Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **502,** 4815–4841 (2021)
Advance Access publication 2021 February 20

doi:10.1093/mnras/staa4008

The properties and environment of very young galaxies in the local Universe

M. Trevisan,[1]★ G. A. Mamon,[2] T. X. Thuan,[2,3] F. Ferrari,[4] L. S. Pilyugin[5] and A. Ranjan[2,6]

[1]*Departamento de Astronomia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS 91501-970, Brazil*
[2]*Institut d'Astrophysique de Paris (UMR 7095 – CNRS and Sorbonne Université), 98 Bis Boulevard Arago, F-75014 Paris, France*
[3]*Astronomy Department, University of Virginia, PO Box 400325, Charlottesville, VA 22904-4325, USA*
[4]*Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande, Rio Grande, RS 96201-900, Brazil*
[5]*Main Astronomical Observatory of National Academy of Sciences of Ukraine, 27 Zabolotnogo Str, 03680 Kiev, Ukraine*
[6]*Korea Astronomy and Space Science Institute, 776 Daedeok-daero, Yuseong-gu, Daejeon 34055, Republic of Korea*

Accepted 2020 December 23. Received 2020 December 7; in original form 2020 September 28

ABSTRACT

In the local Universe, there are a handful of dwarf compact star-forming galaxies with extremely low oxygen abundances. It has been proposed that they are young, having formed a large fraction of their stellar mass during their last few hundred Myr. However, little is known about the fraction of young stellar populations in more massive galaxies. In a previous article, we analysed 404 000 Sloan Digital Sky Survey spectra to identify a surprisingly large sample of more massive very young galaxies (VYGs), defined to have formed at least 50 per cent of their stellar mass within the last 1 Gyr. Here, we investigate in detail the properties of a subsample of 207 galaxies that are VYGs according to all three of our spectral models. We compare their properties with those of control sample galaxies (CSGs). We find that VYGs tend to have higher surface brightness and to be more compact, dusty, asymmetric, and clumpy than CSGs. Analysis of a subsample with H I detections reveals that VYGs are more gas rich than CSGs. VYGs tend to reside more in the inner parts of low-mass groups and are twice as likely as CSGs to be interacting with a neighbour galaxy. On the other hand, VYGs and CSGs have similar gas metallicities and large-scale environments (relative to filaments and voids). These results suggest that gas-rich interactions and mergers are the main mechanisms responsible for the recent triggering of star formation in low-redshift VYGs, except for the lowest mass VYGs, where the starbursts would arise from a mixture of mergers and gas infall.

Key words: galaxies: dwarf – galaxies: evolution – galaxies: stellar content.

1 INTRODUCTION

In the past decades, very deep photometric and spectroscopic surveys have enabled astronomers to trace the cosmic star formation history (CSFH) from these early epochs to the present (Madau & Dickinson 2014, and references therein). These observations show that the cosmic star formation rate (SFR) density peaked approximately 3.5 Gyr after the big bang ($z \sim 1.9$) and declined exponentially afterwards, with less than ∼2 per cent of the stars in the Universe being formed within the last 1 Gyr. However, the CSFH is averaged over large volumes and might not represent the star formation histories (SFHs) of individual galaxies. On the contrary, the SFHs of galaxies are observed to vary noticeably with galaxy properties such as mass (Balogh et al. 2009; McGee et al. 2011; Trevisan et al. 2012; Woo et al. 2013), morphology (e.g. Ferrari, de Carvalho & Trevisan 2015), and with the environment where the galaxies reside (Weinmann et al. 2006; Peng et al. 2010; von der Linden et al. 2010). Among the large variety of assembly histories, the systems that formed more than half of their stellar mass in the last Gyr, which we call hereafter *very young galaxies* (VYGs), are of particular interest because they offer us a very close view of galaxy formation, allowing us to identify the physical mechanisms that govern the recent growth of stellar mass via gas accretion or gas-rich mergers.

In the local Universe, there are a few low-mass star-forming galaxies that have extremely low oxygen abundances, an indication that they could have formed most of their stars only recently (<1 Gyr). A well-studied example is the galaxy I Zw 18, with $12 + \log(\mathrm{O/H}) \sim 7.2$ (Skillman & Kennicutt 1993; Izotov & Thuan 1998), i.e. ∼3 per cent of solar metallicity, adopting $12 + \log(\mathrm{O/H}) = 8.7$ for the Sun (Asplund et al. 2009). Using the *Hubble Space Telescope* (*HST*) to resolve the stellar content of I Zw 18 and construct colour–magnitude diagrams, Izotov & Thuan (2004) estimated that most of its stellar mass was formed within the last 500 Myr. However, deeper *HST* images revealed the presence of an older stellar component (ages $\gtrsim$1–2 Gyr; Aloisi et al. 2007; Contreras Ramos et al. 2011). Nevertheless, the fraction of the total stellar mass corresponding to this old population remains undetermined, and if it is less than 50 per cent, galaxies such as I Zw 18 could still be classified as VYGs. Other examples of star-forming compact dwarf galaxies that are extremely metal poor and that have had a considerable fraction of their stellar mass formed during their most recent burst of star formation (SF), a few million years ago, are J0811+4730, the most metal-deficient star-forming galaxy known (Izotov et al. 2018), and J1234+3901 (Izotov, Thuan & Guseva 2019).

★ E-mail: marina.trevisan@ufrgs.br

Recent works have suggested that not only dwarf galaxies like I Zw 18, with stellar masses of $M_\star \sim 10^7$–$10^8\,\mathrm{M}_\odot$, but also more massive systems can be very young. Using low-resolution ($R \sim 30$) optical spectral energy distributions (SEDs) combined with *ugrizJK* broad-band photometry, Dressler, Kelson & Abramson (2018) derived the SFHs of galaxies at $0.45 < z < 0.75$ and identified a population of objects with $M_\star > 10^{10}\,\mathrm{M}_\odot$ that formed at least 50 per cent of their stars within the last ∼2 Gyr of the time of observation. They predicted that these *late bloomers* account for about 30 per cent of the Milky-Way-sized galaxies ($M_\star \sim 10^{10.5}\,\mathrm{M}_\odot$) at $z = 0.7$, but their frequency declines after $z = 0.3$, and they effectively disappear at the present epoch.

Motivated by the debate on the possible VYG nature of very low metallicity galaxies such as I Zw 18, we (Tweed et al. 2018, hereafter Paper I) independently predicted the frequency of VYGs in the local Universe using analytical and semi-analytic models of galaxy formation. We found that the predicted fraction of VYGs depends on galaxy stellar mass, as well as on the model of galaxy formation adopted. The fraction of galaxies with $10^8 < M_\star < 10^{10}\,\mathrm{M}_\odot$ that are VYGs ranges between ∼0.4 per cent and 4 per cent depending on the model, but the fraction of massive galaxies ($M_\star \gtrsim 10^{11}\,\mathrm{M}_\odot$) that are VYGs is less than ∼0.01 per cent for all models.

In a second article (Mamon et al. 2020, hereafter Paper II), we used the galaxy SFRs inferred from the Sloan Digital Sky Survey Data Release 12 (SDSS-DR12) spectra to identify the VYGs and computed the fraction of these systems in the local Universe. We found that the observed VYG fractions decrease more gradually with stellar mass when compared to the predictions from Paper I. We also found that the fractions of VYGs strongly depend on the spectral model used in the spectral fitting procedure, with differences of up to 1 dex between different models, and that the number of VYGs in the SDSS can be as high as a few tens of thousands, depending on the model. Finally, Paper II discusses the possibility that old stellar populations can dominate the mass despite (1) the spectral fitting and (2) the conservative choice of selecting VYGs that have blue colour gradients.

In this work, we explore in detail the properties of a subsample of the large sample of VYG candidates identified in Paper II, as well as the environment where they reside. To minimize the dependence of SFHs on the spectral model used and select a reliable sample of VYGs, we require that each galaxy in the subsample has more than half of its stellar mass younger than 1 Gyr, according to each of three different spectral models.

Our article is organized as follows: In Section 2, we describe how we select the VYGs and the control sample galaxies (CSGs). In the following sections, we describe the properties (Section 3), gas content (Section 4), and environment (Section 5) of the VYGs, and compare them with those of the CSGs. In Section 6, we summarize how the VYGs differ from other galaxies and discuss our results. Finally, our conclusions are given in Section 7. Throughout the paper, we adopt the 7-yr Wilkinson Microwave Anisotropy Probe cosmological parameters of a flat Lambda cold dark matter Universe with $\Omega_\mathrm{m} = 0.275$, $\Omega_\Lambda = 0.725$, and $H_0 = 70.2\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Komatsu et al. 2011).

2 DATA AND SAMPLE SELECTION

2.1 Ages from galaxy spectra

Galaxy SFHs and ages were derived from the SDSS-DR12 spectra using stellar population synthesis analysis. Because they have a fairly high spectral resolution ($R \approx 2000$), a good signal to noise (S/N $\gtrsim$ 10), cover a wide spectral range (from below 3800 to 9200 Å), and are flux calibrated, the SDSS spectra allow the derivation of the SFH of each galaxy, by matching the observed SED with a linear combination of single stellar population (SSP) spectra with non-negative coefficients.

As described in Paper II, we considered two non-parametric algorithms to estimate the SFH: the STARLIGHT (Cid Fernandes et al. 2005) algorithm and the VESPA data base (Tojeiro et al. 2009). The two algorithms have been run using the Bruzual & Charlot (2003, hereafter BC03) model, calculated with Padova 1994 stellar evolution tracks (Bressan et al. 1993; Fagotto et al. 1994a,b; Girardi et al. 1996) and assuming 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, hereafter M05) based on the Kroupa (2001) IMF. We have also run STARLIGHT with the Medium resolution Isaac Newton Telescope Library of Empirical Spectra (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). The V15 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 and Cardelli, Clayton & Mathis (1989) reddening curve, 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 2000).

We ran STARLIGHT considering 15 bins of ages, ranging from 30 Myr up to 13.5 Gyr, and 6 metallicity bins between [M/H] = −1.3 and +0.4. The SFHs from VESPA were obtained considering 16 bins of ages (from 20 Myr to 14 Gyr) and 4 (for BC03 models) or 5 (for M05 models) bins of metallicity. For each galaxy, and for each one of the models, we compute the *median* age when the stellar mass was half of its final value by interpolating the cumulative fractional mass as a function of the logarithm of age.

In summary, we have six different SFH estimates for each galaxy: those determined with VESPA using BC03 and M05 models, and assuming two different dust modelling in each case, and those obtained with STARLIGHT using BC03 and V15 models. We have investigated how these six SFHs correlate with other indicators of recent SF activity, such as the H α equivalent width (EW). We found that, among all models, three (V15 with STARLIGHT, and the two-component dust BC03 and M05 models with VESPA) lead to a tighter correlation between f_young and H α EW (see Paper II for details). Therefore, we have selected these three models as the benchmark for the rest of our analysis. Moreover, because individual SFHs are not fully reliable, we classify a galaxy as a VYG only if all these three models agree on the youth of the galaxy (see Section 2.3).

2.2 Handling aperture effects

One of the main issues when dealing with SFHs derived from SDSS spectra is the finite aperture of the SDSS fibres (3-arcsec diameter). These spectra sample only the inner regions of most galaxies. To account for the stars lying outside of the fibre aperture, the stellar masses obtained through the stellar population synthesis (SPS) analysis of the SDSS spectra, as described in Section 2.1, are corrected by a factor $\mathrm{dex}[0.4(m_{\mathrm{fibre},z} - m_{\mathrm{model},z}]$, where $m_{\mathrm{fibre},z}$ and $m_{\mathrm{model},z}$ are the fibre and model magnitudes in the z band, respectively. This correction assumes that the SFH is homogeneous throughout the galaxy. We use the z-band magnitude because it traces the stellar mass of the galaxy better than bluer filters.

Although the value of the total stellar mass is not strongly affected by assuming the same SFH throughout the galaxy, the finite aperture

of the SDSS fibres has a more important consequence on the selection of the VYGs. Many galaxies with blue nuclei have redder envelopes. In fact, 17 per cent of all galaxies in our sample have redder global colours than their fibre colour. Therefore, using spectra within fibres that probe only the nuclei would make us classify mistakenly such galaxies as VYGs, even though the bulk of their stellar mass (in the outer regions) is old.

As discussed in Paper II, this aperture effect is potentially serious, since the median fraction of galaxy light subtended by the fibre, f_L, is only 26 per cent, and only 3 per cent of our sample has fibres collecting more than half of the galaxy light. Therefore, instead of limiting our sample to galaxies with high f_L values, we follow the approach of Paper II and require a VYG galaxy to be one where the g–i colour of the global galaxy (using model magnitudes) is bluer than the corresponding fibre colour.

2.3 Selection of VYG and control samples

2.3.1 *Selection of VYGs*

For our comparison of the properties of VYGs to that of a control sample, we started with the sample of 404 931 SDSS galaxies with $0.005 < z < 0.12$ that had good spectra and colours (the `clean` sample described in section 2 of Paper II).[1] We then assembled a sample of VYGs that satisfy the following six additional conditions:

(i) their median ages are younger than 1 Gyr according to all 3 spectral models (V15 with STARLIGHT, and the two-component dust BC03 and M05 models with VESPA; 1214 galaxies);
(ii) their stellar masses derived with STARLIGHT using V15 models are $\geq 10^8$ M$_\odot$ (838 galaxies);
(iii) the signal to noise of their SDSS spectra satisfies S/N ≥ 10, where S/N is computed within a window of 50 Å centred at 4755 Å (rest frame; 654 galaxies);
(iv) their specific SF rates (sSFRs) are available in the MPA/JHU `SpecLineExtra` table of the SDSS data base (634 galaxies);
(v) they do not lie in the active galactic nucleus (AGN) region of the BPT diagram (Baldwin, Phillips & Terlevich 1981) – we adopted the relation by Kewley et al. (2001) to separate the star-forming and AGN galaxies (633 galaxies);
(vi) they show blue colour gradients: $\Delta(g\text{–}i) = (g\text{–}i)_{\rm model} - (g\text{–}i)_{\rm fibre} \leq 0$ (207 galaxies).

The number of galaxies that remain in the sample after each cut is indicated in parentheses. The list of criteria above is very conservative, and yields a relatively small sample of 207 VYGs. As shown in Paper II, the number of VYGs in the SDSS can be as high as a few tens of thousands, but to avoid possible contaminants in the sample, we opted for this conservative approach. The images of these 207 VYGs are shown in Figs C1–C4 in the supplementary material.

2.3.2 *Control sample*

The CSGs were also selected from the *clean* sample after removing the AGNs using the same criteria adopted for the VYG sample (item v above). The CSGs were chosen to be older than 1 Gyr according to all the three models. We matched the CSG galaxies to have the same distributions of redshift, stellar mass (according to STARLIGHT ran with V15), angular size, and the total-fibre colour difference, $\Delta(g\text{–}i)$ (see item vi above). Matching the radii θ_{50} containing 50 per cent of the Petrosian flux in the r band ensures that the aperture effects are similar for both samples. The redshifts were included in the matching procedure to ensure that VYGs and CSGs potentially equally populate the same features of the large-scale distribution of galaxies.

We applied the Propensity Score Matching (PSM) technique (Rosenbaum & Rubin 1983; de Souza et al. 2016) to build the control sample. We used the MATCHIT package (Ho et al. 2011), written in R (R Core Team 2015).[2] This technique allows us to select from the sample of normal galaxies (all minus VYGs) a control sample in which the distribution of observed properties is as close as possible to that of the VYGs. We adopted the Mahalanobis distance approach (Mahalanobis 1936; Bishop 2006) and the nearest neighbour method to perform the matching.

This procedure yielded a sample of 1242 CSGs, i.e. 6 times the size of the VYG sample. We have also tried control samples with different sizes, spanning the range from 1 to 10 times the size of the VYG sample. We found that the results on the comparison between the samples are not affected. The control sample with 6 times the size of the VYG sample was chosen to provide good statistics while keeping the distribution of stellar masses within the same range for both samples.

The distributions of redshifts, stellar masses, $\Delta(g\text{–}i)$, and θ_{50} of the VYGs and CSGs are shown in Fig. 1. Also shown are the p-values derived from the Kolmogorov–Smirnov (KS) test. The VYG and CSG distributions of redshifts, $\Delta(g\text{–}i)$ and θ_{50} (Figs 1a, c, and d) are very similar. On the other hand, we can see in Fig. 1(b) that the VYG sample contains more low-mass galaxies than the CSG sample. It is particularly difficult to draw a control sample of SDSS galaxies with masses below $\sim 10^9$ M$_\odot$, since only $\sim$3.5 per cent of them are in this mass range. Besides, 95 per cent of them are at $z \lesssim 0.02$, the redshift completeness cut-off limit of the SDSS sample defined in Paper II for VYGs within this mass range. This makes it very difficult to select CSGs with the same stellar mass versus redshift distribution as that of the VYGs. This is illustrated in Fig. 2(a), where we see that, at a given mass, the VYGs tend to lie at higher redshifts compared to the CSGs. We have tried to overcome this issue by changing the PSM parameters and the size of the control sample, but we always end up with a similar $M_\star$ versus z distribution. In any event, the results and conclusions presented in this work are not significantly affected by the difference between the VYG and CSG $M_\star$ versus z relations, as we discuss in Section 6.7.3. The other panels in Fig. 2 will be discussed in later sections.

3 PROPERTIES OF VYGS

In this section, we investigate some basic properties of the VYGs by comparing their sSFRs, colours, positions in the BPT diagram (Baldwin et al. 1981), and morphologies with those of the CSGs.

3.1 SFRs and colours

We extracted from the SDSS-DR12 data base the optical magnitudes as well as the sSFRs (`specsfr_tot_p50` from the `galSpecExtra` table, obtained with the algorithm of Brinchmann et al. 2004). To compute the g–i colours, we used the extinction- and k-corrected Petrosian magnitudes. The k-correction was obtained with the

[1] The table with the properties of the galaxies of the `clean` sample is available at http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/MNRAS/492/1791.

[2] https://cran.r-project.org/

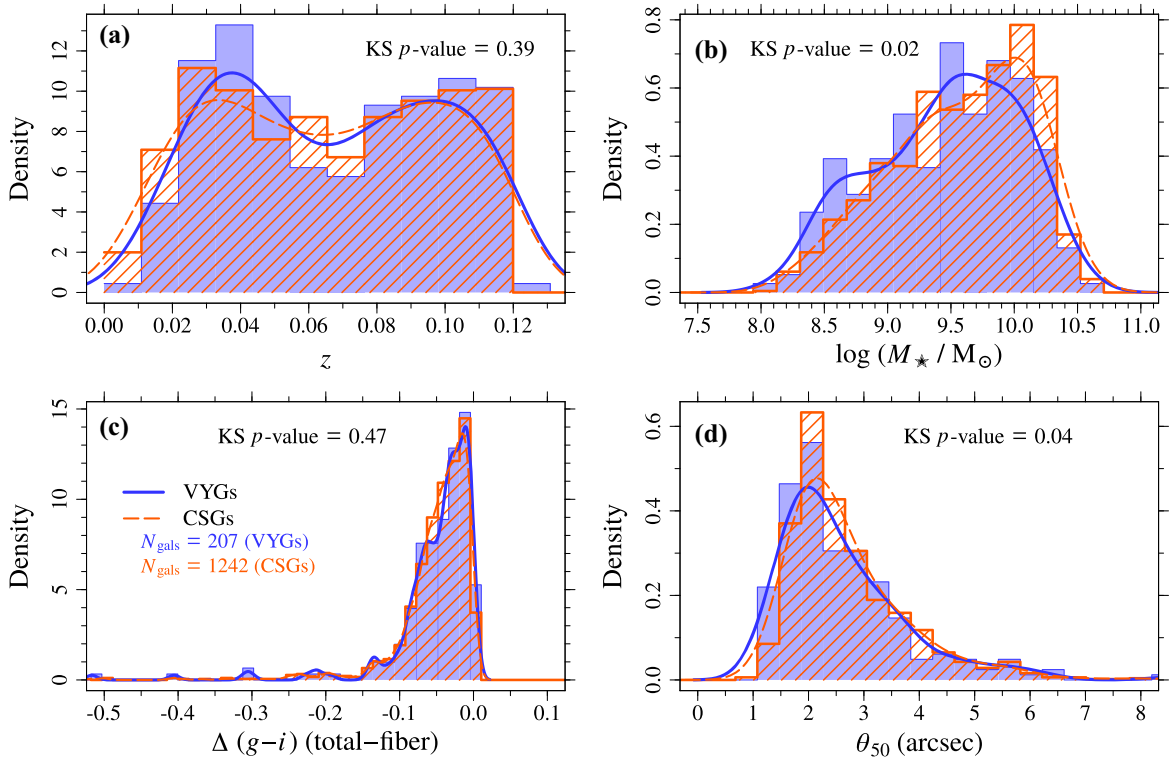

Figure 1. Redshifts (a), stellar masses (b), difference between the total and the fibre g–i colours (c), and angular radius containing 50 per cent of the Petrosian flux, θ_{50} (d). The blue histograms show the distributions of z, $\log(M_\star/\mathrm{M}_\odot)$ (from STARLIGHT V15), $\Delta(g$–$i)$, and θ_{50} of the VYG sample. The control sample (CSG), drawn from the general sample of galaxies by applying the PSM technique (see the text for details), is shown as the orange histogram. The curves are obtained by smoothing the positions of the data points (not the histograms) using a Gaussian kernel with the standard deviation equal to one-third of the standard deviation of the data points. In each panel, we indicate the p-value of the KS test.

`kcorrect` code (version 4_2) of Blanton et al. (2003), choosing as reference the median redshift of the SDSS sample ($z = 0.1$).

In Fig. 3, we compare the sSFRs and colours of the VYGs with those of normal galaxies. Although we allow passive galaxies to be included in the control sample, the CSGs are naturally more likely to be star forming for two reasons. First, ∼90 per cent of the galaxies with stellar masses in the range of those of the VYGs are star-forming systems (the line separating the star-forming and passive galaxies determined by Knobel et al. 2015 is shown in Fig. 3c). Secondly, requiring CSGs to be at similar redshifts and to have similar stellar masses as those of VYGs favours the selection of star-forming systems, since they are brighter and more likely to be seen at distances where VYGs are observed. As described in Section 2, the VYG sample contains slightly more low-mass galaxies than the control sample, even after the PSM, since the SDSS-MGS is highly incomplete for stellar masses below $\sim 10^9$. To minimize the effects of this difference on our comparison results, we have fitted the relations between the galaxy properties and the stellar mass (using only the galaxies in the control sample) and analysed the residuals from these fits.

The curvature that we observe in the best fits to the g–i and log sSFR versus stellar mass relations (shown as solid lines in Fig. 3) is, again, a consequence of requiring the CSGs to have a similar redshift distribution as that of the VYGs. If we build a control sample without including the redshift in the PSM procedure, i.e. using only $\log M_\star$, θ_{50}, and $\Delta(g$–$i)$, the g–i colours (sSFRs) of the selected galaxies systematically increase (decrease) with stellar mass, and the best-fitting relations (indicated by the dashed curves in Fig. 3) do not show the curvature.

Not surprisingly, the VYGs, which have made more than half of their stellar mass in the last Gyr, show higher sSFRs than the CSGs. The effect is especially evident for the low-mass galaxies [$\log(M_\star/\mathrm{M}_\odot) < 9.0$]. The statistical significance of the difference between the two samples is $p < 10^{-10}$ (Figs 3a and b). As expected, the VYGs are bluer than normal galaxies, and those with high Hα EWs have more extreme sSFRs. Again, the differences are more significant for low-mass galaxies.

The BPT diagram (Baldwin et al. 1981) of the VYG and control samples is shown in Fig. 4, where the emission-line fluxes are taken from the MPA-JHU catalogue (table `galSpecLine` based on the SDSS-DR12; Brinchmann et al. 2004; Tremonti et al. 2004). We find a higher fraction of VYGs in the composite region of the diagram compared to the normal galaxies (4.5 times higher). By construction of the VYG and CSG samples, there are no galaxies in the AGN region of the diagram. We used the CSGs to derive a fit for the $\log([\mathrm{O\,III}]/\mathrm{H}\beta)$ versus $\log([\mathrm{N\,II}]/\mathrm{H}\alpha)$ relation, and computed the residuals (Fig. 4b). The distribution of the residuals for the VYGs is shifted towards higher ionization levels compared to the CSGs, with a KS test indicating that the difference is statistically significant (p-value $< 10^{-10}$).

3.2 Galaxy morphologies and structural parameters

To characterize the VYG and CSG morphologies, we have used the morphological classification from the Galaxy Zoo 1 project (GZ1; Lintott et al. 2011) and that by Domínguez Sánchez et al. (2018, hereafter DS18). The GZ1 catalogue is based on simple visually

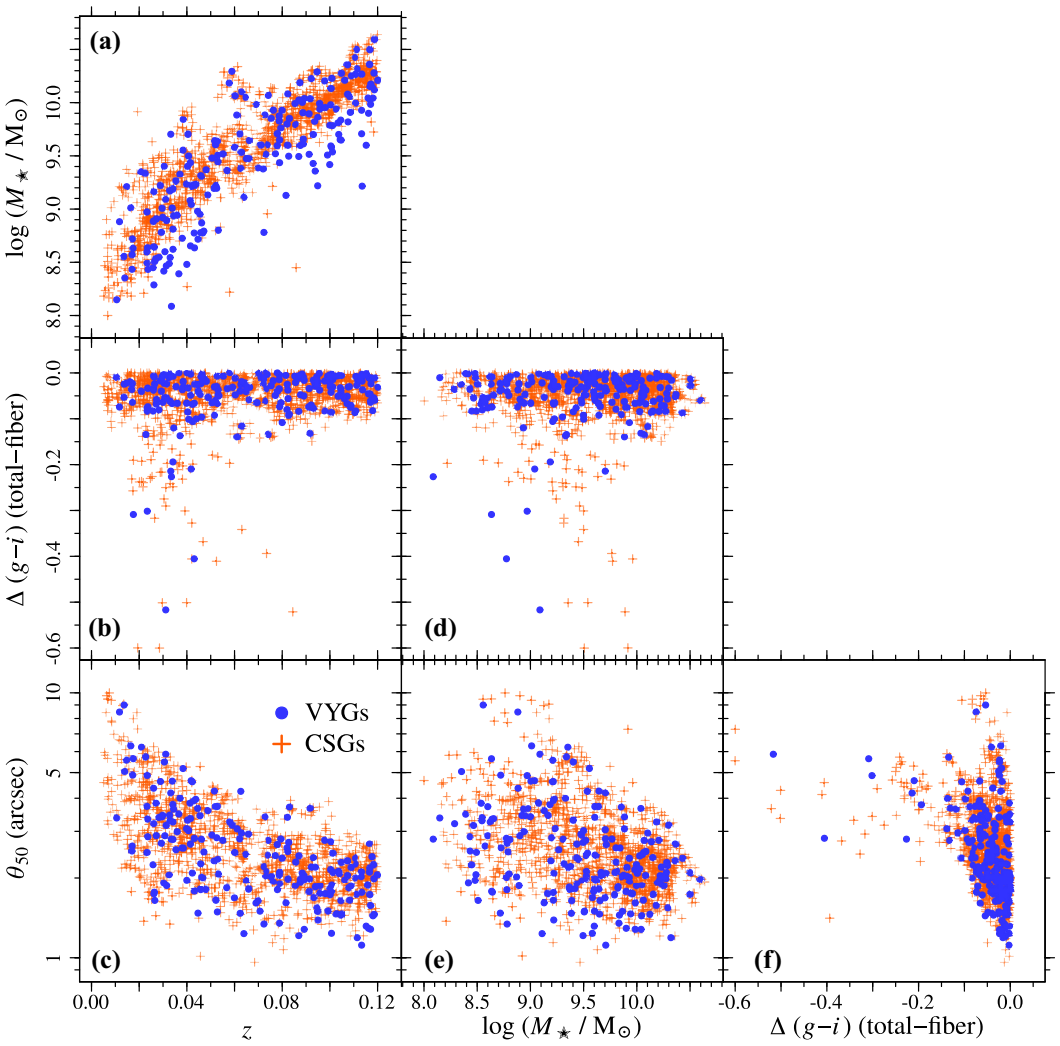

Figure 2. Relations between redshift, stellar mass, difference between the total and the fibre g–i colours, and angular radius containing 50 per cent of the Petrosian flux for the VYGs (*blue symbols*) and for the CSGs (*orange*).

inspected classifications by hundreds of thousands of volunteers, and for each galaxy, it provides the debiased fraction of votes for each morphological type and classifies the galaxies in three categories: elliptical, spiral, and uncertain. The DS18 catalogue was obtained with deep learning algorithms, using convolutional neural networks and visual classification catalogues from Galaxy Zoo 2 project (GZ2; Willett et al. 2013) and by Nair & Abraham (2010) as training sets. Besides providing classifications following a scheme similar to that of the GZ2, it also computes the numerical T type of each galaxy (de Vaucouleurs 1963).

We cross-matched our samples with the GZ1 and DS18 catalogues, and found morphological data for nearly all our galaxies: 1384 of our galaxies are in GZ1 (205 VYGs, 1179 CSGs) and 1441 in DS18 (203 VYGs, 1238 CSGs). Most of the VYGs and CSGs are classified as *uncertain* according to GZ1 (76 per cent and 74 per cent, respectively); all the other VYGs and CSGs are classified as *spirals*, except for one CSG that is an *elliptical*. We also compared the distributions of probabilities of being disc- (`P_CS_DEBIASED` and $p_{\rm disc}$ in the GZ1 and DS18 catalogues, respectively) or bulge-dominated galaxies (`P_EL_DEBIASED`, $p_{\rm bulge}$), and found no significant difference between the VYGs and CSGs. Finally, the VYG and CSG distributions of T types from the DS18 catalogue are also very similar. Fig. 5 shows that most of the galaxies (97.6 per cent) have $T \geq 0$ (94.1 per cent VYGS, 98.1 per cent CSGs), and a KS test confirms the low statistical differences between the two samples (KS p-value = 0.54). We find an excess of VYGs with low T types compared to the CSGs. The number of VYGs and CSGs that lie below 3σ of the T type versus stellar mass relation is 15 (7.4 per cent) and 27 (2.2 per cent), respectively.

The high fraction of *uncertain* classifications in GZ1 indicates that most of our galaxies cannot be simply classified as ellipticals or spirals. Even when discarding galaxies with small angular sizes, the fractions of VYGs and CSGs that are classified as *uncertain* are very high (61 per cent and 50 per cent of VYGs and CSGs with $\theta_{50} > 3$ arcsec, respectively). The morphological classification by DS18 also points to this conclusion, since only a small fraction of VYGs and CSGs could be reliably classified as one of these two morphological types (15 per cent of spirals with $p_{\rm disc} \geq 0.8$ and 0.5 per cent of ellipticals with $p_{\rm bulge} \geq 0.8$). This is confirmed by examination of Figs C1–C4 in the supplementary material, which display the VYG images. On the other hand, the distribution of T types (Fig. 5) is not compatible with irregular morphology, which

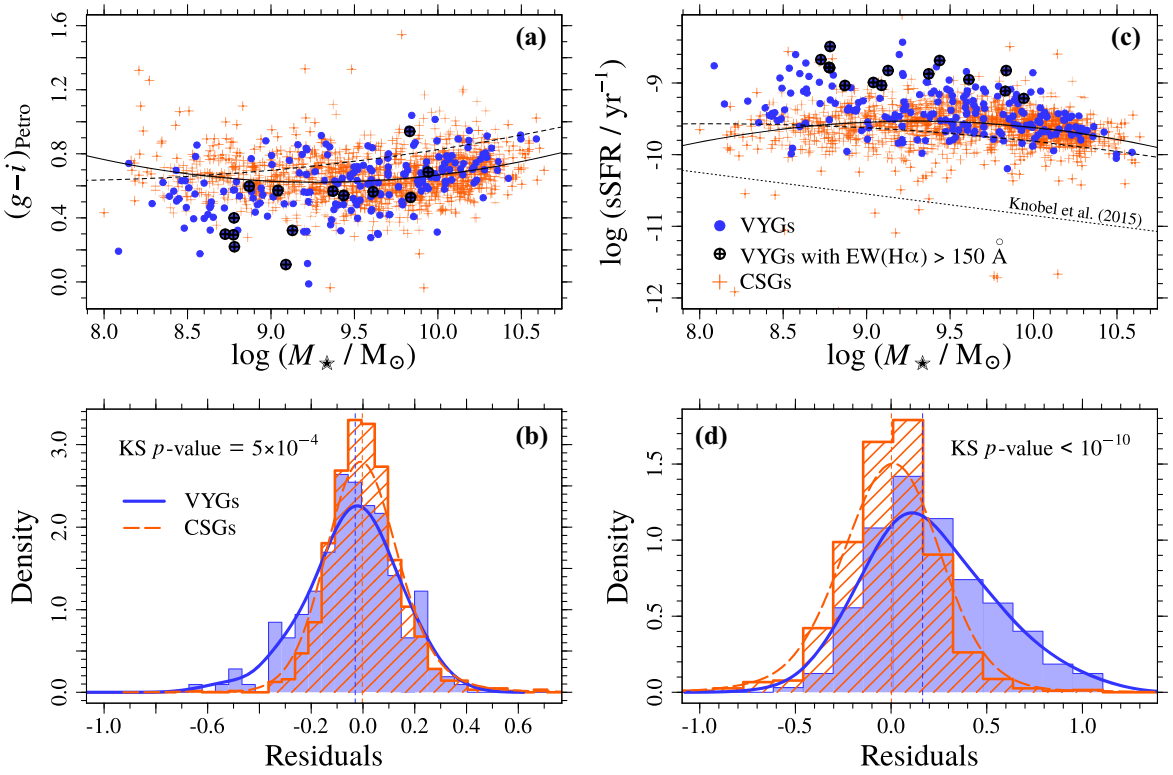

Figure 3. *Top:* Petrosian *g–i* colours (a) and sSFRs (b, from the MPA/JHU table in SDSS) of VYGs (*blue circles*) and normal galaxies (*orange symbols*). The VYGs with EW(Hα) $\geq$ 150 Å are indicated by the *black crossed circles*. The *solid lines* in both panels correspond to a second-order polynomial fit to the colour and log sSFR versus log stellar mass relations for galaxies in the control sample. The *dashed lines* show the best-fitting relations for a sample of galaxies selected using the same approach adopted to build the control sample, but without including the redshift in the PSM procedure (see the text for details). The difference between these two curves illustrates the effect of requiring the CSGs to be at similar distances as the VYGs. The line separating the star-forming and passive galaxies determined by Knobel et al. (2015) is indicated by the *dotted line*. *Bottom:* Residuals from the second-order polynomial fits shown for the VYG (*blue histograms*) and control (*orange*) samples. In each panel, we indicate the KS test *p*-values.

is characterized by high *T*-type values (9–10). However, the range of *T*-types values in the DS18 catalogue goes from ~ -3 to ~ 6, and only 0.4 per cent of the galaxies have $T > 6$ (2861 out of 670 722 galaxies).

To understand what morphological types correspond these *T*-type values, we used the parameter `gz2_class` from GZ2 to select elliptical, spiral, and irregular galaxies and analysed the *T*-type distributions of these three morphological classes. The `gz2_class` parameter is a string that indicates the most common consensus classification for the galaxy, and is composed by the letter E (galaxies that are smooth) or S (galaxies with discs and/or features), followed by other letters indicating several galaxy features (see the appendix in Willett et al. 2013 for details). We assumed to be ellipticals all galaxies with `gz2_class` = Er or Ei, where the letters 'r' and 'i' indicate that the galaxy is round or in between round and cigar shaped. We used the classification for bulge prominence as an indication for spiral morphology, and to avoid selecting lenticulars, we selected galaxies with no bulge or systems for which the bulge is just noticeable (`gz2_class` =SBc+, SBd+, Sc+, or Sd+, where '+' indicates other features). The same selection was made for the irregulars, but `gz2_class` contains the identifier for irregular morphology '(i)', i.e. `gz2class` = Sc+(i), SBc+(i), Sd+(i), and SBd+(i).

Fig. 6 shows that the VYG and CSG *T*-type distributions (Fig. 6a) are similar to that of irregulars, as seen in Fig. 6(b). The irregulars have median $T = 4.7$, and the 10th and 90th percentiles are 2.4 and 5.8, respectively. The spirals have lower *T* types, with median 3.0, and 10th and 90th percentiles of -0.7 and 4.8, respectively. The median of the VYG and CSG *T* type values is 4.4 and 4.3, respectively, with 10th and 90th percentiles of 1.2 and 5.6 (VYGs) and 2.2 and 5.5 (CSGs), respectively. As already indicated in the results shown in Fig. 5, we find an excess of VYGs with small *T* types, and the number of VYGs with negative values is 3.2 times higher than the number of CSGs with $T < 0$ (5.9 per cent of the VYGs and 1.9 per cent of the CSGs). Fisher's (Fisher 1935) and Barnard's (Barnard 1945) tests, performed using the R packages `stats` (R Core Team 2015) and `Barnard` (Erguler 2016), indicate that this difference between the VYG and CSG early-type fractions is statistically significant, with *p*-values = 0.002 and 0.03, respectively. The inspection of the images of these galaxies reveals that both VYGs and CSGs with negative *T* types have spheroid-like morphology, but VYG spheroids are bluer than the CSG spheroids, with median *g–i* colours of 0.7 (VYGs) and 0.9 (CSGs).

Parametric models of bulges and/or discs may not give a good representation of these systems. So, to obtain more information and characterize the galaxy morphology, we have adopted a non-parametric approach. We used the popular CAS system (concentration, asymmetry, and clumpiness) presented in Abraham et al. (1994, 1996) and Conselice, Bershady & Jangren (2000).

We measured the concentration as the ratio of radii containing 90 and 50 per cent of the Petrosian flux in the *r* band, θ_{90}/θ_{50}, which

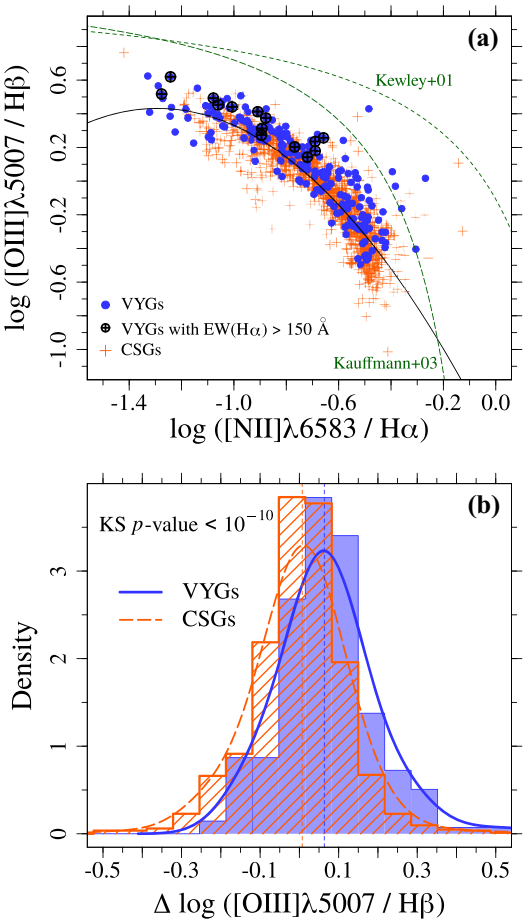

Figure 4. *Top:* BPT diagram showing the VYG (*blue circles*) and control (*orange symbols*) samples. The VYGs with EW(Hα) $\geq$ 150 Å are indicated by the *black crossed circles*. The relations by Kewley et al. (2001) and Kauffmann et al. (2003) separating star-forming galaxies and AGNs are indicated as short- and long-dashed lines, respectively. The relation between the flux ratios of the CSGs was fitted with a second-order polynomial, as indicated as the *black solid line*. *Bottom:* The *blue* and *orange* histograms show the distribution of residuals from the polynomial fit for the VYGs and CSGs, respectively.

we retrieved from the SDSS-DR12 data base as `petroR90_r` and `petroR50_r`, respectively.[3]

We also compute the galaxy surface brightness as

$$\mu_{50} = m_{\rm r} - k_{0.1} + 2.5\log\left(2\pi\theta_{50}^2\right) - 10\log(1+z), \quad (1)$$

where $m_{\rm r}$ is the extinction-corrected Petrosian magnitude in the r band, $k_{0.1}$ is the k-correction, and θ_{50} is the radius (in arcsec) containing 50 per cent of the Petrosian flux in the r band. The k-correction was obtained with the code `kcorrect` with the SDSS filters shifted to $z = 0.1$.

The galaxy asymmetry and clumpiness were measured using MORFOMETRYKA[4] (Ferrari et al. 2015), which is a code to perform structural and morphometric measurements on galaxy images. The asymmetry coefficient is determined by comparing the galaxy image with a rotated version of itself. Three different asymmetry estimates

[3]SDSS does not provide radii at less than half the flux, which are more sensitive to variations in the point spread function.
[4]http://morfometryka.ferrari.pro.br

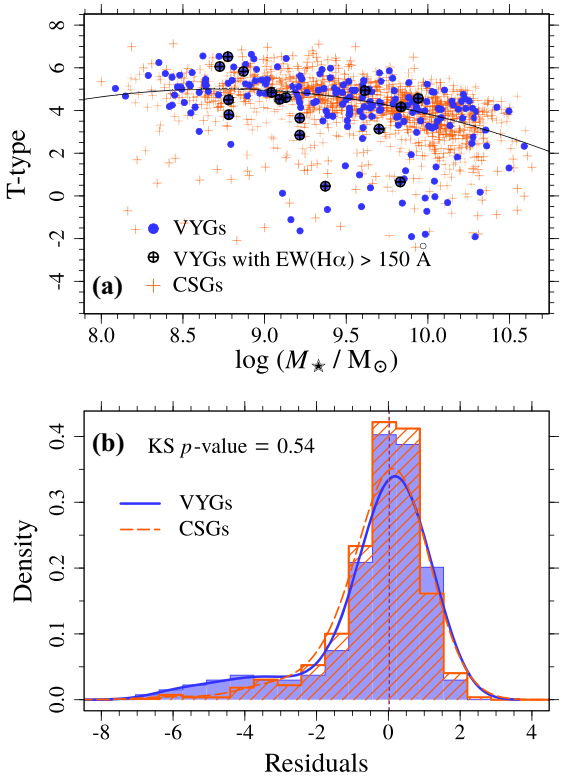

Figure 5. *Top:* Morphological T types from DS18 as a function of stellar mass. The best second-order polynomial fit to the CSG T type versus log ($M_\star$/M$_\odot$) relation is indicated by the *black solid line*. *Bottom:* Distribution of VYG and CSG residuals from the second-order polynomial fit. The notations in both panels are the same as in Fig. 3.

are computed by the code. Here, we adopt the standard one defined by Abraham et al. (1996), A_1, with the difference that we do not subtract the background asymmetry, since this procedure leads to an unstable estimate of the coefficient A_1 and makes it dependent of the region selected to measure the sky asymmetry. Instead, only the region within the Petrosian radius is used.

Similarly, MORFOMETRYKA also computes two different clumpiness coefficients (confusingly denoted 'smoothness' and denoted by S_1 and S_2, even though higher S_i values correspond to more clumpy distributions) by comparing the galaxy image with a smoothed version of itself. In this work, we adopt S_1 defined by Lotz, Primack & Madau (2004) using a Hamming window (Hamming 1998) with size $\theta_{\rm Petro}/4$, where $\theta_{\rm Petro}$ is the galaxy Petrosian radius.

Although the MORFOMETRYKA code returns many other parameters, including modified versions of asymmetry and clumpiness as described in detail in Ferrari et al. (2015), we show only A_1 and S_1 because they are less dependent on the S/N of the images. In particular, the Gini and M20 parameters, which are an extension of the CAS system (Lotz et al. 2004), are also measured by MORFOMETRYKA. However, we found them to be very dependent on the S/N of the image. Besides, it is not straightforward to physically interpret the Gini and M20 parameters.

All the MORFOMETRYKA fits were visually inspected, and we excluded 27 (13 per cent) VYGs and 132 (11 per cent) CSGs with fits affected by foreground stars or very close objects (distances $\lesssim$5 arcsec).

In Fig. 7, we show the galaxy concentration (panel a) and surface brightness (panel c) as a function of the stellar mass. The

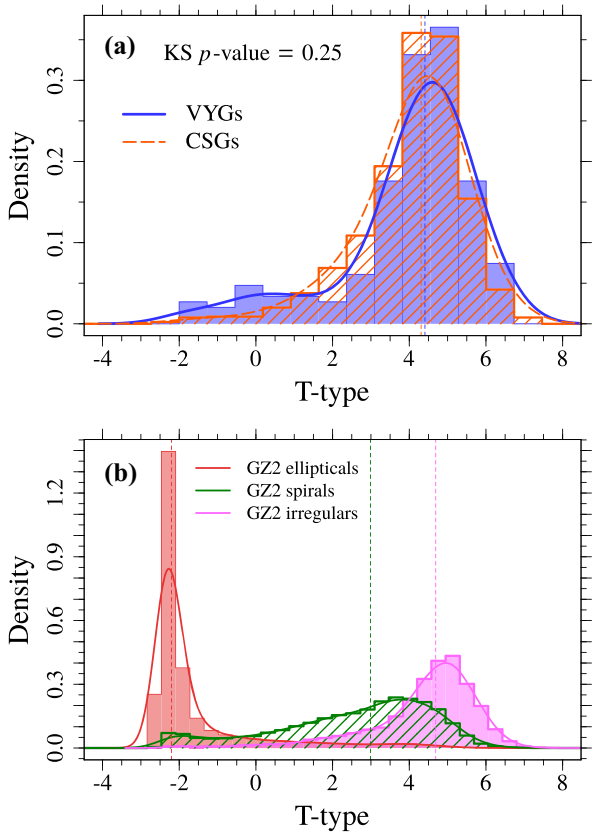

Figure 6. *Top:* Distributions of T types of the VYGs (*blue histogram*) and CSGs (*orange*). *Bottom:* Distribution of T types of galaxies classified as ellipticals (*red*), spirals (*green*), and irregulars (*magenta*) in GZ2.

residuals derived from fitting the data show that the VYGs are more concentrated and have higher surface brightness than normal galaxies (Figs 7b and d). The KS tests indicate that these results have high statistical significance, with $p = 10^{-7}$ (concentration) and $p < 10^{-10}$ (surface brightness).

Since the structural parameters, such as asymmetry and clumpiness, may be sensitive to the S/N of the images, we show in Figs 8(a) and (c) the A_1 and S_1 parameters as a function of the signal-to-noise ratios instead of stellar masses. We confirm the weak dependence of A_1 with S/N, with Kendall and Spearman correlation coefficients $\tau = 0.02$ and $\rho = 0.03$. On the other hand, S_1 shows an anticorrelation with S/N ($\tau = -0.11$ and p-value $= 2 \times 10^{-9}$; $\rho = -0.17$ and p-value $= 8 \times 10^{-11}$). In summary, the comparison between the S_1 distributions of the VYGs and control samples should be considered with caution. The distributions of these parameters are shown in panels b and d. We can see that the VYGs tend to be more asymmetric and clumpy when compared to normal galaxies, with high statistical significance ($p = 2 \times 10^{-7}$ for the asymmetry parameter and $p = 3 \times 10^{-3}$ for the clumpiness). In Fig. 8, we show only galaxies with good MORFOMETRYKA fits; i.e. galaxies with nearby objects (in projection) affecting the estimate of the structural parameters are excluded from the plots (27 VYGs and 132 CSGs, corresponding to 13 per cent and 11 per cent, respectively).

To obtain meaningful structural parameter measurements, the angular size of the galaxy must be greater than the resolution of the SDSS image and the atmospheric seeing. It can be seen in Fig. 1(d) that many of our galaxies have small angular sizes and might be unresolved. However, we note that, since the VYG and CSG θ_{50} distributions are similar by construction, differences between the VYG and CSG structural parameters are unlikely to arise from differences in the angular sizes of galaxies in these two samples.

4 IONIZED GAS, NEUTRAL GAS, AND DUST CONTENT

4.1 Oxygen abundances of the ionized gas

The gas oxygen abundance, an indicator of metallicity, was derived from the emission-line fluxes retrieved from the SDSS data base. Two sets of measurements are available: one given by the MPA-JHU catalogue (table `galSpecLine` in SDSS-DR12; Brinchmann et al. 2004; Tremonti et al. 2004) and the other by the catalogue of the Portsmouth group (table `emissionLinesPort` in SDSS-DR12). The first one employs the BC03 models to fit the stellar continuum, while models from Maraston & Strömbäck (2011) and Thomas, Maraston & Johansson (2011) are used in the Portsmouth measurements. The use of line measurements from two different catalogues allows us to estimate the uncertainties in the abundances of the target galaxies due to uncertainties in the line flux measurements.

We extracted from those catalogues the measurements of the [O II] λ3727, [O II] λ3729, H β, [O III] λ4959, [O III] λ5007, [N II] λ6548, H α, [N II] λ6584, [S II] λ6717, and [S II] λ6731 emission lines in the spectra. We restricted our analysis to galaxies whose emission lines were measured with S/N > 3 for all lines. We corrected the emission-line fluxes for interstellar reddening using the observed H α/H β ratio and the reddening function from Cardelli et al. (1989) for $R_V = 3.1$, which leads to $A_{H\beta} = 1.164\, A_V$.

The 'direct T_e' method (e.g. Dinerstein 1990) is believed to provide the most reliable abundance determinations in SF regions from the emission lines in their spectra. However, it requires high-precision spectroscopy to detect weak auroral lines such as [O III] λ4363 and [N II] λ5755. Unfortunately, these auroral lines are usually not detected or they are measured with a large uncertainty in the SDSS spectra. The abundances in SDSS objects with emission-line spectra are thus usually estimated using 'the strong line method' of Alloin et al. (1979) and Pagel et al. (1979), using intensity ratios of the strong emission lines in the spectra.

Two families of calibrations are widely used. The family of the calibrations following Pagel et al. (1979) is based on the oxygen lines. Both 1D (e.g. Zaritsky, Kennicutt & Huchra 1994; Tremonti et al. 2004) and 2D (or parametric, p-method) calibrations (Pilyugin 2000, 2001; Pilyugin & Thuan 2005) have been suggested. The family of the calibrations following Alloin et al. (1979) is based on the nitrogen lines. There are only 1D calibrations of this kind for abundance determinations in high-metallicity objects (e.g. Pettini & Pagel 2004; Marino et al. 2013). There is no unique relation applicable across the whole range of metallicities of H II regions. Instead, distinct calibration relations are constructed for high (upper branch) and low metallicities (lower branch). However, these branches are not applicable to objects lying in the transition zone (e.g. Pilyugin & Thuan 2005, and discussion there). Moreover, to choose the relevant calibration relation, one has to know a priori to which interval of metallicity the H II region belongs. A wrong choice would lead to a wrong abundance. Many of our galaxies may have moderate to low metallicities, i.e. they belong to the transition zone and to the lower branch, making abundance determinations for our objects particularly difficult.

We used the simple 3D calibration relations proposed by Pilyugin & Grebel (2016). The oxygen abundances $(O/H)_R$ are determined using the R calibration, i.e. $(O/H)_R = f(R_2, R_3, N_2)$, where the oxygen

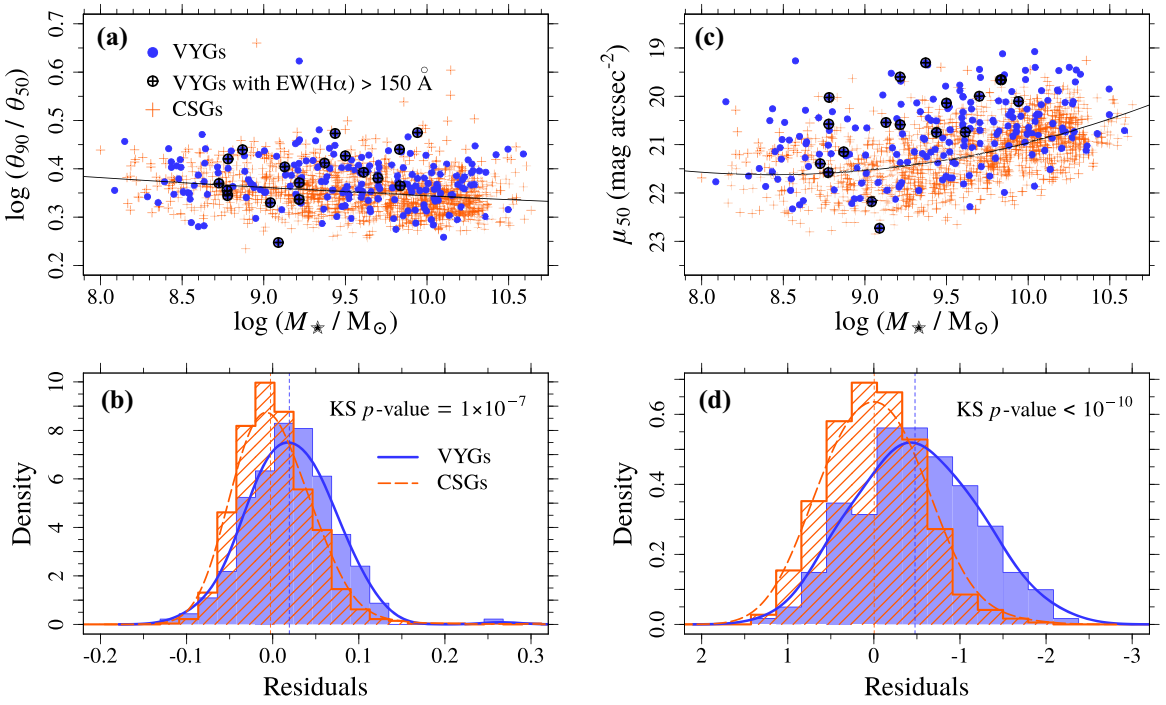

Figure 7. *Top:* Concentration (*left*) and surface brightness (*right*) of VYGs (*blue circles*) and normal galaxies (*orange symbols*). The VYGs with EW(Hα) $\geq$ 150 Å are indicated by the *black crossed circles*. The solid lines in both panels correspond to a second-order polynomial fit to the concentration and μ_{50} versus stellar mass relations for galaxies in the control sample. *Bottom:* Distribution of residuals from the second-order polynomial fit for the VYG (*blue histograms*) and control (*orange*) samples. In each panel, we indicate the *p*-value of KS tests.

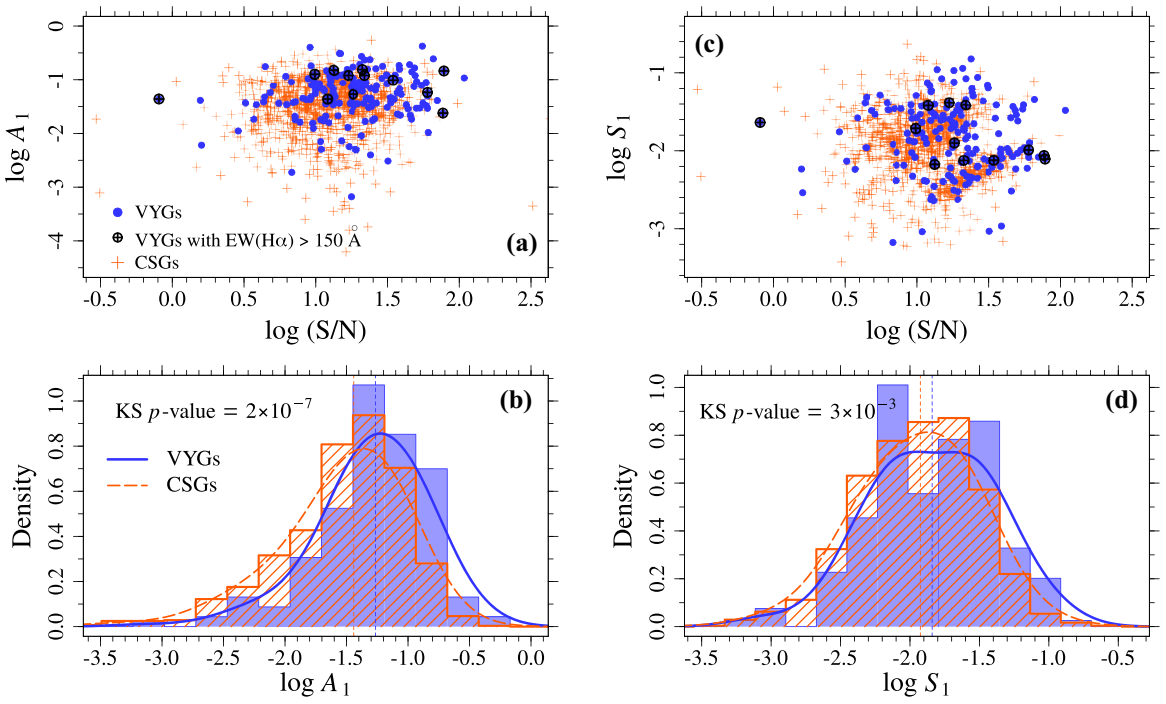

Figure 8. *Top:* Asymmetry (A_1) and clumpiness (S_1) parameters as a function of signal-to-noise ratio. The notation is the same as in Fig. 3. *Bottom:* Distribution of $\log A_1$ and $\log S_1$ for the VYG (*blue histograms*) and control (*orange*) samples. In each panel, we show the KS test *p*-values.

R_2, R_3 and the nitrogen N_2 lines are, respectively, given by

$$R_2 = I_{[\mathrm{O\,II}]\lambda3727+\lambda3729}/I_{\mathrm{H}\,\beta},$$

$$R_3 = I_{[\mathrm{O\,III}]\lambda4959+\lambda5007}/I_{\mathrm{H}\,\beta}, \text{ and}$$

$$N_2 = I_{[\mathrm{N\,II}]\lambda6548+\lambda6584}/I_{\mathrm{H}\,\beta}.$$

One advantage of this calibration is that it is applicable over the entire metallicity range of H II regions. Although distinct relations for high- and low-metallicity objects are constructed, the separation between these two can be simply obtained from the intensity of the N_2 line. Moreover, the applicability ranges of the high- and low-metallicity relations overlap, making the transition zone disappear.

Since the [O III] λ5007 and λ4959 lines originate from transitions from the same energy level, their flux ratio is determined only by the transition probability ratio, close to 3 (Storey & Zeippen 2000). Therefore, the value of R_3 can be estimated as $R_3 = I_{[\mathrm{O\,III}]\lambda4959+\lambda5007}/I_{\mathrm{H}\,\beta}$ or $R_3 = 1.33 I_{[\mathrm{O\,III}]\lambda5007}/I_{\mathrm{H}\,\beta}$. The stronger line [O III] λ5007 is usually measured with higher precision than the weaker line [O III] λ4959. We thus used the latter expression to calculate R_3, instead of the sum of the [O III] line fluxes. The same applies to the nitrogen lines [N II] λ6584 and λ6548. They also originate from transitions from the same energy level and the transition probability ratio for those lines is again close to 3 (Storey & Zeippen 2000). We again calculate N_2 as $N_2 = 1.33$ [N II] λ6584/H β, instead of using the sum of the [N II] lines.

The radiation of the diffuse ionized gas is believed to contribute significantly to some fibre SDSS spectra, and may increase the strength of the low-ionization lines [N II], [O II], and [S II] relative to the Balmer lines. If this is the case, then the abundances derived from the SDSS spectra by strong-line methods may have large errors (Belfiore et al. 2015, 2017; Sanders et al. 2017; Zhang et al. 2017). Pilyugin et al. (2018). found that the mean increase of R_2 and N_2 is less than a factor of ∼1.3, and the R calibration produces reliable abundances.

For each galaxy in our sample with an available oxygen line R_2 and an S/N greater than 3 for the emission lines of interest, we have determined two values of the oxygen abundance, one based on the MPA-JHU line measurements and the other on the Portsmouth ones. We have found that the two sets of abundances are in excellent agreement ($\Delta = 0.005 \pm 0.02$ dex). Hence, we only present here the results based on the MPA-JHU catalogue.

In Fig. 9(a), we show how the VYG and CSG oxygen abundances vary with galaxy stellar mass. Only galaxies with good MPA-JHU emission-line measurements (i.e. S/N ≥ 3) are shown (183 VYGs and 1031 CSGs, corresponding to 88 per cent and 83 per cent of the galaxies in these samples, respectively). We fitted the 12 + log (O/H) versus log ($M_\star$/M$_\odot$) relation of the CSG, and show the distribution of residuals for both samples in Fig. 9(b). The residuals from the fit indicate that there is no significant difference between the oxygen abundances of the VYGs and the CSGs as a function of stellar mass (KS test p-value = 0.09). We also compared the best fit to the VYG 12 + log (O/H) versus log ($M_\star$/M$_\odot$) relation to that of the CSGs, and find that the difference between the two fits does not exceed 0.015 dex within the range $8.0 \leq \log(M_\star/\mathrm{M}_\odot) \leq 10.6$. Using a first-order polynomial to fit the VYG and CSG relations, we also get very similar slopes for both samples (0.23 and 0.22 for the VYGs and CSGs, respectively).

On the other hand, when we limited our analysis to galaxies with $M_\star \leq 10^9$ M$_\odot$, we found that low-mass VYGs have lower oxygen abundances by 0.07 dex, on average, compared to those of the CSGs, with a KS test indicating marginal statistical significance (p-value = 0.02). However, a large fraction of the low-mass CSGs are excluded from this analysis when we require emission-line measurements with S/N ≥ 3. Among the 210 CSGs with $M_\star \leq 10^9$ M$_\odot$, only 78 (37 per cent) pass the S/N criteria, while 73 per cent of the VYGs within the same mass range have good emission-line measurements (36 out of 49). The higher sSFRs of low-mass VYGs compared to the those of the CSGs might explain why a larger fraction of low-mass CSGs is excluded from our analysis of the oxygen abundances. Galaxies with higher sSFRs have stronger emission lines that result in flux measurements with higher S/N ratios. Therefore, the selection effects introduced by this cut in S/N affect the low-mass VYG and CSG samples differently, and the comparison between their oxygen abundances must be seen with caution.

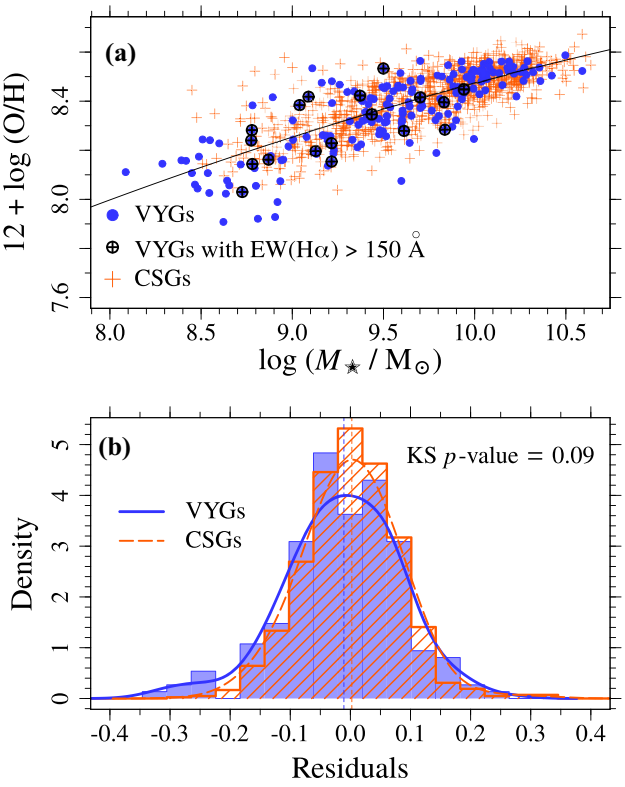

Figure 9. Ionized gas oxygen abundances as a function of stellar mass. The abundances were derived using emission-line fluxes from the SDSS table `GalSpecLine` (MPA-JHU). The notation is the same as in Fig. 3. The bottom panel shows the distribution of residuals from the second-order polynomial fit for the VYG (*blue histograms*) and control (*orange*) samples.

Finally, there are six low-metallicity VYGs with oxygen abundances 3σ below the log (O/H + 12) versus stellar mass relation: They have lower metallicities for their stellar masses. However, this result is not statistically significant.

4.2 Neutral hydrogen content

The atomic gas content of our galaxies was investigated using the data from the Arecibo Legacy Fast ALFA (ALFALFA) survey (Haynes et al. 2011, 2018). The ALFALFA H I source catalogue contains 31 502 detections, and approximately half of our objects lie in the ALFALFA–SDSS overlap region (∼4000 deg^2). The coordinates of the most probable optical counterpart (OC) of each H I detection are available in the catalogue. We cross-matched our VYG and CSG samples by adopting a maximum angular separation

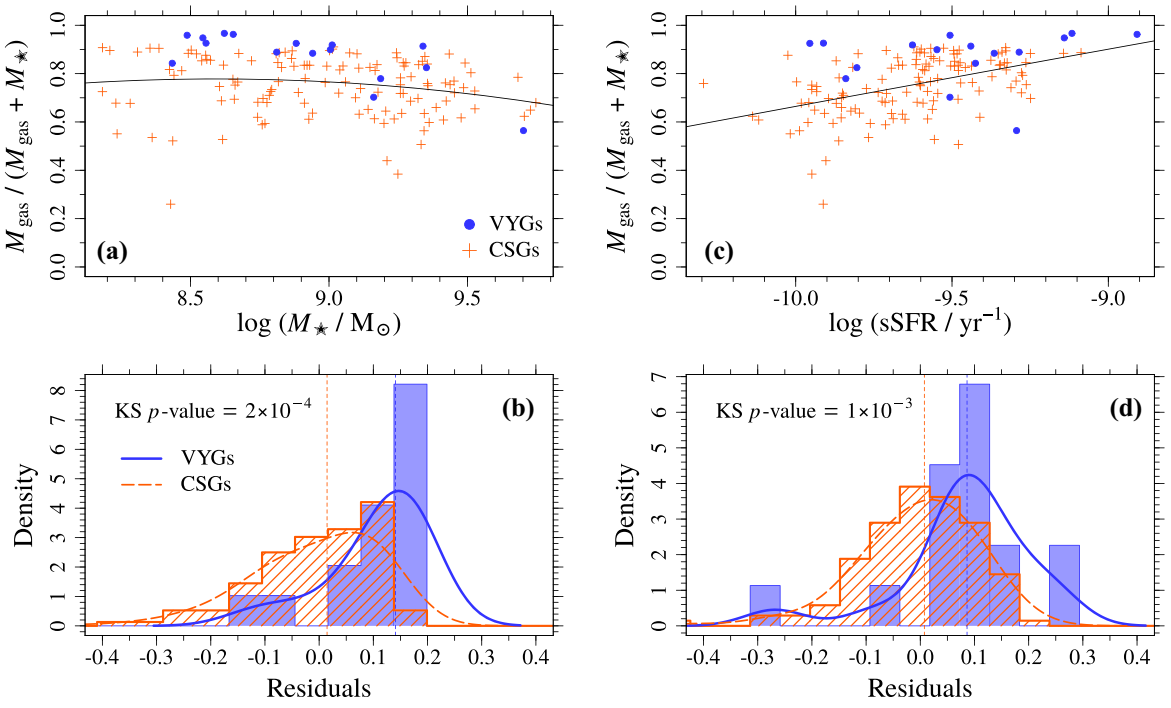

Figure 10. Fraction of atomic gas in galaxies with H I detections in the ALFALFA survey as a function of galaxy stellar mass (*upper left*) and log (sSFR) (*upper right*). The residuals from the linear fits to the CSG $f_{\rm H\,I}$ versus stellar mass and versus log (sSFR) relations are shown in the *bottom panels*. The *blue* and *orange* histograms correspond to the VYG and CSGs, respectively and, in each panel, we indicate the p-values of the KS tests.

of 5 arcsec between the OC and the SDSS galaxy. We found 149 galaxies with H I detections, among which 19 are VYGs and 130 are CSGs, corresponding to 9.2 per cent and 10.5 per cent of these samples, respectively. For each galaxy, we computed the atomic gas mass fraction $f_{\rm gas}$, i.e. the gas mass divided by the sum of the gas and stellar masses, $f_{\rm gas} = M_{\rm gas}/(M_{\rm gas} + M_\star)$. We accounted for the He gas by assuming that the total gas mass satisfies the relation $M_{\rm gas} = 1.33\,M_{\rm H\,I}$. The stellar masses were obtained from our SPS analysis, using the V15 spectral model with STARLIGHT.

Measurements of H I masses may be contaminated by the neutral gas of nearby galaxies because of the large Arecibo beam size (FWHM = 3.5 arcmin). We checked for such neighbour contamination by using the SDSS spectroscopic catalogue to identify the VYGs and CSGs that have other galaxies within a radius of 1.8 arcmin (∼90 per cent of the beam flux) and $\Delta z \leq 2\,W_{20}/c$, where W_{20} is the velocity width of the H I line profile in km s^{-1}, measured at the 20 per cent level of each of the two peaks on the low- and high-velocity horns of the profile (see Springob et al. 2005), and c is the speed of light. Although this procedure misses galaxies fainter than the magnitude limit of the SDSS spectroscopic observations ($m_{\rm r} \leq$ 17.77), we use the spectra to estimate the total stellar mass within the Arecibo beam. The stellar masses of the nearby galaxies were computed using the STARLIGHT code with the V15 spectral models as described in Section 2.

We found that three VYGs and five CSGs (15.8 per cent and 3.8 per cent, respectively) have other galaxies at distances ≤ 1.8 arcmin and within the Δz limit described above. The fraction of VYGs with nearby galaxies is 4.1 times higher than that of the CSGs. However, the stellar masses of the galaxies around the VYGs are lower than those of galaxies in the CSG surroundings. The total stellar mass within the beam is 0.26 dex higher (median) than the mass of the VYGs, while for the CSGs, the total mass is ∼0.9 dex higher (median). The median gas fraction computed with the total stellar mass within the beam, $M_{\rm gas}/(M_{\rm gas} + M_{\star,\rm beam})$, is 0.88 and 0.66 for these VYGs and CSGs, respectively. When comparing these values with the median gas mass fractions of VYGs and CSGs with no nearby galaxies within 1.8 arcmin ($\langle f_{\rm gas}\rangle = 0.91$ and 0.76, respectively), we see that the CSG gas fractions are more affected by nearby galaxies. Therefore, we conclude that the presence of other galaxies does not lead to higher VYGs gas fractions compared to the CSGs.

In Fig. 10, we show how $f_{\rm gas}$ varies with galaxy stellar mass (Fig. 10a) and sSFR (Fig. 10c). The three VYGs and five CSGs that have nearby galaxies within the Arecibo beam, identified as described above, are excluded from the plot. For a given stellar mass and sSFR, the fraction of atomic gas in VYGs is systematically larger compared to the one in CSGs. About 80 per cent of the VYGs with H I detections and no nearby galaxies (13 out of 16 VYGs) have $f_{\rm gas} > 0.8$, while only 42 per cent of the CSGs have such high gas mass fractions (52 out of 125 CSGs). We fitted the CSG $f_{\rm gas}$ versus stellar mass and versus log (sSFR/yr^{-1}) relations, and found that 14 out of 16 VYGs (87.5 per cent) lie above the best-fitting relation. The KS tests applied to the distributions of the residuals provide p-values = 2×10^{-4} and 1×10^{-3} (Figs 10b and d).

The gas mass fractions of VYGs are ∼ 20 per cent higher than those in CSGs. These large values might be due to the fact that VYGs have not managed to convert the neutral gas into stars before the last Gyr. We explore in the next section whether environmental effects can be the cause of that delayed SF.

It is important to note that the gas mass fractions and the properties inferred from the SDSS spectra are measured within different volumes of the galaxy, since the ALFALFA beam is ∼70 times the size of the SDSS fibre.

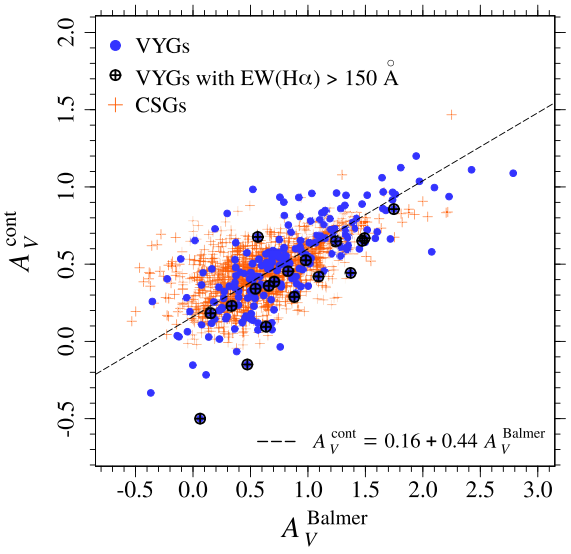

Figure 11. Internal extinction estimated from spectral fitting analysis using STARLIGHT versus extinction obtained from the Balmer decrement. The notation is the same as in Fig. 3. The *dashed line* indicates the relation $A_V^{\rm cont} = 0.16 + 0.44\,A_V^{\rm Balmer}$, which is the relation found by Calzetti et al. (2000) plus an offset of 0.16 mag.

4.3 Dust content and internal extinction

We estimated the amount of dust in our galaxies using two measures. We first derived the extinction in the V band, $A_V^{\rm cont}$, from the SPS analysis using the V15 spectral model with STARLIGHT and assuming the Cardelli et al. (1989) reddening law. We also measured the dust extinction from the Balmer decrement, $A_V^{\rm Balmer}$, which is inferred from the ratio of the H α and H β emission-line fluxes $F_{\rm H\,\alpha}/F_{\rm H\,\beta}$. This ratio should be 2.87 in the case of no extinction and a temperature $T = 10\,000$ K (Savage & Mathis 1979). To derive $A_V^{\rm Balmer}$, we first compute the colour excess $E(B-V)$ given by (see e.g. Momcheva et al. 2013; Domínguez et al. 2013, for details)

$$E(B-V) = \frac{2.5}{k(\lambda_{\rm H\,\beta}) - k(\lambda_{\rm H\,\alpha})} \log\left[\frac{F_{\rm H\,\alpha}}{2.87\,F_{\rm H\,\beta}}\right],$$

where $k(\lambda_{\rm H\,\beta})$ and $k(\lambda_{\rm H\,\alpha})$ are the values of the reddening curves at H β and H α wavelengths, respectively. The factor $2.5/[k(\lambda_{\rm H\,\beta}) - k(\lambda_{\rm H\,\alpha})] = 2.33$ for the Cardelli, Clayton & Mathis (1989) reddening law. The extinction in the V band is then computed as $A_V^{\rm Balmer} = R_V\,E(B-V)$, with $R_V = 3.1$.

We compared the two extinction estimates in Fig. 11, and it can be seen that $A_V^{\rm Balmer}$ tends to be greater than $A_V^{\rm cont}$. This result has been already found in previous studies (e.g. Calzetti et al. 2000; Asari et al. 2007; Kreckel et al. 2013; Florido et al. 2015; Riffel et al. 2021) and suggests that $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ are probing dust in different components of the galaxy ISM. The $A_V^{\rm Balmer}$ value is dominated by dust absorption within molecular clouds in active star-forming regions, while the reddening of the stellar continuum is mainly due to absorption by dust distributed more homogeneously through the galaxy ISM. The slope of the relation between $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ found by Calzetti et al. (2000) for starburst galaxies is consistent with our measurements, as indicated by the line shown in Fig. 11. However, we find an offset of 0.16 mag relative to the Calzetti et al. relation, which might be due to more diffuse dust in our galaxies compared to the starburst galaxies studied by these authors.

In Figs 12(a) and (c), we show $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ as a function of stellar mass for the VYGs and CSGs. We fitted the $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ versus $\log M_\star$ relations for the CSGs and the residuals from these fits are shown in Figs 12(b) and (d). The VYGs have systematically higher internal extinctions, with KS tests indicating high statistical significance (p-values $<10^{-6}$). However, Figs 12(a) and (c) suggest that the higher dust content of VYGs relative to CSGs is most pronounced at high galaxy masses and is non-existent ($A_V^{\rm Balmer}$) or reversed ($A_V^{\rm cont}$) at low masses. We will discuss the differences between the VYGs and CSGs in the high- and low-mass regimes in Section 6.4.

The amount of light absorbed in the interstellar medium depends on the size, geometry, and inclination of the galaxy. By construction, the distribution of the CSG physical sizes is close to that of the VYGs, since the angular sizes and redshifts were included in the matching procedure to construct the control sample, leading to similar θ_{50} versus z distributions shown in Fig. 2(c). In any event, we compared the VYG and CSG $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$, both normalized by galaxy radius: $A_V^{\rm Balmer}/R_{50}$ and $10^{-0.4\,A_V}/R_{50}$. After fitting the normalized quantities versus stellar mass relations and analysing the residuals, we still find significantly higher extinction in VYGs, with KS p-values $= 2\times10^{-8}$ (for $A_V^{\rm Balmer}/R_{50}$) and 2×10^{-7} (for $10^{-0.4\,A_V}/R_{50}$).

We assessed the dependence of the internal extinction on the galaxy geometry and inclination by using the minor-to-major axial ratio b/a as a proxy for these properties. Admittedly, this is a very simplified approach, given the complexity of the ISM structure and geometry, but the detailed characterization of the dust content of the ISM is beyond the scope of this paper. We find that $A_V^{\rm cont}$ increases with decreasing b/a, as expected if b/a indicates the inclination of disc galaxies. The Kendall and Spearman tests show that this anticorrelation is statistically significant, with coefficients $\tau = -0.15$ and $\rho = -0.22$, and p-values $< 10^{10}$ for both tests. On the other hand, $A_V^{\rm Balmer}$ does not show any dependence with the axial ratio, with Kendall $\tau = 0.03$ and Spearman $\rho = 0.04$ (p-values $= 0.12$ and 0.13). The difference in the behaviour of $A_V^{\rm cont}$ and $A_V^{\rm Balmer}$ with galaxy inclination reinforces the idea that they are probing dust in different components of the galaxy ISM.

After taking into account the dependence of $A_V^{\rm cont}$ with b/a by fitting the relation for the CSGs and analysing the residuals, the VYGs still have more extinction, with KS p-values $= 1\times10^{-5}$ (for $A_V^{\rm Balmer}$) and 5×10^{-6} (for $A_V^{\rm cont}$). Therefore, the differences between $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ of VYGs and CSGs do not appear to be due to distinct sizes, geometries, and/or inclination angles, but must be truly related to the amount of dust within these systems, which is significantly higher in VYGs compared to the general population of galaxies.

We extracted the IR magnitudes from the *Wide-field Infrared Survey Explorer* (*WISE*; Wright et al. 2010) data base, and computed the r–$W1$ and r–$W2$ colours for the VYGs and CSGs, where $W1$ and $W2$ are the *WISE* bands at 3.4 and 4.6 μm, respectively. We fitted the r–$W1$ and r–$W2$ versus stellar mass relations, and the residuals show that VYGs tend to be redder than the CSGs, which is consistent with an excess of dust in VYGs.

The higher amount of dust in VYGs compared to the CSGs may be a consequence of their higher SFR, since there is a correlation between the dust mass and SFR (e.g. da Cunha et al. 2010; Hjorth, Gall & Michałowski 2014). In Fig. 13, we show $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ as a function of $\log$ SFR. Both $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ increase with increasing SFR, but the correlation is much stronger for $A_V^{\rm Balmer}$.

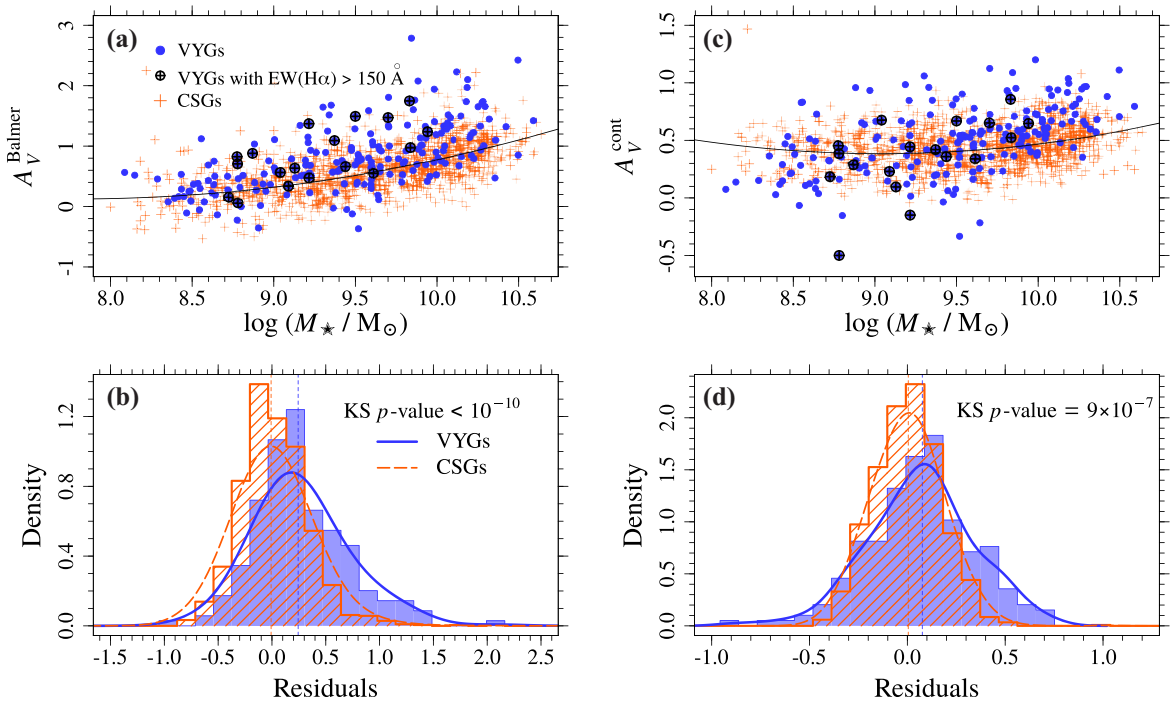

Figure 12. *Top:* Internal extinction estimated from the Balmer decrement (*left*) and from spectral fitting analysis using STARLIGHT (*right*), both as a function of stellar mass. The solid lines in both panels indicate the second-order polynomial fits to the $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ versus stellar mass relations. The notation is the same as in Fig. 3. *Bottom:* Residuals from the second-order polynomial fit for the VYG (*blue histograms*) and control (*orange*) samples.

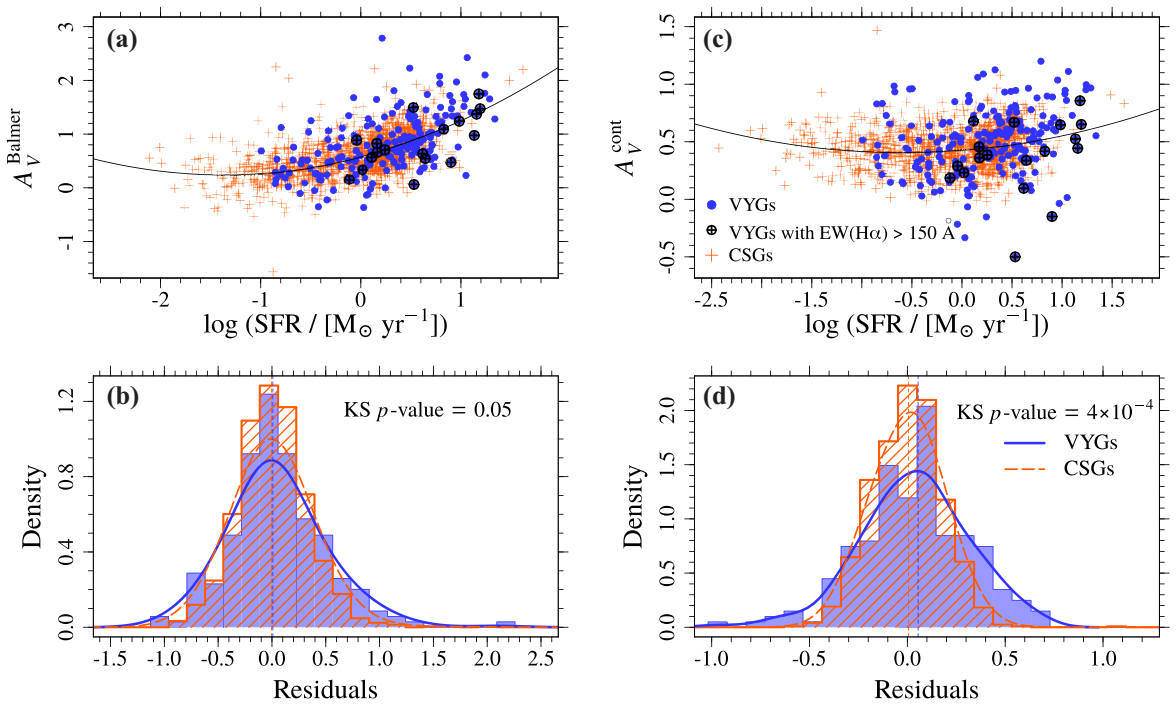

Figure 13. *Top:* Internal extinction estimated from the Balmer decrement (*left*) and from spectral fitting analysis using STARLIGHT (*right*), both as a function of the SFR. The solid lines in both panels indicate the second-order polynomial fits to the $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ versus log SFR relations. The notation is the same as in Fig. 3. *Bottom:* Residuals from the second-order polynomial fit for the VYG (*blue histograms*) and control (*orange*) samples.

Kendall and Spearman correlation tests confirm the tighter relation between $A_V^{\rm Balmer}$ and SFR compared to the $A_V^{\rm cont}$–log SFR relation. We obtained the following coefficients and p-values: $\tau = 0.45$, $\rho = 0.63$, and p-value $< 10^{-15}$ ($A_V^{\rm Balmer}$–log SFR relation); and $\tau = 0.08$, $\rho = 0.13$, and p-value $< 10^{-5}$ ($A_V^{\rm cont}$–log SFR relation). The different behaviour between $A_V^{\rm Balmer}$ and $A_V^{\rm cont}$ suggests, again, that they trace the absorption by dust of different ISM components.

As shown in Fig. 13(b), after removing the dependence with SFR, we see only a marginally significant difference between the VYG and CSG $A_V^{\rm Balmer}$. This result shows that the amount of dust in VYGs is higher due to their higher SFRs compared to those of CSGs. On the other hand, after taking the $A_V^{\rm cont}$–log SFR relation into account, VYGs have higher $A_V^{\rm cont}$ values than CSGs (Fig. 13d). These results suggest that, while the dust mass versus SFR relations for H II regions in VYGs and CSGs are similar, the amount of diffuse ISM dust in VYGs is higher compared to that of CSGs. Given that VYGs have a significant fraction of $\lesssim$1 Gyr stellar populations, they are expected to have a large number of TP-AGB stars polluting the ISM with dust. Another possibility is that the time-scales for the destruction of the ISM dust by sputtering or other processes are longer than 1 Gyr. We will discuss the VYG and CSG internal extinction and their amount of dust in Section 6.6.1.

5 THE ENVIRONMENT OF VYGS

If physical processes external to the VYGs are responsible for the recent SF activity, their environment is expected to be different from that of the CSGs. To investigate this possibility, we analyse successively the *local*, *group*, and *large-scale* environments of our sample galaxies. We identified possible local effects due to interactions with other galaxies in the immediate surroundings, by searching for neighbours within 200 kpc from each of our sample galaxies. We also visually inspected the galaxy images and classified them according to the presence of tidal features indicating interactions with neighbours or recent mergers. We then characterized the group environment of a galaxy by the halo mass of the group nearest to it and its position within that group, i.e. its distance relative to the group centre. Lastly, we determine the large-scale environment by the position of the galaxies relative to voids and filaments. We describe our approach in detail below.

5.1 Local environment

To investigate the local environment, we have computed for each galaxy in the VYG and control samples the distance to its closest neighbour, as follows. We retrieved from the SDSS-DR12 photometric catalogue all galaxies within 200 kpc from each of our sample galaxies, not requiring that the neighbour has spectroscopic observations. We adopt this approach because the great majority of the neighbour galaxies do not have listed redshifts in the SDSS. We selected objects with magnitudes brighter than $m_r + 1$, where m_r is the extinction-corrected Petrosian magnitude in the r band of our sample galaxy.

We ensured not missing galaxies by requesting that at least 95 per cent of the region within 200 kpc from each sample galaxy lies within the SDSS coverage area. For this purpose, we adopted the SDSS-DR7 spectroscopic angular selection function mask[5] provided by the NYU Value-Added Galaxy Catalog team (Blanton et al. 2005), and assembled with the package `MANGLE 2.1` (Hamilton & Tegmark 2004; Swanson et al. 2008). After excluding the galaxies close to bright stars or to the borders of the survey, we get a sample of 883 galaxies, among which 117 are VYGs and 766 are CSGs.

To minimize the contamination by background galaxies, we discard objects with g–i colours 2σ above the red sequence, at the redshift of our sample galaxies. The method used to identify the red sequence at different redshifts is described in Appendix A in the supplementary material.

We found that 428 (out of 883) galaxies do not have any neighbour with magnitude $<m_r + 1$ within 200 kpc (54 VYGs and 374 CSGs, corresponding to 46.2 per cent and 48.8 per cent). Fig. 14 shows the distribution of distances to the closest neighbour with magnitude $<m_r + 1$ for the 455 galaxies that have a nearby object at distances $\leq$200 kpc. The median distances are 71 and 98.6 kpc for the VYG and control sample, respectively.

We find that 9 out of 117 VYGs (7.7 per cent) have a close companion at distances smaller than 5 times R_{50}, while the number for CSGs is 22 out of 766 (2.9 per cent). Fisher's and Barnard's tests indicate that this difference has marginal statistical significance, with a p-values of 0.01 and 0.05, respectively. These results indicate that the VYGs are more likely to be interacting or merging with nearby galaxies than the CSGs.

5.2 Interactions and mergers with neighbour galaxies

Besides identifying the closest neighbour, as described above, we have also looked for signs of interactions or features indicating recent merger events. To identify merging systems, we use the classification from the GZ1 project (Lintott et al. 2011). The method consists in converting a set of visually inspected classifications by hundreds of thousands of volunteers into a single parameter, $p_{\rm merger}$, which corresponds to the weighted fraction of votes in the 'merger' category. Darg et al. (2010) have shown that galaxies with $p_{\rm merger} \gtrsim 0.4$ are, in fact, true mergers. Although $p_{\rm merger}$ is certainly related to the probability that a galaxy is part of an ongoing merger, adopting a single critical $p_{\rm merger}$ threshold to classify a galaxy as merging may be oversimplistic.

To avoid choosing a specific threshold, we computed the fraction of VYGs and CSGs that could be classified as merging systems for different $p_{\rm merger}$ threshold values, as shown in Fig. 15. Regardless of the threshold adopted, the fraction of VYGs that are merging is $\sim$2 times higher than that of the CSGs (KS p-value $= 5 \times 10^{-8}$). In the same figure, we also show the merger fraction according to the morphological classification by DS18. The fractions are higher than those based on Galaxy Zoo, and the difference between the VYGs and CSGs is smaller (1.5 times higher), but still with a high statistical significance (KS p-value $= 0.004$).

However, the Galaxy Zoo classification is biased towards a specific phase of the merger, i.e. when the two galaxies are close enough to be classified as merging. Interacting systems in the early phases of the merger and those that may never merge will be missed. Besides, the features indicating the post-merger phase might be only identifiable by expert astronomers.

Therefore, we performed our own visual classification based on the presence of tidal features indicating interactions with neighbour galaxies. The classification was performed independently by five members of our team. Each participant classified 414 galaxies (207 VYGs and 207 CSGs for comparison), and was asked to identify

[5]We used the file `sdss_dr72safe0_res6d.pol`, which can be downloaded from https://space.mit.edu/~molly/mangle/download/data.html.

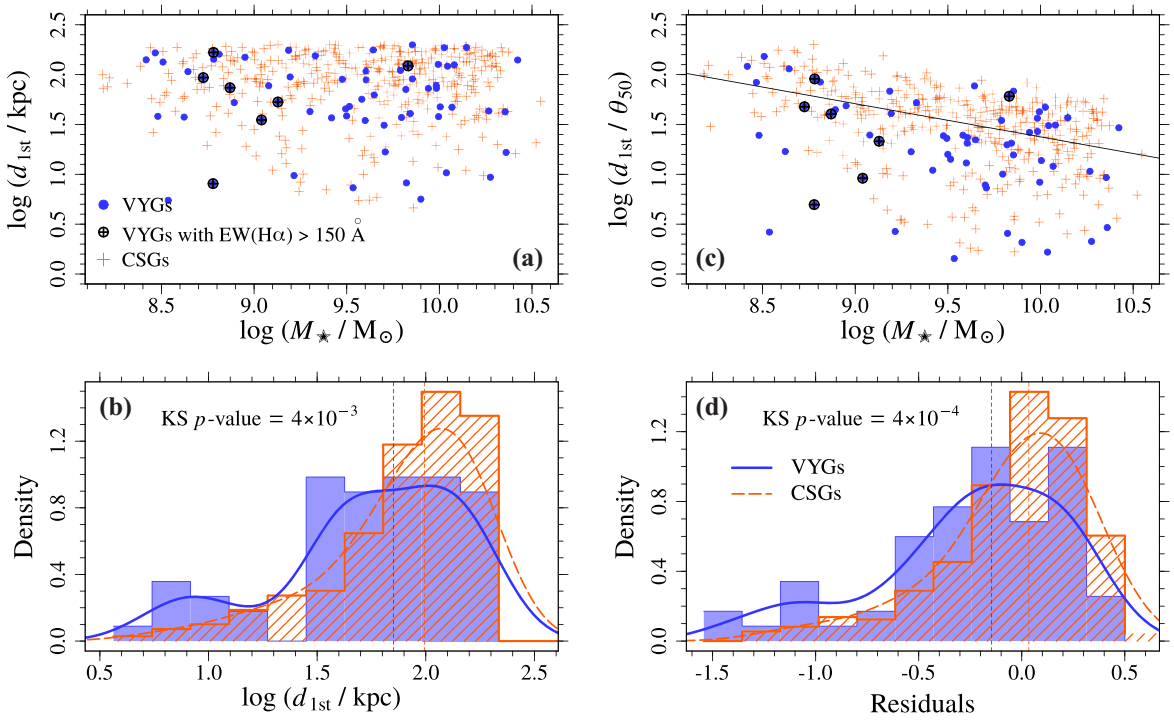

Figure 14. *Top:* Distance to the closest neighbour galaxy with magnitude $<m_r + 1$, where m_r is the r-band Petrosian magnitude of our galaxies. We show the distances in kpc (*left*) and normalized by the galaxy radius containing 50 per cent of the Petrosian flux, θ_{50} (*right*). We show only galaxies that have a neighbour within 200 kpc and that are far (> 200 kpc) from bright stars and the borders of the SDSS coverage area (63 VYGs and 392 CSGs). The notation is the same as in Fig. 3. *Bottom:* Distribution of distances in kpc (*left*) and the residuals from the linear fit to the normalized distance versus stellar mass relation of the CSGs (*right*). The *blue* and *orange* histograms correspond to the VYG and control sample, respectively, and, in each panel, we indicate the p-values of the KS tests.

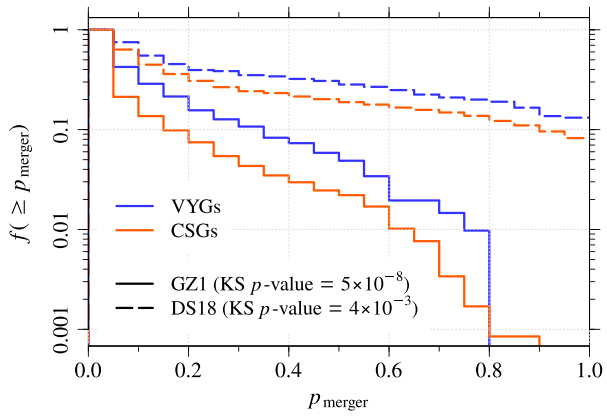

Figure 15. Fraction of galaxies that are classified as mergers in Galaxy Zoo (GZ1; *solid lines*) and by DS18 (*dashed lines*) for different threshold values of p_{merger}. The *blue* and *orange lines* correspond to the VYGs and CSGs, respectively.

tidal features and signs of interactions in the test galaxies as well as their neighbours. To classify the galaxies, each vote received a value v, depending on how visible the interaction features are. The participants could answer `yes`, if they clearly see interaction signs ($v_{yes} = 1$), `maybe` if they are not sure ($v_{mb} = 0.5$), or `no` if they see none ($v_{no} = 0$). For the neighbours, the vote for `no neighbours` received $v_0 = -1$. We then computed the mean of the values for the galaxy, $p_{i,g}$, and its neighbour, $p_{i,n}$, as follows:

$$p_{i,g} = \frac{v_{yes}\, N_{yes,g} + v_{mb}\, N_{mb,g}}{N_{votes}},$$

$$p_{i,n} = \frac{v_{yes}\, N_{yes,n} + v_{mb}\, N_{mb,n} + v_0\, N_{0,n}}{N_{votes}},$$

where $N_{yes,g}$ and $N_{mb,g}$ are the number of `yes` and `maybe` votes for the galaxy and $N_{yes,n}$, $N_{mb,n}$, and $N_{0,n}$ are the number of `yes`, `maybe`, and `no-neighbour` votes for the galaxy neighbour. If the mean of the votes is $p_i \geq 0.8$, the galaxy is classified as 'interacting'. More details of our interaction classification scheme are given in Appendix B in the supplementary material.

The advantage of our classification scheme of interactions is that it also considers the neighbouring galaxies. In Table 1, we compare the fractions of interacting VYGs and CSGs regardless of whether

Table 1. Fractions of VYGs and CSGs that show signs of interactions.

Signs of interaction	VYGs	Control	p-value	
			Fisher	Barnard
(1)	(2)	(3)	(4)	(5)
with neighbours or not	40.6 ± 3.0%	23.2 ± 3.8%	0.0002	0.0002
galaxy and neighbour	12.6 ± 2.0%	6.3 ± 1.2%	0.04	0.04
galaxy and close neighbour	12.1 ± 1.8%	6.3 ± 1.2%	0.06	0.05
with no neighbour	8.2 ± 1.3%	3.9 ± 2.5%	0.10	0.09

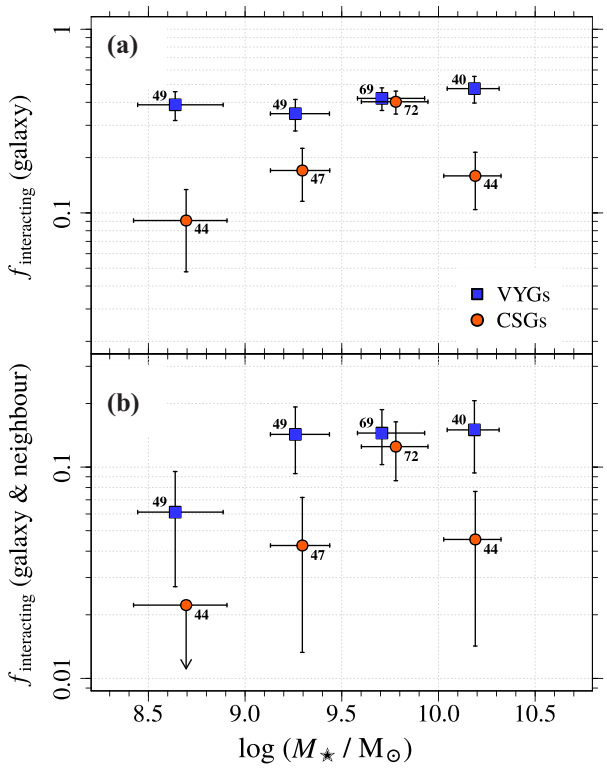

Figure 16. Fraction of galaxies with visually detected signs of interaction (from our five-expert analysis) among VYGs (*blue squares*) and CSGs (*orange circles*), as a function of their stellar mass. In the *upper panel*, we include all galaxies classified as interacting, regardless of neighbours. In the *bottom panel*, we show the fractions of systems where the neighbour galaxy also shows signs of interactions. The *error bars* indicate the 1σ binomial errors for $f_{\rm interacting}$ and the 16th and 84th percentiles of log $M_\star$.

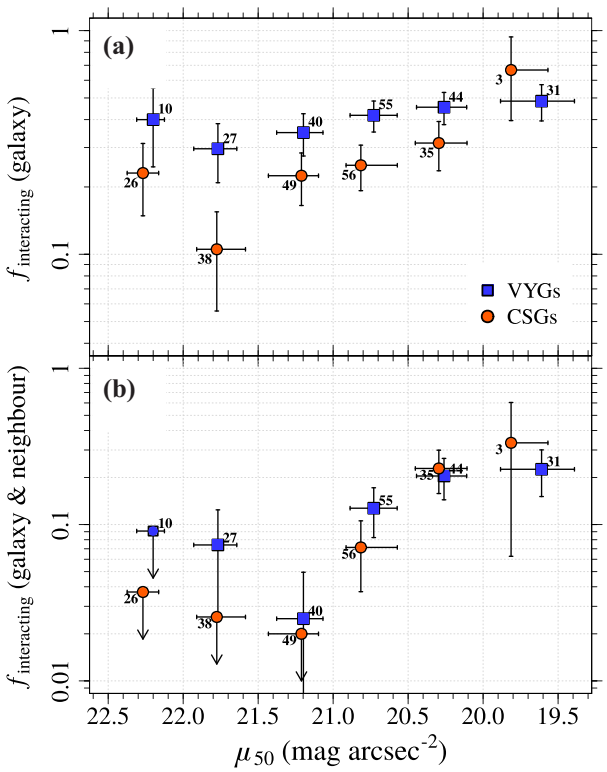

Figure 17. Fraction of VYGs (*blue squares*) and CSGs (*orange circles*) that show signs of interactions in bins of galaxy surface brightness. The notation is the same as in Fig. 16.

there is a neighbour, and also those of galaxies that have neighbours classified as interacting. We also computed the fractions of interacting galaxies that have neighbours with no sign of interaction or no neighbours at all. In all three cases, the fraction of VYGs classified as interacting is roughly twice that of CSGs, in agreement with the results obtained with the Galaxy Zoo classification.

In Fig. 16, we show how the fractions of galaxies, classified as interacting according to our scheme, vary as a function of galaxy stellar mass. In the upper panel, we show the fractions of all interacting galaxies regardless whether there is a neighbour or not. The fraction of VYGs that are interacting is roughly independent of stellar mass at ∼30-40 per cent, while the corresponding fractions of CSGs increase with $M_\star$ up to $M_\star \sim 10^{10}$ M$_\odot$, and decrease again for $M_\star \gtrsim 10^{10}$ M$_\odot$. The difference between these two samples is larger for low-mass galaxies: 38.8 per cent of the VYGs with $M_\star < 10^9$ M$_\odot$ are classified as interacting, a fraction that is 8.5 times higher than that of CSGs in the same mass range ($f_{\rm interacting} = 4.5$ per cent).

The fraction of interacting galaxies with a companion also classified as interacting is shown in the lower panel of Fig. 16. For the VYG sample, $f_{\rm interacting} = 6.1$ per cent for galaxies with $M_\star < 10^9$ M$_\odot$ and varies between 10 and 15 per cent for more massive objects ($M_\star > 10^9$ M$_\odot$). These fractions are >6.7 times higher when compared to those of the CSGs, except for the mass bin of $9.5 < \log(M_\star/{\rm M}_\odot) < 10$, where the $f_{\rm interacting}$ values are similar for the two samples.

Could this higher fraction of interacting galaxies among VYGs be due to a surface brightness effect? The VYGs have higher surface brightness compared to the CSGs, as shown in Fig. 7. This could naturally lead to higher fractions of galaxies classified as interacting, since it would be easier to identify structures and features in the galaxy images. To test this hypothesis, we compared $f_{\rm interacting}$ in bins of μ_{50}. As we can see in Fig. 17(a), the fraction of interacting VYGs is always higher than that of CSGs with similar μ_{50}, with differences ranging from 1.5 to 5.2 times higher, except for the brightest μ_{50} bin. However, the difference is within the errors and might be a result of poor statistics, since there are only three CSGs in this bin. We conclude that the higher fraction of interacting galaxies among VYGs is not caused by surface brightness effects.

Finally, in Fig. 17(b), we show $f_{\rm interacting}$ when both the galaxy and its neighbour(s) show signs of interactions. For the faintest μ_{50} bin ($\mu_{50} > 22$), no systems are classified as interacting. The CSG fractions are lower than those of the VYGs for galaxies with $20.5 < \mu_{50} < 22$, but they are similar for brighter objects ($\mu_{50} < 20.5$). However, the interpretation of these results is not straightforward, since these fractions also depend on the surface brightness of the neighbours that can be fainter than that of our galaxies.

In summary, VYGs exhibit a significantly higher fraction of mergers and interactions than do CSGs.

5.3 Group environment

We now analyse the group environment of the VYGs and CSGs. We selected groups and clusters from the updated version of the catalogue compiled by Yang et al. (2007). The catalogue contains 473 482

Table 2. KS p-values for different comparisons between VYGs and CSGs.

	$12.3 \leq \log M_{\rm halo} < 13.5$				$13.5 \leq \log M_{\rm halo}$			Field	Total
	Centrals	0–0.5 $r_{\rm vir}$	0.5–1 $r_{\rm vir}$	1–3 $r_{\rm vir}$	0–0.5 $r_{\rm vir}$	0.5–1 $r_{\rm vir}$	1–3 $r_{\rm vir}$	$> 3\,r_{\rm vir}$	
VYGs	**17 (23.6%)**	14 (19.4%)	10 (13.9%)	**31 (43.1%)**	6 (25%)	7 (29.2%)	**11 (45.8%)**		
		72 (36%)				24 (12%)		104 (52%)	200
Control	**33 (11.5%)**	38 (13.2%)	36 (12.5%)	**180 (62.7%)**	35 (16.7%)	33 (15.7%)	**142 (67.6%)**		
		287 (24%)				210 (17.6%)		697 (58.4%)	1194
				Fisher's tests					
p-value	**0.013**	0.192	0.843	**0.003**	0.392	0.146	**0.042**		
		0.0006				0.052		0.104	
				Barnard's tests					
p-value	**0.017**	0.108	0.535	**0.005**	0.213	0.095	**0.044**		
		0.002				0.065		0.065	

Notes. Galaxies in groups are divided into bins of projected distance to the group centre, R. The number of central galaxies is shown only for low-mass groups, since none of our VYG and CSGs are found in the centre of high-mass groups. The last lines show the p-values of Fisher's and Barnard's tests comparing the fractions of VYGs and CSGs in each bin of R and in the field. Results with p-values <0.05 are highlighted in boldface.

groups drawn from a sample of 601 751 galaxies, mostly from the SDSS-DR7 (Abazajian et al. 2009). Only groups with $\log(M_{\rm halo}/{\rm M}_\odot) \geq 12.3$ were selected, and we assigned our sample galaxies to the nearest group, following the method described in Trevisan, Mamon & Stalder (2017b). A galaxy is assigned to the group that gravitationally attracts it the most, i.e. the group with the lowest distance, in units of virial radius, $r_{\rm vir}$.

We define $r_{\rm vir}$ of a group as the radius, r_{100}, of a sphere that is $\Delta_{\rm v} = 100$ times denser than the critical density of the Universe. We obtained $r_{\rm vir}$ by first deducing $r_{200,\rm m}$ (of spheres that are 200 times denser than the *mean* density of the Universe) from the $M_{200,\rm m}$ masses given in the Yang et al. (2007) catalogue (which are based on abundance matching with the group luminosities). We then calculated the $r_{\rm vir}$, following appendix A of Trevisan, Mamon & Khosroshahi (2017a) for the conversion from quantities relative to the mean density to those relative to the critical density, and the corresponding virial masses, $M_{\rm vir} = (\Delta_{\rm v}/2)\, H^2(z)\, r_{\rm vir}^3/G$.

To assign the galaxies to their nearest group, we compute the distances, d, between the galaxies and the group centres, assuming two regimes. For galaxies far away from the group, d is given by the redshift–space distance. For a galaxy close to a group, the strong redshift distortions are taken into account by using the overdensity in projected phase space introduced by Yang et al. (2005, 2007). We convert this overdensity to an equivalent redshift–space distance by joining the two estimators at a given radius, $R_{\rm n}$, which marks the transition between the non-linear and the linear regimes. We adopt $R_{\rm n} = 2\,r_{\rm vir}$, and galaxies can be assigned to distances up to $20\,r_{\rm vir}$ from a group. A small fraction (< 4 per cent) of the galaxies was not assigned to any group (7 VYGs and 48 CSGs). Among these galaxies, 22 of them were not assigned because they lie at $z < 0.01$, which is the lower redshift limit of the group catalogue. The other 33 galaxies lie close to the borders of the SDSS Survey or a bright star, and the group catalogue may be incomplete in these regions. Using the package `MANGLE 2.1`, as described in Section 5.1, we estimate that more than 20 per cent of the region within 500 kpc from these galaxies lies outside the SDSS coverage area.

We computed the fractions of VYGs and CSGs at different distances from the group/cluster centre. Table 2 indicates that roughly half of galaxies are found at distances $R > 3\,r_{\rm vir}$ from the group centres (52 per cent and 58.4 per cent of the VYGs and CSGs, respectively). Analysing the galaxies within groups, we see that VYGs reside preferentially in low-mass haloes compared to CSGs.

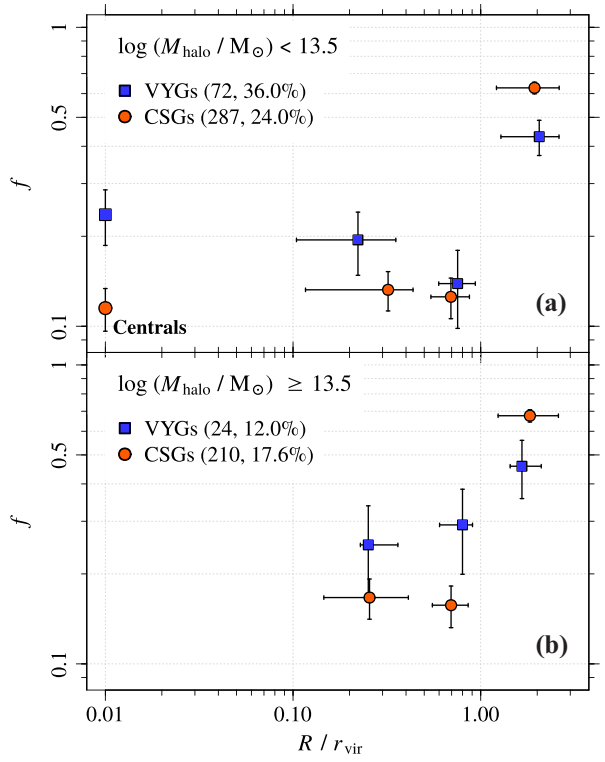

Figure 18. Distribution of radial positions relative to nearest group among VYGs (*blue squares*) and CSGs (*orange circles*). The *upper* and *lower panels* show the radial distributions of galaxies relative to low- and high-mass groups, respectively (there are no centrals among VYGs and CSGs in the higher mass groups, because nearly all centrals of high-mass groups are passive galaxies). The error bars indicate the 1σ binomial errors for f and the 16th and 84th percentiles of $R/r_{\rm vir}$.

As shown in Fig. 18 and Table 2, the fraction of VYGs that reside in groups with $\log(M_{\rm halo}/{\rm M}_\odot) < 13.5$ (close to the median group mass) is ∼50 per cent higher than the corresponding fraction of CSGs. Fisher and Barnard tests indicate a high statistical significance for this result (p-values of 0.0006 and 0.002, respectively).

The way that the VYGs are distributed within and around the group haloes is also different from that of the CSGs. By comparison

to CSGs, VYGs are more likely to lie within groups.[6] As shown in Fig. 18 and Table 2, 63 per cent of the CSGs in low-mass haloes are found between 1 and $3r_{\rm vir}$, while only 43 per cent of the VYGs reside in this region. The fraction of VYGs that are centrals is double that of CSGs, with statistical tests indicating a marginal significance (p-values $\sim$0.02).

Only a small fraction of VYGs and CSGs are in more massive haloes with $\log[M_{\rm halo}/{\rm M}_\odot] \geq 13.5$ (12 per cent of the VYGs and 17.6 per cent of the control sample). However, we find that VYGs are also more likely to be found in the inner parts of the these high-mass groups, with 54.2 per cent of the VYGs lying at distances $< 1r_{\rm vir}$. For the CSGs, this fraction is only 32.4 per cent, with statistical tests indicating that this difference is marginally significant (p-values $\sim$0.04).

From the results shown in Fig. 18 and Table 2, we see that the VYGs are more likely to be found in the inner parts of low-mass groups when compared to the CSGs. We find 41 VYGs that lie within $1r_{\rm vir}$ from the centre of low-mass groups. These numbers correspond to 20.5 per cent of all 200 VYGs that were assigned to a halo. This fraction is much higher than that of CSGs residing in the inner regions of low-mass groups: Only 9 per cent (107 out of 1194) of the CSGs are at distances $< 1r_{\rm vir}$ from the centre of haloes with $\log(M_{\rm halo}/{\rm M}_\odot) < 13.5$. Fisher's and Barnard's tests indicate high statistical significance, with p-values of 6×10^{-6} and 6×10^{-5}, respectively.

We checked if the definition of the group and cluster centres could affect our results. In our group assignment scheme, we assume the position of the brightest group galaxy to be the centre of the group. Since VYGs have young stellar populations, they are expected to be brighter than non-VYGs with similar masses, and this could lead to higher fractions of VYGs that are centrals compared to normal galaxies. Therefore, we repeated our assignment procedure using the most massive galaxy as the centre of the group haloes. We find that the number of VYGs that are central galaxies in haloes with $\log(M_{\rm halo}/{\rm M}_\odot) < 13.5$ remains unchanged (17 VYGs). On the other hand, the number of centrals in the control sample decreases from 33 (11.5 per cent) to 27 (9.2 per cent), and the statistical significance of the result of a larger fraction of centrals in VYGs increases (p-values = 0.002–0.003).

5.4 Large-scale environment: filaments and voids

The positions of our galaxies with respect to the large-scale filamentary structure were obtained from the catalogue of filaments established by Tempel et al. (2014). For each SDSS-DR7 galaxy, the catalogue provides the distance to the closest filament, $D_{\rm fil}$. Since the catalogue is based on SDSS-DR7 (while our samples were drawn from SDSS-DR12) and contains only galaxies in the main survey area ($6.5 \lesssim {\rm RA} \lesssim 18$ h), $D_{\rm fil}$ is available for only 185 VYGs and 1126 CSGs.

There is no statistical difference between the VYG and CSG distributions of distances to the nearest filament (KS test p-value is 0.22), and similar fractions of VYGs and CSGs are within $D_{\rm fil} \leq D_{\rm max} = 1$ Mpc: 35.7 per cent of the VYGs and 31.6 per cent of the CSGs (p-values = 0.27 and 0.31 for Fisher's and Barnard's tests, respectively). In addition, the luminosities of the filaments containing VYGs and CSGs within 1 Mpc are similar. We used the luminosities $L_{0.5}$ and $L_{1.0}$ computed by Tempel et al. (2014) as the sum of luminosities of observed galaxies that are closer than 0.5 and 1.0 h^{-1} Mpc from the filament axis. After removing the dependence of $L_{0.5}$ and $L_{1.0}$ on the stellar mass of the VYGs and CSGs, we see no significant differences between the distribution of luminosities of filaments containing VYGs and those containing CSGs (KS test p-values = 0.22 and 0.08 for $L_{0.5}$ and $L_{1.0}$, respectively). Since roughly one-half of VYGs and CSGs are within the virial spheres of groups, we investigated if the statistics of distance to nearest filaments are blurred by these galaxies. We repeated the analysis using only galaxies that are at distances $R > 3r_{\rm vir}$ from all groups, and we still do not see any significant differences between the VYG and CSG distributions of $D_{\rm fil}$, $L_{0.5}$, and $L_{1.0}$.

We repeated the analysis using different maximum distances to the filaments ($D_{\rm max} = 0.2$ up to 3 Mpc), but we find no significant evidence for higher (or lower) fractions of VYGs lying along the filaments compared to those of CSGs. The distributions of luminosities of filaments containing VYGs within $D_{\rm max}$ are also very similar to those of filaments containing CSGs, with KS tests indicating low statistical significance regardless of the $D_{\rm max}$ value adopted.

We also determined the position of our galaxies relative to voids by using the void catalogue by Sutter et al. (2012),[7] which is also based on SDSS-DR7. To cover the redshift range of our sample, we also used three catalogues of voids that were identified using different samples of SDSS galaxies: `dim1` (for galaxies at $0 \leq z < 0.05$), `dim2` ($0.05 \leq z < 0.1$), and `bright1` ($0.1 \leq z < 0.15$). We used the subcatalogues called '*centrals*' by Sutter et al. (2012). To avoid the biases introduced by the survey boundaries and masks, the *centrals* subcatalogues exclude voids that, when rotated in any direction about its barycentre, intersect a boundary galaxy.

For each VYG and CSG, we computed the distance to the centre of the closest void in units of the void radius, $D_{\rm void}$, assuming that the voids are spherical. The distances $D_{\rm void}$ are given by

$$D_{\rm void} = \frac{1}{r_{\rm void}\,(1+z_{\rm void})}\sqrt{[D_{\rm c}(z) - D_{\rm c}(z_{\rm void})]^2 + \left[\theta\,\overline{D_{\rm c}}\right]^2},$$

where z and $z_{\rm void}$ are the redshifts of the galaxy and the centre of the void, respectively; θ is the angular separation between the void centre and the galaxy; $D_{\rm c}(z)$ and $D_{\rm c}(z_{\rm void})$ are the comoving distances to the galaxy and to the void centre, respectively; $\overline{D_{\rm c}}$ is the mean comoving distance, $\overline{D_{\rm c}} = [D_{\rm c}(z) + D_{\rm c}(z_{\rm void})]/2$; and $r_{\rm void}$ is the void radius in Mpc. Since the catalogue is restricted to voids in the main survey area, we compute the distances for our galaxies within $6.5 \lesssim {\rm RA} \lesssim 18$ h only (187 VYGs and 1155 CSGs).

Almost all of our galaxies (both VYGs and CSGs) are located far from the centre of the voids. Only 8.4 per cent of the galaxies (113 out of 1342) are at distances $D_{\rm void} \leq 1$, and the p-values of statistical tests do not provide any evidence that the VYGs prefer (or avoid) these low-density environments. We obtain p-values of 0.48 and 0.67 when applying Fisher's and Barnard's tests to the fractions of VYGs and CSGs that are at distances $D_{\rm void} \leq 1$ (18 VYGs and 95 CSGs, corresponding to 9.6 per cent and 8.2 per cent, respectively). Both samples have similar $D_{\rm void}$ distributions, with median and standard deviations $D_{\rm void} = 2.52 \pm 1.38$ (VYGs) and 2.54 ± 1.27 (CSGs).

[6] Of course, projection effects limit us to a cylindrical view of groups, preventing us from knowing which galaxies are within or outside groups defined as spheres. These projection effects are much more severe for star-forming galaxies: Nearly half of those that are inside the virial cylinder actually lie outside the virial sphere (Mahajan, Mamon & Raychaudhury 2011). However, both VYGs and CSGs are star-forming systems, and it is unlikely that projection effects would affect these two samples differently.

[7] http://www.cosmicvoids.net/

We applied a KS test to compare the distributions and obtained a p-value of 0.22. In summary, the large-scale environments of VYGs and CSGs are similar.

6 DISCUSSION

6.1 How do VYGs differ from others?

Our comparison of the properties of VYGs and CSGs reveals that the VYGs are different from the general population of galaxies in many aspects. The results presented in Sections 3–5 can be summarized as follows:

(i) *VYGs are bluer and have higher sSFRs than the CSGs* (Fig. 3), which confirms that our sample of VYGs indeed has younger stellar populations.

(ii) *In VYGs, the gas has higher ionization ratios* (Fig. 4), which might be simply a consequence of higher sSFRs.

(iii) *VYGs contain a higher fraction of spheroidal systems compared to CSGs* (Figs 5 and 6). These VYG spheroids correspond to ∼6 per cent of sample and they are bluer than the CSG spheroids. On the other hand, we did not find significant differences between the overall distributions of VYG and CSG T types, but most of our galaxies appear to be irregulars, so these catalogues may not provide a good description of their morphologies.

(iv) *VYGs have higher concentrations and surface brightness* (Fig. 7), indicating that the SF activity in VYGs is occurring in the inner parts of the galaxy. However, we cannot determine how the size of the VYGs compares to the general population of galaxies, since our control sample was defined by using the redshifts and angular effective radii in the PSM procedure; therefore, the VYG and CSG distributions of physical radii are similar by construction.

(v) *VYGs are more asymmetric and more clumpy than CSGs* (Fig. 8), which may be a consequence of interactions with neighbour galaxies.

(vi) Among galaxies detected in H I, *VYGs have significantly higher fractions of atomic gas than CSGs* (Fig. 10). Around 80 per cent of the VYGs with H I detections have $f_{\rm gas} > 0.8$, while only 42 per cent of the CSGs have such high amounts of H I gas.

(vii) *The internal extinction in VYGs is higher than that in the CSGs* (Fig. 12), indicating that these young systems have a higher amount of dust compared to the general population of star-forming galaxies. We discuss the VYG internal extinction and dust content in Section 6.6.1.

(viii) *Compared to the CSGs, the VYGs are at smaller distances from their nearest companions and are more likely to be interacting/merging with a neighbour galaxy*. These results are shown in Figs 14–17 and Table 1, and will be discussed in Section 6.6.

(ix) Roughly half of the VYGs lie outside groups; however, compared to CSGs, *VYGs are more likely to be found in the inner parts of low-mass groups* (Fig. 18 and Table 2). We discuss this result in Section 6.6.

Our sample of VYGs includes some starburst galaxies with very large Hα EWs, which are expected to have different properties. So, one could argue that the differences that we find between the VYG and control samples could be due to these objects. However, even when excluding these galaxies from our sample, the differences between the VYGs and the CSGs are still statistically significant, as indicated by the p-values in columns 3 and 4 of Table 3.

Although VYGs differ from the general population of galaxies in all the properties listed above, the VYGs are very similar to the CSGs in the following aspects:

Table 3. p-values of KS tests applied to the control and VYG samples.

Property	KS p-values		
	All VYGs	VYGs with EW(Hα) $\leq$ 150 Å	VYGs with EW(Hα) $\leq$ 100 Å
(1)	(2)	(3)	(4)
log sSFR	$<10^{-10}$	9×10^{-10}	2×10^{-5}
$(g-i)_{\rm Petro}$	5×10^{-4}	2×10^{-3}	0.01
BPT diagram	$<10^{-10}$	7×10^{-8}	3×10^{-5}
log (O/H)	0.09	0.18	0.55
$f_{\rm gas}$	2×10^{-4}	2×10^{-4}	7×10^{-5}
$A_V^{\rm Balmer}$	$<10^{-10}$	$<10^{-10}$	2×10^{-10}
$A_V^{\rm cont}$	9×10^{-7}	3×10^{-7}	3×10^{-8}
C_1	1×10^{-7}	3×10^{-6}	2×10^{-6}
μ_{50}	$<10^{-10}$	$<10^{-10}$	6×10^{-8}
A_1	2×10^{-7}	3×10^{-5}	2×10^{-5}
S_1	3×10^{-3}	0.02	4×10^{-3}
$d_{1\rm st}$/kpc	4×10^{-3}	0.01	0.03
$d_{1\rm st}/\theta_{50}$	4×10^{-4}	5×10^{-3}	0.03

(i) *VYGs have the same gas metallicity as the CSGs* (Fig. 9), except for hints of lower gas metallicity of VYGs at the low-mass end. We discuss in Section 6.2 the implications of the lack of differences in the gas metallicities of VYGs versus CSG.

(ii) *The distribution of VYGs relative to cosmic filaments and voids is very similar to that of CSGs.*

6.2 The VYG metallicities and relation with other young galaxies in the local Universe

The similar gas metallicities of VYGs and CSGs suggest that SF in VYGs is not being fuelled by infalling metal-poor gas, but by gas that was already enriched by previous generations of stars. This result has a consequence when comparing the VYGs with other populations of VYGs. As already discussed in Section 1, a few low-mass star-forming galaxies with extremely low metallicities in the local Universe, such as the blue compact dwarf galaxies I Zw 18 and SBS 0335–052, are strong VYG candidates.[8] As shown by Izotov et al. (2019), these objects strongly deviate from the oxygen abundance versus $M_\star$ relation defined by the bulk of star-forming galaxies. Therefore, the normal metallicities of VYGs at a given mass suggest that VYGs do not resemble objects like I Zw 18.

Although we find that the gas mass fractions of VYGs with H I detections are very high, I Zw 18 has an even more extreme $f_{\rm gas}$ value. The H I mass of I Zw 18 is 50 times higher than its stellar mass, and the total neutral gas mass fraction is $f_{\rm gas} = 0.98$.[9] The most gas-rich VYG in our ALFALFA sample has $f_{\rm gas} = 0.97$, and only 3 out of 16 VYGs with H I detection have $f_{\rm gas} \geq 0.95$. However, the stellar mass of I Zw 18 is 40 times lower than the lowest masses of our VYG sample, so it is difficult to make a meaningful comparison between this galaxy and the VYGs.

Furthermore, while our VYGs tend to have close companions, I Zw 18 appears to be a very isolated system. Indeed, using the same approach described in Section 5.1 to investigate the local environment of our VYG sample, and assuming $m_{\rm r,I\,Zw\,18} = 16.4$ as the extinction-

[8] I Zw 18 and SBS 0335–052 are too close to lie in the redshift range of the parent `clean` galaxy sample.

[9] To compute the gas mass fraction of I Zw 18, we adopted $M_{\rm H\,I} = 1.25 \times 10^8$ M$_\odot$ (Engelbracht et al. 2008; Thuan et al. 2016) and $M_\star = 2.6 \times 10^6$ M$_\odot$ (Izotov et al. 2014).

corrected Petrosian magnitude in the r band of I Zw 18, we find that there is no other galaxy brighter than $m_r = m_{r,IZw\,18} + 1$ up to 200 kpc from I Zw 18. This suggests that interactions with nearby galaxies cannot be the mechanism triggering the burst of SF in I Zw 18, and other processes must be invoked to explain the SF activity in this galaxy. However, one cannot rule out the possibility that I Zw 18 has very recently grown by gas infall and/or mergers, as evidenced by its irregular morphology and the complex kinematics of its atomic gas (van Zee et al. 1998).

In any event, it is difficult to compare this population of extremely metal-deficient dwarf galaxies to the VYGs studied here. First, while these dwarfs have stellar masses of $\sim 10^{5.5}$–$10^{8.0}$ $M_\odot$ (Izotov et al. 2014, 2019), our sample is restricted to more massive objects with $M_\star > 10^8\,M_\odot$. Moreover, the techniques employed to study the stellar populations of nearby dwarfs are different from those used here. Our ages were determined through SPS analysis of the integrated galaxy spectra, while the age of objects like I Zw 18 is inferred from colour–magnitude diagrams of resolved stellar populations. In Paper II, we determined the age of I Zw 18 using the same method (with STARLIGHT using the V15 spectral model) and found that 100 per cent of its stellar mass was formed in the last 100 Myr, so this galaxy easily meets the VYG classification.

Since these metal-poor dwarfs are very nearby objects, with distances $\lesssim 20$ Mpc, it is very difficult to investigate their global environment. Catalogues of groups and clusters are not reliable at very low redshifts due to uncertainties introduced by peculiar velocities of galaxies. In addition, images of close objects contain more spatial information than those of galaxies at higher redshifts. Hence, morphological classification and morphometry of these galaxies (e.g. with MORFOMETRYKA) depend on galaxy distance, and correcting for this dependence to compare I Zw 18 analogues with our VYGs is beyond the scope of this paper.

6.3 VYGs versus late bloomer galaxies at intermediate redshifts

A sample of young systems at intermediate redshifts was recently studied by Dressler et al. (2018, hereafter DKA18). They derived the SFHs of galaxies at $0.45 < z < 0.75$ and with $M_\star > 10^{10}\,M_\odot$, and identified a galaxy population of *late bloomers* (LBs), i.e. galaxies that formed at least 50 per cent of their stellar mass within 2 Gyr of the epoch of observations. Their SFHs were inferred from the Carnegie-*Spitzer-IMACS* Survey photometry.

DKA18 found that LBs account for ∼20 per cent of galaxies at $z \sim 0.6$, and their fractions systematically decrease with decreasing redshift, and the most massive LBs ($M_\star \gtrsim 10^{11}\,M_\odot$) practically disappear at $z \sim 0$. In Paper II, we showed that the fractions of VYGs in the local Universe are consistent with the extrapolation of the fractions of LBs versus redshift determined by DKA18 (see fig. 14 in Paper II).

Are the properties of VYGs similar to those of LBs? It is difficult to answer this question, given the stellar mass ranges of the VYG and LB samples; while LBs are, by definition, more massive than $M_\star > 10^{10}\,M_\odot$, most of our VYGs have $M_\star \lesssim 10^{10}\,M_\odot$, and it is well known that galaxy properties correlate with stellar mass. A significant fraction of LBs appear to be spiral galaxies, but DKA18 also find some LBs with early-type morphology, with SEDs that are consistent with that of a post-starburst galaxy. Although VYGs appear to be mostly irregular, it is interesting that we also find an excess of spheroidal systems, suggesting that some common mechanism is producing young spheroidal systems at different redshifts.

One very interesting result obtained by DKA18 is that LBs avoid lying close to galaxies that are not LBs up to distances of

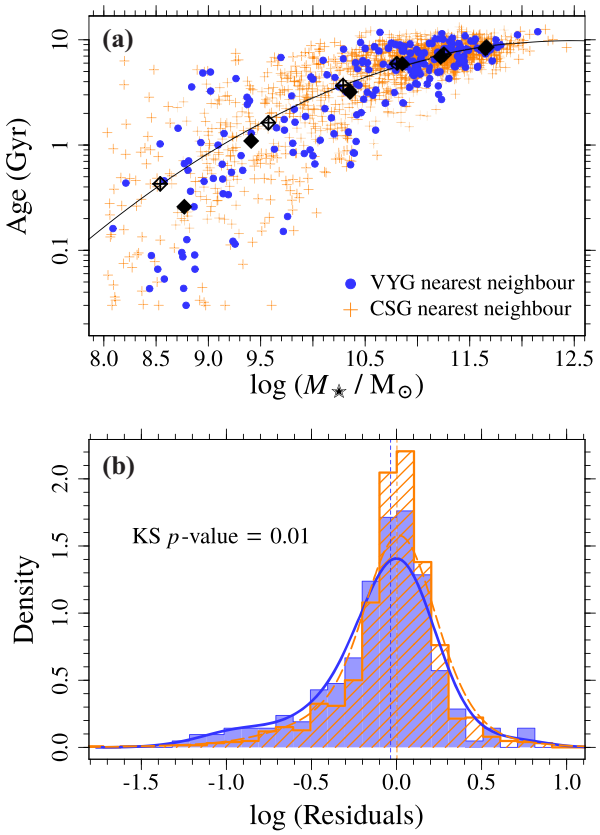

Figure 19. *Top:* Mass-weighted age of the VYG and CSG nearest SDSS spectroscopic galaxy (*blue* and *orange* symbols, respectively). The median ages in bins of stellar mass are indicated by the *black diamonds*, and the best fit to the age–mass relation for the CSG neighbours is shown as the *black solid line*. *Bottom:* Distribution of the logarithm of the residuals from the best-fitting relation shown in the top panel. The *blue* and *orange histograms* correspond to the VYG and CSG nearest neighbours, respectively.

8 Mpc, indicating that galaxy SFHs can trace local and large-scale environmental histories. In Section 5.4, we found no difference between the VYG and CSG positions relative to filaments and voids, and the luminosities of the filaments close to VYGs and CSGs are also similar.

However, the distributions of VYGs and CSGs relative to the large-scale structures being similar does not mean that the properties of galaxies within these structures are also similar. To investigate this, we used the SDSS spectroscopic catalogue to identify the VYG and CSG nearest galaxies (in comoving units). As shown in Fig. 19, the nearest neighbours of VYGs tend to be younger compared to those that are close to CSGs, in agreement with DKA18. The difference is more pronounced for low-mass neighbours; VYG neighbours with $M_\star \leq 10^{10}\,M_\odot$ are ∼0.17 dex, on average, younger than CSG neighbours in the same stellar mass range (p-value = 0.02). Since VYGs tend to have more close companions than CSGs, we checked if the neighbours in the immediate surroundings are leading to this result by comparing the ages of closest galaxies at distances > 1 Mpc only. We confirm that we see the effect in both small (< 1 Mpc, p-value = 0.02) and large (> 1 Mpc, p-value = 0.04) scales. This is also in agreement with DKA18, who found that LBs avoid non-LBs up to ∼8 Mpc, which means that the ages of the LB nearest galaxies are younger compared to the general population of galaxies.

Table 4. Comparison of VYGs to control galaxies, splitting by mass.

	$\log(M_\star/\mathrm{M}_\odot) \leq 9.5$		$\log(M_\star/\mathrm{M}_\odot) > 9.5$	
Property	VYG median	KS p-value	VYG median	KS p-value
(1)	(2)	(3)	(4)	(5)
sSFR	higher	$<10^{-10}$	higher	2×10^{-3}
$(g-i)_{\mathrm{Petro}}$	**lower**	1×10^{-9}	**same**	0.10
BPT diagram	higher	2×10^{-4}	higher	6×10^{-8}
log(O/H)	**lower**	0.06	**same**	0.86
f_{gas}	higher	2×10^{-4}	–	–
A_V^{Balmer}	higher	6×10^{-6}	higher	$<10^{-10}$
A_V^{cont}	**lower**	0.01	**higher**	1×10^{-7}
C_1	higher	3×10^{-3}	higher	3×10^{-5}
μ_{50}	higher	1×10^{-6}	higher	$<10^{-10}$
A_1	higher	4×10^{-5}	higher	3×10^{-3}
S_1	**higher**	0.08	**same**	0.05
$d_{1\mathrm{st}}$/kpc	lower	0.06	lower	0.02
$d_{1\mathrm{st}}/\theta_{50}$	lower	0.05	lower	0.01

Notes. For each property, columns 2 and 4 indicate how the median VYG value compares with that of the CSGs, while columns 3 and 5 show the respective p-values for that property in the corresponding mass bin. Changes in the trend between low and high masses are indicated in bold.

6.4 VYGs and CSGs in the low- and high-mass regimes

Our method of comparing the residuals of the VYG and CSG properties from the CSG properties vs. mass relations may hide how differences between VYGs and CSGs vary from low to high stellar masses. To address this question, we repeated our analysis, separating the galaxies into two mass bins, above and below $\log(M_\star/\mathrm{M}_\odot) = 9.5$, which is close to the median of the VYG sample. Table 4 shows the results.

For several properties, we see similar trends for the VYGs compared to the CSGs in the low- and high-stellar mass regimes. On the other hand, differences between the VYG and CSG colours and asymmetries are seen only for low-mass galaxies. Furthermore, as already mentioned in Section 6.2, low-mass VYGs appear to be more metal poor than CSGs with similar masses, and this is not observed for the high-mass VYGs. According to our merger diagnostics, merger events appear to be an important mechanism for VYGs of all masses, but the low gas metallicity in less massive VYGs suggests that both mergers and gas infall from the cosmic web are operating in the low-mass regime.[10] This qualitatively agrees with our understanding that only high-mass galaxies typically grow by mergers, which are too infrequent in low-mass ones, which thus must usually grow by gas infall (Cattaneo et al. 2011). Following this argument, one would not expect that low-mass VYGs also have more neighbours, in contrast with the marginal indication of more neighbours around low-mass VYGs compared to low-mass CSGs shown in Table 4. Besides, all of the galaxies in our VYG and CSG samples have lower mass than the critical $z = 0$ mass of $M_\star = 2 \times 10^{11}\,\mathrm{M}_\odot$ that Cattaneo et al. found for the growth by mergers (see Fig. 1). Nevertheless, one can reason that VYGs recently formed their stars by mergers, but will subsequently live a more quiet period of gas infall.

From low- to high-mass galaxies, we see a reversal of the trend of the VYG A_V^{cont} with respect to that of the CSGs; i.e. low-mass VYGs have lower A_V^{cont} values than the CSG values, while for high-mass VYGs we see the opposite. These high A_V^{cont} values might be responsible for the lack of bluer colours among high-mass VYGs. The behaviour of A_V^{cont} with galaxy mass, and the larger scatter that we observe in the A_V^{cont}–$M_\star$ and A_V^{cont}–SFR relations compared to the $\tau_{\alpha\beta}$ relations suggest that the amount of extinction of the stellar continuum might be determined by different – and maybe more – factors than the amount of extinction of the gas emission. We will further discuss the dust content in Section 6.6.1.

[10]However, in a recent study, Tiwari, Mahajan & Singh (2020) find that galaxies in filaments are enriched relative to field galaxies.

6.5 How can a galaxy retain so much gas for so long?

As shown in Fig. 10, among galaxies whose atomic gas is detected with ALFALFA, VYGs have very high atomic gas mass fractions, falling in the range of 0.8–1, the same as found by Thuan et al. (2016) for extremely metal deficient blue compact dwarf starburst galaxies. It is not clear if the ALFALFA sample is representative of the whole VYG population, but it seems that at least the low-mass VYGs with H I detections have the amount of gas necessary to fuel the intense SF activity required to form so many stars in such a small time-scale. *However, how can a galaxy retain this gas until recent times?*

A possible scenario comes from a recent study by Zhang et al. (2019), who, using data from SDSS, ALFALFA, GASS (Catinella et al. 2010), and COLD GASS (Saintonge et al. 2011), found that massive quiescent central disc galaxies contain as much atomic gas as corresponding star-forming galaxies ($M_{\mathrm{H\,I}}/M_\star > 0.1$). Moreover, both galaxy classes have identical H I spectra indicating regularly rotating H I discs with a radius of ∼30 kpc and little kinematic perturbations. Following Renzini (2020), the H I gas in quenched discs can be stored in an outer ring with high angular momentum and large infall time-scales. These galaxies are quenched because of the reduced molecular gas content and lower SF efficiency. However, perturbations could drive the outer atomic gas inwards and trigger SF.

This scenario could be extended to the lower masses of our VYGs. However, Zhang et al. only considered central galaxies, and only one out of the seventeen VYGs that are centrals in low-mass groups ($12.3 \leq \log[M_{\mathrm{halo}}/\mathrm{M}_\odot] < 13.5$) has H I detection in the ALFALFA survey. All the other central VYGs are beyond the $z = 0.05$ redshift limit of ALFALFA. We thus cannot determine if our central VYGs also have high fractions of atomic gas.

To investigate the gas fraction of low-mass central galaxies, we computed f_{gas} for a general sample of SDSS galaxies that have a counterpart in the ALFALFA catalogue. To avoid measurements of H I masses that may be contaminated by the neutral gas of nearby galaxies because of the large Arecibo beam size, we excluded galaxies that have companions within a radius of 1.8 arcmin and $\Delta z \leq 2\,W_{20}/c$; i.e. we adopt the same approach as described in Section 4.2. Fig. 20(a) shows f_{gas} versus stellar mass of centrals and satellite galaxies in groups with $12.3 \leq \log[M_{\mathrm{halo}}/\mathrm{M}_\odot] < 13.5$, the same halo mass range of the central VYG haloes. We fitted the satellite f_{gas} versus $\log M_\star$ relation assuming that f_{gas} can be described as

$$f_{\mathrm{gas}} = \frac{1}{1 + \exp\left\{\sum_{i=0}^{4} a_i \left[\log(M_\star/\mathrm{M}_\odot)\right]^i\right\}}.$$

Fig. 20(b) shows the distributions of the residuals from the best-fitting relation for the centrals and satellite galaxies with $M_\star < 10^{11}\,\mathrm{M}_\odot$, which is the mass range of our VYG sample. Central galaxies tend to have higher fractions of atomic gas than satellites. This suggests that the higher gas fractions of VYGs may be related to their excess of centrals (Fig. 18).

Instead of retaining their gas, some VYGs may be experiencing gas infall from their surroundings. It could be pristine gas from the

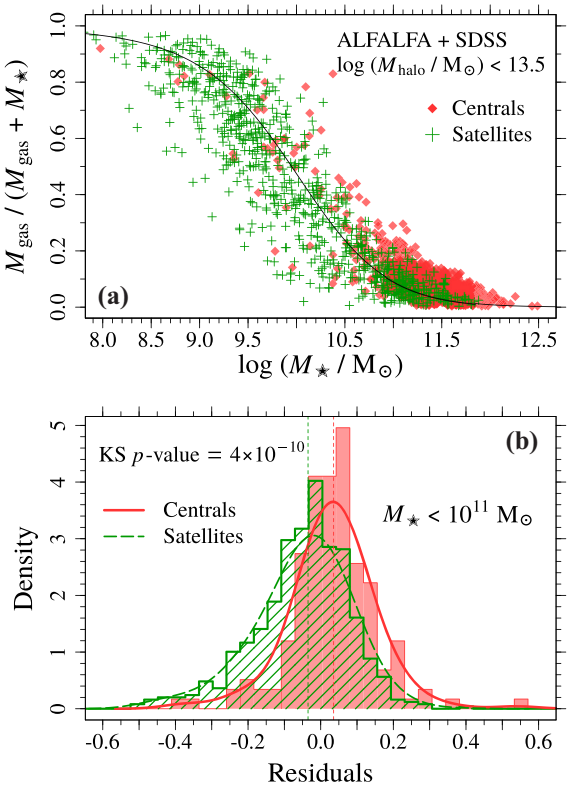

Figure 20. *Top:* Fraction of atomic gas in SDSS galaxies with H I detections in the ALFALFA survey as a function of galaxy stellar mass. The *red* and *green* symbols indicate the central and satellite galaxies in low-mass groups (log [M_{halo}/M$_\odot$] ≤ 13.5). The *black solid line* is the best fit to the satellite f_{gas} versus log $M_\star$ relation (see the text for details). *Bottom:* The residuals from the best fit indicated in panel **a** for galaxies with $M_\star < 10^{11}$ M$_\odot$. The *red* and *green histograms* correspond to low-mass centrals and satellites, respectively.

cosmic web, but, as already mentioned in Section 6.2, this only affects low-mass VYGs, at best, since high-mass VYGs have similar metallicities as CSGs of similar mass, while low-mass VYGs have a marginal indication of lower metallicity. Since a large fraction of the VYGs have companions, the VYG gas can be replenished by the gas from the merging neighbour galaxies. This is more likely than coming from the cosmic web of gas, since the gas brought in by the neighbouring galaxies is already somewhat enriched.

Although we find that VYGs appear to have more atomic gas, differences in the fraction of atomic (and molecular) gas alone cannot account for the observed distribution of objects above the main sequence of star-forming galaxies. Indeed, using data from the ALFALFA, GASS, and COLD GASS surveys to quantify how the mean atomic and molecular gas mass fractions vary in the SFR–$M_\star$ plane, Saintonge et al. (2016) showed that galaxies with very high sSFR, such as those that are necessary to produce a VYG, not only have large gas fractions, but also have high SF efficiencies. It is known that mergers and interactions might enhance the efficiency in which the gas is converted into stars. We explore this scenario in the following section.

6.6 Are VYGs the products of mergers?

Many studies have shown that mergers and interactions induce an intense SF activity in galaxies (Joseph & Wright 1985), with a recent exception (Pearson et al. 2019). Using hydrodynamical simulations, Di Matteo et al. (2008) found that the SF activity of starbursts triggered by mergers is enhanced by a factor of 3, on average, compared to starbursts in normal galaxies, and is greater than 5 in about 15 per cent of major galaxy interactions and mergers. Moreover, multiple bursts can occur, with the bulk of SF among merging galaxies occurring at the second passage, even though there is continuous SF, as well as a first burst at the first pericentre.

Our results suggest that mergers and interactions are important processes triggering recent SF activity. Compared to the control sample, the VYGs are twice as likely to have close companions at distances < 5 θ_{50} (Fig. 14). Moreover, the fraction of VYGs showing signs of interactions is double that of CSGs. Using the parameters p_{merger} from the DS18 and GZ1 catalogues, we have shown that the fraction of VYGs classified as mergers is always twice the corresponding fraction for CSGs, regardless of the threshold adopted for p_{merger} (Fig. 15). Our visual inspection of the SDSS images also shows that the fraction of VYGs showing signs of interaction is twice that of CSGs (Fig. 16 and Table 1), and this result does not depend on the galaxy surface brightness (Fig. 17).

The preference of VYGs to lie in the inner parts of low-mass groups (Fig. 18 and Table 2) also supports that scenario, since interactions and mergers are more likely to occur in this environment, where encounter velocities are slower; hence, the encounters are more efficient (e.g. Mamon 1992, 2000).

The VYG morphologies also point to the merging nature of a significant fraction of these systems, and may be related to different phases of mergers. VYGs are more clumpy and asymmetric than CSGs (Fig. 8), which may be signs of ongoing interactions with neighbour galaxies. In addition, VYGs have higher concentrations than CSGs (Fig. 7), indicating central SF activity. Hydrodynamical simulations show that global torques induced by mergers drive gas inwards, leading to more compact systems (Hopkins et al. 2010, 2013).

These compact galaxies are also seen as an excess of VYGs with low T type values compared to those of the CSGs (Figs 5 and 6). These early-type VYGs are bluer than CSGs with $T \leq 0$, and through a visual inspection of their images, we see that these systems are blue compact spheroids. These very compact VYGs could represent recent post-merger galaxies. Indeed, post-merger galaxies have 3 times higher gas fractions than control galaxies (Ellison, Catinella & Cortese 2018). However, high gas fractions – as in our VYGs – reduce the loss of the gas angular momentum and increase the chances of disc survival (Hopkins et al. 2013). Therefore, the lack of discs in these galaxies suggests that the progenitor galaxies already had low angular momentum, such as irregular galaxies, or the orbital angular momentum of the merging system is low (i.e. head-on merging encounters). Otherwise, other mechanisms should be invoked to account for the formation of these blue compact VYGs.

6.6.1 *Internal extinction, dust mass, and relation with mergers*

We found that VYGs have more internal extinction than CSGs, and this may be related to a merging process. As discussed in Section 4, at a fixed stellar mass, VYGs have higher $\tau_{\alpha\beta}$ (which probes the extinction in H II regions) and A_V^{cont} (which is more related to dust in the diffuse ISM component) values compared to those of the CSGs (Fig. 12). While A_V^{cont} is inferred from the stellar continuum through SPS analysis, $\tau_{\alpha\beta}$ is obtained from H α and H β emission lines. Since both parameters indicate that VYGs have higher internal extinction than CSGs, it is unlikely that this result is a consequence of biases

and uncertainties in the SPS analysis or in the measurements of the emission-line fluxes. Therefore, our results might be truly related to a higher amount of dust in VYGs compared to the general population of star-forming galaxies.

The high VYG $\tau_{\alpha\beta}$ values appear to be related to their high SFRs (Figs 13a and b), since VYGs and CSGs follow a similar $\tau_{\alpha\beta}$–log SFR relation. On the other hand, differences between the VYG and CSG $A_V^{\rm cont}$ do not appear to be fully accounted by differences in the galaxy sizes, morphology, inclination, or SFRs. Differences between VYG and CSG gas metallicities cannot explain the higher amount of dust in VYGs either, since VYGs and CSGs have similar oxygen abundances (Fig. 9). Other factors can also contribute to higher internal extinction in VYGs, such as the H I gas mass (Casasola et al. 2020). However, the VYGs have higher $A_V^{\rm cont}$ and $\tau_{\alpha\beta}$ values compared to those of the CSGs at a fixed $f_{\rm gas}$. *So, why do VYGs have more dust in their diffuse ISM component than CSGs?* Ultra-luminous infrared galaxies, whose very strong infrared emission is caused by dust, are observed to have one (Armus, Heckman & Miley 1987) or several (Borne et al. 2000) companions, implying that mergers triggered the extreme starburst leading to extreme amounts of dust.

Does the connection between SFRs, dust, and mergers extend to low-mass galaxies? To address this question, we analysed a sample of SDSS galaxies with very high SFRs and used the parameter $p_{\rm merg}$ from the catalogue of morphology by DS18 (see Section 5.2) to access the probability of these systems of being mergers. A significant fraction of VYGs in our sample have $\log[{\rm SFR}/({\rm M}_\odot\,{\rm yr}^{-1})] \geq 0.8$ (14.5 per cent, i.e. 30 VYGs), and these galaxies are the ones associated with high internal extinction (see Fig. 13a). However, in a general sample of ∼323 087 SDSS galaxies at $z \leq 0.1$, only 4107 (∼1.3 per cent) have these extreme SFR values, and they are typically more massive than 10^{10} M$_\odot$. We analysed the internal extinction and merger probabilities of these 4107 SDSS galaxies with very high SFRs by comparing them to a sample of galaxies with similar stellar masses and $\log[{\rm SFR}/({\rm M}_\odot\,{\rm yr}^{-1})] < 0.8$.

As expected, systems with high SFRs have larger $\tau_{\alpha\beta}$ and $A_V^{\rm cont}$ values compared to low-SFR galaxies with similar masses. In addition, we find that the distribution of $p_{\rm merg}$ is shifted towards higher values compared to those of the low-SFR galaxies, and the fraction of high-SFR galaxies with $p_{\rm merg} > 0.5$ is almost two times larger than that of low-SFR objects (26 per cent and 15 per cent, respectively). The same results are observed when we consider only galaxies with $\log(M_\star/{\rm M}_\odot) < 10.5$; i.e. at lower stellar masses, galaxies with high SFRs also have more internal extinction and higher probabilities of being merging systems than low-SFR galaxies. Besides, in the low-mass regime, the differences between $p_{\rm merg}$ of high- and low-SFR galaxies are more pronounced when compared to the results obtained for all galaxies. The fraction of low-SFR systems with $p_{\rm merg} > 0.5$ remains unchanged (15 per cent), but this fraction increases to 42 per cent among the high-SFR low-mass galaxies. These results suggest that the relation between SFR, dust, and merger events can be, at least at some level, extended to the low-stellar mass regime, and the high SFRs and internal extinction that we find for our VYGs can be an indication of mergers and interactions.

The internal extinction $\tau_{\alpha\beta}$ inferred from emission lines appears to be more related to the SFR and merger events than $A_V^{\rm cont}$. Once we take the $\tau_{\alpha\beta}$–SFR relation into account, we see no differences between the VYGs and CSGs. On the other hand, differences are still seen for $A_V^{\rm cont}$. Looking at Table 4, we see that the trend of VYG $A_V^{\rm cont}$ values compared to those of CSGs is reverted from low to high mass. Besides, the scatter in the $A_V^{\rm cont}$–$M_\star$ and $A_V^{\rm cont}$–SFR relations is larger compared to the relations for $\tau_{\alpha\beta}$. All these results suggest that dust production in the diffuse ISM is a complex process with several variables. While the properties of the molecular clouds in the star-forming regions of the galaxy appear to be dominant for $\tau_{\alpha\beta}$, $A_V^{\rm cont}$, tracing a more diffuse dust component, should depend on other galaxy properties, such as galaxy size, inclination, and the balance between the production and destruction of dust grains in the ISM. In other words, while the small-scale physics within molecular clouds appears to be roughly independent of the galaxy mass and properties, the processes regulating the production and destruction of dust in the ISM might be more dependent on the galaxy integrated properties.

6.6.2 *Predictions from models of galaxy formation*

In Paper I, we used analytical models of galaxy formation, which were applied to high-mass resolution Monte Carlo halo merger trees, to predict the fraction of VYGs in the local Universe. Four models were considered: one physically motivated model by Cattaneo et al. (2011), including a more realistic cut-off at low halo masses due to reionization feedback (Gnedin 2000; C+G); two empirical models by Moster, Naab & White (2013; MNW) and Behroozi, Wechsler & Conroy (2013; BWC), which use abundance matching to link the stellar-mass and halo-mass functions; and an empirical model by Mutch, Croton & Poole (2013; MCP), where the stellar mass growth rate is proportional to the halo mass growth rate times a function of halo mass and redshift. The models were run with either bursty or quiet halo merging schemes.

The effects of the bursty and quiet schemes on the predictions of VYG fractions depend on the model used. The merging scheme has virtually no effect on the VYG fractions versus stellar mass in the MNW and MCP models (fig. 10 of Tweed et al. 2018). On the other hand, the VYG fractions in the quiet merging scheme are reduced by over one order of magnitude in the C+G model at intermediate mass and in the BWC model at low mass. Tweed et al. (2018) also showed that, with the bursty galaxy merging, VYGs are associated with a recent major halo merger, in contrast to the quiet merging scheme.

In Paper II, we compared the predictions from these models to the observed fractions of VYGs. We found that the observations support the bursty scheme for C+G and BWC models, but little can be said about the bursty and quiet MNW and MCP models due to the small differences between these two schemes for these two models.

Since these models do not provide any prediction on the properties of the VYGs, no direct comparison between our results and these models is possible. However, we do present new constraints in favour of the bursty merging scheme, as our results indicate that, in comparison to CSGs, VYGs are more likely to be associated with gas-rich mergers.

6.6.3 *Are most or all VYGs associated with mergers?*

While we find many indications that mergers and interactions are responsible for triggering an intense SF activity in VYGs, it is not clear if these processes account for most or all of the VYGs. Our visual inspection of the VYG images revealed that at least half of our galaxies do not show any signs of interactions. However, the SDSS images are not deep enough, and we could be missing faint tidal features. So, future studies using deeper imaging data could help understanding the nature of those VYGs for which no signs of interactions are observed. In addition, since mergers leave signatures in the gas and stellar kinematics, spatially resolved spectroscopic observations of VYGs might also help constraining this scenario, in particular to identify the isolated VYGs that are in a post-merger phase.

6.7 Caveats

6.7.1 Colour gradients and outer halo of old stellar populations

As described in Section 2.2, to minimize aperture effects due to the limited size of the SDSS fibre, we required that the total *g–i* VYG colour be bluer than the corresponding fibre colour. We showed in Paper II that the requirement of a blue gradient does not rule out the possibility of a hidden old stellar population in the VYG, but at least it mitigates the aperture effect. However, one must keep in mind the following: (i) a blue gradient in a galaxy can also be a consequence of a negative metallicity gradient due to the age–metallicity degeneracy and (ii) this selection criterion can produce biased samples, as we discuss below.

It is known that different galaxy formation scenarios and physical mechanisms leave typical imprints on how the stellar population properties vary with the distance to the galaxy centre (Kobayashi 2004; Hirschmann et al. 2015; Ferreras et al. 2019). Therefore, when imposing a blue gradient to select our galaxies, we favour physical mechanisms and formation scenarios that leave negative age and/or metallicity gradients. Several studies (e.g. Di Matteo et al. 2009; La Barbera et al. 2010, 2012; Bernardi et al. 2019; Zhuang et al. 2019; Zibetti et al. 2020) have shown that connecting observed galaxy gradients to the physical processes that produce them is not a straightforward task, so it is difficult to determine how this selection criterion affects our results and conclusions. Although the CSGs were selected to have negative (blue) colour gradients as for the VYGs, we cannot claim that our results and conclusions can be extended to VYGs that have positive (red) colour gradients.

6.7.2 Uncertainties in the spectral analysis

Young stellar populations are much more luminous than old ones. So, even within the fibre, the young population can outshine the old stars, making it hard to be detected through SPS analysis of the SDSS spectra and possibly leading to the overestimation of the fraction of mass in populations younger than 1 Gyr, $f_{M_\star,\mathrm{y}}$.

To investigate the uncertainties in the determination of $f_{M_\star,\mathrm{y}}$ using the STARLIGHT code, we used the V15 spectral models to create a set of simulated spectra assuming two bursts of SF, one occurring within the last 1 Gyr and the other 12 Gyr ago. The age of the young population was randomly selected among 21 ages between 0.03 and 1 Gyr, while the old population has a fixed age of 12 Gyr. We used as input $f^{\mathrm{in}}_{M_\star,\mathrm{y}} = 0.1, 0.2, 0.3$, and 0.4, and created 1000 simulated spectra for each $f^{\mathrm{in}}_{M_\star,\mathrm{y}}$. The internal extinction and S/N of the simulated spectra were selected from the distributions of A_V^{cont} and S/N of the VYG sample.

We ran STARLIGHT with the same configuration used in the SDSS spectral analysis of Paper II, and computed the fraction of runs with output $f^{\mathrm{out}}_{M_\star,\mathrm{y}} \geq 0.5$, i.e. the fraction of simulated galaxies that would be wrongly classified as VYGs. We show the results in Table 5, where we also present the median of the $f^{\mathrm{out}}_{M_\star,\mathrm{y}}$ values obtained.

We find a good agreement between the median estimated fraction of young mass and the input fraction. The fraction of incorrect VYG classifications rises very rapidly with the fraction of young stars: from rare with 10 per cent of young mass in the input spectra to 44 per cent with 40 per cent of young mass in the spectrum. This indicates that the reliability of the VYG classification is not too low.

It is difficult to estimate the number of false VYGs in our sample, because we do not know the true distribution of $f_{M_\star,\mathrm{y}}$ or the age and duration of the past and recent bursts of SF. Besides, our two-burst simulations are very simplistic. On the other hand, they represent a

Table 5. False positive VYGs from simulated spectra.

$f^{\mathrm{in}}_{M_\star,\mathrm{y}}$	Median $f^{\mathrm{out}}_{M_\star,\mathrm{y}}$	Fraction of runs with $f^{\mathrm{out}}_{M_\star,\mathrm{y}} \geq 0.5$
(1)	(2)	(3)
0.10	$0.10^{+0.04}_{-0.07}$	0.1%
0.20	$0.19^{+0.11}_{-0.12}$	6.9%
0.30	$0.31^{+0.38}_{-0.18}$	20.7%
0.40	$0.46^{+0.48}_{-0.29}$	44.0%

Notes. Columns are as follows: (1) input mass fraction of the young stellar population (age < 1 Gyr); 1000 simulated spectra were created for each $f^{\mathrm{in}}_{M_\star,\mathrm{y}}$ value; (2) median of the output $f^{\mathrm{out}}_{M_\star,\mathrm{y}}$ values; the upper and lower errors were obtained from the 84th and 16th percentiles of the $f^{\mathrm{out}}_{M_\star,\mathrm{y}}$ distributions; and (3) fraction of runs with $f^{\mathrm{out}}_{M_\star,\mathrm{y}} \geq 0.5$.

'worst' possible scenario of a very old and low-luminosity population (12 Gyr) mixed with a very recent burst (<1 Gyr). Many studies show that the SFHs of real low-mass galaxies are more extended (De Lucia et al. 2006; Thomas et al. 2010; Trevisan et al. 2012), and their old (> 1 Gyr) stellar component would be brighter and thus easier to detect than the population in our simulations, which was formed in single burst 12 Gyr ago. Nevertheless, we argued in Paper II that a conservative lower limit on the reliability of the VYG classification combined with the observed cosmic SFH implies that present-day galaxies with $M_\star > 10^8\,\mathrm{M}_\odot$ experienced, on average, at most four major starbursts over cosmic time.

In summary, our simulations show that it is possible that the fraction of mass in populations younger than 1 Gyr in some of our VYGs is, in fact, less than 50 per cent. However, our results suggest that, even if $f_{M_\star,\mathrm{y}}$ is less than 50 per cent, the young population still accounts for a significant fraction of the total stellar mass of the galaxy ($f_{M_\star,\mathrm{y}} \gtrsim 30$ per cent). In addition, the results from three different spectral models have been combined to increase the reliability of our VYG sample.

Finally, one may think that the SPS analysis is not the proper method to derive the fraction of mass of young populations in galaxies, due to the large uncertainties associated with this method, in particular for galaxies with few absorption features in their spectra (such as galaxies with a very recent starburst). However, the differences in the properties of VYGs relative to our control sample confirm that these systems are different, i.e. younger. These differences cannot be attributed to the presence of starburst galaxies with large EWs of Hα in the VYG sample, since our conclusions remain unchanged when we exclude these galaxies from the VYG sample (see Table 3). Besides, we found differences in properties that were derived through methods that are independent of the SPS analysis, such as morphology, environment, the $\tau_{\alpha\beta}$ measure of interstellar dust, and the gas mass fraction.

6.7.3 Imperfect control sample

As discussed in Section 2 and shown in Fig. 2(a), the $\log M_\star$ versus z distribution of the CSGs is slightly different from that of the VYGs. To investigate if it affects our results, we repeated all the analysis presented in this work after excluding the galaxies that are more than 2σ away from the CSG $\log M_\star$ versus z relation (29 VYGs and 53 CSGs). We used a second-order polynomial to fit the relation. All the results presented in Table 3 remain statistically significant with p-values below $<5 \times 10^{-3}$, except for the parameters $(g-i)_{\mathrm{Petro}}$, S_1, and $d_{1\mathrm{st}}$/kpc, for which KS tests indicate marginal statistical significance

with p-values of 0.04, 0.01, and 0.03, respectively. Therefore, we conclude that the outliers in the $\log M_\star$ versus z relation do not affect our results and conclusions.

6.7.4 *Nebular continuum emission*

The SPS fitting of the spectra with STARLIGHT and VESPA both ignored the nebular contribution to the continuum, which becomes important in strongly star-forming galaxies. The ratio between the ionized gas and the continuum emissions increases with wavelength, ranging from ∼20 per cent at ∼4000 Å to more than 40 per cent around 7000 Å in galaxies with very high sSFRs (Izotov, Guseva & Thuan 2011). As discussed in Paper II, the neglect of this red nebular component causes the SPS analysis to overestimate the ages of starburst galaxies. This was confirmed by Cardoso, Gomes & Papaderos (2019), who found that the FADO SPS code, which includes the nebular continuum emission in the fit, obtains lower ages than does STARLIGHT. Therefore, including the nebular continuum emission in our analysis would lead to even higher fractions of stellar mass formed within the last 1 Gyr, and would not affect the results and conclusions presented in this work.

7 SUMMARY AND CONCLUSIONS

In this work, we investigated the properties and environments of VYGs. The SFHs of galaxies were inferred from SDSS spectra through SPS analysis, and we selected as VYGs those galaxies with over half their stellar mass formed within the last 1 Gyr according to each of three different stellar population models. We built a control sample of normal galaxies with similar stellar masses, redshifts, radii, and colour gradients, and the comparison between these two samples revealed that VYGs have different properties and reside in different environments. In Section 6.1, we summarize all the VYGs properties that we investigated and discuss how they compare to those of CSGs. From our results, we conclude the following:

(i) *VYGs and other young systems like I Zw 18 are likely to be different in nature*, as discussed in Section 6.5. I Zw 18 analogues are low-mass very metal poor systems that deviate from the mass–metallicity relation of normal galaxies, while VYGs follow it.

(ii) *SF in the more massive VYGs is not being fed by infalling metal-poor gas from the cosmic web*, since these galaxies have similar gas oxygen abundances as the control galaxies. *However, gas infall and mergers may both be important triggers of SF in the VYGs at the low-mass end*, given hints of their lower metallicities and their excess of near-neighbours.

(iii) *Gas-rich mergers and interactions are important mechanisms for producing very young systems in the local Universe* (Section 6.6). Differences in the morphology (VYGs are more clumpy, more asymmetric, and a larger fraction show tidal features and other signs of interactions compared to the CSGs) and the environment (VYGs are more likely to have companions and to be found in the inner parts of low-mass groups) of the VYGs support this scenario, in particular for the massive VYGs. These results agree with the predictions by Tweed et al. (2018, Paper I) for their subset of analytical models of galaxy formation where halo mergers are associated with starbursts.

Since mergers leave typical signatures in the gas and stellar kinematics of the galaxy, future studies of VYGs using spatially resolved spectroscopic data can help constraining the merger scenario. Moreover, deep imaging data might help understanding the nature of the VYGs for which there is no evident sign of interactions in the SDSS images. We have also obtained H I radio observations using the Giant Metrewave Radio Telescope (GMRT) and the Very Large Array (VLA) to investigate the gas distribution and kinematics in and around VYGs.

Finally, in this study, we showed that the VYGs differ from normal galaxies in various ways, but *did their progenitors have peculiar properties before becoming VYGs? Or can any galaxy become a VYG?* In a future study, we will address the nature of the VYG progenitors by identifying these systems in hydrodynamical simulations and tracing back the assembly history of these systems.

We provide an electronic table with observed and derived properties of our VYGs and CSGs. The description of the table columns is given in Appendix C in the supplementary material.

ACKNOWLEDGEMENTS

The authors thank the referee Daniel Kelson for his excellent comments and suggestions, which led to an improved version of the manuscript. We also thank Mojtaba Raouf for useful suggestions. MT thanks the support of CNPq (process #307675/2018-1), the Institut Lagrange de Paris, and the program L'Oréal UNESCO ABC *Para Mulheres na Ciência*. MT and TXT are grateful for the hospitality of the Institut d'Astrophysique de Paris where part of this work was carried out. GAM acknowledges the Brazilian CNPq (grant #451451/2019-8) awarded to MT and thanks the Universidade Federal do Rio Grande do Sul for its hospitality. LSP acknowledges support of the programme of the NAS of Ukraine for the development of priority fields of scientific research (CPCEL 6541230). AR acknowledges support from the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (NRF-2019R1C1C1010279). We acknowledge Roberto Cid Fernandes for making his STARLIGHT code public and Rita Tojeiro for making the VESPA output publicly available. We acknowledge the use of the SDSS data (http://www.sdss.org/collaboration/credits.html), TOPCAT Table/VOTable Processing Software (Taylor 2005, http://www.star.bris.ac.uk/mbt/topcat/), and R language and environment for statistical computing (R Core Team 2015). 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.

DATA AVAILABILITY

The data underlying this article are available in the article and in its online supplementary material.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Abraham R. G., Valdes F., Yee H. K. C., van den Bergh S., 1994, ApJ, 432, 75
Abraham R. G., van den Bergh S., Glazebrook K., Ellis R. S., Santiago B. X., Surma P., Griffiths R. E., 1996, ApJS, 107, 1
Alloin D., Collin-Souffrin S., Joly M., Vigroux L., 1979, A&A, 78, 200
Aloisi A. et al., 2007, ApJ, 667, L151
Armus L., Heckman T., Miley G., 1987, AJ, 94, 831
Asari N. V., Cid Fernandes R., Stasińska G., Torres-Papaqui J. P., Mateus A., Sodré L., Schoenell W., Gomes J. M., 2007, MNRAS, 381, 263
Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Balogh M. L. et al., 2009, MNRAS, 398, 754
Barnard G. A., 1945, Nature, 156, 783

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57
Belfiore F. et al., 2015, MNRAS, 449, 867
Belfiore F. et al., 2017, MNRAS, 469, 151
Bernardi M., Domínguez Sánchez H., Brownstein J. R., Drory N., Sheth R. K., 2019, MNRAS, 489, 5633
Bishop C. M., 2006, Pattern Recognition and Machine Learning. Springer, Berlin. Available at: http://research.microsoft.com/en-us/um/people/cmbishop/prml/
Blanton M. R. et al., 2003, AJ, 125, 2348
Blanton M. R. et al., 2005, AJ, 129, 2562
Borne K. D., Bushouse H., Lucas R. A., Colina L., 2000, ApJ, 529, L77
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 (BC03)
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Cardoso L. S. M., Gomes J. M., Papaderos P., 2019, A&A, 622, A56
Casasola V. et al., 2020, A&A, 633, A100
Catinella B. et al., 2010, MNRAS, 403, 683
Cattaneo A., Mamon G. A., Warnick K., Knebe A., 2011, A&A, 533, A5
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Cid Fernandes R., Mateus A., Sodré L., Stasińska G., Gomes J. M., 2005, MNRAS, 358, 363
Conselice C. J., Bershady M. A., Jangren A., 2000, ApJ, 529, 886
Contreras Ramos R. et al., 2011, ApJ, 739, 74
da Cunha E., Eminian C., Charlot S., Blaizot J., 2010, MNRAS, 403, 1894
Darg D. W. et al., 2010, MNRAS, 401, 1043
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
de Souza R. S. et al., 2016, MNRAS, 461, 2115
de Vaucouleurs G., 1963, ApJS, 8, 31
Di Matteo P., Bournaud F., Martig M., Combes F., Melchior A.-L., Semelin B., 2008, A&A, 492, 31
Di Matteo P., Pipino A., Lehnert M. D., Combes F., Semelin B., 2009, A&A, 499, 427
Dinerstein H. L., 1990, in Thronson H. A., Jr, Shull J. M., eds, Astrophysics and Space Science Library, Vol. 161, The Interstellar Medium in Galaxies. Springer-Verlag, Berlin, p. 257
Domínguez A. et al., 2013, ApJ, 763, 145
Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, MNRAS, 476, 3661 (DS18)
Dressler A., Kelson D. D., Abramson L. E., 2018, ApJ, 869, 152 (DKA18)
Ellison S. L., Catinella B., Cortese L., 2018, MNRAS, 478, 3447
Engelbracht C. W., Rieke G. H., Gordon K. D., Smith J. D. T., Werner M. W., Moustakas J., Willmer C. N. A., Vanzi L., 2008, ApJ, 678, 804
Erguler K., 2016, Barnard: Barnard's Unconditional Test. Available at: https://CRAN.R-project.org/package=Barnard
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
Ferrari F., de Carvalho R. R., Trevisan M., 2015, ApJ, 814, 55
Ferreras I. et al., 2019, MNRAS, 489, 608
Fisher R. A., 1935, J. R. Stat. Soc., 98, 39
Florido E., Zurita A., Pérez I., Pérez-Montero E., Coelho P. R. T., Gadotti D. A., 2015, A&A, 584, A88
Girardi L., Bressan A., Chiosi C., Bertelli G., Nasi E., 1996, A&AS, 117, 113
Gnedin N. Y., 2000, ApJ, 542, 535
Hamilton A. J. S., Tegmark M., 2004, MNRAS, 349, 115
Hamming R. W., 1998, Digital Filters, 3rd edn. Dover, Mineola, NY
Haynes M. P. et al., 2011, AJ, 142, 170
Haynes M. P. et al., 2018, ApJ, 861, 49
Hirschmann M., Naab T., Ostriker J. P., Forbes D. A., Duc P.-A., Davé R., Oser L., Karabal E., 2015, MNRAS, 449, 528
Hjorth J., Gall C., Michałowski M. J., 2014, ApJ, 782, L23
Ho D. E., Imai K., King G., Stuart E. A., 2011, J. Stat. Softw., 42, 1
Hopkins P. F. et al., 2010, ApJ, 715, 202
Hopkins P. F., Cox T. J., Hernquist L., Narayanan D., Hayward C. C., Murray N., 2013, MNRAS, 430, 1901
Izotov Y. I., Thuan T. X., 1998, ApJ, 497, 227
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., Guseva N. G., Fricke K. J., Henkel C., 2014, A&A, 561, A33
Izotov Y. I., Thuan T. X., Guseva N. G., Liss S. E., 2018, MNRAS, 473, 1956
Izotov Y. I., Thuan T. X., Guseva N. G., 2019, MNRAS, 483, 5491
Joseph R. D., Wright G. S., 1985, MNRAS, 214, 87
Kauffmann G. et al., 2003, MNRAS, 341, 33
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121
Knobel C., Lilly S. J., Woo J., Kovač K., 2015, ApJ, 800, 24
Kobayashi C., 2004, MNRAS, 347, 740
Komatsu E. et al., 2011, ApJS, 192, 18
Kreckel K. et al., 2013, ApJ, 771, 62
Kroupa P., 2001, MNRAS, 322, 231
La Barbera F., De Carvalho R. R., De La Rosa I. G., Gal R. R., Swindle R., Lopes P. A. A., 2010, AJ, 140, 1528
La Barbera F., Ferreras I., de Carvalho R. R., Bruzual G., Charlot S., Pasquali A., Merlin E., 2012, MNRAS, 426, 2300
Le Borgne J.-F. et al., 2003, A&A, 402, 433
Lintott C. et al., 2011, MNRAS, 410, 166
Lotz J. M., Primack J., Madau P., 2004, AJ, 128, 163
McGee S. L., Balogh M. L., Wilman D. J., Bower R. G., Mulchaey J. S., Parker L. C., Oemler A., 2011, MNRAS, 413, 996
Madau P., Dickinson M., 2014, ARA&A, 52, 415
Mahajan S., Mamon G. A., Raychaudhury S., 2011, MNRAS, 416, 2882
Mahalanobis P. C., 1936, Proc. Natl. Inst. Sci. (Calcutta), 2, 49
Mamon G. A., 1992, ApJ, 401, L3
Mamon G. A., 2000, in Combes F., Mamon G. A., Charmandaris V., eds, ASP Conf. Ser. Vol. 197, Dynamics of Galaxies: From the Early Universe to the Present. Astron. Soc. Pac., San Francisco, p. 377
Mamon G. A., Trevisan M., Thuan T. X., Gallazzi A., Davé R., 2020, MNRAS, 492, 1791 (Paper II)
Maraston C., 2005, MNRAS, 362, 799 (M05)
Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785
Marino R. A. et al., 2013, A&A, 559, A114
Momcheva I. G., Lee J. C., Ly C., Salim S., Dale D. A., Ouchi M., Finn R., Ono Y., 2013, AJ, 145, 47
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
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Pagel B. E. J., Edmunds M. G., Blackwell D. E., Chun M. S., Smith G., 1979, MNRAS, 189, 95
Pearson W. J. et al., 2019, A&A, 631, A51
Peng Y.-j. et al., 2010, ApJ, 721, 193
Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168
Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2006, ApJ, 642, 797
Pilyugin L. S., 2000, A&A, 362, 325
Pilyugin L. S., 2001, A&A, 369, 594
Pilyugin L. S., Grebel E. K., 2016, MNRAS, 457, 3678
Pilyugin L. S., Thuan T. X., 2005, ApJ, 631, 231
Pilyugin L. S., Grebel E. K., Zinchenko I. A., Nefedyev Y. A., Shulga V. M., Wei H., Berczik P. P., 2018, A&A, 613, A1
R Core Team, 2015, R: A Language and Environment for Statistical Computing. R Found. Stat. Comput., Vienna, Austria
Renzini A., 2020, MNRAS, 495, L42
Riffel R. et al., 2021, MNRAS, 501, 4064
Rosenbaum P. R., Rubin D. B., 1983, Biometrika, 70, 41
Saintonge A. et al., 2011, MNRAS, 415, 61
Saintonge A. et al., 2016, MNRAS, 462, 1749
Sánchez-Blázquez P. et al., 2006, MNRAS, 371, 703
Sanders R. L., Shapley A. E., Zhang K., Yan R., 2017, ApJ, 850, 136
Savage B. D., Mathis J. S., 1979, ARA&A, 17, 73
Skillman E. D., Kennicutt R. C., Jr, 1993, ApJ, 411, 655

Springob C. M., Haynes M. P., Giovanelli R., Kent B. R., 2005, ApJS, 160, 149
Storey P. J., Zeippen C. J., 2000, MNRAS, 312, 813
Sutter P. M., Lavaux G., Wandelt B. D., Weinberg D. H., 2012, ApJ, 761, 44
Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391
Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, ASP Conf. Ser. Vol. 347, Astronomical Data Analysis Software and Systems XIV. Astron. Soc. Pac., San Francisco, p. 29
Tempel E., Stoica R. S., Martínez V. J., Liivamägi L. J., Castellan G., Saar E., 2014, MNRAS, 438, 3465
Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775
Thomas D., Maraston C., Johansson J., 2011, MNRAS, 412, 2183
Thuan T. X., Goehring K. M., Hibbard J. E., Izotov Y. I., Hunt L. K., 2016, MNRAS, 463, 4268
Tiwari J., Mahajan S., Singh K. P., 2020, New Astron., 81, 101417
Tojeiro R., Wilkins S., Heavens A. F., Panter B., Jimenez R., 2009, ApJS, 185, 1
Tremonti C. A. et al., 2004, ApJ, 613, 898
Trevisan M., Ferreras I., de La Rosa I. G., La Barbera F., de Carvalho R. R., 2012, ApJ, 752, L27
Trevisan M., Mamon G. A., Khosroshahi H. G., 2017a, MNRAS, 464, 4593
Trevisan M., Mamon G. A., Stalder D. H., 2017b, MNRAS, 471, L47
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)
van Zee L., Westpfahl D., Haynes M. P., Salzer J. J., 1998, AJ, 115, 1000
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 (V15)
von der Linden A., Wild V., Kauffmann G., White S. D. M., Weinmann S., 2010, MNRAS, 404, 1231
Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2
Willett K. W. et al., 2013, MNRAS, 435, 2835
Woo J. et al., 2013, MNRAS, 428, 3306
Wright E. L. et al., 2010, AJ, 140, 1868
Yang X., Mo H. J., van den Bosch F. C., Jing Y. P., 2005, MNRAS, 356, 1293
Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153
Zaritsky D., Kennicutt R. C., Jr, Huchra J. P., 1994, ApJ, 420, 87
Zhang K. et al., 2017, MNRAS, 466, 3217
Zhang C. et al., 2019, ApJ, 884, L52
Zhuang Y., Leaman R., van de Ven G., Zibetti S., Gallazzi A., Zhu L., Falcón-Barroso J., Lyubenova M., 2019, MNRAS, 483, 1862
Zibetti S., Gallazzi A. R., Hirschmann M., Consolandi G., Falcón-Barroso J., van de Ven G., Lyubenova M., 2020, MNRAS, 491, 3562

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Appendix A. Determination of the red sequence at different redshifts.
Appendix B. Visual classification of interacting and merging galaxies.
Appendix C. VYG images and data.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

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