A&A 602, A35 (2017)
DOI: 10.1051/0004-6361/201628141

Astronomy & Astrophysics

The extended epoch of galaxy formation: Age dating of ∼3600 galaxies with 2 < z < 6.5 in the VIMOS Ultra-Deep Survey⋆

R. Thomas[1,2], O. Le Fèvre[1], M. Scodeggio[3], P. Cassata[2], B. Garilli[3], V. Le Brun[1], B. C. Lemaux[1], D. Maccagni[3], J. Pforr[1], L. A. M. Tasca[1], G. Zamorani[4], S. Bardelli[4], N. P. Hathi[1], L. Tresse[1], E. Zucca[4], and A. M. Koekemoer[5]

[1] Aix-Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388 Marseille, France
[2] Instituto de Fisica y Astronomía, Facultad de Ciencias, Universidad de Valparaíso, Gran Bretaña 1111, Playa Ancha, 2360102 Valparaíso, Chile
e-mail: romain.thomas@uv.cl
[3] INAF–IASF Milano, via Bassini 15, 20133 Milano, Italy
[4] INAF–Osservatorio Astronomico di Bologna, via Ranzani, 40127 Bologna, Italy
[5] Space Telescope Science Institute 3700 San Martin Drive, Baltimore, MD 21218, USA

Received 15 January 2016 / Accepted 8 December 2016

ABSTRACT

In this paper we aim at improving constraints on the epoch of galaxy formation by measuring the ages of 3597 galaxies with reliable spectroscopic redshifts $2 \leq z \leq 6.5$ in the VIMOS Ultra Deep Survey (VUDS). We derive ages and other physical parameters from the simultaneous fitting with the GOSSIP+ software of observed UV rest-frame spectra and photometric data from the u band up to 4.5 μm using model spectra from composite stellar populations. We perform extensive simulations and conclude that at $z \geq 2$ the joint analysis of spectroscopy and photometry, combined with restricted age possibilities when taking the age of the Universe into account, substantially reduces systematic uncertainties and degeneracies in the age derivation; we find that age measurements from this process are reliable. We find that galaxy ages range from very young with a few tens of million years to substantially evolved with ages up to 1.5 Gyr or more. This large age spread is similar for different age definitions including ages corresponding to the last major star formation event, stellar mass-weighted ages, and ages corresponding to the time since the formation of 25% of the stellar mass. We derive the formation redshift z_f from the measured ages and find galaxies that may have started forming stars as early as $z_f \sim 15$. We produce the formation redshift function (FzF), the number of galaxies per unit volume formed at a redshift z_f, and compare the FzF in increasing observed redshift bins finding a remarkably constant FzF. The FzF is parametrized with $(1+z)^{\zeta}$, where $\zeta \simeq 0.58 \pm 0.06$, indicating a smooth increase of about 2 dex from the earliest redshifts, $z \sim 15$, to the lowest redshifts of our sample at $z \sim 2$. Remarkably, this observed increase in the number of forming galaxies is of the same order as the observed rise in the star formation rate density (SFRD). The ratio of the comoving SFRD with the FzF gives an average SFR per galaxy of ∼7−17 $M_{\odot}$/yr at $z \sim 4-6$, in agreement with the measured SFR for galaxies at these redshifts. From the smooth rise in the FzF we infer that the period of galaxy formation extends all the way from the highest possible formation redshifts that we can probe at $z \sim 15$ down to redshifts $z \sim 2$. This indicates that galaxy formation is a continuous process over cosmic time, with a higher number of galaxies forming at the peak in SFRD at $z \sim 2$ than at earlier epochs.

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

1. Introduction

The time when galaxies formed remains poorly constrained. In the current ΛCDM hierarchical structure formation paradigm, dark matter (DM) halos form from the growth under gravity of early fluctuations in the matter density field, occasionally merging to form increasingly larger DM halos. Galaxies form as matter collapses in these deep potential wells, with gas cooling and fragmentation triggering star formation on galaxy scales (White & Rees 1978; Bromm et al. 2009), as simulated in cosmological volumes (e.g. Vogelsberger et al. 2014).

The epoch when the first galaxies are born still remains difficult to constrain observationally. This process is assumed to have begun when hydrogen in the Universe was being reionized, which is a time when photons emitted from young stars could hardly escape the neutral hydrogen still surrounding them. As young stars are born, it is assumed that they ionize their immediate surroundings rapidly and create a bubble inside which the medium is fully ionized (Bromm et al. 2009). Growing bubbles then eventually overlap, fully ionizing the Universe. This reionization process is a most fundamental step in galaxy evolution and its exact duration and end is a matter of considerable debate. The most recent results of cosmic microwave background (CMB) observations with the *Planck* satellite report that the reionization optical depth is $\tau = 0.066 \pm 0.016$, leading to a redshift at which half of the Universe is reionized of $z_{re} = 8.8^{+1.7}_{-1.4}$ (Planck Collaboration XVI 2014), which is significantly later in cosmic time than was estimated from WMAP with $z_{re} = 10.4$ (Hinshaw et al. 2013). However, the CMB results do not tell us when the main sources responsible for reionization formed, and these results do not contain information about galaxies forming after reionization is completed.

Understanding when the populations of galaxies at different epochs in the Universe formed requires a complete census of galaxies at the highest possible redshifts. The highest redshift

⋆ Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, under Large Programme 185.A–0791.

galaxy candidates identified so far are at $z \simeq 10$ (Bouwens et al. 2014, 2015). But beyond $z \sim 8.7$ (Zitrin et al. 2015), galaxies are identified solely based on photometric properties and await spectroscopic confirmation. Moreover, this confirmation rests mainly on the Lyman-α emission line, which is the only emission feature that is accessible with current facilities; this emission line is rare at those redshifts, which makes these confirmations very difficult.

Other than the direct identification of galaxies in the reionization epoch, another way to probe the formation of the first galaxies is to measure the ages of galaxies securely identified from spectroscopy at redshifts close to, but not necessarily into, the reionization epoch, and infer their redshift of formation. From the properties of observed galaxies with a formation redshift in the reionization era it is then possible to perform some form of galaxy archaeology by estimating what would have been the bulk properties of these galaxies during the reionization epoch.

Measuring ages of galaxies and associated limitations at increasingly higher redshifts was performed in a number of studies. Using the SDSS sample Thomas et al. (2005) derived ages, total metallicities, and element ratios of 124 early-type galaxies in high and low density environments. These results show that most star formation activity in early-type galaxies happened between redshifts ∼3 and 5 in high density and between redshifts 1 and 2 in low density environments. Cimatti et al. (2008) identified passive galaxies at $1.4 < z < 2$ and inferred ages of ∼1 Gyr from a stacked spectrum, hence placing the formation redshift beyond $z \simeq 2$. Kaviraj et al. (2013) selected ∼330 spheroids at redshifts $1 < z < 3$ and derived their age using spectral energy distribution (SED) fitting. They found that the star formation in these galaxies likely peaked in the redshift range $2 < z < 5$ with a median of $z \sim 3$. At higher redshifts individual galaxy studies have reported ages placing their formation redshift beyond the end of reionization (e.g. Ouchi et al. 2013). Beyond these exploratory studies the distribution of ages, and hence of formation redshifts, of the general population of galaxies at a given epoch is still unclear. We do not know if galaxies formed coevally or over an extended period of cosmic time. The downsizing trends observed in galaxy studies (Cowie et al. 1996; De Lucia et al. 2006) seem to indicate that the most massive galaxies ended their star formation first, and that this cessation of star formation progresses to lower masses as cosmic time passes. A more extensive assessment of when galaxies formed their first stars would consolidate this picture.

Comparing the observed SED to stellar synthesis population models has long been recognized as a powerful method to measure galaxy ages, along with other important physical parameters such as the stellar mass $M_\star$ and star formation rate (SFR) as implemented in SED-fitting codes; these codes are extensively described in Bolzonella et al. (2000), Thomas et al. (2005), Ilbert et al. (2006), Franzetti et al. (2008), Brammer et al. (2008). While the process of measuring galaxy ages follows the same method as for $M_\star$ and SFR measurements, measuring the age of galaxies is traditionally considered a more uncertain parameter. However, a study of galaxy ages at $z \gtrsim 2$ would allow the use of the natural age upper limit given by the age of the Universe. Thus, one may expect that getting to very high redshifts considerably helps limit the degeneracies plaguing galaxy age investigations at $z \lesssim 2$.

Here we present a study of the ages of an unprecedented sample of 3597 galaxies selected from the VIMOS Ultra Deep Survey (VUDS) with spectroscopic redshifts $2 < z < 6.5$. These galaxies are being observed at early cosmic epochs thus they must have formed at even earlier times. We may therefore identify galaxies that have formed most of their stars when the Universe was being reionized. The age of galaxies in our sample is derived using a novel technique applying SED fitting to the combination of spectra observed with VIMOS on the VLT with broadband photometry. We use extensive simulations to estimate uncertainties related to this process. From the ages and star formation histories (SFHs) we study the distribution of ages, derive the formation redshift distribution, and discuss consequences for the epoch of galaxy formation.

The paper is organized as follows. In Sect. 2 we summarize the VUDS sample. We present the technique developed to fit observed spectra and photometric data simultaneously with the GOSSIP+ software and discuss the benefit of this approach in Sect. 3. We present several age definitions in Sect. 4. The reliability of galaxy age estimates and tests for possible degeneracies at $z > 2$ are discussed in Sect. 5. The age distribution of galaxies with $2 \leq z \leq 6.5$ is presented in Sect. 6 and we derive and discuss the distribution in formation redshifts in Sect. 7. Results are discussed in Sect. 8 before concluding.

We use a cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{0,\Lambda} = 0.7$ and $\Omega_{0,\mathrm{m}} = 0.3$. All magnitudes are given in the AB system.

2. Data: the VUDS survey and associated photometry

We draw our sample from the VUDS, described in detail in Le Fèvre et al. (2015); a short summary is provided below.

The VIMOS Ultra Deep Survey was designed to study the evolution of galaxies in the very early Universe. The main selection criterion is based on photometric redshifts with $z_{\mathrm{phot}} + 1\sigma \geq 2.4$, and $i_{AB} \leq 25$. The secondary peak of the photometric redshift probability function is also taken into account if it satisfies $z_{\mathrm{phot}} + 1\sigma \geq 2.4$. In addition, objects in the Lyman break area of colour-colour plots (e.g. *g-r*, *r-i*) are also selected in case they are not already picked up by the photometric redshift selection; this sample represents about 10% of the whole target sample.

Observations are conducted on the ESO-VLT with the VIMOS spectrograph (Le Fèvre et al. 2003). Spectra are observed for ∼14 h in each of the LRBLUE and LRRED grisms with a spectral resolution $R \sim 230$ and cover a wavelength range from 3650 Å to 9350 Å.

Data are reduced with the VIPGI software (Scodeggio et al. 2005) and spectroscopic redshifts are measured with the EZ software (Garilli et al. 2010). The EZ software is based on the cross-correlation of observed spectra with reference spectra, followed by visual inspection of each spectrum by two independent observers who then compare their measurements to set the final "best" redshift measurement. A reliability flag is assigned to each redshift measurement, and gives the probability level for the redshift to be right, as described in Le Fèvre et al. (2015).

All VUDS galaxies are matched to the deep photometric catalogues available in each of the three VUDS fields: COSMOS, ECDFS, and VVDS-02h (Le Fèvre et al. 2015). In the following we use all the broadband optical and near-infrared photometric data available in these fields. In the COSMOS field this includes *u* band data from the CFHT Legacy Survey, griz bands from Subaru (Taniguchi et al. 2007), *YJHK* photometry from the UltraVista survey reaching $K_{AB} = 24.8$ at 5σ (McCracken et al. 2012), as well as 3.6 and 4.5 micron *Spitzer* data from SPLASH (Steinhardt et al. 2014; Laigle et al., in prep.). The ECDFS has deep *UBVRI* imaging down to $R_{AB} = 25.3$ (5σ, Cardamone et al. 2010, and reference therein), WFC3

near-IR imaging that reaches as deep as $H_{AB} = 27.3-27.6$ in the CANDELS survey (Grogin et al. 2011; Koekemoer et al. 2011), and *Spitzer*-warm 3.6 and 4.5 μm imaging data down to $AB = 23.1$ from the warm *Spitzer* survey SERVS (Mauduit et al. 2012). In the VVDS-02h field u', g, r, i observations are available from the CFHTLS survey reaching $i_{AB} = 25.44$ at 50% completeness in the latest DR7 (Cuillandre et al. 2012). Deep infrared imaging was obtained with WIRCAM at CFHT in *YJHK* bands down to $Ks_{AB} = 24.8$ also at 50% completeness (Bielby et al. 2012).

Spectra are calibrated in flux in a two-step process. First a flux calibration is performed using the standard ESO calibration observations of spectrophotometric standard stars using the same VIMOS setup. A further iteration is applied on the spectra to correct for two well-known, wavelength-dependent additional effects: atmospheric transmission that depends on the airmass of the observations and atmospheric refraction that drives a fraction of an object's light out of the slit. Assuming that the light of the galaxy entering the slit is representative of the host object these corrections are applied on an object-by-object basis, as described in Thomas et al. (2017). The flux scale of each spectrum is normalized to the observed i-band photometric flux derived from the total magnitude estimated using the mag-auto parameter from SExtractor (Bertin & Arnouts 1996). This corrects for slit losses due to the 1 arcsecond slit width used for the observations, which is generally smaller than observed object sizes. At the end of the calibration process we compare broadband magnitudes derived from the imaging data and broadband magnitudes computed by integrating the flux of the spectrum given the same filter transmission curves. We find that the differences between the photometric and spectroscopic flux in the broadbands u filters is 0.02 ± 0.3, while the difference in i band is null by definition.

The sample of VUDS galaxies that we are using in this paper is presented in Sect. 3.4.

3. Combining spectra and photometry to measure ages and other physical parameters of galaxies: the GOSSIP+ tool

The SED fitting has reached a maturity level such that SFR and $M_\star$ are extensively used in the literature (e.g. Ilbert et al. 2006, 2009, 2013). From the stellar population models one can also recover the age(s) of the stellar population(s) assembled in the model.

At $z < 2$, age estimations are prone to degeneracies and ages are used only when derived from extensive analysis combining SED fitting with some specific spectral indicators that are able to break the degeneracies (see e.g. Thomas et al. 2005). Using photometry alone the degeneracies between age, metallicity, dust extinction, and SFH appear to be significant (e.g. Wuyts et al. 2009; Pforr et al. 2012). Therefore age measurements derived from SED fitting alone are often taken with skepticism. As a consequence, while the computation of $M_\star$ and SFR is well documented and those parameters are extensively used in the literature, ages of galaxies are only rarely the subject of detailed analysis.

In this section we present an improved method to determine ages based on the joint fitting of observed spectra and multi-band photometry.

3.1. Method

The classical approach to measure photometric redshifts together with physical properties of galaxies ($M_\star$, SFR, dust extinction, age, and metallicity) is to perform SED fitting of broadband photometric measurements with stellar population synthesis models (e.g. Bolzonella et al. 2000; Ilbert et al. 2006; Brammer et al. 2008; Maraston et al. 2010; Pforr et al. 2013). This process makes use of a small number of data points, ideally a dozen of broadband measurements ranging from the u band to the K band or beyond (see e.g. Ilbert et al. 2013), nevertheless providing remarkable constraints on main parameters such as stellar mass.

When spectroscopy is available in addition to photometric data, the information on the SED is significantly increased. The VIMOS spectra of VUDS are sampled by about 1000 spectral data points. The wavelength coverage ranges from 3650 Å to 9350 Å. At $z > 2$, this is sampling the UV rest frame and has the potential to improve constraints on recent star formation and dust extinction. With the strong constraint provided by the strength and slope of the UV continuum (Hathi et al. 2016) and intergalactic medium (IGM) transmission (Thomas et al. 2017), UV rest-frame spectroscopy can help reduce degeneracies that are typical when using the more coarse wavelength sampling of broadband photometry.

Based on these considerations, we developed a methodology that is able to benefit from both a set of photometric data points covering a large wavelength base, but with a poor wavelength sampling, and spectroscopic data covering a smaller wavelength range but with high sampling. This method is applied to the VUDS data but is of general interest when both spectroscopic and photometric datasets are available and complement each other in wavelength.

3.2. Joint spectroscopic and photometric fitting with stellar population synthesis models with GOSSIP+

The Galaxy Observed-Simulated SED Interactive Program (GOSSIP) is a tool built to perform the fitting of both spectroscopy and photometry with stellar population models. It was initially developed (Franzetti et al. 2008) to be used in the framework of the VVDS (Le Fèvre et al. 2005, 2013) and zCOSMOS (Lilly et al. 2007) surveys. The novelty of this software is the fact that it allows, for a given galaxy with a known redshift, for the combination of photometric measurements in different bands with spectroscopic data and performs a χ^2 minimization fitting against a set of synthetic galaxy spectra based on emission from stellar populations. The result of the fit is used to estimate several parameters including SFR, age, and $M_\star$ of the observed galaxy through the probability distribution function (PDF) of each parameter. The GOSSIP tool can be used with different synthetic templates, which can be computed from reference models such as Bruzual & Charlot (2003; hereafter BC03) as well as Maraston models (Maraston 2005; Maraston & Strömbäck 2011, hereafter M05 and M11, respectively). Other user-defined models can also be implemented in GOSSIP.

We developed a new version of the software, called GOSSIP+ to add new functionalities and to modify pre-existing functions. We describe the main modifications that have been implemented in the next subsections.

3.2.1. Use of all spectral points

In the original version of the software (Franzetti et al. 2008), the spectroscopic fitting was performed through a rebinning of the spectrum in larger wavelength bins. The new version of the software, GOSSIP+, now makes use of all the observed

spectroscopic points directly when performing the fit. The only constraint is that reference models need to have a spectral resolution equal or better to the spectral resolution of the observed data otherwise the data may need to be resampled to the lower resolution of the models.

3.2.2. Template builder

A template builder is implemented. This allows us to compute composite stellar population (CSP) models from a library of simple stellar populations (SSP) provided by BC03, M05, or M11. The SSPs are defined by two main parameters: the metallicity and initial mass function (IMF). The BC03 model allows us to explore metallicities from $Z = 0.0004$ to $Z = 0.05$, while M05 metallicities range from $Z = 0.001$ to $Z = 0.04$. The IMF can be either Salpeter (1955), Chabrier (2003), or Kroupa (2001). From those SSPs, GOSSIP+ creates a CSP by applying a SFH. Various SFH can be used, including exponentially declining SFH ($SFH \propto \tau^{-1} \exp(-t/\tau)$) or delayed τ models SFH ($SFH \propto t \times \tau^{-2} \exp(-t/\tau)$). The extinction by dust in a galaxy is then applied using a Calzetti law (Calzetti et al. 2000). The age of each extracted template is defined by the user between a few hundred thousands years up to 20 Gyr[1].

3.2.3. IGM prescription

Another new feature implemented in GOSSIP+ is the treatment of the IGM. The UV rest-frame spectra of distant galaxies combine intrinsic emission with absorption produced by intergalactic gas clouds present along the line of sight. The transmission of the IGM is an important property to be taken into account when analysing the SED of galaxies at $z > 1.5$ (Madau 1995), and models are systematically used when searching for and analysing distant galaxy spectra. Below this redshift the transmission of the IGM reaches ~100% because the intervening clouds along the line of sight are not numerous enough and the IGM contribution can be neglected, but above this redshift the IGM transmission becomes substantial and severely modifies the spectral shape blueward of the Lyman-α line.

Several models have been proposed to estimate the transmission of the IGM (e.g. Madau 1995; Meiksin 2006). Those models provide, for a given redshift, one single average IGM transmission as a function of wavelength, and this mean value is used by SED-fitting algorithms. However, using a mean value for the IGM transmission is a simplifying assumption for ensemble averages of large samples of galaxies and does not necessarily translate correctly to the measurements of individual galaxies as not all lines of sight on the sky are populated with the same density or distribution of HI clouds. In Thomas et al. (2017) we show that the IGM transmission towards distant galaxies presents a large range of values at any given redshift. While we find a mean transmission as predicted by Meiksin (2006) and in agreement with quasar studies (e.g., Beker et al. 2015), we find that the dispersion is large at any given redshift $2.5 < z < 5.5$.

To take the range of possible IGM transmission into account, we consider the IGM as a free parameter in GOSSIP+ adding to the mean IGM transmission curve from Meiksin (2006) six additional IGM transmission curves that span the observed range of transmission at any given redshift (Thomas et al. 2017). With this prescription the IGM transmission at 1100 Å can vary from 20% to 100% at $z = 3.0$ and from 5% to 50% at $z = 5.0$.

[1] BC03 and M05/M11 templates are intrinsically defined on an age grid of 221 ages.

3.2.4. Emission lines

It is now widely documented (e.g. de Barros et al. 2014) that emission lines in star-forming galaxies (SFGs), the dominant population at high redshifts, must be taken into account in the model template spectra used in the cross-correlation with the observed SED. Strong emission lines present in SFGs such as Hα, the Hβ and [OIII]4959/5007 Å doublet, [OII]3737 Å, and Lyα, may change observed broadband magnitudes by several tenths of magnitudes. Physical parameters derived from SED fitting may therefore be wrongly estimated if model spectra used in the SED fitting do not include these emission lines (e.g. Ilbert et al. 2009; de Barros et al. 2014).

We therefore implemented a treatment of emission lines in GOSSIP+. For the purpose of this analysis we opted to model emission lines using standard photoionization case-B recombination. Six emission lines are implemented: Lyα, [OII], [OIII$_{a,b}$], [Hβ], and [Hα]. The procedure implemented in GOSSIP+ works as follows. The rescaling factor between the synthetic model and the observation gives access to the SFR of the template. This SFR is then converted into [OII] luminosity with the Kennicut law (Kennicutt 1998). The emission lines are then estimated using the following line ratios: [OIII/OII] = 0.36, [Hβ/OII] = 0.61, [Hα/OII] = 1.77, and [Lyα/OII] = 2, which is similar to what is used in the LePhare software (Ilbert et al. 2009). As the Lyα line can be either in absorption or in emission this receives a particular treatment. A rough measurement of the equivalent width of the Lyα line is computed directly on the observed spectrum; if $EW(\mathrm{Ly}\alpha) < -10$ Å (here the minus sign means emission) the emission line is created following the above ratios. If $EW(\mathrm{Ly}\alpha) > -10$ Å the emission line is not taken into account as it is not expected to affect broadband magnitude measurements. One may argue that taking into account the Lyα line may not be correctly carried out since it does not linearly trace the SFR. To test for this effect, we fitted a sample of strong Lyα emitters ($EW(\mathrm{Ly}\alpha) < -25$ Å) with and without taking the Lyα line into account. The results for $M_\star$ and SFR are not significantly modified while age measurements are not affected at all.

3.3. Computation of χ^2 and PDF

The fit of a spectrum or photometric data follows the same recipe as standard SED-fitting codes (Bolzonella et al. 2000; Ilbert et al. 2006; Brammer et al. 2008; Maraston et al. 2010). We first create libraries of template models from BC03 and M05 to fit the spectra and photometry of VUDS galaxies. Then for each observed galaxy, we perform a χ^2 minimization over the entire template library. For a given galaxy and a given template the χ^2 is computed with

$$\chi_i^2 = \sum_{i=1}^{N} \frac{F_{\mathrm{obs},i} - A \times F_{\mathrm{temp},i}}{\sigma_i}, \quad (1)$$

where $F_{\mathrm{obs},i}$, $F_{\mathrm{temp},i}$, σ_i, and A are the observed flux (or magnitude), synthetic flux density (or magnitude) from the template, observed error, and normalization factor applied to the template, respectively. The normalization factor is computed from the comparison of the observed broadband photometry and the photometry computed on the galaxy template. The reduced χ^2, χ_ν^2, is then computed as in Salim et al. (2007) by considering that the number of parameters that are fitted is not linked to the number of physical parameters (i.e. properties) of the galaxies. The parameter that is fitted is then A. Consequently, the number of

degrees of freedom is given by $N - 1$, where N is the number of data points.

For a given observation GOSSIP+ computes the χ^2 of all the templates. The set of χ^2 values are then used to create the PDF. To build this PDF we compute for each template the probability

$$P_i = \exp\left(\frac{-\chi_i^2}{2}\right). \quad (2)$$

We use the median of the normalized PDF to estimate the parameter value. The errors on each parameter are taken from effective $\pm 1\sigma$ values of the PDF derived from the values including 68% of the PDF of each parameter centred on the median value of the PDF.

When GOSSIP+ combines both a photometry dataset and a spectrum it has to take into account that the weight of both datasets are not the same. The spectroscopy contains ~1000 points for VUDS-like spectra while the photometry is made of ~10 points. Then if both observations are not weighted one photometric point has the same importance as a spectroscopic point and the spectrum dominates the fit. We decided to give the same weight to both photometry and spectroscopy. Therefore, the combination of the full spectrum and full photometry in one single fit is performed with the combined $\chi^2_{\rm comb}$ defined as

$$\chi^2_{\rm comb} = \chi^2_\nu({\rm phot}) + \chi^2_\nu({\rm Spec}), \quad (3)$$

where, χ^2_ν(phot) and χ^2_ν(Spec) are the reduced χ^2 of the photometric and spectral fits, respectively.

To compute the PDF of the combined fit, we therefore choose to give the same weight to the large ensemble of points of the spectrum and to the smaller number of photometric points. This choice rests on the spectral coverage and S/N of the VUDS spectra, on the one hand, and on the broad wavelength coverage of the photometry, on the other hand. Different weighting schemes may be applied for different datasets, giving more or less relative weight to the photometry and spectroscopy depending, for example on the wavelength range and spectral resolution of the spectra or the wavelength coverage and depth of the photometric data. While a complete statistical treatment of the relative weights between spectroscopy and photometry remains to be defined, we emphasize that changes of our weighting scheme by a factor of two do not significantly modify the results presented in this paper.

3.4. Selection of reliable fits

In addition to selecting VUDS galaxies with spectroscopic reliability level of 50, 75, 95, 100, and 80% (flags 1–4, and 9, and the like, respectively; see Le Fèvre et al. 2015), we perform a selection based on the quality of the fit as described below.

Our fitting method is based on the ability to combine both the photometry and spectroscopy. The agreement between spectroscopy and photometry is therefore crucial to produce a good fit. As there can be a mismatch between these two datasets affecting the quality of fit, we implement a visual inspection of each fit to define its quality.

Four fit quality flags are used, as described in Table 1. The two weakest flags, 0 and 1, correspond to bad and poor fits, respectively. These flags indicate a significant departure of the fit compared to either of the photometric or spectroscopic data points, or both. A bad fit can also be the result of an incorrect spectroscopic redshift assignment, which becomes evident when using a broad wavelength range and the fit. Galaxies with fit quality flags 2 and 3 correspond to good to excellent fits, and these constitute the working sample for our study.

Table 1. Fit quality flag used in our study to select the best galaxy sample.

Fit flag	Meaning
0	Bad fit
1	Poor fit
2	Good fit
3	Excellent fit

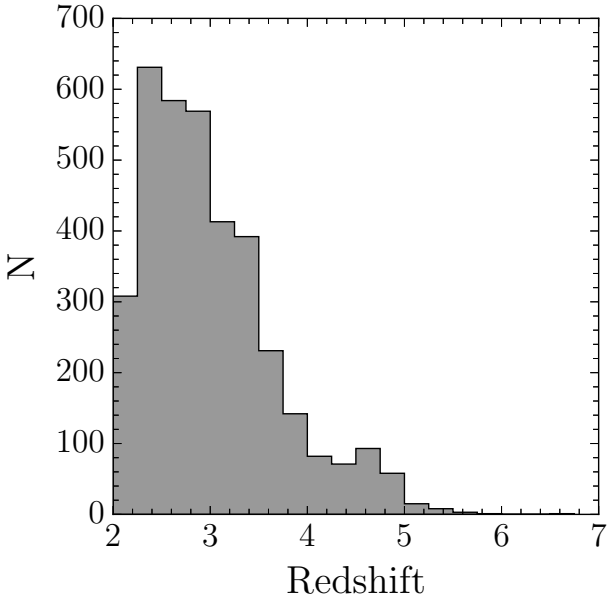

Fig. 1. Redshift distribution of the 3597 VUDS galaxies at $z > 2$ used in this study.

We verified that the visual selection mainly selects the fits with the smallest $\chi^2_{\rm comb}$, for which both the photometry and spectroscopy are well reproduced by a galaxy model (see Sect. 6.1 for examples). The fraction of selected galaxies with $\chi^2_{\rm comb} > 10$, corresponding to increasingly bad fits, is negligible. As expected the visual classification follows the redshift reliability classification of the survey (spectroscopic flag; see Le Fèvre et al. 2015). The visual classification therefore retains galaxies with both a reliable spectrum and SED fit and a reliable spectroscopic redshift. After this selection we selected ~67% of the total $z \geq 2$ VUDS galaxies. Our working catalogue is composed of 3597 VUDS galaxies with redshift between $z = 2$ and $z = 6.5$ that have a fit flag of 2 or 3. The redshift distribution of our sample is presented in Fig. 1. This sample is a representative sample of the general star-forming population in the VUDS survey with $M_\star$ and SFR distributions similar to the full sample.

4. The ages of galaxies: definitions

The age of a galaxy is a parameter that can be defined in many different and complementary ways. One may seek to find the time when the first stars were formed or the time in the life of a galaxy when a fraction of its stellar mass was assembled, as defined below.

Figure 2 provides a visual interpretation of the three definitions listed below.

We define $\mathcal{A}_{\rm onset}$ as the age corresponding to the onset of star formation, i.e. the time since the beginning of the SFH that corresponds to the beginning of the stellar mass assembly of the

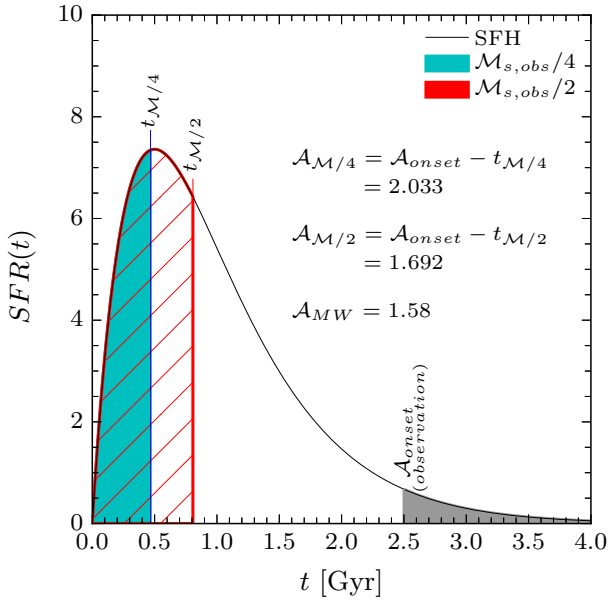

Fig. 2. Different age definitions and their relation to stellar mass on an example. The grey line represents the SFH of the galaxy while the grey shaded area is in the future of the observation (the "0" of the x-axis represents the formation of the galaxy). The blue and red areas represent one-quarter and half of the stellar mass, respectively. Finally, $t_{M/4}$ and $t_{M/2}$ show the time since the onset of the star formation and the time to build up 25% and 50% of the stellar mass, respectively. In this example the galaxy is observed 2.5 Gyr after the start of star formation, $\mathcal{A}_{\rm onset}$ = 2.5 Gyr, and the stellar mass is 10^{10} $\mathcal{M}_{\odot}$. This induces $\mathcal{A}_{M/4}$ = 2.033 Gyr, $\mathcal{A}_{M/2}$ = 1.692 Gyr, and $\mathcal{A}_{\rm MW}$ = 1.58 Gyr.

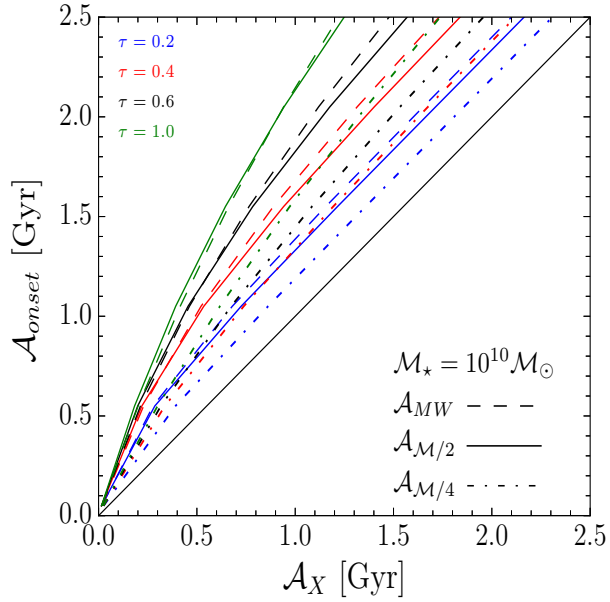

Fig. 3. Comparison of different age definitions to the age defined as the onset of star formation $\mathcal{A}_{\rm onset}$. Different colours indicate different timescales for the SFH: $\tau_{\rm SFH}$ = 0.4 Gyr in red, $\tau_{\rm SFH}$ = 0.6 Gyr in black, and $\tau_{\rm SFH}$ = 1.0 Gyr in green. The different line styles correspond to three age definitions: solid line for half-mass ages ($\mathcal{A}_{\mathcal{M}/2}$), dashed line for mass weighted ages ($\mathcal{A}_{\rm MW}$), and dot-dashed line for quarter-mass ages ($\mathcal{A}_{\mathcal{M}/4}$).

galaxy. This definition corresponds to the output of the standard stellar population models with smooth SFH used to perform SED fitting. The parameter $\mathcal{A}_{\rm onset}$ is related to the observed $M_{\star}$ by the following relation (see Fig. 2):

$$\mathcal{M}_{\star,\rm obs} = \int_0^{\mathcal{A}_{\rm onset}} SFH(t)\mathrm{d}t, \tag{4}$$

where $\mathcal{M}_{\star,\rm obs}$ and $SFH(t)$ are the observed $M_{\star}$ and SFH of the galaxy, respectively.

One may argue that defining the starting point of the life of a galaxy is difficult. It is then useful to define the age of a galaxy as the time when it has built up a large portion of its observed stellar mass. We therefore use two other age definitions that represent the time since the galaxy has built up 25% and 50% of its current stellar mass at the time of observation. We refer to these ages as the quarter mass age noted $\mathcal{A}_{M/4}$ and the half mass age, $\mathcal{A}_{M/2}$, respectively. Mathematically they are defined as

$$\mathcal{A}_{M/4} = \mathcal{A}_{\rm onset} - t_{M/4} \tag{5}$$

$$\text{with } 1/4 \times \mathcal{M}_{\star,\rm obs} = \int_0^{t_{\mathcal{M}/4}} SFH(t)\mathrm{d}t,$$

$$\mathcal{A}_{M/2} = \mathcal{A}_{\rm onset} - t_{M/2} \tag{6}$$

$$\text{with } 1/2 \times \mathcal{M}_{\star,\rm obs} = \int_0^{t_{\mathcal{M}/2}} SFH(t)\mathrm{d}t,$$

where $t_{M/4}$ and $t_{M/2}$ are the times where the quarter and the half of the stellar mass are built up and are represented in blue and red in Fig. 2, respectively.

Another common definition of age is the mass weighted age (MW for mass weighted age), used for example in cosmological simulations. This age is related to the relative importance of each population of stars that composes the galaxy. Then, MWA is defined as the sum of the age of each of the population of stars in the galaxy weighted by their contribution to the stellar mass of the galaxy,

$$\mathcal{A}_{\rm MW} = \sum_{n=1}^{N} \frac{\mathcal{A}_{{\rm onset},i}\mathcal{M}_i}{\mathcal{M}_{\rm obs}}, \tag{7}$$

where $\mathcal{A}_{{\rm onset},i}$, $\mathcal{M}_i$ and $\mathcal{M}_{\star,\rm obs}$ are the age of the ith population, its mass, and the observed stellar mass of the galaxy. In practice the computation of the MWA is performed considering the SFH from the onset of the star formation truncated at the $\mathcal{A}_{\rm onset}$ age. This truncated SFH is then divided in equal time intervals representing individual star populations. We tested splitting the SFH in $N_{\rm bin}$ = 10, 100, and 1000 intervals (the length of each interval is then given by $\mathcal{A}_{\rm onset}/N_{\rm bin}$) and found that above a binning of 100 (hence for a typical bin of ~10 Myr) the mass weighted ages remain unchanged and we adopt this binning for this study.

Each age definition depends on the SFH that is used to compute $M_{\star}$ and its timescale parameter $\tau_{\rm SFH}$ as shown in Fig. 3. The larger the timescale, the larger the differences between different age definitions: as $\tau_{\rm SFH}$ increases, the width of the star formation burst is getting broader and the time to accumulate one-half or one-quarter of the stellar mass is getting much larger. The quarter-mass age is evidently the closest to $\mathcal{A}_{\rm onset}$ while the half-mass age shows a larger difference with $\mathcal{A}_{\rm onset}$ since $t_{\mathcal{M}/2} > t_{\mathcal{M}/4}$. Figure 3 also shows that the half-mass age is very close to the mass-weighted age. Therefore, we do not use the half-mass age in the following but study only the mass weighted age, the quarter-mass age and the age corresponding to the onset of the SFH, $\mathcal{A}_{\rm onset}$.

The robustness against the spectro-photometric fitting with GOSSIP+ of each of these definitions ($\mathcal{A}_{\rm onset}$, $\mathcal{A}_{\mathcal{M}/4}$ and $\mathcal{A}_{\rm MW}$) is presented in the next section.

5. Measuring ages and degeneracies at $z > 2$: simulations

For the relatively young stellar populations expected in galaxies at high redshifts it is important to identify possible degeneracies between age measurements and other parameters.

At low redshifts, the main degeneracies linked to the age parameter are the age-metallicity degeneracy and the age-dust degeneracy (see Sect. 5.5). At $z > 2$ degeneracies are expected to be less severe than at lower redshifts as the age of a galaxy is strongly restricted to be smaller than the age of the Universe at the observed redshift. Assuming the parameters of the cosmological world model, the ages of galaxies are limited to a maximum of ∼0.8 Gyr at $z = 6$ and ∼3.3 Gyr at $z = 2.0$, which are the redshift boundaries of our sample.

In the following we perform extensive simulations of galaxies with similar properties to the VUDS sample, and use spectroscopy and photometry jointly when performing model fitting. We analyse possible degeneracies that could affect age estimations and conclude with the accuracy of age measurements at these redshifts.

5.1. Simulations: 180 000 galaxies representative of VUDS properties

We produce a large set of mock galaxies to test the accuracy and robustness of age measurements of galaxies at $z > 2$. Then, we run the SED fitting with GOSSIP+ to measure their ages and compare these measurements with input ages in the simulation.

We simulate both broadband photometry and the associated spectra covering the same wavelength range as VUDS observations ($3650 < \lambda < 9350$ Å). We create a sample of mock galaxies with the same properties as our VUDS sample ($2 < z < 6.5$, $22 < i_{AB} < 25$ and signal-to-noise ratio $2 < S/N < 25$), i.e. with the same redshift, magnitude i_{AB} and S/N combinations. For each of these VUDS combinations, 50 mock galaxies are created, which each have a different galaxy synthetic model. The final simulated sample contains 180 000 objects.

The synthetic model is randomly chosen in a BC03 library composed of 10 000 templates (with 400 Å $< \lambda <$ 60 000 Å rest-frame). In this library, the dust extinction is computed from the Calzetti law and $E(B-V)$ values range from 0.0 to 0.5 (5 equal steps). The SFH is an exponentially delayed model that contains both increasing and declining SFH sections with a timescale $\tau_{\rm SFR}$ that ranges from 0.1 to 5.0 Gyr (in 10 steps); the timescale $\tau_{\rm SFR}$ represents the time between the onset of the star formation and the peak of the SFH. We use the Chabrier IMF Chabrier (2003). Ages are allowed to be in the interval [0.05; 4.0] Gyr (0.05 Gyr steps between 0.05 Gyr and 2.5 Gyr and 0.25 above), and the selected model has to be younger than the age of the Universe at the redshift of the galaxy. The metallicity is in the range [$Z = 0.004$; $Z = 0.05$] (4 values, $Z_\odot = 0.02$). Finally, an IGM template is chosen randomly among seven possibilities at the considered redshift (see Sect. 3 and Thomas et al. 2017 for details).

After the selection of a triplet (z, i_{AB}, S/N) we randomly select a synthetic model, redshifted at a redshift z. Then, the IGM transmission is applied. The model is normalized to i_{AB} and resampled to the same spectral resolution as the observations. The wavelength of the final mock spectra ranges from 3600 Å to 9500 Å with a $\Delta\lambda = 5.35$ Å to mimic the VUDS spectra. Poisson noise is then added to the data as the square root of the simulated flux. Finally, the magnitudes are computed by convolving the model with the transmission of several photometric filters: *ugriz* from the Megacam camera, *JHK* from the WIRCAM camera and two *Spitzer* IRAC band at 3.6 and 4.5 μm.

Table 2. Parameter space used during the fit of the mock galaxies with both BC03 and M05 models.

Parameter	BC03	M05
IMF	Chabrier	Chabrier
Metallicity [Z]	0.004, 0.008, 0.02, 0.05	0.001, 0.01, 0.02, 0.04
$\tau_{\rm SFH}$	0.1 to 5.0 Gyr (10 steps)	
$E(B-V)$	0.0, 0.1, 0.2, 0.3, 0.4, 0.5	
Ages (Gyr)	0.05 to 4.0 (30 steps)	
IGM	7 templates per redshift	

5.2. Fitting process

Once the simulations are created we run GOSSIP+ three times: on spectra alone, on the photometry alone, and on both combined. For this process we use two different fitting libraries: one with the BC03 population synthesis model and the other with M05. The parameter space used for the two libraries is presented in Table 2.

This provides a large statistical basis to examine the robustness of the age computation and various possible associated degeneracies. We present the results in the next two subsections, with first a visual inspection of which degeneracies are present, followed by a quantitative analysis of age estimates.

5.3. Stellar mass and star formation rate

Before focusing on age measurements in the next section we discuss the reliability of GOSSIP+ in measuring $M_\star$ and SFR from the combined fitting of spectra and broadband photometry. We use the simulations as defined above and we compare the $M_\star$ and SFR from the three types of fit, on the set of simulated broadband photometric magnitudes alone, on simulated spectra only, and on the combined photometry and spectroscopy datasets.

Figures 4 and 5 compares $M_\star$ and SFR obtained from each type of fit with the input values of the simulation. In both figures, the top panel shows the density map of combined data with its 1σ contour, while the bottom panel, shows the evolution of $\Delta\mathrm{Param} = \mathrm{Param}_{\rm in} - \mathrm{Param}_{\rm out}$ as a function of the parameter Param for the three types of fits. Table 3 gives the measurement of ΔParam for each parameter combination and fit type for the full simulation with the evolution of this quantity as a function of redshift.

5.3.1. Stellar mass, $M_\star$

The comparison of $M_\star$ computed with GOSSIP+ to the simulated mass is shown in Fig. 4 and Table 3 for the three cases considered here.

The spectroscopic data covers the UV rest frame as VUDS recovers $M_\star$ on average, but with a relatively poor accuracy. The difference between the input and measured values at $\log_{10} M_\star > 10.2$ is $\Delta\mathcal{M}_\star \sim 0.3$ dex on average, while at $\log_{10} M_\star < 10$ it is ∼−0.20 dex. The median absolute deviation is of ∼0.4 dex. With UV rest frame only the spectroscopic data obviously do not constrain the near-infrared (NIR), which is most sensitive to the stellar mass. Therefore, the resulting errors on $M_\star$ are large, as shown by the evolution of $\Delta\mathcal{M}$ particularly for high mass galaxies. This is because high mass galaxies generally host older and then redder stellar populations that are difficult to constrain

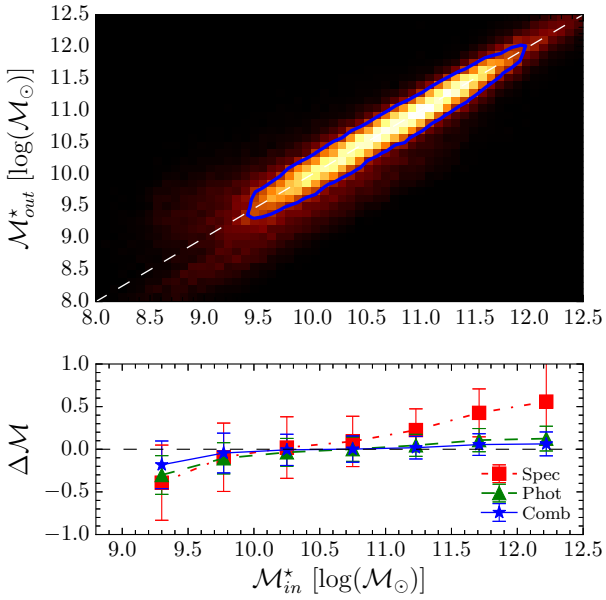

Fig. 4. Comparison of the stellar mass measurements $M^{\star}_{\rm out}$ obtained using GOSSIP+ on the 180 000 simulated $M^{\star}_{\rm in}$ sample with the input values. *Top panel*: the density plot represents the measurement of $M_{\star}$ from the combined fit versus the input $M_{\star}$. The highest density is represented by the brightest colour, while the lowest density is represented in black. The blue line represents the 1σ contour of the distribution. The combined fit is able to find the correct $M_{\star}$ at any mass in the simulated range. *Bottom panel*: evolution of the quantity $\Delta\mathcal{M} = \mathcal{M}^{\star}_{\rm in} - \mathcal{M}^{\star}_{\rm out}$ over the stellar mass range. The points in red, green, and blue represent the fit on the spectroscopy only, photometry only and on the combined data, respectively. The combined fit leads to the most accurate $M_{\star}$ computation with $\Delta M_{\star} < 0.1$ dex.

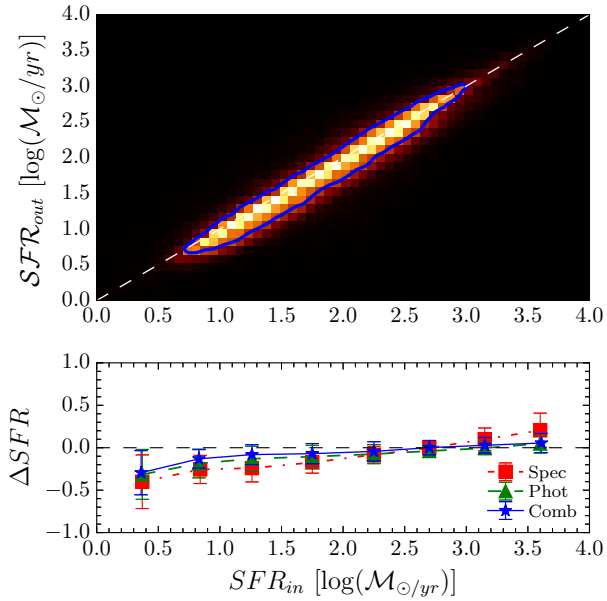

Fig. 5. Comparison of SFR recovered using SED fitting with GOSSIP+ to the simulated dataset of 180 000 galaxies. Colours and line types are as for Fig. 4

based only on UV rest-frame spectroscopy. As expected, the UV rest-frame spectroscopy alone is not the right observable to constrain the stellar mass.

The fit on the photometric data alone provides $M_{\star}$ measurements with good accuracy. This is widely used in the literature, for example from photometric redshift codes (Ilbert et al. 2013, and reference therein). We note that the accuracy in measuring $M_{\star}$ slightly decreases with redshift, going from –0.09 dex at $z \sim 2.4$ to −0.21 dex at $z > 4.25$. This can be explained by the fact that at high redshifts, the NIR photometry corresponds to the UV rest-frame part of the data, then imposing fewer constraints on the older stellar populations tracing the stellar mass. GOSSIP+ is able to compute $M_{\star}$ from a photometric SED as reliably as most standard codes in the literature.

The fit on the combined spectroscopic and photometric data shows an excellent accuracy in $M_{\star}$ measurements at all redshifts. At most redshifts explored here the accuracy in $M_{\star}$ is typically ∼0.02 dex, which is almost five times better than with photometry alone at $z \sim 2$ and always twice better at any higher redshift. This can be explained by the fact that these data combine the IR constraints from the photometry with the UV rest frame from spectroscopy leading to a finer separation of model templates in the parameter space explored. The evolution of $\Delta\mathcal{M}$ from the fit on combined data shows that for low mass galaxies, the addition of the photometry allows us to limit an overestimation of the mass, while for high mass galaxies we observe the opposite behaviour.

This analysis of the stellar mass shows that the combined fit of broadband photometric data covering from the u band to 4.5 μm with UV rest-frame spectroscopic data leads to a significant improvement in the estimation of this important physical parameter.

5.3.2. Star formation rate

Figure 5 and Table 3 show the estimation of the SFR against the input values of our 180 000 simulated galaxy sample. For this parameter the three types of fit give closer results than for $M_{\star}$.

The fit on the spectroscopic data alone is the least accurate fit with an accuracy of 0.2 dex on average. The accuracy improves at all redshifts when using the photometric data alone. The estimation of the SFR with the fit on the spectroscopy and photometry combined is particularly efficient up to $z \sim 4.25$, with an average $\Delta\mathcal{SFR} \sim -0.15$ dex and a median absolute dispersion of −0.13. The combination of UV rest-frame spectroscopic data, which is sensitive to recent star formation, with optical and NIR broad-band photometric data, which is sensitive to star formation accumulated over a long time baseline, is particularly efficient in recovering the SFR. Shifting to higher redshifts, the accuracy become worse at ∼−0.18 dex since half of the observed spectrum is composed by the region below the Lyman limit, which does not bring constraints on the SFR; but the combined fit recovers the SFR more accurately than the fit on spectroscopy or photometry taken separately.

5.4. Estimation of galaxy ages, quantitative analysis

In this section we analyse our simulations to assess the robustness of age measurement for the three types of fit, following a similar methodology as for $M_{\star}$ and SFR in the previous sections. We analyse results for the three definitions of age presented in Sect. 4, $\mathcal{A}_{\rm onset}$, $\mathcal{A}_{\mathcal{M}/4}$, and $\mathcal{A}_{\rm MW}$. We therefore create in both input (output) the mass-weighted age and half-mass age from the input (output) SFH and $\mathcal{A}_{\rm onset}$ as described in Sect. 4.

Results are presented from Figs. 6 to 8 and in Table 4. In each figure, we present the $\Delta\mathcal{A}$ in four $\mathcal{A}_{\rm in}$ bins and for the three age definitions.

Table 3. Estimation of the stellar mass and SFR from the three types of fit.

Parameter/Fit	Full simulation	$2.0 \leq z < 2.75$	$2.75 \leq z < 3.5$	$3.5 \leq z < 4.25$	$z \geq 4.25$
Mass/SPEC	0.18 ± 0.39	0.21 ± 0.39	0.20 ± 0.39	0.19 ± 0.40	0.16 ± 0.49
Mass/MAGS	−0.09 ± 0.18	−0.05 ± 0.16	−0.11 ± 0.19	−0.11 ± 0.21	−0.21 ± 0.28
Mass/COMB	**−0.02 ± 0.19**	**0.01 ± 0.16**	**−0.03 ± 0.20**	**−0.05 ±0.21**	**−0.12 ± 0.27**
SFR/SPEC	−0.20 ± 0.13	−0.20 ± 0.12	−0.20 ± 0.14	−0.18 ± 0.16	−0.14 ± 0.26
SFR/MAGS	−0.19 ± 0.14	−0.16 ± 0.13	−0.17 ± 0.12	−0.13 ± 0.13	−0.20 ± 0.19
SFR/COMB	**−0.13 ± 0.12**	**−0.14 ± 0.12**	**−0.13 ± 0.13**	**−0.10 ± 0.11**	**−0.18 ± 0.16**

Notes. For each combination of parameter and fit type, we give the mean of the quantity $\Delta\text{Param} = \text{Param}_{\text{in}} - \text{Param}_{\text{out}}$ and the associated median absolute deviation.

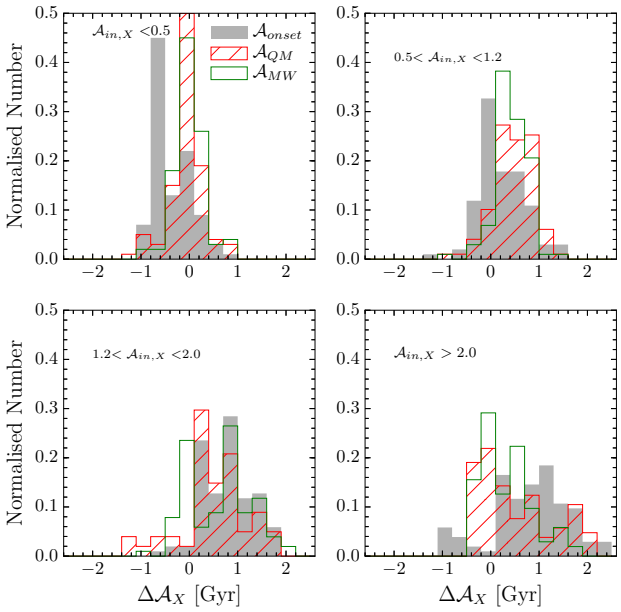

Fig. 6. Comparison of the $\Delta\mathcal{A}$ for the three definitions of ages from the fit on the spectroscopy only. In grey, we show the $\mathcal{A}_{\text{onset}}$ definition in green $\mathcal{A}_{\text{MW}}$ and in red $\mathcal{A}_{M/4}$. We show three input age bins. *From top left to bottom right*, $\mathcal{A}_{\text{in}} < 0.5$, $0.5 < \mathcal{A}_{\text{in}} < 1.2$, $1.2 < \mathcal{A}_{\text{in}} < 2.0$, and $\mathcal{A}_{\text{in}} > 2.0$. All the ages are given in Gyr.

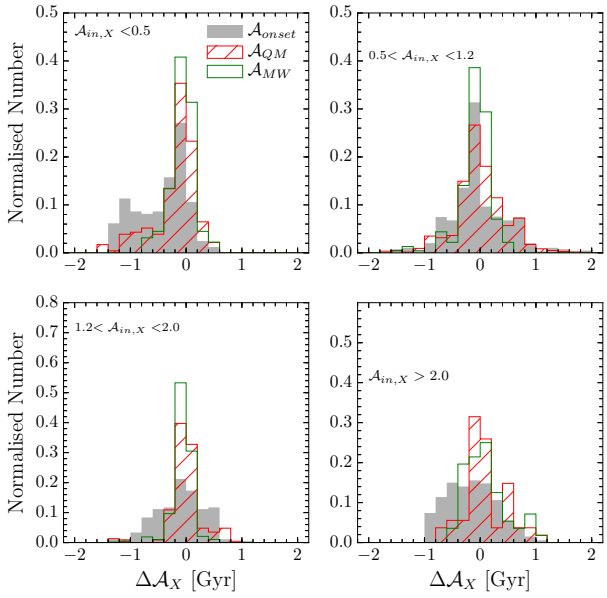

Fig. 7. Same as 6 but only for the fit on the photometry.

Not surprisingly, the fit on the spectroscopy alone provides the worst estimate of galaxy ages. Using $\mathcal{A}_{\text{onset}}$, the difference between the input and the measured ages on the whole simulation is around 0.3 Gyr and the median absolute deviation is of 0.50 Gyr. The age is generally underestimated. The accuracy is improving with redshift as expected since the upper age limit is given by the age of the Universe and therefore decreases with redshift. This is confirmed from the evolution of ΔAge as presented in Fig. 6 as a function of the input age. When the simulated age is higher than 2.0 Gyr (only at redshift $z \lesssim 3$) ΔAge is 1.0 Gyr, but ΔAge ∼ 0.25 when the input age is smaller than 1.2 Gyr (corresponding to a redshift $z \gtrsim 4.5$).

For $\mathcal{A}_{M/4}$ we note that ΔAge is slightly improved with respect to the $\mathcal{A}_{\text{onset}}$ definition. This is particularly the case for the highest input age for which the accuracy reaches 0.5 Gyr against 0.8 for $\mathcal{A}_{\text{onset}}$. This is similar for the median absolute deviation. This behaviour is expected since ages are artificially reduced in this definition, leading to lower value of ΔAge and median absolute deviations. Finally, $\mathcal{A}_{\text{MW}}$ gives a better estimation of the age with ΔAge ∼ 0.16 Gyr for the full simulation. As shown in Sect. 4, this definition gives the smallest age, leading to a smaller ΔAge.

Using the photometry alone (Fig. 7) improves the estimation of the age compared to only using of the spectroscopic data. The fit is better constrained over a large wavelength range by the full set of photometric points; the NIR and IR regions provide constraints on the age of the oldest stars. The difference in age for $\mathcal{A}_{\text{onset}}$ is typically ΔAge ∼ 0.15 Gyr, with a median absolute deviation of 0.33 Gyr. The age parameter with the lowest ΔAge is the mass-weighted-age. As described in the previous paragraph, the artificial reduction of the age with this definition leads to an artificial reduction of the results of the simulation. For this parameter ΔAge = −0.08 Gyr.

Combining photometry and spectroscopy clearly provides the best age estimates for any of the three age definitions considered (Fig. 8). A strong constraint on young massive stars is imposed by the UV rest-frame spectroscopic data and the optical and NIR photometric data constrains the population of oldest stars. For $\mathcal{A}_{\text{onset}}$ the age difference between the simulated input and the measurements is better than ΔAge ∼ 0.1 Gyr for the whole simulation with a median absolute deviation of 0.32 Gyr. The age difference between input and output is the smallest for the mass weighted age $\mathcal{A}_{\text{MW}}$ in agreement with the findings of Pforr et al. (2012) at all redshifts; the average ΔAge is 0.03 Gyr with a median absolute deviation of 0.15 Gyr. This level of accuracy in measuring galaxy ages at these high redshifts is unprecedented.

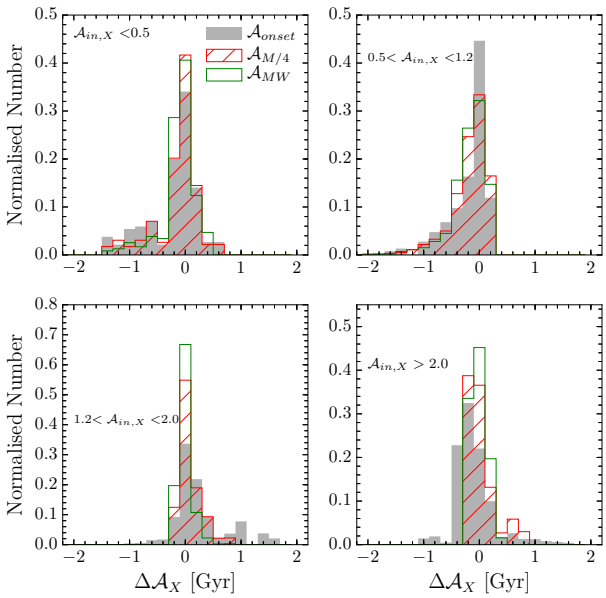

Fig. 8. Same as 6 but for the combined fit of spectroscopy and photometry.

To summarize, this analysis demonstrates that the combined use, on high redshift galaxies ($z > 2$), of high quality UV rest-frame spectra and deep multiwavelength photometry, including rest-frame optical and NIR data, enables significantly improve galaxy age measurements. This combination of different information allows the increase of accuracy by a factor 2 for the $\mathcal{A}_{\rm onset}$ definition with respect to the classical SED fitting and by factor more than 3 with respect to the fit on the spectroscopic data. For the parameter that gives the smallest age, $\mathcal{A}_{\rm MW}$, the combined fit shows an accuracy almost three times better than the photometric fit and more than five times better than the spectroscopic fit. Age measurements using this method show an uncertainty of ∼9−10% (1σ) comparable to the uncertainty in $M_\star$ measurements.

With this level of accuracy it becomes possible to investigate age-related properties to get a complementary point of view on galaxy evolution compared to using a more restricted analysis of $M_\star$ or SFR alone.

5.5. Studying age-related degeneracies

5.5.1. Method

Degeneracies are the major problem when computing galaxy ages. We developed a method based on PDFs to study the influence of both the age limit imposed by the age of the Universe and the addition of spectroscopy to photometry. We created 2D PDFs, called probability density maps (PDM), for any parameter in combination to age; we study here only $\mathcal{A}_{\rm onset}$, but the conclusions derived for this age definition are qualitatively similar for the other age definitions.

A PDM represents a probability density in 2D space and clearly highlights how the PDFs of the two parameters relate to each other and the possible degeneracies that are at play. We create PDMs as follows. For two parameters (δ_1; δ_2), which can take N_1 and N_2 values, the 2D space is built with $N_1 \times N_2$ points. During the fit, a χ^2 value is assigned to each model template in the library and is used to derive a probability for the goodness of fit (see Eq. (2)). Therefore, the probability P of a given point in the PDM ($\delta_{1,i}$; $\delta_{2,j}$ for example) corresponds to the sum of the probability p of the template that satisfies $\delta 1 = \delta_{1,j}$ and $\delta_2 = \delta_{2;j}$, i.e.

$$P(\delta_1;\delta_2) = \sum_k p[\text{template}_k(\delta_1 = \delta_{1,j} : \delta_2 = \delta_{2;j})]. \quad (8)$$

Since the two-parameter space of the different PDMs is not continuous, the PDMs are linearly interpolated between parameter values (see next section for numerous examples). To compare the quality of the fits we analyse three quantities as follows:

- $N_{\rm peak}$: the number of peaks included in the 1σ contour. This indicates remaining degeneracies as each peak could host the right parameter values.
- The position of the input parameters inside or outside the 1σ contour. Obviously if the input parameters are in the 1σ contour this indicates that the fit has properly converged towards the correct value.
- The ratio of the area of the 1σ contour between two different fits (spectrum only, photometry only, or the two combined). Ratios smaller than one indicate an improvement of the fit quality because the parameter space allowed by the fit is reduced. Since maps are constructed on the same grid, the area of the 1σ contour is calculated as the number of points (parameter pairs) that fall in this contour.

We study in the next two subsections the age–dust and age–metallicity degeneracies using the PDM formalism.

5.5.2. Age-dust degeneracy

One of the main difficulties in measuring ages is the age-dust degeneracy. Dust strongly attenuates flux at $\lambda \lesssim 4500$ Å (rest frame), re-emitting it in the far-infrared. A strongly SFG galaxy with a large amount of dust has a spectrum where the blue part is strongly attenuated and the bright red part may mimic old stellar populations and hence an old galaxy. From the PDMs in our simulations, we find that fitting spectroscopy or photometry alone is not able to reproduce the right pair of age and dust parameters. Both fit modes exhibit rather large 1σ contours in the PDMs enclosing young age values with high dust content as well as older ages with lower $E(B-V)$.

Three representative examples are presented in Fig. 9 with the PDMs of three simulated galaxies at $z = 2.20$, $z = 3.15$, and $z = 4.26$, with ($E(B-V)$, Age) pairs of (0.2, 0.40), (0.4, 0.32) and (0.2, 0.4), respectively. The shape of the 1σ contours are clearly the result of the age-dust degeneracy. Indeed, to keep a similar template shape, an increase of the dust extinction has to come with a lower age.

As discussed in the previous section, the spectroscopy only is not able to measure the age with a good accuracy. Since, age and dust are closely linked together, the accuracy on the dust estimate is not so high either. We clearly see that the 1σ contour at low age corresponds to higher values of $E(B-V)$ and the contours at ages higher than the simulated age also correspond to smaller dust extinction values. This is the result of the age-dust degeneracy. Moreover, over the three examples presented here only one clearly shows an input pair in the 1σ contour, while the two others are at the limit.

The presence of strong secondary peaks in the PDMs built from the fit of the photometric data alone confirms that the photometry alone is not able to resolve the age-dust degeneracy. The map based on the photometry only shows that the spectral range is large enough to be sensitive to dust extinction variations but remains a poor constraint in the blue part of the spectral domain

Table 4. Estimation of galaxy ages for different age definitions from the three types of fit.

Age definition/Fit	Full sample	$2 < z \leq 2.75$	$2.75 < z \leq 3.5$	$3.5 < z \leq 4.25$	$z \geq 4.25$
$\mathcal{A}_{\rm onset}$/SPEC	0.28 ± 0.50	0.33 ± 0.61	0.28 ± 0.50	0.25 ± 0.40	0.23 ± 0.34
$\mathcal{A}_{\rm onset}$/MAGS	0.14 ± 0.33	0.18 ± 0.37	0.14 ± 0.32	0.12 ± 0.29	0.10 ± 0.28
$\mathcal{A}_{\rm onset}$/COMB	**0.08 ± 0.32**	**0.14 ± 0.34**	**0.07 ± 0.31**	**0.05 ± 0.27**	**0.04 ± 0.29**
$\mathcal{A}_{\mathcal{M}/4}$/SPEC	0.25 ± 0.32	0.29 ± 0.39	0.25 ± 0.33	0.18 ± 0.25	0.15 ± 0.23
$\mathcal{A}_{\mathcal{M}/4}$/MAGS	−0.12 ± 0.22	−0.16 ± 0.24	−0.12 ± 0.21	−0.13 ± 0.17	−0.13 ± 0.18
$\mathcal{A}_{\mathcal{M}/4}$/COMB	**–0.04 ± 0.20**	**–0.09 ± 0.19**	**–0.005 ± 0.18**	**0.00 ± 0.19**	**0.04 ± 0.18**
$\mathcal{A}_{\rm MW}$/SPEC	0.16 ± 0.23	0.19 ± 0.29	0.17 ± 0.23	0.13 ± 0.18	0.11 ± 0.16
$\mathcal{A}_{\rm MW}$/MAGS	−0.08 ± 0.17	−0.10 ± 0.20	−0.07 ± 0.17	−0.07 ± 0.13	−0.07 ± 0.14
$\mathcal{A}_{\rm MW}$/COMB	**0.03 ± 0.15**	**–0.04 ± 0.16**	**0.03 ± 0.15**	**0.02 ± 0.13**	**0.04 ± 0.13**

Notes. For each combination of parameter and fit type, we give the mean of the quantity $\Delta\text{Param} = \text{Param}_{\rm in} - \text{Param}_{\rm out}$ and the associated median absolute deviation.

because of the limited spectral resolution of the photometry (typically one point per ∼1000 Å).

Finally, the PDMs based on the combined fit of UV rest-frame spectroscopy and photometry shows only one peak with a much smaller 1σ contour, as shown in the examples presented in Fig. 9. Solutions with lower age and higher dust extinction or with higher dust extinction and lower ages are mostly ruled out. It is remarkable that the combination of photometry and spectroscopy data allows us to break the age-dust degeneracy and to produce a pronounced probability peak allowing to recover the simulated (Age, dust) pair. Rest-frame UV spectra provide a stronger constraint than the photometry in a wavelength domain where the effect of dust extinction is strongly varying.

We turn now into the analysis of the three properties described in the previous section to study the age-dust degeneracy over the entire simulated sample. Table 5 shows two of these properties for the age-dust PDMs. We give the mean Npeak and the fraction of simulations in which the input couple $(\text{Age} - E(B-V))_{\rm in}$ is contained in the 1σ contour. We compare here only the two best-fit modes, photometry, and the combined photometry plus spectroscopy datasets.

This table shows that in the combined fit case, the 1σ contour of the PDM contain the input couple more than 80% of the time. This number drops below 70% when only photometry is used. The combined fit is therefore the fit mode that is more likely to retrieve the right parameters. Then, the median number of peaks encompassed by the 1σ contour is of 1 for the combined fit and of 2 for the pure photometric fitting. This means that when we analyse the PDMs of the combined fit the output couple is in general the only one possible while the photometric fitting gives another possible peak due to the degeneracy.

Finally, we compute the area limited by the 1σ contours of the individual PDMs. We compare the size of the 1σ area between the fit on the photometry only and the fit on the combined data. Figure 10 shows the ratio between the size of the 1σ area of the PDMs on the photometry alone and on the combined data. We find that on average, 1σ contours based on the photometry cover 3.9 times larger area than 1σ contours based on the combined fit.

Our simulations therefore demonstrate that, at high redshifts $z > 2$ when the age of the Universe is significantly reduced, the combination of spectroscopy and photometry generally results in probability distributions maps with a single peak including the simulated pair of parameters. This probability peak is significantly narrower than when using photometry only. We conclude that the age-dust degeneracy is significantly reduced when using the combined spectroscopic and photometric data.

Table 5. Global properties of the age-dust PDMs in our simulations.

Property	Photometry	Combined
$(\text{Age} - E(B-V))_{\rm in} \in 1\sigma$	66%	81%
Median $N_{\rm peak}$	2	1

5.5.3. Age-metallicity degeneracy

We now study the age-metallicity relation in the same way as for the age-dust degeneracy. The degeneracy between age and metallicity has long been recognized as a limitation in computing ages (Worthey 1994). As a stellar population in a galaxy evolves the metallicity increases steadily, therefore to reproduce the SED of a galaxy it is possible to compensate the general reddening of the SED from metallicity enhancement by a decrease of the age of the galaxy and still get a very similar SED. Figure 11 shows the individual maps for three representative simulated galaxy at $z = 2.25$, $z = 3.45$, and $z = 4.47$. The age-metallicity pair of these mock galaxies are (0.7 $Z_\odot$), (0.7, 2.5 $Z_\odot$), and (0.1, $Z_\odot/5$). The colour coding is the same as Fig. 9.

The PDMs resulting from the fit on the spectroscopy show that the maximum of the map is on average 0.7 Gyr away from the input value and the 1σ contour spans all the age range possible and hence the resulting parameters are poorly constrained. The spectroscopy alone shows that it is almost impossible to put a constraint on the age. The peak of the PDMs are always at a different place as the input pair and both parameter are not retrieved simultaneously.

The photometry alone does not allow for the recovery of the input parameters either. The PDMs show that ages are closer to the input, but the metallicity does not correspond to the input value. Moreover, we observe that the probability for other metallicities values remains high and hence the discrimination power is limited. While the peak of the PDM are in general close in terms of age, the metallicity is always different as the input. The 1σ contours are very extended, signaling poor constraints on the final parameters.

Finally the combined fit of the photometry and spectroscopy is the only mode that correctly recovers the input parameters. The 1σ contours are well centred on the input values and narrower than those on the fit of the photometry only.

The analysis of the age-metallicity PDMs is summarized in Table 6 and Fig. 12. The median number of peaks in the contour produced by the combined fit is unity while it increases to 2 when fitting only the photometry. In the case of the combined fit we retrieve the input couple in the contour for 78% of the cases, which is a small improvement compared to the photometry only

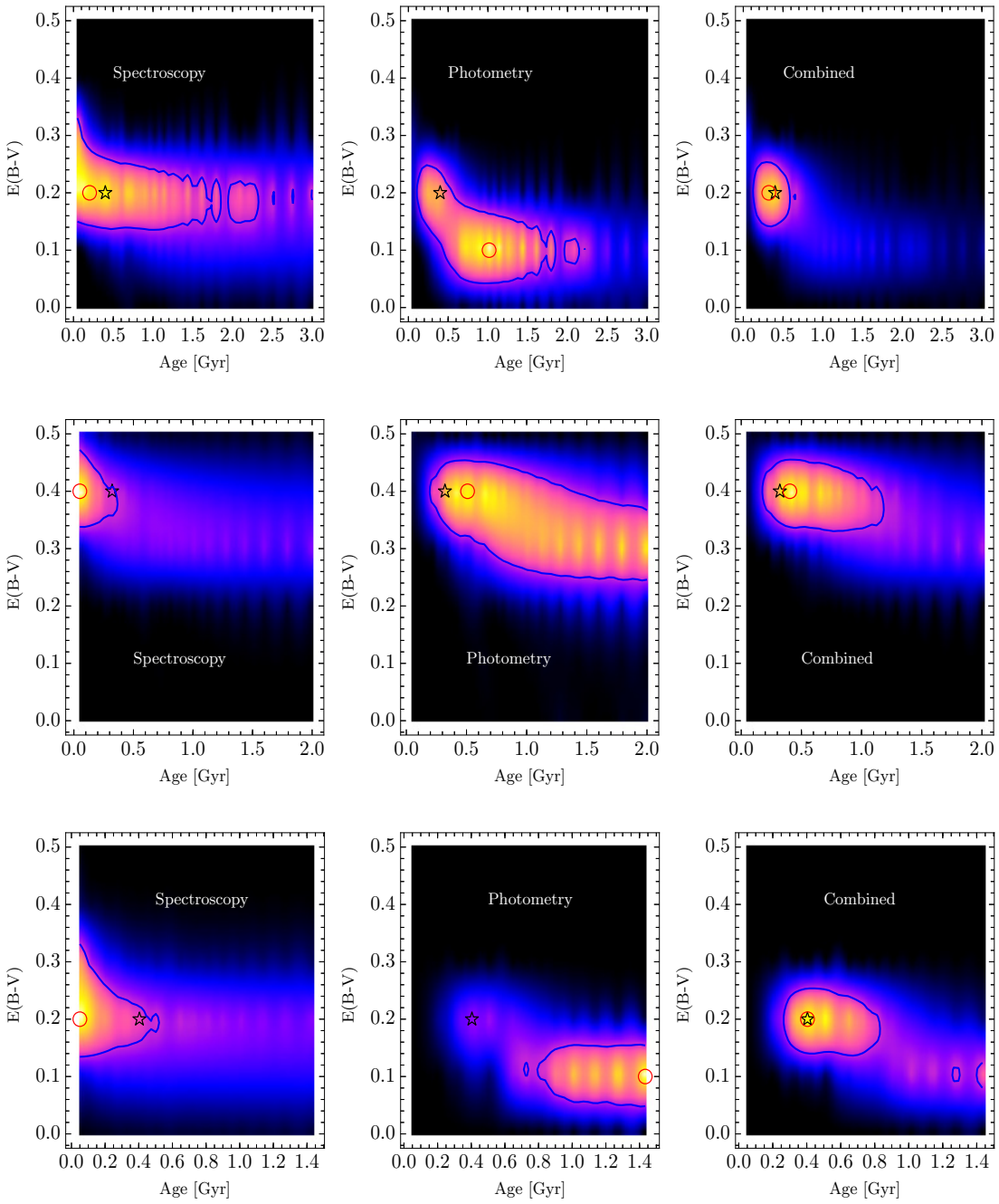

Fig. 9. Nine examples of probability density maps in our simulation. *From left to right*: fitting the photometry only, the spectroscopy only, and the combined fit. *From top to bottom*: simulated galaxy at $z = 2.20$ with a dust extinction of 0.2 and age of 0.404 Gyr, simulated galaxy at $z = 3.15$ with $E(B-V) = 0.4$ and age = 0.321 Gyr, and a simulated galaxy at $z = 4.26$, $E(B-V) = 0.2$ and age = 0.4 Gyr. In each map, the red circle shows the maximum of probability, the 1σ contour is shown by the blue line and the input couple ($E(B-V)$, age) is shown by the black star. The age corresponds to the $\mathcal{A}_{\rm onset}$ definition. In this example, the combined fit is the only fit that is able to reproduce the input values and also the only fit in which the 1σ is the smallest.

(75%). However the ratio between the area of the 1σ contour from the photometry only and the combined fit is in on average ~13 (Fig. 12). This analysis indicates that the combined fit is able to retrieve the right simulated parameter couple with a significantly better significance than when using the photometry alone.

It is evident but well worth saying that the accuracy of physical parameters measurements will keep improving as more and more of the spectral domain is covered with spectroscopy. The main limitation of the SED-fitting method will then likely remain the systematic uncertainties linked to the simplifying assumptions used to produce model spectra representative of the true galaxy population.

6. The age distribution of galaxies with $2 \leq z \leq 6.5$ in VUDS

6.1. Fitting process

We fit the 3597 VUDS galaxies with the GOSSIP+ software (Sect. 3) using the combined fit of the available broad-band photometry and the spectroscopy. We use two different

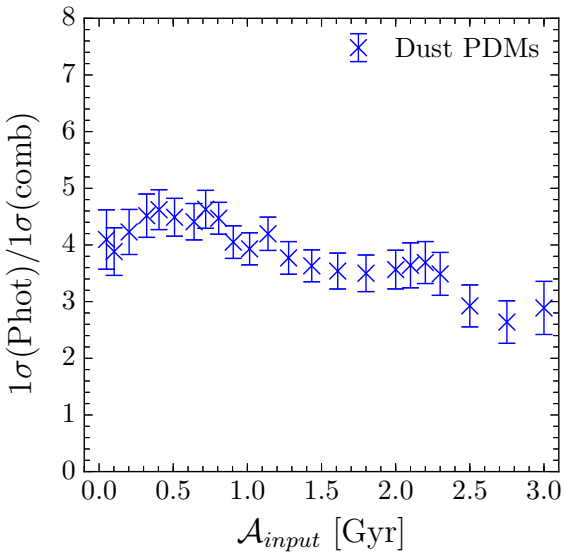

Fig. 10. Comparison of the 1σ area from the fit on the photometry and the fit on the combined data as a function of the input $\mathcal{A}_{\rm onset}$ for the age-dust relation. At any age, the 1σ(phot) area based on the photometry is on average 4 times bigger than the 1σ(comb) area based on the combined fit. The error bars are computed from the median absolute deviation.

population synthesis models: BC03 (Bruzual & Charlot 2003); these libraries are presented in Table 7.

As an example of the fitting result we present the fit of ten galaxies in our sample based on both photometry and spectroscopy in Figs. 13 and 14, covering the redshift range of this study. The agreement between the data and the best-fit model is good in all cases, representative of the whole sample.

An example of the effect of combining photometry and spectroscopy with a comparison of each type of fit (spectroscopy only, photometry only, and combined data) is presented in Fig. 15. The fit on the spectroscopy data represents the data in the wavelength range of the spectroscopic data very well, but the redder part ($\lambda > 9500$ Å) is not well constrained, by definition, and the best-fit model is therefore far from the photometry in this region. The best-fit model based on the photometric data does indeed fit the photometry very well. Nevertheless, the wavelength range $3500 < \lambda < 9500$ Å (corresponding to the range of the spectroscopy data) is poorly constrained and the Lyman$-\alpha$ region is not well reproduced because the sampling by photometric data is quite coarse in this domain. Finally, this example shows that the fit on the combined data is the best able to reproduce both the infrared photometry and the UV-rest-frame spectroscopy.

6.2. Stellar mass and SFR

As the quarter-mass-age and mass-weighted-age measurements depend strongly on mass estimation, we first present the $M_\star$ and SFR distributions. The stellar mass distribution is presented in the bottom panel of Fig. 16 and the SFR is shown in the upper panel. Stellar masses range from $\log M_\star \sim 8.3$ to $\sim$11.9 with a median $\log M_\star = 9.8$. The SFR ranges from $\log SFR \sim -0.25$ to $\sim$3.50, with a median $\log SFR = 1.45$. The $M_\star$ and SFR distributions in the sample used in this paper (Sect. 3.4) are similar to the distributions of all VUDS galaxies and hence our selection does not produce any particular population bias.

6.3. Ages from the combined fit of spectra and photometry

The age distribution of VUDS galaxies is presented in the top panel of Fig. 17 for the three definitions of age $\mathcal{A}_{\rm onset}$, $\mathcal{A}_{\rm MW}$, and $\mathcal{A}_{\mathcal{M}/4}$ (see Sect. 4).

For $\mathcal{A}_{\rm onset}$, the ages range from 0.05 Gyr (the lowest age allowed in the fit) to ages higher than 2 Gyr. The former represents a very low fraction of our galaxies around ~1%. Around 10% of galaxies have ages in the range 1 Gyr $< \mathcal{A}_{\rm onset} <$ 2 Gyr, while ~89% are younger than 1 Gyr in age. The number of galaxies with a lower 1-sigma limit of 0.05 Gyr, which is the youngest age in our library, is of 13%; our choice of limiting the age to 0.05 Gyr, therefore, does not affect the vast majority of our galaxies. The $\mathcal{A}_{\mathcal{M}/4}$ ages are younger than 2 Gyr with ~5% in the range 1 Gyr $< \mathcal{A}_{\rm onset} <$ 2 Gyr, as the rest of the sample is younger than 1 Gyr. Using the $\mathcal{A}_{\rm MW}$ definition, ~2% of galaxies are between 1 and 2 Gyr and most of the sample is younger than 1 Gyr old.

Differences observed between age distributions with the three different age definitions are as expected since the mass-weighted age is smaller than the quarter-mass age, the latter being smaller than the onset age.

The bottom panel of Fig. 17 shows the distributions of age as a function of redshift. Ages continuously cover the range from the lowest ages allowed in our fit (0.05 Gyr) to ages up to 2 Gyr at $z \sim 2-3$, which is the lowest redshift end of our sample. We do not observe galaxies right at the age corresponding to the age of the Universe in ΛCDM. This indicates that our data and measurement process provide enough age resolution and accuracy to avoid a saturation effect on the largest ages; less than ten objects are close to the age of Universe at the observed redshift.

We performed a number of tests to determine the dependency of these ages with input assumptions. Dependencies with dust prescription, input models, and idealized SFH are presented in Sect. 7.2.2. Appendix A.1 discusses the effects of using a SFH library with two bursts instead of using a single burst delayed SFH. While it is clear that when fitting a single burst model we measure the age of the last *major* burst and, therefore, the ages that we derive are most likely lower limits, a single burst model is capable of recovering ages reasonably well when the first burst dominates the SFH.

7. The formation redshift function (FzF)

7.1. Determining the redshift of formation from galaxy ages

We now compute the formation redshift $z_{\rm f}$ of each galaxy from the age distributions presented in Sect. 6, combining the age of the galaxy and the observed redshift $z_{\rm obs}$. The way $z_{\rm f}$ is measured is illustrated in Fig. 18, where we link the redshift at which the galaxy is observed, $z_{\rm obs}$, the age of the galaxy, and the formation redshift $z_{\rm f}$. The formation redshift is calculated by subtracting the age of the galaxy (determined with GOSSIP+, ΔA in the figure), from the age of the Universe $A_{\rm U}(z_{\rm obs})$ at the observed redshift $z_{\rm obs}$. The redshift corresponding to the resulting age, $A_{\rm U}(z_{\rm f})$, is the formation redshift.

The computation of the formation redshift is quite sensitive to age errors. The uncertainty on $z_{\rm f}$ is defined as

$$\Delta z_{\rm f} = z_{\rm f}({\rm Age} + \delta{\rm Age}) - z_{\rm f}({\rm Age}), \quad (9)$$

where $\Delta z_{\rm f}$ ranges from ~0.1 at $z \sim 2$ to ~2 at $z \sim 6$ for an uncertainty in age measurement of δAge = 0.1 Gyr; this more than doubles for δAge = 0.3 Gyr. A small change of age at high

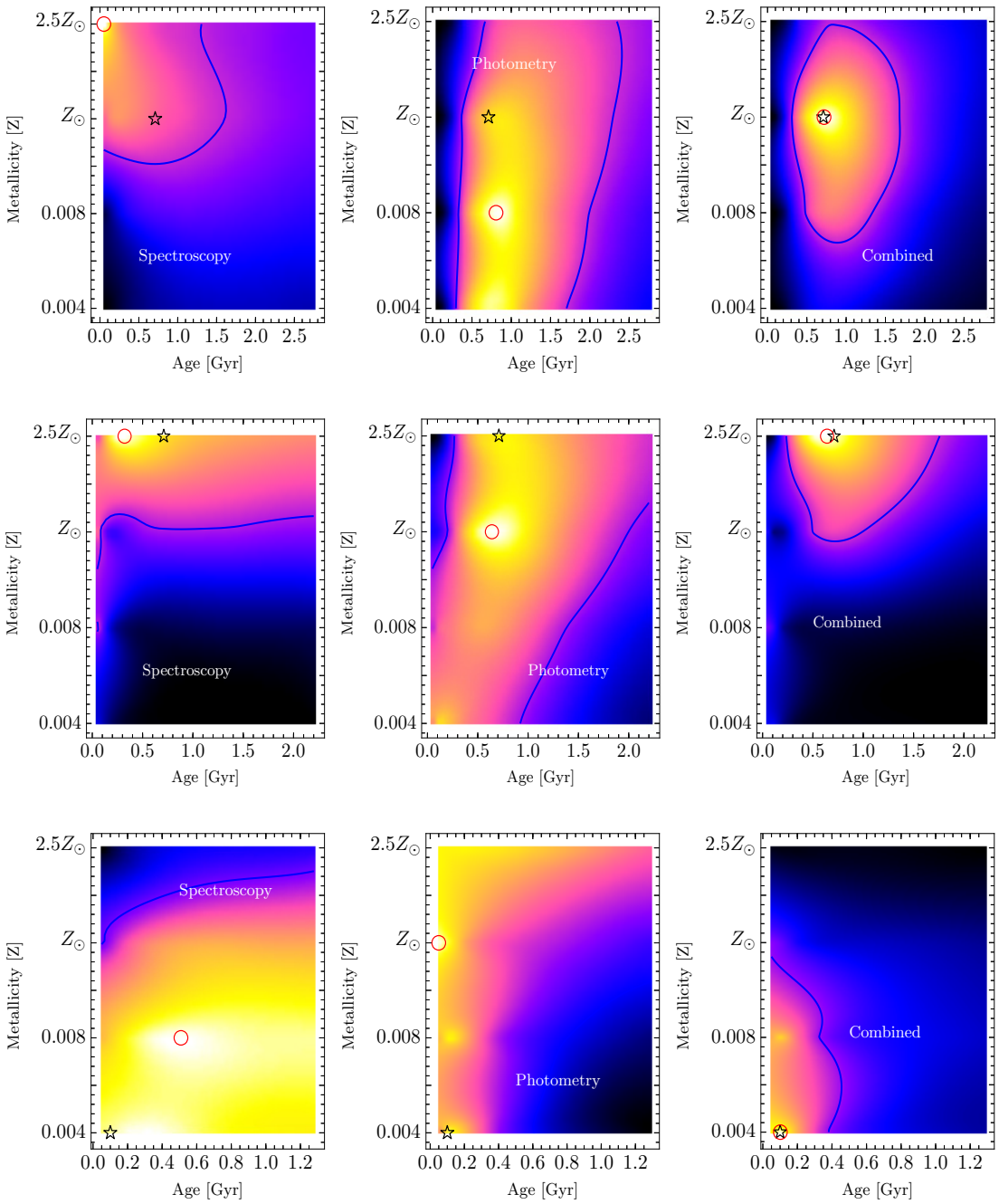

Fig. 11. Same as Fig. 9 but for the age-metallicity relation. *From top to bottom*: Simulated galaxy at $z = 2.25$ with a metallicity of $Z_\odot$ and an age of 0.7 Gyr, simulated galaxy at $z = 3.45$ with a metallicity of 2.5 $Z_\odot$ and an age of 0.7, and finally a simulated galaxy at $z = 4.47$ with a low metallicity of $Z_\odot/5$ and an age of 0.1 Gyr.

redshift may therefore significantly change the formation redshift estimation. However, the relative distribution of z_f in a star-forming galaxy population should not be affected, as discussed below. As mentioned in the previous section, one should keep in mind that a galaxy age maybe a lower limit when the SFH is dominated by a recent strong burst hiding older burst in their past. For such galaxies, the z_f we are presenting here are also lower limits to the formation redshift.

We apply this method to compute z_f for the 3597 galaxies in our VUDS sample and obtain the distribution of formation redshifts presented in Fig. 19. In the top panel the maximum formation redshift defined from $\mathcal{A}_{\rm onset}$ is $z_f = 19.5$ and a significant rise in the number of galaxies occurs at $z_f < 10$. The median of the distribution is $\langle z_f(\mathcal{A}_{\rm onset})\rangle \sim 3.7$. With the $\mathcal{A}_{\mathcal{M}/4}$ definition, the highest formation redshift is at $z_f = 14.8$, the median formation redshift is $\langle z_f(\mathcal{A}_{\mathcal{M}/4})\rangle \sim 3.36$, and the histogram starts to increase significantly at $z_f < 9$. Using $\mathcal{A}_{\rm MW}$, the highest formation redshift is at $z_f = 11.9$, the median formation redshift is $\langle z_f(\mathcal{A}_{\mathcal{MW}})\rangle \sim 3.22$ with a significant fraction of galaxies with $z_f \sim 8$.

In the bottom panel of Fig. 19 we present the distribution of formation redshifts for galaxies in three observed redshift bins, i.e. $2 < z_{\rm obs} \leq 3$, $3 < z_{\rm obs} \leq 4$, and $z_{\rm obs} > 4$. We observe that the formation redshift distribution shifts to higher redshifts for galaxy observed at increasing redshifts. This is easily understood when applying an increasing lower redshift bound to a continuous age distribution when going to higher redshifts, cutting out the galaxies that form at the lower redshifts.

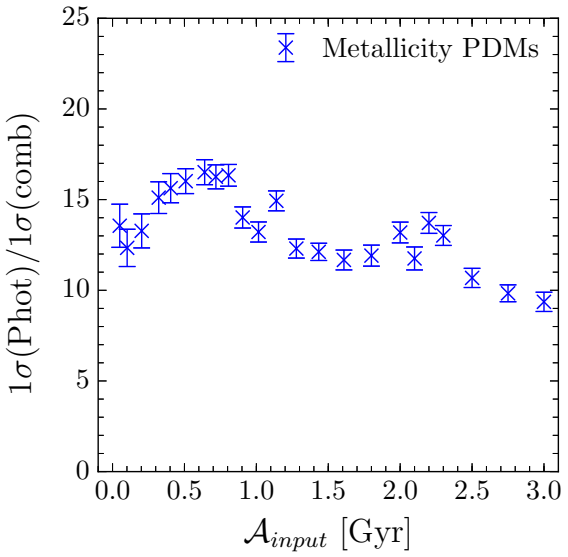

Fig. 12. Comparison of the 1σ area from fit on the photometry and the fit on the combined data as function of the input $\mathcal{A}_{\rm onset}$ for the age-metallicity relation. At any age, the 1σ(phot) area based on the photometry is on average 13.3 times bigger than the 1σ(comb) area, based on the combined fit. The error bars are computed from the median absolute deviation.

Table 6. Global properties of the age-metallicity PDMs in our simulations.

Property	Photometry	Combined
Median $N_{\rm peak}$	2	1
$(\mathrm{Age}-Z)_{\rm in} \in 1\sigma$	75%	78%

Table 7. Libraries from population synthesis models BC03 and M05 used to estimate ages of VUDS galaxies.

Parameter	BC03
IMF	Chabrier
Metallicity	0.004, 0.008, 0.02, 0.05
SFH	Delayed with $\tau_{\rm SFH} \in [0.1; 5.0]$ Gyr
$E(B-V)$	0.0 to 0.5
Ages (Gyr)	0.05 to 4.0
IGM	7 models/z

The observed $z_{\rm f}$ distributions reported in Fig. 19 are shaped by the redshift distribution of the VUDS sample (Le Fèvre et al. 2015) produced by the survey selection function, and therefore do not directly represent the volume-average distribution of formation redshifts of the underlying population. This is derived in the next section.

7.2. The distribution of formation redshifts: Formation redshift function (FzF)

7.2.1. Formation redshift function (FzF)

We introduce a new statistical description of a galaxy population to investigate when galaxies preferentially form: the formation redshift function (FzF). The FzF describes the number of galaxies formed per unit volume at a formation redshift $z_{\rm f}$. From the definition of $z_{\rm f}$ the FzF indicates how many galaxies started forming their stars at a given $z_{\rm f}$.

The observed $z_{\rm f}$ distribution needs to be corrected for the selection function of the VUDS sample applying the target sampling rate (TSR) and spectroscopic redshift success rate (SSR; for a definition of TSR and SSR; see Cucciati et al. 2012; Le Fèvre et al. 2015) to the observed data to recover volume quantities (Tasca et al., in prep.). From the formation redshifts, $z_{\rm f}$, which were determined in the previous section, we compute the number of galaxies formed at $z_{\rm f}$ in a formation redshift bin Δz using the $V_{\rm max}$ formalism (Schmidt 1968); we compute for each galaxy, $V_{\rm max}$, the volume that the galaxy could lie in without dropping outside the survey selection limits.

To approximately follow the evolution of the same galaxy population with redshift and keep the highest number of galaxies in the FzF estimation, we select galaxies that are more massive than $\log_{10} M_\star \geq 9.3\ M_\odot$ at $z \sim 5$. We then evolve this mass limit with redshift using the evolution of the characteristic mass of the stellar mass function as derived by Ilbert et al. (2013) and the evolution of the sSFR with redshift in the VUDS survey (Tasca et al. 2015). This leads to the following mass threshold:

$$\log_{10} M_\star > 10.30 - z \times 0.2. \quad (10)$$

The selection of galaxies is shown in Fig. 20.

Applying this mass cutoff we selected ~2350 VUDS galaxies to compute the FzF in three observed redshift ranges, $2 < z \leq 3$, $3 < z \leq 4$, and $z \geq 4.0$, as shown in Fig. 21. We find that the FzF has a very similar shape in the three observed redshift ranges with a continuously rising number of forming galaxies from high $z_{\rm f}$ down to the lower redshift considered in this study (top left panel). From redshift 10 to redshift 4 there are ~60 times more galaxies forming that later develop into galaxies with $M_\star > 10.30 - z \times 0.2$. The number of galaxies starting to form their stars at $z \sim 10$ is already substantial with a value at $\sim 2.88 \times 10^{-4}$ galaxies $\times$ Mpc^{-3}, while at $z \sim 3$ the number of forming galaxies is $\sim 1.4 \times 10^{-2}$ galaxies $\times$ Mpc^{-3}.

The comparison of the FzF produced from different age definitions is presented in the three other panels of Fig. 21. Different age definitions do not produce a significant change in the two highest redshift bins. While in the lowest redshift bins, we see some difference between the FzF given by $\mathcal{A}_{\rm onset}$, on the one hand, and $\mathcal{A}_{\mathcal{M}/4}$ or $\mathcal{A}_{\rm MW}$. This is in line with the formation redshift distributions presented in the previous section.

We parametrize the FzF by a power-law $\log_{10}\mathrm{FzF} = \alpha(1+z)^{\zeta}$. The results of the fit for the FzF derived from $\mathcal{A}_{\rm onset}$ are shown in Fig. 22 and we list the values of α and ζ for different observed redshift bins in Table 8.

The parameters resulting from the fit of the FzF are very close to each other from one observed redshift bin to another. This strong similarity of the FzF independent of the observed redshift seems to indicate that the FzF is a universal function for this population of SFGs selected to have $\log_{10} M_\star > 10.30 - z \times 0.2$.

7.2.2. Influence of the SFH, stellar population models, and dust prescription on the FzF

The FzFs presented above were computed from the combined fit using exponentially delayed SFH. Such galaxy SFHs present both rising and declining parts. Nevertheless it is possible to study other kind of SFHs such as the purely exponentially declining SFHs. We therefore reran GOSSIP+ using only exponentially declining SFHs. The resulting FzF (and associated fits) are presented in Fig. 23 and Table 9. This figure presents the comparison of the SFH from the exponentially delayed SFHs

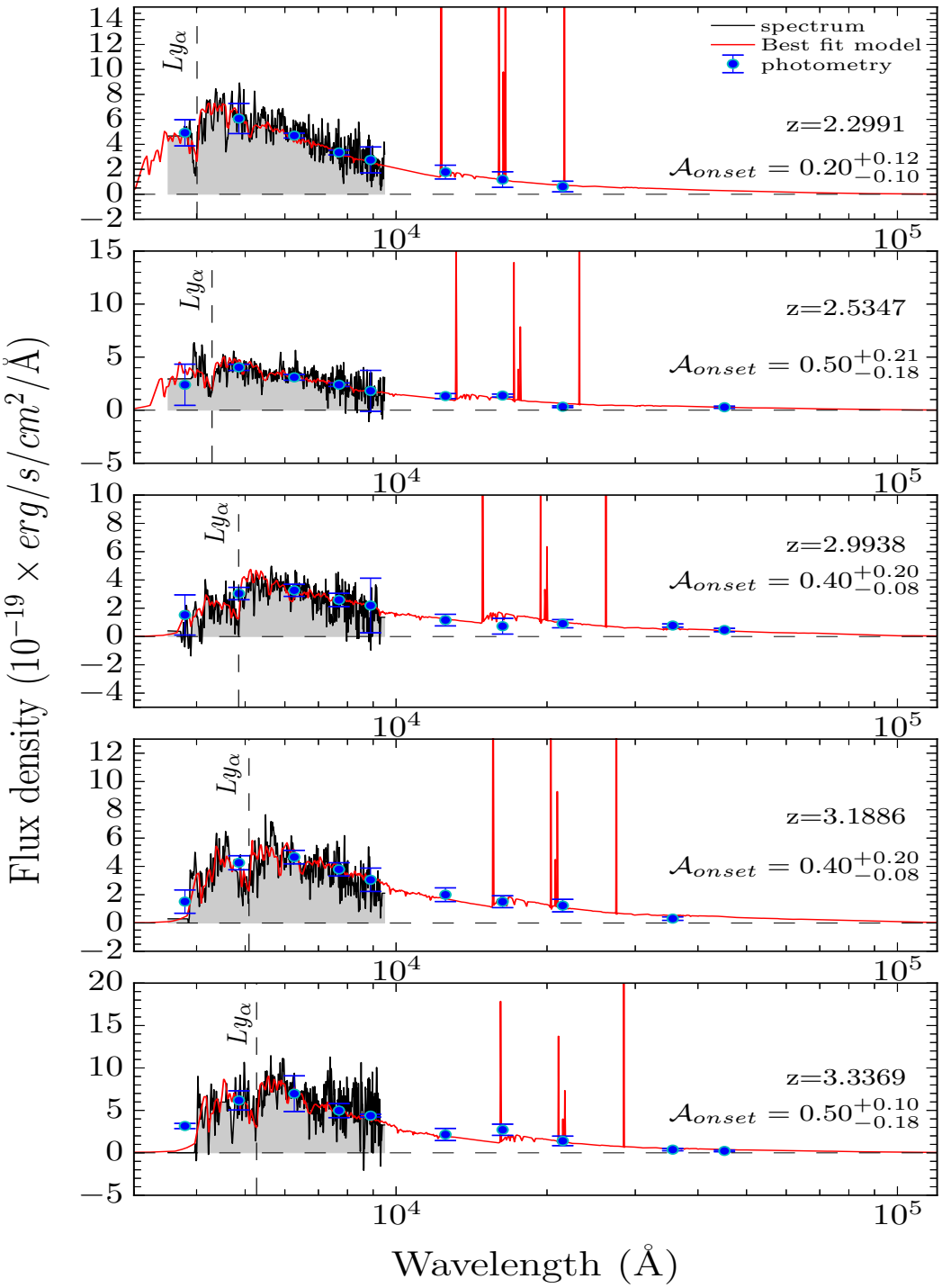

Fig. 13. Five examples of fits in our galaxy sample between $z = 2.2991$ and $z = 3.3369$. In each panel the black line represents the VUDS spectroscopy, the blue points are the photometric points, and the best-fit template is in red. For each case, the template fit represents both photometric and spectroscopic data well.

and the exponentially declining SFHs. The new FzFs are on average 0.7 dex below those computed from SFH exponentially delayed. This behaviour is expected. Because of the presence of only a declining part in the SFH, the ages from this type of SFH are artificially lowered. This leads to smaller formation redshifts. Consequently, at a given formation redshift in the FzF, the number of created galaxies is lower when using exponentially declining SFH. A discussion on the use of multi-burst SFH to compute ages is presented in Appendix A.1. Ages computed from single burst delayed SFH and from two-burst SFH are in good agreement because of the relatively short time available at these redshifts to build two or more bursts, or because the first burst is also the strongest and dominates the SFH at the cosmic time the galaxy is observed. Therefore the FzF, which is constructed with ages derived from delayed SFH or two or more burst SFH, do not exhibit strong differences. If the first bursts are systematically hidden by later strong bursts, then ages would be higher than we are computing and z_f of individual galaxies would be larger; the FzF would be systematically shifted towards even higher formation redshifts.

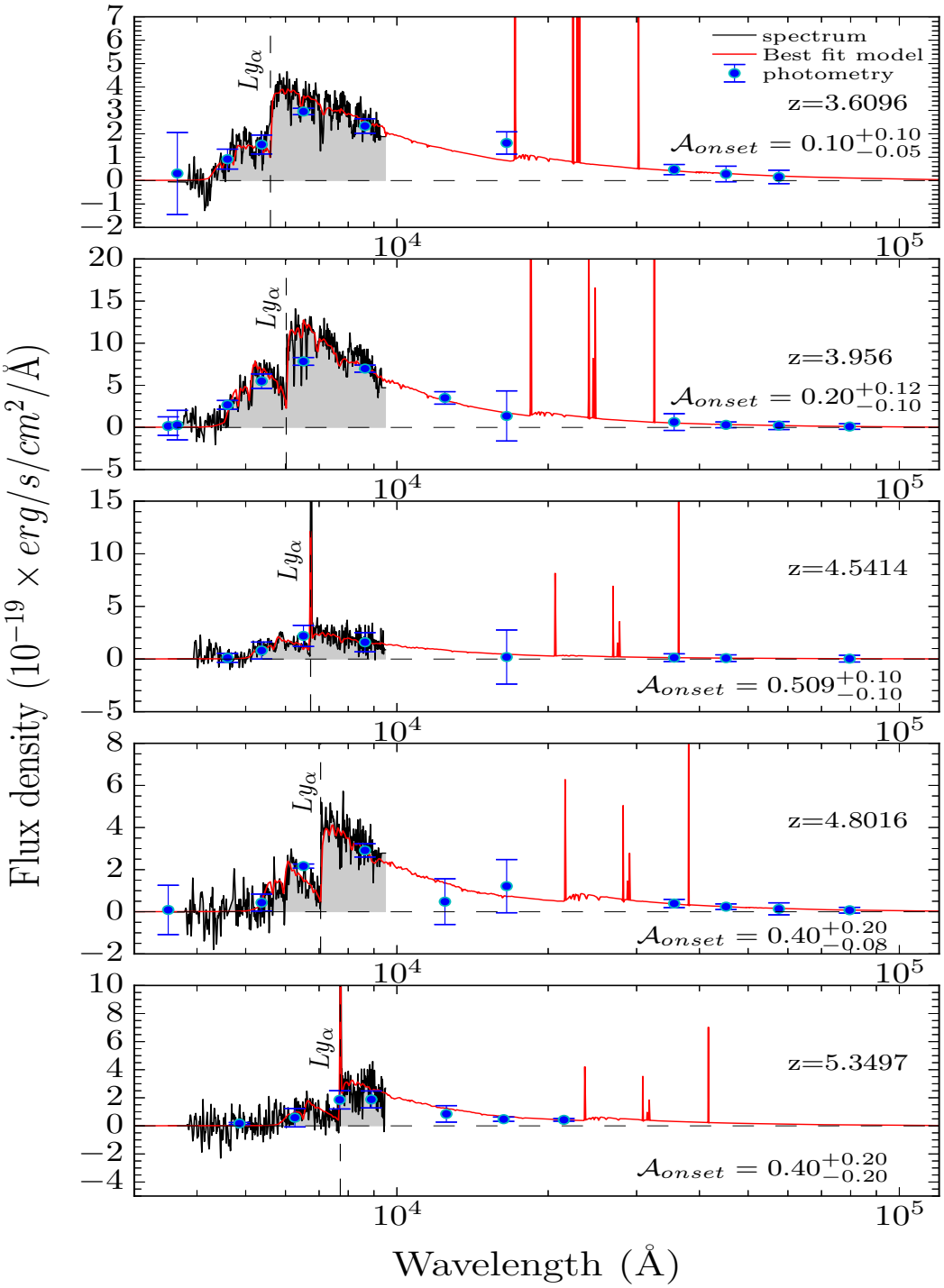

Fig. 14. Same as Fig. 13 for galaxies between $z = 3.6$ and $z = 5.34$.

The impact of a change of stellar population models is presented in Fig. 24 and Table 10 using Maraston (2005) models instead of BC03 models (the age we are using here are based on the $\mathcal{A}_{\rm onset}$ definition). With these models we use the same parameters as set for BC03 models but with slightly different metallicities of 0.001, 0.01, 0.02 (solar), and 0.04.

At low redshift using the M05 models leads to more forming galaxies than with the BC03 models. At high redshift, we observe the opposite: the M05 models produce fewer forming galaxies than the BC03 models. This change is expected as the M05 models tend to give lower ages than BC03 models (see Sect. 5). Consequently the number of galaxies with high formation redshifts should be lower in the case of M05, as observed.

We also test the impact of the dust extinction law on the FzF used in the fitting process. We produce a new fit using the SMC extinction from Prevot et al. (1984). The results are shown in Fig. 25 and Table 11. All the other parameters are the same as presented in Sect. 6.1 with the use of BC03 models.

In most of the redshift domain we are studying, the extinction from the Calzetti law is smaller than for the SMC law. Therefore, to compensate for this change, stellar populations need to be slightly redder and, therefore, ages must be slightly older

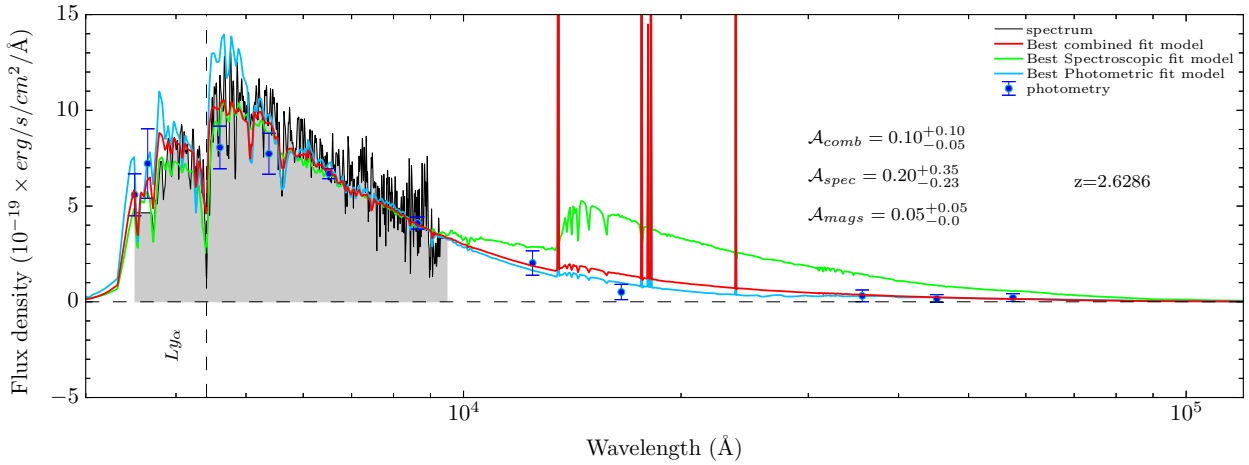

Fig. 15. Comparison of the fits from the three types of fitting modes on a galaxy at $z = 2.6286$. The black line represents the observed spectroscopy and the blue points represent the observed photometry. We show the best fit model of the fit on the photometry, combined data and spectroscopy in light blue, red and green, respectively. The fit on the combined data is only able to reproduce both types of data.

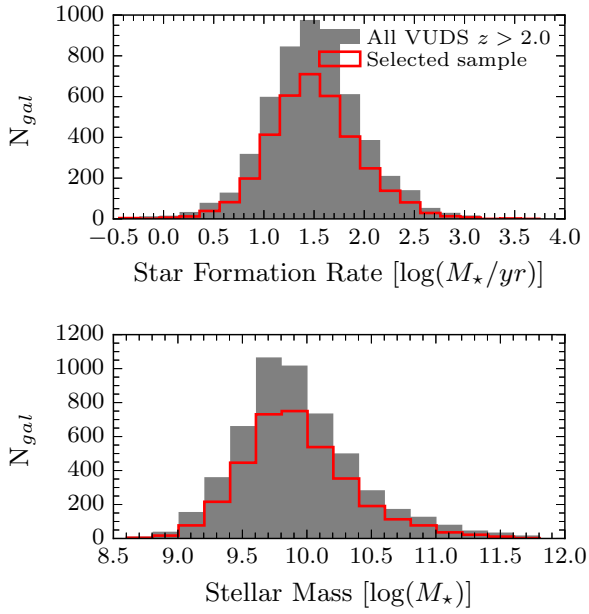

Fig. 16. Stellar mass and SFR distributions for galaxies in the selected VUDS sample (red histograms) compared to the full VUDS sample (shaded grey histograms). The two distributions are similar showing that our selection is not imposing any particular bias.

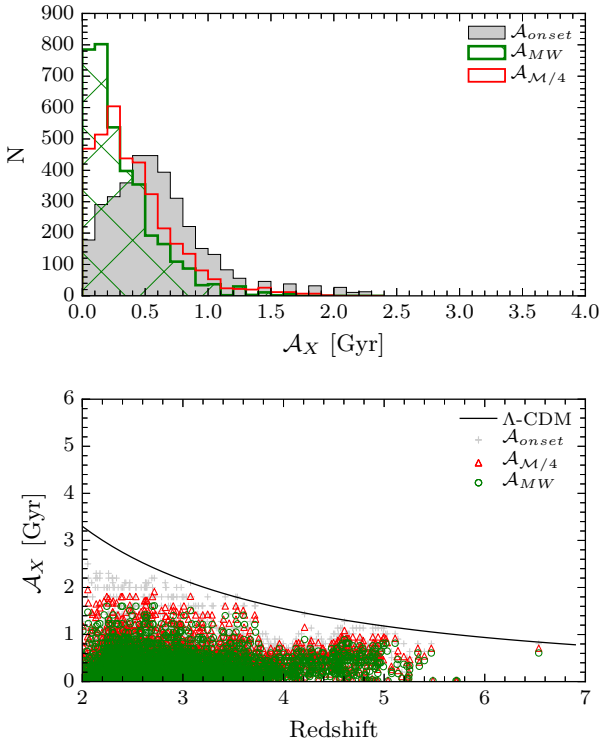

Fig. 17. Age distributions of 3597 VUDS galaxies computed from the combined fit of the spectroscopy and the photometry as described in Sect. 3. *Top panel*: the three distributions in red, hatched green, and filled grey histograms correspond to the three definitions of age we use in our paper and defined in Sect. 4, $\mathcal{A}_{\rm onset}$, $\mathcal{A}_{\rm MW}$, and $\mathcal{A}_{M/4}$, respectively. *Bottom panel*: age distributions of VUDS galaxies as a function of observed redshift $z_{\rm obs}$. The three age definitions, $\mathcal{A}_{\rm onset}$, $\mathcal{A}_{\rm MW}$, and $\mathcal{A}_{M/4}$ are presented in red, green, and grey, respectively. The line represents the age of the Universe in the ΛCDM cosmological model. The lowest age corresponds to the smallest possible age allowed by our model library (50 Myr). The upper limit is in general lower than the age of the Universe.

when using the SMC law than when using the Calzetti law. Consequently, the FzF is marginally flatter when using the SMC extinction law with more galaxies with old ages leading to more high-$z_{\rm f}$ galaxies.

A last test is performed to examine how variations in the [OII]/[OIII] line ratio would change the FzF. To check this we multiply the [OII]/[OIII] ratio defined in Sect. 3 by 10 and rerun the fit. The change in age results over the sample is of 0.05 Gyr with an error on the mean of 4×10^{-3}. Therefore we conclude that uncertainties on the [OII]/[OIII] ratio only marginally affect the computation of the FzF.

8. Discussion and conclusions: the epoch of formation of bright massive SFGs

The distribution of ages of distant galaxies is an important piece of information in the galaxy formation puzzle.

Our study demonstrates that the computation of galaxy ages at $z > 2$ can be impressively improved when performing SED fitting on combined photometry and spectroscopy data, given the knowledge of the age of the Universe. Performing fits on both spectroscopy and photometry, together with the knowledge

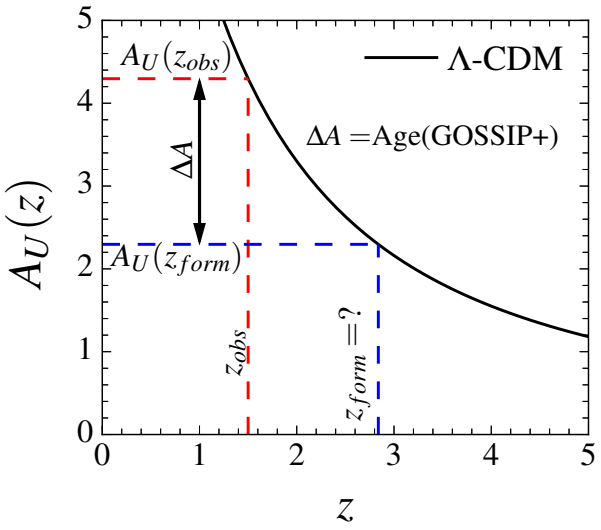

Fig. 18. Method used to compute the formation redshift. The observed redshift is z_{obs} and the age of the Universe at this redshift is given by $A_U(z_{obs})$. The age of the galaxy measured by GOSSIP+ is Age(GOSSIP+). $A_U(z_f)$ is the age of the Universe at z_f. The black curve represents the age of the Universe as a function of redshift in the ΛCDM model. From the measurement of z_{obs} and the age given by GOSSIP+, we can infer the formation redshift by successive iterations (see text).

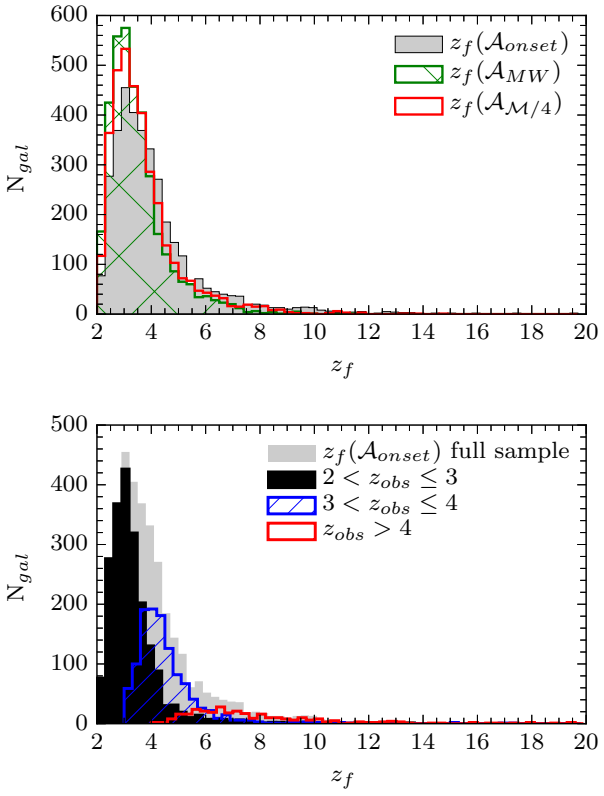

Fig. 19. Formation redshifts of the VUDS sample. *Top panel*: formation redshift distributions from the three age definitions $\mathcal{A}_{\rm onset}$, $\mathcal{A}_{\rm MW}$, and $\mathcal{A}_{\mathcal{M}/4}$ (grey filled, hatched green, and red empty histograms, respectively). *Bottom panel*: comparison of the formation redshift distribution (from $\mathcal{A}_{\rm onset}$) in three observed redshift bins: $2 < z_{\rm obs} \leq 3$, $3 < z_{\rm obs} \leq 4$, and $z_{\rm obs} > 4$, shown by the filled black histograms, blue hatched histogram, and red empty histogram.

of the age limit imposed by the age of the Universe at the observed redshift considerably restricts possible age models, and

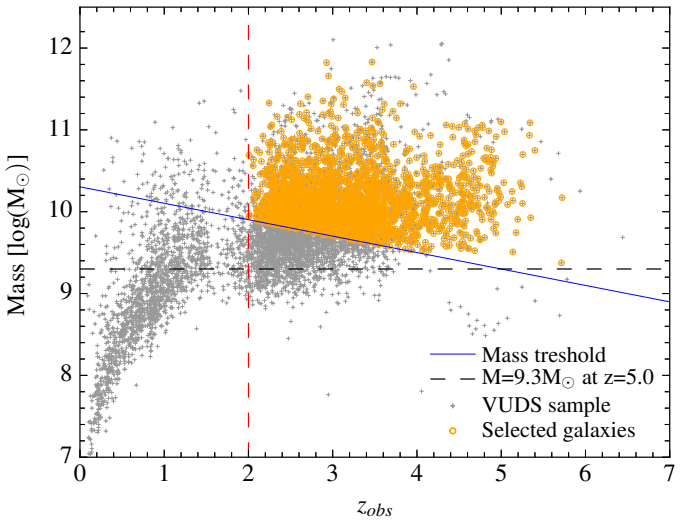

Fig. 20. Mass cutoff used to select the galaxy population for the construction of the formation redshift function (FzF). The grey crosses represent the full VUDS sample, while the orange circles are the galaxies selected for the computation of the FzF. These galaxies are selected to be above the blue line. The blue line represents the evolution of a galaxy with $\log M_\star = 9.3$ at $z = 5$ with redshift, following the evolution of the characteristic mass of the stellar mass function as derived by Ilbert et al. (2013).

leads to reliable age measurements. From the analysis of PDMs, we find that degeneracies present at low redshift between age, metallicity, and dust extinction tend to be considerably reduced at high redshifts.

We emphasize the benefit of combining UV rest-frame spectra and photometry when performing the SED fitting. While the redder photometric bands help recover the stellar mass $M_\star$, the detailed information available from UV spectra significantly improves the measurement of the SFR and dust extinction. We conclude from our analysis and simulations that the age of a galaxy at $z > 2$ can be computed with a typical uncertainty of 10% provided that important assumptions on, for example the shape of the SFH, are valid. We find that age measurements using SED fitting at these redshifts are of an accuracy comparable to that of $M_\star$ and SFR measurements and therefore can be taken into full consideration in our quest to understand galaxy evolution.

Using this formalism, we use rest-frame UV spectra from the VUDS spectroscopic survey in combination with extensive deep photometry in the COSMOS, ECDFS, and VVDS02h fields to compute ages of ~3600 galaxies with $2 < z_{\rm obs} < 6.5$ to study the epoch of galaxy formation. The large VUDS sample with a broad selection function allows us to probe a large range of ages. At any observed redshift we find ages ranging from the youngest possible allowed by our model library at 50 Myr to the oldest allowed by the age of the Universe at that redshift.

Assuming that the age of the Universe is known from the current best cosmological model, we then derive formation redshifts $z_{\rm f}$. We explore the impact on $z_{\rm f}$ of different age definitions including the age defined by the onset of star formation, the mass-weighted age, and the quarter-mass age. We find that these different definitions do not change the general $z_{\rm f}$ distribution.

It is striking to observe that massive galaxies with $\log_{10} M_\star \geq 10.30 - z \times 0.2$ at $z > 2$ present a continuous distribution of formation redshifts (Fig. 19). This is an indication that galaxies can start forming their stars at any redshift and that there is no preferred epoch of galaxy formation. This is clearly evident

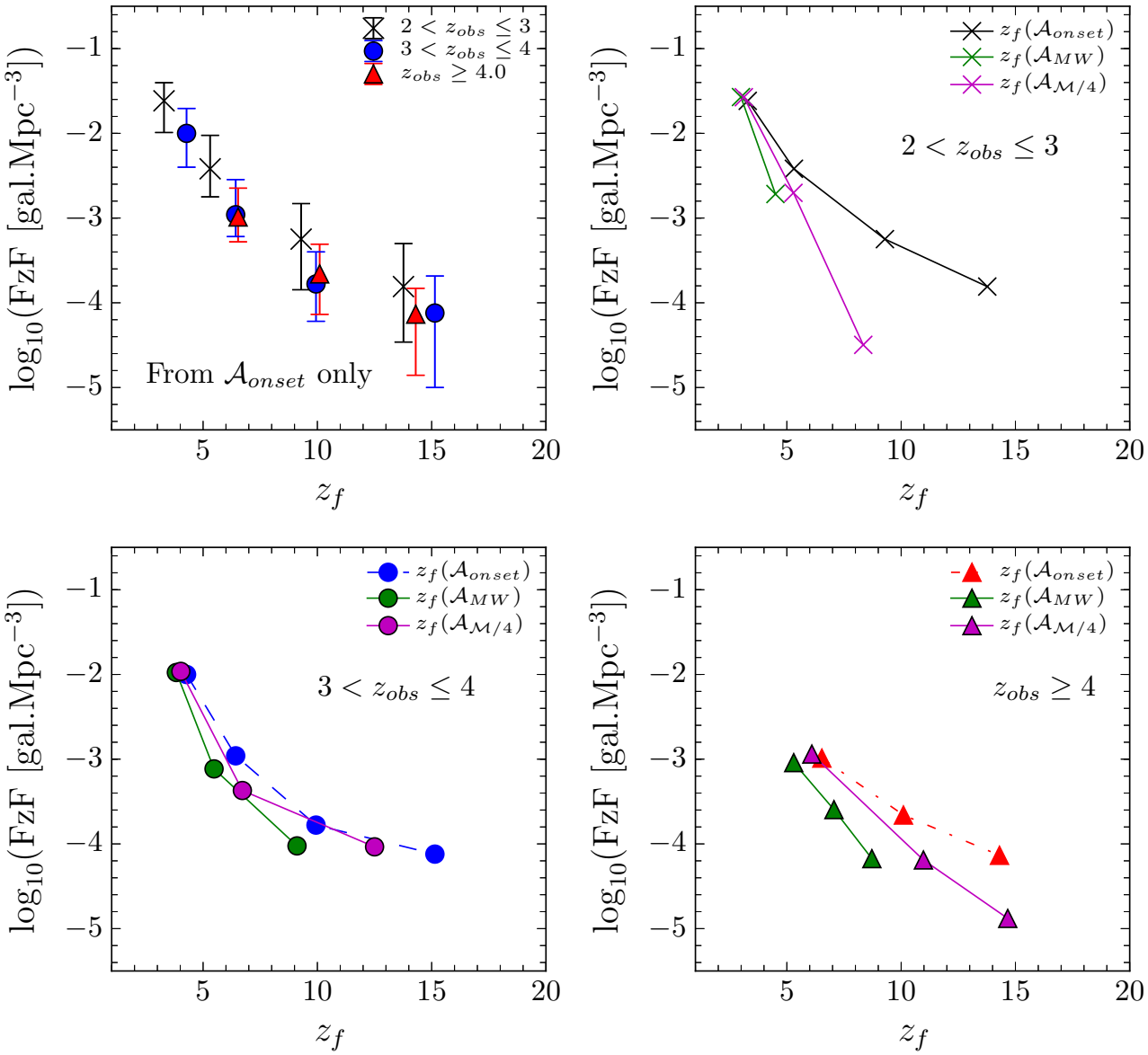

Fig. 21. Formation redshift functions (FzF). *Top left*: formation redshift functions computed from the galaxies in 3 observed redshift bins: $2 < z_{obs} \leq 3$ (black crosses), $3 < z_{obs} \leq 4$ (blue points), and $z_{obs} \geq 4.0$ (red triangles). The *three other panels* correspond to the comparison of the FzFs corresponding to the 3 different age definitions used in this paper in increasing redshift bins. *From top right to bottom right* we compare the FzFs in the following observed redshift bins: $2 < z \leq 3$, $3 < z \leq 4$ and $z \geq 4.0$.

when computing the formation redshift function FzF representing the number of galaxies per unit volume that start forming their stars as a function of formation redshift (Fig. 21). The FzF is continuously rising as redshift decreases, which is well represented by a power-law $\log_{10}\mathrm{FzF}(z_f) = \alpha(1+z_f)^{\zeta}$ with ζ in the range from -0.61 to -1.06 at $z \sim 2.6$ to $z \sim 4.5$; no preferred formation redshift can be identified.

While galaxies appear to form at any redshift there are a lot more galaxies forming at $z \sim 3$ than at $z \sim 10$. We find that the number of forming galaxies continuously increases from $\sim 2.88 \times 10^{-4}$ galaxies Mpc^{-3} at $z = 10$, to $\sim 1.4 \times 10^{-2}$ galaxies Mpc^{-3} at $z \sim 3$, which is an increase of almost 2 dex. The SFRD is directly related to the number of newly formed galaxies in the Universe, which is expressed by the FzF. Assuming a constant average SFR for galaxies at the same stage of evolution across the redshift $2 < z < 6$, the global comoving SFRD must be proportional to the number of forming galaxies traced by the FzF, and we therefore infer from the FzF that the SFRD should also increase by $\sim$2 dex from $z \sim 10$ to $z \sim 3$. This qualitative estimate of the SFRD evolution from the FzF is in excellent agreement with the evolution of the SFRD over the same redshift range of $\sim$1.5–2 dex as measured, for example from the UV luminosity density (Madau & Dickinson 2014; Bouwens et al. 2015). Moreover, the ratio between the SFRD and the FzF gives a SFR at $z \sim 4$–6 of $\sim$8–17 $M_{\odot}$/yr (up to 20 $M_{\odot}$/yr from the mass-weighted age FzF). This value is of the same order as the SFR of galaxies measured in our VUDS sample (Tasca et al. 2015; this paper).

With this first detailed and systematic exploration of galaxy ages in a representative sample of SFGs in the redshift range $2 < z < 6.5$, we demonstrated that the age distribution of galaxies provides important clues on the formation of galaxies in the first few billion years of evolution after the Big Bang. Reliable

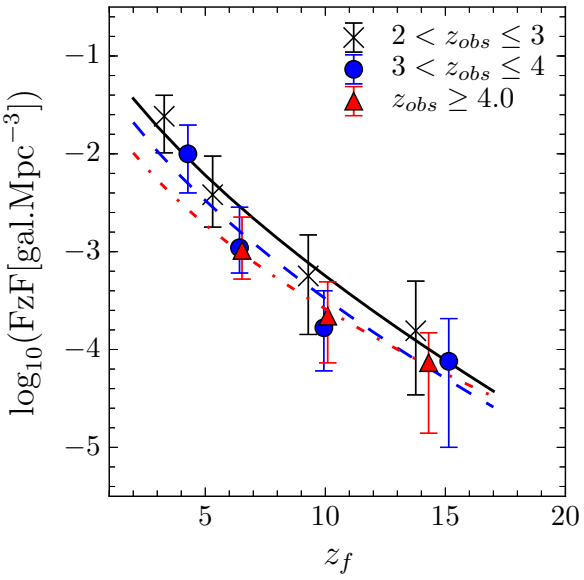

Fig. 22. Fit of the formation redshift functions in three observed redshift bins using $\log_{10}(f) = \alpha(1+z)^{\zeta}$.

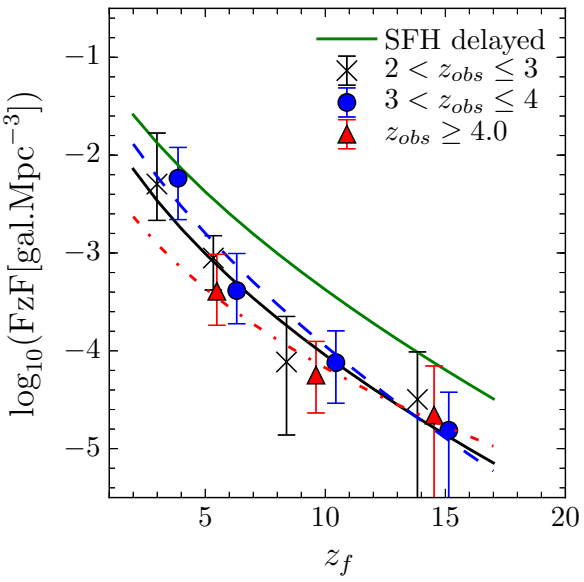

Fig. 23. Comparison of the FzFs from different SFHs. In green, global FzF computed from all the points presented in Fig. 22 derived with exponentially delayed SFH. In black, blue, and red, we show the FzFs derived from exponentially delayed SFH as well as their fits in 3 observed redshift bins: $2 < z_{\rm obs} \leq 3$, $3 < z_{\rm obs} \leq 4$, and $z_{\rm obs} \geq 4.0$, respectively.

Table 8. Parameters of the formation redshift function (FzF) fit in different observed redshift bins.

$z_{\rm obs}$	$\langle z_{\rm obs} \rangle$	α	$\delta\alpha$	ζ	$\delta\zeta$
$2.0 < z < 3.0$	2.63	−0.71	19%	0.62	12%
$3.0 < z < 4.0$	3.40	−0.90	28%	0.56	23%
$z > 4.0$	4.48	−1.21	10%	0.45	8%
Full sample	–	−0.84	14%	0.58	10%

Notes. The fit function is parametrized as $\log_{10}\mathrm{FZF}(z) = \alpha(1+z)^{\zeta}$.

age estimates derived from the method proposed in this paper further expands on our tool box to characterize galaxies at any epoch towards a better understanding of galaxy formation and evolution at early cosmic times.

Adding rest-frame optical spectra to UV rest-frame spectra, combined with broadband photometry, will be an important part of next generation high-redshift surveys as this will further improve the accuracy of parameters from SED fitting. The accuracy of physical parameter measurements such as Age, $M_\star$, SFR, or dust content will keep improving as wider spectroscopic coverage becomes available from NIR spectrographs on ground-based telescopes or with the *James Webb* Space Telescope. This will be an important contribution to further improve our understanding of galaxy formation.

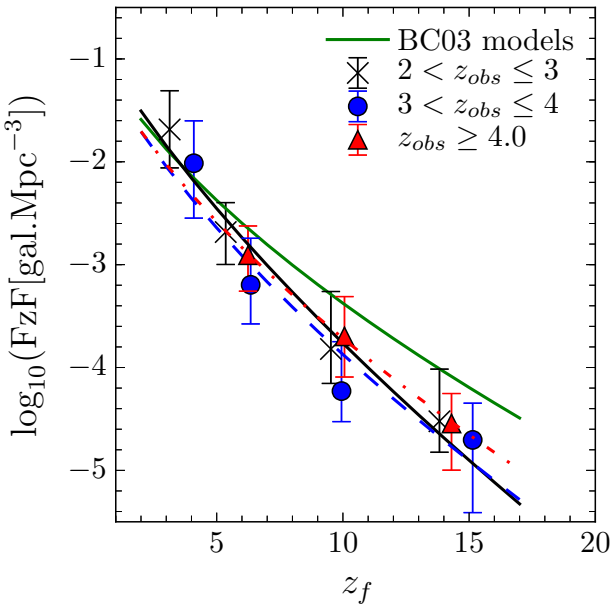

Fig. 24. Comparison of the FzFs from different stellar population synthesis models. In green, a global FzF computed from all the point presented in Fig. 22, derived with exponentially delayed SFH and BC03 models. In black, blue and red, we show the FzFs derived from M05 models as well as their fits in 3 observed redshift bins: $2 < z_{\rm obs} \leq 3$, $3 < z_{\rm obs} \leq 4$, and $z_{\rm obs} \geq 4.0$, respectively.

Table 9. Parameters of the formation redshift function (FzF) fit in different observed redshift bins for the FzFs derived from exponentially declining SFH.

$z_{\rm obs}$	$\langle z_{\rm obs} \rangle$	α	$\delta\alpha$	ζ	$\delta\zeta$
$2.0 < z < 3.0$	2.98	−1.25	22%	0.48	18%
$3.0 < z < 4.0$	3.40	−1.00	21%	0.56	14%
$z > 4.0$	4.48	−1.75	15%	0.35	17%

Notes. The fit function is parametrized as $\log_{10}\mathrm{FZF}(z) = \alpha(1+z)^{\zeta}$.

Table 10. Parameters of the formation redshift function (FzF) fit in different observed redshift bins for the FzFs derived from M05 models.

$z_{\rm obs}$	$\langle z_{\rm obs} \rangle$	α	$\delta\alpha$	ζ	$\delta\zeta$
$2.0 < z < 3.0$	2.98	−0.69	18%	0.70	11%
$3.0 < z < 4.0$	3.40	−0.85	25%	0.63	22%
$z > 4.0$	4.48	−0.88	5%	0.60	4%

Notes. The fit function is parametrized as $\log_{10}\mathrm{FZF}(z) = \alpha(1+z)^{\zeta}$.

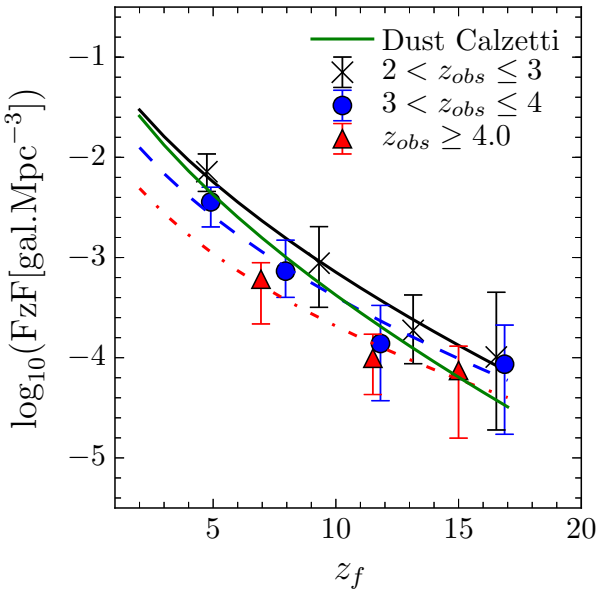

Fig. 25. Comparison of the FzFs using different dust extinction prescriptions. The FzF presented in Fig. 22 derived with Calzetti's prescription and BC03 models is shown in green. In black, blue and red, we show the FzFs derived using the SMC extinction law of Prevot et al. (1984), including fits of the FzFs in 3 observed redshift bins: $2 < z_{\rm obs} \leq 3$, $3 < z_{\rm obs} \leq 4$ and $z_{\rm obs} \geq 4.0$, respectively.

Table 11. Parameters of the formation redshift function (FzF) fit in different observed redshift bins for the FzFs derived using the SMC extinction law.

$z_{\rm obs}$	$\langle z_{\rm obs} \rangle$	α	$\delta\alpha$	ζ	$\delta\zeta$
$2.0 < z < 3.0$	2.98	−0.82	15%	0.55	7%
$3.0 < z < 4.0$	3.40	−1.16	19%	0.44	18%
$z > 4.0$	4.48	−1.55	25%	0.35	13%

Notes. The fit function is parametrized as $\log_{10}\mathrm{FZF}(z) = \alpha(1+z)^{\zeta}$.

Acknowledgements. We thank Claudia Maraston and Roser Pellò for useful discussions as well as the anonymous referee for useful suggestions. This work is supported by funding from the European Research Council Advanced Grant ERC–2010–AdG–268107–EARLY, by INAF Grants PRIN 2010, PRIN 2012 and PICS 2013 AC, OC, MT, and VS acknowledge the grant MIUR PRIN 2010–2011 and by GEMINI-CONICYT FUND No. 32140033. This work is based on data products made available at the CESAM data centre, Laboratoire d'Astrophysique de Marseille, France.

References

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bielby, R., Hudelot, P., McCracken, H. J., et al. 2012, A&A, 545, A23
Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, A&A, 363, 476
Bouwens, R. J., Bradley, L., Zitrin, A., et al. 2014, ApJ, 795, 126
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2015, ApJ, 803, 34
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Bromm, V., Yoshida, N., Hernquist, L., & McKee, C. F. 2009, Nature, 459, 49
Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000
Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682
Cardamone, C. N., van Dokkum, P. G., Urry, C. M., et al. 2010, ApJS, 189, 270
Chabrier, G. 2003, PASP, 115, 763
Cimatti, A., Cassata, P., Pozzetti, L., et al. 2008, A&A, 482, 21
Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 839
Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31
Cuillandre, J.-C. J., Withington, K., Hudelot, P., et al. 2012, in SPIE Conf. Ser., 8448, 84480M
de Barros, S., Schaerer, D., & Stark, D. P. 2014, A&A, 563, A81
De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499
Franzetti, P., Scodeggio, M., Garilli, B., Fumana, M., & Paioro, L. 2008, in Astronomical Data Analysis Software and Systems XVII, eds. R. W. Argyle, P. S. Bunclark, & J. R. Lewis, ASP Conf. Ser., 394, 642
Garilli, B., Fumana, M., Franzetti, P., et al. 2010, PASP, 122, 827
Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35
Hathi, N. P., Le Fèvre, O., Ilbert, O., et al. 2016, A&A, 588, A26
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Ilbert, O., Capak, P., Salvato, M., et al. 2009, ApJ, 690, 1236
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Kaviraj, S., Cohen, S., Ellis, R. S., et al. 2013, MNRAS, 428, 925
Kennicutt, Jr., R. C. 1998, ApJ, 498, 541
Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36
Kroupa, P. 2001, MNRAS, 322, 231
Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003, in Instrument Design and Performance for Optical/Infrared Ground-based Telescopes, eds. M. Iye, & A. F. M. Moorwood, SPIE Conf. Ser., 4841, 1670
Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845
Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14
Le Fèvre, O., Tasca, L. A. M., Cassata, P., et al. 2015, A&A, 576, A79
Lilly, S. J., Le Fèvre, O., Renzini, A., et al. 2007, ApJS, 172, 70
Madau, P. 1995, ApJ, 441, 18
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maraston, C. 2005, MNRAS, 362, 799
Maraston, C., & Strömbäck, G. 2011, MNRAS, 418, 2785
Maraston, C., Pforr, J., Renzini, A., et al. 2010, MNRAS, 407, 830
Mauduit, J.-C., Lacy, M., Farrah, D., et al. 2012, PASP, 124, 714
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
Meiksin, A. 2006, MNRAS, 365, 807
Ouchi, M., Ellis, R., Ono, Y., et al. 2013, ApJ, 778, 102
Pacifici, C., Kassin, S. A., Weiner, B., Charlot, S., & Gardner, J. P. 2013, ApJ, 762, L15
Pforr, J., Maraston, C., & Tonini, C. 2012, MNRAS, 422, 3285
Pforr, J., Maraston, C., & Tonini, C. 2013, MNRAS, 435, 1389
Planck Collaboration XVI. 2014, A&A, 571, A16
Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E., & Rocca-Volmerange, B. 1984, A&A, 132, 389
Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267
Salpeter, E. E. 1955, ApJ, 121, 161
Schmidt, M. 1968, ApJ, 151, 393
Scodeggio, M., Franzetti, P., Garilli, B., et al. 2005, PASP, 117, 1284
Steinhardt, C. L., Speagle, J. S., Capak, P., et al. 2014, ApJ, 791, L25
Taniguchi, Y., Scoville, N., Murayama, T., et al. 2007, ApJS, 172, 9
Tasca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, A&A, 581, A54
Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, ApJ, 621, 673
Thomas, R., Le Fèvre, O., Le Brun, V., et al. 2017, A&A, 597, A88
Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
Worthey, G. 1994, ApJS, 95, 107
Wuyts, S., Franx, M., Cox, T. J., et al. 2009, ApJ, 696, 348
Zitrin, A., Labbé, I., Belli, S., et al. 2015, ApJ, 810, L12

Appendix A: The impact of star formation histories on age measurements

A.1. Simulations with two-bursts SFHs

The SFH we choose to use in this study is a delayed exponential with a range of timescales. This type of SFH has the advantage of having both a rising and a declining part. We find that 40% of our galaxies are on the rising part of our model SFH, while 60% are on the declining part. Our choice is supported by recent studies (e.g. Pacifici et al. 2013) showing that the SFH of SFGs can be reasonably well reproduced by delayed τ-model and by the fact that the elapsed time in the history of the galaxies is short because it is limited by the age of the Universe at these high redshifts. However, true SFHs might be more complex, as shown in computer simulations. In particular, it is always possible that the last strong burst of star formation would dominate the emitted spectrum and hide any earlier populations, and that therefore the age measured would be an underestimate of the real age. To test under which conditions this may happen, we consider SFHs with two exponentially delayed bursts, varying the relative strengths of the two bursts and the time delay between them. Such an SFH has an increased number of five free parameters including the two single burst timescales, τ_1 and τ_2[2], the two single burst strengths, f_1 and f_2 and the time delay, Δt, between the peak of the first burst and the start of the second burst. The possible values considered for each parameters in our simulation are listed in Table A.1.

This library is constructed from the mono-burst library defined in Sect. 6, adding one additional burst between $t = 0$ and the already existing burst. Therefore possible timescales for the second burst have the same values as for the mono-burst libraries. Because of the exponentially increasing computational time for a fit with an increase number of free parameters, we somewhat restrain the SFH parameter space. The added burst can take three different possible timescales (τ_1), 0.1, 0.4, and 0.7 Gyr. Its strength (f_1) can vary from 0.5 to twice the strength of the second burst. The time delay Δt between the two bursts is allowed to take four different values, 0.1, 0.4, 0.7, and 1.0 Gyr.

We create a mock sample of 8000 galaxies following the method presented in Sect. 5.1 using the two-burst library presented above. In order to test only the uncertainties related to the SFH, we fixed the $E(B-V)$ and metallicity to mean values $E(B-V) = 0.1$ and $Z = 0.008$, and we set the IGM extinction to the mean values of Meiksin (2006) computed by the fit in Sect. 6.3. Then, we conducted two different fitting runs with GOSSIP. First, to simulate a case in which the SFH of the observed object is more complex than a library of templates built only with single bursts, we made a fit of the 8000 simulated dataset using the library of templates as created in Sect. 5.1 only with single bursts. For the full run, the median age difference (using the $\mathcal{A}_{\rm onset}$ definition) is $\mathrm{Age}_{\rm 2-burst} - \mathrm{Age}_{\rm delayed} = -0.05 \pm 0.21$ Gyr, which shows that we are able to recover a reasonable estimation of galaxy ages for galaxies with a multi-burst SFH even using a single burst library. The galaxies for which the fit locks on the first peak of the SFH are similar to the simulations presented in Sect. 5.1. For those galaxies for which the fit indicates that they already passed the second peak in the SFH, we see that the difference in age is strongly dependent on the time delay Δt between the two bursts. When Δt is small we obtain a good estimate of the age, with $\mathrm{Age}_{\rm 2-burst} - \mathrm{Age}_{\rm delayed} = 0.00 \pm 0.23$ Gyr.

[2] Quantities with 1 as indices refer to the first burst that the galaxy experiences, while quantities with 2 as indices refer to the second burst in the history.

Table A.1. Two-burst star formation history parameters

Parameter	Possible values
τ_1 [Gyr]	0.1, 0.4, 0.7
Relative strength of the first burst f_1/f_2	0.5, 1.0, 1.5, 2.0
Time delay between the two bursts Δt [Gyr]	0.1, 0.4, 0.7, 1.0
τ_2 [Gyr]	$\tau_{\rm SFH} \in [0.1; 5.0]$ Gyr
Strength of the second burst f_2	1.0

As Δt increases, and for galaxies that have passed the second peak, the fit shows that it becomes increasingly difficult to estimate the age of the simulated galaxies with a single burst library. For the highest value of 1.0 Gyr that we consider for Δt, $\mathrm{Age}_{\rm 2-burst} - \mathrm{Age}_{\rm delayed} = 0.41 \pm 0.30$ Gyr. This case only concerns the oldest galaxies of our simulated sample that have a long enough lifetime to form two bursts separated by 1.0 Gyr. Assuming that we measure a galaxy of 1.1 Gyr at $z = 4.0$, the formation redshift would be $\sim z_{\rm f} = 10.5$. If this galaxy happens to have experienced an earlier burst 1.0 Gyr before that identified with our single burst library (and therefore an age on average ~0.4 Gyr older than that computed by the fit), the redshift of formation $z_{\rm f}$ would have to be re-estimated at a much earlier formation redshift $z_{\rm f} > 30$; this may prove to be unrealistic. For simulated galaxies with a high Δt, the age actually measured by the fit corresponds to the time since the start of the second burst in the SFH (the last burst experienced by the galaxy).

In general whatever the age of a galaxy measured using single bursts with a delayed SFH, an older burst is hard to identify if the most recent burst dominates the SFH. In most cases our method accurately measures ages when the first burst dominates the spectro-photometry at the time when the galaxy is observed. Ages measured with single burst delayed exponential SFH are therefore always lower limits to the real ages of galaxies. For a fraction of galaxies in the VUDS sample, their formation redshift could therefore be even higher than what we have measured, pushing the formation even earlier in the reionization epoch.

A second fitting run was then performed on the 8000 simulated galaxies using a fitting library including templates with two bursts SFHs. From this simulation, we observe that when the input galaxy has already experienced two bursts and is on the decaying part of the second burst, the fit also identifies a SFH with two bursts in 80% of the cases, with $\mathrm{Age}_{\rm 2-burst,in} - \mathrm{Age}_{\rm 2-burst,out} = 0.02 \pm 0.20$. When a galaxy has not yet experienced a second peak in its SFH, the fit also identifies that the galaxy is observed before a second peak in ~70% of the cases. In this case, $\mathrm{Age}_{\rm 2-burst,in} - \mathrm{Age}_{\rm 2-burst,out} = -0.07 \pm 0.37$ Gyr. We also see that when the time between the two peaks is important $\Delta t \gtrsim 0.7$ Gyr it is much more difficult to retrieve the right age, as previously concluded. This makes us conclude that fits using a two-burst SFH library are able to recover the right position of a galaxy on the SFH. It is important to realize that these conclusions are linked to the high redshift of the observations considered here. The fact that the age of the Universe is only 1/7th of its present age at $z = 3$ greatly helps limit the time that a galaxy has to form bursts. At lower redshift, where the age of the Universe is much higher, the possible elapsed time in the past of a galaxy is much higher and therefore greatly increases the possible SFH types.

A.2. Age results using two-bursts template libraries on the VUDS data

The VUDS galaxies are fitted with the SFH presented in the previous section with all the other parameters remaining as

presented in Sect. 6.1; we analyse the results in three different categories, depending on the number of peaks in the SFH (0, 1 or 2) that a galaxy experienced during its past history. The results of this fitting run show that 10% of the galaxies did not yet go through a first peak of the SFH, while for 69% they are after the first peak and before a second peak and 21% experienced two peaks and are on the declining part of the second peak. The majority of the VUDS galaxies at these early redshifts are better fit by a single burst SFH than by a simple multi-burst SFH with two peaks. For the remaining 21% of our sample, the galaxies that cross the second peak of the SFH and are on its decreasing part. While the second burst is passed, we can have a look at the typical properties of the two different bursts. The median timescale of the first peak is $\tau_1 = 0.1$ Gyr, which is also the typical timescale of the second burst τ_2 that was experienced by the galaxy. Interestingly, the typical Δt between the two bursts is 0.1 Gyr, which is the smallest possible value available in the two-burst library. This shows that, while the two bursts are passed, the properties of the typical SFH represents the most compressed SFH of the library with two strong and very short bursts separated by the minimum possible time. This is easily understandable since at such high redshifts two more extended bursts cannot fit in the available cosmic time since formation. Finally, in this case, the first burst that the galaxy experienced is 1.5 times stronger than the last burst. Finally, the mean age of these galaxies is ∼1.0 Gyr with a typical mass of $10^{10.13}$ $M_\odot$. Comparing the age of these galaxies with the age of the classical single burst library, we see that the multi-burst library gives an age that is slightly older with $\Delta\mathrm{Age}_{\mathrm{multiburst}} = 0.11 \pm 0.16$ Gyr. Consequently, the galaxies that passed the 2nd burst are also the oldest galaxies in our sample; they are the only galaxies that are old enough to experience two distinguishable bursts. Finally, the comparison of the χ^2_ν between the two libraries shows that the χ^2_ν of the two-burst library are on average higher ($\Delta\chi^2_\nu = 0.7$) than that given by the single burst library. Formally, this means that the latter gives slightly better fits than when we use a two-burst library.

To conclude, when using more complex SFH than a single burst library to fit galaxies at these redshifts, we find small differences in ages. The ages reported in this paper, based on single burst delayed exponential SFH, are therefore robust against major unknowns on the past SFH of these galaxies.