



# Quiescent Galaxies 1.5 Billion Years after the Big Bang and Their Progenitors

Francesco Valentino<sup>1,2</sup>, Masayuki Tanaka<sup>3,4</sup>, Iary Davidzon<sup>1,2</sup>, Sune Toft<sup>1,2</sup>, Carlos Gómez-Guijarro<sup>1,2,5</sup>, Mikkel Stockmann<sup>1,2</sup>, Masato Onodera<sup>6,7</sup>, Gabriel Brammer<sup>1,2</sup>, Daniel Ceverino<sup>1,2</sup>, Andreas L. Faist<sup>8</sup>, Anna Gallazzi<sup>9</sup>, Christopher C. Hayward<sup>10</sup>, Olivier Ilbert<sup>11</sup>, Mariko Kubo<sup>3</sup>, Georgios E. Magdis<sup>1,2,12,13</sup>, Jonatan Selsing<sup>1,2</sup>, Rhythm Shimakawa<sup>6</sup>, Martin Sparre<sup>14,15</sup>, Charles Steinhardt<sup>1,2</sup>, Kiyoto Yabe<sup>16</sup>, and Johannes Zabl<sup>17,18</sup>

<sup>1</sup> Cosmic Dawn Center (DAWN), Denmark; [francesco.valentino@nbi.ku.dk](mailto:francesco.valentino@nbi.ku.dk)

<sup>2</sup> Niels Bohr Institute, University of Copenhagen, Lyngbyvej 2, DK-2100 Copenhagen Ø, Denmark

<sup>3</sup> National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan

<sup>4</sup> Department of Astronomical Science, The Graduate University for Advanced Studies, SOKENDAI, Mishima, 411-8540, Japan

<sup>5</sup> Laboratoire AIM-Paris-Saclay, CEA/DSM-CNRS-Université Paris Diderot, Irfu Département d’Astrophysique, CEA Saclay, Orme des Merisiers, F-91191 Gif sur Yvette, France

<sup>6</sup> Subaru Telescope, National Astronomical Observatory of Japan, National Institutes of Natural Sciences (NINS), 650 North A’ohoku Place, Hilo, HI 96720, USA

<sup>7</sup> Department of Astronomical Science, SOKENDAI (The Graduate University for Advanced Studies), 650 North A’ohoku Place, Hilo, HI 96720, USA

<sup>8</sup> IPAC, California Institute of Technology, 1200 East California Boulevard, Pasadena, CA 91125, USA

<sup>9</sup> INAF—Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy

<sup>10</sup> Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

<sup>11</sup> Aix Marseille Univ, CNRS, CNES, LAM, Laboratoire d’Astrophysique de Marseille, Marseille, France

<sup>12</sup> DTU-Space, Technical University of Denmark, Elektrovej 327, DK-2800 Kgs. Lyngby, Denmark

<sup>13</sup> Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, GR-15236 Athens, Greece

<sup>14</sup> Institut für Physik und Astronomie, Universität Potsdam, Karl-Liebknecht-Str. 24/25, D-14476 Golm, Germany

<sup>15</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany

<sup>16</sup> Kavli Institute for the Physics and Mathematics of the Universe (WPI), UTIAS, The University of Tokyo, Kashiwa, Chiba 277-8583, Japan

<sup>17</sup> Institut de Recherche en Astrophysique et Planetologie (IRAP), Université de Toulouse, CNRS, UPS, F-31400 Toulouse, France

<sup>18</sup> Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France

Received 2019 September 23; revised 2019 December 19; accepted 2019 December 20; published 2020 January 29

## Abstract

We report two secure ( $z = 3.775, 4.012$ ) and one tentative ( $z \approx 3.767$ ) spectroscopic confirmations of massive and quiescent galaxies through  $K$ -band observations with Keck/MOSFIRE and Very Large Telescope/X-Shooter. The stellar continuum emission, absence of strong nebular emission lines, and lack of significant far-infrared detections confirm the passive nature of these objects, disfavoring the alternative solution of low-redshift dusty star-forming interlopers. We derive stellar masses of  $\log(M_*/M_\odot) \sim 11$  and ongoing star formation rates placing these galaxies  $\gtrsim 1$ –2 dex below the main sequence at their redshifts. The adopted parameterization of the star formation history suggests that these sources experienced a strong ( $\langle \text{SFR} \rangle \sim 1200$ – $3500 M_\odot \text{ yr}^{-1}$ ) and short ( $\sim 50$  Myr) burst of star formation, peaking  $\sim 150$ – $500$  Myr before the time of observation, all properties reminiscent of the characteristics of submillimeter galaxies (SMGs) at  $z > 4$ . We investigate this connection by comparing the comoving number densities and the properties of these two populations. We find a fair agreement only with the deepest submillimeter surveys detecting not only the most extreme starbursts but also more normal galaxies. We support these findings by further exploring the *Illustris* TNG cosmological simulation, retrieving populations of both fully quenched massive galaxies at  $z \sim 3$ –4 and SMGs at  $z \sim 4$ –5, with number densities and properties in agreement with the observations at  $z \sim 3$  but in increasing tension at higher redshift. Nevertheless, as suggested by the observations, not all of the progenitors of quiescent galaxies at these redshifts shine as bright SMGs in their past, and, similarly, not all bright SMGs quench by  $z \sim 3$ , both fractions depending on the threshold assumed to define the SMGs themselves.

*Unified Astronomy Thesaurus concepts:* Galaxy evolution (594); Elliptical galaxies (456); Galaxy stellar content (621); Star formation (1569); High-redshift galaxies (734); Starburst galaxies (1570)

## 1. Introduction

Decades of investigation allowed astrophysicists to clearly define a class of “quiescent” galaxies in the local universe. These systems are typically characterized by their large stellar masses and sizes, several billion years old stellar populations, red colors, little to no active formation of new stars, very limited amount of cold gas and dust, and overdense surrounding environment (Renzini 2006). The inferred old ages at  $z = 0$  suggest that these galaxies were already in place at high redshift (e.g., Gallazzi et al. 2005; Thomas et al. 2005), and indeed, a numerous population of massive, compact ( $\sim 1$  kpc) quiescent galaxies that are Gyr old already at  $z > 2$  has been

firmly established from photometry (e.g., Daddi et al. 2005; Trujillo et al. 2006; Toft et al. 2007; Cimatti et al. 2008; Cassata et al. 2013; Straatman et al. 2014) and now securely detected with spectrographs (e.g., Kriek et al. 2009; van Dokkum et al. 2009; Gobat et al. 2012; Toft et al. 2012; van de Sande et al. 2013; Belli et al. 2014, 2017a, 2017b; Stockmann et al. 2020). Their extreme stellar densities suggest that these compact quiescent systems at  $z \sim 2$  might be the remnants of an intense burst of star formation triggered by the rapid collapse of a large amount of gas at  $z > 4$ . In this scenario, dissipative gas-rich mergers, counterrotating gas streams, or disk instabilities would ignite star formation in high-redshift and dusty star-forming galaxies (DSFGs) detectable at

submillimeter wavelengths, quickly consuming the gas and leaving compact and passive remnants (e.g., Cimatti et al. 2008; Barro et al. 2013; Toft et al. 2014; Zolotov et al. 2015; Gómez-Guijarro et al. 2018, 2019; Tadaki et al. 2018). The matching number densities, sizes, masses, and formation timescales of submillimeter galaxies (SMGs) at  $z \sim 4\text{--}4.5$  and quiescent galaxies at  $z \sim 2$  may support this picture (Toft et al. 2014, and references therein). On the other hand, the mechanism physically responsible for the cessation of star formation in these massive systems is still a matter of discussion, with several scenarios still competing (see Man & Belli 2018 for a recent discussion).

This evolutionary scheme has been recently challenged by the spectroscopic confirmation of quiescent systems with  $M_\star \sim 10^{11} M_\odot$  above  $z > 3$  and up to  $z = 3.717$  (Gobat et al. 2012; Glazebrook et al. 2017; Simpson et al. 2017; Schreiber et al. 2018a, 2018b, hereafter S18b; Forrest et al. 2019; C. D’Eugenio et al. 2020, in preparation) as part of a substantial population of photometrically selected red galaxies (e.g., Fontana et al. 2009; Ilbert et al. 2013; Muzzin et al. 2013; Straatman et al. 2014; Mawatari et al. 2016; Davidzon et al. 2017; Deshmukh et al. 2018; Merlin et al. 2018, 2019; Girelli et al. 2019; Guarnieri et al. 2019, to mention recent results). In at least one case, their quiescent nature was initially challenged by submillimeter observations (Simpson et al. 2017) but later confirmed with a high spatial resolution follow-up, necessary to disentangle the emission of these galaxies from nearby companions (Schreiber et al. 2018b). Systematic studies of larger samples of  $z > 3$  photometric candidates in the submillimeter further support their average quiescence (Santini et al. 2019). The extreme masses, stellar densities, old ages, low star formation rates (SFRs), and number densities appear to be hardly reproducible by hydrodynamical simulations and semianalytical models at  $z > 3$  (Steinhardt et al. 2016, S18b; Cecchi et al. 2019). Catching quenched and quenching galaxies at the highest possible redshifts thus represents a formidable tool to test our galaxy formation models and simulations and, ultimately, the cosmology. For this scope, studying high-redshift galaxies allows us to better estimate their ages, as they are limited by the age of the universe (Belli et al. 2019).

Here we report the discovery of three massive objects at  $z = 3.77\text{--}4.01$  with suppressed star formation, followed up with Keck/Multi-Object Spectrometer For Infra-Red Exploration (MOSFIRE) and Very Large Telescope (VLT)/X-Shooter long-slit spectroscopy. In one case, the high quality of the data allowed us to estimate the stellar velocity dispersion in the highest-redshift target, opening the way to the study of its stellar dynamics and structure. We explored such properties in a dedicated companion paper (Tanaka et al. 2019). Based on the observed properties, we further investigated the expected characteristics of the progenitors of our sample of massive quiescent galaxies and compared them with dusty star-forming objects from surveys at  $z > 4$  selected based on their submillimeter fluxes, testing the evolutionary connection suggested for lower-redshift systems. Incorporating information about the previously confirmed quiescent galaxies at  $z \sim 3.5$ , we compared the number densities of this population and the putative SMG progenitors. Finally, we explored the content of the recent *Illustris* TNG cosmological simulation in order to look for rare quenched systems at high redshift and study their connection with their star-forming progenitors.

This paper is structured as follows. In Section 2 we introduce the sample that we follow up spectroscopically with the

observations described in Section 3. We present the data analysis in Sections 4–6. In Section 7 we explore the connection between our sample of quiescent galaxies and their progenitors at higher redshift, including the view offered by cosmological simulations (Section 8). Concluding remarks are collected in Section 9. We assumed a  $\Lambda$ CDM cosmology with  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$ , and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  and a Chabrier initial mass function (IMF; Chabrier 2003). All magnitudes are expressed in the AB system.

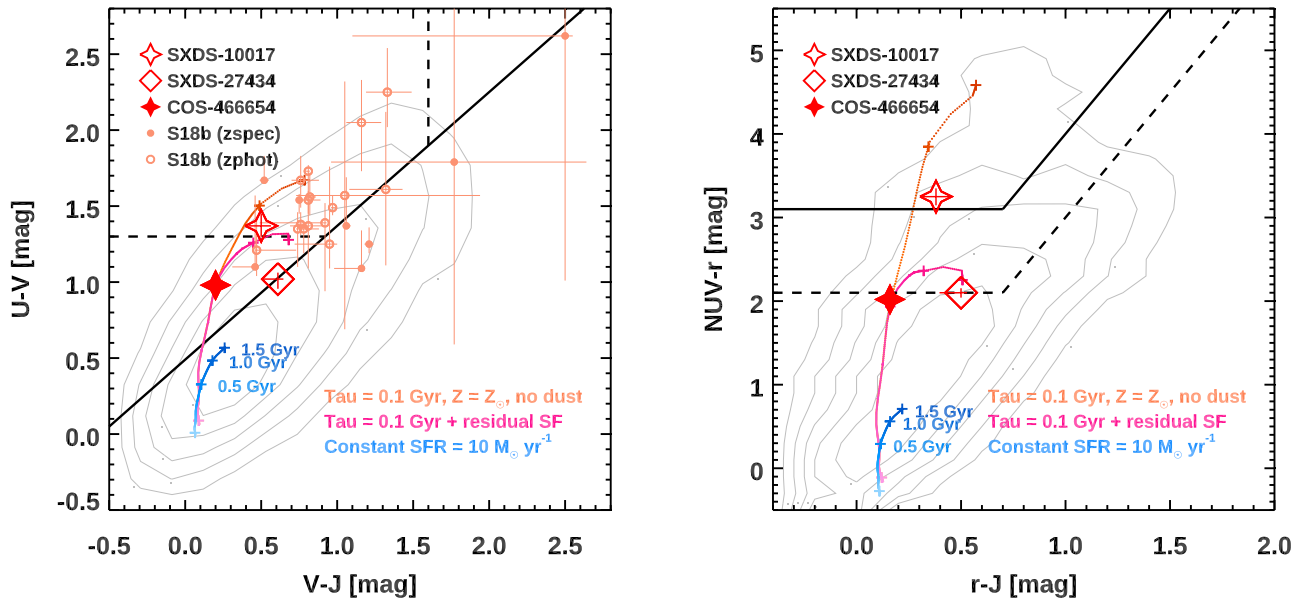
## 2. Selection

We based our search for suitable quiescent galaxy candidates on a combination of full modeling of the optical and near-infrared light and a rest-frame color–color selection. The former has proved to be a trustable way to select quiescent objects up to  $z \sim 2$  (Stockmann et al. 2020), and it naturally incorporates all of the information available from the photometry while relying on a set of assumptions on models and templates. Color–color diagrams are flexible instruments to broadly separate galaxy populations by capturing the main features with limited observations, but valuable information from the rest of the spectrum might be discarded from the analysis (see Merlin et al. 2018 for a recent detailed analysis).

Two members of our team fitted the optical/near-infrared spectral energy distributions (SEDs) of potential targets: M.T. modeled galaxies in the Subaru–*XMM-Newton* Deep Field (SXDS; Furusawa et al. 2008), and I.D. used the COSMOS field (Scoville et al. 2007). Having been designed for independent spectroscopic runs at different facilities and epochs, the original selection of candidate quiescent galaxies in the two fields was comparable but not identical.

### 2.1. SXDS/UDS Field

We performed the SED modeling for galaxies in the SXDS field as described in Kubo et al. (2018). Briefly, we ran the custom Bayesian photometric code MIZUKI (Tanaka 2015) on a multiwavelength catalog comprising *u*-band observations from CFHT/Megacam, optical *BVRiz* imaging from Subaru/Suprime-Cam (Furusawa et al. 2008), *JHK* bands from the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al. 2007, DR10), and *Spitzer* coverage from the UKIDSS Ultra Deep Survey (SpUDS; PI: J. Dunlop), covering an area of  $\sim 0.8 \text{ deg}^2$ . We modeled the SED adopting the Bruzual & Charlot (2003) models, exponentially declining star formation histories (SFHs), solar metallicities, the Calzetti et al. (1994) dust attenuation law, and the Chabrier (2003) IMF. Emission lines are included using the intensity ratios from Inoue (2011). The  $\text{Ly}\alpha$  escape fraction is assumed to be 0.1. Given the constraining power of the data covering the full SEDs of galaxies, we applied top-hat priors, assigning zero probability to templates outside observationally motivated ranges. The parameters include  $\tau$  ( $\in [1, 100]$  Gyr for the SFH model), dust attenuation (Calzetti law, with optical depth in the *V* band  $\in [0, 5]$ ), age ( $\in [0.001, 14]$  Gyr, excluding ages larger than the age of the universe at each redshift), and redshift ( $\in [0, 6]$ ), while the metallicity was fixed to solar. We confirmed that our results do not significantly change if we apply the physical priors discussed in Tanaka (2015). We then selected a sample of five candidate “quiescent” and 12 “quenching” (or “post-starburst”) galaxies based on their specific SFRs ( $\text{sSFR} \leq 10^{-11} \text{ yr}^{-1}$  and  $10^{-11} \text{ yr}^{-1} \leq \text{sSFR} \leq 10^{-10} \text{ yr}^{-1}$ ), redshift ( $4 < z_{\text{phot}} < 4.6$ ),



**Figure 1.** Rest-frame colors. Left:  $UVJ$  colors for a sample of  $3 < z < 4$  galaxies in the UDS (7277 objects) and COSMOS (11,653) fields (background gray contours), our observed  $z = 3.77\text{--}4.01$  quiescent objects (red filled and open stars, red open diamond), and the spectroscopic and photometric sample from Schreiber et al. (2018a; filled and open orange circles, respectively). The background contours show the density of UDS+COSMOS points from the parent samples in a bin of 0.2 mag on both axes, smoothed over  $2 \times 2$  bins with a boxcar average filter. Right:  $NUVrJ$  colors for the same objects in the left panel. The same symbols and colors apply here. The background contours trace the density of UDS+COSMOS points from the parent samples in a bin of 0.2 mag on both axes. The tracks in both panels show the evolution of synthetic Bruzual & Charlot (2003) models of rapidly quenched galaxies (exponentially declining SFH with  $\tau = 0.1$  Gyr,  $Z = Z_{\odot}$ , no dust; orange line) with possible residuals of late star formation ( $1 M_{\odot} \text{ yr}^{-1}$  after quenching for a galaxy of  $10^{10} M_{\odot}$ ; pink line) or systems characterized by a constant SFH (blue line), the latter simulating an active SFG. The plus signs mark the time steps as indicated by the blue labels, identical for all tracks.

and reliability of the fit (reduced  $\chi^2 < 4.5$ ). The adopted redshift cut would have allowed us to observe the 4000 Å break in the  $K$  band, providing a constraint on the age of the stellar populations. In order to minimize the exposure time for a spectroscopic follow-up, we finally selected the quiescent candidate SXDS-10017 at  $z_{\text{phot}} = 4.07^{+0.07}_{-0.07}$  and the quenching object SXDS-27434 at  $z_{\text{phot}} = 4.12^{+0.03}_{-0.05}$ , the brightest galaxies among our initial pool of sources ( $K = 22.5$  and 21.9 mag, respectively). Consistent with the sSFR cut and quality checks, SXDS-10017 falls in the quiescent region of both the  $UVJ$  (Williams et al. 2009) and  $NUVrJ$  (Ilbert et al. 2010) rest-frame color diagrams (Figure 1). On the other hand, consistent with the looser constraint on the sSFR, SXDS-27434 falls outside the canonical  $UVJ$  and  $NUVrJ$  limits for quiescent galaxies (Ilbert et al. 2010; Williams et al. 2009), showing rather blue  $U - V$  and  $NUV - r$  colors. However, this is not unexpected for recently and abruptly quenched galaxies at high redshift (Merlin et al. 2018). A posteriori, using the spectroscopic or photometric redshift estimates does not significantly change the location of our targets in the color-color diagrams. In Figure 1, we show the expected color evolution for galaxies with an exponentially declining SFH ( $\tau = 0.1$  Gyr) compared with objects with the same SFH but residuals of late star formation (a constant  $1 M_{\odot} \text{ yr}^{-1}$  after quenching for a galaxy of  $10^{10} M_{\odot}$ ) and active SFGs with a constant SFH of  $10 M_{\odot} \text{ yr}^{-1}$ . The tracks are based on Bruzual & Charlot (2003) models with no dust and a  $Z_{\odot}$  metallicity (see Belli et al. 2019 for a similar attempt with different parameters and at lower redshift). We therefore expected SXDS-27434 to show younger stellar populations than SXDS-10017, allowing us to probe the post-starburst or quenching epoch at  $z \sim 4$ .

## 2.2. COSMOS Field

We remodeled the SEDs of galaxies in the COSMOS field ( $\sim 1.8 \text{ deg}^2$ ) with LEPHARE (Arnouts et al. 1999; Ilbert et al. 2006) following Davidzon et al. (2017) based on the 30-band photometric catalog by Laigle et al. (2016). We adopted Bruzual & Charlot (2003) stellar population models, exponentially declining and delayed SFHs, solar and subsolar metallicities ( $Z = 0.4 Z_{\odot}$ ), and a Chabrier (2003) IMF. For the estimate of the photometric redshifts, we adopted a Small Magellanic Cloud extinction law (Prevot et al. 1984) and three different flavors of the Calzetti et al. (2000) prescription, including the bump at 2700 Å. On the other hand, we used the Calzetti et al. (2000; allowing for the bump) and Arnouts et al. (2013) extinction laws when estimating the physical parameters. We then selected two quenching candidates with  $K_s < 23$  mag and  $NUVrJ$  colors consistent with an sSFR  $< 10^{-10} \text{ yr}^{-1}$  (Figure 1), excluding solutions at  $\chi^2 > 10$ . We finally chose the brightest among our candidates at  $z > 4$  (COS-466654;  $z_{\text{phot}} = 4.11^{+0.04}_{-0.1}$  and  $K_s = 22.26$ ) so as to target the 4000 Å break in the  $K$  band. In agreement with the initial constraint on the  $NUVrJ$  colors, COS-466654 falls in the post-starburst or quenching regions of the rest-frame color diagrams, as in the case of SXDS-27434 (Figure 1, Section 2.1).

## 2.3. Consistency of the Applied Selection Criteria

As shown in Figure 1, the rest-frame color and sSFR selection criteria are consistent. In fact, the same galaxies would have been targeted if we also applied the COSMOS selection criteria in SXDS, and vice versa. However, we used different SED fitting codes to implement the two selections. To test possible systematics introduced by this choice, we applied



the modeling that we used for the SXDS field to the COSMOS catalog and vice versa. A small difference in redshifts emerged from this test. For the sources in SXDS, LEPHARE returns  $z_{\text{phot}} = 3.58$  and  $4.12$  for SXDS-10017 and SXDS-27434, respectively. However, in contrast to the MIZUKI selection, the one with LEPHARE does not strictly exclude candidates at  $z_{\text{phot}} < 4$ , considering the typical accuracy of redshift measurements in the distant universe. On the other hand, we retrieve a slightly lower redshift for COS-466654 with MIZUKI ( $z_{\text{phot}} = 3.72^{+0.08}_{-0.07}$ ), but its quiescent state is confirmed. Notice that the spectroscopic constraints for SXDS-10017 and COS-466654 fall in between the photometric estimates above (Section 4). This kind of discrepancy may affect statistical studies that purely rely on the parent photometric samples, but they are less relevant for the present analysis, since we focus on quiescent galaxies individually confirmed via spectroscopy.

### 3. Spectroscopic Observations

#### 3.1. Keck/MOSFIRE Observations of the UDS Field

We observed the two targets in the UDS field with MOSFIRE (McLean et al. 2012) at the Keck I telescope on two separate runs. On 2017 November 23, we collected a total of 4 hr of observations of SXDS-10017 out of a complete night initially granted, with an average seeing of  $\text{FWHM} = 0''.8$  during the night, as estimated from individual stars in the field. We observed SXDS-27434 on 2018 December 20–21 for 7.75 hr, with an average seeing of  $\text{FWHM} = 0''.7$  over 2 half-nights. In both cases, we observed the targets in the  $K$  band with a slit width of  $0''.7$ , ensuring a nominal initial spectral resolution of  $R \sim 3600$ . We adopted the standard ABBA nodding technique with a dithering of  $1''.5$  and 180 s exposures to allow for an optimal background subtraction. We reduced the data with the MOSFIRE pipeline and obtained a final flux-calibrated, optimally combined 2D spectrum. We corrected for aperture losses by modeling the galaxies as 2D Gaussian curves with  $\text{FWHM} = \text{FWHM}_{\text{seeing}}$  and calculating the light lost outside a  $0''.7$  wide rectangular slit. The measured effective major axis of the targets ( $R_{\text{eff,maj}} = (0.95 \pm 0.32)$  and  $(0.76 \pm 0.20)$  kpc for SXDS-10017 and 27434, respectively; Kubo et al. 2018; Tanaka et al. 2019) effectively allows us to treat the targets as point sources and adopt this simple approach. We further checked for possible residual telluric absorptions due to the varying airmass during the night by calibrating the spectra of several stars observed simultaneously with our science target, resulting in a negligible effect. We rebinned the 2D spectra by  $2\times$ ,  $4\times$ , and  $30\times$  to a final resolution of 4.3, 8.7, and  $65 \text{ \AA}$  for testing and displaying purposes. We applied a running optimal weighted mean, providing the maximal signal-to-noise ratio (S/N) among the several approaches we attempted (median, mean, and clipped mean). We simultaneously increased the noise to account for the possible correlation among adjacent spectral elements by forcing the reduced  $\chi^2_{\text{red}} = 1$  computed in regions of the 2D frame of pure background. This approach is applicable, since the noise variations due to sky lines occur on scales smaller than the absorption features we aim to detect. We optimally extracted the spectrum and its associated noise following Horne (1986). We finally corrected the 1D spectrum for possible residual flux losses by computing the synthetic photometry and anchoring it to the best model reproducing the photometry (Section 5). The data reduction process for SXDS-10017 resulted in a final median of  $\text{S/N} = 6.7$  and a maximum

of  $\text{S/N} = 11.6$  for bins of  $65 \text{ \AA}$ , comparable with the performances reported in S18b for similar integrations and  $K$ -band luminosities. Consistent with the brighter  $K$ -band magnitude and  $\sim 2\times$  longer integration, we find a final median of  $\text{S/N} = 16.4$  and a maximum of  $\text{S/N} = 30.0$  for bins of  $65 \text{ \AA}$  for SXDS-27434. The 2D frames and optimally extracted spectra are shown in Figure 2.

#### 3.2. VLT/X-Shooter Observations of the COSMOS Field

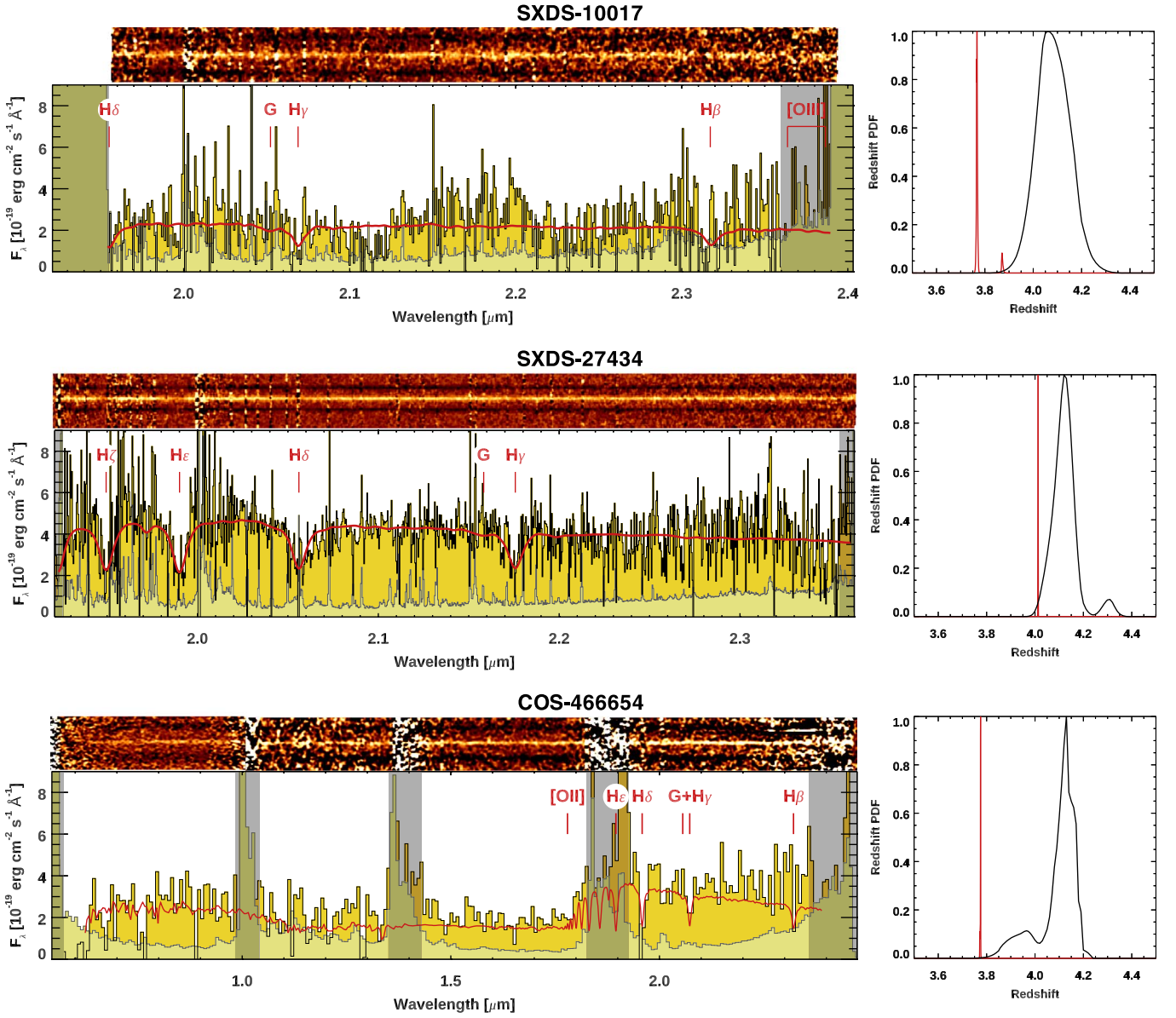
We observed the target in the COSMOS field with the cross-dispersed echelle spectrograph X-Shooter (Vernet et al. 2011) mounted on UT2 at the VLT. The observations were carried out in service mode over 2018 March–April. A total of 8.6 hr of spectroscopic integration were spent on target, with an average seeing of  $\text{FWHM} = 0''.66$ . We observed the target with slit widths of  $1''.0$ ,  $0''.9$ , and  $0''.9$  in the UVB, VIS, and NIR arms, respectively, ensuring a lower limit on the resolving power of 4350, 7450, and 5300 in the three arms, given the average seeing smaller than the slit widths. The chosen configuration and seeing conditions allowed us to cover the wavelength range from 3000 to 24800  $\text{\AA}$  with minimal slit losses. We adopted the standard ABBA dithering technique with a nod throw of  $4''.5$  and jitter width of  $1''$ . We optimized the observing time for 1 hr observing blocks (OBs), exposing for 420, 448, and 480 s in the UVB, VIS, and NIR arms, respectively. For the only OB of 0.5 hr, we integrated for 534, 563, and 600 s in the three arms. The spectra have been bias-corrected, flat-fielded, wavelength-calibrated, rectified, and flux-calibrated using observations of spectrophotometric standards with the VLT/X-Shooter pipeline (Modigliani et al. 2010). We supplemented the pipeline with optimal frame combination, telluric correction, and a slit flux loss estimate using the customized scripts described in Selsing et al. (2019) and publicly available online.<sup>19</sup> We rebinned the 2D spectrum using a weighted mean to a final wavelength step of  $1.8 \text{ \AA}$  in the NIR arm to match the resolution of MOSFIRE and ensure a minimum S/N to measure the redshift. We further rebinned the spectrum up to  $72 \text{ \AA}$  for display purposes. We optimally extracted the spectrum (Horne 1986) and corrected the residual flux losses due to seeing variations over the epochs of observations by computing the synthetic photometry in the available optical and near-infrared bands and tying them to the best model representing the observed photometry (Section 5). The  $72 \text{ \AA}$  rebinned combined frame and the optimally extracted spectrum are shown in Figure 2. The final median S/N over bins of  $72 \text{ \AA}$  is 3.3, 1.7, 3.8, and 4.1 in the VIS arm and  $J$ ,  $H$ , and  $K$  bands, with a maximum of  $\text{S/N} = 7.4$ , 3.7, 6.7, and 9.4, respectively. We did not detect significant emission in the UVB arm.

### 4. Redshift Estimate

We estimated the redshift using SLINEFIT,<sup>20</sup> a flexible algorithm based on  $\chi^2$  minimization that allows for simultaneous continuum template matching and emission line measurements. We initially fit the spectra with the best SED models obtained fixing  $z = z_{\text{phot}}$ . We convolved these models with a Gaussian curve to reach a stellar velocity dispersion of 230 and  $214 \text{ km s}^{-1}$  for SXDS-10017 and COS-466654, respectively, following the  $\sigma_{\text{vel}}-M_{\star}$  relation as in S18b (Belli et al. 2017a). This is an

<sup>19</sup> [https://github.com/jselsing/XSGBR\\_reduction\\_scripts](https://github.com/jselsing/XSGBR_reduction_scripts)

<sup>20</sup> <https://github.com/cscreib/slinefit>



**Figure 2.** Spectra of the quiescent galaxies. Top: Keck/MOSFIRE  $K$ -band spectrum of SXDS-10017, rebinned to a wavelength bin of  $8.7 \text{ \AA}$ . Middle: Keck/MOSFIRE  $K$ -band spectrum of SXDS-27434 at its original  $2.17 \text{ \AA}$  resolution. Bottom: VLT/X-Shooter VIS+NIR spectrum of COS-466654, rebinned to a wavelength bin of  $72 \text{ \AA}$ . In every panel, the 2D frame is smoothed with a Gaussian kernel of 2 pixels in width. The orange and gold areas mark the optimally extracted 1D spectrum and its noise, respectively. The spectra are rebinned differently to reach comparable levels of S/N. The best SED models with  $z = z_{\text{spec}}$  obtained including the  $\sim 2 \text{ \AA}$  resolution spectra are shown in red. Each SED model is broadened as mentioned in Section 4. The locations of the main absorption features are labeled. The darker areas indicate the wavelengths of poor atmospheric transmission between the observed bands. Right: The black and red solid lines indicate the probability distribution functions for the photometric and spectroscopic redshift, respectively.

assumption, as we cannot constrain the stellar velocity dispersion given the available combination of S/N and spectral resolution. On the contrary, for SXDS-27434, we used the measured dispersion of  $\sigma_{\text{vel}} = 268 \pm 59 \text{ km s}^{-1}$  (Tanaka et al. 2019), consistent with the  $\sigma_{\text{vel}}-M_*$  relation we adopted for the other two sources. We further took into account the possible presence of emission lines with fixed velocity widths ranging between 60 and  $300 \text{ km s}^{-1}$ . However, the redshift search returns the same results with or without the presence of emission lines. We explored the redshift range  $2 < z < 5$  and then refined the redshift measurement within  $\pm 0.2$  from the best fit over the large interval. We ran the code on the  $\sim 2 \text{ \AA}$  resolution spectra and  $2\times$  and  $3\times$  binned frames. This did not impact the final solution of  $4.0127^{+0.0005}_{-0.0005}$  for SXDS-27434 consistently with the one-peak redshift probability

distribution (RPD) and the high probability associated with it ( $p = 100\%$ , integrating the RPD within  $\pm 0.01$  from the best fit; Benítez 2000; Brammer et al. 2008; reduced  $\chi^2 = 1.3$ ). Using the penalized pixel fitting algorithm (pPXF; Cappellari & Emsellem 2004; Cappellari 2017) and a slightly different set of assumptions returns a fully consistent estimate (Tanaka et al. 2019). Similarly secure is the solution of  $z = 3.775^{+0.002}_{-0.003}$  for COS-466654 ( $p = 98\%$ , reduced  $\chi^2 = 1.1$ ). On the other hand, the solution for SXDS-10017,  $z = 3.767^{+0.103}_{-0.001}$ , is more uncertain, and it varies by 0.015 when using the native resolution or mildly binned spectra. This uncertainty manifests itself as a second peak of the RPD at  $z = 3.871$  and with a lower probability associated with the best solution ( $p = 81\%$ , reduced  $\chi^2 = 1.6$ ). We conservatively adopt the solution derived

with the native resolution, despite the significant improvement in probability ( $p \sim 100\%$ ) when running SLINEFIT on the  $3 \times$  binned spectrum. In every case, we rescaled the RPD by the empirical factor  $C = 2$ , as described in S18b ( $P(z) \propto \exp[(\chi^2(z) - \chi_{\min}^2)/2C]$ ), in order to take into account the noise on scales of a few spectral elements relevant for the template matching. We derived independent symmetrical uncertainties randomly perturbing and refitting the spectrum 1000 times, obtaining consistent results ( $z = 4.0127_{-0.0004}^{+0.0004}$ ,  $z = 3.775_{-0.004}^{+0.004}$ , and  $z = 3.767_{-0.051}^{+0.051}$  for SXDS-27434, COS-466654, and SXDS-10017, respectively). We finally refit the spectra with the best SED models obtained, fixing  $z = z_{\text{spec}}$ . In every case, we find that our initial photometric redshift overestimated the spectroscopic determination (Figure 2). As a result, the initial choice of  $z_{\text{phot}} > 4$  candidates to measure the 4000 Å break in the  $K$  band did not have success. This is likely due to the choice of following up the brightest targets, biasing against breaks fully enclosed in the  $K$  band. Based on the criteria defined in S18b, we can consider the redshifts for SXDS-27434 and COS-466654 “robust,” while the estimate for SXDS-10017 is “uncertain.”

### 5. Modeling of the SED

In order to derive detailed physical properties, we remodeled the photometry and the rebinned spectra simultaneously with FAST++,<sup>21</sup> fixing  $z = z_{\text{phot,spec}}$  and the maximum possible age to the age of the universe  $t = t_{\text{obs}} (z = z_{\text{phot,spec}})$ . The results are robust against the use of the photometric or spectroscopic redshift. We assumed Bruzual & Charlot (2003) stellar population models, the Chabrier (2003) IMF, and the Calzetti et al. (2000) dust attenuation law, allowing for extinction values in the range  $A_V = 0\text{--}6$  mag. We fixed the metallicity to solar  $Z = Z_{\odot} = 0.02$ , as is reasonable for very massive objects. We then adopted multiple analytical parameterizations of the SFHs.

1. The delayed exponentially declining form  $\text{SFR}(t) \propto te^{-t/\tau}$ , where  $t$  is time, widely adopted in the literature. We allowed  $\tau$  to vary within steps of 0.1 dex within  $\log(\tau/\text{yr}^{-1}) = [6.5, 10]$  and set a minimum age of 100 Myr.
2. A truncated model, consisting of a constant SFR ( $t$ ) over an interval  $t_{\text{CSF}}$  starting at an onset time  $t_{\text{onset}}$  and then instantaneously switching to  $\text{SFR} = 0 \text{ } M_{\odot} \text{ yr}^{-1}$ . The duration  $t_{\text{CSF}}$  is free to vary within  $\log(t_{\text{CSF}}/\text{yr}^{-1}) = [6.5, t_{\text{obs}}]$  in steps of 0.1 dex.
3. The composite SFH described in S18b, to allow direct comparison with their sample of massive quiescent galaxies at similar redshifts. This SFH consists of exponentially rising and declining phases with e-folding times free to vary,

$$\text{SFR}_{\text{base}}(t) \propto \begin{cases} e^{(t_{\text{burst}}-t)/\tau_{\text{rise}}} & \text{for } t > t_{\text{burst}} \\ e^{-(t-t_{\text{burst}})/\tau_{\text{decl}}} & \text{for } t \leq t_{\text{burst}} \end{cases} \quad (1)$$

and  $t$  is the look-back time. We adopted the same grid of possible parameters as in S18b:  $t_{\text{burst}} = [10 \text{ Myr}, t_{\text{obs}}]$  with (logarithmic) steps of 0.05 dex, and  $\tau_{\text{rise}}, \tau_{\text{decl}} = [10 \text{ Myr}, 3 \text{ Gyr}]$  with steps of 0.1 dex. As in S18b, we further included an extra degree of freedom to decouple the current SFR from the previous history of formation, allowing for a burst or abrupt quenching on a short period

of duration  $t_{\text{free}}$  (Ciesla et al. 2016),

$$\text{SFR}(t) = \text{SFR}_{\text{base}}(t) \times \begin{cases} 1 & \text{for } t > t_{\text{free}}, \\ R_{\text{SFR}} & \text{for } t \leq t_{\text{free}}, \end{cases} \quad (2)$$

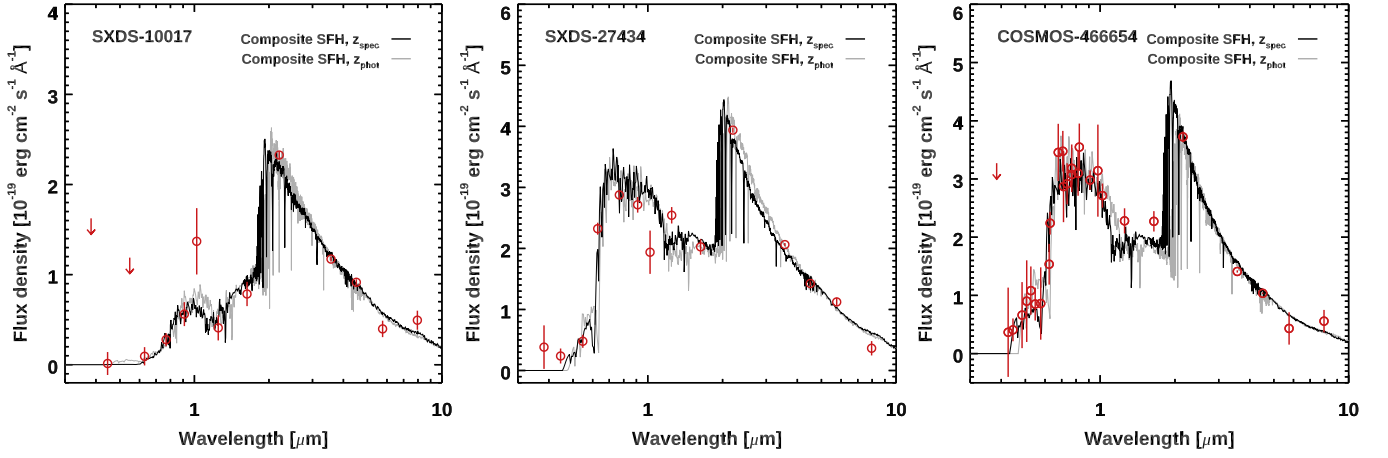
where  $t_{\text{free}}$  is free to vary between 10 and 300 Myr with steps of 0.5 dex, and  $R_{\text{SFR}}$  is within  $10^{-2}$  and  $10^5$  with steps of 0.2 dex.

In order to properly compare the results from the different SFH parameterizations, we computed several integrated quantities (e.g., Pacifci et al. 2016; Belli et al. 2019; S18b). We adopt the same terminology as in S18b to simplify the comparison between the two works. We define the epoch of assembly as the “half-mass formation time”  $t_{\text{form}}$ , i.e., the time at which 50% of the total stellar mass was formed, excluding mass loss and recycling, obtained integrating  $\text{SFR}(t)$  over time. The “duration of the main formation epoch”  $\Delta t_{\text{form}}$  is the contiguous period enclosing  $t_{\text{form}}$  and 68% of the total integrated SFR, i.e., limited by the 16th and 84th percentiles of the integral of  $\text{SFR}(t)$  over time. We assumed the mean  $\langle \text{SFR} \rangle_{\text{main}}$  during this period as representative of “the typical SFR during the main mass assembly episode.” We computed “the quenching epoch”  $t_{\text{quench}}$  as the initial point of the longest contiguous time interval starting from the time of observation  $t_{\text{obs}}$  (at  $z = z_{\text{spec}}$ ) and going backward, where  $\text{SFR} < 10\% \langle \text{SFR} \rangle_{\text{main}}$ . We finally adopted the uncertainties on the individual parameters estimated by FAST++ following the  $\Delta\chi^2 = (\chi^2 - \min\{\chi^2\}) < 2.71$  criterion to encompass the 90% confidence interval (Avni 1976; Schreiber et al. 2018b, S18b). We further cross-checked the uncertainties by bootstrapping 100 (1000) Monte Carlo simulations for COS-466654 and SXDS-27434 (SXDS-10017) with the same code. The number of simulations for COS-466654 and SXDS-27434 was limited by the available computational time, as their spectra have  $>2000$  individual elements. The numerical approach results in less conservative error bars than the analytical one, as previously found for similar high-redshift quiescent objects (S18b). We therefore adopted the  $\chi^2$  criterion as the final estimate for uncertainties derived from the modeling of the SED.

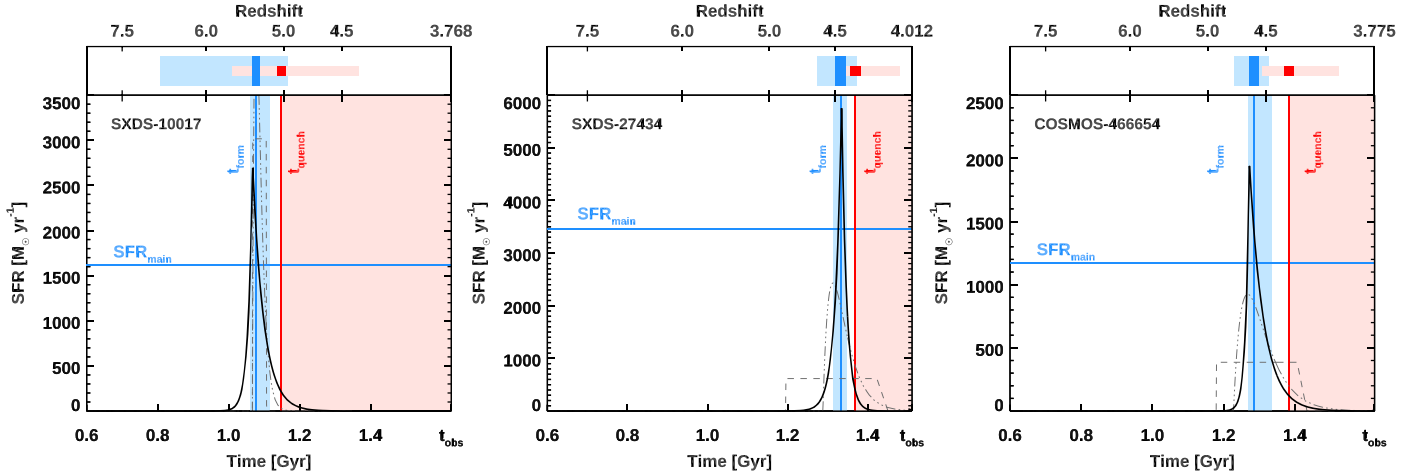
We show the best-fit SEDs in Figure 3, resulting from the composite SFHs in Figure 4. The best models based on the three parameterizations of the SFHs are indistinguishable, and the resulting parameters are consistent with each other within the uncertainties (Appendix A). Similarly, we tested the choice of the Bruzual & Charlot (2003) models by comparing with the set from Conroy & Gunn (2010), retrieving consistent results within the error bars (Appendix A). From here on, we therefore adopt the double exponential SFH as a reference in order to facilitate the comparison with the sample in Schreiber et al. (2018a) using the Bruzual & Charlot (2003) templates. The best-fit parameters are reported in Table 1. We find that all three sources went through a major burst of star formation with  $\log(\langle \text{SFR} \rangle_{\text{main}} [M_{\odot} \text{ yr}^{-1}]) \sim 3.07\text{--}3.54$  ( $\langle \text{SFR} \rangle_{\text{main}} \sim 1200\text{--}3500 \text{ } M_{\odot} \text{ yr}^{-1}$ ) over a short period of time,  $\Delta t_{\text{form}} \sim 50 \text{ Myr}$ . Notice that this refers only to the last episode of star formation. The abrupt quenching following the rapid formation occurred earlier in time for SXDS-10017 ( $z_{\text{quench}} = 5.02_{-0.61}^{+0.62}$ ) than for SXDS-27434 and COS-466654 ( $z_{\text{quench}} = 4.41_{-0.31}^{+0.21}$  and  $4.37_{-0.39}^{+0.21}$ , respectively). This naturally follows our initial selection and the rest-frame colors, once we have ascertained

<sup>21</sup> <https://github.com/cschreib/fastpp>





**Figure 3.** The SED of the quiescent galaxies. The gray and black solid lines show the best modeling of the SED with fixed  $z_{\text{phot}}$  and  $z_{\text{spec}}$ , respectively. In both cases, we show the results obtained with the composite SFH parameterization by Schreiber et al. (2018a). Red open circles and arrows mark the photometric points and  $3\sigma$  upper limits.



**Figure 4.** The SFH from modeling of the SED. The black line shows the composite SFH (Equations (1) and (2)) corresponding to the best model representing the SED in Figure 3. The blue vertical line and shaded area mark the main formation epoch  $t_{\text{form}}$  and duration of the star formation episode  $\Delta t_{\text{form}}$  of the galaxy. The red line indicates the quenching time  $t_{\text{quench}}$ . The blue horizontal line shows the mean  $\langle \text{SFR} \rangle_{\text{main}}$  during the main formation epoch. The dashed and dotted-dashed gray lines mark the best truncated and delayed SFHs. The blue and red bands in the top insets mark the formation and quenching redshifts  $z_{\text{form}}$  and  $z_{\text{quench}}$  and their 90% confidence intervals, respectively (Table 1).

the spectroscopic redshift of the sources and excluded the contamination of low-redshift interlopers.

## 6. Quiescence

From the SED modeling, we estimate stellar masses of  $\log(M_*/M_\odot) = 10.89^{+0.05}_{-0.06}$ ,  $10.82^{+0.03}_{-0.03}$ , and  $11.06^{+0.04}_{-0.04}$  and SFRs of  $\log(\text{SFR}/M_\odot \text{ yr}^{-1}) < 0.05$  (90% upper limit),  $0.46^{+0.13}_{-0.69}$ , and  $1.38^{+0.28}_{-1.25}$  for SXDS-10017, COS-466654, and SXDS-27434, respectively. These estimates place the galaxies  $< 2.1$ ,  $1.6$ , and  $1.0$  dex below the main sequence of galaxies at their redshift, adopting the parameterization of Schreiber et al. (2015). We show the location of our targets in the  $M_*$ –SFR plane in Figure 5, along with the sample of similarly selected massive quiescent galaxies at  $3 < z < 4$  from Schreiber et al. (2018a), the massive object at  $z = 3.493$  reported in Forrest et al. (2019), and the passive galaxy at  $z = 2.99$  confirmed through *Hubble Space Telescope* WFC3 slitless spectroscopy by Gobat et al. (2012).

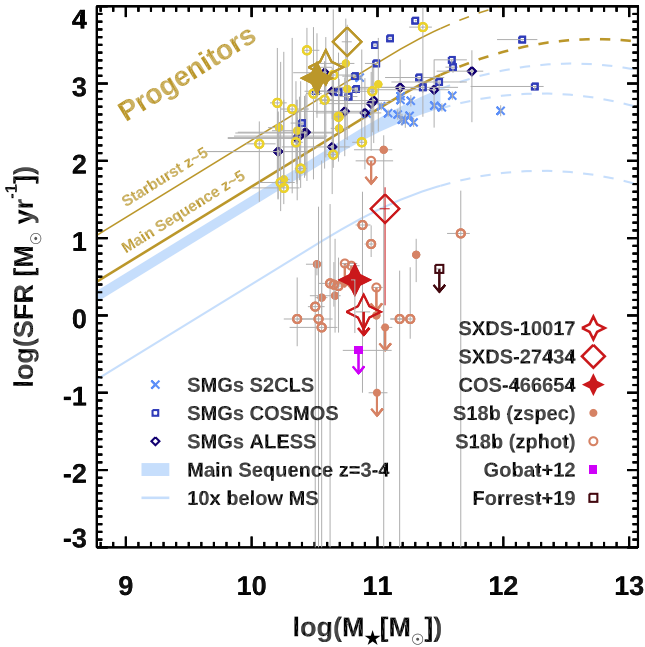
We further estimated an upper limit on the SFR from  $\text{H}\beta$  and  $[\text{O II}]$  emission lines, when covered by our observations. We do not identify any significant detections from the line search in the original resolution and the  $\sim 2 \text{ \AA}$  binned spectra. No evident residual emission appears when subtracting the best stellar SED continuum model from the spectra. We therefore put upper limits on the line fluxes as  $\sqrt{\sum_i \sigma_i^2}$ , where  $\sigma_i$  is the noise per  $i$ th spectral bin covered by the potential line. As line widths, we assumed the FWHM of the stellar models we adopted to estimate the redshift (Section 4). We then converted the upper limits into SFRs following Kennicutt (1998), modified according to our Chabrier (2003) IMF. We adopted the  $A_V$  extinction from the best SED model to correct for the dust attenuation. To be more conservative, we also computed a final upper limit using the 90% upper limit on  $A_V$  and including a possible extra absorption for emission lines following the correction described in Kashino et al. (2019;  $E_{\text{neb}}(B - V) = E_{\text{star}}(B - V)/0.69$ , adopting a Calzetti et al. 2000 extinction law for both nebular and stellar emission). For SXDS-10017,

**Table 1**  
Physical Properties of the Quiescent Galaxies

Properties <sup>a</sup>	SXDS-10017	SXDS-27434	COS-466654
Coordinates (R.A., decl.)/deg.	(34.756250, −5.308038)	(34.29871, −4.98987)	(149.419583, 2.007550)
$z_{\text{phot}}$	$4.07^{+0.12}_{-0.10}$	$4.12^{+0.09}_{-0.08}$	$4.11^{+0.06}_{-0.21}$
$z_{\text{spec}}$	$3.767^{+0.103}_{-0.001}$	$4.0127^{+0.0005}_{-0.0005}$	$3.775^{+0.002}_{-0.003}$
$\log(M_*/M_\odot)$	$10.89^{+0.05}_{-0.06}$	$11.06^{+0.04}_{-0.04}$	$10.82^{+0.03}_{-0.03}$
$\log(\text{SFR}_{\text{SED}}/M_\odot \text{ yr}^{-1})$	$<0.05$	$1.38^{+0.28}_{-1.25}$	$0.46^{+0.13}_{-0.69}$
$\log(\text{SFR}_{\text{H}\beta}/M_\odot \text{ yr}^{-1})$	$<0.53$	...	$<0.92$
$\log(\text{SFR}_{[\text{O II}]} / M_\odot \text{ yr}^{-1})$	...	...	$<0.41$
$A_V/\text{mag}$	$0.2^{+0.2}_{-0.2}$	$0.7^{+0.1}_{-0.1}$	$0^{+0}_{-0}$
$t_{\text{form}}/\text{Gyr}$	$1.08^{+0.09}_{-0.27}$	$1.33^{+0.04}_{-0.06}$	$1.28^{+0.04}_{-0.05}$
$z_{\text{form}}$	$5.34^{+1.34}_{-0.37}$	$4.51^{+0.16}_{-0.11}$	$4.64^{+0.15}_{-0.12}$
$t_{\text{quench}}/\text{Gyr}$	$1.15^{+0.22}_{-0.14}$	$1.37^{+0.11}_{-0.05}$	$1.38^{+0.14}_{-0.08}$
$z_{\text{quench}}$	$5.02^{+0.62}_{-0.61}$	$4.41^{+0.13}_{-0.31}$	$4.37^{+0.21}_{-0.39}$
$\Delta t_{\text{form}}/\text{Myr}$	$50^{+730}_{-27}$	$32^{+140}_{-10}$	$56^{+140}_{-41}$
$\log(\langle \text{SFR} \rangle_{\text{main}}/M_\odot \text{ yr}^{-1})$	$3.21^{+0.64}_{-0.40}$	$3.54^{+0.19}_{-0.35}$	$3.07^{+1.22}_{-0.30}$
$\log(\tau_{\text{decl}} \text{ yr}^{-1})$	$7.1^{+1.3}_{-0.1}$	$7.2^{+0.2}_{-0.2}$	$7.0^{+0.6}_{-0.0}$

**Note.**

<sup>a</sup> The uncertainties on the quantities derived from the SED modeling represent the 90% confidence interval computed following Avni (1976; and Section 5 of this work). These values are computed by adopting the composite SFH from Schreiber et al. (2018a; Equations (1) and (2) of this work).



**Figure 5.** Location with respect to the main sequence of galaxies. The red stars and diamond indicate the location of our  $z = 3.77\text{--}4.01$  quiescent galaxies in the  $M_*$ -SFR plane. The golden symbols mark the  $\langle \text{SFR}_{\text{main}} \rangle$  of their progenitors at  $z_{\text{form}}$ , fixing the mass to 50% of the total final stellar mass. The magenta filled square and brown open square indicate the quiescent galaxies at  $z = 2.99$  and  $3.493$  reported in Gobat et al. (2012) and Forrest et al. (2019), respectively. The filled and open orange circles mark spectroscopically confirmed and unconfirmed quiescent sources at  $3 \lesssim z \lesssim 4$  from Schreiber et al. (2018a). The yellow circles show the  $\langle \text{SFR}_{\text{main}} \rangle$  of the progenitors at  $z_{\text{form}}$  from S18b. The location of the main sequence at  $z = 3\text{--}4$  as parameterized in Schreiber et al. (2015) is shown by the blue shaded area. The thin blue line marks the position of sources  $10\times$  below the main sequence at  $z = 3\text{--}4$ . The golden solid lines indicate the position of the main sequence and  $4\times$  above it at  $z = 5 (\sim z_{\text{form}})$ . The dashed lines show the extrapolation of the main sequence to masses larger than  $\log(M_*/M_\odot) = 11.5$ . The blue diamonds mark SMGs at  $z > 4$  from da Cunha et al. (2015), blue crosses from Michałowski et al. (2017), and blue open squares from Miettinen et al. (2017).

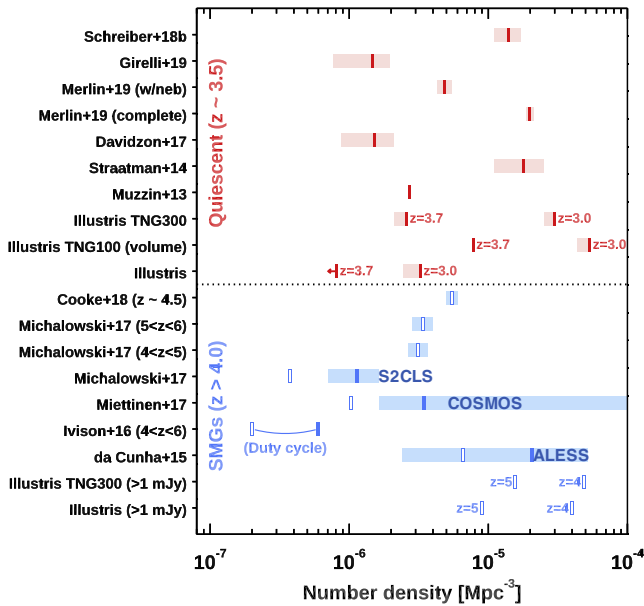
we derive  $\text{SFR}(\text{H}\beta) < 2.3 M_\odot \text{ yr}^{-1}$  ( $< 3.4 M_\odot \text{ yr}^{-1}$  for the 90% conservative upper limit), placing the galaxy 1.8 ( $> 1.6$ ) dex below the main sequence at its redshift. For COS-466654, we estimate  $\text{SFR}(\text{H}\beta) < 8.3$  and  $\text{SFR}([\text{O II}]) < 2.6 M_\odot \text{ yr}^{-1}$  (90% confidence interval), corresponding to  $> 1.2$  and  $> 1.7$  dex below the main sequence. The  $\text{H}\beta$  and  $[\text{O II}]$  emission lines are not covered for SXDS-27434, but a similar attempt for  $\text{H}\gamma$  returns an upper limit of  $\text{SFR} < 8 M_\odot \text{ yr}^{-1}$  (Tanaka et al. 2019).

We finally looked for possible far-infrared/submillimeter emission associated with the three galaxies. The galaxies SXDS-10017 and SXDS-27434 are not detected in the *Spitzer*/MIPS  $24 \mu\text{m}$  (SpUDS; PI: James Dunlop); *Herschel*/SPIRE 250, 350, and  $500 \mu\text{m}$  bands from HerMES (Oliver et al. 2012); or SCUBA-2  $870 \mu\text{m}$  maps from the SCUBA-2 Cosmology Legacy Survey (S2CLS; Geach et al. 2017) or at the Very Large Array 1.4 GHz (Simpson et al. 2006). Similarly, COS-466654 is not detected in any of the mid-infrared ( $24 \mu\text{m}$ ) to radio (1.4 GHz) bands collected in the “super-deblended” catalog of the COSMOS field by Jin et al. (2018), resulting in a combined infrared S/N of  $S/N_{\text{IR}} = 1.8$ . At the current sensitivity and spatial resolution limits, this further confirms the quiescence of the two galaxies and excludes the presence of bright dusty star-forming companions in their immediate proximity, at odds with at least one previously reported case (Glazebrook et al. 2017; Simpson et al. 2017; Schreiber et al. 2018b).

## 7. Progenitors

After putting the existence and properties of our targets on solid ground, we now explore their past history. The number of quiescent galaxies at  $3 < z < 4$ , their epoch of formation and its duration, the quenching time, and the average SFR can be used to look for plausible progenitors. The short formation intervals  $\Delta t$  and the large  $\langle \text{SFR} \rangle_{\text{main}}$  are reminiscent of the depletion timescales and observed SFRs of DSFGs at high redshift (Daddi et al. 2010; Tacconi et al. 2010; Casey et al. 2014). To quantify





**Figure 6.** Number densities of high-redshift quiescent galaxies and SMGs. The red ticks mark the observed number densities of massive ( $M_* \gtrsim 4 \times 10^{10} M_\odot$ ) quiescent galaxies at  $3 \lesssim z \lesssim 4$  from the works and projects reported on the Y-axis. The rose areas indicate the uncertainties, when available. The blue open ticks mark the number densities of variously selected  $z > 4$  SMGs available in the literature. The blue filled ticks include a correction of the duty cycle of SMGs ( $\rho_{\text{corr}} = \rho \times t_{\text{obs}}/t_{\text{burst}}$ , where  $t_{\text{burst}} = 200$  Myr) following Ivison et al. (2016). The blue shaded areas indicate the uncertainties reported in the original works (Cooke et al. 2018; Michałowski et al. 2017 ( $4 < z < 5$ ); Michałowski +17 ( $5 < z < 6$ )) or recomputed in this work based on the  $z_{\text{phot}}$  uncertainties of the samples (Michałowski et al. 2017; Miettinen et al. 2017; da Cunha et al. 2015). The calculations for the *Illustris* and *Illustris* TNG (300-1) simulations were performed in the snapshots corresponding to the labeled redshifts. The comoving number densities of SMGs in both suites are based on the catalogs by C. Hayward et al. (2020, in preparation) down to a threshold of 1 mJy, similar to the ALESS limiting flux and  $\sim 3$ – $4\times$  higher than for S2CLS/COSMOS. No duty cycle correction is applied to the simulations. For the simulated quiescent galaxies, the rose area indicates the variation between estimates including only “quenched” ( $\text{sSFR} < 10^{-11} \text{ yr}^{-1}$ ) or also “quenching/post-starburst” objects ( $\text{sSFR} < 10^{-10} \text{ yr}^{-1}$ ; red ticks). The values for the *Illustris* TNG (100-1) have been corrected by  $1.2\times$  for the box volume (Section 8.2).

such a possible connection, we first computed the abundance of these two populations in terms of their comoving number densities. Figure 6 shows a compilation of values from recent works in the literature.

### 7.1. Number Densities of Quiescent Galaxies at $3 < z < 4$

For the quiescent samples, we report the number densities for UVJ-selected galaxies at  $M_* > 4 \times 10^{10} M_\odot$  and  $3 < z < 4$  from UltraVISTA (Muzzin et al. 2013) and ZFOURGE as in Straatman et al. (2014; see Spitler et al. 2014 for an earlier estimate). The latter was updated by S18b after their spectroscopic follow-up and corrected for contamination of low-redshift interlopers. Notice that S18b computed the number density to a 25% smaller threshold in stellar mass, compatible with their completeness limit ( $M_* = 3 \times 10^{10} M_\odot$ ). We then included the calculation based on the latest version of the COSMOS catalog as in Davidzon et al. (2017), who presented an extensive comparison with previous works (see references therein); the extended sample at  $3 < z < 4$  from Girelli et al. (2019); and the combined CANDELS fields by Merlin et al. (2019), expanding their previous work on

GOODS-South (Merlin et al. 2018). The authors corrected their estimates for incompleteness and computed the impact of the emission lines on the photometry (see Santini et al. 2019 for the recent confirmation of the quiescence of a subsample of these galaxies). For consistency, we report the estimates with similar mass and redshift cuts as in Straatman et al. (2014) and S18b. We recomputed the number densities and their uncertainties whenever necessary to match the criteria above, i.e., by integrating the stellar mass functions in Muzzin et al. (2013) and Davidzon et al. (2017). For the other works, we reported the original values.

Figure 6 shows a wide range of measurements for the quiescent population, with variations up to a factor of  $10\times$ . The values derived integrating the stellar mass functions over large-area surveys (Muzzin et al. 2013; Davidzon et al. 2017) are systematically lower than those computed by counting red galaxies in smaller fields (Straatman et al. 2014; Schreiber et al. 2018a; Merlin et al. 2019). We estimated the impact of the cosmic variance on COSMOS ( $1.8 \text{ deg}^2$ ) and ZFOURGE-like ( $0.1 \text{ deg}^2$ ) areas as in Davidzon et al. (2017), both by adopting the analytic approach by Moster et al. (2011) and by comparing with the mock galaxy catalogs from 24 realizations in the Millennium simulations (Springel et al. 2005; Boylan-Kolchin et al. 2009; Henriques et al. 2015), the latter including Poisson noise. For masses  $\gtrsim 3 \times 10^{10} M_\odot$ , we estimate a  $\sigma_{\text{cv}} = 9.3\%$  (9.8%) uncertainty due to cosmic variance from the analytical (heuristic, including Poissonian noise) approach in the COSMOS field at  $3 < z < 4$ . For the area covered by the ZFOURGE survey for the same redshift range and  $M_*$  threshold, we compute a  $\sigma_{\text{cv}} = 15.3\%$  (a factor of  $1.5\times$ ) uncertainty from the analytical (heuristic, including Poissonian noise) approach. Besides the variations induced by cosmic and sample variance, the difference among the various estimates of number densities is affected by the classification method, based on colors and/or sSFR with different thresholds (Davidzon et al. 2017; Merlin et al. 2019); the contamination of lower-redshift interlopers (see the discussion in S18b) and active galactic nuclei (AGNs; Davidzon et al. 2017); the slightly different lower stellar mass integration limits and redshift intervals considered; and the depth of the observations.

### 7.2. A Conservative Lower Limit from Spectroscopy

For the sake of completeness, we finally derived a conservative estimate of the comoving number densities of quiescent objects at  $3.219 < z_{\text{spec}} < 4.012$  purely based on spectroscopically confirmed galaxies by combining our sample and the objects in S18b. Considering the eight robust detections over an area of  $2.6 \text{ deg}^2$  covered by the COSMOS, UDS, and ZFOURGE fields,<sup>22</sup> we obtain  $n = 3.4 \times 10^{-7} \text{ Mpc}^{-3}$  at face value and a  $3\sigma$  lower limit of  $> 8.8 \times 10^{-8} \text{ Mpc}^{-3}$  assuming a Poissonian distribution (Table 2 in Gehrels 1986). Including five extra uncertain redshift estimates, we derive  $n = 5.5 \times 10^{-7}$  and  $> 2.0 \times 10^{-7} \text{ Mpc}^{-3}$  at face value and as a  $3\sigma$  lower limit, respectively. These estimates represent the most conservative limits on the number of quiescent objects at these redshifts, being purposely uncorrected for any uncertain completeness effect. With the steady growth of the number of confirmed quiescent objects, the constraining power of these lower limits is destined to rapidly increase in the immediate future.

<sup>22</sup> Two of the ZFOURGE fields are included in COSMOS and UDS. Therefore, we did not account for their area in this calculation.

### 7.3. Number Densities of SMGs at $z > 4$

We collected recent results from large surveys of SMGs with detailed modeling of the optical and near-infrared counterparts, a necessary step to derive at least a photometric estimate of the redshift and stellar masses. The definition of an SMG is purely observational, and it hides a certain degree of diversity of the underlying population. However, it overlaps with the physical definition of DSFGs at the highest redshifts explored so far, therefore capturing suitable candidates to be the progenitors of massive quiescent galaxies. In this work, we compiled results from recent surveys on large fields.

1. The (ALMA) LABOCA Extended *Chandra* Deep Field South Survey ((A)LESS; Weiß et al. 2009; Hodge et al. 2013; Simpson et al. 2014; da Cunha et al. 2015; Danielson et al. 2017): 99 securely detected sources down to an rms of  $0.4 \text{ mJy beam}^{-1}$  with the Atacama Large Millimeter/submillimeter Array (ALMA) Band 7, originally selected at  $870 \mu\text{m}$  down to rms =  $1.2 \text{ mJy}$  with the Large Apex BOlometer CAmera (LABOCA) on the APEX telescope in the Extended *Chandra* Deep Field South. The optical/near-infrared follow-up includes 19 bands with a  $3\sigma$  detection limit in  $K_s = 24.4 \text{ mag}$  (Simpson et al. 2014). Here we use the results of the full SED modeling with MAGPHYS (da Cunha et al. 2008) presented in da Cunha et al. (2015). Including the uncertainties on the  $z_{\text{phot}}$ ,  $17_{-15}^{+26}$  sources lie at  $z_{\text{phot}} \geq 4$  over a  $0.25 \text{ deg}^2$  area (Simpson et al. 2014). We further take into account the  $2\times$  underdensity of SMGs in the field (Weiß et al. 2009) for the number density calculation. Note that this area is formally correct only for fluxes covered by the original LESS survey. Below this limit, the number counts might not be fully representative of the overall population of fainter SMGs down to the ALESS detection limit but rather of dim sources in the vicinity of previously known bright galaxies (Karim et al. 2013). In this sense, the number density for the faintest sources is, in principle, biased (but see Simpson et al. 2014). However, in the absence of a blind survey with such wide coverage, we proceed with the calculation.
2. The S2CLS (Geach et al. 2017; Michałowski et al. 2017):  $\simeq 650$  sources detected at  $\geq 4\sigma$  at  $850 \mu\text{m}$  with SCUBA-2 at the James Clerk Maxwell Telescope with secure fluxes  $\gtrsim 4 \text{ mJy}$  in the COSMOS and UDS fields (area of  $2.17 \text{ deg}^2$ ) and  $1.1 \text{ mm}$  coverage from ASTE AzTEC (see Dudzevičiūtė et al. 2019 for a recent reimaging of the UDS field with ALMA). Roughly 70% of this sample has a mass estimate obtained by modeling the optical/near-infrared photometry with MAGPHYS, notably assuming a double-component SFH (Michałowski et al. 2014). The limiting  $3\sigma$  depth in the  $K_s$  band is  $24.0/24.9 \text{ mag}$  from the UltraVISTA (McCracken et al. 2012) DR3 deep and ultra-deep stripes in COSMOS and  $25.2 \text{ mag}$  in UDS. Considering only the objects above the completeness limit of  $\text{SFR} = 300 M_{\odot} \text{ yr}^{-1}$  and taking into account the uncertainties on  $z_{\text{phot}}$ ,  $91_{-77}^{+63}$  sources lie at  $z_{\text{phot}} \geq 4$ , but there are only  $16_{-7}^{+7}$  with a photometric redshift from the optical/near-infrared, the rest being determined from the far-infrared SED only.

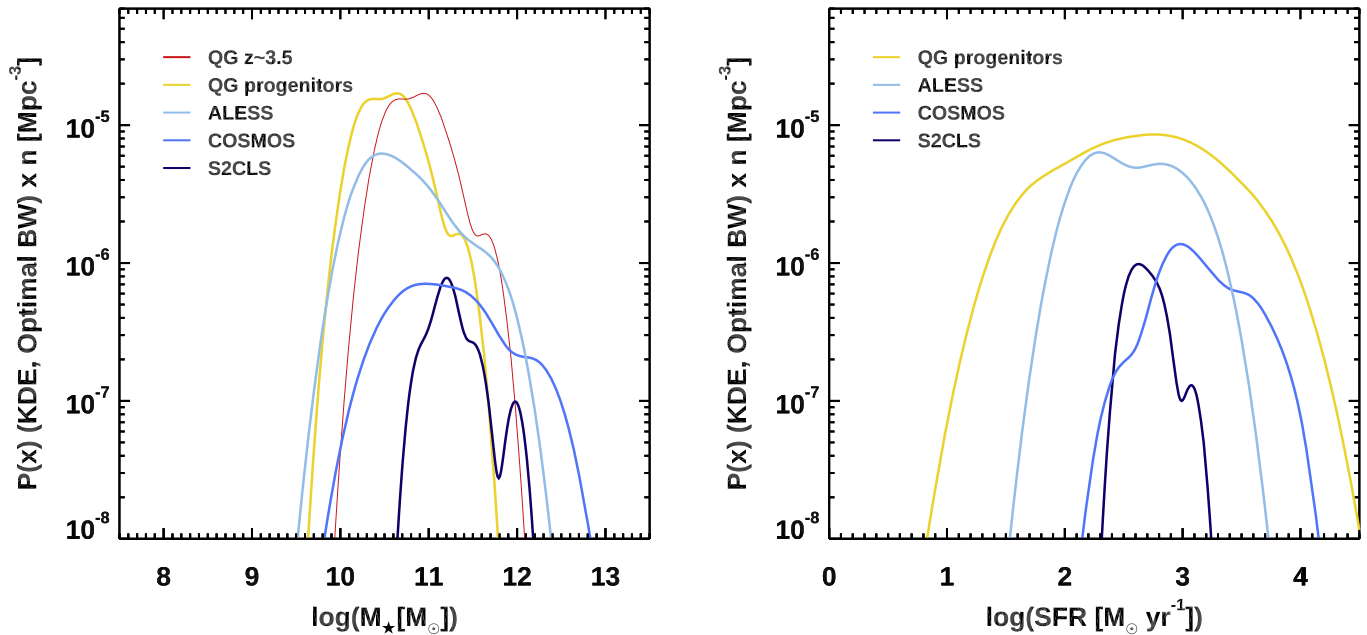
3. The ALMA follow-up of 124 SMGs in COSMOS, selected at  $1.1 \text{ mm}$  with ASTE AzTEC down to  $S_{1.1 \text{ mm}}^{\text{AzTEC}} = 3.5 \text{ mJy}$  at  $>4\sigma$  and with optical/near-infrared counterparts (Brisbin et al. 2017; Miettinen et al. 2017). The ALMA follow-up at  $1.3 \text{ mm}$  reaches an rms of  $\sim 0.1 \text{ mJy beam}^{-1}$ . The SED has been modeled with MAGPHYS to obtain stellar masses and SFRs (Miettinen et al. 2017) and with HYPERZ (Bolzonella et al. 2000) to estimate the photometric redshifts (Brisbin et al. 2017). Five sources are spectroscopically confirmed above  $z > 4$ , consistent with their photometric redshifts (Smolčić et al. 2015; Gómez-Guijarro et al. 2018). Over a covered area of  $0.72 \text{ deg}^2$ , we count  $17_{-9}^{+16}$  sources with  $z_{\text{phot}} \geq 4$ .

We then computed the number densities of SMGs above a redshift threshold  $z = z_{\text{thresh}} = 4$  as described in Ivison et al. (2016),

$$n = \frac{N}{V_{\text{com}}} C_{\text{duty}} [\text{Mpc}^{-3}], \quad (3)$$

where  $N$  is the number of galaxies detected at  $z > z_{\text{thresh}}$  and  $V_{\text{com}}$  is the comoving volume spanned by the observations detecting the SMG population. Here  $C_{\text{duty}} = t_{\text{obs}}/t_{\text{burst}}$  corrects for the typically short duty cycle of SMGs ( $t_{\text{burst}}$ ), which shine in the far-infrared/submillimeter regime for only a fraction of their whole existence or, in this case, the cosmic time probed by the observations  $t_{\text{obs}}$  (e.g., Toft et al. 2014). In Figure 6, we show the number densities both uncorrected and introducing  $C_{\text{duty}}(t_{\text{burst}} = 200 \text{ Myr})$  as a reference (see Section 7.4.2 below). We included the uncertainties on the photometric redshift estimates by counting galaxies with  $z_{\text{up}}$  and  $z_{\text{low}} > z_{\text{thresh}}$ , where  $z_{\text{up}}$  and  $z_{\text{low}}$  are the upper and lower boundaries of the redshift uncertainties as reported in the original works. This uncertainty dominates the error budget. We did not include further corrections for completeness of the surveys beyond what was reported in the literature (see Miettinen et al. 2017 for a caveat for the COSMOS sources). For reference, we also reported two recent results from the literature, namely number densities for ultrared galaxies selected from the *Herschel* Astrophysical Terahertz Large Area Survey (*H-ATLAS*; Ivison et al. 2016) and a subsample of spectroscopically confirmed SMGs at  $z \sim 4.5$  from the S2CLS (Cooke et al. 2018), along with the number densities originally reported in Michałowski et al. (2017).

Figure 6 shows that large variations are present in the observed number densities of SMGs. This is likely due to the variety of selection criteria and the depths, spatial resolution, and completeness of the various surveys, plus the intrinsic diversity of the SMG population. The choice of the submillimeter band for the initial selection corresponds to the sampling of a specific portion of the far-infrared SED. A selection based on *Herschel* bands or at  $\sim 870 \mu\text{m}$  ( $170 \mu\text{m}$  rest frame at  $z = 4$ ) maps the SED closer to the peak of the dust emission, being sensitive to the temperature and total IR luminosity ( $\propto \text{SFR}$ ). Opting for a cut at  $1.1 \text{ mm}$  ( $220 \mu\text{m}$  rest frame at the same redshift) results in the sampling of the optically thin dust emission in the Rayleigh–Jeans tail of the



**Figure 7.** Stellar masses and SFRs of candidate progenitors of high-redshift quiescent galaxies. Shown are the KDEs of the density probability distributions of stellar masses (left panel) and SFRs (right panel) for the progenitors at  $z_{\text{form}}$  of massive quiescent galaxies at  $3 < z < 4$  (golden line; this work; S18b) and SMGs at  $z > 4$  from ALESS, COSMOS, and S2CLS (blue lines; da Cunha et al. 2015; Miettinen et al. 2017; Michałowski et al. 2017). For reference, we show the stellar mass distribution of the quiescent galaxies (red line). The areas under the curves are normalized to the number density of each population (Figure 6).

SED, thus being mildly sensitive to the effective temperature and privileging large dust masses over large SFRs, even if the two quantities are ultimately correlated. The depth and spatial resolution of the observations likely have an even stronger impact than the selection criterion: shallow limits from single-dish observations allow us to capture only the rarest and strongest starbursting systems, likely affected by source blending. Moreover, the necessity of identifying a counterpart in the optical/near-infrared to estimate stellar masses and redshifts biases the results against the dustiest and highest-redshift objects. This is evident for the S2CLS sample, for which only  $\sim 20\%$  of the galaxies with  $z_{\text{phot}} > 4$  have an optical/near-infrared redshift estimate, the rest being determined from the far-infrared SED only (Michałowski et al. 2017). However, adopting the best available photometric redshifts allows us to improve previous estimates based on the simple assumption of a fixed fraction of SMGs above  $z = 4$  (e.g., Straatman et al. 2014; Ivison et al. 2016).

#### 7.4. Standard Evolutionary Connection

We now explore the possible evolutionary connection between SMGs and quiescent galaxies (see previous results in the literature at similar or lower redshift; Cimatti et al. 2008; Simpson et al. 2014; Straatman et al. 2014; Toft et al. 2014; Miettinen et al. 2017; Gómez-Guijarro et al. 2018). Here we test the common and extreme assumption that all SMGs at  $z > 4$  turn into quiescent objects at  $3 < z < 4$  by matching their number densities (e.g., Toft et al. 2014). For the calculation, we adopt a final number density for quiescent galaxies at  $3 < z < 4$  of  $n = 1.4 \times 10^{-5} \text{ Mpc}^{-3}$  from S18b. We compute results for the samples of SMGs with optical/near-infrared counterparts from ALESS (da Cunha et al. 2015), S2CLS (Michałowski et al. 2017), and COSMOS (Miettinen et al. 2017).

##### 7.4.1. Stellar Masses and SFR Distributions

First, we can compare the distributions of the stellar masses and SFRs of SMGs and the progenitors of quiescent galaxies. In Figure 7, we present the (Gaussian) kernel density estimation (KDE) of the observed distributions. We derived the expected properties of the quiescent galaxy progenitors from the SED modeling of SXDS-10017, SXDS-27434, and COS-466654 and the sample in S18b. We adopted  $\langle \text{SFR} \rangle_{\text{main}}$  as the average SFR during the main formation epoch and 50% of the final stellar mass as  $M_*$  for the progenitors (excluding mass loss and recycling; Section 5). Notice that the choice of the SFH does not impact the  $\langle \text{SFR} \rangle_{\text{main}}$  within the uncertainties (Figure 9). For the SMGs, we compiled the stellar masses from da Cunha et al. (2015), Michałowski et al. (2017), and Miettinen et al. (2017) and the SFRs they derived from the far-infrared luminosities. Given the best-fit values for  $\Delta t_{\text{form}}$ , the far-infrared luminosities and  $\langle \text{SFR} \rangle_{\text{main}}$  probe similar timescales for star formation. Not all of the SMGs have an optical/near-infrared counterpart, necessary to derive a stellar mass. Therefore, the results are biased toward the less dust-obscured objects, either at lower masses or more advanced stages of formation, when the dust shrouding the cocoons of star formation starts fading away.

By definition, the integral of each probability density curve is equal to  $\int_{-\infty}^{+\infty} P(x) dx = 1$ . In order to compare the various populations, we normalized the KDE to the number density of each sample ( $\int_{-\infty}^{+\infty} n P(x) dx = n [\text{Mpc}^{-3}]$ ). Therefore, the area under each curve in Figure 7 is exactly equal to the number density of each sample, with  $C_{\text{duty}} = 1$  for the SMGs (no correction for duty cycle).

The left sides of the distributions are in first approximation related to the depth of the observations. The  $K_s$ -band limits are similar for all of the surveys, and, indeed, the stellar mass distribution for the quiescent galaxies and their progenitors, ALESS (da Cunha et al. 2015) and COSMOS (Miettinen et al. 2017), show



a consistent lower mass limit. The only exception is the S2CLS (Michałowski et al. 2017), which—even sharing a similar  $K_s$ -band depth with all of the other surveys—results in a nonnegligible probability of low stellar masses. However, this tail of low-mass galaxies disappears when considering only  $z_{\text{phot}}$  estimates from the optical/near-infrared. Its existence may therefore be a spurious effect due to wrong redshift estimates from the far-infrared SED, which we therefore discarded. For clarity, we show the stellar mass distribution without this low-mass tail in Figure 7. For what concerns the SFRs, the effect of the cut imposed by the sensitivity limits of the SMG surveys is evident, resulting in a lower limit on the  $L_{\text{IR}} \sim \text{SFR}$ . On the other hand, the lower limit of the distribution of the progenitors mainly depends on the modeling of the SED and the best SFH.

The peaks of the distributions of the stellar masses are consistent for ALESS and the progenitors of the quiescent galaxies, while the modeling of the S2CLS and COSMOS sources results in larger stellar masses than the quiescent (and star-forming) population at lower redshift, as already reported in Miettinen et al. (2017) and Michałowski et al. (2017). Under the initial assumption that all SMGs at  $z > 4$  become quiescent at  $3 < z < 4$ , the massive SMGs cannot be considered only a tail of the overall distribution (Appendix B). Notice that here we do not include the stellar mass that the ongoing episode of star formation will add to SMGs, which would further increase the discrepancy with the quiescent population. On the other hand, all of the distributions of SFRs roughly peak at the same value, even if the KDE for the progenitors of quiescent galaxies allows for significant probability densities at low SFRs, not being hampered by an observational limit. The extension of the distribution of the progenitors at low SFRs points toward the necessity of including less extreme systems to explain the existence of red, quiescent galaxies at  $3 < z < 4$ . Independent structural and dynamical arguments suggest a similar conclusion on the origin of quiescent galaxies at  $z \sim 2$ , whose precursors might be compact blue SFGs with properties typical of objects on the main sequence (Barro et al. 2013, 2017; Popping et al. 2017; Gómez-Guijarro et al. 2019) or a dust-obscured population so far overlooked (Wang et al. 2019; Williams et al. 2019). This shows the relevance of pushing our search for DSFGs at lower SFRs and not limiting it to classical starbursting SMGs. We come back to this point later, providing supporting evidence from cosmological simulations (Section 8).

#### 7.4.2. The Duty Cycle of SMGs

As mentioned in Section 7.3, only a fraction of the global star-forming population shines as SMGs at a specific time, and this fraction depends on the submillimeter flux threshold to define an SMG. With the usual assumption that all such systems at  $z > 4$  turn into the quiescent population at  $3 < z < 4$ , we can derive the duty cycle correction  $\mathcal{C}_{\text{duty}}$  necessary to make the number densities of the two populations exactly equal (e.g., Toft et al. 2014). Here we do this by matching the number densities (i.e., the area under the curves in Figure 7) and the stellar mass or SFR distributions (i.e., the shape of the curves) of the quiescent progenitors and the samples of SMGs (see Appendix B for details). We derive  $t_{\text{burst}} \sim 200\text{--}400$  Myr for the sources in ALESS. For COSMOS and S2CLS, we find similar  $t_{\text{burst}} \sim 100$  Myr when matching the SFR distributions but substantially shorter values down to  $t_{\text{burst}} \sim 10$  Myr when matching the  $M_*$  KDEs. The values for

$t_{\text{burst}}$  are consistent with the typical depletion timescales of strongly star-forming galaxies at high redshift ( $t_{\text{duty}} \sim 100\text{--}200$  Myr; Daddi et al. 2010; Tacconi et al. 2010; Casey et al. 2014) and with the formation timescales via the archeological approach (e.g., Thomas et al. 2005; Renzini 2006; Choi et al. 2014; Onodera et al. 2015). However, at least for the self-consistent values of the ALESS sources, we find duty cycle corrections  $\mathcal{C}_{\text{duty}} \sim 5 \times$  smaller than estimated by Toft et al. (2014) for the range  $2 < z < 3$ . Notice that this is partly due to the longer  $t_{\text{burst}}$  we estimate, and it is partly a natural consequence of observing higher-redshift sources, as the time span  $4 < z < 6$  is half of  $2 < z < 3$ .

These calculations depend on a set of assumptions that we specified at each step, and they are affected by several sources of uncertainties. An extended discussion is reported in Appendix B. We stress once again that the evolutionary connection we tested here relies on the extreme assumption that all SMGs at  $z > 4$  become quiescent at  $3 < z < 4$ , driving toward the introduction of a duty cycle correction. Looser conditions (e.g., only  $X\%$  of the SMGs turn into quiescent galaxies) result to first order in longer duty cycles by a similar amount. Similar rescaling factors apply when considering a value different than our reference estimate of the comoving number density of quiescent galaxies.

## 8. Can Cosmological Simulations Capture the Formation of Quiescent Galaxies at $z \sim 4$ ?

### 8.1. Realistic Comoving Number Densities of Quiescent Galaxies

Previous attempts at reproducing the population of quenched galaxies at  $z > 3$  with cosmological simulations and semi-analytical models fell short in producing enough systems by up to an order of magnitude in the majority of cases (S18b; Cecchi et al. 2019). Here we explore the content of the recent *Illustris* TNG cosmological simulation public release (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018, 2019; Pillepich et al. 2018; Springel et al. 2018) in comparison with the previous *Illustris* project (Vogelsberger et al. 2014a, 2014b; Genel et al. 2014; Nelson et al. 2015; see Merlin et al. 2019 for a similar attempt). As for the observations presented above, we structured our search in two steps. We first looked for the quiescent galaxies in the largest boxes available, TNG 300 (205/ $h$  comoving Mpc) and TNG 100 (75/ $h$  comoving Mpc), so as to build enough statistics of these rare systems. We subsequently studied the progenitors of the quenched systems possibly shining as SMGs in these boxes (C. Hayward et al. 2020, in preparation; Hayward et al. 2013). *Illustris* TNG 100 has a similar box size to the old *Illustris*-1, so we drew comparisons among these two to test the performances of the new simulations.

To mimic our observational selection, we identified quiescent galaxies in the  $z = 3.7$  snapshot based on the sSFR within twice the stellar half-mass radius (INRAD quantities in the catalogs). The SFR is averaged over 10 Myr, but we checked for consistent quiescence in the descendant subhalos down to  $z = 2$  in order to exclude contamination of temporary low-activity galaxies. We selected both “quenched” and “quenching” (or post-starburst) galaxies fixing a threshold of  $\text{sSFR} \leq 10^{-11}$  and  $10^{-11} < \text{sSFR} \leq 10^{-10} \text{ yr}^{-1}$ , respectively, and imposing a minimum  $M_* = 4 \times 10^{10} M_\odot$  similar to the mass completeness limits in the observations (Section 7.1).

**Table 2**  
Number and Number Densities of Quiescent Galaxies and Their Progenitors in the *Illustris* and *Illustris* TNG Cosmological Simulations

	<i>Illustris</i> -1	<i>Illustris</i> TNG 100-1	<i>Illustris</i> TNG 100-2	<i>Illustris</i> TNG 300-1
Box size [cMpc]	75/ $h$	75/ $h$	75/ $h$	205/ $h$
Dark matter mass resolution [ $10^6 M_\odot$ ]	6.3	7.5	59.7	59
Quiescent Galaxies <sup>a</sup>				
$N_{<-11}, N_{[-11,-10]} (z = 3)$	1, 3	45, 10	25, 6	631, 123
$n_{<-11}, n_{[-11,-10]} (z = 3) [10^{-6} \text{ Mpc}^{-3}]$	0.8, 2.4	36.6, 8.1 <sup>c</sup>	20.3, 4.9	25.1, 4.9
$N_{<-11}, N_{[-11,-10]} (z = 3.7)$	0, 1	8, 0	0, 1	53, 12
$n_{<-11}, n_{[-11,-10]} (z = 3.7) [10^{-6} \text{ Mpc}^{-3}]$	–, 0.8	6.5 <sup>c</sup> , ...	..., 0.8	2.1, 0.5
SMGs <sup>b</sup>				
$N_{>1 \text{ mJy}}, N_{>3.5 \text{ mJy}} (z = 4)$	49, 0	51, 2	59, 2	1224, 48
$n_{>1 \text{ mJy}}, n_{>3.5 \text{ mJy}} (z = 4) [10^{-6} \text{ Mpc}^{-3}]$	39.8, –	41.5, 1.6	48.0, 1.6	48.7, 1.9
$N_{>1 \text{ mJy}}, N_{>3.5 \text{ mJy}} (z = 5)$	11, 1	15, 0	17, 1	393, 11
$n_{>1 \text{ mJy}}, n_{>3.5 \text{ mJy}} (z = 5) [10^{-6} \text{ Mpc}^{-3}]$	8.9, 0.8	12.2, –	13.8, 0.8	15.6, 0.4

#### Notes.

<sup>a</sup> Number of galaxies in the  $z = 3$  and  $3.7$  snapshots with  $M_\star \geq 4 \times 10^{10} M_\odot$  and  $\text{sSFR} \leq 10^{-11} \text{ yr}^{-1}$  ( $N_{<-11}$ , “quenched”) or  $10^{-11} \text{ yr}^{-1} < \text{sSFR} \leq 10^{-10} \text{ yr}^{-1}$  ( $N_{[-11,-10]}$ , “quenching”). All quantities are computed within twice the stellar half-mass radius (INRAD in the *Illustris* data releases). The comoving number densities  $n$  of quenched and quenching galaxies are obtained by dividing  $N_{<-11}, N_{[-11,-10]}$  by the comoving volume of the box (size<sup>3</sup>, adopting  $h = 0.7$ ).

<sup>b</sup> Number of galaxies in the  $z = 4$  and  $5$  snapshots with  $S_{850}$  flux  $> 1$  and  $> 3.5$  mJy within 25 kpc, computed following Hayward et al. (2013) and C. Hayward et al. (2020, in preparation).

<sup>c</sup> These values are not corrected by the factor  $1.2\times$  for the volume of the box (Section 8.2).

Note that such a selection in sSFR is robust against variations of the timescale over which the star formation is averaged in simulations in the range  $\sim 10$ – $200$  Myr and the measurement of quantities in different apertures (Sparre et al. 2015; Davidzon et al. 2018). Moreover, the quiescent fractions are similar when separating galaxies based on colors or distance from the main sequence (Donnari et al. 2019).

We summarize the results of the search in Table 2 and Figure 6. The comoving number densities  $n$  of quenched galaxies in the *Illustris* TNG boxes are consistent only with the lowest observational estimates at  $z = 3.7$ , while the old *Illustris* run does not contain enough of these objects, as previously noted (Wellons et al. 2015; S18b) and in line with similar previous attempts in smaller boxes (e.g., MUFASA; Davé et al. 2016). On the contrary, we retrieve a numerous enough population of quiescent galaxies in the  $z = 3$  snapshot (i.e., the lower limit of the redshift range explored here with spectroscopy) of the *Illustris* TNG simulations, while the old *Illustris* still fails at reproducing the observed number densities. This is likely due to the new feedback scheme implemented in the TNG simulations.

### 8.2. The Impact of the Mass Resolution

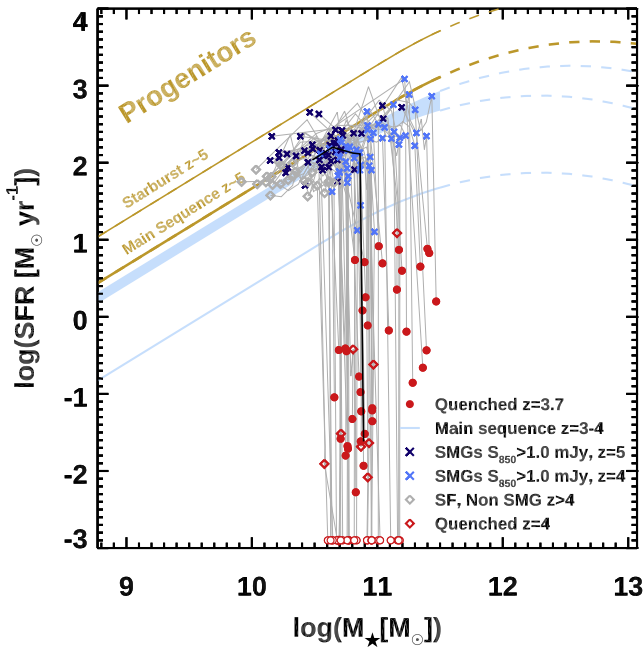
Interestingly, the largest TNG 300 box contains  $2.5\times$  and  $1.5\times$  less quenched/quenching galaxies per unit comoving volume than TNG 100 at  $z = 3.7$  and  $3$ , respectively (Table 2). We checked for the effect of the  $\sim 8\times$  lower mass resolution in TNG 300 than in TNG 100, which might not result in full convergence. As discussed in Pillepich et al. (2018), the lower resolution of TNG 300(-1) translates into lower stellar masses and SFR than TNG 100(-1), which might bias our number densities. Therefore, we compared the number of galaxies above our mass threshold ( $M_\star \geq 4 \times 10^{10} M_\odot$ ) in TNG 100-1 and TNG 100-2, the latter having a resolution similar to TNG 300-1. We retrieve  $\sim 25\%$  less galaxies in TNG 100-2 than in TNG 100-1; this fraction increases when

selecting quenched/quenching galaxies, so we find only one (31) objects with  $\text{sSFR} \leq 10^{-10} \text{ yr}^{-1}$  at  $z = 3.7$  ( $z = 3$ ) in TNG 100-2, a factor of  $8\times$  ( $1.7\times$ ) less than TNG 100-1. Therefore, while at  $z = 3.7$ , the low number statistics and cosmic variance likely dominate the difference between TNG 300-1 and TNG 100-1, the mass resolution explains the discrepancy at  $z = 3$  for our selection.

From the comparison between the number densities of quenched/quenching objects in TNG 300-1 and TNG 100-2, we derive a simple factor to correct the comoving number densities from the smaller box of the TNG 100-1 run while retaining the advantages of the high resolution. We compute such a correction as the ratio of the number densities  $n_{\text{TNG300}}/n_{\text{TNG100(2)}} = 1.2\times$  at  $z = 3$ , where we have enough statistics, and we assume that this factor also applies at the previous snapshot at  $z = 3.7$ . Therefore, our fiducial number densities of quenched/quenching objects for the TNG 100-1 box are  $5.4 \times 10^{-5}$  and  $7.8 \times 10^{-6} \text{ Mpc}^{-3}$  at  $z = 3$ – $3.7$ , respectively.

### 8.3. Simulated SMGs and Their Properties

We then traced the evolution of the main subhalo progenitor of each quiescent galaxy back to the main formation epoch derived for the observed targets ( $z \sim 5$ ; Section 5). In Figure 8, we show the evolutionary tracks in the  $M_\star$ –SFR plane for TNG 300, but similar conclusions hold for TNG 100. In order to directly compare with the operational definition of SMGs widely spread among observers, we also show the modeling of the  $850 \mu\text{m}$  emission of  $z = 4$ – $5$  galaxies by C. Hayward et al. (2020, in preparation). The authors computed mock  $850 \mu\text{m}$  fluxes for all sources with  $M_\star > 4 \times 10^9 M_\odot$  based on a relation derived from full radiative transfer calculations on idealized disks and mergers at a spatial resolution inaccessible for cosmological simulations (Hayward et al. 2011, 2013). The calculation is performed by integrating the interstellar medium and SFR within 25 kpc (and not within twice the half-mass



**Figure 8.** Stellar mass and SFR plane for the *Illustris* TNG 300 simulation. The red circles indicate the location of quiescent galaxies ( $sSFR \leq 10^{-10} \text{ yr}^{-1}$  within twice the half-mass radius) in the  $z = 3.7$  snapshot (the open symbols represent upper limits fixed at  $SFR = 10^{-3} M_{\odot} \text{ yr}^{-1}$  for display purpose). The location of the main sequence at  $z = 3-4$  as parameterized in Schreiber et al. (2015) is shown by the blue shaded area. The blue solid line marks the position of sources  $10\times$  below the main sequence at  $z = 3-4$ . The golden solid lines indicate the position of the main sequence and  $4\times$  above it at  $z = 5$  ( $\sim z_{\text{form}}$ ). The dashed lines show the extrapolation of the main sequence to masses larger than  $\log(M_*/M_{\odot}) = 11.5$  in order to facilitate the comparison with Figure 5. Dark and light blue crosses mark SMGs with  $S_{850 \mu\text{m}} \geq 1 \text{ mJy}$  at  $z = 5$  and  $4$ , respectively, as computed by C. Hayward et al. (2020, in preparation). The gray and red diamonds show SFGs with  $S_{850 \mu\text{m}} \leq 1 \text{ mJy}$  at  $z > 4$  and quiescent objects at  $z = 4$ , respectively. The gray tracks show the evolution of  $z = 3.7$  quiescent objects back to  $z = 5$ , while the black solid line shows the median evolution in the same interval.

radius), excluding the particles in satellites, in order to compare with typically resolved SMGs (C. Hayward et al. 2020, in preparation). A significant difference between SFR(25 kpc) and SFR(INRAD) is the reason why a minority of objects display bright  $S_{850}$  fluxes while being formally quiescent in our selection (Figure 8). Notice that the  $850 \mu\text{m}$  emission depends not only on the SFR but also on the dust mass and its temperature: highly star-forming galaxies thus might not shine as SMGs due to a lower dust content and, thus, hotter dust temperature (e.g., Hayward et al. 2012; Safarzadeh et al. 2016). As a reference, here we adopt cuts at  $S_{850 \mu\text{m}} > 1$  and  $> 3.5 \text{ mJy}$ , comparable with the ALESS and COSMOS/S2CLS limiting sensitivities, respectively.

While differing in principle, the SMGs have similar  $M_*$  and SFR to the progenitors of the quenched galaxies at  $z = 3-3.7$ . However, they are consistent with being the tail of the most massive galaxies on the main sequence at their redshift, without reaching the extreme values of observed starbursts (Figure 5). The choice of the  $S_{850}$  flux thresholds has a strong impact on the stellar mass of the selected SMGs, while the redshift does not. Flux cuts at  $> 1$  and  $> 3.5 \text{ mJy}$  result in median  $\log(M_*/M_{\odot}) = 10.5$  and  $11.1$ , respectively, at  $z = 4$  and  $5$  in TNG 300. This naturally arises from the fact that, by definition, a higher  $S_{850}$  threshold selects for higher SFR and  $M_{\text{dust}}$  values, both of which correlate with  $M_*$ . Therefore, an  $S_{850}-M_*$  relation is in place at least for main-sequence galaxies, matching the observations (Hayward et al. 2013).

#### 8.4. Number Densities of Simulated SMGs

We further calculated the number densities for the SMGs with  $S_{850} > 1$  and  $> 3.5 \text{ mJy}$  in the snapshots at  $z = 4$  and  $5$  (Table 2 and Figure 6). To avoid overpopulating the latter, we show only the number densities for a threshold of  $S_{850} = 1 \text{ mJy}$ , similar to the ALESS limiting flux. At face value, the simulations slightly overshoot the empirical number densities uncorrected for the duty cycle, but we find an overall good agreement within the observational uncertainties. Notice that no correction is needed for the simulated number densities. Similar conclusions hold when comparing simulations with COSMOS/S2CLS at the corresponding depth. The comoving number densities for SMGs with  $> 1 \text{ mJy}$  are all consistent within 20% among the three suites of simulations, while we find a sizable sample of objects with  $> 3.5 \text{ mJy}$  only in TNG 300. We further checked for the dependence on the mass resolutions of the simulations as for the quenched galaxies (Section 8.2) by comparing TNG 100-1 and TNG 100-2, which have a similar resolution to TNG 300-1. We retrieve  $\sim 10\%-15\%$  less galaxies with  $S_{850}$  fluxes  $> 0.1-0.5 \text{ mJy}$  but  $\sim 15\%-30\%$  more SMGs in TNG 100-2 at  $z = 3-5$  than in TNG 100-1. The low number statistics dominate the comparison for higher  $S_{850}$  flux thresholds. These fractions should be considered as the typical uncertainties on the number densities we derived for TNG 300-1. We also attempted a calculation of a volume correction for the SMGs in TNG 100-1 by comparing TNG 300 and TNG 100-2 (Section 8.2) at  $z = 4-5$  down to our reference flux threshold of  $S_{850} = 1 \text{ mJy}$ . However, such a factor is negligible (2%–10%).

Moreover, C. Hayward et al. (2020, in preparation) found that the *Illustris* TNG systematically underpredicts the  $850 \mu\text{m}$  cumulative number counts (therefore, integrated over redshift) with respect to observations and the old *Illustris* run at a similar mass resolution, owing to the increased efficiency of AGN feedback reducing the number of massive, strongly star-forming systems. The difference between the two simulation sets is not evident at  $z \geq 4$ . This is likely due to the fact that SMGs at these redshifts have not yet reached the critical black hole mass to trigger efficient AGN feedback. Therefore, the discrepancy between *Illustris* and *Illustris* TNG reported by Hayward et al. is due to the later evolution, while systems at  $z > 4$  constitute only a minor fraction of the SMG population.

#### 8.5. Not All Progenitors Are SMGs

The comparison between the comoving number densities of quiescent galaxies at  $z = 3-3.7$  and the SMGs at  $z = 4-5$  (Table 2) shows that only a fraction of the SMGs at high redshift turn quiescent at  $z = 3-3.7$ , and this fraction depends on the submillimeter flux threshold. In TNG 300, 88%–90% of the SMGs with  $S_{850} > 3.5 \text{ mJy}$  at  $z = 4-5$  are quenched ( $sSFR < 10^{-10} \text{ yr}^{-1}$ ) at  $z = 3$  but only 20%–30% at  $z = 3.7$ . These fractions drop significantly when considering a lower flux threshold of  $S_{850} > 1 \text{ mJy}$ : 45%–60% of the SMGs above this limit at  $z = 4-5$  are quenched at  $z = 3$  (5%–15% at  $z = 3.7$ ). We do not attempt the comparison in TNG 100 and *Illustris*-1 due to the low or even absent statistics. The difference due to the flux threshold likely mirrors the mass selection mentioned above: larger  $S_{850}$  fluxes correspond to larger masses, therefore closer to the threshold to ignite efficient AGN feedback and rapid quenching. The shorter time interval between  $z = 4$  and  $3.7$  concurs with the drop of the



fractions of SMGs quenching between these two limits, compared with quenching occurring in the  $z = 3\text{--}5$  period. Finally, at the spatial and time resolution of *Illustris* TNG, lower  $S_{850}$  appears to be sustainable for longer timescales. Dedicated simulations at higher resolution, necessary to capture the stochasticity of the processes igniting the brightest SMGs, could test this result, but they are beyond the scope of this paper.

Inverting the order of the terms of comparison, we find that only a fraction of the progenitors of quenched galaxies at  $z = 3\text{--}3.7$  shine as SMGs at  $z = 4\text{--}5$ , and this fraction depends on the submillimeter flux threshold. As Figure 8 shows, 80% of the quenched galaxies at  $z = 3.7$  have  $S_{850} > 1$  mJy at  $z = 4$  (70% at  $z = 5$ ) in TNG 300, while 12% (0%) were already quenched at  $z = 4$  ( $z = 5$ ). The fraction of submillimeter bright progenitors drops when considering higher flux thresholds: 14% of the quiescent galaxies at  $z = 3.7$  have  $S_{850} > 3.5$  mJy at  $z = 4$  (5% at  $z = 5$ ).

This comparison suggests a more complex connection between the progenitors of high-redshift quiescent galaxies and SMGs than previously assumed (e.g., Section 7.4). While the majority of extremely bright SMGs at  $z = 4\text{--}5$  detectable by shallow surveys quench by  $z = 3$ , a substantial fraction of the highest-redshift quenched systems have less extreme progenitors, lying on the main sequence during their epoch of main stellar mass assembly and emitting submillimeter fluxes partially detectable only by the deepest surveys currently available.

## 9. Conclusions

In this work, we report the discovery and detailed analysis of three quiescent galaxies at 1.5 Gyr after the Big Bang. Dedicated spectroscopic follow-up in the optical/near-infrared allowed us to securely confirm the redshift of two of the sources ( $z = 3.775$ , 4.012) and put tentative constraints on the third one ( $z \approx 3.767$ ). Their quiescence is supported by the modeling of their SEDs, the absence of emission lines, and the nondetections in the far-infrared/submillimeter regimes. Given their large stellar masses of  $\sim 10^{11} M_{\odot}$ , these objects are located  $\gtrsim 1\text{--}2$  dex below the main sequence at their redshifts. The combined modeling of the SED and the spectra suggests that these galaxies went through a short phase ( $\sim 50$  Myr) of intense star formation ( $\sim 1200\text{--}3500 M_{\odot} \text{ yr}^{-1}$ ) peaking  $\sim 150\text{--}500$  Myr prior to the time of observation, followed by an abrupt decrease of their SFRs and the cessation of any relevant formation of new stars.

We then explored their connection with star-forming progenitors at higher redshifts, testing previous suggestions of a direct link with strongly starbursting SMGs. We compared the comoving number densities  $n$  of  $850 \mu\text{m}$  and/or  $1.1 \text{ mm}$ -selected SMGs at  $z > 4$  and quiescent galaxies at  $3 < z < 4$ , compiling recent results in the literature. In general, a large scatter affects the estimates of  $n$  for both populations, predominantly due to a combination of factors dominated by different classification schemes and selection criteria and the uncertainties on the redshift. We find the number densities of SMGs at  $z > 4$  from the deepest surveys ( $S_{850 \mu\text{m}} \gtrsim 1$  mJy) to be in broad agreement with the estimates for quiescent galaxies at  $3 < z < 4$ . Brighter SMGs ( $S_{850 \mu\text{m}} \gtrsim 3.5$  mJy) are  $6\text{--}20\times$  less numerous. Adopting the assumption of a univocal correspondence between SMGs and the progenitors of quiescent galaxies, we attempted to estimate the duty cycle correction necessary to match their comoving number densities.

The resulting duty cycle is of the order of the depletion timescale estimated for strongly star-forming galaxies at high redshift ( $t_{\text{burst}} \sim 100\text{--}400$  Myr), but this value drops significantly, to  $\sim 10$  Myr, for the brightest SMGs, questioning the underlying assumption. This is reinforced by the comparison of the stellar masses and SFRs of SMGs and quiescent galaxies. We find that the current SED modeling tends to overpredict  $M_{\star}$  for the brightest SMGs with respect to quiescent galaxies at lower redshift, while objects with lower submillimeter fluxes can better reproduce the distributions of  $M_{\star}$  and SFR expected for the progenitors of quiescent galaxies. This points toward the necessity of including less extreme systems to explain the existence of red quiescent galaxies at  $3 < z < 4$ .

Finally, we tested our assumptions on the evolutionary path of high-redshift quiescent galaxies by comparing our results with the recent *Illustris* TNG simulation. We retrieved comoving number densities of massive quiescent galaxies at  $z = 3$  that are in fair agreement with the broad range of observed values, surpassing the performance of the previous generation of cosmological simulations. This is especially due to the large box size necessary to study these rare systems and the improved feedback scheme. However, we report a growing inconsistency in the comoving number densities at increasing redshift, so that at  $z = 3.7$ , the simulated populations match only the lowest among the observational estimates. Moreover, we traced their progenitors at  $z = 4\text{--}5$  and their mock submillimeter fluxes, finding a population of SMGs as numerous as in the observations. We find that  $\sim 90\%$  of the SMGs with  $S_{850} > 3.5$  mJy at  $z = 4\text{--}5$  are quenched by  $z = 3$ , but this fraction drastically decreases to  $\sim 45\%\text{--}60\%$  for dimmer sources with  $S_{850} > 1$  mJy. In other words, we showed that not all  $z = 3\text{--}3.7$  quiescent galaxies have an SMG progenitor, and, similarly, that not all SMGs at  $z = 4\text{--}5$  quench by  $z \sim 3\text{--}3.7$ ; the fractions mainly depend on the submillimeter flux cut to select the DSFGs. Moreover, as suggested by the observations mentioned above, the simulations indicate that the progenitors of massive quiescent galaxies at  $z > 3$  are not necessarily prototypical extreme starbursting systems but rather more normal star-forming galaxies at the most massive end of the main sequence. This highlights the importance of obtaining deep submillimeter observations of high-redshift galaxies, abandoning the original definition of “SMGs” to focus on more physically meaningful categories.

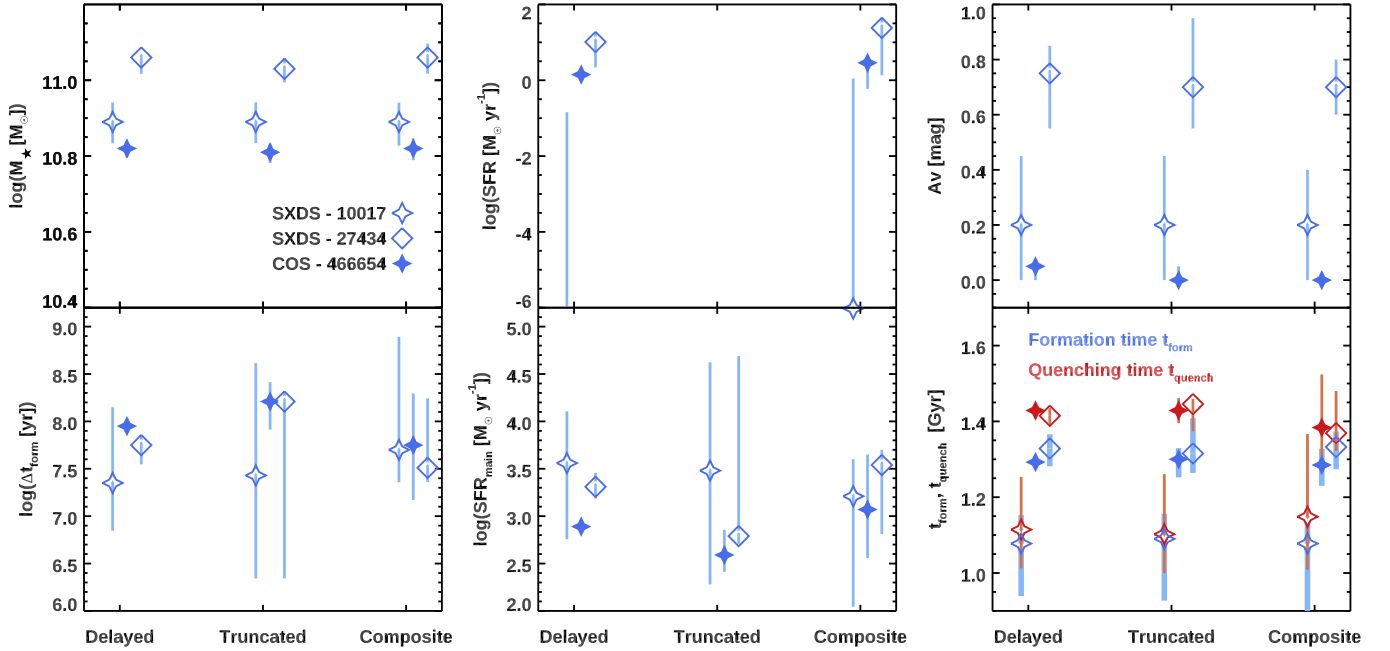
We acknowledge the constructive comments from the anonymous referee that improved the content and presentation of the results. We warmly thank Corentin Schreiber for the useful discussions and guidance in the use of SLINEFIT and FAST++. F.V. thanks Davide Martizzi for discussions about the *Illustris* simulations and Tao Wang and Helmut Dannerbauer for comments about the submillimeter galaxy samples. F.V. and G.M. acknowledge Villum Fonden research grant 13160, “Gas to stars, stars to dust: tracing star formation across cosmic time,” and Carlsbergfondet research grant CF18-0388, “Galaxies: Rise and Death.” S.T., M.S., C.G.G., and G.M. acknowledge support from the European Research Council (ERC) Consolidator Grant funding scheme (project “ConTEExt,” grant No. 648179). The Cosmic Dawn Center (DAWN) is funded by the Danish National Research Foundation under grant No. 140. M.O. acknowledges support from JSPS KAKENHI grant No. JP17K14257. M.S. acknowledges support by the European Research Council under ERC-CoG

grant CRAGSMAN-646955. O.I. acknowledges the funding of the French Agence Nationale de la Recherche for the project “SAGACE” and the Centre National d’Etudes Spatiales (CNES). K.Y. was supported by JSPS KAKENHI grant No. JP18K13578. The Flatiron Institute is supported by the Simons Foundation. Based on observations collected at the European Southern Observatory under ESO program 0100.B-0922(A). Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation. The observations were carried out within the framework of the Subaru–Keck time exchange program, where the travel expense was supported by the Subaru Telescope, which is operated by the National Astronomical Observatory of Japan. The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the indigenous Hawaiian

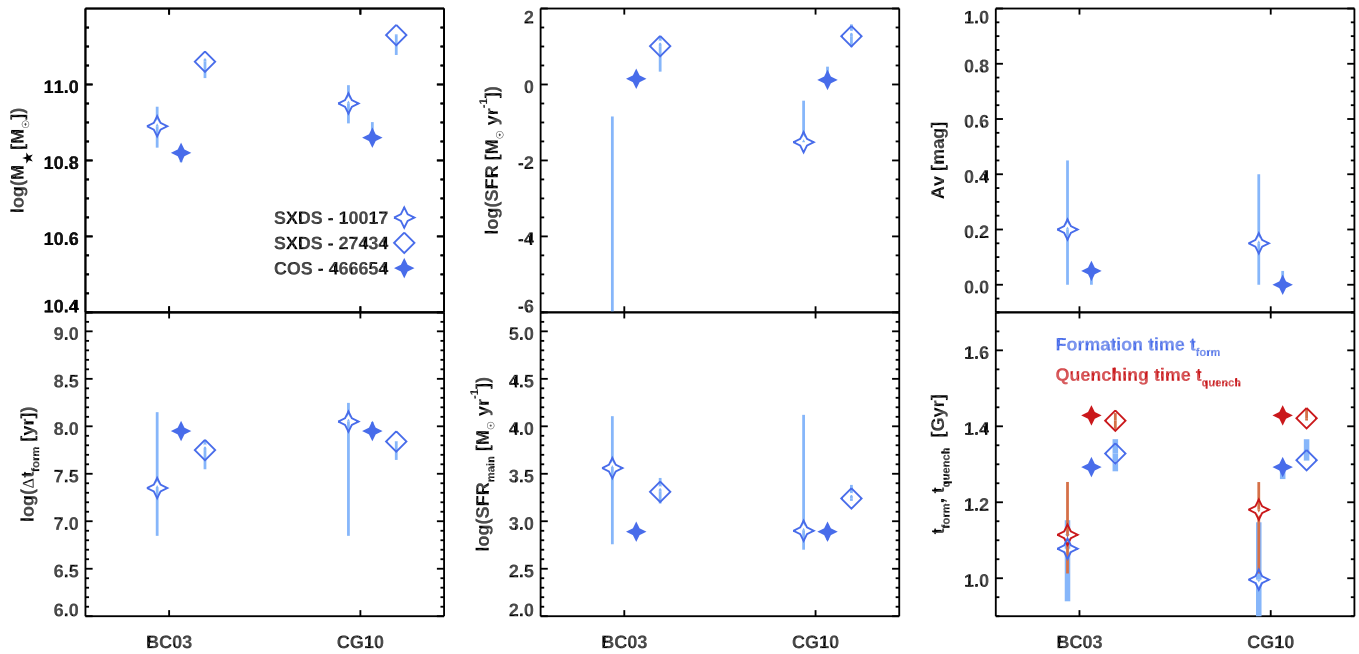
community. We are most fortunate to have the opportunity to conduct observations from this mountain.

## Appendix A Stellar Population Models, SFHs, and Their Parameterization

In Figure 9, we show the main quantities derived from the SED modeling that we used in this work, obtained by fitting the same set of data but varying the parameterization of the SFH. The quantities we adopted to describe the formation and quenching epochs of the studied objects are robust against the choice of the SFH. Notice that by definition, the current SFR for the truncated SFH is null. The error bars represent the 90% confidence intervals computed following the  $\chi^2$  criterion by Avni (1976; Section 5). In Figure 10, we present the same quantities computed by varying the stellar templates (Bruzual & Charlot 2003; Conroy & Gunn 2010) while fixing the SFH to a delayed model. The choice of the stellar templates does not affect our results.



**Figure 9.** Main quantities derived from SED modeling varying the SFH. From top left (clockwise): stellar mass, SFR,  $A_V$  dust attenuation,  $t_{\text{quench}}$  and  $t_{\text{form}}$ ,  $\langle \text{SFR} \rangle_{\text{main}}$ , and duration of the main formation epoch as a function of the SFH (delayed, truncated, or composite as in Schreiber et al. 2018a) for the three sources studied in this work (open stars: SXDS-10017; open diamonds: SXDS-27434; filled stars: COS-466654). The definition of each quantity and the analytical expressions of the SFHs are described in Section 5. The error bars correspond to the 90% confidence intervals computed following Avni (1976).



**Figure 10.** Main quantities derived from SED modeling varying the stellar models. From top left (clockwise): stellar mass, SFR,  $A_V$  dust attenuation,  $t_{\text{quench}}$  and  $t_{\text{form}}$ , ( $\text{SFR}$ )<sub>main</sub>, and duration of the main formation epoch as a function of the stellar population models (Bruzual & Charlot 2003; Conroy & Gunn 2010) for the three quiescent galaxies at fixed delayed SFH. The error bars correspond to the 90% confidence intervals computed following Avni (1976).

## Appendix B

### Properties of Candidate Progenitors of Quiescent Galaxies

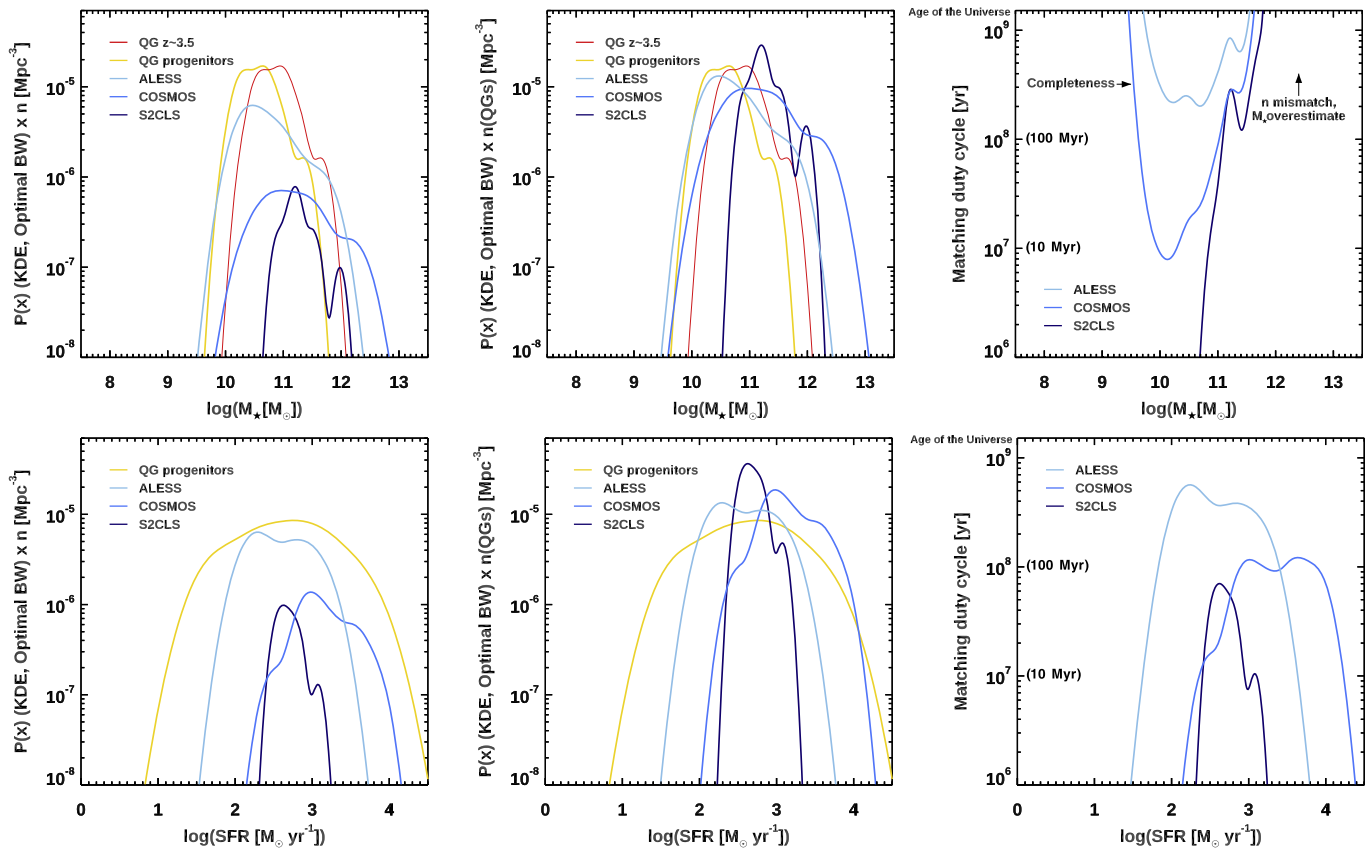
In Figure 7, we presented the (Gaussian) KDE of the observed stellar mass and SFR distributions of quiescent galaxies at  $3 < z < 4$  and SMGs at  $z > 4$ , each one normalized by the comoving number density of the corresponding sample. In the middle panels of Figure 11, we normalized all of the distributions to the same final number density of quiescent galaxies,  $n = 1.4 \times 10^{-5} \text{ Mpc}^{-1}$  (S18b). This value is only for reference, as the final results scale proportionally with the choice of this parameter (Section 7.4.2). To facilitate the comparison and appreciate the difference between the different normalizations, we show the same panels of Figure 7 in Figure 11. The shapes of the KDEs are identical in the left and middle panels, but the areas are different.

First, the renormalization exacerbates the tensions among the stellar mass distributions discussed in Section 7.4.1, pointing at inconsistencies in the modeling of the SMG population yet to be solved. Second, as mentioned in Section 7.4.2, we can derive a duty cycle for SMGs by assuming that all SMGs at  $z > 4$  become quiescent at  $3 < z < 4$ . The right panels of Figure 11 show the  $t_{\text{burst}}$  necessary to match the number density (i.e., the area under the curves) and the stellar mass or SFR distributions (i.e., the shape of the curves) of the quiescent progenitors and the samples of SMGs. Matching the distributions of stellar masses or SFRs provides consistent results for ALESS as a consequence of the overall agreement of both sets of distributions with the progenitor properties. On the other hand, as noted above, matching the  $M_*$  KDEs for the COSMOS and S2CLS sources results in  $>10\times$  shorter  $t_{\text{burst}}$  than matching the SFR distributions. Notice that the optical/near-infrared SED modeling to derive the stellar masses for SMGs is, in principle, rather independent of the derivation of the SFRs, the latter being connected to the far-infrared portion of the spectrum, allowing us to separate the two approaches to estimate  $t_{\text{burst}}$ .

### B.1. Caveats

We mentioned that the derivation of the duty cycle correction is valid only under the assumption that all SMGs at  $z > 4$  become quiescent at  $3 < z < 4$ , but more hypotheses and limitations enter this analysis. First, the calculation of the number densities depends on the depth and completeness of the surveys and the redshift estimates. In particular, the latter are hard to obtain for SMGs, relying on the correct identification of an optical/near-infrared counterpart and being naturally complicated by the extreme obscuration of these sources. However, our attempt to include the photometric redshift information and its uncertainty represents an improvement toward the determination of reliable number densities, rather than arbitrarily rescaling luminosity functions and the ensuing number counts by fixed fudge factors (see Section 8). The comparison of the  $M_*$  distributions of SMGs and quiescent galaxies suffers from similar uncertainties, relying on the modeling of the optical/near-infrared part of the spectrum. Moreover, the details of the modeling and the assumptions behind it notoriously generate systematic differences among different results. An accurate reconstruction of the stellar masses of SMGs is beyond the scope of this paper; thus, we assumed the results presented in the literature (see Gómez-Guijarro et al. 2018 for a detailed assessment of the optical/near-infrared properties of some of the  $z \sim 4.5$  SMGs in Miettinen et al. 2017 and the importance of spatial high-resolution observations). Here we stress only two salient differences among the various samples we presented. First, we use different stellar population synthesis codes for the quiescent population and the SMGs. In particular, the SEDs of all of the SMGs have been modeled with MAGPHYS, which enforces a global energy balance between the UV/optical and the infrared part of the electromagnetic spectrum. On the other hand, the quiescent galaxies, including the three sources presented in this paper, have been modeled with classical codes not assuming any energy balance (FAST++, LEHARE,





**Figure 11.** Stellar masses, SFRs, and duty cycle of candidate progenitors of high-redshift quiescent galaxies. Left: KDE of the density probability distributions of stellar masses (top panel) and SFR (bottom panel) for the progenitors at  $z_{\text{form}}$  of massive quiescent galaxies at  $3 < z < 4$  (golden line; this work; S18b) and SMGs at  $z > 4$  from ALESS, COSMOS, S2CLS (blue lines; da Cunha et al. 2015; Miettinen et al. 2017; Michałowski et al. 2017). For reference, we show the stellar mass distribution of the quiescent galaxies (red line). The areas under the curves are normalized to the number density of each population (Figure 6). Middle: KDE of the distributions of stellar masses (top panel) and SFR (bottom panel) for the same populations as the left panel. The area under every curve is normalized to the number density of the quiescent population at  $3 < z < 4$  by Schreiber et al. (2018a) (i.e., assuming that all SMGs at  $z > 4$  turn into quiescent galaxies at  $3 < z < 4$ ). Right: blue lines indicate the  $t_{\text{burst}}$  to assume as a function of  $M_*$  to match the normalization (i.e., the number density) and the shape of the stellar mass (top panel) and SFR (bottom panel) distributions of high-redshift SMGs (da Cunha et al. 2015; Michałowski et al. 2017; Miettinen et al. 2017) and the quiescent descendants at  $3 < z < 4$  (this work; S18b), assuming that all SMGs become quiescent.

MIZUKI), but this should not play a major role in dust-poor quiescent galaxies. The choice of the SFH is a second element of relevance when using MAGPHYS for SMGs, driving variations up to  $\sim 0.5$  dex in the final stellar mass, as evident from Figures 5 and 11 (Michałowski et al. 2014). Overall, we find that the current estimates of  $M_*$  for SMGs are systematically larger than in lower-redshift galaxies, quiescent or not (see also Miettinen et al. 2017 and Michałowski et al. 2017). Notice that we did not take into account the contribution of the stellar mass assembled during the ongoing episode of star formation in SMGs at  $z > 4$ , which would increase the differences even with the lower-redshift objects. We further remark that the expected properties of the progenitors of quiescent galaxies also fully depend on our implementation of the SED modeling, therefore being susceptible to possible systematics due to the current assumptions. Finally, our calculations assume fixed comoving number densities at the explored redshifts, while mergers might a priori change the mass ranking at the base of this technique (Torrey et al. 2017).

### ORCID iDs

Francesco Valentino <https://orcid.org/0000-0001-6477-4011>  
 Masayuki Tanaka <https://orcid.org/0000-0002-5011-5178>  
 Iary Davidzon <https://orcid.org/0000-0002-2951-7519>

Sune Toft <https://orcid.org/0000-0003-3631-7176>  
 Carlos Gómez-Guijarro <https://orcid.org/0000-0002-4085-9165>  
 Mikkel Stockmann <https://orcid.org/0000-0001-5983-6273>  
 Masato Onodera <https://orcid.org/0000-0003-3228-7264>  
 Gabriel Brammer <https://orcid.org/0000-0003-2680-005X>  
 Daniel Ceverino <https://orcid.org/0000-0002-8680-248X>  
 Andreas L. Faist <https://orcid.org/0000-0002-9382-9832>  
 Anna Gallazzi <https://orcid.org/0000-0002-9656-1800>  
 Christopher C. Hayward <https://orcid.org/0000-0003-4073-3236>  
 Olivier Ilbert <https://orcid.org/0000-0002-7303-4397>  
 Mariko Kubo <https://orcid.org/0000-0002-7598-5292>  
 Georgios E. Magdis <https://orcid.org/0000-0002-4872-2294>  
 Jonatan Selsing <https://orcid.org/0000-0001-9058-3892>  
 Rhythm Shimakawa <https://orcid.org/0000-0003-4442-2750>  
 Martin Sparre <https://orcid.org/0000-0002-9735-3851>  
 Charles Steinhardt <https://orcid.org/0000-0003-3780-6801>  
 Kiyoto Yabe <https://orcid.org/0000-0001-6229-4858>  
 Johannes Zabl <https://orcid.org/0000-0002-9842-6354>

### References

Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, *MNRAS*, 310, 540  
 Arnouts, S., Le Floc'h, E., Chevallard, J., et al. 2013, *A&A*, 558, A67

- Avni, Y. 1976, *ApJ*, **210**, 642
- Barro, G., Faber, S. M., Koo, D. C., et al. 2017, *ApJ*, **840**, 47
- Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, *ApJ*, **765**, 104
- Belli, S., Genzel, R., Förster Schreiber, N. M., et al. 2017a, *ApJL*, **841**, L6
- Belli, S., Newman, A. B., & Ellis, R. S. 2017b, *ApJ*, **834**, 18
- Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, **874**, 17
- Belli, S., Newman, A. B., Ellis, R. S., & Konidakis, N. P. 2014, *ApJL*, **788**, L29
- Benítez, N. 2000, *ApJ*, **536**, 571
- Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, *A&A*, **363**, 476
- Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, *MNRAS*, **398**, 1150
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, **686**, 1503
- Brisbin, D., Miettinen, O., Aravena, M., et al. 2017, *A&A*, **608**, A15
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
- Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, *ApJ*, **429**, 582
- Cappellari, M. 2017, *MNRAS*, **466**, 798
- Cappellari, M., & Emsellem, E. 2004, *PASP*, **116**, 138
- Casey, C. M., Narayanan, D., & Cooray, A. 2014, *PhR*, **541**, 45
- Cassata, P., Giavalisco, M., Williams, C. C., et al. 2013, *ApJ*, **775**, 106
- Cecchi, R., Bolzonella, M., Cimatti, A., & Girelli, G. 2019, *ApJL*, **880**, L14
- Chabrier, G. 2003, *PASP*, **115**, 763
- Choi, J., Conroy, C., Moustakas, J., et al. 2014, *ApJ*, **792**, 95
- Ciesla, L., Boselli, A., Elbaz, D., et al. 2016, *A&A*, **585**, A43
- Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, *A&A*, **482**, 21
- Conroy, C., & Gunn, J. E. 2010, *ApJ*, **712**, 833
- Cooke, E. A., Smail, I., Swinbank, A. M., et al. 2018, *ApJ*, **861**, 100
- da Cunha, E., Charlot, S., & Elbaz, D. 2008, *MNRAS*, **388**, 1595
- da Cunha, E., Walter, F., Smail, I. R., et al. 2015, *ApJ*, **806**, 110
- Daddi, E., Bournaud, F., Walter, F., et al. 2010, *ApJ*, **713**, 686
- Daddi, E., Renzini, A., Pirzkal, N., et al. 2005, *ApJ*, **626**, 680
- Danielson, A. L. R., Swinbank, A. M., Smail, I., et al. 2017, *ApJ*, **840**, 78
- Davé, R., Thompson, R., & Hopkins, P. F. 2016, *MNRAS*, **462**, 3265
- Davidzon, I., Ilbert, O., Faist, A. L., Sparre, M., & Capak, P. L. 2018, *ApJ*, **852**, 107
- Davidzon, I., Ilbert, O., Laigle, C., et al. 2017, *A&A*, **605**, A70
- Deshmukh, S., Caputi, K. I., Ashby, M. L. N., et al. 2018, *ApJ*, **864**, 166
- Donnari, M., Pillepich, A., Nelson, D., et al. 2019, *MNRAS*, **485**, 4817
- Dudzevičiūtė, U., Smail, I., Swinbank, A. M., et al. 2019, arXiv:1910.07524
- Fontana, A., Santini, P., Grazian, A., et al. 2009, *A&A*, **501**, 15
- Forrest, B., Annunziatella, M., Wilson, G., et al. 2019, arXiv:1910.10158
- Furusawa, H., Kosugi, G., Akiyama, M., et al. 2008, *ApJS*, **176**, 1
- Gallazzi, A., Charlot, S., Brinchmann, J., White, S. D. M., & Tremonti, C. A. 2005, *MNRAS*, **362**, 41
- Geach, J. E., Dunlop, J. S., Halpern, M., et al. 2017, *MNRAS*, **465**, 1789
- Gehrels, N. 1986, *ApJ*, **303**, 336
- Genel, S., Vogelsberger, M., Springel, V., et al. 2014, *MNRAS*, **445**, 175
- Girelli, G., Bolzonella, M., & Cimatti, A. 2019, *A&A*, **632**, 80
- Glazebrook, K., Schreiber, C., Labbé, I., et al. 2017, *Natur*, **544**, 71
- Gobat, R., Strazzullo, V., Daddi, E., et al. 2012, *ApJL*, **759**, L44
- Gómez-Guijarro, C., Magdis, G. E., Valentino, F., et al. 2019, *ApJ*, **886**, 88
- Gómez-Guijarro, C., Toft, S., Karim, A., et al. 2018, *ApJ*, **856**, 121
- Guarnieri, P., Maraston, C., Thomas, D., et al. 2019, *MNRAS*, **483**, 3060
- Hayward, C. C., Jonsson, P., Kereš, D., et al. 2012, *MNRAS*, **424**, 951
- Hayward, C. C., Kereš, D., Jonsson, P., et al. 2011, *ApJ*, **743**, 159
- Hayward, C. C., Narayanan, D., Kereš, D., et al. 2013, *MNRAS*, **428**, 2529
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, *MNRAS*, **451**, 2663
- Hodge, J. A., Karim, A., Smail, I., et al. 2013, *ApJ*, **768**, 91
- Home, K. 1986, *PASP*, **98**, 609
- Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, *A&A*, **457**, 841
- Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, *A&A*, **556**, A55
- Ilbert, O., Salvato, M., Le Floc'h, E., et al. 2010, *ApJ*, **709**, 644
- Inoue, A. K. 2011, *MNRAS*, **415**, 2920
- Iverson, R. J., Lewis, A. J. R., Weiss, A., et al. 2016, *ApJ*, **832**, 78
- Jin, S., Daddi, E., Liu, D., et al. 2018, *ApJ*, **864**, 56
- Karim, A., Swinbank, A. M., Hodge, J. A., et al. 2013, *MNRAS*, **432**, 2
- Kashino, D., Silverman, J. D., Sanders, D., et al. 2019, *ApJS*, **241**, 10
- Kennicutt, R. C., Jr. 1998, *ARA&A*, **36**, 189
- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, **700**, 221
- Kubo, M., Tanaka, M., Yabe, K., et al. 2018, *ApJ*, **867**, 1
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, **224**, 24
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, *MNRAS*, **379**, 1599
- Man, A., & Belli, S. 2018, *NatAs*, **2**, 695
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, *MNRAS*, **480**, 5113
- Mawatari, K., Yamada, T., Fazio, G. G., Huang, J.-S., & Ashby, M. L. N. 2016, *PASJ*, **68**, 46
- McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, *A&A*, **544**, A156
- McLean, I. S., Steidel, C. C., Epps, H. W., et al. 2012, *Proc. SPIE*, **8446**, 84460J
- Merlin, E., Fontana, A., Castellano, M., et al. 2018, *MNRAS*, **473**, 2098
- Merlin, E., Fortuni, F., Torelli, M., et al. 2019, *MNRAS*, **490**, 3309
- Michałowski, M. J., Dunlop, J. S., Koprowski, M. P., et al. 2017, *MNRAS*, **469**, 492
- Michałowski, M. J., Hayward, C. C., Dunlop, J. S., et al. 2014, *A&A*, **571**, A75
- Miettinen, O., Delvecchio, I., Smolčić, V., et al. 2017, *A&A*, **606**, A17
- Modigliani, A., Goldoni, P., Royer, F., et al. 2010, *Proc. SPIE*, **7737**, 773728
- Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, *ApJ*, **731**, 113
- Muzzin, A., Wilson, G., Demarco, R., et al. 2013, *ApJ*, **767**, 39
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **477**, 1206
- Nelson, D., Pillepich, A., Genel, S., et al. 2015, *A&C*, **13**, 12
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *ComAC*, **6**, 2
- Oliver, S. J., Bock, J., Altieri, B., et al. 2012, *MNRAS*, **424**, 1614
- Onodera, M., Carollo, C. M., Renzini, A., et al. 2015, *ApJ*, **808**, 161
- Pacifici, C., Oh, S., Oh, K., Lee, J., & Yi, S. K. 2016, *ApJ*, **824**, 45
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, *MNRAS*, **475**, 648
- Popping, G., Decarli, R., Man, A. W. S., et al. 2017, *A&A*, **602**, A11
- Prevot, M. L., Lequeux, J., Maurice, E., Prevot, L., & Rocca-Volmerange, B. 1984, *A&A*, **132**, 389
- Renzini, A. 2006, *ARA&A*, **44**, 141
- Safarzadeh, M., Hayward, C. C., Ferguson, H. C., & Somerville, R. S. 2016, *ApJ*, **818**, 62
- Santini, P., Merlin, E., Fontana, A., et al. 2019, *MNRAS*, **486**, 560
- Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018a, *A&A*, **618**, A85
- Schreiber, C., Labbé, I., Glazebrook, K., et al. 2018b, *A&A*, **611**, A22
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, **575**, A74
- Scoville, N., Aussel, H., Brusa, M., et al. 2007, *ApJS*, **172**, 1
- Selsing, J., Malesani, D., Goldoni, P., et al. 2019, *A&A*, **623**, 92
- Simpson, C., Martínez-Sansigre, A., Rawlings, S., et al. 2006, *MNRAS*, **372**, 741
- Simpson, J. M., Smail, I., Wang, W.-H., et al. 2017, *ApJL*, **844**, L10
- Simpson, J. M., Swinbank, A. M., Smail, I., et al. 2014, *ApJ*, **788**, 125
- Smolčić, V., Karim, A., Miettinen, O., et al. 2015, *A&A*, **576**, A127
- Sparre, M., Hayward, C. C., Springel, V., et al. 2015, *MNRAS*, **447**, 3548
- Spitler, L. R., Straatman, C. M. S., Labbé, I., et al. 2014, *ApJL*, **787**, L36
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, *MNRAS*, **475**, 676
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629
- Steinhardt, C. L., Capak, P., Masters, D., & Speagle, J. S. 2016, *ApJ*, **824**, 21
- Stockmann, M., Toft, S., Gallazzi, A., et al. 2020, *ApJ*, **888**, 4
- Straatman, C. M. S., Labbé, I., Spitler, L. R., et al. 2014, *ApJL*, **783**, L14
- Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, *Natur*, **463**, 781
- Tadaki, K., Iono, D., Yun, M. S., et al. 2018, *Natur*, **560**, 613
- Tanaka, M. 2015, *ApJ*, **801**, 20
- Tanaka, M., Valentino, F., Toft, S., et al. 2019, *ApJL*, **885**, L34
- Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, *ApJ*, **621**, 673
- Toft, S., Gallazzi, A., Zirm, A., et al. 2012, *ApJ*, **754**, 3
- Toft, S., Smolčić, V., Magnelli, B., et al. 2014, *ApJ*, **782**, 68
- Toft, S., van Dokkum, P., Franx, M., et al. 2007, *ApJ*, **671**, 285
- Torrey, P., Wellons, S., Ma, C.-P., Hopkins, P. F., & Vogelsberger, M. 2017, *MNRAS*, **467**, 4872
- Trujillo, I., Förster Schreiber, N. M., Rudnick, G., et al. 2006, *ApJ*, **650**, 18
- van de Sande, J., Kriek, M., Franx, M., et al. 2013, *ApJ*, **771**, 85
- van Dokkum, P. G., Kriek, M., & Franx, M. 2009, *Natur*, **460**, 717
- Vernet, J., Dekker, H., D'Odorico, S., et al. 2011, *A&A*, **536**, A105
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, *MNRAS*, **444**, 1518
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014b, *Natur*, **509**, 177
- Wang, T., Schreiber, C., Elbaz, D., et al. 2019, *Natur*, **572**, 211
- Weiß, A., Kovács, A., Coppin, K., et al. 2009, *ApJ*, **707**, 1201
- Wellons, S., Torrey, P., Ma, C.-P., et al. 2015, *MNRAS*, **449**, 361
- Williams, C. C., Labbe, I., Spilker, J., et al. 2019, *ApJ*, **884**, 154
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, **691**, 1879
- Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, *MNRAS*, **450**, 2327