



**HAL**  
open science

## The role of stellar mass in the cosmic history of star formation as seen by Herschel and ALMA

L. Leroy, D. Elbaz, B. Magnelli, C. Gómez-Guijarro, E. Daddi, M. Xiao, M. Dickinson

► **To cite this version:**

L. Leroy, D. Elbaz, B. Magnelli, C. Gómez-Guijarro, E. Daddi, et al.. The role of stellar mass in the cosmic history of star formation as seen by Herschel and ALMA. *Astronomy & Astrophysics - A&A*, 2024, 691, pp.A248. 10.1051/0004-6361/202348341 . cea-04789621

**HAL Id: cea-04789621**

<https://cea.hal.science/cea-04789621v1>

Submitted on 18 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# The role of stellar mass in the cosmic history of star formation as seen by *Herschel* and ALMA

L. Leroy<sup>1,\*</sup>, D. Elbaz<sup>1</sup>, B. Magnelli<sup>1</sup>, C. Gómez-Guijarro<sup>1</sup>, E. Daddi<sup>1</sup>, M. Xiao<sup>1,2</sup>, and M. Dickinson<sup>3</sup>

<sup>1</sup> Université Paris Cité, Université Paris-Saclay, CEA, CNRS, AIM, F-91191 Gif-sur-Yvette, France

<sup>2</sup> Department of Astronomy, University of Geneva, Chemin Pegasi 51, 1290 Versoix, Switzerland

<sup>3</sup> NSF's National Optical-Infrared Astronomy Research Laboratory, 950 N. Cherry Ave., Tucson, AZ 85719, USA

Received 20 October 2023 / Accepted 27 September 2024

## ABSTRACT

**Aims.** We explore the contribution of galaxies, as a function of their stellar mass, to the cosmic star formation history (CSFH). In order to avoid uncertain extrapolations of the infrared luminosity function, which is often polluted by the contribution of starbursts, we base our analysis on stellar mass. Attenuation by dust is accounted for thanks to the combination of deep surveys by *Herschel* and the Atacama Large Millimeter/submillimeter array (ALMA).

**Methods.** We combined for the first time the deepest *Herschel* (GOODS-South, GOODS-North, COSMOS and UDS) and ALMA (GOODS-South) surveys. We constrained the star formation rate (SFR), dust mass ( $M_{\text{dust}}$ ), dust temperature ( $T_{\text{dust}}$ ) and gas mass ( $M_{\text{gas}}$ ) of galaxies as a function of their stellar mass ( $M_{\star}$ ) from  $z \sim 5$  to  $z \sim 0$  by performing a stacking analysis of over 128 000 *Hubble* Space Telescope (HST) *H*-band selected galaxies. We studied the evolution of the star formation efficiency of galaxies as a function of redshift and  $M_{\star}$ .

**Results.** We show that the addition of ALMA to *Herschel* allows us to reach lower  $M_{\star}$  and higher redshifts. We confirm that the SFR- $M_{\star}$  star formation main sequence (MS) follows a linear evolution with a slope close to unity with a bending at the high-mass end at  $z < 2$ . The mean  $T_{\text{dust}}$  of MS galaxies evolves linearly with redshift, with no apparent correlation with  $M_{\star}$ . We show that, up to  $z \sim 5$ , massive galaxies (i.e.  $M_{\star} \geq 10^{10} M_{\odot}$ ) account for most of the total SFR density ( $\rho_{\text{SFR}}$ ), while the contribution of lower-mass galaxies (i.e.  $M_{\star} \leq 10^{10} M_{\odot}$ ) is rather constant. We compare the evolution of star-forming galaxy (SFGs) to the cosmological simulation TNG100. We find that TNG100 exhibits a noticeable difference in the evolution of the CSFH, that is, the marked evolution of massive galaxies found in the observations appears to be smoothed in the simulation, possibly due to feedback that is too efficient. In this mass complete analysis, *H*-dropout (also called HST-dark) galaxies account for  $\sim 23\%$  of the CSFH in massive galaxies at  $z > 3$ . Finally, we find hints that the star formation efficiency of distant galaxies ( $z = 3-5$ ) is stronger (shorter depletion time) as compared to low-redshift galaxies.

**Key words.** galaxies: evolution – galaxies: high-redshift – galaxies: photometry – galaxies: star formation – infrared: galaxies – submillimeter: galaxies

## 1. Introduction

The evolution of the star formation rate density ( $\rho_{\text{SFR}}$ ) over time follows a gradual growth from the Big Bang to cosmic noon (i.e.  $z \sim 2$ ), followed by a continuous decline by a factor  $\sim 10$  to the present day (e.g. Madau & Dickinson 2014; Bouwens et al. 2015b; Liu et al. 2018; Leslie et al. 2020 and Schreiber et al. 2015, hereafter S15). This evolution alone raises many questions about the growth and death of galaxies. For example, the contribution to the  $\rho_{\text{SFR}}$  of galaxies of different stellar masses ( $M_{\star}$ ), or the  $\rho_{\text{SFR}}$  at high redshift (i.e.  $z \geq 4$ ) are still largely uncertain. Our understanding of the high-redshift part (i.e.  $z \geq 4$ ) is mainly built from the ultraviolet (UV), and then deduced by correcting these UV measurements from dust attenuation (e.g. Bouwens et al. 2012b,a; Schenker et al. 2013; Bouwens et al. 2015a,b; Oesch et al. 2018). However, it has recently been claimed, through studies using dust-unbiased measurements (i.e. radio or far-infrared emission), that  $\rho_{\text{SFR}}$  is actually higher, by a factor of two to six, at high redshift (i.e.  $z \geq 4$ , e.g. Rowan-Robinson et al. 2016; Novak et al. 2017;

Lagache 2018; Gruppioni et al. 2020; Khusanova et al. 2021) compared to Madau & Dickinson (2014).

In order to deduce the cosmic star formation history, it is necessary to correctly infer the star formation rate (SFR) of galaxies. It has been shown that the SFR of star-forming galaxies (SFGs) is positively correlated with their stellar mass (e.g. Elbaz et al. 2007; Daddi et al. 2007; Whitaker et al. 2012, 2014; Speagle et al. 2014; S15; Lee et al. 2015; Delvecchio et al. 2021; Leslie et al. 2020; Popesso et al. 2023) with a small scatter of  $\sim 0.2-0.3$  dex (e.g. Elbaz et al. 2007; Whitaker et al. 2012; Speagle et al. 2014; S15). This correlation is called the main sequence (MS; Noeske et al. 2007) of SFGs. Although the specifics of this correlation remain a matter of debate, most recent studies tend to favour a linear MS in logarithmic space with a bending appearing at high  $M_{\star}$  and for  $z \leq 2-3$  (e.g. Whitaker et al. 2014; Magnelli et al. 2014; S15; Delvecchio et al. 2021; Leslie et al. 2020), rather than a strictly linear correlation between SFR and  $M_{\star}$  (e.g. Whitaker et al. 2012; Speagle et al. 2014).

The fraction of star formation obscured by dust is still highly uncertain at high redshift (i.e.  $z \geq 3$ ), as it could be higher up to a factor of ten above unobscured formation (Casey et al. 2018). As the correction factors for dust extinction are quite

\* Corresponding author; [lucas.a.leroy@gmail.com](mailto:lucas.a.leroy@gmail.com)

large (e.g. Calzetti et al. 1994; Madau et al. 1998; Steidel et al. 1999), this leads studies to combine SFR from the UV (uncorrected for dust extinction) and SFR deduced directly from the infrared (IR) (e.g. S15; Delvecchio et al. 2021) when seeking the total SFR and  $\rho_{\text{SFR}}$ . Although most of them are based on *Hubble* Space Telescope (HST) detected catalogues to infer the evolution of the SFR, they do not take into account the contribution of so-called HST-dark galaxies in the study of the  $\rho_{\text{SFR}}$ . *H*-dropout (Wang et al. 2019), HST-dark (Zhou et al. 2020) or optically dark and faint (Gómez-Guijarro et al. 2022a; Xiao et al. 2023b) galaxies, represent massive and highly obscured galaxies usually detected with a low significance or not at all in the optical. Although initially thought to be a marginal population with little effect on the  $\rho_{\text{SFR}}$ , over time it has been shown that such obscured galaxies could significantly contribute to the  $\rho_{\text{SFR}}$  above  $M_{\star} \sim 10^{10.5} M_{\odot}$  (e.g. Wang et al. 2019; Xiao et al. 2023b).

Since the Atacama Large Millimeter/submillimeter array (ALMA) came online in 2013n it has enabled large and deep surveys, with a better resolution, in the millimetre and submillimetre range (e.g. Franco et al. 2018; Gómez-Guijarro et al. 2022a). Studies can now combine higher-quality millimetre and submillimetre measurements with IR measurements. ALMA has proven to be a powerful tool to probe the gas content of galaxies (e.g. Scoville et al. 2014; Groves et al. 2015; Scoville et al. 2016; Schinnerer et al. 2016; Kaasinen et al. 2019; Liu et al. 2019; Millard et al. 2020; Magnelli et al. 2020; Wang et al. 2022), and developing our global understanding of high-redshift galaxies (see Hodge & da Cunha 2020 for a review). Studies have also been using the discrete ALMA archive pointings to study the  $\rho_{\text{SFR}}$  (e.g. Traina et al. 2024).

Spatially resolved studies show that the SFR mostly correlates with the molecular gas ( $\text{H}_2$ ) surface density, and very little with the atomic gas (HI) surface density (e.g. Bigiel et al. 2008; Leroy et al. 2008). The gas content of galaxies is usually probed through CO emission lines because of the difficulty of directly observing the  $\text{H}_2$  content (see Bolatto et al. 2013 for a review). However, other techniques have been developed to infer the gas content of galaxies through the study of multi-wavelength dust spectral energy distribution (SED) fits, and by applying the gas-to-dust ratio to infer gas masses ( $M_{\text{gas}}$ ) from dust masses ( $M_{\text{dust}}$ ; e.g. Leroy et al. 2011; Magdis et al. 2011, 2012; Magnelli et al. 2012; Rémy-Ruyer et al. 2014; Genzel et al. 2015).

By combining the SFR and the gas content of galaxies, we can infer the star formation efficiency ( $SFE \equiv SFR/M_{\text{gas}}$ ) of galaxies, which gives a straightforward indicator of how galaxies form stars at a different moment in cosmic time. The SFE has been shown to evolve quickly at a low redshift (i.e.  $0 \leq z \leq 1$ ), but to be rather constant at higher redshifts (i.e.  $z > 1$ ; e.g. Saintonge et al. 2017; Scoville et al. 2017; Tacconi et al. 2018; Liu et al. 2019; Wang et al. 2022). Another way of looking at this is to use the Kennicutt-Schmidt relation (Kennicutt 1998b). This relation has been shown to follow the power-law correlation between the SFR and gas surface density with a slope of  $\sim 1.0$ – $1.5$  (e.g. Kennicutt 1998b; de los Reyes & Kennicutt 2019; Wang et al. 2022).

In this paper, combining *Spitzer*, *Herschel* and ALMA, we study the evolution of galaxy properties from  $z \sim 0$  to 5 through stacking on four fields: COSMOS, UDS, GOODS-South, and GOODS-North. We try to answer a few open questions about the evolution and global history of galaxies. In particular, we seek to constrain the main sequence,  $\rho_{\text{SFR}}$  at a high redshift (i.e.  $z \geq 4$ ), the different contributions to  $\rho_{\text{SFR}}$  as a function of  $M_{\star}$ , and the impact of the *H*-dropout galaxy population on the

**Table 1.** Number of galaxies in the final sample from each field.

Field	Number of galaxies	Area (arcmin <sup>2</sup> )
GOODS South	26 811	134
GOODS North	35 322	142
COSMOS	34 124	388
UDS	32 695	412
Total	128 952	1077
GOODS-ALMA	8659	69

properties of galaxies as a whole. A similar study using *Spitzer* and *Herschel* measurements has been carried out by S15. However, the addition of ALMA measurements to *Herschel* using stacking has not been carried out yet. The reason is that this requires access to an extensive study programme on ALMA. This has only recently been possible thanks to blind surveys such as the GOODS-ALMA survey (Franco et al. 2018, 2020; Gómez-Guijarro et al. 2022a). A summary of the data used in this study is given in Sect. 2. The stacking method is described in Sect. 3. Section 4 is devoted to the SED fitting procedure. Section 5 reviews the properties of galaxies that can be deduced from this analysis. Sections 6 and 7 present the cosmic star formation history and the cosmic evolution of the gas mass density deduced for this work, respectively. We interpret and discuss our results in a global cosmological context in Sect. 8, and in Sect. 9 we summarise the main results and conclusions.

For this work, we adopted a Salpeter (1955) initial mass function (IMF) and the cosmological parameters ( $\Omega_{\text{M}}, \Omega_{\Lambda}, h$ ) = [0.30, 0.70, 0.70]. We used a factor of 1.7 to convert  $M_{\star}$  and SFR from a Chabrier (2003) to a Salpeter (1955) IMF whenever necessary (e.g. Reddy et al. 2006; Santini et al. 2012; Elbaz et al. 2018; Gómez-Guijarro et al. 2022b). When mentioned, magnitudes are in the AB system, such that  $M_{\text{AB}} = 23.9 - 2.5 \log_{10}(S_{\nu} [\mu\text{Jy}])$ .

## 2. Sample and observations

### 2.1. Sample

In this study, we worked from catalogues of *H*-band selected SFGs on four fields: GOODS-South, GOODS North, COSMOS and UDS. We used the ultra-deep *H*-band catalogue of the CANDELS-HST team (Grogin et al. 2011; Koekemoer et al. 2011) for GOODS-South (Guo et al. 2013), COSMOS (Nayyeri et al. 2017) and UDS (Galametz et al. 2013) fields, while in GOODS-North we use the catalogue from Barro et al. (2019). The  $5\sigma$  limiting magnitude range from  $H \sim 27.4$  to 29.7 for GOODS-South,  $H \sim 27.4$  to 28.8 for COSMOS,  $H \sim 27.1$  to 27.6 for UDS, and  $H \sim 27.8$  to 28.7 for GOODS-North. The photometric redshifts and  $M_{\star}$  of the galaxies in these catalogues were derived in S15 (GOODS-South, COSMOS and UDS) and Barro et al. (2019) (GOODS-North). Photometric data were fitted up to IRAC 4.5  $\mu\text{m}$ , with *EAZY* (Brammer et al. 2008), by assuming a delayed exponentially declining star formation history with the Bruzual & Charlot (2003) stellar population synthesis model. The SFGs are UVJ selected, following the definition from Muzzin et al. (2013). The final number of SFGs in the sample is given in Table 1. *H*-dropout galaxies are by definition not included in these samples, as they are not detected in *H*-band. However, we used the sample from Wang et al. (2019) to add and discuss their impact in any necessary analyses.

## 2.2. Observations

Observations with *Spitzer*-MIPS at 24  $\mu\text{m}$  include maps of the COSMOS field (PI: D. Sanders; [LeFloc'h et al. 2009](#)), the GOODS-South and GOODS-North fields (GOODS Legacy programme; PI: M. Dickinson), and the UDS field (SpUDS *Spitzer* Legacy programme; PI: J. Dunlop). The *Herschel* PACS and SPIRE maps of the four fields come mainly from the CANDELS-*Herschel* programme. The PACS GOODS-North and GOODS-South maps are the combined *Herschel*-PACS data from the PEP ([Lutz et al. 2011](#)) and GOODS-*Herschel* ([Elbaz et al. 2011](#)) programmes, as described in [Magnelli et al. \(2013\)](#). For ALMA, we used the 1.13 mm low-resolution GOODS-ALMA map (sensitivity reaching an average of  $\sigma = 95.2 \mu\text{Jy beam}^{-1}$ , [Gómez-Guijarro et al. 2022a](#)), within the GOODS-South field with the 1''1 circularised point spread function (PSF) full width half maximum (FWHM). We note that there is a global and local offset between the position of sources in the ALMA and HST images ([Franco et al. 2018, 2020](#)). We corrected them using the offsets provided by [Franco et al. \(2020\)](#).

## 3. Stacking

Our method consists in stacking several images of galaxies on top of each other in order to increase the overall signal-to-noise ratio (S/N). The result of a stacking procedure is the mean or median flux for all galaxies stacked together. The main advantage of this method is to be able to recover reliable flux measurements for populations of galaxies whose S/N is too low when studied individually. In practice, this means being able to measure the properties of galaxies with lower  $M_\star$  and higher redshift. An effective way to take advantage of the stacking method is to group galaxies into sub-populations with similar properties. To this end, in this study, we stack our sample of galaxies over different redshift and  $M_\star$  bins. The total numbers of galaxies per bin are displayed in the Table 2.

In order to access the completeness of the catalogues that we stacked in this study, we followed the method of [S15](#). In summary, we assumed, for each redshift bin, that the observed luminosity at 1.6  $\mu\text{m}$  ( $L_{1.6\mu\text{m}}^{\text{obs}}$ ) can be related to  $M_\star$  simply by  $M_\star = C \times L_{1.6\mu\text{m}}^{\alpha}$  with a scatter. The relation and dispersion are fitted from the *H*-band photometry of the catalogues. We derive the completeness from Monte Carlo simulations by generating, from a uniform redshift distribution in the redshift bin and a given  $M_\star$ , the corresponding  $L_{1.6\mu\text{m}}^{\text{obs}}$  taking into account the dispersion, and comparing it to the corresponding *H*-band detection limit for each catalogue. We assume that catalogues are complete when the completeness is above 90% ([S15](#)). The corresponding  $M_\star^{90\%}$ , above which all catalogues are assumed to be mass complete, are listed in Table 3. All but two of the bins which yield significant detection (i.e. with  $S/N \geq 3\sigma$  in at least one band from 24  $\mu\text{m}$  to 1.13 mm) considered in this study are mass complete: for  $9.5 \leq \log_{10}(M_\star/M_\odot) \leq 10.0$  at  $2.3 \leq z \leq 3.1$ , and  $10.0 \leq \log_{10}(M_\star/M_\odot) \leq 10.5$  at  $3.1 \leq z \leq 3.9$ . The latter bin is, however, almost mass complete as  $M_\star^{90\%} = 10^{10.1} M_\odot$  and is consistent with a  $M_\star^{90\%} = 10^{10} M_\odot$  within the uncertainties of the method used to compute the completeness.

The completeness computed in this paper is determined assuming that the *H*-band is a good proxy for the stellar mass. This hypothesis may however become less robust at high redshifts (i.e.  $z > 3$ ). To assess whether our derivation of the completeness limit remains relevant for the redshift range  $3 < z < 5$ ,

**Table 2.** Number of galaxies in each bin of  $M_\star$  and redshift.

Redshift	$\log_{10}(M_\star/M_\odot)$				
	8.5	9.5	10	10.5	11
0.1	954 86	182 10	125 13	56 3	11 1
0.4	3,122 298	746 67	497 56	239 21	54 7
0.7	5,686 403	1,449 98	872 71	458 21	95 5
1.0	5,325 382	1,425 115	782 70	403 44	106 10
1.3	12,091 905	2,742 188	1,389 86	611 47	166 10
1.8	5,365 410	1,770 145	1,015 83	590 47	261 17
2.3	5,656 475	2,459 187	1,250 125	520 55	182 16
3.1	2,352 352	1,305 115	656 55	173 7	53 3
3.9	1,063 131	692 69	406 35	133 10	43 2
5.0					

**Notes.** Number of galaxies in each bin of  $M_\star$  and redshift for the stacking in the *Herschel* 100  $\mu\text{m}$ , 160  $\mu\text{m}$ , 250  $\mu\text{m}$ , 350  $\mu\text{m}$  and 500  $\mu\text{m}$  bands, and for the stacking in the ALMA band are displayed in the top and bottom row of each cell, respectively. The two axes represent the boundaries of the bins in redshift and  $M_\star$ . Bins with at least a FIR, and a sub-millimeter stack detection (i.e. with a  $S/N \geq 3\sigma$  in at least one band) are highlighted in faded blue and red, respectively. Bins that are mass complete are noted with a black bar.

**Table 3.** Completeness limits.

Redshift bin	$\log_{10}(M_\star^{90\%}/M_\odot)$	<i>H</i> -dropout/arcmin <sup>2</sup>
$0.1 \leq z \leq 0.4$	7.5	0
$0.4 \leq z \leq 0.7$	7.9	0
$0.7 \leq z \leq 1.0$	8.1	0
$1.0 \leq z \leq 1.3$	8.2	0
$1.3 \leq z \leq 1.8$	8.6	0
$1.8 \leq z \leq 2.3$	9.2	$1.67 \times 10^{-3}$
$2.3 \leq z \leq 3.1$	9.7	$1.17 \times 10^{-2}$
$3.1 \leq z \leq 3.9$	10.1 <sup>(*)</sup>	$3.67 \times 10^{-2}$
$3.9 \leq z \leq 5.0$	10.0	$3 \times 10^{-2}$

**Notes.** Stellar mass for which we obtain a completeness of 90% in the *H*-band catalogues used in this work; and the number of *H*-dropout/arcmin<sup>2</sup> that can be found at these redshifts from [Wang et al. \(2019\)](#). <sup>(\*)</sup>Consistent with a  $M_\star^{90\%} = 10^{10} M_\odot$  within the uncertainties of the method used to compute the completeness.

we looked at a deep JWST-NIRCAM survey of the EGS field ([Gómez-Guijarro et al. 2023](#)). We found that 89.7% of the galaxies with  $\log_{10}(M_\star/M_\odot) > 10$  (209 galaxies detected, 233 overall) detected in the JWST bands are also detected in the HST *H*-band (the  $5\sigma$  limiting magnitude is 27.37 in CEERS – see [Gómez-Guijarro et al. 2023](#), Sect. 2.2 – that is comparable to 27.4, 27.4, 27.1, 27.8 in the GOODS-South, COSMOS, UDS GOODS-North fields respectively). This is consistent with our 90% completeness of  $\log_{10}(M_\star^{90\%}/M_\odot) \sim 10$  for  $3.1 \leq z \leq 5$  and  $3.9 \leq z \leq 5.0$ , which gives us confidence that our sample is not strongly biased.

We choose to use mean stacking for our work, rather than median stacking. Indeed, although median stacking better

**Table 4.** Clustering bias correction factors.

Wavelength ( $\mu\text{m}$ )	Correction
100	0% <sup>+7%</sup> <sub>-7%</sub>
160	3% <sup>+9%</sup> <sub>-8%</sub>
250	8% <sup>+12%</sup> <sub>-8%</sub>
350	13% <sup>+12%</sup> <sub>-10%</sub>
500	25% <sup>+19%</sup> <sub>-18%</sub>

**Notes.** The clustering bias correction factors in this Table are from S15.

suppresses secondary sources in the stacked image (i.e. bright sources which are close to the stacked target and might appear in the final stacked image), it has been shown that it also yields to systematically biased measurements at low S/N or in the presence of a flux distribution skewed towards low or high values (White et al. 2007; S15). It comes from the fact that the median ( $\langle \cdot \rangle$ ) is not a linear operation:  $\langle a + b \rangle \neq \langle a \rangle + \langle b \rangle$ . As shown in S15, correcting for these systematic biases involves making strong assumptions on the actual flux distribution of the stacked sources, which is non-trivial and very uncertain. On the contrary, the mean stack gives us access to the total flux in the stack, which we currently miss to infer  $\rho_{\text{SFR}}^{\text{IR}}$ .

During the stacking procedure, it is possible to treat detected and undetected galaxies separately. The main way to do this is to stack only the undetected sources on the residual map, and then add them to the fluxes of the detected galaxies via a weighted mean (e.g. Magnelli et al. 2009). Although this method reduces the confusion noise of faint sources and removes most of the contamination from bright neighbours, it can also introduce some biases (S15). Following S15, we have therefore chosen to treat detected and undetected sources in the same way for consistency (i.e. directly from the image and not from the residual). The contamination of bright neighbours is dealt with later in the study (see Sect. 3.2).

We noticed the presence of a global background gradient in most *Herschel* images. In order to deal with it, we have decided to rotate the stacked postage stamp images successively by  $90^\circ$ .

The resulting flux density in each stacked image was obtained via standard aperture measurements or PSF fitting method (see detail for each wavelength in Sects. 3.1, 3.2 and 3.3). The corresponding S/N was then derived using a simple Monte Carlo approach, that is,  $S/N = S_v/\sigma_{\text{MC}}$ , where  $S_v$  is the flux density measured at the centre of the stacked image within an aperture of radius,  $r$ , while  $\sigma_{\text{MC}}$  is the standard deviation of the signal in 100 circles, of the same radius  $r$ , randomly positioned on the edges of the stacked stamp. Depending on the method used to compute the flux,  $r$  corresponds to either the radius used for the aperture measurement method, or the radius within which the PSF fitting method was performed.

In addition to this photometric noise, there is an uncertainty in the recovered flux density due to the intrinsic dispersion of the underlying stacked population (i.e. all galaxies in the stack do not have the exact same flux density). This flux dispersion can be quantified using a bootstrap analysis (e.g. S15). The method consists in repeating several times the full stacking and flux density measurement process, picking, for each realisation, the galaxies from the origin sample, with replacement, until the number of galaxies in the original sample is reached. In this study, we compute a 100 runs. The error is then deduced from the standard deviation of the computed flux densities for this bin.

### 3.1. Spitzer

As the  $24\ \mu\text{m}$  MIPS-*Spitzer* map ( $\sim 5''.8$  PSF FWHM) has a better resolution than *Herschel*-PACS ( $\sim 7''$  and  $\sim 11''.3$  PSF FWHM at  $100\ \mu\text{m}$  and  $160\ \mu\text{m}$ , respectively), we do not expect the clustering bias (see Sect. 3.2) to be a dominant effect on the stacked stamp. The fluxes are calculated using a classical aperture photometry method. The aperture is chosen to be 4 pixels (i.e.  $4''.8$ ) in radius, which contains about  $\sim 60\%$  of the total flux for a point source.

### 3.2. Herschel

When stacking galaxies, a ‘clustering bias’ can occur due to the neighbouring galaxies of the main stacking targets. Such contamination from neighbouring sources can become significant when the size of the PSF becomes comparable to the typical cluster length of SFGs. It has been shown that the clustering bias has a non-negligible impact on the results when stacking galaxies in the *Herschel* bands (e.g. Bavouzet et al. 2008; Béthermin et al. 2010; Kurczynski & Gawiser 2010; Bourne et al. 2012; Béthermin et al. 2012; Viero et al. 2013; Béthermin et al. 2014; S15; Béthermin et al. 2015; Delvecchio et al. 2021) due to the large PSF of *Herschel*. Correcting for this bias is crucial to accurately measure the peak of SED of SFGs, as this bias tends to cause the fluxes of *Herschel* to be increasingly overestimated with increasing wavelength (as the size of the PSF increases with the wavelength).

To mitigate this clustering signal contamination, we choose to follow the method presented in S15. It consists in fitting only a PSF and a local background term:

$$S(x, y) = \varphi \times \text{PSF}(x, y) + \varepsilon, \quad (1)$$

where  $\varphi$  and  $\varepsilon$  are the normalisation of the source flux and background, respectively. The fit is performed on a fixed aperture of radius of  $0.9 \times \text{FWHM}$ , as this was found to minimise the clustering contamination to  $\varphi$  (S15). The clustering bias signal is in this case largely included in the background term. Nevertheless, even with this radius the contamination of  $\varphi$  is not null and we still need to apply a correction for what remains of the clustering signal in the mean flux term ( $\varphi$ ). The correction factors were calculated by simulating the stacking procedure on mock images (see S15 for more details). The correction factors of S15 are listed in Table 4. For the PACS maps, we used the truncated PSF derived from Vesta, while for *Herschel*-SPIRE the PSF is assumed to be Gaussian with a FWHM of  $18''.15$ ,  $25''.15$  and  $36''.3$  respectively at 250, 350 and  $500\ \mu\text{m}$  according to Griffin et al. (2010) (see also Shirley et al. 2021).

Finally, for *Herschel*-PACS, it was shown (Popesso et al. 2012; Magnelli et al. 2013) that the high-pass filter data reduction technique that was used to remove low-frequency noise in the maps could induce an underestimation of the photometric measurements of the unmasked faint sources. It was shown in Popesso et al. (2012) that a correction factor of 17%, and 10% should be taken into account when stacking undetected faint, and detected, sources in the PACS maps ( $100\ \mu\text{m}$  and  $160\ \mu\text{m}$ ), respectively.

### 3.3. ALMA

For the purpose of this study, we decided to work in the image plane when dealing with ALMA. Another option would have been to work in the  $uw$  plane (i.e. in Fourier space), as ALMA provides us with measurements in the  $uw$  space. Although this

may give more robust results, since we do not need to go through the conversion between the  $uv$  plane and the image plane first, it is very computationally intensive. Furthermore, on few test bins we only observed a 10% difference between fluxes calculated from stacks in the  $uv$  and image planes. We decided not to stack directly in the  $uv$  plane but in the image plane in order to save computing time as this should not impact the results of this study.

As the ALMA 1.1mm map has a much better resolution ( $\sim 1''.1$  PSF FWHM; which is comparable to typical sizes of individual SFGs, [Suess et al. 2019](#); [Wang et al. 2022](#)) than *Herschel*-PACS ( $\sim 7''$  and  $\sim 11''.3$  PSF FWHM at  $100\mu\text{m}$  and  $160\mu\text{m}$ , respectively), the clustering bias is negligible on the stacked stamp. Due to the high angular resolution of the ALMA data, and to possible small offsets between the optical centroids (on which our stack positions are based) and millimetre centroids, the stacked images are extended on scales larger than that of the ALMA PSF. However, this has no significant impact of the measure flux density via the aperture photometry method. To measure ALMA flux densities, we applied an aperture photometry within a radius of  $1''$ , containing 85% of the total flux density (i.e.  $S_{\text{AP}}^{\text{PSF}}/S_{\text{tot}}^{\text{PSF}} = 0.85$ ).

## 4. SED fitting

### 4.1. SED fitting procedure

One of the objectives of this paper is to retrieve several properties such as SFR, dust temperature ( $T_{\text{dust}}$ ), dust masses and gas masses from our far-infrared (FIR) to submillimetre stacking analysis, and compare them to the literature. We performed an SED fit from the measured fluxes for each bin of redshift and  $M_{\star}$ . Because we do not have many points on the FIR at low  $M_{\star}$  and high redshifts, we used the library from [Schreiber et al. \(2018, hereafter S18\)](#), which is well suited for our study given its small number of free parameters. Choosing a model with more parameters, such as [Draine & Li \(2007\)](#) and [Draine et al. \(2014\)](#), would imply fixing parameters on some bins. The library of S15 is calibrated on galaxies from  $z = 0.5$  to  $z = 4$ , which allows to have a realistic SED while reducing the number of free parameters. We have assumed a form of SED that we would expect for the main sequence galaxies, and no active galactic nuclei (AGN) contribution to the stacked SED.

To fit the SED, we only considered fluxes with  $S/N \geq 3$ . Stacked flux measurements with  $S/N < 3$  are replaced by conservative  $5\sigma$  upper limits. At  $z > 4$  MIPS-24 stacked fluxes, if available, have been transformed to  $5\sigma$  upper limits as they are no longer dominated by dust and polycyclic aromatic hydrocarbon (PAH) emissions, but rather by stellar emission. From these SED fits, we measured the corresponding infrared luminosity ( $L_{\text{IR}}$ ) by integrating the best-fit SED in the range 8–1000  $\mu\text{m}$  rest frame. The error on  $L_{\text{IR}}$  was obtained by varying their stacked photometry randomly within their uncertainties.

The  $T_{\text{dust}}$  can be defined in different ways (i.e. weighted by mass or luminosity). The luminosity weighted  $T_{\text{dust}}^{\text{L}}$  for the library from S18, is calculated from a grey body of effective emissivity  $\beta = 1.5$  (S18). This means that  $T_{\text{dust}}^{\text{L}}$  follows Wien's law (see Eq. (2)).

$$T_{\text{dust}}^{\text{L}}[\text{K}] = 2.897 \times 10^3 / (\lambda_{\text{max}}[\mu\text{m}]), \quad (2)$$

where  $\lambda_{\text{max}}$  is the wavelength corresponding to the peak of  $\lambda^{\beta} L_{\lambda}$ . A mass weighted  $T_{\text{dust}}^{\text{M}}$  was also calculated for each template by mass weighted averaging each individual template of [Galliano et al. \(2011\)](#) (see S18 for more details). In the cases where we only had one or two points to perform our fit (i.e. the

SED peak was not well defined in this case), we chose to restrict the  $T_{\text{dust}}^{\text{M}}$  during the fit to the  $T_{\text{dust}}^{\text{M}}$  evolution from  $S18 \pm 10$  K. This is a reasonable way to reduce the error on the deduced  $L_{\text{IR}}$  by slightly restricting the  $T_{\text{dust}}^{\text{M}}$  to reasonable values. Our best-fits of SEDs are displayed in Fig. 1.

The templates from the library of S18 are built using the amorphous carbon model from [Galliano et al. \(2011\)](#). This differs from the model from [Draine & Li \(2007\)](#) which takes into account amorphous silicate and graphite grains. This change of model was shown in [Galliano et al. \(2011\)](#) to lower the tension, in the Large Magellanic Cloud and the Milky Way, between the observed dust-to-gas ratio and the stellar abundances. The differences between the two models are mainly reflected in the different emissivity. The choice of a different emissivity does not affect the dust temperature or the  $L_{\text{IR}}$ , as these properties correspond to the peak and the area under the SED, respectively. However, the  $M_{\text{dust}}$  deduced from [Galliano et al. \(2011\)](#) is about a factor of 2 lower than those deduced from [Draine & Li \(2007\)](#), without affecting the  $M_{\text{gas}}$  as the factor cancels out when converting  $M_{\text{dust}}$  to  $M_{\text{gas}}$ . The debate on the composition of dust grains goes far beyond the scope of this study. We chose to work with  $M_{\text{dust}}$  derived from the [Draine & Li \(2007\)](#) model, as it will ease comparison with the literature, which is widely based on the later model. In practice, we have re-fitted the model library of S18 with a [Draine & Li \(2007\)](#) model to associate the  $M_{\text{dust}}$ . The impact on  $M_{\text{gas}}$  of choosing an amorphous carbon model instead of an amorphous silicate and graphite grains model is briefly discussed at the end of the Sect. 5.3.

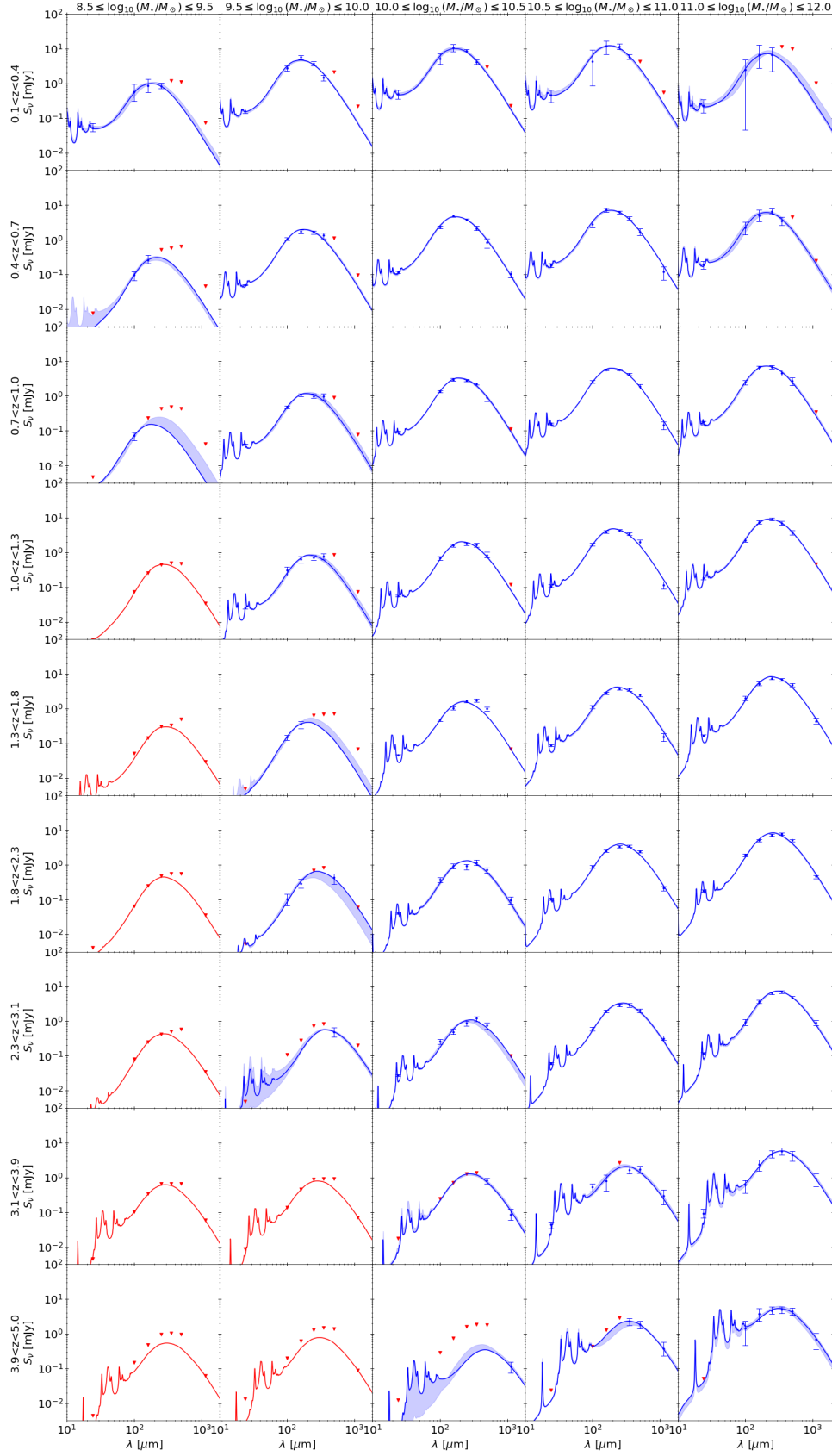
Overall we verified that fitting our stacked flux with the S18 or [Draine & Li \(2007\)](#) libraries has little to no impact on our results, as the differences on some key properties such as  $L_{\text{IR}}$ ,  $T_{\text{dust}}$  and  $M_{\text{dust}}$  (once corrected for the emissivity chosen in S18) are quite small:  $-7\%_{-4\%}^{+7\%}$  for  $L_{\text{IR}}$ ,  $1\%_{-6\%}^{+12\%}$  for  $T_{\text{dust}}$  and  $4\%_{-17\%}^{+33\%}$  for  $M_{\text{dust}}$ . An essential consequence of this observation is that the models of S18 reproduce well the global SED shape of stacked galaxies. This reinforces the choice we made to use the S18 templates instead of [Draine & Li \(2007\)](#) for this study.

### 4.2. Adding ALMA to Herschel

The addition of ALMA to this study results in an improvement of the SED at high redshift (i.e. at  $3.1 \leq z \leq 5.0$ ). More specifically, it allows to get an ALMA measurement at  $10.0 \leq \log_{10}(M_{\star}) \leq 10.5$  for  $3.1 \leq z \leq 3.9$  and  $3.9 \leq z \leq 5$  that provides information on some of the properties of galaxies (i.e.  $L_{\text{IR}}$  and  $M_{\text{gas}}$ ) at these redshifts and  $M_{\star}$  instead of what would have just been an upper limit on these properties. At  $10.5 \leq \log_{10}(M_{\star}) \leq 11.0$  for  $3.1 \leq z \leq 3.9$  and  $3.9 \leq z \leq 5$ , and  $11.0 \leq \log_{10}(M_{\star}) \leq 12.0$  for  $3.9 \leq z \leq 5$ , this also allows the SED peak to be constrained more effectively, as it is only weakly constrained, if at all, by the *Herschel* data alone. The SED peak is crucial for determining properties such as  $T_{\text{dust}}^{\text{M}}$  and  $M_{\text{gas}}$ . Elsewhere it also provides an extra point or upper limit that significantly reduces the uncertainty in the properties inferred from the SED fit. These extra ALMA points are thus decisive for this study, which aims to probe the properties of galaxies over a wide dynamic range, extending up to  $z \sim 5$ .

### 4.3. Simulations to correct for averaging biases

Because a mean stacking procedure is luminosity weighted, it can have non-linear effects on the shape of the resulting SED (e.g. [Elbaz et al. 2011](#); S15; S18). These effects include the widening of the FIR bump and a bias of the peak towards warmer



**Fig. 1.** Best-fit SED for each bin of redshift and stellar mass. Blue dots correspond to the flux measurements, red triangles represent the  $5\sigma$  upper limits. Blue line is the best-fit SED, the blue shaded area shows the 68% uncertainty of the fit. Red line is the SED maximising  $L_{\text{IR}}$  in case only upper limits are available.

$T_{\text{dust}}^L$ . It is mainly a result of mixing galaxies of different redshifts and  $T_{\text{dust}}^L$  (e.g. Elbaz et al. 2011; S15; S18). The broadening of the SED increases the difficulty of determining precisely  $T_{\text{dust}}^L$  (S18). For the rest of the paper, we chose to work with mass-weighted dust temperatures, and we will refer to  $T_{\text{dust}}^M$  simply as  $T_{\text{dust}}$ . To ensure that our conclusions are not biased by these effects, we performed simulations to identify any systematic bias due to our stacking procedure.

To this end, for each bin of  $M_*$  and redshift, we simulated the biases coming from stacking starting from the distribution of the galaxies in the bin. This has the advantage of taking into account the specificity of the  $M_*$  and redshift distributions within our bin. For each set of galaxies, we created a mock counterpart in order to compare the properties resulting from the stacking with those expected.

To each galaxy of mass  $M_*$ , we assigned the SFR starting from  $M_*$  and following  $SFR = R_{\text{SB}} \times SFR_{\text{MS}}$ . The  $SFR_{\text{MS}}$  was calculated from the main sequence trend found in this work (See Eq. (12) and parameters Table 5) as a function of redshift and  $M_*$ .  $R_{\text{SB}}$  represents the starburstiness and is defined by  $R_{\text{SB}} = SFR/SFR_{\text{MS}}$ . We want to generate both mock main sequence galaxies and starburst galaxies. It was shown in S15 that both the main sequence width ( $\sim 0.3$  dex, see S15) and the starburst fraction do not evolve with redshift and  $M_*$ . We can therefore reasonably assume that the distribution function of  $R_{\text{SB}}$  does not vary (S15). This assumption still allows the luminosity functions to be reconstructed properly (e.g. Sargent et al. 2012; S15). Following Sargent et al. (2012), we have modelled the probability density function of  $R_{\text{SB}}$  by a double log-normal distribution (see Eq. (3)):

$$\Phi_{R_{\text{SB}}}(x) = \frac{1 - f_{\text{SB}} - f_{\text{miss}}}{\sqrt{2\pi}\sigma_{\text{MS}}} \exp\left(\frac{-\log_{10}(x/x_0)^2}{2\sigma_{\text{MS}}^2}\right) + \frac{f_{\text{SB}}}{\sqrt{2\pi}\sigma_{\text{SB}}} \exp\left(\frac{-\log_{10}(x/B_{\text{SB}})^2}{2\sigma_{\text{SB}}^2}\right), \quad (3)$$

where  $f_{\text{SB}}$  is the fraction of starbursts,  $f_{\text{miss}}$  is the fraction of galaxies missed by such distribution (neither starburst nor main sequence galaxies),  $\sigma_{\text{MS}}$  and  $\sigma_{\text{SB}}$  are the widths of the main sequence and starburst distributions,  $B_{\text{SB}}$  is the median multiplicative boost of star formation that can be expected for a starburst compared to a main sequence galaxy (i.e. the median of starburst galaxies), and  $x_0$  is the median  $R_{\text{SB}}$  of main sequence galaxies. We note that with this parametrisation, we expect  $f_{\text{miss}}$  and  $x_0$  to be close to 0 and 1 respectively, by construction. We have chosen here to use the parametrisation of S15:  $\sigma_{\text{MS}} = \sigma_{\text{SB}} = 0.31 \pm 0.02$  dex,  $f_{\text{SB}} = 3.3\% \pm 1.5\%$ ,  $B_{\text{SB}} = 5.3 \pm 0.4$ ,  $f_{\text{miss}} = 0\% \pm 2\%$ , and  $x_0 = 0.87 \pm 0.04$ .

To each galaxy we assign  $L_{\text{IR}}$ , deduced from the SFR by subtracting the UV SFR assuming the UV dust attenuation ( $A_{\text{UV}}$ ) derived from  $M_*$  as in Pannella et al. (2015) (see Eqs. (4) and (5)) and Kennicutt (1998a) (see Eq. (6)):

$$A_{\text{UV}} = 1.6 \times \log_{10}(M_*) - 13.5, \quad (4)$$

$$SFR_{\text{IR}} = SFR - SFR_{\text{UV}} \quad (5)$$

where  $SFR_{\text{UV}} = SFR \times 10^{-0.4 \times A_{\text{UV}}}$ ,

$$L_{\text{IR}}[L_{\odot}] = 5.8 \times 10^9 \times SFR_{\text{IR}}[M_{\odot}\text{yr}^{-1}]. \quad (6)$$

Then  $T_{\text{dust}}$  was calculated using the best-fit of our work (see Eq. (7)) for the main sequence trend, and we followed

Magnelli et al. (2014) (see Eq. (9)) to take into account the impact on  $T_{\text{dust}}$  of the distance of the mock galaxy from the main sequence ( $R_{\text{SB}}$ ). The SED, for each mock galaxy, was then calculated using  $T_{\text{dust}}$ ,  $L_{\text{IR}}$  and redshift with the template library from S18 assuming a contribution of PAH molecules  $f_{\text{PAH}} \equiv M_{\text{dust}}^{\text{PAH}}/M_{\text{dust}}$ . The value of  $f_{\text{PAH}}$  was set to follow a Gaussian distribution with a mean of 0.039 and a scatter 2.5/100 (S18, roughly for main sequence galaxies). Next, the  $M_{\text{dust}}$  were obtained from the selected SED template and the  $M_{\text{gas}}$  were calculated according to the different methods presented in Sect. 5.3. The individual SEDs were then stacked using a mean stacking method. Fluxes at 24  $\mu\text{m}$ , 100  $\mu\text{m}$ , 160  $\mu\text{m}$ , 250  $\mu\text{m}$ , 350  $\mu\text{m}$ , 500  $\mu\text{m}$  and 1130  $\mu\text{m}$  were deduced from the stacked SED. The fluxes were then fitted with the template library from S18. The properties of the stacked SED were deduced using the method presented in the corresponding section of this paper: see Sect. 5.1 for  $T_{\text{dust}}$ , Sect. 5.2 for  $L_{\text{IR}}$  and Sect. 5.3 for  $M_{\text{gas}}$ . Potential biases were then investigated by comparing the actual average properties with those deduced from our stacking analysis.

The relative differences between the actual average properties with those deduced from our stacking analysis are quite small:  $-1\%_{-4\%}^{+2\%}$  for  $L_{\text{IR}}$ ,  $-6\%_{-2\%}^{+1\%}$  for  $T_{\text{dust}}$  and  $7\%_{-4\%}^{+2\%}$  for  $M_{\text{dust}}$ . As a result, we find no clear evidence of significant averaging bias and thus decided not to apply any correction.

#### 4.4. Active galactic nuclei bias

It has been reported that AGN can have a major contribution to the total outgoing light of a galaxy (e.g. Hao et al. 2005; Richards et al. 2006). However, most of this emission is radiated at wavelengths shorter than 24  $\mu\text{m}$  and thus will not affect our FIR measurement. Most extreme AGN may still have an impact on the mid-to-far infrared ratio (in particular 24  $\mu\text{m}$  in our case), but should not affect the FIR colour compared to normal star-forming galaxies (Hatiminaoglou et al. 2010). We checked for any AGN contribution by fitting our SED with a combination of Draine & Li (2007) dust model, and Fritz et al. (2006) AGN model. No conclusive evidence for a major contribution (i.e.  $L_{\text{IR}}^{\text{AGN}}/L_{\text{IR}}^{\text{TOT}} \geq 10\%$ ) from AGNs to the rest-frame FIR was found in any of our  $M_*$  and redshift bins.

## 5. Stacked galaxies properties

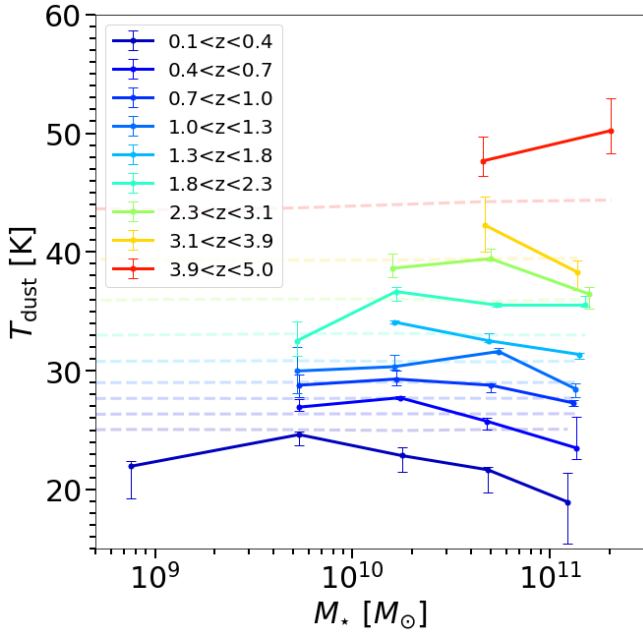
### 5.1. Dust temperature

In this section, we examine the evolution of  $T_{\text{dust}}$ . In Fig. 2, we display the corresponding  $T_{\text{dust}}$  for our SED as a function of redshift and  $M_*$ . We do not see a significant dependence of  $T_{\text{dust}}$  on  $M_*$ . It seems that, overall, the  $T_{\text{dust}}$  of a main sequence galaxy is mainly determined by its redshift, independently of its  $M_*$  (see also Magdis et al. 2012; Magnelli et al. 2014).

The temperature of the cosmic microwave background (CMB) ranges from 2.73 K at  $z = 0$ , up to 16.4 K at  $z = 5$ , and, thus, it could become a significant source of heating at high redshift. We, however, verified, following da Cunha et al. (2013) that this is not the case. Indeed, the observed  $T_{\text{dust}}$  of galaxies at high redshift is significantly higher than that of the CMB.

Because our results come from  $H$ -band selected galaxies, our stacking analysis does not take into account the  $H$ -dropout galaxies (Wang et al. 2019). There are 63  $H$ -dropout galaxies that have been detected in Wang et al. (2019) over  $\sim 600$  arcmin<sup>2</sup>. In this work, we have 1464 galaxies with  $M_* \geq 10^{10} M_{\odot}$  and  $z \geq 3$  over  $\sim 1077$  arcmin<sup>2</sup>. The  $H$ -dropout would then only account for  $\sim 7.2\%$  of the total sample. We re-fitted the stack





**Fig. 2.**  $T_{\text{dust}}$  as a function of  $M_*$  colour-coded by the redshift bin. Average trend of  $T_{\text{dust}}$  as a function redshift from S18, is shown as a reference in faded coloured dashed line.

of  $H$ -dropout from Wang et al. (2019) with S18 templates and deduced a  $T_{\text{dust}}^{H\text{-drop}} = 37.34^{+1.2}_{-1.27}$  K. The contribution of  $H$ -dropout from Wang et al. (2019) was then added to our last bin of redshift (i.e.  $3.9 \leq z \leq 5$ ).

In Fig. 2, our result seems to be globally consistent with the trend of S18. We see only slight evidence that the highest  $M_*$  bin ( $11 \leq \log_{10}(M_*) \leq 12$ ) could be cooler than lower-mass galaxies at a fixed redshift. This effect can be observed for  $z \leq 2.5$ . This could simply show that high-mass galaxies are actually starting to slowly reduce their SFE on their way to become quiescent. This is particularly apparent in our first redshift bin ( $0.1 \leq z \leq 0.4$ ) in which the bending of the main sequence is also the strongest (see Sect. 5.2; S15).

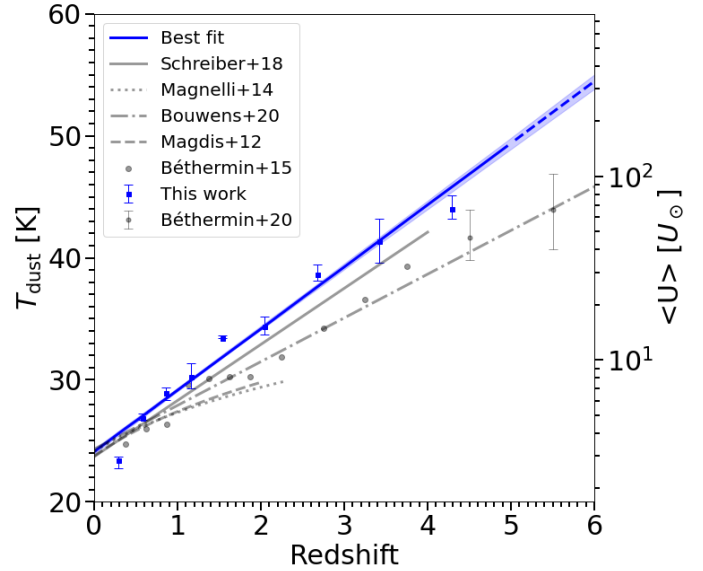
In Fig. 3, we display the mean  $T_{\text{dust}}$  over each redshift bin, weighted by the number of galaxies in each bin of  $M_*$ , as a function of redshift. Our analysis suggests a linear evolution of  $T_{\text{dust}}$  as a function of redshift. Our best linear fit, for  $0.1 \leq z \leq 5$ , follows:

$$T_{\text{dust}}[\text{K}] = 34.20^{+0.14}_{-0.15} + 5.06^{+0.03}_{-0.03}(z - 2). \quad (7)$$

We compare our results with those in the literature: S18, Magnelli et al. (2014), Bouwens et al. (2020), Magdis et al. (2012) and Béthermin et al. (2015). We also re-fitted the two stacks of Béthermin et al. (2020) using the template library from S18. For consistency, we converted all  $T_{\text{dust}}^{\text{L}}$  to  $T_{\text{dust}}^{\text{M}}$  when comparing our results with the literature (i.e. for Magnelli et al. 2014; Bouwens et al. 2020). To do this, we adopt the conversion factor, between mass and light weighted dust temperature, given by S18:

$$T_{\text{dust}}^{\text{M}}[\text{K}] = 0.91 \times T_{\text{dust}}^{\text{L}}[\text{K}]. \quad (8)$$

This conversion factor represents the average conversion factor between  $T_{\text{dust}}^{\text{M}}$  and  $T_{\text{dust}}^{\text{L}}$  for each individual template. Some studies (i.e. Magdis et al. 2012; Béthermin et al. 2015) consider the mean starlight heating rate ( $\langle U \rangle$ ) (Draine & Li 2007; Draine et al. 2014) instead of  $T_{\text{dust}}$  as defined here. To compare



**Fig. 3.**  $T_{\text{dust}}$  and  $\langle U \rangle$  as a function of redshift. The blue dots represent the  $T_{\text{dust}}$  of this work, the blue line is the best-fit (up to  $z = 5$ ), the dashed blue line is the best-fit extrapolation (for  $z \geq 5$ ), and the blue shaded area represents the 68% uncertainty of the fit. From the literature: S18 (grey solid line), Magnelli et al. (2014) (grey dotted line) and Bouwens et al. (2020) (grey dash-dotted line) converted using Eq. (8). Magdis et al. (2012) (grey, dashed line) and Béthermin et al. (2015) (grey dots) converted using Eq. (9). We also re-fitted the two stacks from Béthermin et al. (2020) (grey dotted error bars) using template library from S18.

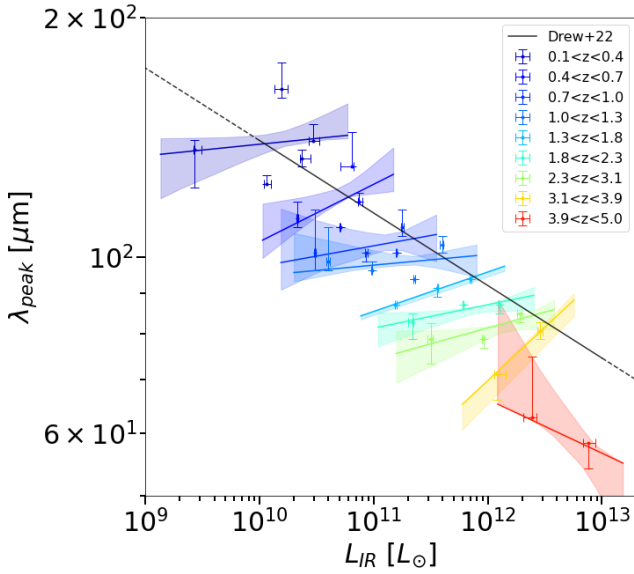
our results with those, we have chosen to use the  $\langle U \rangle$ -to- $T_{\text{dust}}$  conversion formula of S18 (see Eq. (9)):

$$T_{\text{dust}}[\text{K}] = (\langle U \rangle [U_{\odot}])^{1/5.57} \times 18.2 \text{ K}. \quad (9)$$

As  $\langle U \rangle$  is only a proxy of  $T_{\text{dust}}$ , any comparison between the two quantities should be essentially qualitative.

Our  $T_{\text{dust}}$  as a function of redshift is consistent within the uncertainties with S18, extended to  $z = 5$ . We find no clear evidence for the softening of the  $T_{\text{dust}}$  as the redshift increases reported by Magnelli et al. (2014), Magdis et al. (2012). This could stem from the lack of clustering bias correction in their study, as previously reported in S18. Clustering bias correction is a quite important step when stacking in *Herschel* passbands. Especially in the SPIRE wavelengths, as this effect can account for up to 50%, of the total signal, on average at  $500 \mu\text{m}$  (e.g. S15; Béthermin et al. 2015; Delvecchio et al. 2021). Ignoring this could result in a cooler SED and a lower  $T_{\text{dust}}$ . Comparing our trend to other works by Béthermin et al. (2015, 2020), Bouwens et al. (2020), all agree on a linear trend. The differences with our trend may again stem from the way in which the correction for clustering bias is handled in the two works, as it is essential to correctly determine  $T_{\text{dust}}$ .

Another way to look at the redshift evolution of the FIR SED is to consider the wavelength of the IR SED bump ( $\lambda_{\text{peak}}$ ) as a function of  $L_{\text{IR}}$  (see Fig. 4). We observe no dependence of  $\lambda_{\text{peak}}$  as a function of  $L_{\text{IR}}$  at fixed redshift, up to  $z \approx 4$ . Only our last two redshift bins (i.e.  $3.1 \leq z \leq 3.9$  and  $3.9 \leq z \leq 5.0$ ) shows some evolution. However, for the same reason as above, this could result from a bias due to selection effects, as well as the fact that we only have two points in these redshift intervals. We conclude that  $\lambda_{\text{peak}}$  of main sequence galaxies does not depend significantly on  $L_{\text{IR}}$  at a given redshift. This property has



**Fig. 4.**  $\lambda_{\text{peak}}$  as a function of  $L_{\text{IR}}$ . The colour-coded error bars (per redshift bins) represent the results of this work, coloured lines represent the best-fit and shaded area the 68% uncertainty of the fit. The solid black line corresponds to the trend of Drew & Casey (2022), while the dotted black line is an extrapolation of their relation.

already been observed by Magnelli et al. (2014), up to  $z \simeq 2$ , where no clear dependence of  $T_{\text{dust}}$  has been observed along the main sequence in the  $SFR - M_{\star}$  plane. However, the normalisation seems to evolve with redshift, reflecting the smooth increase of  $\lambda_{\text{peak}}$  with redshift presented in Drew & Casey (2022). But our results differ from those of Drew & Casey (2022) who studied a different sample built on individual detections and found an evolution of  $T_{\text{dust}}$ , as probed by  $\lambda_{\text{peak}}$ , with  $L_{\text{IR}}$ . This discrepancy comes from the probable incompleteness of the sample of Drew & Casey (2022). In this paper, we stack galaxies to recover the main sequence, and we look at a mass complete sample, where non-detections are accounted for thanks to the stacking technique. On the other hand, Drew & Casey (2022) worked with individually detected galaxies, and thus might be biased towards the brightest galaxies at all redshifts (i.e. starbursts). In addition, their flux-limited sample suffer from selection effect: at low  $L_{\text{IR}}$ , their  $\lambda_{\text{peak}}$  comes from low redshift galaxies, while at high  $L_{\text{IR}}$ , their values are coming from distant galaxies. What Drew & Casey (2022) find is an effect of redshift not an effect of  $L_{\text{IR}}$ .

## 5.2. The main sequence of star-forming galaxies

Here we constrain the  $SFR - M_{\star}$  correlation, also called the MS of star-forming galaxies (Noeske et al. 2007). The preferred method for estimating the SFR of a galaxy is to study the light from OB stars, because of the close link between their short lifetimes and the instantaneous SFR of galaxies. Although most of their light is emitted in the UV, it can be largely absorbed by dust and then re-emitted as thermal radiation in the IR. To obtain the total SFR associated with a galaxy, or in our case the stack of several galaxies, it is necessary to combine the SFRs deduced from both the UV and the IR as:

$$SFR_{\text{Tot}} = SFR_{\text{UV}} + SFR_{\text{IR}}. \quad (10)$$

The  $SFR_{\text{UV}}$ , uncorrected for dust attenuation, is computed from  $L_{\text{UV}}$  (rest-frame 1500 Å UV luminosity) following

Daddi et al. (2004):

$$SFR_{\text{UV}} [M_{\odot} \text{yr}^{-1}] = 2.17 \times 10^{-10} L_{\text{UV}} [L_{\odot}]. \quad (11)$$

For our catalogues, the  $L_{\text{UV}}$  were calculated for all individual galaxies from EAZY (Pannella et al. 2015). In this study, the  $SFR_{\text{UV}}$  was derived for each redshift and  $M_{\star}$  bin by averaging the  $SFR_{\text{UV}}$  of all individual galaxies in the bin. The  $L_{\text{IR}}$  was obtained from the SED fit of our stacks, by integrating the best-fit SED in the range 8–1000 μm rest frame, we deduced the  $SFR_{\text{IR}}$  following Kennicutt (1998a) (see Eq. (6)).

The left panel of Fig. 5 displays our SFR for each bin of redshift and  $M_{\star}$ . The SFR follows a monotonic dependence with increasing  $M_{\star}$ , at fixed redshift; and with increasing redshift, at fixed  $M_{\star}$ . The high-mass end presents a bending of the main sequence slope, which is more prominent as the redshift decreases.

Regarding the contribution of  $H$ -dropouts to our main sequence estimate, Wang et al. (2019) reported that ALMA-detected  $H$ -dropouts mostly fall within the main sequence at  $z \sim 4$  (the version from S15). Moreover, as they represent only  $\sim 7.1\%$  of the galaxies at equivalent  $M_{\star}$ , their omission should not alter the overall shape of the main sequence.

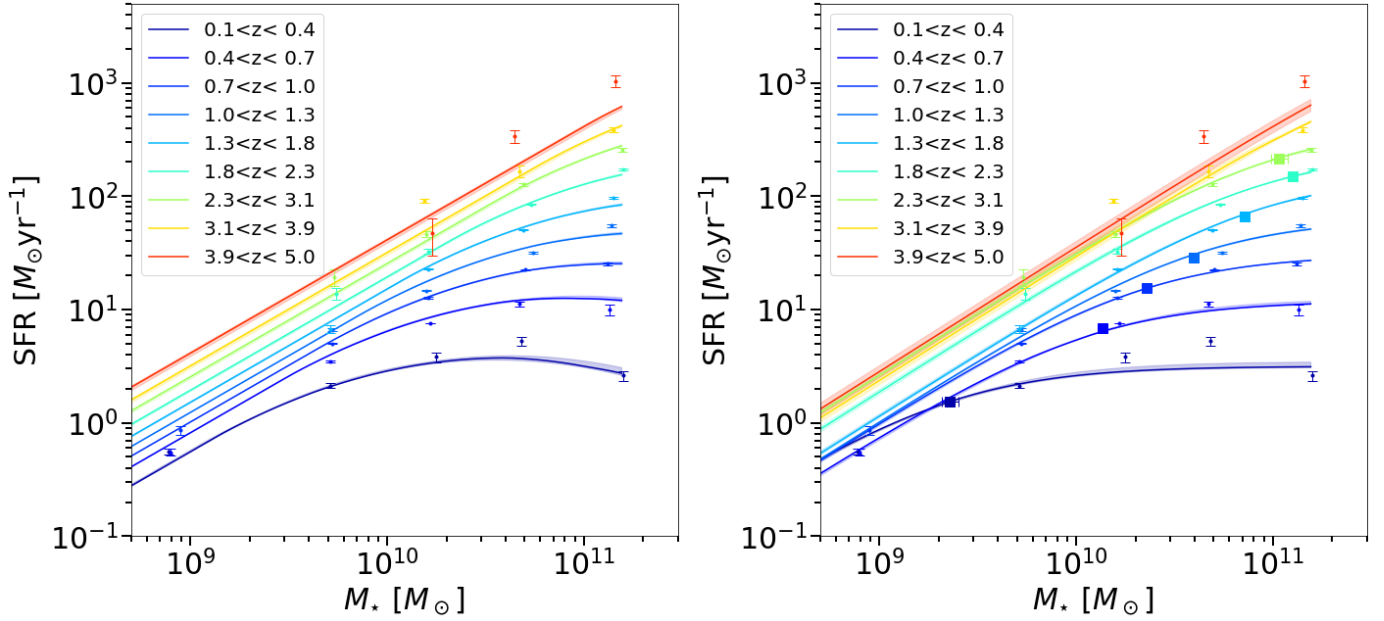
The  $SFR - M_{\star}$  correlation was then fitted using the formula introduced in S15 as it is able to capture the bending of the main sequence at the high-mass end (see Eq. (12)),

$$\log_{10}(SFR_{\text{MS}} [M_{\odot} \text{yr}^{-1}]) = m - m_0 + a_0 r - a_1 [\max(0, m - m_1 - a_2 r)]^2, \quad (12)$$

where  $r \equiv \log_{10}(1+z)$  and  $m \equiv \log_{10}(M_{\star}/10^9 M_{\odot})$ . It was shown in S15 that the main sequence has the shape in logarithmic space of a Gaussian distribution with a homogeneous scatter of  $\sigma \simeq 0.3$  dex. Because we stacked our galaxies, via mean stacking, we actually recover the  $\langle L_{\text{IR}} \rangle$  of our sample which is different from the mean of a Gaussian distribution in logarithmic space with a dispersion  $\sigma_{\text{dist}}$ , with  $\langle 10^X \rangle = \exp((\sigma_X \times \ln(10))^2/2) \times 10^{\langle X \rangle}$ ,  $X = \log_{10}(SFR)$  following a Gaussian distribution. To correct for this, we assume that the dispersion of this Gaussian distribution in logarithmic space is  $\sigma_{\text{dist}} = 0.3$  dex (S15). Our best-fit parameters are given in Table 5, and displayed on the left panel of Fig. 5 (left panel).

The shape chosen for the fit sets a slope of one, which is broadly consistent with our data points. The normalisation of the main sequence increases significantly with redshift. As we have already pointed out in the study of  $T_{\text{dust}}$ , at the highest redshift (i.e.  $3.9 \leq z \leq 5.0$ ) and the two highest  $M_{\star}$  bins (i.e.  $10.5 \leq \log_{10}(M_{\star}) \leq 11$  and  $11 \leq \log_{10}(M_{\star}) \leq 12$ ), appear to be highly star-forming and are above our main sequence by a factor of two. This suggests that these bins include a significant fraction of galaxies with high SFR relative to the main sequence (i.e. high  $R_{\text{SB}}$ ).

We compare our results with several versions of the main sequence from the literature in Fig. 6. Compared to Speagle et al. (2014), who did not fit any bending, our best-fit is consistent with the general trend and evolution of the normalisation. However, our work suggests that the bending is real and becomes stronger at low redshift. Overall, our best-fit remains close to what was found by S15 at  $z \geq 0.7$ . As already confirmed by Delvecchio et al. (2021), the trend of S15 main sequence holds when constructed mainly from FIR data. Nevertheless, we find a stronger curvature of the main sequence at  $z \leq 0.7$  compared to what we could extrapolate from S15. This is probably due in part to the fact that they did not probe the main sequence for  $z \leq 0.3$ . The version of the main sequence by



**Fig. 5.**  $SFR_{MS}$  as a function of  $M_*$  over different redshift bins. The dots represent data from this work. The upper limits were used to perform the fit but are not shown here to avoid overloading the figure. Left panel: fitted by the Eq. (12), the shaded area the 68% uncertainty of the fit. Right panel: fitted by the Eq. (13), the shaded area the 68% uncertainty of the fit. The squares represent the best-fit parameter  $M_0$ . No bending was detected for  $z > 3.1$  and therefore no square error bars are displayed.

Leslie et al. (2020) was derived from a study of the radio continuum at 3 GHz, which may explain the differences in slope, normalisation and bending. But both roughly agree on the same evolutionary trend and on a bending of the main sequence with marginal differences. The bending of the main sequence being at the high-mass end, it suggests that it is triggered by some mass driven physical processes.

We have also followed the evolution of the stellar mass knee  $M_0$  marking the bending point of the main sequence by following Daddi et al. (2022) (see Eq. (13)):

$$\frac{SFR}{SFR_0} = \frac{1}{1 + (M_0/M_*)^\gamma}. \quad (13)$$

Following recommendations from Daddi et al. (2022), we set  $\gamma = 1.1$ , which should help to reduce the errors and should not affect the result much. If  $\gamma$  was defined as a free parameter in the fit, the  $\langle \gamma \rangle$  would be close to 1.1 anyway. Our best-fit parameters are given in Table 6, and displayed on the right panel of Fig. 5. Overall, our results are quite comparable to the evolution of the bending found by Lee et al. (2015) and Daddi et al. (2022) (i.e. a decrease in  $M_0$  as we move to a lower redshift). However, we see little or no evidence of a bending for  $z > 3$ . We suggest that this may be because the physical processes that trigger the main sequence bending have not had enough time to impact the main sequence trend at the high-mass end at  $z > 3$ . Another possibility is that the bending still occurs at high redshifts, but only for very high-mass galaxies ( $M_* > 2 - 3 \times 10^{11} M_\odot$ ), which we could not probe with this study because they are extremely rare.

### 5.3. Gas mass

We chose to calculate  $M_{gas}$  from  $M_{dust}$  deduced from our stacked SEDs. We can link  $M_{gas}$  to  $M_{dust}$  by a gas-to-dust mass ratio ( $\delta_{GDR}$ ) that depends only on the metallicity  $Z$  ( $\log Z = 12 + \log_{10}(O/H)$ ) of the galaxy,

$$M_{gas} = \delta_{GDR}(Z)M_{dust}. \quad (14)$$

**Table 5.** Best-fit parameters of the main sequence of star-forming galaxies.

$m_0$	$m_1$	$a_0$	$a_1$	$a_2$
$9.36^{+0.02}_{-0.02}$	$8.52^{+0.26}_{-0.06}$	$1.34^{+0.09}_{-0.04}$	$0.241^{+0.018}_{-0.010}$	$3.42^{+0.11}_{-0.12}$

**Notes.** Best-fit parameters of the main sequence of star-forming galaxies using Eq. (12).

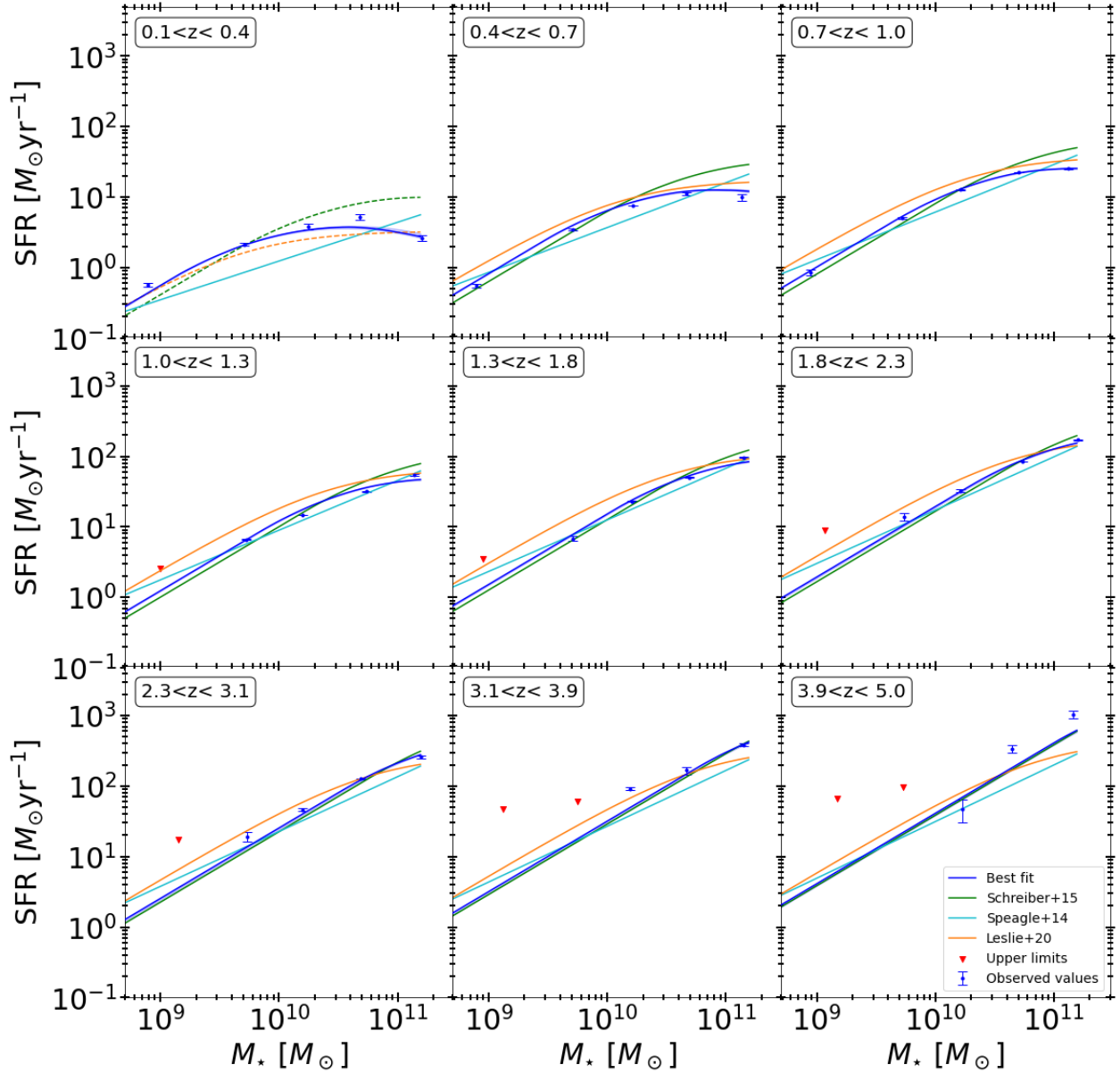
Multiple studies (e.g. Leroy et al. 2011; Magdis et al. 2012; Rémy-Ruyer et al. 2014; Genzel et al. 2015) show consistent  $\delta_{GDR} - Z$  relations that hold at both low and high redshift. We chose here to use the relation from Magdis et al. (2012):

$$\log \delta_{GDR} = 10.54 - 0.99 \times (12 + \log_{10}(O/H)). \quad (15)$$

As for most galaxies we do not have UV or optical spectra, we do not have a direct measurement of their metallicity. To go around this problem, we use the mass-metallicity relation (MZR; Erb et al. 2006) and follow its redshift dependent version from Genzel et al. (2015),

$$12 + \log_{10}(O/H) = a - 0.087 \times (\log_{10}(M_*/1.7) - b)^2, \quad (16)$$

where  $a = 8.74$  and  $b = 10.4 + 4.46 \log_{10}(1+z) - 1.78(\log_{10}(1+z))^2$ . We note that Genzel et al. (2015) use a Chabrier (2003) IMF to define the MZR. In order to use it correctly, our  $M_*$  must first be divided by a factor of 1.7. We chose to adopt an uncertainty of 0.2 dex for our metallicity following the recommendation from Magdis et al. (2012). These latter metallicities are calibrated in the PP04 N2 scale ( $N_2 = [N_{II}] \lambda 6583/H_\alpha$ ; Pettini & Pagel 2004). We note that by using the  $\delta_{GDR}$  relation of Leroy et al. (2011), instead of Magdis et al. (2012), would not change our conclusions because for our sample, the relative median difference in linear scale between the two estimates is  $(M_{gas}^{MZR-Leroy11} - M_{gas}^{MZR-Magdis12})/M_{gas}^{MZR-Magdis12} = 0.14^{+0.05}_{-0.08}$ .



**Fig. 6.**  $SFR_{MS}$  as a function of  $M_*$ . The red triangles represent upper limits, the blue dots are the observed values from this work, and shaded area the 68% uncertainty of the fit. The blue lines represent the best-fit from this work. The green, cyan and orange lines represents the main sequence from S15, Speagle et al. (2014) and Leslie et al. (2020), respectively. The dashed lines represent extrapolated main sequences to redshifts that were not investigated in their respective studies.

Another way to calculate the metallicity is the fundamental metallicity relation (FMR; Mannucci et al. 2010, see Eq. (17)). It differs from the MZR by adding some dependence on the SFR,

$$12 + \log_{10}(\text{O}/\text{H}) = 8.90 + 0.37m - 0.14s - 0.19s^2 + 0.12ms - 0.054s^2, \quad (17)$$

where  $m = \log_{10}(M_*/1.7) - 10$  and  $s = \log_{10}(SFR/1.7)$ . These metallicities are calibrated for the KD02 photoionisation models (Kewley & Dopita 2002). We used the recipe from Kewley & Ellison (2008) to convert it into a PP04 N2 scale. But once again the differences in metallicities between the MZR and FMR method does not impact our conclusions, as for our sample, the relative median is  $(M_{\text{gas}}^{\text{FMR-Magdis12}} - M_{\text{gas}}^{\text{MZR-Magdis12}})/M_{\text{gas}}^{\text{MZR-Magdis12}} = -0.16^{+0.16}_{-0.06}$ .

The previous methods are only reliable when the SED peak is well defined (i.e.  $T_{\text{dust}}$  and  $M_{\text{dust}}$ ), which is not the case in our bins. The only two bins concerned are  $10.0 \leq \log_{10}(M_*) \leq 10.5$  for  $3.1 \leq z \leq 3.9$  and  $3.9 \leq z \leq 5.0$ . In this situation, we can also estimate  $M_{\text{gas}}$  from a single band measurement located in the Rayleigh-Jeans part of the SED (e.g. Scoville et al. 2014; Groves et al. 2015; Schinnerer et al. 2016). The main limitation is that it does not take into account the evolution with redshift of the  $M_{\text{gas}}$  metallicity and  $T_{\text{dust}}$  (e.g. Genzel et al. 2015; Berta et al. 2016; Schinnerer et al. 2016; Magdis et al. 2017; Harrington et al. 2021). However, this is a reliable way of estimating  $M_{\text{gas}}$  at low cost (25% uncertainties; Scoville et al. 2016) and especially when there is only one band measured in the Rayleigh-Jeans tail. For these reasons, while we favoured  $\delta_{\text{GDR}} - \text{MRZ}$  methods to deduce  $M_{\text{gas}}$  as long as the  $T_{\text{dust}}$  was well defined. In the opposite case, and when an ALMA measurement

**Table 6.** Best-fit parameters of the main sequence of star-forming galaxies.

Redshift bin	$\langle z \rangle$	$\log_{10}(M_0/M_\odot)$	$\log_{10}(SFR_0/M_\odot \text{ yr}^{-1})$
$0.1 \leq z \leq 0.4$	0.19	$9.36^{+0.04}_{-0.04}$	$0.49^{+0.03}_{-0.03}$
$0.4 \leq z \leq 0.7$	0.59	$10.13^{+0.02}_{-0.02}$	$1.13^{+0.01}_{-0.01}$
$0.7 \leq z \leq 1.0$	0.87	$10.36^{+0.01}_{-0.01}$	$1.49^{+0.01}_{-0.01}$
$1.0 \leq z \leq 1.3$	1.16	$10.60^{+0.03}_{-0.02}$	$1.76^{+0.02}_{-0.02}$
$1.3 \leq z \leq 1.8$	1.52	$10.86^{+0.02}_{-0.02}$	$2.12^{+0.02}_{-0.02}$
$1.8 \leq z \leq 2.3$	2.02	$11.10^{+0.02}_{-0.02}$	$2.48^{+0.02}_{-0.02}$
$2.3 \leq z \leq 3.1$	2.70	$11.03^{+0.05}_{-0.04}$	$2.63^{+0.04}_{-0.03}$
$3.1 \leq z \leq 3.9$	3.37	$11.65^{+0.35}_{-0.23}$	$3.25^{+0.08}_{-0.11}$
$3.9 \leq z \leq 5.0$	4.30	$11.92^{+0.07}_{-0.11}$	$3.84^{+0.08}_{-0.11}$

**Notes.** Best-fit parameters of the main sequence of star-forming galaxies using Eq. (13).

is available, we calculated  $M_{\text{gas}}$  by the method described in Scoville et al. (2016) (see their Eqs. (16) and (6)), that is

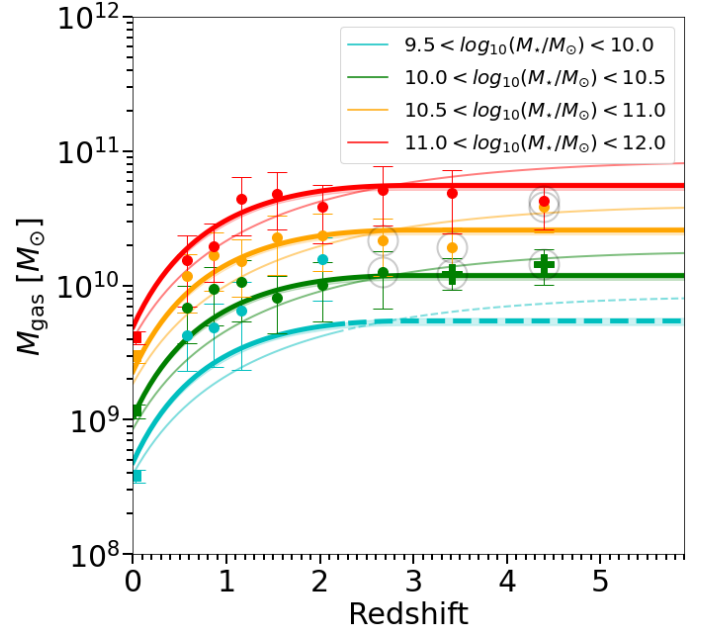
$$M_{\text{gas}} = 1.78 S_{\nu_{\text{obs}}} [\text{mJy}] (1+z)^{-4.8} \left( \frac{\nu_{850 \mu\text{m}}}{\nu_{\text{obs}}} \right)^{3.8} (d_L [\text{Gpc}])^2 \times \frac{6.7 \times 10^{19} \Gamma_0}{\alpha_{850} \Gamma_{\text{RJ}}} 10^{10} M_\odot \quad \text{for } \lambda_{\text{rest}} > 250 \mu\text{m}, \quad (18)$$

$$\Gamma(z, T_d, \nu) = \frac{h \nu_{\text{obs}} (1+z)}{k_B T_d} \frac{1}{\left( \exp\left( \frac{h \nu_{\text{obs}} (1+z)}{k_B T_d} \right) - 1 \right)}, \quad (19)$$

where  $d_L$  is the luminosity distance,  $\alpha_{850} = 6.2 \times 10^{19} \text{ erg sec}^{-1} \text{ Hz}^{-1} \text{ M}_\odot^{-1}$ ,  $\Gamma_0 = \Gamma(z=0, T_d=25 \text{ K}, \nu=\nu_{850 \mu\text{m}})$  and  $\Gamma_{\text{RJ}} = \Gamma(z, T_d=25 \text{ K}, \nu=\nu_{1.1 \text{ mm}})$ ,  $h$  is Planck's constant,  $k_B$  is Boltzmann's constant,  $T_d = 25 \text{ K}$  and  $\nu_\lambda = \frac{c}{\lambda}$ . We note that this technique provides, for our sample, consistent results with those of the MZR method, with a relative median  $(M_{\text{gas}}^{\text{Scoville16}} - M_{\text{gas}}^{\text{MZR-Magdis12}}) / M_{\text{gas}}^{\text{MZR-Magdis12}} = 0.19^{+0.35}_{-0.32}$  when both  $\delta_{\text{GDR}} - \text{MRZ}$  and Scoville et al. (2016) could be performed.

In summary, we chose to calculate our  $M_{\text{gas}}$  using Magdis et al. (2012) for  $\delta_{\text{GDR}}(Z)$  (Eq. (15)) and Genzel et al. (2015) MZR for the metallicities (Eq. (16)). When the peak of the SED is poorly constrained and an ALMA flux is available, we chose to calculate  $M_{\text{gas}}$  from Scoville et al. (2016) method. For the rest of the paper, we will refer to  $M_{\text{gas}}^{\text{MZR-Magdis12}}$ , or  $M_{\text{gas}}^{\text{Scoville16}}$  simply as  $M_{\text{gas}}$ .

Here  $M_{\text{gas}}$  represents the total gas budget of a galaxy including both molecular gas ( $M_{\text{H}_2}$ ) and atomic gas ( $M_{\text{HI}}$ ), that is,  $M_{\text{gas}} = M_{\text{H}_2} + M_{\text{HI}}$ . The HI content of galaxies is still poorly known outside the local universe because the 21cm emission line is difficult to detect with current facilities. However, Bauermeister et al. (2010) showed that it is unlikely that the HI content of galaxies varies strongly with redshift. In contrast, the  $\text{H}_2$  content evolves strongly with redshift (e.g. Daddi et al. 2010; Tacconi et al. 2010, 2013; Lagos et al. 2015; Genzel et al. 2015; Tacconi et al. 2018). In Tacconi et al. (2018), it was concluded that the assumption  $M_{\text{gas}} \sim M_{\text{H}_2}$  should hold for  $z > 0.4$ . Therefore, we decided not to consider our  $M_{\text{gas}}$  for  $z < 0.4$  (i.e. our first redshift bin), as we could not probe the HI content of the galaxy at these redshifts. In order to add some reliable measurements of  $M_{\text{H}_2}$  to our study at  $z < 0.4$ , we used data points from Sainonge et al. (2017) that come from very local measurements ( $0.01 \leq z \leq 0.05$ ).



**Fig. 7.**  $M_{\text{gas}}$  as a function of redshift for different  $M_\star$ . The dots represent  $M_{\text{gas}}$  estimates from this work using the  $\delta_{\text{GDR}} - \text{MRZ}$  method. The crosses represent  $M_{\text{gas}}$  estimates from this work following the method from Scoville et al. (2016). The squares are measurements from Sainonge et al. (2017). The bins that have been corrected for the  $H$ -dropout contribution are circled. The thick and thin solid colour lines represent the best-fit from this work, using Eq. (21) and Eq. (20), respectively. The dashed lines are the fits extrapolation. The shaded area the 68% uncertainty of the fit.

Our  $M_{\text{gas}}$  measurements are displayed in Fig. 7. We observe a rapid rise with redshift from Sainonge et al. (2017) data at low redshift, to our first data points at  $z \geq 0.4$ . Then, at fixed  $M_\star$ ,  $M_{\text{gas}}$  reach a maximum at  $z \sim 1-2$  and remains relatively constant as the redshift increases. At fixed redshift,  $M_{\text{gas}}$  gradually increase with  $M_\star$ . All values of our gas masses are provided in Appendix A.

We examined the contribution of the  $H$ -dropouts to  $M_{\text{gas}}$  in order to have the most unbiased view possible. Here, we have simply calculated the  $M_{\text{gas}}$ , associated with each  $H$ -dropout in the sample of Wang et al. (2019), using their ALMA measurement at  $870 \mu\text{m}$  and following Scoville et al. (2016) method. The final contribution of the  $H$ -dropout to the total  $M_{\text{gas}}$  within each bin is  $M_{\text{gas}}^{\text{H-dropout}} \times N_{\text{gal}}^{\text{H-dropout}} \times \epsilon / (M_{\text{gas}}^{\text{H-dropout}} \times N_{\text{gal}}^{\text{H-dropout}} \times \epsilon + M_{\text{gas}} \times N_{\text{gal}}) = 19^{+7\%}_{-17\%}$  on average (only the bins with at least one  $H$ -dropout are taken into account), where  $\epsilon = \text{Area}_{\text{This work}} / \text{Area}_{\text{Wang+19}} \sim 1.8$ . The contribution of  $H$ -dropouts can represent up to  $\sim 32\%$  for some bins, hence we chose to add it to our data. The seven bins that have been corrected for the  $H$ -dropout contribution are circled Fig. 7.

Tacconi et al. (2018) combined  $M_{\text{gas}}$  from stacks in the IR and  $M_{\text{gas}}$  from CO emission. We try to fit our data using the formula of Tacconi et al. (2018):

$$\log_{10}(M_{\text{gas}}) = A + B \times (\log_{10}(1+z) - F)^\beta + D \times (\log_{10}(M_\star/1.7) - 10.7) + \log_{10}(M_\star/1.7) + [C \times \log_{10}(R_{\text{SB}}(z, M_\star)) + E \times \log_{10}(R_e/R_{e0}(z, M_\star))], \quad (20)$$

where the factor 1.7 represents the conversion of  $M_\star$  from a Salpeter (1955) to Chabrier (2003) IMF. As our data points represent main sequence galaxies, we cannot probe the  $C_\mu$  term in

**Table 7.** Best-fit parameters of  $M_{\text{gas}}$  evolution of the main sequence.

$A$	$B$	$F$	$D$	$\beta^{(*)}$
$0.0453^{+0.146}_{-0.036}$	$-1.61^{+0.73}_{-2.87}$	$0.901^{+0.362}_{-0.373}$	$-0.306^{+0.929}_{-0.718}$	2

**Notes.** Best-fit parameters of  $M_{\text{gas}}$  evolution of the main sequence using Eq. (20).  $(*)$  Indicates that this parameter was fixed during the fit.

**Table 8.** Best-fit parameters of  $M_{\text{gas}}$  evolution of the main sequence.

$m_0$	$m_1$	$a_1$	$a_2$
$1.93^{+0.18}_{-0.23}$	$0.69^{+0.02}_{-0.01}$	$3.53^{+0.33}_{-0.56}$	$-2.96^{+0.92}_{-0.71}$

**Notes.** Best-fit parameters of  $M_{\text{gas}}$  evolution of the main sequence using Eq. (21).

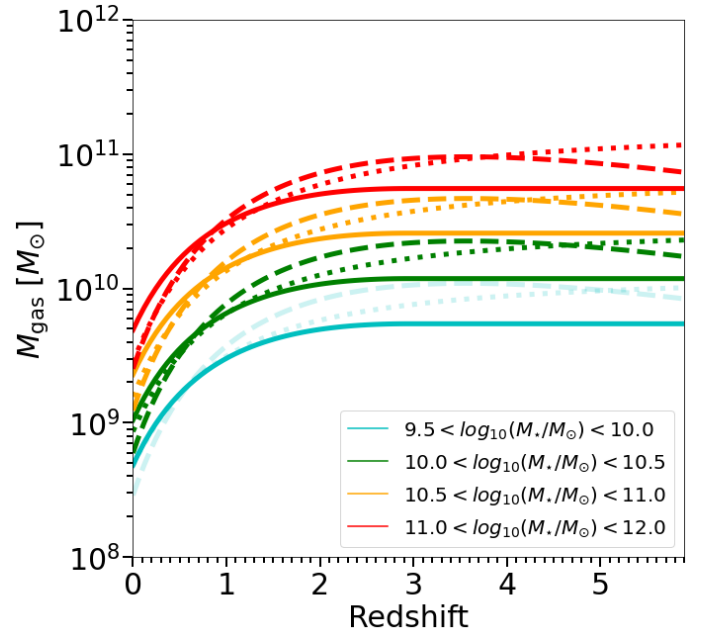
Tacconi et al. (2018), that represents the evolution with respect to the distance to the main sequence  $sSFR/sSFR_{\text{MS}}$ , where  $sSFR$  is the specific SFR ( $sSFR = SFR/M_{\star}$ ). And we do not explore the morphology of galaxies in this study, which makes it impossible to probe the term  $E_{\mu}$ , which compares the effective radius of galaxies ( $R_e$ ) to the mean effective radius of the star forming population  $R_{e_0}$ . The results are displayed in Fig. 7, and our best-fit parameters are given in Table 7.

Looking at Fig. 7, we notice that  $M_{\text{gas}}$  observed at low redshift ( $0.4 \leq z \leq 1$ ) tends to be higher than the best-fit trend (even though it is in most cases within the error bars), and the trends from literature such as Tacconi et al. (2018). A similar effect has been observed by Tacconi et al. (2018) who measured higher depletion times ( $\tau_{\text{dep}}$ ) deduced from dust observations compared to the ones deduced from CO line fluxes. The effect observed in Tacconi et al. (2018) decreases slowly with redshift (about 0.3 dex at  $z \sim 0.4$  to 0 at  $z \sim 1.5$ ) and roughly matches the effect observed here. We conclude that this effect, which was solved in Tacconi et al. (2018) by matching the zero-point for each method, arises when deducing  $M_{\text{gas}}$  from dust observations. It feels that the form use in Tacconi et al. (2018) to fit  $M_{\text{gas}}$  fails to properly recover the form of our measurements. In a way to provide for a better fit, we chose to also fit our data using a formula of the form:

$$\log_{10}(M_{\text{gas}}) = \begin{cases} m_0 + m_1 m + a_1 r + a_2 r^2 & \text{for } z < z_0 \\ m_0 + m_1 m + a_1 r_0 + a_2 r_0^2 & \text{for } z \geq z_0, \end{cases} \quad (21)$$

where  $m = \log_{10}(M_{\star})$ ,  $r = \log_{10}(1 + z)$ ,  $r_0 = \log_{10}(1 + z_0)$  and  $z_0 = -a_1/2a_2$ . The best-fit parameters, using Eq. (21) are given in Table 8.

We compare our results to the trends observed by Tacconi et al. (2018) and Wang et al. (2022) in Fig. 8. In Fig. 8, the trends are displayed for stellar mass bins in the Salpeter (1955) IMF. Differences between our results and previous works (e.g. Scoville et al. 2016, 2017; Tacconi et al. 2018; Liu et al. 2019; Wang et al. 2022) may result from selection effects and from the method used to derive the gas masses. The present sample was built using an  $H$ -band selection that is the best proxy for the stellar mass, although not perfect for the highest redshifts. Other studies generally rely on a CO-line and/or dust continuum selection. This may lead to a potential bias favouring galaxies with higher specific star formation rate, that is, higher



**Fig. 8.**  $M_{\text{gas}}$  as a function of redshift and  $M_{\star}$ . The solid lines represent the best-fit from this work colour-coded by  $M_{\star}$  (using Eq. (21)), the dashed lines represent the trend from Tacconi et al. (2018), the dotted lines are from Wang et al. (2022). The faded lines represent an extrapolation from their respective laws.

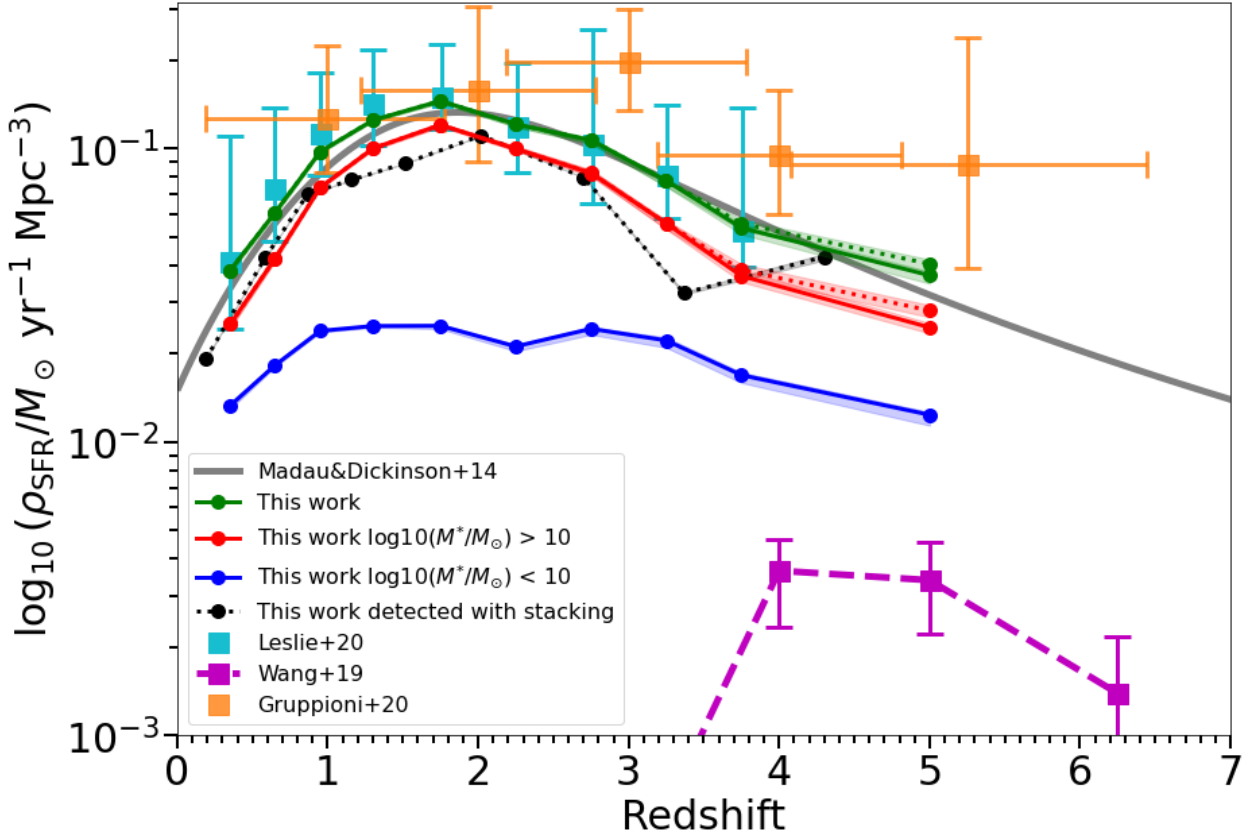
$sSFR = SFR/M_{\star}$ , hence galaxies above the main sequence. It has been found that such galaxies in general exhibited higher gas fractions. In our analysis, we tried to avoid such bias by checking that our stacked samples followed the median of the  $SFR - M_{\star}$  main sequence.

The trends observed by Wang et al. (2022) were obtained as well using a mass-selected sample, but the gas masses were derived from millimetre continuum fluxes using the method from Scoville et al. (2016). We produced a version of our analysis using the same method for consistency check, it is presented in Appendix B. We find that we are able to reproduce the same trends as Wang et al. (2022) if we were to use the same technique based only on the formula of Scoville et al. (2016). Hence, the differences found in the present study can be explained by the use of a different method to derive gas masses. We favoured the present method because it takes into account the variation of metallicity and the associated change in dust-to-gas ratio as a function of stellar mass. This is one of the reasons why we opted for a mass-selection in the first place.

As mentioned in Sect. 4, we chose to work with  $M_{\text{dust}}$  from amorphous silicate and graphitic grains (Draine & Li 2007) instead of amorphous carbon (Galliano et al. 2011; S18). We investigated what the impact would have been on  $M_{\text{gas}}$  if we had chosen a model based on amorphous carbon such as Galliano et al. (2021). Although the results are quite comparable, using the model from Galliano et al. (2021), would translate into slightly lower  $M_{\text{gas}}$  ( $\sim -10\%$  at  $z = 4$ ). These differences would not have changed the conclusions drawn in this paper.

## 6. Cosmic star formation history

In this section, we present the redshift evolution of  $\rho_{\text{SFR}}$ , that is, the cosmic star formation history. To calculate the  $\rho_{\text{SFR}}$ , we start from the stellar mass function of SFGs of Davidzon et al. (2017), that is given for different redshift bins. For each redshift bin, we



**Fig. 9.** Cosmic star formation rate density ( $\rho_{\text{SFR}}$ ) as a function of redshift. The green line represents the total  $\rho_{\text{SFR}}$  trend from this work (i.e. integrating the stellar mass function down to  $3 \times 10^9 M_{\odot}$ ). The red and blue lines represent the contribution of galaxies with  $\log_{10}(M_{\star}/M_{\odot}) > 10$  and  $\log_{10}(M_{\star}/M_{\odot}) < 10$ , respectively. The black dotted line represents the  $\rho_{\text{SFR}}$  detected with stacking. The purple squares show the contribution of  $H$ -dropout from Wang et al. (2019). The dotted green and red lines show the trend once  $H$ -dropout contribution from Wang et al. (2019) has been added to the respective coloured solid lines. The grey line represents the trend from Madau & Dickinson (2014). The cyan and orange squares represent  $\rho_{\text{SFR}}$  estimate from Leslie et al. (2020) and Gruppioni et al. (2020), respectively.

generated galaxies with uniform redshift distribution within the bin, and a  $M_{\star}$  distribution following these stellar mass functions. From the redshift and  $M_{\star}$ , we then assigned a SFR using the same method as in Sect. 4.3. In summary,  $SFR = R_{\text{SB}} \times SFR_{\text{MS}}$ , where  $SFR_{\text{MS}}$  is calculated from the main sequence evolution found in this work (see Eq. (12) and parameters Table 5). And a  $R_{\text{SB}}$  is randomly drawn from a double Gaussian distribution representing the position of normal and starbursting galaxies relative to the main sequence (Eq. (3)). Then  $\rho_{\text{SFR}}$  is calculated by summing the SFR of galaxies down to  $M_{\star}^{\text{lim}} = 3 \times 10^9 M_{\odot}$ . Errors were generated by varying the  $SFR_{\text{MS}}$  trend of Eq. (12) within the errors of the fit, a 100 times.

The study of Wang et al. (2019) presents the contribution to  $\rho_{\text{SFR}}$  of  $H$ -dropout galaxies, which can reach up to 10% at  $z \sim 4$ –5. We have therefore added the contribution of  $H$ -dropout galaxies of Wang et al. (2019) to our  $\rho_{\text{SFR}}$  considering the sample of  $H$ -dropout galaxies have a median stellar mass of  $M_{\star} \sim 10^{10.6} M_{\odot}$  (Wang et al. 2019).

We compare the  $\rho_{\text{SFR}}$  evolution with that of Madau & Dickinson (2014) who integrate UV and IR luminosity functions down to  $L_{\text{min}} = 0.03L_{\star}$ . Deducing  $\rho_{\text{SFR}}$  by integrating luminosity functions down to  $0.03L_{\star}$ , or the mass function down to  $3 \times 10^9 M_{\odot}$  should yield, to the first order, similar results (Schreiber et al. 2015). We chose for the rest of the paper to derive our  $\rho_{\text{SFR}}$  by integrating the mass function down to  $3 \times 10^9 M_{\odot}$ . Thus, all conclusions on the evolution of the

$\rho_{\text{SFR}}$  in this paper should be understood in this framework, as integrating the luminosity function down to lower luminosities, or the mass function down to lower masses, must yield a higher total  $\rho_{\text{SFR}}$ .

The total  $\rho_{\text{SFR}}$  is displayed in Fig. 9 along with some examples from the literature, and the data points from this work are summarised in Table 9. We observe a rise of  $\rho_{\text{SFR}}$  from  $z \sim 5$  to  $z \sim 2$ , and then it gradually decreases down to  $z \sim 0.35$ . The high masses (i.e.  $\log_{10}(M_{\star}/M_{\odot}) > 10$ ) account for most of the  $\rho_{\text{SFR}}$  until  $z \sim 4$ . On the other hand the  $\rho_{\text{SFR}}$  associated to low-mass galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) < 10$ ) is roughly constant over  $0 \leq z \leq 5$ .

We can see that our estimate of the total  $\rho_{\text{SFR}}$  is close to what has been observed by Leslie et al. (2020) and to the evolution of Madau & Dickinson (2014). On the other hand, the measurements from Gruppioni et al. (2020) are mostly in disagreement with our results at high redshift (i.e.  $z \geq 2$ ). The  $\rho_{\text{SFR}}$  from Gruppioni et al. (2020) is deduced from the integration of the IR-luminosity function (which make it  $\rho_{\text{SFR}}^{\text{IR}}$ ) down to  $10^8 L_{\odot}$ .

In Fig. 10, we show the contribution to the total  $\rho_{\text{SFR}}$  of the full range of stellar masses. We observe that the increase in the total  $\rho_{\text{SFR}}$ , from  $z \sim 5$  to  $z \sim 2$ , comes from the growing number of massive galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) > 10$ ), which can be seen in the evolution of the stellar mass function at these redshifts (Davidzon et al. 2017). Downsizing and the bending of the main sequence explain the fall of the contribution of massive galaxies

**Table 9.**  $\rho_{\text{SFR}}$  as a function of redshift.

Redshift	0.35	0.65	0.95	1.3	1.75	2.25	2.75	3.25	3.75	5.00
$\log_{10} \left( \frac{\rho_{\text{SFR}}}{M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}} \right)$	$-1.41^{+0.01}_{-0.01}$	$-1.22^{+0.01}_{-0.01}$	$-1.01^{+0.01}_{-0.01}$	$-0.90^{+0.01}_{-0.01}$	$-0.84^{+0.01}_{-0.01}$	$-0.92^{+0.01}_{-0.01}$	$-0.97^{+0.01}_{-0.01}$	$-1.11^{+0.01}_{-0.02}$	$-1.26^{+0.01}_{-0.02}$	$-1.39^{+0.02}_{-0.03}$
$\log_{10} \left( \frac{\rho_{\text{SFR}}^{M_{\star} > 10^{10} M_{\odot}}}{M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}} \right)$	$-1.60^{+0.01}_{-0.01}$	$-1.37^{+0.01}_{-0.01}$	$-1.13^{+0.01}_{-0.01}$	$-1.00^{+0.01}_{-0.01}$	$-0.92^{+0.01}_{-0.01}$	$-1.00^{+0.01}_{-0.01}$	$-1.08^{+0.01}_{-0.01}$	$-1.25^{+0.01}_{-0.02}$	$-1.41^{+0.02}_{-0.02}$	$-1.55^{+0.02}_{-0.02}$
$\log_{10} \left( \frac{\rho_{\text{SFR}}^{M_{\star} < 10^{10} M_{\odot}}}{M_{\odot} \text{ yr}^{-1} \text{ Mpc}^{-3}} \right)$	$-1.88^{+0.03}_{-0.03}$	$-1.74^{+0.02}_{-0.03}$	$-1.62^{+0.02}_{-0.02}$	$-1.60^{+0.02}_{-0.02}$	$-1.60^{+0.03}_{-0.04}$	$-1.67^{+0.04}_{-0.05}$	$-1.61^{+0.04}_{-0.05}$	$-1.65^{+0.03}_{-0.07}$	$-1.77^{+0.05}_{-0.07}$	$-1.91^{+0.05}_{-0.09}$

**Notes.**  $\rho_{\text{SFR}}$  as a function of redshift from this work. Here,  $\rho_{\text{SFR}}$  values include the  $H$ -dropout contribution from Wang et al. (2019).

from  $z \sim 2$  to  $z \sim 0$ . We again observe that massive galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) > 10$ ) dominate the total  $\rho_{\text{SFR}}$  at all redshifts (i.e. they account for more than 70%). In particular, it appears that galaxies with  $10.5 \leq \log_{10}(M_{\star}/M_{\odot}) \leq 11.25$  account for more than  $\sim 55\%$  of the total  $\rho_{\text{SFR}}$  at  $z = 2$ , making them the main driver of the peak in the observed cosmic star formation history at this redshift.

## 7. Cosmic evolution of the gas mass density

In Sect. 6, we have simulated a catalogue of galaxies where for each of them, we calculated  $M_{\star}$  and  $R_{\text{SB}}$ . From these properties, we now infer  $M_{\text{gas}}$  by expanding Eq. (21) as follows:

$$\log_{10}(M_{\text{gas}}) = \begin{cases} m_0 + m_1 m + a_1 r + a_2 r^2 \\ + C \times \log_{10}(R_{\text{SB}}(z, M_{\star})) & \text{for } z < z_0 \\ m_0 + m_1 m + a_1 r_0 + a_2 r_0^2 \\ + C \times \log_{10}(R_{\text{SB}}(z, M_{\star})) & \text{for } z \geq z_0, \end{cases} \quad (22)$$

where  $m = \log_{10}(M_{\star})$ ,  $r = \log_{10}(1 + z)$ ,  $r_0 = \log_{10}(1 + z_0)$  and  $z_0 = -a_1/2a_2$ . Here,  $m_0$ ,  $m_1$ ,  $a_1$  and  $a_2$  are taken from Table 8, and  $C = 0.53$  is taken from Tacconi et al. (2018). Next, the cosmic evolution of gas density ( $\rho_{\text{gas}}$ ) is calculated by summing the  $M_{\text{gas}}$  of galaxies down to  $M_{\star}^{\text{lim}} = 3 \times 10^9 M_{\odot}$ . Errors were generated by varying our best-fit of  $M_{\text{gas}}^{\text{MS}}$  within its errors, one hundred times. Defined this way,  $\rho_{\text{gas}}$  directly represents the gas content of galaxies contributing to the  $\rho_{\text{SFR}}$  presented in Fig. 9. We display the evolution of  $\rho_{\text{gas}}$  as a function of redshift in Fig. 11. We can see the same kind of features as for the evolution of  $\rho_{\text{SFR}}$ : a rise and fall with redshift with a maximum around  $z \sim 2$ ; a dominance across all redshifts of the high-mass galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) > 10$ ) contribution; and a relatively flat evolution of the low-mass galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) < 10$ ). This shows that the SFE of both high and low-mass galaxies are not drastically changing with time, and thus that the gas content of galaxies (i.e. the accretion) is the primary driver of their SFRs: high-mass galaxies have higher SFR at fixed redshift because they have more gas. However, comparing the relative evolution of  $\rho_{\text{gas}}$  (see Fig. 11) and  $\rho_{\text{SFR}}$  (see Fig. 9) shows that  $\rho_{\text{gas}}^{z=5}/\rho_{\text{gas}}^{z=0}$  (and in particular the high-mass contribution) is lower by a factor  $\sim 3$ – $4$  compared to  $\rho_{\text{SFR}}^{z=5}/\rho_{\text{SFR}}^{z=0}$ . This hints that one unit of gas leads to more stars being formed at  $z \sim 5$  compared to  $z \sim 0$  (i.e. a higher SFE at  $z \sim 5$  compared to  $z \sim 0$ ).

We substituted, in our later method to estimate  $\rho_{\text{gas}}$ ,  $M_{\text{gas}}$  from this work, for best-fit of  $M_{\text{gas}}$  from the literature and compared it to our estimate of  $\rho_{\text{gas}}$ . We can see in Fig. 11, that the redshift evolution of  $\rho_{\text{gas}}$  from Tacconi et al. (2018) or Wang et al. (2022) are higher compared to the one from this work. These are simply resulting from discrepancies already observed in the respective  $M_{\text{gas}}$  trend they were built from, and goes together with a stronger contribution of low-mass galaxies for which the

gas mass discrepancy discussed in Fig. 8 is the largest at high redshifts (i.e.  $z > 3$ ). The latter part is a potential weakness of all studies including ours, since the low stellar mass – high redshift of the cosmic gas mass density range relies on rather uncertain extrapolations. As already pointed out in Liu et al. (2019), the form of the formula chosen to fit  $M_{\text{gas}}$  can have a significant impact on the resulting  $\rho_{\text{gas}}$  trend, and could be the cause of what is observed here.

We also computed  $\rho_{\text{gas}}$  by summing  $M_{\text{gas}}$  of all galaxies down to several different  $M_{\star}^{\text{lim}}$  ( $M_{\star}^{\text{lim}} = 1.7 \times 10^8$ ,  $1.7 \times 10^9$  and  $1.7 \times 10^{10} M_{\odot}$ ). We compared it to Magnelli et al. (2020) where a  $\rho_{\text{gas}}$  was computed for the same  $M_{\star}^{\text{lim}}$  from the stacking of  $H$ -band selected galaxies in ALMA and through the method from Scoville et al. (2016). The results are displayed in Fig. 12. Here, our work agrees quite well with that of Magnelli et al. (2020) for the total  $\rho_{\text{gas}}$  (i.e.  $M_{\star}^{\text{lim}} = 1.7 \times 10^8$ ). However, the mass distribution is not similar: massive galaxies (i.e.  $M_{\star}^{\text{lim}} = 1.7 \times 10^{10}$ ) contribute more at low redshift (i.e.  $z \leq 0.7$ ) in this work as compared to Magnelli et al. (2020); conversely, their contribution is smaller at higher redshift (lower by a factor  $\sim 3$  at  $z \sim 2.8$  compared to Magnelli et al. 2020).

## 8. Discussion

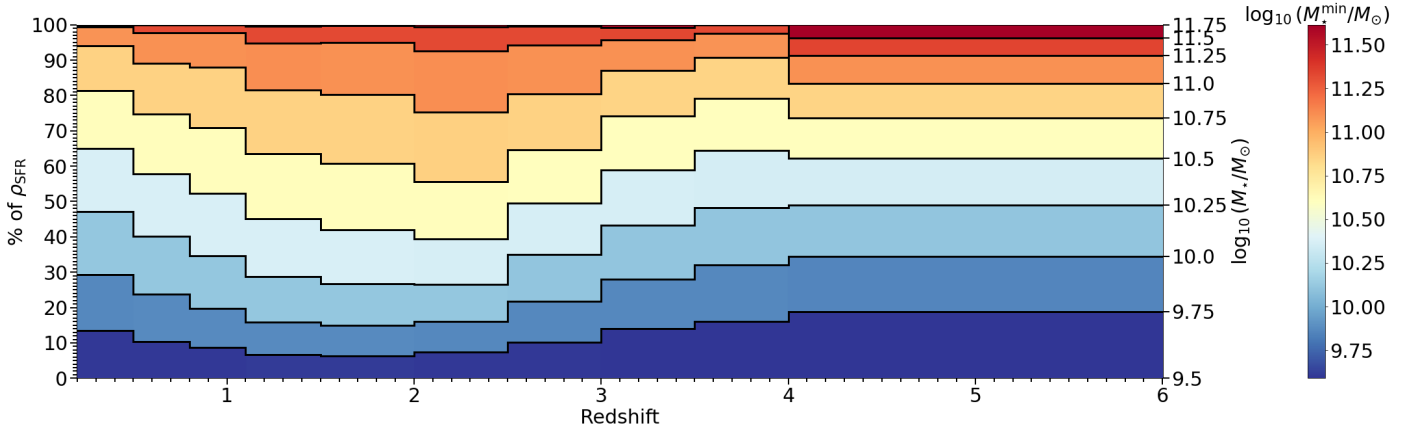
Here, when we discuss the evolution of galaxies, it should be understood that we are discussing the evolution of main sequence galaxies, as all scaling relations of SFR,  $M_{\text{gas}}$  and  $T_{\text{dust}}$  with  $M_{\star}$  and redshift have been deduced for main sequence galaxies.

### 8.1. Cosmic densities of star formation and gas density as a function of stellar mass

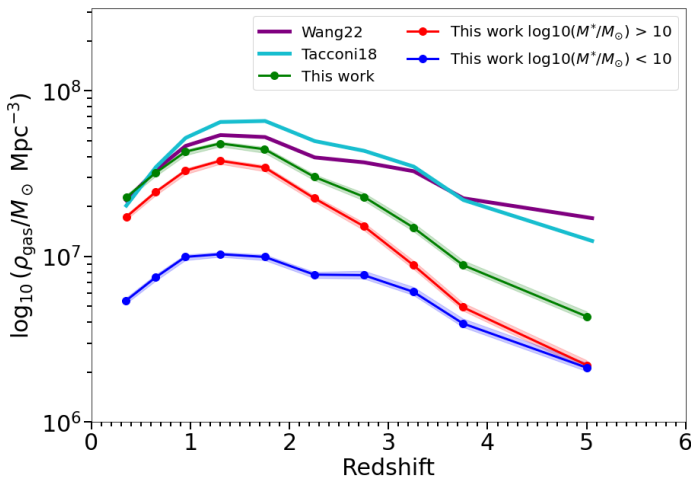
An interesting feature of Fig. 9 is the fact that the contribution to  $\rho_{\text{SFR}}$  of galaxies with  $M_{\star} \leq 10^{10} M_{\odot}$  appears to be constant over  $1 \leq z \leq 3$  and slightly decreasing for  $z \geq 3$ . This means that the balance of the number of these galaxies and their efficiency in producing stars remains constant with cosmic time. This feature can also be seen in Fig. 11, where the contribution to  $\rho_{\text{gas}}$  is also quite stable for galaxies with  $M_{\star} \leq 10^{10} M_{\odot}$ .

On the other hand, galaxies with  $M_{\star} \geq 10^{10} M_{\odot}$  account for most stars formed up to  $z \sim 5$ . This shows that it is galaxies with  $M_{\star} \geq 10^{10} M_{\odot}$  that are responsible for the observable shape on the  $\rho_{\text{SFR}}$ , especially at the cosmic noon at  $z \sim 2$ . In Fig. 13a, we can see that the 10% (of the total number of galaxies contributing to  $\rho_{\text{SFR}}$ ) most massive galaxies contribute to a large fraction of the  $\rho_{\text{SFR}}$  at all redshifts. This contribution goes from  $\sim 24\%$  at  $z \sim 0$  to  $\sim 48\%$  at  $z \sim 5$ . In particular, this means that by considering only 10% (in number) of the most massive galaxies, we can deduce a relatively good estimation of the total  $\rho_{\text{SFR}}$

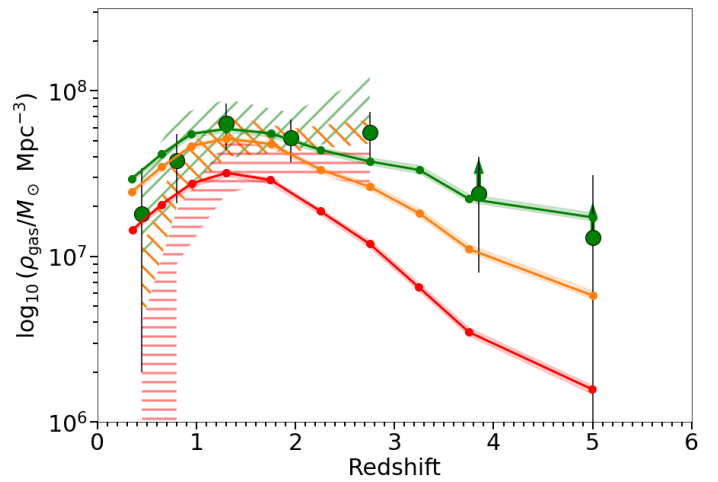




**Fig. 10.** Contributions over the whole range of stellar masses to the total  $\rho_{\text{SFR}}$  as a function of redshift for SFGs.



**Fig. 11.**  $\rho_{\text{gas}}$  as a function of redshift. The green line represents the total  $\rho_{\text{gas}}$  inferred from this work (i.e. integrating the stellar mass function down to  $3 \times 10^9 M_{\odot}$ ). The red and blue lines represent the contribution to the total  $\rho_{\text{gas}}$  of galaxies with  $\log_{10}(M_{\star}) > 10$  and  $\log_{10}(M_{\star}) < 10$ , respectively. The purple and cyan lines represent  $\rho_{\text{gas}}$  deduced from the  $M_{\text{gas}}$  definition of Wang et al. (2022) and Tacconi et al. (2018), respectively.



**Fig. 12.**  $\rho_{\text{gas}}$  as a function of redshift. The green, orange and red lines represent  $\rho_{\text{gas}}$  derived from this work with  $M_{\star}^{\text{lim}}$  of  $1.7 \times 10^8$ ,  $1.7 \times 10^9$  and  $1.7 \times 10^{10} M_{\odot}$ , respectively. The green, orange and red hatches represent estimates of  $\rho_{\text{gas}}$  from Magnelli et al. (2020) with  $M_{\star}^{\text{lim}}$  of  $1.7 \times 10^8$ ,  $1.7 \times 10^9$  and  $1.7 \times 10^{10} M_{\odot}$ , respectively. The green dots with black circles depict the estimate of  $\rho_{\text{gas}}$  from Magnelli et al. (2020) for  $M_{\star} > M_{\text{limit}}$ ,  $M_{\text{limit}}$  being the stellar mass completeness limit from Magnelli et al. (2020).

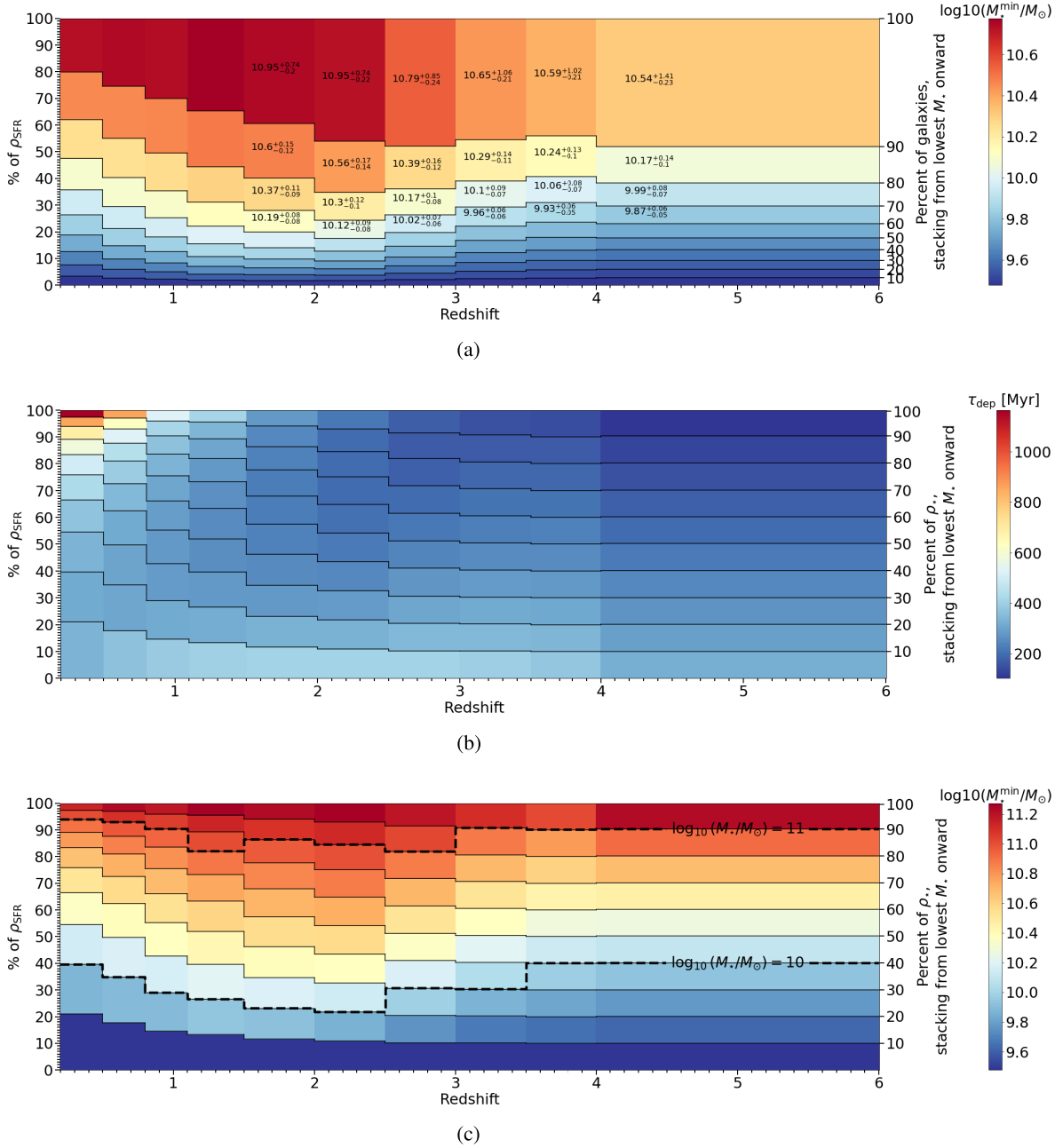
at intermediate and high redshifts. From these numbers, we can also recognise some hierarchical growth effect, as bins gradually rise in  $M_{\star}$  over time.

We find that  $H$ -dropout galaxies account for 13% of the total cosmic SFR density (CSFD) at  $z = 4-6$ . This is in agreement with the value of 10% found in Wang et al. (2019). We note that  $H$ -dropout galaxies account for  $\sim 23\%$  of the stars formed in massive galaxies (i.e.  $M_{\star} \geq 10^{10.3} M_{\odot}$ ). This is a large contribution but not a dominant one, as opposed to the claim of Wang et al. (2019) that  $H$ -dropouts dominate the SFR density in the most massive galaxies. This discrepancy results from our use of a fully complete sample of galaxies selected in stellar mass, whereas Wang et al. (2019) compared  $H$ -dropouts to UV-selected and similarly massive Lyman-break (LBG) galaxies.

We observe a decline in star formation in massive galaxies that mirrors that of the total star formation density. This is an illustration of the bending effect of the MS at high mass (see Fig. 13a). The use of the depletion time in Fig. 13b illustrates this clearly: the depletion time of massive galaxies varies from  $\tau_{\text{dep}} \lesssim 200$  Myr at  $z > 4$  to  $\tau_{\text{dep}} \geq 600$  Myr at  $z < 1$ .

In Fig. 13c, we show that the contribution of the most massive galaxies (e.g.  $M_{\star} \sim 10^{11} M_{\odot}$ ) to the CSFD is nearly flat (i.e. their contribution to the CSFD is roughly constant), meaning that they tightly follow the global history of cosmic star formation (i.e. the global shape of the CSFD consisting of a rise, a cosmic noon and a fall).

Let us consider a typical galaxy with  $M_{\star} = M_{\text{star}}$ , such that 50% of  $\rho_{\star}$  (cosmic stellar mass density) is made of galaxies above and below  $M_{\text{star}}$ . We can follow the contribution of such typical galaxy to the cosmic SFR density by following the line where the  $Y$ -axis on the right of Fig. 13c equals 50%. We find that galaxies with  $M_{\star} > M_{\text{star}}$  contribute to  $\sim 50\%$  of  $\rho_{\text{SFR}}$  at  $z > 3$  and drops to only  $\sim 25\%$  at  $z \sim 0$ . This decrease of the relative contribution of massive galaxies to  $\rho_{\star}$  illustrates the impact on the cosmic SFR history of the bending of the MS. The fact that this happens continuously with cosmic time, supports a scenario of a slow downfall of star formation rather than a rapid quenching of the most massive galaxies.



**Fig. 13.** Contribution to  $\rho_{\text{SFR}}$  of SFGs as a function of redshift. Catalogues were binned through different methods and a specific property is displayed through colours: see each sub-figure for specifics. (a) Each bin includes 10% of the galaxies in number, picking from the lowest  $M_*$  onwards. The minimum  $M_*$  in the bin defines the colour. The numbers give the median  $M_*$  within the bin along with the lower and maximum extension of the bin. (b) Each bin includes 10% of the total  $M_*$  of all galaxies, picking from the lowest  $M_*$  onwards. The median  $\tau_{\text{dep}}$  in the bin defines colour. (c) Each bin includes 10% of the total  $M_*$  of all galaxies, picking from the lowest  $M_*$  onwards. The minimum  $M_*$  in the bin defines the colour. Contours of  $\log_{10}(M_*/M_{\odot}) = 10$  and  $\log_{10}(M_*/M_{\odot}) = 11$  are added as black dashed lines.

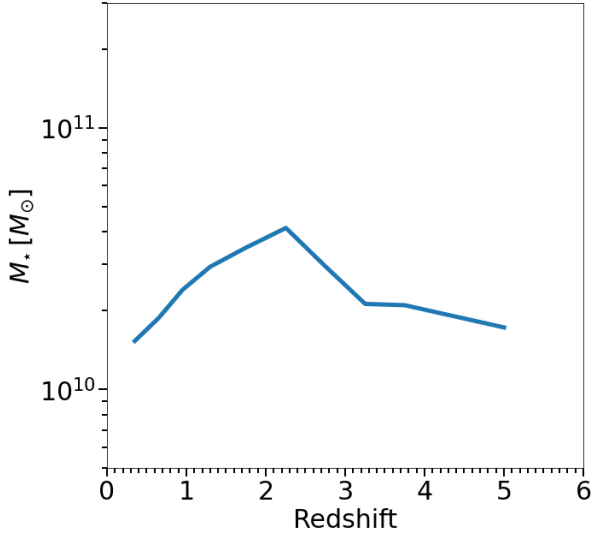
In Fig. 14, we track the evolution of the stellar mass above which galaxies contribute to exactly 50% of the CSFD. We can see in Fig. 14 that this particular stellar mass falls around  $\sim 10^{10.2} M_{\odot}$  with an evolution with redshift that mirror the one of the CSFD by peaking at the cosmic noon at  $\sim 10^{10.4} M_{\odot}$ .

## 8.2. Comparison with the TNG100 simulation

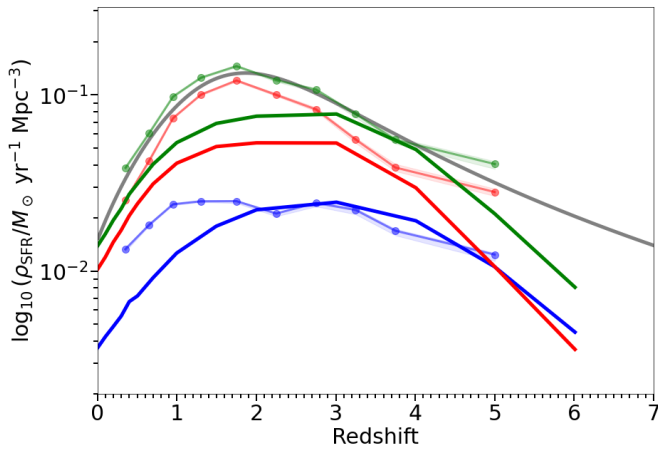
Here we compare our results with cosmological simulations. In particular, we have investigated whether current simulations are able to reproduce the contribution of the different stellar

mass bins to the total  $\rho_{\text{SFR}}$ . We chose to examine the TNG100 simulation (Nelson et al. 2019, 2018; Springel et al. 2018; Marinacci et al. 2018; Naiman et al. 2018; Pillepich et al. 2018b,a; Weinberger et al. 2017) of the IllustrisTNG project.

We retrieved the  $\rho_{\text{SFR}}$  of TNG100 using the same method as the one with which we calculate the  $\rho_{\text{SFR}}$  in this work, that is, by integrating down to  $3 \times 10^9 M_{\odot}$ . In order to compare the trends with our work, this was also done by integrating only galaxies with  $\log_{10}(M_*/M_{\odot}) > 10$  and  $\log_{10}(M_*/M_{\odot}) < 10$ . We display the evolution of  $\rho_{\text{SFR}}$  with redshift retrieved from the

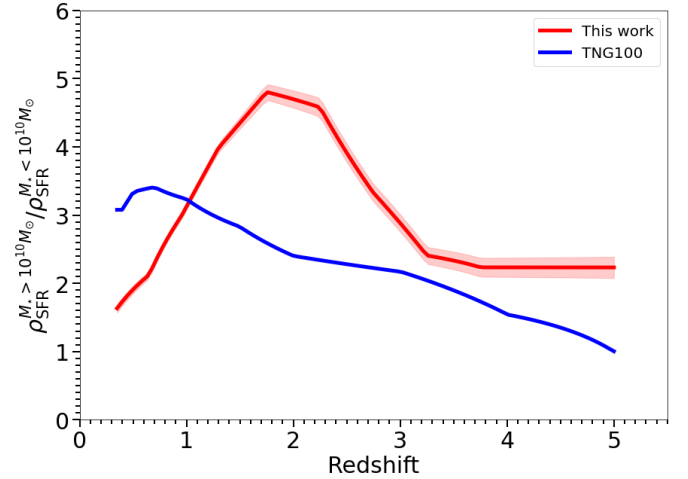


**Fig. 14.** Stellar mass above which galaxies contribute exactly 50% of the CSFD as a function of redshift.



**Fig. 15.** Cosmic star formation rate density ( $\rho_{\text{SFR}}$ ) as a function of redshift. The green, red and blue lines represent  $\rho_{\text{SFR}}$  deduced from the TNG100 simulation integrated from  $3 \times 10^9 M_{\odot}$ , including all galaxies, galaxies with  $\log_{10}(M_{\star}/M_{\odot}) > 10$ , and galaxies with  $\log_{10}(M_{\star}/M_{\odot}) < 10$ , respectively. The green, red and green faded dots represent the total  $\rho_{\text{SFR}}$  trend,  $\rho_{\text{SFR}}$  of galaxies with  $\log_{10}(M_{\star}/M_{\odot}) > 10$ , and galaxies with  $\log_{10}(M_{\star}/M_{\odot}) < 10$ , respectively (including  $H$ -dropout contribution from Wang et al. 2019), from this work. The grey line represents the trend from Madau & Dickinson (2014) for reference.

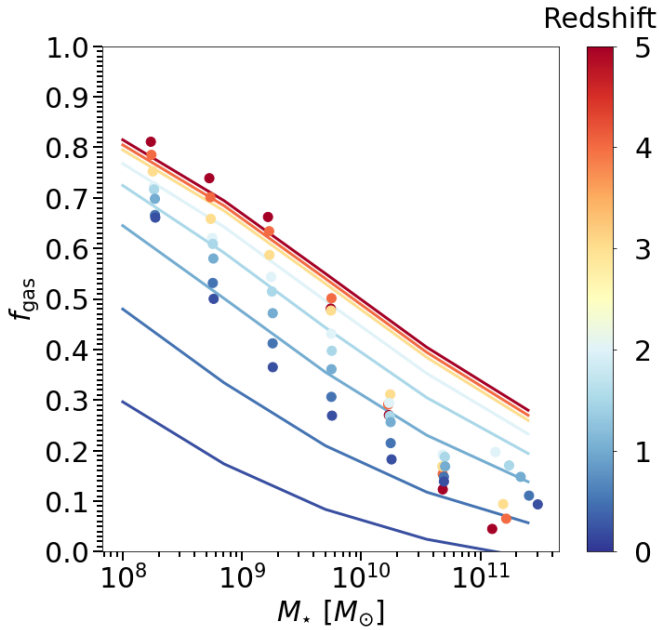
TNG100 simulation in Fig. 15. By comparing the trends of the TNG100 simulation with our results, we can see clear discrepancies as also noticed and discussed in Pillepich et al. (2018a). The  $\rho_{\text{SFR}}$  of the low masses in TNG100 (i.e.  $\log_{10}(M_{\star}/M_{\odot}) < 10$ ) is not nearly as flat as observed (especially at  $z \leq 2$ ). On the other hand, the  $\rho_{\text{SFR}}$  of the high masses in TNG100 (i.e.  $\log_{10}(M_{\star}/M_{\odot}) > 10$ ) does not account for as large a part of the total  $\rho_{\text{SFR}}$  for  $z < 3$ . The high masses account for less than half of the total  $\rho_{\text{SFR}}$  at  $z \sim 1.7$  in the simulation compared to the observations. As a result, the  $\rho_{\text{SFR}}$  peak is reached too early (i.e. at  $z \sim 3$  instead of the observed at  $z \sim 1.7$ ) in the simulation. The disparities between the two total  $\rho_{\text{SFR}}$  trends can be almost exclusively associated with the contribution of massive galaxies being off in the TNG100 simulation.



**Fig. 16.** Ratio of the  $\rho_{\text{SFR}}$  of high-mass ( $M_{\star} > 10^{10} M_{\odot}$ ) over low-mass ( $M_{\star} < 10^{10} M_{\odot}$ ) galaxies, as a function of redshift. The red line represents the ratio deduced from this work (i.e. from observations). The blue line represents the ratio deduced from the TNG100 simulation.

Another way to show the difference in behaviour between the TNG100 simulation and our observations for the two mass bins, is to look at the evolution over redshift of  $\rho_{\text{SFR}}^{M_{\star} > 10^{10}} / \rho_{\text{SFR}}^{M_{\star} < 10^{10}}$  (see Fig. 16). We can see that the ratio decreases continuously with increasing redshift in the simulation, while observations show a clear peak in the ratio around  $z \sim 1.7$  (i.e. the cosmic noon) where the amount of stars formed in high-mass galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) > 10$ ) exceeds by a factor up to  $\sim 5.6$  the one from low-mass galaxies (i.e.  $\log_{10}(M_{\star}/M_{\odot}) < 10$ ).

In addition to underestimating the contribution of massive galaxies to  $\rho_{\text{SFR}}$ , these simulations tend to recover, for SFGs, lower gas fractions (see Fig. 17) and SFR (see Fig. 18) compared to observations. In TNG100,  $f_{\text{gas}} < 15\%$  for  $M_{\star} > 10^{10} M_{\odot}$  at all redshifts (Fig. 17), as it was already addressed in Pillepich et al. (2019). Moreover, the SFR of massive individual galaxies is lower in the TNG100 simulation than in observations at, for example,  $z > 4$  (see Fig. 18). Lewis et al. (2024) shows a lack of metal content in IllustrisTNG simulations at  $z = 1$ , that might be due to AGN feedback that removes large quantities of metal-rich gas from the centres of massive galaxies. The lower metallicities observed by Lewis et al. (2024) in IllustrisTNG simulations at  $z = 1$  could be a residual memory of how AGN affected star formation at higher  $z$ , which matches the lack of star formation we observe in Fig. 18. This could indicate flaws in the way gas accretion, the efficiency of galaxies to form stars from their gas content, or the ejection of gas through galactic winds is treated in the simulations. This shows that the problems encountered in the simulations may be due to the feedback processes used to regulate star formation and the gas reservoir of galaxies. In particular, the impact of AGN feedback may be overestimated to regulate star formation and kill massive galaxies. It has been shown that AGN feedback does not directly expel the gas and quench star formation at large scales. High-resolution simulations (i.e. with maximum spatial resolution  $\sim 5\text{--}100$  pc) show that the AGN feedback tends to have little or no effect on the dense gas inside the galactic disc, because most of the out-flowing winds escape perpendicular to the galactic disc (Gabor & Bournaud 2014). AGN-driven outflows (i.e. ejective feedback), if sustained, could only quench a galaxy after a long time scale, that is, more than one Gyr (Gabor & Bournaud 2014; Biernacki & Teyssier 2018). These high-resolution simulations favour preemptive feedback



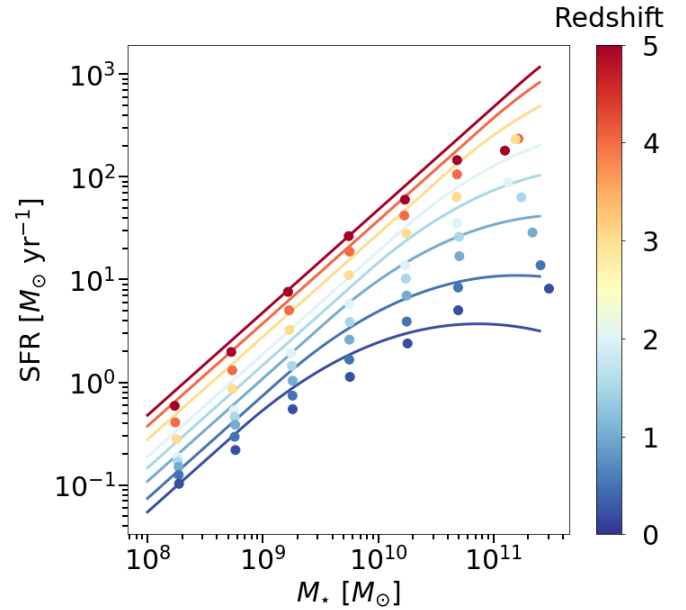
**Fig. 17.**  $f_{\text{gas}}$  as a function of  $M_*$  for SFGs in TNG100. Coloured lines represent the evolution deduced from the best-fit of  $M_{\text{gas}}$  from this work (see Table 7).

(i.e. cutting out inflows of gas into the disc), though AGN-driven winds, to quench a galaxy (Gabor & Bournaud 2014; DeGraf et al. 2017; Biernacki & Teyssier 2018). However, even strong AGN-driven winds (i.e.  $\sim 2500$  km/s) would only reduce star formation in the galaxy by a factor of 2 (DeGraf et al. 2017). The TNG100 simulation has a much lower spatial resolution (i.e. a softening length of  $\sim 0.74$  kpc at  $z = 1$ ), and is therefore not able to resolve the AGN-feedback interactions correctly. The sub-grid model of AGN feedback used in the *IllustrisTNG* has already been advocated to be responsible for a lower submillimetre galaxies (SMGs) number counts compared to observations (Hayward et al. 2021). The first *James Webb* Space Telescope (*JWST*) results interestingly go in the same direction: the number of bright, possibly massive, galaxies is found to exceed predictions at high  $z$  (Finkelstein et al. 2022; Donnan et al. 2023; Mason et al. 2023). This discrepancy on  $\rho_{\text{SFR}}$  could be a starting point to the underestimation of the global  $\rho_{\text{SFR}}$  in TNG100 for ( $1 \leq z \leq 2$ ) reported early by Donnari et al. (2019) or non-trivial issues in the simulated galaxy and halo populations found in the *Illustris* simulation (Nelson et al. 2015).

### 8.3. The SFE and Kennicutt-Schmidt relation

Through this work, we have obtained a measurement of both SFR and  $M_{\text{gas}}$  of main sequence galaxies over various redshift and  $M_*$  bins. We can thus deduce the corresponding SFE of main sequence galaxies (i.e. examine the SFR- $M_{\text{gas}}$  plane) and its evolution as a function of redshift and  $M_*$ . We display the SFE deduced from this work in Fig. 19, where, bins of redshift were regrouped into 3 bins to empathise results discussed later on.

The SFE we measured shows an increase with redshift, which is accentuated at the highest masses. To investigate whether the evolution of SFE can be explained by the gas content of galaxies at high redshift we look at the Kennicutt-Schmidt relation (Kennicutt 1998b). The Kennicutt-Schmidt relation illustrates similar properties by examining SFR and  $M_{\text{gas}}$



**Fig. 18.** SFR as a function of  $M_*$  for SFGs in TNG100. Coloured lines represent the best-fit from this work (see Table 5).

surface density (i.e.  $\Sigma_{\text{SFR}}$  and  $\Sigma_{\text{gas}}$  respectively):

$$\Sigma_{\text{SFR}} = SFR / (2\pi R_e^2), \quad (23)$$

$$\Sigma_{\text{gas}} = M_{\text{gas}} / (2\pi R_e^2). \quad (24)$$

It also links them as:

$$\log_{10}(\Sigma_{\text{SFR}}) = A + N \times \log_{10}(\Sigma_{\text{gas}}). \quad (25)$$

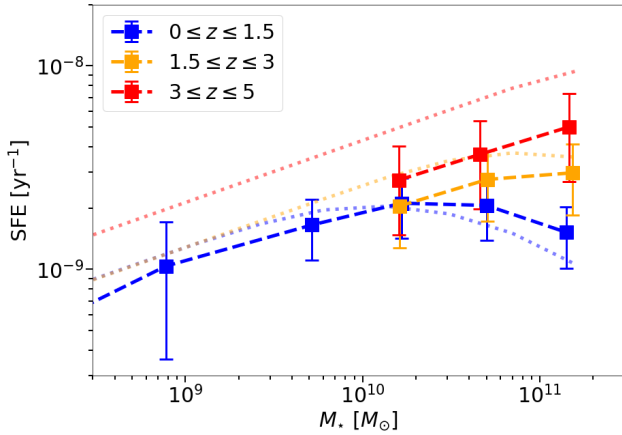
We do not probe  $R_e$  in this study, thus, we chose to use  $R_e(z, M_*)$ , for galaxies with  $M_* \geq 10^{10}$ , from Wang et al. (2022). We divided galaxies into bins of redshift ( $0 \leq z \leq 6$ ) and  $M_*$  ( $10 \leq \log_{10}(M_*) \leq 12$ ). When fitting all bins, we added the sample with low redshift and low  $M_*$  from Kennicutt (1998b). The error on the fit was generated by randomly varying the bin values within the uncertainties of the corresponding bin. We display the Schmidt-Kennicutt relation deduced from this work in Fig. 20.

All studies agree on a linear dependence of  $\log_{10}(\Sigma_{\text{SFR}})$  as a function of  $\log_{10}(\Sigma_{\text{gas}})$  with an overall slope ranging from 1 to 1.5 (e.g. Kennicutt 1998b; Wang et al. 2022), with no evolution of the slope or normalisation with redshift or  $M_*$ . Our best-fit, over all redshifts and  $M_*$  bins, also suggests a power-law scaling with a slope of  $N = 1.18^{+0.01}_{-0.01}$ .

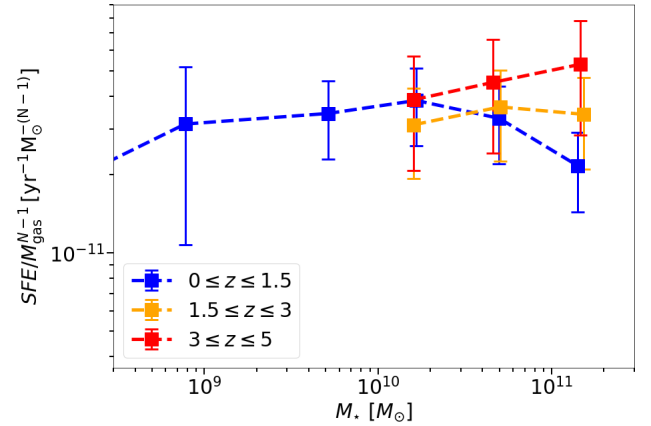
We found that  $\Sigma_{\text{SFR}} = A \times (\Sigma_{\text{SFR}})^N$  with  $N \sim 1.18$  which means that an increase in  $M_{\text{gas}}$  induces some increase in SFE (i.e.  $\text{SFE} \sim (M_{\text{gas}})^{0.18}$ ). It would not be the case if the slope of the Kennicutt-Schmidt relation was  $N = 1$  as in Tacconi et al. (2013). We can correct Fig. 19 from this effect by plotting  $\text{SFE}/M_{\text{gas}}^{(N-1)}$  instead of SFE (see Fig. 21).

At  $z \leq 2$ , in Fig. 21, the normalisation does not evolve much with redshift. On the other hand, in Fig. 7,  $M_{\text{gas}}$  increases with redshift and  $M_*$  at  $z \leq 2$ . This means that an increase in SFR as a function of redshift or stellar mass, at  $z \leq 2$ , is mainly due to the variation of  $M_{\text{gas}}$ .

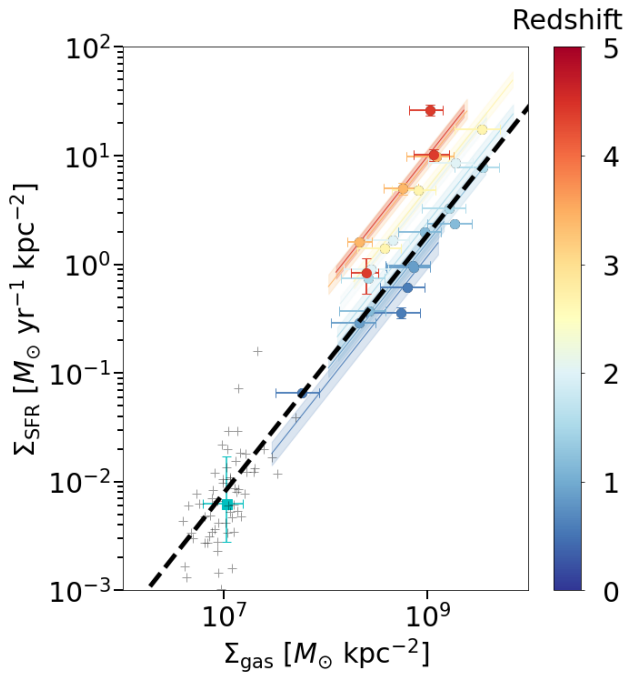
At  $z > 2$ , in Fig. 21, the normalisation increases with redshift: by a factor  $\sim 3$  between  $z \sim 2$  and  $z \sim 4.3$ . On the other hand, in Fig. 7,  $M_{\text{gas}}$  remains roughly constant for  $z > 2$ . This



**Fig. 19.** SFE as a function of stellar mass. Squares linked with a dashed line represent the measured from this work, colour coded by bin of redshift. Dotted lines are the value SFE deduced from our best fit law of the SFR and  $M_{\text{gas}}$  main sequences.



**Fig. 21.**  $\text{SFE}/M_{\text{gas}}^{(N-1)}$  ( $N = 1.18$ ) as a function of stellar mass. Squares linked with a dashed line represent the measured from this work, colour coded by bin of redshift.



**Fig. 20.** Kennicutt-Schmidt relation:  $\Sigma_{\text{SFR}}$  as a function of  $\Sigma_{\text{gas}}$ . The dots represent the data from this work, and are colour coded by redshift. The cyan squared error bar represents the median of the sample distribution from Kennicutt (1998b), with individual galaxies displayed as grey crosses. The dashed black line represents the best-fit of all points. Coloured lines represent the best-fit, by fixing the slope to the all points best-fit value (i.e.  $N = 1.18_{-0.01}^{+0.01}$ ), at fix redshift.

means that the SFE plays a non-negligible role in increasing the normalisation of the main sequence SFR between  $2 \leq z \leq 4.3$ .

In Fig. 21, we can see that the increase of SFE seen in Fig. 19 cannot be explained by the amount of gas as it is still appearing in Fig. 21.

Overall, it means that, we need to differentiate two effects: an increase of SFE due to the increase of  $M_{\text{gas}}$ , and an intrinsic increase in SFE. For example, at fixed stellar mass for  $10.5 \leq \log_{10}(M_{\star}/M_{\odot}) \leq 11$  between  $z = 1.15$  and  $z = 2$ , SFR increases by a factor  $\sim 4$  (see Fig. 5), while  $M_{\text{gas}}$  increase by a factor  $\sim 1.6$  (see Fig. 7): it means that SFE increases by

a factor  $\sim 2.5$ , which reflects the increase of the normalisation in Fig. 19 by  $\sim 2.5$  for the same redshift range. The first effect in an increase of SFE due to the increase of  $M_{\text{gas}}$  (i.e. resulting from the Kennicutt-Schmidt relation of slope  $N \sim 1.18$ ):  $\alpha_{\text{SFE}}^{\text{due to } M_{\text{gas}}} \sim \alpha_{M_{\text{gas}}}^{N-1} = (1.6)^{0.18} \sim 1.09$ . The second effect is an intrinsic increase in SFE between the two redshifts (i.e. independent of the increase of  $M_{\text{gas}}$ ):  $\alpha_{\text{SFE}}^{\text{intrinsic}} \sim \alpha_{\text{SFE}}^{\text{due to } M_{\text{gas}}} \sim 2.3$ . Finally, we could conclude by saying that a factor  $\sim 2.3$  is solely due to an increase of SFE, and a factor  $\sim \alpha_{M_{\text{gas}}} \times \alpha_{\text{SFE}}^{\text{due to } M_{\text{gas}}} = 1.75$  is due to the increase of  $M_{\text{gas}}$ .

We note that this effect is not inherent to our specific data points, or even best-fit evolution laws of SFR and  $M_{\text{gas}}$ , as a similar evolution could be observed if we were to generate catalogues from the laws of S15 for SFR, and Wang et al. (2022) or Tacconi et al. (2018) for  $M_{\text{gas}}$  (following the method used to generate  $\rho_{\text{SFR}}$  and  $\rho_{\text{gas}}$  see Sects. 6 and 7). If we were to choose laws from Speagle et al. (2014) for SFR and Tacconi et al. (2018) for  $M_{\text{gas}}$ , this effect would disappear completely, and we would end up observing an evolution of the SFE with redshift. We conclude, in this case, that the correct normalisation and the presence of the bending of the main sequence (which are lacking in the study of Speagle et al. 2014) are essential to observe this effect. One could advocate that the observation of an increase in efficiency with redshift when generating catalogues, for example, from S15 for SFR, and Tacconi et al. (2018) for  $M_{\text{gas}}$ , comes from the fact that the  $M_{\text{gas}}$  fitting form from Tacconi et al. (2018), or the one use in this work, do not allow a bending or an evolution of it. However, we found no evidence of a bending or its evolution allowing them. Overall, it suggests that some physical processes limit the star formation efficiency in low-redshift galaxies, as compared to high-redshift galaxies, beyond the simple fact that they have less available gas.

Overall, even though the evolution of the SFE with redshift is not well constrained, this study suggests that galaxies at high-redshift form stars more efficiently at a fixed gas surface density. This excess cannot be explained by a universal Kennicutt-Schmidt relation nor SFE over all redshifts. This may partly explain the early results of JWST: a higher  $\rho_{\text{UV}}$  at  $z \geq 8$  (Finkelstein et al. 2023; Donnan et al. 2023; Mason et al. 2023) compared to pre-JWST extrapolation Oesch et al. (2018); more massive galaxies at  $z \geq 5$  (Labbe et al. 2023; Xiao et al. 2023a). This may

be explained by the model proposed by Dekel et al. (2023) who showed that the environment of the most massive dark-matter haloes at  $z \geq 10$  favour high star formation efficiencies.

In conclusion, the gradual appearance of the bending of the main sequence with redshift appears to result from a reduction of the SFE, that is, a gradual disappearance of this extra factor affecting the SFE (observed at  $z \sim 4.3$  and absent at  $z \sim 2$ ). Then, at a fixed lower redshift (i.e.  $z < 2$ ), the bending of the main sequence (as  $M_\star$  increases) is mainly due to an increase in  $M_{\text{gas}}$  with stellar mass, and to second order, to a decrease in SFE with stellar mass. This characteristic is also observed in the local universe (i.e. for  $0.01 < z < 0.05$ ) in Saintonge et al. (2016): when the stellar mass increases by a factor 10, the SFR increases by a factor  $\sim 1.4$ , which is due to an increase of  $M_{\text{gas}}$  by a factor  $\sim 1.8$  and a decrease in SFE by a factor  $\sim 1.25$ . Saintonge et al. (2016) reach the same conclusion in the local universe: at a fixed redshift,  $M_{\text{gas}}$  is the main reason for the bending of the main sequence. The slow downfall of the star formation efficiency in the most massive galaxies (i.e.  $M_\star \geq 10^{11} M_\odot$ ) observed between  $0.1 \leq z \leq 4.3$ , was also observed in Schreiber et al. (2016). It seems that a discontinuation of gas accretion cannot be the only reason for the bending of the SFR main sequence (Daddi et al. 2022). It must be coupled with other effects that prevent the gas in high  $M_\star$  galaxies from cooling and collapsing into stars (i.e. an evolution of the SFE). The main driver of the bending could be the quenching of galaxies due to environmental effects, as massive galaxies populate denser environments where they suffer from ram pressure stripping, galaxy harassment (e.g. Kalita et al. 2022).

## 9. Summary and conclusions

For this work we gathered catalogues of  $H$ -band selected SFGs over four fields (GOODS-South, GOODS-North, COSMOS, UDS). This sample is stacked over seven wavelengths (24  $\mu\text{m}$ , 100  $\mu\text{m}$ , 160  $\mu\text{m}$ , 250  $\mu\text{m}$ , 350  $\mu\text{m}$ , 500  $\mu\text{m}$  and 1.13 mm). The main addition to this work, compared to previous studies, is the use of the GOODS-ALMA survey (Gómez-Guijarro et al. 2022b), which helps to reach lower  $M_\star$  and higher redshifts, when combined with *Herschel*. We combined the stacking of  $H$ -band selected galaxies with  $H$ -dropout galaxy properties to obtain a complete view of the cosmic history of galaxies. We derived the evolution of MS galaxy properties such as SFR,  $T_{\text{dust}}$ , and  $M_{\text{gas}}$  through a consistent analysis. We generated a detailed view of  $\rho_{\text{SFR}}$  and  $\rho_{\text{gas}}$  depending on key galaxy properties such as  $M_\star$ ,  $L_{\text{IR}}$  and  $\tau_{\text{dep}}$ .

This study allowed us to confirm several features of galaxies, the main ones being summarised as follows:

- The  $T_{\text{dust}}$  evolves linearly with redshift. There is no clear evidence for a dependence of  $T_{\text{dust}}$  depending on  $M_\star$  at a fixed redshift, when considering MS galaxies.
- The SFR- $M_\star$  MS has a slope close to unity at low masses (i.e.  $M_\star \lesssim 10^{10} - 10^{11} M_\odot$ ), with a bending appearing below  $z \sim 2$  on the high-mass end. The bending of the MS corresponds to the slow downfall of SFE in massive galaxies.
- The total  $\rho_{\text{SFR}}$  follows a form close to the one presented in Madau & Dickinson (2014), at least up to  $z = 5$ . As a result, the contribution of  $H$ -dropout galaxies to the total  $\rho_{\text{SFR}}$  is  $\sim 13\%$  at  $z = 5$ .

The main new results from this paper can be summarised as follows:

- We measured the contribution of galaxies of different  $M_\star$  to the total  $\rho_{\text{SFR}}$  over cosmic time. We find that massive galaxies (i.e.  $M_\star \geq 10^{10} M_\odot$ ) account for most of the  $\rho_{\text{SFR}}$  up to

$z \sim 5$ . Low-mass galaxies (i.e.  $M_\star \leq 10^{10} M_\odot$ ) contribute a roughly constant fraction to the total  $\rho_{\text{SFR}}$  at  $0.4 \leq z \leq 5$ .

- The global discrepancy between the cosmic SFR density observed and the TNG100 one is discussed in Pillepich et al. (2018b). However, we show that the TNG100 simulation fails to reproduce the contributions of the different bins of  $M_\star$  to the total  $\rho_{\text{SFR}}$ . In particular, in the TNG100 simulation, massive galaxies do not form enough stars around cosmic noon, when the cosmic SFR density peaks (i.e.  $\rho_{\text{SFR-TNG100}}^{M_\star \geq 10^{10} M_\odot} \simeq 50\% \rho_{\text{SFR-observed}}^{M_\star \geq 10^{10} M_\odot}$  at  $z = 2$ ). We conclude that the mechanisms used to regulate star formation (in particular feedback from AGNs) are possibly too strong in this simulation, namely too efficient at producing a drop in the gas content in massive galaxies.
- We estimate that  $H$ -dropout galaxies (as defined in Wang et al. 2019) contribute to  $\sim 23\%$  of the  $\rho_{\text{SFR}}$  of galaxies with  $M_\star \geq 10^{10.3} M_\odot$  at  $z = 4-6$ .
- We find hints that high-redshift galaxies convert their gas reservoirs more efficiently than local galaxies, that is, suggesting a possible evolution of SFE with redshift.

*Acknowledgements.* This work was supported by the Programme National Cosmology et Galaxies (PNCG) of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES. C.G.G. acknowledges support from CNES.

## References

- Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, *ApJS*, 243, 22
- Bauermeister, A., Blitz, L., & Ma, C.-P. 2010, *ApJ*, 717, 323
- Bavouzet, N., Dole, H., Le Floch, E., et al. 2008, *A&A*, 479, 83
- Berta, S., Lutz, D., Genzel, R., Förster-Schreiber, N. M., & Tacconi, L. J. 2016, *A&A*, 587, A73
- Béthermin, M., Dole, H., Cousin, M., & Bavouzet, N. 2010, *A&A*, 516, A43
- Béthermin, M., Le Floch, E., Ilbert, O., et al. 2012, *A&A*, 542, A58
- Béthermin, M., Kilbinger, M., Daddi, E., et al. 2014, *A&A*, 567, A103
- Béthermin, M., Daddi, E., Magdis, G., et al. 2015, *A&A*, 573, A113
- Béthermin, M., Fudamoto, Y., Ginolfi, M., et al. 2020, *A&A*, 643, A2
- Biernacki, P., & Teyssier, R. 2018, *MNRAS*, 475, 5688
- Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, 136, 2846
- Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, *ARA&A*, 51, 207
- Bourne, N., Maddox, S. J., Dunne, L., et al. 2012, *MNRAS*, 421, 3027
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012a, *ApJ*, 754, 83
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012b, *ApJ*, 752, L5
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015a, *ApJ*, 811, 140
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015b, *ApJ*, 803, 34
- Bouwens, R., González-López, J., Aravena, M., et al. 2020, *ApJ*, 902, 112
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, *ApJ*, 429, 582
- Casey, C. M., Zavala, J. A., Spilker, J., et al. 2018, *ApJ*, 862, 77
- Chabrier, G. 2003, *PASP*, 115, 763
- da Cunha, E., Groves, B., Walter, F., et al. 2013, *ApJ*, 766, 13
- Daddi, E., Cimatti, A., Renzini, A., et al. 2004, *ApJ*, 617, 746
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156
- Daddi, E., Elbaz, D., Walter, F., et al. 2010, *ApJ*, 714, L118
- Daddi, E., Delvecchio, I., Dimauro, P., et al. 2022, *A&A*, 661, L7
- Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, *A&A*, 605, A70
- DeGraf, C., Dekel, A., Gabor, J., & Bournaud, F. 2017, *MNRAS*, 466, 1462
- Dekel, A., Sarkar, K. S., Birnboim, Y., Mandelker, N., & Li, Z. 2023, *MNRAS*, 523, 3201
- de los Reyes, M. A. C., & Kennicutt, R. C. Jr 2019, *ApJ*, 872, 16
- Delvecchio, I., Daddi, E., Sargent, M. T., et al. 2021, *A&A*, 647, A123
- Donnan, C. T., McLeod, D. J., Dunlop, J. S., et al. 2023, *MNRAS*, 518, 6011
- Donnari, M., Pillepich, A., Nelson, D., et al. 2019, *MNRAS*, 485, 4817
- Draine, B. T., & Li, A. 2007, *ApJ*, 657, 810
- Draine, B. T., Aniano, G., Krause, O., et al. 2014, *ApJ*, 780, 172
- Drew, P. M., & Casey, C. M. 2022, *ApJ*, 930, 142
- Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, *A&A*, 468, 33
- Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, *A&A*, 533, A119
- Elbaz, D., Leiton, R., Nagar, N., et al. 2018, *A&A*, 616, A110
- Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, *ApJ*, 644, 813

- Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2022, CEERS Key Paper I: An Early Look into the First 500 Myr of Galaxy Formation with JWST
- Finkelstein, S. L., Bagley, M. B., Ferguson, H. C., et al. 2023, *ApJ*, **946**, L13
- Franco, M., Elbaz, D., Béthermin, M., et al. 2018, *A&A*, **620**, A152
- Franco, M., Elbaz, D., Zhou, L., et al. 2020, *A&A*, **643**, A53
- Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, *MNRAS*, **366**, 767
- Gabor, J. M., & Bournaud, F. 2014, *MNRAS*, **441**, 1615
- Galametz, A., Grazian, A., Fontana, A., et al. 2013, *ApJS*, **206**, 10
- Galliano, F., Hony, S., Bernard, J. P., et al. 2011, *A&A*, **536**, A88
- Galliano, F., Nersesian, A., Bianchi, S., et al. 2021, *A&A*, **649**, A18
- Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, *ApJ*, **800**, 20
- Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022a, *A&A*, **658**, A43
- Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022b, *A&A*, **659**, A196
- Gómez-Guijarro, C., Magnelli, B., Elbaz, D., et al. 2023, *A&A*, **677**, A34
- Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *A&A*, **518**, L3
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35
- Groves, B. A., Schinnerer, E., Leroy, A., et al. 2015, *ApJ*, **799**, 96
- Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, *A&A*, **643**, A8
- Guo, K., Zheng, X. Z., & Fu, H. 2013, *ApJ*, **778**, 23
- Hao, L., Strauss, M. A., Fan, X., et al. 2005, *AJ*, **129**, 1795
- Harrington, K. C., Weiss, A., Yun, M. S., et al. 2021, *ApJ*, **908**, 95
- Hatziminaoglou, E., Omont, A., Stevens, J. A., et al. 2010, *A&A*, **518**, L33
- Hayward, C. C., Sparre, M., Chapman, S. C., et al. 2021, *MNRAS*, **502**, 2922
- Hodge, J. A., & da Cunha, E. 2020, *R. Soc. Open Sci.*, **7**, 200556
- Kaasinen, M., Scoville, N., Walter, F., et al. 2019, *ApJ*, **880**, 15
- Kalita, B. S., Daddi, E., Bournaud, F., et al. 2022, *A&A*, **666**, A44
- Kennicutt, R. C., Jr 1998a, *ARA&A*, **36**, 189
- Kennicutt, R. C., Jr 1998b, *ApJ*, **498**, 541
- Kewley, L. J., & Dopita, M. A. 2002, *ApJS*, **142**, 35
- Kewley, L. J., & Ellison, S. L. 2008, *ApJ*, **681**, 1183
- Khusanova, Y., Béthermin, M., Le Fèvre, O., et al. 2021, *A&A*, **649**, A152
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, **197**, 36
- Kurczynski, P., & Gawiser, E. 2010, *AJ*, **139**, 1592
- Labbe, I., van Dokkum, P., Nelson, E., et al. 2023, *Nature*, **616**, 266
- Lagache, G. 2018, in *Peering towards Cosmic Dawn*, eds. V. Jelić, & T. van der Hulst, 333, 228
- Lagos, C. d. P., Crain, R. A., Schaye, J., et al. 2015, *MNRAS*, **452**, 3815
- Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, *ApJ*, **801**, 80
- LeFloch, E., Aussel, H., Ilbert, O., et al. 2009, *ApJ*, **703**, 222
- Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, **136**, 2782
- Leroy, A. K., Bolatto, A., Gordon, K., et al. 2011, *ApJ*, **737**, 12
- Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, *ApJ*, **899**, 58
- Lewis, Z. J., Andrews, B. H., Bezanson, R., et al. 2024, *ApJ*, **964**, 59
- Liu, D., Daddi, E., Dickinson, M., et al. 2018, *ApJ*, **853**, 172
- Liu, D., Schinnerer, E., Groves, B., et al. 2019, *ApJ*, **887**, 235
- Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, *A&A*, **532**, A90
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Madau, P., Pozzetti, L., & Dickinson, M. 1998, *ApJ*, **498**, 106
- Magdis, G. E., Daddi, E., Elbaz, D., et al. 2011, *ApJ*, **740**, L15
- Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, *ApJ*, **760**, 6
- Magdis, G. E., Rigopoulou, D., Daddi, E., et al. 2017, *A&A*, **603**, A93
- Magnelli, B., Elbaz, D., Chary, R. R., et al. 2009, *A&A*, **496**, 57
- Magnelli, B., Saintonge, A., Lutz, D., et al. 2012, *A&A*, **548**, A22
- Magnelli, B., Popesso, P., Berta, S., et al. 2013, *A&A*, **553**, A132
- Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, *A&A*, **561**, A86
- Magnelli, B., Boogaard, L., Decarli, R., et al. 2020, *ApJ*, **892**, 66
- Mannucci, F., Cresci, G., Maiolino, R., Marconi, A., & Gnerucci, A. 2010, *MNRAS*, **408**, 2115
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, *MNRAS*, **480**, 5113
- Mason, C. A., Trenti, M., & Treu, T. 2023, *MNRAS*, **521**, 497
- Millard, J. S., Eales, S. A., Smith, M. W. L., et al. 2020, *MNRAS*, **494**, 293
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013, *ApJ*, **777**, 18
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **477**, 1206
- Nayyeri, H., Hemmati, S., Mobasher, B., et al. 2017, *ApJS*, **228**, 7
- Nelson, D., Pillepich, A., Genel, S., et al. 2015, *Astron. Comput.*, **13**, 12
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *Comput. Astrophys. Cosmol.*, **6**, 2
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJ*, **660**, L43
- Novak, M., Smolčić, V., Delhaize, J., et al. 2017, *A&A*, **602**, A5
- Oesch, P. A., Bouwens, R. J., Illingworth, G. D., Labbé, I., & Stefanon, M. 2018, *ApJ*, **855**, 105
- Pannella, M., Elbaz, D., Daddi, E., et al. 2015, *ApJ*, **807**, 141
- Pettini, M., & Pagel, B. E. J. 2004, *MNRAS*, **348**, L59
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, *MNRAS*, **475**, 648
- Pillepich, A., Springel, V., Nelson, D., et al. 2018b, *MNRAS*, **473**, 4077
- Pillepich, A., Nelson, D., Springel, V., et al. 2019, *MNRAS*, **490**, 3196
- Popesso, P., Magnelli, B., Buttiglione, S., et al. 2012, *A&A*, submitted [arXiv:1211.4257]
- Popesso, P., Concas, A., Cresci, G., et al. 2023, *MNRAS*, **519**, 1526
- Reddy, N. A., Steidel, C. C., Erb, D. K., Shapley, A. E., & Pettini, M. 2006, *ApJ*, **653**, 1004
- Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2014, *A&A*, **563**, A31
- Richards, G. T., Lacy, M., Storrie-Lombardi, L. J., et al. 2006, *ApJS*, **166**, 470
- Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016, *MNRAS*, **461**, 1100
- Saintonge, A., Catinella, B., Cortese, L., et al. 2016, *MNRAS*, **462**, 1749
- Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, *ApJS*, **233**, 22
- Salpeter, E. E. 1955, *ApJ*, **121**, 161
- Santini, P., Fontana, A., Grazian, A., et al. 2012, *A&A*, **538**, A33
- Sargent, M. T., Béthermin, M., Daddi, E., & Elbaz, D. 2012, *ApJ*, **747**, L31
- Schenker, M. A., Robertson, B. E., Ellis, R. S., et al. 2013, *ApJ*, **768**, 196
- Schinnerer, E., Groves, B., Sargent, M. T., et al. 2016, *ApJ*, **833**, 112
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, **575**, A74
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2016, *A&A*, **589**, A35
- Schreiber, C., Elbaz, D., Pannella, M., et al. 2018, *A&A*, **609**, A30
- Scoville, N., Aussel, H., Sheth, K., et al. 2014, *ApJ*, **783**, 84
- Scoville, N., Sheth, K., Aussel, H., et al. 2016, *ApJ*, **820**, 83
- Scoville, N., Lee, N., Vanden Bout, P., et al. 2017, *ApJ*, **837**, 150
- Shirley, R., Duncan, K., Campos Varillas, M. C., et al. 2021, *MNRAS*, **507**, 129
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, **214**, 15
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, **475**, 676
- Steidel, C. C., Adelberger, K. L., Giavalisco, M., Dickinson, M., & Pettini, M. 1999, *ApJ*, **519**, 1
- Suess, K. A., Kriek, M., Price, S. H., & Barro, G. 2019, *ApJ*, **885**, L22
- Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, *Nature*, **463**, 781
- Tacconi, L. J., Neri, R., Genzel, R., et al. 2013, *ApJ*, **768**, 74
- Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, *ApJ*, **853**, 179
- Traina, A., Gruppioni, C., Delvecchio, I., et al. 2024, *A&A*, **681**, A118
- Viero, M. P., Moncelsi, L., Quadri, R. F., et al. 2013, *ApJ*, **779**, 32
- Wang, T., Schreiber, C., Elbaz, D., et al. 2019, *Nature*, **572**, 211
- Wang, T.-M., Magnelli, B., Schinnerer, E., et al. 2022, *A&A*, **660**, A142
- Weinberger, R., Springel, V., Hernquist, L., et al. 2017, *MNRAS*, **465**, 3291
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJ*, **754**, L29
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *ApJ*, **795**, 104
- White, R. L., Helfand, D. J., Becker, R. H., Glikman, E., & de Vries, W. 2007, *ApJ*, **654**, 99
- Xiao, M., Oesch, P., Elbaz, D., et al. 2023a, ArXiv e-prints [arXiv:2309.02492]
- Xiao, M. Y., Elbaz, D., Gómez-Guijarro, C., et al. 2023b, *A&A*, **672**, A18
- Zhou, L., Elbaz, D., Franco, M., et al. 2020, *A&A*, **642**, A155

## Appendix A: Table of gas masses

Table A.1. Summary of the gas masses derived in this work.

Redshift bin	$M_\star$ bin	$S_{1.13\text{ mm}}$ [ $\mu\text{Jy}$ ]	$M_{\text{gas}}^{\text{Scoville16}}$ [ $M_\odot$ ]	$M_{\text{dust}}$ [ $M_\odot$ ]	$12 + \log_{10}(O/H)$	$M_{\text{gas}}^{\text{MZR}}$ [ $M_\odot$ ]
$0.1 \leq z \leq 0.4$	$8.5 \leq \log_{10}(M_\star) \leq 9.5$			$6.99^{+0.28}_{-0.06}$	$8.31^{+0.2}_{-0.2}$	$9.31^{+0.34}_{-0.21}$
$0.1 \leq z \leq 0.4$	$9.5 \leq \log_{10}(M_\star) \leq 10.0$			$7.5^{+0.04}_{-0.06}$	$8.58^{+0.2}_{-0.2}$	$9.55^{+0.2}_{-0.21}$
$0.1 \leq z \leq 0.4$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$			$7.88^{+0.06}_{-0.09}$	$8.68^{+0.2}_{-0.2}$	$9.83^{+0.21}_{-0.22}$
$0.1 \leq z \leq 0.4$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$			$8.1^{+0.1}_{-0.03}$	$8.72^{+0.2}_{-0.2}$	$10.01^{+0.22}_{-0.2}$
$0.1 \leq z \leq 0.4$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$			$8.12^{+0.55}_{-0.11}$	$8.74^{+0.2}_{-0.2}$	$10.01^{+0.58}_{-0.23}$
$0.4 \leq z \leq 0.7$	$9.5 \leq \log_{10}(M_\star) \leq 10.0$			$7.49^{+0.04}_{-0.03}$	$8.48^{+0.2}_{-0.2}$	$9.63^{+0.2}_{-0.2}$
$0.4 \leq z \leq 0.7$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$	$0.1^{+0.03}_{-0.03}$	$9.91^{+0.13}_{-0.11}$	$7.82^{+0.02}_{-0.01}$	$8.61^{+0.2}_{-0.2}$	$9.83^{+0.2}_{-0.2}$
$0.4 \leq z \leq 0.7$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.12^{+0.05}_{-0.04}$	$9.98^{+0.19}_{-0.16}$	$8.13^{+0.02}_{-0.04}$	$8.69^{+0.2}_{-0.2}$	$10.07^{+0.2}_{-0.2}$
$0.4 \leq z \leq 0.7$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$			$8.29^{+0.11}_{-0.13}$	$8.73^{+0.2}_{-0.2}$	$10.19^{+0.22}_{-0.24}$
$0.7 \leq z \leq 1.0$	$9.5 \leq \log_{10}(M_\star) \leq 10.0$			$7.47^{+0.09}_{-0.08}$	$8.4^{+0.2}_{-0.2}$	$9.69^{+0.22}_{-0.21}$
$0.7 \leq z \leq 1.0$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$			$7.89^{+0.03}_{-0.04}$	$8.55^{+0.2}_{-0.2}$	$9.97^{+0.2}_{-0.2}$
$0.7 \leq z \leq 1.0$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.14^{+0.04}_{-0.05}$	$10.16^{+0.13}_{-0.14}$	$8.25^{+0.04}_{-0.02}$	$8.65^{+0.2}_{-0.2}$	$10.23^{+0.2}_{-0.2}$
$0.7 \leq z \leq 1.0$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$			$8.38^{+0.04}_{-0.02}$	$8.71^{+0.2}_{-0.2}$	$10.29^{+0.2}_{-0.2}$
$1.0 \leq z \leq 1.3$	$9.5 \leq \log_{10}(M_\star) \leq 10.0$			$7.51^{+0.2}_{-0.19}$	$8.32^{+0.2}_{-0.2}$	$9.81^{+0.28}_{-0.28}$
$1.0 \leq z \leq 1.3$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$			$7.88^{+0.02}_{-0.07}$	$8.48^{+0.2}_{-0.2}$	$10.02^{+0.2}_{-0.21}$
$1.0 \leq z \leq 1.3$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.12^{+0.02}_{-0.03}$	$10.11^{+0.09}_{-0.1}$	$8.17^{+0.02}_{-0.01}$	$8.62^{+0.2}_{-0.2}$	$10.18^{+0.2}_{-0.2}$
$1.0 \leq z \leq 1.3$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$			$8.7^{+0.02}_{-0.04}$	$8.69^{+0.2}_{-0.2}$	$10.64^{+0.2}_{-0.2}$
$1.3 \leq z \leq 1.8$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$			$7.7^{+0.01}_{-0.0}$	$8.41^{+0.2}_{-0.2}$	$9.91^{+0.2}_{-0.2}$
$1.3 \leq z \leq 1.8$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.15^{+0.03}_{-0.03}$	$10.25^{+0.09}_{-0.09}$	$8.29^{+0.01}_{-0.05}$	$8.56^{+0.2}_{-0.2}$	$10.36^{+0.2}_{-0.2}$
$1.3 \leq z \leq 1.8$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$	$0.43^{+0.09}_{-0.1}$	$10.71^{+0.09}_{-0.1}$	$8.71^{+0.02}_{-0.01}$	$8.65^{+0.2}_{-0.2}$	$10.68^{+0.2}_{-0.2}$
$1.8 \leq z \leq 2.3$	$9.5 \leq \log_{10}(M_\star) \leq 10.0$			$7.7^{+0.04}_{-0.1}$	$8.13^{+0.2}_{-0.2}$	$10.19^{+0.2}_{-0.22}$
$1.8 \leq z \leq 2.3$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$	$0.1^{+0.03}_{-0.02}$	$10.04^{+0.11}_{-0.09}$	$7.72^{+0.06}_{-0.03}$	$8.34^{+0.2}_{-0.2}$	$10.0^{+0.21}_{-0.2}$
$1.8 \leq z \leq 2.3$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.21^{+0.03}_{-0.03}$	$10.39^{+0.06}_{-0.06}$	$8.25^{+0.01}_{-0.02}$	$8.51^{+0.2}_{-0.2}$	$10.37^{+0.2}_{-0.2}$
$1.8 \leq z \leq 2.3$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$	$0.46^{+0.08}_{-0.04}$	$10.73^{+0.07}_{-0.04}$	$8.58^{+0.02}_{-0.04}$	$8.62^{+0.2}_{-0.2}$	$10.58^{+0.2}_{-0.2}$
$2.3 \leq z \leq 3.1$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$			$7.71^{+0.02}_{-0.06}$	$8.25^{+0.2}_{-0.2}$	$10.09^{+0.2}_{-0.21}$
$2.3 \leq z \leq 3.1$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.32^{+0.05}_{-0.04}$	$10.55^{+0.07}_{-0.05}$	$8.14^{+0.03}_{-0.04}$	$8.43^{+0.2}_{-0.2}$	$10.33^{+0.2}_{-0.2}$
$2.3 \leq z \leq 3.1$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$	$0.9^{+0.18}_{-0.15}$	$11.0^{+0.09}_{-0.07}$	$8.66^{+0.08}_{-0.03}$	$8.57^{+0.2}_{-0.2}$	$10.71^{+0.21}_{-0.2}$
$3.1 \leq z \leq 3.9$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$	$0.08^{+0.02}_{-0.02}$	$9.95^{+0.1}_{-0.12}$			
$3.1 \leq z \leq 3.9$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.29^{+0.16}_{-0.21}$	$10.49^{+0.23}_{-0.32}$	$7.91^{+0.12}_{-0.02}$	$8.36^{+0.2}_{-0.2}$	$10.17^{+0.23}_{-0.2}$
$3.1 \leq z \leq 3.9$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$	$0.91^{+0.37}_{-0.37}$	$10.98^{+0.17}_{-0.17}$	$8.57^{+0.08}_{-0.09}$	$8.51^{+0.2}_{-0.2}$	$10.69^{+0.21}_{-0.22}$
$3.9 \leq z \leq 5.0$	$10.0 \leq \log_{10}(M_\star) \leq 10.5$	$0.11^{+0.03}_{-0.04}$	$10.07^{+0.13}_{-0.17}$			
$3.9 \leq z \leq 5.0$	$10.5 \leq \log_{10}(M_\star) \leq 11.0$	$0.37^{+0.19}_{-0.15}$	$10.59^{+0.23}_{-0.18}$	$8.06^{+0.2}_{-0.12}$	$8.28^{+0.2}_{-0.2}$	$10.4^{+0.28}_{-0.23}$
$3.9 \leq z \leq 5.0$	$11.0 \leq \log_{10}(M_\star) \leq 12.0$	$0.7^{+0.34}_{-0.34}$	$10.86^{+0.21}_{-0.21}$	$8.35^{+0.01}_{-0.13}$	$8.5^{+0.2}_{-0.2}$	$10.47^{+0.2}_{-0.24}$

Notes. These gas masses do not include the possible contribution of  $H$ -dropout.



## Appendix B: Using only the Scoville16 method

In this appendix, we look at the results that would be deduced if we only considered gas masses that could be deduced from our ALMA measurements and using the [Scoville et al. \(2016\)](#) method to derive  $M_{\text{gas}}$ . This appendix goes through the accordingly updated Figures 7, 11, 12, 19, 20 and 21.

We follow the same method we went through in Sect. 5.3, we add the contribution of the  $H$ -dropouts to  $M_{\text{gas}}$ . In that case, Eq. 20 allows to get an acceptable fit. Therefore, we chose to fit the final  $M_{\text{gas}}$  with Eq. 20. Our best fit parameters are given in Table B.1.

**Table B.1.** Best-fit parameters of  $M_{\text{gas}}$  evolution of the main sequence.

$A$	$B$	$F$	$D$	$\beta^*$
$0.136^{+0.235}_{-0.103}$	$-1.68^{+0.56}_{-0.63}$	$0.897^{+0.151}_{-0.282}$	$-0.321^{+0.090}_{-0.088}$	2

**Notes.** Best-fit parameters of  $M_{\text{gas}}$  evolution of the main sequence using Eq. 20 when using the method from [Scoville et al. \(2016\)](#) only. \* indicates that this parameter was fixed during the fit.

Our new gas masses measurements are deduced from the method from [Scoville et al. \(2016\)](#) are displayed in Fig. B.1. In Fig. B.1, we also compare our results to the trends observed by [Tacconi et al. \(2018\)](#) and [Wang et al. \(2022\)](#). In Fig. B.1, the trends are displayed for stellar mass bins in the [Salpeter \(1955\)](#) IMF. Using this method, our results coincide with the one from [Wang et al. \(2022\)](#), which is to be expected as they are both coming from a mass selected sample, deducing gas masses from millimetre measurements through the method of [Scoville et al. \(2016\)](#).

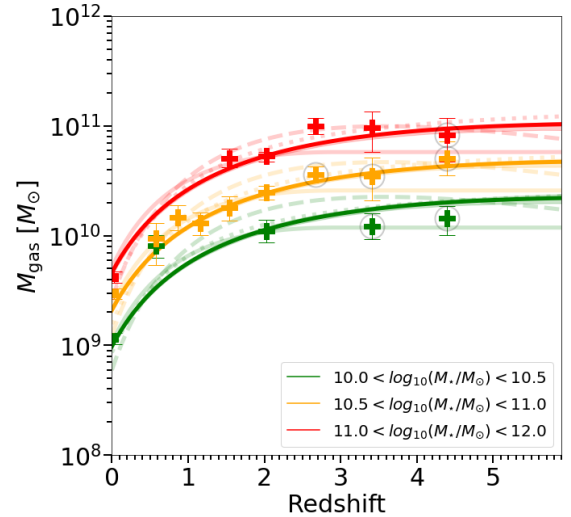
Next, the resulting gas mass densities are deduced through the method we describe in Sect. 7. We display the results in Fig. B.2. When deducing gas masses using the method from [Scoville et al. \(2016\)](#) only we end up with results much closer to what is found in the literature (e.g. [Tacconi et al. 2018](#); [Wang et al. 2022](#)).

We compare our results to [Magnelli et al. \(2020\)](#) in Fig. B.3. Again, results are much more in line with what is found in [Magnelli et al. \(2020\)](#).

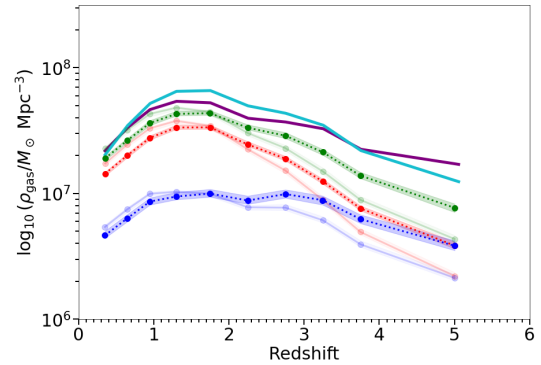
Following the method we describe in Sect. 8.3, we computed the SFE corresponding to the gas masses deduced from the method of [Scoville et al. \(2016\)](#). We display the results in Fig. B.4. In that case, we can still observe an increase of the SFE with redshift.

We also investigate whether the evolution of SFE can be explained by the gas content of galaxies at high redshift by looking at the Kennicutt-Schmidt relation ([Kennicutt 1998b](#)). We follow the same method presented throughout Sect. 8.3. In that case, our best-fit of the linear dependence of  $\log_{10}(\Sigma_{\text{SFR}})$  as a function of  $\log_{10}(\Sigma_{\text{gas}})$  suggests a power-law scaling with a slope of  $N = 1.29^{+0.01}_{-0.01}$ . The results of the Kennicutt-Schmidt relation are shown in Fig. B.5. In Fig. B.5, an increase of the normalisation can still be observed at  $z > 3$ . This effect is, however, much more limited compared to the one observed though the main studied (see Fig. 20).

We corrected the SFE from the gas contribution by looking at  $\text{SFE}/M_{\text{gas}}^{(N-1)}$  instead of SFE (see Fig. B.6). In Fig. B.6, we can still observe an evolution with redshift (i.e. SFE cannot be explain only by the amount of  $M_{\text{gas}}$ ).

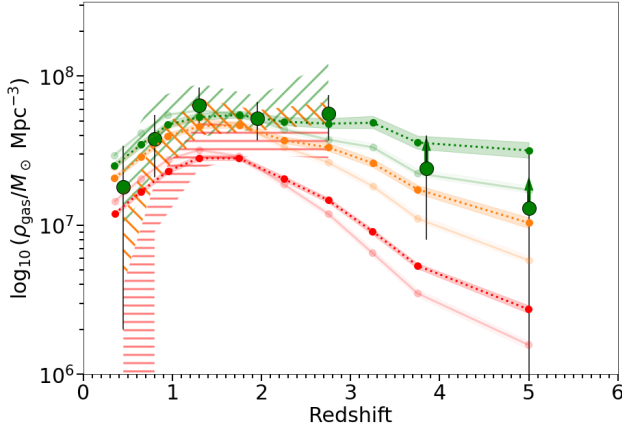


**Fig. B.1.**  $M_{\text{gas}}$  as a function of redshift for different  $M_{\star}$ . The crosses represent  $M_{\text{gas}}$  estimates from this work following the method from [Scoville et al. \(2016\)](#). The squares are measurements from [Saintonge et al. \(2017\)](#). The bins that have been corrected for the  $H$ -dropout contribution are circled. The thick solid colour lines represent the best-fit from this work using Eq. 20. The shaded area the 68% uncertainty of the fit. The faded solid lines represent the best-fit from this work colour-coded by  $M_{\star}$  (using Eq. 21), the faded dashed lines represent the trend from [Tacconi et al. \(2018\)](#), the faded dotted lines are from [Wang et al. \(2022\)](#).

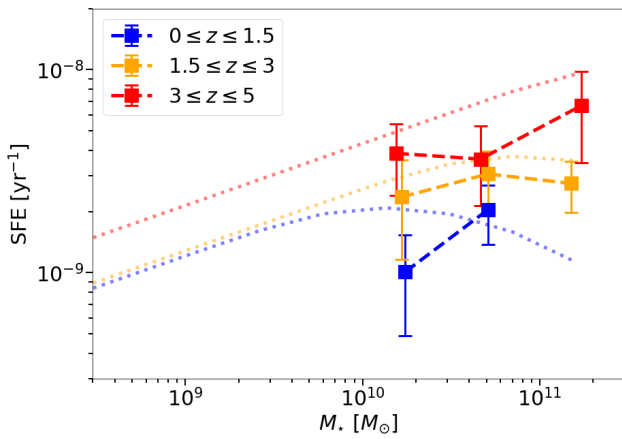


**Fig. B.2.**  $\rho_{\text{gas}}$  as a function of redshift. The green dotted lines represent the total  $\rho_{\text{gas}}$  inferred from this work (i.e. integrating the stellar mass function down to  $3 \times 10^9 M_{\odot}$ ) using  $M_{\text{gas}}$  deduced with the [Scoville et al. \(2016\)](#) method only. The red and blue dotted lines represent the contribution to the total  $\rho_{\text{gas}}$  of galaxies with  $\log_{10}(M_{\star}) > 10$  and  $\log_{10}(M_{\star}) < 10$ , respectively. The corresponding result deduced throughout this paper are displayed in faded solid red, blue and green coloured lines. The purple and cyan lines represent  $\rho_{\text{gas}}$  deduced from the  $M_{\text{gas}}$  definition of [Wang et al. \(2022\)](#) and [Tacconi et al. \(2018\)](#), respectively.

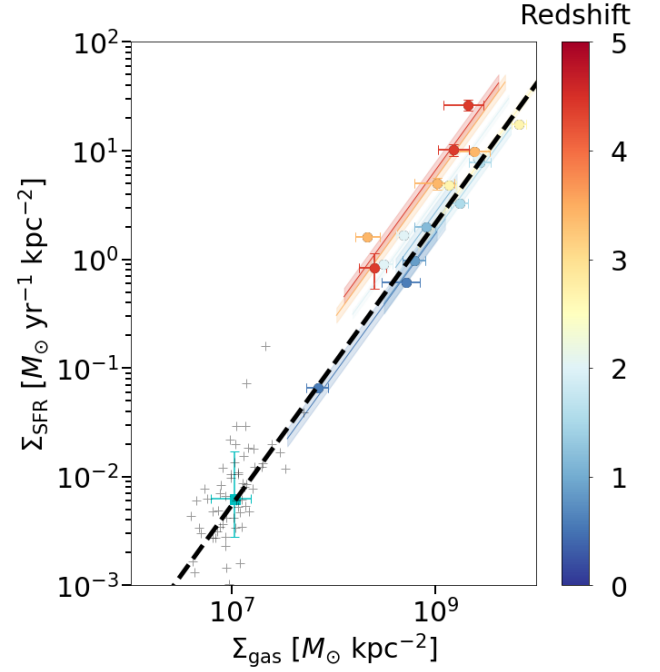
Our conclusion on these results (i.e. using the method from [Scoville et al. \(2016\)](#) to deduce gas masses) and the discrepancies observed from the ones of the main work of this paper (i.e. using an MZR method to deduce gas masses from the IR to millimetre) is that the main reason these higher and evolving SFE were not observed is mostly due to the method used to infer the gas masses: IR to millimetre compared to millimetre only methods.



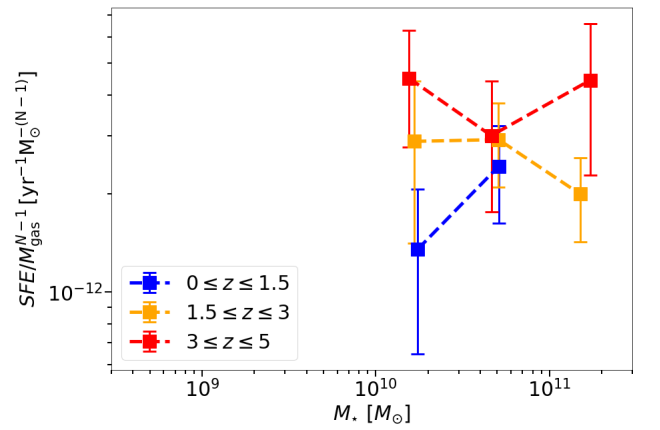
**Fig. B.3.**  $\rho_{\text{gas}}$  as a function of redshift. The green, orange and red dotted lines represent  $\rho_{\text{gas}}$  derived from this work with  $M_{\star}^{\text{lim}}$  of  $1.7 \times 10^8$ ,  $1.7 \times 10^9$  and  $1.7 \times 10^{10} M_{\odot}$ , respectively, using  $M_{\text{gas}}$  deduced with the Scoville et al. (2016) method only. The corresponding result deduced throughout this paper are displayed in faded solid red, orange and green coloured lines. The green, orange and red hatches represent estimates of  $\rho_{\text{gas}}$  from Magnelli et al. (2020) with  $M_{\star}^{\text{lim}}$  of  $1.7 \times 10^8$ ,  $1.7 \times 10^9$  and  $1.7 \times 10^{10} M_{\odot}$ , respectively. The green dots with black circles depict the estimate of  $\rho_{\text{gas}}$  from Magnelli et al. (2020) for  $M_{\star} > M_{\text{limit}}$ ,  $M_{\text{limit}}$  being the stellar mass completeness limit from Magnelli et al. (2020).



**Fig. B.4.** SFE as a function of stellar mass. Squares linked with a dashed line represent the measured from this work when computed using  $M_{\text{gas}}$  deduced with the Scoville et al. (2016) method only, colour coded by bin of redshift. Dotted lines are the value SFE deduced from our best fit law of the SFR and  $M_{\text{gas}}$  main sequences.



**Fig. B.5.** Kennicutt-Schmidt relation:  $\Sigma_{\text{SFR}}$  as a function of  $\Sigma_{\text{gas}}$ . The dots represent the data from this work when computed using  $M_{\text{gas}}$  deduced with the Scoville et al. (2016) method only, and are colour coded by redshift. The cyan squared error bar represents the median of the sample distribution from Kennicutt (1998b), with individual galaxies displayed as grey crosses. The dashed black line represents the best-fit of all points. Coloured lines represent the best-fit, by fixing the slope to the all points best-fit value (i.e.  $N = 1.29^{+0.01}_{-0.01}$ ), at fix redshift.



**Fig. B.6.**  $SFE/M_{\text{gas}}^{(N-1)}$  ( $N = 1.29$ ) as a function of stellar mass. Squares linked with a dashed line represent the measured from this work when computed using  $M_{\text{gas}}$  deduced with the Scoville et al. (2016) method only, colour coded by bin of redshift.