Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **468,** 1747–1759 (2017)
Advance Access publication 2017 February 22

doi:10.1093/mnras/stx453

On the robustness of the Hβ Lick index as a cosmic clock in passive early-type galaxies

Alice Concas,[1,2,3]★ L. Pozzetti,[3] M. Moresco[2,3] and A. Cimatti[2]

[1]*Excellence Cluster Universe, Boltzmannstr. 2, D-85748 Garching, Germany*
[2]*Dipartimento di Fisica e Astronomia, Universitá degli Studi di Bologna, V.le Berti Pichat 6/2, I-40127 Bologna, Italy*
[3]*INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, I-40127 Bologna, Italy*

Accepted 2017 February 20. Received 2017 February 17; in original form 2016 September 2

ABSTRACT

We examine the Hβ Lick index in a sample of ∼24 000 massive ($\log(M/M_\odot) > 10.75$) and passive early-type galaxies extracted from the Sloan Digital Sky Survey at $z < 0.3$, in order to assess the reliability of this index to constrain the epoch of formation and age evolution of these systems. We further investigate the possibility of exploiting this index as 'cosmic chronometer', i.e. to derive the Hubble parameter from its differential evolution with redshift, hence constraining cosmological models independently of other probes. We find that the Hβ strength increases with redshift as expected in passive evolution models, and shows at each redshift weaker values in more massive galaxies. However, a detailed comparison of the observed index with the predictions of stellar population synthesis models highlights a significant tension, with the observed index being systematically lower than expected. By analysing the stacked spectra, we find a weak [N II] λ6584 emission line (not detectable in the single spectra) that anti-correlates with the mass, which can be interpreted as a hint of the presence of ionized gas. We estimated the correction of the Hβ index by the residual emission component exploiting different approaches, but find it very uncertain and model dependent. We conclude that, while the qualitative trends of the observed Hβ–z relations are consistent with the expected passive and downsizing scenario, the possible presence of ionized gas even in the most massive and passive galaxies prevents us to use this index for a quantitative estimate of the age evolution and for cosmological applications.

Key words: galaxies: evolution – galaxies: formation – galaxies: fundamental parameters – galaxies: general – galaxies: stellar content – cosmological parameters.

1 INTRODUCTION

Early-type galaxies (ETGs) are perfect candidates to investigate both galaxy formation and evolution theories and, at the same time, to put constraints on the expansion history of the Universe. Many recent observations show that most massive galaxies contain the oldest stellar populations up to $z \sim 1$–2 (e.g. Spinrad et al. 1997; Cowie, Songaila & Barger 1999; Heavens et al. 2004; Thomas et al. 2005, 2010; Cimatti et al. 2008; Moresco et al. 2011) and that only few per cent of their current stellar mass was formed after $z \sim 1$. Moreover, studies on the stellar mass function at different redshifts (e.g. Bundy, Ellis & Conselice 2005, 2006; Borch et al. 2006; Ilbert et al. 2010; Pozzetti et al. 2010; Maraston et al. 2013) showed that most massive galaxies ($M \sim 10^{11}\,M_\odot$) are characterized by an increase of their stellar mass density at $z \sim 1$ (e.g. Ilbert et al. 2010 based on the COSMOS 2 deg^2 field) and a subsequent negligible evolution in number density from $z \sim 1$ to the present (Cimatti, Daddi & Renzini 2006; Pozzetti et al. 2010; Moresco et al. 2013). All these pieces of observational evidence suggest that these massive, red and passively evolving galaxies represent the oldest objects in the Universe below $z \sim 1$. For this reason, a study based on the variation of ETG properties, such as age, chemical composition, dynamic, at different redshifts, could rightly help to expand our current knowledge of the physical process that drive the formation and evolution of galaxies. Moreover, as first pointed out by Jimenez & Loeb (2002), the differential age evolution of ETGs can also be used to put several cosmological constrains. In particular, by estimating the relative differential age ($\Delta t_{\rm age}$) between ETGs, formed at the same cosmic time, but observed at different redshifts (Δz), it is possible to determine the derivative of redshift with respect to cosmic time $\mathrm{d}z/\mathrm{d}t$.

This quantity is directly related to the cosmological expansion history described by the Hubble parameter: $H(z) = -\frac{1}{1+z}\frac{\mathrm{d}z}{\mathrm{d}t}$, which can be used to estimate many cosmological parameters such as H_0, $\Omega_{0,M}$, $\Omega_{0,\Lambda}$ and w_Λ. This technique, known as 'cosmic

★ E-mail: alice.concas@tum.de

chronometers' method, has been used recently in several works and obtained promising results in the estimate of the Hubble parameter and of the dark energy equation of state (Simon, Verde & Jimenez 2005; Stern et al. 2010; Moresco et al. 2012a,b, 2016b; Zhang et al. 2014). The basic problem of this method is the fact that it relies on the measure of stellar population age, which presents, when estimated from a fit of the spectrum and/or spectral energy distribution (SED), significant degenerations with other parameters, such as metallicity, star formation history (SFH) and dust content. An alternative method to mitigate these degeneracies is to identify a particular spectral feature that is sensitive to the ageing of the stellar population. An example of this approach is provided by Moresco et al. (2011, 2012b, 2016a) and Moresco (2015), in which the expansion history of the Universe is determined by studying the redshift dependence of the 4000 Å break. Despite the high potential of the method, the break at 4000 Å is particularly sensitive not only to the age, but also to the variation of the metallicity of the stellar population, which has to be taken into account as a systematic error in the estimate of cosmological parameters. The need to find a new spectral feature that minimizes the dependence systematic and able to translate the amount $\mathrm{d}z/\mathrm{d}t$ in the observable $\mathrm{d}z/\mathrm{d}(\mathrm{feature})$ and then able to determine the cosmological parameter with higher accuracy is evident.

In the effort of exploiting reliable age indicators, we present a new study of ETGs' relative age evolution based on the variation of the Hβ Lick index with redshift. This index, first introduced by Burstein et al. (1984) and redefined by Worthey et al. (1994), has been identified as the one with the maximum dependence on the variation of the age of the stellar population, and the minimal dependence on metallicity and chemical composition (Trager et al. 1998; Lee & Worthey 2005; Thomas, Maraston & Johansson 2011). In detail, we measure the Hβ Lick index in a sample of passive galaxies selected from the Sloan Digital Sky Survey (SDSS), to check if this particular spectral feature is actually able to provide a reliable differential dating of galaxy stellar ages, in order to derive constraints both on galaxy evolution and formation and on the expansion history of the Universe.

This paper is organized as follows. In Section 2, we introduce the data sample. We describe the physical parameters of the passive galaxy sample, and our method for the Hβ Lick measurement. In Section 3, we present the results about the Hβ–z and Hβ–mass relations and compare them with the theoretical predictions of the stellar population synthesis (SPS) models. We discuss the effect of possible contamination by the emission lines in the observed Hβ absorption line in Section 4, and finally in Section 5 we list our conclusions. Throughout the paper, we assume $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_{0,M} = 0.25$ and $\Omega_\Lambda = 0.75$ cosmology.

2 DATA

The galaxies analysed in this work are selected from the spectroscopic catalogue of the Sixth Data Release of Sloan Digital Sky Survey (SDSS-DR6). The SDSS-DR6 provides photometry in the u, g, r, i, z bands and the spectrum of about 800 000 galaxies, extracted with Petrosian magnitude $r < 17.77$. Each spectrum is measured from 3800 to 9200 Å with a resolution $\lambda/\Delta\lambda \approx 1800$–2000.[1] The SDSS data are provided in vacuum wavelengths. To be meaningfully compared with atomic transitions, we converted them in the air system.[2] To have a wider photometric coverage, we decided to analyse the sample provided by Moresco et al. (2011), where SDSS galaxies are matched with the Two Micron All Sky Survey, with photometry in the J, H and K bands. This wider coverage allows a better accuracy in the stellar mass estimates from the SED fitting, and a more precise selection based on photometric data (see Section 2.1).

2.1 The sample of passive ETGs

This study strongly relies on the selection of a sample of passively evolving galaxies that, having assembled most of their mass at high redshifts, are able to trace homogeneously the age evolution of the Universe. To select the purest possible sample of passive galaxies, with no evidence of recent episodes of star formation, we followed the approach described in Moresco et al. (2011), and selected galaxies with:

(i) *no strong emission lines:* starting from the equivalent width (EW) measurements provided by the analysis of the MPA-JHU DR7,[3] we select only the objects with a rest-frame equivalent width $\mathrm{EW(H}\alpha) > -5$ Å and $\mathrm{EW([O\,II]}\,\lambda3727) > -5$ Å, where, by convention, the measures are given in negative values when in emission.

(ii) *SED matching the reddest passive ETG templates*: following the approach of Zucca et al. (2006), the SED of each galaxy has been compared with a library of 62 empiric SEDs described in Ilbert et al. (2006), comprising galaxies with both old stellar populations and intense star formation activity. The galaxies are divided according to their SED, and we select those galaxies matching an early-type template, corresponding to the four reddest templates.

With this selection, we obtain a sample of ∼100 000 galaxies. For more details, we refer to Moresco et al. (2011).

In order to investigate the evolution of our sample, we study their physical properties, focusing in particular on the stellar mass, stellar metallicity and velocity dispersion. We use the stellar mass estimates from Moresco et al. (2011), obtained by a multi-colour SED-fitting procedure (eight photometric bands from u to K). We apply a further mass cut, selecting only the most massive galaxies with a stellar mass $\log(M/\mathrm{M}_\odot) \geq 10.75$. This choice is made to select only the most massive objects that, according to the downsizing scenario, should correspond to the oldest objects (Thomas et al. 2010). In the *mass-downsizing* scenario, the most massive galaxies were formed earlier and over a shorter period than those in less massive galaxies. For this reason, we decide to split our sample into four narrow mass bins ($\Delta\log(M/\mathrm{M}_\odot) = 0.25$), to have a more homogeneous sampling of the redshift of formation: $10.75 < \log(M/\mathrm{M}_\odot) < 11$ (BinI), $11 < \log(M/\mathrm{M}_\odot) < 11.25$ (BinII), $11.25 < \log(M/\mathrm{M}_\odot) < 11.5$ (BinIII) and $\log(M/\mathrm{M}_\odot) > 11.5$ (BinIV) (see Fig. 1). Furthermore, for this sample, we have a measure of the SFH estimated by Moresco et al. (2011) with a delayed exponential SFH: $\mathrm{SFR}(t,\tau) \propto (t/\tau^2)\exp(-t/\tau)$ (with τ in the range of values $0.05 \leq \tau \leq 1$ Gyr). The median value of the our total sample is $\tau = 0.2$ Gyr.

[1] http://www.sdss3.org/dr9/scope.php# specstats

[2] The conversion relation from vacuum wavelengths to air wavelengths, proposed in Morton (1991), is the following:

$$\lambda_{\mathrm{AIR}} = \frac{\lambda_{\mathrm{VAC}}}{\left(1.0 + 2.735\,182\mathrm{E}-4 + \frac{131,4182}{\lambda^2_{\mathrm{VAC}}} + \frac{2.762\,49\mathrm{E}8}{\lambda^4_{\mathrm{VAC}}}\right)} \quad (1)$$

with λ_{AIR} and λ_{VAC} expressed in Å.

[3] http://www.mpa-garching.mpg.de/SDSS/DR7/

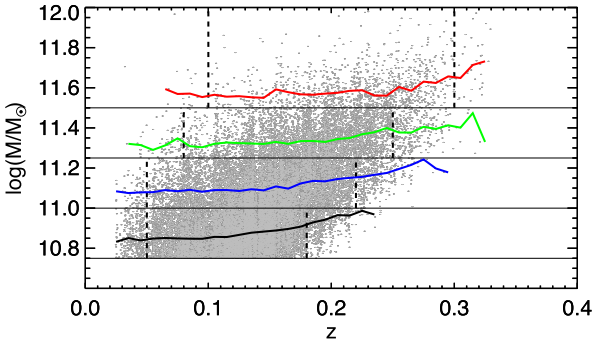

Figure 1. Stellar mass–redshift distributions of the four mass subsamples (red, green, blue and black curves, respectively, for BinIV, BinIII, BinII and BinI, from top to bottom). The black dotted vertical lines indicate the redshift division of the sample (see Table 1 for the stellar mass median values).

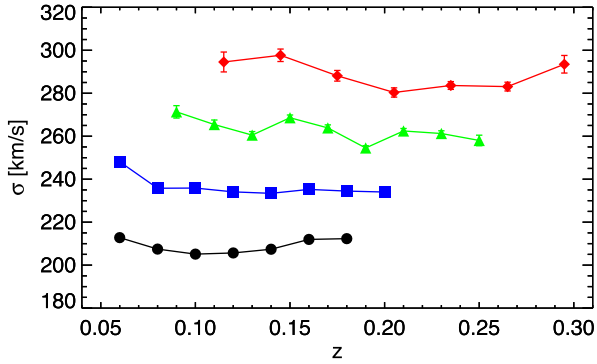

Figure 2. Velocity dispersion–redshift relations, averaged in bin of redshift, for the different mass subsamples, BinIV, BinIII, BinII and BinI from top to bottom (see Table 1 for the median values)

In Fig. 1, we present the mass–redshift distribution. It is evident that the median mass increases with the redshift due to magnitude limit effect of the survey. Therefore, to avoid possible biases, we select the galaxies having $z \leq 0.18$, $z \leq 0.22$, $z \leq 0.25$ and $z \leq 0.3$, respectively, for BinI, BinII, BinIII and BinIV. In this way, we limit the difference in the median mass along the redshift range to 0.1 dex on average (see Fig. 1).

As mentioned in the previous section, the Hβ Lick index is the less sensitive to metallicity variations compared to all other indices of the Lick system. However, in the wavelength range where the index is defined, there may be weak metal absorption lines and, for this reason, the metallicity dependence of the Hβ index is not totally negligible. The sample has been therefore cross-matched with the SDSS-DR4 subsample obtained by Gallazzi et al. (2005), for which the stellar metallicities of the galaxies are estimated from the simultaneous fit of several spectral features (D4000, Hβ, Hδ_a + Hγ_a, [Mg_2Fe] and [MgFe]$'$). We study the metallicity–redshift relation. In order to minimize the possible metallicity effect, we made a further cut in redshift, selecting only galaxies with redshift $z \geq 0.05$ for BinI and BinII, $z \geq 0.08$ for BinIII and $z \geq 0.1$ for BinIV. The median metallicity of the final sample is $Z/Z_{\odot} \sim 1.2$ (with a total spread of about 0.3 within the total redshift range analysed and of ∼0.2 in narrow redshift bins $\Delta z \sim 0.02$–0.03), showing only a slight variation with redshift, in agreement with the recent metallicity measurements obtained by total spectral fitting of our stacked spectra (Citro et al. 2015). We note, further, that there is no significant difference in the median metallicity of the different mass subsamples.

In this paper, we adopt the velocity dispersion derived by the Princeton group.[4] Fig. 2 shows the velocity dispersion relations for all mass subsamples. We find that the velocity dispersion is constant in each subsample (see Table 1 for the median values).

Finally, the observed optical spectrum can be contaminated by the presence of night sky emission lines (see Hanuschik 2003). Particularly, the night sky emission lines at 5577 Å ([O I]) and at 5889.950, 5895.920 (Na I) Å are very strong, and leave significant contamination in Hβ measurement.

In order to avoid this source of contamination, we decide to exclude the galaxies having a redshift between $0.14 < z < 0.155$ and $0.20 < z < 0.22$, which are the redshifts for which those lines fall into the Hβ definition range.

[4] The data are available on the website http://spectro.princeton.edu/

After all the cuts adopted, the final sample consists of ∼23 914 galaxies. The median values and the respective errors[5] of mass and velocity dispersion for the different mass subsamples are reported in Table 1. In the same table, we also report the number of galaxies in each bin.

Fig. 3 shows the median stacked spectra for the four mass subsamples. From the left-hand panel, it is evident that all spectra show typical features of a passive population, with (i) absence of emission lines ([O II], Hβ, [O III] and Hα), (ii) strong absorption features (Hβ, Hα, *G* band, Ca H&K) and (iii) strong D4000 break. We notice that there is still present a weak emission at 6584 Å [N II], and that the amplitude of this feature is mass dependent, being most evident in the less massive subsample (BinI). For more details, see Section 4. We highlight the region near Hβ line (right-hand panel); all the four spectra show a typical and unperturbed absorption profile.

2.2 Hβ index

We measure the Hβ Lick absorption line index for all 23 914 galaxies. We implement an IDL (Interactive Data Language[6]) code in order to perform measures coherently on observational data and models. We use the Lick Hβ index definition by Worthey et al. (1994); the bandpasses are shown in Fig. 3. In the standard Lick system, the indices are defined on low-resolution spectra [∼9 Å full width at half-maximum (FWHM) in Hβ region]. However, the modern spectroscopic surveys, and in particular the SDSS, provide spectra with a much higher resolution (∼2.4–2.7 Å FWHM), comparable with the new SPS models (for example Bruzual & Charlot 2003 and Maraston & Strömbäck 2011). In order to measure coherently this feature in the models and observations, we decide to perform the analysis directly on the observed spectra, without altering the resolution.

Before applying our code to real data, we test its reliability to reproduce the Hβ Lick values tabulated in the Bruzual & Charlot (2003) templates. We find a very good agreement between the two measurements with a negligible median offset of 2 per cent. We also compare our measurements with those derived by the MPA-JHU group for the same galaxy sample (http://www.mpa-garching.mpg.de/SDSS/DR7/). We find a good agreement between the values, with a mean difference of −0.03 Å and a standard deviation

[5] The error on the median is evaluated as the median absolute deviation/$\sqrt{N}$, MAD = 1482∗median(|x − median(x)|) (see Hoaglin, Mosteller & Tukey 1983).

[6] http://www.exelisvis.com/ProductsServices/IDL.aspx

Table 1. Redshift range, mass range, median mass and median velocity dispersion of the passive galaxies in different mass subsamples.

	Redshift z	Mass range $\log(M/M_\odot)$	Median mass $\log(M/M_\odot)$	Median velocity dispersion σ (km s^{-1})	No. of galaxies
BinI	0.05–0.18	10.75–11.0	10.87 ± 0.001	207.3 ± 0.3	11 631
BinII	0.05–0.22	11.0–11.25	11.10 ± 0.001	234.5 ± 0.4	7990
BinIII	0.08–0.25	11.25–11.5	11.34 ± 0.001	259.9 ± 0.6	3314
BinIV	0.1–0.3	>11.5	11.58 ± 0.002	286.1 ± 1.1	979
Total	0.05–0.3	>10.75	11.01 ± 0.001	226.7 ± 0.2	23 914

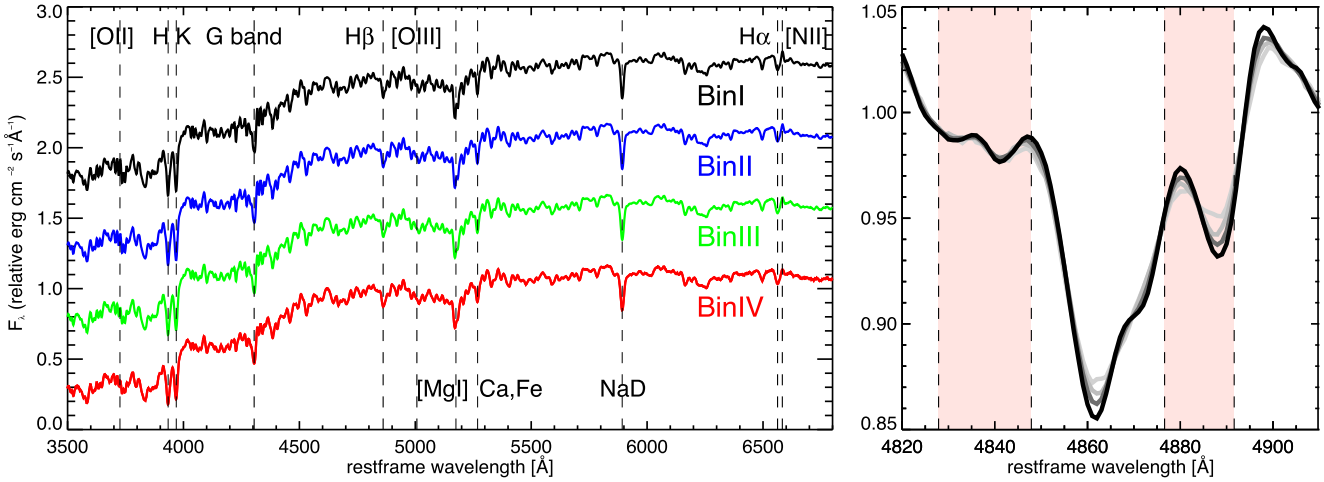

Figure 3. Median stacked spectra of the four passive galaxies' mass subsamples (from top to bottom, respectively, for BinI, BinII, BinIII and BinIV, left-hand panel). In the range between 3500 and 6800 Å, there are not significant emission lines ([O II], [O III], Hα and [N II] λ6584). The Hβ line shows a typical absorption feature (right-hand panel) in all the mass bins (from black to light grey for BinI, BinII, BinIII and BinIV, respectively). The spectra are normalized between 5000 and 5500 Å rest frame.

of 0.033 Å, which is comparable with the average measured error. Therefore, we find that the two independent methods are consistent, and without any significant biases.

3 TRENDS OF Hβ

We derive the median Hβ–redshift relations in each mass subsample; we use narrow redshift bins ($\Delta z \sim 0.02$–0.03) for the three most populated mass subsamples, while in the more massive subsample we use $\Delta z \sim 0.04$ because of the lower statistic with respect to the other samples.

The results are shown in Fig. 4. From this figure, we note two main trends.

(i) A clear Hβ–redshift relation for each subsample, with an increase in the index strength of ∼10 per cent along the redshift range.

(ii) A clear trend with mass, i.e. at each redshift more massive galaxies present a lower Hβ index with respect to less massive ones (∼9 per cent).

It is interesting to note that both effects can easily be seen even by direct inspection of the median spectra. As an example, we show the stacked spectra at different redshifts for the sample with $11 < \log(M/M_\odot) < 11.25$ (BinII, see Fig. 5, left-hand panel), from which a deeper absorption line with redshift is evident. The trend with the mass is also visible in the stacked spectra of galaxies at fixed redshift, with most massive galaxies showing a smaller absorption line than lower mass ones (see Fig. 5, right-hand panel).

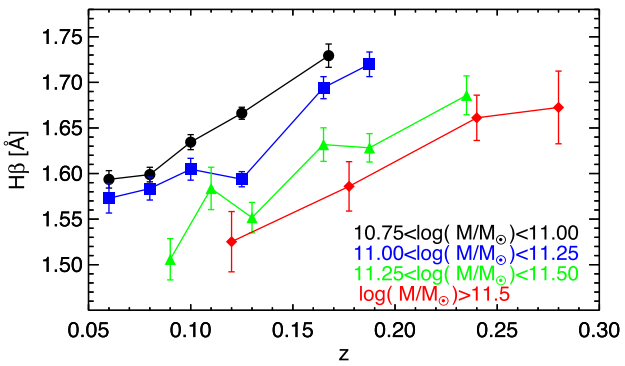

Figure 4. Median Hβ–redshift relation for passive galaxies in different mass subsamples (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV, respectively).

Both these trends are qualitatively consistent with a general ageing of the stellar population and a mass–age relation ('mass downsizing'), also found by Thomas et al. (2010). Our results also concur with a recent work of Choi et al. (2014). By modelling the full optical spectrum of a quiescent galaxy sample at low and intermediate redshift ($0.07 < z < 0.09$ from the SDSS and $0.1 < z < 0.7$ from AGN and Galaxy Evolution Survey), Choi et al. (2014) found that the best-fitting single stellar population (SSP) equivalent age is higher for the most massive galaxies and it increases with decreasing redshift at fixed stellar mass. This correspondence between two different methods highlights the potential of our Hβ line as a tracer

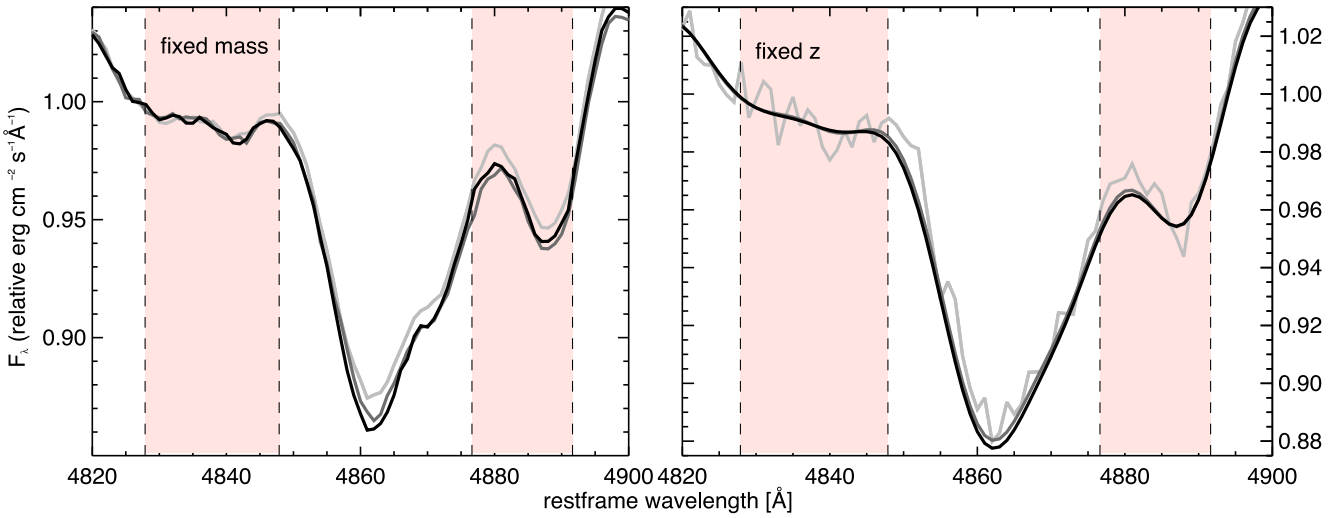

Figure 5. Median stacked spectra in the Hβ line region. Left-hand panel: stacked spectra with the same mass ($11 < \log(M/M_{\odot}) < 11.25$, BinII) and different mean redshift $z = 0.08$, 0.10 and 0.1875 (from light grey to black, respectively). Right-hand panel: stacked spectra with the same redshift, $z = 0.125$, for different mass subsamples (black, grey and light grey for BinI, BinII and BinIV, respectively). For illustrative purposes, the BinI and BinII spectra are smoothed to the common velocity dispersion of BinIV.

of ageing in the stellar population. We test this interpretation in Section 3.1.

3.1 Sensitivity to metallicity and velocity dispersion

The Hβ index is not totally immune to metallicity ($Z/Z_{\odot}$) effects for the integrated stellar population (Worthey et al. 1994). Furthermore, this index is sensitive to the galaxy velocity dispersion (σ). To exclude the possibility that the trends found are due to a σ or $Z/Z_{\odot}$ variation, we study the impact of these parameters using SPS models. We consider Bruzual & Charlot (2003) and Maraston & Strömbäck (2011) models (hereafter BC03 and MaStro, respectively). BC03 models are available at a resolution of 3 Å FWHM in the wavelength range between 3200 and 9500 Å; MaStro models are available with a resolution of ∼2.3 Å FWHM[7] between 3525 and 7500 Å. Both models share a resolution very similar to the one of SDSS spectra (≈1800–2000 between 3800 and 9200 Å). This is particularly useful since it allows a direct comparison between the observed spectra with the theoretical SED without having to modify the spectral resolution.

These models are used to create libraries of synthetic spectra with different velocity dispersions ($\sigma = 0$, 200, 250 and 300 km s^{-1}) and different metallicity ($Z/Z_{\odot} = 1$, 1.4, 1.5 obtained by interpolating quadratically between $Z/Z_{\odot} = 0.4$, 1, 2.5 available in BC03 and $Z/Z_{\odot} = 0.5$, 1, 2 in MaStro). The grid has been created with age $0.5 \leq t \leq 15$ Gyr, with delayed exponential SFH, $\mathrm{SFR}(t, \tau) \propto (t/\tau^2)\exp(-t/\tau)$, where τ is chosen in the range $0.05 \leq \tau \leq 1$ Gyr.

3.1.1 Effects of the velocity dispersion

To check the effect of the velocity dispersion in our Hβ Lick measures, we perform several tests. By using the theoretical spectra described in the previous section, we verify that an increase of velocity dispersion causes an apparent decrease of the observed Hβ index strength only due to a broadening effect. We further estimate the maximum percentage difference in Hβ between our mass bins that can be attributed to σ from the analysis of SPS models; we find that the percentage difference in Hβ between an SSP with $\sigma = 200$ and 250 km s^{-1} is <1.5 per cent, and is <3 per cent between $\sigma = 200$ and 300 km s^{-1}, for both models. We underline that given the range of σ of our data, these two estimates represent the maximum difference in Hβ that can be attributed to a σ effect between our mass subsamples, and they are always lower than the mean percentage difference in our subsamples with similar σ differences. Indeed, the mean percentage difference in Hβ median values between BinI and BinIII is ∼6 per cent, and ∼9 per cent between BinI and BinIV (similarly, we get a difference of ∼4 per cent between BinII and BinIII, and ∼6 per cent between BinII and BinIV). Moreover, we note that all SPS models analysed show that SEDs with different velocity dispersions have a difference in flux in the red pseudo-continuum region (4876.625–4891.625), but the stacked spectra of the same mass subsample do not show this trend (see Fig. 5).

For this reason, we exclude that the velocity dispersion by itself can cause the observed mass segregation between the Hβ–z relation in different mass subsamples.

Finally, we verified that in our data the median velocity dispersion (see Table 1 for median values of BinI, BinII, BinIII and BinIV) shows almost no redshift evolution in each mass bin (see Fig. 2). For this reason, we exclude that the observed Hβ index increase with redshift can be due to a velocity dispersion effect.

3.1.2 Effects of the metallicity

The models are also used to establish the effect in the Hβ index due to the stellar metallicity. While there is no significant difference in the median metallicity of the different mass subsamples, each mass bin instead shows a slight decrease of the median metallicity with redshift.

We create a library of Hβ indices as a function of age using stellar population models with different metallicity: 1, 1.4 and 1.5 $Z/Z_{\odot}$. The study of these Hβ synthetic indices shows that even an extreme variation of metallicity, between $Z/Z_{\odot} = 1$ and $Z/Z_{\odot} = 1.4$, 1.5,

[7] More precisely, in Beifiori et al. (2011) it is shown that the spectral resolution of the MaStro models, based on the stellar library MILES, is about 2.54 Å FWHM, instead of the nominal 2.3 Å.

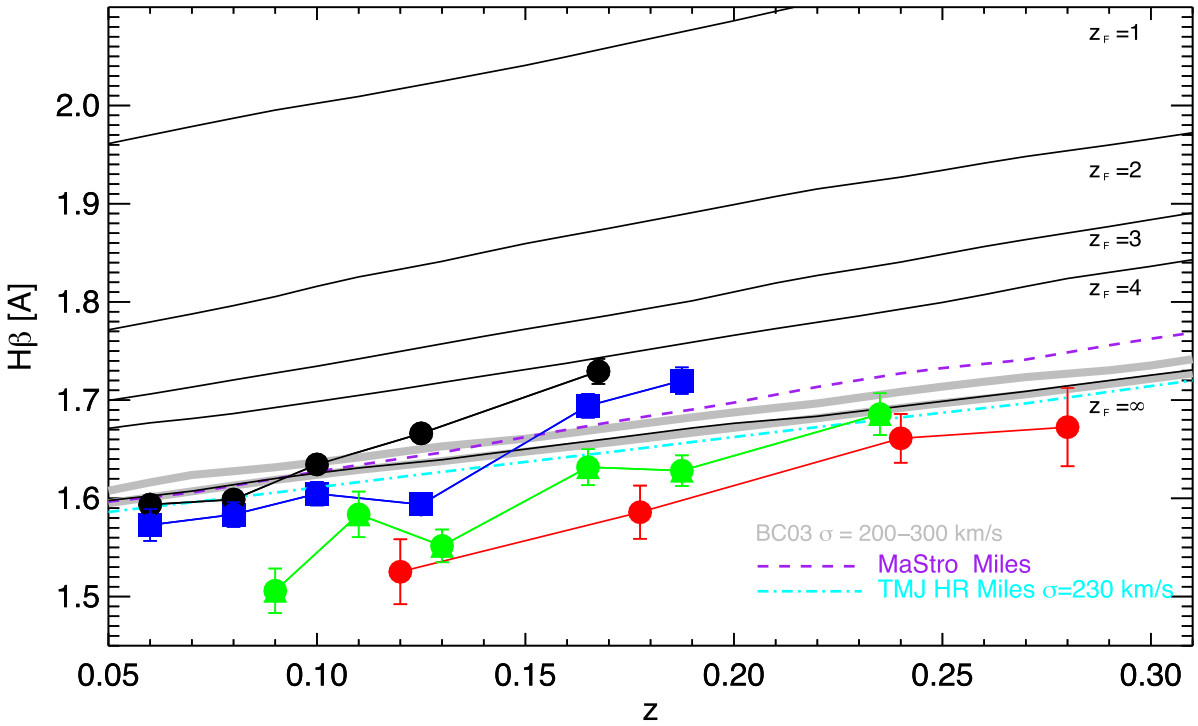

Figure 6. Comparison between the observed Hβ–z relations (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV, respectively) and different models. The black solid lines are the BC03 models with $\sigma = 230$ km s^{-1}, $Z/Z_{\odot} = 1.2$ and different formation redshifts ($z_F = 1$, 2, 3, 4 and ∞). For $z_F = \infty$, we also show the models with $\sigma = 200$ and 300 km s^{-1}, gray top and bottom regions. The MaStro STELIB and TMJ MILES models are shown for $z_F = \infty$ (dashed and dash dot cyan lines, respectively). The MaStro models are showed at $\sigma = 230$ km s^{-1} and $Z/Z_{\odot} = 1.2$. The TMJ models are rescaled at $\sigma = 230$ km s^{-1} using the velocity dispersion corrections proposed by Carson & Nichol (2010). (A colour version of this figure is available in the online version.)

implies an Hβ variation of only ∼4–5 per cent on average, always lower than Hβ trend observed. Moreover, to further check the metallicity effect, not only in the models but also on our data, we create a control sample of galaxies with a narrow metallicity range (e.g. $\Delta(Z/Z_{\odot}) = 0.5$ for BinI).

Even in this subsample, we find a steady trend of growth in the Hβ–z relations.

From these analyses, we find that the increase of the Hβ EW with redshift (at fixed mass) and decrease with mass (at fixed redshift) of passive galaxies cannot be only due to σ or $Z/Z_{\odot}$ variation, and hence can mostly be explained by an age-evolution effect, in particular with the age of the stellar population decreasing with increasing redshift, and increasing with increasing mass.

3.2 Hβ age: comparison with SPS models

To interpret the evolution of Hβ with redshift and its dependence on mass in the context of passive galaxies evolution, we proceed with a direct comparison between the observed data and the index values calculated on the SED provided by SPS models, hereafter theoretical Hβ. Both the observed and theoretical measurements are obtained by the same method, described in Section 2.2. The theoretical Hβ curves as a function of age are transformed into Hβ–z relations assuming a Λ cold dark matter cosmology (see Section 1 for the parameters used), and a formation redshift.

For our studies, we chose to probe different redshifts of formation for our galaxies, namely $z_F = 1$, 2, 3, 4 and ∞. For this comparison, we use two different SPS models, BC03 and MaStro and the theoretical models of Hβ Lick indices of Thomas et al. (2011, hereafter TMJ), calculated by theoretically manipulating the index response functions. For BC03 and MaStro models, we adopt a Chabrier initial mass function (IMF) instead the TMJ models have a Salpeter IMF. However, using BC03 models, we verify that an IMF variation does not affect significantly the Hβ value. In particular, in the case of solar metallicity, the choice of a Salpeter IMF instead of a Chabrier causes an index decrease of <1 per cent. Furthermore, in the BC03 and MaStro models, we use a delayed exponential SFH, in particular we use models with $\tau = 0.2$ Gyr that correspond to the τ median value estimate for the our sample (see Section 2.1). We also investigate the possible variation of Hβ value with different τ values. We compare SPS models with $\tau = 0.05$ and 0.2 Gyr, and we find that for SPS older than 6 Gyr the Hβ Lick index decreases by ≤1 per cent in the SPS with $\tau = 0.05$ Gyr to respect at SPS with $\tau = 0.2$ Gyr. Since the median metallicity of the sample is $Z/Z_{\odot} \sim 1.2$, for the comparison with the data we interpolated quadratically the Hβ values of all models in the three metallicities provided, in order to have an estimate of the index at $Z/Z_{\odot} = 1.2$. The BC03 and MaStro models are convolved with a velocity dispersion $\sigma = 230$ km s^{-1}, comparable to the median value of our sample. In Fig. 6, we show as an example the observed Hβ–z relations and the theoretical curves (BC03 models) for different formation redshifts. For MaStro and TMJ models, we find similar results, as shown in the case of formation redshift $z_F = \infty$.

From the comparison with theoretical models, as can be clearly seen from Fig. 6, we find that they do not reproduce the observed Hβ median relations, since the models predict higher Hβ values than the median Hβ measured on our samples, so that in most cases the observed values would require an age greater than the age of the Universe. For the most massive sample, the minimum shift between data and models is of the order of Δ H$\beta = 0.1$.

Since the models are built with the same metallicity and velocity dispersion found in the data, we exclude that this inconsistency is due to one of these two parameters. We verify that also by imposing an extreme velocity dispersion for our sample ($\sigma = 300$ km s^{-1}), the models do not reproduce the observations. Regarding the metallicity, we instead find that the Hβ–z relations can be reproduced only by considering models with metallicity approximately $Z/Z_{\odot} = 2$, which are inconsistent with the metallicity estimated for this sample.

Furthermore, comparing the metallicity estimated by Thomas et al. (2010) for a morphological selected sample of ETGs at $0.05 \leq z \leq 0.06$, we notice that for a subsample with velocity dispersion similar to our values they estimate a slightly higher metallicity than those considered in this work (Gallazzi et al. 2005). In particular, for $\sigma = 207$ km s^{-1} (median value of the less massive sample, BinI), Thomas et al. (2010) find a metallicity $Z/Z_{\odot} \sim 1.5$, and for the median value of the total sample, $\sigma = 230$ km s^{-1}, they estimate a metallicity of the order of $Z/Z_{\odot} \sim 1.6$ (for more details, refer to Thomas et al. 2010). These higher metallicity values are also confirmed in a more recent work by Citro et al. (2015). In this study, the authors find a median $Z/Z_{\odot} \sim 1.450 \pm 0.075$. Even these higher values are not enough to completely reconcile models with data.

4 INVESTIGATING THE ORIGIN OF THE OFFSET

In this section, we further investigate possible causes of the offset between the model predictions and the measures of the Hβ Lick index shown in Section 3.2. In particular, we explore the possibility that our sample, despite the selection criteria (see Section 2.1), is still contaminated by an emission line component. This potential emission component would contaminate the measurement of the Hβ index by filling the line and producing a less intense absorption feature, consistent with the recent results shown by Serven & Worthey (2010) for an SDSS low-redshift ($0.06 < z < 0.08$) galaxy sample.

4.1 Improving sample selection: the [N II] emission line contamination

As discussed in Section 2, we select only the galaxies that do not show significant emission lines in Hα and [O II] (EW < -5 Å). We check that also cutting the sample by adopting a more stringent Hα limit (S/N < 3), the offset between the Hβ Lick index in our data and models is not removed. However, a more detailed analysis of the stacked spectra shows the presence of a weak [N II] emission at 6584 Å (see Fig. 3). Since the first ionization potential of the nitrogen is very similar to that of hydrogen, the presence of an emission line in [N II] λ6584 may indicate the presence of an emission component also in the H lines. In order to study the possible emission line contamination in our sample, we use the EW([N II] λ6584) values provided by the MPA-JHU group. From this analysis, about 43 per cent of the total sample have an equivalent width EW([N II] λ6584) with S/N > 3, consistent with the clearly visible [N II] λ6584 emission line in the median spectra (see Fig. 3, left-hand panel). The median and dispersion[8] EW of the total sample are EW([N II] λ6584) $= -0.56 \pm 0.58$ Å, in agreement with the EW measured on the median stacked spectrum (EW([N II] λ6584) $= -0.68$ Å).

[8] The dispersion is evaluated with the median absolute deviation, MAD.

Moreover, by analysing separately each mass subsample, we identify a relationship of the [N II] λ6584 emission line as a function of stellar mass. In Fig. 7, bottom-right panel, we show the stacked spectra of the four mass subsamples, from which it is evident that the [N II] λ6584 emission line increases with decreasing of the median mass. This evidence is confirmed also by analysing the EW([N II] λ6584)–z relations: we note that the EW([N II] λ6584) median values are almost constant with redshift, and they decrease with increasing mass (see Fig. 8).

To verify if the presence of a [N II] λ6584 emission is related to the presence of Hβ emission line, we analyse the galaxies most affected by this contamination. We select the galaxies with a [N II] λ6584 clearly detected, i.e. with S/N([N II] λ6584) > 10; in this way, we select 1145 galaxies[9] (5 per cent of the total sample). Fig. 7 shows the median spectrum relative to this subsample. It can be seen that, differently from the stacked spectrum of the global sample (see Fig. 3), in this spectrum, together with a strong [N II] λ6584, there are clearly detectable several emission lines, i.e. [O II], [O III] λ5007, [N II] λ6548, Hα and [S II] λλ6717, 6731 (Fig. 7). Nevertheless, also in this case, in the Hβ region, there is not a clear emission component (Fig. 7, bottom-left panel). The Hβ absorption line only shows a particular shape, different from that expected for a simple absorption, probably caused by an overlapping with a weak emission line contamination.

Following this result, we decide to apply a further cut to our sample, by selecting only the galaxies with a [N II] λ6584 line detected with S/N([N II] λ6584) < 3; this sample is composed of 13 626 galaxies (∼57 per cent of the total sample). We find that, even for this sample there is a residual [N II] λ6584 line, with a median and dispersion (see footnote 8) EW([N II] λ6584) $\sim -0.26 \pm 0.27$ Å. The residual emission can be seen from the comparison between the stacked spectrum of this sample and a pure absorption SPS model (see Fig. 7, bottom panels).

On this new subsample with S/N([N II] λ6584) < 3, we calculate also the median Hβ–z relations at different masses; the results are shown in Fig. 9. We note that the cut in [N II] λ6584 produces a general increase in median values of the median Hβ index (on average to ∼0.05 Å for BinI and BinII, ∼0.02 for BinIII and ∼0.01 for BinIV). The offset found between data and observations of the total sample is reduced. However, there is still a significant disagreement between data and theoretical Hβ–z curves, especially for the more massive sample (BinIV). However, despite this discrepancy, we find similar Hβ–z and Hβ–mass relationships as seen in the total sample, but with slightly different slopes and normalizations. These variations make impossible to use the median Hβ values for an accurate estimate of the absolute and differential age evolution of the stellar populations with the redshift. Consequently, it is not possible to place constraints on the redshift of formation of the galaxies and even the use of the Hβ–z relation in the method of the cosmic chronometers appears inadequate.

4.2 Methods to remove nebular emission

The SDSS galaxy spectra are extensively used for many purposes, such as stellar mass and SFH estimates (Kauffmann et al. 2003), age and metallicity estimates in the local universe (Gallazzi et al. 2005),

[9] In this work, we decide to cut the sample above this limit (S/N([N II] λ6584) > 10) because this enables us to select only those objects in which the [N II] line is detected effectively while maintaining a good statistic (1145 galaxies).

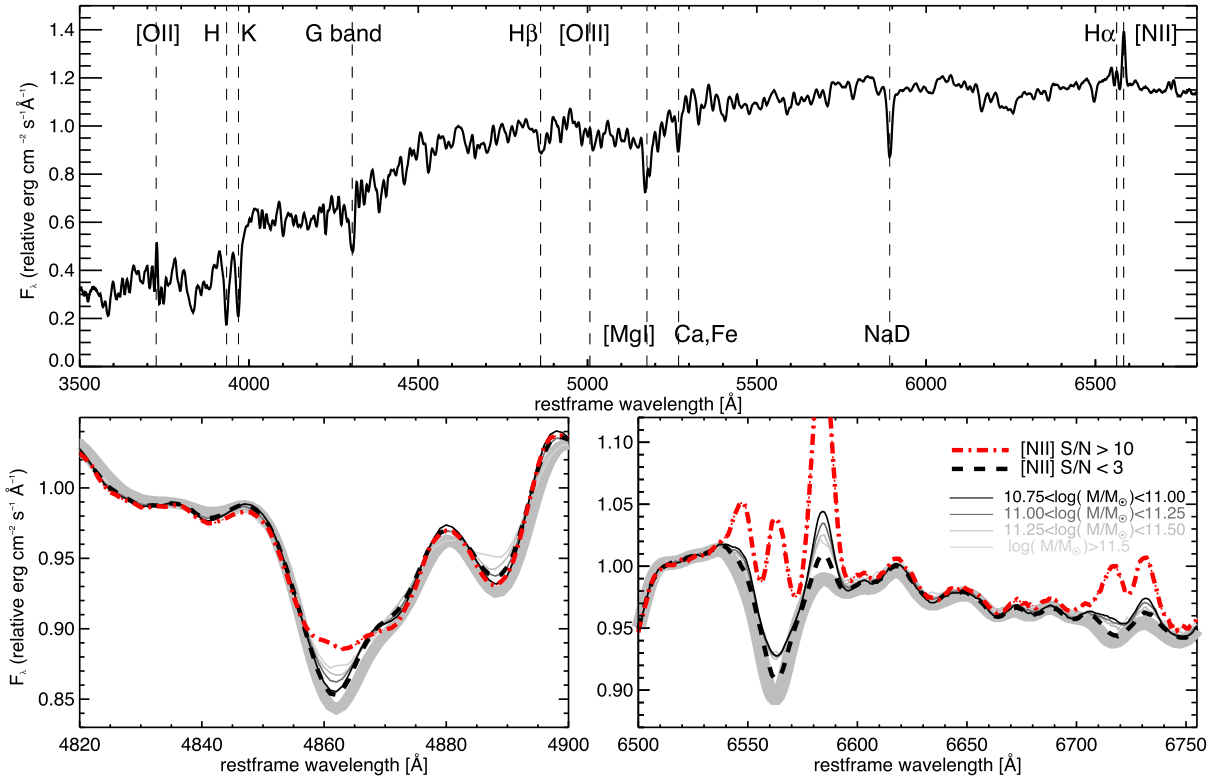

Figure 7. Top panel: stacked spectrum of the sample with [N II] λ6584 S/N > 10 (1145 galaxies). It shows some evident emission lines: [O II], [O III], Hα and [N II] λ6584. Bottom panels: comparison between the [N II] λ6584 S/N > 10 stacked spectrum (red dash–dotted lines), [N II] λ6584 S/N < 3 stacked spectra (black dashed lines) and the stacked spectra of the original sample in different mass subsamples (from black to light grey for BinI, BinII, BinIII and BinIV, respectively). In the Hβ region (left-hand panel), emission lines are not evident; instead, in the Hα region there are evident emissions ([N II] λλ6548, 6584, Hα, [S II] λλ6717, 6731). The grey area is the SSP solar model with age = 10 Gyr, σ = 250 km s^{-1} calculated by BC03 model.

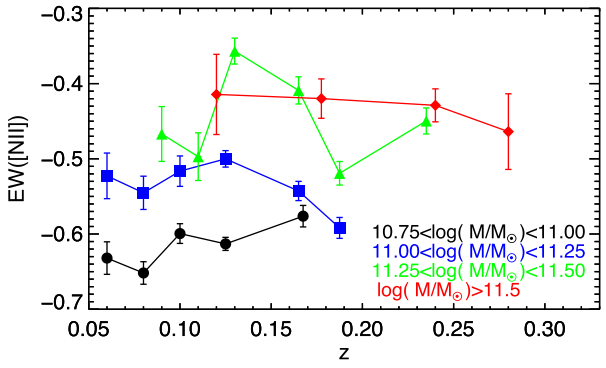

Figure 8. [N II] λ6584–redshift relations for the four mass subsamples (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV, respectively).

environment effects in the galaxy formation (Thomas et al. 2010) and many others. In these previous works, before interpreting the observed spectra, and therefore before using the stellar absorption line indices, standard procedures are followed to remove the contamination by nebular emission lines. In this section, we try to apply these cleaning methods both in the individual and in the stacked galaxy spectra.

The method proposed by Tremonti et al. (2004), for example, makes a non-negative least-squares fit of the emission line free observed regions of the spectrum, using a spectral library built using BC03 models. Then, the fitted spectrum is subtracted from the observed spectrum, and the residuals can be fitted to Gaussian broadened emission line templates. Finally, the fitted emission lines are subtracted from the original observed spectrum to produce a 'pure' absorption line spectrum (for further details, we refer to Tremonti et al. 2004). To verify if this method of decontamination is actually able to produce a clean sample in the Hβ region, we use Hβ values measured by the MPA-JHU DR7 group, which are available both for the original observed spectra and for the 'pure' absorption line spectra. We select a sample consisting of galaxies without emission lines, i.e. with no correction for emission lines based on the method of Tremonti et al. (2004), which contains 19 904 objects. We measure the median Hβ–z relations also for this subsample, with the same procedure as described in Section 3. The Hβ–z relations obtained are consistent with those found for the subsample with S/N([N II] λ6584) < 3. This suggests that, using objects with no emission lines detected by standard correction method, the offset between the observed Hβ values and the predicted ones is not totally removed.

In addition, we test the presence of weak Hβ emission line directly in the stacked spectra. We split our sample in terms of stellar mass and redshift as we have done in Section 3 for the Hβ values, and then we built the median spectra for each mass and redshift bin. In order to separate the stellar continuum from the nebular emission component, we decide to use a combination of the publicly available IDL codes: Penalized Pixel-Fitting (PPXF), developed by Cappellari & Emsellem (2004) and gas and absorption line fitting (GANDALF) written by Sarzi et al. (2006). We measure the line-of-sight velocity distribution by using PPXF code. Then, we perform the

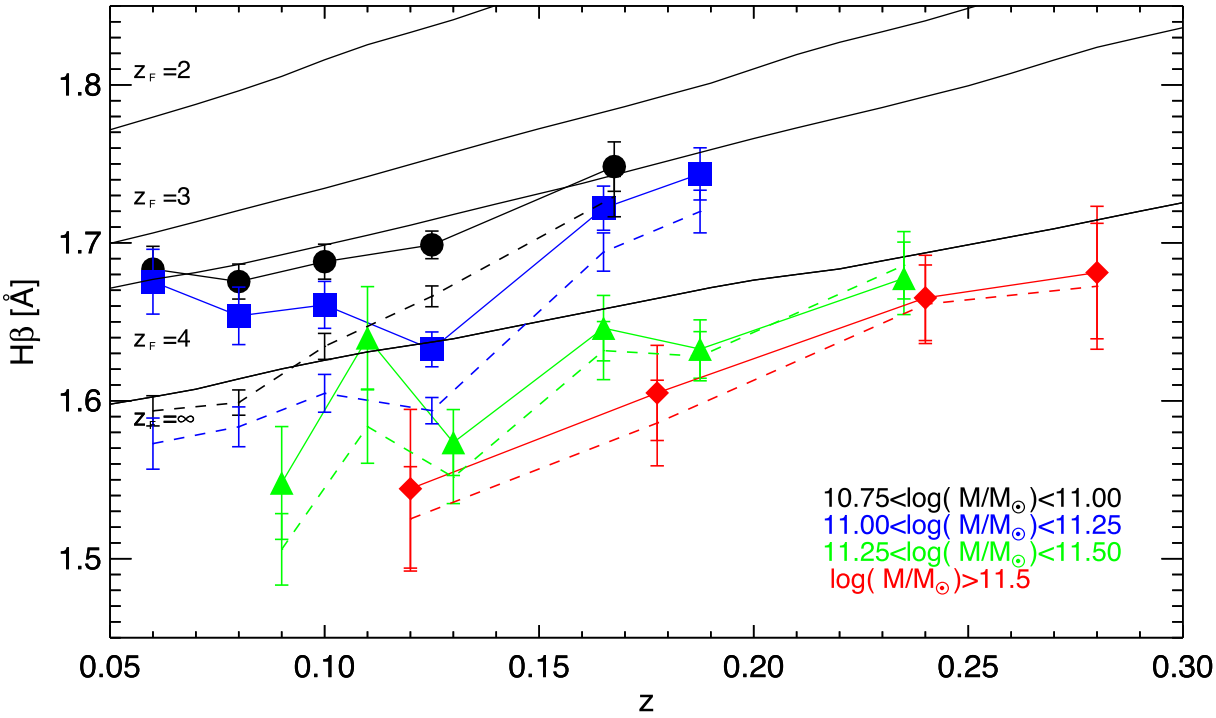

Figure 9. Comparison between Hβ–z relations (circles, squares, triangles and diamonds for BinI, BinII, BinIII and BinIV, respectively) of the total sample (dotted curves) and the sample with EW([N II] λ6584) S/N $<$ 3 (continuous curves).

GANDALF analysis to convolve a set of input synthetic spectra with the previous kinematic and to fit the observed stacked spectrum simultaneously with the models and a Gaussian emission line template. The result is a superposition of an optimal combination of the SSP templates with a set of Gaussians that represent the emission lines. Through the subtraction of the emission line spectrum from the observed one, we get the clean absorption line spectrum free from emission line contamination. First, we analyse the obtained Hβ emission lines and then we explore the impact of this cleaning method on the observed spectra. In order to test the robustness of the emission line extraction, we repeat the methodology by using three different spectral libraries: two based on the BC03 SSP[10] with different age $0.01 \leq t \leq 14$ Gyr (hereafter BC03_14 Gyr) and a more extended age $0.01 \leq t \leq 20$ Gyr (hereafter BC03_20 Gyr); and the third library built with MaStro models with age $0.01 \leq t \leq 14$ Gyr. We find that the recovered Hβ emission lines are model dependent: the EW is systematically higher for the BC03_20 Gyr library than the BC03_14 Gyr (ΔEW$\sim$0.02 Å) and the differences increase if we compare BC03 with MaStro models, ΔEW $\sim$0.2 Å. The median EWs are: EW $\sim -0.09 \pm 0.005$, $\sim -0.07 \pm 0.01$ and $\sim -0.27 \pm 0.01$ Å for BC03_14 Gyr, BC03_20 Gyr and MaStro library, respectively. This result is consistent with the study shown in Singh et al. (2013); their fig. 4 shows that at low Hβ fluxes there are differences in the flux extraction by using GANDALF code with SSP instead of the stellar templates. Therefore, the correction for very low emission in the absorption Hβ line depends on the templates library used in the continuum fit. Furthermore, we find that, in all the three libraries, the Hβ emission lines are detected with a very low confidence level, and the amplitude over noise is always A/N $\leq$ 3 (the noise is defined as the dispersion of fluctuations in the fit residuals).

[10] The BC03 and MaStro stellar population templates have a Chabrier IMF and metallicity $Z/Z_{\odot} = 0.2, 0.4, 1, 2.5$ and $Z/Z_{\odot} = 0.5, 1, 2$, respectively.

Finally, we test the more recent emission correction for the hydrogen features proposed by Serven & Worthey (2010). Serven & Worthey (2010) derived emission corrections of the Balmer series Lick indices for an SDSS quiescent galaxy spectral sample (Graves et al. 2007) by comparing the Hα–Mg b diagram from the SDSS stacked spectra with the measurements obtained for 13 Virgo galaxies. By using the same prescriptions shown in Serven & Worthey (2010), we re-calibrate the Hα–Mg b diagram and then the Hβ emission correction in our stacked spectra. We find that for the more massive sample (BinIV), the mean Mg b value is 4.23 ± 0.15 Å and the mean Hβ emission correction factor is of the order of -0.17 ± 0.08 Å, that it is in good agreement with the emission values obtained with the GANDALF method mentioned before. However, we stress that, also in this case, the method is model dependent, since the continuum correction was determined by using the Worthey et al. (1994) and Trager et al. (1998) models (see section 2 of Serven & Worthey 2010). Moreover, rescaling the Hα–Mg b relation from the SDSS values to the Virgo data could be age dependent if the two samples have different mean ages. We conclude that for our sample, also in the stacked spectra, it is very difficult to measure a well detected and model-age-indipendent Hβ emission line and than correct our observed spectra in order to obtain a 'pure' absorption Hβ line.

4.3 Hβ self-diluted emission line

In order to have a qualitative estimate of the hidden EW(Hβ) emission related to the observed [N II] λ6584 emission line, we study the emission line ratios of our spectra in the diagnostic diagrams. We adopt the classical (Baldwin, Phillips & Terlevich 1981) diagram, hereafter BPT diagram. The EW measurements for [O III], Hβ, Hα and [N II] λ6584 are taken by the MPA-JHU group analysis. In Fig. 10, we show the BPT diagrams for the subsample with all EWs detected with S/N $>$ 3, the subsample with EW([N II] λ6584) with S/N $>$ 10 and the subsample with EW([N II] λ6584) with S/N $<$ 3.

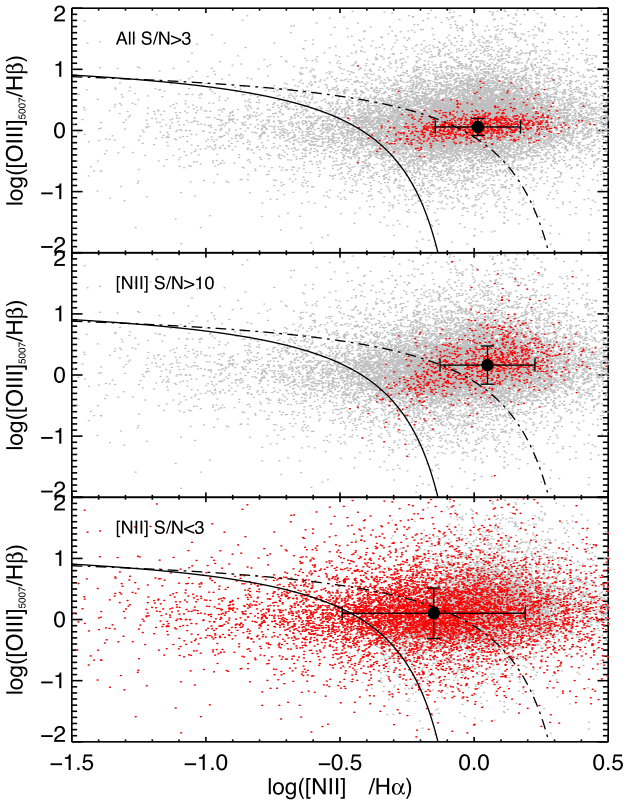

Figure 10. Distribution of the galaxies in the BPT diagram for the sample with all EW S/N > 3, EW([N II] λ6584) S/N > 10 and EW([N II] λ6584) S/N < 3, respectively, from the top downwards. The grey cloud is the total sample. The dashed curve is the theoretical demarcation of Kewley et al. (2001), which separates star-forming galaxies and composites from AGN. The solid curve indicates the empirical division between pure star-forming galaxies from composite and AGN of Kauffmann et al. (2003). The median value and dispersion of the distribution are shown with the black symbols, in each subsample.

In the first two cases, we find that the majority of objects are located in the LINER (low-ionization nuclear emission line regions; Heckman 1980). For the first and the second samples, the median and dispersion for log(EW([N II])/EW(Hα)) are -0.014 ± 0.159 and 0.050 ± 0.177, and for log(EW([O III])/EW(Hβ)) are 0.058 ± 0.139 and 0.162 ± 0.312, respectively. In the last subsample, the dispersion in the data is larger because of the lower signal-to-noise ratio (see Fig. 10). In this last case, the median value of the sample is located in the composite region [log(EW([N II])/EW(Hα)) $\sim -0.150 \pm 0.341$ and log(EW([O III])/EW(Hβ)) $\sim 0.103 \pm 0.413$].

By using the median EW([N II] λ6584)/EW(Hα) values and the median EW([N II] λ6584), we estimate qualitatively the EW(Hα) contribution. Then, we estimate the EW(Hβ) by assuming no absorption by dust, and hence an emission line ratio, Hα/Hβ ∼2.86 (Osterbrock 1989), for electron density of $n = 100$ cm^{-3} and electron temperature $T_e = 10^4$ K.

In the sample with EW([N II] λ6584) with S/N < 3, the EW([N II] λ6584) is $\simeq -0.26 \pm 0.27$ Å. For this sample, we expect an EW(Hβ) emission in the range −0.28 to −0.06 Å [by considering the EW([N II] λ6584)/EW(Hα) dispersion], or weaker in the case of dust extinction.

The same conclusions can be drawn for the sample with the larger [N II] λ6584 emission lines (S/N([N II] λ6584) > 10). In this case, the median EW([N II] λ6584) is -1.99 ± 0.58 Å, and the interval of variability of log(EW([N II])/EW(Hα) is smaller, $\sim 0.050 \pm 0.177$. Then, we expect an Hβ emission line EW between −0.9 and −0.4 Å in the case of no absorption by dust, or weaker otherwise.

In order to verify whether the Hβ emission line values, expected in the two different samples, S/N([N II] λ6584) < 3 and > 10, can be effectively detected in the observed spectra, we perform some simulations, described in the following section.

4.4 Hβ emission line simulations

Starting from the theoretical SED (in which only the absorption component is taken into account) of SPS models (BC03), we simulate the presence of an emission line by adding a Gaussian component with variable amplitude and FWHM.

Since different physical processes can give rise to emission lines with different FWHM, we perform two different simulations.

(i) Simul A, lines with FWHM = 5 Å, mainly due to star formation. The analysis of a spectrum of a typical star-forming galaxy from SDSS has shown a typical FWHM of the emission lines of ∼5 Å.

(ii) Simul B, lines with FWHM = 9 Å. In the case of emission lines due to different processes than star formation, such as AGN activity, the lines are broader than the ones found for star-forming galaxies.

By changing the FWHM and the EW line in theoretical emission lines, we study the variation of the absorption line profile with the contamination of these different emissions. The results of the simulations (for the case of an SSP SED with age = 10 Gyr) are shown in Fig. 11. In both simulations (A and B), it is possible to identify a threshold of emission below which it is not possible to detect the Hβ emission component on the absorption line. We can appreciate the distortion of the Hβ line due to the introduction of an emission line, with increasing EW, only after EW ≤ -0.2 Å or EW ≤ -0.9 Å for Simul A or B, respectively. Emissions smaller than these values are not able to change the shape of the line, but, however, they cause an intensity decrease in the absorption line, producing a drastic variation in the Hβ Lick index measure, with $\Delta(\mathrm{H}\beta)_{\max} \sim 0.2$ Å or $\Delta(\mathrm{H}\beta)_{\max} \sim 0.9$ Å in Simul A and B, respectively. From this analysis, we can see that, for galaxy spectra with a resolution of the order of 3 Å FWHM, there is a threshold detection limit of the Hβ emission line. As seen in the previous section, the presence of [N II] λ6584 emission in our samples may indicate the presence of a weak Hβ emission line with EW ≥ -0.28 Å for the sample with S/N([N II] λ6584) < 3 and EW ≥ -0.9 Å for the sample with S/N([N II] λ6584) > 10. These values are near to the threshold limit identified by simulations, therefore consistent with not being detectable in emission in our spectra. However, this potential hidden emission line can explain the detected offset between observational data and theoretical models discussed in Section 3.2.

Performing the same simulations for SSP models with different ages (5 and 13 Gyr), we find that the detection limit for the Hβ emission line does not change significantly.

4.5 Ionization sources

We find that even the less contaminated passive galaxies (sample with EW([N II] λ6584) S/N < 3) may present an emission line contamination in the Hβ absorption line. As discussed in Section 4.3 using the BPT diagrams, we found that by selecting samples with higher S/N (sample with all lines S/N > 3 and sample with EW([N II]

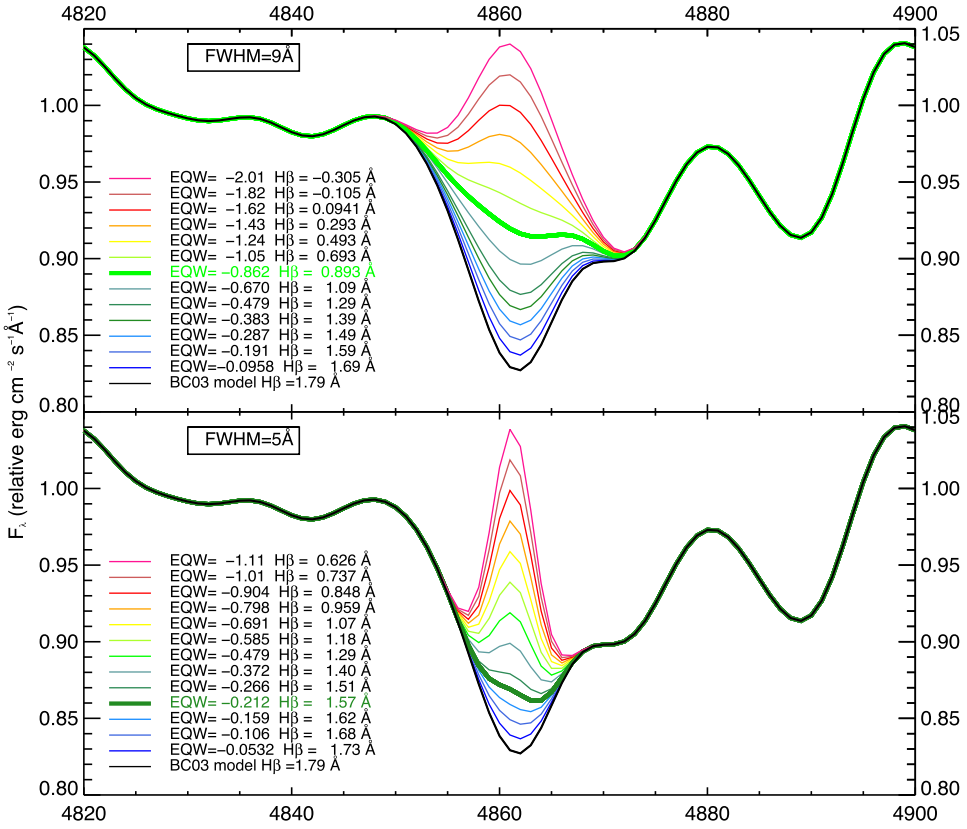

Figure 11. Hβ emission line simulations. We show the emission line effect within the Hβ absorption line (Simul A in the lower panel and Simul B in the upper panel). The different EW contaminations are plotted according with the labels. In both figures, we can identify a threshold limit above which it is possible to detect an emission contamination (thick curves). The absorption model is an SSP with $Z/Z_{\odot} = 1$, age = 10 Gyr and $\sigma = 200$ km s^{-1}.

λ6584) S/N > 10), our galaxies show typically LINER-like rather than star formation emission line ratios. These results are consistent with several imaging and spectroscopic observations (Phillips et al. 1986; Goudfrooij 1999; Sarzi et al. 2006; Annibali et al. 2010; Yan & Blanton 2012), which show a strong indication of the presence of warm gas (typically with $T \sim 10^4$ K) in ETGs, with line ratios typically classified as LINERs. Despite the number of studies, two fundamental questions remain open: what is the origin of this interstellar medium (ISM) in the ETGs and what is the physical mechanism that determines its ionization?

Studies carried out on the kinematics of the gas, in most cases, show a misalignment between gas and stellar component, suggesting an external origin of the gas (Caon, Macchetto & Pastoriza 2000). This observational evidence is in agreement with recent estimates of the gas emission metallicity, which seems to indicate an external source (Annibali et al. 2010).

However, as reported in Sarzi et al. (2006), the angular momenta measured do not seem to be consistent with a purely external origin.

The nature of the mechanism of ionization in LINERs is still not clear, but currently, the most accredited are photoionization by AGN activity (e.g. Ho 1999; Ho et al. 2000; Kewley et al. 2006), photoionization by post-asymptotic giant branch (post-AGB) stars (Binette et al. 1994) and fast shocks (e.g. Dopita & Sutherland 1995). The AGN activity, in fact, produces a source of energetic photons (X-rays and UV) able to ionize the ISM; this is confirmed by some observed bubbles in LINER galaxies (Baldi et al. 2009). For some objects, the presence of AGN activity was also confirmed by observations of the radio core and X-rays point sources in their centres and UV variability (e.g. Ho 2008). Nevertheless, recent works have shown that in many cases the gas emissions are not only neighbouring to the centre, but an extended emission component also exists (Sarzi et al. 2006; Annibali et al. 2010; Yan & Blanton 2012; Singh et al. 2013). For this reason, the AGNs are not considered as the main cause of the LINER-like emission lines in ETGs in favour of spatially extended ionizing sources, which in some cases follow the profile of stellar density (Yan & Blanton 2012). These properties are found in another source proposed for the first time by Binette et al. (1994): the post-AGB stars in the old stellar populations. These stars, after the AGB phase, are often surrounded by material ejected during their thermal pulses. After that, this material is dispersed in the ISM, and the AGB stars, having very high temperatures ($\sim 10^5$ K), are capable of producing a diffuse radiation field. Recent spatially resolved spectroscopy observations obtained with a prototype of Mapping Nearby Galaxies at Apache Point Observatory (MaNGA) instrument confirm the presence of an extended LINER-like emission associated with spectral features of old and metal-rich stars (see Belfiore et al. 2015). This scenario is also supported by the fact that post-AGBs are able to reproduce the variation of the parameter of ionization (defined as the ratio of the ionizing photon flux density to the electron density) with the radius of the galaxy (Yan & Blanton 2012). As expected in Binette et al. (1994) and Cid Fernandes et al. (2009), the post-AGBs are able to produce weak emission lines: H$\alpha \sim -0.6, -1.7$ Å for a stellar population of 8 and 13 Gyr, respectively. This is very interesting because if this emission is expressed in terms of Hβ emissions, similar to what was seen in Section 4.3, we find a measure of EW(Hβ)$\sim -$ 0.2, -0.6 Å emission line perfectly consistent with what is expected for our sample, especially with regard to the subsample with EW([N II]

λ6584) S/N < 3 Å. However, as shown in Yan & Blanton (2012), the post-AGB stars, while representing the ideal candidate for the creation of a photoionizing source, according to the most recent models, produce an ionization parameter too small compared to what is required by the observations. It should be noted that the current knowledge of the number density of post-AGB stars is still very uncertain, and therefore the question if these stars may be the main cause of emissions in older populations remains open. Finally, the emission lines with LINER-like ratios can be produced by fast shock. However, this mechanism cannot be mainly responsible for the emission in ETGs, since, as suggested in Yan & Blanton (2012), it provides emission lines with similar velocities between the different elements, contrary to what emerges from the recent estimates of velocity in different lines.

The data analysed here do not give a direct indication of how to disentangle the three possible photoionization source candidates. However, the observed anti-correlation between the [N II] emission and the mass (see Section 4.1, Fig. 8) could disfavour the AGN hypothesis as a photoionization source, since its presence should instead be correlated with the galaxy mass (see Best et al. 2005; Brusa et al. 2009).

Finally, it is particularly important that the detection of emission lines in our sample seems to suggest a more frequent presence of ionized gas in these galaxies. This observation has significant effects on knowledge of this particular class of galaxies. The ETGs are not the simplest systems, but they are more complex in which only knowledge of both the stellar contribution and of the ISM (and of the link between the two components) can shed light on their formation history and evolution.

5 CONCLUSIONS

In this paper, we explore the properties of the Hβ Lick index of massive and passive ETGs, to estimate its robustness as an age indicator, being the Lick index more sensitive to the stellar population age and less affected by stellar metallicity, as suggested since the pioneering work of Worthey et al. (1994). The aim of this work is to establish its reliability as 'cosmic chronometer' to trace the age evolution of the Universe as a function of redshift and to provide new constraints on the age of formation and evolution of galaxies' stellar populations, eventually allowing measurement of the Hubble parameter $H(z)$ through the 'cosmic chronometers' approach (Jimenez & Loeb 2002; Moresco et al. 2012b, 2016a; Moresco 2015).

Using photometric and spectroscopic information, we select the most massive, passive and red ETGs in the SDSS-DR6 survey. The final sample consists of about 23 914 galaxies with stellar mass $\log(M/\mathrm{M}_\odot > 10.75)$ in the redshift range $0.05 \leq z \leq 0.3$. We divide the sample into four mass subsamples ($\Delta\log(M/\mathrm{M}_\odot) = 0.25$) in order to avoid possible biases due to mass-downsizing effect, and obtained four homogeneous subsamples in redshift of formation. All the samples are further divided into redshift bins and the spectra analysed to measure the Hβ Lick index, obtaining a median Hβ–z relation in each mass subsample.

The main results of this analysis may be summarized as follows.

(i) Despite the rather strict selection criteria, all the spectra present characteristic features of a passive population, we find in the median stacked spectra of each mass bin clear evidence of a weak [N II] emission line, with EW of the order of −0.6 Å, which can be interpreted as a hint of the presence of ionized gas. This residual emission line contamination may have a significant impact on the Hβ Lick index measures. To address this issue, we split our samples on the base of the [N II] line, having a purer sample of 13 626 galaxies selected with SN([N II]) < 3, and a more contaminated sample with SN([N II]) > 10. The analysis of these two subsamples confirmed the first hypothesis of a possible hidden contamination, with the sample with SN([N II]) > 10 presenting clear emission lines also in Hα, [O III] and Hβ. We also find that the amplitude of the [N II] emission line anti-correlates with stellar mass.

(ii) The analysis of all mass subsamples reveals a clear evolution of the Hβ Lick index as a function of redshift, with the index decreasing with cosmic time. These trends are qualitatively consistent with a passive evolution, and are proven to be independent of many effects that can affect the analysis. These trends are also confirmed on both the original samples and on the 'purest' one obtained after selecting only galaxies with SN([N II]) < 3.

(iii) At each redshift, more massive galaxies present a median Hβ index lower than less massive ones, confirming a mass-downsizing scenario for which more massive systems have assembled their stars earlier and faster. This result has been demonstrated not to depend on a selection effect due to the different velocity dispersions and metallicities of the samples; also in this case, this trend is found in both the original and the 'purest' samples.

(iv) The comparison with SPS models highlights an inconsistency with observable data, for which observed galaxies appear, in most cases, to be older than the age of the Universe at the given redshifts. These differences are greater than $\Delta\mathrm{H}\beta \sim 0.1$ Å and do not seem to be reconcilable with any possible further re-selection of the samples. Only a stellar metallicity systematically higher than the one found in these samples ($Z/\mathrm{Z}_\odot \sim 2$) may alleviate the tension between the data and models.

(v) We tested the presence of a weak Hβ emission line in our stacked spectra by using GANDALF code and the emission corrections for the hydrogen features proposed by Serven & Worthey (2010). We find that the recovered emission lines are very uncertain, detected with a very low confidence level and are model dependent. The median emission EW ranges from ∼−0.1 to ∼−0.27 Å.

(vi) We also find that in the stacked spectrum obtained from the 'purest' sample with SN([N II]) < 3, there is a residual [N II] emission line contamination, even if very weak (EW ∼−0.26).

(vii) Throughout simulations, we demonstrate that there exists a threshold limit below which an emission line component within the Hβ absorption features would not be detectable; this threshold depends on the FWHM of the line, being EW(Hβ) ≤ −0.9 Å or ≤−0.2 Å, respectively, for a broader and narrower FWHM (characteristic of star formation or AGN activity). We note that this EW is compatible with the observed offset between models and data.

In order to obtain a quantitative estimate of the age of formation of the galaxies and to test the feasibility of using this index as a 'cosmic chronometers' indicator, the observed Hβ–z relations should be calibrated on SPS models. However, the dependence of the normalization and of the slope of these relations on the different possible selections, and the apparent inconsistency with theoretical models, makes Hβ an index that is difficult to rely on to estimate both absolute and relative ages, and this, as discussed, is due to a possible contamination of the line by an undetectable emission component. We discuss the possible candidates of ionization sources, finding a better agreement with post-AGB and slightly higher tension with AGN.

Despite that this index has been historically identified as the best suited to constrain the age of a galaxy population, all the highlighted issues do not allow us to use it for an accurate estimate of the

absolute and differential age evolution of the stellar populations with redshift. Therefore, it is not possible to place constraints on the galaxies' redshift of formation, and even the use of the Hβ–z relation in the 'cosmic chronometers' approach appears inadequate.

Finally, higher SN and resolution spectra may help in mitigating these problems during the selection phase, and to better disentangle a narrow emission line component for a less biased measurement. Another possible option is to study higher order Balmer lines (Hγ, Hδ), as e.g. suggested by the work of Vazdekis & Arimoto (1999), since those lines should be less affected by an underlying emission component. This analysis will be further exploited in a following paper.

ACKNOWLEDGEMENTS

We would like to thank the referee, Guy Worthey, for useful suggestions. We also thank Marco Mignoli, Claudia Maraston, Daniel Thomas, Jonas Johansson, Marcella Brusa and Gianni Zamorani for the stimulating discussions. We acknowledge Anna Gallazzi and Jarle Brinchmann for the SDSS data availability and their clarifications. AC is also grateful to Annalisa Citro and Salvatore Quai for sharing a preview of their results on SDSS stacked spectra analysis. Part of this work was supported by the INAF, Osservatorio Astronomico di Bologna and by the Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna. We acknowledge the grants ASI n. I/023/12/0 'Attività relative alla fase B2/C per la missione Euclid' and PRIN MIUR 2010-2011 'The dark Universe and the cosmic evolution of baryons: from current surveys to Euclid'. ACi, LP and MM acknowledge the PRIN MIUR 2015 'Cosmology and Fundamental Physics: illuminating the Dark Universe with Euclid'.

REFERENCES

Annibali F., Bressan A., Rampazzo R., Zeilinger W. W., Vega O., Panuzzo P., 2010, A&A, 519, A40
Baldi A., Forman W., Jones C., Kraft R., Nulsen P., Churazov E., David L., Giacintucci S., 2009, ApJ, 707, 1034
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Beifiori A., Maraston C., Thomas D., Johansson J., 2011, A&A, 531, A109
Belfiore F., Maiolino R., Bundy K., Thomas D., 2015, MNRAS, 449, 867
Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS 362, 25
Binette L., Magris C. G., Stasińska G., Bruzual A. G., 1994, A&A, 292, 13
Borch A. et al., 2006, A&A, 453, 869
Brusa M. et al., 2009, A&A, 507, 1277
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bundy K., Ellis R. S., Conselice C. J., 2005, ApJ, 625, 621
Bundy K., Ellis R. S., Conselice C. J., 2006, ApJ, 651, 120
Burstein D., Faber S. M., Gaskell C. M., Krumm N., 1984, ApJ, 287, 586
Caon N., Macchetto D., Pastoriza M., 2000, ApJS, 127, 39
Cappellari M., Emsellem E., 2004, PASP, 116, 138
Carson D. P., Nichol R. C., 2010, MNRAS, 408, 213
Choi J., Conroy C., Moustakes J., Graves G. J., Holden B. P., Brodwin M., Brown M. J. I., van Dokkum P. G., 2014, ApJ, 792, 95
Cid Fernandes R., Schlickmann M., Stasinska G., Asari N. V., Gomes J. M., Schoenell W., Mateus A., Sodré L., Jr., 2009, in Wang W., Yang Z., Luo Z., Chen Z., eds, ASP Conf. Ser. Vol. 408, The Starburst-AGN Connection. Astron. Soc. Pac., San Francisco, p. 122
Cimatti A., Daddi E., Renzini A., 2006, A&A, 453, L29
Cimatti A. et al., 2008, A&A, 482, 21
Citro A., Pozzetti L., Moresco M., Cimatti A., 2015, A&A, 592, A19
Cowie L. L., Songaila A., Barger A. J., 1999, AJ, 118, 603
Dopita M. A., Sutherland R. S., 1995, ApJ, 455, 468
Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, MNRAS, 362, 41
Goudfrooij P., 1999, in Carral P., Cepa J., eds, ASP Conf. Ser. Vol. 163, Star Formation in Early Type Galaxies. Astron. Soc. Pac., San Francisco, p. 55
Graves G. J., Faber S. M., Schiavon R. P., Yan R., 2007, ApJ, 671, 243
Hanuschik R. W., 2003, A&A, 407, 1157
Heavens A., Panter B., Jimenez R., Dunlop J., 2004, Nature, 428, 625
Heckman T. M., 1980, A&A, 87, 152
Ho L. C., 1999, ApJ, 510, 631
Ho L. C., 2008, ARA&A, 46, 475
Ho L. C., Rudnick G., Rix H.-W., Shields J. C., McIntosh D. H., Filippenko A. V., Sargent W. L. W., Eracleous M., 2000, ApJ, 541, 120
Hoaglin D. C., Mosteller F., Tukey J. W., 1983, Wiley Series in Probability and Mathematical Statistics. Wiley, New York, p. 404
Ilbert O. et al., 2006, A&A, 457, 841
Ilbert O. et al., 2010, ApJ, 709, 644
Jimenez R., Loeb A., 2002, ApJ, 573, 37
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
Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961
Lee H.-c., Worthey G., 2005, ApJS, 160, 176
Maraston C., Strömbäck G., 2011, MNRAS, 418, 2785
Maraston C., Pforr J., Henriques B. M., Thomas D., 2013, MNRAS, 435, 2764
Moresco M., 2015, MNRAS, 450, L16
Moresco M., Jimenez R., Cimatti A., Pozzetti L., 2011, J. Cosmol. Astropart. Phys., 3, 45
Moresco M., Verde L., Pozzetti L., Jimenez R., Cimatti A., 2012a, J. Cosmol. Astropart. Phys., 7, 53
Moresco M. et al., 2012b, J. Cosmol. Astropart. Phys., 8, 6
Moresco M. et al., 2013, A&A, 558, A61
Moresco M. et al., 2016a, J. Cosmol. Astropart. Phys., 05, 014
Moresco M., Jimenez R., Verde L., Cimatti A., Pozzetti L., Maraston C., Thomas D., 2016b, J. Cosmol. Astropart. Phys., 12, 39
Morton D. C., 1991, ApJS, 77, 119
Osterbrock D. E., 1989, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. Univ. Sci. Books, Mill Valley, CA
Phillips M. M., Jenkins C. R., Dopita M. A., Sadler E. M., Binette L., 1986, AJ, 91, 1062
Pozzetti L. et al., 2010, A&A, 523, A13
Sarzi M. et al., 2006, MNRAS, 366, 1151
Serven J., Worthey G., 2010, AJ, 140, 152
Simon J., Verde L., Jimenez R., 2005, Phys. Rev. D, 71, 123001
Singh R. et al., 2013, A&A, 558, A43
Spinrad H., Dey A., Stern D., Dunlop J., Peacock J., Jimenez R., Windhorst R., 1997, ApJ, 484, 581
Stern D., Jimenez R., Verde L., Kamionkowski M., Stanford S. A, 2010, J. Cosmol. Astropart. Phys., 2, 8
Thomas D., Maraston C., Bender R., Mendes de Oliveira C., 2005, ApJ, 621, 673
Thomas D., Maraston C., Schawinski K., Sarzi M., Silk J., 2010, MNRAS, 404, 1775
Thomas D., Maraston C., Johansson J., 2011, MNRAS, 412, 2183
Trager S. C., Worthey G., Faber S. M., Burstein D., González J. J., 1998, ApJS, 116, 1
Tremonti C. et al., 2004, ApJ, 613, 898
Vazdekis A., Arimoto N., 1999, ApJ, 525, 144
Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, ApJS, 94, 687
Yan R., Blanton M. R., 2012, ApJ, 747, 61
Zhang C., Zhang H., Yuan S., Liu S., Zhang T.-J., Sun Y.-C., 2014, Res. Astron. Astrophys., 14, 1221
Zucca E. et al., 2006, A&A, 455, 879

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