Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **486,** 560–569 (2019)
Advance Access publication 2019 March 19

doi:10.1093/mnras/stz801

Passive galaxies in the early Universe: ALMA confirmation of $z \sim 3$–5 candidates in the CANDELS GOODS-South field

P. Santini,[1]★ E. Merlin,[1] A. Fontana,[1] B. Magnelli,[2] D. Paris,[1] M. Castellano,[1] A. Grazian,[1] L. Pentericci,[1] S. Pilo[1] and M. Torelli[1]

[1]*INAF – Osservatorio Astronomico di Roma, via di Frascati 33, I-00078 Monte Porzio Catone (Roma), Italy*
[2]*Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*

Accepted 2019 February 21. Received 2019 January 30; in original form 2018 May 21

ABSTRACT

The selection of red, passive galaxies in the early Universe is very challenging, especially beyond $z \sim 3$, and it is crucial to constrain theoretical modelling of the processes responsible for their rapid assembly and abrupt shut-down of the star formation. We present here the analysis of ALMA archival observations of 26 out of the 30 galaxies in the deep CANDELS GOODS-South field that we identified as passive at $z \sim 3$–5 by means of a careful and conservative SED fitting analysis. ALMA data are used to verify the potential contamination from red, dusty but star-forming sources that could enter the sample due to similar optical–near-IR colours. With the exception of a few marginal detections at $<3\sigma$, we could only infer upper limits, both on individual sources and on the stacks. We translated the ALMA continuum measurements into corresponding SFRs, using a variety of far-IR models. These SFRs are compared with those predicted by secondary star-forming solutions of the optical fits and with the expected position of the star formation main sequence. This analysis confirms the passive nature of nine candidates with high confidence and suggests that the classification is correct for at least half of the sample in a statistical sense. For the remaining sources the analysis remain inconclusive because available ALMA data are not deep enough, although the stacking results corroborate their passive nature. Despite the uncertainties, this work provides decisive support to the existence of passive galaxies beyond $z \sim 3$.

Key words: methods: data analysis – galaxies: evolution – galaxies: formation – galaxies: high-redshift.

1 INTRODUCTION

The existence of massive, passively evolving galaxies at high redshift represents an arduous challenge to theoretical models of galaxy formation that struggle to reproduce the observations (Fontana et al. 2009; Vogelsberger et al. 2014; Feldmann et al. 2016). The abundance of these galaxies at different epochs is a crucial observable to constrain the different physical processes responsible for their rapid assembly and for the abrupt shut-down of their star formation activity. Theoretical simulations are indeed very sensitive to the detailed modelling of processes such as merger-driven starbursts or feedback (e.g. Menci et al. 2006; Hopkins et al. 2008; Choi et al. 2015).

To better understand these delicate physical processes, it is important to define reliable samples to which compare theoretical predictions. This is not a straightforward task, especially at high redshift, where highly dust-enshrouded galaxies are much more abundant and well mimic the emission of red evolved ones (Brammer et al. 2009). A number of selection criteria have been developed to this aim, such as colour – colour diagrams (e.g. Franx et al. 2003; Daddi et al. 2004; Wuyts et al. 2007; Martis et al. 2016), SED fitting (e.g. Grazian et al. 2007; Fontana et al. 2009), or a combination of SED fitting and a colour cut (e.g. Deshmukh et al. 2007), and are mostly limited to relatively low-intermediate redshifts. Some of the candidates have been spectroscopically confirmed (e.g. Cimatti et al. 2004; Onodera et al. 2012; Whitaker et al. 2013). In particular, the recent detection of a quiescent galaxy at $z \simeq 3.7$ provided crucial evidence of the existence of such objects even at $z > 3$ (Glazebrook et al. 2017; Schreiber et al. 2018b).

While being relatively easy to implement, especially at high redshift where other techniques become hard and sometimes not applicable, selections based on colours may suffer from incompleteness. This is clearly demonstrated by our previous work (Merlin et al. 2018, M18 hereafter), where we show that galaxies that have undergone an abrupt truncation of their star formation activity can

★ E-mail: paola.santini@inaf.it

remain outside the passive selection region of the $U - V$ versus $V - J$ diagram (Williams et al. 2009) for a few hundreds Myr. In M18, we performed a very accurate and conservative selection based on SED fitting, with an appropriate choice of the model star formation history (SFH), and selected a sample of 30 passive, red, and dead galaxies at $z > 3$ in the GOODS-South field. Our analysis showed that the reliability of the selection depends also crucially on the details of the SED fitting method, such as the inclusion of emission lines or the full inclusion of redshift uncertainties that may decrease the size of the sample by large factors. For this reason, and taking into account also other possible degeneracies between red dusty/star-forming and passive solutions, a more stringent verdict upon the lack of star formation is achievable through far-IR/submillimetre observations, which are able to sample the cold dust emission, expected to be prominent in star-forming galaxies. In M18, we performed a sanity check by means of *Herschel* data, and found detection for 2 out of 30 candidates, potentially caused by hot dust emission from an AGN hosted in these two galaxies. However, *Herschel* observations only allow the detection of few, extremely star-forming galaxies at these redshifts, while normal, main sequence (MS hereafter) galaxies would remain undetected in any case. In this paper, we make use of the rich ALMA archive to search for cold dust emission, hinting on-going star formation, around our candidates, and exploit the submillimetre inferred (limits on the) SFR to validate our classification, both on an individual basis and in a statistical sense.

The paper is organized as follows. We summarize our previous work and candidate selection in Section 2, describe the ALMA observations in Section 3, and derive the expected SFR based on these observations in Section 4. Finally, we present our results in Section 5 and draw our conclusions in Section 6. In the following, we adopt the Λ cold dark matter concordance cosmological model ($H_0 = 70\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.3$, and $\Omega_\Lambda = 0.7$) and a Salpeter (1955) Initial Mass Function (IMF). All magnitudes are in the AB system.

2 CANDIDATES SELECTION

We briefly summarize here the strategy pursued in M18 to select candidate passive galaxies in the CANDELS GOODS-S field by means of SED fitting.

Our selection takes advantage of the deep and high-quality photometry available in this field as well as of the sophisticated photometric measure techniques adopted (Merlin et al. 2015, 2016). In M18, we demonstrated that the known criterion based on the rest-frame $U - V$ versus $V - J$ colours (*UVJ* in the following) suffers from uncertainties due to the high redshift and extremely red colours of the desired candidates. In addition, we showed that the *UVJ* criterion is physically inappropriate to take into account the short time-scales for galaxies to become quiescent at $z > 3$, especially using the standard exponentially declining laws to model star formation histories.

To face these issues, we adopted a SED fitting technique assuming a 'top-hat' SFH, characterized by a period of constant star formation followed by an abrupt truncation of the star formation, that is set to zero thereafter. To estimate the reliability of our candidates we adopted a full statistical analysis, implementing a strict criterion based on the probability P of the χ^2 resulting from the fitted solution. For a galaxy to be selected as passive we require that the best-fitting solution is passive and has a probability $P(\chi^2_Q) > 30$ per cent *and* that no star-forming solution with $P(\chi^2_\mathrm{SF}) > 5$ per cent exists.

Table 1. List of ALMA programs used in this analysis and corresponding PI, ALMA band, and resolution.

ALMA program	PI	Band	Beam (arcsec × arcsec)
2012.1.00173.S	J. Dunlop	6	0.62 × 0.52
2012.1.00869.S	J. Mullaney	7	0.74 × 0.62
2013.1.00718.S	M. Aravena	6	1.44 × 0.73
2013.1.01292.S	R. Leiton	7	0.61 × 0.58
2015.1.00098.S	K. Kohno	6	0.60 × 0.60
2015.1.00543.S	D. Elbaz	6	0.60 × 0.60
2015.1.00664.S	K. Tadaki	6	0.72 × 0.58
2015.1.00870.S	T. Wiklind	7	0.70 × 0.61
2015.1.01074.S	H. Inami	7	0.67 × 0.59
2015.1.01495.S	T. Wang	7	0.63 × 0.58

We have implemented this approach in three different flavours. First, we have adopted naked Bruzual & Charlot (2003) models without any inclusion of nebular emission, and fixing the redshift to the photometric one (e.g. Grazian et al. 2007; Fontana et al. 2009). This choice results in a sample of 30 objects (dubbed S_0 hereafter). We have then added emission lines, self-computed on the basis of the ionizing flux of each template, as described in Castellano et al. (2014) and Schaerer & de Barros (2009). The inclusion of solutions with strong emission lines changes the predicted shape of the spectral slope in the reddest bands, and strongly decreases the number of candidates to 10 (S_1 sample). Finally, we have also let the redshift free to vary, and removed from the sample the objects that have a plausible star-forming solution at a different redshift. This way we are left with only two objects in the sample (S_2 sample).

This drastic reduction of the number of the 'bona-fide' candidates as we vary the spectral models used for the star-forming templates is found also in the whole CANDELS field, as we present in a companion paper (Merlin et al., in preparation). We remark that our method is quite conservative, as it retains only the objects that have both a quiescent best-fitting solution and no plausible star-forming solutions.

Far-infrared data are fundamental to exclude the potential star-forming solutions for our candidates. In M18, we have searched for *Herschel* counterparts of our candidates and found detections for two of them. As discussed in Section 5, they turned out to be potentially obscured AGNs rather than star-forming galaxies. *Herschel* images are, however, not deep enough to probe normal star-forming galaxies at $z > 3$. Deep ALMA observations therefore provide, at present, the only tool for validating the passive solutions for our candidates. For this reason we have searched the ALMA archive for observations of the whole S_0 sample.

3 ALMA OBSERVATIONS AND DATA ANALYSIS

We have inspected the ALMA archive at the positions of our candidates, and found public observations for 26 out of 30 sources, belonging to a number of different programs, listed in Table 1.

Among the 26 sources observed by ALMA are one of the two objects belonging to the S_2 sample (ID10578) and further seven belonging to the S_1 sample (ID2782, ID3912, ID8785, ID9209, ID17749, ID18180, and ID23626). The observed sources are listed in Table 2.

Table 2. Passive galaxy candidates covered by ALMA observations, ordered by their reliability (S_2, S_1, and S_0 sample – see the text – separated by lines). The SFRs are obtained from the M10 template. 1σ upper limits are provided for $<1\sigma$ detections.

ID	z	Stellar mass log $(M/M_\odot)$	Band	ALMA program(s)	Measured flux (mJy beam^{-1})	Sensitivity per beam (mJy beam^{-1})	SFR ($M_\odot$ yr^{-1})
10578 C	3.06	$11.38^{+0.16}_{-0.19}$	6	2015.1.00098.S + 2015.1.00543.S	0.14	0.09	38.2 ± 24.2
			7	2015.1.01074.S	0.04	0.84	<130.0
2782 C	3.47	$10.84^{+0.09}_{-0.20}$	7	2015.1.00870.S	0.04	0.11	<22.5
3912	4.08	$10.56^{+0.23}_{-0.25}$	6	2015.1.00543.S	−0.08	0.17	<42.7
8785	3.98	$10.59^{+0.16}_{-0.22}$	6	2015.1.00543.S	−0.39	0.22	<56.4
9209 C	4.55	$10.96^{+0.11}_{-0.83}$	6	2015.1.00543.S	−0.03	0.16	<41.2
			7	2015.1.01074.S + 2013.1.01292.S	0.10	0.28	<58.8
17749 C	3.73	$11.04^{+0.11}_{-0.28}$	6	2015.1.00098.S + 2015.1.00543.S	0.12	0.08	32.4 ± 20.4
			7	2013.1.01292.S	0.04	0.30	<48.7
18180 C	3.61	$10.95^{+0.10}_{-0.24}$	6	2015.1.00098.S + 2015.1.00543.S	0.08	0.08	21.5 ± 20.3
			7	2013.1.01292.S	−0.48	0.28	<40.6
23626	4.64	$10.88^{+0.14}_{-0.18}$	7	2013.1.01292.S	0.16	0.30	<74.0
2608	3.58	$9.65^{+0.10}_{-0.23}$	7	2015.1.00870.S	−0.10	0.19	<27.5
3973 C	3.67	$11.27^{+0.04}_{-0.27}$	7	2013.1.01292.S	0.74	0.29	108.8 ± 43.3
4503 C	3.52	$11.15^{+0.10}_{-0.23}$	6	2015.1.00543.S	−0.56	0.31	<77.4
			7	2013.1.01292.S	−0.04	0.29	<42.3
4587	3.58	$9.74^{+0.25}_{-0.16}$	6	2015.1.00543.S	−0.08	0.17	<43.6
5592	4.45	$10.48^{+0.20}_{-0.34}$	6	2015.1.00543.S + 2015.1.00870.S + 2015.1.00664.S	0.05	0.05	<29.6
6407	4.74	$10.20^{+0.23}_{-0.11}$	6	2015.1.00543.S	−0.13	0.17	<42.9
			7	2012.1.00869.S	−0.20	0.37	<59.2
7526	3.42	$10.56^{+0.17}_{-0.29}$	6	2015.1.00543.S	0.10	0.23	<84.5
			7	2015.1.00870.S	0.10	0.12	<31.5
7688	3.35	$10.36^{+0.19}_{-0.31}$	6	2015.1.00543.S	0.25	0.22	63.1 ± 56.6
			7	2015.1.00870.S	0.09	0.14	<34.3
8242	3.18	$9.82^{+0.11}_{-0.18}$	6	2015.1.00543.S	0.18	0.23	<104.4
9091	3.30	$9.45^{+0.29}_{-0.15}$	6	2015.1.00543.S	0.07	0.17	<62.6
			7	2012.1.00869.S	−0.08	0.11	<16.0
10759	3.07	$8.96^{+0.36}_{-0.53}$	6	2015.1.00543.S	0.07	0.22	<75.4
12178	3.28	$10.61^{+0.15}_{-0.13}$	6	2015.1.00543.S	0.13	0.17	<78.1
15457	3.41	$9.64^{+0.22}_{-0.07}$	6	2015.1.00543.S + 2015.1.00098.S + 2012.1.00173.S	0.03	0.03	11.2 ± 11.0
			7	2015.1.01074.S	1.79	0.79	265.0 ± 117.3
16506 C	3.34	$9.70^{+0.23}_{-0.06}$	6	2015.1.00543.S + 2015.1.00098.S + 2012.1.00173.S + 2013.1.00718.S	0.01	0.02	<9.2
19301	3.60	$10.06^{+0.21}_{-0.27}$	6	2015.1.00098.S + 2015.1.00543.S	0.01	0.08	<26.2
19446	3.25	$10.30^{+0.07}_{-0.33}$	6	2015.1.00098.S + 2015.1.00543.S	0.03	0.09	<31.9
19505 C	3.33	$10.67^{+0.05}_{-0.17}$	6	2015.1.00870.S	0.02	0.04	<20.6
			7	2015.1.00870.S	−0.05	0.11	<16.8
22610	3.22	$9.98^{+0.17}_{-0.17}$	7	2015.1.01495.S + 2015.1.01074.S	0.09	0.24	<48.9
	3.44	$10.31^{+0.09}_{-0.11}$	6	Stack all Band 6 sources at $3 < z < 5$	0.02	0.01	7.4 ± 3.8
	3.37	$10.29^{+0.08}_{-0.10}$	6	Stack all Band 6 sources at $3 < z < 4$	0.02	0.01	7.6 ± 4.0
	4.45	$10.54^{+0.19}_{-0.33}$	6	Stack all Band 6 sources at $4 < z < 5$	0.02	0.05	<18.9
	3.49	$10.66^{+0.15}_{-0.22}$	7	Stack all Band 7 sources at $3 < z < 5$	0.02	0.05	<9.5
	3.41	$10.64^{+0.14}_{-0.21}$	7	Stack all Band 7 sources at $3 < z < 4$	0.02	0.05	<9.4
	4.63	$10.84^{+0.18}_{-0.33}$	7	Stack all Band 7 sources at $4 < z < 5$	0.05	0.18	<36.5
	3.44	$10.15^{+0.08}_{-0.11}$	6	Stack undetected Band 6 sources at $3 < z < 5$	0.01	0.01	4.3 ± 4.3
	3.35	$10.09^{+0.07}_{-0.09}$	6	Stack undetected Band 6 sources at $3 < z < 4$	0.01	0.01	<8.7

Table 2 – *continued*

ID	z	Stellar mass log ($M/M_\odot$)	Band	ALMA program(s)	Measured flux (mJy beam^{-1})	Sensitivity per beam (mJy beam^{-1})	SFR ($M_\odot$ yr^{-1})
	4.45	$10.54^{+0.19}_{-0.33}$	6	Stack undetected Band 6 sources at $4 < z < 5$	0.02	0.05	<18.9
	3.48	$10.63^{+0.15}_{-0.22}$	7	Stack undetected Band 7 sources at $3 < z < 5$	−0.00	0.05	<6.7
	3.40	$10.61^{+0.14}_{-0.21}$	7	Stack undetected Band 7 sources at $3 < z < 4$	−0.01	0.05	<6.9
	4.63	$10.84^{+0.18}_{-0.33}$	7	Stack undetected Band 7 sources at $4 < z < 5$	0.05	0.18	<36.5

Note. The 'C' next to the ID mark candidates that are individually confirmed with high confidence. Redshift and stellar masses for the stacks have been obtained as a weighted average. The uncertainty on the stacked stellar mass has been computed from the weighted average relative error.

The observations have been carried out in Band 6 and Band 7, with different sensitivities, setups, and configurations. We have not used Band 3 and Band 4 observations as they do not add further information to the analysis (due to their shallowness and/or their dearth). Most of the sources are covered by more than one program, either in the same band or in a different one. We have stacked the sources observed more than once in the same band and we have combined the results from the different bands, as explained below.

Observations were calibrated with CASA (McMullin et al. 2007) using scripts provided by the ALMA project. Imaging was then performed using the multifrequency synthesis algorithm implemented within the CASA task CLEAN. We used 'natural' weightings and *uv*-tapers when needed, producing images with spatial resolution >0.6 arcsec (see Table 1). This resolution was chosen following the results from Franco et al. (2018), who demonstrated that all sources are unresolved in their GOODS-S 0.6 arcsec-tapered mosaic (from which most of our sources are taken) and that the average size of galaxies at submillimetre wavelengths is 0.3 ± 0.1 arcsec. Our galaxies being in addition relatively compact in the near-IR (average $R_{1/2} \sim 0.22$ arcsec in the H band), we could securely assume that all were point-like sources in our ALMA images.

With this point-like source assumption, the flux of each candidate corresponds simply to the pixel value at its position in the primary beam-corrected ALMA image (in unit of mJy beam^{-1}). The associated flux error was measured by taking the standard deviation of all pixels in the map with similar coverage, i.e. pixels corresponding to primary beam corrections within ±10 per cent to that at the position of the candidate. In this procedure, we excluded pixels within one full width at half-maximum of the candidate and we applied a 3σ clipping method to avoid biases from the candidate and any other sources.

We report in Table 2 the flux measured for each object. As can be seen, we found no high-confidence (i.e. $>3\sigma$) detection on ALMA images for any of the candidates. For sources observed more than once in the same band, we have stacked the inferred flux densities by averaging the fluxes measured from different programs weighting them with the associated errors (the final sensitivity per beam was inferred as the standard error on the weighted mean). Despite the improved sensitivity achieved by stacking, none of the sources is detected at a significant confidence level, with only two sources marginally detected at $\sim 2-2.5\sigma$ in Band 7. With the exception of additional five measurements barely above the noise level ($1-1.5\sigma$), for the rest we could only infer upper limits.

We note that the number of 1 and 2σ detections is consistent with a normal distribution of the signal-to-noise ratio, i.e. is consistent with a sample of undetected sources: indeed, out of our 37 measurements (26 sources, some of which observed in both bands), one would expect ∼6 sources in the upper ($>1\sigma$) tail of the S/N distribution, of which slightly less than one at $>2\sigma$.

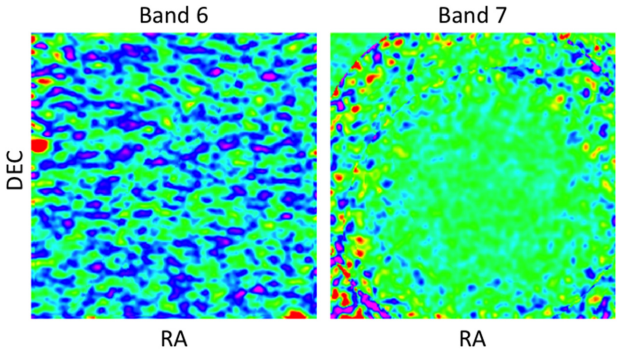

Figure 1. Stacked continuum images (25 arcsec × 25 arcsec) in Band 6 (*left*) and Band 7 (*right*) of 21 and 16 sources, respectively, at $3 < z < 5$.

Finally, we stacked all sources observed in the same band both over the entire redshift range and divided in two redshift bins ($3 < z < 4$ and $4 < z < 5$). No detection is obtained in Band 7. Sources in Band 6 are only marginally detected ($\lesssim 2\sigma$, in both the $3 < z < 4$ and $3 < z < 5$ redshift bins). When stacking only sources that are individually undetected, a flux comparable with the noise level is measured at $3 < z < 5$ in Band 6, while no detection is obtained at $3 < z < 4$. The fact that no detection emerges even from the stacks allows us to exclude the possibility that a significant fraction of the undetected objects has flux at $>1\sigma$.

We show in Fig. 1 the stacked images in Bands 6 and 7 obtained by averaging all individual images weighting them with the corresponding rms. The lack of detection at centre can be clearly seen.

Table 2 lists the ALMA program, band, actual measured flux, and image sensitivity per beam for each of the 26 sources, together with their best-fitting redshift and stellar mass. We also list the stacking results for the sources observed in the two ALMA bands, both considering all sources and only the undetected ones.

4 ALMA PREDICTIONS ON THE SFR

To quantify whether these observations are effective in discriminating against star-forming solutions, we have used ALMA observations to infer estimates of, or in most cases upper limits on, the dust-obscured SFR. To this aim, we computed the 1σ upper limits on the flux as the rms for any source with zero or negative measured flux, and as measured flux plus the rms for any source with a small amount of positive flux. If the measured flux is larger than the rms, the measurement is treated as a marginal detection. When using ALMA observations to individually validate the candidates, we adopt more conservative 3σ limits.

For candidates whose flux has been obtained by combining different ALMA images with slightly different frequency settings

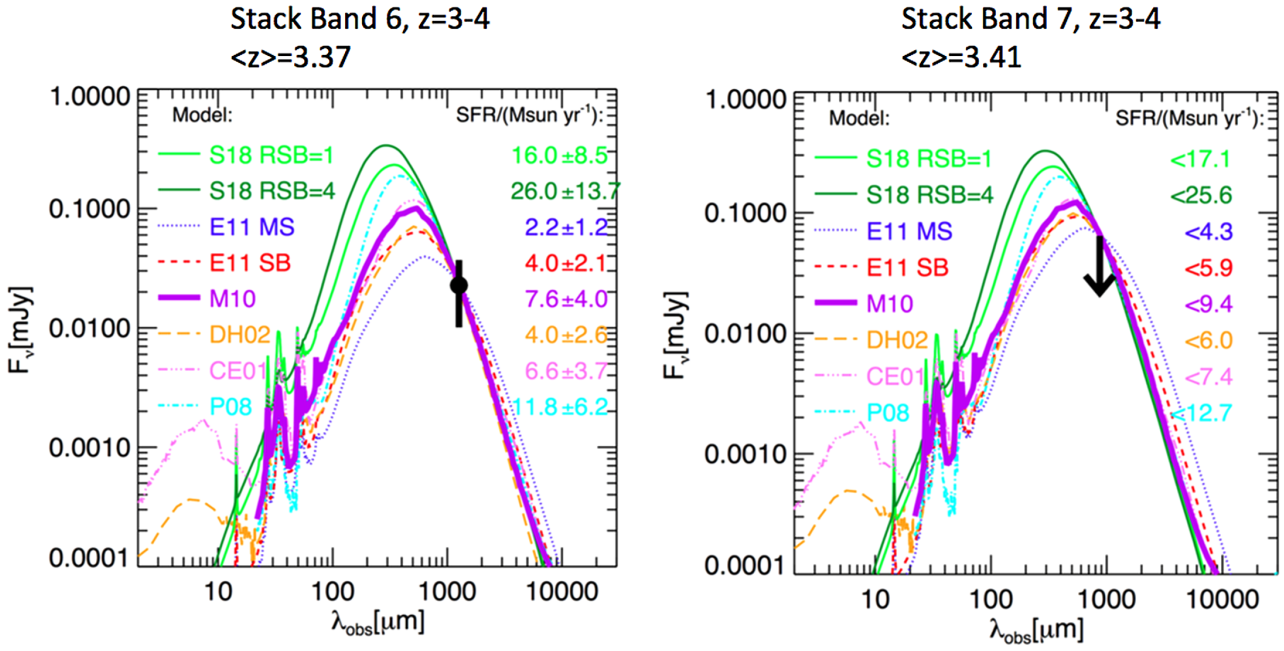

Figure 2. Best fit to the ALMA (limiting) flux measured by stacking all $3 < z < 4$ sources in Band 6 (*left*) and Band 7 (*right*). Our reference model (M10) is shown by the thick purple lines, while the others are represented by thin lines colour-coded and characterized by different linestyles according to the legend on the left. On the right-hand side of each panel, we report the corresponding SFR.

within the same band, we have computed the final band as a weighted average of the individual spectral windows covered by the different observations.

The SFR has been computed from the total infrared luminosity between 3 and 1100 μm adopting the calibration of Kennicutt & Evans (2012), adjusted to a Salpeter IMF using their conversion factor. We note that we do not take into account the contribution from old stellar populations to dust heating that may be not negligible if our candidates are truly passive. The true SFRs are therefore likely to be lower than those inferred by us, hence our results have to be considered as conservative.

To obtain an estimate of, or an upper limit on, the total infrared luminosity we have adopted a number of different models available in the literature. As reference model we used the average SED of Michałowski, Hjorth & Watson (2010) (M10), based on a sample of 76 submillimetre galaxies (SMGs) at $0.01 < z < 3.6$ fitted with the GRASIL (Silva et al. 1998) model. We then considered the average SMG template of Pope et al. (2008) (P08), the two average SEDs fitted by Elbaz et al. (2011) (E11) for MS and starburst (SB) galaxies, and the full libraries of Chary & Elbaz (2001) (CE01), Dale & Helou (2002) (DH02), and Schreiber et al. (2018a) (S18). To reduce the number of free parameters of the latter library, we constrained the dust temperature and IR8 parameter ($= L_{\rm IR}/L_{8\,\mu{\rm m}}$) based on the source redshift following the recipes provided by the authors; we considered a template for a MS galaxy ($R_{\rm SB} = {\rm SFR/SFR_{MS}} = 1$) and one for a starburst galaxy with $R_{\rm SB} = 4$. The results obtained from the M10 model are reported in Table 2. As we show later, the submillimetre-based SFRs vary only mildly with redshift. This allows us to compute the SFR also for our stacked sources at the average redshift of each group of them, that we report in the lower part of Table 2.

The application of different models yields a range of resulting SFRs, which are typically in the range $10-50\,{\rm M_\odot\,yr^{-1}}$, but that may span nearly a decade if we consider the full range of adopted models, especially when the two extreme S18 models are considered. This is clearly shown in Fig. 2 where we show the different IR SEDs obtained after normalizing the models to the observed (limiting) flux of the average source, obtained by stacking all candidates at $3 < z < 4$ in the two ALMA bands. For a given ALMA flux, the total IR spectrum resulting from each model is clearly different, and hence the resulting SFR that is derived from the total IR luminosity. We note in particular that the two S18 models predict a significantly higher total IR luminosity because the peak of their SED is at lower wavelengths than the other models. The S18 templates, built on *Herschel* and ALMA $0.5 < z < 4$ detections, assume an increasing dust temperature with redshift, resulting in a mass-weighted temperature around 40–50 K at the redshift of our galaxies (and an even higher expected luminosity-weighted temperature, as discussed at length in Scoville et al. 2016). We note, however, that our sources are undetected (or only marginally detected) at submillimetre wavelengths, hence not representative of the sample adopted to build the library. Moreover, the recent work of Gobat et al. (2018) fits a dust temperature of 21–25 K for $z \sim 1.8$ quiescent galaxies. For these reasons, we believe that the S18 library is likely inappropriate and overestimates the SFR for our candidates. Finally, it is also clear from Fig. 2 that longer wavelength bands yield the most uncertain total IR luminosity, as they fall far from the peak of the grey body emission.

In the following, we will refer to M10 as our reference model since it was built on red submillimetre sources (as could be the case of our candidates, though mostly below the sensitivity of ALMA images) at high redshift. Moreover, this model predicts SFRs that are somewhat intermediate with respect to all other models considered. Using our reference M10 model we find that the ALMA observations constrain the SFR of our objects to be typically below 40–50 $\rm M_\odot\,yr^{-1}$.

Table 3. ALMA based (limiting) SFR obtained by combining the results from observations in different bands, when available. The lower part lists the stacking results when only considering the final (i.e. obtained after combining all data) undetected sources.

ID	Combined SFR ($M_\odot$ yr^{-1})	Band
10578	38.2 ± 24.2	6
2782	<22.5	7
3912	<42.7	6
8785	<56.4	6
9209	<41.2	6
17749	32.4 ± 20.4	6
18180	21.5 ± 20.3	6
23626	<74.0	7
2608	<27.5	7
3973	108.8 ± 43.3	7
4503	<42.3	7
4587	<43.6	6
5592	<29.6	6
6407	<42.9	6
7526	<31.5	7
7688	63.1 ± 56.6	6
8242	<104.4	6
9091	<16.0	7
10759	<75.4	6
12178	<78.1	6
15457	13.4 ± 11.0	6 and 7
16506	<9.2	6
19301	<26.2	6
19446	<31.9	6
19505	<16.8	7
22610	<48.9	7
Stack undetected Band 6 sources at $3 < z < 4$	<8.7	6
Stack undetected Band 6 sources at $4 < z < 5$	<18.9	6
Stack undetected Band 7 sources at $3 < z < 4$	<7.6	7

We combined measurements obtained from different bands in order to end up with a single value of the SFR for each source and for each model. When available, we used the detections or a weighted average of them, albeit we remind that these detections are in any case below 3σ. We verified their consistency with the limits, when limits and detections are available for the same source. When only upper limits are available, we took the most stringent one. The final SFRs used for the analysis are listed in Table 3.

5 VALIDATING THE PASSIVE SOLUTIONS

We finally use the SFRs inferred above to validate the quiescent nature of our candidates, both individually and of the whole sample in a statistical sense.

5.1 Validation of robust individual candidates

We compare the ALMA-based SFRs to the SFRs predicted by the optical fit. Basically, to test whether our candidates were erroneously best fitted by passive templates, we can check whether the alternative star-forming solutions are compatible or not with ALMA results.

To this aim, we have performed the SFR computation at any redshift between 0 and 6 to account for possible uncertainties in the photo-z fitting. The outcome of the analysis is reported in Fig. 3, where we show the resulting (limiting) SFR at all possible redshifts for all 26 sources. As mentioned above, it is clear that the resulting SFR is an almost flat function of redshift (unsurprisingly, as the negative K-correction at submillimetre and millimetre wavelengths almost compensates for cosmological dimming). We compare the ALMA-based SFRs (at any possible redshift) to the SED fitting SFRs predicted by the star-forming solutions, ranked with their probability of reproducing the observed SED, as measured by the probability $P(\chi^2_{\rm SF})$ of yielding the observed χ^2 in the fit to the optical–near-IR bands. The values with $P(\chi^2_{\rm SF}) > 5$ per cent are shown for each object in Fig. 3 in blue shades, while in grey are shown those with lower probability. Clearly, for each object the SED fitting SFRs are spread over a large range of potential values, so that clear-cut conclusions are difficult to reach. In addition, the values inferred from the far-IR models span almost a decade. For the sake of clarity, we show the full range of far-IR models only for the stacks, shown in Fig. 4, but the same uncertainty applies to each individual source. Given the uncertainty associated with the choice of the IR template and the arbitrariness of the 5 per cent threshold chosen to accept a solution as 'plausible', it is almost impossible to extract a well-defined statistics. However, it is clear from the comparison reported in Fig. 3 that ALMA-based SFRs lend crucial support to the passive nature of our candidates.

Given these uncertainties, we apply here conservative criteria to confirm our candidates on an individual basis. We discard the S18 FIR model as, on the basis of the motivations above, it is likely inappropriate to describe our sources. We adopt our reference model M10, but we note that the results are unchanged with the assumption of the P08 FIR template, which predicts the highest SFR (with the exception of S18). We assume 3σ submillimetric limits for all sources. We apply a criterion based on the requirement that all the star-forming solutions with $P(\chi^2_{\rm SF}) > 5$ per cent are above the ALMA predictions. We find that in 9 out of the total sample of 26 candidates (i.e. 35 per cent), 5 of which belonging to the most secure S_1 subsample (i.e. 63 per cent), the star-forming fits predict star formation rates that are above, and often significantly larger than, the estimates derived from ALMA, implying that such solutions are implausible. We mark these individually confirmed candidates with a 'C' in Table 2. We note that additional two sources (ID23626 and ID5592) just do not pass the selection, but the bulk of their optical fit solutions are anyhow above the 3σ curve. We also note that even the adoption of the most extreme, though likely inappropriate, model of S18 with $R_{\rm SB} = 4$, results in the exclusion of only two of the nine selected sources (not belonging to S_1). The fact that most SED solutions are so high is explained by the very red SEDs of these objects, which demand large amounts of dust to be fitted with a star-forming template. For this reason, this test is conservative itself as the SFRs obtained from the optical fits are known to be likely underestimated for extremely red and dusty sources (e.g. Santini et al. 2009). For the remaining, unconfirmed candidates the comparison is inconclusive, as the values of SFR predicted from the fit are lower than for the other sources, often because of the lower amount of dust necessary to fit the observed colour, and cannot be excluded by ALMA limits. We only find a couple of exceptions (ID7688 and ID15457) where ALMA-based SFRs are formally consistent with the star-forming solutions of the optical fit.

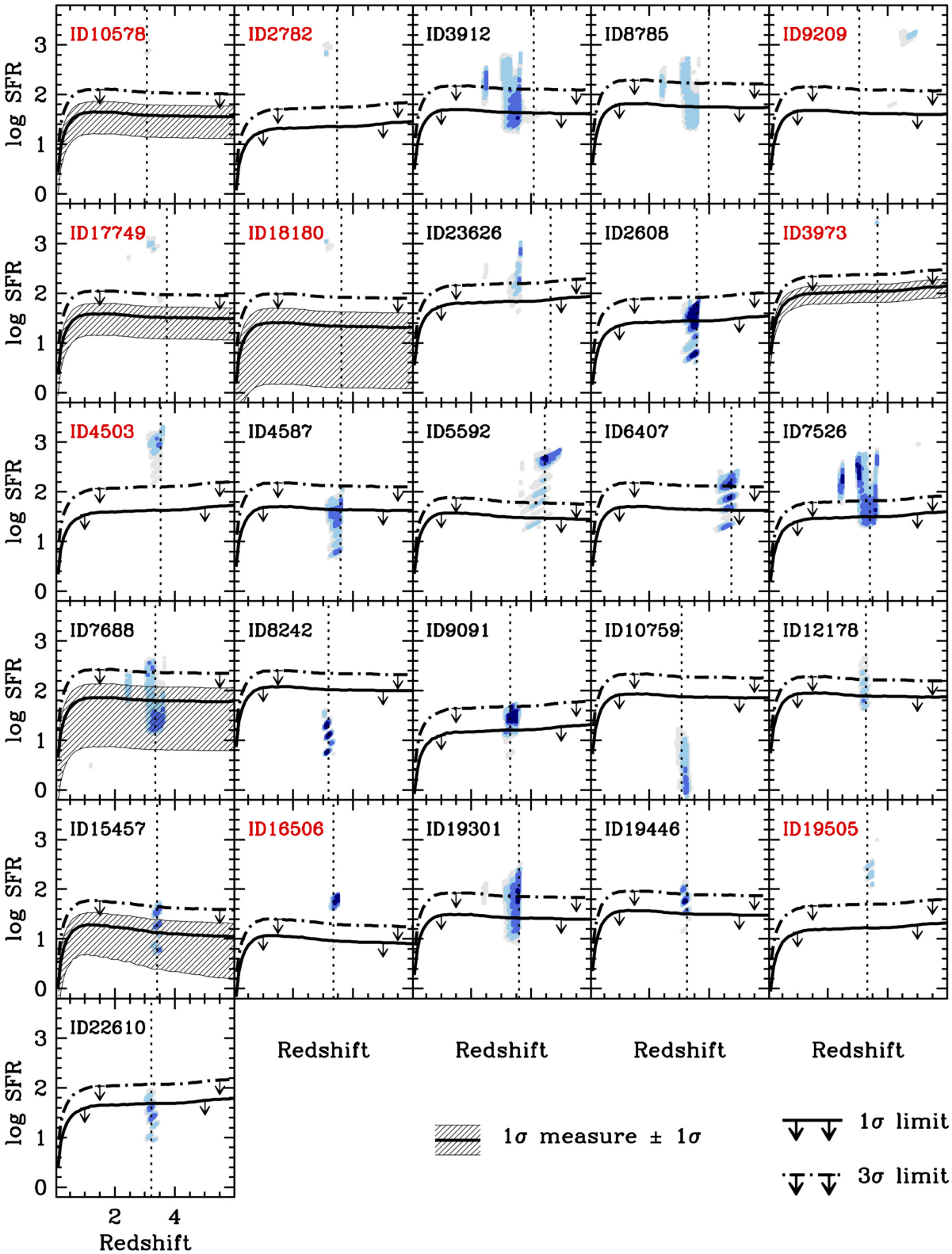

Figure 3. SFR confidence intervals or upper limits on the SFR at different redshifts based on ALMA observations according to our reference model M10. We show 1σ results as solid curves and 3σ ones as dot–dashed curves. The vertical dotted line indicates the best-fitting CANDELS photo-z. Blues dots, from lighter to darker shades, show the SFR inferred by the optical best fit by considering star-forming solutions at different redshifts (see the text), associated with probabilities 5–20 per cent, 20–50 per cent, and >50 per cent, respectively. Light grey dots show lower probability solutions ($P < 5$ per cent). The ID is printed in red colour for sources individually confirmed with high confidence.

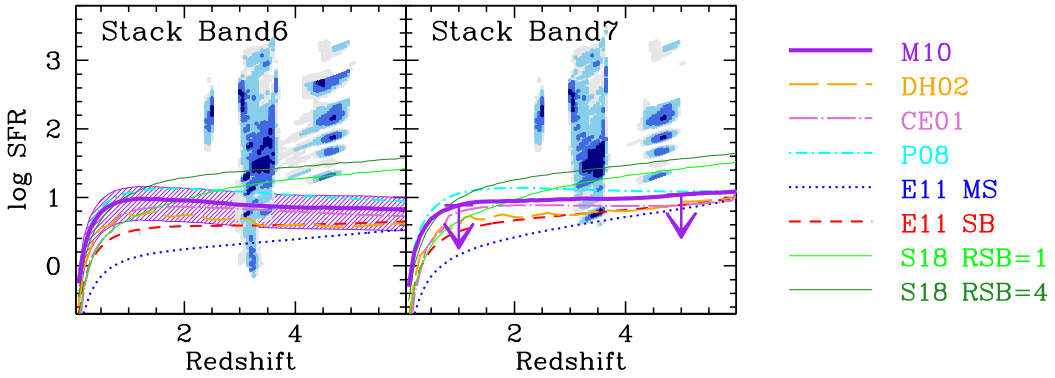

Figure 4. SFR confidence interval (*left*) and 1σ upper limits on the SFR (*right*) at different redshifts given the ALMA observations, according to the predictions of the different templates adopted, for stacked sources in Band 6 (*left*) and Band 7 (*right*) at all redshifts. Our reference model (M10) is shown by thick purple lines, while the others are represented by thin lines colour-coded and characterized by different linestyles according to the legend. The solutions of the optical fit (colour-coded as in Fig. 3) have been plotted for all stacked sources.

5.2 Validation of the whole population in a statistical sense

While only nine candidates can be individually confirmed with high confidence, more information can be drawn from the data that can be used to extract a statistical evaluation of the sample. To this aim, we adopt 1σ limits. We notice that, according to a Gaussian statistics, in 16 per cent of the undetected sources the limit may be too optimistic, and these sources may be erroneously classified as passive. However, this would not change the global results.

First of all, we take advantage from the stacking results to study the population as a whole. We perform a similar test on the stacked sources in the two ALMA bands. The stacked SFRs are compared with the collection of star-forming solutions of all sources included in the stacks. The results are shown in Fig. 4. We can claim that our candidates are on average consistent with being passive, i.e. the ALMA (limiting) SFR is lower than the star-forming solutions of the optical fit (we note that the low SFR tail in the solutions are essentially given by one single source in each of the stacks, i.e. ID10759 for Band 6 and ID2608 for Band 7).

A second, somewhat independent approach to statistically validate the sample is the comparison of the ALMA derived SFRs with those predicted by placing the objects on the star-forming MS. Rather than relying on the SFR derived from the fit to the optical–near-IR bands, in this case we use the stellar mass estimated for our candidates (that is usually considered a more robust measurement than the star formation rate; Santini et al. 2015) and evaluate whether these objects have a SFR lower than their siblings of the same stellar mass – i.e. if they lie on the observed MS for star-forming galaxies, or below it. The result of this comparison is shown in Fig. 5, where we plot the observed (i.e. not corrected for the Eddington bias) MS of star-forming galaxies at the same redshifts as inferred from the *HST* Frontier Fields data by Santini et al. (2017). Of the six marginally detected sources, three fall in the quiescent area, i.e. are located below the lower 3σ percentile of the distribution of star-forming galaxies, one is 1σ below the MS and two (ID7688 and ID15457, none of the two belonging to the most secure S_1 subsample) are consistent with the MS. However, the latter have with huge error bars, especially extending into the passive region of the diagram, that prevent any conclusion. Three/nine among the upper limits indicate that the candidates are at least 3σ/1σ below the MS. For only a couple of sources the limit on the SFR falls significantly above the MS, basically because of their low stellar mass. In total, given their stellar mass, the SFRs predicted by ALMA place 13 (6) candidates, 6 (4) of which belonging to the S_1 sample, at least 1σ (3σ) below the MS or around this threshold. The passive classification is on average confirmed at 1σ for at least 50 per cent of the sample. We note that, with the exception of the two extreme templates of S18 (see discussion above), these numbers are solid against the uncertainty in the modelling of the FIR spectrum.

On the upper panels of Fig. 5, we also show the average SFR derived from stacking all undetected sources observed in the same ALMA band (whose values are listed in the lower part of Table 3) as large red symbols. The average sources observed in both bands lie below the 2σ lower envelope of the distribution and in one case below 3σ, suggesting that our candidates undetected by ALMA are on average correctly classified as passive even when submillimetre data are not deep enough to draw conclusions.

5.3 Final considerations

It is important to remark that the two analyses yield consistent results, in that they both identify a subsample of more secure passive candidates and one made of sources for which the inferred limits on the SFR are not stringent enough to draw firm conclusions, although the stacking results seem to corroborate their passive nature. This can be clearly seen on the lower panels of Fig. 5: the objects for which the submillimetre-based SFRs are (much) lower than those allowed by the star-forming solutions of the optical fit (green symbols) populate the passive region of the SFR–stellar mass diagram while those whose ALMA limits do not exclude the star-forming solutions (purple symbols) lie around (or above) the MS. As mentioned, with both approaches, the results of the stacks yield tighter constraints.

It is interesting to note that the present analysis corroborates the interpretation of galaxies recently quenched by the emission of their still actively radiating nucleus for two of our candidates. Indeed, to exclude contamination from red, dusty sources, in M18 we searched for FIR emission on *Herschel* observations, and found a detection for two of the strongest candidates (ID10578 and ID3973). After a careful analysis of their optical and X-ray emission, we attributed the FIR emission to a dust-obscured AGN hosted at their centre. The much fainter and marginal detection at submillimetre wavelengths confirms that *Herschel* fluxes are likely caused by host dust heated by the central nucleus rather than cold dust heated by newly formed stars.

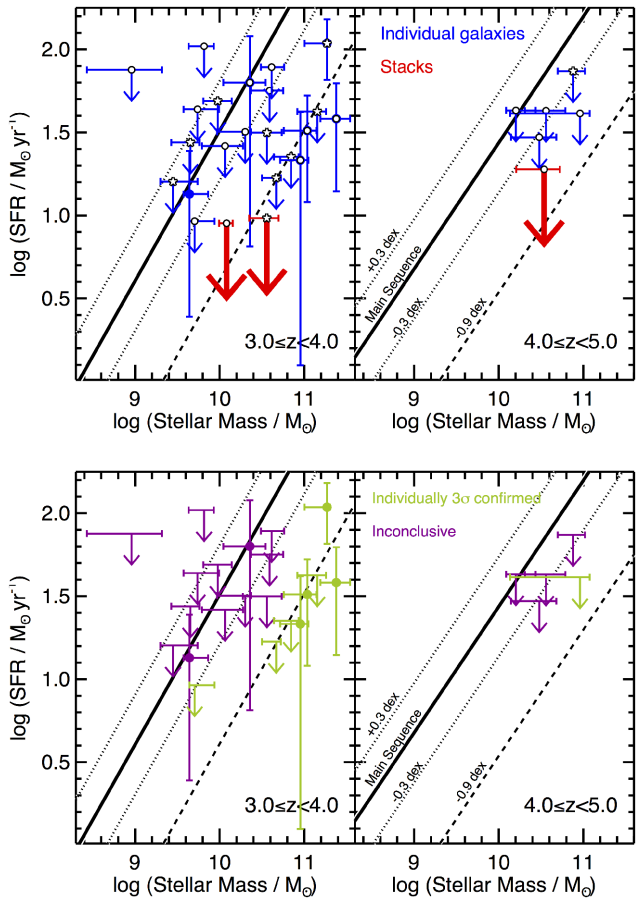

Figure 5. Location of the passive candidates on the SFR–stellar mass diagram, in two redshift bins, based on their ALMA SFR. Arrows represent 1σ upper limits on the SFR. Solid lines show the observed MS (i.e. not corrected for the Eddington bias) inferred from *HST* Frontier Field data by Santini et al. (2017). Dotted lines are 1σ above and below the MS (estimated from the observed 0.3 dex scatter), while the dashed line is 3σ below it. In the *upper* panels, white circles and stars denote observations in Band 6 and Band 7, respectively. Blue symbols show the individual sources and large and thick red symbols show the stacks of the undetected sources in each redshift bin and ALMA band. In the *lower* panels, green symbols represent objects whose passive nature is individually and robustly confirmed by ALMA at $\geq 3\sigma$, i.e. objects whose SFR predicted by ALMA at any redshift is below any possible star-forming solutions of the optical fit with an acceptable probability ($P(\chi^2_{\rm SF}) > 5$ per cent). Purple symbols show objects for which ALMA observations may be consistent with the star-forming solutions with $P(\chi^2_{\rm SF}) > 5$ per cent, but that are inconclusive because of too shallow submillimetre data. See the text and Fig. 3 for details.

6 SUMMARY AND CONCLUSIONS

In this paper, we have presented a follow-up analysis of the passive galaxy candidates selected at high redshift ($z = 3$–5) by our previous study (M18) by means of SED fitting of optical–near-IR bands. Starting from a sample of 30 candidates in the GOODS-S field, we found ALMA archival observations for 26 of them, in Band 6 and Band 7. None of the individual sources is robustly (i.e. $>3\sigma$) detected in the submillimetre, the large majority of sources being below the noise level. Indeed, the few marginally detected sources are consistent with belonging to the tail of the Gaussian noise distribution. No significant detection is obtained even by stacking sources observed in the same ALMA band. This allows us to exclude the possibility that a significant fraction of undetected objects has actually flux at $>1\sigma$. From the flux and rms measured on the ALMA images, we derived estimates of, or in most cases upper limits on, the SFR, which we use to validate the passive nature of our candidates, both individually and in a statistical sense.

First, we compared the submillimetre-based SFRs with the star-forming secondary solutions of the optical fits. For nine candidates the star-forming solutions are rejected by the ALMA observations adopting the most conservative assumptions, i.e. adopting 3σ upper limits and the FIR models providing the highest SFRs (with the exception of S18 model that we deem inappropriate to describe our sources). These sources are individually confirmed with high confidence. Secondly, we used the ALMA-based (limits on the) SFRs to compare the location of our candidates with respect to the star formation MS, given their stellar mass: 50 per cent of the candidates are placed below the 1σ distribution of star-forming galaxies, and 23 per cent (6 out of 26) fall in the quiescent area (i.e. 3σ below the MS). The results of the two tests overlap very nicely. While for the remaining, unconfirmed, candidates the comparison is inconclusive because the available submillimetre data are too shallow to draw firm conclusion, the stacking results suggest an overall passive nature for our sample.

Although the exact quantification of the fraction of confirmed candidates depends on the details of the analysis and of the models, we can reach the following conclusions from our study:

(i) ALMA observations lend decisive evidence to the quiescent nature of our passive candidates that clearly show a distribution of SFR that is inconsistent with the typical one at these redshifts;

(ii) currently available ALMA archive observations are not deep enough to individually confirm most of our candidates with high confidence; however,

(iii) we can individually confirm 9 candidates out of 26 (35 per cent) adopting conservative assumptions;

(iv) the stacking analysis and the lack of reliable detections corroborate the passive nature of the remaining part of the sample, at least in a statistical sense;

(v) at least half of the sample is located at least 1σ below the MS;

(vi) these results confirm the existence of passive galaxies in the early Universe ($z > 3$);

and

(vii) validate the robustness and reliability of the selection technique developed by our previous analysis (M18), in particular when the most conservative selection criteria are adopted.

In the next future, *JWST* observations not only will improve the selection of passive galaxy candidates at high z (M18), but will also make it possible to finally confirm them by means of a spectroscopic analysis.

ACKNOWLEDGEMENTS

We thank the anonymous referee for his/her helpful comments, and we thank M. Dickinson and M. Ginolfi for useful discussion regarding ALMA data reduction. The research leading to these results has received funding from the European Union Seventh Framework Programme ASTRODEEP (FP7/2007-2013) under grant agreement No. 312725. This paper makes use of the following ALMA data: ADS/JAO.ALMA#2012.1.00173.S,

ADS/JAO.ALMA#2012.1.00869.S,
ADS/JAO.ALMA#2013.1.00718.S,
ADS/JAO.ALMA#2013.1.01292.S,
ADS/JAO.ALMA#2015.1.00098.S,
ADS/JAO.ALMA#2015.1.00543.S,
ADS/JAO.ALMA#2015.1.00664.S,
ADS/JAO.ALMA#2015.1.00870.S,
ADS/JAO.ALMA#2015.1.01074.S, and
ADS/JAO.ALMA#2015.1.01495.S. ALMA is a partnership of ESO (representing its member states), NSF (USA), and NINS (Japan), together with NRC (Canada), NSC and ASIAA (Taiwan), and KASI (Republic of Korea), in cooperation with the Republic of Chile. The Joint ALMA Observatory is operated by ESO, AUI/NRAO and NAOJ.

REFERENCES

Brammer G. B. et al., 2009, ApJ, 706, L173
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Castellano M. et al., 2014, A&A, 566, A19
Chary R., Elbaz D., 2001, ApJ, 556, 562
Choi E., Ostriker J. P., Naab T., Oser L., Moster B. P., 2015, MNRAS, 449, 4105
Cimatti A. et al., 2004, Nature, 430, 184
Daddi E., Cimatti A., Renzini A., Fontana A., Mignoli M., Pozzetti L., Tozzi P., Zamorani G., 2004, ApJ, 617, 746
Dale D. A., Helou G., 2002, ApJ, 576, 159
Deshmukh S. et al., 2018, ApJ, 864, 166
Elbaz D. et al., 2011, A&A, 533, A119
Feldmann R., Hopkins P. F., Quataert E., Faucher-Giguère C.-A., Kereš D., 2016, MNRAS, 458, L14
Fontana A. et al., 2009, A&A, 501, 15
Franco M. et al., 2018, A&A, 620, A152
Franx M. et al., 2003, ApJ, 587, L79
Glazebrook K. et al., 2017, Nature, 544, 71
Gobat R. et al., 2018, Nat. Astron., 2, 239
Grazian A. et al., 2007, A&A, 465, 393
Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356
Kennicutt R. C., Evans N. J., 2012, ARA&A, 50, 531
Martis N. S. et al., 2016, ApJ, 827, L25
McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Francisco, p. 127
Menci N., Fontana A., Giallongo E., Grazian A., Salimbeni S., 2006, ApJ, 647, 753
Merlin E. et al., 2015, A&A, 582, A15
Merlin E. et al., 2016, A&A, 590, A30
Merlin E. et al., 2018, MNRAS, 473, 2098 (M18)
Michałowski M., Hjorth J., Watson D., 2010, A&A, 514, A67 +
Onodera M. et al., 2012, ApJ, 755, 26
Pope A. et al., 2008, ApJ, 675, 1171
Salpeter E. E., 1955, ApJ, 121, 161
Santini P. et al., 2009, A&A, 504, 751
Santini P. et al., 2015, ApJ, 801, 97
Santini P. et al., 2017, ApJ, 847, 76
Schaerer D., de Barros S., 2009, A&A, 502, 423
Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Franco M., 2018a, A&A, 609, A30
Schreiber C. et al., 2018b, A&A, 611, A22
Scoville N. et al., 2016, ApJ, 820, 83
Silva L., Granato G. L., Bressan A., Danese L., 1998, ApJ, 509, 103
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Whitaker K. E. et al., 2013, ApJ, 770, L39
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Wuyts S. et al., 2007, ApJ, 655, 51

This paper has been typeset from a TEX/LATEX file prepared by the author.