Red & Dead CANDELS: massive passive galaxies at the dawn of the Universe

E. Merlin[1], F. Fortuni[1], M. Torelli[1], P. Santini[1], M. Castellano[1], A. Fontana[1], A. Grazian[2], L. Pentericci[1], S. Pilo[1] and K. B. Schmidt[3]

[1] *INAF - Osservatorio Astronomico di Roma, via Frascati 33, 00078 Monte Porzio Catone (RM), Italy*
[2] *INAF - Osservatorio Astronomico di Padova, Vicolo Osservatorio 5, 35122, Padova, Italy*
[3] *Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, 14482 Potsdam, Germany*

ABSTRACT

We search the five CANDELS fields (COSMOS, EGS, GOODS-North, GOODS-South and UDS) for passively evolving a.k.a. "red and dead" massive galaxies in the first 2 Gyr after the Big Bang, integrating and updating the work on GOODS-South presented in a previous paper. We perform SED-fitting on photometric data, with top-hat star-formation histories to model an early and abrupt quenching, and using a probabilistic approach to select only robust candidates. Using libraries without (with) spectral lines emission, starting from a total of more than 20,000 $z > 3$ sources we end up with 102 (40) candidates, including one at $z = 6.7$. This implies a minimal number density of $1.73 \pm 0.17 \times 10^{-5}$ $(6.69 \pm 1.08 \times 10^{-6})$ Mpc^{-3} for $3 < z < 5$; applying a correction factor to account for incompleteness yields $2.30 \pm 0.20 \times 10^{-5}$. We compare these values with those from five recent hydrodynamical cosmological simulations, finding a reasonable agreement at $z < 4$; tensions arise at earlier epochs. Finally, we use the star-formation histories from the best-fit models to estimate the contribution of the high-redshift passive galaxies to the global Star Formation Rate Density during their phase of activity, finding that they account for ∼5-10% of the total star formation at $3 < z < 8$, despite being only ∼ 0.5% of the total in number. The resulting picture is that early and strong star formation activity, building massive galaxies on short timescales and followed by a quick and abrupt quenching, is a rare but crucial phenomenon in the early Universe: the evolution of the cosmos must be heavily influenced by the short but powerful activity of these pristine monsters.

Key words: Galaxies

1 INTRODUCTION

Quantifying the abundance of passively evolving ("red and dead") galaxies in the early Universe is a difficult but crucial task. We know that massive galaxies typically have red colors at all epochs: while in the local Universe this is mostly caused by the absence of young stellar populations (with a degeneracy caused by metallicity, e.g. Worthey 1994), at high redshift this is more often a consequence of high star-formation rates (SFRs) coupled with strong dust obscuration (e.g. Cimatti et al. 2002; Dunlop et al. 2007). However, it is established that a non-negligible fraction of massive galaxies in the first ∼ 2 Gyr after the Big Bang is intrinsically red because of passive evolution following the quenching of the star-formation (SF) activity (e.g. Labbé et al. 2005; Mobasher et al. 2005; Fontana et al. 2009; Grazian et al. 2015).

The very existence of such early red and dead massive galaxies is a challenge to our present understanding of the cosmos. The formation of the structures in the concordance Λ-CDM cosmological scenario is inherently hierarchical (Press & Schechter 1974; Lacey & Cole 1993), with large structures assembling at later times with ongoing bursts of star formation (e.g. White & Rees 1978; De Lucia & Blaizot 2007). On the other hand, the so-called *downsizing* trend is a well-established evidence, with massive galaxies assembling their stellar content earlier, and typically on shorter timescales, than smaller ones (Matteucci 1994; Cowie et al. 1996; Thomas et al. 2005; Bundy et al. 2006; Cimatti et al. 2006). In the last decades, theoretical models and hydrody-

namical simulations have struggled to reproduce the properties of the observed galactic populations at all epochs (e.g. Vogelsberger et al. 2014a; Feldmann et al. 2017); however, to date there is no consensus yet on a robust theoretical approach capable to accurately reconcile the observational data with the models.

The issue can also be viewed under a different perspective. In the last 15 years the tight correlation between galaxies SFRs and stellar masses, the so-called "main sequence", has become a thoroughly studied topic (Brinchmann et al. 2004; Noeske et al. 2007; Elbaz et al. 2007; Daddi et al. 2007). The sequence is now confirmed to exist to high redshifts (Rodighiero et al. 2014; Schreiber et al. 2015). However, at any epoch some (typically compact, bulge-dominated) galaxies fall below it, indicating little or no star formation activity and implying the occurred action of some quenching mechanism (Wuyts et al. 2011; Tacchella et al. 2018). The time-scales of such processes, and the physical drivers behind them, remain largely unclear to date (see e.g. Man & Belli 2018); the usual suspects include AGN-driven outflows (Brennan et al. 2017), stellar feedback (Kawata 1999; Chiosi & Carraro 2002; Ceverino & Klypin 2009; Merlin et al. 2012), gas strangulation (Peng et al. 2015) or starvation (Feldmann & Mayer 2015), virial shocking of the circum-galactic medium (Dekel & Birnboim 2006), or a combination of all these. Whatever the cause, an abrupt halt of the star formation activity makes the galaxy colors turn redder, but blue light from young stellar object can outshine the old populations for several Myrs after the quenching. This makes a simple color-based selection prone to bias, even when using rest-frame inferred magnitudes as in the *UVJ* diagram (Labbé et al. 2005; Wuyts et al. 2007). Galaxies that have quenched shortly before being observed will not enter the selection regions until later times unless an ad-hoc modeling is adopted, and as we have shown in our previous paper (Merlin et al. 2018, M18 hereafter) in which we exploited the CANDELS photometric data for the GOODS-South field (Grogin et al. 2011; Koekemoer et al. 2011; Guo et al. 2013), this is particularly true at very high redshifts, when the timescales of the events are comparable to the life span of the Universe. Davidzon et al. (2017) and Ichikawa & Matsuoka (2017) argue that the *NUVrJ* diagram is better suited to identify recently quenched galaxies. However, the CANDELS catalogue does not include a *NUV* band (0.23 μm), so we could not counter-check the reliability of such technique. Of course, other more refined approaches - e.g. the analysis of the the main sequence where recently quenched objects are found in a transient position between star forming and passive objects - could be investigated, but the analysis would be in any case posterior to the SED-fitting, since the knowledge of the physical properties of the sources would be required.

In M18 we also showed that tailoring a reliable method to identify high-redshift passive objects is arduous anyway, because of the low signal-to-noise ratio (SNR) of such distant sources, even of the bright ones. This makes it challenging to compare the observed fluxes with template models of Spectral Energy Distributions (SEDs) or colors. In that work, we took advantage of the photometric data from the GOODS-South catalogue, complemented with new observations (Fontana et al. 2014) and new deep *Spitzer* mosaics; we exploited state-of-the-art techniques (T-PHOT, Merlin et al.

2015, 2016) and ad-hoc SED-fitting libraries built with constant (a.k.a. "top-hat") star formation histories (SFHs); and we used a stringent statistical approach to exclude potential false positives. In this way we ended up with 30 passive candidates at $z > 3$. However, we also showed how changing some properties of the stellar libraries, or letting the redshifts of the solutions vary, dramatically impacted the results, reducing the sample to 10 (including nebular line emission in the models) or even only two (letting the redshifts free in the fitting process) candidates. To strengthen the robustness of our selection and validate the basic assumptions, in Santini et al. (2019, S19 hereafter) we further checked the nature of the M18 candidates by means of archival ALMA data (available for 26 out of the 30 sources), statistically corroborating the passive classification of the sample from the lack of on-going star formation as seen at sub-mm wavelengths, free from the parameter degeneracies (especially the age-dust degeneracy) typical of the optical domain. Moreover, we could robustly confirm the individual passive nature of 35% of our candidates, adopting conservative assumptions. In M18 we also showed that upcoming facilities such as the *James Webb Space Telescope* will propel a leap forward, allowing for a much more robust photometric precision and, consequently, determination of physical properties. However, for now we can only trust the predictive and analytic power of the currently available instrumentation, and enlarge the statistical significance including more data.

To this aim, in the present paper we discuss the results from the joint analysis of the remaining four CANDELS fields (COSMOS, EGS, GOODS-North, and UDS). Since we used a refined grid of SED models, we also repeat the processing on GOODS-South. The five fields have different typical depths, therefore mixing the analysis might be risky, but we can safely consider the results as a lower limit to the actual number of passive objects above $z > 3$. Furthermore, in this work we address two more points: the concordance of the observations with the predictions from numerical models, and the impact that these early monsters had on the global SFH of the Universe.

The paper is organized as follows. In Section 2 we describe the dataset and we briefly summarize the method we used to single out the passive sample. In Section 3 we discuss the confirmation of the candidates by means of the available far-infrared (FIR) and spectroscopic data, and in Section 4 we discuss some properties of our candidate galaxies. In Section 5 we compute the number densities of our passive sample, and we compare our findings to the predictions of five state-of-the-art hydrodynamical models: ILLUSTRIS (Vogelsberger et al. 2014a), ILLUSTRIS-TNG100 and TNG300 (Pillepich et al. 2018), EAGLE (Schaye et al. 2015) and SIMBA (Davé et al. 2019). In Section 6 we present a method to compute the Star Formation Rate Density (SFRD) from the fitted SEDs, and compare the contribution of the red and dead populations to the total. Finally, in Section 7 we summarize and discuss the main findings of the work.

Throughout the paper, we assume a Λ-CDM cosmology ($H_0 = 70.0$, $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$), a Salpeter (1959) Initial Mass Function (IMF) except where noted otherwise, and AB magnitudes.

2 DATASET AND METHODS

For the GOODS-South field we use again the 19-bands catalog already discussed in M18, which improves on the original catalog published by Guo et al. (2013) as it includes three more bands (WFC3 *F*140*W* from the *Hubble Space Telescope* and VIMOS *B*, plus the deep HAWK-I *Ks* band presented in Fontana et al. 2014), and it has improved photometry on the *Spitzer* bands thanks to new mosaics (IRAC CH1 and CH2, by R. McLure) and new software (all four channels were re-processed using T-PHOT). As anticipated we decided to re-analyze the GOODS-South field taking advantage of refined SED libraries and redshift estimates (see below).

For the remaining four fields, we exploited the published CANDELS photometric catalogs, released in 2015 and presented in Nayyeri et al. (2017), Stefanon et al. (2017), Barro et al. (2019) and Galametz et al. (2013) for COSMOS, EGS, GOODS-North and UDS respectively. All catalogs are based on ∼ 20 wide bands, and in COSMOS and EGS they are complemented with some narrow and/or medium bands. Fluxes have been typically measured by means of SExtractor (Bertin & Arnouts 1996) aperture photometry for *Hubble* bands, after PSF-matching to the detection band *H*160; and template-fitting for ground-based and *Spitzer* bands, with TFIT (Laidler et al. 2007) or T-PHOT (Merlin et al. 2015, 2016)[1]. The typical depth of the detection band, WFC3 *F*160*W*, is ∼ 27.5 (5σ in 2 FWHM diameter). The properties of the five fields are summarized in Table 1; the cumulative area is ∼969.7 sq. arcmin.

As for the redshifts, we took advantage of the latest CANDELS estimates, to be presented in Kodra et al. (in preparation) which improve upon the original Dahlen et al. (2013) estimates; the new photo-*z*'s (where spec-*z*'s are not available) are obtained combining four independent estimations, using the minimum Frechet distance combination method.

In COSMOS, ID-16676 corresponds to the Z-FOURGE source 20115, a recognized passive source first discussed in Glazebrook et al. (2017), and object of a thorough study by Schreiber et al. (2018a), who showed how the presence of sub-mm flux is actually due to a strongly obscured close companion. Glazebrook et al. (2017) assigns to this source a spectroscopic redshift of 3.7172, different from both the CANDELS and the 3D-HST photo-*z*'s (4.127 and 3.545, respectively). We take their spec-*z* as the reference redshift of the object and use it in all our subsequent analysis. In other cases of spectroscopically confirmed redshift we kept

---

[1] In template-fitting tecniques, cutouts from the high-resolution detection band are used as priors to build low-resolution templates of the sources, by means of a convolution kernel that matches the PSFs of the two images. The templates are then used to solve a linear system minimizing the difference between a model collage and the real low-resoluion image, assigning to each source a multiplicative factor that best matches the observed flux. The method has proven to yield great improvements especially when the blending of the sources becomes important, as it is the case for ground based and mid/far-infrared bands. T-PHOT is the heir of TFIT; it improves on it in terms of accuracy, robustness and computational performance, and it includes a number of additional options. For a detailed description of the techniques and the codes, plase refer to the cited publications.

Table 1. Summary of the five CANDELS catalogs: number of photometric bands, 5σ limiting magnitudes in WFC3 F160W (the detection *H* band), area in sq. arcmin. In COSMOS, 20 out of 43 bands are medium/narrow bands from SuprimeCam and NEWFIRM; in EGS, 6 out of 23 bands are medium bands from NEWFIRM. In GOODS-North the estimate of the limiting magnitude corresponds to the Wide area; corresponding limiting magnitudes for the Intermediate and Deep areas are 28.2 and 28.7, respectively. A similar pattern applies to GOODS-South, where the estimate of the limiting magnitude corresponds to the Wide area; corresponding limiting magnitudes for the Deep and HUDF areas are 28.16 and 29.74, respectively. $m_{\rm lim}$ are typically defined as 5 × the average standard deviation within 2 FWHM circular apertures in empty regions of the fields.

| Field	Bands	$H_{\rm lim}$	Area
COSMOS	43 (20)	27.56	216.0
EGS	23 (6)	27.6	206.0
GOODS-N	18	27.8	173.0
GOODS-S	19	27.36	173.0
UDS	19	27.45	201.7

the CANDELS photo-*z* since the estimate was always sufficiently close (see Section 3.2).

2.1 Selection technique and results

As we did in M18, we proceed as follows to single out our red and dead candidates:

• we build a library of SED templates with top-hat SFHs, to better model the abrupt quenching of star formation in the very early Universe. The rationale for this choice is extensively discussed in M18. We improved upon the library used in the previous work by extensively refining the grid of models; the new library is described in detail in Appendix A;

• we consider all the objects from the *H*-detected catalogs of the five CANDELS fields at $z_{\rm CANDELS} \geq 3.0$, and we select the objects with *H* < 27 and with SNR > 1 in *Ks*, IRAC-CH1 and IRAC-CH2;

• on these lists of sources, we perform SED-fitting using our code ZPHOT (Fontana et al. 2000), in two flavors: (i) without the inclusion of nebular emission lines (we dub resulting selection "reference" sample), and (ii) with the inclusion of the lines (we dub this selection "lines" sample). We also made a third run, (iii) including the lines and letting the redshift of the fit vary ("*z*-free" sample), as in M18. However, S19 have shown that the third criterion is too conservative, given that in M18 only 2 over 30 galaxies survived this selection for the GOODS-South field, while 9 out of 26 sources have been robustly confirmed as passive by means of the follow-up analysis on ALMA data. Therefore, we only cite it here for the sake of completeness, but we will not discuss it further in the paper;

• we include in our lists of passive candidates only the objects having: (i) a passive best-fit model with probability $p_{\rm best} > 30\%$; and (ii) only star-forming solutions (if they have any) with probability $p_{\rm SF} < 5\%$ (these figures come from tailored simulations which are discussed in M18).
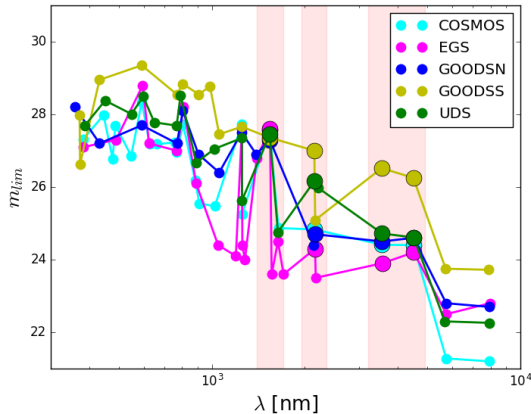
Figure 1. Compared 5σ limiting magnitudes for all the wide bands in the five fields, as given in the CANDELS papers cited in the text. Highlighted with vertical red strips and larger symbols are the bands that are crucial in our analysis, i.e. from left to right WFC3 $H160$ (the detection band, which has similar depths in all the fields), Ks, and the first two IRAC channels.

In this way we end up with the three selections summarized in Table 2. In total, we find 102 candidates in the "reference" sample, which become 40 in the "lines" selection. In Fig. 2 we show the positions of the "reference" candidates on the five CANDELS fields.

We point out that while the properties of the detection band ($H160$) are quite similar in the five fields, the same does not hold for what concerns the other bands. In Fig. 1 the nominal 5σ limiting magnitudes across the spectrum, as reported in the papers describing the catalogues, are compared: they show that the quality is far from uniform. This is particularly true for the K and *Spitzer* bands, which are crucial for the characterization of high-z objects:

- the K bands come from various surveys and facilities, so that their properties are very different in the five fields. For example, the HAWK-I Ks mosaic in GOODS-South is very deep ($m_{\lim} \simeq 27.0$, 5σ in 2 FWHM) with the finest seeing (FWHM$\simeq$0.4"), ensuring exquisite quality data. On the other hand, UDS HAWK-I Ks has FWHM$\simeq$0.4" and $m_{\lim} \simeq$ 25.9, the COSMOS VISTA Ks band has FWHM$\simeq$0.98" and $m_{\lim} \simeq 24.8$, EGS WIRCAM has FWHM$\simeq$0.65" and $m_{\lim} \simeq$ 24.3, and GOODS-North CHFT WIRCam has FWHM$\simeq$0.6" and $m_{\lim} \simeq 24.7$. The filter response functions are different as well;

- the IRAC bands also reach very different depths in the various fields: for example, CH1 reaches $m_{\lim} = 26.5$ at 5σ in GOODS-S, while it is limited to 24.7 in UDS, 24.5 in GOODS-N, 24.4 in COSMOS, and 23.9 in EGS.

Of course, these differences have a strong impact on the efficiency of our methods in the five fields, resulting in significant variations of the number of candidates, as Tab. 2 shows.

It is worth pointing out that the new redshift estimates change the selection for the GOODS-South field with respect to our sample in M18. Of the present 33 candidates, 26 are in common with the previous selection; 4 are now excluded (IDs 5592, 9091, 26802, 10759) and 7 new are in-

cluded (IDs 3718, 4202, 4949, 5934, 13394, 16526, 19883). All of the changes are due to variations in the probabilities of star forming solutions in the SED-fitting procedure caused by the different photo-z. Since we checked that our candidates were reliable both in M18 and in S19, this seem to imply that our selection criteria are conservative, and the true number of passive objects is probably higher than these estimates.

We also note that IDs 2075 in COSMOS; 24177 in EGS; 157, 643, 6620, 9626, 13007, 24572 and 27251 in GOODS-North; 4949, 6407, 12178 and 19446 in GOODS-South, show complex morphologies in the optical, H or K bands, and possibly have close companions which might cause strong contamination despite the robustness of the adopted photometric methods.

3   CONFIRMATION OF CANDIDATES WITH FAR-INFRARED AND SPECTROSCOPIC DATA

In this Section we study the properties of our selected candidates using their far-infrared photometry and spectroscopic data, when available.

3.1   *Herschel* fluxes

As in M18, we have cross-matched the positions of our selected candidates with *Herschel* public catalogs, to check for degeneracies and possible misinterpretations of low-z dusty galaxies as high-z passive dust-free sources. FIR data is available for all CANDELS fields.

For the two GOODS fields we take advantage of the new, deep Astrodeep catalogs by Wang et al. (in preparation), which combine data from PEP (Lutz et al. 2011) and GOODS-Herschel (Elbaz et al. 2011) surveys with the PACS camera, and the HerMES survey (Oliver et al. 2012) with SPIRE. The catalogs are extracted using H-band CANDELS priors, making the cross-correlation very handy. We find that all the new candidates in GOODS-S have no match, while we refer the reader to M18 for the discussion on the two candidates having a FIR-counterpart (IDs 3973 and 10578). The fact that these two sources, also showing X-ray emission, are among the ones robustly confirmed by ALMA (see S19), consolidates their interpretation as galaxies that have been recently quenched by strong feedback from the central active nucleus, the latter likely being the responsible for the detected FIR emission. We find one possible match among GOODS-N candidates (ID-35028), having a 3.5σ flux at 100 μm and a 1.5σ flux at 160 μm. However, its flux may be strongly blended with that of a very close-by source, only 0.8" apart, with a lower redshift ($z \sim 1.8$, more common for *Herschel* detections) and a low level of on-going SF (0.2 $M_{\odot}$/yr according to the fit with standard exponentially declining star formation histories, or τ-models).

For the other fields we used the HerMES DR4 (COSMOS and EGS) and DR3 (UDS) catalogs (Roseboom et al. 2010, 2012) for the SPIRE bands, while for PACS bands we used the PEP prior-based catalogs (COSMOS and EGS) and the catalogs compiled by the HerMES team (UDS). The only source possibly showing some FIR emission is ID-2075 in COSMOS: it is associated to a source at a distance of

Table 2. Number of sources in the five CANDELS fields, and in total. Left to right, we list: the total number of detected sources, the number of sources at $z_{\text{CANDELS}} > 3$, the number of sources at $z_{\text{CANDELS}} > 3$ with SNR>1 in K, IRAC1 and IRAC2; and the number of selected candidates in the two passive selections: the "reference" sample and the "lines" sample obtained including nebular lines in the library.

Field/Sample	Total	$z > 3$	$S/N_{z>3} > 1$	Reference	Lines
COSMOS	38671	3778	1525	4	2
EGS	41457	4830	1775	13	5
GOODS-N	35445	3953	1793	36	11
GOODS-S	34930	5029	2884	33	13
UDS	35932	4018	2540	16	9
All fields	186435	21608	10517	102	40

2" having a 3σ flux at 250 μm, but the flux estimate could be contaminated by a brighter, star-forming source 6" apart (according to the fit with τ-models; we remind that the PSF at 250 μm is 18").

Given these results, and considering the high fraction of potential mis-associations due to the large *Herschel* PSFs, and the contamination from nearby sources, we can conclude that we did not find any clear evidence for FIR emission for any of our "reference" candidates, with only a few moderately uncertain cases.

3.2 Spectroscopic data

Checking the new CANDELS catalogs to be presented in Kodra et al., we find that the following candidates have spectroscopic redshifts (which were used in our analysis): IDs 2490 and 6539 in EGS, IDs 20589 in GOODS-North, IDs 10578 and 16526 in GOODS-South, and ID 8689 in UDS. Among these, only GOODSS-10578 enters our "lines" selection, while all the other are only in the "reference" sample.

We searched the VANDELS (Pentericci et al. 2018; McLure et al. 2018) and VUDS (Le Fèvre et al. 2015) databases to visually inspect the available spectra for GOODS-South, finding data for additional sources with respect to the CANDELS catalogue. All the spec-z's are consistent with the CANDELS estimates, unless explicitly specified. We already discussed IDs 4503, 9209 and 10578 in M18 (Sect. 4.4). To these, we can now add from the DR2: IDs 4949 and 5934, which show no evident features; ID-12178, which has a strong emission line and is classified at $z_{\text{spec}} = 0.56$ (while $z_{\text{CANDELS}} = 3.29$), but might be spectroscopically contaminated by a very close companion (see Appendix C); and IDs 16526 and 19505, both showing moderate Lyman-α emission, but no other evident features.

Many candidates in GOODS-South were also observed with MUSE, as part of the MUSE-Wide (Urrutia et al. 2018) and MUSE Deep (Bacon et al. 2017; Inami et al. 2017) GTO programs (while all our COSMOS candidates all fall outside the MUSE-Wide COSMOS footprint). Most of the candidates covered by the 1 hour deep extended MUSE-Wide data (GOODS-South IDs: 3912, 4587, 4949, 5934, 6407, 7526, 7688, 8242, 8785, 9209, 12178, 17749, 18180, 19301, 19446) show no signs of Lyα emission, in support of these sources candidacy as passive galaxies. Although in general the spectra are not deep enough to provide conclusive evidence, the absence of detectable emission lines strengthen our conclu-

sions. For IDs 9091, 12178 (already mentioned above) and 13394 there is line emission detected within 1.0", 0.5" and 0.5" in the MUSE-Wide data, respectively. These emission lines are however associated with different foreground objects at redshifts 1.33, 0.56 and 0.42 (MUSE-Wide DR1 ID: 143003008). Three candidates in GOODS-South fall in the MUSE Deep footprint. ID-15457 has no emission detected in the MUSE Deep data; approximately 0.5" from ID-16506, a Lyα emitter is detected at $z = 3.33$. Lastly, ID-10578 is confirmed as an AGN by the MUSE Deep data, as already discussed in M18.

Based on these results, we conclude that we find no strong spectroscopic evidence to exclude any of our candidates from the selected sample.

4 PROPERTIES OF THE SELECTED SAMPLE

The main physical properties of all the candidate objects, as obtained in the SED-fitting procedure with the top-hat library, are given in Appendix B. We provide snapshots and SEDs of all the candidates in the on-line supplementary material, and we show a few significant examples in Appendix C. Here we give a brief summary of the global properties of the candidates.

First of all, we compared the properties of our 102 "reference" candidates with those from the 3D-HST catalogs (Skelton et al. 2014). Most of the 3D-HST photometric redshifts, obtained with EAzY (Brammer et al. 2008), are in reasonably good agreement with the CANDELS ones ($\Delta z < 0.3$), with the exceptions of the 19 sources listed in Tab. 3 (inconsistencies between the two catalogs can be due to many factors, including the usage of different photometric methods and photo-z codes). None of the discrepant objects has a spectroscopic redshift estimate; most of them have lower photo-z in 3D-HST than in CANDELS, with 8 candidates that would be excluded from our selections having $z_{\text{3D-HST}} < 3$ and one (ID-13 in GOODS-North) lacking a reliable photo-z estimate.

We note that in GOODS-South 6 out of 9 discrepant objects are among our strongest candidates (i.e. they belong to the "lines" selection), with three among them having been individually confirmed as passive sources by the analysis in S19 on ALMA data: IDs 9209, 17749 and 18180. However, in 3D-HST their photo-z is still above 3, and all have $sSFR < 10^{-11}$ yr^{-1} (estimated with FAST, Kriek et al.
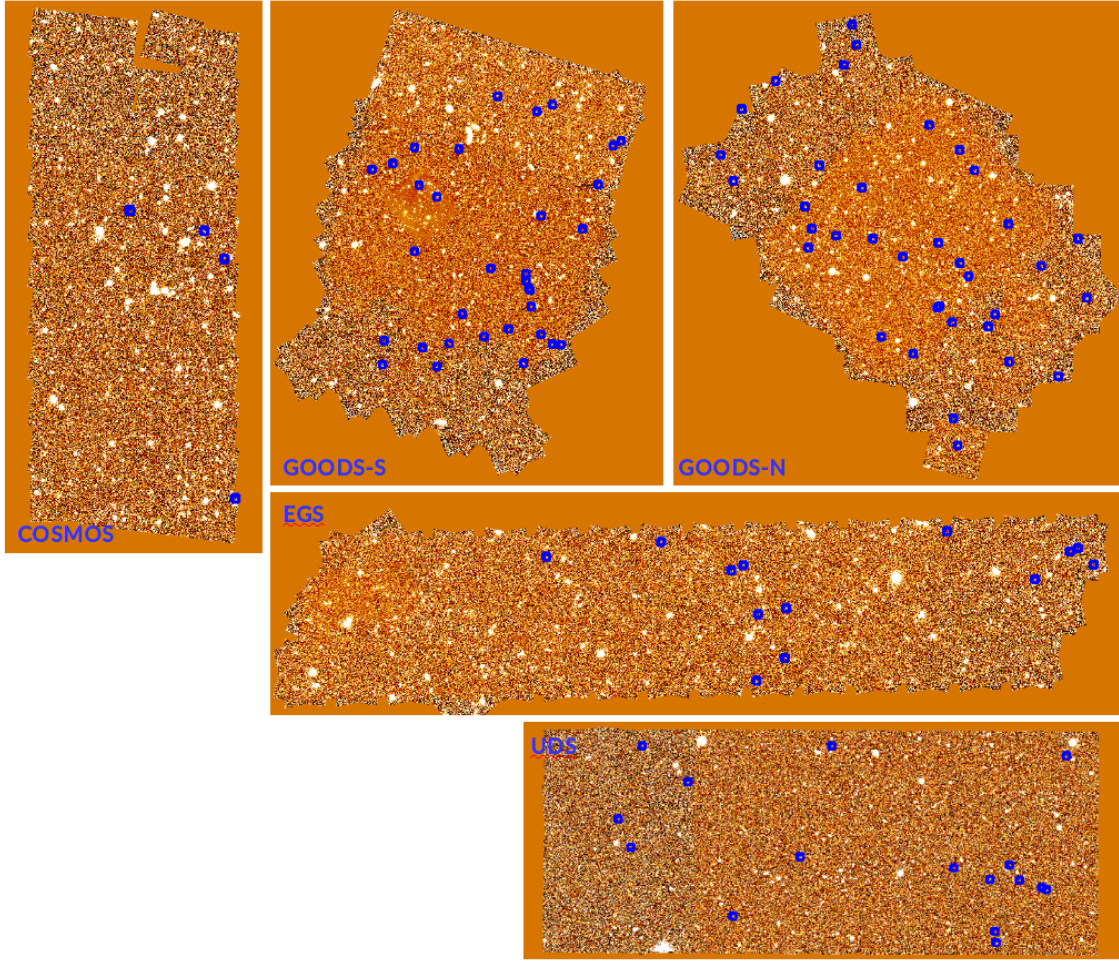
Figure 2. Positions of the passive candidates in the five CANDELS fields (shown is the detection band, WFC3 *H*160). The dimensions of the images are not scaled to the real dimensions of the fields. We checked that the sources close to borders have complete image covering. Snapshots of all the candidate galaxies are given in the on-line supplementary material.

2009), therefore being robust candidates in their analysis as well. More in general, while $\sim 38\%$ of our candidates having $z_{\text{3D-HST}} > 3$ can be labelled as passive using 3D-HST FAST estimates and the hard threshold $sSFR < 10^{-11}$ yr^{-1}, in $\sim 40\%$ of cases the sSFR is above 10^{-10} yr^{-1}. We note that in contrast, using CANDELS estimates on SFR and masses, we find $\sim 14\%$ and $\sim 47\%$, respectively.

4.1 UVJ diagram

Fig. 3 shows the position of our passive candidates on the *UVJ* diagram. The region of passively evolving objects is delimited as in Whitaker et al. (2011). Grey small dots correspond to the whole sample of $z > 3$ galaxies in the CANDELS fields, while the passive candidates in the "reference" sample are plotted as large dots; empty squares mark the candidates belonging to the "lines" selection as well.

We point out that for this plot we use the rest-frame colors obtained fitting their observed SEDs with a standard exponentially declining SFH (τ-models): this choice is motivated by the fact that we want to check whether the *UVJ* color selection method, straightforwardly applied, can be considered a reliable approach. However, in M18 we showed

that the colors obtained with our top-hat SFHs are quite similar, with small shifts in the $V-J$ color which we ascribe to the less constrained photometry in the observed redder part of the spectrum (the two 5.6/8.0 μm IRAC bands have the poorest SNR), with respect to the visible and near-infrared (NIR) bands which straddle the rest-frame $U - V$ break at $z \sim 3$. We thus showed how the *UVJ* selection criteria is certainly powerful, but at these high redshifts it can miss a number of interesting candidates, in particular many recently quenched objects which still show bluer colors than the typical red passive galaxies. On the same note, Schreiber et al. (2018b) claimed that the *UVJ* selection tends to be pure (although with a $\sim 20\%$ failure rate) but incomplete, in that a fraction of $\sim 40\%$ more quiescent galaxies can be identified using e.g. their specific star-formation rate (sSFR) fitted value.

Here too we see that many candidates lie outside the passive region of the diagram. It is interesting to note that most of the outliers belong to two fields, GOODS-S and GOODS-N, which have high quality data in the infrared bands, in particular considering IRAC CH3 and CH4: this seems reasonable, since their SED-fitting is better constrained, yielding less possible solutions and therefore ex-

Table 3. CANDELS vs. 3D-HST photometric redshift outliers ($\Delta z \geq 0.3$). The symbol † highlight the sources that have $z_{\text{3D-HST}} < 3$.
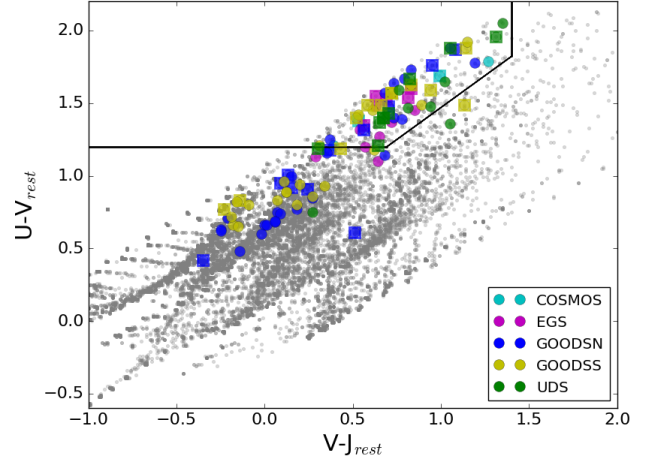
Field	ID_{CAND}	z_{CAND}	ID_{3D}	$z_{\text{3D-HST}}$
COSMOS	16676	4.127	19670	3.5446
EGS	25724 †	3.795	33354	2.8674
GOODSN	13	3.014	5	-
GOODSN	1570 †	3.226	2212	0.1497
GOODSN	5744	3.459	8249	3.0668
GOODSN	10672 †	6.713	15083	1.7364
GOODSN	13403 †	3.793	18781	0.5635
GOODSN	21034	3.328	28997	4.2929
GOODSN	28344	4.758	3484	4.3757
GOODSN	35028 †	3.642	34844	1.2655
GOODSS	2608	3.720	4857	3.4116
GOODSS	3912	3.897	7177	4.9983
GOODSS	7526 †	3.317	15494	2.647
GOODSS	8785 †	3.852	17894	2.5811
GOODSS	9209	4.486	18684	4.8638
GOODSS	16506 †	3.382	30821	0.3737
GOODSS	17749	3.697	32872	3.1606
GOODSS	18180	3.650	33566	3.3366
GOODSS	19301	3.592	35502	4.0423
UDS	25893	4.491	41274	3.8881

Figure 3. UVJ selection diagram. The solid lines enclose the region of passive sources, as identified in Whitaker et al. (2011). Grey dots: all CANDELS $z > 3$ sources. Large coloured dots: passive "reference" selection; the dots with a square are also part of the "lines" selection. The rest-frame colors are obtained with τ-models fits. The points follow regular patterns because they correspond to the colors of the templates in the SED-fitting library, which have well determined values.

cluding potential star-forming fits. This seems to indicate that in the other fields we are probably underestimating the real number of passive sources.

More generally, we also note that many objects fall inside the passive region of the diagram, but do not belong to our samples. We ascribe this fact to our stringent criteria, which are tailored to select strictly *passive* objects, rather than *quiescent* ones (we emphasize that we use the term "quiescent" to dub sources that retain a weak star-formation activity, as opposed to completely "passive" ones). As already noted, it might be too conservative excluding potentially reliable candidates because of the mere existence of a few possible, albeit improbable, star-forming solutions in the SED-fitting procedure.

4.2 Diagnostic diagrams

Figure 4 shows the stellar mass vs. redshift plane. In this case, for the CANDELS full catalog we consider the masses obtained in the best fit with the "delayed" τ-models, for which SFR $\propto (t^2/\tau) \times exp(-t/\tau)$ (Santini et al. 2015), while for the red and dead sample we use the ones of the best top-hat fit from this work. We applied a correction factor of 1.75 to make the CANDELS mass estimates, which assume a Chabrier (2003) IMF, consistent with ours, which assume a Salpeter IMF. Most of the "reference" candidates

have $M \geq 10^{10} M_\odot$, which is unsurprising since our selection criteria only select luminous (i.e. high SNR) sources. In particular, we find two candidates at $z > 4$ with masses above $10^{11} M_\odot$: UDS-10430 and UDS-25893.

We then consider the half-light radii of $3 < z < 5$ sources, as measured by the SEXTRACTOR detection runs in the $H160$ band, converting them from pixel space to arcseconds and then to proper kpc at the relevant redshifts by means of appropriate Astropy (Astropy Collaboration et al. 2013; Price-Whelan et al. 2018) routines.

We do not take into account the k-correction factor which would be necessary wince we consider galaxies at different redshifts, implying that the observed H band samples different regions of the rest-frame spectrum (i.e., 400 nm at $z = 3$, 320 nm at $z = 4$ and 267 nm at z=5). However, we have checked that the sizes measured on images corresponding to bands sampling the same region of the spectrum (i.e., Y-band at $z = 3$ and J-band at $z = 4$, at 267 nm rest-frame; all images were PSF-matched to H) are very similar to the H band sizes, and considering that the Y-band data is not always available we preferred to simply use H band data. A possible drawback of this choice is that in this way we are looking at the rest-frame B band, which is not particularly well suited to study the size of a red object.

The resulting mass vs. radius relation is plotted in Figure 5. In general, the passive candidates appear to be typically massive but mostly concentrating towards the region of small radii. Also, a dependence on the redshift seems to be present, as highlighted by the size-coding of the dots and in the lower panel of the Figure: earlier sources appear to be more compact than later ones. The typical radii of these objects are roughly consistent with the ones found for other high-redshift samples of passive objects, $R_e \sim 1$ kpc, therefore showing a general compactness with respect to local

galaxies (e.g. Cimatti et al. 2008, for $1.4 < z < 2$ passive galaxies in a similar mass range).

4.3 A massive red and dead galaxy at $z \sim 6.7$?

We find a particularly interesting object among the galaxies in our selection: GOODS-North ID-10672 has $z_{CANDELS} = 6.713$, and it enters both the "reference" and "lines" samples. Its mass inferred from the SED-fitting is $4 \times 10^{10} M_\odot$, with an age of ~ 500 Myr, implying a formation redshift of ~ 14, i.e. ~ 300 Myr after the Big Bang; the fitted burst duration is of just 300 Myr, implying an average SFR of $\simeq 130 M_\odot/\mathrm{yr}$. With a typical τ-model SFH, it is fitted with $sSFR = 5.21877 \times 10^{-11}$ yr^{-1}, so it would be classified as quiescent, but it would fail a strictly passive selection based on a standard $sSFR < 10^{-11}$ yr^{-1} criterion. The object has been observed and catalogued in many surveys (e.g. Finkelstein et al. 2015; Bouwens et al. 2015; Harikane et al. 2016) as a $z \sim 6 - 7$ mildly star forming or quiescent galaxy. As already pointed out (Tab. 3), in 3D-HST it is instead classified a $z = 1.73$ source by means of EAzY photometric redshift estimate (without grism information). We find that no detectable signal is associated to this galaxy in MIPS, PACS and SPIRE (i.e. 24 to 500 μm): while some flux is detected within a few arcsec from the source position, it can quite confidently be attributed to other nearby star-forming objects. However, from a SED-fit with redshift fixed at the 3D-HST value of 1.73 we obtain SFR values of $\sim 3 - 8 M_\odot/\mathrm{yr}$ (depending on the details of the adopted models), which is too low to be detected in MIPS or *Herschel*. Another possibility might be that the source is actually a cold star, possibly a brown dwarf; instead, we tend to rule out the options of a young and obscured star or of an AGB star, because the source is very isolated, which would be unusual for these kind of objects. A spectroscopic observation would be important to definitively rule out alternative possibilities, but the very faint continuous emission would make the analysis very difficult and perhaps unfruitful. Snapshots and best-fitting SED of this objects are show in Appendix C; a more extended analysis to exclude the possibility of degeneracies in the fitting would be required, but is beyond the scope of the present work and we leave it to future work.

4.4 Comparison with other selections

In M18 we compared our results on GOODS-South to the ones by Straatman et al. (2014, S14) and Nayyeri et al. (2014, N14). We repeat the analysis now, given that our sample has sligthly changed because of the new redshift estimates. Cross-correlating the CANDELS catalogue with the ZFOURGE selection in S14, we now find that four out of six sources in the S14 selection belong to our sample as well: ID-19883 has now entered both our "reference" and "lines" selections, together with IDs 4503, 17749 and 18180 which were already present in M18. As for the N14 selection, the situation in unchanged, with five sources in common and the other 11 rejected in our analysis. We refer the reader to M18 for other considerations about these comparisons.

More recently, Schreiber et al. (2018b) used MOSFIRE H and K spectra to confirm the quiescent nature of 22 galaxies (out of 24 reduced spectra) detected in COSMOS, EGS

Table 4. Matches between the ZFOURGE/3D-HST and CANDELS IDs in the Schreiber et al. (2018b) sample.

ZFOURGE / 3D-HST	CANDELS
ZF-COS-20115	COS 16676
3D-EGS-18996	EGS 14727
3D-EGS-31322	EGS 24177
ZF-UDS-3651	UDS 1244
ZF-UDS-4347	UDS 2571
ZF-UDS-7329	UDS 7520
ZF-UDS-7542	UDS 7779
ZF-UDS-8197	UDS 8682
ZF-UDS-41232	UDS 25688
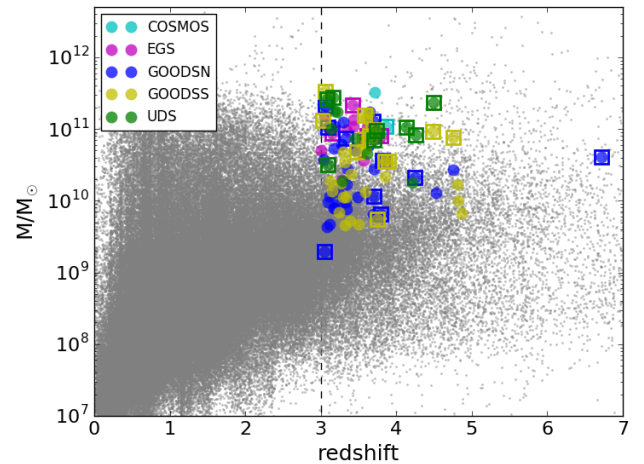
Figure 4. Stellar mass vs. redshift of the selected candidates in all CANDELS fields. Grey points show the full CANDELS catalogues, with "delayed" τ SFH mass estimates from Santini et al. (2015), corrected to account for the Chabrier IMF (our fit assumes a Salpeter IMF). Full circles correspond to the objects from "reference" sample; empty squares mark the objects from the "lines" selections. For the passive candidates we plot the masses from the best top-hat fit in the "reference" run.

and UDS, based on ZFOURGE and 3D-HST catalogs. These sources had been previously singled-out as quiescent candidates via UVJ selection. We cross-matched the ZFOURGE, 3D-HST and CANDELS catalogs RA-DEC coordinates, and found that out of the 22 objects, only 9 are in our selection as well (see Tab. 4). We could not find matching objects for three spectroscopic sources (ZF-COS-17779 is outside the HST footprint, ZF-UDS-35168 and ZF-UDS-39102 are not detected in CANDELS); among the other 9, two have $z_{CANDELS} < 3$, while the other 7 do not match our selection criteria (having some SF solutions; we must point out that they select quiescent rather than totally passive sources).
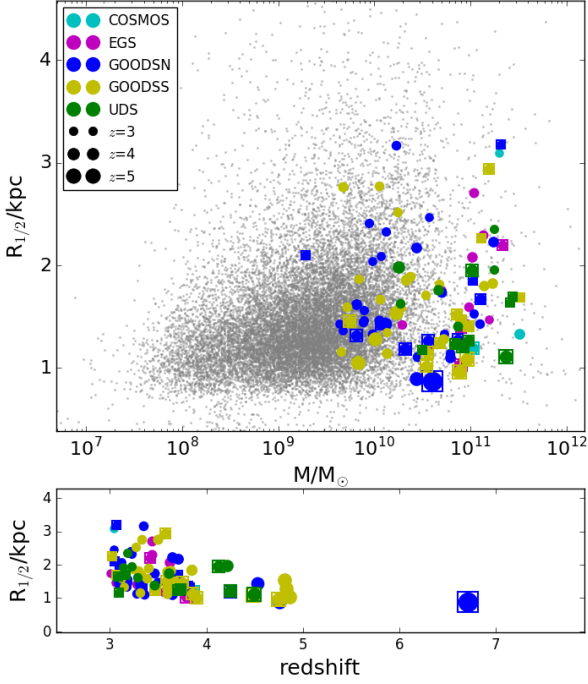
Figure 5. Top: mass-radius relation. Grey dots: all CANDELS sources at $3 < z < 5$; colored circles: passive candidates, size-coded for redshift (see legend); empty squares mark the objects in the "lines" selection. The radii are the half-light radius in H band from the SEXTRACTOR runs, converted into proper kpc. The masses are the "delayed" τ SFH estimates from Santini et al. (2015), corrected to account for the Chabrier IMF (our fit assumes a Salpeter IMF). Most of the candidates are on the compact tail of the distribution, with a clear trend with redshift: high redshift sources are smaller compared to $z \sim 3$ sources, as show in the bottom panel, where we plot the radii of the candidates against their redshift.

5 NUMBER DENSITY

We now try to estimate the number density of passive galaxies, dividing the number N of our red and dead candidates observed within a redshift interval of interest Δz by the cosmological volume V corresponding to such interval, re-scaled by the ratio between the total area of the survey (~ 969.7 sq. arcmin) and that of the full sky. We compute the errors as the poissonian uncertainties $\sqrt{N}/V$, which we assume do take into account cosmic variance given the fact that we are considering five different realizations (fields). As a reference result, we first compute the number densities yielded by our list of candidates, without considering any correction for incompleteness; we will discuss and complement this in the next subsection.

The resulting number density of passive galaxies in the redshift interval $3 < z \leq 5$, considering the sum of our samples in the five fields, is $1.73 \pm 0.17 \times 10^{-5}$ for the "reference" selection ($6.69 \pm 1.08 \times 10^{-6}$ for the "lines" selection) Mpc^{-3}. These values would change to $2.16 \pm 0.22 \times 10^{-5}$ ($8.52 \pm 1.54 \times 10^{-6}$) Mpc^{-3} if we excluded COSMOS from the average, a potentially reasonable choice given that this field yields significantly lower estimates with respect to the other
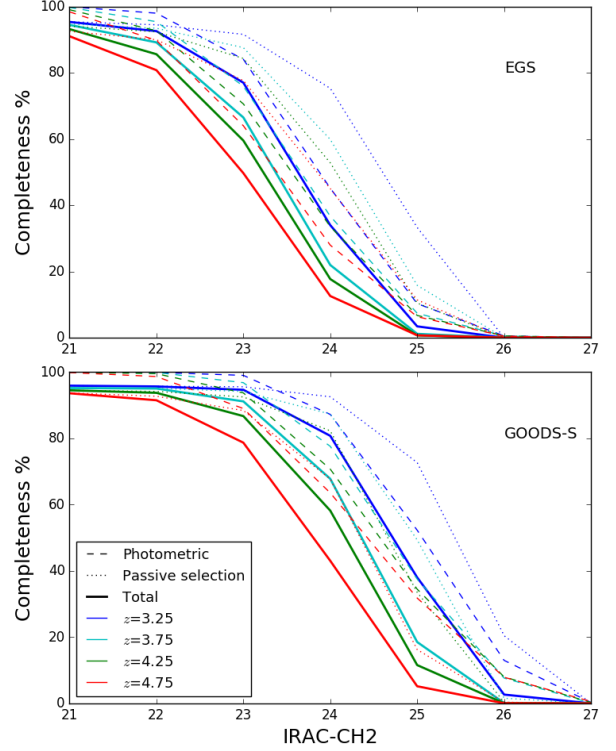

Figure 6. Estimated completeness of the passive selection as a function of IRAC-CH2 observed magnitude, as estimated from a set of dedicated simulations, for the two fields with the best and worst data quality (upper panel: GOODS-South; lower panel: EGS). In each panel, the dashed lines correspond to the photometric selection completeness, the dotted lines to the passive candidates selection completeness (computed considering only the photometrically selected galaxies), and the solid lines to the total completeness given by the product of the two. The colors of the lines correspond to different redshift bins. See text for more details.

Table 5. Number densities (Mpc^{-3}) of passive candidates in the total CANDELS area (~ 969.7 sq. arcsec), considering three redshift intervals, for our two selections "reference" (both as observed and after correction for completeness) and "lines".

Δz	Reference		Lines
	Observed	Corrected	
$3 < z \leq 5$	1.73×10^{-5}	2.30×10^{-5}	6.69×10^{-6}
$3 < z \leq 4$	2.90×10^{-5}	3.66×10^{-5}	1.08×10^{-5}
$4 < z \leq 5$	4.34×10^{-6}	7.94×10^{-6}	2.17×10^{-6}

four (which is probably due to cosmic variance, but also to worse photometric properties like e.g. the broad FWHM in Ks, see Sect. 2.1). However, in the following we will stick to the estimates obtained averaging on all the five fields.
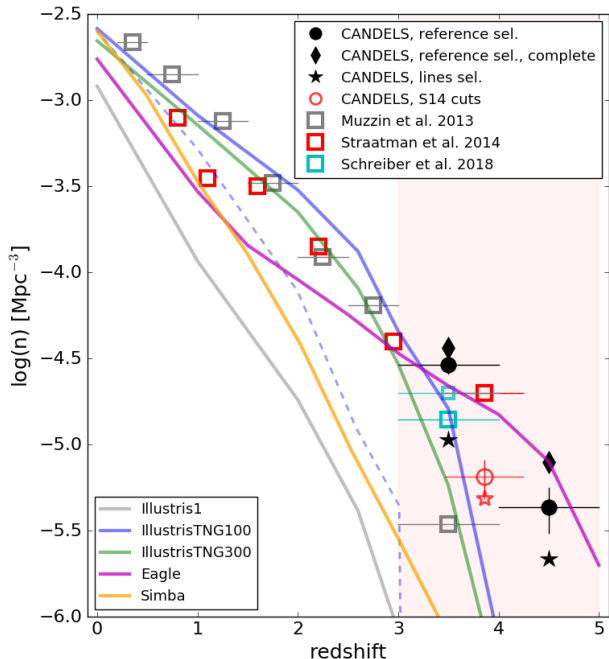
Figure 7. Number density of passive galaxies as a function of redshift. We plot many different estimates, both from observations (Muzzin et al. 2013; Straatman et al. 2014; Schreiber et al. 2018b) and from hydrodynamical simulations (from which we take the galaxies with $M_*/M_\odot > 5 \times 10^9$): ILLUSTRIS-1 (Vogelsberger et al. 2014a), ILLUSTRIS-TNG100 and ILLUSTRIS-TNG300 (Pillepich et al. 2018), EAGLE (Schaye et al. 2015, run RefL0100N1504 with full physics included), and SIMBA (Davé et al. 2019). See text for relevant information on these datasets and models. We then plot the results based on our SED-fitting method for the "reference" sample, considering the cuts in mass and redshift adopted by S14 (red empty circle), and the full $3 < z < 5$ interval divided in two bins considering all masses, before (black circles) or after (black diamonds) correcting for completeness as described in Sect. 5.1. We also plot the values corresponding to the "lines" selections (black or red stars). The shadowed area highlights the relevant redshift interval. The agreement between all observations and the ILLUSTRIS-TNG models is reasonable up to $z \simeq 4$, while it worsens at earlier epochs; on the other hand, EAGLE yields better results at $z > 4$, but underestimates the number of passive objects at $z < 3$. Finally, SIMBA seem to fall short at all epochs, however doing better than the original ILLUSTRIS-1 model. See text for more details.

Table 6. Number densities (Mpc^{-3}) of passive candidates in the five CANDELS fields, considering the cosmological volume $3 < z \le 5$, for our two selections "reference" (both as observed and after correction for completeness) and "lines".

Field	Reference		Lines
	Observed	Corrected	
COSMOS	3.08×10^{-6}	3.50×10^{-6}	1.54×10^{-6}
EGS	1.05×10^{-5}	1.34×10^{-5}	4.04×10^{-6}
GOODS-N	3.36×10^{-5}	4.38×10^{-5}	9.62×10^{-6}
GOODS-S	3.17×10^{-5}	4.44×10^{-5}	1.25×10^{-5}
UDS	1.32×10^{-5}	1.75×10^{-5}	7.42×10^{-6}

5.1 Completeness

It is not easy to try and quantify the completeness of our sample of passive candidates, given the different depths of the five fields, and the subtleties of the technique we adopted. We attempt to do so proceeding as follows. Using ZPHOT, we create a library of 3000 $3 < z < 5$ synthetic spectra with top-hat SFHs, having different ages and duration of SF bursts so that 1740 are passive and 1260 are star forming. Then, we consider the two fields with the highest and the lowest quality of data in terms of depth, namely GOODS-South and EGS, respectively; by means of the in-house software SIMULCAT, we use the synthetic models to create two simulated observed catalogues, having the properties of the two fields in terms of filters and depths (i.e., signal-to-noise ratios) at all magnitudes and in all the observed bands, and fixing a reference filter (we choose IRAC-CH2) to predefined magnitude values, i.e. $m_{I2} = 21, 22, 23, 24, 25, 26, 27$. In both cases, each model is replicated ten times with a slightly different noise realization, so each catalogue finally contains 17400 passive objects. Finally, we fit these mock observed catalogs with our top-hat library, and proceed to select passive candidate following exactly the same procedure we used on the real data.

The results for the two fields are show in Fig. 6. In both panels, the solid lines show the inferred completeness in four redshift bins, as a function of the reference magnitude in IRAC-CH2. This completeness is the result of the product of two factors: the photometric completeness, which comes from the pre-selection we perform on the observed catalogue, as described in the second bullet of the list at the beginning of Sect. 2.1 (namely: $H < 27$, $SNR_{K, IRAC1, IRAC2} > 1$, plus an additional condition on the SNR of the detection band to take into account the typical detection cut in the original input catalogs, $SNR_H > 4$), shown as dashed lines in the Figure; and the passive-selection completeness, which we computed considering only the sources that survived the photometric selection, and is shown with dotted lines. Therefore, the product of the two quantities is self-consistent as a total completeness.

As expected, the values drop as a function of the magnitude and of the redshift. In the case of GOODS-South, a 50% completeness is reached at $m_{IRAC2} \simeq 24.75$ at $3 < z < 3.5$, and at $m_{IRAC2} \simeq 23.75$ at $4.5 < z < 5$; in EGS, the 50% completeness is reached at ~ 1 magnitude brighter. We note that we are $\sim 95\%$ complete in GOODS-South at $m_{IRAC2} \simeq 23$ at $z = 3$.

We now use these values to estimate an inferred "true" number densities of sources (we only do so for the "reference" selection, for the sake of simplicity). We multiply the number of objects we actually find in each magnitude and redshift bin by the inverse of the corresponding completeness estimate; as a reasonable approximation of the complex features of the different fields, we use the values that we obtained using GOODS-South for the two GOODS catalogues, and the estimates computed on EGS for the other three. To obtain a fair estimate, we take 50% as the minimum reliable value, and when the completeness drops below this threshold we continue to use it to compute the actual multiplicative factor. Also, if a bin of magnitude and redshift contains zero objects, its counts will of course remain zero. These two points imply that once again we are be-

ing conservative, and the obtained number density can still be underestimated. With this approach, we end up with a corrected total number density of $2.30 \pm 0.20 \times 10^{-5}$ for the whole redshift interval $3 < z < 5$ (for the sake of reference, we point out that if we decided not use the 50% completeness threshold, and instead we used the full completeness functions as obtained from the simulations, we would get a number density of $3.25 \pm 0.24 \times 10^{-5}$).

In Tab. 5 we list the number densities both before and after the completeness correction, for the whole survey area, in three redshift intervals; while in Tab. 6 we give the number densities corresponding to the individual cosmological volumes of the five fields, in the full redshift interval $3 < z \leq 5$. GOODS-South yields the higher value (by a factor of ~ 2 with respect to the average of the five fields, in the "reference" selection), likely because of a better constrained photometry in the infrared bands leading to more robust SED fits. Cosmic variance also plays a role: the number of all $z > 3$ sources varies by a factor of ~ 1.5 between the five fields (see Tab. 2).

5.2 Comparison with predictions from numerical simulations

To understand how the values we have obtained fit in the current theoretical scenario, we have compared the number density of our passive candidates at high redshift with the estimates obtained in five recent cosmological hydrodynamical simulations: ILLUSTRIS (Vogelsberger et al. 2014b,a; Genel et al. 2014; Nelson et al. 2015), ILLUSTRIS-TNG100 and TNG300 (Pillepich et al. 2018; Nelson et al. 2019), EAGLE (Schaye et al. 2015), and SIMBA (Davé et al. 2019). Full simulations data containing particles and groups information are publicly available and downloadable for the first four, while SIMBA data were privately provided by the authors.

The ILLUSTRIS and ILLUSTRIS-TNG simulations exploit the moving-mesh code AREPO (Springel 2010); simulated volumes and baryonic mass resolutions are as follows: 106.5^3 Mpc^3, $1.26 \times 10^6 M_\odot$ for ILLUSTRIS; 110.7^3 Mpc^3, $1.4 \times 10^6 M_\odot$ for ILLUSTRIS-TNG100; 302.6^3 Mpc^3, $1.1 \times 10^7 M_\odot$ for ILLUSTRIS-TNG300. The EAGLE simulations exploit GADGET-3, the the latest incarnation of the original Tree-SPH (Smoothed Particle Hydrodynamics) code developed by Springel (2005); we used data from the RefL0100N1504 simulation, which is the most complete in the suite of runs in terms of physics included in the code; the run simulates a volume of 100.0^3 Mpc^3, and has a baryonic mass resolution of $1.81 \times 10^6 M_\odot$. Finally, SIMBA is based on the GIZMO code by Hopkins (2015), a mesh-free finite-mass hydrodynamic code which handles shocks via Riemann solvers, with no need for artificial viscosity; the simulation also includes detailed and novel recipes for AGN feedback, and has a volume of 100^3 Mpc^3, with a baryonic mass initial resolution of $1.82 \times 10^7 M_\odot$, i.e. comparable to the TNG-300 simulation. All the models include baryonic sub-grid physics to simulate star and black hole formation, stellar and AGN feedback, and metal enrichment; while ILLUSTRIS, ILLUSTRIS-TNG and SIMBA model subgrid physics from first principles, EAGLE use empirical relations to match observed properties of galaxies at $z = 0$. All simulations assume a Chabrier IMF, and we applied again the corrective factor of 1.75 to com-

pare with the Salpeter IMF adopted in our SED-fitting procedure. We also paid attention to consider the cosmological factor $h = H_0/100$ in the computation of masses and SFRs, when necessary (some simulations define masses and lengths in units of h, others do not).

For the models, we considered simulated galaxies with $M_*/M_\odot > 5 \times 10^9$, and we adopted the usual selection criterion sSFR$< 10^{-11}$ yr^{-1}. Since in the ILLUSTRIS and ILLUSTRIS-TNG data releases multiple values of masses and SFRs are given for each simulated object, depending on the different radii within which they are computed, we tried to mimic as accurately as possible the observational approach: to this aim, we used the values estimated within twice the half mass radius of each object (see also the analysis in Donnari et al. 2019). On the other hand, EAGLE outputs masses and SFRs within a set of apertures (defined as the diameters of spheres centered on the position of the object) of fixed proper lengths, from 1 to 100 proper kpc; and two estimations of the half mass radius, $R_{1/2,30}$ and $R_{1/2,100}$, computed within 30 and 100 kpc respectively. To obtain a fair comparison with the ILLUSTRIS and TNG cases, in this case we compute the sSFR by considering, for each simulated galaxy in a snapshot, the mass and SFR within the aperture which is closest to $4 \times R_{1/2,100}$. Finally, the values of the sSFR within twice the half mass radius of each object for SIMBA were directly communicated by the authors.

In Figure 7 we summarize the results of our analylsis on number densities. We show three estimates from other studies on observed data (empty squares): Muzzin et al. (2013) use the color-color UVJ diagram to select quiescent galaxies in COSMOS/UltraVISTA, in 7 redshift bins up to $z \sim 4$ (we show the values corresponding to the $M_*/M_\odot > 10^{10}$ selection, in grey); Straatman et al. (2014) select quiescent sources in the ZFOURGE survey, again using the UVJ criterion, and focusing on the mass range $\log(M/M_\odot) > 10.6$ and the redshift range $3.47 < z < 4.25$ (in red - but we report their lower redshifts estimates as well); and Schreiber et al. (2018b) estimate the number density of quiescent galaxies at $3 < z < 4$ and $M_*/M_\odot > 3 \times 10^{10}$ by combining the UVJ selection on ZFOURGE photometric data with MOSFIRE spectral analysis, and we plot their two estimates for strictly color-selected objects (large empty cyan square) and sSFR-threshold candidates, which include recently quenched galaxies, as in M18 (smaller cyan square). We then plot the results based on our SED-fitting method for the "reference" (both before and after the correction for incompleteness, as circles or diamonds, respectively) and the "lines" selections (as stars), considering both the cuts in mass and redshift adopted by S14 (red empty symbols), and the $3 < z < 5$ interval without any mass cut (dividing in two redshift bins, $3 < z < 4$ and $4 < z < 5$ - black solid symbols). Finally, we plot the results from the five hydrodynamical simulations (solid colored lines), as described above, showing also a case in which we consider the total sSFR of the simulated galaxies to select the passive ones, rather than the sSFR of the central regions, for the sake of comparison.

Looking at the resulting plot, at $3 < z \leq 4$ we find a reasonable agreement between our number density estimations, the ones obtained by other observational studies, and the EAGLE and TNG-100 models; TNG-300 is close enough as well, although its number density is slightly lower, perhaps because of the lower resolution of the simulation with respect

to TNG100. Remarkably, while the original ILLUSTRIS simulation is not able to reproduce the properties of the galactic populations at high redshift and falls short at all epochs above $z > 1$ in reproducing the number of quenched galaxies, the TNG runs have largely cured this issue. On a side note, we point out that if we only considered the CANDELS fields yielding the highest densities the agreement both with other observations and with the simulations would be much less satisfying. We also see that, consistently with what we found in M18, our estimate is lower than the one by S14 using their cuts for mass and redshift, and we ascribe this to their shallower selection criteria, which include also mildly star-forming objects in the selection.

On the other hand, at $z > 4$ the TNG models still show a clear tension with the observational data, whereas the EAGLE model performs better; however, the EAGLE trend is too flat and struggles at reproducing the observed number densities at $z < 3$. It is difficult to trace back the origin of this different trends to the various physical mechanisms in the models, because hydrodynamical simulations are highly non-linear by construction. It is worth pointing out that the ILLUSTRIS and TNG mass functions are typically close to the observed ones (e.g. Genel et al. 2014; Pillepich et al. 2018), pointing to the conclusion that the discrepancy at $z > 4$ is not due to a poor sampling of the whole high-redshift galaxy populations, but more likely to enduring issues with the modeling of quenching mechanism. Gas hydrodymancs, radiative cooling, star formation and feedback from stars and AGNs are indeed implemented in different ways in TNG and EAGLE. We speculate that details of the AGN feedback implementation play a major role in regulating the activity of the simulated galaxies, and therefore in the defining the properties of the passive populations at different redshifts. In particular, the thermal feedback implemented in EAGLE is more efficient at high redhisft, while in TNG the most effective mechanism is the kinetic feedback, which has a larger impact at low redshifts (Weinberger et al. 2018), consistently with the global trends we have found. However, these are basic speculations. A detailed work specifically dedicated to the analysis of this topic is currently in the making (Fortuni et al., in preparation).

We point out that the values of SFR and masses that we choose to consider play a major role in reconciling models and observations: indeed, for TNG and EAGLE the sSFRs computed over limited central areas turn out to be typically lower than those computed considering the full extension of the objects, and in some cases they are exactly zero, thus allowing the inclusion of more sources in the passive selection. On the other hand, using the simplest approach of summing on all the particles of a simulated galaxy would lead to strong tensions at all redshifts, as in manu cases the sSFRs would be higher, excluding objects from the selection criteria and finally resulting in lower number density estimates (see e.g. the blue dashed line in Fig. 7, which shows the case for TNG100 model). Note that this implies that in the models the stellar mass profiles are more centrally concentrated than SFR profiles (at least at high-z), so that a significant amount of star formation takes place in the outskirts of the galaxies (see also Donnari et al. 2019): therefore, enlarging the radius in which the SFR and mass are computed yields higher sSFR values that eventually exclude many objects from the passive selection. If this is the case, current obser-

vations might underestimate the actual cosmic SFR, missing some amount of peripheral activity.

Finally, the number densities in SIMBA are closer to the observed values than the ones from ILLUSTRIS-1, but fall short with respect to TNG and EAGLE. Remarkably, in this model the values remain almost unchanged varying the aperture radius over which SFRs and masses are computed, at variance with the other simulations. Apparently, the peripheral star formation activity is not present in SIMBA. This might be due to the different star-formation prescriptions in the codes: while ILLUSTRIS and TNG assume that stars can form at gas densities $n > 0.13$ cm^{-3} (Vogelsberger et al. 2013), in SIMBA a subgrid criteria based on the estimated local density of the H2 molecule is adopted (Davé et al. 2019), and this tends to require significantly global higher densities to trigger the activity (Davé, priv. comm.). As a result, TNG might form more stars in the outskirts of galaxies than SIMBA, perhaps in particular as a result of shock-induced cooling and collapsing of gas. On the contrary, the resolution of the simulations seem to play a minor role: while the TNG-100 and EAGLE ones are finer ($\sim 10^6 M_\odot$ per gas particle), the TNG-300 one is comparable to SIMBA ($\sim 10^7 M_\odot$ per gas particle), but its number density of passive sources is close to the one in TNG-100.

As final remarks, it must be pointed out that: (i) we are applying demanding criteria which only include in the sample very robust candidates; (ii) since we are working with the CANDELS catalogs, we are not considering H-dropouts; K or IRAC-detected sources might increase the actual number of high-redshift passive galaxies, since sources that quenched long before the observation have negligible UV rest-frame flux, thus failing the H-band detection at high redshifts; (iii) as shown in Sect. 5.1, the observed values are most likely far from completeness, and although we also considered "corrected" values, we might anyway still be underestimating the total number of real passive sources. These points suggest that the tension with models cannot be considered completely ruled out.

We point out that the volumes of the simulations are comparable to the two considered cosmological volumes for $3 < z < 4$ and $4 < z < 5$, which are of the order of 3×10^6 Mpc^3; in particular, while TNG-100, EAGLE and SIMBA are a factor of ~3 smaller than the observed volumes, TNG-300 is a factor of ~9 larger, and is therefore statistically significant.

6 THE CONTRIBUTION OF THE PASSIVE POPULATION TO THE UNIVERSAL STAR FORMATION RATE DENSITY

The SED-fitting approach that we adopted in this study has another advantageous outcome. The best-fit template of each galaxy can be used to infer the complete description of its SFH through the cosmic history, by means of two of its free parameters: τ (the e-folding time for the exponentially declining models, or the burst duration for the top-hat models) and C (the normalization factor of the solution). They are the only two parameters needed to constrain the functional form of the SFH, $\Psi(z)$. From it, we can compute the average SFR of a galaxy in any redshift interval, and then

obtain an estimate of the universal SFRD summing up all the contributions from individual galaxies[2].

As a cautional remark, it is worth mentioning that this approach implies the assumption that the best template in the SED-fitting process (the one with the lowest χ^2) is the "correct" one, while in reality many more or less plausible solutions can be associated to each source. We do not take into account the uncertainties associated to this complication. In addition to this, we should also recall that it is assumed that galaxies evolve in isolation; in other words, mergers are not considered in this approach. This is by construction inherent to the SED-fitting method. However, we do not consider this point to be invalidating: the past evolution of an object and its SFH can be caused by a number of causes/events, but the amount of stars formed per year is the same, independently from the merger history.

To assess the reliability of the method, we first compare the reconstruction of the SFRD based on the SED-fitted SFHs of the CANDELS catalogs with the observed ones; we consider the well known one by Madau & Dickinson (2014, M14), and the one by Yu & Wang (2016, Y16, obtained from a MCMC technique applied to the observed mass functions rather than from the direct summation of UV and IR fluxes at the observed redshifts). For the CANDELS data, we use the delayed τ-models (in which the star formation is parametrized as $SFR(t) = t^2/\tau \times \exp(-t/\tau)$) and the mass estimates with method 6a_delτ, presented in Santini et al. (2015).

We point out that we cannot apply any correction for incompleteness here, since the approach is based on the computation of individual SFHs, which by definition cannot be inferred for undetected objects. In this exercise, the analysis is necessarily limited to galaxies with $H160 < 27.5$, i.e. to a certain mass cut, while the SFRD is typically computed by integrating the observed luminosity and mass function down to very faint limits in SFR and mass, far below the observed one, by deep extrapolation. Therefore, our estimation is actually a lower limit.

We proceed as follows. First, we define two grids of redshift bins:

- $dz_{\rm obs}$, which we use to bin the observed catalogue. Each bin can represent a slice of Universe which evolves comovingly, and therefore its global SFH can be considered an independent realization of an universal SFH, for $z > z_{\rm obs}$;
- dz_H, which we use to bin the cosmic history: we will obtain the global SFH of each $dz_{\rm obs}$, as the sum of the SFRs of individual galaxies (computed as described above) in each dz_H.

---

[2] Starting from the fitted quantities, the procedure is as follows. Each source's best-fit model is characterized by the stellar mass M_* at the age the galaxy has when it is observed, t_f. From these two fitted quantities, one can infer the normalization factor C as

$$C = \frac{M_*}{\int_0^{t_f} f_*(t) \times \Psi'(t)dt},\qquad(1)$$

where $f_*(t) = 1 - f_{\rm rec}(t)$ is the dimensionless parameter that accounts for the recycled fraction of gas, and $\Psi'(t)$ is the instantaneous *normalized* SFR (given by the SFH model, for example $\Psi'(t) = \exp(-t/\tau)$ for the exponentially declining models). The true instantaneous SFR at any epoch t is then $\Psi(t) = C \times \Psi'(t)$.

In other words, we select from the CANDELS catalogs the galaxies observed in a bin $dz_{\rm obs}$, and we trace back their SFH, summing all their contributions in each dz_H bins. To obtain the SFR *density*, we then divide the result by the comoving volume of $dz_{\rm obs}$: i.e., the volume of the spherical shell centered on the observer, extending from $z_{\rm obs}$ to $z_{\rm obs} + dz_{\rm obs}$, and re-scaled to the total angular area of the five fields. We then repeat the procedure for all the $dz_{\rm obs}$; finally, in each dz_H we compute the mean of the various SFRDs corresponding to each $dz_{\rm obs}$. In formulae:

$$\forall dz_{\rm obs} \longrightarrow SFR_{dz_H} = \sum_{i=1}^{n} SFR_{dz_H,i}\qquad(2)$$

where n is the number of galaxies in the considered $dz_{\rm obs}$; then,

$$\forall dz_{\rm obs} \longrightarrow SFRD_{dz_H} = SFR_{dz_H}/V_{dz_{\rm obs}}\qquad(3)$$

and finally

$$\langle SFRD_{dz_H}\rangle = \frac{\sum_{j=1}^{N_{dz_{\rm obs}}} SFRD_{dz_H,j}}{N_{dz_{\rm obs}}}\qquad(4)$$

The method is sketched in Figure 8. In Fig. 9 we plot the result of this approach, together with the two cited observational ones.

Our points are in good agreement with the curves, and particularly with Y16, although they are consistent with both estimations considering the uncertainties (the error bars are the standard deviations of the distribution of the SFRD from each $dz_{\rm obs}$). We speculate that while the discrepancy with the M14 curve can be understood in terms of the caveats on the incompleteness of our sample discussed above, the good consistency at $z < 5$ with the Y16 estimate is justified by the fact that we are basically adopting their same method to compute the SFRD (see their Eq. 5). Therefore, the SFR estimated from the observed UV/IR fluxes seems not strictly equivalent to the SFR inferred from the observed stellar mass, which might be due to several factors discussed in Y16, among which the IMF, the metallicity, the outshining of young stars, or overestimated absorption, as well as possible differences in the computation of the recycled fraction. However, the tension with MD14 is not dramatic, and furthermore, we note that the precise values of the averaged SFRDs depend on the choice of $dz_{\rm obs}$ binning (and *binning is sinning...*).

At $z > 5$ our estimation apparently diverges from the other two, but the scatter between the SFRDs from the various $dz_{\rm obs}$ bins is large enough to make the global estimate consistent with the curve within the error budget. We recall that the bars correspond the standard deviation of the distribution of the individual SFRDs of each $dz_{\rm obs}$, so that their extension at high z is likely due to poor statistics (few detected objects, plus cosmic variance); on the other hand, at low redshifts the $dz_{\rm obs}$ bins contributing to the summation are less, so their dispersion is lower.

Having assessed that this approach can reasonably reproduce known observational constraints, we can now exploit it to estimate the contribution to the cosmic SFRD of the galaxies that we have selected as passively evolving at
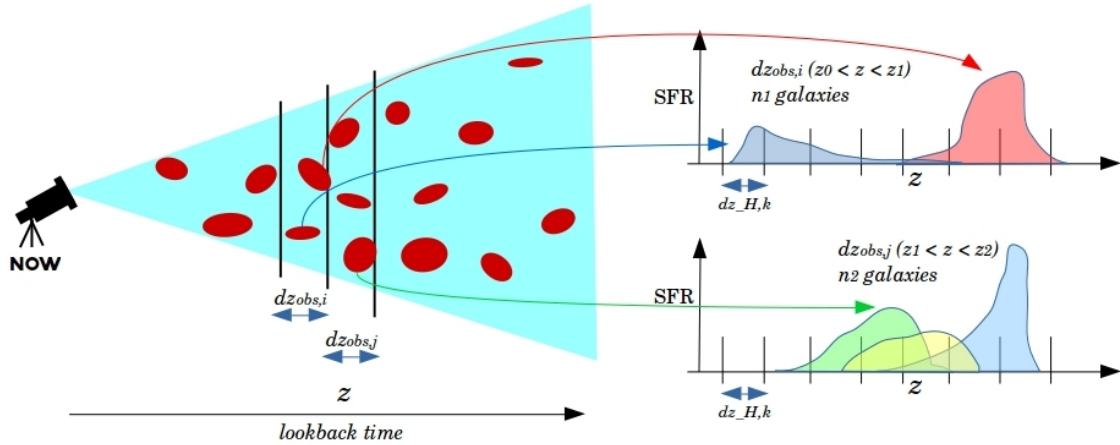
Figure 8. Sketch of the method adopted to compute the global SFRD across the cosmic time. We sum the contributions of all the galaxies observed in a given redshift interval $dz_{\rm obs}$ to the SFR in each redshift bin dz_H across the cosmic history. Such contributions can be evaluated from the shape and normalisation of the SFH of each galaxy, which in turn come from the SED-fitting procedure. Then, we divide the obtained total SFR in each bin dz_H by the $dz_{\rm obs}$ comoving volume to obtain the SFRD in that $dz_{\rm obs}$. Finally, we average over the $dz_{\rm obs}$ bins to obtain the SFRD of each dz_H.

the time of their observation, during their previous phase of SF activity. Applying the very same technique to our sample (taking $dz_{\rm obs} = 3 - 5$, and considering the top-hat SFHs we used in the SED-fitting procedure rather than the standard τ-models) we obtain the thick black (for the "reference" selection, without nebular emission lines) and thin grey (for the "lines" selection) curves plotted in Fig. 9. We note that the sharp decrease in the plotted SFRD of passive galaxies towards $z = 3$ is of course an artifact due to the fact that we are only looking for $z > 3$ objects, so all our candidates cease to form stars at earlier epochs, by construction. In this case, we must recall that as discussed in Sect. 5.1 we are probably missing passive sources due to our demanding passive selection criteria, on top of the mentioned photometric incompleteness. As already discussed, we cannot apply any correction here, so again we must take the result as a lower limit of the true values.

Considering the "reference" sample, it turns out that the contribution of the passive candidates to the global SFRD is $5 - 10\%$ of the total, at $z < 8$. We can compare this result to their fractional abundance in number: from our estimates summarized in Tab. 2, they are $102/21608 \simeq 0.5\%$ of the total number of all $z > 3$ detected sources. Their SF activity must therefore be $\sim 10 - 20$ times higher than the average.

In Fig. 9 we also plot the contribution to the SFRD from the individual SFHs of the passive candidates (thin lines, color coded from blue to red according to their final stellar mass; they are straight horizontal lines, because the SFR is constant for the top-hat models). A few objects have a very high impact on the total SFRD and this is reflected in the sudden jumps in the the global curve. However, the bulk of the contribution is given by longer, intermediate intensity bursts.

As a final note we also remark that the most massive objects (redder lines) typically have higher SFRs. Since the lines are plotted as a function of redshift, it might not be immediate to realize that the actual duration of the bursts of SF activity are not dramatically different (e.g., from $\simeq 1.9$ Gyr for a burst starting at $z = 20$ to $\simeq 1.2$ Gyr for one start-

ing at $z = 6$, both ending at $z = 3$). This clarifies the direct correlation between the mass of the galaxies (traced by the color of the lines) and the value of their SFRD. Nevertheless, it is worth noticing that the galaxies with the highest SFRDs ($> 10^{-4}$ $M_\odot/\rm{yr}/\rm{Mpc}^3$) have very short bursts, implying an extremely fast and efficient star formation activity abruptly quenched either by gas consumption or by a very effective feedback mechanism.

7 SUMMARY AND CONCLUSIONS

Constraining the number and the features of high-redshift passive galaxies is crucial to understand the evolution of the Universe. To this aim, we have searched for red and dead objects at $z > 3$ in the five CANDELS fields, improving upon the first work on GOODS-South presented in M18. We performed SED-fitting with ad-hoc, abruptly quenched (top-hat) SFHs rather than standard τ-models, and exploited a probabilistic technique to select robust candidates considering all the possible solutions of the fit and only keeping sources with zero or very low probability star-forming ones ($p_{\rm SF} < 5\%$).

Our selection method allows us to find galaxies that would not be identified as passive with more traditional techniques such as the widely used UVJ color-color diagram. In particular, we can single out objects that fail the $U - V$ color selection, having abruptly quenched their SF activity only shortly before the observation. The quenching might be caused by a strong quasar activity episode, which typically happens on very short timescales; moreover, feedback from young stars in the form of UV radiation, winds and supernovae can prevent gas from infalling back to the galaxy avoiding re-juvenation processes, at the same time preventing the local gas (which has been observed to exist in quenched galaxies) to cool down and collapse into new stars. Other possible explanations might include low nuclear activity, shock heating of infalling gas, strangulation/removing of gas, or low SF efficiency. The relative compactness of the
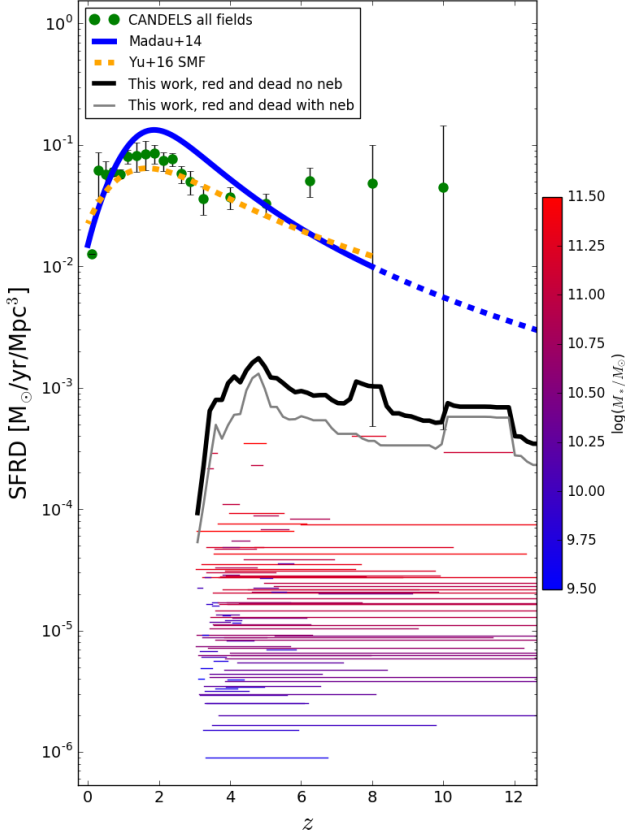
Figure 9. Universal star formation rate densities. The blue solid thick line is the Madau & Dickinson (2014) SFRD inferred from UV and FIR observational data. The dashed yellow thick line is the SFRD by Yu & Wang (2016), who infer the instantaneous SFR from the stellar mass functions at various redshifts by means of a MCMC approach, and argue that a substantial offset from the observed curve is evident. The green dots are our determination of the SFRD using the five CANDELS fields as described in the text: the mean of the SFRDs at all epochs obtained from samples of objects at various observed bins of redshifts $dz_{\rm obs}$ (the error bars represent the standard deviations of the distribution of the SFRD from each $dz_{\rm obs}$, see text for definitions). The points are in good agreement with observations (particularly with the Yu et al. curve), reassuring that the method is reliable. Finally, the black and grey solid lines are the SFRDs of our selections of passive galaxies in the CANDELS fields (respectively with and without nebular emission lines), and the thin blue to red horizontal lines are the individual contribution of the candidates to the SFRD, color-coded as a function of their observed mass.

observed passive sample (Fig. 5) hints at an inside-out formation process, in which peripheral star formation due to minor mergers might happen at lower z.

Depending on the chosen library of models and SED-fitting technique, we end up with the selections summarized in Table 2. We individuate 102 red and dead candidates, which decrease to 40 if we include the emission lines in the libraries. The physical properties of all the candidates, as inferred from their best fit, are given in Table B1; snapshots and SEDs are available on-line as complementary material.

We find the number densities given in Tabs. 5 and 6, with a global estimation of $\simeq 1.73 \pm 0.17 \times 10^{-5}$ (6.69 ±

1.08×10^{-6}) Mpc^{-3} at $3 < z \leq 5$, for the "reference" ("lines") selection. The estimate rises to $\simeq 2.30 \pm 0.20 \times 10^{-5}$ after multiplying by a correction factor to take into account the estimated incompleteness of our selected sample. We compare these numbers with the ones obtained by five cosmological hydrodynamical simulations, for which we adopt the $sSFR < 10^{-11}$ yr^{-1} threshold to identify passive galaxies with $M_*/M_\odot > 5 \times 10^9$. We find a reasonable agreement with the predictions from the ILLUSTRIS-TNG and EAGLE models at $3 < z \leq 4$, provided that a wise choice of the relevant data is made, i.e. including in the computations only the stellar particles within some typical radius reminiscent of an observed quantity (e.g. twice the half mass radii), rather than summing on all the particles formally belonging to a simulated galaxy (if we do so, many objects formerly passing the selection would instead be classified as star forming). Interestingly, this would indicate that the observations might be missing to detect some level peripheral star formation activity. However, a non negligible tension remains at higher redshifts, where all the models fail to produce a significant number of quenched objects, whereas we find 14 galaxies in the redshift bin $4 < z \leq 5$. The other models (SIMBA and ILLUSTRIS-1) provide less accurate matching at all epochs. An important remark is that some underestimation of the real number density of passive sources is probably present and unavoidable, due to many factors (e.g., the demanding selection criteria, the conservative corrections for incompleteness, the atypical number density of the COSMOS field which lowers the average), so that we cannot claim that the tension between models and observation is ruled out.

Finally, we computed the contribution of passive galaxies, as inferred from our selection, to the global cosmic SFRD during their past active phase, finding that despite being a tiny fraction ($\sim 0.5\%$) of the whole galactic population, they provide a $\sim 5-10\%$ fraction of the total activity from $z \simeq 8$ to $z \simeq 3$. Therefore, they play a crucial role in the cosmic star formation history.

As discussed in M18, upcoming observations with telescopes of the new generation will provide more stringent constraints on the observational data. In Fig. 10 we give an example of the current and future possibilities. We plot, with arbitrary units, the SEDs of top-hat models at $z = 3$, 5 and 8, each one at three different epochs: the time of the quenching of star-formation $t_{\rm quench}$, 50 Myr later, and 300 Myr later (the latter except for the $z = 8$ case since it is not included in our library of models). The models are dust-free, but for comparison we also show the case of the $z = 3$ models with $E(B - V) = 0.2$: its rest-frame spectrum blueward of the 4000Å (D_{4000}) break at $t_{\rm quench}$ is very similar to the spectrum at $t_{\rm quench}$+50Myr of the corresponding dust-free model. This creates a degeneracy that can be broken with a better sampling of the break, looking at FIR wavelengths, or using well tailored spectral indexes such as Mg$_{UV}$ (Daddi et al. 2005).

As it can be seen looking at the leftmost red filter curve, *HST H*160, the current work is based on a NIR selection which corresponds to UV/optical rest-frame at $z > 3$. Since soon after the quenching of the star formation activity the luminosity drops by 1-2 magnitudes in this range of wavelenghts, current selections are most likely biased towards recently quenched sources at $z > 4$, where the detection falls blueward of the break: at such redshifts we are probably
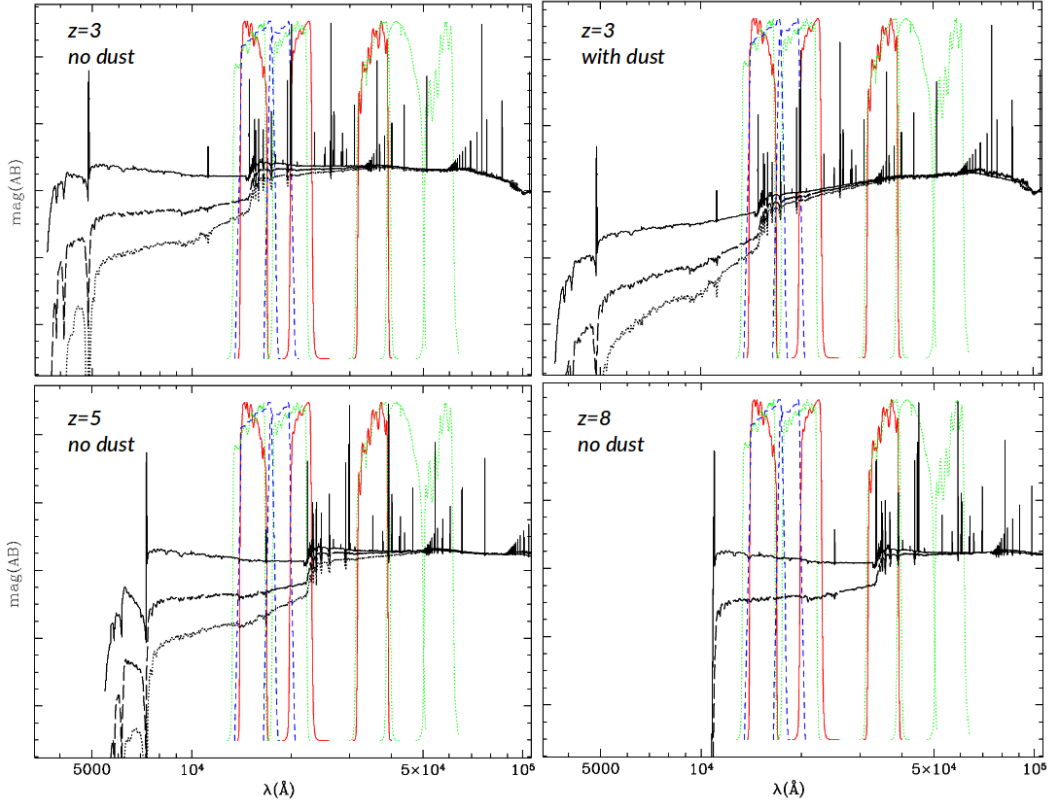
Figure 10. Compared capabilities of present and future facilities in detecting passive sources at different redshifts. The black lines are the SEDs of galaxies that quench at different epochs (redshifts), as indicated in the panels (the magnitude scales on the y axis are arbitrary). The SEDs were simulated with the top-hat library used in this paper. The solid line in each panel, which has brighter UV fluxes, shows the SED at the epoch of quenching $t_{\rm quench}$, while the dashed and dotted lines with fainter and fainter UV fluxes correspond to 50 Myr and 300 Myr later, respectively. The $z = 8$ panel does not include the $t_{\rm quench} + 300$Myr case, because it is not modeled in our libraries. Note that the upper right panel shows the SEDs of a galaxy with dust ($E(B - V) = 0.2$), while the other cases are dust-free: the UV region of its spectrum at $t_{\rm quench}$ epoch is very similar to the $t_{\rm quench} + 50$Myr epoch of the corresponding dust-free case, a known degeneracy that can be broken sampling the D_{4000} break, obtaining the Mg_{UV} spectral index, or looking at the FIR region of the spectra. Superposed to the SEDs in arbitrary units are the the curves of some significant pass-band filters of CANDELS (solid red: left to right, F160, Ks and IRAC-CH1), WFIRST (dashed blue: left to right, H158 and F184), and JWST (dotted green: left to right, NIRCAM F150, F200, F356, F444, and MIRI F560). See text for discussion.

missing galaxies that have quenched more than a few tens of Myr before the observation epoch. On the other hand, *JWST* will straddle the optical/NIR rest-frame wavelengths with the highest resolution, sampling the D_{4000} break for $z \sim 8$ sources with the F356 NIRCAM filter (which corresponds to the IRAC-CH1 bandpass), at the same time enhancing the SNR reaching AB magnitudes $\sim 29 - 30$ and the resolution of the images. Furthermore, while we are now forced to work on a very restricted sky area, WFIRST will observe a ~ 2500 sq. degrees reaching $j \sim 27$ (29 in the deep survey), allowing for much grander statistic significance, although its filter coverage is similar to the CANDELS *HST* passbands so that it will likely be useful only for detection.

In sinergy with FIR facilities like ALMA, which can confirm the nature of the candidates analyzing the gas and dust content as we did in S19, or EELT which will provide spectral data for extremely faint objects (together with robust morphological analysis thanks to adaptive optics technology), the accuracy and statistical significance of data will quite soon be enormously enhanced. This will hopefully allow to finally understand whether the currently determined

numbers, which the models seem to be able to broadly reproduce up to $z \simeq 4$ but fail to match at earlier epochs, are indeed a good estimation of the real number of passive objects, or they are actually lower limits - in which case a tension with the theoretical predictions would reappear even at lower redshifts.

References

Astropy Collaboration et al., 2013, A&A, 558, A33
Bacon R., et al., 2017, A&A, 608
Barro G., et al., 2019, ApJS, 243, 22
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bouwens R. J., et al., 2015, ApJ, 803, 34
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Brennan R., et al., 2017, MNRAS, 465, 619
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bundy K., et al., 2006, ApJ, 651, 120
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Castellano M., et al., 2014, A&A, 566, A19
Ceverino D., Klypin A., 2009, ApJ, 695, 292
Chabrier G., 2003, PASP, 115, 763
Chiosi C., Carraro G., 2002, MNRAS, 335, 335
Cimatti A., et al., 2002, A&A, 391, L1
Cimatti A., Daddi E., Renzini A., 2006, doi:10.1051/0004-6361:20065155, 453, L29
Cimatti A., et al., 2008, doi:10.1051/0004-6361:20078739, 482, 21
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Daddi E., et al., 2005, ApJ, 626, 680
Daddi E., et al., 2007, ApJ, 670, 156
Dahlen T., et al., 2013, ApJ, 775, 93
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Davidzon I., et al., 2017, A&A, 605, A70
De Lucia G., Blaizot J., 2007, MNRAS, 375, 2
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Donnari M., et al., 2019, MNRAS, 485, 4817
Dunlop J. S., Cirasuolo M., McLure R. J., 2007, MNRAS, 376, 1054
Elbaz D., et al., 2007, A&A, 468, 33
Elbaz D., et al., 2011, A&A, 533, A119
Feldmann R., Mayer L., 2015, MNRAS, 446, 1939
Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., 2017, MNRAS, 470, 1050
Finkelstein S. L., Dunlop J., Le Fevre O., Wilkins S., 2015, preprint, (arXiv:1512.04530)
Fontana A., D'Odorico S., Poli F., Giallongo E., Arnouts S., Cristiani S., Moorwood A., Saracco P., 2000, AJ, 120, 2206
Fontana A., et al., 2009, A&A, 501, 15
Fontana A., et al., 2014, A&A, 570, A11
Galametz A., et al., 2013, ApJS, 206, 10
Genel S., et al., 2014, MNRAS, 445, 175
Glazebrook K., et al., 2017, Nature, 544, 71
Grazian A., et al., 2015, A&A, 575, A96
Grogin N. A., et al., 2011, ApJS, 197, 35
Guo Y., et al., 2013, ApJS, 207, 24
Harikane Y., et al., 2016, ApJ, 821, 123
Hopkins P. F., 2015, MNRAS, 450, 53
Ichikawa A., Matsuoka Y., 2017, ApJ, 843, L7
Inami H., et al., 2017, A&A, 608, A2
Kawata D., 1999, PASJ, 51, 931
Koekemoer A. M., et al., 2011, ApJS, 197, 36
Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, ApJ, 700, 221
Labbé I., et al., 2005, ApJ, 624, L81
Lacey C., Cole S., 1993, MNRAS, 262, 627
Laidler V. G., et al., 2007, PASP, 119, 1325
Le Fèvre O., et al., 2015, A&A, 576, A79
Lutz D., et al., 2011, A&A, 532, A90
Madau P., Dickinson M., 2014, ARA&A, 52, 415

Man A., Belli S., 2018, Nature Astronomy, 2, 695
Matteucci F., 1994, 288, 57
McLure R. J., et al., 2018, MNRAS, 479, 25
Merlin E., Chiosi C., Piovan L., Grassi T., Buonomo U., Barbera F. L., 2012, MNRAS, 427, 1530
Merlin E., et al., 2015, A&A, 582, A15
Merlin E., et al., 2016, preprint, (arXiv:1609.00146)
Merlin E., et al., 2018, MNRAS, 473, 2098
Mobasher B., et al., 2005, ApJ, 635, 832
Muzzin A., et al., 2013, ApJ, 777, 18
Nayyeri H., et al., 2014, ApJ, 794, 68
Nayyeri H., et al., 2017, ApJS, 228, 7
Nelson D., et al., 2015, Astronomy and Computing, 13, 12
Nelson D., et al., 2019, Computational Astrophysics and Cosmology, 6, 2
Noeske K. G., et al., 2007, ApJ, 660, L43
Oliver S. J., et al., 2012, MNRAS, 424, 1614
Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192
Pentericci L., et al., 2018, A&A, 616, A174
Pillepich A., et al., 2018, MNRAS, 473, 4077
Press W. H., Schechter P., 1974, ApJ, 187, 425
Prevot M. L., Lequeux J., Prevot L., Maurice E., Rocca-Volmerange B., 1984, A&A, 132, 389
Price-Whelan A. M., et al., 2018, AJ, 156, 123
Rodighiero G., et al., 2014, MNRAS, 443, 19
Roseboom I. G., et al., 2010, MNRAS, 409, 48
Roseboom I. G., et al., 2012, MNRAS, 419, 2758
Salpeter E. E., 1959, ApJ, 129, 608
Santini P., et al., 2015, ApJ, 801, 97
Santini P., et al., 2019, arXiv e-prints,
Schaerer D., de Barros S., 2009, A&A, 502, 423
Schaye J., et al., 2015, MNRAS, 446, 521
Schreiber C., Merlin E., Elbaz D., Wang T., Shu X., Fontana A., Castellano M., Santini P., 2015
Schreiber C., et al., 2018a, A&A, 611, A22
Schreiber C., et al., 2018b, A&A, 618, A85
Skelton R. E., et al., 2014, ApJS, 214, 24
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Stefanon M., et al., 2017, ApJS, 229, 32
Straatman C. M. S., et al., 2014, ApJ, 783, L14
Tacchella S., et al., 2018, ApJ, 859, 56
Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673
Urrutia T., et al., 2018, arXiv e-prints,
Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
Vogelsberger M., et al., 2014a, MNRAS, 444, 1518
Vogelsberger M., et al., 2014b, Nature, 509, 177
Weinberger R., et al., 2018, MNRAS, 479, 4056
Whitaker K. E., et al., 2011, ApJ, 735, 86
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Worthey G., 1994, ApJS, 95, 107
Wuyts S., et al., 2007, ApJ, 655, 51
Wuyts S., et al., 2011, ApJ, 742, 96
Yu H., Wang F. Y., 2016, ApJ, 820, 114

APPENDIX A: THE TOP-HAT LIBRARY

The library used to fit the observed photometry with ZPHOT consists of $\sim 7.9 \times 10^8$ models. The grid of models is built considering the following criteria:

- we assume a Salpeter (1959) IMF;
- redshifts vary from 0.0 to 8.0. These are the redshift at which the objects are observed, and at which the templates are shifted applying the K-correction;
- the burst durations $\Delta t_{\mathrm{burst}}$ vary from 100 Myr to 1.5 Gyr (with linear steps of 100 Myr); note that at $z = 3$ the Universe has an age of $\simeq 2.2$ Gyr;
- we considered two extinction laws, namely the Calzetti et al. (2000) and the Small Magellanic Cloud by Prevot et al. (1984) ones;
- the ages and extinctions values vary as a function of $\Delta t_{\mathrm{burst}}$. We considered three possible regimes:

  – for $t \leq \Delta t_{\mathrm{burst}}$, we built a model for each 50 Myr of age, and allowed for values of $E(B-V)$ from 0 to 1;
  – for the first 300 Myr after the quenching, we again built a models for each 50 Myr of age, but restricted the allowed values of $E(B-V)$ to the range $0-0.4$;
  – for later ages, we built models for each 50 Myr of age up to 1 Gyr after the quenching, and then we used logarithmic steps of 0.1 to reach the age of the Universe (at $z = 2$, again to allow for more relaxed fitting), and considered $E(B-V) < 0.2$.

These choices for the $E(B-V)$ values were made to mimic the expected drop of dust content in a quenched galaxy after the end of the star-forming activity.

- the emission lines are included using the prescription described in Castellano et al. (2014); Schaerer & de Barros (2009).

APPENDIX B: PHYSICAL PROPERTIES OF THE SELECTED SAMPLE OF RED AND DEAD CANDIDATES

ID_{CAND}	RA DEC [deg]	z_{CAND}	Age [Gyr]	M_* [10^9 M$_\odot$]	Δt_{burst} [Gyr]
COSMOS-16676* ⊕	150.0615 2.3787	3.72	$0.50^{+0.65}_{-0.15}$	$322.70^{+153.3}_{-165.00}$	0.2
COSMOS-19502 ⊕	150.1309 2.4136	3.87	$0.40^{+1.00}_{-0.15}$	$107.40^{+73.20}_{-31.04}$	0.1
COSMOS-2075	150.0535 2.2045	3.35	$1.00^{+0.70}_{-0.99}$	$45.59^{+29.71}_{-32.39}$	0.9
COSMOS-18286	150.0767 2.3986	3.04	$1.40^{+0.25}_{-1.15}$	$199.40^{+72.30}_{-89.30}$	1.4
EGS-14727 ⊕	214.8956 52.8566	3.05	$0.90^{+1.00}_{-0.60}$	$155.90^{+67.30}_{-60.74}$	0.7
EGS-21351 ⊕	214.6736 52.7326	3.61	$0.65^{+0.90}_{-0.35}$	$80.76^{+37.04}_{-23.84}$	0.2
EGS-24177 ⊕	214.8661 52.8843	3.42	$0.80^{+0.85}_{-0.45}$	$214.40^{+83.00}_{-59.60}$	0.5
EGS-25724 ⊕	214.9978 52.9862	3.79	$1.10^{+0.35}_{-0.90}$	$77.77^{+57.93}_{-35.44}$	0.8
EGS-29547 ⊕	214.6953 52.7969	3.15	$1.70^{+0.15}_{-1.40}$	$86.37^{+25.43}_{-42.06}$	1.5
EGS-2490	214.9515 52.8292	3.10	$1.40^{+0.20}_{-1.25}$	$39.16^{+10.60}_{-19.80}$	1.4
EGS-6539	214.9132 52.8247	3.44	$1.55^{+0.10}_{-1.40}$	$136.50^{+51.40}_{-52.55}$	1.5
EGS-15868	214.8712 52.8451	3.61	$0.30^{+1.25}_{-0.15}$	$105.10^{+80.60}_{-31.18}$	0.2
EGS-23036	214.8791 52.8881	3.57	$1.55^{+0.05}_{-1.40}$	$36.73^{+17.00}_{-19.49}$	1.4
EGS-24356	214.6201 52.7096	3.43	$1.65^{+0.00}_{-1.45}$	$109.60^{+38.10}_{-49.35}$	1.5
EGS-26762	214.6261 52.7268	3.28	$1.10^{+0.65}_{-0.90}$	$19.41^{+10.78}_{-9.34}$	1. 0
EGS-27491	214.6177 52.7242	3.34	$1.65^{+0.05}_{-1.45}$	$96.93^{+51.07}_{-55.25}$	1.5
EGS-30675	214.9049 52.9354	3.01	$1.25^{+0.60}_{-1.10}$	$50.26^{+17.50}_{-22.48}$	1.2
GOODSN-13 ⊕	189.1544 62.0947	3.01	$0.15^{+1.20}_{-0.05}$	$11.03^{+11.54}_{-3.20}$	0.1
GOODSN-2901 ⊕	189.2199 62.1569	3.70	$0.40^{+1.10}_{-0.25}$	$11.65^{+10.30}_{-3.73}$	0.3
GOODSN-4004 ⊕	189.2657 62.1684	3.81	$1.40^{+0.05}_{-1.20}$	$36.48^{+27.85}_{-13.20}$	1.3
GOODSN-5059 ⊕	189.1623 62.1782	3.69	$1.30^{+0.20}_{-0.95}$	$126.70^{+63.40}_{-54.97}$	0.1
GOODSN-10672 ⊕	189.0325 62.2164	6.71	$0.50^{+0.15}_{-0.35}$	$40.29^{+42.77}_{-15.96}$	0.3
GOODSN-12446 ⊕	189.3730 62.2287	3.05	$0.15^{+0.35}_{-0.10}$	$1.94^{+0.99}_{-0.59}$	0.1
GOODSN-13403 ⊕	189.2779 62.2350	3.79	$0.45^{+1.00}_{-0.30}$	$6.49^{+4.50}_{-3.06}$	0.4
GOODSN-13800 ⊕	189.3323 62.2370	3.33	$0.65^{+1.05}_{-0.40}$	$74.62^{+21.80}_{-28.39}$	0.5
GOODSN-15054 ⊕	189.0799 62.2448	3.06	$1.40^{+0.35}_{-1.15}$	$208.80^{+77.40}_{-58.00}$	1.3
GOODSN-19580 ⊕	189.4824 62.2739	3.10	$0.80^{+1.10}_{-0.50}$	$106.40^{+35.70}_{-37.85}$	0.5
GOODSN-24501 ⊕	189.3026 62.3665	4.25	$0.35^{+0.85}_{-0.15}$	$20.98^{+18.80}_{-5.01}$	0.2
GOODSN-357	189.1611 62.1129	3.09	$0.35^{+1.15}_{-0.20}$	$9.52^{+7.22}_{-2.29}$	0.3
GOODSN-1570	189.0083 62.1412	3.23	$0.15^{+0.70}_{-0.00}$	$13.15^{+10.67}_{-0.45}$	0.1
GOODSN-4691	189.1099 62.1752	3.18	$0.90^{+0.90}_{-0.55}$	$108.90^{+56.40}_{-43.43}$	0.8
GOODSN-5744	189.1001 62.1836	3.46	$0.50^{+1.15}_{-0.25}$	$50.56^{+35.84}_{-21.30}$	0.2
GOODSN-6430	189.1844 62.1882	3.21	$0.65^{+0.40}_{-0.50}$	$8.79^{+2.34}_{-3.16}$	0.6
GOODSN-6620	189.1814 62.1893	3.70	$0.25^{+1.10}_{-0.10}$	$27.56^{+18.88}_{-9.58}$	0.2
GOODSN-7385	188.9659 62.1945	3.18	$0.30^{+1.05}_{-0.15}$	$7.87^{+3.06}_{-3.71}$	0.1
GOODSN-9626	189.1384 62.2095	3.18	$0.20^{+1.15}_{-0.05}$	$7.83^{+6.12}_{-2.36}$	0.1

GOODSN-10956	189.1512	62.2184	3.09	$0.20^{+1.30}_{-0.05}$	$4.33^{+4.48}_{-0.52}$	0.1
GOODSN-11579	189.2347	62.2227	3.17	$1.00^{+0.85}_{-0.65}$	$53.22^{+30.88}_{-22.52}$	0.7
GOODSN-13007	189.1825	62.2320	3.04	$1.10^{+0.50}_{-0.95}$	$37.04^{+9.83}_{-19.58}$	1.1
GOODSN-13435	188.9787	62.2350	3.65	$0.70^{+0.85}_{-0.50}$	$61.89^{+34.91}_{-26.32}$	0.2
GOODSN-14482	189.3675	62.2418	3.49	$0.40^{+1.10}_{-0.20}$	$11.27^{+6.16}_{-2.83}$	0.2
GOODSN-16817	189.3778	62.2569	3.69	$0.15^{+1.05}_{-0.00}$	$6.50^{+7.48}_{-1.47}$	0.1
GOODSN-18860	189.2944	62.2694	4.53	$0.20^{+0.85}_{-0.05}$	$12.83^{+17.96}_{-3.33}$	0.1
GOODSN-20589	189.1298	62.2812	3.34	$0.20^{+1.00}_{-0.05}$	$7.62^{+5.97}_{-0.86}$	0.1
GOODSN-21034	189.3571	62.2847	3.33	$1.40^{+0.30}_{-1.25}$	$9.52^{+5.35}_{-4.46}$	1.3
GOODSN-21961	189.5017	62.2916	3.36	$1.55^{+0.15}_{-1.40}$	$27.71^{+13.54}_{-10.44}$	1.5
GOODSN-22398	189.1517	62.2950	3.11	$0.25^{+0.60}_{-0.10}$	$4.72^{+2.15}_{-1.33}$	0.2
GOODSN-24092	189.3094	62.3801	3.30	$1.25^{+0.50}_{-0.95}$	$124.30^{+39.90}_{-44.58}$	1. 0
GOODSN-24572	189.3214	62.3529	3.34	$0.35^{+0.60}_{-0.20}$	$16.75^{+6.27}_{-5.13}$	0.3
GOODSN-25209	189.4217	62.3420	3.27	$1.40^{+0.35}_{-1.15}$	$62.06^{+19.13}_{-29.56}$	1.3
GOODSN-27251	189.1960	62.3121	3.12	$0.45^{+0.75}_{-0.30}$	$11.67^{+5.65}_{-5.61}$	0.4
GOODSN-28344	189.0797	62.1513	4.76	$0.70^{+0.35}_{-0.69}$	$27.27^{+26.72}_{-15.78}$	0.3
GOODSN-35028	189.4713	62.3227	3.64	$1.10^{+0.45}_{-0.90}$	$174.50^{+90.50}_{-106.59}$	0.1
GOODSS-2608 ⊕	53.1486	-27.8896	3.72	$0.30^{+0.75}_{-0.15}$	$3.78^{+2.34}_{-1.17}$	0.2
GOODSS-2717 ⊕	53.1893	-27.8885	3.02	$0.40^{+1.55}_{-0.05}$	$131.10^{+89.50}_{-37.88}$	0.1
GOODSS-2782 ⊕	53.0836	-27.8875	3.58	$1.25^{+0.30}_{-1.00}$	$72.30^{+19.26}_{-30.09}$	0.8
GOODSS-3912 ⊕	53.0622	-27.8750	3.90	$1.25^{+0.15}_{-1.05}$	$35.16^{+22.62}_{-16.54}$	1.2
GOODSS-4587 ⊕	53.0705	-27.8686	3.75	$0.25^{+1.20}_{-0.10}$	$5.49^{+4.35}_{-2.01}$	0.1
GOODSS-8785 ⊕	53.0818	-27.8287	3.85	$1.40^{+0.00}_{-1.20}$	$35.57^{+18.63}_{-17.21}$	0.9
GOODSS-9209 ⊕	53.1082	-27.8251	4.49	$1.10^{+0.05}_{-0.90}$	$91.80^{+5.54}_{-25.15}$	1.1
GOODSS-10578 ⊕	53.1653	-27.8141	3.06	$1.10^{+0.30}_{-0.85}$	$333.60^{+43.20}_{-138.60}$	1.1
GOODSS-17749 ⊕	53.1969	-27.7605	3.70	$1.40^{+0.10}_{-1.15}$	$96.49^{+36.91}_{-38.16}$	0.9
GOODSS-18180 ⊕	53.1812	-27.7564	3.65	$1.40^{+0.15}_{-1.15}$	$80.81^{+30.69}_{-25.87}$	0.8
GOODSS-19883 ⊕	53.0107	-27.7416	3.57	$1.25^{+0.35}_{-1.00}$	$156.80^{+39.80}_{-75.91}$	1.1
GOODSS-22085 ⊕	53.0739	-27.7222	3.47	$1.40^{+0.25}_{-1.05}$	$48.20^{+15.77}_{-16.13}$	1.2
GOODSS-23626 ⊕	53.1030	-27.7123	4.75	$0.45^{+0.60}_{-0.25}$	$77.19^{+38.31}_{-23.69}$	0.2
GOODSS-3718	53.1593	-27.8772	3.85	$1.00^{+0.40}_{-0.85}$	$21.51^{+20.85}_{-9.09}$	1. 0
GOODSS-3897	53.0554	-27.8753	3.12	$1.40^{+0.10}_{-1.25}$	$19.12^{+1.82}_{-9.58}$	1.4
GOODSS-3973	53.1393	-27.8745	3.63	$1.25^{+0.30}_{-0.90}$	$170.00^{+82.70}_{-69.10}$	0.8
GOODSS-4202	53.1881	-27.8725	3.31	$1.10^{+0.40}_{-0.90}$	$4.53^{+1.56}_{-2.43}$	1.1
GOODSS-4503	53.1133	-27.8699	3.59	$1.25^{+0.30}_{-1.00}$	$139.00^{+57.50}_{-55.82}$	1.1
GOODSS-4949	53.0947	-27.8651	4.83	$0.20^{+0.85}_{-0.05}$	$10.08^{+6.25}_{-3.36}$	0.1
GOODSS-5934	53.1295	-27.8550	4.86	$0.25^{+0.80}_{-0.10}$	$6.73^{+6.25}_{-1.95}$	0.2
GOODSS-6407	53.0778	-27.8501	4.81	$0.25^{+0.80}_{-0.10}$	$16.85^{+12.93}_{-4.34}$	0.1
GOODSS-7526	53.0787	-27.8395	3.32	$1.00^{+0.75}_{-0.80}$	$34.40^{+15.19}_{-18.18}$	0.6

ID	RA	DEC	z	Age	M_*	Δt_{burst}
GOODSS-7688	53.0796	-27.8382	3.40	$1.65^{+0.00}_{-1.50}$	$23.70^{+10.64}_{-12.16}$	1.4
GOODSS-8242	53.0816	-27.8334	3.24	$1.00^{+0.50}_{-0.50}$	$6.91^{+1.26}_{-1.79}$	1.0
GOODSS-12178	53.0393	-27.7993	3.29	$1.55^{+0.20}_{-1.35}$	$47.44^{+16.17}_{-19.33}$	1.5
GOODSS-13394	53.0705	-27.7909	3.29	$1.00^{+0.50}_{-0.60}$	$11.36^{+6.49}_{-2.83}$	1.0
GOODSS-15457	53.1483	-27.7784	3.50	$0.35^{+0.70}_{-0.20}$	$4.69^{+3.57}_{-1.53}$	0.3
GOODSS-16506	53.1618	-27.7706	3.38	$0.30^{+0.75}_{-0.15}$	$5.23^{+3.82}_{-1.38}$	0.2
GOODSS-16526	53.0276	-27.7703	3.15	$1.00^{+0.05}_{-0.60}$	$13.57^{+1.24}_{-4.08}$	1.0
GOODSS-19301	53.1320	-27.7468	3.59	$1.55^{+0.00}_{-1.40}$	$13.42^{+7.83}_{-7.21}$	1.5
GOODSS-19446	53.1652	-27.7458	3.27	$1.10^{+0.40}_{-0.90}$	$17.60^{+4.90}_{-7.02}$	1.1
GOODSS-19505	53.0166	-27.7448	3.59	$1.40^{+0.15}_{-1.25}$	$52.36^{+23.19}_{-21.13}$	1.4
GOODSS-22610	53.0620	-27.7176	3.33	$1.00^{+0.50}_{-0.85}$	$11.29^{+4.01}_{-4.72}$	1.0
UDS-1244 ⊕	34.2895	-5.2698	3.79	$1.40^{+0.05}_{-1.25}$	$137.60^{+63.40}_{-72.10}$	1.4
UDS-2571 ⊕	34.2904	-5.2621	3.70	$1.00^{+0.50}_{-0.80}$	$69.39^{+43.01}_{-26.50}$	0.9
UDS-7520 ⊕	34.2559	-5.2338	3.17	$1.10^{+0.75}_{-0.75}$	$274.00^{+46.60}_{-149.50}$	0.8
UDS-10086 ⊕	34.3179	-5.2192	3.09	$0.90^{+0.95}_{-0.75}$	$31.87^{+20.59}_{-15.19}$	0.8
UDS-10430 ⊕	34.2806	-5.2172	4.13	$0.70^{+0.60}_{-0.50}$	$105.30^{+79.60}_{-56.97}$	0.6
UDS-20843 ⊕	34.4961	-5.1610	3.73	$1.40^{+0.05}_{-1.15}$	$95.64^{+53.56}_{-26.41}$	1.3
UDS-23628 ⊕	34.2426	-5.1431	4.25	$1.10^{+0.15}_{-0.90}$	$83.09^{+76.71}_{-32.42}$	1.1
UDS-25688 ⊕	34.5266	-5.1360	3.08	$1.70^{+0.20}_{-1.20}$	$261.70^{+50.40}_{-56.70}$	1.4
UDS-25893 ⊕	34.3996	-5.1363	4.49	$1.10^{+0.05}_{-0.75}$	$233.80^{+173.50}_{-117.10}$	0.7
UDS-4332	34.4657	-5.2519	3.18	$1.80^{+0.00}_{-1.79}$	$178.50^{+98.60}_{-128.59}$	1.5
UDS-7779	34.2589	-5.2323	3.14	$0.25^{+1.60}_{-0.00}$	$101.30^{+66.10}_{-35.36}$	0.1
UDS-8682	34.2937	-5.2270	3.46	$1.55^{+0.10}_{-1.40}$	$73.66^{+49.94}_{-38.80}$	1.5
UDS-8689	34.2741	-5.2274	3.22	$1.40^{+0.25}_{-1.15}$	$175.50^{+59.90}_{-79.74}$	1.4
UDS-11532	34.4207	-5.2116	4.21	$0.70^{+0.55}_{-0.55}$	$17.76^{+13.04}_{-7.94}$	0.7
UDS-12640	34.5341	-5.2050	3.61	$0.30^{+1.25}_{-0.15}$	$45.98^{+39.41}_{-14.58}$	0.2
UDS-32406	34.5426	-5.1861	3.28	$0.50^{+1.25}_{-0.35}$	$18.67^{+9.99}_{-9.51}$	0.3

Table B1: Physical properties of the red and dead candidates belonging to the "reference" sample, as obtained from their best fit with the top-hat SFH library without emission lines. ID$_{CAND}$ is the identification number in the CANDELS catalogues, followed by the WCS RA-DEC coordinates; z_{CAND} is the official CANDELS redshift; Age is the time span from the onset of the star formation activity to the epoch of observation; M_* is the stellar mass, which assumes a Salpeter IMF; Δt_{burst} is the duration of the star formation activity, during which the SFR is constant ($SFR = M*/\Delta t_{burst}$). The SFR at the epoch of observation is always zero, by definition. For each field, the table lists first the candidates which have also passed the "lines" selection (marked with a ⊕), and then the remaining objects in the "reference" sample. COSMOS ID-16676 is marked with a * because its properties have been obtained fitting at $z_{spec} = 3.72$ rather than at $z_{CANDELS} = 4.13$, which would yield: Age= $0.65^{+0.65}_{-0.35}$, $M_* = 362.20^{+231.10}_{-178.10}$, $\Delta t_{burst} = 0.3$.

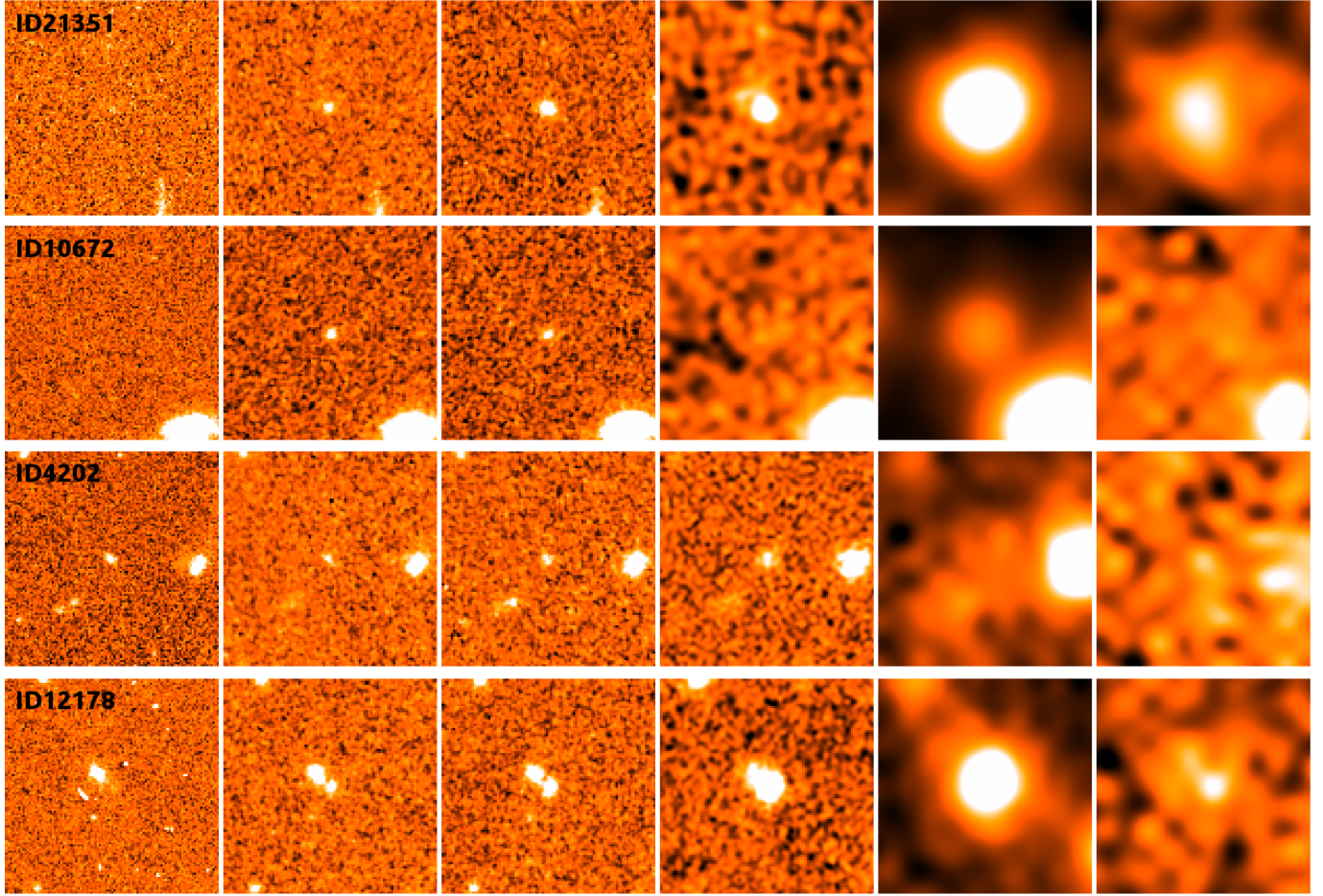## APPENDIX C: EXAMPLES OF SNAPSHOTS AND SEDS OF RED AND DEAD CANDIDATES

Figure C1: Illustrative snapshots of some passive candidates. Left to right: *ACS B*435 + *V*606 + *I*814 stack, *WFC*3 *J*125, *WFC*3 *H*160, *Ks*, IRAC 3.6 + 4.5 μm stack, IRAC 5.8 + 8.0 μm stack. Top to bottom: EGS-21351 (a robust candidate from the "lines" selection); GOODSN-10672 (the z_{CANDELS} = 6.7 candidate discussed in Sect. 4.3, again in the "lines" selection); GOODSS-4202 (a candidate from the "reference" selection, with a star-forming best fit in the run including emission lines; see also Fig. C2); and GOODSS-12178 (the candidate from the "lines" selection which has VANDELS z_{spec} = 0.56, but has a close companion which might contaminate the spectrum, as discussed in Sect. 3.2; this object has a passive best-fit in both the runs with and without emission lines, but it is excluded from the "lines" selection because it also has star-forming solutions with $p_{\mathrm{SF}} > 5\%$).
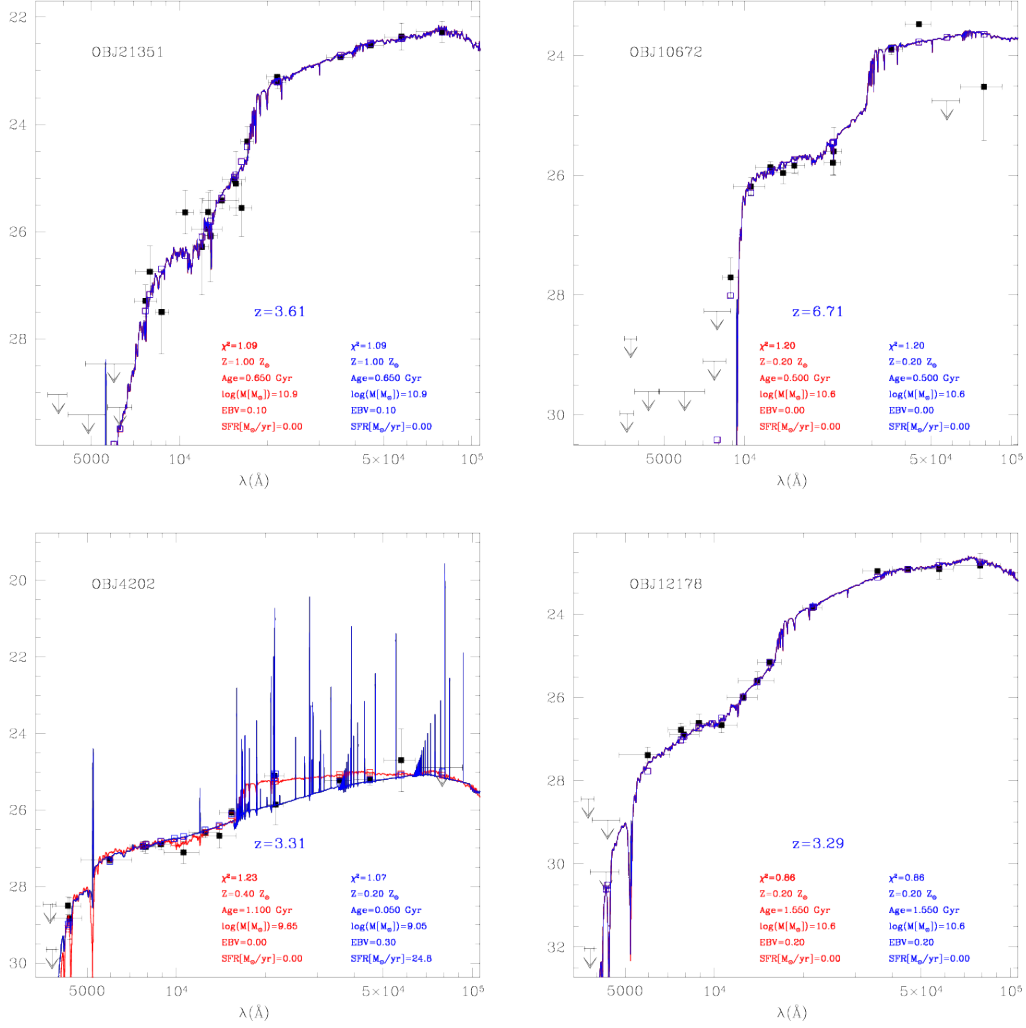
Figure C2: SEDs of the four passive candidates shown above. Black filled squares with error bars are the observed magnitudes. The red line is the SED of the best fit without emission lines, whereas the blue one corresponds the fit including the lines; empty squares of the same colors are the model magnitudes corresponding to the observed ones.

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