sample. *Left:* Full sample. *Right:* Result with the mass ratio requirement of $1/4 < M_{*,1}/M_{*,2} < 4$.

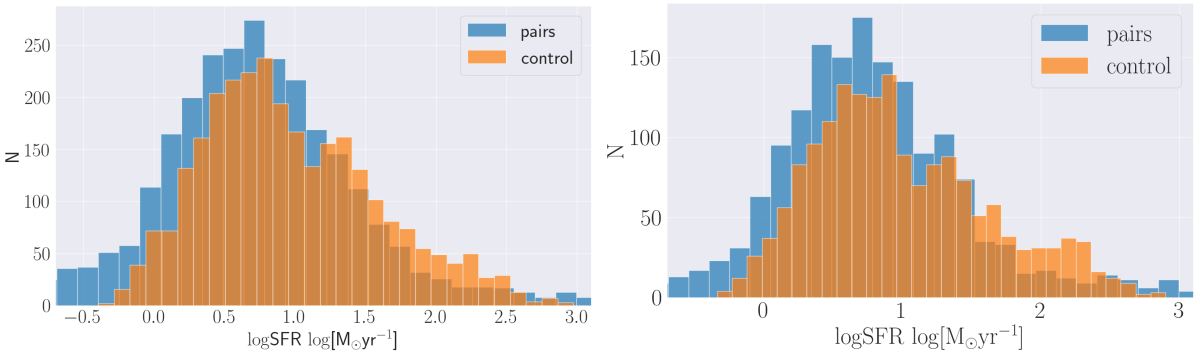

Figure 20: Histograms of logSFR in $\log \mathrm{M}_\odot \mathrm{yr}^{-1}$ for paired galaxies (blue) and their control sample (orange) in the COSMOS sample. *Left:* Full sample. *Right:* Result with the mass ratio requirement of $1/4 < M_{*,1}/M_{*,2} < 4$.

These criteria are adapted from La Marca et al. (2024), because they use the same KiDS sample. Paired galaxies that had fewer than ten control galaxies in the control group were ignored for this analysis. For the KiDS sample mergers were removed from the set that control galaxies were selected from. From the possible control galaxies we randomly select ten for each paired galaxy to form the control group. We then calculate the difference in SFR between a single paired galaxy and its control sample as

$$\Delta \log SFR = \log SFR_{sample} - \log \mathrm{mean}\,(SFR)\,. \tag{13}$$

Just like the distance to the MS, this quantity has units of dex. We plot the distributions of logSFR_{sample} and $\log \mathrm{mean}\,(\mathrm{SFR})$ in Fig. 19, and Fig. 20, for the KiDS and COSMOS samples respectively. We exclude galaxies for which $\sigma_{\log \mathrm{SFR}} > 1$ dex, the median error in the KiDS sample is 0.28 dex and 0.11 dex in the COSMOS sample after applying the aforementioned constraint. We find that KiDS sample galaxies have a significantly lower SFR compared to the control sample, we examine this further below. For the COSMOS sample galaxies we find no clearly visible difference in these plots, this is also examined further below.

We calculate difference between the two distributions in Fig. 19, and 20 and take the median of it. We repeat this 1000 times and plot the distribution of the average difference in SFR in Fig. 21. For the KiDS sample we find a difference of -0.768 ± 0.006 dex in $\log \mathrm{SFR}$, meaning that the SFR in $\mathrm{M}_* \,\mathrm{yr}^{-1}$ is $82.9 \pm 0.2\%$ lower in paired galaxies. For the COSMOS sample we find a difference of -0.182 ± 0.009 dex in $\log \mathrm{SFR}$, meaning that the SFR in $\mathrm{M}_* \,\mathrm{yr}^{-1}$ is $34.3 \pm 1.4\%$ lower in paired galaxies. Compared to Scudder et al. (2012) this is the opposite result. They find an increase of $30 - 60\%$, depending on the criteria for r_p, Δv, and the mass ratio.

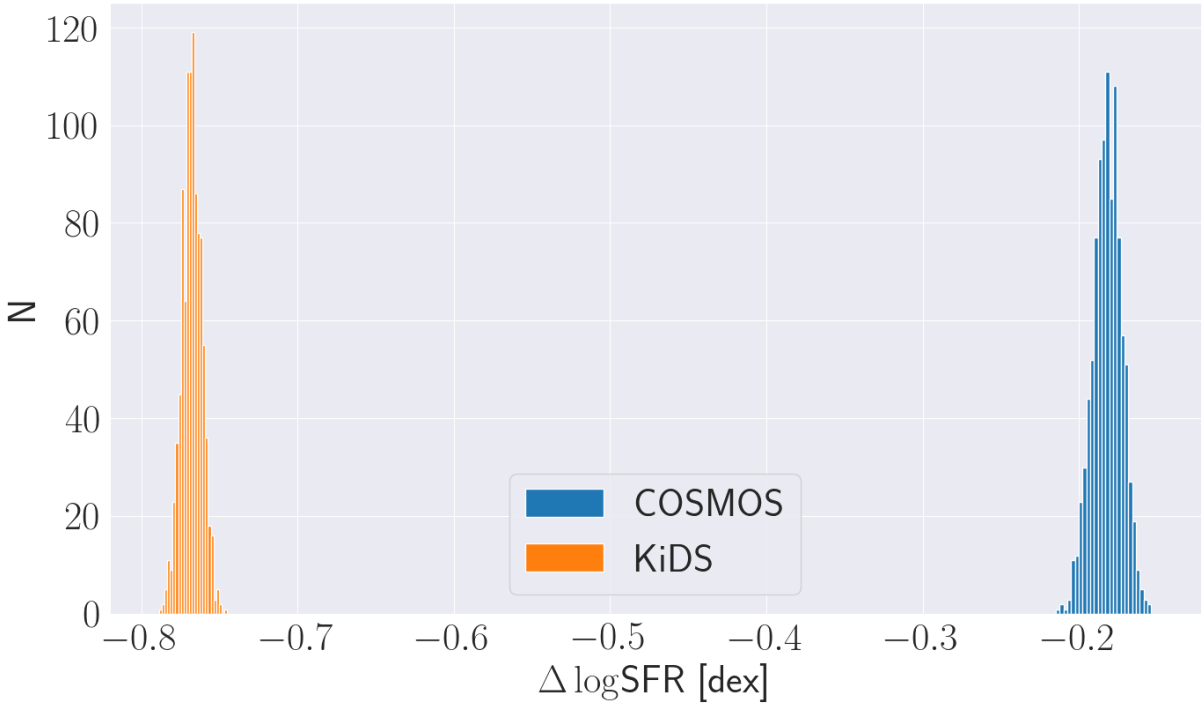

Figure 21: Histograms of the median difference between paired galaxies in the KiDS (orange) and COSMOS (blue) samples and their control samples, for 1000 different control samples. In the KiDS sample we find a median difference of -0.768 ± 0.006 dex. In the COSMOS sample the mean of the distribution is -0.182 ± 0.009 dex.

To constrain our results we also applied a third criteria for our pairs, outlined in Sec. 3.1. To reiterate, we constrain our pairs to a mass ratio of $1/4 < M_{*,1}/M_{*,2} < 4$. This results with this constraint are shown in Fig. 22, the results without this constraint are shown in Fig. 21. In the KiDS data we find a mean difference of -0.797 ± 0.006 dex, thus the SFR of pairs is $84.0 \pm 0.2\%$ lower on average than the control sample. In the COSMOS2020 field we find a mean difference of -0.17 ± 0.01 dex, thus the SFR of pairs is $32 \pm 2\%$ lower on average than the control sample. Compared to the results without a mass constraint we have a slightly higher SFR in the COSMOS sample, but lower in the KiDS sample. However, it still decreases relative to the control sample. The difference with the results of Scudder et al. (2012) could be explained by the different methods of determining the SFR. Their SFR is calibrated by the fit of emission lines to SDSS spectra, while our SFR values are modelled using CIGALE. These are dependent on the parameters of that model and the number of wavelength bands in the data.

4.3 Distance to Galaxy MS vs Merger Fraction

To compare the SFR of mergers we again calculated the distance to the MS outlined in Sec. 4.1. The resulting distribution is shown in Fig. 23, the relative number of galaxies above the MS has decreased compared to the full KiDS sample in Fig. 13. Either the selection of the subsample is biased towards MS galaxies or mergers are more likely to lie on the MS.

The same binning of the distance to the MS as in Sec. 4.1 was performed for the mergers, resulting in Fig. 25. The left plot shows the relation between $\Delta \log \mathrm{SFR}$ and the merger fraction, $f_{\mathrm{merger}} = N_{\mathrm{merger}}/N_{\mathrm{non-merger}}$, for the KiDS subsample. The right plot includes the $\sigma_{\Delta \log \mathrm{SFR}} < 1$ constraint. In both plots we find a higher f_{merger} below the MS, the merger fraction decreases only slightly above the main sequence. A similar result was seen in Sec. 4.1 with the pair fraction,

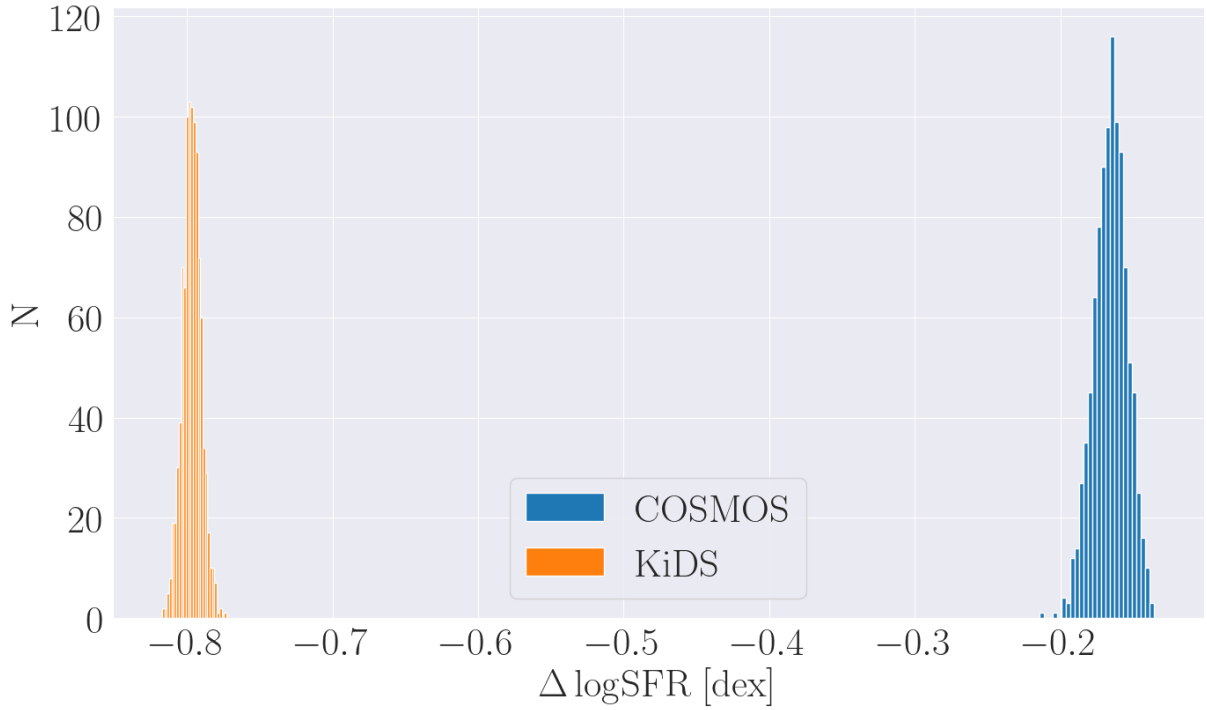

Figure 22: Histograms of the median difference between paired galaxies in the KiDS (orange) and COSMOS (blue) samples and their control samples, for 1000 different control samples. The paired galaxies have mass ratios of $1/4 < M_{*,1}/M_{*,2} < 4$. In the KiDS sample we find a mean difference of -0.797 ± 0.006 dex. In the COSMOS sample we find a mean difference of -0.17 ± 0.01 dex.

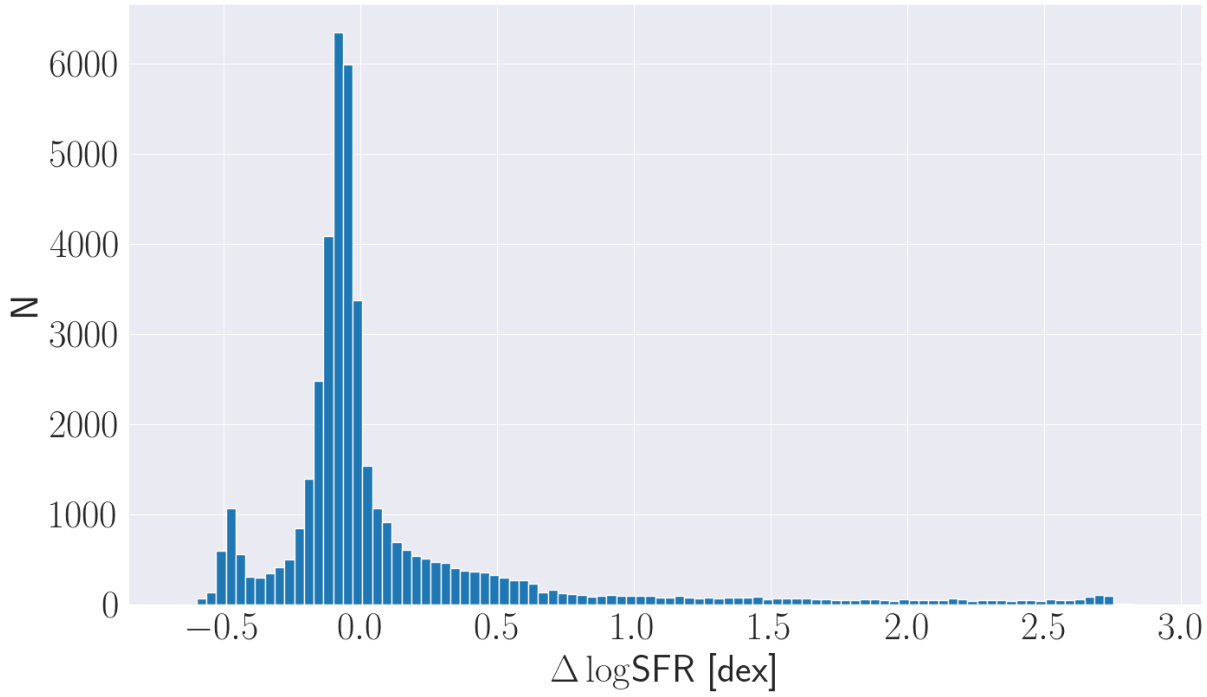

Figure 23: Distribution of the distance to the MS for the merger subsample of the KiDS sample.

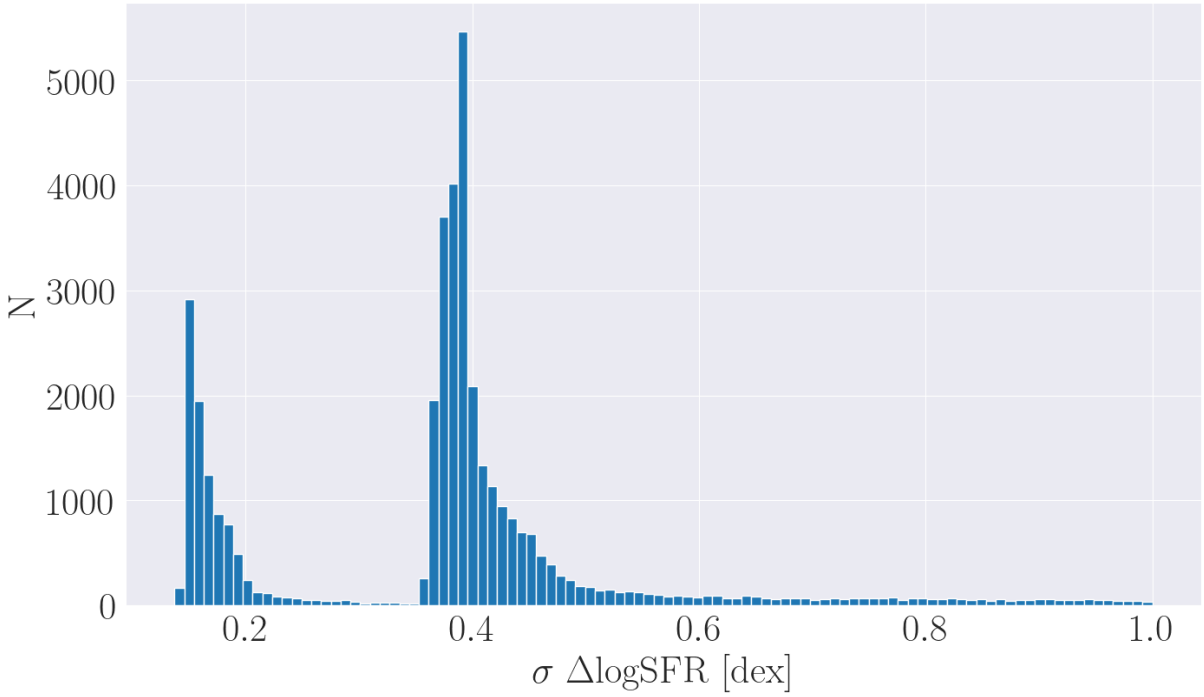

Figure 24: Histogram of the error in $\Delta \log \mathrm{SFR}$ for the KiDS subsample. We find that the error tapers off faster than for the full sample, as seen in Fig. 14. We limit the range of this plot to $\sigma_{\Delta \log \mathrm{SFR}} < 1$.

indicating that this relation is consistent between the different methods of classifying mergers that we use. The larger fraction of mergers below the MS could be older (post-)mergers, because star formation can rapidly shut down in mergers after they have coalesced, as is shown by Ellison et al. (2022). Due to the lower quality of the SFR determined with CIGALE, as mentioned in Secs. 2 and 3.2, this result should be taken with a grain of salt. Another caveat is that the subsample of the KiDS sample in which La Marca et al. (2024) classified the galaxies as mergers, indeterminate, or non-mergers was created by selecting galaxies with a fitted AGN component. However, it is not guaranteed that this subsample is representative of the entire population. In the right plot of Fig. 25 we find the same result as the left plot, however the constraint on $\sigma_{\Delta \log \mathrm{SFR}}$ has caused several significant outliers around $\Delta \log \mathrm{SFR} \approx 1$. The sample size in these bins is lower than around $\Delta \log \mathrm{SFR} = 0$, but the bins at $\Delta \log \mathrm{SFR} > 2$ have similar sample sizes without any large outliers.

4.4 Comparison of the SFR of Mergers and Control Group

A similar analysis as in Sec. 4.2 was performed on the merger sample. We again exclude galaxies for which $\sigma_{\log \mathrm{SFR}} > 1$ dex, the median error is 0.07 dex after applying the aforementioned constraint. We compare the $\log \mathrm{SFR}$ of mergers to a control sample in Fig. 26. We find that both have a very similar distribution, but the control sample has a larger population of galaxies around $\log \mathrm{SFR} \approx 3$. We calculate the median difference between the merger $\log \mathrm{SFR}$ and the control sample for 1000 control samples. The resulting distribution is shown in Fig. 27. We find a mean difference between mergers and non-mergers of -0.014 ± 0.001 dex. The SFR of mergers in our sample is $3.2 \pm 0.3\%$ lower on average than non-mergers in our sample.

In the comparison of mergers/close pairs and their control samples there are several possible sources of errors. It should first be noted that the error given for the mean difference between

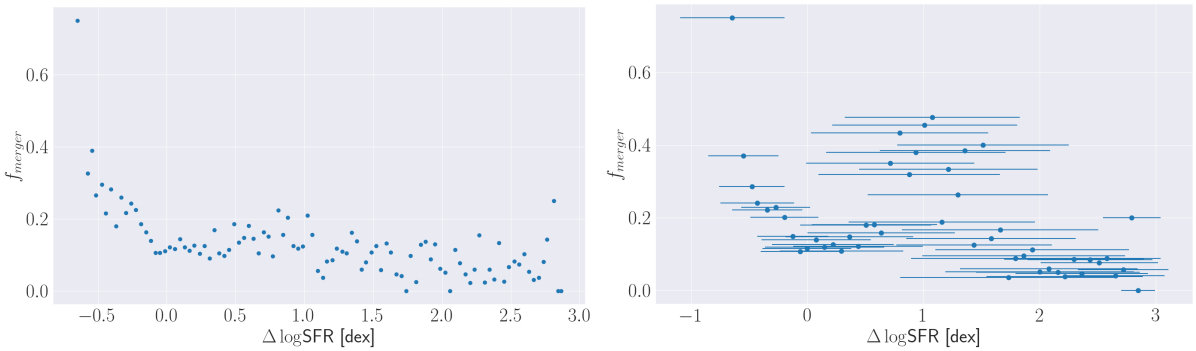

Figure 25: Plots of the distance to the MS against the merger fraction in the $\Delta \log \mathrm{SFR}$ bins in the KiDS subsample. The sample size of each bin is shown in Fig. 23. We find a decrease in the merger fraction as $\Delta \log \mathrm{SFR}$ increases. The left plot is derived from the KiDS subsample, while the right plot includes the requirement on the error in $\Delta \log \mathrm{SFR}$ of $\sigma_{\Delta \log \mathrm{SFR}} < 1$ dex.

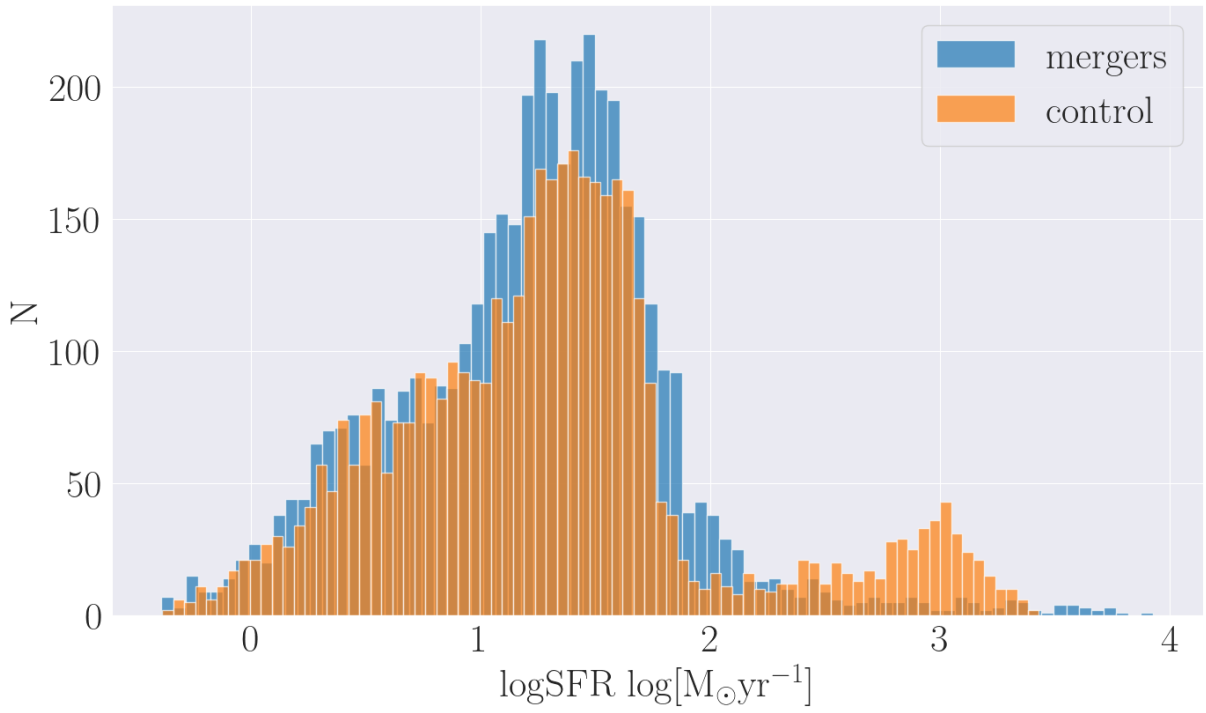

Figure 26: Comparison of the logSFR of mergers (blue) in the KiDS sample and a control sample (orange). The control sample shows a relatively larger population of galaxies around $\log \mathrm{SFR} \approx 3$ dex.

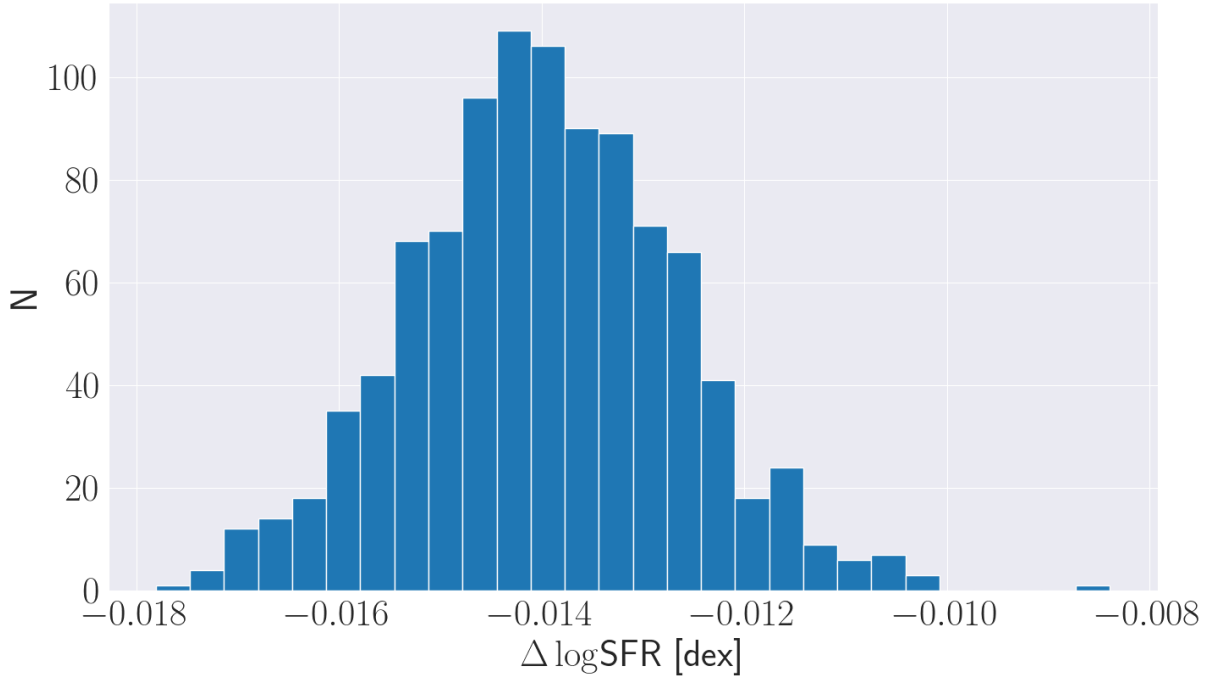

Figure 27: Distribution of the median difference in SFR between mergers in the KiDS subsample and the control sample, calculated for 1000 control samples. We find a mean difference of -0.014 ± 0.001 dex.

mergers/close pairs and 1000 control samples is the standard deviation of the distributions shown in Fig. 21, 22, and 27. The median of the errors in $\log \mathrm{SFR}$ is 0.28, 0.07, and 0.11 dex for the KiDS sample, subsample and COSMOS sample respectively. The error is especially large for the KiDS sample. However, this error is not large enough to sway our result for that sample: a decrease of $83 - 84\%$ (-0.797 dex) in close pairs w.r.t. their control samples. The result for the merger subsample was heavily dependent on whether we use the mean or the median when calculating the average difference of the distributions in Fig. 26 and the average of the distribution in Fig. 27. The choices caused the result of -0.014 ± 0.001 dex to vary from ≈ -0.04 dex to ≈ 0.02 dex. The last potential impact on the results is any inherent bias in the SFR derived with CIGALE. We have seen that the KiDS sample has higher errors and its SFR-M_* diagram has a less physically reasonable shape than the COSMOS sample. The SFR-M_* diagrams for the KiDS and COSMOS samples can be compared in Fig. 2, and 4 respectively.

5 Conclusion

We analysed the relation between the SFR and merger/close pair fraction in the KiDS and COSMOS samples. We find 11023 unique galaxies in close pairs in the KiDS sample. The close pair fraction was $f_{pair} = 0.021$ within the redshift range of $0.10 < z < 0.76$. We compared these results to the literature and found that our value is lower, however, Desmons et al. (2023) use a lower redshift range of $0.04 < z < 0.20$. They also use the smaller GAMA G02 field, so our results are not directly comparable. The lower redshift can cause galaxies which would not be detected at higher redshifts to be detected. This makes it possible that the number density of galaxies is larger in their data, and therefore that we find a lower pair fraction.

In the COSMOS sample we find 2889 unique galaxies in close pairs, with $f_{pair} = 0.11$, this is significantly higher than the pair fraction in the KiDS sample. One explanation is that the imaging in this field is much deeper than the KiDS sample, as this field was observed by the Hubble Space Telescope. Ferreira et al. (2020) find $f_{pair} = 0.041 - 0.181$ for the same redshift range and field. They use a lower projected separation criterion of $r_p < 20\,h^{-1}\,\mathrm{kpc}$, which would lead to a lower pair fraction. Man et al. (2012) find $f_{pair} = 0.08 - 0.15$ with $r_p < 30\,\mathrm{kpc}$ in a redshift range of $0.0 < z < 3.0$, they also examine the COSMOS field. Our COSMOS sample doesn't include the entire COSMOS field, which might be another factor in the difference in f_{pair} when compared with redshift in Fig. 6.

When we require a mass ratio of $1/4 < M_{*,1}/M_{*,2} < 4$ for our pairs we find 7587 unique paired galaxies in the KiDS sample and 1706 in the COSMOS sample. The new pair fractions are thus $f_{pair} = 0.015$ and $f_{pair} = 0.065$ for the KiDS and COSMOS sample respectively.

Fig. 6 shows the pair fraction as a function of redshift in the KiDS and COSMOS samples. The redshift bins on the x-axis are defined in Sec. 2. The trend in f_{pair} is similar for all but one z-bin of both fields, the last z-bin of the COSMOS sample has a relatively higher pair fraction when we apply the mass ratio criterion compared to previous bins. We conjecture that this is because of the smaller mass range detected at higher redshifts, which in turn makes it more likely for two galaxies in a pair to satisfy the mass ratio criterion. The pair fractions from Ferreira et al. (2020), which also include the same mass ratio criterion, seem to line up for the first two z-bins. However, they deviate for the last two z-bins. Especially z-bin 4 shows a significant discrepancy in f_{pair} between our findings and theirs.

We also compare the functional form of the galaxy MS from Popesso et al. (2023) to the COSMOS data, the result is shown in Fig. 7. This function did not fit well to the KiDS data, nor did any other function from the literature, so we fit Eq. 9 in each z-bin separately. The results are shown in Fig. 31, 32, and 33 in Appendix A. The value of the parameters $a(t)$ and $b(t)$ are shown in Fig. 12, they do not show the linear relation we expect from Eqs. 5, 6, and 7 (Speagle et al., 2014; Leslie et al., 2020; Popesso et al., 2023).

With the functional forms of the galaxy MS we determined the distance to the MS as the difference between the logSFR of our galaxies as measured and if they were to lie on the MS. This was converted to a SFR ratio to determine the increase or decrease in SFR. We show the results of the analysis of the pair fraction and projected separation in Fig. 15, 16, 17, and 18. The last two include the mass ratio constraint of $1/4 < M_{*,1}/M_{*,2} < 4$ and the second and fourth include the error constraint of $\sigma_{\Delta \log \mathrm{SFR}} < 1$ dex. We found that the pair fraction decreases with distance to the MS in both fields and with or without the mass ratio constraint. However, when we constrain the error, we find that f_{pair} remains constant in both samples. The samples do show significant outliers, in the COSMOS sample these could be caused by a small sample size, but this is not the case for the KiDS sample. We also found that the distance to the MS increases with r_p of close pairs in the KiDS sample, in the COSMOS sample we found no such trend.

There is a large spread in the data and the trend in the KiDS sample is not very strong, so no clear conclusion can be drawn from it. In the COSMOS sample we find a decrease in distance to the MS as r_p increases, but this trend disappears when we apply the mass ratio constraint. Once we apply both the mass ratio constraint and $\sigma\Delta \log \mathrm{SFR}$ constraint, we find that $\Delta \log \mathrm{SFR}$ is constant with r_p in both samples, except for three outliers in the KiDS sample. The SFR of close pairs is 2 dex above the MS in the KiDS sample, but 0.2 dex below the MS in the COSMOS sample.

We also construct control samples and compare their SFR against that of the close pairs. We find that, without the mass ratio constraint, paired galaxies in the KiDS sample have a SFR that is $82.9 \pm 0.2\%$ lower than the control sample and $84.0 \pm 0.2\%$ with the mass ratio constraint. For the COSMOS sample the SFR is $34.3 \pm 1.4\%$ and $32 \pm 2\%$ lower, with and without the mass ratio requirement respectively. The error is derived from the random selection of our control sample, by repeating the selection 1000 times we found a Gaussian shape for the percentage increase and use the standard deviation as the error. The pairs always show a decrease, which was unexpected compared to the literature. Scudder et al. (2012) find an increase of $30 - 60\%$, but they calibrate their SFR by fitting emission lines to SDSS spectra, while we use CIGALE.

With the mergers found by La Marca et al. (2024) we plot the merger fraction as a function of distance to the MS in Fig. 25. We find a larger merger fraction below the main sequence than above. This might be a population of mergers in which star formation has already quenched. It is important to keep in mind the lower quality of the data in the KiDS sample, as is outlined in Secs. 2 and 4. It is also important to remember that the mergers were identified in a subsample of the full dataset in which only galaxies with a fitted AGN component are present. This subsample is likely not representative of the complete sample, as is outlined in Sec 4.

We also compare the difference in SFR between mergers and non-mergers, as identified by La Marca et al. (2024). We find that the SFR in mergers is $3.2 \pm 0.3\%$ lower than in non-mergers. The error is also derived from repeating the random control sample selection.

Overall our results show an unexpected decrease in SFR in close pairs compared to the literature. We also find a decrease in SFR in mergers, but this a very small decrease which can fall within the margin of error from the CIGALE SED fitting as discussed at the end of Sec. 4. Others sources of error are discussed at the end of Sec. 4. We also find significant outliers in the KiDS sample. More research is needed to uncover why these are present in the data.

References

[1] Aihara, H., Armstrong, R., Bickerton, S., Bosch, J., et al. (2018). First data release of the Hyper Suprime-Cam Subaru Strategic Program. *PASJ*, 70:S8.

[2] Bickley, R. W., Ellison, S. L., Patton, D. R., Bottrell, C., Gwyn, S., and Hudson, M. J. (2022). Star formation characteristics of CNN-identified post-mergers in the Ultraviolet Near Infrared Optical Northern Survey (UNIONS). *MNRAS*, 514(3):3294–3307.

[3] Bock, D. C. J., Bolatto, A. D., Hawkins, D. W., Kemball, A. J., et al. (2006). First results from CARMA: the combined array for research in millimeter-wave astronomy. In Stepp, L. M., editor, *Ground-based and Airborne Telescopes*, volume 6267 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 626713.

[4] Boquien, M., Burgarella, D., Roehlly, Y., Buat, V., et al. (2019). CIGALE: a python Code Investigating GALaxy Emission. *AAp*, 622:A103.

[5] Brunner, H., Liu, T., Lamer, G., Georgakakis, A., et al. (2022). The eROSITA Final Equatorial Depth Survey (eFEDS). X-ray catalogue. *AAp*, 661:A1.

[6] Cutri, R. M., Wright, E. L., Conrow, T., Fowler, J. W., et al. (2013). Explanatory Supplement to the AllWISE Data Release Products. Explanatory Supplement to the AllWISE Data Release Products, by R. M. Cutri et al.

[7] de Jong, J. T. A., Kuijken, K., Applegate, D., Begeman, K., et al. (2013). The Kilo-Degree Survey. *The Messenger*, 154:44–46.

[8] Desmons, A., Brough, S., Martínez-Lombilla, C., De Propris, R., et al. (2023). Galaxy and mass assembly (GAMA): comparing visually and spectroscopically identified galaxy merger samples. *MNRAS*, 523(3):4381–4393.

[9] Driver, S. P., Hill, D. T., Kelvin, L. S., Robotham, A. S. G., et al. (2011). Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *MNRAS*, 413(2):971–995.

[10] Dubois, Y., Pichon, C., Welker, C., Le Borgne, D., et al. (2014). Dancing in the dark: galactic properties trace spin swings along the cosmic web. *MNRAS*, 444(2):1453–1468.

[11] Duncan, K., Conselice, C. J., Mundy, C., Bell, E., et al. (2019). Observational Constraints on the Merger History of Galaxies since z $\approx$ 6: Probabilistic Galaxy Pair Counts in the CANDELS Fields. *ApJ*, 876(2):110.

[12] Edge, A., Sutherland, W., Kuijken, K., Driver, S., et al. (2013). The VISTA Kilo-degree Infrared Galaxy (VIKING) Survey: Bridging the Gap between Low and High Redshift. *The Messenger*, 154:32–34.

[13] Ellison, S. L., Wilkinson, S., Woo, J., Leung, H.-H., et al. (2022). Galaxy mergers can rapidly shut down star formation. *MNRAS*, 517(1):L92–L96.

[14] Euclid Collaboration, Moneti, A., McCracken, H. J., Shuntov, M., et al. (2022). Euclid preparation. XVII. Cosmic Dawn Survey: Spitzer Space Telescope observations of the Euclid deep fields and calibration fields. *AAp*, 658:A126.

[15] Ferreira, L., Conselice, C. J., Duncan, K., Cheng, T.-Y., et al. (2020). Galaxy Merger Rates up to z $\sim$ 3 Using a Bayesian Deep Learning Model: A Major-merger Classifier Using IllustrisTNG Simulation Data. *ApJ*, 895(2):115.

[16] Garay-Solis, Y., Barrera-Ballesteros, J. K., Colombo, D., Sánchez, S. F., et al. (2023). Exploring the Impact of Galactic Interactions and Mergers on the Central Star Formation of APEX/EDGE-CALIFA Galaxies. *ApJ*, 952(2):122.

[17] Güsten, R., Nyman, L. Å., Schilke, P., et al. (2006). The Atacama Pathfinder EXperiment (APEX) - a new submillimeter facility for southern skies -. *AAp*, 454(2):L13–L16.

[18] Kuijken, K., Heymans, C., Dvornik, A., Hildebrandt, H., et al. (2019). The fourth data release of the Kilo-Degree Survey: ugri imaging and nine-band optical-IR photometry over 1000 square degrees. *AAp*, 625:A2.

[19] La Marca, A., Margalef-Betnabol, B., Wang, L., Gao, F., et al. (2024). Dust and Power: Unravelling the merger-AGN connection in the second half of the cosmic history.

[20] Leauthaud, A., Massey, R., Kneib, J.-P., Rhodes, J., et al. (2007). Weak Gravitational Lensing with COSMOS: Galaxy Selection and Shape Measurements. *ApJS*, 172(1):219–238.

[21] Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

[22] Leslie, S. K., Schinnerer, E., Liu, D., Magnelli, B., et al. (2020). The VLA-COSMOS 3 GHz Large Project: Evolution of Specific Star Formation Rates out to z $\sim$ 5. *ApJ*, 899(1):58.

[23] Man, A. W. S., Toft, S., Zirm, A. W., Wuyts, S., and van der Wel, A. (2012). The Pair Fraction of Massive Galaxies at 0 <= z <= 3. *ApJ*, 744(2):85.

[24] McCracken, H. J., Milvang-Jensen, B., Dunlop, J., Franx, M., et al. (2012). UltraVISTA: a new ultra-deep near-infrared survey in COSMOS. *AAp*, 544:A156.

[25] Mundy, C. J., Conselice, C. J., Duncan, K. J., Almaini, O., et al. (2017). A consistent measure of the merger histories of massive galaxies using close-pair statistics - I. Major mergers at z $<$ 3.5. *MNRAS*, 470(3):3507–3531.

[26] Newman, J. A. and Gruen, D. (2022). Photometric Redshifts for Next-Generation Surveys. *ARAA*, 60:363–414.

[27] Pillepich, A., Springel, V., Nelson, D., Genel, S., et al. (2018). Simulating galaxy formation with the IllustrisTNG model. *MNRAS*, 473(3):4077–4106.

[28] Popesso, P., Concas, A., Cresci, G., Belli, S., et al. (2023). The main sequence of star-forming galaxies across cosmic times. *MNRAS*, 519(1):1526–1544.

[29] Predehl, P., Andritschke, R., Arefiev, V., Babyshkin, V., et al. (2021). The eROSITA X-ray telescope on SRG. *AAp*, 647:A1.

[30] Saintonge, A., Kauffmann, G., Kramer, C., Tacconi, L. J., et al. (2011). COLD GASS, an IRAM legacy survey of molecular gas in massive galaxies - I. Relations between H_2, H I, stellar content and structural properties. *MNRAS*, 415(1):32–60.

[31] Saintonge, A., Tacconi, L. J., Fabello, S., Wang, J., et al. (2012). The Impact of Interactions, Bars, Bulges, and Active Galactic Nuclei on Star Formation Efficiency in Local Massive Galaxies. *ApJ*, 758(2):73.

[32] Sánchez, S. F., Kennicutt, R. C., Gil de Paz, A., van de Ven, G., et al. (2012). CALIFA, the Calar Alto Legacy Integral Field Area survey. I. Survey presentation. *AAp*, 538:A8.

[33] Sawicki, M., Arnouts, S., Huang, J., Coupon, J., et al. (2019). The CFHT large area U-band deep survey (CLAUDS). *MNRAS*, 489(4):5202–5217.

[34] Scudder, J. M., Ellison, S. L., Torrey, P., Patton, D. R., and Mendel, J. T. (2012). Galaxy pairs in the Sloan Digital Sky Survey - V. Tracing changes in star formation rate and metallicity out to separations of 80 kpc. *MNRAS*, 426(1):549–565.

[35] Speagle, J. S., Steinhardt, C. L., Capak, P. L., and Silverman, J. D. (2014). A Highly Consistent Framework for the Evolution of the Star-Forming "Main Sequence" from z ~0-6. *ApJS*, 214(2):15.

[36] Taniguchi, Y., Kajisawa, M., Kobayashi, M. A. R., Shioya, Y., et al. (2015). The Subaru COSMOS 20: Subaru optical imaging of the HST COSMOS field with 20 filters*. *PASJ*, 67(6):104.

[37] Taniguchi, Y., Scoville, N., Murayama, T., Sanders, D. B., et al. (2007). The Cosmic Evolution Survey (COSMOS): Subaru Observations of the HST Cosmos Field. *ApJS*, 172(1):9–28.

[38] Valiante, E., Smith, M. W. L., Eales, S., Maddox, S. J., et al. (2016). The Herschel-ATLAS data release 1 - I. Maps, catalogues and number counts. *MNRAS*, 462(3):3146–3179.

[39] Ventou, E., Contini, T., Bouché, N., Epinat, B., et al. (2019). New criteria for the selection of galaxy close pairs from cosmological simulations: evolution of the major and minor merger fraction in MUSE deep fields. *AAp*, 631:A87.

[40] Violino, G., Ellison, S. L., Sargent, M., Coppin, K. E. K., et al. (2018). Galaxy pairs in the SDSS - XIII. The connection between enhanced star formation and molecular gas properties in galaxy mergers. *MNRAS*, 476(2):2591–2604.

[41] Wang, L., La Marca, A., Gao, F., Pearson, W. J., et al. (2024). Probabilistic and progressive deblended far-infrared and sub-millimetre point source catalogues I. Methodology and first application in the COSMOS field. *arXiv e-prints*, page arXiv:2405.18290.

[42] Weaver, J. R., Kauffmann, O. B., Ilbert, O., McCracken, H. J., et al. (2022). COSMOS2020: A Panchromatic View of the Universe to z∼10 from Two Complementary Catalogs. *ApJS*, 258(1):11.

[43] Wilkinson, S., Ellison, S. L., Bottrell, C., Bickley, R. W., et al. (2022). The merger fraction of post-starburst galaxies in UNIONS. *MNRAS*, 516(3):4354–4372.

[44] Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., Ressler, M. E., et al. (2010). The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *AJ*, 140(6):1868–1881.

[45] York, D. G., Adelman, J., Anderson, John E., J., Anderson, S. F., et al. (2000). The Sloan Digital Sky Survey: Technical Summary. *AJ*, 120(3):1579–1587.

[46] Zamojski, M. A., Schiminovich, D., Rich, R. M., Mobasher, B., et al. (2007). Deep GALEX Imaging of the COSMOS HST Field: A First Look at the Morphology of z ~0.7 Star-forming Galaxies. *ApJS*, 172(1):468–493.

A Additional Plots

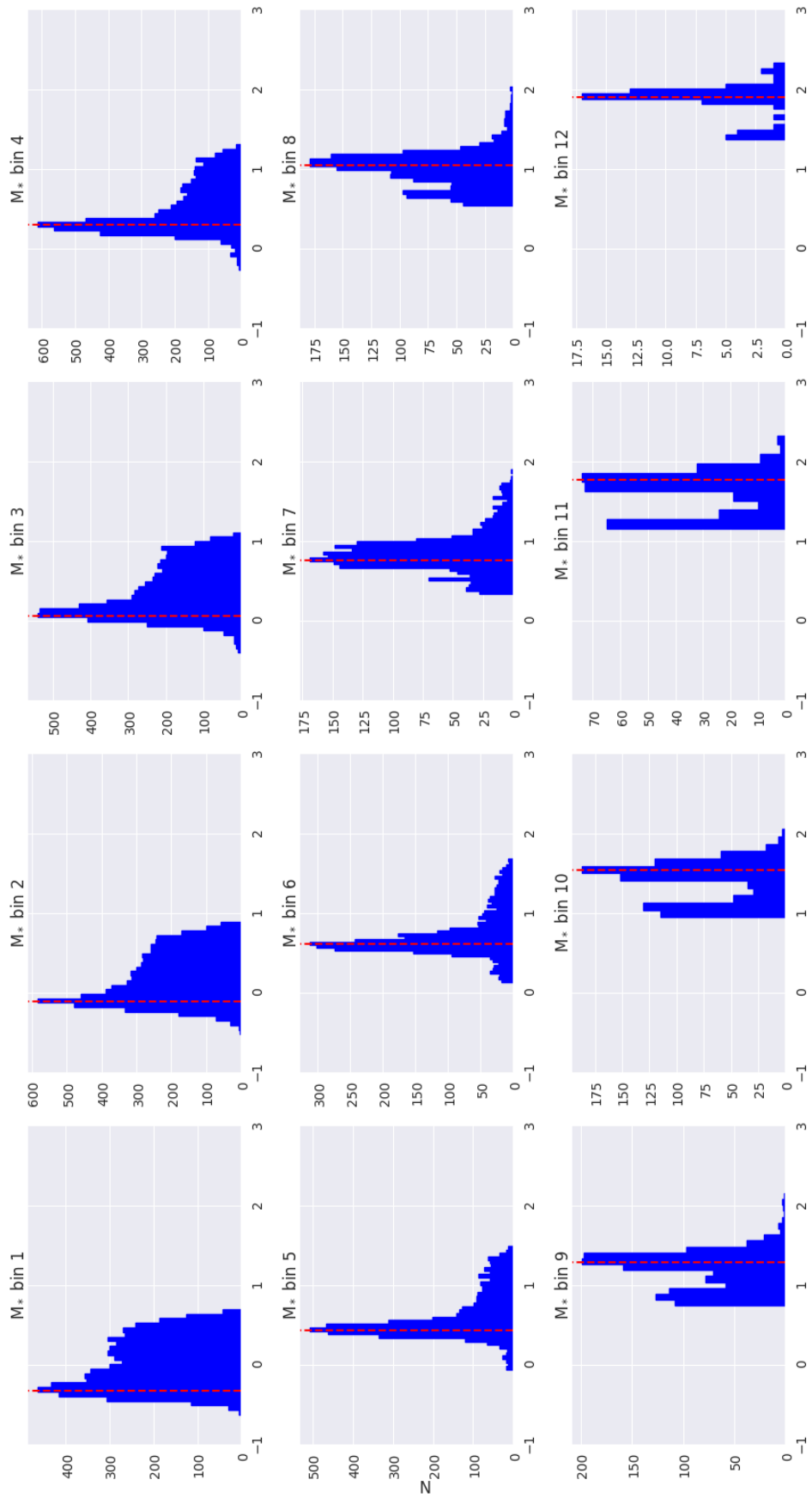

Figure 28: These plots show the distribution of the SFR in the individual M_* bins in redshift bin 1. The red dashed line indicates the mode of the distribution, which is used to fit the galaxy MS.

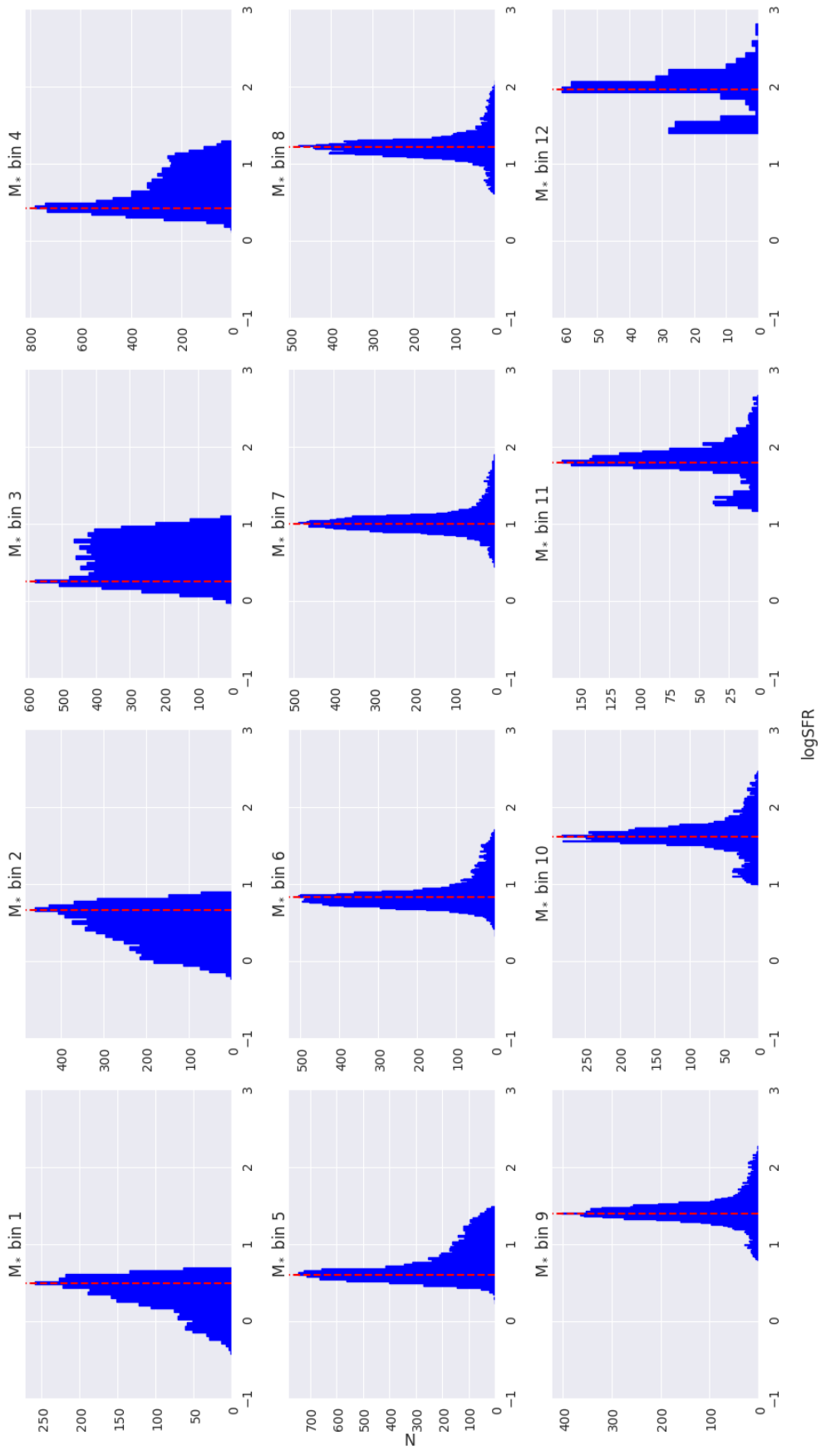

Figure 29: These plots show the distribution of the SFR in the individual M_* bins in redshift bin 2. The red dashed line indicates the mode of the distribution, which is used to fit the galaxy MS.

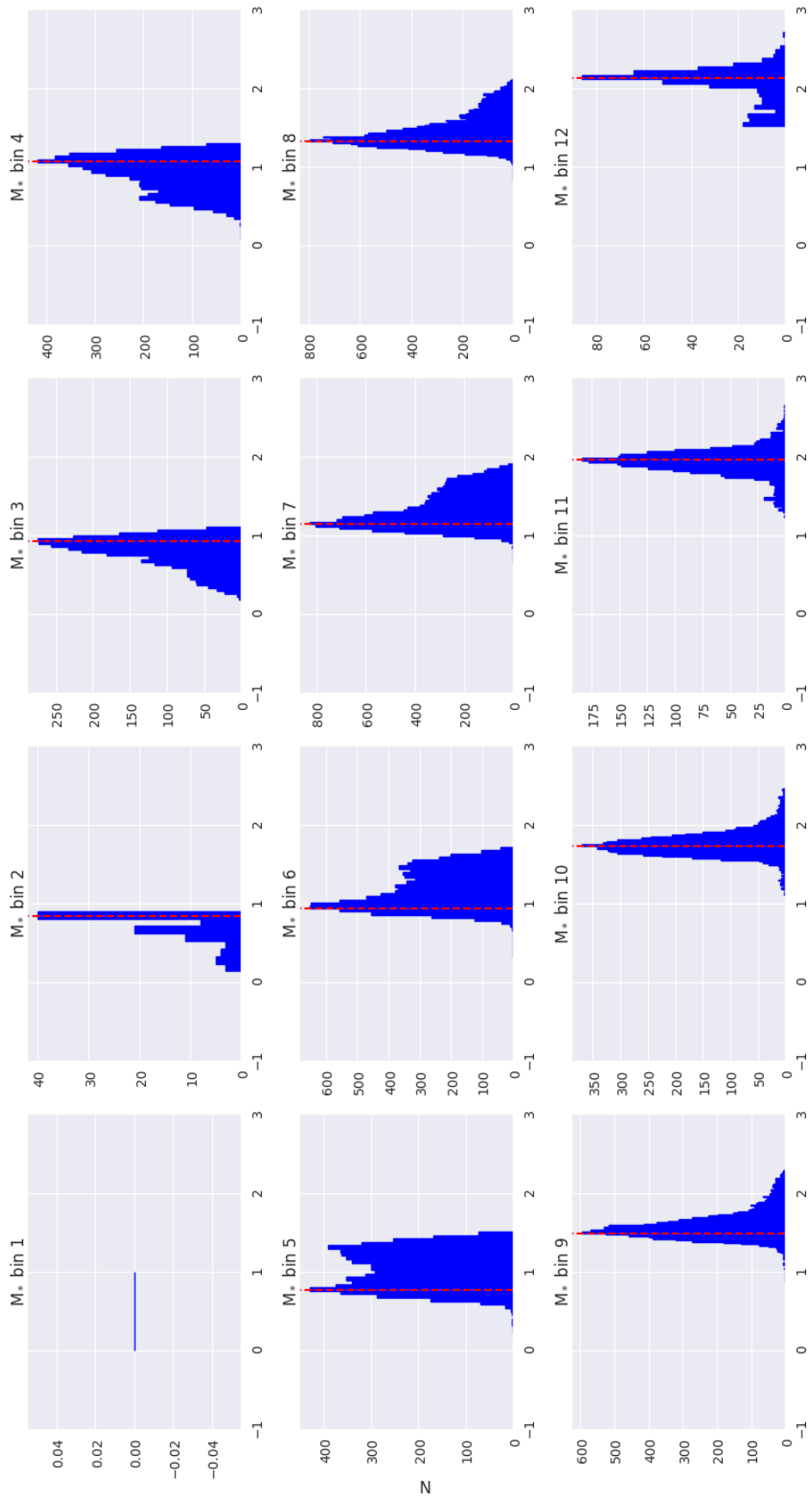

Figure 30: These plots show the distribution of the SFR in the individual M_* bins in redshift bin 3. The red dashed line indicates the mode of the distribution, which is used to fit the galaxy MS. The first bin is empty because the mass limit in redshift bin 3 is higher than that of 1 & 2.

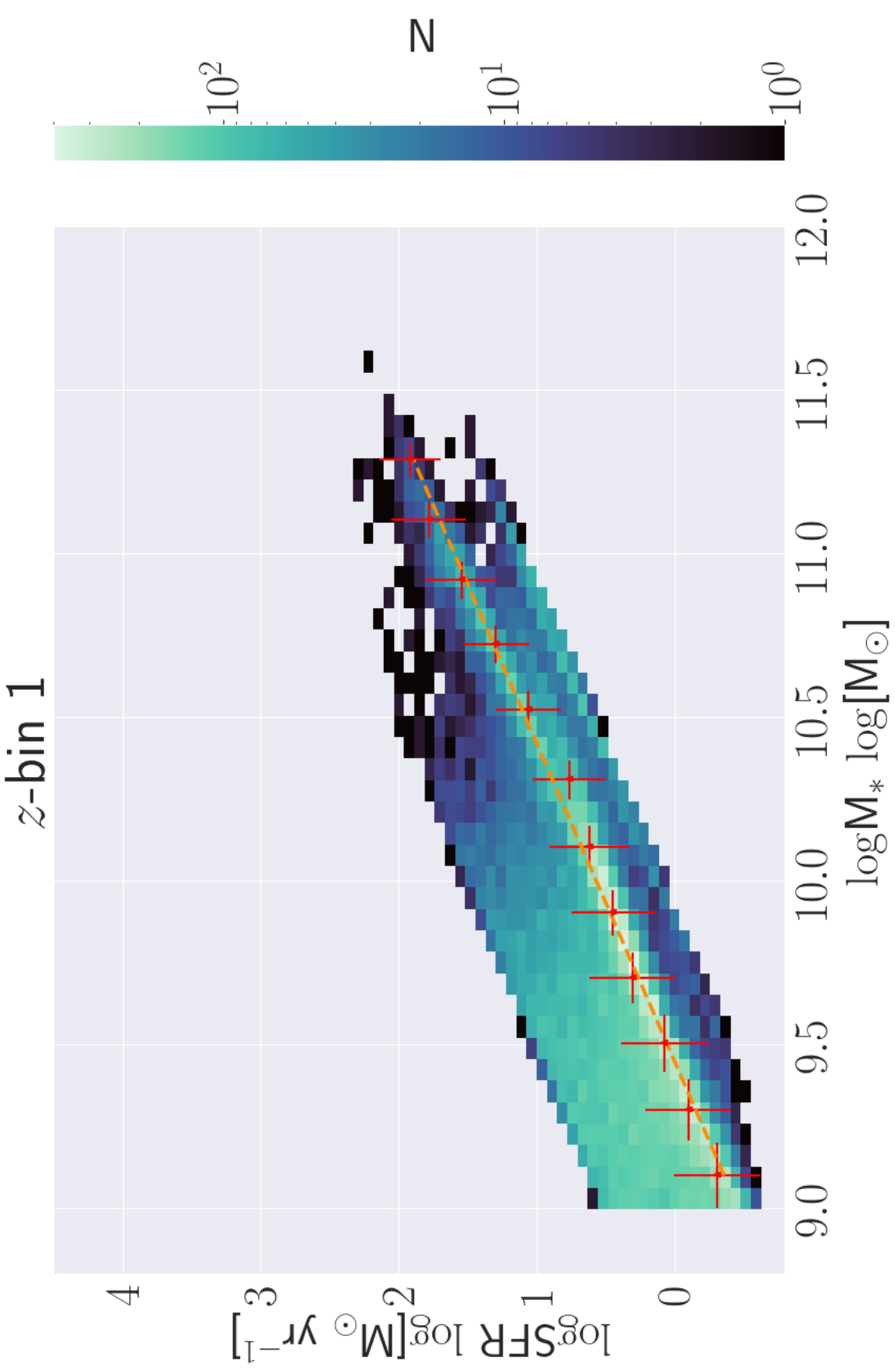

Figure 31: SFR-M_* plot of the fit of the galaxy MS in z-bin 1. The colour intensity indicates the number of galaxies in each bin, the x axis indicates the log M_* in $M_\odot$ and the y axis indicates the SFR in $M_\odot yr^{-1}$. The red points indicate the mode of the SFR distribution in each M_* bin with their associated errors, shown in Fig 28, the orange dashed line shows the fit of Eq. 9.

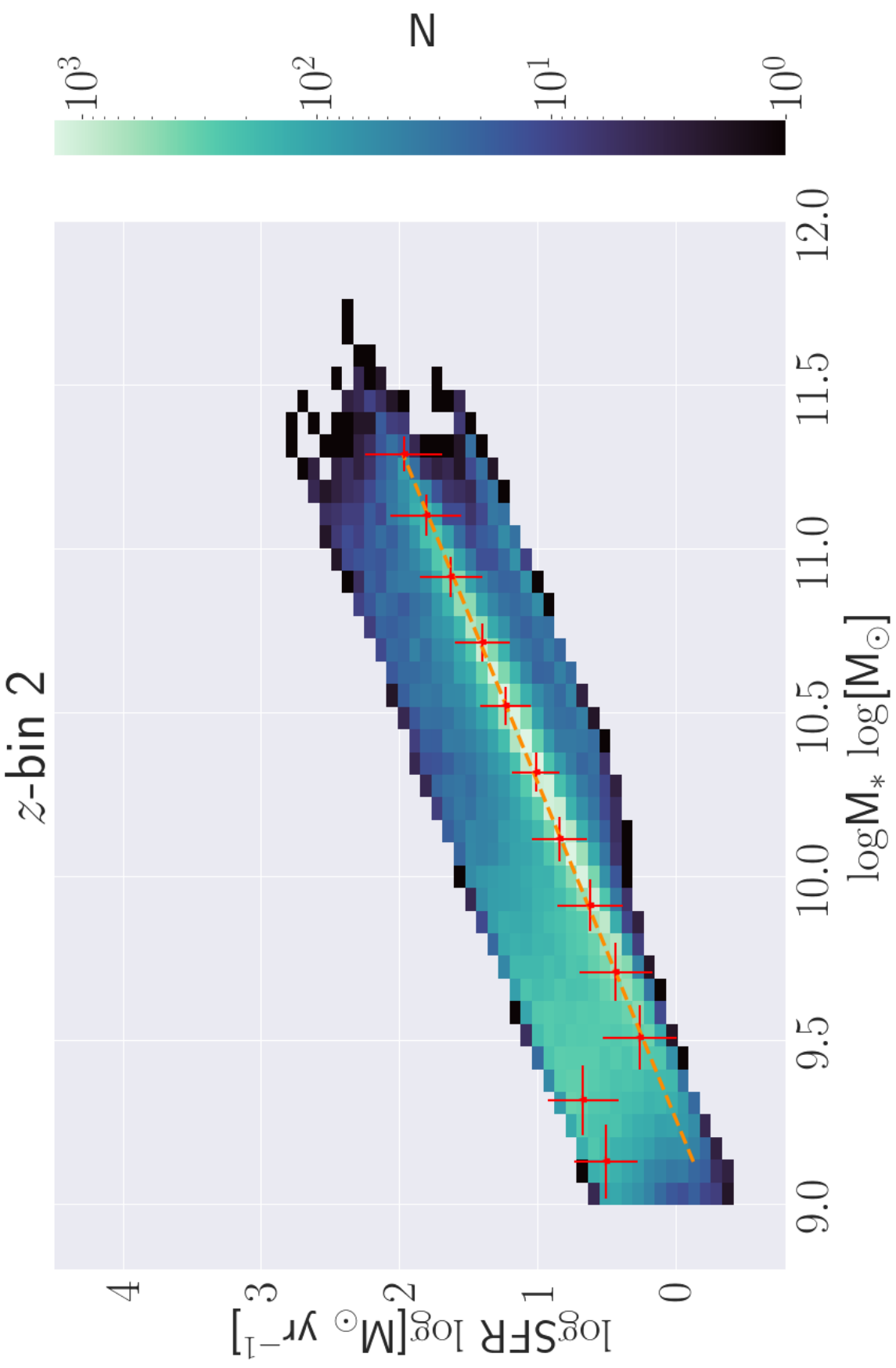

Figure 32: SFR-M$_*$ plot of the fit of the galaxy MS in z-bin 2. The colour intensity indicates the number of galaxies in each bin, the x axis indicates the log M$_*$ in M$_\odot$ and the y axis indicates the SFR in M$_\odot$yr^{-1}. The red points indicate the mode of the SFR distribution in each M$_*$ bin with their associated errors, shown in Fig 29, the orange dashed line shows the fit of Eq. 9. The first two points are ignored when fitting, because they clearly do not lie on the MS.

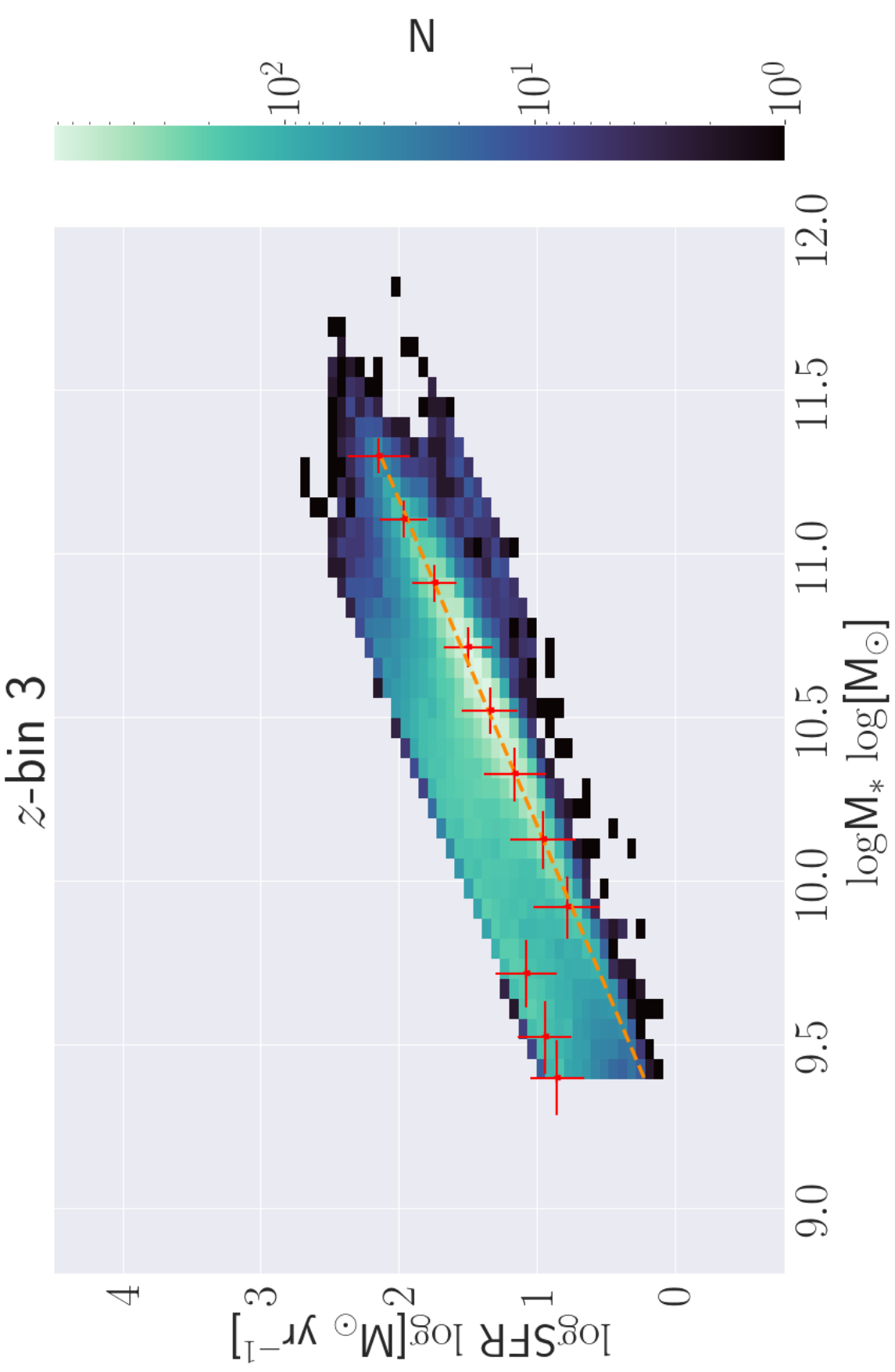

Figure 33: SFR-M$_*$ plot of the fit of the galaxy MS in z-bin 3. The colour intensity indicates the number of galaxies in each bin, the x axis indicates the log M$_*$ in M$_\odot$ and the y axis indicates the SFR in M$_\odot$yr^{-1}. The red points indicate the mode of the SFR distribution in each M$_*$ bin with their associated errors, shown in Fig 30, the orange dashed line shows the fit of Eq. 9. The first three points are ignored when fitting, because they clearly do not lie on the MS.