The frequency of very young galaxies in the local Universe: II. The view from SDSS spectra

Gary. A. Mamon1⋆, Marina Trevisan21, Trinh X. Thuan31, Anna Gallazzi4, and Romeel Davé5,6,7

1 Institut d’Astrophysique de Paris (UMR 7095: CNRS & Sorbonne Université), 98 bis Bd Arago, F-75014 Paris, France
2 Universidade Federal do Rio Grande do Sul – Departamento de Astronomia – 91501-970, Porto Alegre-RS, Brazil
3 Astronomy Department, University of Virginia, P.O. Box 400325, Charlottesville, VA 22904-4325
4 INAF – Osservatorio Astrofisico di Arcetri, Largo Enrico Fermi 5, I-50125 Firenze, Italy
5 Institute for Astronomy, Royal Observatory, University of Edinburgh, Edinburgh EH9 3HJ, UK
6 University of the Western Cape, Bellville, Cape Town 7535, South Africa
7 South African Astronomical Observatories, Observatory, Cape Town 7925, South Africa

ABSTRACT

Only a handful of galaxies in the local Universe appear to be very young. We estimate the fraction of very young galaxies (VYGs), defined as those with more than half their stellar masses formed within the last Gyr. We fit non-parametric star formation histories (SFHs) to ∼280,000 galaxy spectra from a flux- and volume-limited subsample of the Main Galaxy Sample (MGS) of the SDSS, which is also complete in mass-to-light ratio, thus properly accounting for passive galaxies of a given mass. The VYG fractions decrease with increasing galaxy stellar mass, from ∼50 per cent at $m = 10^{8} M_{\odot}$ to ∼0.1 per cent at $m = 10^{11.5} M_{\odot}$, with differences of up to 1 dex between the different spectral models used to estimate the SFH and on how we treat aperture effects. But old stellar populations may hide in our VYG sample built with galaxies that are globally bluer than within the region viewed by the SDSS fibre. The VYG fractions versus mass decrease more gradually compared to the Tweed et al. predictions using analytical and semi-analytical models of galaxy formation, but agree better with the SIMBA hydrodynamical simulation. These discrepancies highlight the usefulness of VYGs in constraining the strong uncertainties in both galaxy formation models and spectral modelling of galaxy SFHs. Given the log-normal cosmic SFH, these mean VYG fractions suggest that galaxies with $m > 10^{8} M_{\odot}$ undergo at most 4 major starbursts on average.

Key words: galaxies: evolution – galaxies: dwarf – galaxies: statistics – methods: numerical

1 INTRODUCTION

Among the wide variety of galaxies in the local Universe, a handful of low-mass star-forming galaxies (SFGs) have drawn attention because of their extremely low metal abundance, $12 + \log O/H \leq 7.3$. The blue compact dwarf (BCD) galaxy I Zw 18, first observed spectroscopically by Searle & Sargent (1972), with oxygen abundances $12 + \log O/H \sim 7.17-7.26$ (Skillman & Kennicutt 1993; Izotov & Thuan 1998) long stood as the lowest metallicity SFG known. The most metal-poor dwarf SFG known to date has $12 + \log O/H = 6.98 \pm 0.02$ (Izotov et al. 2018), i.e. 1/50th the Sun’s metallicity if $12 + \log O/H = 8.7$ is adopted for the Sun. In their discussion of I Zw 18, Searle & Sargent (1972) posed the question: are the most metal-deficient BCDs young galaxies, i.e. they formed their first stars at a relatively recent time ($\leq 1$ Gyr ago) or are they old galaxies which formed most of their stars several Gyr ago, in which case the present starburst is just the last episode of a series of such events? For the vast majority of BCDs (more than 99 per cent), the second case appears to be true: deep CCD images clearly show an extended low-surface brightness stellar component, indicative of an older stellar population, whereas the most metal-deficient BCDs, those with $12 + \log O/H \leq 7.3$, do not show such an extended low surface brightness stellar component (e.g. Papaderos et al. 2002). This suggests that
the most metal-deficient galaxies in the local Universe have very young stellar populations, i.e. are very young galaxies (hereafter, VYGs).

Using colour-magnitude diagrams (CMDs) of stars in I Zw 18, resolved through observations with the Hubble Space Telescope (HST), Izotov & Thuan (2004) estimated the bulk of the stellar mass to have formed within the last 500 Myr. Their age was based on the lack of red giant stars. However, the young age of I Zw 18 was disputed by Aloisi et al. (2007) and Contreras Ramos et al. (2011) who obtained deeper CMDs with additional HST observations. They found more red giant stars and attributed an age of at least 1–2 Gyr to I Zw 18. However, it was not clear from their work whether that older stellar population found constituted only a small fraction of the stellar mass of I Zw 18, or accounted for the bulk of it. In the first case, it could still be called a VYG.

In the last few years, there has been an independent development that also suggest young ages, not for low-mass dwarf galaxies like BCDs, but for galaxies with considerably larger stellar masses. While the BCDs have stellar masses in the range $\sim 10^{-8}$M$_\odot$, Dressler et al. (2018) have found a population of galaxies in the redshift range $0.45<z<0.75$ and with stellar masses $>10^9$M$_\odot$, that formed the majority of their stars within $\sim 2$ Gyr of the epoch of observation, which they called Late Bloomers (LBs). These LBs account for $\sim 20$ per cent of $z\sim 0.6$ galaxies with Milky Way masses, with a moderate dependence on mass ($\sim 30$ per cent at half the Milky Way mass and 5 to 10 per cent at masses greater than that of the Milky Way). According to Dressler et al. (2018), major star formation in about 1/5th of the massive galaxies somehow got delayed by billions of years. The LB population declines rapidly at $z<0.3$ and is essentially extinct today for $M>10^{10}$M$_\odot$. These LBs are like VYGs except they have higher stellar masses and are at larger redshifts.

These observations have motivated us to look at the question of the frequency of VYGs in the local Universe. Tweed et al. (2018), hereafter Paper I, have studied the predicted frequency of VYGs at $z=0$, defined to have at least half their stellar mass formed in the last 1 Gyr ($z<0.08$). They have used various models of galaxy formation, four analytical ones and one semi-analytical model (SAM), to predict the VYG fraction as a function of galaxy stellar mass (hereafter ‘galaxy mass’ or simply ‘mass’). In Paper I, we have found that the predicted VYG fraction as a function of galaxy mass can differ between the models, by up to 3 orders of magnitude. While all models predict low fractions of VYGs at stellar masses $m>10^{11}$M$_\odot$, some predict a few percent of VYGs at masses below $10^{10}$M$_\odot$, while another found a peak in the VYG fraction at $10^9$M$_\odot$. Finally, the SAM predicts only 0.01 per cent of VYGs at intermediate masses, increasing somewhat at lower masses.

These wildly discrepant model results have led us to ask whether we can constrain, directly from observations, the fraction of VYGs in the local Universe. We defined VYGs following Paper I to have at least half their stellar mass formed within 1 Gyr from the epoch corresponding to their redshift.

The ideal way to measure a galaxy’s star formation history (SFH) is by combining a high spectral resolution spectrum, covering a wide spectral range, with a colour-magnitude diagram (CMD). Unfortunately, CMDs have been measured in only 186 nearby galaxies.¹ This sample is much too small to study the mass dependence of the VYG fraction, especially if the latter reaches peak values of only a few percent. However, one can derive SFHs of galaxies from their spectral energy distributions (SEDs). In fact, that was the method used by Dressler, Kelson & Abraham (2018) whose work was published during the course of our own. From SEDs defined by broadband optical and infrared (IR) photometry and optical prism observations with $R \approx 30$, those authors derived SFHs for 20 000 galaxies at $0.25<z<0.75$ and with stellar mass $m>10^{10}$M$_\odot$.

In this article, we make use of the spectral database of the Main Galaxy Sample (MGS) of the Sloan Digital Sky Survey (SDSS) Data Release 12 (DR12), comparing the SFHs of over 400 000 galaxies between two SFH codes, using a total of 7 models of single stellar populations. This allows us to estimate the fraction of VYGs as a function of galaxy stellar mass for each spectral model.

We describe our sample in Sect. 2 and explain our methods in Sect. 3. In Sect. 4, we present the fractions of VYGs inferred from the SDSS spectra. These results are summarized and discussed in Sect. 5.

2 DATA AND SAMPLE SELECTION

2.1 Summary

We extracted our sample from the MGS of the SDSS data release 12 (DR12), according to the following criteria:

(i) flux limit: $r^0_{\text{Petro}} \leq 17.77$;
(ii) object spectra obtained with original 3 arcsec fibre;
(iii) redshift range: 0.005 < $z$ < 0.12;
(iv) the STARLIGHT (Cid Fernandes et al. 2005) SFH code does not fail when applied to the object;
(v) object is in the VErsatile SPectral Analysis database (VESPA)² of SFHs;
(vi) single spectrum for each galaxy;
(vii) surface brightness limit: $\mu^0_{r,50} \leq 23.0$;
(viii) stellar mass range: 6 < log$(m/M_\odot)$ < 12.5 for all spectral models;
(ix) reasonable magnitudes: $g_{\text{Petro}} > 0$, $r_{\text{Petro}} > 0$, $i_{\text{Petro}} > 0$;
(x) colours are not extreme: $-1 < (g-i)_{\text{model}} < 2.5$ AND $-1 < (g-i)_{\text{fibrec}} < 2.5$;
(xi) fibre colours are not too different from model colours: $|(g-i)_{\text{fibrec}} - (g-i)_{\text{model}}| < 1$;
(xii) all 6 spectral fits yield $\chi^2 > 0$;
(xiii) redshift is not too large to fail to see passive galaxies: $z < z_{\text{max}}(m)$.

In addition, we allowed ourselves to apply the following criteria to the VYG candidates (but not to the parent sample):

(xiv) galaxy does not contain an Active Galactic Nucleus

¹ According to the NASA Extragalactic Database at http://ned.ipac.caltech.edu/Library/Distances/
² The VESPA database is publicly accessible at http://www-wsfau.roe.ac.uk/vespa.

MNRAS 000, 1–22 (2017)
(AGN, using the curve of Kauffmann et al. 2003 that conservatively separates AGN from SFGs in the Baldwin, Phillips & Terlevich 1981, hereafter BPT, diagram); (xv) fibre colour is not bluer than the model colour: $(g-i)_{\text{fibre}} > (g-i)_{\text{model}}$.

### 2.2 Basic selection criteria

According to Strauss et al. (2002), the MGS sample is defined to be limited in both flux ($r_{\text{Petro}} \leq 17.77$) and surface brightness $\mu_{r,50} = r_{\text{Petro}} + 2.5 \log(2\pi\theta_{50}^2) \leq 24.5$, where $r_{\text{Petro}}$ is the extinction-corrected Petrosian magnitude, while $\theta_{50}$ is the angular half-light radius.

We first selected galaxies, according to our first 3 criteria, with

```
SELECT (...) FROM SpecObj as s, PhotoObj as p WHERE 
s.bestobjid = p.objid 
AND s.instrument = 'sdss' 
AND s.Class = 'GALAXY' 
AND (p.petroMag_r - p.extinction_r) < 17.77 
AND s.z >= 0.005 & s.z <= 0.12
```

This provided us an initial sample of 424,506 galaxies.

The limit in the apparent magnitude (criterion (i)), corrected for Galactic extinction, is the limit of the SDSS MGS. By selecting the ‘sdss’ instrument, we are avoiding the BOSS spectra that employ smaller fibres (criterion (ii)). The lower redshift limit ensures a sufficiently accurate distance based on the redshift (the peculiar velocities contribute negligibly to the redshift), while the upper redshift limit ensures that we are not limited to the most luminous galaxies, given the flux limit (criterion (iii)).

We added extra requirements based on our stellar population synthesis analysis (see Sect. 3). The STARLIGHT SFH code (see Sect. 3) did not provide SFHs (criterion (iv)) for 842 spectra (2 per cent), leaving us with 423,664 galaxies. The STARLIGHT V15 model (see Table 1). The plus signs display the galaxies whose fibre colours are bluer than –1 or redder than 2.5. The crosses show the galaxies whose fibre or model colours fail to meet equation (1). The blue points are roughly Gaussian-distributed for $-1 < (g-i)^{\text{fibre}} < 2.5$. The brown region shows very young galaxies from SDSS spectra (both corrected for Galactic extinction) for the sample built with criteria (i)-(ix). The magenta region in Fig. 1 indicates our colour-filtered sample. We then visually inspected all galaxies whose colours that do not satisfy equation (1), but whose difference between model and fibre colours is greater than 1 magnitude in absolute value. The red points are roughly Gaussian-distributed for $-1 < (g-i)^{\text{model}} < 2.5$, but the fibre colours display much fewer outliers. We visually inspected all galaxies whose colours that do not satisfy equation (1) and found that nearly all were contaminated by neighbouring saturated stars or large galaxies. We thus removed from our sample those galaxies whose model or fibre colours do not obey equation (1). The square shaded box in Fig. 1 indicates our colour-filtered sample. We then visually inspected all galaxies whose model and fibre colours both satisfied equation (1), but differed by over 1 magnitude in absolute value. Most of these galaxies were again contaminated by nearby saturated stars or large galaxies. We thus made a second colour selection, requesting that the absolute difference of model and fibre colours be less than unity (magenta region in Fig. 1). This led us to discard an additional 336 galaxies, leaving us with 405,358 galaxies.

We excluded objects with unreliable photometry, analysing their colours (criteria (x) and (xi)) as follows. Fig. 1 displays the model $(g-i)^{\text{model}}$ colours (corrected for Galactic extinction) as a function of the corresponding fibre colours, as well as the distributions of model and fibre colours. Both the model (brown) and fibre (purple) colours are roughly Gaussian-distributed for $-1 < (g-i)^{\text{model}} < 2.5$.
2.3 Handling of selection effects

Several observational selection effects can influence our analysis of VYG fractions.

2.3.1 Mass-to-light ratio selection effects

When estimating fractions of VYGs, we must also worry that, at a given stellar mass, galaxies with young stellar populations, which have low mass-to-light ratios, will be more luminous. We must thus ensure that the galaxies with old stellar populations are not missed, hence the fraction of VYGs not overestimated because of this selection effect against high mass-to-light ratio galaxies. Since we are measuring VYG fractions as a function of galaxy stellar mass, a doubly complete survey in volume and luminosity is not complete for the passive galaxies. A sample complete in stellar mass would be desirable, but many more galaxies are found by requesting that the sample is complete in mass-to-light-ratio for given mass bins.

For this, we estimated the maximum redshift for which the subsample of galaxies of a given stellar mass is complete in luminosity (or in mass-to-light-ratio) at our flux limit. This maximum redshift for given mass, $z_{\text{max}}(m)$, is estimated in an analogous fashion as Garilli, Maccagni & Andreon (1999), La Barbera et al. (2010), and Trevisan, Mamon & Khosroshahi (2017). Here, we considered a fine grid of maximum redshifts, $z_{\text{max}}$ (in steps of $\Delta z_{\text{max}} = 0.005$), estimated for each one of the 95th percentile in galaxy stellar mass, $m_{95}(r_{\text{lim}})$, in narrow bins of apparent magnitude $r$, and deduced $m_{95}(r_{\text{lim}})$ using a linear fit of $m_{95}(r)$ vs. $r$.

As seen in the inset of Fig. 2, the relation between maximum redshift and stellar mass is a power law, $\log z_{\text{max}} = a_0 + a_1 \log \left(\frac{m}{M_\odot}\right)$, whose coefficients are given in Table 1. This constitutes criterion (xiii). This procedure is illustrated in Fig. 2 for 3 values of $z_{\text{max}}$, where the oblique lines indicate the linear fits to $m_{95}(r)$ vs. $r$ and the big circles show the linear extrapolation for $m_{95}(r_{\text{lim}})$. The inset of Fig. 2 shows the resulting maximum redshift as a function of stellar mass. There are usually no solutions $z_{\text{max}} < 0.005$ (below which we do not trust our stellar masses, given the importance of peculiar velocities that reduce the accuracy of redshifts as distance indicators) for $\log (m/M_\odot) \leq 8$. The fits of equation (2) and Table 1 are thus only adequate for $\log (m/M_\odot) \geq 8$.

Fig. 3 illustrates this procedure in another fashion. It shows how the variation with redshift of the mass-to-light ratios of galaxies in the clean sample are affected by the flux limit. This is shown for different stellar masses (shown as different colours). The 95th percentiles in equal number of the 6 spectral models (criterion (xii)), leaving us with our clean sample of 404 931 galaxies.
The frequency of very young galaxies from SDSS spectra

Table 1. Star formation history models

| SFH code | SSP acronym | reference | stellar library | evol. tracks | IMF model | age bins | metal. bins | a0 | a1 |
|----------|-------------|-----------|-----------------|-------------|-----------|---------|-------------|----|----|
| SL       | BC03        | Bruzual & Charlot (2003) | STELIB | Padova | Chabrier | screen | 15 | 5 | –4.677 | 0.331 |
| SL       | V15         | Vazdekis et al. (2015) | BaSTI | MILES | Kroupa | mixed | 15 | 6 | –4.814 | 0.344 |
| VESPA    | BC03        | Bruzual & Charlot (2003) | STELIB | Padova | Chabrier | mixed | 16 | 4 | –5.218 | 0.391 |
| VESPA    | M05         | Maraston (2005) | Basel | Padova | Kroupa | mixed | 16 | 4 | –5.121 | 0.382 |
| VESPA    | M05         | Maraston (2005) | Basel | Padova | Kroupa | 2-comp | 16 | 4 | –5.202 | 0.390 |

Notes: The columns are: (1): star formation history code (‘SL’ stands for STARLIGHT); (2): single stellar population code; (3): spectral model acronym; (4): reference; (5): stellar library; (6): evolutionary tracks; (7): initial mass function (Chabrier 2003 and Kroupa 2003); (8): dust model (screen, mixed or mixed with additional component around young stars); (9): number of log age bins; (10): number of metallicity bins; (11) and (12): parameters of z – max vs. m (eq. [2]).

bins (circles) show that m/Lr is constant with redshift up to some limiting redshift, then decreases as the flux limit increases. This behaviour is seen for all masses except the lowest mass (10^7 M☉), where constant mass-to-light ratio is not seen at low redshifts, nor at the highest mass in 10^12 M☉, where the flux limit allows for much higher m/Lr values than found.

We fitted a broken power-law to each run of m/Lr vs. redshift, with slopes of 0 and –2. This was obtained with a first estimate of the turning point height using the median log(m/Lr) of the 25 per cent of the lowest redshifts. Our initial estimate of the turning point redshift used Median(log z + 0.5 log m/Lr) – Median(log m/Lr), computed on the 25 per cent highest redshifts. We then minimised \( \chi^2 = \sum [\log(m/L_r)_{\text{data}} - \log(m/L_r)_{\text{model}}]^2 \) for the 2 parameters of the broken line break, allowing for \( \log m/L_r \) and \( z \) to lie within 0.3 from the first guess. The best-fit broken lines are displayed in black in Fig. 3. The inset to the figure shows the best fit values of zmax, the absissa of the breaks of the broken lines in the main figure. One notices that log zmax increases linearly with log m, as log zmax = –4.849 + 0.340 log(m/M☉) for the S_V15 model.

Given the similarity between the fits obtained with these two methods, we adopted the relation of equation (2) with the coefficients shown in Table 1, which we assume extrapolate to \( m < 10^8 M☉ \).

2.3.2 Discarding AGN

Although this is not a selection effect, we allow ourselves to discard AGN from our candidate VYGs (criterion (xiv)). For this, we discard VYG candidates that lie above the Kauffmann et al. (2003) curve of the BPT diagram.

2.3.3 Aperture effects

The finite (3 arcsec) aperture of the SDSS fibres limits the spectra to the inner regions of many galaxies. Fig. 1 shows that the galaxies that are considered VYGs by one of the spectral models (STARLIGHT V15, see Sect. 3) before filtering by colour and the \( \chi^2 \) of the spectral fits) are associated with blue fibre colours. But there is very little relation between global (model) colour and galaxy youth (blue symbols) estimated from the spectrum measured within the fibre. Many (21 per cent of) galaxies in the clean sample have bluer nuclei than their global bodies (lying below the solid green curve of Fig. 1), leading to spectral fits pointing to a younger stellar population than averaged over the galaxy. For example, the galaxy NGC 838 in the HCG 16 compact group, whose stellar mass is \( \approx 5 \times 10^9 M☉ \) (O’Sullivan et al. 2014 and references therein), has a nucleus (the SDSS fibre subtends an angular radius of 400 pc) whose stellar population has an important very young component (younger than 300 Myr), which accounts for nearly half of the stellar mass of the nucleus, depending on the SFH and SSP model (O’Sullivan et al.), while the bulk of the galaxy is older (Vogt et al. 2013). Some of these galaxies with blue nuclei may thus be classified as VYGs even if the bulk of their stellar mass is old.

This aperture effect is potentially serious, since the median fraction of galaxy light subtended by the fibre is only 26 per cent, and only 3 per cent of our sample have fibres covering more than half of the galaxy light. Moreover, building a galaxy sample where the fibres see over, say, 30 per cent of the galaxy light (39 per cent of the sample), leads to an effective cut at higher surface brightness than our limit of \( \mu_r = 23 \), leading us to lose 99.8 per cent of all galaxies with 21.5 < \( \mu_r < 23 \).

Instead, we handle the aperture selection effect by restricting our VYGs to galaxies with blue colour gradients, \( d(g-i)/dR < 0 \) (where \( R \) is the projected distance to the center), i.e. with a global (model) colour that is bluer than the fibre colour. This constitutes criterion (xv).

2.4 Final samples

The sample extracted with criteria (i)-(xii) is our clean sample (404931 galaxies). The sample with criteria (i)-(xiii) is our complete sample (typically 280 000 galaxies, depending on the spectral model). We consider two samples of VYGs: liberal samples (one for each spectral model) and a conservative sample that excludes AGN (criterion (xiv)) as well as galaxies with red colour gradients (criterion (xv)). The two parent samples and the two VYG samples are shown in Table 2.
3 AGES FROM GALAXY SPECTRA

Galaxy SFHs and ages were inferred from the SDSS-DR12 spectra through stellar population synthesis analysis. These spectra have the advantage of fairly high spectral resolution ($R > 2000$), excellent signal to noise (above ~20), wide spectral range (from below 3800 Å to 9200 Å) and they are flux calibrated. They allow to estimate the SFH of each galaxy, by fitting the observed spectrum with the predicted spectrum obtained by the combination of single stellar population (SSP) spectra convolved by the SFH.

We considered three algorithms to estimate the SFH. We used two non-parametric codes: STARLIGHT (Cid Fernandes et al. 2005), which we ran ourselves, and the VESPA database (Tojeiro et al. 2009). Both codes fit non-parametric SFHs and metal histories with an internal extinction, considering typically 16 (roughly equally spaced) bins of log age, as well as 4 to 6 bins of metallicity, and a single extinction (see Table 1). This leads to a total of typically $16 \times 4 + 1 = 65$ free parameters.

We also used the method of Gallazzi et al. (2005), who analysed the SFHs of the 175 128 galaxies in the 2nd data release (DR2) of the SDSS. Gallazzi et al. adopted a Bayesian approach to derive median likelihood estimates of stellar population properties such as $r$-band weighted and mass-weighted mean stellar age, stellar metallicity and stellar mass. For each galaxy, observed absorption features were compared with the predictions from a large library of metallicities and SFHs, described by an exponential declining law on top of which random bursts are added, convolved with BC03 SSP models. The probability distribution functions (from which the 16, 50, 84 percentiles are measured) of the light- and mass-weighted mean ages are obtained by marginalising over all the other parameters of the models. We limited our analysis of the Gallazzi et al. sample to the intersection with our clean and complete galaxy samples discarding the 18 per cent of galaxies with no age measurements, with numbers given in Table 1.

All four algorithms have been run using the Bruzual & Charlot (2003, hereafter BC03) model to produce synthetic spectra of SSPs to fit the observed SDSS spectra. The SSPs were calculated with the Padova 1994 stellar evolution tracks (Bressan et al. 1993; Fagotto et al. 1994a,b; Girardi et al. 1996) and with the Chabrier (2003) initial mass function (IMF). The BC03 model employs the STELIB stellar library (Le Borgne et al. 2003). VESPA has also been run using the SSPs of Maraston (2005, M05) using the Basel stellar library (Lejeune et al. 1998) and the Kroupa (2001) IMF. We also ran STARLIGHT on 424 506 SDSS spectra. STARLIGHT uses the Medium resolution Isaac Newton Telescope Library of Empirical Spectra (MILES, Sánchez-Blázquez et al. 2006), using the updated version 10.0 (Vazdekis et al. 2015, hereafter V15) of the code presented in Vazdekis et al. (2010). These MILES models were computed with the Kroupa (2001) IMF, and stellar evolution tracks from BaSTI (Bag of Stellar Tracks and Isochrones, Pietrinferni et al. 2004, 2006). While STARLIGHT was run assuming a screen dust model, VESPA assumed either a mixed slab interstellar dust model (Charlot & Fall 2000) or combined it with extra dust around young stars (also from Charlot & Fall). Gallazzi et al. (2005) used a screen dust, with the $A^{-0.7}$ extinction curve used later by VESPA.

The advantage of the STARLIGHT model with the V15 SSP is its recent stellar library and evolutionary tracks, while the VESPA models involve a more refined treatment of dust extinction. The age and metallicity bins of the two SFH codes are quite similar: The lowest age bins are 25 Myr for STARLIGHT and less than 20 Myr for VESPA, while the metallicity bins ranged from $Z = 0.0004$ to 0.05 (VESPA BC03) or 0.04 (VESPA M05) or from $Z = 0.001$ to 0.048 (STARLIGHT). Note that the lower limit of metallicity of $Z = 0.001$ corresponds to the limit of “safe” spectral modeling with the MILES based models (see http://research.iac.es/proyecto/miles/pages/ssp-models/safe-ranges.php).

For each galaxy and for each spectral (SFH and SSP) model (see Table 1), we compute the median age, $\text{ages}_5$, when the stellar mass was half its final value, by linear interpolation of the fractional mass as a function of log age, or use the arithmetic mean mass-weighted age for the Gallazzi model. VYGs are defined as galaxies whose median (or mean for Gallazzi) ages are less than 1 Gyr.

3.1 Tests of the spectral models

We tested the VESPA and STARLIGHT spectral models in several ways. We first check that the fraction of VYGs is very high for the bluest fibre colours. Fig. 4 compares the variation of the VYG fraction with fibre colour for the 6 spectral models. One sees that the VESPA M05 model with 1-component dust fails to achieve a high fraction of VYGs for the bluest colours. Only the STARLIGHT V15 model and the VESPA models with 2-component dust reach high fractions of VYGs at the bluest fibre colours, i.e. over 82 per cent of galaxies with $(g - i)^{0.1}_{\text{lim}} < -0.1$ are VYGs according to these 3 models, while less than 45 per cent of them are VYGs according to each of the other 3 models.

We then checked that galaxies with the largest Hα equivalent widths (EWs) have the greatest fractions of VYGs, close to unity. Indeed, EW measures the ratio of line flux over continuum intensity, which is proportional to the ratio of ongoing or very recent star formation (SF) rate over luminosity, and hence roughly proportional to the ratio of recently formed stellar mass over total stellar mass. Therefore, strong EWs are a sign of high fractions of recently formed stars. This interpretation of the EW is over-

---

**Table 2. Galaxy samples**

| Sample  | Ages (Gyr) | Size |
|---------|-----------|------|
| clean   | V15       | VcGaMo5,Gr2 | VcMo5,Gr2 | Gallazzi |
| liberal VYG | < 1      | 29775 | 5974 | 5780 | 303 |
| conserv. VYG | < 1   | 16212 | 3041 | 3011 | 124 |
| complete | all      | 264847 | 287748 | 291317 | 74836 |
| liberal VYG | < 1    | 1461 | 1196 | 1187 | 39 |
| conserv. VYG | < 1   | 649  | 729 | 746 | 15 |

Notes: The clean galaxy sample is filtered against bad measurements, while the complete one is also complete in mass-to-light ratio. The conservative VYG sample is like the liberal one, except it excludes AGN and galaxies with red colour gradients. Only the main spectral models are shown.
simplified, as it depends on the evolution of the stellar mass-to-light ratio, i.e. on the SFH, as well as on the geometrical effects of dust opacity. Nevertheless, according to the STARBURST99 model (Leitherer et al. 1999), assuming a Salpeter (1955) IMF extending to star masses of 100 M⊙ and metallicities of Z⊙/20, EW(Hα) > 500 Å indicates ages lower than 50 Myr for a continuous SFH (fig. 84 of Leitherer et al.), and lower than only 5 Myr for an instantaneous starburst (fig. 83 of Leitherer et al.). These maximum ages are lower than 50 Myr for a continuous SFH (fig. 84 of Leitherer et al.), and lower than only 5 Myr for an instantaneous starburst (fig. 83 of Leitherer et al.).

Figure 4. Fraction of very young galaxies (more than half the stellar mass formed in the last Gyr) versus fibre colour, for the STARLIGHT and VESPA spectral models (using the acronyms given in Table 1), for our clean galaxy sample. The error bars are binomial with Wilson 95 percent confidence (± 1σ) upper limits. The abscissa are slightly shifted for clarity.

Figure 5. Comparison of the Hα equivalent widths between the Portsmouth emission line code (absorption line galaxies have EW = 0) and the MPA/JHU code, which uses the opposite nomenclature (negative EW for emission). Each point is a galaxy (one point in 10 is plotted for clarity) from the clean sample, only keeping galaxies whose EW uncertainties are less than one-third of the absolute value of the EW. 9 galaxies with emission lines according to Portsmouth have absorption lines according to MPA/JHU (thus not shown). The red line denotes equality, while the blue points show all galaxies with very blue fibre colours (g−i)< −0.3.

and assigns astronomically high values of EW (from over 10^4 Å to over 10^5 Å) for another 120 galaxies.

Fig. 5 compares the EWs between the MPA/JHU and Portsmouth for the 304 418 galaxies for which 0.01 Å < EW(Hα)/[Portsmouth] < 5000 Å and for which the absolute value of the EW is at least 3 times its uncertainty for both models. Note that the MPA values use a nomenclature opposite to that of the Portsmouth values: the MPA EWs are negative for emission lines and positive for absorption lines. There is good general agreement between the two measures of Hα EW. In particular, all galaxies with EW(Hα)/[Portsmouth] > 100 Å have −EW(Hα)/[MPA/JHU] > 44 Å. However, the converse is not true: among galaxies with −EW(Hα)/[MPA/JHU] > 100 Å, 12 have EW(Hα)/[Portsmouth] < 44 Å, yet most of them have very blue fibre colours, i.e. (g−i) ≤ −0.3 (thick blue points), suggesting very recent SF (at least in the nucleus). While most such galaxies with very blue fibre colours lie on the ridge of equality (taking into account the different sign conventions) between the two measures, and all of them but one have −EW(Hα)/[MPA/JHU] > 200 Å, several galaxies with blue fibre colours have EW(Hα)/[Portsmouth] < 200 Å, and 7 have EW(Hα)/[Portsmouth] lower than 20 Å. This led us to adopt the MPA/JHU values for the Hα EWs.

Fig. 6 displays the VYG fractions as a function of the
The models with the best behaving VYG fractions as a function of fibre colour and Hα EWs thus appear to be V15 with STARLIGHT, and the 2-component dust BC03 and M05 models with VESPA. We therefore adopt these 3 spectral models in our subsequent analysis of VYGs, and will no longer discuss the other 3 models. We now refer to our adopted spectral models as V15 STARLIGHT, VESPA BC03 and VESPA M05, respectively.

3.2 Comparison between different spectral models

Fig. 7 shows the distribution of median ages of galaxies according to our 3 adopted spectral models. It confirms the well-known downsizing trend of typically lower masses of galaxies with younger stellar populations. Moreover, the figure shows that the lower age bins are over-represented in the clean sample (open histograms) compared to the complete sample built with $z_{\text{max}}(m)$ as described in Sect. 2.3.1 (filled histograms). This effect is especially evident at low mass, where the maximum redshift for completeness is considerably lower (eq. [2], Table 1, Figs. 2 and 3).

Comparing the first three panels of Fig. 7, it seems odd that the two VESPA SFH models, despite their more realistic treatment of dust (a mixed slab instead of a simple screen), assign extremely young ages (less than 80 Myr) to many galaxies at all masses, yet yield only few galaxies with slightly older ages (between 80 and 250 Myr), especially with the BC03 SSP. It is difficult to imagine a universe where the log age distribution of galaxies is so bimodal at all masses. This odd behaviour is not seen in the STARLIGHT V15 model, which shows, within the statistical uncertainties, more continuous distributions of log age for all galaxy masses. The Gallazzi model does not show the extension to very low ages seen in the other models, but this is based on the (arithmetic) mass-weighted mean ages, which are higher than the median ages.

Fig. 8 summarises the differences between the SFH/SSP models by showing the median (over mass bins) of the median age versus galaxy stellar mass. One notices important differences between the models at low galaxy stellar mass. In the clean sample (small points), the 2 VESPA models show a jump in age$_{\alpha}$ at low masses, while the increase is more gradual with STARLIGHT V15. On the other hand, the complete sample shows high ages at low mass with the VESPA models compared to low ages at low mass for the STARLIGHT model. In other words, STARLIGHT V15 predicts a strong downsizing down to low masses, while the downsizing trend is weaker for the VESPA models, with signs of reaching a plateau at low mass (and perhaps even upsizing with VESPA BC03). In any event, the median ages of galaxies are enhanced when switching from the clean sample to the complete sample, especially for the STARLIGHT V15 model, but least so for VESPA M05.

Fig. 8 also displays the mass-weighted arithmetic mean ages for the intersection of our complete sample (with S V15) with the sample extracted from Gallazzi et al. (2005)\(^4\), shown as the brown shaded region.

4 RESULTS

4.1 Fraction of very young SDSS galaxies versus mass

We now arrive at the focus of our analysis: the fraction of VYGs as a function of galaxy stellar mass. Fig. 9 shows the fraction of VYGs versus galaxy mass for the complete sample of SDSS galaxies, using the 3 spectral models, as well as that of Gallazzi et al. (2005). These models are shown for both the liberal and conservative (not AGN & V$_{(g-i)}$ < 0) VYG samples.

For all 3 spectral models, the liberal VYG fractions (open symbols in Fig. 9) decrease with increasing mass. The

\(^{4}\) This should not be confused with the luminosity-weighted ages reported in Gallazzi et al. (2005), which are lower than mass-weighted ages, because young stars are so much more luminous for their mass.

\[ \text{fraction of very young galaxies} \]

\[ \text{against galaxy mass (in log) with clean and complete samples.} \]

\[ \text{The abscissa are slightly shifted for clarity.} \]

\[ \text{Fig. 8 shows the mass-weighted arithmetic mean ages for the intersection of our complete sample (with S V15) with the sample extracted from Gallazzi et al. (2005), shown as the brown shaded region.} \]

\[ \text{4 RESULTS} \]

\[ \text{4.1 Fraction of very young SDSS galaxies versus mass} \]

\[ \text{We now arrive at the focus of our analysis: the fraction of VYGs as a function of galaxy stellar mass. Fig. 9 shows the fraction of VYGs versus galaxy mass for the complete sample of SDSS galaxies, using the 3 spectral models, as well as that of Gallazzi et al. (2005). These models are shown for both the liberal and conservative (not AGN & V$_{(g-i)}$ < 0) VYG samples.} \]

\[ \text{For all 3 spectral models, the liberal VYG fractions (open symbols in Fig. 9) decrease with increasing mass. The} \]

\[ \text{\(^{4}\) This should not be confused with the luminosity-weighted ages reported in Gallazzi et al. (2005), which are lower than mass-weighted ages, because young stars are so much more luminous for their mass.} \]
Figure 7. Distribution of the median (mass-weighted mean for Gallazzi) ages of galaxies in the clean SDSS Main Galaxy Sample in bins of stellar mass, for our four spectral models. The open and filled histograms respectively indicate the clean and complete (i.e. limited to $z < z_{\text{max}}(m)$, using the fit given in eq. [2] and Table 1) samples. Note that there are no galaxies in the complete sample in the lowest mass bin of all 4 spectral models. The intervals between ticks of log frequency are unity (1 dex).

STARLIGHT V15 model gives VYG fractions >3 per cent at low masses ($m < 10^{10} \, M_\odot$) and 50 per cent at the lowest securely complete mass, $m = 10^{8.5} \, M_\odot$ (see Sect. 2.3.1). In comparison, the two VESPA models lead to VYG fractions over 1 per cent at low masses ($m < 10^{10} \, M_\odot$) and to 12 per cent at $m=10^{8.5} \, M_\odot$. However at high masses, the VESPA models produce a slowly decreasing VYG fraction with increasing mass, while the STARLIGHT V15 model gives a sharply decreasing VYG fraction for $m > 10^{10} \, M_\odot$, 1–2 orders of magnitude lower than the VYG fractions found in VESPA models.

As discussed in Sect. 2.3.3, a serious concern is that the SDSS 3 arcsec diameter fibre is often limited to the light from the central regions of extended galaxies. We may thus mistakenly identify a galaxy as very young, while it may be an old galaxy with a very young nucleus. Hence, in Fig. 9, we also plot the fraction of galaxies in the complete sample that are conservative VYG candidates, i.e. that are not AGNs, and with global colours that are bluer than their fibre colours. The liberal VYG sample may be contaminated by AGN and by galaxies that are young according to their fibre colours.

The black points in Fig. 9 represent our most conservative estimate of the VYG fraction. We obtained these ‘3-way VYG’ fractions by requesting that all 3 spectral models agree that the median age of conservative VYG candidates is less than 1 Gyr, leading to only 9 3-way VYGs over all masses. The shaded regions in Fig. 9 show that the fractions of VYGs derived from the Gallazzi model (using mass-weighted ages) are typically 10 times lower than with the STARLIGHT and VESPA models.

Since the SFHs of VYGs are not Gaussian in time but have long high-age tails, the arithmetic mean mass-weighted age will be greater than the corresponding median age. This implies that the VYG fractions defined with arithmetic mean mass-weighted ages should be smaller than with median ages. Fig. 10 confirms this idea using the STARLIGHT

---

\(^5\) Removing AGN from the VYGs turns out to have a negligible effect on the VYG fraction at low and intermediate masses, while it decreases the VYG fraction by less than 0.2 dex at high masses. However, imposing a blue colour gradient has a very small effect at high masses, but a huge one at low masses, as clearly shown in Fig. 9.

---

The frequency of very young galaxies from SDSS spectra
V15 model: the VYG fractions are indeed lower with arithmetic mean mass-weighted ages instead of median ages, in particular at intermediate and high masses for the complete sample, where the decrease reaches a factor 4 (while the VYG fractions obtained with the mass-weighted geometric mean ages are only slightly lower than those found with the median ages). Nevertheless, the Gallazzi model (using mean ages) yields even lower VYG fractions than the STARLIGHT V15 model using similarly defined mean mass-weighted ages.

One question of interest is how the VYG fractions vary when the maximum median age is changed to values below or above 1 Gyr. Fig. 11 illustrates the distribution of median ages for the 3 spectral models and in 3 bins of stellar mass. This allows us to estimate the fraction of galaxies younger than any chosen age. For example, with STARLIGHT V15, the fraction of galaxies younger than 100 Myr is typically one order of magnitude lower than the fraction of galaxies younger than 1 Gyr (our VYGs). In contrast, the VESPA models indicate nearly identical fractions of galaxies younger than 100 Myr and younger than 1 Gyr. This difference is related to the wide gap between 30 Myr and 500 Myr in the distribution of ages seen in Fig. 7 for these models (see the filled histograms corresponding to the complete sample). Finally, whereas the Gallazzi model produces only rare galaxies younger than 1 Gyr, it produces roughly the same fraction of galaxies less than 2 Gyr than the other 2 SFH codes.

5 DISCUSSION

5.1 Summary of results

The spectral analyses of SDSS galaxies with different SFH codes and different single stellar population models lead to different age-mass relations (Fig. 8). STARLIGHT shows downsizing at all masses down to the limit of \( \log (m/M_\odot) = 8.5 \) where we are confident that the data is complete, while VESPA gives a rough plateau in age at masses below \( 10^{10} M_\odot \) and Gallazzi is in between.

The derived VYG fractions fall with increasing mass, with possibly as many as 50 per cent (STARLIGHT V15, liberal) at \( \log (m/M_\odot) = 8.5 \) to less than half a per cent at \( \log (m/M_\odot) = 10.5 \). The VESPA models show a shallower slope (\( \approx -1.5 \)) with stellar mass than STARLIGHT (\( \approx -3 \)). Thus, VESPA models predict that VYGs should occur at high masses, while STARLIGHT V15 does not. The spectral model of Gallazzi et al. (2005) yields VYG fractions that resemble the lower VESPA fractions at low mass and the lower STARLIGHT fractions at high mass.
The frequency of very young galaxies from SDSS spectra

Figure 10. Effect of the definition of age in the fraction of liberal very young galaxies in the complete sample as a function of stellar mass. The galaxy ages are computed as medians (circles), or mass-weighted with geometric means (diamonds) or arithmetic means (triangles), for STARLIGHT V15 (red) and the Gallazzi model (brown). The error bars are binomial with 90 per cent Wilson upper limits. The error bars are binomial with 90 per cent Wilson upper limits. The abscissa are slightly shifted for clarity.

Figure 11. Cumulative distribution functions of median (mass-weighted mean for Gallazzi) age for galaxies in the complete sample, for the four spectral models and for 3 mass bins of width ±0.5 dex. The vertical line illustrates our choice of 1 Gyr for defining VYGs.

Figure 12. Effect of considering mass loss in the fraction of VYGs versus stellar mass, using the STARLIGHT V15 model on the complete sample. The error bars are binomial, with 90 per cent Wilson upper limits. The abscissa are shifted for clarity.

5.2 Caveats

Estimating the fraction of VYGs from SDSS spectra is difficult for several reasons.

(i) The SDSS spectra are limited to the inner regions of the galaxies, hence do not provide a global view of the SFH of each galaxy. We considered this aperture effect in our conservative VYG sample, which discards not only AGN, but also VYG candidates with red colour gradients. The true VYG fractions ought to lie in between our liberal and conservative estimates.

(ii) Individual SFHs of galaxies suffer from degeneracies caused by their different epochs of SF, their metallicities, and their galactic extinction in the inner region sampled by the SDSS fibre. In fact, old stellar populations may not be revealed by the SFH modelling, as we will discuss in Sect. 5.8.

(iii) Our SFHs neglect stellar evolution and the disappearance of high-mass stars exploding as type II supernovae. Indeed, the VESPA code does not consider these effects, and for consistency we have chosen to use the same approach with STARLIGHT. Supernova explosions affect massive stars that are short-lived, thus depleting more the stars formed earlier. Similarly, mass loss from stellar evolution takes time, thus decreases more the stars formed earlier. So both type II supernovae and stellar evolution should result in enhanced VYG fractions. Fig. 12 confirms that mass loss does raise the VYG fractions, but never more than by 0.2 dex for low and intermediate galaxy stellar masses (log(m/M⊙) ≤ 10).

(iv) There is disagreement between different spectral models using the same SFH code (e.g. BC03 vs. M05 with VESPA) and using different SFH codes (e.g. STARLIGHT vs. VESPA).

(v) The spectral models may lead to worse fits for low-mass galaxies, which show the highest fractions of VYGs.
The median reduced \( \chi^2 \) of the STARLIGHT V15 fits is 0.74 for the complete sample instead of unity, suggesting that the uncertainties of the fit are overestimated.

\( \chi^2 \) is plotted versus galaxy stellar mass. The high young galaxy fractions found with STARLIGHT V15, the median age of I Zw 18 should be even lower than 30 Myr.

Regardless of the quality of the fits, it is very difficult to extract SFHs of galaxies with non-negligible young stellar populations, since young stars have so high luminosity/mass ratios compared to older ones.

### 5.3 Comparison to the literature

The only other study known to us in estimating fractions of very young galaxies per bin of stellar mass is that of Dressler et al. (2018). These authors derived SFHs from galaxy SEDs defined by \( r = 20–50 \) prism observations in the optical band, combined with \( ugrizJK \) broadband photometry, following the prescriptions described in Dressler et al. (2016). They used single stellar population models from Conroy, Gunn & White (2009), based on the BaSeL 3.1 library (Lejeune et al. 1998), the Padova isochrones (Marigo & Girardi 2007; Marigo et al. 2008), and an IMF that is virtually identical to that of Chabrier (2003).

Note that Dressler et al. selected their galaxies at 3.6 \( \mu m \), i.e. at typically 2.2 \( \mu m \) in the rest frame. While galaxies evolve more slowly in the near-IR than in the optical \( r \) band used in SDSS, there is still significant evolution of the mass-to-light ratio in the \( K \) band between 1 and 13 Gyr for single stellar populations, by a similar fraction as in the \( r \) band (see fig. 5 of Charlot & Bruzual 1991), but there is no evolution of \( M/L_r \) after 4 Gyr, contrary to \( M/L_r \). Therefore, the sample of Dressler et al. (2018) may be incomplete for high mass-to-light ratio galaxies. However, we noted in Sect. 2.3.1 that this incompleteness is smaller at the high stellar masses probed by Dressler et al. (2018).

As seen in Fig. 14, the young galaxy fractions measured by Dressler et al. decrease with decreasing redshift, as expected, since the redshifts considered are past the cosmic noon (\( z \approx 2 \)), the epoch when the cosmic star SF rate density is highest. The comparison of their fractions with ours involves an extrapolation of their fractions to the low redshifts probed by the complete subsamples of the SDSS/MGS (\( z \approx 0.07 \)). The high young galaxy fractions found with the VESPA BC03 spectral model (also seen in Fig. 11 for I Zw 18 is too close to satisfy our criterion (iii).
The frequency of very young galaxies from SDSS spectra

Figure 14. Comparison of our analysis of the SDSS Main Galaxy Sample with the analysis of Dressler et al. (2018) of a higher redshift sample of the mass dependence of the fraction of ‘young’ galaxies, defined here to have over half their stellar mass formed within 2 Gyr before the epoch of observation, versus redshift in 3 bins of stellar mass. The SDSS fractions (i.e. low redshift) are for the conservative definition using the complete sample, with slightly shifted abscissas for clarity.

Figure 15. Star formation histories for VYGs (blue), intermediate age galaxies (1 < age < 6.1 Gyr, green) and old galaxies (ages > 6.1 Gyr, red) for 3 bins of final stellar mass, for the STARLIGHT V15 model, using the complete sample. These star formation histories account for stellar evolution and the disappearance of massive stars by supernova explosions. The points indicate the median star formation history, while the error bars indicate 16th and 84th percentiles. The vertical dashed lines delimit the maximum allowed median ages for the very young and intermediate galaxy classes.

5.4 Star formation histories

One can go further and ask whether the SFHs of VYGs are such that the distribution of stellar ages has a narrow peak slightly below 1 Gyr or is more spread out. We also explore how these SFHs vary with final galaxy stellar mass, and how varied are the SFHs of VYGs of given stellar mass.

Fig. 15 displays, for 3 bins of z=0 stellar mass, the SFHs of VYGs (blue symbols) as well as of intermediate-age galaxies, which we define as having median ages between 1 and 6.1 Gyr, and of old galaxies (older than 6.1 Gyr). The figure shows the STARLIGHT SFHs, accounting for stellar evolution and the disappearance of massive stars from type II supernova explosions.

At high mass, the VYGs display an intermediate-age (1-4 Gyr) stellar population, and virtually no population younger than 600 Myr, except for an extremely young population of 30 Myr. This is probably an unavoidable artefact of the over-sensitivity of SFH codes to very young stellar populations. At lower masses, the VYGs no longer display the 1-4 Gyr population, but progressively enhance their younger 0.4-0.6 Gyr population. In other words, the SFHs of VYGs, expressed in log age, are positively skewed at high masses and negatively skewed at low masses. One also notices an increase in the diversity of the SFHs of VYGs towards lower galaxy stellar masses. Similar trends are seen for intermediate-age galaxies (green symbols), indicating a general trend of higher age diversity at lower masses (see Fig. 7).

5.5 Comparison to predictions from models of galaxy formation

We now compare the spectral models in this paper with the galaxy formation models described in Paper I (Tweed et al. 2018). There, we had considered three analytical models where the galaxy stellar mass is a function of z=0 halo mass and redshift: one physical model (Cattaneo et al. 2011), improved with a more realistic cut off at low halo masses from the hydrodynamical simulations of Gnedin (2000) [C+G], and two empirical models (Moster et al. 2013 [MNW], Behroozi et al. 2013 [BWC]) that use abundance matching to link galaxy stellar masses with their halo masses. We also considered a model where the galaxy stellar mass growth rate is equal to its halo mass growth rate times a function of the z=0 halo mass and of redshift (Mutch et al. 2013 [MCP]). Finally, we considered the SAM of Henriques et al. (2015), which has the advantage of having environmen-
Figure 16. Median (over each mass bin) of median stellar ages versus galaxy stellar mass, measured in the complete sample of the SDSS Main Galaxy Sample with 3 spectral models (symbols) and predicted in galaxy formation models (lines): the physical model of Cattaneo et al. (2011) with a more realistic truncation at low mass from Gnedin (2000) [C+G] (royal blue), the empirical (abundance matching) models of Moster et al. (2013) [MNW] (olive green) and Behroozi et al. (2013) [BWC] (violet), where all three models the stellar mass a a function of halo mass and redshift; the model of Mutch et al. [MCP] (salmon), which has the stellar mass growth rate as the halo mass growth rate times a function of halo mass and redshift, and the Henriques et al. (2015) [Henriques] SAM (blue-green) run on the Millennium II simulation (Boylan-Kolchin et al. 2009) for centrals only. The solid and dashed lines show the analytical galaxy formation models with the bursty and quiet halo merger schemes, respectively. The orange shaded region highlights very young galaxies.

Fig. 16 compares the age-mass relations of the galaxy formation models derived in Paper I with those of the spectral models (also shown in Fig. 8). The spectral models exhibit a shallower decrease of ages with decreasing stellar mass (i.e. shallower downsizing of mass) between log(m/M⊙) = 10 to 11.5. While several galaxy formation models (C+G, MNW, and Henriques) predict upsizing at masses below 10^9 M⊙, this is not seen in the spectral models run on SDSS galaxies, except in the VESPA BC03 model.

Can the differences in the VYG fractions of SDSS galaxies obtained in Paper I, the only study known to us that uses galaxy formation/evolution models to predict the fractions of VYGs versus stellar mass. The two panels are for the two schemes of halo merging, one that comes with starbursts, and the other one that does not.

One first notices that the different galaxy formation models display a much wider range of predicted VYG fractions than the spectral models do, although in the bursty halo merging scheme the analytical models agree fairly well with one another. It is the SAM that disagrees the most by predicting orders of magnitude fewer VYGs. The VYG fractions predicted by the galaxy formation models vary little with galaxy stellar mass, while the spectral models predict decreasing liberal VYG fractions with increasing stellar mass. One exception is the BWC model.

Figure 17. Fraction of very young galaxies versus stellar mass: constraints from the complete sample of SDSS galaxy spectra (shaded regions and lines) confronted to the predictions from the galaxy formation models (symbols, with same acronyms as in Fig. 16) of Paper I, for bursty (top) and quiet (bottom) halo merging. The upper and lower envelopes of the shaded regions are respectively the 84th percentiles of the liberal VYG sample and the 16th percentiles of the conservative VYG sample (no AGN and blue colour gradients). The lines are the 84th percentiles of the liberal VYG sample. The error bars for the ‘Hen’ (Henriques et al. 2015) SAM and the SIMBA simulation (Dave et al. 2019) are binomial and the upper limits are 1σ. In the first 4 models, the galaxies are weighted, hence the error bars are estimated from 100 bootstraps.
in the quiet merging scheme, which shows a peak in VYG fraction at \( \log (m/M_\odot) = 9.5 \).

The match between the VYG fractions versus mass based on galaxy formation model predictions and on spectral modelling is generally not good. In particular, the STARLIGHT spectral model yields much higher conservative VYG fractions at \( m < 10^9 \)M_\odot than any galaxy formation model predicts. Also, the Henriques SAM predicts at least 100 times lower VYG fractions than even the conservative spectral models. At \( m = 10^{10.5} \)M_\odot, the Henriques SAM predicts 1000 times fewer VYGs than the conservative estimate of the STARLIGHT V15 model. However a closer look reveals that, in the bursty halo merging scheme, the BWC model predicts VYG fractions that match well the VESPA VYG fractions.

Fig. 17 also shows the first predictions of the fractions of VYGs versus stellar mass at \( z = 0 \) from a cosmological hydrodynamical simulation, here the 100 kpc SIMBA simulation (Davé et al. 2019). This simulation has a more refined treatment of AGN feedback for low and intermediate mass galaxies, and reproduces better observational properties of galaxies at \( z = 0 \) than analogous simulations of the same resolution in space and mass (1 h\(^{-1}\)kpc). Moreover, hydrodynamical simulations should be more realistic than the analytical models and the SAM, thanks to their much better spatial resolution and their more refined subgrid physics (in particular the AGN feedback in SIMBA). The VYG fractions found in SIMBA are much higher than in the SAM (up to over 300 times at \( \log (m/M_\odot) = 9 \)), but comparable to the analytical models, and quite similar to the upper envelope of the Gallazzi spectral models. However, the VYG fractions are overestimated at the lowest masses, because galaxies near the lowest resolved mass cannot form early galaxy formation is delayed, as we found in Paper I for the Henriques SAM run on the lower resolution Millennium simulation. In the SIMBA simulation, the SFHs appear to have converged for 256 star particles, corresponding to \( \log (m/M_\odot) = 9.3 \).

### 5.6 Comparison to emission-line galaxies

Strong emission lines in galaxy spectra are another indicator of very young stellar populations. As already mentioned in Sect. 3.1, by measuring the ratio of the line flux over the continuum, the EW of an emission line essentially measures the ongoing SF rate over its integral over time, which to first order is the specific SF rate. According to Leitherer et al. (1999) using the STARBURST99 code, a single burst of SF is characterised by an (H\(\alpha\)) EW that decreases 4 Myr after the burst as \( r^{-4} \), becoming negligible (< 0.1 \AA\) after 1 Gyr. On the other hand, if the SF rate is uniform in time, the EW falls off much more slowly, roughly as \( r^{-1/3} \), leading to values of 30 and 150 \AA, depending on whether the initial mass function is truncated at 30 or 100 M_\odot at the high end.

Fig. 18 displays the fraction of galaxies with high EWs; 1) those with EW > 50 \AA\ and 2) those with EW > 100 \AA. These fractions decrease with increasing mass, roughly in the same way as the fractions of VYGs. In particular, the conservative case of the fraction of galaxies in the complete sample that have high H\(\alpha\) EWs > 50 \AA\, are not AGN, and have blue colour gradients (solid symbols), shows the same plateau at \( \sim 1.5 \) per cent as does the fraction of VYGs with the same constraints in the two VESPA models (Fig. 9).

Fig. 19 compares the ages and H\(\alpha\) EWs of the clean galaxy sample. One notices 3 classes of galaxies in this figure: a class of low EW with high ages (upper left clump of points), a dominant class of moderately high EW with high ages (upper right clump), and a class of somewhat higher EW and low ages (lower clump). Admittedly, the bimodal distribution of galaxy ages (even with STARLIGHT, see Fig. 7) may be an artifact of the spectral modeling of the SFHs. Note that the dominant clump of high-EW old galaxies cannot be AGN, since we had filtered these out of the clean sample for this figure. Fig. 19 displays a clear anti-correlation between ages and EWs (\( r = 0.37 \) from a Spearman rank correlation test, with a null \( p \) value). However, one might have hoped for an even stronger correlation, as can be seen by the disagreement in the trajectories of median age vs EW (brown) and its transpose (green).

Telles & Melnick (2018) analysed a sample of galaxies with EW(H\(\alpha\)) > 50 \AA, EW(H\(\beta\)) > 30 \AA\ and a rectangular selection of the upper left portion of the BPT diagram. They fit parametric three-burst SFHs to the spectral energy distributions of these galaxies from the ultraviolet (UV) to the mid-IR. They found that the ‘young’ stellar population contributes to typically less than 2 per cent of the stellar mass of their emission-line galaxies, and never more than 8 per
cent. However, their ‘young’ population is defined to be less than 10 Myr old, and unfortunately they do not provide the fraction of stellar mass in galaxies in their ‘intermediate’ age population which spans 100 Myr to 1 Gyr in age.

5.7 Number of very young galaxies in the SDSS Main Galaxy Sample

Given the non-negligible fractions of VYGs, how many VYGs may lurk in the SDSS Main Galaxy Sample? Fig. 20 displays the numbers of VYGs in the clean sample for the different spectral models, as well as the numbers of galaxies that are VYGs for all three spectral models (black bars).

The SDSS spectra analysed by STARLIGHT and VESPA lead to thousands of conservative VYGs in the clean samples of the SDSS/MGS (16211, 3040, and 3010 for the STARLIGHT V15, VESPA BC03, and VESPA M05 spectral models, respectively). These spectral estimates lead to peaks in the number of VYG galaxies in the SDSS/MGS (16211, 3040, and 3010 clean MPA/JHU) for the non-AGN galaxies in the clean sample. Only 1 random galaxy among 20 is shown, while the contours (in steps of 0.5 dex in density – and 2 intermediate ones shown by dashes – estimated in 40x40 cells) are for the full galaxy sample (the limits of the figure correspond to the limits of the data, after discarding systems with –EW(Hα) < 0). The horizontal blue line corresponds to a median age of 1 Gyr (7th percentile), while the vertical purple line corresponds to –EW(Hα) = 41 (93rd percentile). The brown circles show the medians of the median ages in equal number bins of EW (1000/bin), while the green circles show the medians of EW in equal number bins of median age.

5.8 Can VYGs contain hidden old stellar populations?

The higher fraction of VYGs found in our analysis of SFHs derived from SDSS spectra, in comparison with the predictions from nearly all galaxy formation models (Fig. 17), suggest that old stellar populations may be present in VYGs, but are hidden from our view and do not contribute strongly to the SED. By selecting conservative VYGs that have bluer global colours than their fibre colours, we have attempted to eliminate galaxies that possess an old extended stellar population underlying the young stellar population. However, the old stellar population may have no effect on the global colour within the extent of the galaxy in the SDSS image, and yet become increasingly important with radius, dominating the integrated stellar mass at large radii.

We examine this situation by considering the following simple model. Assume that the galaxy is composed of two stellar populations, one young and one old, which both follow Sérsic surface brightness profiles,

$$\Sigma(\theta) = \Sigma(0) \exp \left[ -b(n) \left( \frac{\theta}{\theta_e} \right)^{1/n} \right],$$

where \( b(n) = 2 n - 1/3 + 0.009876/n \) (Prugniel & Simien 1997), with effective (half-light) radii, \( \theta_{\text{young}} \) and \( \theta_{\text{old}} \), and indices \( n_{\text{young}} \) and \( n_{\text{old}} \).
Figure 21. Ratios of old over young total masses, \( m_{\text{old}} / m_{\text{young}} \) (eq. [6]) as a function of the Sérsic index of the old population, \( n_{\text{old}} \) and the ratio of effective radii \( \theta_{\text{old}} / \theta_{\text{young}} \), assuming \( n_{\text{young}} = 1 \) and \( \theta_{\text{young}} = 3'' \). The white curve denotes a ratio of unity.

and \( n_{\text{old}} \) and with very different mass-to-light ratios, \( T_{\text{young}} \) and \( T_{\text{old}} \), both assumed to be independent of radius. Our model includes the additional constraint that each population contributes to half the total stellar mass within the fibre. This is a conservative assumption, since our VYGs are selected to contain at least half their stellar mass younger than 1 Gyr. We also assume that the young population has a Sérsic index \( n_{\text{young}} = 1 \) as expected for exponential disks and choose \( \theta_{\text{young}} = 3'' \) (close to the median of the conservative VYGs of 2.9 arcsec in the clean sample).

Given that the integrated stellar mass (or luminosity), in cylindrical apertures, of the Sérsic model follows (see Graham & Colless 1997)

\[
m(\theta) = \frac{2 \pi n}{b^n} \gamma \left( 2n, b(n) \left( \frac{\theta}{\theta_c} \right)^{1/n} \right) \int_0^\infty \theta^2 d\theta_c,
\]

where \( \Sigma \) is the central surface mass density, \( d \) is the (cosmological angular) distance and \( \gamma(a, x) \) is the lower incomplete gamma function, the ratio of total old mass over total mass is

\[
m_{\text{old}} \quad m_{\text{young}} = \frac{m(\theta_{\text{fib}})}{m(\theta_{\text{fib}})} \left[ \frac{\theta_{\text{old}}}{\theta_{\text{young}}} \right]^{1/n_{\text{old}}},
\]

\[
P(\theta) \left[ 2n_{\text{young}}, b(n_{\text{young}}) \left( \theta_{\text{fib}}/\theta_{\text{young}} \right)^{1/n_{\text{young}}} \right] \int_0^\infty \theta^2 d\theta_c.
\]

where \( P(a, x) = \gamma(a, x)/\Gamma(a) \) is the lower regularised gamma function and where equation (5) comes from the equal contributions of the young and old populations to the mass within the fibre.

Fig. 21 indicates that the ratio of total masses of the old versus young populations can be much larger than unity for sufficiently large effective radius of the old stellar population (the region above the white curve).

Could this old stellar population hide in our conservative VYG sample whose galaxies have bluer global colours than their fibre colours? Since we assume that the mass-to-light ratio of each stellar population is independent of the aperture, their \((g-i)^0\) colours are also constants. Therefore, the integrated colour of the mix of the two populations at aperture \( \theta \) is

\[
(g-i)^0(\theta) = -2.5 \log \left[ \frac{\gamma_{\text{young}} / \gamma_{\text{old}} \left( \theta / \theta_{\text{old}} \right)^{1/n_{\text{old}}} \left( \theta / \theta_{\text{young}} \right)^{1/n_{\text{young}}} \right] = -2.5 \log \left[ R(\theta) \right],
\]

\[
\text{where}
\]

\[
R(\theta) = \frac{1 + A_i(\theta) \text{dex} \left[ -0.4 (g-i)^0_{\text{young}} \right]}{1 + A_i(\theta)}.
\]

In equation (8), \( A_i(\theta) \) is the ratio of old to young i-band luminosities at the aperture of angle \( \theta \):

\[
A_i(\theta) = \frac{L_{\text{old}}(\theta)}{L_{\text{young}}(\theta)} = \frac{\gamma \left( 2n_{\text{old}}, b(n_{\text{old}}) \left( \theta / \theta_{\text{old}} \right)^{1/n_{\text{old}}} \right)}{\gamma \left( 2n_{\text{young}}, b(n_{\text{young}}) \left( \theta / \theta_{\text{young}} \right)^{1/n_{\text{young}}} \right)}.
\]

for our adopted Sérsic models and our usual assumption of equal young and old contributions to the stellar mass within the fibre.

We illustrate our toy model by adopting the STARLIGHT V15 spectral model, assuming that the young and old stellar populations have respective ages of 1 and 10 Gyr, both with solar metallicity. Table 3 displays the i-band mass-to-light ratios and \((g-i)^0\) colours of the young and old populations for the V15 single stellar population model, determined by cubic spline interpolation of both log \( T \) and \((g-i)^0\) vs. (log) metallicity in the table for the BaSTI isochrones and the Kroupa Universal IMF (see Table 1) in Magnitudes, colours and mass-to-light ratios for the SDSS filters (AB system).8

Fig. 22 shows that, contrary to intuition, the mix of a young stellar population with an old one does not necessarily lead to a red colour gradient. This can be understood by the differences in the low and high index Sérsic models: in a log-log plot the surface brightness profile of an \( n = 1 \) Sérsic model is flat in the inner regions and falls rapidly at large

Table 3. Parameters of two-population model

| Population  | Age (Gyr) | Metallicity | \( n \) | \( \theta_c \) (Solar) | \( \gamma_i \) | \( (g-i)^0 \) |
|-------------|----------|-------------|------|-----------------|----------|-------------|
| Young       | 1        | Solar       | 1    | 3               | 0.56     | 0.52        |
| Old         | 10       | Free        | 2.5  | 1.16            | 2.52     | 1.16        |

Notes: The columns are: 1) population; 2) age; 3) Sérsic index; 4) effective radius; 5) mass-to-light ratio in the i band; 6) \((g-i)^0\) colour.

8 http://www.iac.es/proyecto/miles/pages/photometric-predictions-based-on-e-miles-seds.php
nential profiles out to 4 effective radii\(^9\)) and the fibre colour, again as a function of the old population Sérsic index and of the ratio of old to young effective radii. Fig. 23 clearly shows that there is a range of effective radii of the old population that allows this old population to dominate the total mass (between the thick white curve and either of the yellow curves) without making the colour redder than \((g−i)_\text{fibre}\). The region between the white and solid yellow curves only allows bluer global colours, while extending to the dashed yellow curve allows slightly redder global colours (by 0.05 mag), consistent with typical photometric errors. For example, if \(n_{\text{old}} = 3\), one can have an old stellar population if \(1.5 < \theta_{\text{old}}/\theta_{\text{young}} < 2.1\) (i.e. 0.18 to 0.32 in log) for no red gradient, or up to 4.9 for \(0 < (g−i)(\theta_{\text{max}})−(g−i)(\theta_{\text{fib}}) < 0.05\), if we allow for photometric errors. The same model applied to the VESPA models, using the same online calculator to recompute the colours and mass-to-light ratios, yields very similar curves as in Fig. 23.

We conclude that forcing a blue gradient does not rule out an underlying old stellar population of high Sérsic index coupled with an effective radius that is greater but not very much greater than that of the young stellar population. Note that in high EW galaxies, the colours are strongly affected by the emission lines, so that bluer does not necessarily mean younger, as is clear in I Zw 18 (Papaderos & Östlin 2012). This reinforces the possibility of an old stellar population hiding in our conservative sample of VYGs.

We finally ask how deep will observations be required to detect a possible underlying old stellar population. Fig. 24 shows the colour and surface brightness profiles for the young stellar population and mixtures of this young population with four extreme old stellar populations, two with Sérsic index \(n_{\text{old}} = 2\) and two with \(n_{\text{old}} = 4\). For each Sérsic index, we adopt the extreme values for the effective radii, \(\theta_{\text{old}}\), lying on the thick white (minimum effective radius for old population dominating the mass) and yellow (maximum effective radius to allow for a blue gradient up to 4 \(\theta_{\text{young}}\)) curves of Fig. 23. Assuming 0.1 mag photometric uncertainties for very deep observations, colours will have 0.14 mag uncertainties, and a 2σ detection of a red colour gradient will require a redward shift of 0.3 mag, i.e. an outer colour of \((g−i) = 0.9\) given the typical inner colour of \((g−i) = 0.6\). The top panel of Fig. 24 indicates \((g−i) = 0.9\) is reached between 12 and 15 arcsec. Reading off the corresponding values for the surface brightness in the bottom panel of Fig. 24, we deduce that observations reaching a depth in the range from 26 mag arcsec\(^{-2}\) (\(n_{\text{old}} = 4\), largest possible \(\theta_{\text{old}}\)) to 28 mag arcsec\(^{-2}\) (\(n_{\text{old}} = 2\), smallest possible \(\theta_{\text{old}}\)) will be required to detect the possible hidden old stellar population. But Sérsic indices of \(n = 4\) pertain to giant ellipticals, and a hidden old stellar population is likely to be less luminous and with a lower index, making it more difficult to detect (the surface brightness will need to reach 27 mag arcsec\(^{-2}\)). Repeating this analysis allowing for the 2σ detection of a 0.1 mag redward shift (instead of 0.3 mag) brings the required surface brightness values to 25 to 27 mag arcsec\(^{-2}\) (26 to 27 mag arcsec\(^{-2}\) for \(n_{\text{old}} \leq 3\)).

\(^9\) https://www.sdss.org/dr12/algorithms/magnitudes/
5.9 The number of major starbursts and the reliability of the VYG classification

Can the VYG fractions constrain the SFHs of galaxies? The most realistic assumption is that the distribution of lookback times should follow the cosmic star formation history (CSFH). We first note the CSFH (denoted \( \psi \)) is very close to a log-normal in \( 1+z \), i.e. \( \psi(z) \approx \psi_{\text{max}} \exp[-(\log(1+z) - \mu)^2/(2\sigma^2)] \). This is shown in Figure 25 for our best-fit \( \mu = 0.44 \) and \( \sigma = 0.21 \) (with a peak CSFR \( \psi = 0.13 M_\odot \text{yr}^{-1}\text{Mpc}^{-3} \) at \( z = 10^{0.44} - 1 = 1.76 \), with 84 per cent of the CSFR occurring at redshifts \( 0.7 < z < 3.5 \)). The lognormal CSFH produces a slightly smaller \( \chi^2 \) (in log \( \psi \)) than the fit of Madau & Dickinson (2014), despite having one less parameter (leading to a reduced \( \chi^2 \) of 1.46 instead of 1.58).

The mean redshift of our complete sample is \( z_{\text{obs}} = 0.07 \), i.e. a lookback time of 0.93 Gyr. This means that the critical (maximum) redshift for our galaxies to be VYGs is the redshift corresponding to lookback time of 1.93 Gyr, i.e. \( z_{\text{crit}} = 0.154 \). According to the lognormal CSFH, we expect the probability \( p_{\text{young}} \) that a star in a galaxy at \( z = 0.07 \) is younger than 1 Gyr will be

\[
p_{\text{young}} = P(z_{\text{obs}} < z < z_{\text{crit}} | z > z_{\text{obs}}) = \frac{\int_{z_{\text{obs}}}^{z_{\text{crit}}} \psi(z) / [(1+z)E(z)] \text{d}z}{\int_{z_{\text{obs}}}^{\infty} \psi(z) / [(1+z)E(z)] \text{d}z} = 0.026. \tag{10}
\]

In one extreme situation, suppose that all galaxies have single burst SFHs. If we draw random SF redshifts from our lognormal CSFH, we expect that VYGs will comprise 2.6 per cent of our sample of galaxies (eq. [10]). At the other extreme, if the SFH of galaxies is continuous, the probability of \( > 10^8 \) (unrelated) stars to be younger than 1 Gyr, will be negligible. More precisely, if each star has a probability \( p_{\text{young}} \) of being younger than 1 Gyr, the probability of \( N \) unrelated stars or starbursts (of equal mass) leading to at least half the stellar mass being younger than 1 Gyr will be the cumulative distribution function of the binomial distribution

\[
f_{\text{young}} = \sum_{k=k_{\text{min}}}^{N} \binom{N}{k} p_{\text{young}}^k (1-p_{\text{young}})^{N-k}, \tag{11}
\]

where \( k_{\text{min}} = N/2 \) and \( (N+1)/2 \) for even and odd \( N \), respectively. As shown in Figure 26 (symbols), Eq. (11) yields a VYG fraction of \( 2p_{\text{young}} (1-p_{\text{young}})/2 = 5 \) per cent for \( N = 2 \) starbursts, but falls very rapidly to 0.4 per cent for \( N = 4 \) starbursts, \( 3 \times 10^{-4} \) for \( N = 6 \) starbursts (and below \( 10^{-10} \) for \( N = 20 \) bursts of SF).

We compare these predicted VYG fractions with the
mean fractions predicted from a mass-limited sample:

$$\langle f_{\text{young}} \rangle = \frac{\int_{m_{\text{min}}}^{\infty} \phi(m) f_{\text{young}}(m) \, dm}{\int_{m_{\text{min}}}^{\infty} \phi(m) \, dm},$$

(12)

where $f_{\text{young}}(m)$ is the VYG fraction that we obtained for the complete SDSS/MGS sample and $\phi(m)$ is the $z = 0.07$ cosmic stellar mass function. Adopting the double Schechter stellar mass function of Baldry et al. (2008) and integrating down to minimum galaxy stellar mass $m_{\text{min}} = 10^8 M_\odot$, we find liberal mean VYG fractions of 32, 9 and 3 per cent for STARLIGHT V15, VESPA and Gallazzi, respectively, and conservative fractions of 8 per cent for STARLIGHT V15 and 1 per cent for the VESPA and Gallazzi models, as shown in Figure 26 (shaded regions). The geometric means of these liberal and conservative mean VYG fractions are 16, 3, and 2 per cent for per cent with the STARLIGHT, VESPA, and Gallazzi models, respectively. The VYG classification with STARLIGHT is, at best, 30 per cent reliable for the 2-burst model ($f_{\text{young}} = 0.05$), while with the other spectral models it is either incomplete and fully reliable for the 2-burst model, or nearly complete and reliable for the single-burst model. For 4 major starbursts, the VYG classification would have reliability of only 2.5, 13 and 20 percent for STARLIGHT, VESPA and Gallazzi, respectively. We now focus on STARLIGHT, which predicts $\sim 8$ times more frequent conservative VYGs than the other two models. Adopting the ansatz that the conservative VYG classification of STARLIGHT is only 1 per cent reliable, we deduce that $m > 10^8 M_\odot$ galaxies undergo on average at most 4 major starbursts.

This discussion is based on the full galaxy population. It cannot be applied to specific populations. It suggests that the decrease of the fraction of VYGs with increasing galaxy stellar mass is a natural outcome of the more bursty SFHs of low mass galaxies (as seen in cosmological hydrodynamical simulations, e.g. Tollet et al. 2019).

Figure 26. Comparison of prediction of mean fractions of very young galaxies of stellar mass $m > 10^8 M_\odot$ predicted by the cosmic star formation history (from eq. [11], filled circles) and obtained from SDSS (from eq. [12], shaded regions extending from conservative to liberal).

5.10 Concluding remarks

Our galaxy SFHs from SDSS spectra using different spectral models all indicate that the fraction of VYGs decrease with stellar mass. Nevertheless, the fraction of VYGs remains an unsolved problem for two reasons: 1) the wide variation in estimates of VYG fractions, both among the spectral models, and among the galaxy formation models; 2) the possibility of our missing old stellar populations. Turning this around, we believe that the very strong sensitivity of the fraction of VYGs, to the galaxy formation modelling on one hand (and the number of major starbursts per galaxy) and to the spectral modelling of observed spectra on the other hand, make the study of the mass variation of the fraction of VYGs a strong constraint for both models of galaxy formation and evolution and for spectral models.

It will be worthwhile to confirm the existence of VYGs (some nearly as massive than our Milky Way), in several ways:

(i) Analyze the spectra of the VYG candidates with Bayesian modelling (e.g. BEAGLE, Chevallard & Charlot 2016), for example to penalise against decreasing metal histories.

(ii) Perform spectral fits to stacked spectra of VYG candidates.

(iii) Analyze jointly SDSS spectra and photometry in the far- and near-UV from the Galaxy Evolution Explorer (GALEX) and in the near-IR from the Wide-field Infrared Survey Explorer (WISE), accounting for aperture effects (e.g. with BEAGLE).

(iv) Obtain integral field spectroscopy of a large statistical sample of the VYG candidates to test whether they are globally young or if their youth is restricted to their inner regions probed by the SDSS fibres (instead of relying on the blue colour gradient as a way to avoid aperture effects, since the presence of such a gradient is insufficient to exclude old stellar populations).

(v) Obtain deep images of VYG candidates, e.g. with E-ELT, to search for low-surface brightness, extended old stellar populations that may dominate the stellar mass (as Papaderos et al. 2002 have done for I Zw 18).

(vi) Stack hundreds or thousands of images of VYG candidates to achieve comparable depths to search for old stellar populations.

(vii) Obtain deep near-IR images of nearby VYG candidates with the future James Web Space Telescope (JWST) to derive colour-magnitude diagrams of resolved stars deeper than previously done with HST.

While we cannot be certain that the majority of the thousands of VYGs that we find in the SDSS MGS truly have median ages below 1 Gyr, these VYGs may still hold clues to important recent star formation. Indeed, in a forthcoming article (Trevisan et al. in prep.), we will highlight the specific properties that distinguish VYGs from other star forming galaxies and discuss the impact of galaxy mergers in the formation of VYGs at low redshift. Moreover, we have begun radio observations at GMRT and VLA to understand the kinematics and accretion history of neutral gas around VYGs.

An electronic table of the galaxies that are VYGs according to one of our 3 preferred spectral models is provided in electronic form, with the top 10 lines shown in Table A1.
ACKNOWLEDGMENTS

We thank Andrea Cattaneo, Daniel Kunth, Sophia Lian-nou, and Polychronis Papaderos for enlightening discussions and the referee, Erik Tollerud, for his numerous constructive comments that strengthened this article. M.T. and T.X.T. are grateful to the hospitality of the Institut d’Astrophysique de Paris, where a large part of this work was performed. G.A.M. acknowledges the Brazilian CNpq (grant #451451/2019-8) and the Universidade Federal do Rio Grande do Sul for its hospitality. We are grateful to Roberto Cid Fernandes for making his STARLIGHT code publicly available and Rita Tojeiro for making the VESPA output publicly available.

This research has made use of the NASA/IPAC Extragalactic Database (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions.

REFERENCES

Aloisi A., et al., 2007, ApJ, 667, L151
Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, MNRAS, 398, 1150
Bressan A., Fagotto F., Bertelli G., Chiosi C., 1993, A&AS, 100, 647
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Cattaneo A., Mamon G. A., Warnick K., Knebe A., 2011, A&A, 533, A5
Chabrier G., 2003, PASP, 115, 763
Charlot S., Bruzual G., 1991, ApJ, 367, 126
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Chevallard J., Charlot S., 2016, MNRAS, 462, 1415
Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, MNRAS, 358, 363
Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
Contreras Ramos R., et al., 2011, ApJ, 739, 74
Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827
Dressler A., Kelson D. D., Abramson L. E., 2018, ApJ, 869, 152
Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994a, A&AS, 104, 365
Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994b, A&AS, 105, 29
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Garilli B., Maccagni D., Andreon S., 1999, A&A, 342, 408
Girardi L., Bressan A., Chiosi C., Bertelli G., Nasi E., 1996, A&AS, 117, 113
Gnedin N. Y., 2000, ApJ, 542, 535
Graham A., Colless M., 1997, MNRAS, 287, 221

Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Izotov Y. I., Thuan T. X., 2004, ApJ, 616, 768
Izotov Y. I., Guseva N. G., Thuan T. X., 2011, ApJ, 728, 161
Izotov Y. I., Thuan T. X., Guseva N. G., Liss S. E., 2018, MNRAS, 473, 1956
Kauffmann G., et al., 2003, MNRAS, 341, 33
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Tremena J., 2001, ApJ, 556, 121
Kroupa P., 2001, MNRAS, 322, 231
La Barbera F., de Carvalho R. R., de La Rosa I. G., Lopes P. A. A., Kohl-Moreira J. L., Capelatto H. V., 2010, MNRAS, 408, 1313
Le Borgne J.-F., et al., 2003, A&A, 402, 433
Leitherer C., et al., 1999, ApJS, 123, 3
Lejeune T., Cuisinier F., Buser R., 1998, A&AS, 130, 65
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Maraston C., 2005, MNRAS, 362, 799
Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785
Marigo P., Girardi L., 2007, A&A, 469, 239
Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883
Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121
Mutch S. J., Croton D. J., Poole G. B., 2013, MNRAS, 435, 2445
O’Sullivan E., et al., 2014, ApJ, 793, 73
Papaderos P., Östlin G., 2012, A&A, 537, A126
Papaderos P., Izotov Y. I., Thuan T. X., Noeske K. G., Fricke K. J., Guseva N. G., Green R. F., 2002, A&A, 393, 461
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2006, ApJ, 642, 797
Prugniel P., Simien F., 1997, A&A, 321, 111
Salpeter E. E., 1955, ApJ, 121, 161
Sánchez-Blázquez P., et al., 2006, MNRAS, 371, 703
Sarzi M., et al., 2006, MNRAS, 366, 1151
Strauss M. A., et al., 2002, AJ, 124, 1810
Telles E., Mennick J., 2018, A&A, 615, A55
Thomas D., et al., 2013, MNRAS, 431, 1383
Tojeiro R., Wilkins S., Heavens A. F., Panter B., Jimenez R., 2009, ApJS, 185, 1
Tollet É., Cattaneo A., Macciò A. V., Dutton A. A., Kang X., 2019, MNRAS, 485, 2511
Tremonti C. A., et al., 2004, ApJ, 613, 898
Trevisan M., Mamon G. A., Khosroshahi H. G., 2017, MNRAS, 464, 4593
Tweed D. P., Mamon G. A., Thuan T. X., Cattaneo A., Dekel A., Menci N., Calura F., Silk J., 2018, MNRAS, 477, 1427 (Paper I)
Vazdekis A., Sánchez-Blázquez P., Falcón-Barroso J., Cenarro A. J., Beasley M. A., Cardiel N., Gorgas J., Peletier R. F., 2010, MNRAS, 404, 1639
Vazdekis A., et al., 2015, MNRAS, 449, 1177
Vogt F. P. A., Dopita M. A., Kewley L. J., 2013, ApJ, 768, 151
Table A1. Table of galaxies that are very young according to at least one of our three preferred spectral models

| ID | RA deg J2000 | Dec | plate | MJD | fiber | Compl. blue gradient | AGN | log(m/M⊙) age00 (Gyr) | log(m/M⊙) age40 (Gyr) | log(m/M⊙) age90 (Gyr) | SFH (SL V15 - no evol) 0–1 Gyr | 1–6.1 Gyr | >6.1 Gyr | SFH (SL V15 - evol) 0–1 Gyr | 1–6.1 Gyr | >6.1 Gyr |
|----|-------------|-----|-------|-----|------|-----------------------|-----|-----------------------|-----------------------|-----------------------|-----------------------------|-----------|----------|-----------------------------|-----------|----------|
| 1  | 0.0065      | −0.0926 | 387 | 51791 | 186  | Y   | N   | T      | 10.96    | 5.84 | 10.67 | 10.33 | 10.70 | 11.43 | 0.00 | 0.56 | 0.44 | 0.00 | 0.57 | 0.53 |
| 2  | 0.0083      | 15.6972 | 750 | 52235 | 456  | Y   | Y   | T      | 11.04    | 6.58 | 10.91 | 9.44 | 10.63 | 7.06 | 0.01 | 0.43 | 0.56 | 0.01 | 0.45 | 0.51 |
| 3  | 0.0087      | 15.8817 | 750 | 52235 | 459  | Y   | Y   | N      | 10.05    | 12.93 | 10.20 | 3.28 | 10.25 | 11.33 | 0.08 | 0.04 | 0.89 | 0.12 | 0.04 | 0.88 |
| 4  | 0.0118      | 14.7155 | 750 | 52235 | 174  | N   | N   | N      | 9.62     | 1.36 | 9.63 | 3.31 | 9.66 | 2.74 | 0.42 | 0.36 | 0.21 | 0.51 | 0.34 | 0.20 |
| 5  | 0.0134      | −1.1130 | 685 | 52203 | 232  | N   | Y   | N      | 8.75     | 0.79 | 8.92 | 7.23 | 8.74 | 2.67 | 0.95 | 0.04 | 0.01 | 0.96 | 0.03 | 0.01 |
| 6  | 0.0139      | −10.7211 | 650 | 52143 | 211  | Y   | Y   | T      | 11.78    | 12.00 | 10.99 | 5.12 | 10.91 | 8.19 | 0.01 | 0.18 | 0.81 | 0.01 | 0.20 | 0.79 |
| 7  | 0.0144      | 14.1982 | 750 | 52235 | 92   | Y   | Y   | T      | 11.61    | 7.43 | 11.23 | 9.61 | 11.17 | 9.07 | 0.01 | 0.38 | 0.61 | 0.01 | 0.42 | 0.57 |
| 8  | 0.0174      | −8.7342 | 650 | 52143 | 446  | Y   | Y   | T      | 11.27    | 9.31 | 10.95 | 10.89 | 10.83 | 7.05 | 0.00 | 0.21 | 0.79 | 0.00 | 0.22 | 0.78 |
| 9  | 0.0192      | −8.9438 | 650 | 52143 | 447  | N   | Y   | N      | 8.70     | 0.47 | 9.31 | 5.35 | 8.82 | 3.89 | 0.74 | 0.26 | 0.00 | 0.79 | 0.21 | 0.00 |
| 10 | 0.0198      | 0.7817  | 387 | 51791 | 446  | Y   | Y   | T      | 11.92    | 8.64 | 11.53 | 11.24 | 11.57 | 11.09 | 0.00 | 0.21 | 0.79 | 0.00 | 0.22 | 0.77 |

Notes: the columns are: (1): unique identifier; (2) and (3): equatorial coordinates; (4), (5) and (6): SDSS plate, MJD and fiber; (7): is galaxy in complete sample? (8): does galaxy have a blue colour gradient? \((g−i)_0^\text{model} < (g−i)_0^\text{fiber}\); (9): is galaxy an AGN: 'Y' (yes, above Kewley et al. 2001 curve), 'T' (transition between curves of Kewley et al. and Kauffmann et al. 2003), 'N' (no, below Kauffmann et al. curve); (10): stellar mass according to STARLIGHT with V15 model; (11): median age according to STARLIGHT with V15 model; (12): stellar mass according to VESPA with BC03 2-component dust model; (13): median age according to VESPA with BC03 2-component model; (14): stellar mass according to VESPA with M05 2-component dust model; (15): median age according to VESPA with M05 2-component dust model; (16): median age according to VESPA with M05 2-component model; (17): star formation history in respective bins of 0–1 Gyr, 1–6.12 Gyr, and over 6.12 Gyr, using STARLIGHT with the V15 model, with no correction for stellar evolution; (18) to (21): star formation history in respective bins of 0–1 Gyr, 1–6.12 Gyr, and over 6.12 Gyr, using STARLIGHT with the V15 model, now correcting for stellar evolution. The full table for the 404931 galaxies of the clean sample is available online.

**APPENDIX A: TABLE OF VERY YOUNG GALAXIES**

This paper has been typeset from a TeX/LaTeX file prepared by the author.