The hidden side of cosmic star formation at $z > 3$

M.-Y. Xiao, D. Elbaz, C. Gómez-Guijarro, L. Leroy, L.-J. Bing, E. Daddi, B. Magnelli, M. Franco, L. Zhou, M. Dickinson, et al.

A&A 672, A18 (2023)
https://doi.org/10.1051/0004-6361/202245100

Astronomy & Astrophysics

The hidden side of cosmic star formation at $z > 3$

Bridging optically dark and Lyman-break galaxies with GOODS-ALMA

M.-Y. Xiao[1,2,3], D. Elbaz[2], C. Gómez-Guijarro[2], L. Leroy[2], L.-J. Bing[4], E. Daddi[2], B. Magnelli[2], M. Franco[5], L. Zhou[1,3], M. Dickinson[6], T. Wang[1,3], W. Rujopakarn[7,8,9], G. E. Magdis[10,11,12], E. Treister[13], H. Inami[14], R. Demarco[15], M. T. Sargent[16,17], X. Shu[18], J. S. Kartaltepe[19], D. M. Alexander[20], M. Béthermin[4], F. Bournaud[2], L. Ciesla[4], H. C. Ferguson[21], S. L. Finkelstein[5], M. Giavalisco[22], Q.-S. Gu[1,3], D. Iono[23,24], S. Juneau[6], G. Lagache[4], R. Leiton[15], H. Messias[25,26], K. Motohara[27], J. Mullaney[28], N. Nagar[15,29], M. Pannella[30,31], C. Papovich[32,33], A. Pope[25], C. Schreiber[34], and J. Silverman[9]

(Affiliations can be found after the references)

Received 30 September 2022 / Accepted 1 February 2023

ABSTRACT

Our current understanding of the cosmic star formation history at $z > 3$ is primarily based on UV-selected galaxies (Lyman-break galaxies, i.e., LBGs). Recent studies of H-dropouts (HST-dark galaxies) have revealed that we may be missing a large proportion of star formation that is taking place in massive galaxies at $z > 3$. In this work, we extend the H-dropout criterion to lower masses to select optically dark or faint galaxies (OFGs) at high redshifts in order to complete the census between LBGs and H-dropouts. Our criterion ($H > 26.5$ mag & $[4.5] < 25$ mag) combined with a de-blending technique is designed to select not only extremely dust-obscured massive galaxies but also normal star-forming galaxies (typically $E(B-V) > 0.4$) with lower stellar masses at high redshifts. In addition, with this criterion, our sample is not contaminated by massive passive or old galaxies. In total, we identified 27 OFGs at $z_{\rm phot} > 3$ (with a median of $z_{\rm med} = 4.1$) in the GOODS-ALMA field, covering a wide distribution of stellar masses with $\log(M_\star/M_\odot) = 9.4-11.1$ (with a median of $\log(M_{\star\rm med}/M_\odot) = 10.3$). We find that up to 75% of the OFGs with $\log(M_\star/M_\odot) = 9.5-10.5$ were neglected by previous LBGs and H-dropout selection techniques. After performing an optical-to-millimeter stacking analysis of the OFGs, we find that rather than being limited to a rare population of extreme starbursts, these OFGs represent a normal population of dusty star-forming galaxies at $z > 3$. The OFGs exhibit shorter gas depletion timescales, slightly lower gas fractions, and lower dust temperatures than the scaling relation of typical star-forming galaxies. Additionally, the total star formation rate ($\rm SFR_{tot} = SFR_{IR} + SFR_{UV}$) of the stacked OFGs is much higher than the $\rm SFR_{UV}^{corr}$ ($\rm SFR_{UV}$ corrected for dust extinction), with an average $\rm SFR_{tot}/SFR_{UV}^{corr} = 8 \pm 1$, which lies above (~0.3 dex) the 16–84th percentile range of typical star-forming galaxies at $3 \leq z \leq 6$. All of the above suggests the presence of hidden dust regions in the OFGs that absorb all UV photons, which cannot be reproduced with dust extinction corrections. The effective radius of the average dust size measured by a circular Gaussian model fit in the *uv* plane is $R_{\rm e(1.13\,mm)} = 1.01 \pm 0.05$ kpc. After excluding the five LBGs in the OFG sample, we investigated their contributions to the cosmic star formation rate density (SFRD). We found that the SFRD at $z > 3$ contributed by massive OFGs ($\log(M_\star/M_\odot) > 10.3$) is at least two orders of magnitude higher than the one contributed by equivalently massive LBGs. Finally, we calculated the combined contribution of OFGs and LBGs to the cosmic SFRD at $z = 4$–5 to be $4 \times 10^{-2}\ M_\odot\ \rm yr^{-1}\ Mpc^{-3}$, which is about 0.15 dex (43%) higher than the SFRD derived from UV-selected samples alone at the same redshift. This value could be even larger, as our calculations were performed in a very conservative way.

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

1. Introduction

Our current knowledge of the first two billion years of cosmic star formation history is based mainly on (i) UV-selected galaxies, such as Lyman-break galaxies (LBGs; e.g., Giavalisco et al. 2004; Bouwens et al. 2012a, 2015, 2020; Oesch et al. 2014, 2015, 2018; Madau & Dickinson 2014), which are known to be biased against massive galaxies; and (ii) the most massive and extremely dusty starburst galaxies (e.g., Walter et al. 2012; Marrone et al. 2018), which are limited to a rare population and are not representative of the most common galaxies typically on the star-formation main sequence (SFMS; e.g., Elbaz et al. 2007, 2011; Noeske et al. 2007; Magdis et al. 2010; Whitaker et al. 2012, 2014; Speagle et al. 2014; Schreiber et al. 2015; Lee et al. 2015; Leslie et al. 2020). Recent Atacama Large Millimeter and Submillimeter Array (ALMA) and *Spitzer* observations have identified a more abundant and less extreme population of obscured galaxies at $z > 3$ (e.g., H-dropouts in Wang et al. 2019; HST-dark galaxies in Zhou et al. 2020, optically dark or faint galaxies in Gómez-Guijarro et al. 2022a), revealing that a significant population of high-z optically dark or faint galaxies have been missed, and they may dominate the massive end of the stellar mass function. The contribution of these optically dark or faint galaxies to the cosmic star formation rate density (SFRD) at $z > 3$ could be substantial, corresponding up to ~10–25% of the SFRD from LBGs, or even up to ~40%, depending on the methodology (e.g., Wang et al. 2019; Williams et al. 2019; Gruppioni et al. 2020; Fudamoto et al. 2021; Talia et al. 2021; Enia et al. 2022; Shu et al. 2022; Barrufet et al. 2022).

Optically dark or faint galaxies have generally been completely undetected or tentatively detected with very low significance even in the deepest HST/WFC3 images (typical 5σ depth of $H > 27$ mag), but brighter at longer wavelengths such as *Spitzer*/IRAC 3.6 and 4.5 μm, (e.g., Franco et al. 2018; Yamaguchi et al. 2019; Zhou et al. 2020; Smail et al. 2021; Gómez-Guijarro et al. 2022a). In GOODS-ALMA 1.0,

Franco et al. (2020a) reported six optically dark galaxies (i.e., HST-dark galaxies) out of 35 galaxies detected above 3.5σ at 1.13 mm. With the ALMA spectroscopic follow-up, Zhou et al. (2020) further analyzed these six optically dark galaxies in detail and found that four (∼70%) could be associated with a $z \sim 3.5$ overdensity (corresponding to OFG1, 2, 25, and 27 in the south-west region of Fig. 1). Afterward, in the deeper GOODS-ALMA 2.0, Gómez-Guijarro et al. (2022a) updated the sample with 13 optically dark or faint galaxies (including six in the GOODS-ALMA 1.0), among a total of 88 sources detected above 3.5σ at 1.13 mm. So far, we do not have a unified and clear definition of optically dark or faint galaxies. The six optically dark galaxies in GOODS-ALMA 1.0 have no optical counterparts in the deepest H-band based on the CANDELS catalog down to $H = 28.16$ AB (5σ limiting depth in CANDELS-deep field). However, two of them show H-band magnitudes of approximately 25 mag and 27 mag following a de-blending process (Zhou et al. 2020). The remaining seven sources were classified as optically dark or faint galaxies because they are currently undetected or very faint in the H-band of the deepest fields and other shorter wavelength bands (Gómez-Guijarro et al. 2022a).

Therefore, the purpose of our work is to first make a clear definition of the selection of optically dark or faint galaxies. Furthermore, by systematically studying optically dark or faint galaxies in the GOODS-ALMA field, we aim to obtain a more complete picture of the cosmic star formation history in the $z > 3$ Universe. In this work, our sample includes not only sources detected by ALMA 1.13 mm, but also those that are currently undetected (i.e., no millimeter counterparts in the GOODS-ALMA 2.0 catalog) to obtain a somewhat complete sample of optically dark or faint galaxies. By stacking their optical to millimeter emission, we can, however, investigate the differences between the optically dark or faint galaxies detected by ALMA 1.13 mm and those that remain undetected.

This paper is organized as follows. In Sect. 2, we describe the GOODS-ALMA survey and the multiwavelength data used. In Sect. 3, we present our selection criterion for optically dark or faint galaxies at $z > 3$. In Sect. 4, we study the properties of individual sources in our sample, such as the redshift, stellar mass, star formation rate (SFR), molecular gas mass, and dust temperature. In Sect. 5, we present and discuss the properties of optically dark or faint galaxies mainly based on our optical to millimeter stacking analysis. In Sect. 6, we calculate the cosmic SFRD contributed by optically dark or faint galaxies and discuss the level of the incompleteness of our understanding of dust-obscured star formation in the $z > 3$ Universe. Finally, we summarize our main conclusions in Sect. 7.

Throughout this paper, we adopt a Chabrier initial mass function (IMF; Chabrier 2003) to estimate SFR and stellar mass. We assume cosmological parameters of $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. When necessary, data from the literature have been converted with a conversion factor of $M_\star$ (Salpeter 1955, IMF) $= 1.7 \times M_\star$ (Chabrier 2003, IMF). All magnitudes are in the AB system (Oke & Gunn 1983), such that $m_{\rm AB} = 23.9 - 2.5 \times \log(S_\nu\,[\mu\mathrm{Jy}])$.

2. Data

2.1. GOODS-ALMA survey

GOODS-ALMA is an ALMA 1.13 mm survey in the deepest part of the Great Observatories Origins Deep Survey South field (GOODS-South; Dickinson et al. 2003; Giavalisco et al. 2004). It covers a continuous area of 72.42 arcmin^2 (effective area of ∼69 arcmin^2 if the shallower areas at the edges are trimmed off) with ALMA band 6 receivers, centered at $\alpha = 3^{\rm h}\,32^{\rm m}\,30.0^{\rm s}$, $\delta = -27^\circ\,48'\,00''$ (J2000). The observations were obtained from Cycle 3 and Cycle 5, with two different array configurations to include both small and large spatial scales. The ALMA Cycle 3 observations (high-resolution dataset; Project ID: 2015.1.00543.S; PI: D. Elbaz) were conducted between August and September 2016 in the C40-5 array configuration with a total on-source exposure time of approximately 14.06 h, providing a high-resolution image with the longest baseline of 1808 m. The ALMA Cycle 5 observations (low-resolution dataset; Project ID: 2017.1.00755.S; PI: D. Elbaz) were performed between July 2018 and March 2019 with the C43-2 array configuration with a total on-source exposure time of 14.39 h, providing a low-resolution image with the longest baseline of 360.5 m.

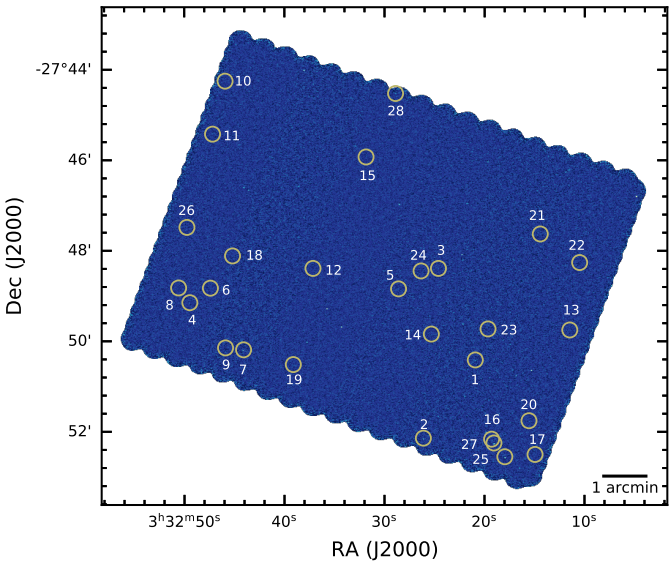

Fig. 1. Sky distributions of the optically dark or faint galaxies (OFGs) in the GOODS-ALMA 2.0 map at 1.13 mm. North is up and east is to the left.

The calibration was processed using the Common Astronomy Software Application package (CASA; McMullin et al. 2007) with the standard pipeline. We systematically inspected calibrated visibilities and added a few additional flags to the original calibration scripts. The calibrated visibilities were then time- and frequency-averaged over 120 s and 8 channels, respectively, to reduce the computational time for subsequent continuum imaging. Given the excellent coverage of the uv plane and the absence of very bright sources, we used the task `TCLEAN` in CASA version 5.6.1–8 to produce a dirty map with 0.05″ pixels and a natural weighting scheme to avoid potential biases from the `CLEAN` algorithm. The resulting high- and low-resolution 1.13 mm continuum maps have similar root mean square (rms) sensitivities, that is, $\sigma \simeq 89.0$ and 95.2 μJy beam^{-1}, with spatial resolutions of full width at half maximum ($FWHM$) $\simeq 0''.251 \times 0''.232$ and $1''.330 \times 0''.935$, respectively. To improve the sensitivity, we concatenated these two data configurations in the uv plane with visibility weights proportional to 1:1. The combined map achieves an rms sensitivity of $\sigma \simeq 68.4$ μJy beam^{-1} with a spatial resolution of $0''.447 \times 0''.418$ (see Fig. 1). For more details on the same data reduction, we refer to Franco et al. (2018) for the high-resolution dataset (GOODS-ALMA 1.0) and

Gómez-Guijarro et al. (2022a) for the low-resolution dataset and the combined dataset (GOODS-ALMA 2.0).

2.2. Multiwavelength images

Here we list the multiwavelength data we used for ultraviolet (UV) to mid-infrared (MIR) and MIR to millimeter (mm) spectral energy distribution (SED) fitting, as well as those used for the stacking analysis (see Sects. 4.1, 4.2, and 5.2): (*i*) X-ray data: *Chandra* 7 Ms (0.5–7.0 keV, 0.5–2.0 keV, and 2–7 keV bands) images in the *Chandra* Deep Field-South (CDF-S) field (Luo et al. 2017); (*ii*) UV, optical (OPT), and near-infrared (NIR) data: HST/ACS (*F*435*W*, *F*606*W*, *F*775*W*, *F*814*W*, *F*850*LP*) and HST/WFC3 (*F*105*W*, *F*125*W*, *F*140*W*, *F*160*W*) images from the *Hubble* Legacy Fields Program (HLF v2.0; Whitaker et al. 2019), VLT/VIMOS (*U*, *R*; Nonino et al. 2009) images, VLT/ISAAC (*H*, *J*, K_s; Retzlaff et al. 2010) images, VLT/HAWK-I (K_s; Fontana et al. 2014) images, *Magellan*/FourStar (H_s, H_l, J_1, J_2, J_3, K_s) images from ZFOURGE (Straatman et al. 2016), and CFHT/WIRCAM (*J*, K_s; Hsieh et al. 2012) images; (*iii*) MIR data: deepest *Spitzer*/IRAC images from the GREATS program (3.6 µm, 4.5 µm, 5.8 µm, 8 µm; Stefanon et al. 2021), which were obtained by combining programs of the IUDF (PI: I. Labbé), IGOODS (PI: P. Oesch), GOODS (PI: M. Dickinson), ERS (PI: G. Fazio), S-CANDELS (PI: G. Fazio), SEDS (PI: G. Fazio), UDF2 (PI: R. Bouwens), and GREATS (PI: I. Labbé); (*iv*) far-infrared (FIR) data: *Spitzer*/MIPS (24 µm) images from GOODS (PI: M. Dickinson). We used *Herschel*/PACS images (100 µm, 160 µm; Magnelli et al. 2013) combined from the PEP (Lutz et al. 2011) and GOODS-*Herschel* (Elbaz et al. 2011) programs and *Herschel*/SPIRE (250 µm, 350 µm, 500 µm; Elbaz et al. 2011) images; (*v*) millimeter data: 1.13 mm map of GOODS-ALMA 2.0, which is a combination of high- and low-resolution 1.13 mm continuum maps (see details in Sect. 2.1); and (*vi*) radio data: radio image at 3 GHz (10 cm) from the VLA (PI: W. Rujopakarn, private communication), which covers the entire GOODS-ALMA field (Rujopakarn et al. 2016, and in prep.). In Table 1, we summarize the multiwavelength dataset used in this work from UV to mm.

Table 1. Broad and intermediate bands (UV to mm) in this work.

Telescope/Camera	Filter	λ_c (µm)	Ref.
VLT/VIMOS	*U*	0.3759	(a)
	R	0.6481	
HST/ACS	*F*435*W*	0.4347	(b)
	*F*606*W*	0.6033	
	*F*775*W*	0.7730	
	*F*814*W*	0.8143	
	*F*850*LP*	0.9085	
HST/WFC3	*F*105*W*	1.0644	(b)
	*F*125*W*	1.2561	
	*F*140*W*	1.4064	
	*F*160*W*	1.5463	
CFHT/WIRCAM	*J*	1.2554	(c)
	K_s	2.1630	
VLT/ISAAC	*J*	1.2423	(d)
	H	1.6560	
	K_s	2.1709	
VLT/HAWK-I	K_s	2.1586	(e)
Magellan/FourStar	J_1	1.0552	(f)
	J_2	1.1472	
	J_3	1.2819	
	H_s	1.5564	
	H_l	1.7038	
	K_s	2.1599	
Spitzer/IRAC	CH1	3.5763	(g)
	CH2	4.5289	
	CH3	5.7875	
	CH4	8.0449	
Spitzer/MIPS	MIPS	24	(h)
Herschel/PACS	blue	70	(i)
	green	100	
	red	160	
Herschel/SPIRE	PSW	250	
	PMW	350	
	PLW	500	
ALMA	band 6	1130	(j)

Notes. These bands are used for the UV-to-MIR and MIR-to-mm SED fitting. From left to right: Telescope and/or instrument, the filter name, the central wavelength of the filters, and the reference for the survey, including the images and catalogs we used. [(a)]Nonino et al. (2009). [(b)]HLF program (Whitaker et al. 2019). [(c)]Hsieh et al. (2012). [(d)]Retzlaff et al. (2010). [(e)]Fontana et al. (2014). [(f)]ZFOURGE program (Straatman et al. 2016). [(g)]Image: the GREATS program (Stefanon et al. 2021); catalog: this work (see Appendix A). [(h)]PI: M. Dickinson; catalog: Magnelli et al. (2011). [(i)]Image: Magnelli et al. (2013) and Elbaz et al. (2011); catalog: T. Wang (priv. comm.) and Elbaz et al. (2011). [(j)]Image: this work (see Sect. 2.1); catalog (and also image): Gómez-Guijarro et al. (2022a).

2.3. The multiwavelength catalogs

We list below the main catalogs we used in this work. Briefly, we combined catalogs from the HST (HLF v2.0; Whitaker et al. 2019) and IRAC to select our sample, and from *Herschel* and ALMA (GOODS-ALMA 2.0; Gómez-Guijarro et al. 2022a) for the MIR-to-mm SED fitting. In addition, we also used the GOODS-ALMA 2.0 catalog to classify our sources detected at ALMA 1.13 mm and the X-ray catalog (CDF-S 7 Ms; Luo et al. 2017) to help identify candidate X-ray active galactic nuclei (AGN). We note that here we have corrected for systematic and local astrometric offsets in different catalogs following Franco et al. (2020a) and Whitaker et al. (2019) to ensure a consistent astrometric reference frame.

1. The HST *H*-band catalog: The HLF v2.0 in the GOODS-South region (HLF-GOODS-S) uses a deep noise-equalized combination of four HST bands (*F*850*LP*, *F*125*W*, *F*140*W*, *F*160*W*) for detection (Whitaker et al. 2019). The catalog includes all UV, optical, and NIR data (13 filters in total) taken by HST over 18 years across the field. The 5σ point-source depth in the *H*-band is approximately 27.0−29.8 mag. We note that the GOODS-South field includes the HUDF, which is much deeper than other parts of the HST data, but covers only a small portion of the field.
2. IRAC catalog: We built the IRAC catalog in the GOODS-ALMA field using the deepest IRAC 3.6 and 4.5 µm images from the GREATS program (Stefanon et al. 2021). The source detection was performed using `Source Extractor` (SE version 2.25.0; Bertin & Arnouts 1996) on the background-subtracted 3.6 and 4.5 µm images.

The total sources in the catalogs are 125 338 for 3.6 µm and 154 234 for 4.5 µm in the GOODS-S field (∼150 arcmin^2).

To ensure the purity of detections, we then cross-matched these two catalogs with a radius of 1.0″ (∼0.5 FWHM). Furthermore, for sources in the GOODS-ALMA 2.0 catalog (Gómez-Guijarro et al. 2022a) that were detected at least in one IRAC band, we also considered them to be real sources and kept them in the final catalog. Finally, we end up with 71 899 sources in the IRAC catalog, of which 5127 are in the GOODS-ALMA region (see Appendix A for more details).

3. *Spitzer*/MIPS 24 μm and *Herschel* catalog: We mainly used the catalog of Wang et al. (in prep.), which was built based on a state-of-the-art de-blending method, similar to that used in the 'super-deblended' catalogs in GOODS-North (Liu et al. 2018) and COSMOS (Jin et al. 2018). Briefly, Wang et al. used *Spitzer*/MIPS 24 μm detections as priors for source extraction on the PACS and SPIRE images. They performed super-deblending in one band each time, from shorter to longer wavelengths, and predicted fluxes at longer wavelengths based on the redshift and photometry information given by the shorter wavelengths. Then, the faint priors at longer wavelengths were removed, which helped break the blending degeneracies.
   For one source (OFG27 in our work) still affected by blending with problematic photometry from MIR to FIR, we used the catalog of Elbaz et al. (2011).
4. GOODS-ALMA 2.0 catalog: This catalog contains 88 sources detected by ALMA at 1.13 mm in the GOODS-ALMA 2.0 field (Gómez-Guijarro et al. 2022a). These include 50% of sources detected above 5σ with a purity of 100% and 50% detected within the 3.5 and 5σ range aided by priors. The median redshift and stellar mass of the 88 sources are $z_{\rm med} = 2.46$ and $\log(M_{\star\rm med}/M_\odot) = 10.56$ (Chabrier 2003 IMF), respectively.
5. X-ray catalog: We used the CDF-S 7 Ms catalog (Luo et al. 2017). It contains 1008 sources in the main source catalog, observed in three X-ray bands (0.5–7.0 keV, 0.5–2.0 keV, and 2–7 keV) and 47 lower-significance sources in a supplementary catalog. This catalog includes the candidate X-ray AGNs identified by Luo et al. (2017), which we used in this work to search for X-ray AGN counterparts of our sources.

3. Selection of optically dark or faint galaxies at $z > 3$

We aim to obtain a more complete picture of the cosmic star formation history, that is, to bridge the extreme population of optically dark galaxies (e.g., Wang et al. 2019) with the most common population of lower-mass, less-attenuated galaxies, such as those selected using the LBG selection technique (e.g., Bouwens et al. 2012a, 2015, 2020). To reach this goal, we chose to select our sample with a less strict cut than the one used in Wang et al. (2019) for the H-band and 4.5 μm magnitudes, which would allow for our sample to encompass lower-mass and less-attenuated galaxies, while still including extremely dust-obscured galaxies. Here, we call them optically dark or faint galaxies (hereafter, OFGs) and we select them with the criteria defined below (see Sect. 3.1).

3.1. A selection criterion for OFGs uncontaminated by massive passive galaxies

To start, we used a combination of H-band and IRAC 4.5 μm photometry, where the H-band measures rest-frame UV light ($\lambda < 4000$ Å) for galaxies at $z > 3$, while the 4.5 μm band probes the rest-frame J-band. Galaxies with red color in $H-[4.5]$ are either quiescent or passive galaxies or dusty star-forming galaxies with significant dust extinction (UV attenuation).

For optically dark galaxies at high redshifts, a common selection approach in the literature is to target H-dropouts, which are defined to be undetected in the H-band (i.e., absent in the H-band catalog) and bright in the IRAC band (e.g., $[4.5] < 24$ mag; Wang et al. 2019; see the blue triangular region in Fig. 2), and/or extremely red in color (e.g., $H-[4.5] > 4$ in Shu et al. 2022). These methods can help to select extremely dust-obscured massive galaxies. However, the detection of an H-dropout obviously depends on the depth of the H-band. For example, in the HLF-GOODS-S field, the 5σ depth in the H-band ranges from 27.0 mag to 29.8 mag (Whitaker et al. 2019). To avoid this imprecise selection and to extend the selection in view of bridging the heavily obscured star-forming galaxies with more common galaxies, we defined OFGs based on the following characteristics: 1) $H > 26.5$ mag; 2) $[4.5] < 25$ mag (see the red triangular region in Fig. 2). Instead of only selecting galaxies undetected in the H-band and/or extremely red in H-[4.5], we used the criterion of $H > 26.5$ mag to select not only optically dark sources but also optically faint galaxies with less dust obscuration. The $H > 26.5$ mag cut also helps to distinguish massive passive galaxies with stellar masses $\log(M_\star/M_\odot) > 10$ (the grey region in Fig. 2; see more details afterward) from our selected OFGs (the red triangular region in Fig. 2). In Fig. 2, the faintest modeled passive galaxy in the H-band ($H = 26.0$ mag) has a similar magnitude to the brightest OFGs (OFG25 and OFG26: $H = 26.3 \pm 0.2$ mag), considering the 1σ uncertainty in the flux measurements. The criterion of $[4.5] < 25$ mag can help to select not only massive galaxies, but also galaxies with intermediate stellar masses.

To verify the reliability of our selection criteria, we investigated the evolutionary tracks of theoretical galaxy templates at $z = 2$–6 in the color-magnitude diagram (H-[4.5] vs. [4.5]; Fig. 2). The templates are based on the BC03 models (Bruzual & Charlot 2003), including an instantaneous burst model (i.e., simple stellar population; SSP) formed at $z = 10$ and a non-evolving constant star formation (CSF) model with an age of 300 Myr with varying degrees of reddening. The templates have stellar masses in the range of $\log(M_\star/M_\odot) = 10$–$11$, with the Calzetti et al. (2000) attenuation law and solar metallicity. The SSP model corresponds to passive or old galaxies with an age of 1 Gyr at $z = 3.5$, while the CSF models represent star-forming galaxies with different dust obscurations.

Our selection criteria for OFGs are shown in Fig. 2 as the red triangular region, which encompasses high redshift galaxies with lower stellar mass and less dust attenuation that were excluded by previous H-dropout selection criteria such as those of Wang et al. (2019; the blue triangular region in Fig. 2). For instance, the selected OFGs include those with $E(B-V) = 0.4$, $\log(M_\star/M_\odot) = 10$, and $z = 4$–5, as well as those with $E(B-V) = 0.6$, $\log(M_\star/M_\odot) = 10$, and $z = 3$–4. Similarly, extremely dust-obscured massive galaxies, such as those with $E(B-V) = 0.8$, $\log(M_\star/M_\odot) = 11$, and $z > 4$ can also be selected by our criteria. We note that although a few OFGs ($E(B-V) = 0.8$ and $\log(M_\star/M_\odot) = 10$) at $z = 2.5$–3 were selected, the total OFGs dominate at $z > 3$. Overall, in our selection, the majority of OFGs have $E(B-V) > 0.4$ and are at $z > 3$. In addition, with these criteria, our sample is not contaminated by massive passive or old galaxies ($\log(M_\star/M_\odot) > 10$; the grey region in Fig. 2). Therefore, the selection of OFGs at high redshifts with $H > 26.5$ mag and $[4.5] < 25$ mag is a reliable approach. In summary, the selected OFGs contain not only extremely

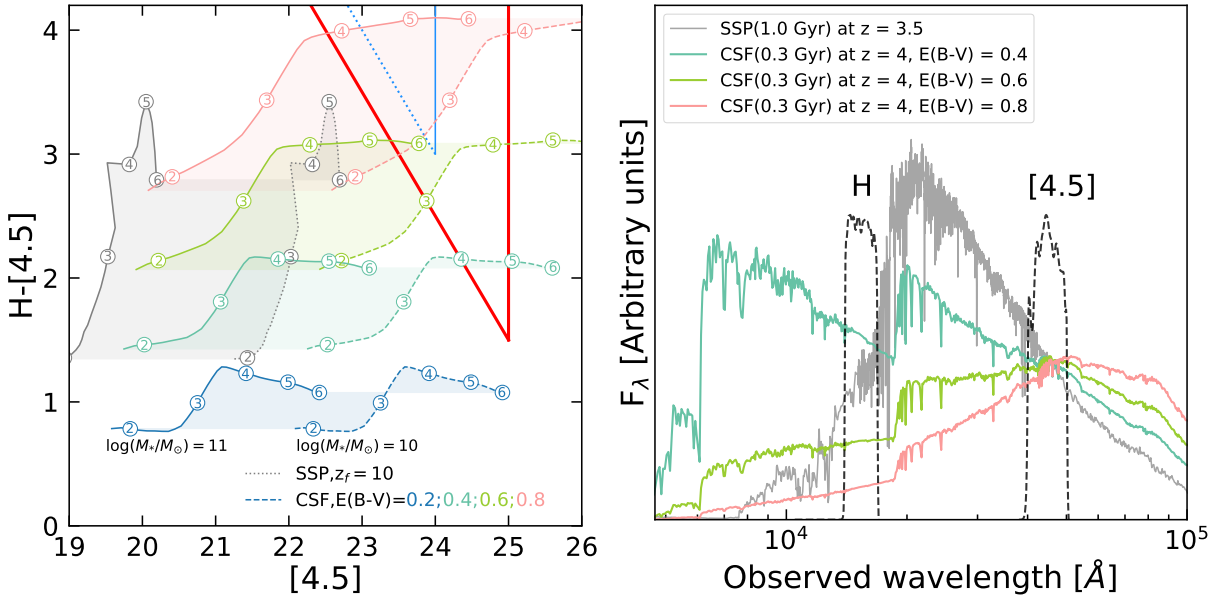

Fig. 2. Selection criteria for OFGs at $z \gtrsim 3$. *Left*: color-magnitude diagram for the selection of OFGs. This diagram shows the evolutionary tracks of theoretical galaxy templates at $z = 2$–6, with stellar mass $\log(M_\star/M_\odot) = 10$–$11$ and solar metallicity. The numbers in the circles indicate the redshift. The solid and dashed evolutionary tracks correspond to stellar masses with $\log(M_\star/M_\odot) = 11$ and $\log(M_\star/M_\odot) = 10$, respectively. These templates are based on the BC03 models (Bruzual & Charlot 2003), including an instantaneous burst (i.e., simple stellar population; SSP) model formed at $z = 10$ (grey area) and a non-evolving constant star formation (CSF) model with an age of 300 Myr with varying degrees of reddening (blue, green, olive, and red areas). The SSP model corresponds to a passive or old galaxy with an age of 1 Gyr at $z = 3.5$. The blue triangular region shows the selection criteria for H-dropouts in Wang et al. (2019), i.e., no H-band flux above 5σ ($H > 27$ mag; blue dashed line) and [4.5] < 24 mag (blue solid line). The red triangular region shows the selection criteria for OFGs in this work. *Right*: different theoretical galaxy templates in HST/$F160W$ (H-band) and IRAC 4.5 μm filters. The templates include a passive galaxy with an age of 1 Gyr at $z = 3.5$ and three star-forming galaxies with different degrees of reddening at $z = 4$. At a similar [4.5], the three star-forming galaxies with different dust extinction values exhibit different H-band magnitudes.

dust-obscured massive galaxies at $z > 4$, but also lower-mass and less-attenuated (typically $E(B-V) > 0.4$) galaxies at $z > 3$, without contamination from massive passive galaxies.

3.2. Sample selection

We selected candidate OFGs by cross-matching our IRAC catalog at 3.6 and 4.5 μm (see Appendix A for details on the catalog construction) with the HLF catalog (Whitaker et al. 2019) in the GOODS-ALMA field. Candidates were required to have none or very faint HST counterparts ($H > 26.5$ mag) within a 0.6″ radius around the IRAC positions. The small radius of 0.6″ (roughly one-third of the point spread function size of IRAC 4.5 μm) was taken to select as many individual candidates as possible, while avoiding excessive contamination of our final sample by fake candidates. After cross-matching, we had 88 candidates.

3.3. Photometry

We visually inspected all the candidates and noted that blending was common in the IRAC images due to their relatively worse spatial resolution (e.g., ∼2″ at 4.5 μm). To obtain photometric values in different bands without contamination from neighboring galaxies, we simultaneously de-blended sources in the multi-wavelength images (from UV to 8 μm) by applying the de-blending code[1] described in Schreiber et al. (2018b). The de-blending method is briefly summarized in the following three steps. Step 1: for each OF candidate, to save computational time, we first cut the stacked HST image (four bands of HST/$F105W$, $F125W$, $F140W$, and $F160W$) into a 10″ × 10″ area around the IRAC position of the candidate. Then, we detected all sources in the clipped stacked HST image.

For some optically dark galaxies undetected even in the stacked HST image, we used positions of their K_s-band, 4.5 μm, or ALMA counterparts and modeled them as point sources. Step 2: following Schreiber et al. (2018b), we fitted all the sources detected in the clipped stacked HST image simultaneously with a single Sérsic profile to obtain a best-fit deconvolved model (intrinsic light profile) for each source. Step 3: these models were then convolved with a point spread function (PSF) for each image at all wavelengths (up to 8 μm). We then used the positions and the PSF-convolved models as priors for all objects to fit the multi-wavelength images. The uncertainties of the fluxes were calculated by Monte Carlo simulations. We find that 60% (53/88) of the sources in our candidate sample needed to be de-blended at 4.5 μm.

3.4. Incompleteness correction and final sample

To identify high-z OFGs, we used the simple selection techniques discussed in Sect. 3.1: (1) $H > 26.5$ mag and (2) [4.5] < 25 mag. Considering the 1σ uncertainty of the flux measurements, we finally identified 26 individual OFGs in total in the GOODS-ALMA field (see Figs. 1 and 3).

Here, we discuss the corrections for the incompleteness of our sample selection approach. Considering the criterion of no/very faint HST counterparts within the search radius of 0.6″ at the IRAC position, we may have missed some target sources simply due to random bright HST sources falling within the

[1] qdeblend: https://github.com/cschreib/qdeblend

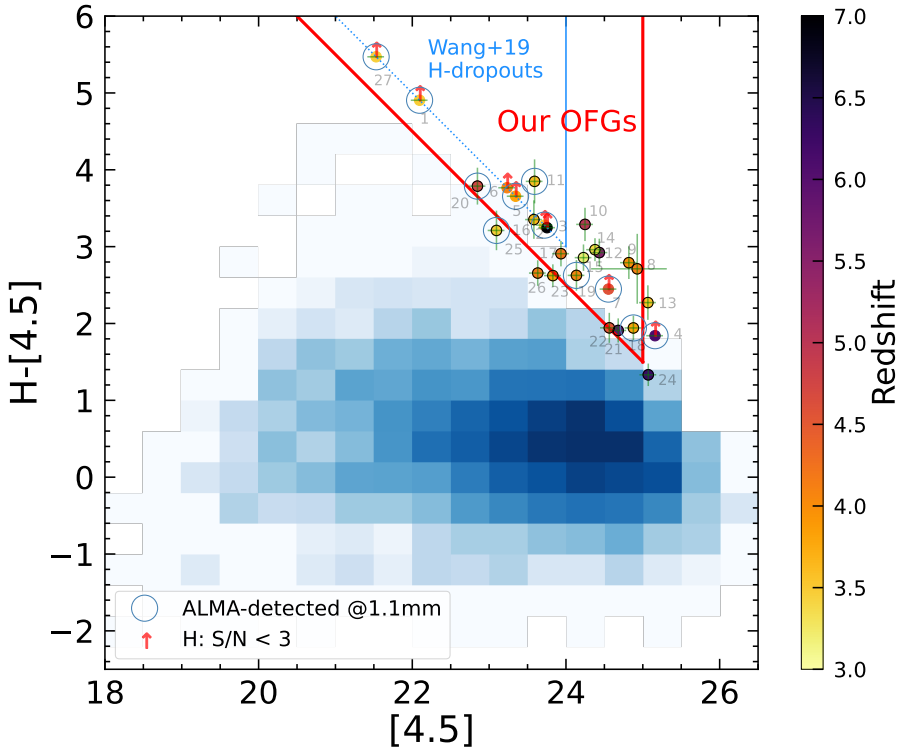

Fig. 3. Color-magnitude diagram color-coded by photometric redshift. Our criteria (in red) for selecting OFGs are: $H > 26.5$ mag & $[4.5] < 25$ mag. We note that we include sources outside the wedge whose 1σ photometric uncertainties overlap the wedge, so we have points outside the red triangle. The arrows show our H-dropouts ($S/N < 3$), with the typical depth of $H = 27$ mag (5σ) in the shallowest region of the HLF survey as their lower limits. The blue shaded area describes the distribution of all the IRAC detected sources (see Appendix A for more details) in the GOODS-ALMA field. The blue and red triangular regions are the same as in Fig. 2.

radius. Following the same method as in Lilly et al. (1999; also in Wang et al. 2019), at a given position, the probability of finding one or more random galaxies within a given radius is defined as:

$$p(n \geq 1) = 1 - \exp(-N \times \pi \times \mathrm{radius}^2), \quad (1)$$

where N represents the surface density of bright HST sources ($H < 26.5$ mag) in our case. In the GOODS-ALMA field, N is 0.05 galaxies arcsec^{-2}, so the derived $p(n \geq 1) = 0.056$. It suggests that using this selection approach, we may have missed around 5% of OFGs that we wrongly associated with a counterpart due to projection effects. That is, one source could have been missed due to the serendipitous presence of a bright H detection within our 0.6″ radius search circle.

After comparing our sample with 13 ALMA-detected OFGs in the GOODS-ALMA 2.0 catalog, we confirmed that the missing galaxy is OFG27 (A2GS7 in Gómez-Guijarro et al. 2022a). A random bright source (ID55970 in the HLF catalog) with $H = 24.5$ mag is located at a distance of 0.33″ (<0.6″ searching radius) from OFG27. This bright source has been confirmed not to be the H-band counterpart of OFG27 (AGS17 in Zhou et al. 2020). Therefore, we included OFG27 in our catalog (see Table 2) to correct the incompleteness of our selection approach. Also, we included OFG27 in the analysis of the main discussion presented in this paper.

In addition, an extra IRAC 4.5 μm dropout candidate was detected only in the longer wavelength images: JCMT/SCUBA-2 850 μm and ALMA 870 μm and 1.13 mm and 1.2 mm (OFG28; see Table 2). Including this one, we have 28 OFGs in our final catalog (Table 2). Considering that OFG28 is an IRAC 4.5 μm dropout candidate with $[4.5] \gg 25$ mag, in the following analysis, we focus only on the first 27 OFGs, which meet our criteria: $H > 26.5$ mag & $[4.5] < 25$ mag.

3.5. Lyman-break galaxies and H-dropouts in final sample

To compare our sample with LBGs (e.g., Bouwens et al. 2012a) and H-dropouts (e.g., Wang et al. 2019), we need to determine how many OFGs are LBGs or H-dropouts in our sample.

3.5.1. Lyman-break galaxies

LBGs are a UV-selected population of high-z star-forming galaxies. To understand how many galaxies in our OFG catalog are missed by this UV-selected approach and to further know their contribution to the cosmic SFRD, we identify LBGs from our OFG catalog by employing the Lyman-break color criteria used in Bouwens et al. (2020; also see similar methods in Bouwens et al. 2012a, 2015). The redshift range of our OFGs is $z = 3$–7 (see Sect. 4.1, Fig. 4, and Table 2). The Lyman-break color criteria are as follows:

$$z \sim 3 : (U_{336} - B_{435} > 1) \wedge (B_{435} - V_{606} < 1.2) \wedge (i_{775} - Y_{105} < 0.7) \wedge (\chi^2_{\mathrm{UV}_{225},\mathrm{UV}_{275}} < 2),$$

$$z \sim 4 : (B_{435} - V_{606} > 1) \wedge (i_{775} - J_{125} < 1) \wedge (B_{435} - V_{606} > 1.6(i_{775} - J_{125}) + 1),$$

$$z \sim 5 : (V_{606} - i_{775} > 1.2) \wedge (z_{850} - H_{160} < 1.3) \wedge (V_{606} - i_{775} > 0.8(z_{850} - H_{160}) + 1.2),$$

Table 2. Derived properties of the OFGs.

ID	RA (deg)	Dec (deg)	H (mag)	[4.5] (mag)	$S_{1.13\,\mathrm{mm}}$ (mJy)	z	$\log(M_\star)$ $\log(M_\odot)$	$\log(L_{\mathrm{IR}})$ $\log(L_\odot)$	SFR ($M_\odot$ yr^{-1})	Other ID
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
OFG1	53.087184 (†)	−27.840242 (†)	(...)	$22.10^{+0.11}_{-0.10}$	0.85 ± 0.11	3.47^{o}	$10.79^{+0.11}_{-0.16}$	12.41 ± 0.05	384 ± 48	AGS24, A2GS29
OFG2	53.108810 (†)	−27.869037 (†)	(...)	$23.72^{+0.11}_{-0.10}$	1.24 ± 0.10	3.47^{o}	$10.76^{+0.11}_{-0.17}$	12.39 ± 0.03	365 ± 30	AGS11, A2GS15
OFG3	53.102536	−27.806531	(...)	$23.75^{+0.11}_{-0.10}$	(...)	$7.04^{+1.01}_{-1.19}$	$10.91^{+0.22}_{-0.14}$	(...)	(...)	(...)
OFG4	53.206064 (†)	−27.819142 (†)	(...)	$25.16^{+0.11}_{-0.10}$	0.70 ± 0.13	$6.13^{+1.02}_{-1.36}$	$10.60^{+0.16}_{-0.36}$	12.56 ± 0.07	537 ± 100	A2GS38
OFG5	53.119150 (†)	−27.814066 (†)	(...)	$23.34^{+0.11}_{-0.10}$	0.43 ± 0.13	$3.81^{+0.74}_{-0.94}$	$10.17^{+0.34}_{-0.27}$	12.26 ± 0.09	268 ± 61	A2GS87, GDS44539
OFG6	53.197493	−27.813789	(...)	$23.23^{+0.11}_{-0.10}$	(...)	$4.10^{+0.31}_{-0.32}$	$10.93^{+0.06}_{-0.06}$	(...)	(...)	(...)
OFG7	53.183697 (†)	−27.836500 (†)	(...)	$24.55^{+0.11}_{-0.10}$	1.23 ± 0.12	$4.58^{+0.42}_{-0.42}$	$10.34^{+0.08}_{-0.09}$	12.86 ± 0.04	1070 ± 101	AGS25, A2GS17
OFG8	53.210736	−27.813706	$27.64^{+0.31}_{-0.24}$	$24.93^{+0.60}_{-0.40}$	(...)	$4.14^{+0.41}_{-0.44}$	$9.92^{+0.18}_{-0.27}$	(...)	(...)	(...)
OFG9	53.191204	−27.835791	$27.61^{+0.22}_{-0.19}$	$24.82^{+0.11}_{-0.10}$	(...)	$3.99^{+0.18}_{-0.17}$	$9.59^{+0.08}_{-0.08}$	(...)	(...)	(...)
OFG10	53.191348	−27.737554	$27.54^{+0.23}_{-0.19}$	$24.25^{+0.11}_{-0.10}$	(...)	$5.04^{+0.34}_{-0.31}$	$10.37^{+0.06}_{-0.08}$	(...)	(...)	(...)
OFG11	53.196569 (†)	−27.757065 (†)	$27.44^{+0.35}_{-0.27}$	$23.59^{+0.11}_{-0.10}$	0.62 ± 0.12	$3.60^{+0.27}_{-0.27}$	$10.41^{+0.06}_{-0.07}$	12.41 ± 0.09	384 ± 90	A2GS40, GDS48885
OFG12	53.154787	−27.806529	$27.36^{+0.17}_{-0.15}$	$24.43^{+0.11}_{-0.10}$	(...)	$5.55^{+0.29}_{-0.28}$	$10.31^{+0.05}_{-0.06}$	(...)	(...)	(...)
OFG13	53.047834	−27.829186	$27.34^{+0.24}_{-0.20}$	$25.07^{+0.11}_{-0.10}$	(...)	$3.56^{+0.09}_{-0.09}$	$9.50^{+0.05}_{-0.03}$	(...)	(...)	(...)
OFG14	53.105489	−27.830711	$27.34^{+0.12}_{-0.11}$	$24.38^{+0.11}_{-0.10}$	(...)	$3.39^{+0.92}_{-0.57}$	$9.86^{+0.23}_{-0.22}$	(...)	(...)	(...)
OFG15	53.132675	−27.765496	$27.08^{+0.16}_{-0.14}$	$24.23^{+0.11}_{-0.10}$	(...)	3.192^{sp}	$9.96^{+0.04}_{-0.05}$	(...)	(...)	(...)
OFG16	53.080379	−27.869420	$26.93^{+0.29}_{-0.23}$	$23.58^{+0.11}_{-0.10}$	(...)	$3.69^{+0.27}_{-0.25}$	$10.38^{+0.08}_{-0.08}$	(...)	(...)	(...)
OFG17	53.062276	−27.875036	$26.84^{+0.15}_{-0.13}$	$23.93^{+0.11}_{-0.10}$	(...)	$4.23^{+0.23}_{-0.21}$	$10.29^{+0.07}_{-0.05}$	(...)	(...)	(...)
OFG18	53.188278 (†)	−27.801928 (†)	$26.82^{+0.14}_{-0.12}$	$24.88^{+0.11}_{-0.10}$	0.37 ± 0.11	$3.81^{+0.11}_{-0.10}$	$9.44^{+0.03}_{-0.11}$	12.28 ± 0.11	286 ± 86	A2GS47
OFG19	53.162978 (†)	−27.841940 (†)	$26.76^{+0.20}_{-0.17}$	$24.14^{+0.11}_{-0.10}$	0.40 ± 0.12	$4.09^{+0.35}_{-0.32}$	$10.30^{+0.06}_{-0.09}$	12.26 ± 0.09	271 ± 59	A2GS82
OFG20	53.064807 (†)	−27.862613 (†)	$26.64^{+0.27}_{-0.22}$	$22.85^{+0.11}_{-0.10}$	0.54 ± 0.10	$4.74^{+0.42}_{-0.50}$	$10.88^{+0.16}_{-0.26}$	12.79 ± 0.04	913 ± 86	A2GS57
OFG21	53.060144	−27.793838	$26.59^{+0.13}_{-0.12}$	$24.68^{+0.11}_{-0.10}$	(...)	$5.89^{+0.36}_{-0.38}$	$9.75^{+0.32}_{-0.13}$	(...)	(...)	(...)
OFG22	53.043745	−27.804347	$26.51^{+0.19}_{-0.16}$	$24.56^{+0.12}_{-0.11}$	(...)	$4.56^{+0.26}_{-0.24}$	$9.97^{+0.07}_{-0.07}$	(...)	(...)	(...)
OFG23	53.081890	−27.828815	$26.45^{+0.11}_{-0.10}$	$23.83^{+0.11}_{-0.10}$	(...)	$3.88^{+0.18}_{-0.18}$	$10.25^{+0.04}_{-0.04}$	(...)	(...)	(...)
OFG24	53.109771	−27.807466	$26.40^{+0.11}_{-0.10}$	$25.07^{+0.11}_{-0.10}$	(...)	$6.27^{+0.16}_{-0.16}$	$9.87^{+0.10}_{-0.08}$	(...)	(...)	(...)
OFG25	53.074868 (†)	−27.875889 (†)	$26.31^{+0.30}_{-0.24}$	$23.10^{+0.11}_{-0.10}$	1.67 ± 0.10	3.47^{o}	$9.99^{+0.46}_{-0.11}$	12.92 ± 0.03	1227 ± 74	AGS15, A2GS10
OFG26	53.207252	−27.791408	$26.29^{+0.22}_{-0.19}$	$23.63^{+0.11}_{-0.10}$	(...)	$4.16^{+0.25}_{-0.25}$	$10.33^{+0.08}_{-0.09}$	(...)	(...)	(...)
OFG27 (*)	53.079416 (†)	−27.870820 (†)	(...)	$21.53^{+0.11}_{-0.10}$	2.05 ± 0.12	3.467^{sp}	$11.11^{+0.15}_{-0.19}$	13.08 ± 0.02	1795 ± 90	AGS17, A2GS7
OFG28 (**)	53.120402 (†)	−27.742111 (†)	(...)	(...)	0.95 ± 0.12	(...)	(...)	(...)	(...)	A2GS33

Notes. (1) Source ID; (2)(3) Right ascension and declination (J2000) of sources. Coordinates detected in the ALMA 1.13 mm image are marked with a "(†)" exponent; (4)(5) H-band and IRAC 4.5 μm AB magnitudes. These magnitudes are given for the best-fitting Sérsic profile during de-blending procedure (see Sect. 3.3); (6) ALMA 1.13 mm flux density: obtained from the GOODS-ALMA 2.0 catalog (Gómez-Guijarro et al. 2022a); (7) Photometric redshifts: determined with the EAzY code (see Sect. 4.1; spectroscopic redshifts expressed in three decimal places and flagged with a "sp" exponent). The spectroscopic redshifts of OFG15 and OFG27 are from Herenz et al. (2017) and Zhou et al. (2020), respectively. OFG1, 2, 25, 27 were discovered in an overdensity region with the redshift peak to be 3.47 (flagged with a "o" exponent; Zhou et al. 2020); (8) Stellar masses: determined with the FAST++ (see Sect. 4.1); (9) Infrared luminosities: derived from CIGALE for three sources (OFG2, OFG20, and OFG27) with a *Herschel* counterpart or from the IR template library (Schreiber et al. 2018c) for the galaxies without a *Herschel* counterpart (see Sect. 4.2); (10) SFR = $\mathrm{SFR_{IR}} + \mathrm{SFR_{UV}}$ (see Sect. 4.3); (11) Source IDs in other work: AGS (GOODS-ALMA 1.0 catalog; Franco et al. 2018, 2020a); A2GS (GOODS-ALMA 2.0 catalog; Gómez-Guijarro et al. 2022a); GDS (H-dropouts catalog; Wang et al. 2019). (*)The missing galaxy in our selection approach is due to incompleteness of the search radius of 0.6″. We add it to our catalog to correct this incompleteness. (**)OFG28 is a candidate IRAC 4.5 μm dropout, which is only detected in longer wavelength images, e.g., 850 μm from the JCMT/SCUBA-2 and 870 μm from the ALMA (ID68 in Cowie et al. 2018), 1.13 mm from the GOODS-ALMA 2.0 100% pure source catalog (A2GS33 in Gómez-Guijarro et al. 2022a), and 1.2 mm from the ALMA (ID20 in Yamaguchi et al. 2019). We add it here to refine our OFG catalog in the GOODS-ALMA field. We note that OFG28 is not used in our analysis. We also note that OFG9, OFG13, OFG18, OFG22, and OFG24 are identified as LBGs (see Sect. 3.5.1). In this catalog, eight sources (OFG1-OFG7, and OFG27) are H-dropouts, which are not detected in the H-band ($<3\sigma$).

$$z\sim 6: (i_{775}-z_{850}>1.0)\wedge(Y_{105}-H_{160}<1.0)\wedge (i_{775}-z_{850}>0.777(Y_{105}-H_{160})+1.0),$$

$$z\sim 7: (z_{850}-Y_{105}>0.7)\wedge(J_{125}-H_{160}<0.45)\wedge (z_{850}-Y_{105}>0.8(J_{125}-H_{160})+0.7),$$

here, $\wedge$ and $\vee$ represent the logical AND and OR symbols, respectively. The $\chi^2_{\mathrm{UV_{225},UV_{275}}}=\Sigma_{\mathrm{i}}\,\mathrm{SGN}(f_{\mathrm{i}})(f_{\mathrm{i}}/\sigma_{\mathrm{i}})^2$, where f_{i} (σ_{i}) is the flux (uncertainty) in the i-band and $\mathrm{SGN}(f_{\mathrm{i}})$ is equal to 1 if $f_{\mathrm{i}}>0$ and −1 if $f_{\mathrm{i}}<0$. As in Bouwens et al. (2015), we use a 1σ upper limit as the flux in the dropout band in the case of non-detection. The selected sources are required to be undetected ($<2\sigma$) in all bands blueward of the Lyman break and detected ($>3\sigma$) in all of the above bands redward of the break. We note that we do not include U_{336}, UV_{225}, and UV_{275} in our work because our OFGs are undetected in these bands. Even if we only use the criteria of $z\sim 3$: $(B_{435}-V_{606}<1.2)\wedge(i_{775}-Y_{105}<0.7)$, no galaxy at $z\sim 3$ in our catalog is classified as LBG. The color criteria for Lyman-break of Bouwens et al. (2020) are slightly different from those of Bouwens et al. (2012a, 2015), thus we also used the criteria of Bouwens et al. (2012a, 2015) to select LBGs. All three methods identified the same 5 LBGs in our OFG catalog: OFG9, OFG13, and OFG18 at $z\sim 4$; OFG22 at $z\sim 5$; and OFG24 at $z\sim 6$.

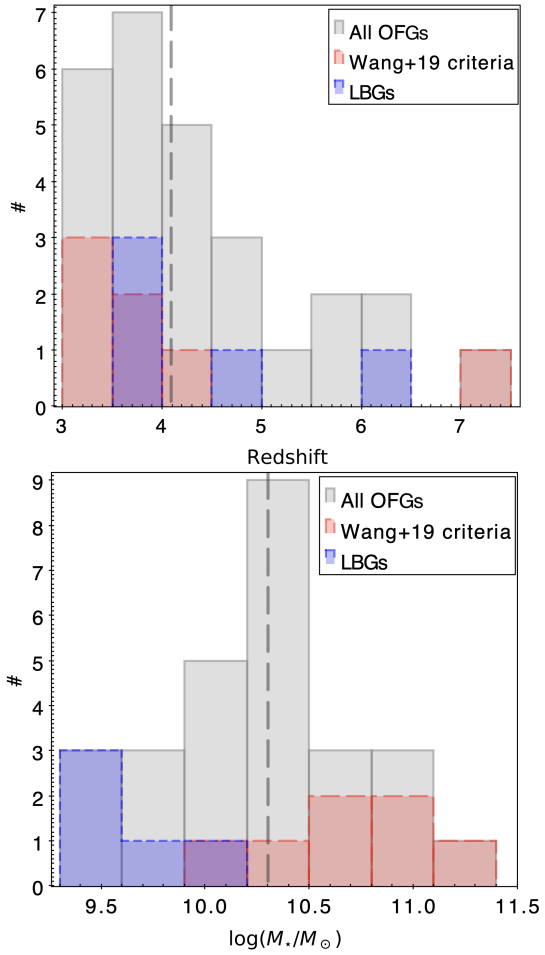

Fig. 4. Photometric redshift and stellar mass distributions for our total 27 OFGs (grey). Among them, five OFGs are identified as LBGs (blue; see Sect. 3.5.1) and seven OFGs meet the criteria of H-dropouts ($H >$ 27 mag & [4.5] < 24 mag) in Wang et al. (2019; red; see Sect. 3.5.2). *Top*: photometric redshifts of our OFGs (including two sources with spectroscopic redshifts). *Bottom*: LBGs cover the low stellar mass end and H-dropouts cover the high stellar mass end. Our OFGs have a wide range of stellar masses with $\log(M_\star/M_\odot) = 9.4–11.1$. The median redshift and stellar mass of the OFGs (dashed lines) are $z_{\rm med} = 4.1$ and $\log(M_{\star\rm med}/M_\odot) = 10.3$, respectively.

3.5.2. H-dropouts

Galaxies bright in IRAC but not detected in the H-band are commonly referred to as H-dropouts. However, this definition can be confusing since different fields have been observed at different depths in the H-band. Here, we used the deepest H-band image to date in the GOODS-South field (HLF; Whitaker et al. 2019), with a 5σ point-source depth of approximately 27.0–29.8 mag, to identify H-dropouts and to extend the sample to our more general definition of OFGs. In our sample, eight galaxies (OFG1-OFG7 and OFG27) are classified as H-dropouts, that is, there is no detection above 3σ in the H-band (see Table 2). We note that in Wang et al. (2019), H-dropouts include all sources with no H-band flux above 5σ, that is, $H > 27$ mag – instead of 3σ here – and [4.5] < 24 mag. If we apply the same criterion, we find seven OFGs that meet this definition: OFG1, OFG2, OFG3, OFG5, OFG6, OFG11, and OFG27 (see Fig. 3).

In the GOODS-ALMA 2.0 catalog, Gómez-Guijarro et al. (2022a) reported 13 OFGs out of 88 galaxies detected above 3.5σ at 1.13 mm in the GOODS-ALMA field. Among them, 12 OFGs are included in our sample (see Table 2 with a "†" exponent). The remaining one (A2GS2 in Gómez-Guijarro et al. 2022a or AGS4 in Franco et al. 2018, 2020a; Zhou et al. 2020) does not meet our criterion of $H > 26.5$ mag, since after applying our de-blending procedure (Sect. 3.3), we measured an H-band magnitude of $H = 24.76$ mag. This value is consistent with the findings of Zhou et al. (2020), who measured $H = 25.23$ mag.

4. Properties of individual galaxies

In this section, we focus on a set of properties of individual galaxies: redshift, stellar mass, infrared luminosity, star formation rate, gas mass, dust mass, and dust temperature. The methodologies used to derive these properties are also used for the stacked samples, as described in Sect. 5.

4.1. Redshifts and stellar masses

We fit the SED from the UV to MIR (rest-frame UV to NIR) to measure the photometric redshifts ($z_{\rm phot}$) and stellar masses ($M_\star$). The photometric redshifts were determined with the code `EAzY`[2] (Brammer et al. 2008). Then we fixed the redshift and derived $M_\star$ with the code `FAST++`[3]. The setup is described below.

Photometric redshifts were obtained with the galaxy template set "`eazy_v1.3`", which includes, in particular, a dusty starburst model to account for extremely dusty galaxies. We did not apply the redshift prior based on K-band magnitudes, as this prior is based on models that do not reproduce high-redshift mass functions (see discussion in Schreiber et al. 2018a).

Two sources, OFG15 and OFG27, have spectroscopic redshifts ($z_{\rm spec}$) confirmed by one-line detections. The galaxy OFG15 has a $z_{\rm spec}$ measured by the Lyman α line ($\sim 7\sigma$) from the MUSE-Wide survey (ID139013229 in Herenz et al. 2017; Urrutia et al. 2019). The galaxy OFG27 has a $z_{\rm spec}$ identified from the CO(6-5) line detection ($\sim 10\sigma$) with ALMA (AGS17 in Zhou et al. 2020). In addition, OFG1, 2, 25, and 27 were discovered in an overdensity region with a peak redshift of $z_{\rm overdensity} = 3.47$ (AGS24, 11, 15, 17 in Zhou et al. 2020, where they were studied in detail). For the above five galaxies, we used their $z_{\rm spec}$ or $z_{\rm overdensity}$ in the following analysis.

Stellar masses were then derived using the code `FAST++`, assuming a delayed, exponentially declining star formation history (SFH), with Bruzual & Charlot (2003) stellar population models and a Calzetti et al. (2000) dust attenuation law. The parameters used in `FAST++` are shown in Table 3.

In Fig. 4, we show the distributions of derived redshifts and stellar masses of our OFGs. The redshift distribution confirms that our OFGs exhibit redshifts of $z_{\rm phot} > 3$, which are consistent with the theoretical galaxy templates (see Fig. 2). The median redshift of the distribution is $z_{\rm med} = 4.1$. Compared to the LBGs covering the low stellar mass end and the H-dropouts in Wang et al. (2019) covering the high stellar mass end, our sample presents a broad distribution of stellar masses with $\log(M_\star/M_\odot) = 9.4–11.1$ and a median value of $\log(M_{\star\rm med}/M_\odot) = 10.3$. The individual redshift and stellar mass values are listed in Table 2, and the individual SEDs are presented in Figs. C.1 and C.2.

We investigate in Fig. 5, the proportions of LBGs, H-dropouts, and remaining OFGs (after removing LBGs and

[2] `EAzY`: `https://github.com/gbrammer/eazy-photoz`
[3] `FAST++`: `https://github.com/cschreib/fastpp`, an updated version of the SED fitting code `FAST` (Kriek et al. 2009).

Table 3. Input parameters used in the UV to 8 μm SED fitting procedures with FAST++ to derive stellar masses.

Parameter	Value
	Delayed SFH
Age [log(yr^{-1})]	6.0–10.2, step 0.1
τ [log(yr^{-1})]	6.5–11, step 0.1
Metallicity	0.02 (solar)
	Dust attenuation: Calzetti et al. (2000)
A_v	0–6, step 0.02

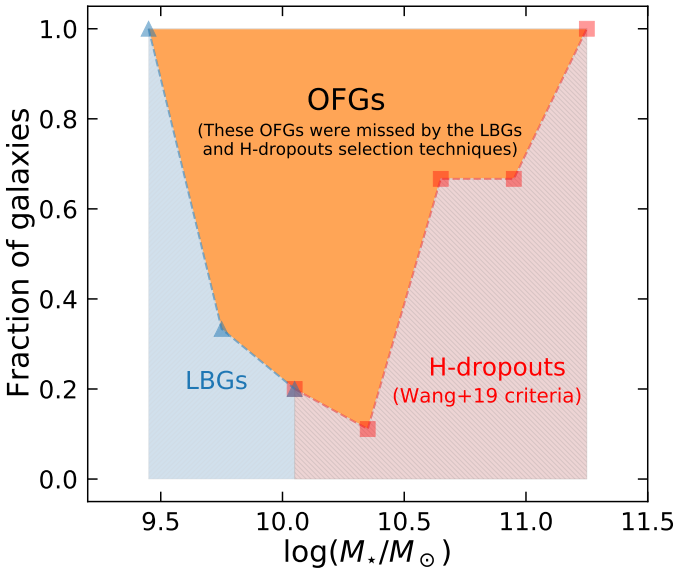

Fig. 5. Number fractions of galaxies in our OFG sample that are identified as LBGs (blue shaded region), H-dropouts (red shaded region), and the remaining OFGs (orange filled region) that are missed by the first two selection techniques. The OFGs undetected by the LBGs and H-dropouts criteria represent approximately 75% at stellar masses between $\log(M_\star/M_\odot) = 9.5–10.5$.

Table 4. Input parameters used in the 24 μm to mm SED fitting procedures with CIGALE.

Parameter	Value
	Dust emission: Draine et al. (2014)
$q_{\rm PAH}$	0.47, 1.12, 1.77, 2.50, 3.19, 3.90, 4.58, 5.26, 5.95, 6.63, 7.32
$U_{\rm min}$	0.1, 0.5, 1, 5, 12, 15, 20, 25, 30, 35, 40, 50
α	1, 1.5, 2, 2.5, 3.0
γ	0.0001, 0.001, 0.01, 0.1, 0.5, 1

H-dropouts) in our sample at different stellar masses. At stellar masses of $\log(M_\star/M_\odot) = 9.5–10.5$, the fraction of OFGs is about three times the sum of LBGs and H-dropouts. In other words, up to 75% of the galaxies with a stellar mass of $\log(M_\star/M_\odot) = 9.5–10.5$ at $z > 3$ are missed by the previous LBG and H-dropout selection techniques.

4.2. Infrared SED fitting

The infrared luminosities (8–1000 μm; $L_{\rm IR}$) and dust mass ($M_{\rm dust}$) are derived from the FIR SED-fitting. We used two methodologies in the fit, depending on whether the galaxies have a *Herschel* counterpart or not (see Figs. C.1 and C.2 for the entire sample).

For galaxies with a *Herschel* counterpart (3/27), we performed the FIR SED-fitting with CIGALE[4] (Code Investigating GALaxy Emission; Burgarella et al. 2005; Noll et al. 2009; Boquien et al. 2019). We fit data from 24 μm up to millimeter wavelengths from the catalogs of Wang et al. (in prep.), Elbaz et al. (2011), and GOODS-ALMA v2.0 1.13 mm (Gómez-Guijarro et al. 2022a). The dust infrared emission model is the one of Draine et al. (2014). The parameters used in CIGALE are shown in Table 4.

For galaxies without a *Herschel* counterpart but with ALMA detections at 1.13 mm (8/27; Gómez-Guijarro et al. 2022a), applying the dust emission model (Draine et al. 2014) of CIGALE would fit a single point in the FIR with four parameters, which would leave us with much less meaningful results. As a compromise, we used the IR template library[5] described in Schreiber et al. (2018c). In brief, it consists of two ingredients: *i*) dust continuum created by big dust grains (silicate + amorphous carbon grains) and *ii*) mid-infrared features contributed by polycyclic aromatic hydrocarbon (PAH) molecules. To form a full dust spectrum, the relative contribution of big grains and PAHs are characterized by the mid-to-total infrared color (IR8 = $L_{\rm IR}/L_8$), which is related to the redshift and starburstiness of a galaxy. The starburstiness is defined by $R_{\rm SB} \equiv \Delta{\rm MS} \equiv {\rm SFR/SFR_{MS}}$ (Elbaz et al. 2011), where $\rm SFR_{MS}$ is the average SFR of the main sequence galaxies presented in Schreiber et al. (2015).

We fit the data iteratively with two templates: the star formation main sequence (MS) template and the starburst (SB) template, with fixed $R_{\rm SB}$ values of 1 and 5, respectively. For each template, given the known redshift and fixed $R_{\rm SB}$, we calculated the dust temperature ($T_{\rm dust}$) and IR8 using Eqs. (18) and (19) from Schreiber et al. (2018c). The templates we used were normalized to $M_{\rm dust} = 1\,M_\odot$. After re-normalizing the SED to the ALMA flux density at 1.13 mm, we obtained the total $M_{\rm dust}$ and total $L_{\rm IR}$ by integrating the SED in the 8–1000 μm rest-frame range. Then we computed $R_{\rm SB}$ using the output $L_{\rm IR}$ (Kennicutt & Evans 2012; the contribution of UV to the SFR is negligible, as shown later in Sect. 5.3.1). We computed two $R_{\rm SB}$ values for each galaxy derived from both templates (MS and SB). If both $R_{\rm SB}$ values are less (greater) than 3, we consider a galaxy a MS (SB) galaxy. The best-fit SED was then generated with the typical value $R_{\rm SB} = 1$ (MS) or $R_{\rm SB} = 5$ (SB). Otherwise, that is, if both templates do not agree with each other for $R_{\rm SB} < 3$ or >3, we kept the two SEDs given by both templates as upper and lower limits and used the average template as the best SED. This approach is similar to the one used by Gómez-Guijarro et al. (2022b) but slightly more conservative. Compared to the $L_{\rm IR}$ in Gómez-Guijarro et al. (2022b) for the same sources, the results are generally consistent, with a median relative difference of $(L_{\rm IR}^{\rm This\,work} - L_{\rm IR}^{\rm G22})/L_{\rm IR}^{\rm G22} = 0.04 \pm 0.18$. The relatively large dispersion was expected because the IR template fit is based on only one observed data point of ALMA 1.13 mm with large uncertainty in $z_{\rm phot}$.

As a consistency test, for the three sources detected by *Herschel* and ALMA, we performed a SED fit using the IR template library normalized only to the ALMA point as if they had no *Herschel* values. We then derived $L_{\rm IR}$. The ratios of the IR luminosities, $L_{\rm IR}^{\rm CIGALE}/L_{\rm IR}^{\rm IR\,template}$ (where $L_{\rm IR}^{\rm IR\,template}$ is derived

[4] CIGALE: https://cigale.lam.fr
[5] S17 library: http://cschreib.github.io/s17-irlib/

with only one photometric point), are 0.49, 2.95, and 1.95 for OFG2, OFG20, and OFG27, respectively. The sample is obviously statistically limited but we do not find a systematic offset when using only one photometric point.

One caveat for the $M_{\rm dust}$ estimates from CIGALE or the IR template library is that they are based on different dust models. Compared to the more standard dust models of Draine et al. (2014; an updated version of Draine & Li 2007) that we adopted in CIGALE, the one used in the IR template library of Schreiber et al. (2018c) assumes that the carbonated grains are amorphous carbon grains rather than graphites. Schreiber et al. (2018c) stated that different dust grain species from the IR template library have different emissivities, systematically lowering the derived $M_{\rm dust}$ by a factor of about two. Therefore, to have comparable $M_{\rm dust}$ for galaxies with and without a *Herschel* counterpart, we have corrected the differences in $M_{\rm dust}$ obtained using the IR template library in Table 5 and also in the following sections.

4.3. SFRs

The total SFR was measured from the contributions of dust-obscured star formation ($\rm SFR_{IR}$) and unobscured star formation ($\rm SFR_{UV}$). The $\rm SFR_{IR}$ was calculated based on the total infrared luminosity ($L_{\rm IR}$), derived from integrating the best-fitted SED between 8 and 1000 µm in the rest frame, following Kennicutt & Evans (2012). The $\rm SFR_{UV}$ was derived from the luminosity emitted in the UV ($L_{\rm UV}$), which was not corrected for dust attenuation, following Daddi et al. (2004) (scaled to a Chabrier 2003 IMF). We calculated the total SFR:

$$\begin{aligned}\rm SFR_{tot}\ [M_\odot\,yr^{-1}] &= \rm SFR_{IR} + SFR_{UV}\\ &= 1.49\times10^{-10}L_{\rm IR} + 1.27\times10^{-10}L_{\rm UV},\end{aligned} \quad (2)$$

both $L_{\rm IR}$ and $L_{\rm UV}$ in units of $L_\odot$, and

$$L_{\rm IR}\ [L_\odot] = 4\pi D_L^2 \int_{8\,\mu\rm m}^{1000\,\mu\rm m} F_\nu(\lambda)\times\frac{c}{\lambda^2}{\rm d}\lambda, \quad (3)$$

$$L_{\rm UV}\ [L_\odot] = 4\pi D_L^2 \frac{\nu_{1500}}{(1+z)}\frac{10^{-0.4(48.6+m_{1500})}}{3.826\times10^{33}}, \quad (4)$$

where D_L is the luminosity distance (cm), ν_{1500} is the frequency (Hz) corresponding to the rest-frame wavelength 1500 Å, and m_{1500} is the AB magnitude at the rest-frame 1500 Å. Here, the value for m_{1500} was derived from the best-fitting templates using EAzY, with a top-hat filter centered at 1500 Å and a width of 350 Å.

For individual OFGs with a *Herschel* and/or ALMA counterpart (and therefore $\rm SFR_{IR}$), we present their SFR ($\rm SFR_{tot} = SFR_{UV}+SFR_{IR}$) in Table 2. Further discussion will be provided later in Sect. 5.3.1 for the stacked samples.

4.4. Molecular gas mass

The gas mass, $M_{\rm gas}$, can be determined from $M_{\rm dust}$ by employing the gas-to-dust ratio ($\delta_{\rm GDR}$) with a metallicity dependency (e.g., Magdis et al. 2012):

$$M_{\rm gas} = M_{\rm H_2} + M_{\rm H_I} = \delta_{\rm GDR} M_{\rm dust}, \quad (5)$$

$$\log(\delta_{\rm GDR}) = (10.54\pm1.0) - (0.99\pm0.12)\times(12+\log({\rm O/H})). \quad (6)$$

Table 5. Dust and gas properties of the OFGs.

ID	$\log(M_{\rm dust})$ $\log(M_\odot)$	$T_{\rm dust}$ (K)	$\log(M_{\rm gas})$ $\log(M_\odot)$	$f_{\rm gas}$	$\tau_{\rm dep}$ (Myr)
(1)	(2)	(3)	(4)	(5)	(6)
OFG1	$8.49^{+0.05}_{-0.06}$	(...)	$10.64^{+0.05}_{-0.06}$	$0.42^{+0.08}_{-0.08}$	115 ± 20
OFG2	$8.80^{+0.10}_{-0.13}$	$37.0^{+1.7}_{-1.6}$	$10.96^{+0.10}_{-0.13}$	$0.61^{+0.09}_{-0.10}$	249 ± 67
OFG4	$8.04^{+0.07}_{-0.09}$	(...)	$10.39^{+0.07}_{-0.09}$	$0.38^{+0.11}_{-0.14}$	46 ± 12
OFG5	$8.09^{+0.06}_{-0.07}$	(...)	$10.49^{+0.06}_{-0.07}$	$0.67^{+0.26}_{-0.11}$	114 ± 31
OFG7	$8.37^{+0.04}_{-0.04}$	(...)	$10.75^{+0.04}_{-0.04}$	$0.72^{+0.05}_{-0.04}$	52 ± 7
OFG11	$8.28^{+0.06}_{-0.07}$	(...)	$10.57^{+0.06}_{-0.07}$	$0.59^{+0.05}_{-0.05}$	97 ± 27
OFG18	$7.96^{+0.11}_{-0.16}$	(...)	$10.71^{+0.11}_{-0.16}$	$0.95^{+0.02}_{-0.02}$	180 ± 77
OFG19	$8.02^{+0.06}_{-0.06}$	(...)	$10.39^{+0.06}_{-0.06}$	$0.55^{+0.05}_{-0.06}$	90 ± 23
OFG20	$8.40^{+0.04}_{-0.05}$	$68.5^{+4.2}_{-3.7}$	$10.59^{+0.04}_{-0.05}$	$0.34^{+0.10}_{-0.10}$	43 ± 6
OFG25	$8.66^{+0.03}_{-0.03}$	(...)	$11.11^{+0.03}_{-0.03}$	$0.93^{+0.12}_{-0.02}$	105 ± 9
OFG27	$9.01^{+0.04}_{-0.05}$	$51.4^{+2.2}_{-2.1}$	$11.07^{+0.04}_{-0.05}$	$0.48^{+0.11}_{-0.09}$	66 ± 8

Notes. (1) Source ID; (2) Dust mass obtained from CIGALE for the galaxies with a *Herschel* counterpart or from the IR template library (Schreiber et al. 2018c) for the galaxies without a *Herschel* counterpart but with an ALMA counterpart (see Sect. 4.2). Since the dust emissivity used in Schreiber et al. (2018c) is different from the one of Draine et al. (2014; used in CIGALE), resulting in a systematic twice lower $M_{\rm dust}$, here we have corrected the $M_{\rm dust}$ derived from the IR template library by multiplying by two. (3) Dust temperature obtained from a single temperature MBB model for the galaxies with a *Herschel* counterpart (see Sect. 4.5); (4) Gas mass obtained from the metallicity-dependent gas-to-dust mass ratio technique (see Sect. 4.4); (5) Gas fraction: $f_{\rm gas} = M_{\rm gas}/(M_\star+M_{\rm gas})$; (6) Gas depletion time: $\tau_{\rm dep} = M_{\rm gas}/(\rm SFR_{IR}+SFR_{UV})$, which is the inverse of the star formation efficiency (SFE = $1/\tau_{\rm dep}$).

The metallicity was determined from the redshift-dependent mass-metallicity relation (MZR; Genzel et al. 2015):

$$12+\log({\rm O/H}) = a - 0.087[\log(M_*)-b]^2, \quad (7)$$

where $a = 8.74$ and $b = 10.4+4.46\times\log(1+z)-1.78\times\log(1+z)^2$. We adopted an uncertainty of 0.2 dex in the metallicities (Magdis et al. 2012).

With the estimates of $M_{\rm gas}$, we can calculate the gas fraction ($f_{\rm gas}$) and gas depletion time ($\tau_{\rm dep}$) as $f_{\rm gas} = M_{\rm gas}/(M_\star + M_{\rm gas})$ and $\tau_{\rm dep} = M_{\rm gas}/(\rm SFR_{IR}+SFR_{UV})$. The $\tau_{\rm dep}$ is the inverse of the star formation efficiency (SFE = $1/\tau_{\rm dep}$). The $M_{\rm gas}$, $f_{\rm gas}$, and $\tau_{\rm dep}$ for the individual sources are presented in Table 5. We underline that only three OFGs have a *Herschel* counterpart, and the rest have $M_{\rm dust}$ only based on ALMA 1.13 mm. Thus, there is a large uncertainty in the values of $M_{\rm dust}$ and, consequently, $M_{\rm gas}$, $f_{\rm gas}$, and $\tau_{\rm dep}$. Therefore, this paper does not go deeper into the $M_{\rm dust}$ and gas properties of individual galaxies. Instead, for the study of gas properties of the OFGs, we focus on the stacked sample, described in Sect. 5.4.

4.5. Dust temperatures

For the comparison with previous studies, we measured the effective dust temperatures ($T_{\rm dust}$) by fitting single-temperature modified black-body (MBB) models to the FIR to mm photometry of the individual galaxies with a *Herschel* counterpart, following:

$$S_\nu \propto \frac{\nu^{3+\beta}}{\exp(\frac{h\nu}{k_{\rm B}T_{\rm dust}})-1}, \quad (8)$$

under the assumption of optically thin dust, where S_ν is the flux density, $k_{\rm B}$ is Boltzmann constant, h is the Planck constant,

and β is the dust emissivity index. We assumed $\beta = 1.5$, a typical value for dusty star-forming galaxies (e.g., Hildebrand 1983; Kovács et al. 2006; Gordon et al. 2010). We note that changing β does not have a significant effect on $T_{\rm dust}$, as β is affecting the slope of the Rayleigh-Jeans (RJ) tail of the dust emission at the rest-frame $\lambda \geq 200\,\mu$m, while the peak of the dust SED is what determines $T_{\rm dust}$ (e.g., Casey 2012; Jin et al. 2019).

Following the criteria used in Hwang et al. (2010; also used in Franco et al. 2020b; Gómez-Guijarro et al. 2022b), we only fit the observed data points at $\lambda \geq 0.55\lambda_{\rm peak}$ to avoid contamination from small dust grains, polycyclic aromatic hydrocarbon (PAH) molecules, and/or AGNs in the MIR, where $\lambda_{\rm peak}$ is the peak of IR SEDs of the CIGALE best fit. The fitted galaxies should satisfy the following conditions: (i) at least one data point at $0.55\lambda_{\rm peak} \leq \lambda < \lambda_{\rm peak}$; and (ii) at least one data point at $\lambda > \lambda_{\rm peak}$ (to exclude the synchrotron contribution from radio data). In our case, we finally fitted the photometry from *Herschel*/SPIRE bands (250 μm, 350 μm, and 500 μm) and ALMA 1.13 mm. We note that we did not consider the CMB effect in the MBB fit because of the lack of data points in the RJ-tail where the CMB plays an important role (e.g., Jin et al. 2019).

The MBB fit was performed using a Markov Chain Monte Carlo (MCMC) approach with 12000 iterations using the Python package PyMC3[6]. The derived $T_{\rm dust}$ for individual galaxies with a *Herschel* counterpart are listed in Table 5. For individual galaxies (three galaxies with *Herschel* counterparts were fitted with MBB), their $T_{\rm dust}$ exhibit a large dispersion, that is, $T_{\rm dust} = 37-69$ K, and only one ($T_{\rm dust} = 51 \pm 2$ K) is in agreement with the expected value from the redshift evolution in MS galaxies from the literature (Schreiber et al. 2018c). The fitting results are also shown in Figs. C.1 and C.2. Considering that only three OFGs have $T_{\rm dust}$ values, we do not go deeper into the discussion of individual galaxies. Instead, we further discuss the stacked OFG samples in Sect. 5.4.

5. Properties of the stacked OFGs

5.1. Stacking analysis

We performed a stacking analysis to study the global properties of our sample. Considering that OFGs are very faint in the optical/NIR ($H > 26.5$ mag) and only 11/27 have *Herschel* and/or ALMA counterparts, performing a stacking analysis helps improving the accuracy of the median photometric redshift and SFR measurements. To build the SED of our sample, we generated a median and mean stacked image in each filter, from the optical to 1.13 mm. Specifically, we used images from the HST/ACS ($F435W$, $F606W$, $F775W$, $F814W$, $F850LP$), HST/WFC3 ($F105W$, $F125W$, $F140W$, $F160W$), ZFOURGE $K_{\rm s}$-band, *Spitzer*/IRAC (3.6, 4.5, 5.8, and 8 μm), *Spitzer*/MIPS (24 μm), *Herschel*/PACS&SPIRE (100, 160, 250, 350 and 500 μm), and ALMA 1.13 mm maps.

The photometry was obtained mostly using aperture photometry techniques, except for the *Herschel* bands, where appropriate aperture corrections were applied to account for flux losses outside the aperture. This procedure is very similar to that used previously in the deep surveys, which we summarize here. In the HST/ACS and HST/WFC3 bands, fluxes were extracted on the PSF-matched images (to the F160W) using the same aperture of 0.7″-diameter as in Whitaker et al. (2019), which maximizes the signal-to-noise ratio (S/N) of the resulting aperture photometry. In the $K_{\rm s}$-band, we used a 1.2″ diameter circular aperture to measure flux on the ZFOURGE $K_{\rm s}$-band image following Straatman et al. (2016), whose PSF was matched to a Moffat profile with $FWHM = 0.9''$. In the IRAC bands, fluxes were extracted separately without PSF matching due to the broader PSFs. We adopted a 2.2″ diameter aperture to maximize the S/N of the resulting aperture photometry. In the MIPS 24 μm band, we used a large aperture of 6″ in diameter corresponding to its full width at half maximum. At 1.13 mm, we used a diameter of 1.6″ to measure the flux, which is the optimal trade-off between total flux and SNR.

Uncertainties on the photometry were derived from the Monte Carlo simulations. For each band, we carried out the same stacking analysis as above, but at random positions, and measured the flux value on the stacked image. This was repeated 1000 times. We then calculated the 16th and 84th percentiles of the distribution of values as flux uncertainties.

For the *Herschel*/PACS and SPIRE bands, we used the PSF fitting with a free background to fit the stacked image following Schreiber et al. (2015). The uncertainties were obtained using the following methods: 1) a bootstrap approach; specifically, as an example, we generated a sample of 27 sources from 27 OFGs, allowing the same galaxy to be picked repeatedly, and measured the stacked flux. This procedure was repeated 100 times, and we calculated their standard deviation as the flux uncertainty; and 2) a Monte Carlo simulation approach, which is the same as that used for lower wavelength images and 1.13 mm images. Here we adopted a $0.9 \times FWHM$ diameter circular aperture. We note that the results given from the bootstrap approach include the uncertainties from i) the PSF fitting, ii) the clustering bias effect, and iii) background fluctuation. Thus, the derived values of uncertainties from bootstrap are larger than those from the Monte Carlo simulation. We conservatively take the former values as our uncertainties.

5.2. Fitting of the stacked SEDs

We obtained the stacked full-wavelength SEDs in the same way as for the individual galaxies (see Sects. 4.1, 4.2, Tables 3, 4). In brief, first, we fitted the broad photometry at OPT to MIR with the EAzY code to obtain photometric redshifts. Then, we independently performed the OPT to MIR SED fitting with FAST++ and the MIR to mm SED fitting with CIGALE, respectively, at the previously obtained redshifts of the stacked sources. This approach helps to 1) disentangle the degeneracy between redshift and other parameters, such as stellar age and dust temperature; and 2) break the energy balance principle (the total energy emitted in the MIR and FIR is determined by the attenuation of observed starlight in the UV and optical) used in CIGALE. For dusty star-forming galaxies, especially for H-dropouts and $K_{\rm s}$-dropouts with strong dust obscuration, there could exist regions with strong UV extinction due to strong dust obscuration, which may not participate in the UV to optical part, but emit FIR light (e.g., Simpson et al. 2015; Gómez-Guijarro et al. 2018; Elbaz et al. 2018). Assuming an energy balance with a fixed redshift will lead to an underestimation of the $L_{\rm IR}$, hence, the SFR.

The best-fit SED is shown in Fig. 6. The median redshift for the total sample is $z_{\rm med,stacked} = 4.5 \pm 0.2$, which is consistent with $z_{\rm med} = 4.1$ derived from the median value of individual OFGs with a wide distribution (see Fig. 4, top). In addition, the median stacked SED peak (and the mean stacked SED peak; see Fig. 7) is between 350 and 500 μm, also in agreement with being at $z \sim 4$. Thus, these agreements confirm that the bulk population of OFGs consists of dusty star-forming galaxies at $z \sim 4-5$. Remarkably, most fluxes in the stacked images are above the 3σ

[6] PyMC3 is available at: https://docs.pymc.io/en/v3/

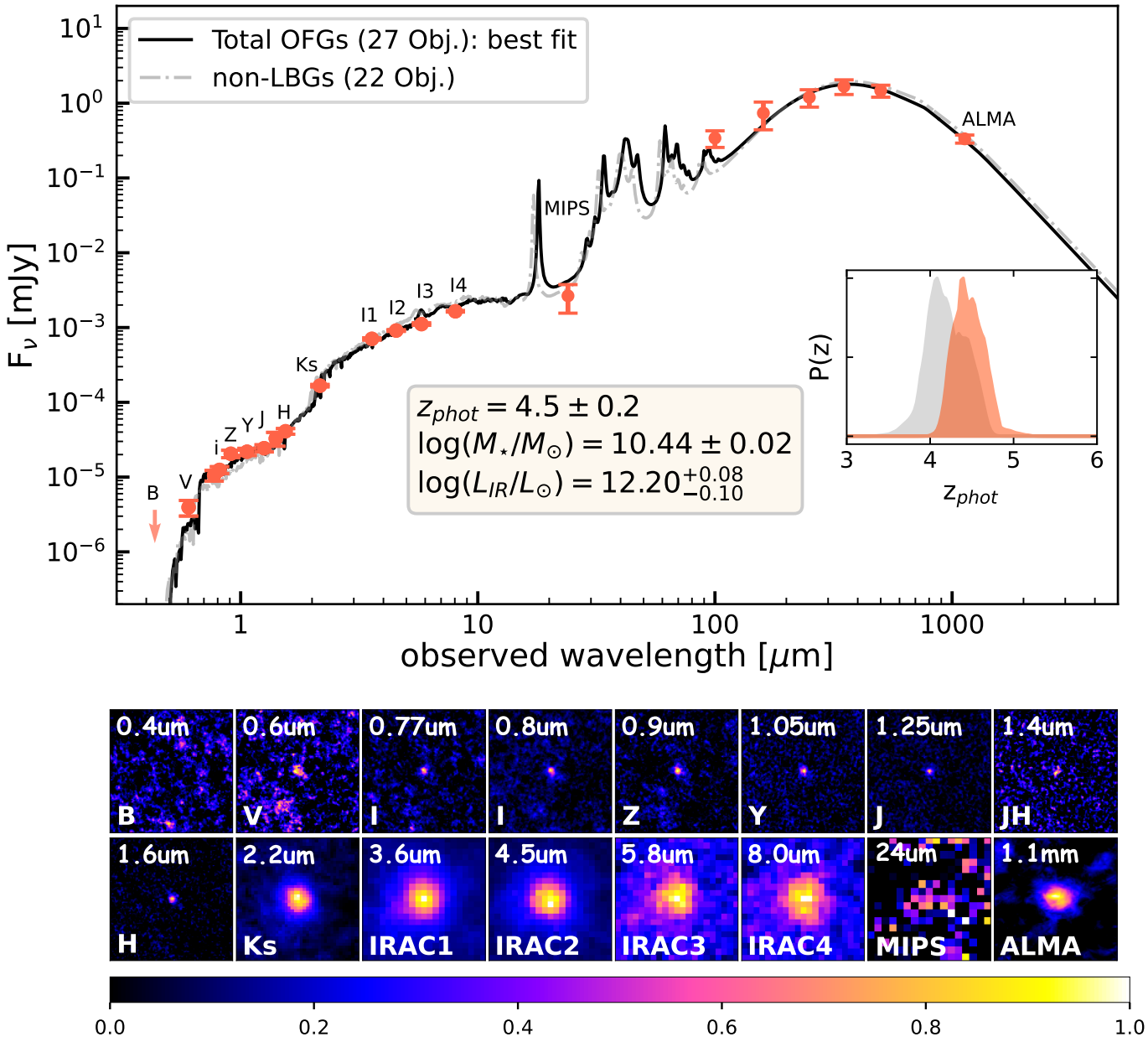

Fig. 6. Median stacked SED and images of the total sample of 27 OFGs in this work. *Top*: best-fit SED of the total sample (black line). The measured fluxes (red points) are derived from the stacked images. Error bars (1σ) and upper limits (3σ) are obtained from the Monte Carlo simulation (except *Herschel*) and bootstrap approach (*Herschel*; see Sect. 5.2). We also show the best-fit SED for 22 non-LBGs (grey line). These 22 non-LBGs will be used to calculate the cosmic SFRD. The inset shows the likelihood distributions of the photometric redshift of our samples (total sample in red, 22 non-LBGs in grey), based on the UV to MIR SED fitting from `EAzY`, which is normalized to the peak value. The redshift obtained from the maximized likelihood is $z \sim 4.5$ for the total 27 OFGs and $z \sim 4.2$ for the 22 non-LBGs. *Bottom*: stacked images of the total sample with peak fluxes normalized. Each panel is $6'' \times 6''$ except for the MIPS 24 µm, which is $24'' \times 24''$.

confidence level, especially in the H, K_s, and IRAC bands, helping to establish the position of the Balmer and 4000 Å breaks very well, hence determining a robust redshift. The median properties derived from the stacked SED for the total sample are summarized in Table 6.

In addition, to further investigate the characteristics of different subpopulations of our OFGs, we performed median and mean stacked SED fitting for four sub-samples of OFGs. The four subsamples and the purpose of our investigation are: 1) OFGs that are LBGs: in our sample, five OFGs are classified as LBGs. Given that the traditional approach to estimating the cosmic SFRD at $z > 3$ is mainly based on the LBGs (e.g., Madau & Dickinson 2014; Bouwens et al. 2020), studying the differences in the properties of LBGs and OFGs (after removing 5 LBGs) in our sample can help us understand the importance of OFGs in the cosmic SFRD; 2) OFGs that are not LBGs: 22 OFGs are not classified as LBGs. To quantity the level of underestimation of the cosmic SFRD at $z > 3$ (Madau & Dickinson 2014) based mainly on LBGs, we used the mean SFR of this sub-sample to calculate the SFRD to avoid contamination by LBGs (describe later in Sect. 6.1); 3) Massive OFGs that are not LBGs: there are 16 OFGs not classified as LBGs with $\log(M_\star/M_\odot) > 10.3$. To compare with the results of Wang et al. (2019) on the SFRD, here we

Table 6. Median physical properties of the total sample of 27 OFGs.

Derived from SED fitting with median stacked photometry [(*)]		
$z_{\rm phot}$ [(a)]		4.5 ± 0.2
$M_\star$ [(b)]	$M_\odot$	$(2.8^{+0.2}_{-0.1}) \times 10^{10}$
$L_{\rm IR}$ [(c)]	$L_\odot$	$(1.6 \pm 0.3) \times 10^{12}$
$A_{\rm V}$ [(b)]	mag	$0.9^{+0.3}_{-0.0}$
$M_{\rm dust}$ [(c)]	$M_\odot$	$(1.2 \pm 0.2) \times 10^8$
$T_{\rm dust}$ [(d)]	K	$45.5^{+2.1}_{-2.1}$
Median stacked photometry		
H	mag	27.4 ± 0.1
[4.5]	mag	23.92 ± 0.04
$S_{1.13\,\rm mm}$	μJy	334 ± 24
$S_{3\,\rm GHz}$	μJy	4.1 ± 0.7
Derived quantities [(**)]		
$L_{1.4\,\rm GHz}$ [(e)]	erg s^{-1} Hz^{-1}	$(9.53 \pm 1.53) \times 10^{30}$
$q_{\rm TIR}$ [(f)]		2.23 ± 0.03
$\rm SFR_{rad,med}$ [(g)]	$M_\odot$ yr^{-1}	287.59 ± 53.86
$\rm SFR_{IR,med}$ [(h)]	$M_\odot$ yr^{-1}	$235.33^{+47.37}_{-49.77}$
$\rm SFR_{UV,med}$ [(i)]	$M_\odot$ yr^{-1}	0.33 ± 0.02
ΔMS [(j)]		$1.45^{+0.29}_{-0.31}$
$M_{\rm gas}$ [(k)]	$M_\odot$	$(2.6 \pm 0.4) \times 10^{10}$
$f_{\rm gas}$ [(l)]		0.48 ± 0.05
$\tau_{\rm dep}$ [(m)]	Myr	110^{+29}_{-30}

Notes. [(*)]Uncertainties are the 16–84th percentile ranges of the probability distribution function given by the SED fitting. [(**)]Uncertainties on derived quantities were calculated from the propagation of the errors in the parameter values. [(a)]Photometric redshift, determined with the code EAzY. [(b)]$M_\star$ and $A_{\rm V}$, derived from the UV to MIR SED fitting with the code FAST++. [(c)]Given by IR SED fitting with CIGALE. [(d)]Measured by MBB model fit (see Sect. 4.5). [(e)]Derived from $S_{3\,\rm GHz}$ assuming a radio spectral index $\alpha = -0.75$ (see Eq. (10)). [(f)]Calculated from the IR-radio correlation (see Eq. (12)). [(g)]Calculated following Delhaize et al. (2017; see Eq. (9)), which was simply estimated from the radio emission without correction for AGN. [(h)]Derived following Kennicutt & Evans (2012; see Eq. 2). [(i)]Derived following Daddi et al. (2004), scaled to a Chabrier (2003) IMF (see Eq. (2)). [(j)]Distance to the SFMS: ΔMS = SFR/SFR$_{\rm MS}$, where SFR$_{\rm MS}$ is the average SFR of MS galaxies at fixed stellar mass and redshift (Schreiber et al. 2015, see Fig. 9 and Sect. 5.4). [(k)]$M_{\rm gas}$, computed based on gas-to-dust ratio. [(l)]Gas fraction: $f_{\rm gas} = M_{\rm gas}/(M_\star + M_{\rm gas})$. [(m)]Gas depletion time: $\tau_{\rm dep} = M_{\rm gas}/(\rm SFR_{IR} + SFR_{UV})$, which is the inverse of the star formation efficiency (SFE = $1/\tau_{\rm dep}$).

used the same stellar mass cut for this sub-sample; and 4) ALMA-detected OFGs: 11 OFGs in our sample are detected by ALMA at 1.13 mm (>3.5σ; Gómez-Guijarro et al. 2022a). Since using ALMA detections to select OF sources is a very efficient method (e.g., Franco et al. 2020a; Zhou et al. 2020; Gómez-Guijarro et al. 2022a), we study the properties of the ALMA-detected OFGs and compare them with other sub-samples to understand whether there is a selection bias using this approach.

The best-fit mean SEDs of the total sample and the four sub-samples are shown in Fig. 7. For the OFGs that are LBGs (Fig. 7b), there is a 3.3σ detection at 24 μm, no detection in all the *Herschel* bands, and a 1.2σ detection at 1.13 mm. To successfully perform the FIR SED fitting, we fit the fluxes at 24 μm and 1.13 mm. We then compared the best-fit model with the 3σ upper limits in the *Herschel* bands (red arrows in Fig. 7b). The best-fit model is below the red arrows, showing a good consistency with the *Herschel* no detections. The mean properties derived from the mean stacked SEDs for the sub-samples are summarized in Table 7.

For the mean stacked SED fit, one hypothesis here is that all OFGs have a similar SED shape. This is because the stacked SED we used to derive the SFR is a flux-weighted average in each band and if the brightest galaxy has a different SED shape, then the fitting results will be biased towards the properties of the brightest galaxy. For example, it has been shown that the dust temperature derived from the mean stacked SED is biased by 1.5 K to a higher temperature than the true value since the starburst galaxies in the sample are warmer and brighter (Schreiber et al. 2018c). However, we do not yet know the true IR SED shapes of most OFGs in our sample because of the lack of the *Herschel* detections. It is also unclear whether the brightest OFGs have a different SED shape compared to the remaining OFGs, therefore causing a bias. Hence, we cannot correct this potential bias here. Instead, we performed a median stacked SED fitting as a comparison. Although the median stacked SED fitting exhibits a lower confidence level (because it is less influenced by the brightest sources) compared to the mean one, their properties are more robust against outliers and are representative of the vast majority of galaxies in the sample. On the other hand, and most importantly, we need the SFR derived from the mean stacked SED fitting to calculate the cosmic SFRD (described later in Sect. 6.1).

5.3. SFRs and AGN

5.3.1. SFRs

We obtained the SFR$_{\rm tot}$ of the stacked OFG (sub)samples using the same method as for individual galaxies (see Sect. 4.3), following Eq. (2). With the 3 GHz VLA observations in the GOODS-South, we can also calculate the radio-based SFR (SFR$_{\rm rad}$; assuming a Chabrier 2003 IMF) following Delhaize et al. (2017):

$$\rm SFR_{rad}\,[M_\odot\,yr^{-1}] = 10^{-24} \times 10^{q_{TIR}(z)}\,L_{1.4\,GHz}\,[W\,Hz^{-1}], \quad (9)$$

where $L_{1.4\,\rm GHz}$ is the rest-frame 1.4 GHz luminosity converted from the 3 GHz flux density ($S_{3\,\rm GHz}$ at observed-frame; W m^{-2} Hz^{-1}) using:

$$L_{1.4\,\rm GHz}\,[\rm W\,Hz^{-1}] = \frac{4\pi D_L^2}{(1+z)^{\alpha+1}} \left(\frac{1.4\ [\rm GHz]}{3\ [\rm GHz]}\right)^{\alpha} S_{3\,\rm GHz}, \quad (10)$$

here, the radio spectral index $\alpha^{3\,\rm GHz}_{1.4\,\rm GHz}$ is assumed to be $\alpha = -0.75$. The $q_{\rm TIR}(z)$ in Eq. (9) is the IR-to-radio luminosity ratio, which was recently found to evolve primarily with the stellar mass and depend secondarily on the redshift (Delvecchio et al. 2021):

$$q_{\rm TIR}(M_\star, z) = (2.646 \pm 0.024) \times (1+z)^{-0.023 \pm 0.008} - (0.148 \pm 0.013) \times (\log M_\star/M_\odot - 10). \quad (11)$$

The derived IR-based and radio-based SFR values (SFR$_{\rm IR}$ and SFR$_{\rm rad}$) are in good agreement (except for the sub-sample of OFGs that are LBGs), as shown in Tables 6 and 7. For the sub-sample of OFGs that are LBGs, the SFR$_{\rm rad}$ is about four times higher than the SFR$_{\rm IR}$, although with large uncertainty, hinting at the existence of radio AGNs in the sub-sample of OFGs that are LBGs. The median SFRs for our total 27 OFGs are given in Table 6, while the mean SFRs for our OF sub-samples are summarized in Table 7. For our entire sample, the median contribution from SFR$_{\rm UV}$ to SFR$_{\rm tot}$ is only 0.1%, which is negligible.

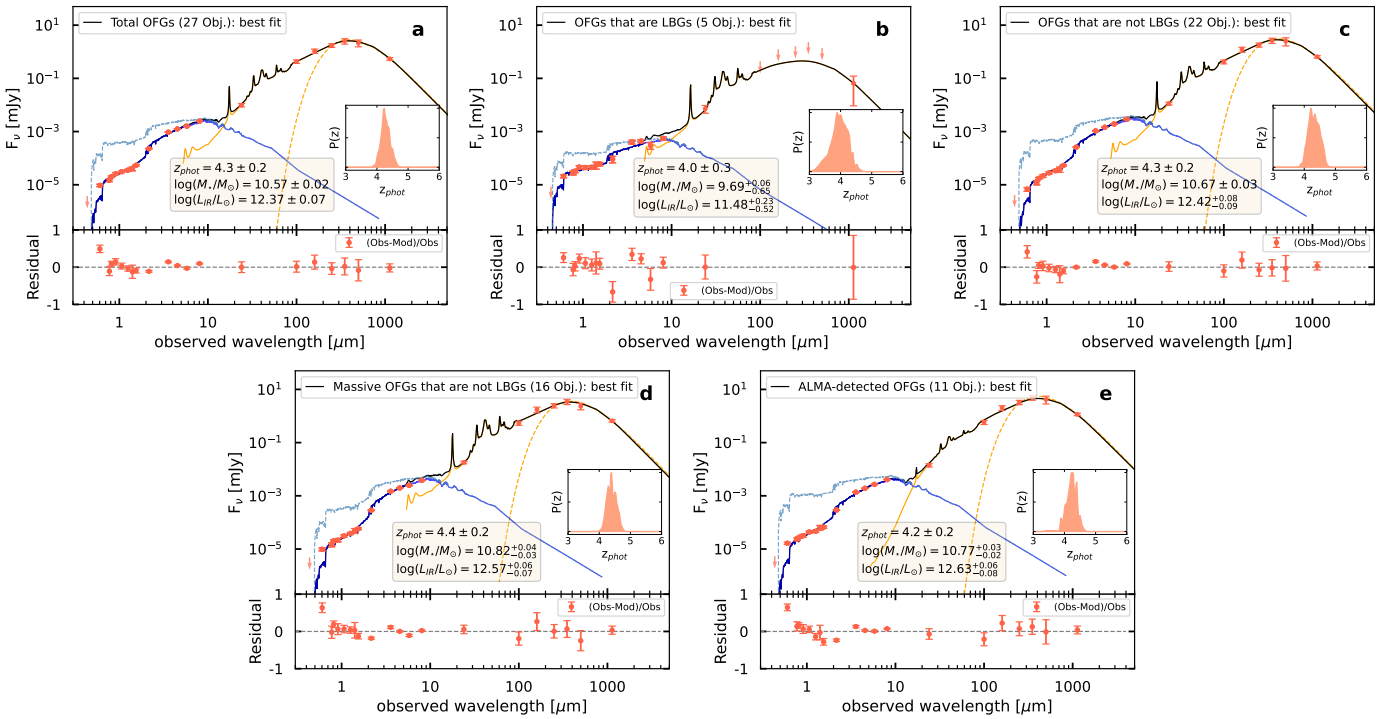

Fig. 7. Mean stacked SEDs for the total sample and four sub-samples of our OFGs. *Panel a*: total sample of 27 OFGs; (*b*)–(*e*): sub-samples of OFGs that are LBGs, OFGs that are not LBGs, massive OFGs ($\log(M_\star/M_\odot) > 10.3$) that are not LBGs, and OFGs with ALMA detections, respectively. We use the sub-samples in *panel c* and *panel d* to calculate the SFRD. *Top*: best-fit SED of the stacked OFGs (black line), which is composed of the uncorrected dust-attenuated stellar component (dark blue line) and the IR dust contribution (orange solid line). The corrected UV emission is shown as a faint blue line. We also plot the best fit of a MBB (orange dashed line) with a fixed dust emissivity index $\beta = 1.5$. We note that there is no significant AGN contribution in our sample, except for the five OFGs that are LBGs displaying the presence of radio AGNs (see Sect. 5.3.2). The measured fluxes (red points) are derived from the stacked images. Error bars (1σ) and upper limits (3σ) are obtained from the Monte Carlo simulation (except *Herschel*) and bootstrap approach (*Herschel*; see Sect. 5.2). The inset shows the likelihood distribution of the photometric redshift based on the UV to MIR SED fitting from `EAzY` (see Sect. 4.1), which is normalized to the peak value. *Bottom*: residuals calculated from (observation – model)/observation. The SEDs for individual OFGs are given in the appendix (see Figs. C.1 and C.2).

5.3.2. AGN

As our selection criterion has been designed to avoid selecting passive galaxies, the OFGs in our sample are mainly dusty star-forming galaxies. Studying the presence of AGN in our sample can help us understand the co-evolution between AGN and star formation activity in the early Universe. It is also crucial for ensuring that our calculations of the SFR and, eventually, the cosmic SFRD are correct (uncontaminated by AGN). Here, we examine our sample for AGN contributions using three different methods, that is, studying their IR, radio, and X-ray excesses.

First, we fit the stacked IR SED with an additional AGN template using `CIGALE` (Fritz et al. 2006, see Sect. 5.2 for the infrared SED fitting). The contribution of the IR-bright AGN to the total IR luminosity ($f_{\rm AGN}$) can lead to an overestimation of the dust IR emission and thus of the total SFR. We found that the SED fitting yields a $f_{\rm AGN} < 0.01$, indicating the absence of IR-bright AGN in our sample.

Secondly, the $q_{\rm TIR}$ is defined as the IR-to-radio luminosity ratio (e.g., Helou et al. 1985; Yun et al. 2001):

$$q_{\rm TIR} \equiv \log\left(\frac{L_{\rm IR}\,[\rm W]}{3.75\times10^{12}\,[\rm Hz]}\right) - \log\left(L_{1.4\,\rm GHz}\,[\rm W\,Hz^{-1}]\right). \quad (12)$$

The derived $q_{\rm TIR}$ from the IR-radio correlation (Eq. (12)) is presented in Tables 6 and 7. Except for the sub-sample of OFGs that are LBGs, $q_{\rm TIR}$ values of the OFGs are consistent with those in Delvecchio et al. (2021) for star-forming galaxies at the same redshift and stellar mass. These agreements suggest that for the OFGs that are not LBGs, there is a lack of strong AGN activity in the radio band. On the other hand, the sub-sample of OFGs that are LBGs present a mean $q_{\rm TIR} = 1.8$, much smaller than the typical $q_{\rm TIR}$ of 2.6 for star-forming galaxies at the same redshift and stellar mass, and would thus be classified as radio AGNs (Fig. 12 in Delvecchio et al. 2021).

We also searched for X-ray-bright AGN in the CDF-S 7 Ms catalog (Luo et al. 2017). Among 1008 sources in the main catalog and 47 lower-significance sources in the supplementary catalog, we did not find any X-ray counterpart for the individual OFGs within a 0.6″ radius. None of the sources in our catalog exhibit a total X-ray luminosity integrated over the entire 0.5–7 keV range larger than $L_X = 10^{42.5}$ erg s^{-1} (AGN definition in Luo et al. 2017). Hence, we find no evidence for any bright X-ray AGN in our catalog. We also performed mean and median stacking for the 27 OFGs in 0.5–7 keV images and did not find any significant detections ($\ll 3\sigma$) in either of the stacked images.

In addition, we considered the MIR-AGN selection criterion developed by Donley et al. (2012) to diagnose the presence of a power-law AGN based on IRAC colors. However, this criterion does not apply to our high-z OFGs. At $z > 3$, the IRAC bands mainly collect emission from stars below 2 μm in the rest frame, outside the typical domain where power-law AGNs contribute. In summary, we do not find evidence for significant contamination by AGNs in our OFG sample, except for the five OFGs that are LBGs displaying the presence of radio AGNs.

Table 7. Mean physical properties of sub-samples of OFGs.

Parameter	Unit	OFGs that are LBGs (5 Obj.)	OFGs that are not LBGs (22 Obj.)	Massive OFGs that are not LBGs (16 Obj.)	ALMA-detected OFGs (11 Obj.)
Derived from SED fitting with mean stacked photometry					
$z_{\rm phot}$		4.0 ± 0.3	4.3 ± 0.2	4.4 ± 0.2	4.2 ± 0.2
$M_\star$	$M_\odot$	$(4.9^{+0.7}_{-3.8}) \times 10^9$	$(4.7 \pm 0.3) \times 10^{10}$	$(6.6^{+0.6}_{-0.4}) \times 10^{10}$	$(5.9^{+0.4}_{-0.3}) \times 10^{10}$
$L_{\rm IR}$	$L_\odot$	$(2.9 \pm 2.1) \times 10^{11}$	$(2.6^{+0.6}_{-0.5}) \times 10^{12}$	$(3.7 \pm 0.6) \times 10^{12}$	$(4.3^{+0.6}_{-0.7}) \times 10^{12}$
$A_{\rm V}$	mag	$0.7^{+0.4}_{-0.3}$	$1.4^{+0.1}_{-0.3}$	$1.2^{+0.3}_{-0.1}$	$1.7^{+0.1}_{-0.2}$
$M_{\rm dust}$	$M_\odot$	$(0.3 \pm 1.3) \times 10^8$	$(2.9 \pm 0.7) \times 10^8$	$(2.8 \pm 0.6) \times 10^8$	$(5.3 \pm 1.1) \times 10^8$
$T_{\rm dust}$	K	(...)	$42.3^{+1.2}_{-1.2}$	$45.5^{+1.4}_{-1.5}$	$41.5^{+1.2}_{-1.2}$
Mean stacked photometry					
H	mag	26.95 ± 0.17	27.09 ± 0.08	26.99 ± 0.08	26.85 ± 0.10
[4.5]	mag	24.74 ± 0.31	23.34 ± 0.03	23.33 ± 0.02	22.99 ± 0.03
$S_{1.13\,\rm mm}$	μJy	65 ± 56	650 ± 26	660 ± 30	1153 ± 40
$S_{3\,\rm GHz}$	μJy	2.7 ± 0.5	6.5 ± 0.5	6.9 ± 0.5	10.6 ± 0.8
Derived quantities					
$L_{1.4\,\rm GHz}$	erg s^{-1} Hz^{-1}	$(4.81 \pm 0.92) \times 10^{30}$	$(1.40 \pm 0.11) \times 10^{31}$	$(1.54 \pm 0.11) \times 10^{31}$	$(2.18 \pm 0.17) \times 10^{31}$
$q_{\rm TIR}$		$1.80^{+0.15}_{-0.43}$	2.28 ± 0.03	2.39 ± 0.02	2.30 ± 0.03
$\rm SFR_{rad,avg}$	$M_\odot$ yr^{-1}	189.95 ± 58.40	392.17 ± 49.02	408.91 ± 49.54	590.15 ± 74.07
$\rm SFR_{IR,avg}$	$M_\odot$ yr^{-1}	44.52 ± 33.08	$387.93^{+82.03}_{-72.08}$	$556.32^{+84.94}_{-85.59}$	$635.15^{+96.61}_{-104.83}$
$\rm SFR_{UV,avg}$	$M_\odot$ yr^{-1}	0.63 ± 0.07	0.37 ± 0.03	0.45 ± 0.03	0.59 ± 0.04
ΔMS		1.81 ± 1.32	$1.46^{+0.31}_{-0.27}$	1.45 ± 0.22	$1.96^{+0.30}_{-0.32}$
$M_{\rm gas}$	$M_\odot$	$(1.3 \pm 5.7) \times 10^{10}$	$(5.0 \pm 1.2) \times 10^{10}$	$(4.4 \pm 0.9) \times 10^{10}$	$(8.4 \pm 1.8) \times 10^{10}$
$f_{\rm gas}$		0.73 ± 0.87	0.52 ± 0.06	0.40 ± 0.06	0.59 ± 0.05
$\tau_{\rm dep}$	Myr	291 ± 1277	130^{+42}_{-40}	79 ± 21	133^{+34}_{-35}
$R_{e(1.13\,\rm mm)}$	kpc	(...)	1.09 ± 0.05	1.05 ± 0.06	0.80 ± 0.03
$\Sigma_{\rm SFR}$ [(a)]	$M_\odot$ yr^{-1} kpc^{-2}	(...)	52 ± 11	80 ± 15	158 ± 28
Cosmic SFRD[(*)]					
V [(**)]	Mpc^3	(...)	7.4×10^5	7.4×10^5	7.4×10^5
SFRD	$M_\odot$ yr^{-1}Mpc^{-3}	(...)	$(1.2 \pm 0.2) \times 10^{-2}$	$(1.2 \pm 0.2) \times 10^{-2}$	$(0.9 \pm 0.2) \times 10^{-2}$

Notes. Same as Table 6 but for the four sub-samples of OFGs. [(a)]SFR surface density: $\Sigma_{\rm SFR} = 0.5{\rm SFR_{tot}}/(\pi R^2_{e(1.13\,\rm mm)})$, where $\rm SFR_{tot} = SFR_{IR,avg} + SFR_{UV,avg}$. [(*)]The SFRD is discussed with details in Sect. 6. [(**)]Survey volume, calculated using Eq. (16) with a broad redshift range of $z = 3.2$–7.0.

5.4. The main sequence of star-forming galaxies and the properties of gas and dust for the stacked OFGs

In this section, we investigate the properties of the stacked samples derived from the SED fitting. We examine their locations in the star-formation main sequence, their gas depletion timescales, gas fractions, and dust temperatures in the framework of the scaling relations for galaxy evolution, and their dust sizes.

In Fig. 8, we place the stacked total OFG sample and the four sub-samples in the SFR-$M_\star$ plane, showing the locations compared to the SFMS. In the SFR-$M_\star$ plane, it is well known that the SFMS as a whole moves to higher SFRs with increasing redshift (e.g., Elbaz et al. 2007, 2011; Noeske et al. 2007; Magdis et al. 2010; Whitaker et al. 2012, 2014; Speagle et al. 2014; Schreiber et al. 2015; Lee et al. 2015; Leslie et al. 2020). We adopted a fixed $z = 4$ for the SFMS (Schreiber et al. 2015) as a comparison since it is close to $z_{\rm med} = 4.1$ from the median value of the individual OFGs and $z_{\rm med,stacked} = 4.5$ from the median stacked total OFGs. This figure shows that all the (sub-)samples are located within the SFMS region (∼0.5 dex) at $z = 4$, and most of them lie within the 1σ scatter of the SFMS ($0.5 < \Delta{\rm MS} < 2$, i.e., ∼0.3 dex), consistent with being normal star-forming galaxies at the same redshift. It suggests that unlike studies limited to a rare population of extreme starburst galaxies (e.g., Riechers et al. 2013; Strandet et al. 2017; Marrone et al. 2018; Dudzevičiūtė et al. 2020; Riechers et al. 2020), our OFGs represent a less extreme population of dusty star-forming galaxies at $z > 3$.

Furthermore, we study the gas and dust properties of the OFGs by focusing on their gas depletion timescales, gas fractions, dust temperatures, and dust sizes. In Fig. 9, we show the normalized $\tau_{\rm dep}$, $f_{\rm gas}$, and $T_{\rm dust}$ by scaling them to the observed relation (scl; which is the median of the MS) of $\tau_{\rm dep,scl}(z, M_\star, \Delta{\rm MS})$ and $f_{\rm gas,scl}(z, M_\star, \Delta{\rm MS})$ from Tacconi et al. (2018) and of $T_{\rm dust,scl}(z, \Delta{\rm MS})$ from Schreiber et al. (2018c) as a function of ΔMS and $M_\star$. The ΔMS is the SFR of each stacked (sub)sample normalized by the SFR of the SFMS (ΔMS = SFR/$\rm SFR_{MS}$; Schreiber et al. 2015) at its own redshift and stellar mass. The $\tau_{\rm dep,scl}$, $f_{\rm gas,scl}$, and $T_{\rm dust,scl}$ are calculated for each data point at fixed redshift, stellar mass, and ΔMS. We also present dust continuum sizes at 1.13 mm ($R_{e(1.13\,\rm mm)}$) of the mean stacked OFG (sub)samples as a function of ΔMS and $M_\star$. Here, the half-light radius $R_{e(1.13\,\rm mm)}$ was measured in the *uv* plane by fitting a circular Gaussian (task *uvmodelfit*) after performing *uv* plane stacking according to the method described by Gómez-Guijarro et al. (2022a). We note that we did not scale $R_{e(1.13\,\rm mm)}$ to the observed relations (e.g., van der Wel et al. 2014) because the redshifts ($z > 3$) of our OFGs exceed the limits of these relations.

In Fig. 9, there is no global offset between mean and median results of normalized $\tau_{\rm dep}$, $f_{\rm gas}$, and $T_{\rm dust}$ for the stacked (sub)samples. It indicates no significant differences in the SED shapes of the brightest OFGs compared to the remaining ones, which would otherwise cause a strong bias (as discussed in Sect. 5.2) and further show a global offset even for the different stacked samples. The mean SFR is larger than the median SFR (see Fig. 8), which is expected since the former is influenced by the brightest sources in the flux-weighted average in each band (as discussed in Sect. 5.2).

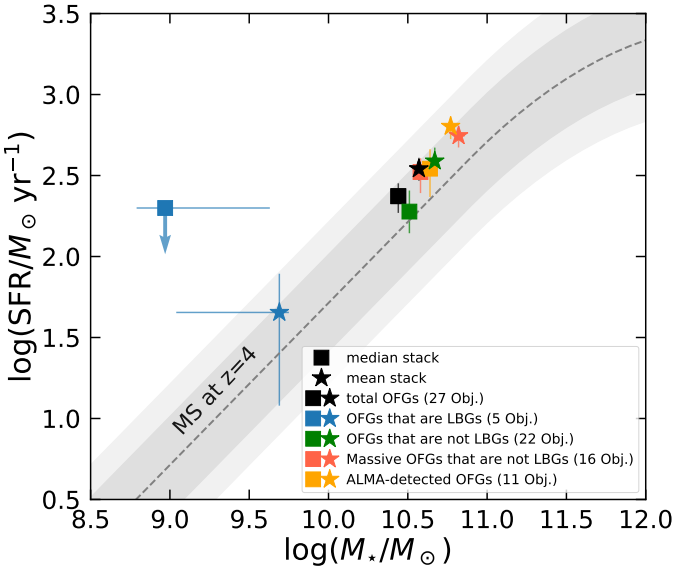

Fig. 8. Locations of OFGs compared to the SFMS in the SFR-$M_\star$ plane. The SFMS at $z = 4$ (Schreiber et al. 2015), 1σ scatter ($0.5 < \Delta\mathrm{MS} < 2$, i.e., ∼0.3 dex), and $\pm3\times\Delta$MS region ($0.33 < \Delta\mathrm{MS} < 3$, i.e., ∼0.5 dex) are highlighted with a grey dashed line, a grey shaded area, and a light grey shaded area, respectively. $\Delta\mathrm{MS} > 3$ is commonly used to separate MS and SB galaxies. Squares and stars respectively represent the median and mean stacking results for our total sample (black) and four sub-samples. The four sub-samples are the OFGs that are LBGs (blue), the OFGs that are not LBGs(green), the massive OFGs that are not LBGs(red), and the OFGs with ALMA detections (orange). When necessary, data from the literature have been converted to a Chabrier (2003) IMF.

In the first and second columns of Fig. 9, the stacked (sub)samples show $\tau_{\rm dep}$ values below the scatter of the scaling relation, while $f_{\rm gas}$ is at the lower boundary of the scaling relation. That is to say, the OFGs have shorter $\tau_{\rm dep}$ and slightly lower $f_{\rm gas}$ values compared to normal star-forming galaxies. This indicates that galaxies with stronger dust obscuration tend to have lower gas fractions and shorter gas depletion times. Their gas is consumed more rapidly, hence, they form their stars with a high efficiency, which sets them in the so-called class of starbursts in the main sequence (Elbaz et al. 2018; Gómez-Guijarro et al. 2022b).

Among all the stacked (sub)samples, the ALMA-detected OFGs have the longest gas depletion timescale and the highest gas fraction. We believe this is due to a selection effect, as galaxies with higher dust content are more easily detected by ALMA at 1.13 mm. The $M_{\rm gas}$ was derived from $M_{\rm dust}$ in our study by employing the gas-to-dust ratio (see Sect. 4.4). Thus, the ALMA-detected galaxies tend to have higher $M_{\rm gas}$ and, consequently, higher values of $\tau_{\rm dep}$ and $f_{\rm gas}$ as well. Furthermore, the SFR is positively correlated with $M_{\rm dust}$ for star-forming galaxies at fixed $T_{\rm dust}$ (e.g., Genzel et al. 2015; Orellana et al. 2017; Donevski et al. 2020). It explains why they show a higher SFR in the SFR-$M_\star$ plane compared to the total stacked OFGs (in Fig. 8). Notably, it raises the caveat that the approach of selecting only ALMA-detected galaxies in studies of OFGs will end up biasing the sample toward larger SFRs, longer $\tau_{\rm dep}$, and larger $f_{\rm gas}$.

In addition, the massive OFGs (excluding LBGs) present the lowest gas fraction and the shortest gas depletion timescale of all stacked (sub)samples in Fig. 9. Yet, we did not see any significant difference in the ΔMS of the massive OFGs compared with the other stacked (sub)samples. This suggests that these galaxies are observed just before becoming passive.

The median $T_{\rm dust} = 46 \pm 2$ K for the stacked total OFGs (see Table 6) is consistent with the scaling relation of $T_{\rm dust}(z, \Delta\mathrm{MS})$ (Schreiber et al. 2018c, black squares in third column of Fig. 9). However, surprisingly, most of the stacked (sub-)samples show slightly colder dust temperatures compared to the scaling relation. In particular, the ALMA-detected OFGs have the most abundant dust but show the lowest $T_{\rm dust}$, indicating that the dust is colder in the more obscured sources ($A_{\rm V} = 1.7^{+0.1}_{-0.2}$ in Table 7). The mean $T_{\rm dust} = 42 \pm 1$ K of the ALMA-detected OFGs (see Table 7) is consistent with the $T_{\rm dust} = 40 \pm 2$ K[7] of the median stacked ALMA-detected massive H-dropouts at $z = 4$ (Wang et al. 2019). The median $T_{\rm dust}$ of the ALMA-detected OFGs is much lower, with $T_{\rm dust} = 33 \pm 3$ K. The low $T_{\rm dust}$ of the ALMA-detected OFGs cannot be explained by current studies (see, e.g., Magnelli et al. 2014; Schreiber et al. 2018c), which suggest that an increasing $T_{\rm dust}$ is correlated to an enhanced specific star formation rate. Furthermore, this is contrary to the findings of Sommovigo et al. (2022), for instance, where the authors conclude that dust is warmer in obscured sources because a larger obscuration leads to more efficient dust heating. However, quite intriguingly, cases of cold dusty star-forming galaxies at high redshifts have already been reported in the literature, such as GN20 at $z = 4.05$ with $T_{\rm dust} = 33$ K (Magdis et al. 2012; Cortzen et al. 2020) and four ALMA-detected sources at 3 mm at $z \sim 5$ (Jin et al. 2019). A possible reason for the cold dust temperature is that the dust emission in the FIR of the dust-obscured sources may be optically thick rather than optically thin, where a warm and compact dust core is hidden (Jin et al. 2019, 2022). Indeed, the compact dust core is shown in the last column of Fig. 9. Among (sub-)samples of our OFGs, the ALMA-detected OFGs with the highest dust obscuration (largest $A_{\rm V}$) present the most compact dust core with a half-light radius $R_{e(1.13\,\mathrm{mm})} = 0.80 \pm 0.03$ kpc. The SFR surface density ($\Sigma_{\rm SFR}$) of the ALMA-detected OFGs is about two to three times higher than the others. This would imply that the measured dust temperature underestimates the actual dust temperature, that would be higher after correcting for the attenuation in the shorter FIR bands. Making this correction is out of the scope of this paper due to the limited information that we have on those galaxies.

5.5. The hidden side of the dust region

An important check is to test whether UV continuum emission alone (after correcting for dust extinction) provides a robust estimate of the total SFR, especially for those highly dust-obscured galaxies. We again used the stacked total sample and four sub-samples of OFGs.

We derived the $\mathrm{SFR_{UV}}$ corrected for dust extinction (i.e., $\mathrm{SFR^{corr}_{UV}}$) using the Calzetti et al. (2000) reddening law and assuming a constant star formation history from the UV to MIR SED fitting with the code FAST++. Specifically, similar to Sect. 4.3, the $\mathrm{SFR^{corr}_{UV}}$ was obtained from the $L^{\rm corr}_{\rm UV}$ following Daddi et al. (2004), which was calculated based on the AB magnitude at the rest-frame 1500 Å (see Eq. (4)). The intrinsic flux was derived with

$$f_{\rm int}(\lambda) = f_{\rm obs}(\lambda)\, 10^{0.4\,A_\lambda}, \tag{13}$$

[7] $T_{\rm dust}$ is derived using the IR template library (Schreiber et al. 2018c). To compare with our results, it has been scaled to the light-weighted dust temperature by applying Eq. (6) in Schreiber et al. (2018c).

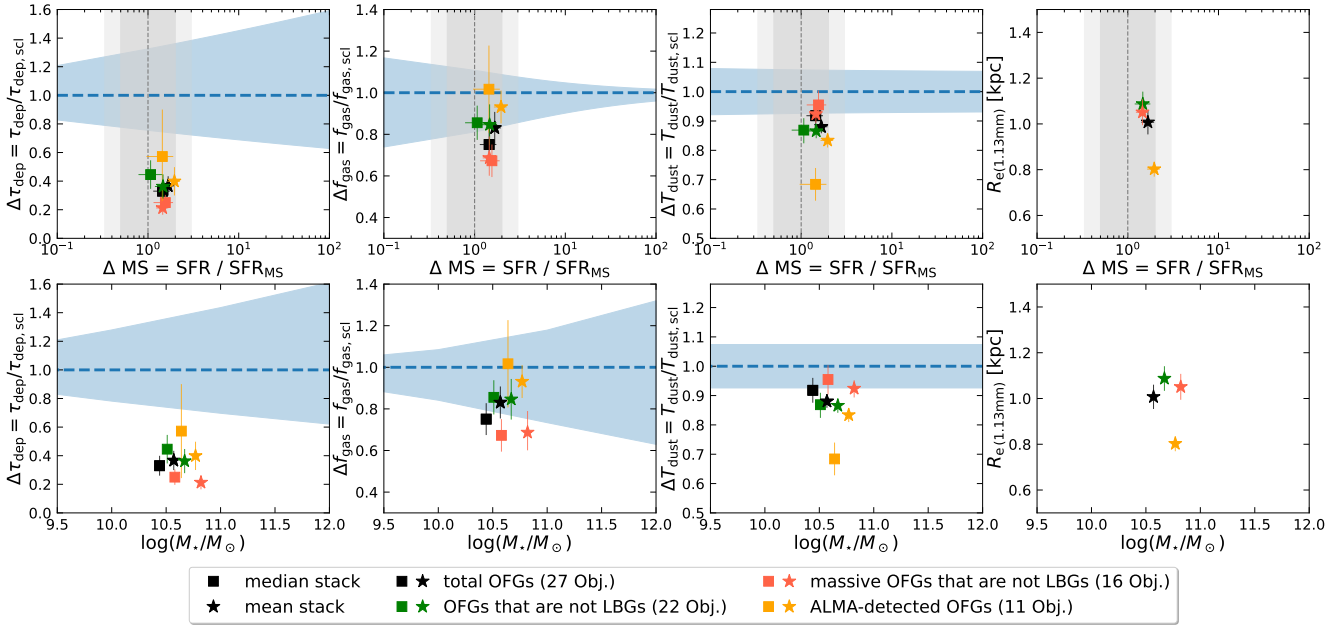

Fig. 9. Normalized $\tau_{\rm dep}$ (*left*), $f_{\rm gas}$ (*mid-left*), and $T_{\rm dust}$ (*mid-right*) by the scaling relation of $\tau_{\rm dep,scl}(z, M_\star, \Delta{\rm MS})$ and $f_{\rm gas,scl}(z, M_\star, \Delta{\rm MS})$ from Tacconi et al. (2018) and of $T_{\rm dust,scl}(z, \Delta{\rm MS})$ from Schreiber et al. (2018c) as well as the half-light radius $R_{\rm e}$ (*right*) at ALMA 1.13 mm as a function of ΔMS (*first row*) and $M_\star$ (*second row*). Here, ΔMS is the distance to the SFMS (Schreiber et al. 2015), $\Delta{\rm MS} = {\rm SFR/SFR_{MS}}$, at its own stellar mass and redshift for each sample. The blue dashed line and shaded area show the scaling relation and scatter. For each data point, the $\tau_{\rm dep,scl}$, $f_{\rm gas,scl}$, and $T_{\rm dust,scl}$ are calculated at a fixed redshift, stellar mass, and ΔMS. Squares, stars, and shaded grey areas are the same as in Fig. 8.

where $f_{\rm int}$ and $f_{\rm obs}$ are the intrinsic and observed fluxes, respectively. The extinction A_λ is related to the reddening curve $k(\lambda)$:

$$A_\lambda = \frac{k(\lambda)\,A_{\rm V}}{R_{\rm V}}. \tag{14}$$

From the Calzetti et al. (2000) reddening law, we have $R_{\rm V} = 4.05$ and

$$K_\lambda = 2.659\left(-2.156 + \frac{1.509}{\lambda} - \frac{0.198}{\lambda^2} + \frac{0.011}{\lambda^3}\right) + R_{\rm V}, \tag{15}$$

with $0.12\,\mu{\rm m} \le \lambda \le 0.63\,\mu{\rm m}$.

We then compared the derived $\rm SFR_{UV}^{corr}$ with the total SFR ($\rm SFR_{tot} = SFR_{IR} + SFR_{UV}$) and calculated their ratios. The results are presented in Table 8. We find that all the stacked (sub-)samples have $\rm SFR_{tot}$ much larger than $\rm SFR_{UV}^{corr}$. In addition, they have similar $\rm SFR_{tot}/SFR_{UV}^{corr}$ ratios within uncertainties, with $\rm SFR_{tot}/SFR_{UV}^{corr} = 8 \pm 1$ for the mean stacked total OFGs. It suggests: (*i*) the existence of a hidden dust region in the OFGs (even for the LBGs) that absorbs all the UV photons, which cannot be reproduced with a dust extinction correction, indicating that the dust emission in these OFGs might be optically thick; and (*ii*) that it is fundamental to include IR/mm band observations when studying extremely dusty star-forming galaxies. Otherwise, the total SFR and, therefore, the cosmic SFRD will be strongly underestimated.

Furthermore, we compared the stacked OFG (sub)samples with the star-forming galaxies from the ZFOURGE catalog (Straatman et al. 2016). These star-forming galaxies are from the CDF-S field with a *Herschel* detection, and are split into two redshift bins ($0 \le z \le 3$ and $3 \le z \le 6$). We calculated their $\rm SFR_{tot}/SFR_{UV}^{corr}$ ratios using the same method as for the stacked OFG (sub)samples. As shown in Fig. 10, the $\rm SFR_{tot}/SFR_{UV}^{corr}$ ratio increases with increasing starburstiness, indicating the presence

Table 8. UV-corrected SFR vs. total SFR for the mean stacked OFGs.

Mean stacked OFGs	$\rm SFR_{UV}^{corr}$ ($M_\odot\,{\rm yr}^{-1}$)	$\rm SFR_{tot}$ ($M_\odot\,{\rm yr}^{-1}$)	$\rm SFR_{tot}/SFR_{UV}^{corr}$
(1)	(2)	(3)	(4)
Total OFGs	46^{+3}_{-4}	348 ± 59	8 ± 1
OFGs that are LBGs	5^{+1}_{-1}	45 ± 33	9 ± 7
OFGs that are not LBGs	57^{+6}_{-4}	388 ± 82	7 ± 2
Massive OFGs that are not LBGs	88^{+9}_{-8}	557 ± 86	6 ± 1
ALMA-detected OFGs	83^{+10}_{-7}	636 ± 105	8± 2

Notes. (1) Mean stacked total sample and four sub-samples of OFGs; (2) $\rm SFR_{UV}$ corrected for dust extinction (see Sect. 5.5); (3) $\rm SFR_{tot} = SFR_{IR} + SFR_{UV}$; (4) Ratio of total SFR to $\rm SFR_{UV}^{corr}$.

of more hidden dust regions in galaxies that are likely to be optically thick. It suggests that using the UV emission alone to determine the total SFR of starburst galaxies, even after dust attenuation correction, could result in strong underestimates, consistent with the findings of Elbaz et al. (2018) and Puglisi et al. (2017). We further find that the strong underestimations appear at both redshift bins, suggesting that this may be a general phenomenon for starburst galaxies, regardless of the redshift. In addition, for MS galaxies ($\Delta{\rm MS} \sim 1$) with $0 \le z \le 3$, their $\rm SFR_{tot}$ and $\rm SFR_{UV}^{corr}$ are very similar, showing that both SFR estimators agree with each other for typical MS galaxies at low redshifts. However, for MS galaxies with $3 \le z \le 6$, their $\rm SFR_{tot}$ is about twice (~0.3 dex) larger than the $\rm SFR_{UV}^{corr}$. Generally, the median $\rm SFR_{tot}/SFR_{UV}^{corr}$ ratio is about twice higher for star-forming galaxies with $3 \le z \le 6$ than those with $0 \le z \le 3$. We note that in Fig. 10, we did not perform a stellar-mass cut for the star-forming galaxies from the ZFOURGE catalog due to the small number of galaxies. As a test, we selected galaxies with $\log(M_\star/M_\odot) > 9$ and obtained similar results but with a larger dispersion of the galaxy distribution because of their

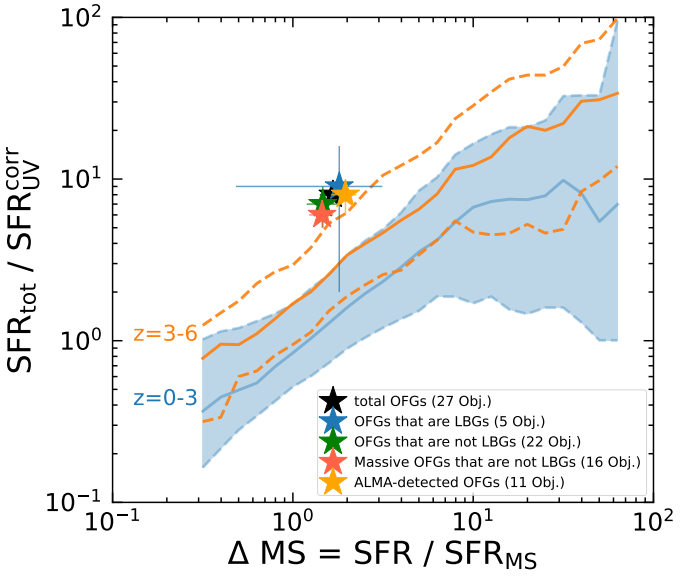

Fig. 10. Ratio of the total SFR (SFR_{tot} = SFR_{IR} + SFR_{UV}) and the SFR_{UV} corrected for the dust extinction (SFR_{UV}^{corr}) as a function of the starburstiness ($\Delta MS = SFR_{tot}/SFR_{MS}$). Stars represent the mean stacked OFG (sub)samples as in Fig. 8. The solid and dashed lines show the sliding median and 16–84th percentile range of star-forming galaxies from the CDF-S field with a *Herschel* detection at $0 \le z \le 3$ (blue) and $3 \le z \le 6$ (orange) from the ZFOURGE catalog (Straatman et al. 2016).

small number. A more detailed study of the stellar masses, and the $SFR_{tot}/SFR_{UV}^{corr}$ ratio is beyond the scope of this paper.

In Fig. 10, the stacked OFG (sub)samples, with $z_{med,stacked}$ = 4.5, lie above the 16–84th percentile range of the star-forming galaxies at $3 \le z \le 6$. This is consistent with the fact that these dusty star-forming galaxies are the more extreme cases (more dust-obscured), with lower dust temperatures compared to typical star-forming galaxies at similar redshifts and with relatively compact dust sizes ($R_{e(1.13\,mm)}$ = 1.01 ± 0.05 kpc; black star in Fig. 9).

6. Cosmic star formation rate density

6.1. Star formation rate density

In this section, we calculate the SFRD in a very conservative way. In our calculation, the derived value would be a lower limit for the OFGs.

As our OFGs were discovered randomly within a blind GOODS-ALMA survey area, we can simply calculate their SFRD by using their total SFR divided by the survey volume. The survey volume is the volume between the shells defined by the redshift range of the sources and within a solid angle:

$$V = \frac{\Omega}{4\pi}(V_{z_1} - V_{z_0}) = \frac{\Omega}{4\pi}\left[\frac{4\pi}{3}(d_{z_1}^3 - d_{z_0}^3)\right] = \frac{\Omega}{3}(d_{z_1}^3 - d_{z_0}^3), \quad (16)$$

where the solid angle Ω corresponding to the effective area of 72.42 arcmin^2, in units of steradian (6.1×10^{-6} sr), of the GOODS-ALMA survey; d_{z_0} and d_{z_1} are the comoving distances at given redshifts of z_0 and z_1. To be conservative, here we use a broad redshift range of z = 3.2–7.0 from the whole OFG sample (see Fig. 4) for our total 22 OFGs that are not LBGs, instead of using the redshift and 1σ confidence interval given by the stacked SED.

The cosmic star-formation-rate density for our total OFG sample is:

$$\psi\,[M_\odot \mathrm{yr}^{-1}\mathrm{Mpc}^{-3}] = \frac{\mathrm{SFR}_{\mathrm{mean}} \times 22}{V}, \quad (17)$$

where the SFR_{mean} is the UV+IR-based SFR given from the mean stacked SED fitting of the 22 OFGs that are not LBGs (Table 7). Assuming that the intrinsic infrared SED of our OFGs is the same as the SED derived from mean stacking, the SFRD of our 22 OFGs (excluding LBGs) reaches approximately $1.2 \times 10^{-2}\,M_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$ with the mean redshift of z_{mean} = 4.3 (and z_{median} = 4.1). The survey volume, V, is 7.4×10^5 Mpc^3. The number density is $n \sim 3 \times 10^{-5}$ Mpc^{-3}, which is slightly higher than the one of massive H-dropouts.

6.2. Incompleteness of our understanding of the cosmic star-formation history

The reference cosmic star-formation history at $z \gtrsim 4$ (Madau & Dickinson 2014) is based on measurements of the LBGs corrected for dust attenuation in the rest-frame UV (Bouwens et al. 2012a,b). The study of OFGs (excluding LBGs) can help us quantify the incompleteness of our understanding of the cosmic star-formation history at high redshifts.

The left panel of Fig. 11 shows the contribution of OFGs (z_{mean} = 4.3) to the cosmic SFRD. At z = 4–5, the contribution of OFGs (red-filled circles) reaches about 52% of the reference SFRD from LBGs (Madau & Dickinson 2014, black solid line with grey open triangles). If we combine the contributions of OFGs (excluding LBGs) and LBGs (Madau & Dickinson 2014) as total SFRD, the OFGs contribute about 34% and the LBGs contribute about 66% of the total SFRD. It shows that the LBGs dominate the total SFRD, while the OFGs (excluding LBGs) make up about a third of the total SFRD.

In addition, we investigate the contribution of massive OFGs to the cosmic SFRD. Out of a total of 22 OFGs (excluding LBGs), 16 massive OFGs have $\log(M_\star/M_\odot) > 10.3$, the same stellar mass cut as the so-called ALMA-detected massive H-dropouts (Wang et al. 2019). The SFRD of massive OFGs is approximately $1.2 \times 10^{-2}\,M_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$ with z_{mean} = 4.4 (and z_{median} = 4.1), shown as the blue-shaded area in the left panel of Fig. 11 (see also Table 7). If we limit the redshift range to z = 4–5, the SFRD of massive OFGs is approximately $1.6 \times 10^{-2}\,M_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}$, as illustrated by the blue-filled circles. This value is consistent with the SFRD of the total 22 OFGs within errors (red-filled circles in Fig. 11). Since the massive and total OFGs have similar SFRD values, it indicates that among the OFGs, the massive ones dominant the SFRD compared to the remaining less massive ones.

We further compare the massive OFGs with equivalently massive LBGs and H-dropouts. Firstly, comparing the massive OFGs with equivalently massive LBGs, the SFRD of massive OFGs is at least two orders of magnitude higher. Thus, we conclude that for massive galaxies ($\log(M_\star/M_\odot) > 10.3$), the cosmic SFRD is dominated by massive OFGs (99%) rather than massive LBGs (1%) at high redshifts. Secondly, the SFRD of massive OFGs is more than four times higher than that of massive H-dropouts detected by ALMA (Wang et al. 2019). We recall that the selection criteria for H-dropouts are $H > 27$ mag & [4.5] < 24 mag (Wang et al. 2019), whereas our OFGs are selected with $H > 26.5$ mag & [4.5] < 25 mag. It implies that the optically faint galaxies that contribute significantly to the SFRD of massive galaxies have been neglected in previous

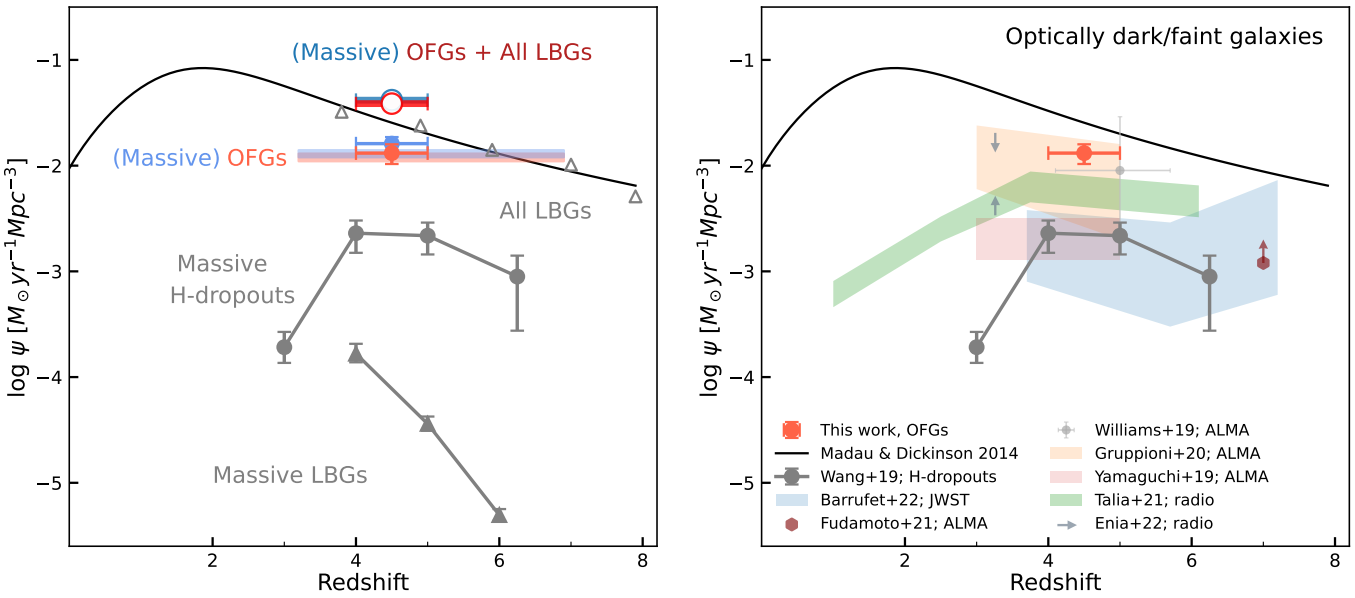

Fig. 11. Cosmic star formation history of the Universe. *Left*: contribution of our OFGs to the cosmic SFRD. The black line is the cosmic SFRD, ψ, as a function of redshift from Madau & Dickinson (2014), which is based on LBGs at $z \gtrsim 4$ with a dust correction applied (grey open triangles; Bouwens et al. 2012a,b). The 22 OFGs that are not LBGs are shaded in red and the 16 massive OFGs that are not LBGs ($\log(M_\star/M_\odot) > 10.3$) are shaded in blue, where the sources at $z = 4-5$ are shown in faint red and blue filled circles, respectively. Grey filled circles are ALMA-detected massive H-dropouts with $\log(M_\star/M_\odot) > 10.3$ (Wang et al. 2019). Grey filled triangles are massive LBGs with $\log(M_\star/M_\odot) > 10.3$ whose SFRD is based on the dust-corrected UV (Wang et al. 2019). We also show the combined contribution of OFGs and LBGs to the cosmic SFRD. The red (blue) open circle indicates the combined contribution of OFGs (massive OFGs) and all LBGs (Madau & Dickinson 2014) at $z = 4-5$. *Right*: summary of the contribution of optically dark or faint galaxies to the cosmic SFRD from the literature. These optically dark or faint galaxies were selected in different methods, including our OFGs (red point; this work), ALMA-detected massive H-dropouts (grey filled circles; Wang et al. 2019), JWST-selected HST-dark galaxies from CEERS (blue area; Barrufet et al. 2022), ALMA-detected optically dark sources from REBELS (brown point and arrow; Fudamoto et al. 2021), an ALMA-detected NIR-dark galaxy in COSMOS (light grey point; Williams et al. 2019), ALMA-detected HST+near-IR dark galaxies from ALPINE (orange area; Gruppioni et al. 2020), ALMA-detected K-dropouts from ASAGAO (red area; Yamaguchi et al. 2019), radio 3 GHz-selected UV-dark galaxies in COSMOS (green area; Talia et al. 2021), and radio 1.4 GHz-selected HST-dark galaxies in GOODS-N (grey arrows; Enia et al. 2022). All data from the literature have been converted to a Chabrier (2003) IMF.

studies of LBGs as well as H-dropouts. Therefore, we emphasize the importance of considering moderately obscured, non-LBG sources in complementing the high-z SFRD measurements, rather than focusing only on those extremely dusty star-forming galaxies.

Table 7 presents the SFRD of the total OFGs, the massive OFGs, and the ALMA-detected OFGs, with LBGs excluded. Table 7 shows that these three (sub)-samples have the same SFRD values within errors. A comparison of the total OFGs and the massive OFGs has already been discussed above and shown in the left panel of Fig. 11. For the ALMA-detected OFGs, their SFRD value is lower than the total OFGs but still within errors. This implies that the investigation of the ALMA-detected sources alone might be sufficient to account for the contribution of OFGs in the SFRD. However, limited by the number of sources, this result needs to be further confirmed with a larger OFG sample in future studies. Besides, we would like to remind that the dust and gas properties of the ALMA-detected OFGs are not representative of all OFGs, as described in Sect. 5.4.

In our work, we find that the OFGs contribute ~52% to the reference SFRD at $z = 4-5$. Many works have also studied optically dark or faint galaxies selected or originally detected with different methodologies. The right panel of Fig. 11 shows a summary of the cosmic SFRD of optically dark or faint galaxies from the literature. Specifically, in Wang et al. (2019), the contribution of the ALMA-detected massive H-dropouts accounts for ~10% of the reference SFRD from LBGs at $z \sim 4-6$. At $z \sim 5$, this value is broadly consistent with that from the ALMA-detected NIR-dark galaxy in COSMOS (Williams et al. 2019) and the ALMA-detected HST+near-IR dark galaxies from ALPINE (Gruppioni et al. 2020). Similar results have also been found for the ALMA-detected K-dropouts from ASAGAO (~10% at $z \sim 3-5$; Yamaguchi et al. 2019) and the JWST-selected HST-dark galaxies from CEERS (~10% at $z \sim 4-6$, but increasing to ~36% at $z \sim 7$; Barrufet et al. 2022). At $z \sim 7$, Fudamoto et al. (2021) identified two optically dark galaxies in two separate REBELS pointings that contribute ~10–25% to the reference SFRD. In radio, the 3 GHz-selected UV-dark galaxies from VLA-COSMOS (Talia et al. 2021) contribute 10–25% at $z \sim 3-4.5$ and 25–40% at $z > 4.5$ to the reference SFRD. More recently, Enia et al. (2022) found a 7–58% contribution from 1.4 GHz radio-selected HST-dark galaxies in GOODS-N at $z \sim 3$. Despite the broad range of contributions from these galaxies, the high values all agree that the contribution of optically dark or faint galaxies to the cosmic SFRD cannot been neglected, suggesting that highly dust-obscured star formation is relatively common in the $z > 3$ Universe.

Finally, we calculated the combined contribution of the OFGs and LBGs (values given from Madau & Dickinson 2014) to the cosmic SFRD. As shown by the red open circle in the left panel of Fig. 11, the total SFRD is $4.0 \times 10^{-2}\ M_\odot\ \mathrm{yr}^{-1}\ \mathrm{Mpc}^{-3}$ at $z = 4-5$, about 0.15 dex higher (43%) than the SFRD at the same redshift from Madau & Dickinson (2014). As we mention previously, the calculation of the SFRD values was very conservative and the true total SFRD of the OFGs and LBGs could be greater.

7. Conclusions

This work aims to obtain a more complete picture of the cosmic star formation history in the $z > 3$ Universe, that is, to bridge the extreme population of optically dark or faint galaxies (or H-dropouts) with the most common population of lower mass, less attenuated galaxies, such as LBGs. We use a more permissive criterion ($H > 26.5$ mag & $[4.5] < 25$ mag) to select optically dark or faint galaxies (i.e., OFGs) at high redshifts, which avoids limiting the sample to the most extreme cases. This criterion selects extremely dust-obscured massive galaxies that are normal star-forming galaxies, with dust obscuration typically at $E(B-V) > 0.4$, with lower stellar masses at high redshifts than H-dropouts. In addition, our selection method has the capacity to select OFGs without contamination from passive or old galaxies. In the GOODS-ALMA region, we have a total of 28 OFGs (including a candidate IRAC 4.5 μm dropout). We analyzed the properties of individual and stacked OFGs, respectively. We calculated their SFRD and quantified the incompleteness of our understanding of the cosmic star-formation history in the $z > 3$ Universe.

Here are the main results of this work:

1. After performing SED analyses with the code `EAzY` and `FAST++`, we find that the OFGs cover a redshift of $z_{\rm phot} > 3$, as indicated by the theoretical galaxy templates. The median redshift of individual OFGs is $z_{\rm med} = 4.1$, with a wide distribution ($z = 3.2-7.0$), consistent with $z_{\rm med,stacked} = 4.5 \pm 0.2$ derived from the median stacked SED. The OFGs have a broad stellar mass distribution with $\log(M_\star/M_\odot) = 9.4-11.1$, with a median of $\log(M_{\star\rm med}/M_\odot) = 10.3$.
2. We investigate the proportions of LBGs, H-dropouts, and remaining OFGs (after removing LBGs and H-dropouts) in our sample. Remarkably, at stellar masses of $\log(M_\star/M_\odot) = 9.5-10.5$, the fraction of remaining OFGs is about three times the sum of LBGs and H-dropouts. In other words, up to 75% of the OFGs with $\log(M_\star/M_\odot) = 9.5-10.5$ at $z > 3$ are neglected by the previous LBGs and H-dropout selection techniques.
3. All stacked OFGs are located within the SFMS region ($0.33 < \Delta{\rm MS} < 3$, i.e., ~0.5 dex), which is consistent with being normal star-forming galaxies at the same redshift. It suggests that rather than being limited to a rare population of extreme starburst galaxies, our OFGs represent a normal population of dusty star-forming galaxies at $z > 3$.
4. The gas properties of the OFGs imply that the OFGs have shorter $\tau_{\rm dep}$ and slightly lower $f_{\rm gas}$ values compared to the scaling relation followed by typical main sequence galaxies. Their gas is consumed more rapidly, hence they form their stars with a high efficiency, setting them in the so-called class of starbursts in the main sequence. In addition, the massive OFGs have the shortest $\tau_{\rm dep}$ and lowest $f_{\rm gas}$ of all stacked (sub)samples, indicating that they are in the process of becoming passive. Finally, we point out that the approach of selecting only ALMA-detected galaxies in studies of OFGs will end up biasing the sample toward larger SFRs, longer $\tau_{\rm dep}$, and larger $f_{\rm gas}$.
5. Studying dust temperatures in the OFGs, we find that, surprisingly, most of the stacked (sub)samples show colder $T_{\rm dust}$ than the scaling relation. In particular, the ALMA-detected OFGs have the most abundant dust but show the lowest $T_{\rm dust}$, indicating that the dust is colder in more obscured sources. A possible reason for the cold dust temperature is that the dust emission in the FIR of the dust-obscured sources may be optically thick rather than optically thin, where a warm and compact dust core is hidden.
6. In the comparison of $\rm SFR_{UV}^{corr}$ and $\rm SFR_{tot}$, we find that (*i*) all the stacked (sub-)samples have $\rm SFR_{tot}$ much larger than $\rm SFR_{UV}^{corr}$. There could be a hidden dust region of OFGs (even for LBGs as well) that absorbs all the UV photons, which cannot be reproduced with a dust extinction correction; and (*ii*) it is fundamental to include IR/mm band observations when studying extremely dusty star-forming galaxies; otherwise, the total SFR and, therefore, the cosmic SFRD will be underestimated.
7. After excluding five LBGs in the OFG sample, we study the hidden side of cosmic SFRD at high redshift ($z \gtrsim 3$) and find that: (*i*) among all galaxies, the total SFRD is dominated by LBGs, followed by OFGs. At $z = 4-5$, the contribution of OFGs reaches about 52% of the SFRD (Madau & Dickinson 2014), which is calculated mainly based on the LBGs (Bouwens et al. 2012a,b); (*ii*) for the OFGs, the massive and total OFGs have similar SFRD values, indicating that the massive OFGs make a major contribution to the SFRD compared to the remaining low-mass OFGs; (*iii*) for massive galaxies, the SFRD is dominated by massive OFGs rather than massive LBGs. The SFRD contributed by massive OFGs is at least two orders of magnitude higher than the one contributed by massive LBGs; (*iv*) the contribution of massive OFGs to the SFRD is more than four times higher than that of H-dropouts (Wang et al. 2019). It implies that optically faint galaxies also contribute significantly to the SFRD, which has been neglected in previous studies of LBGs and H-dropouts; and (*v*) the ALMA-detected OFGs and the total OFGs have similar SFRD values, but with different gas and dust properties (as we mentioned in point 4), such as $\tau_{\rm dep}$, $f_{\rm gas}$, $T_{\rm dust}$, and $R_{e(1.13\,\rm mm)}$, which require attention when studying the OFGs selected by ALMA detection alone.
8. Finally, we calculate the combined contribution of the OFGs and LBGs to the cosmic SFRD at $z = 4-5$, which is $4.0 \times 10^{-2}\ M_\odot\ \rm yr^{-1}\ Mpc^{-3}$, about 0.15 dex higher (43%) than the SFRD derived from LBGs alone (Madau & Dickinson 2014) at the same redshift. This value could be even larger as our calculation was very conservative.

Acknowledgements. We are very grateful to the anonymous referee for instructive comments, which helped to improve the overall quality and strengthened the analyses of this work. We thank Alain Omont, Boris Sindhu Kalita, Dangning Hu, Shuowen Jin, Pascal Oesch, and Bing Luo for valuable discussions and suggestions that improved this paper. This paper makes use of the following ALMA data: ADS/JAO.ALMA #2015.1.00543.S and ADS/JAO.ALMA #2017.1.00755.S. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSC, 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 work is supported by the National Key Research and Development Program of China (No. 2017YFA0402703), and by the National Natural Science Foundation of China (No. 11733002, 12121003, 12192220, 12192222, 12173017, and 12141301). This work is also supported by the Programme National Cosmology et Galaxies (PNCG) of CNRS/INSU with INP and IN2P3, co-funded by CEA and CNES. M.-Y. X. acknowledges the support by China Scholarship Council (CSC). M.F. acknowledges NSF grant AST-2009577 and NASA JWST GO Program 1727. G.E.M. acknowledges the Villum Fonden research grants 13160 and 37440 and the Cosmic Dawn Center of Excellence funded by the Danish National Research Foundation under the grant No. 140. R.D. gratefully acknowledges support by the ANID BASAL projects ACE210002 and FB210003.

References

Barrufet, L., Oesch, P. A., Weibel, A., et al. 2022, MNRAS, submitted, [arXiv:2207.14733]
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012a, ApJ, 754, 83
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012b, ApJ, 752, L5
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Bouwens, R., González-López, J., Aravena, M., et al. 2020, ApJ, 902, 112
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Casey, C. M. 2012, MNRAS, 425, 3094
Chabrier, G. 2003, PASP, 115, 763
Cortzen, I., Magdis, G. E., Valentino, F., et al. 2020, A&A, 634, L14
Cowie, L. L., González-López, J., Barger, A. J., et al. 2018, ApJ, 865, 106
Daddi, E., Cimatti, A., Renzini, A., et al. 2004, ApJ, 617, 746
Delhaize, J., Smolčić, V., Delvecchio, I., et al. 2017, A&A, 602, A4
Delvecchio, I., Daddi, E., Sargent, M. T., et al. 2021, A&A, 647, A123
Dickinson, M., Giavalisco, M., & GOODS Team 2003, The Mass of Galaxies at Low and High Redshift, 324
Donevski, D., Lapi, A., Małek, K., et al. 2020, A&A, 644, A144
Donley, J. L., Koekemoer, A. M., Brusa, M., et al. 2012, ApJ, 748, 142
Draine, B. T., & Li, A. 2007, ApJ, 657, 810
Draine, B. T., Aniano, G., Krause, O., et al. 2014, ApJ, 780, 172
Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2020, MNRAS, 494, 3828
Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Elbaz, D., Leiton, R., Nagar, N., et al. 2018, A&A, 616, A110
Enia, A., Talia, M., Pozzi, F., et al. 2022, ApJ, 927, 204
Fontana, A., Dunlop, J. S., Paris, D., et al. 2014, A&A, 570, A11
Franco, M., Elbaz, D., Béthermin, M., et al. 2018, A&A, 620, A152
Franco, M., Elbaz, D., Zhou, L., et al. 2020a, A&A, 643, A53
Franco, M., Elbaz, D., Zhou, L., et al. 2020b, A&A, 643, A30
Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767
Fudamoto, Y., Oesch, P. A., Schouws, S., et al. 2021, Nature, 597, 489
Genzel, R., Tacconi, L. J., Lutz, D., et al. 2015, ApJ, 800, 20
Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, ApJ, 600, L93
Gómez-Guijarro, C., Toft, S., Karim, A., et al. 2018, ApJ, 856, 121
Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022a, A&A, 658, A43
Gómez-Guijarro, C., Elbaz, D., Xiao, M., et al. 2022b, A&A, 659, A196
Gordon, K. D., Galliano, F., Hony, S., et al. 2010, A&A, 518, L89
Gruppioni, C., Béthermin, M., Loiacono, F., et al. 2020, A&A, 643, A8
Helou, G., Soifer, B. T., & Rowan-Robinson, M. 1985, ApJ, 298, L7
Herenz, E. C., Urrutia, T., Wisotzki, L., et al. 2017, A&A, 606, A12
Hildebrand, R. H. 1983, QJRAS, 24, 267
Hsieh, B.-C., Wang, W.-H., Hsieh, C.-C., et al. 2012, ApJS, 203, 23
Hwang, H. S., Elbaz, D., Magdis, G., et al. 2010, MNRAS, 409, 75
Jin, S., Daddi, E., Liu, D., et al. 2018, ApJ, 864, 56
Jin, S., Daddi, E., Magdis, G. E., et al. 2019, ApJ, 887, 144
Jin, S., Daddi, E., Magdis, G. E., et al. 2022, A&A, 665, A3
Kennicutt, R. C., & Evans, N. J. 2012, ARA&A, 50, 531
Kovács, A., Chapman, S. C., Dowell, C. D., et al. 2006, ApJ, 650, 592
Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, ApJ, 700, 221
Lee, N., Sanders, D. B., Casey, C. M., et al. 2015, ApJ, 801, 80
Leslie, S. K., Schinnerer, E., Liu, D., et al. 2020, ApJ, 899, 58
Lilly, S. J., Eales, S. A., Gear, W. K. P., et al. 1999, ApJ, 518, 641
Liu, D., Daddi, E., Dickinson, M., et al. 2018, ApJ, 853, 172
Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2
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., Elbaz, D., Chary, R. R., et al. 2011, A&A, 528, A35
Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132
Magnelli, B., Lutz, D., Saintonge, A., et al. 2014, A&A, 561, A86
Marrone, D. P., Spilker, J. S., Hayward, C. C., et al. 2018, Nature, 553, 51
McMullin, J. P., Waters, B., Schiebel, D., Young, W., & Golap, K. 2007, ASP Conf. Ser., 376, 127
Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43
Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793
Nonino, M., Dickinson, M., Rosati, P., et al. 2009, ApJS, 183, 244
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2014, ApJ, 786, 108
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2015, ApJ, 808, 104
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2018, ApJ, 855, 105
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Orellana, G., Nagar, N. M., Elbaz, D., et al. 2017, A&A, 602, A68
Puglisi, A., Daddi, E., Renzini, A., et al. 2017, ApJ, 838, L18
Retzlaff, J., Rosati, P., Dickinson, M., et al. 2010, A&A, 511, A50
Riechers, D. A., Bradford, C. M., Clements, D. L., et al. 2013, Nature, 496, 329
Riechers, D. A., Hodge, J. A., Pavesi, R., et al. 2020, ApJ, 895, 81
Rujopakarn, W., Dunlop, J. S., Rieke, G. H., et al. 2016, ApJ, 833, 12
Salpeter, E. E. 1955, ApJ, 121, 161
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Schreiber, C., Labbé, I., Glazebrook, K., et al. 2018a, A&A, 611, A22
Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018b, A&A, 618, A85
Schreiber, C., Elbaz, D., Pannella, M., et al. 2018c, A&A, 609, A30
Shu, X., Yang, L., Liu, D., et al. 2022, ApJ, 926, 155
Simpson, J. M., Smail, I., Swinbank, A. M., et al. 2015, ApJ, 799, 81
Smail, I., Dudzevičiūtė, U., Stach, S. M., et al. 2021, MNRAS, 502, 3426
Sommovigo, L., Ferrara, A., Pallottini, A., et al. 2022, MNRAS, 513, 3122
Speagle, J. S., Steinhardt, C. L., Capak, P. L., et al. 2014, ApJS, 214, 15
Stefanon, M., Labbé, I., Oesch, P. A., et al. 2021, ApJS, accepted [arXiv:2110.06226]
Straatman, C. M. S., Spitler, L. R., Quadri, R. F., et al. 2016, ApJ, 830, 51
Strandet, M. L., Weiss, A., De Breuck, C., et al. 2017, ApJ, 842, L15
Tacconi, L. J., Genzel, R., Saintonge, A., et al. 2018, ApJ, 853, 179
Talia, M., Cimatti, A., Giulietti, M., et al. 2021, ApJ, 909, 23
Urrutia, T., Wisotzki, L., Kerutt, J., et al. 2019, A&A, 624, A141
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Walter, F., Decarli, R., Carilli, C., et al. 2012, Nature, 486, 233
Wang, T., Schreiber, C., Elbaz, D., et al. 2019, Nature, 572, 211
Whitaker, K. E., van Dokkum, P. G., Brammer, G., et al. 2012, ApJ, 754, L29
Whitaker, K. E., Franx, M., Leja, J., et al. 2014, ApJ, 795, 104
Whitaker, K. E., Ashas, M., Illingworth, G., et al. 2019, ApJS, 244, 16
Williams, C. C., Labbe, I., Spilker, J., et al. 2019, ApJ, 884, 154
Yamaguchi, Y., Kohno, K., Hatsukade, B., et al. 2019, ApJ, 878, 73
Yun, M. S., Reddy, N. A., & Condon, J. J. 2001, ApJ, 554, 803
Zhou, L., Elbaz, D., Franco, M., et al. 2020, A&A, 642, A155

[1] School of Astronomy and Space Science, Nanjing University, Nanjing 210093, PR China
e-mail: my.xiao@smail.nju.edu.cn
[2] AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France
[3] Key Laboratory of Modern Astronomy and Astrophysics (Nanjing University), Ministry of Education, Nanjing 210093, PR China
[4] Aix- Marseille Université, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, 13013 Marseille, France
[5] Department of Astronomy, The University of Texas at Austin, 2515 Speedway Blvd Stop C1400, Austin, TX 78712, USA
[6] Community Science and Data Center/NSF's NOIRLab., 950 N. Cherry Ave., Tucson, AZ 85719, USA
[7] Department of Physics, Faculty of Science, Chulalongkorn University, 254 Phayathai Road, Pathumwan, Bangkok 10330, Thailand
[8] National Astronomical Research Institute of Thailand (Public Organization), Don Kaeo, Mae Rim, Chiang Mai 50180, Thailand
[9] Kavli IPMU (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan
[10] Cosmic Dawn Center (DAWN), Jagtvej 128, 2200 Copenhagen N, Denmark
[11] DTU-Space, Technical University of Denmark, Elektrovej 327, 2800 Kgs. Lyngby, Denmark
[12] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, 2200 Copenhagen N, Denmark
[13] Instituto de Astrofísica, Facultad de Física, Pontificia Universidad Católica de Chile, Casilla 306, Santiago 22, Chile
[14] Hiroshima Astrophysical Science Center, Hiroshima University, 1-3-1 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8526, Japan
[15] Departamento de Astronomía, Facultad de Ciencias Fśicas y Matemáticas, Universidad de Concepción, Concepción, Chile
[16] Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK
[17] International Space Science Institute (ISSI), Hallerstrasse 6, CH-3012 Bern, Switzerland
[18] Department of Physics, Anhui Normal University, Wuhu, Anhui 241000, PR China
[19] Laboratory for Multiwavelength Astrophysics, School of Physics and Astronomy, Rochester Institute of Technology, 84 Lomb Memorial Drive, Rochester, NY 14623, USA

[20] Centre for Extragalactic Astronomy, Department of Physics, Durham University, Durham DH1 3LE, UK
[21] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[22] Astronomy Department, University of Massachusetts, Amherst, MA 01003, USA
[23] National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan
[24] Department of Astronomical Science, SOKENDAI (The Graduate University for Advanced Studies), Mitaka, Tokyo 181-8588, Japan
[25] Joint ALMA Observatory, Alonso de Córdova 3107, Vitacura 763-0355, Santiago, Chile
[26] European Southern Observatory, Alonso de Córdova 3107, Vitacura, Casilla 19001, 19 Santiago, Chile
[27] Institute of Astronomy, Graduate School of Science, The University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan
[28] Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, UK
[29] Department of Astronomy, Universidad de Concepción, Casilla 160-C, Concepción, Chile
[30] Astronomy Unit, Department of Physics, University of Trieste, via Tiepolo 11, 34131 Trieste, Italy
[31] Fakultät für Physik der Ludwig-Maximilians-Universität, 81679 München, Germany
[32] Department of Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242, USA
[33] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics and Astronomy, Texas A&M University, College Station, TX 77843-4242, USA
[34] Astrophysics, Department of Physics, Keble Road,, Oxford OX1 3RH, UK

Appendix A: IRAC catalog compilation

We constructed the IRAC catalog using the deepest IRAC 3.6 and 4.5 μm images from the GREATS program (Stefanon et al. 2021). Source detection was performed using `Source Extractor` (SE version 2.25.0; Bertin & Arnouts 1996) on the background-subtracted 3.6 and 4.5 μm images, respectively. This is to help us exclude as many false detections as possible and get pure sources after cross-matching two catalogs from 3.6 and 4.5 μm images. Each detected source was required to have a minimum area of 4 pixels, with each pixel achieving the threshold of 0.25σ. The deblending parameters were optimized by setting `DEBLEND_THRESH` = 64 and `DEBLEND_MINCONT` = 0.0001, which were a compromise between deblending neighboring sources and minimizing a potential division of a larger object into multiple components. The sky subtraction was performed with SE, using a bicubic interpolation of the background with an adopted mesh size of 512 pixels and a median filter size of 5 pixels. A Mexhat filter was used to smooth the images before detection, to help detect faint and extended objects. The Mexhat filter can very well de-blend faint sources around the bright source, but it also extracts artifacts around the bright source caused by diffraction, which needs to be cleaned. We set `CLEAN_PARAM` = 1 to clean spurious detections with SE. Then, we visually inspected images to further remove artifacts around the bight sources. For the 3.6 and 4.5 μm images, we obtained their source catalogs respectively. The total sources in the catalogs are 125,338 for 3.6 μm and 154,234 for 4.5 μm in the entire GOODS-S field.

To ensure the purity of detections, we then cross-matched two catalogs from 3.6 and 4.5 μm images with a radius of 1.0″ (~0.5×FWHM). For sources that are simultaneously in the GOODS-ALMA 2.0 catalog (Gómez-Guijarro et al. 2022a) and detected in at least one IRAC band, we considered them as real sources and kept them in the final catalog. Finally, we have 5,127 pure sources detected by both 3.6 and 4.5 μm bands and/or ALMA 1.13 mm in the GOODS-ALMA field.

To estimate the completeness of our detection strategy, we employed a Monte Carlo approach where we simulated 10,000 artificial model sources with random magnitudes between 14.5 mag and 29 mag. They are point sources with the same PSF profiles as 3.6 and 4.5 μm images, respectively. The simulated sources were allowed to fall at random positions on the real image, including on top of other sources, to account for the impact of source blending in the real image. We injected ten sources each time on the IRAC 3.6um and 4.5um images to avoid excessively artificial source confusion caused by bright fake sources. After each injection, we performed the same blind source detection procedure using SE.

In Fig. A.1 (left), we show the completeness as a function of the input flux density ($S_{\rm in}$) for the simulated sources in 3.6 and 4.5 μm images. The survey reaches a 100% completeness for all simulated sources for flux densities $S_{\rm in} \gtrsim 20$ μJy at both 3.6 and 4.5 μm wavelength. To understand the incompleteness of our OFG sample, we need to know the behavior of completeness as a function of output flux density ($S_{\rm out}$) shown in Fig. A.1 (right). Here, $S_{\rm out}$ is the parameter being measured directly.

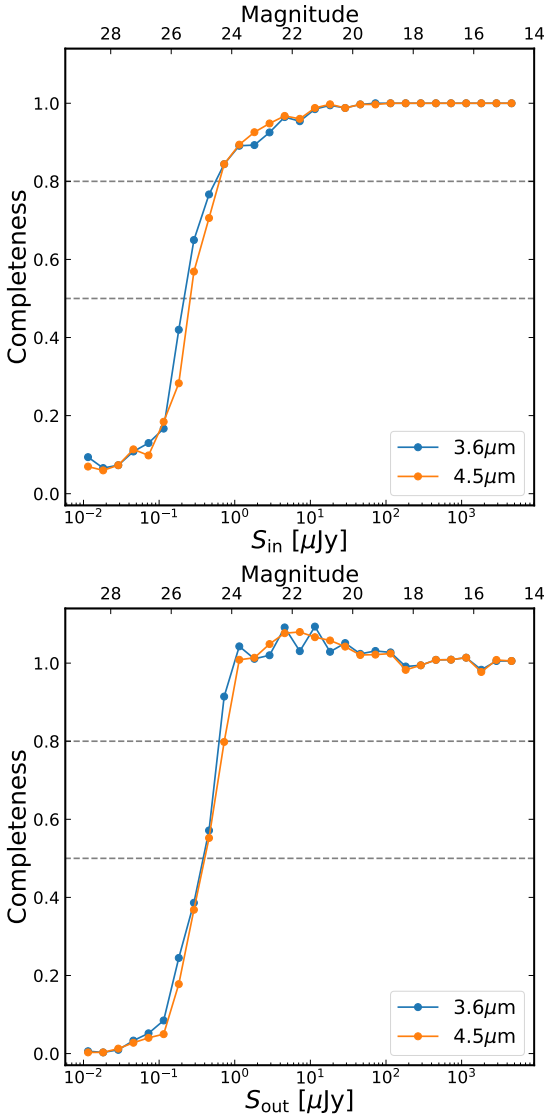

Fig. A.1. Completeness as a function of the input flux density ($S_{\rm in}$; top) and the output flux density ($S_{\rm out}$; bottom). The dashed lines represent the 50% and 80% completeness limits as a reference.

Appendix B: Multiwavelength Postage-stamp

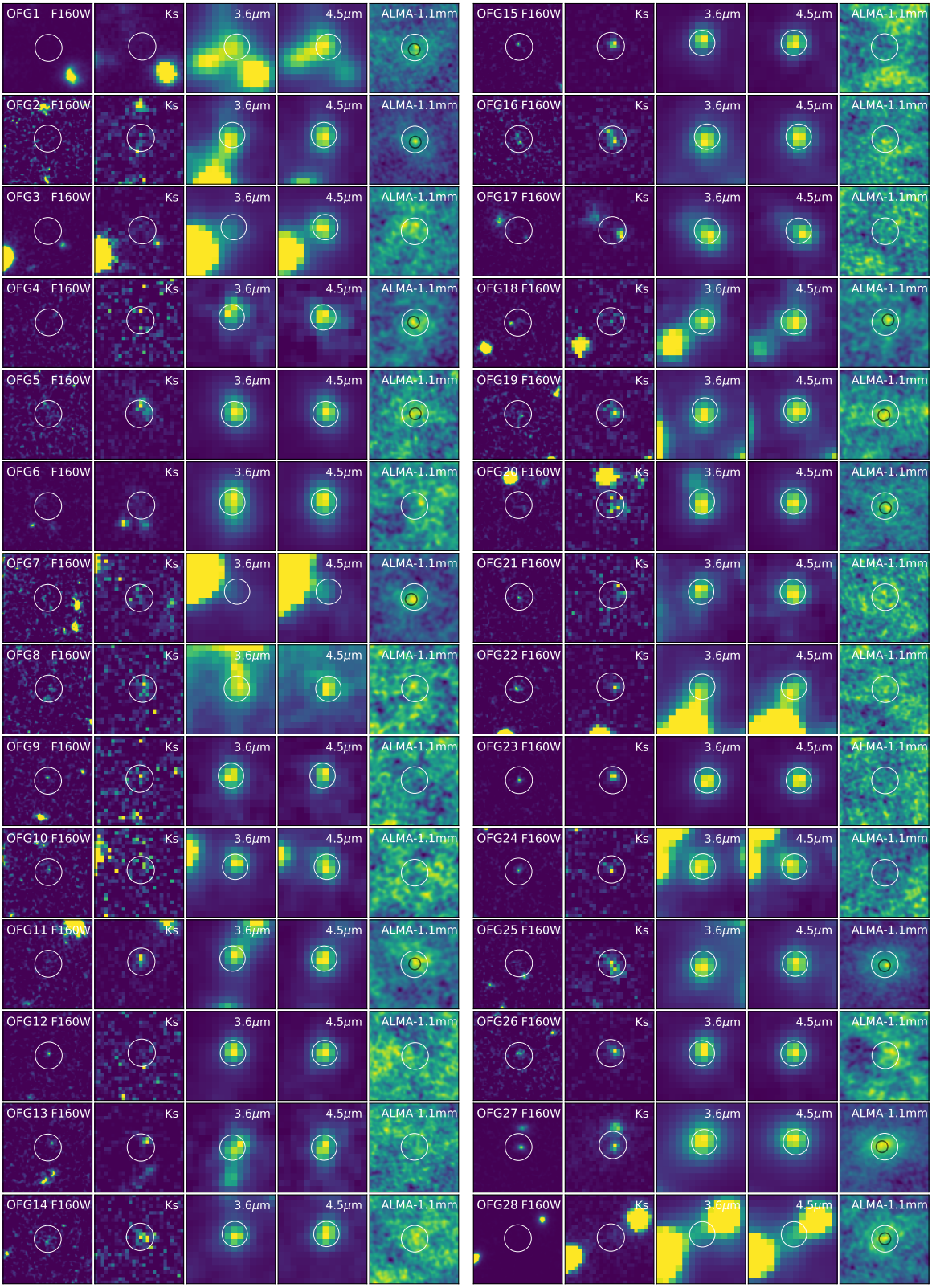

Fig. B.1. Postage-stamps (4″ × 4″) of our OFGs. From left to right: HST/WFC3 (F160W), ZFOURGE (K_s), *Spitzer*/IRAC (3.6 μm and 4.5 μm), and ALMA band 6 (1.13 mm). The white circles mark our target galaxies with 0.6″ in radius. The black circles indicate the positions of the ALMA detections. North is up and east is to the left.

Appendix C: SEDs for the individual OFGs

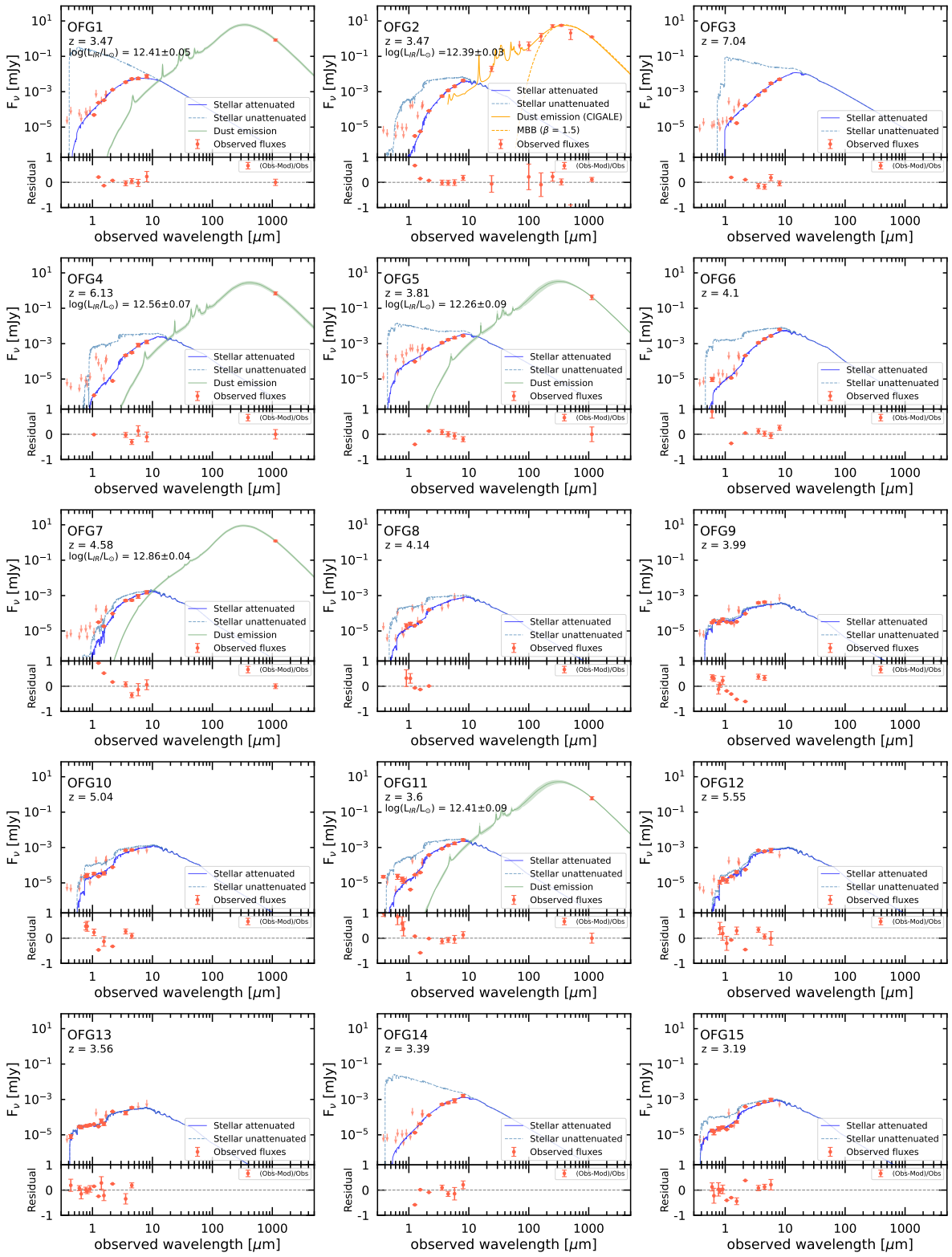

Fig. C.1. UV to millimeter SEDs for 27 OFGs. The symbol convention and lines are the same as those in Figs. 7. The data from the UV to MIR (16 μm) bands are fitted with the FAST++ code (see §4.1). From 24 μm up to millimeter wavelengths, we fitted galaxies with a *Herschel* counterpart using the CIGALE code (orange solid line) and galaxies without a *Herschel* counterpart but with the ALMA 1.13 mm detection using the dust template libraries (green line; Schreiber et al. 2018c, see §4.2).

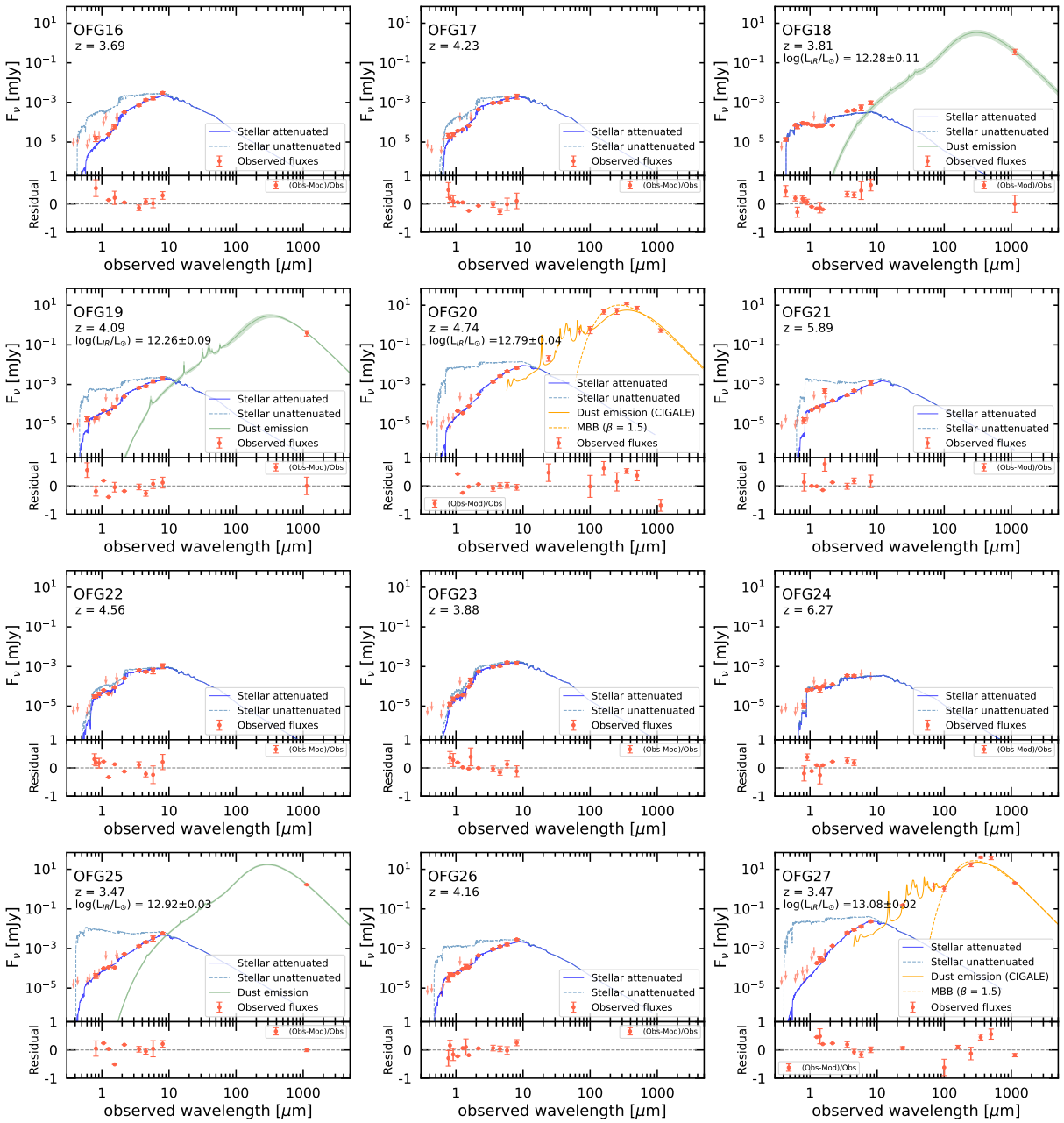

Fig. C.2. continued.