A3COSMOS: A census on the molecular gas mass and extent of main-sequence galaxies across cosmic time

Tsan-Ming Wang, Benjamin Magnelli, Eva Schinnerer, Daizhong Liu, Ziad Aziz Modak, Eric Faustino Jiménez-Andrade, Christos Karoumpis, Vasily Kokorev, Frank Bertoldi

A&A 660, A142 (2022)
https://doi.org/10.1051/0004-6361/202142299

Astronomy & Astrophysics

A^{3}COSMOS: A census on the molecular gas mass and extent of main-sequence galaxies across cosmic time

Tsan-Ming Wang (王贊銘)[1], Benjamin Magnelli[1,2], Eva Schinnerer[3], Daizhong Liu[4], Ziad Aziz Modak[1], Eric Faustino Jiménez-Andrade[5], Christos Karoumpis[1], Vasily Kokorev[6,7], and Frank Bertoldi[1]

[1] Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
e-mail: twan@uni-bonn.de
[2] AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France
[3] Max-Planck-Institut für Astronomie, Königstuhl 17, 69117 Heidelberg, Germany
[4] Max-Planck-Institut für extraterrestrische Physik, Gießenbachstraße 1, 85748 Garching b. München, Germany
[5] National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville 22903, USA
[6] Cosmic Dawn Center (DAWN), Copenhagen, Denmark
[7] Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, 2100 Copenhagen Ø, Denmark

Received 24 September 2021 / Accepted 17 January 2022

ABSTRACT

Aims. We aim to constrain for the first time the mean mass and extent of the molecular gas of a mass-complete sample of normal $>10^{10}\,M_\odot$ star-forming galaxies at $0.4 < z < 3.6$.
Methods. We apply an innovative *uv*-based stacking analysis to a large set of archival Atacama Large Millimeter/submillimeter Array (ALMA) observations using a mass-complete sample of main-sequence (MS) galaxies. This stacking analysis, performed on the Rayleigh-Jeans dust continuum emission, provides accurate measurements of the mean mass and extent of the molecular gas of galaxy populations, which are otherwise individually undetected.
Results. The molecular gas mass of MS galaxies evolves with redshift and stellar mass. At all stellar masses, the molecular gas fraction decreases by a factor of ∼24 from $z \sim 3.2$ to $z \sim 0$. At a given redshift, the molecular gas fraction of MS galaxies decreases with stellar mass at roughly the same rate that their specific star-formation rate (SFR/$M_\star$) decreases. The molecular gas depletion time of MS galaxies remains roughly constant at $z > 0.5$ with a value of 300–500 Myr, but increases by a factor of ∼3 from $z \sim 0.5$ to $z \sim 0$. This evolution of the molecular gas depletion time of MS galaxies can be predicted from the evolution of their molecular gas surface density and a seemingly universal MS-only $\Sigma_{M_{\rm mol}} - \Sigma_{\rm SFR}$ relation with an inferred slope of ∼1.13, the so-called Kennicutt–Schmidt (KS) relation. The far-infrared size of MS galaxies shows no significant evolution with redshift or stellar mass, with a mean circularized half-light radius of ∼2.2 kpc. Finally, our mean molecular gas masses are generally lower than previous estimates, likely due to the fact that literature studies were largely biased toward individually detected MS galaxies with massive gas reservoirs.
Conclusions. To first order, the molecular gas content of MS galaxies regulates their star formation across cosmic time, while variation in their star-formation efficiency plays a secondary role. Despite a large evolution of their gas content and star-formation rates, MS galaxies have evolved along a seemingly universal MS-only KS relation.

Key words. galaxies: evolution – galaxies: high-redshift – galaxies: ISM

1. Introduction

Understanding galaxy evolution across cosmic time is one of the key topics of modern astronomy. One very successful approach for addressing this vast and important question is to assemble and study large and representative samples of galaxies through deep multiwavelength extragalactic surveys. Using this approach, much has been learned over the last few decades about the global star-formation history of the Universe. The cosmic star-formation rate density (SFRD) increases from early cosmic times, $z \sim 2$, and decreases by a factor of 10 by $z \sim 0$ (Madau & Dickinson 2014). About 80% of this star formation takes place in relatively massive galaxies ($>10^{10}\,M_\odot$) that reside on the so-called main sequence (MS) of star-forming galaxies (SFGs; e.g., Noeske et al. 2007; Rodighiero et al. 2011; Sargent et al. 2012). This MS denotes the tight correlation that exists between the stellar mass ($M_\star$) and star-formation rate (SFR) of galaxies, which is observed up to $z \sim 4$ (e.g., Brinchmann et al. 2004; Pannella et al. 2009; Magdis et al. 2010; Zahid et al. 2012; Kashino et al. 2013; Whitaker et al. 2014; Sobral et al. 2014; Speagle et al. 2014; Johnston et al. 2015; Tomczak et al. 2016; Bourne et al. 2017; Pearson et al. 2018; Popesso et al. 2019; Leslie et al. 2020). The existence of the MS, with its constant scatter of 0.3 dex and a normalization that decreases by a factor of 20 from $z \sim 2$ to $z \sim 0$, suggests that most SFGs are isolated and secularly evolving with long (>1 Gyr) star-forming duty cycles. On the contrary, galaxies above the MS (∼5% of the SFG population; Luo et al. 2014) seem to be mostly associated with short, intense starbursts triggered by major mergers and contribute only 10% to the SFRD at all redshifts (e.g., Sargent et al. 2012). While the evolution of the MS and SFRD across cosmic time is observationally well established up to $z \sim 2$, the mechanisms driving their evolution remain poorly constrained. At $z > 2$, our understanding is even more limited because observations obtained from different rest-frame frequencies (i.e., ultraviolet, far-infrared, or radio) provide a somewhat discrepant view of the exact evolution of the SFRD (e.g., Bouwens et al. 2015; Novak et al. 2017; Liu et al. 2018; Gruppioni et al. 2020).

To shed light on the physical processes that regulate star formation across cosmic time, it is paramount to obtain a precise

measurement of the molecular gas content of local and high-redshift galaxies. Indeed, molecular gas fuels star formation, as revealed by the tight correlation between gas mass and SFR surface densities, the so-called Kennicutt–Schmidt (KS) relation (Kennicutt 1998a). Molecular hydrogen (H_2) is the most abundant constituent of molecular gas, but it is difficult to observe due to its lack of a dipole moment. For this reason, the carbon monoxide (CO) molecule, which is the most abundant and readily observable constituent of molecular gas, is usually used to trace the molecular gas content of galaxies (see Bolatto et al. 2013, for a review). However, even with the Atacama Large Millimeter/submillimeter Array (ALMA), obtaining such measurements for $z > 2.0$ MS galaxies with stellar masses of $\sim 10^{10}\,M_\odot$ still requires an hour of observing time per object. The CO molecule is thus still poorly suited for the study of large and representative samples of high-redshift galaxies. Therefore, in recent years, an alternative approach of focusing on high-redshift galaxies has emerged, which relies on dust mass measurements and a standard gas-to-dust mass ratio calibrated in the local Universe. These gas mass measurements, inferred from either multiwavelength dust spectral energy distribution (SED) fits (e.g., Magdis et al. 2012; Magnelli et al. 2012b, 2014; Santini et al. 2014; Tan et al. 2014; Béthermin et al. 2015; Berta et al. 2016; Hunt et al. 2019) or single Rayleigh–Jeans (RJ) flux density conversion (e.g., Scoville et al. 2014, 2016; Groves et al. 2015; Schinnerer et al. 2016; Kaasinen et al. 2019; Liu et al. 2019b; Magnelli et al. 2020; Millard et al. 2020), were shown to be surprisingly accurate when compared to state-of-the-art CO measurements (e.g., Genzel et al. 2015; Scoville et al. 2016, 2017; Tacconi et al. 2018, 2020).

This dust-based approach has since allowed the measurement of the gas content of hundreds of high-redshift SFGs. It was found that the gas fraction of massive SFGs (i.e., $M_{\rm gas}/M_*$) is relatively constant at $z > 2$ but decreases significantly from $z \sim 2$ to $z \sim 0$ (Carilli & Walter 2013; Sargent et al. 2014; Schinnerer et al. 2016; Miettinen et al. 2017b; Scoville et al. 2017; Tacconi et al. 2018, 2020; Gowardhan et al. 2019; Liu et al. 2019b; Wiklind et al. 2019; Cassata et al. 2020). This evolution follows that of the normalization of the MS and implies that the star-formation efficiency (SFE; i.e., SFR/$M_{\rm gas}$) in these galaxies remains relatively constant across cosmic time. This finding is confirmed by the global evolution of the co-moving gas mass density, which resembles that of the SFRD (Magnelli et al. 2020). At any redshift, the depletion time ($t_{\rm depl} = 1/{\rm SFE}$) of the gas reservoirs of massive SFGs is found to be relatively short, on the order of ∼0.5–1 Gyr. Without continuous replenishment of their gas reservoirs, star formation in massive MS galaxies would thus cease within ∼0.5–1 Gyr, in tension with the existence of the MS itself (i.e., long star-forming duty cycles). The continuous accretion of fresh gas from the intergalactic or circum-galactic medium would thus be the main parameter regulating star formation across cosmic time, as also suggested by hydro-dynamical simulations (e.g., Faucher-Giguère et al. 2011; Walther et al. 2019).

While all these previous studies provided key information for our understanding of galaxy evolution, they all suffer from a set of limitations. Firstly, all relied on samples of a few hundred to at most a thousand galaxies and thus suffered from small number statistics, especially because these samples were further split into numerous redshift, stellar mass, and ΔMS ($\Delta{\rm MS} = \log_{10}({\rm SFR}/{\rm SFR_{MS}})$) bins. Secondly, all these studies were based on subsets of galaxies drawn from a parent sample using complex underlying selection functions. Each subsample could thus still fail to provide a complete and representative view of the gas content of high-redshift galaxies. This likely explains in part why these studies agreed qualitatively but disagree quantitatively on the exact redshift evolution of the gas content of massive galaxies (see Liu et al. 2019b). Finally, and most importantly, these studies relied mainly on individually detected galaxies and were thus limited to the high-mass end ($>10^{10.5}\,M_\odot$) of the SFG population. While constraining the gas content of massive galaxies is important, extending our knowledge toward lower stellar masses is crucial because the bulk of the star-formation activity of the Universe is known to take place in $10^{10\cdots10.5}\,M_\odot$ galaxies (e.g., Karim et al. 2011; Leslie et al. 2020). The gas properties of these crucial low-mass high-redshift SFGs thus remain largely unknown simply because most are individually undetected, even in deep ALMA observations.

To statistically retrieve the faint emission of this SFG population, one can perform a stacking analysis. Indeed, by grouping galaxies in meaningful ways (e.g., in bins of redshift and stellar mass) and by stacking their observations (e.g., summing or averaging), one effectively increases the observing time toward this galaxy population and can thus infer their average properties. The noise in the stacked image decreases as the root square of the number of stacked galaxies, and thus large samples can lead to robust detections of previously individually undetected galaxy populations. Such a statistical approach applied to, for example, *Spitzer*, *Herschel*, or ALMA images has proven to be extremely powerful and to push measurements well below the conventional instrumental and confusion noise limits of these observatories (e.g., Dole et al. 2006; Zheng et al. 2006; Magnelli et al. 2014, 2015, 2020; Scoville et al. 2014; Schreiber et al. 2015; Lindroos et al. 2016). Although stacking over the entire ALMA archive provides a unique opportunity to study the gas mass content of low-mass high-redshift SFGs, it also presents two challenges when compared to standard stacking analyses performed with *Spitzer*, *Herschel*, or single ALMA projects, as the ALMA archival data are heterogeneous in terms of observed frequencies and spatial resolution. While stacking data obtained at different observing frequencies simply implies a rescaling of each individual data set to a common rest-frame luminosity frequency using locally calibrated submillimeter SEDs, stacking data with different spatial resolutions is a more uncommon challenge that has only rarely been tackled in the literature (e.g., Lindroos et al. 2016; Chang et al. 2020). It can, however, be easily addressed thanks to the very nature of ALMA observations. Indeed, while combining observations with different spatial resolutions would involve very uncertain and complex convolutions in the image domain, combining them in the *uv* domain is strictly equivalent to performing aperture synthesis on a single object (e.g., Lindroos et al. 2016; Chang et al. 2020).

In this work, we aim at mitigating most of the limitations that affect current studies of the gas properties of high-redshift SFGs by applying an innovative *uv*-based stacking analysis to a large set of ALMA observations toward a mass-complete sample of $M_\star > 10^{10}\,M_\odot$ MS galaxies. This sample is drawn from one of the largest, yet still deep, multiwavelength extragalactic surveys, the Cosmic Evolution Survey 2015 (COSMOS-2015) catalog (Laigle et al. 2016). The stellar masses and redshifts of our galaxies were taken directly from the COSMOS-2015 catalog, while their SFRs were estimated from their COSMOS-2015 rest-ultraviolet, mid-infrared (MIR), and far-infrared (FIR) photometry following the ladder of SFR indicators of Wuyts et al. (2011). From this mass-complete sample of MS galaxies, we only kept those with an ALMA archival band-6 or band-7 coverage as assembled by the Automated mining of the ALMA Archive in the COSMOS field (A^{3}COSMOS) project (Liu et al. 2019a).

This mass-complete sample of MS galaxies was then subdivided into several redshift and stellar mass bins, and a measurement of their mean molecular gas mass and size was performed using a *uv*-based stacking analysis of their ALMA observations. This stacking analysis allows for accurate mean gas mass and size measurements even at low stellar masses where galaxies are too faint to be individually detected by ALMA. Our results provide, for the first time, robust RJ-based constraints on the mean cold gas mass of a mass-complete sample of $M_\star > 10^{10}\,M_\odot$ galaxies up to $z \sim 3$. Combined with their mean FIR size measurements, this yields the first stringent constraint of the KS relation at high redshift.

The structure of the paper is as follows: in Sect. 2 we introduce the ALMA data used in our study and our mass-complete sample of MS galaxies; in Sect. 3 we describe the method used to estimate the mean gas mass and size of a given galaxy population, stacking their ALMA observations in the *uv* domain; in Sect. 4 we present our results, and we discuss them in Sect. 5; finally, in Sect. 6 we summarize our findings and present our conclusions.

Throughout the paper, we assume a flat Λ cold dark matter cosmology with $H_0 = 67.8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.308$, and $\Omega_\Lambda = 0.692$ (Planck Collaboration XIII 2016). All stellar masses and SFRs are provided assuming a Chabrier (2003) initial mass function.

2. Data

2.1. The A^{3}COSMOS data set

The A^{3}COSMOS project aims at homogeneously processing (i.e., calibration, imaging and source extraction) of all ALMA projects targeting the COSMOS field that are publicly available, and providing these calibrated visibilities, cleaned images, and value-added source catalog via a single access portal (Liu et al. 2019a). In our analysis we use the A^{3}COSMOS 20200310 version[1], that is, all ALMA projects publicly available over the COSMOS field as of 10 March 2020. This database contains 80 independent ALMA projects with band-6 and/or band-7 observations. The interferometric calibration was performed by the A^{3}COSMOS project using the Common Astronomy Software Applications (CASA) package (McMullin et al. 2007) and the calibration scripts provided by the ALMA observatory. During this calibration step, a weight is assigned to each calibrated visibility and this weight is key for the accuracy of our stacking analysis (see Sect. 3.2). Unfortunately, the definition of these weights changed between the CASA versions used for the ALMA cycles 0, 1, and 2, and those used for ALMA cycles >3. For this reason, we excluded from our analysis all cycle 0, 1, and 2 ALMA projects. Our final database contains 64 ALMA projects, 39 in band-6 and 25 in band-7. These projects include 1893 images (equivalently ALMA pointings), which contain a total of 1002 sources with $>4.35\sigma$ (Liu et al. 2019a).

2.2. Our sample

COSMOS is a deep extragalactic blind survey of two square degrees on the sky centered at RA (J2000) = $10^{\mathrm{h}}00^{\mathrm{m}}28.6^{\mathrm{s}}$, Dec = $+02°12'21.0''$ (Scoville et al. 2007). This survey has been carried out over 46 broad and narrow bands probing the entire electromagnetic spectrum, from X-ray (e.g., *XMM-Newton*; Cappelluti et al. 2009), ultraviolet (e.g., GALEX; Zamojski et al. 2007), optical (e.g., Koekemoer et al. 2007; Taniguchi et al. 2007), infrared (e.g., *Spitzer*; Sanders et al. 2007), to radio wavelengths (e.g., VLA; Schinnerer et al. 2010; Smolčić et al. 2017). These observations have triggered numerous spectroscopic follow-up studies, providing nowadays more than 10 000 spectroscopic redshifts for galaxies over this field. From all these photometric and spectroscopic multiwavelength coverage, Laigle et al. (2016) built the reference COSMOS-2015 catalog, providing the photometry, redshift (photometric or spectroscopic), stellar mass, and SFR of more than half a million of galaxies. From their careful analysis, Laigle et al. (2016) classified galaxies into quiescent and star-forming based on a standard rest-frame near-ultraviolet-*r*/*r*-*J* selection method. The mass-completeness of their SFGs is down to stellar masses of $\sim 10^{9.3}\,M_\odot$ at $z < 1.75$ and $\sim 10^{9.9}\,M_\odot$ at $z < 3.50$ (see their Table 6). Here we select only SFGs above their mass-completeness limit. Moreover, to avoid contamination from active galactic nuclei (AGNs), we exclude from our analysis all galaxies classified as AGNs based on their X-Ray luminosity ($L_X \geq 10^{42}\,\mathrm{erg\,s^{-1}}$; Szokoly et al. 2004) using the latest COSMOS X-ray catalog of Marchesi et al. (2016). After the selection of SFGs and exclusion of AGNs, our parent sample is left with 515 465 galaxies (green contours in Fig. 1). We note that photometric redshifts in the COSMOS-2015 catalog are highly reliable even up to the redshift limit of our study (i.e., $z = 3.6$), with a redshift accuracy of $\sigma_{\delta z/(1+z)} \sim 0.028$ (Laigle et al. 2016).

To select from this parent sample galaxies residing within the MS of SFGs, one needs to accurately measured their SFRs. The COSMOS-2015 catalog provides such estimates but those are solely based on optical-to-near-infrared SED fits performed by Laigle et al. (2016). While reliable for stellar masses with $M_\star < 10^{11}\,M_\odot$ and moderately SFGs, observations from the *Herschel* Space Observatory have unambiguously demonstrated that such measurements are inaccurate for starbursting or massive SFGs, in which star formation can be heavily dust-enshrouded (e.g., Wuyts et al. 2011; Qin et al. 2019). To accurately measure the SFR of all galaxies in our parent sample, we thus used the approach advocated by Wuyts et al. (2011): applying to each galaxy the best dust-corrected star-formation indicator available (the so-called ladder of SFR indicator; see below for details). The SFR of galaxies for which infrared observations were available, were obtained by combining their un-obscured and obscured SFRs, following Kennicutt (1998b) for a Chabrier (2003) initial mass function,

$$\mathrm{SFR_{UV+IR}}[M_\odot\,\mathrm{yr^{-1}}] = 1.09\times10^{-10}(L_{\mathrm{IR}}[L_\odot]+3.3\times L_{\mathrm{UV}}[L_\odot]), \quad (1)$$

where the rest-frame L_{UV} at 2300 Å was taken from the COSMOS-2015 catalog, and the rest-frame $L_{\mathrm{IR}} = L(8-1000\,\mu\mathrm{m})$ was calculated from their MIR/FIR photometry[2]. For galaxies with multiple FIR photometry in the COSMOS-2015 catalog[3], we estimated their L_{IR} by fitting their *Herschel*-Photodetector Array Camera and Spectrometer (PACS) and *Herschel*-Spectral and Photometric Imaging Receiver (SPIRE) flux

[1] A^{3}COSMOS 20200310 version: https://sites.google.com/view/a3cosmos/data/dataset_v20200310

[2] Among the 3037 galaxies of our final sample (see below), 972 (32%) have MIR 24 μm photometry and among those 482 (16%) have multiple FIR photometry. Among the 1376 galaxies of our final sample with stellar mass $>10^{10}\,M_\odot$ (those detectable by our stacking analysis; see Sect. 4), 852 (62%) have MIR 24 μm photometry and among those 461 (33%) have multiple FIR photometry.

[3] The *Herschel* photometry in the COSMOS-2015 catalog is based on the 24 μm prior source extraction performed by the PEP (Lutz et al. 2011) and HerMES (Oliver et al. 2012) consortia.

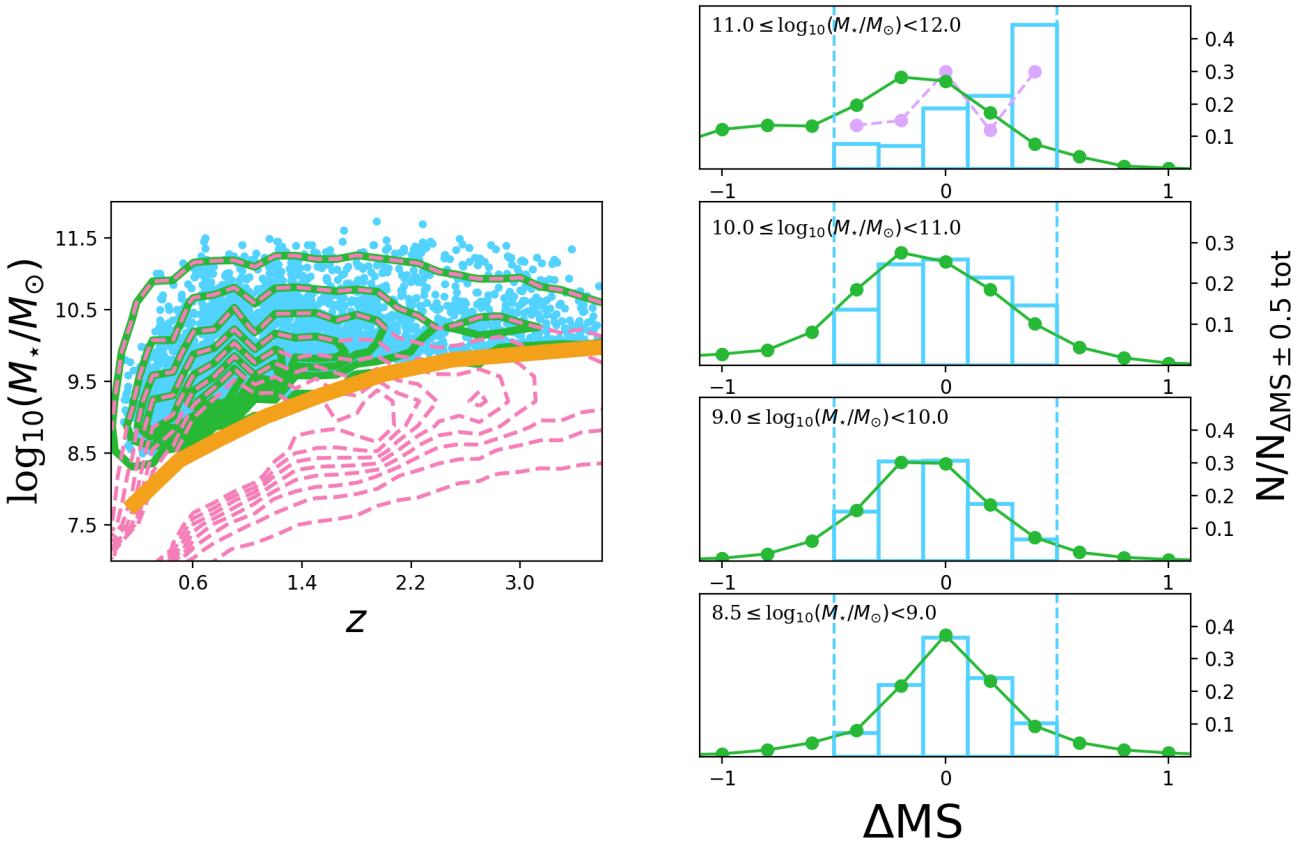

Fig. 1. Completeness and relative ΔMS distribution in our final sample. *Left*: Stellar mass and redshift distribution in our final ALMA-covered mass-complete sample of MS galaxies (blue dots). The dashed pink contours display the number density of SFGs in Laigle et al. (2016), i.e., our parent sample of SFGs. The pink contour levels are in steps of 500 from 200 to 3700 galaxies per z–$\log_{10} M_\star$ bin of size 0.14 and 0.15, respectively. The solid orange line represents the stellar mass completeness limit of SFGs in Laigle et al. (2016). The green contour shows the number density of SFGs in Laigle et al. (2016) above this stellar mass completeness limit, i.e., our parent mass-complete sample. The green contour levels are in steps of 500 from 200 to 3700 galaxies per z–$\log_{10} M_\star$ bin of size 0.14 and 0.15, respectively. *Right*: Relative ΔMS distribution in our final ALMA-covered mass-complete sample of MS galaxies (blue histogram) and our mass-complete parent sample of SFGs (green histogram) in different stellar mass bins. In the highest stellar mass bin, the dashed purple line shows the relative ΔMS distribution after having rejected from our final sample all ALMA primary targets, i.e., galaxies at the phase center of the ALMA observation. The vertical dashed blue lines display the ±0.5 dex interval used to defined MS galaxies. Over this interval, the integral of each histogram is equal to one. This normalization is needed to compare our final and mass-complete parent samples, which contain 3037 and 515 465 galaxies, respectively.

densities (Lutz et al. 2011; Oliver et al. 2012) with the SED template library of Chary & Elbaz (2001). For galaxies without a multiple FIR photometry but an MIR 24 μm detection in the COSMOS-2015 catalog, we estimated their $L_{\rm IR}$ by scaling the MS SED template of Elbaz et al. (2011) to their 24 μm flux densities (Le Floc'h et al. 2009). This particular MS SED template was chosen because it provides accurate 24 μm-to-$L_{\rm IR}$ conversions over the redshift and stellar mass ranges probed in our study (Elbaz et al. 2011). For galaxies without any MIR or FIR photometry, we used the SFRs measured by Laigle et al. (2016) and which were obtained by fitting their optical-to-near-infrared photometry with the Bruzual & Charlot (2003) SED model. We verified that toward intermediate SFRs – that is, where the fraction of galaxies with an MIR/FIR detection starts to decrease (i.e., $0 < \log(\mathrm{SFR_{IR+UV}}) < 1.5$) – our ultraviolet-plus-infrared-based SFR measurements agree with those solely based on this optical-to-near-infrared SED fits, with a median $\log(\mathrm{SFR_{IR+UV}}/\mathrm{SFR_{SED}})$ of $0.09^{+0.39}_{-0.53}$ (Fig. 2). This agreement ensures a smooth transition between the different steps of our ladder of SFR indicators. Also, among the 269 galaxies of our final sample with stellar masses $>10^{10}\,M_\odot$ (that is, those detectable by our stacking analysis; see Sect. 4) and with SFR > $100\,M_\odot\,\mathrm{yr}^{-1}$, only 54 have their SFRs solely based on their SED fits and thus potentially underestimated by ~0.3–0.5 dex (see Fig. 2). Finally, we note that at high SFRs, where a high fraction of galaxies are individually detected by ALMA, our SFRs agree with those from the A^{3}COSMOS catalog, that is, inferred with Multi-wavelength Analysis of Galaxy Physical Properties (MAGPHYS; da Cunha et al. 2008, 2015) SED fitting combining the COSMOS-2015 photometry with super-deblended *Herschel* (Jin et al. 2018) and ALMA photometry.

From their redshift, stellar mass, and SFR, we can measure the offset of each of these galaxies from the MS: $\Delta\mathrm{MS} = \log(\mathrm{SFR}(z, \mathrm{SM})/\mathrm{SFR_{MS}}(z, \mathrm{SM}))$. To this end, we used the MS calibration of Leslie et al. (2020), as it is also based on the mass-complete COSMOS-2015 catalog:

$$\log(\mathrm{SFR_{MS}}(z, \mathrm{SM})) = S_0 - a_1 t - \log\left(1 + \left(\frac{10^{M'_t}}{10^{M}}\right)\right), \qquad (2)$$
$$M'_t = M_0 - a_2 t,$$

where M is $\log(M_\star/M_\odot)$, t is the age of the Universe in Gyr, $S_0 = 2.97$, $M_0 = 11.06$, $a_1 = 0.22$, and $a_2 = 0.12$. Our mass-complete sample of MS galaxies was then constructed by selecting galaxies with ΔMS between −0.5 and 0.5 (e.g., Rodighiero et al. 2014). This sample contains 92 739 galaxies.

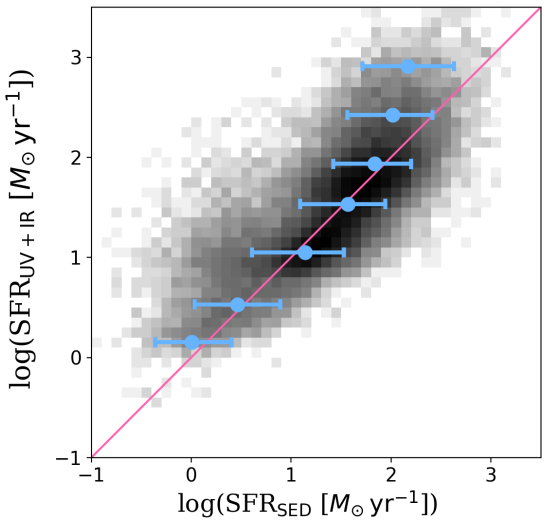

Fig. 2. Comparison of the SFRs obtained from the COSMOS-2015 catalog, i.e., SFR_{SED}, to the SFRs obtained from the ladder of SFR, i.e., SFR_{UV+IR}. Number densities are displayed in log-scale. Blue circles represent the median value of $\log(SFR_{SED})$ in $\log(SFR_{UV+IR})$ bins, starting from −0.25 dex and with a bin size of 0.5 dex. Error bars correspond to the 16th and 84th percentiles. The pink line is the one-to-one relation.

Finally, from this mass-complete sample of MS galaxies, we selected those with an ALMA band-6 (~243 GHz) or band-7 (~324 GHz) coverage in the A^{3}COSMOS database (see Sect. 2.1). Here, we only consider galaxies well within the ALMA primary beam (i.e., where the primary beam response is higher than 0.5). This conservative primary beam cut was used because uncertainties in the primary beam response far from the phase center can significantly affect our stacking analysis (see Sect. 3.2). In addition, to avoid contamination by bright neighboring sources, we excluded from our analysis galaxy pairs (<2.″0) with $S^1_{ALMA}/S^2_{ALMA} > 2$ or $M^1_\star/M^2_\star > 3$ (for ALMA undetected galaxies, assuming a first-order $M_{gas} - M_\star$ correlation). About 8% of our galaxies are excluded by these criteria. However, we note that most of these excluded galaxy pairs (~95%) are due to projection effects ($\Delta z > 0.05$). This implies that the exclusion of these galaxies does not introduce any biases into our final ALMA-covered mass-complete sample of MS galaxies. There are 3,037 galaxies in this final sample. The left panel of Fig. 1 shows the stellar mass and redshift distribution of our parent and final samples. Our final sample probes a broad range in redshifts and stellar masses, similar to that probed by our parent sample. We verified that our parent and final samples have consistent stellar mass, redshift and L_{IR} distributions, with Kolmogorov–Smirvov probabilities of 99%, 99%, and 96% of being drawn from the same distribution, respectively.

The ALMA archive cannot be treated as a real blind survey and thus our ALMA coverage selection criteria could have introduced a bias in our final ALMA-covered mass-complete MS galaxy sample. As an example (though rather unrealistic), if all ALMA projects in COSMOS would have targeted MS galaxies with $\Delta MS = 0.3$ dex, our final sample would naturally be biased toward this population and thus not be representative of the entire MS galaxy population. A simple way to test the presence of such bias is to compare the ΔMS distributions of our final and parent samples for different stellar mass bins (Fig. 1; right panels). As expected, our parent sample (green histogram) exhibits in all stellar mass bins a Gaussian distribution centered at 0 and with a 0.3 dex dispersion. At low stellar masses ($M_\star < 10^{10.0}\,M_\odot$), our final sample follows the same distribution, with a Kolmogorov–Smirvov 99% probability of being drawn from the same sample (this finding remaining true even if we further divide these stellar mass bins into several redshift bins). Indeed, in these low stellar mass bins, only 5% of our galaxies are located at the phase center of the ALMA image and thus were the primary target of the ALMA observations. However, we note that in the highest stellar mass bins the ΔMS distribution of our final sample is significantly skewed toward high ΔMS values (this finding is still true if we further divide these stellar mass bins into several redshift bins). In these stellar mass bins, about 63% of our galaxies are the primary targets of the ALMA observations (i.e., located at the phase center), and thus potentially affected by complex and uncontrollable selection biases. Excluding these primary targets from our galaxy sample yields ΔMS distribution in much better agreement with those of our parent sample. In the rest of our analysis, at high masses, we show our stacking results before and after excluding these primary-target galaxies. In addition, we account for these ΔMS distributions while fitting the cosmic and stellar mass evolution of the mean molecular gas content of MS galaxies.

3. Method

ALMA has revolutionized the study of high-redshift SFGs at (sub)millimeter wavelengths. Nevertheless, even with its unparallel sensitivity, ALMA cannot detect within a reasonable observing time MS galaxies with $M_\star < 10^{10.5}\,M_\odot$ at $z > 0.5$. Consequently, despite including all individually detected galaxies within the A^{3}COSMOS images (i.e., primary targets and serendipitous detections), the final sample of Liu et al. (2019b) is still mostly restricted to the high-mass end of the SFG population. The emission of such low-mass high-redshift SFGs captured within these images is too faint to be individually detected, and thus remains unexploited. To statistically retrieve the faint emission of this SFG population, we need to perform a stacking analysis. As already mentioned, stacking over the entire A^{3}COSMOS data set presents two challenges when compared to standard stacking analysis performed with *Spitzer*, *Herschel*, or individual ALMA projects. Indeed, the A^{3}COSMOS database is heterogeneous in terms of observed frequencies and spatial resolution. The frequency-heterogeneity problem is simply solved by a prior rescaling of each individual data set to a common rest-frame luminosity frequency using locally calibrated submillimeter SEDs (Sect. 3.1), while the spatial resolution-heterogeneity problem is solved by performing our stacking analysis in the *uv* domain (Sect. 3.2).

In the following, we describe in detail the different steps of our stacking analysis, while the validation of this methodology via Monte Carlo simulations is presented in Appendix A.

3.1. From observed-frame flux densities to rest-frame luminosities

The A^{3}COSMOS observations were performed at different frequencies and the galaxies to be stacked also lie at slightly different redshifts. Therefore, prior to proceeding with our stacking analysis, we needed to convert the ALMA observations of a given galaxy from observed flux density to its rest-frame luminosity at 850 μm (i.e., L^{rest}_{850}). To do so, we used the MS SED templates of Béthermin et al. (2012), which accurately capture the monotonic increase in the dust temperature of MS galaxies with redshift (e.g., Magdis et al. 2012; Magnelli et al. 2014). First, we computed the SED template luminosity ratio at rest-frame

850 μm and the observed rest-frame wavelength of the galaxy of interest,

$$\Gamma^{\rm SED} = L^{\rm SED}_{850} / L^{\rm SED}_{\lambda_{\rm obs}/(1+z)}. \quad (3)$$

The observed ALMA visibility amplitudes toward this galaxy (i.e., $|V(u,v,w)|_{\lambda_{\rm obs}}$) – which are in units of flux density – were then converted into rest-frame 850 μm luminosity following

$$|L(u,v,w)|^{\rm rest}_{850} = 4\,\pi\,D_L^2 \times |V(u,v,w)|_{\lambda_{\rm obs}} \times \Gamma^{\rm SED}/(1+z), \quad (4)$$

where D_L is the luminosity distance of the galaxy of interest. This rescaling of the amplitude (and weights) of the ALMA visibilities was performed for each stacked galaxy using the CASA tasks `gencal` and `applycal`.

3.2. Stacking in the uv domain

Stacking in the *uv* domain relies on the exact same principle as aperture synthesis. The only difference is that one combines multiple baselines pointing at the same galaxy population instead of multiple baselines pointing at the same galaxy. The tools or tasks needed to perform stacking in the *uv* domain are thus all readily available in CASA. For each of our stellar mass-redshift bin and each galaxy within these bins, we proceeded as follow. First, we time- and frequency-averaged their measurement set, producing one averaged visibility per ALMA scan (lasting typically 30 s and originally divided into ten 3-second integration bins) and ALMA spectral window (probing typically 2 GHz and originally divided into hundreds of channels). This step, which was performed using the CASA task `split`, is crucial to keep the volume of our final stacked measurement sets within current computing capabilities. These averaged visibilities were then rescaled from observed-frame flux density into rest-frame 850 μm luminosity using the CASA tasks `gencal` and `applycal` (see Sect. 3.1). Finally, the phase center of these averaged and rescaled visibilities were shifted to the coordinate of the stacked galaxy. This step was performed using the CASA package `STACKER` (Lindroos et al. 2015) following

$$L_{\rm shifted}(u,v,w)^{\rm rest}_{850} = L(u,v,w)^{\rm rest}_{850}\,\frac{1}{A_N(\hat{S}_k)}\,e^{\frac{2\pi}{\lambda}iB\cdot(\hat{S}_0-\hat{S}_k)}, \quad (5)$$

where $L(u,v,w)^{\rm rest}_{850}$ is the averaged and rescaled visibility, $\hat{S}_0$ is a unit vector pointing to the original phase center, $\hat{S}_k$ is a unit vector pointing to the position of the stacked galaxy, $A_N(\hat{S}_k)$ is the primary beam attenuation in the direction $\hat{S}_k$, B is the baseline of the visibility. The final stacked measurement set of a given stellar mass-redshift bin was then obtained by concatenating the shifted, rescaled, and averaged visibilities (i.e., $L_{\rm shifted}(u,v,w)^{\rm rest}_{850}$) of all galaxies within this bin using the CASA task `concat`. Because all these steps were performed in CASA, the original weights of all visibilities (i.e., those accounting for their system temperature, channel width, integration time…) were properly renormalized and could thus be used for the forthcoming *uv*-model fit and image processing.

To measure the stacked rest-frame 850 μm luminosity of each of our stellar mass–redshift bins (i.e., $L^{\rm stack}_{850}$), we used two different approaches. First, we extracted this information from the *uv* domain by fitting a single component model to the stacked measurement set. This fit was performed using the CASA task `uvmodelfit`, assuming a single Gaussian component and fixing its position to the stacked phase center. Second, we measured $L^{\rm stack}_{850}$ from the image domain. To do so, we imaged the stacked measurement set with the CASA task `tclean`, using Briggs natural weighting and cleaning the image down to 3σ. Then, we fitted a 2D Gaussian model to the cleaned image using the Python Blob Detector and Source Finder (`PyBDSF`) package (Mohan & Rafferty 2015). For all our stellar mass–redshift bins, these two approaches agreed within the uncertainties.

Our *uv*-domain and image-domain fits provide us also with the mean size (or upper limit) of the galaxy population in a given stellar mass–redshift bin. From the intrinsic (i.e., beam-deconvolved) full width at half maximum (FWHM) of the major axis outputted by `uvmodelfit` or `PyBDSF`, we define the effective – equivalently half-light – radius ($R_{\rm eff}$) of the stacked population following Jiménez-Andrade et al. (2019) (i.e., $R_{\rm eff} \approx FWHM/2.43$). Then, we express these mean size measurements in form of circularized radii, $R^{\rm circ}_{\rm eff}$,

$$R^{\rm circ}_{\rm eff} = R_{\rm eff} \times \sqrt{\frac{b}{a}}, \quad (6)$$

where b/a is the axis ratio measured with `uvmodelfit` or `PyBDSF`.

Finally, to infer the uncertainties associated with these stacked rest-frame 850 μm luminosity and size measurements, we used a standard resampling method. These uncertainties account not only for the instrumental noise in the stacked measurement set (i.e., the detection significance) but also for the intrinsic distribution of L_{850} and size within the stacked galaxy population. For a stellar mass–redshift bin containing N galaxies, we performed N different realizations of our stacking analysis, removing in each realization one galaxy of the stacked sample. The uncertainties on $L^{\rm stack}_{850}$ and size are then given by the standard deviation of these quantities measured over these realizations multiplied by $\sqrt{N}$. We note that because there is a possible mismatch of $\sim$0.″2 between the stacked optical-based position and the actual (sub)millimeter position of the sources (e.g., Elbaz et al. 2018), the average FIR sizes inferred in our study could be slightly overestimated. This is further discussed in Sect. 4.3.

It should be noted that although some studies have used median stacking to mitigate the contribution of bright outliers to the stacked flux densities (e.g., Algera et al. 2020; Feltre et al. 2020; Fudamoto et al. 2020; Gabányi et al. 2021; Johnston et al. 2021), we decided to perform our analysis using a mean stack; that is to say, in the *uv* domain our models are fitted to the weighted mean visibility amplitudes and our images are created by `tclean` using weighted mean visibilities. This choice was made for the following reasons: (i) the impact of bright outliers is already mitigated by our $-0.5 < \Delta{\rm MS} < 0.5$ selection, which by construction excludes gas-rich starbursts; (ii) the impact of bright outliers is accounted for in our uncertainties (i.e., resampling method); and finally (iii) Schreiber et al. (2015) and Leslie et al. (2020), which thoroughly tested mean and median stacking, concluded both that median stacking is biased toward higher values at low S/N because the median is not a linear operation and that the stacked distribution is intrinsically a log-normal distribution skewed toward bright sources. As a result, median stacked fluxes are difficult to interpret and are often not measuring the median nor mean fluxes, but something in between. We note, however, that the median visibility amplitudes of each of our stacked bins are consistent, within the uncertainties, with the mean visibility amplitudes (see open symbols in Fig. 3).

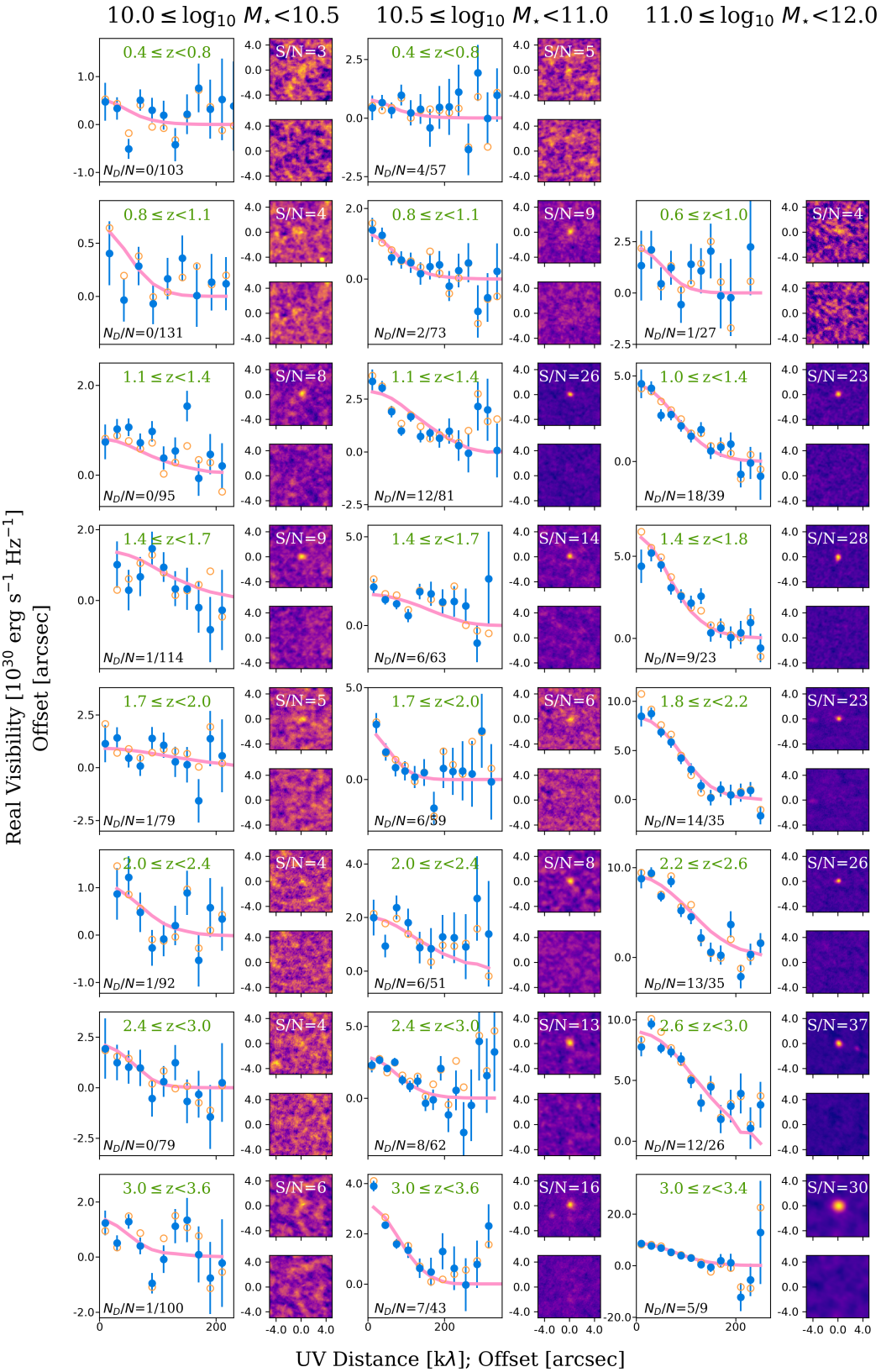

Fig. 3. Results of our stacking analysis for MS galaxies in the *uv* and image domain. For each stellar mass–redshift bin, *the left panel* shows the single component model (solid pink line) fitted to the (stacked) mean visibility amplitudes (filled blue circles) using the `CASA` task `uvmodelfit`. Open orange circles show the median visibility amplitudes, which are consistent, within the uncertainties, with the mean visibility amplitudes. *The top-right and bottom-right panels* show, respectively, the stacked and residual images, the latter being obtained by subtracting from the former the single 2D Gaussian component fitted by `PyBDSF`. The number of individually detected galaxies ($N_{\rm D}$) and the number of stacked galaxies (N) in each stellar mass–redshift bin is reported in the *left panel* (i.e., $N_{\rm D}/N$), while the detection significance i.e., $S/N_{\rm peak}$, is reported in the *upper-right panel*.

3.3. From rest-frame 850 μm luminosities to molecular gas masses

The literature contains a plethora of relations linking molecular gas mass of galaxies with their (sub)millimeter luminosities (e.g., Bourne et al. 2013; Groves et al. 2015; Scoville et al. 2017; Bertemes et al. 2018; Saintonge et al. 2018; Kaasinen et al. 2019). All of them rely on an assumed gas-to-dust mass ratio (or a direct 870 μm luminosity-to-gas mass ratio) that might or might not depend on the metallicity. Liu et al. (2019b) thoroughly studied how these different relations influence our molecular gas mass estimation, using a sample of galaxies down to a stellar mass of $\sim 10^{10.3}\,M_{\odot}$. They found that metallicity-dependent relations (Rémy-Ruyer et al. 2014; Genzel et al. 2015) and the 850 μm luminosity-dependent relation of Hughes et al. (2017) only differ by ∼0.15–0.25 dex (which is comparable to the observed scatter), and that the relation of Hughes et al. (2017) provided the best agreement with local observations (e.g., Bertemes et al. 2018; Saintonge et al. 2018). They concluded that the 850 μm luminosity-dependent relation is thus the most preferable relation for galaxies down to a stellar mass of $M_{\star} \sim 10^{10.3}\,M_{\odot}$ and for which no metallicity measurements are available. Based on their analysis, we decided to use this empirically calibrated relation of Hughes et al. (2017). The mean molecular gas mass of a given galaxy population (i.e., $M_{\rm mol}$) is thus computed from their stacked rest-frame 850 μm luminosities following

$$\log_{10} M_{\rm mol} = (0.93 \pm 0.01) \cdot \log_{10} L_{850} - (17.74 \pm 0.05), \quad (7)$$

where $M_{\rm mol}$ already includes the 1.36 correction factor to account for helium and assumes a CO-to-$M_{\rm mol}$ conversion factor (i.e., $\alpha_{\rm CO}$) of 6.5 $(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$.

In Sect. 4.5.1 and Appendix B, we thoroughly present and discuss the impact on our results of using different gas mass calibration relations. In brief, the main conclusions of our paper are not qualitatively affected by this particular choice; the H17 method yields measurements that are bracket by those inferred from other relations. Finally, measurements obtained using H17 are in good agreement with Tacconi et al. (2020) at high stellar masses, where the Tacconi et al. (2020) study can be considered as the reference.

4. Results

The results of our stacking analysis are shown in Fig. 3 and summarized in Table 1. In our highest stellar mass bin (i.e., $10^{11} \leq M_{\star}/M_{\odot} < 10^{12}$; right-most column), the number of stacked sources per redshift bin varies from 9 to 39, with about 37% of them being individually detected. In this stellar mass bin, our stacking analysis yields high significance detections, with peak signal-to-noise ratios ($S/N_{\rm peak}$) greater than 20, except in our lowest redshift bin with $S/N_{\rm peak} \sim 4$. In the *uv* domain, those high significance detections are characterized by a Gaussian-like decrease of the stacked visibility amplitudes with the *uv*-distance, well fitted by our single component model. These galaxy populations are thus detected and spatially resolved by our stacking analysis. In the image domain, this translates into bright spatially resolved phase-center emission (i.e., with a median synthesized beam FWHM of 0.″5 and an median angular size-to-synthesized beam FWHM ratio of 1.5) that is well described by single 2D Gaussian components. In our intermediate stellar mass bin (i.e., $10^{10.5} \leq M_{\star}/M_{\odot} < 10^{11.0}$), the number of stacked sources per redshift bin increases (43–81), while the fraction of them being individually detected decreases to about 10%. As for our highest stellar mass bin, our stacking analysis yields high significance detections (i.e., $S/N_{\rm peak} > 5$) in all of our redshift bins and those are spatially resolved at our median synthesized beam FWHM of 0.″6. Finally, in our lowest stellar mass bin (i.e., $10^{10} \leq M_{\star}/M_{\odot} < 10^{10.5}$), the number of stacked sources per redshift bin increases even further (79–131) and only few of them are individually detected (1%). In this low stellar mass bin, the same patterns are observed, i.e., spatially resolved detections in the *uv* and image domain (with a median synthesize beam FWHM of 0.″7), though at lower significance, i.e., $3 < S/N_{\rm peak} < 9$. This implies that the number of stacked galaxies (controlled by the stellar mass function of MS galaxies) does not increase sufficiently to fully counterbalance the decrease in their molecular gas content with respect to the most massive population. Nevertheless, even in this low stellar mass bin, our stacking analysis yields clear detection ($S/N_{\rm peak} > 3$), especially when considering both the *uv* domain and image-domain constraints. We note that pushing this stacking analysis to lower stellar masses ($M_{\star} < 10^{10}$) did not produce any significant detection. These results are thus not presented here and not discussed further in the paper.

We conclude that our stacking analysis provides robust mean molecular gas mass and FIR size measurements for $M_{\star} > 10^{10}\,M_{\odot}$ MS galaxies from $z \sim 0.4$ to 3.6. Considering that in our highest and lowest stellar mass bins only 37% and ∼1% of these galaxies were individually detected in the A^{3}COSMOS catalog, respectively, our stacking analysis clearly provides the first unbiased ALMA view on the gas content and size of MS galaxies.

4.1. The molecular gas content of MS galaxies

The redshift evolution of the molecular gas mass of MS galaxies inferred from our stacking analysis is shown in Fig. 4. It is compared to analytical predictions from the literature (Scoville et al. 2017; Liu et al. 2019b; Tacconi et al. 2020), individually detected MS galaxies taken from the A^{3}COSMOS catalog (Liu et al. 2019b) and a local reference (i.e., $z \sim 0.03$) taken from Saintonge et al. (2017). In addition, in Fig. 5, we present the evolution of the molecular gas fraction (i.e., $\mu_{\rm mol} = \langle M_{\rm mol}\rangle/\langle M_{\star}\rangle$) of MS galaxies as a function of redshifts and stellar masses. We note that for our galaxies in common with the A^{3}COSMOS catalog, the stellar masses used here (i.e., those from the COSMOS-2015 catalog) are about 0.22 dex lower than those reported in the A^{3}COSMOS catalog. This offset, which is also discussed in Liu et al. (2019b), is likely explained by the fact that stellar masses in the A^{3}COSMOS catalog rely on full optical-to-millimeter energy-balanced SED fits performed with MAGPHYS. While this offset is observed for massive galaxies, it might not be present at $\lesssim 10^{10.5}\,M_{\odot}$, where the number of galaxies available in the A^{3}COSMOS catalog is too scarce to provide meaningful comparison with the COSMOS-2015 catalog. In any case, when comparing our analytical predictions to those from Liu et al. (2019b), we thus show both their original predictions and those inferred by accounting for this systematic 0.22 dex offset.

Our measurements reveal a significant evolution of the molecular gas mass of MS galaxies with both redshifts and stellar masses. For all stellar mass bins, the molecular gas masses of MS galaxies (equivalently molecular gas fraction) increase by a factor of ∼24 from $z \sim 0$ to $z \sim 3.2$. In addition, at a given redshift, the molecular gas masses of MS galaxies significantly increase with stellar masses. This trend is, however, sublinear in the log-log space, which implies that the molecular gas fraction of MS galaxies decreases with stellar mass at a given redshift (Fig. 5). To obtain a more quantitative constraint on the stellar

Table 1. Molecular gas mass and size properties of MS galaxies.

$M_\star$	z	N	N_D	$\langle z\rangle$	$\langle M_\star\rangle$ ($\log_{10}[M_\odot]$)	$\langle$SFR$\rangle$ ($\log_{10}[M_\odot\,\mathrm{yr}^{-1}]$)	$\langle\Delta$MS$\rangle$	$\langle\nu_{\rm obs}\rangle$ (GHz)	$\langle L_{850-uv}\rangle$ (10^{30} erg s^{-1} Hz^{-1})	$S/N_{\rm peak}$	$M_{\rm mol-uv}$ ($\log_{10}[M_\odot]$)	$M_{\rm mol-py}$ ($\log_{10}[M_\odot]$)	$\theta^{\rm circ}_{\rm beam}$ (arcsec)	$\theta^{\rm circ}_{\rm mol-uv}$ (arcsec)	$\theta^{\rm circ}_{\rm mol-py}$ (arcsec)	$R^{\rm circ}_{\rm eff-uv}$ (kpc)	$R^{\rm circ}_{\rm eff-py}$ (kpc)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
$11.0 \le \log_{10}$	$0.6 \le z < 1.0$	27	1	0.79	11.16	1.58	0.09	298.02	$2.3_{\pm0.7}$	4	$10.5_{\pm0.1}$	$10.4_{\pm0.2}$	0.51	1.10	0.83	$3.0_{\pm1.}$	$2.1_{\pm1.1}$
$M_\star < 12.0$	$1.0 \le z < 1.4$	39	18	1.24	11.17	2.01	0.19	313.30	$4.5_{\pm0.7}$	23	$10.8_{\pm0.1}$	$10.8_{\pm0.1}$	0.50	0.75	0.75	$1.9_{\pm0.2}$	$1.9_{\pm0.2}$
	$1.4 \le z < 1.8$	23	9	1.57	11.22	2.21	0.20	311.27	$6.3_{\pm1.2}$	28	$10.9_{\pm0.1}$	$10.9_{\pm0.1}$	0.47	0.77	0.78	$2.1_{\pm0.1}$	$2.2_{\pm0.2}$
	$1.8 \le z < 2.2$	35	14	1.98	11.22	2.35	0.15	312.98	$8.4_{\pm3.0}$	23	$11.0_{\pm0.2}$	$11.1_{\pm0.1}$	0.42	0.65	0.67	$1.7_{\pm0.1}$	$1.8_{\pm0.2}$
	$2.2 \le z < 2.6$	35	13	2.35	11.26	2.52	0.24	310.17	$9.3_{\pm1.9}$	26	$11.1_{\pm0.1}$	$11.1_{\pm0.1}$	0.42	0.59	0.62	$1.5_{\pm0.1}$	$1.6_{\pm0.1}$
	$2.6 \le z < 3.0$	26	12	2.82	11.22	2.57	0.22	297.10	$9.0_{\pm1.8}$	37	$11.0_{\pm0.1}$	$11.1_{\pm0.1}$	0.72	0.79	0.87	$1.1_{\pm0.1}$	$1.5_{\pm0.1}$
	$3.0 \le z < 3.4$	9	5	3.19	11.16	2.76	0.38	271.01	$8.8_{\pm1.5}$	30	$11.0_{\pm0.1}$	$11.0_{\pm0.1}$	1.39	1.60	1.60	$2.6_{\pm0.2}$	$2.5_{\pm0.2}$
$11.0 \le \log_{10}$	$0.6 \le z < 1.4$	17	0	0.97	11.17	1.62	−0.03	270.21	$2.0_{\pm0.9}$	5	$10.4_{\pm0.2}$	$10.5_{\pm0.2}$	0.83	0.92	1.02	$1.4_{\pm0.7}$	$2.0_{\pm0.9}$
$M_\star < 12.0$	$1.4 \le z < 2.2$	20	7	1.84	11.19	2.18	0.00	301.15	$3.8_{\pm2.1}$	13	$10.7_{\pm0.2}$	$10.7_{\pm0.2}$	0.69	0.81	0.87	$1.5_{\pm0.2}$	$1.9_{\pm0.2}$
(off the phase	$2.2 \le z < 2.6$	13	6	2.39	11.20	2.40	0.11	284.18	$7.7_{\pm1.9}$	17	$11.0_{\pm0.1}$	$11.0_{\pm0.1}$	0.72	0.77	0.80	$0.9_{\pm0.1}$	$1.2_{\pm0.2}$
center)	$2.6 \le z < 3.0$	8	3	2.80	11.20	2.55	0.19	277.93	$7.4_{\pm1.8}$	16	$11.0_{\pm0.1}$	$11.0_{\pm0.1}$	0.98	1.13	1.17	$1.8_{\pm0.2}$	$2.1_{\pm0.3}$
	$3.0 \le z < 3.4$	6	3	3.18	11.19	2.72	0.33	259.30	$9.3_{\pm1.4}$	15	$11.1_{\pm0.1}$	$11.1_{\pm0.1}$	1.29	1.57	1.57	$2.9_{\pm0.2}$	$2.9_{\pm0.3}$
$10.5 \le \log_{10}$	$0.4 \le z < 0.8$	57	4	0.67	10.69	1.43	0.18	316.73	$0.8_{\pm0.5}$	5	$10.1_{\pm0.2}$	$10.1_{\pm0.2}$	0.60	0.82	0.84	$1.6_{\pm0.7}$	$1.7_{\pm0.8}$
$M_\star < 11.0$	$0.8 \le z < 1.1$	73	2	0.92	10.72	1.58	0.06	298.96	$1.3_{\pm0.9}$	9	$10.3_{\pm0.3}$	$10.3_{\pm0.3}$	0.62	0.85	0.87	$1.9_{\pm0.2}$	$2.0_{\pm0.4}$
	$1.1 \le z < 1.4$	81	12	1.25	10.71	1.83	0.08	274.75	$2.9_{\pm0.6}$	26	$10.6_{\pm0.1}$	$10.6_{\pm0.1}$	0.55	0.64	0.68	$1.2_{\pm0.1}$	$1.4_{\pm0.2}$
	$1.4 \le z < 1.7$	63	6	1.55	10.72	1.99	0.11	299.42	$1.7_{\pm0.6}$	14	$10.4_{\pm0.2}$	$10.4_{\pm0.1}$	0.68	0.73	0.73	$1.0_{\pm0.5}$	$1.0_{\pm0.2}$
	$1.7 \le z < 2.0$	59	6	1.85	10.69	1.96	−0.02	284.15	$2.7_{\pm0.6}$	6	$10.6_{\pm0.1}$	$10.6_{\pm0.1}$	0.54	1.08	1.04	$3.3_{\pm0.4}$	$3.1_{\pm0.8}$
	$2.0 \le z < 2.4$	51	6	2.19	10.72	2.07	−0.04	282.95	$2.0_{\pm0.5}$	8	$10.4_{\pm0.1}$	$10.5_{\pm0.1}$	0.80	0.90	0.94	$1.8_{\pm0.4}$	$2.6_{\pm0.7}$
	$2.4 \le z < 3.0$	62	8	2.69	10.69	2.16	−0.02	279.43	$2.8_{\pm0.6}$	13	$10.6_{\pm0.1}$	$10.5_{\pm0.1}$	0.81	0.98	1.01	$2.2_{\pm0.2}$	$2.0_{\pm0.4}$
	$3.0 \le z < 3.6$	43	7	3.22	10.70	2.10	−0.12	269.51	$3.2_{\pm1.1}$	16	$10.6_{\pm0.1}$	$10.7_{\pm0.1}$	0.63	0.88	0.96	$2.1_{\pm0.2}$	$2.3_{\pm0.2}$
$10.0 \le \log_{10}$	$0.4 \le z < 0.8$	103	0	0.67	10.26	1.24	0.11	302.70	$0.5_{\pm0.5}$	3	$9.9_{\pm0.4}$	$10.1_{\pm0.2}$	0.75	1.26	1.65	$3.0_{\pm1.1}$	$4.3_{\pm2.2}$
$M_\star < 10.5$	$0.8 \le z < 1.1$	131	0	0.94	10.23	1.34	0.02	299.47	$0.7_{\pm0.2}$	4	$10.0_{\pm0.1}$	$9.8_{\pm0.3}$	0.63	1.24	0.88	$3.7_{\pm0.7}$	$2.1_{\pm1.1}$
	$1.1 \le z < 1.4$	95	0	1.24	10.24	1.48	−0.04	270.81	$0.8_{\pm0.8}$	8	$10.1_{\pm0.4}$	$10.2_{\pm0.3}$	0.68	0.73	1.03	$0.8_{\pm0.7}$	$2.7_{\pm0.7}$
	$1.4 \le z < 1.7$	114	1	1.54	10.22	1.67	0.06	289.10	$1.4_{\pm0.2}$	9	$10.3_{\pm0.1}$	$10.4_{\pm0.1}$	0.71	0.82	0.94	$1.4_{\pm0.3}$	$2.2_{\pm0.5}$
	$1.7 \le z < 2.0$	79	1	1.85	10.21	1.67	−0.04	278.14	$0.9_{\pm0.2}$	5	$10.1_{\pm0.1}$	$10.2_{\pm0.1}$	0.76	0.91	1.08	$1.7_{\pm0.7}$	$2.7_{\pm1.2}$
	$2.0 \le z < 2.4$	92	1	2.18	10.22	1.62	−0.14	284.62	$1.1_{\pm0.4}$	4	$10.2_{\pm0.1}$	$10.3_{\pm0.1}$	0.40	0.67	0.71	$1.9_{\pm0.4}$	$2.1_{\pm0.7}$
	$2.4 \le z < 3.0$	79	0	2.74	10.24	1.80	−0.10	269.94	$2.2_{\pm0.9}$	4	$10.5_{\pm0.2}$	$10.4_{\pm0.2}$	0.58	1.20	0.91	$3.5_{\pm0.2}$	$2.3_{\pm1.2}$
	$3.0 \le z < 3.6$	100	1	3.30	10.20	1.79	−0.11	261.39	$1.7_{\pm0.3}$	6	$10.4_{\pm0.1}$	$10.4_{\pm0.1}$	0.84	1.11	1.27	$2.3_{\pm0.7}$	$3.0_{\pm1.2}$

Notes. (1) Stellar mass bin, (2) redshift bin, (3) number of stacked galaxies, (4) number of individually detected stacked galaxies, (5) mean redshift, (6) mean stellar mass, (7) mean SFR, (8) mean ΔMS, (9) mean observed frequency, (10) mean L_{850} inferred from the *uv* domain, (11) peak S/N on the image, (12) mean gas mass inferred from the *uv*-domain, and (13) from the image domain, (14) circularized synthesized beam FWHM, (15) circularized intrinsic FWHM from the *uv* domain, (16) circularized intrinsic FWHM from the image domain, (17) circularized half-light radii from the *uv* domain, and (18) from the image domain.

mass and redshift dependences of evolution of the molecular gas fraction of MS galaxies, we fitted our measurements, together with the local reference, following Liu et al. (2019b), that is,

$$\begin{aligned}\log_{10}\mu_{\rm mol} = &(a + ak\times\log_{10}(M_\star/10^{10}))\times\Delta{\rm MS}\\ &+ b\times\log_{10}(M_\star/10^{10})\\ &+ (c + ck\times\log_{10}(M_\star/10^{10}))\times t_{\rm cosmic}\\ &+ d,\end{aligned} \tag{8}$$

where $t_{\rm cosmic}$ is the cosmic time in units of Gyr, and $M_\star$ is in units of $M_\odot$. Because our analysis does not probe a large dynamic range in ΔMS, we fixed a and ak to the values reported by Liu et al. (2019b) (i.e., $a = 0.4195$ and $ak = 0.1195$, respectively). To constrain the remaining parameters of Eq. (8), we then performed a standard Bayesian analysis using the python Markov chain Monte Carlo (MCMC) package `emcee` (Foreman-Mackey et al. 2013). In this analysis, we accounted for the redshift, stellar mass, and ΔMS of each galaxy in a given stacked bin, that is, in each MCMC step, we compared our stacked measurements, $\langle M^i_{\rm gas}\rangle$, to $\langle f(t^i_{\rm cosmic}, M^i_\star, \Delta{\rm MS}^i)\rangle$ and not $f(\langle t^i_{\rm cosmic}\rangle, \langle M^i_\star\rangle, \langle\Delta{\rm MS}^i\rangle)$, where i is the ith galaxy of our stacked bin and f is the fitted function. This avoids averaging biases that could arise if one would simply fit our stacked measurements using $\langle t^i_{\rm cosmic}\rangle$, $\langle M^i_\star\rangle$, and $\langle\Delta{\rm MS}^i\rangle$. Results of this MCMC analysis are shown in Fig. 6, with $b = -0.468^{+0.070}_{-0.070}$, $c = -0.122^{+0.008}_{-0.008}$, $d = 0.572^{+0.059}_{-0.060}$, and $ck = 0.002^{+0.011}_{-0.011}$. These results unambiguously demonstrate that the molecular gas fraction of MS galaxies decreases with stellar masses (i.e., $b < 0$) while it increases with redshifts (i.e., $c < 0$; see blue solid lines in Fig. 4). We note that repeating this MCMC analysis while fixing $a = 0$ and $ak = 0$ (i.e., considering that our measurements are for ΔMS = 0 galaxies), the likelihood of our fit decreases but the inferred gas fraction evolution remains qualitatively consistent with our previous fit, albeit with a somewhat flatter stellar mass dependence (i.e., $b = -0.252^{+0.072}_{-0.072}$, $c = -0.114^{+0.008}_{-0.008}$, $d = 0.481^{+0.061}_{-0.061}$, and $ck = -0.016^{+0.011}_{-0.011}$).

By comparing our results with those from the A^{3}COSMOS catalog, one immediately notices that these individually detected galaxies systematically lie above our measurements. This systematic offset results from an observational bias. First of all, at a given redshift and stellar mass, the A^{3}COSMOS catalog mostly contains galaxies on the upper part of the MS because those galaxies have higher molecular gas mass ($a > 0$ in Eq. 8) and are thus more likely to be individually detected. This observational bias was, however, accounted for when fitting the A^{3}COSMOS population using Eq. (8). This "correction" can be seen in Fig. 4 by noticing that the A^{3}COSMOS analytical predictions for MS galaxies systematically lies below the A^{3}COSMOS data-points. Nevertheless, at a given redshift, stellar mass, and ΔMS, the A^{3}COSMOS catalog could still be biased toward galaxies with a bright millimeter emission and thus high molecular gas mass (see discussion in Liu et al. 2019b). Our measurements, which are not affected by this bias and which lie systematically below those of Scoville et al. (2017), Liu et al. (2019b), and Tacconi et al. (2020), clearly demonstrate the presence of this residual observational bias in these literature studies. By averaging at a given redshift and stellar mass all MS galaxies in the field, our stacking analysis reveals their true mean molecular gas mass. Taken at face value, our findings imply that previous studies might have systematically overestimated by at least 10–40% the gas content of MS galaxies in redshift and stellar mass bins with relatively high detection fraction (i.e., mostly $M_\star > 10^{11}\,M_\odot$),

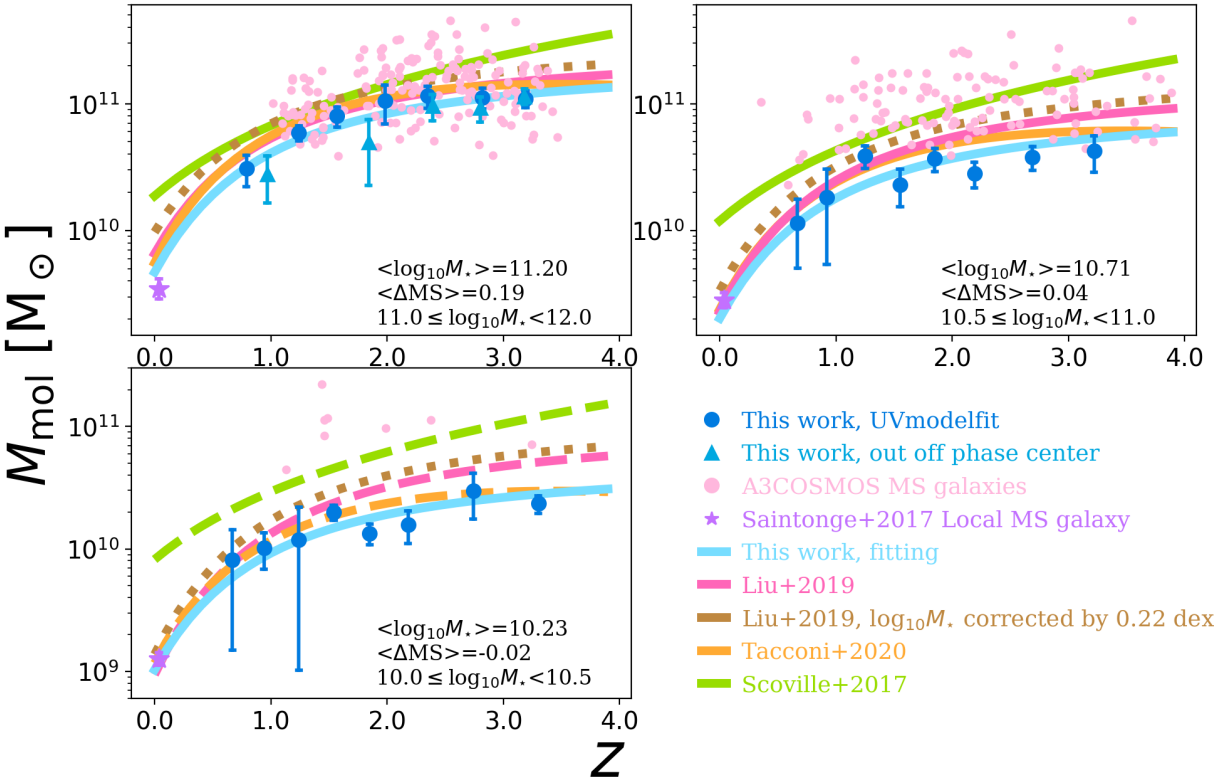

Fig. 4. Redshift evolution of the mean molecular gas mass of MS galaxies in three stellar mass bins, i.e., $10^{11} \leq M_\star/M_\odot < 10^{12}$, $10^{10.5} \leq M_\star/M_\odot < 10^{11}$, and $10^{10} \leq M_\star/M_\odot < 10^{10.5}$. Blue circles show our *uv*-domain measurements, while in the highest stellar mass bin blue triangles show those obtained after excluding ALMA primary-target galaxies from our stacked sample (see Sect. 2.2). Pink circles are individually detected MS galaxies taken from the A^{3}COSMOS catalog (Liu et al. 2019b), while purple stars present the local reference taken from Saintonge et al. (2017). Lines show the analytical evolution of the gas fraction as inferred from our work (blue), from Scoville et al. (2017, green), from Liu et al. (2019b, pink), from Tacconi et al. (2020, orange), and from Liu et al. (2019b, dotted brown line) but this time accounting for the systematic 0.22 dex offset observed between their and our stellar mass estimates. In our lower stellar mass bin, lines from the literature are dashed as they mostly rely on extrapolations. We note that here and in all following figures, the values of $\langle M_\star \rangle$ and $\langle \Delta\mathrm{MS} \rangle$ given in each panel are simply used to plot the analytical evolution of the gas fraction. These values naturally vary for each stacked measurements and are accounted for by our MCMC analysis. This avoids averaging biases that could arise if one simply fit our stacked measurements using $\langle t_\mathrm{cosmic} \rangle$, $\langle M_\star \rangle$, and $\langle \Delta\mathrm{MS} \rangle$.

and by 10–60% in bins with low detection fraction. While significant, one should, however, acknowledge that these offsets remain reasonable considering all the selection biases affecting these previous studies. The impact of this finding for galaxy evolution models is discussed in Sect. 5.

4.2. The molecular gas depletion time of MS galaxies

The redshift evolution of the molecular gas depletion time (i.e., $\tau_\mathrm{mol} = M_\mathrm{mol}/\mathrm{SFR}$) of MS galaxies as inferred from our stacking analysis is shown in Fig. 7, together with analytical predictions from Scoville et al. (2017), Liu et al. (2019b), and Tacconi et al. (2020) as well as the local reference for MS galaxies taken from Saintonge et al. (2017). In addition, in Fig. 8, we compare the redshift evolution of the molecular gas depletion time as inferred for our three stellar mass bins. Again, to obtain a more quantitative constraint on the stellar mass and redshift dependences of the molecular gas depletion time of MS galaxies, we fitted our measurements, together with the local reference, following Liu et al. (2019b), that is,

$$\begin{aligned}\log_{10}\tau_\mathrm{mol} = {} & (a + ak \times \log_{10}(M_\star/10^{10})) \times \Delta\mathrm{MS} \\ & + b \times \log_{10}(M_\star/10^{10}) \\ & + (c + ck \times \log_{10}(M_\star/10^{10})) \times t_\mathrm{cosmic} \\ & + d.\end{aligned} \tag{9}$$

Our analysis does not probe a large dynamic range in ΔMS, we thus fixed a and ak to the values reported by Liu et al. (2019b), that is, $a = -0.5724$ and $ak = 0.1120$. Results of our MCMC analysis are shown in Fig. 9, with $b = 0.055^{+0.069}_{-0.071}$, $c = 0.049^{+0.008}_{-0.008}$, $d = -0.643^{+0.056}_{-0.057}$, and $ck = 0.016^{+0.010}_{-0.010}$. Because the depletion time is the ratio of M_mol by SFR, we also display in Fig. 7 the redshift evolution of depletion time as one would infer by dividing $M_\mathrm{gas}(z, M_*, \Delta\mathrm{MS})$ from Eq. (8) by the $\mathrm{SFR_{MS}}(z, M_*, \Delta\mathrm{MS})$ from Leslie et al. (2020, dash-dotted light-blue line).

In all our stellar mass bins, the molecular gas depletion time of MS galaxies decreases by a factor of ∼3–4 from $z \sim 0$ to $z \sim 3.2$, with, however, most of this decrease happening at $z \lesssim 1.0$. At $z \gtrsim 1$, the molecular gas depletion time of MS galaxies remains instead roughly constant with redshifts and stellar masses with a value of ∼300–500 Myr. While such evolution is qualitatively predicted by all literature studies, its amplitude as well as its exact redshift and stellar mass dependences quantitatively disagree (see Fig. 7). For example, our measurements and those from Liu et al. (2019b) agree at high stellar masses, but differ by ∼30–40% in our lower stellar mass bins. These differences are likely explained by the observational biased discussed in Sect. 4.1, which implies that the mean molecular gas mass and thus depletion time of MS galaxies inferred by Liu et al. (2019b) are slightly overestimated especially at low stellar masses. The same effect likely explains the ∼20–30% overestimation of the

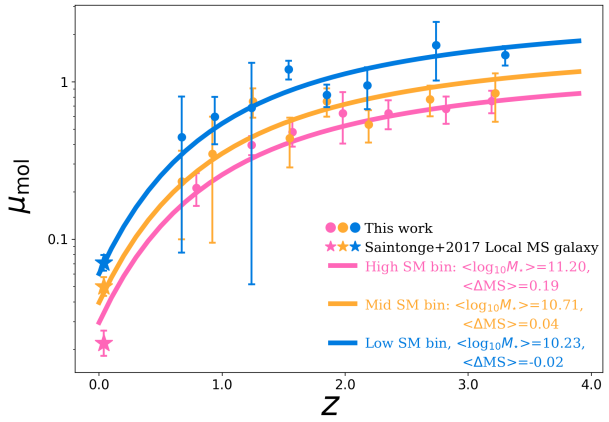

Fig. 5. Redshift evolution of the mean molecular gas fraction of MS galaxies. Circles show the mean molecular gas fraction from our work. Stars present the local reference taken from Saintonge et al. (2017). Lines display the analytical evolution of the molecular gas fraction inferred from our work. Symbols and lines are color-coded by stellar mass, i.e., pink for $10^{11} \leq M_\star/M_\odot < 10^{12}$, orange for $10^{10.5} \leq M_\star/M_\odot < 10^{11}$, and blue for $10^{10} \leq M_\star/M_\odot < 10^{10.5}$.

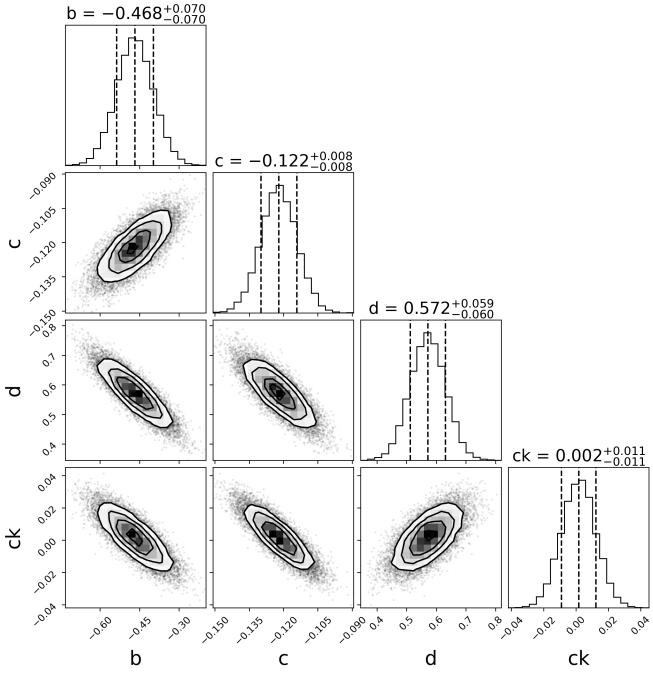

Fig. 6. Probability distributions of the parameters in Eq. (8), as found by fitting our stacked measurements using an MCMC analysis. The dashed vertical lines show the 16th, 50th, and 84th percentiles of each distribution.

molecular gas depletion time inferred in Tacconi et al. (2020) in most redshift–stellar mass bins probed here. While our direct analytical fit of the redshift/stellar mass evolution of the molecular gas depletion time (solid blue lines in Fig. 7) matches relatively well the local reference from Saintonge et al. (2017), this is not the case of our fit inferred by simply dividing $M_{\rm gas}(z, M_*, \Delta{\rm MS})$ from Eq. (8) by ${\rm SFR_{MS}}(z, M_*, \Delta{\rm MS})$ from Leslie et al. (2020, dash-dotted light-blue line). This disagreement between predictions and observations at $z \sim 0$ is entirely attributed to a miss-match in ${\rm SFR_{MS}}(z, M_*, \Delta{\rm MS})$, that is, at a given stellar mass, the mean SFR of $z \sim 0$ MS galaxies as predicted by Leslie et al. (2020) does not match that observed by Saintonge et al. (2017). This disagreement is, however, not unexpected as the sample used in Leslie et al. (2020) was restricted to $z > 0.3$ galaxies.

In general, we conclude that our depletion times agree at high stellar masses with Liu et al. (2019b), that is, where their study relies on a large and robust amount of ALMA-based measurements of MS galaxies; while our depletion time agree better at low stellar masses with Tacconi et al. (2020), that is, where their study, contrary to that of Liu et al. (2019b), still relies on some observational measurements of MS galaxies thanks to their *Herschel* stacking analysis. Like our measurements, those from Tacconi et al. (2020) predict only a minor evolution of the molecular gas depletion time of MS galaxies with stellar masses. This implies that the flattening of the MS at high stellar masses observed in most studies (i.e., $\log_{10}\ {\rm SFR^{MS}} = 0.7 \times \log_{10}\ M_\star^{\rm MS} + {\rm C}$) is not associated with or due to lower SFEs (i.e., $1/\tau_{\rm mol}$) in massive systems but rather lower molecular gas fraction (see Sect. 4.1). This is further discussed in Sect. 5. In addition, we note that extrapolating our molecular gas depletion time predictions to $z \sim 5$, that is, $\tau_{\rm mol}^{\rm pred} = 250$ Myr (from Eq. (9)) or $\tau_{\rm mol}^{\rm pred} = 620$ Myr (from Eq. (8)/specific ${\rm SFR_{MS}}$), our prediction qualitatively agrees with the latest observational constraints from the ALPINE [C II] ALMA large project (i.e., $\tau_{\rm mol}^{\rm obs} = 520 \pm 70$ Myr) (Dessauges-Zavadsky et al. 2020).

Finally, we note that the accuracy of the depletion times relies not only on accurate gas masses but also on accurate SFRs. In our study, the latter were estimated using the so-called ladder of SFR indicators, that is, by applying to each galaxy the best dust-corrected star-formation indicator available (Sect. 2.2). In particular, among the 1376 galaxies in our final sample with stellar mass $>10^{10}\,M_\odot$, 852 (62%) have very robust dust-corrected SFRs based on the combination of infrared and ultraviolet measurements. Of the remaining 524 galaxies whose SFRs are solely based on their ultraviolet-to-optical fits, most (470) should also have robust SFRs, as they falls below the $\sim 100\,M_\odot\,{\rm yr}^{-1}$ limit above which ${\rm SFR_{SED}}$ starts to be systematically underestimated (Fig. 2). We verified that our results remain unchanged (within the uncertainties) when excluding from our stacking analysis these 54 galaxies with ${\rm SFR_{SED}} > 100\,M_\odot\,{\rm yr}^{-1}$ and without infrared detection.

4.3. The FIR sizes of MS galaxies

Our stacking analysis provides the first measurements of the mean FIR size of MS galaxies across cosmic time. These mean FIR (at the observed-frame 850–1300 µm) sizes of MS galaxies are presented in Fig. 10, and compared to optical, FIR and radio sizes measurements from van der Wel et al. (2014), Barro et al. (2016), Rujopakarn et al. (2016), Elbaz et al. (2018), Jiménez-Andrade et al. (2019), Suess et al. (2019), Chang et al. (2020), and Tadaki et al. (2020). Because most of the FIR and radio size measurements from the literature were made on the image-plane, we displayed in Fig. 10 our 2D Gaussian image-plane measurements inferred using PyBDSF. Displaying instead our *uv*-plane size measurements would, however, not change any of our conclusions, as both agree within their uncertainties with no apparent systematic offset between them.

These FIR sizes do not seem to evolve significantly with redshift or stellar mass, with a mean circularized effective – equivalently half-light – radius of 2.2 kpc (i.e., corresponding

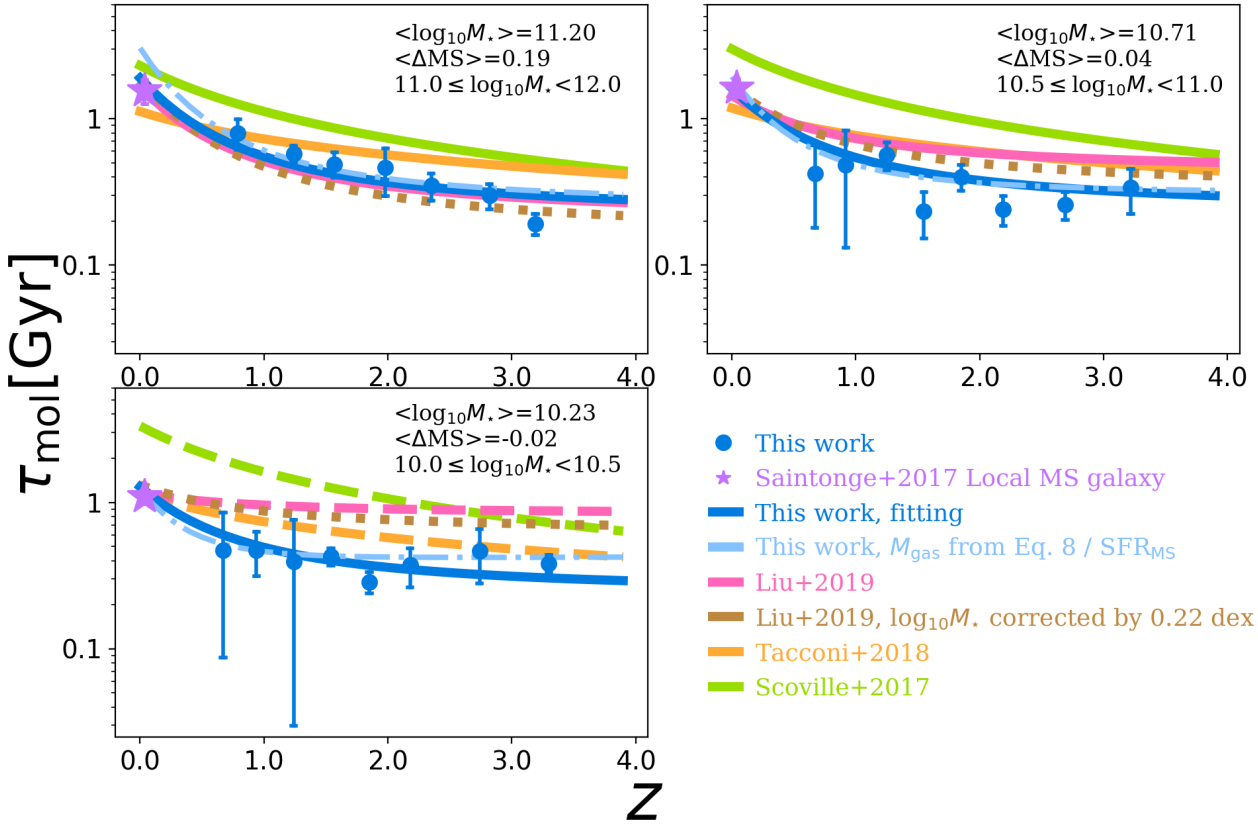

Fig. 7. Redshift evolution of the molecular gas depletion time of MS galaxies in three stellar mass bins, i.e., $10^{11} \leq M_\star/M_\odot < 10^{12}$, $10^{10.5} \leq M_\star/M_\odot < 10^{11}$, and $10^{10} \leq M_\star/M_\odot < 10^{10.5}$. Blue circles show our *uv*-domain molecular gas mass measurements divided by the mean SFR of each of these stacked samples. Purple stars show the local MS reference taken from Saintonge et al. (2017). Lines present the analytical evolution of the molecular gas depletion time as inferred from our work (blue lines; see text for details), from Liu et al. (2019b, pink line), from Scoville et al. (2017, green line), and from Tacconi et al. (2020, orange line). In our lower stellar mass bin, lines from the literature are dashed as they mostly rely on extrapolations.

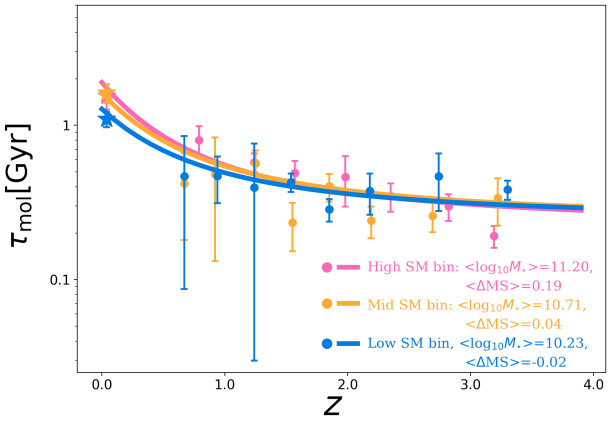

Fig. 8. Redshift evolution of the molecular gas depletion time of MS galaxies. Circles show the mean molecular gas depletion time from our work. Stars show the local MS reference taken from Saintonge et al. (2017). Lines display the analytical evolution of the molecular gas fraction inferred from our work. Symbols and lines are color-coded by stellar mass, i.e., pink for $10^{11} \leq M_\star/M_\odot < 10^{12}$, orange for $10^{10.5} \leq M_\star/M_\odot < 10^{11}$, and blue for $10^{10} \leq M_\star/M_\odot < 10^{10.5}$.

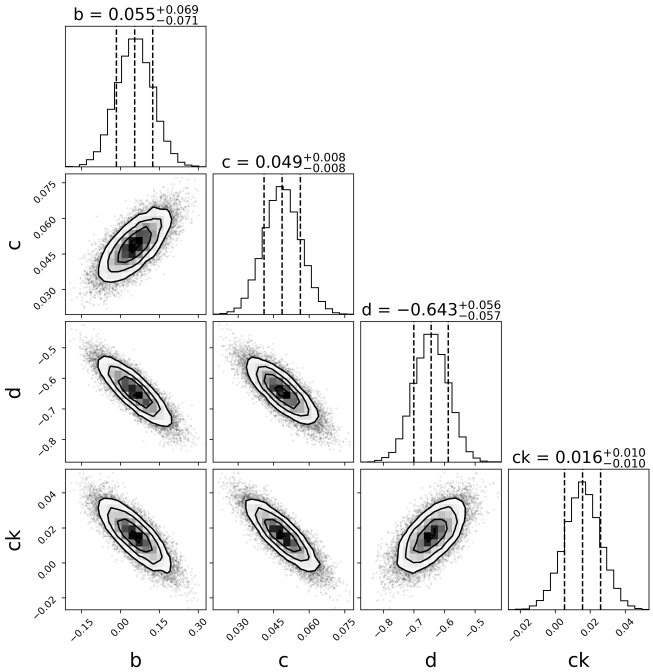

Fig. 9. Probability distributions of the parameters in Eq. (9) as found by fitting our stacked measurements using an MCMC analysis. The dashed vertical lines show the 16th, 50th, and 84th percentiles of each distribution.

to a median angular size-to-synthesized beam FWHM ratio of 1.5). Because there is a possible mismatch between our stacked position and the actual millimeter position of the sources (e.g., Elbaz et al. 2018), these average FIR sizes could, however, be

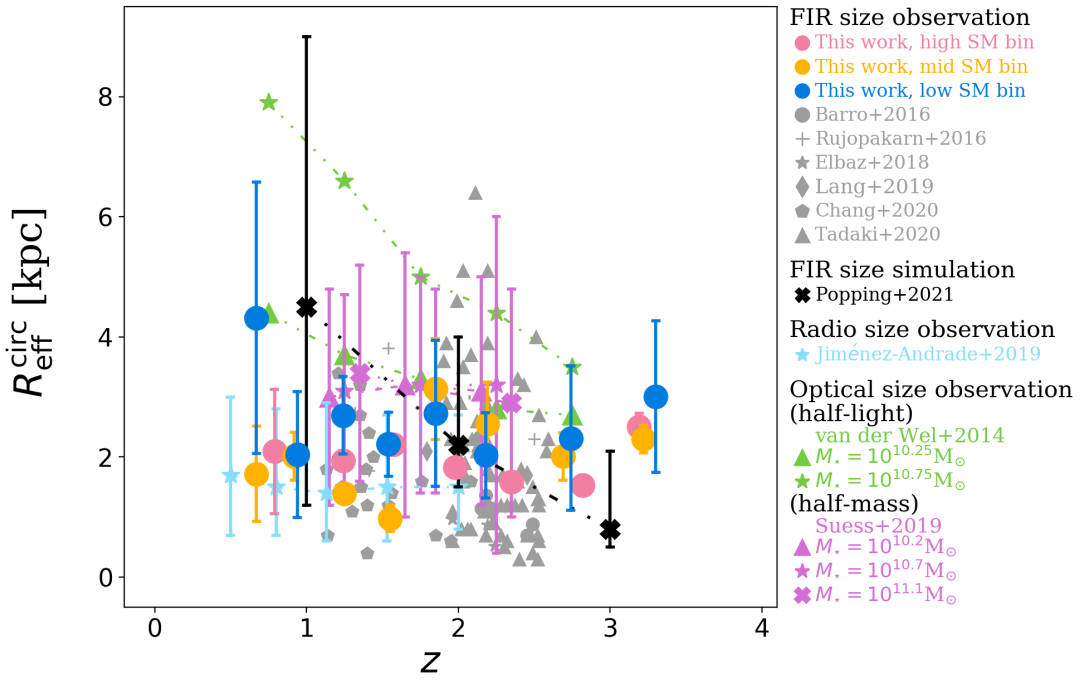

Fig. 10. Redshift evolution of the half-light (or half-mass) radius of MS galaxies. Pink, orange, and blue circles present our stacking results for our high, mid, and low stellar mass bins, i.e., $10^{11} \leq M_\star/M_\odot < 10^{12}$, $10^{10.5} \leq M_\star/M_\odot < 10^{11}$, and $10^{10} \leq M_\star/M_\odot < 10^{10.5}$, respectively. The gray data points are the FIR sizes of MS galaxies from Barro et al. (2016, dots), Rujopakarn et al. (2016, pluses), Elbaz et al. (2018, stars), Lang et al. (2019, diamonds), Chang et al. (2020, pentagons), and Tadaki et al. (2020, triangles); light blue stars are radio sizes of MS galaxies from Jiménez-Andrade et al. (2019). The black crosses are the FIR sizes of $M_\star = 10^{10.5}\,M_\odot$ MS galaxies from the simulations of Popping et al. (2022). The green triangles and stars are the optical half-light sizes of $M_\star = 10^{10.25}\,M_\odot$ and $M_\star = 10^{11.25}\,M_\odot$ MS galaxies from van der Wel et al. (2014). Finally, the purple triangles, stars, and crosses are the optical half-mass sizes of $M_\star = 10^{10.25}\,M_\odot$, $M_\star = 10^{10.75}\,M_\odot$, and $M_\star = 10^{11.25}\,M_\odot$ MS galaxies from Suess et al. (2019). Because most of these literature studies relied on image-plane fits, the stacked FIR sizes displayed here are those from our 2D Gaussian image-plane fits done using `PyBDSF`.

slightly overestimated. Nevertheless, such bias does not seem to be significant as our measurements agree qualitatively and quantitatively with the mean star-forming size of massive ($M_\star \sim 10^{10.7-11.7}\,M_\odot$) MS galaxies inferred by the most recent literature studies. In contrast, the half-light stellar size of massive MS galaxies is typically larger than these FIR extents by a factor of 2 and 4 at $z \sim 3$ and 1, respectively (see Fig. 10; van der Wel et al. 2014). In lower-mass MS galaxies (i.e., $M_\star \sim 10^{10.3}\,M_\odot$), larger half-light stellar size than FIR extents are also observed but mostly at low redshifts. As discussed in Sect. 5, this apparent discrepancy between optical and FIR sizes of MS galaxies does not, however, necessarily translate into stellar half-mass radius discrepancy, as complex obscuration biases need to be accounted for when converting half-light stellar radius into half-mass stellar radius (e.g., Lang et al. 2019; Suess et al. 2019; Popping et al. 2022). For example, the FIR sizes inferred in our study agree quantitatively with the mean redshift-independent half-mass stellar radius of SFGs measured by Suess et al. (2019).

4.4. The Kennicutt-Schmidt relation

Combining our half-light FIR radii (from our image-plane fits), molecular gas mass, and SFR measurements, we study in Fig. 11 the relation between the SFR and gas mass surface densities of MS galaxies (i.e., $\Sigma_{\rm SFR} = {\rm SFR}/(2\pi\,R_{\rm eff-py}^{\rm circ\ 2})$ versus $\Sigma_{M_{\rm mol}} = M_{\rm mol}/(2\pi\,R_{\rm eff-py}^{\rm circ\ 2})$; the so-called KS relation). We compare our estimates with results from the literature: local normal and starburst galaxies from Kennicutt (1998a, hereafter K98) and de los Reyes & Kennicutt (2019) (taking only their molecular gas phase measurements and thus excluding contribution from the atomic gas phase) as well as the global fit of the KS relation from K98; $\langle z \rangle = 1.2$ MS galaxies from Daddi et al. (2010) as well as their MS-only galaxies fit of the KS relation; $\langle z \rangle = 1.5$ MS galaxies from Davis et al. (2007), Noeske et al. (2007), and Tacconi et al. (2010); $\langle z \rangle = 2.3$ MS galaxies from Erb et al. (2006); and finally the MS-only galaxies fit of the KS relation from Genzel et al. (2010). We note that here we take the FIR size of galaxies as a proxy of their SFRs and gas mass distributions (under the hypothesis that the dust and gas are co-spatial). This assumption is justified by recent simulations in which the FIR half-light radius of galaxies is found to be consistent with the radius containing half their star formation and to be only slightly more compact than the radius containing half their molecular gas mass, at least in $z \lesssim 2$ galaxies (Popping et al. 2022).

There is a tight correlation between the $\Sigma_{M_{\rm mol}}$ and $\Sigma_{\rm SFR}$ of MS galaxies, with no significant dependences of this relation on stellar mass or redshift; in other words, at a given $\Sigma_{M_{\rm mol}}$, measurements from different stellar mass or redshift bins agree within their uncertainties. Our measurements are consistent with previous individually detected MS galaxy estimates while they fall below those from individually detected starbursts. In general, at a given redshift, MS galaxies with higher stellar masses are located at the higher end of the $\Sigma_{\rm SFR} - \Sigma_{M_{\rm mol}}$ relation due to the increase in their molecular gas content and the absence of significant size evolution with stellar mass, which translates into an overall increase in their $\Sigma_{M_{\rm mol}}$. Similarly, at a given stellar mass, MS galaxies at higher redshifts are mostly located at the higher end of the $\Sigma_{\rm SFR} - \Sigma_{M_{\rm mol}}$ relation due to the increase in their molecular gas content and the absence of significant size evolution

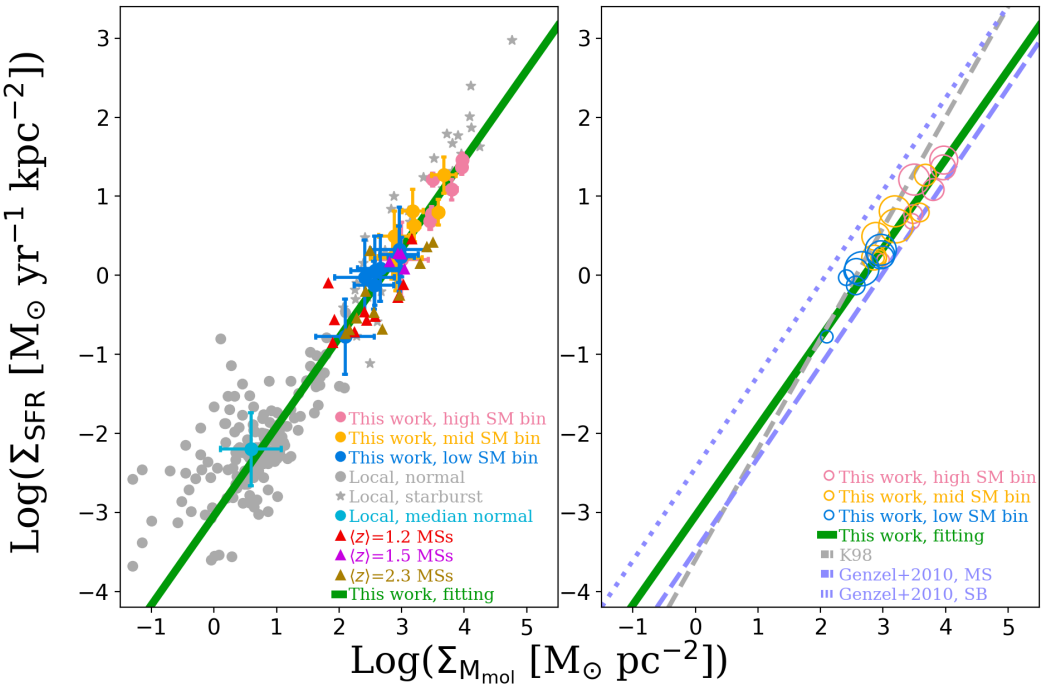

Fig. 11. Relation between the SFR and gas mass densities of SFGs, i.e., the so-called KS relation. *Left*: our stacking results for our high, mid, and low stellar mass bins, i.e., $10^{11} \leq M_\star/M_\odot < 10^{12}$, $10^{10.5} \leq M_\star/M_\odot < 10^{11}$, and $10^{10} \leq M_\star/M_\odot < 10^{10.5}$, respectively, shown by pink, orange, and blue circles. Gray circles and stars are normal and starburst local galaxies from K98 and de los Reyes & Kennicutt (2019, taking only their molecular gas mass estimates, i.e., excluding the atomic phase). Red, purple, and brown triangles are $\langle z \rangle = 1.2$ (Daddi et al. 2010), $\langle z \rangle = 1.5$ (Davis et al. 2007; Noeske et al. 2007; Tacconi et al. 2010), and $\langle z \rangle = 2.3$ (Erb et al. 2006) MS galaxies, respectively. The green line is a fit to the KS relation considering only MS galaxies, i.e., our measurements together with the K98 normal local galaxy average (turquoise dot). *Right*: comparison of our MS-only KS relation to the global fit of K98, the MS-only (long dashed blue line) and starburst-only (dotted blue line) fits of Genzel et al. (2010). Open circles show our measurements, with symbol size increasing with redshift and color-coded by stellar masses.

with redshifts, which also translates into an overall increase in their $\Sigma_{M_{\rm mol}}$.

We performed a linear fit of the KS relation of MS-only galaxies in log-log space, combining our high-redshift MS galaxies measurements with those from the local Universe obtained by K98 and de los Reyes & Kennicutt (2019),

$$\log_{10}\Sigma_{\rm SFR} = (1.13 \pm 0.09) \cdot \log_{10}\Sigma_{M_{\rm mol}} - (3.06 \pm 0.33). \quad (10)$$

The inferred power index of the MS-only KS relation (i.e., $\alpha = 1.13$) is smaller than that found by Daddi et al. (2010) considering MS-only galaxies ($\alpha = 1.42$), but similar to that found by Genzel et al. (2010) for MS-only galaxies ($\alpha = 1.17$). We note that previous high-redshift investigations (i.e., Daddi et al. 2010; Genzel et al. 2010) were only based on relatively small samples of massive high-redshift SFGs (i.e., $N < 50$, $z > 1$, and $M_\star > 10^{11}\,M_\odot$), and are thus likely limited by selection biases. A power law index for the MS-only KS relation that is greater than unity implies that the depletion time (i.e., $\tau_{\rm mol}$) – equivalently, the SFE (i.e., $1/\tau_{\rm mol}$) – of MS galaxies is controlled by their $\Sigma_{M_{\rm mol}}$. In other words, the evolution of the depletion time with redshift and stellar mass seen in Fig. 7 can be predicted from their $\Sigma_{M_{\rm mol}}$ and this universal redshift-independent MS-only KS relation. The KS of MS-only galaxies remains thus one of the most fundamental relation to understand the stellar mass growth of the Universe over the last 10 Gyr.

Finally, as already pointed out by, for example, Daddi et al. (2010) and Genzel et al. (2010), we found that MS galaxies seems to follow a KS relation that at high $\Sigma_{M_{\rm mol}}$ falls below the relation followed by starburst galaxies. In this high $\Sigma_{M_{\rm mol}}$ regime, starbursts exhibit SFEs that are two to three times higher.

We note that in this analysis we implicitly assume that the dust and gas are co-spatial. However, this assumption might not always be verified, as suggested by some ALMA high-resolution observations of submillimeter-selected galaxies (e.g., Chen et al. 2017; Calistro Rivera et al. 2018), which revealed around two times more compact dust continuum emission than gas CO emission. Increasing our FIR sizes by a factor of 2 would shift our data points toward lower surface densities along the one-to-one line in the log-log space but would not significantly change the slope of the inferred KS relation. However, such a large offset or discrepancy in spatial distribution would also translate into very uncertain dust-based gas mass measurements and would thus impact in a more complex way the inferred KS relation. Regardless, submillimeter-selected galaxies are extreme object located far above the MS (e.g., Magnelli et al. 2012a; Casey et al. 2014) and MS galaxies do not seem to exhibit any significant discrepancies between their gas and dust sizes (Puglisi et al. 2019).

4.5. Limitations and uncertainties

Naturally, our analysis suffers from a number of limitations and uncertainties. Those can be mostly divided into two categories: those inherent to all studies measuring molecular gas masses from single RJ dust continuum flux densities; and those specifically associated with our stacking analysis that are related to the averaged nature of our stacked measurements. In the following, we try to exhaustively list these limitations and uncertainties, and discuss their impact on the main conclusions of our analysis.

4.5.1. From observed-frame flux densities to molecular gas masses

To convert observed-frame flux densities into molecular gas masses, we applied a two-step approach, first converting observed-frame flux densities into rest-frame 850 µm luminosities using a standard SED template (the so-called k-correction) and then converting these rest-frame luminosities into molecular gas masses using a standard L_{850}-to-$M_{\rm mol}$ relation. To study how these particular choices of SED templates and L_{850}-to-$M_{\rm mol}$ relations influence our results, we repeated our analysis using alternatives commonly adopted in the literature.

Instead of using the SED template of Béthermin et al. (2012) to perform our k-corrections, we repeated our analysis using the SED template of Schreiber et al. (2018) or a single gray-body emission with $T_{\rm dust}$ = 25 K and β = 1.8 (as it is assumed in, e.g., Scoville et al. 2016). These two k-correction methods yield molecular gas masses that are, respectively, 12% higher and 16% lower at $z \sim 0.6$ than our original calculation and 5% higher and 5% lower at $z \sim 3.2$ than our original calculation. Because our original k-corrections are bracket by these alternatives and because the inferred offsets are in any cases well within the uncertainties of our original constraints, we conclude that the specific choice of this SED template has no significant impact on our results.

As extensively discussed in Liu et al. (2019b), systematic offsets are found between all different metallicity-dependent or -independent L_{850}-to-$M_{\rm mol}$ relations. We evaluate the impact of these relations on our results in Appendix B by repeating our analysis using instead of the metallicy-independent L_{850}-to-$M_{\rm mol}$ relation of Hughes et al. (2017, hereafter H17), (i) the H17 relation inferred assuming a $\alpha_{\rm CO}$ of 4.35 $M_\odot\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$ (hereafter H17$_{\alpha_{\rm CO}=4.35}$) instead of 6.5 $M_\odot\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$; (ii) the metallicity-independent L_{850}-to-$M_{\rm mol}$ relation of Scoville et al. (2016, hereafter S16), (iii) a L_{850}-to-$M_{\rm dust}$ relation assuming $T_{\rm dust}$ = 25 K and β = 1.8 combined to the metallicity-dependent $M_{\rm dust}$-to-$M_{\rm mol}$ relation of Bertemes et al. (2018, hereafter B18) and finally (iv) a L_{850}-to-$M_{\rm dust}$ relation assuming $T_{\rm dust}$ = 25 K and β = 1.8 combined to the metallicity-dependent $M_{\rm dust}$-to-$M_{\rm gas}$ relation of Leroy et al. (2011, hereafter $\delta_{\rm GDR}$). As for our k-corrections, our original calculation (i.e., H17) yields estimates that are bracket by these alternatives: H17$_{\alpha_{\rm CO}=4.35}$ produces estimates that are systematically lower than ours by ∼0.17 dex, B18 gives values that are systematically lower than ours by ∼0.17 dex; S16 yields values that are consistent with those reported here within ∼0.04 dex; while the $\delta_{\rm GDR}$ method produces estimates that are systematically higher than ours by ∼0.13 dex. In addition to this global offsets, the metallicity-dependent methods (i.e., $\delta_{\rm GDR}$ and B18) introduces redshift- and stellar mass-dependent trends, which are due to the fact that lower-mass and higher-redshift galaxies have increasingly lower metallicities. Consequently, the offsets between our measurements and those inferred with the $\delta_{\rm GDR}$ method increase toward lower masses and higher redshifts, while the offsets with B18 decrease toward lower masses and higher redshifts. Even if present, these stellar mass- and redshift-dependent trends do not qualitatively change the main conclusions of our papers: irrespective of the assumed methods: (i) the molecular gas fraction of MS galaxies still increases with redshifts and decreases with stellar masses; (ii) their depletion time still remains mostly dependent on their redshifts and not their stellar masses and finally, (iii) MS galaxies still evolve along a seemingly universal MS-only KS relation. These evolutions are, however, quantitatively changed, with, for example, the molecular gas fraction of MS galaxies increasing by a factor of ∼15, ∼45, ∼33, and ∼17 from $z \sim 0$ to $z \sim 3.2$ for the H17$_{\alpha_{\rm CO}=4.35}$, $\delta_{\rm GDR}$, S16, and B18 gas mass calibrations, respectively, as compared to the factor of 24 found for the H17 method; and their molecular gas depletion time being 200–300, 400–600, 400–700, and 200–500 Myr for the H17$_{\alpha_{\rm CO}=4.35}$, $\delta_{\rm GDR}$, S16, and B18 gas mass calibrations, respectively, instead of 300–500 Myr for our original calculation. Again, our original constraints are roughly bracket by these alternatives and values are consistent within 1–2σ.

We note that combining the local measurements from Saintonge et al. (2017) with our high-redshift H17 estimates yields a redshift evolution of the molecular gas content of SFGs that is in very good agreement with that of Tacconi et al. (2020) at high stellar masses (Fig. 4), where this latter can be considered as the reference as it is based on a fairly complete sample of massive SFGs across cosmic time and a thorough cross-calibration of the CO- and dust-based methods. This agreement could seem surprising as the L_{850}-to-$M_{\rm mol}$ relation of H17 was calibrated using $\alpha_{\rm CO}$ = 6.5, while our local reference (i.e., Saintonge et al. 2017) converted their CO measurements into molecular gas masses using $\alpha_{\rm CO} \sim 4$ (at the high stellar masses of our study). This agreement between our high-redshift H17 measurements and those from Tacconi et al. (2020) is due to the fact that using $\alpha_{\rm CO}$ = 6.5 instead of 4.3 to calibrate the local L_{850}-to-$M_{\rm mol}$ relation corrects indirectly (and to first order) for the fact that at a given stellar mass, high-redshift galaxies have lower metallicities than local galaxies, and thus have a higher gas-to-dust ratio (e.g., Leroy et al. 2011) and consequently should have a lower L_{850}-to-$M_{\rm mol}$ ratio.

Finally, we note that the particular choice of a L_{850}-to-$M_{\rm mol}$ relation cannot explain the differences observed between the molecular gas masses of MS galaxies at a given stellar mass and redshift inferred by our study and that from Liu et al. (2019b) and Scoville et al. (2016). Indeed, Liu et al. (2019b) also used H17 to infer their molecular gas mass estimates, while the method used in Scoville et al. (2017) (i.e., S16) provides consistent results with H17 (within ∼0.04 dex). In both cases, differences between our and their measurements are likely caused by the fact that these literature studies were largely biased toward individually detected MS galaxies with massive gas reservoirs.

4.5.2. Limitation and uncertainties associated with stacking

Stacking in the *uv* domain is a difficult task and could be subject to a series of potential downfalls when applied to the heterogeneous A^{3}COSMOS database. To test the reliability of our stacking analysis, we used realistic simulations, in which mock sources with different flux densities and sizes were introduced in an A^{3}COSMOS-like interferometric database and subsequently stacked using the same procedure as the real sources. The results of these simulations, which are shown in Appendix A, unambiguously demonstrate the reliability of our stacking analysis to accurately retrieve the intrinsic flux densities and sizes of a stacked population. As a reminder, performing such a stacking analysis on images with drastically different spatial resolutions would be virtually impossible or very uncertain.

While our simulations demonstrated that we were able to accurately measure the mean flux density of a galaxy population, one still has to remember that these stacked measurements are averaged values for galaxy populations with intricate stellar mass, SFR, and redshift distributions. How these mean molecular gas measurements, $\langle M_{\rm mol}\rangle$, can be related to $\langle t_{\rm cosmic}\rangle$, $\langle M_\star\rangle$, and $\langle\Delta{\rm MS}\rangle$ to infer $\mu_{\rm gas}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$ and $\tau_{\rm mol}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$ is an none trivial question and depends on

the intrinsic stellar mass, SFR, and redshift distributions of each stacked populations. To address this issue, we applied a Bayesian analysis in which using the true distributions of these stacked populations we estimated the likelihood to measure their stacked $\langle M_{\rm mol}\rangle$ for a given analytical functions of $\mu_{\rm gas}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$ and $\tau_{\rm mol}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$. Without such an approach (i.e., simply inferring $\mu_{\rm gas}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$ and $\tau_{\rm mol}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$ by fitting $\langle M_{\rm mol}\rangle$, $\langle t_{\rm cosmic}\rangle$, $\langle M_\star\rangle$, and $\langle\Delta{\rm MS}\rangle$), those analytical functions would differ from our original calculation by up to 25%. While this approach is currently the best way to deal with this averaging issue, only future high-sensitivity ALMA observations that individually detect all our MS galaxies would be able to definitively constrain $\mu_{\rm gas}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$ and $\tau_{\rm mol}(M_*, t_{\rm cosmic}, \Delta{\rm MS})$.

Recent findings in the literature suggests that there might exist an mean offset of 0.″2 between the optical position and the actual (sub)millimeter position of our sources (e.g., Elbaz et al. 2018). While we verified via simulations that such an offset does not affect significantly our stacked flux density and size measurements, one should at worst consider our molecular gas sizes as upper limits or more realistically keep in mind that those should be corrected from this extra convolution kernel. We decided, however, not to correct our size measurements from this effect in Table 1 because this optical-to-FIR position offset still needs to be confirmed and because deconvolving those intrinsic sizes (i.e., $\theta^{\rm circ}_{\rm mol-}{}_{uv}$ or $\theta^{\rm circ}_{\rm mol-py}$) by an Gaussian kernel with a FWHM of 0.″2 would only have lowered them by ∼6%, leaving all our results unchanged.

5. Discussion

Our analysis reveals that (i) the molecular gas fraction of MS galaxies increases with redshift and decreases with stellar mass; (ii) the depletion time of MS galaxies does not depend on their stellar masses but mostly on their redshift, increasing from 0.4 Gyr at $z \sim 3.6$ to 1.3 Gyr at $z \sim 0$; (iii) the FIR size of MS galaxies does not evolve with redshifts nor stellar masses, with a mean half-light radius of 2.2 kpc; and finally, (iv) MS galaxies evolve along a seemingly universal MS-only KS relation with a slope of ∼1.13.

In the following, we discuss some of these results in light of recent observational findings and galaxy evolution scenarios.

5.1. A universal Kennicutt-Schmidt relation

It is crucial to accurately measure the relation between the SFR and gas (surface) densities of galaxies because it provides theoretical models with key information about the mechanisms and efficiency with which these galaxies turn their gas into stars. The pioneering work of Schmidt (1959) suggests that in the Galactic plane the SFR volume density ($\rho_{\rm SFR}$) is proportional to the gas volume density ($\rho_{\rm gas}$) with a power law index of ∼2. This power law index directly reflects the physical conditions for star formation and can be studied, assuming a constant gas scale height, via the observationally more convenient $\Sigma_{\rm gas}$–$\Sigma_{\rm SFR}$ relation (i.e., the so-called KS relation). At sub-kiloparsec scales, this power law index is found to vary from $\alpha \sim 0.75$ to 2, rendering difficult any theoretical interpretation of this relation at small scales (e.g., Wong & Blitz 2002; Schuster et al. 2007; Bigiel et al. 2008; Leroy et al. 2013; Momose et al. 2013; Miura et al. 2014; Shetty et al. 2014; Wilson et al. 2019; Wang & Hwang 2020; Ellison et al. 2021; Pessa et al. 2021; Sánchez et al. 2021). For example, a power law index of ∼0.75 is expected if giant molecular clouds (GMCs) convert all their gas into stars over a free-fall time (e.g., Krumholz & McKee 2005), while a power index of ∼2.0 is expected if star formation is mostly induced by collisions of small clouds of gas (e.g., Wyse 1986; Wong & Blitz 2002).

K98 provides the first accurate measurement of the $\Sigma_{\rm gas}$–$\Sigma_{\rm SFR}$ relation at global scales by combining data from both normal and starburst galaxies. K98 find $\alpha = 1.4$, which is near the expected value of 1.5 for self-gravitating disks if the SFR scales as the ratio of gas volume density ($\rho_{\rm gas}$) to the free-fall timescale ($\rho_{\rm gas}^{-0.5}$). However, Bigiel et al. (2008) argue that the conditions for star formation in starbursts are too different to be combined with normal galaxies (e.g., Gao & Solomon 2004; Rosolowsky & Blitz 2005) and find $\alpha = 1.0$ when considering only kiloparsec-scale star-forming regions of nearby spirals. They conclude that stars are forming in GMCs with relatively uniform properties and that at supra-kiloparsec scales, star formation remains unresolved and $\Sigma_{\rm SFR}$ becomes thus more a measure of the filling fraction of GMCs than changes in conditions for star formation. Recent studies at high redshift and global scales support this conclusion (e.g., Genzel et al. 2010; Miettinen et al. 2017a). In particular, Genzel et al. (2010) argue that normal and starburst galaxies seem to follow two different KS relations both with a near-unity power law index but with different normalization, the latter galaxy population being more efficient in turning gas into stars. Our analysis also supports the existence of a universal KS relation for MS galaxies with a near-unity power law index. Our results are the first to extend this finding up to $z \sim 3.6$ and using a mass-complete sample of $>10^{10}\,M_\odot$ MS galaxies. We note in particular that the extension of our analysis to very high redshift is crucial because high-redshift MS galaxies have sufficiently high $\Sigma_{\rm gas}$ to provide adequate leverage to accurately constrain the power law index of the KS relation on global scales. The high $\Sigma_{\rm gas}$ of these high-redshift MS galaxies also allow us to compare their SFE with that of local starbursts that have similarly high $\Sigma_{\rm gas}$ (see local starbursts in Fig. 11 from K98 and de los Reyes & Kennicutt 2019).

While the power law index of the KS relation for MS galaxies found in our analysis is consistent with that of Genzel et al. (2010), our normalization differs by about 0.2 dex (see right panel of Fig. 11), with our findings predicting shorter depletion times for MS galaxies than theirs. This difference can most likely be explained by the same limitation as that affecting Liu et al. (2019b) and Tacconi et al. (2020): literature studies on the KS relation are based on individually detected CO galaxy observations and are thus likely biased toward gas-rich galaxies at fixed $\Sigma_{\rm SFR}$. Naturally, one cannot rule out that part of this offset is due to some remaining offset between CO-based and dust-based molecular gas mass estimates (e.g., Tacconi et al. 2020).

5.2. Molecular gas depletion time

The molecular gas depletion time is defined as the time that a galaxy would need to deplete its molecular gas reservoir through star formation, provided a constant SFR and no gas accretion (i.e., $\tau_{\rm mol} = M_{\rm mol}/{\rm SFR}$). It can theoretically be written as $\tau_{\rm mol} = M_{\rm mol}/{\rm SFR} = t_{\rm ff}/\varepsilon_{\rm ff}$, where $t_{\rm ff}$ is the free-fall time and $\varepsilon_{\rm ff}$ is a dimensionless measure of the SFR efficiency, linking the mass of gas available for star formation and that effectively turning into stars (Krumholz et al. 2012). Theoretically, $\varepsilon_{\rm ff}$ is supposed to be roughly constant (≈0.01), rendering any variations in depletion time mostly due to variations in $t_{\rm ff}$ (Krumholz et al. 2012).

Observationally, the molecular depletion time of SFGs was found to follow tight scaling relations with their ΔMS, redshifts,

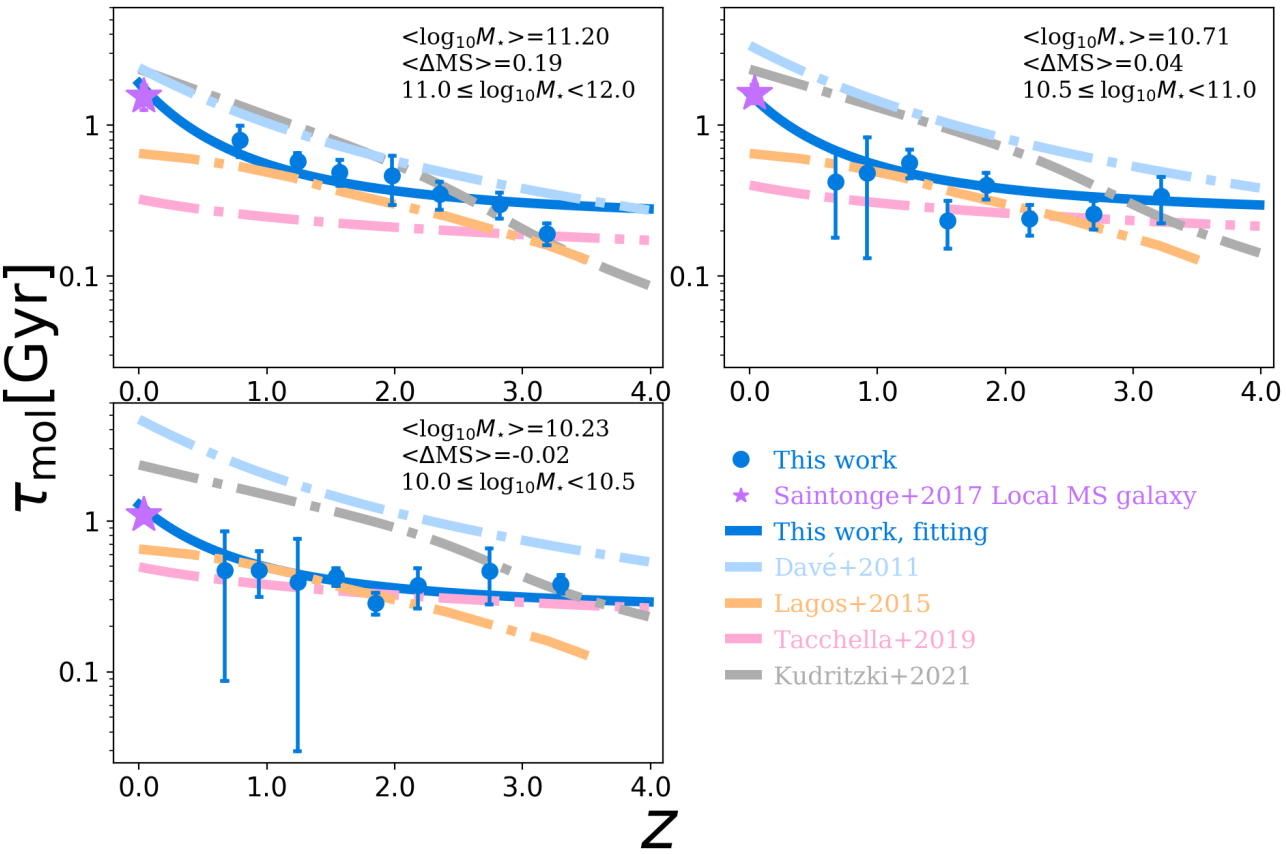

Fig. 12. Same as Fig. 7, but comparing our results to predictions from the cosmological hydrodynamic simulations of Davé et al. (2011, dash-dotted light-blue line), Lagos et al. (2015, dash-dotted orange line), Tacchella et al. (2016, dash-dotted pink line), and Kudritzki et al. (2021, dash-dotted gray line).

and stellar masses, providing thereby key information for models of galaxy evolution (e.g., Saintonge et al. 2016; Tacconi et al. 2018, 2020; Liu et al. 2019b; Hunt et al. 2020; Popesso et al. 2020). For example, the depletion time of galaxies situated well above the MS (i.e., ×4) was found to be significantly shorter (∼0.1 Gyr) than that of MS galaxies, suggesting a different star-formation mode for this galaxy population, likely triggered by the major merger of two gas-rich galaxies (e.g., Genzel et al. 2010; Hayward et al. 2011; Alaghband-Zadeh et al. 2012; Riechers et al. 2013). In addition, the depletion time of MS galaxies was found to slightly decrease with redshift up to $z \sim 3$ (e.g., Schinnerer et al. 2016; Tacchella et al. 2016; Tacconi et al. 2018, 2020; Padmanabhan & Loeb 2020). Our analysis, which confirms this finding, reveals that this apparent evolution is, however, not associated per se with a redshift evolution of the star-formation mode of MS galaxies but rather to the increase in their gas content with redshift, a relatively constant star-forming extent, and a seemingly universal KS relation with a power law index of ∼1.13: at a given stellar mass, the gas content of MS galaxies increases with redshift while their star-forming size remains roughly constant; their $\Sigma_{M_{\rm mol}}$ increases thus smoothly with redshift, shifting toward higher $\Sigma_{\rm SFR}$ to $\Sigma_{M_{\rm mol}}$ ratios. The molecular gas depletion time of MS galaxies was finally found to slightly decrease with stellar masses (e.g., Saintonge et al. 2016; Tacconi et al. 2018, 2020; Padmanabhan & Loeb 2020). On the contrary, our analysis finds that the evolution of their depletion time is mostly independent from their stellar masses. We note that this stellar mass-independent evolution of the depletion time suggests that the flattening of the MS at high stellar masses ($>10^{10.5}\,M_\odot$) and $z \lesssim 2.5$ (Leslie et al. 2020) is mostly due to their lower gas content rather than lower SFEs (see Sect. 5.4).

The scaling relations between the molecular gas depletion time of MS galaxies and their redshifts and stellar masses provide stringent constraints to hydrodynamic simulations performed in a cosmological context. In Fig. 12, we compare our findings to predictions from the simulations of Davé et al. (2011), Lagos et al. (2015), Tacchella et al. (2016), and Kudritzki et al. (2021). Overall, all these simulations predict a decrease in the molecular gas depletion time of MS galaxies with redshift and stellar mass, in qualitative agreement with our observations. However, the slope and overall normalization of these scaling relations vary by at least a factor of 3 between all these simulations and none can accurately reproduce the observed relations. As discussed in Kudritzki et al. (2021), predictions of the molecular depletion time are indeed strongly affected by the exact star formation, accretion and feedback models implemented in these simulations. The large disagreement between simulations and with the observations demonstrates that our understanding of these complex mechanisms across cosmic time is far from being complete and it also demonstrates the power of simple scaling relations to constrain models of galaxy evolution.

Finally, irrespective of the exact slope of these various scaling relations, all observations point toward relatively short molecular gas depletion times (∼0.5–1 Gyr) for MS galaxies of any stellar masses and redshifts. Without a constant replenishment of their gas reservoirs, the population of MS galaxies observed at, for example, $z \sim 2$, would thus have fully disappeared by $z = 1.5$ (see, e.g., Walter et al. 2020). These findings

strongly support the so-called gas regulator models (e.g., Erb 2008; Bouché et al. 2010; Davé et al. 2012; Lilly et al. 2013; Peng & Maiolino 2014; Rathaus & Sternberg 2016), in which galaxy growth is mostly driven by a continuous supply of fresh gas from the cosmic web (Dekel et al. 2009).

5.3. Compact star-forming extent

Our analysis as well as numerous recent studies have revealed that the star-forming half-light radius of MS galaxies is relatively compact (i.e., 1–3 kpc) and does not evolve significantly with redshift nor stellar mass (e.g., Barro et al. 2016; Rujopakarn et al. 2016; Elbaz et al. 2018; Lang et al. 2019; Jiménez-Andrade et al. 2019; Chang et al. 2020; Tadaki et al. 2020). In contrast, Fujimoto et al. (2017) found that the FIR size of MS galaxies evolves slightly with redshift. However, as stressed in their study, their individually detected ALMA sample is SFR-selected and could therefore be biased at high redshift toward compact SFGs (i.e., galaxies with high surface brightness). In any cases, the optical half-light radius of late-type galaxies of similar masses and at $z \lesssim 3$ is found to be about two times larger (∼3–8 kpc) than their star-forming extent (Sect. 4.3 van der Wel et al. 2014; Fujimoto et al. 2017; Elbaz et al. 2018; Lang et al. 2019; Jiménez-Andrade et al. 2019, 2021; Chen et al. 2020; Tadaki et al. 2020). This centrally enhanced star formation is usually interpreted in the literature as a sign that the cold gas accreted by MS galaxies falls preferentially onto their central region and triggers the formation of their bulge (e.g., Fisher 2006; Goldbaum et al. 2016; Tonini et al. 2016). These bulges would therefore grow from inside out and quench in the latest evolutionary stage of MS galaxies, leaving solely the outer disk with star-formation activities (e.g., Tacchella et al. 2015; Ellison et al. 2018; Rowlands et al. 2018; Colombo et al. 2020). While possible for massive SFGs, which are known to have massive central bulges (e.g., Leslie et al. 2020), such an interpretation seems less likely for less massive MS galaxies (i.e., $M_\star < 10^{11.0}\,M_\odot$) that are explored for the first time here. Using high-resolution ALMA and *Hubble* Space Telescope observations of 20 submillimeter-selected galaxies, Lang et al. (2019) argues instead that the discrepancy between FIR and optical sizes is mostly due to observational biases in which important radial color gradients yield very discrepant half-light and half-mass radii. This observational finding has recently been supported by The Next Generation Illustris 50 (TNG50) simulations coupled with state-of-the-art radiative transfer code to study the FIR, optical, and half-mass radius of thousands high-redshift 10^9–$10^{11}\,M_\odot$ MS galaxies (Popping et al. 2022). Indeed, in these simulations it is found that while the FIR half-light radius correlates with the radius containing half the star formation in galaxies, strong and un-corrected obscuration of the stellar light toward the galaxy center increases significantly the apparent extent of the disk sizes in the optical. Popping et al. (2022) conclude that the compact dust-continuum emission of MS galaxies with respect to the optical size is not necessarily evidence of the buildup of a dense central stellar component. Future high-resolution near-infrared observations performed by the *James Webb* Space Telescope will certainly play a key role in validating or invalidating these later findings.

5.4. The flattening of the MS relation at high masses

It is now relatively well established that the slope of the MS of SFGs flattens at high stellar masses, with this flattening becoming more and more prominent at $z \lesssim 2.5$ (e.g., Schreiber et al. 2015; Popesso et al. 2019; Leslie et al. 2020). Such a flattening of the MS is usually associated with the so-called mass-quenching model (e.g., Peng et al. 2010; Nelson et al. 2018; Wright et al. 2019). In this model, massive galaxies with their high gravitational potential hold a large accretion rate, growing their core rapidly in a few gigayears. As the core keeps growing, however, the ever larger gravitational potential could shock heat and/or an AGN could heat the new infalling gas, slowing down the gas accretion rate and reducing thereby the specific SFR of massive galaxies (e.g., Dekel & Birnboim 2006; Rodríguez Montero et al. 2019; Donnari et al. 2021). On the other hand, less massive galaxies could take more than a Hubble time to trigger such feedback and could thus efficiently accrete fresh cold gas even at low redshift. Our analysis reveals that the molecular gas fraction of our MS galaxies decreases with stellar mass at a rate mirroring than that of the flattening of the MS, yielding almost constant depletion time (equivalently SFE) with stellar mass. This implies that the slow downfall of the star formation in massive MS galaxies is principally due to an decrease in their molecular gas content rather than a decrease in their SFE, in agreement with recent observations of low-redshift galaxies (Colombo et al. 2020). Our findings support thus an interpretation in which the flattening of the MS at high masses is primarily controlled by the ability of galaxies to efficiently accrete or not accrete gas from the intergalactic medium.

Finally, we note that Leslie et al. (2020) found that the flattening of the MS must be linked with changes in the morphological composition of galaxies: bulge-dominated late-type galaxies, which dominate the SFG population at high stellar masses, show a flattening of the MS, while disk-dominated late-type galaxies have align on a SFR-M_* sequence with a slightly higher normalization and with a power law index in the log-log space close to unity. Although this result seems to favor a scenario in which the gas in bulge-dominated MS galaxies is stabilized against fragmentation (so-called morphological quenching; Martig et al. 2013), our findings suggest that the gas is instead simply not present to form stars in these galaxies: either it has been removed by feedback or cold gas is no longer able to be accreted efficiently onto bulge-dominated late-type galaxies.

6. Summary

We have investigated the evolution of the molecular gas content of MS galaxies from $z \sim 3.6$ to $z \sim 0.4$. We applied an innovative *uv*-based stacking analysis to a large set of ALMA observations toward a mass-complete sample of $>10^{10}\,M_\odot$ MS galaxies. This *uv*-based stacking analysis, performed on the RJ dust continuum emission of these galaxies, provides an accurate measurement of their mean molecular gas content (Hughes et al. 2017). With this unique data set and innovative approach, we constrain the redshift and stellar mass evolution of the mean molecular gas mass, molecular gas fraction, molecular gas depletion time, and molecular gas size of MS galaxies down to $10^{10}\,M_\odot$ and up to $z \sim 3.4$. Finally, we have also studied for the first time – using a mass-complete sample of MS galaxies – the KS relation at high redshift. Our main findings are:

1. The mean molecular gas mass of MS galaxies evolves significantly with redshift and depends on the stellar mass. At all stellar masses, the molecular gas fraction (i.e., $\mu_{\rm gas} = M_{\rm mol}/M_\star$) decreases by a factor of ∼24 from $z \sim 3.2$ to $z \sim 0$. In addition, at a given redshift, $\mu_{\rm gas}$ decreases with stellar mass at roughly the same rate as the decrease in the specific SFR (i.e., SSFR = SFR/$M_\star$) of MS galaxies.

2. Our mean molecular gas mass measurements are generally lower (∼10–60%) than literature estimates (e.g., Scoville et al. 2017; Liu et al. 2019b; Tacconi et al. 2020), especially at low stellar masses. Literature measurements, which have mostly relied on individually detected galaxies, were likely biased toward gas-rich galaxies.

3. The molecular gas depletion time (i.e., $\tau_{\rm mol} = M_{\rm mol}/{\rm SFR}$) of MS galaxies remains mostly constant at $z > 0.5$ with a value of 300–500 Myr, but increases by a factor of ∼3 by $z \sim 0$.

4. The mean FIR size of MS galaxies does not seem to evolve significantly with redshift or stellar mass, with a mean circularized half-light radius of ∼2.2 kpc. This result agrees qualitatively and quantitatively with the star-forming extent of MS galaxies measured in Jiménez-Andrade et al. (2019) using high-resolution radio observations of the COSMOS field (i.e., $R^{\rm circ}_{\rm radio} \sim 1.5^{+1.5}_{-0.8}$ kpc).

5. The redshift evolution of $\tau_{\rm mol}$ can be accurately predicted from the redshift evolution of the molecular gas surface density (i.e., $\Sigma_{M_{\rm mol}}$) of MS galaxies and a seemingly universal MS-only $\Sigma_{M_{\rm mol}} - \Sigma_{M_{\rm mol}}$ relation with a slope of ∼1.13, the KS relation.

Our findings provide key constraints for galaxy evolution models, as $>10^{10}\,M_\odot$ MS galaxies are known to have been responsible for the bulk of the star-forming activity of the Universe over the last 10 Gyr. To first order, it seems that the molecular gas content of MS galaxies regulates the evolution of their star-formation activity across cosmic time, while variation in their SFE (i.e., $1/\tau_{\rm mol}$) plays only a secondary role. The short depletion time of the molecular gas reservoir of MS galaxies (<1 Gyr) contrasts with the long duty cycle inherent to the existence of the MS itself. This suggests that the continuous replenishment of the molecular gas reservoir of MS galaxies plays a fundamental role in regulating star formation across cosmic time. Finally, despite large variations in the gas content and SFR of MS galaxies over the last 10 Gyr, their star formation seems to take place in their inner 2 kpc radius and to follow a seemingly universal MS-only $\Sigma_{M_{\rm mol}} - \Sigma_{\rm SFR}$ relation.

Acknowledgements. This research was carried out within the Collaborative Research Centre 956, sub-project A1, funded by the Deutsche Forschungsgemeinschaft (DFG) – project ID 184018867. E.S. and D.L. acknowledge funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 694343). ALMA is a partnership of ESO (representing its member states), NSF (USA) and NINS (Japan), together with NRC (Canada), MOST and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2015.1.00026.S, ADS/JAO.ALMA#2015.1.00055.S, ADS/JAO.ALMA#2015.1.00122.S, ADS/JAO.ALMA#2015.1.00137.S, ADS/JAO.ALMA#2015.1.00260.S, ADS/JAO.ALMA#2015.1.00379.S, ADS/JAO.ALMA#2015.1.00388.S, ADS/JAO.ALMA#2015.1.00540.S, ADS/JAO.ALMA#2015.1.00568.S, ADS/JAO.ALMA#2015.1.00664.S, ADS/JAO.ALMA#2015.1.00695.S, ADS/JAO.ALMA#2015.1.00704.S, ADS/JAO.ALMA#2015.1.00928.S, ADS/JAO.ALMA#2015.1.01074.S, ADS/JAO.ALMA#2015.1.01105.S, ADS/JAO.ALMA#2015.1.01111.S, ADS/JAO.ALMA#2015.1.01171.S, ADS/JAO.ALMA#2015.1.01212.S, ADS/JAO.ALMA#2015.1.01345.S, ADS/JAO.ALMA#2015.1.01495.S, ADS/JAO.ALMA#2016.1.00279.S, ADS/JAO.ALMA#2016.1.00463.S, ADS/JAO.ALMA#2016.1.00478.S, ADS/JAO.ALMA#2016.1.00624.S, ADS/JAO.ALMA#2016.1.00646.S, ADS/JAO.ALMA#2016.1.00778.S, ADS/JAO.ALMA#2016.1.00804.S, ADS/JAO.ALMA#2016.1.01012.S, ADS/JAO.ALMA#2016.1.01040.S, ADS/JAO.ALMA#2016.1.01184.S, ADS/JAO.ALMA#2016.1.01208.S, ADS/JAO.ALMA#2016.1.01240.S, ADS/JAO.ALMA#2016.1.01355.S, ADS/JAO.ALMA#2016.1.01426.S, ADS/JAO.ALMA#2016.1.01559.S, ADS/JAO.ALMA#2016.1.01604.S, ADS/JAO.ALMA#2017.1.00326.S, ADS/JAO.ALMA#2017.1.00413.S, ADS/JAO.ALMA#2017.1.00428.L, ADS/JAO.ALMA#2017.1.01020.S, ADS/JAO.ALMA#2017.1.01176.S, ADS/JAO.ALMA#2017.1.01276.S, ADS/JAO.ALMA#2017.1.01358.S, ADS/JAO.ALMA#2017.1.01451.S, ADS/JAO.ALMA#2017.1.01618.S, ADS/JAO.ALMA#2017.1.00046.S, ADS/JAO.ALMA#2017.1.01217.S, ADS/JAO.ALMA#2017.1.01259.S, ADS/JAO.ALMA#2018.1.00085.S, ADS/JAO.ALMA#2018.1.00216.S, ADS/JAO.ALMA#2018.1.00681.S, ADS/JAO.ALMA#2018.1.00992.S, ADS/JAO.ALMA#2018.1.01044.S, ADS/JAO.ALMA#2018.1.01359.S.

References

Alaghband-Zadeh, S., Chapman, S. C., Swinbank, A. M., et al. 2012, MNRAS, 424, 2232
Algera, H. S. B., Smail, I., Dudzevičiūtė, U., et al. 2020, ApJ, 903, 138
Barro, G., Kriek, M., Pérez-González, P. G., et al. 2016, ApJ, 827, L32
Berta, S., Lutz, D., Genzel, R., et al. 2016, A&A, 587, A73
Bertemes, C., Wuyts, S., Lutz, D., et al. 2018, MNRAS, 478, 1442
Béthermin, M., Daddi, E., Magdis, G., et al. 2012, ApJ, 757, L23
Béthermin, M., Daddi, E., Magdis, G., et al. 2015, A&A, 573, A113
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
Bouché, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001
Bourne, N., Dunne, L., Bendo, G. J., et al. 2013, MNRAS, 436, 479
Bourne, N., Dunlop, J. S., Merlin, E., et al. 2017, MNRAS, 467, 1360
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calistro Rivera, G., Hodge, J. A., Smail, I., et al. 2018, ApJ, 863, 56
Cappelluti, N., Brusa, M., Hasinger, G., et al. 2009, A&A, 497, 635
Carilli, C. L., & Walter, F. 2013, ARA&A, 51, 105
Casey, C. M., Narayanan, D., & Cooray, A. 2014, Phys. Rep., 541, 45
Cassata, P., Liu, D., Groves, B., et al. 2020, ApJ, 891, 83
Chabrier, G. 2003, PASP, 115, 763
Chang, Y.-Y., Le Floc'h, E., Juneau, S., et al. 2020, ApJ, 888, 44
Chary, R., & Elbaz, D. 2001, ApJ, 556, 562
Chen, C.-C., Hodge, J. A., Smail, I., et al. 2017, ApJ, 846, 108
Chen, C.-C., Harrison, C. M., Smail, I., et al. 2020, A&A, 635, A119
Colombo, D., Sanchez, S. F., Bolatto, A. D., et al. 2020, A&A, 644, A97
da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595
da Cunha, E., Walter, F., Smail, I. R., et al. 2015, ApJ, 806, 110
Daddi, E., Elbaz, D., Walter, F., et al. 2010, ApJ, 714, L118
Davé, R., Finlator, K., & Oppenheimer, B. D. 2011, MNRAS, 416, 1354
Davé, R., Finlator, K., & Oppenheimer, B. D. 2012, MNRAS, 421, 98
Davis, M., Guhathakurta, P., Konidaris, N. P., et al. 2007, ApJ, 660, L1
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Nature, 457, 451
de los Reyes, M. A. C., & Kennicutt, R. C. 2019, ApJ, 872, 16
Dessauges-Zavadsky, M., Ginolfi, M., Pozzi, F., et al. 2020, A&A, 643, A5
Dole, H., Lagache, G., Puget, J.-L., et al. 2006, A&A, 451, 417
Donnari, M., Pillepich, A., Joshi, G. D., et al. 2021, MNRAS, 500, 4004
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
Ellison, S. L., Sánchez, S. F., Ibarra-Medel, H., et al. 2018, MNRAS, 474, 2039
Ellison, S. L., Lin, L., Thorp, M. D., et al. 2021, MNRAS, 501, 4777
Erb, D. K. 2008, ApJ, 674, 151
Erb, D. K., Steidel, C. C., Shapley, A. E., et al. 2006, ApJ, 647, 128
Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982
Feltre, A., Maseda, M. V., Bacon, R., et al. 2020, A&A, 641, A118
Fisher, D. B. 2006, ApJ, 642, L17
Foreman-Mackey, D., Hogg, D. W., Lang, D., et al. 2013, PASP, 125, 306
Fudamoto, Y., Oesch, P. A., Faisst, A., et al. 2020, A&A, 643, A4
Fujimoto, S., Ouchi, M., Shibuya, T., et al. 2017, ApJ, 850, 83
Gabányi, K. É., Frey, S., & Perger, K. 2021, MNRAS, 506, 3641
Gao, Y., & Solomon, P. M. 2004, ApJ, 606, 271
Genzel, R., Tacconi, L. J., Gracia-Carpio, J., et al. 2010, MNRAS, 407, 2091
Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20
Goldbaum, N. J., Krumholz, M. R., & Forbes, J. C. 2016, ApJ, 827, 28
Gowardhan, A., Riechers, D., Pavesi, R., et al. 2019, ApJ, 875, 6
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
Hayward, C. C., Kereš, D., Jonsson, P., et al. 2011, ApJ, 743, 159
Hughes, T. M., Ibar, E., Villanueva, V., et al. 2017, MNRAS, 468, L103
Hunt, L. K., De Looze, I., Boquien, M., et al. 2019, A&A, 621, A51
Hunt, L. K., Tortora, C., Ginolfi, M., et al. 2020, A&A, 643, A180
Jiménez-Andrade, E. F., Magnelli, B., Karim, A., et al. 2019, A&A, 625, A114
Jiménez-Andrade, E. F., Murphy, E. J., Heywood, I., et al. 2021, ApJ, 910, 106
Jin, S., Daddi, E., Liu, D., et al. 2018, ApJ, 864, 56
Johnston, R., Vaccari, M., Jarvis, M., et al. 2015, MNRAS, 453, 2540
Johnston, R. S., Stil, J. M., & Keller, B. W. 2021, ApJ, 909, 73
Kaasinen, M., Scoville, N., Walter, F., et al. 2019, ApJ, 880, 15
Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61

Kashino, D., Silverman, J. D., Rodighiero, G., et al. 2013, ApJ, 777, L8
Kennicutt, R. C. 1998a, ApJ, 498, 541
Kennicutt, R. C. 1998b, ARA&A, 36, 189
Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196
Krumholz, M. R., & McKee, C. F. 2005, ApJ, 630, 250
Krumholz, M. R., Dekel, A., & McKee, C. F. 2012, ApJ, 745, 69
Kudritzki, R.-P., Teklu, A. F., Schulze, F., et al. 2021, ApJ, 910, 87
Lagos, C. del P., Crain, R. A., Schaye, J., et al. 2015, MNRAS, 452, 3815
Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, ApJS, 224, 24
Lang, P., Schinnerer, E., Smail, I., et al. 2019, ApJ, 879, 54
Le Floc'h, E., Aussel, H., Ilbert, O., et al. 2009, ApJ, 703, 222
Leroy, A. K., Bolatto, A., Gordon, K., et al. 2011, ApJ, 737, 12
Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, AJ, 146, 19
Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, ApJ, 899, 58
Lilly, S. J., Carollo, C. M., Pipino, A., et al. 2013, ApJ, 772, 119
Lindroos, L., Knudsen, K. K., Vlemmings, W., et al. 2015, MNRAS, 446, 3502
Lindroos, L., Knudsen, K. K., Fan, L., et al. 2016, MNRAS, 462, 1192
Liu, D., Daddi, E., Dickinson, M., et al. 2018, ApJ, 853, 172
Liu, D., Schinnerer, E., Groves, B., et al. 2019a, ApJ, 887, 235
Liu, D., Lang, P., Magnelli, B., et al. 2019b, ApJS, 244, 40
Luo, W., Yang, X., & Zhang, Y. 2014, ApJ, 789, L16
Lutz, D., Poglitsch, A., Altieri, B., et al. 2011, A&A, 532, A90
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Magdis, G. E., Rigopoulou, D., Huang, J.-S., et al. 2010, MNRAS, 401, 1521
Magdis, G. E., Daddi, E., Béthermin, M., et al. 2012, ApJ, 760, 6
Magnelli, B., Saintonge, A., Lutz, D., et al. 2012a, A&A, 548, A22
Magnelli, B., Lutz, D., Santini, P., et al. 2012b, A&A, 539, A155
Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, A&A, 561, A86
Magnelli, B., Ivison, R. J., Lutz, D., et al. 2015, A&A, 573, A45
Magnelli, B., Boogaard, L., Decarli, R., et al. 2020, ApJ, 892, 66
Marchesi, S., Civano, F., Elvis, M., et al. 2016, ApJ, 817, 34
Martig, M., Crocker, A. F., Bournaud, F., et al. 2013, MNRAS, 432, 1914
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, in Astronomical Data Analysis Software and Systems XVI, eds. R. A. Shaw, F. Hill, & D. J. Bell, ASP Conf. Ser., 376, 127
Miettinen, O., Delvecchio, I., Smolčić, V., et al. 2017a, A&A, 602, L9
Miettinen, O., Delvecchio, I., Smolčić, V., et al. 2017b, A&A, 606, A17
Millard, J. S., Eales, S. A., Smith, M. W. L., et al. 2020, MNRAS, 494, 293
Miura, R. E., Kohno, K., Tosaki, T., et al. 2014, ApJ, 788, 167
Mohan, N., & Rafferty, D. 2015, PyBDSF: Python Blob Detection and Source Finder, Astrophysics Source Code Library [record ascl:1502.007]
Momose, R., Koda, J., Kennicutt, R. C., et al. 2013, ApJ, 772, L13
Nelson, D., Kauffmann, G., Pillepich, A., et al. 2018, MNRAS, 477, 450
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
Oliver, S. J., Bock, J., Altieri, B., et al. 2012, MNRAS, 424, 1614
Padmanabhan, H., & Loeb, A. 2020, MNRAS, 496, 1124
Pannella, M., Carilli, C. L., Daddi, E., et al. 2009, ApJ, 698, L116
Pearson, W. J., Wang, L., Hurley, P. D., et al. 2018, A&A, 615, A146
Peng, Y. J., & Maiolino, R. 2014, MNRAS, 443, 3643
Peng, Y. J., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193
Pessa, I., Schinnerer, E., Belfiore, F., et al. 2021, A&A, 650, A134
Planck Collaboration XIII 2016, A&A, 594, A13
Popesso, P., Morselli, L., Concas, A., et al. 2019, MNRAS, 490, 5285
Popesso, P., Concas, A., Morselli, L., et al. 2020, MNRAS, 496, 2531
Popping, G., Pillepich, A., Calistro Rivera, G., et al. 2022, MNRAS, 510, 3321
Puglisi, A., Daddi, E., Liu, D., et al. 2019, ApJ, 877, L23
Qin, J., Zheng, X. Z., Wuyts, S., et al. 2019, MNRAS, 485, 5733
Rathaus, B., & Sternberg, A. 2016, MNRAS, 458, 3168
Rémy-Ruyer, A., Madden, S. C., Galliano, F., et al. 2014, A&A, 563, A31
Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, Nature, 496, 329
Rodighiero, G., Daddi, E., Baronchelli, I., et al. 2011, ApJ, 739, L40
Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, MNRAS, 443, 19
Rodríguez Montero, F., Davé, R., Wild, V., et al. 2019, MNRAS, 490, 2139
Rosolowsky, E., & Blitz, L. 2005, ApJ, 623, 826
Rowlands, K., Heckman, T., Wild, V., et al. 2018, MNRAS, 480, 2544
Rujopakarn, W., Dunlop, J. S., Rieke, G. H., et al. 2016, ApJ, 833, 12
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
Saintonge, A., Wilson, C. D., Xiao, T., et al. 2018, MNRAS, 481, 3497
Sánchez, S. F., Barrera-Ballesteros, J. K., Colombo, D., et al. 2021, MNRAS, 503, 1615
Sanders, D. B., Salvato, M., Aussel, H., et al. 2007, ApJS, 172, 86
Santini, P., Maiolino, R., Magnelli, B., et al. 2014, A&A, 562, A30
Sargent, M. T., Béthermin, M., Daddi, E., et al. 2012, ApJ, 747, L31
Sargent, M. T., Daddi, E., Béthermin, M., et al. 2014, ApJ, 793, 19
Schinnerer, E., Sargent, M. T., Bondi, M., et al. 2010, ApJS, 188, 384
Schinnerer, E., Groves, B., Sargent, M. T., et al. 2016, ApJ, 833, 112
Schmidt, M. 1959, ApJ, 129, 243
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Schreiber, C., Elbaz, D., Pannella, M., et al. 2018, A&A, 609, A30
Schuster, K. F., Kramer, C., Hitschfeld, M., et al. 2007, A&A, 461, 143
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
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
Shetty, R., Kelly, B. C., Rahman, N., et al. 2014, MNRAS, 437, L61
Smolčić, V., Novak, M., Bondi, M., et al. 2017, A&A, 602, A1
Sobral, D., Best, P. N., Smail, I., et al. 2014, MNRAS, 437, 3516
Speagle, J. S., Steinhardt, C. L., Capak, P. L., et al. 2014, ApJS, 214, 15
Suess, K. A., Kriek, M., Price, S. H., et al. 2019, ApJ, 877, 103
Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, ApJS, 155, 271
Tacchella, S., Carollo, C. M., Renzini, A., et al. 2015, Science, 348, 314
Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 457, 2790
Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, Nature, 463, 781
Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179
Tacconi, L. J., Genzel, R., & Sternberg, A. 2020, ARA&A, 58, 157
Tadaki, K., Belli, S., Burkert, A., et al. 2020, ApJ, 901, 74
Tan, Q., Daddi, E., Magdis, G., et al. 2014, A&A, 569, A98
Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9
Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2016, ApJ, 817, 118
Tonini, C., Mutch, S. J., Croton, D. J., et al. 2016, MNRAS, 459, 4109
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Walter, F., Carilli, C., Neeleman, M., et al. 2020, ApJ, 902, 111
Walther, M., Oñorbe, J., Hennawi, J. F., et al. 2019, ApJ, 872, 13
Wang, T.-M., & Hwang, C.-Y. 2020, A&A, 641, A24
Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104
Wiklind, T., Ferguson, H. C., Guo, Y., et al. 2019, ApJ, 878, 83
Wilson, C. D., Elmegreen, B. G., Bemis, A., et al. 2019, ApJ, 882, 5
Wong, T., & Blitz, L. 2002, ApJ, 569, 157
Wright, R. J., Lagos, C. del P., Davies, L. J. M., et al. 2019, MNRAS, 487, 3740
Wuyts, S., Förster Schreiber, N. M., Lutz, D., et al. 2011, ApJ, 738, 106
Wyse, R. F. G. 1986, ApJ, 311, L41
Zahid, H. J., Dima, G. I., Kewley, L. J., et al. 2012, ApJ, 757, 54
Zamojski, M. A., Schiminovich, D., Rich, R. M., et al. 2007, ApJS, 172, 468
Zheng, X. Z., Bell, E. F., Rix, H.-W., et al. 2006, ApJ, 640, 784

Appendix A: Reliability of the stacked flux density and size measurements

To test the reliability of the flux densities and sizes measured by stacking in the *uv* and image domain, we used simulations tailored to reproduce the peculiar properties of the A^{3}COSMOS archive (i.e., with a heterogeneous frequency, depth, and spatial resolution). The logic of our approach is to (i) simulate 100 realistic ALMA observations of a galaxy population with a given intrinsic flux density and size properties, (ii) stack these 100 mock observations, and (iii) finally compare the measured stacked flux density and size to the intrinsic ones. For a given set of flux density and size properties, we judge the reliability of our stacking analysis by computing $(S_{\rm in} - S_{\rm out})/S_{\rm in}$ and $(R^{\rm circ}_{\rm in} - R^{\rm circ}_{\rm out})/R^{\rm circ}_{\rm in}$ (i.e., the error on the retrieved stacked flux density and size, respectively).

Our mock ALMA observations were generated using the CASA task `simobserve`. To realistically reproduce the heterogeneity of the A^{3}COSMOS archive, the observing properties of each simulation (i.e., frequency resolution, integration time, and antenna positions) were randomly picked from one of the A^{3}COSMOS data set. Then, the stacked position of the mock galaxies were randomly selected within a radius of 3″from the simulated phase center. Finally, to account for possible mismatches between the stacked position and the actual millimeter position of the stacked sources, we randomly placed these mock galaxies around their stacked position, following a 2D Gaussian distribution with a dispersion of 0.″2 (e.g., Elbaz et al. 2018). We note that here we do not test the effect of the observed-frame flux density to rest-frame luminosities conversion discussed in Sect. 3.1. Indeed, from a technical point of view this conversion is strictly equivalent to a simple multiplication and thus a noise increase or decrease in the initial to-be-stacked data set. Our simulation already test the effect of stacking data set with different noise, redistributing this noise distribution is thus unnecessary.

Results of these Monte Carlo simulations for three realistic angular sizes (i.e., point source, $FWHM$ = 0.″5 and $FWHM$ = 1.″0) and twelves flux densities combinations, are shown in Fig. A.1 and A.2. These results demonstrate the reliability of our stacked flux density and size measurements. Indeed, all size-flux density combinations that yield a detection have $(S_{\rm in} - S_{\rm out})/S_{\rm in}$ and $(R^{\rm circ}_{\rm in} - R^{\rm circ}_{\rm out})/R^{\rm circ}_{\rm in}$ values consistent with 0 within the measured uncertainties. We only notice a slight systematic overestimation of our stacked flux densities by 6% and 4% when inferred from the *uv* and image domain, respectively. Such an offset is virtually insignificant with respect to the measured uncertainties. Lastly, we note that galaxies with larger intrinsic sizes (i.e., 1.″0) do not yield a detection down to our lowest flux density combination. This is not inherent to our stacking analysis but simply to the spreading of the flux densities of these galaxies over several synthesized beams.

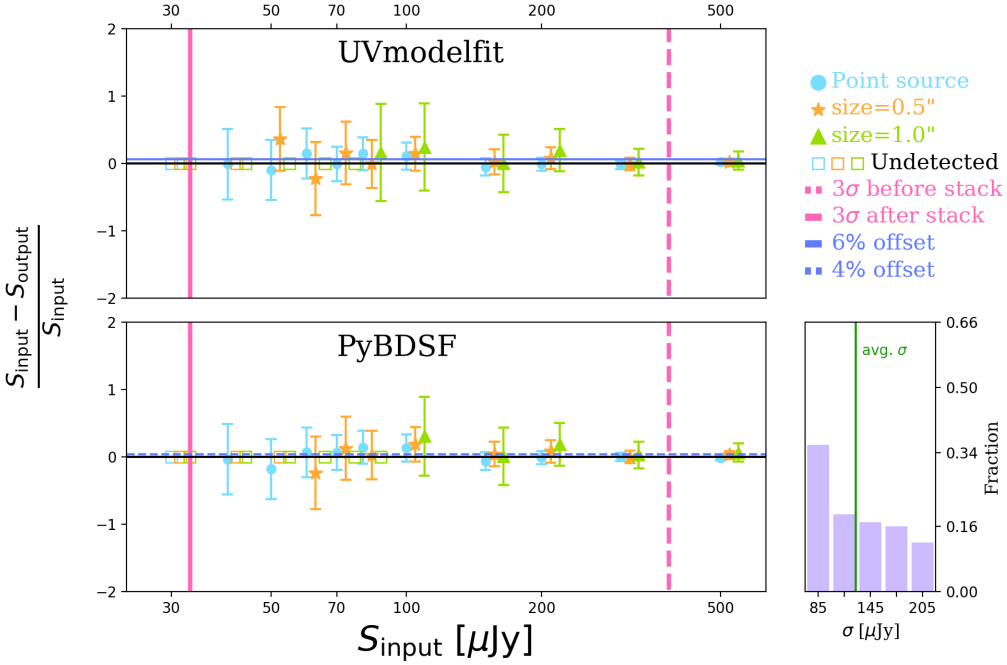

Fig. A.1. Uncertainty on the *uv*-domain (upper panel) and image-domain (lower panel) stacked flux density measurements of 100 simulated galaxies as a function of their intrinsic flux densities and sizes. The vertical pink dashed line shows the average "point source" 3σ detection threshold of the measurement sets to be stacked, while the lower right subpanel displays their 1σ distribution. The vertical solid pink line presents the point source 3σ detection threshold in the stacked measurement set. Our stacking analysis allows accurate mean flux density measurements (i.e., $(S_{\rm in} - S_{\rm out})/S_{\rm in} \sim 0$) for galaxy populations that are otherwise individually undetected, i.e., with intrinsic flux densities lower than the vertical pink dashed line. Stacked flux densities are in average underestimated by 6% and 4% in the *uv* domain and image domain, respectively, as illustrated by the dark blue solid and dashed horizontal lines, respectively.

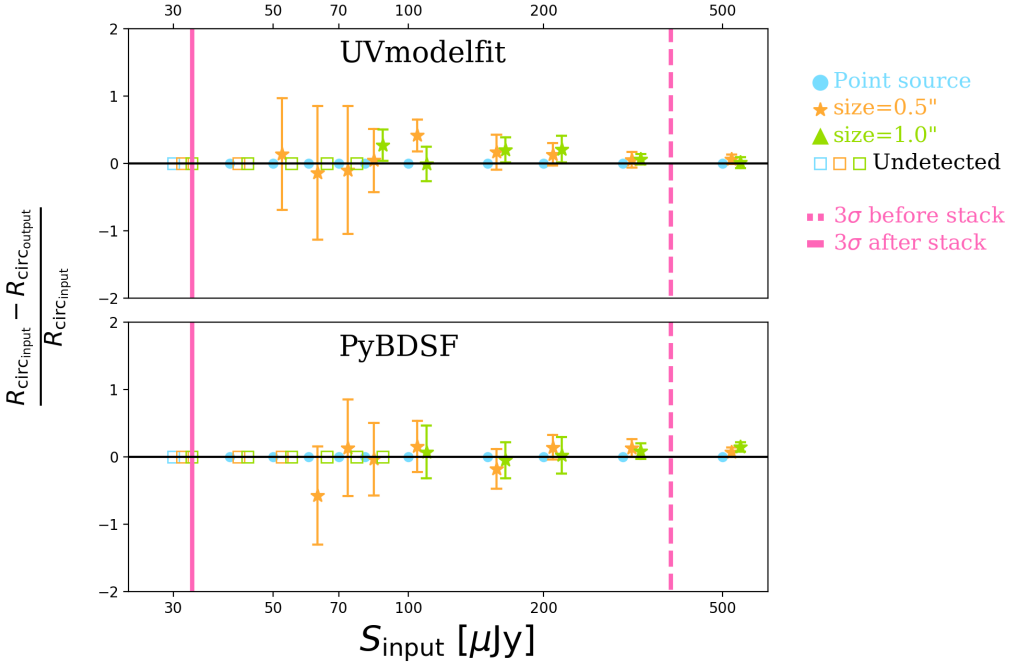

Fig. A.2. Uncertainty on the *uv*-domain (upper panel) and image-domain (lower panel) stacked size measurements of 100 simulated galaxies as a function of their intrinsic flux densities and sizes. Lines and symbols are the same as in Fig. A.1. Our stacking analysis allows accurate mean size measurements (i.e., $(R_{\mathrm{in}}^{\mathrm{circ}} - R_{\mathrm{out}}^{\mathrm{circ}})/R_{\mathrm{in}}^{\mathrm{circ}} \sim 0$) for galaxy populations that are otherwise individually undetected, i.e., with intrinsic flux densities lower than the vertical dashed pink line.

Appendix B: Different gas mass calibrations

The results of our study could naturally be influenced by the particular assumptions made here to convert stacked rest-frame 850 μm luminosities into molecular gas masses (see, e.g., discussion in Liu et al. 2019b). To evaluate the impact of these assumption on our results, we compare molecular gas masses obtained here using the metallicity-independent L_{850}-to-$M_{\rm mol}$ relation of H17 with those obtained using others relations commonly applied in the literature (Fig. B.1).

H17 inferred two L_{850}-to-$M_{\rm mol}$ relations: the one used in the core of our paper was calibrated using a CO-to-$M_{\rm mol}$ conversion factor (i.e., $\alpha_{\rm CO}$) that assumes that the density and properties in the gas reservoirs of high-redshift galaxies are similar to Milky Way GMCs (i.e., $\alpha_{\rm CO}$=6.5 $M_\odot$ (K km s^{-1} pc^2)$^{-1}$); and another relation calibrated instead using $\alpha_{\rm CO}$=4.35 (H17$_{\alpha_{\rm CO}=4.35}$), which is the standard value for the Milky Way. These relations have the same slope and only differ in terms of normalization. Molecular gas masses obtained with H17$_{\alpha_{\rm CO}=4.35}$ are thus shifted to lower values by 0.17 dex (Fig. B.1). This yields a quantitatively different evolution of the molecular gas mass fraction and depletion time with redshift than those inferred with H17 (Fig. B.2). However, the main conclusions of our analysis remains unchanged: the molecular gas fraction of MS galaxies still increases significantly with redshifts and decreases with stellar masses; and their depletion time still decrease more moderately with redshifts and remains mostly independent from their stellar mass.

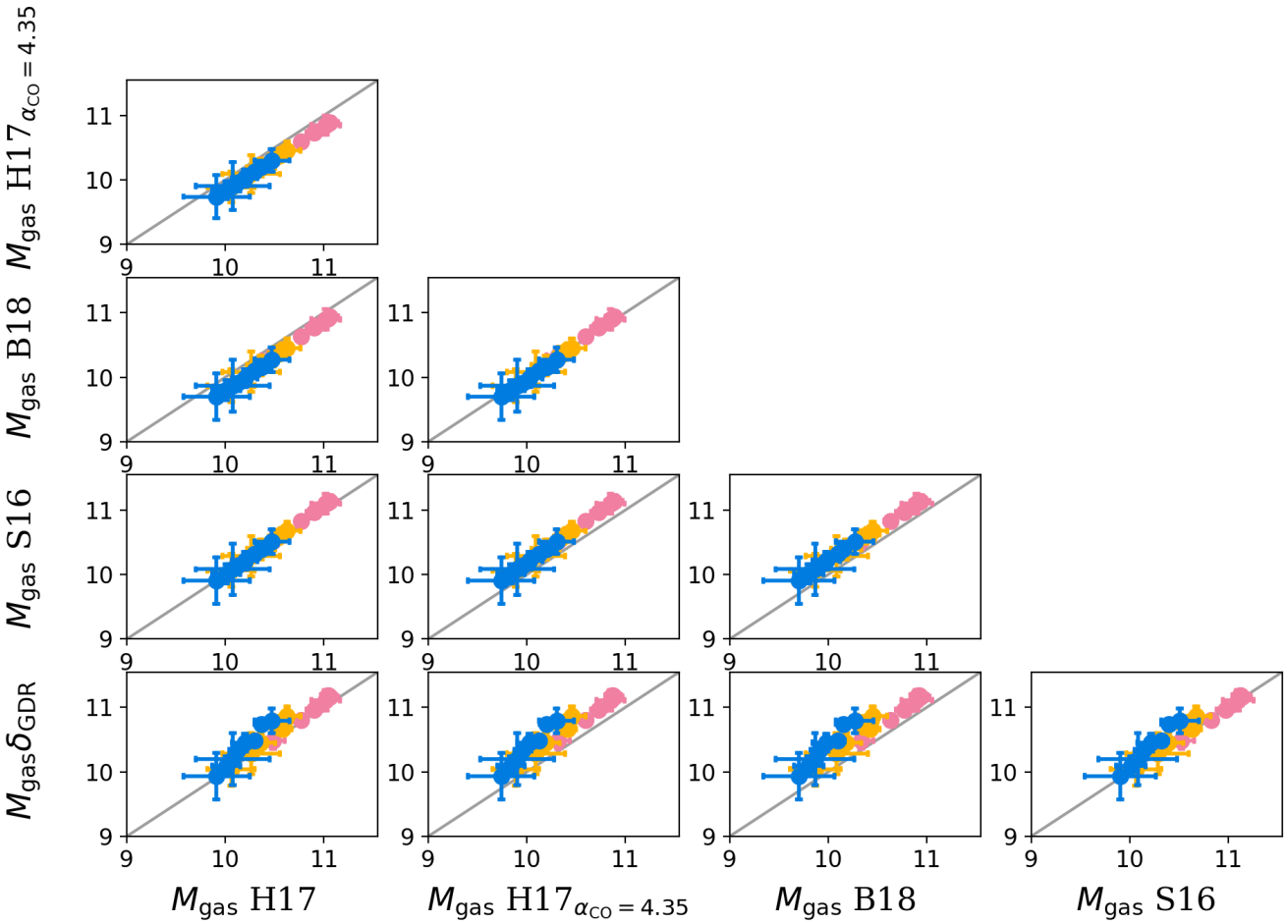

Fig. B.1. Comparisons between four molecular gas mass calibrations, i.e., from Hughes et al. (2017) assuming $\alpha_{\rm CO}$ = 6.5 (H17; used throughout our paper), Hughes et al. (2017) assuming $\alpha_{\rm CO}$ = 4.35 (H17$_{\alpha_{\rm CO}=4.35}$), Scoville et al. (2016, S16), Bertemes et al. (2018, B18), and Leroy et al. (2011, $\delta_{\rm GDR}$). In each panel, the solid gray line is the one-to-one relation. Symbol colors have the same meaning as in Fig. 11.

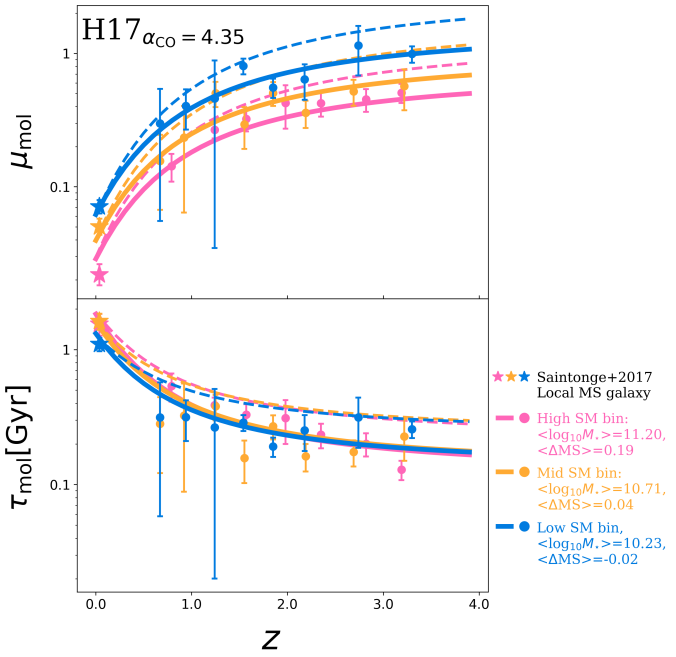

Fig. B.2. Redshift evolution of the mean molecular gas fraction and molecular gas depletion time of MS galaxies, with the M_{gas} from the H17$_{\alpha_{CO}=4.35}$ gas mass calibration. Dots show the mean molecular gas fraction and molecular gas depletion time from our work. Stars represent the local reference taken from Saintonge et al. (2017). Solid lines display the analytical evolution of the molecular gas fraction and molecular gas depletion time inferred with the H17$_{\alpha_{CO}=4.35}$ gas mass calibration. Dashed lines show the analytical evolution of the molecular gas fraction and molecular gas depletion time inferred with the H17 gas mass calibration (used throughout the paper). Symbols and lines are color-coded by stellar mass.

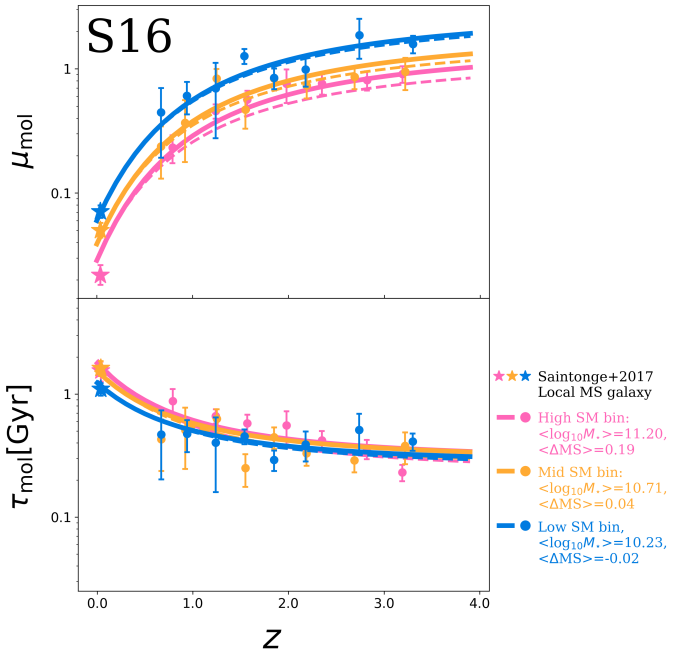

Fig. B.3. Same as Fig. B.2 but for the gas mass calibration from S16.

Besides the H17 gas mass calibration, S16 also provides a metallicity-independent L_{850}-to-M_{mol} relation calibrated on a sample of 70 SFGs with both dust RJ and CO measurements. Then, assuming a constant mass-weighted dust temperature of 25 K, a dust emissivity, β, of 1.8, and α_{CO}=6.5 $M_\odot$ (K km s^{-1} pc^2)$^{-1}$, they calibrated a light-to-mass relation

$$\log(M_{mol}) = \log(L_{850}) - 19.83, \quad \text{(B.1)}$$

where M_{mol} and L_{850} are in units of $M_\odot$ and erg s^{-1} Hz^{-1}, respectively. This relation yields molecular gas masses in very good agreement with those inferred here (Fig. B.1), with only a systematic offset of +0.04 dex for S16. The redshift evolution of the molecular gas mass fraction and depletion time inferred using S16 are shown in Fig. B.3. Those agree qualitatively and quantitatively with the original findings of our study.

One can also measure the molecular gas mass of SFGs by applying first a standard L_{850}-to-M_{dust} relation and then a metallicity-dependent M_{dust}-to-M_{mol} relation. To begin with, we thus convert our stacked rest-frame 850 μm luminosities into M_{dust} following Magnelli et al. (2020), who assumed a constant mass-weighted dust temperature of 25 K, a dust emissivity of 1.8, and a photon cross section to dust mass ratio at rest-frame 850 μm, κ_{850}, of 0.0431 m^2 kg^{-1},

$$\log(M_{dust}) = \log(L_{850}) - 21.86, \quad \text{(B.2)}$$

where M_{dust} is in unit of $M_\odot$. This dust mass can then be converted into molecular gas mass using the metallicity-dependent M_{dust}-to-M_{mol} relation of, for example, B18, which was calibrated using 78 local SFGs with known gas-phase metallicity,

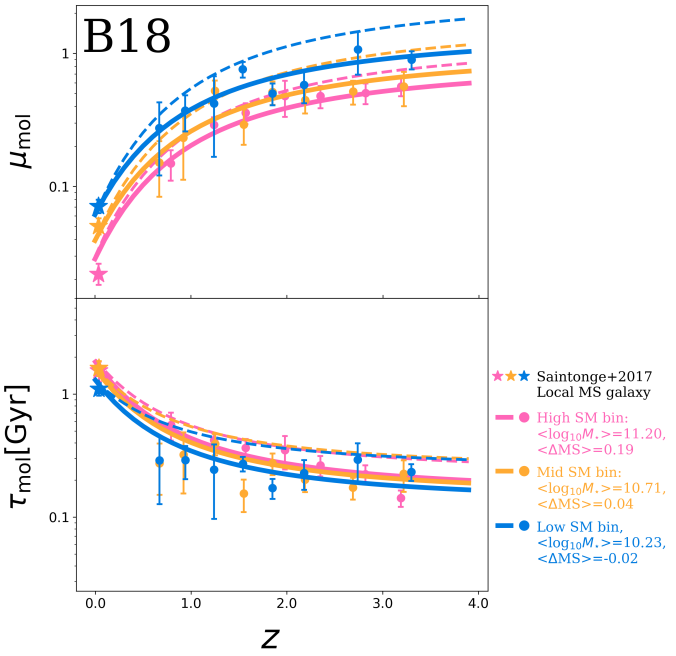

Fig. B.4. Same as Fig. B.2 but for the gas mass calibration from B18.

$$\log(M_{mol}) = \log(M_{dust}) + 1.83 + 0.12 \times ((12 + \log(O/H)) - 8.67), \quad \text{(B.3)}$$

where the gas-phase metallicity, 12+log(O/H), can be inferred using the redshift- and stellar mass-dependent relation given in Liu et al. (2019b),

$$12 + \log(O/H) = \begin{cases} a & \text{if } \log(M_\star/M_\odot) \geq b(z), \\ a - 0.087 \times (\log(M_\star/M_\odot) - b(z))^2, & \text{else,} \end{cases} \quad \text{(B.4)}$$

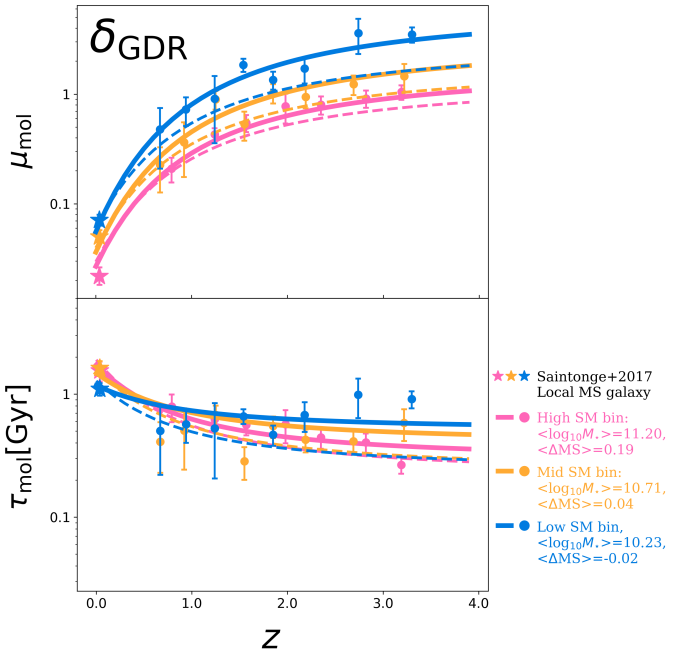

Fig. B.5. Same as Fig. B.2 but for the gas mass calibration from Leroy et al. (2011, $\delta_{\rm GDR}$).

Table B.1. Best-fit coefficients for the molecular gas fraction (Eq. 8) and molecular gas depletion time (Eq. 9) functions.

$\log_{10}\mu_{\rm mol}$ with a = 0.4195 and ak = 0.1195 from Liu et al. (2019b)	b	c	ck	d
H17	$-0.468^{+0.070}_{-0.070}$	$-0.122^{+0.008}_{-0.008}$	$0.572^{+0.059}_{-0.060}$	$0.002^{+0.011}_{-0.011}$
$H17_{\alpha_{\rm CO}=4.35}$	$-0.463^{+0.070}_{-0.071}$	$-0.102^{+0.008}_{-0.008}$	$0.311^{+0.061}_{-0.058}$	$0.000^{+0.010}_{-0.010}$
B18	$-0.353^{+0.069}_{-0.069}$	$-0.099^{+0.008}_{-0.008}$	$0.266^{+0.057}_{-0.057}$	$-0.008^{+0.010}_{-0.010}$
S16	$-0.396^{+0.070}_{-0.070}$	$-0.123^{+0.008}_{-0.008}$	$0.584^{+0.059}_{-0.058}$	$-0.004^{+0.011}_{-0.011}$
L11	$-0.679^{+0.068}_{-0.067}$	$-0.152^{+0.008}_{-0.008}$	$0.948^{+0.057}_{-0.057}$	$0.017^{+0.010}_{-0.010}$

$\log_{10}\tau_{\rm mol}$ [Gy^{-1}] with a = −0.5724 and ak = 0.1120 from Liu et al. (2019b)	b	c	ck	d
H17	$0.055^{+0.069}_{-0.071}$	$0.049^{+0.008}_{-0.008}$	$-0.643^{+0.056}_{-0.057}$	$0.016^{+0.010}_{-0.010}$
$H17_{\alpha_{\rm CO}=4.35}$	$0.054^{+0.071}_{-0.069}$	$0.069^{+0.008}_{-0.008}$	$-0.899^{+0.055}_{-0.057}$	$0.014^{+0.010}_{-0.011}$
B18	$0.168^{+0.069}_{-0.067}$	$0.072^{+0.008}_{-0.008}$	$-0.947^{+0.057}_{-0.057}$	$0.006^{+0.010}_{-0.010}$
S16	$-0.125^{+0.070}_{-0.067}$	$0.047^{+0.008}_{-0.008}$	$-0.628^{+0.058}_{-0.058}$	$0.010^{+0.010}_{-0.010}$
L11	$-0.157^{+0.068}_{-0.070}$	$0.019^{+0.008}_{-0.008}$	$-0.265^{+0.059}_{-0.057}$	$0.031^{+0.010}_{-0.011}$

where a=8.74 and $b(z)$=10.4+4.46×log(1+z)-1.78×(log(1+z))^2. Using this molecular gas mass calibration (i.e., B18) yields estimates that are ∼0.17 dex lower than those from H17 (Fig. B.1). In addition to this global offset, this metallicity-dependent approach yields quantitatively different evolution of the molecular gas mass fraction and depletion time with redshift than those inferred with H17 (Fig. B.4). However, the main conclusions of our analysis remains unchanged: the molecular gas fraction of MS galaxies still increases significantly with redshifts and decreases with stellar masses; and their depletion time still decreases more moderately with redshifts and remains mostly independent from their stellar mass.

Finally, we converted the dust masses from Eq. B.2 into molecular gas masses using instead the standard metallicity-dependent $M_{\rm dust}$-to-$M_{\rm mol}$ relation of Leroy et al. (2011, $\delta_{\rm GDR}$), which was calibrated using high-resolution observations of five local group galaxies,

$$\log(M_{\rm mol}) = \log(M_{\rm dust}) + 9.4 - 0.85 \times (12 + \log({\rm O/H})). \quad \text{(B.5)}$$

We note that while this relation formally account for both molecular and atomic gas masses, we implicitly assume here that the gas in high-redshift ($z > 0.5$) SFGs is dominated by their molecular phase Tacconi et al. (see discussion in 2018); Liu et al. (see discussion in 2019b). Overall, the molecular gas masses inferred using the $\delta_{\rm GDR}$ method are about ∼0.13 dex higher than those obtained from H17 (Fig. B.1). As for B18, this gas mass calibration also yields slightly different, yet qualitatively consistent, molecular gas mass fraction and depletion time redshift evolution (Fig. B.5) than those inferred with H17.

To summarize, the gas mass calibrations are: $M_{\rm mol}^{{\rm H17}_{\alpha_{\rm CO}=4.35}} \approx M_{\rm mol}^{\rm B18} < M_{\rm mol}^{\rm H17} \approx M_{\rm mol}^{\rm S16} < M_{\rm mol}^{\delta_{\rm GDR}}$. The redshift evolution of the molecular gas fraction and depletion time obtained from S16 agree qualitatively and quantitatively with those from H17, while those measured using $H17_{\alpha_{\rm CO}=4.35}$, B18 and $\delta_{\rm GDR}$ agree only qualitatively with those from H17 (see Table B.1). However, because the main conclusion of our paper are not qualitatively affected by the particular choice of a given relation and because the H17 approach yield measurements that are bracket by others, we decided to use H17 in our paper.

The choice of using the L_{850}-to-$M_{\rm mol}$ relation from H17 could seem at odds since this relation was calibrated using $\alpha_{\rm CO}$=6.5, while our local reference (i.e., Saintonge et al. 2017) converted their CO measurements into molecular gas masses using $\alpha_{\rm CO}$ ∼4 (at the high stellar masses of our study). Despite this apparent inconsistency, the redshift evolution of the molecular gas content of massive SFGs inferred by combining this local reference with our high-redshift H17 measurements is in much better agreement with Tacconi et al. (2020) than when combining this local reference with our high-redshift $H17_{\alpha_{\rm CO}=4.35}$ measurements (see Fig. 4 and B.1). Because at high stellar masses results from Tacconi et al. (2020) can be considered as the reference (as they are based in a fairly complete sample of massive SFGs and a thorough cross-calibration of the CO- and dust-based methods), we decided to use in the core of our paper the L_{850}-to-$M_{\rm mol}$ relation from H17. The agreement between these high-redshift H17 measurements and those from Tacconi et al. (2020) is explained by the fact that using $\alpha_{\rm CO}$=6.5 instead of 4.3 to calibrate the local L_{850}-to-$M_{\rm mol}$ relation corrects indirectly (and to first order) for the fact that at a given stellar mass, high-redshift galaxies have lower metallicities than local galaxies, and thus have a higher gas-to-dust ratio (e.g., Leroy et al. 2011) and consequently should have lower L_{850}-to-$M_{\rm mol}$ ratio. This seems also confirmed by the good agreement at high stellar masses between our H17 measurements and those from the metallicity-dependent $\delta_{\rm GDR}$ method.