

The cosmic evolution of the stellar mass–velocity dispersion relation of early-type galaxies

Carlo Cannarozzo^{1,2*}, Alessandro Sonnenfeld³ and Carlo Nipoti¹

¹*Dipartimento di Fisica e Astronomia, Alma Mater Studiorum Università di Bologna, via Piero Gobetti 93/2, I-40129 Bologna, Italy*

²*INAF - Osservatorio di Astrofisica e Scienza dello Spazio di Bologna, Via Piero Gobetti 93/3, I-40129 Bologna, Italy*

³*Leiden Observatory, Leiden University, Niels Bohrweg 2, 2333 CA Leiden, The Netherlands*

Submitted, October 12, 2019

ABSTRACT

Taking advantage of a Bayesian hierarchical inference formalism, we study the evolution of the observed correlation between central stellar velocity dispersion σ_0 and stellar mass M_* of massive ($M_* \gtrsim 10^{10.5} M_\odot$) early-type galaxies (ETGs) out to redshift $z \approx 2.5$. Collecting ETGs from state-of-the-art literature samples, we consider both a *fiducial* ($0 \lesssim z \lesssim 1.2$) and an *extended* ($0 \lesssim z \lesssim 2.5$) samples. Based on the fiducial sample, we find that at $z \lesssim 1.2$ the $M_* - \sigma_0$ relation is well represented by $\sigma_0 \propto M_*^\beta (1+z)^\zeta$, with $\beta \approx 0.23$ (at given z , σ_0 increases with M_* as a power law with slope similar to the classical Faber-Jackson relation) and $\zeta \approx 0.26$ (at given M_* , σ_0 decreases for decreasing z , for instance by $\approx 20\%$ from $z = 1$ to $z = 0$). The analysis of the extended sample suggests that over the wider redshift range $0 \lesssim z \lesssim 2.5$ the slope β could be slightly steeper at higher z ($\beta \approx 0.2 + 0.2 \log(1+z)$ describes the data as well as constant $\beta \approx 0.23$) and that the redshift-dependence of the normalisation could be stronger ($\zeta \approx 0.42$). The intrinsic scatter in σ_0 at given M_* is ≈ 0.07 dex for the fiducial sample at $z \lesssim 1.2$, but is found to increase with redshift in the extended sample, being ≈ 0.11 dex at $z = 2$. Our results suggest that, on average, the velocity dispersion of *individual* massive ($M_* \gtrsim 5 \times 10^{11} M_\odot$) ETGs decreases with time while they evolve from $z \approx 1$ to $z \approx 0$.

Key words: galaxies: elliptical and lenticular, cD – galaxies: evolution – galaxies: formation – galaxies: fundamental parameters – galaxies: kinematics and dynamics

1 INTRODUCTION

Since the late 1970s it was found empirically that present-day early-type galaxies (ETGs) follow scaling relations, i.e. correlations among global observed quantities, such as the Faber-Jackson relation (Faber & Jackson 1976) between luminosity L and central stellar velocity dispersion σ_0 , the Kormendy relation (Kormendy 1977) between effective radius R_e and surface brightness (or luminosity), and the fundamental plane (Djorgovski & Davis 1987; Dressler et al. 1987) relating σ_0 , L and R_e . When estimates of the stellar masses are available, analogous scaling relations are found, replacing L with M_* : the $M_* - R_e$ (stellar mass–size) relation, the $M_* - \sigma_0$ (stellar mass–velocity dispersion) relation and the stellar-mass fundamental plane (e.g., Hyde & Bernardi 2009a,b; Auger et al. 2010; Zahid et al. 2016b). These scaling laws are believed to contain valuable information on the process of formation and evolution of ETGs. Any success-

ful theoretical model of galaxy formation should reproduce these empirical correlations of the present-day population of ETGs (Somerville & Davé 2015; Naab & Ostriker 2017).

The observations strongly indicate that ETGs are not evolving passively. For instance, measurements of sizes and stellar masses of samples of quiescent galaxies at higher redshift imply that the $M_* - R_e$ relation evolves with time: on average, for given stellar mass, galaxies were significantly more compact in the past (e.g. Ferguson et al. 2004; Damjanov et al. 2019). There are also indications that ETGs at higher redshift have, on average, higher stellar velocity dispersion than present-day ETGs of similar M_* (e.g. van de Sande et al. 2013b; Belli et al. 2014; Gargiulo et al. 2016; Belli et al. 2017; Tanaka et al. 2019). Interestingly, the stellar-mass fundamental plane, relating M_* , σ_0 and R_e appears to change little with redshift (Bezanson et al. 2013b, 2015; Zahid et al. 2016a). The observed behaviour of these scaling relations as a function of redshift represents a further challenge to models of galaxy formation and evolution.

In the standard cosmological framework, structure for-

* E-mail: carlo.cannarozzo3@unibo.it

mation in the Universe occurs as a consequence of the collapse and virialisation of the dark matter halos, in which baryons infall and collapse, thus forming galaxies. In this framework, massive ETGs are believed to be the end products of various merging and accretion events. Given the old ages of the stellar populations of present-day ETGs, any relatively recent merger that these galaxies experienced must have had negligible associated star formation. Based on these arguments, a popular scenario for the late ($z \lesssim 2$) evolution of ETGs is the idea that these galaxies grow via dissipationless (or “dry”) mergers. Interestingly, dry mergers make galaxies less compact: for instance, galaxies growing via parabolic dry merging increase their size as $R_e \propto M_*^a$, with $a \gtrsim 1$, while their velocity dispersion evolves as $\sigma_0 \propto M_*^b$, with $b \lesssim 0$ (Nipoti et al. 2003; Naab et al. 2009; Hilz et al. 2013). Thus, the transformation of individual ETGs via dry mergers is a possible explanation of the observed evolution of the $M_* - R_e$, $M_* - \sigma_0$ and stellar-mass fundamental plane relations (Nipoti et al. 2009b, 2012; Posti et al. 2014; Oogi & Habe 2013; Frigo & Balcells 2017). Though this explanation is qualitatively feasible, it is not clear whether and to what extent dry mergers can explain quantitatively the observed evolution of these scaling laws. In this context, the stellar velocity dispersion σ_0 is a very interesting quantity to consider. Even for purely dry mergers of spheroids, σ_0 can increase, decrease or stay constant following a merger, depending on the merger mass ratio and orbital parameters (Boylan-Kolchin et al. 2006; Naab et al. 2009; Nipoti et al. 2009a, 2012; Posti et al. 2014). Moreover, even slight amounts of dissipation and star formation during the merger can produce a non-negligible increase of the central stellar velocity dispersion with respect to the purely dissipationless case (Robertson et al. 2006; Ciotti et al. 2007).

In a cosmological context, the next frontier in the theoretical study of the scaling relations of ETGs is the comparison with observations of the evolution measured in hydrodynamic cosmological simulations. A quantitative characterisation of the evolution of the observed scaling relations of the ETGs is thus crucial to use them as test beds for theoretical models. On the one hand, the evolution of the observed stellar mass–size relation is now well established, being based on relatively large samples of ETGs out to $z \approx 3$ (Cimatti et al. 2012; van der Wel et al. 2014). On the other hand, given that measuring the stellar velocity dispersion requires spectroscopic observations with relatively high resolution and signal-to-noise ratio, the study of the redshift evolution of correlations involving σ_0 , such as the $M_* - \sigma_0$ relation and the stellar-mass fundamental plane, is based on much smaller galaxy samples than those used to study the stellar mass–size relation. This makes it more difficult to characterise quantitatively the evolution of these scaling laws out to significantly high redshift.

In this paper, we focus on the stellar mass–velocity dispersion relation of ETGs with the aim of improving the quantitative characterisation of the observed evolution of this scaling law. We build an up-to-date sample of massive ETGs with measured stellar mass and stellar velocity dispersion by collecting and homogenising as much as possible available state-of-the-art literature data. In particular, we consider galaxies with stellar masses higher than $10^{10.5} M_\odot$ and we correct the observed stellar velocity dispersion to σ_e , the central line-of-sight stellar velocity dispersion within an

aperture of radius R_e , so in our case $\sigma_0 = \sigma_e$. We analyse statistically the evolution of the $M_* - \sigma_e$ relation without resorting to binning in redshift and using a Bayesian hierarchical approach. As a result of this analysis we provide the posterior distributions of the hyper-parameters describing the $M_* - \sigma_e$ relation in the redshift range $0 \lesssim z \lesssim 2.5$, under the assumption that, at given redshift, $\sigma_e \propto M_*^\beta$. We explore both the case of redshift independent β and the case in which β is free to vary with redshift.

The paper is organised as follows. Section 2 describes the galaxy sample and the criteria adopted to select ETGs. We present the statistical method in section 3 and our results in section 4. Our results are compared with previous works in section 5. Section 6 concludes. Throughout this work, we adopt a standard Λ cold dark matter cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. All stellar masses are calculated assuming a Chabrier (2003) initial mass function (IMF).

2 GALAXY SAMPLE

To study the evolution of the stellar mass–velocity dispersion relation of ETGs we build a sample of galaxies consisting in a collection of various subsamples of ETGs in the literature. Our goal is to build a sample spanning a redshift range as large as possible. At the same time, in order to make an accurate inference, it is important to select galaxies and measure their stellar masses and velocity dispersions in a homogeneous way. These two requirements are somewhat conflicting: in order to minimize biases, it would be ideal to use the same selection criteria, data and analysis methods for all galaxies in our study. This, however, is difficult to achieve in practice, due to the lack of large samples of galaxies with measured stellar velocity dispersion at $z > 1$.

In light of this challenge, we find it useful to build our sample as the sum of two distinct samples, each optimized for meeting one of our two criteria. The first is our *fiducial* sample, consisting of quiescent galaxies drawn from the Sloan Digital Sky Survey (SDSS; Eisenstein et al. 2011) and the Large Early Galaxy Astrophysics Census (LEGAC; van der Wel et al. 2016), which cover the redshift range $0 \lesssim z \lesssim 1.2$. For the galaxies in this sample we strictly apply consistent selection criteria, then measure their stellar masses using photometric data from the first data release of the Hyper Suprime-Cam (HSC; Miyazaki et al. 2018) Subaru Strategic Program (Aihara et al. 2018, DR1). The second is a *high-redshift* sample, consisting of stellar mass and velocity dispersion measurements of galaxies at $0.8 \lesssim z \lesssim 2.5$ from various independent studies. For the galaxies in this high-redshift sample, we only require that the definitions of stellar mass and stellar velocity dispersion are the same as those of the fiducial sample. In both cases, we apply a lower limit to the measured stellar mass of $\log(M_*^{\text{obs}}/M_\odot) > 10.5$. We also define an *extended* sample, obtained by combining the fiducial and high-redshift samples.

Our strategy is to carry out our inference on both the fiducial and the extended samples. Given the way the samples are built, we expect our results at $z < 1$ to be more robust (i.e. less prone to observational biases), but it is nevertheless very interesting to examine trends out to $z \approx 2.5$, as probed by our extended sample. In the following two sub-

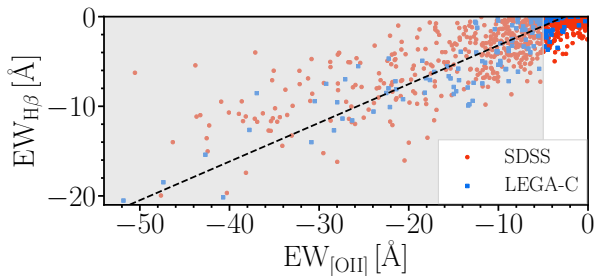


Figure 1. Equivalent width $EW_{H\beta}$ of $H\beta$ as a function of equivalent width $EW_{[OII]}$ of $[OII]$ for galaxies drawn from the original catalogues of SDSS (red dots) and LEGA-C (blue squares). For LEGA-C galaxies, we select objects with signal-to-noise ratio > 10 . The black dashed line represents a linear fit to the data. Galaxies in the shaded region of the diagram ($EW_{[OII]} < -5$) are excluded from our sample of ETGs.

sections we describe in detail how measurements for these samples are obtained.

2.1 The fiducial sample

Our fiducial sample consists of two sets of galaxies. The first set is drawn from the data release 12 (DR12; Alam et al. 2015) of the SDSS. In particular, we consider only objects belonging to the main spectroscopic sample, by applying the condition `programname='legacy'` or `programname='southern'`. The second set is selected from the LEGA-C survey DR2 (Straatman et al. 2018). The LEGA-C DR2 contains spectra of 1,922 objects obtained with the Visible Multi-Object Spectrograph (VIMOS; Le Fèvre et al. 2003) on the Very Large Telescope (VLT). LEGA-C targets were selected by applying a cut in K_s -band magnitude to a parent sample of galaxies with photometric redshift in the range $0.6 < z < 1.0$ drawn from the Ultra Deep Survey with the VISTA telescope (UltraVISTA; Muzzin et al. 2013).

With the goal of selecting only passive galaxies, we apply a selection based on the equivalent width of the forbidden emission line doublet of $[OII]$, $EW_{[OII]}\lambda\lambda 3726, 3729$. Specifically, we include only those galaxies that have $EW_{[OII]} \geq -5 \text{ \AA}$, where $EW_{[OII]}$ of SDSS and LEGA-C galaxies are obtained from the respective data release catalogues. Although $[OII]$ is not a perfect indicator of star formation activity, as it can suffer from contamination from emission by an active galactic nucleus, and other spectral lines could be used in its place ($H\beta$, for example), these lines are in general not accessible in the spectra of most LEGA-C galaxies, as they are redshifted outside of the available spectral range. For the sake of homogeneity in our selection criteria, and in order to keep the high end of the redshift distribution of the LEGA-C galaxies in our sample, we use $[OII]$ as a first step towards obtaining a sample of quiescent galaxies. Nevertheless, we find a good correlation between $EW_{[OII]}$ and $EW_{H\beta}$ for those galaxies drawn from the original catalogues of SDSS and LEGA-C for which both measurements are available (see Figure 1). Although half of the LEGA-C galaxies do not have values of $EW_{[OII]}$ in the DR2 catalogue, these are for the most part objects at the low end of the redshift range, $z < 0.8$.

We then look for imaging data in the Wide layer of the HSC DR1. The Wide layer of HSC covers approximately

108 square degrees. The number of SDSS galaxies present in this dataset is ≈ 3000 , which, while only a small fraction of the total number of SDSS galaxies, is still sufficiently large to carry out a statistical analysis of the stellar mass-velocity dispersion relation. LEGA-C targets are located in a $\approx 1.3 \text{ deg}^2$ region, for the most part overlapping with the Cosmic Evolution Survey (COSMOS; Scoville et al. 2007) area. HSC DR1 data from the Ultra Deep layer are available for most (≈ 1700) of the objects in the LEGA-C DR2.

The motivation for using HSC data is in its high depth (i -band 26 mag detection limit for a point source in the Wide layer) and good image quality (typical i -band seeing is $0.6''$). This is particularly important for the LEGA-C galaxies, which are much fainter and have smaller angular sizes compared to the SDSS ones, due to their higher redshift. For each galaxy with available HSC DR1 data, we obtain cutouts in the g, r, i, z, y filters, then visually inspect them to remove objects showing any presence of discs, spiral arms, contamination by external objects, merging pairs, as well as galaxies for which a single Sérsic model (Sérsic 1968) does not provide a qualitatively good description of the surface-brightness distribution (e.g., irregular galaxies). Roughly 50% of the objects in both samples are rejected at this stage.

We measure stellar masses by, first, fitting a Sérsic surface brightness distribution to g, r, i, z, y imaging data, and, then, fitting stellar population synthesis (SPS) models to the resulting magnitudes, following the same procedure adopted by Sonnenfeld et al. (2019). In particular, we obtain 201×201 pixel ($\approx 34'' \times 34''$) sky-subtracted cutouts of each galaxy in each band, we fit the five-band data simultaneously with a seeing-convolved Sérsic surface brightness profile with elliptical isophotes and spatially uniform colors, while iteratively masking out foreground or background objects using the software SExtractor (Bertin & Arnouts 1996). For the sake of robustness in our flux measurements, we only keep galaxies for which the half-light radius of the best-fit Sérsic profile is smaller than $0.5R_{\text{sky},i}$, where $R_{\text{sky},i}$ is the radius of the i -band isophote with surface brightness equal to the RMS of the sky background. In other words, we only consider galaxies for which most of the flux accounted for by the model is actually detected in HSC data. This step removes $\approx 10\%$ of the objects: these are typically galaxies with a very large best-fit Sérsic index, $n_{\text{Ser}} \approx 10$.

We fit the observed *grizy* fluxes with composite stellar population models, obtained using the BC03 stellar population synthesis code (Bruzual & Charlot 2003), with semi-empirical stellar spectra from the BaSeL 3.1 library (Westera et al. 2002), Padova 1994 stellar evolution tracks (Fagotto et al. 1994a,b,c) and a Chabrier IMF. We consider star formation histories with an exponentially declining star formation rate and we apply a prior on metallicity based on the mass-metallicity relation measured by Gallazzi et al. (2005). We sample the posterior probability distribution of stellar mass, age (time since the initial burst of star formation), star formation rate decline timescale, metallicity and dust attenuation with a Markov Chain Monte Carlo (MCMC), following the method introduced by Auger et al. (2009). We then consider the posterior probability distribution in log-stellar mass, marginalised over the other parameters, and

Table 1. Properties of the subsamples of ETGs used to build our fiducial (SDSS and LEGA-C) and high-redshift (vdS13, B14, G15 and B17) samples (the acronyms are defined in section 2). Column 1: subsample name. Column 2: redshift range. Column 3: stellar mass range in logarithm. Column 4: number of galaxies.

| Sample | z | $\log(M_*/M_\odot)$ | N_{ETG} |
|--------|--------------|---------------------|------------------|
| SDSS | (0.05; 0.55) | (10.50; 12.14) | 1583 |
| LEGA-C | (0.60; 1.18) | (10.50; 11.72) | 127 |
| vdS13 | (0.80; 2.19) | (10.54; 11.93) | 73 |
| B14 | (1.03; 1.60) | (10.59; 11.34) | 26 |
| G15 | (1.26; 1.41) | (11.04; 11.49) | 4 |
| B17 | (1.52; 2.44) | (10.60; 11.68) | 24 |

approximate it as a Gaussian with mean equal to

$$\log M_*^{\text{obs}} = \frac{\log M_*^{(84)} + \log M_*^{(16)}}{2} \quad (1)$$

and standard deviation

$$\sigma_{M_*} = \frac{\log M_*^{(84)} - \log M_*^{(16)}}{2}, \quad (2)$$

where $\log M_*^{(84)}$ and $\log M_*^{(16)}$ are the 84 and 16 percentile of the distribution, respectively. We refer to [Sonnenfeld et al. \(2019\)](#) for more details. Finally, we apply a lower cutoff to the observed stellar mass, selecting only galaxies with $\log(M_*^{\text{obs}}/M_\odot) > 10.5$. The final sample consists of 1583 SDSS galaxies, with mean redshift ≈ 0.203 and redshift standard deviation ≈ 0.106 , and 127 galaxies from the LEGA-C sample, with mean redshift ≈ 0.833 and redshift standard deviation ≈ 0.113 .

For each SDSS galaxy, we obtain, from the DR12 catalogue, the value and relative uncertainty of the line-of-sight stellar velocity dispersion measured in the $1.5''$ radius fiber of the SDSS spectrograph, which we label σ_{ap} . We convert this measurement into an estimate of the central velocity dispersion integrated within an aperture equal to the half-light radius, σ_e , by applying the following correction:

$$\sigma_e = \sigma_{\text{ap}} \times \left(\frac{R_e}{1.5''} \right)^{-\delta}, \quad (3)$$

where R_e is the half-light radius and $\delta = 0.066$ ([Cappellari et al. 2006](#)).

Velocity dispersion measurements provided in the LEGA-C DR2 are converted to values of the central velocity dispersion σ_e applying the aperture correction

$$\sigma_e = 1.05 \sigma_{\text{ap}}, \quad (4)$$

which is a good approximation for galaxies at $z < 1.8$ ([van de Sande et al. 2013a](#); [Belli et al. 2014](#)). The distributions in redshift and in stellar mass of the fiducial sample and of the SDSS and LEGA-C subsamples are shown in [Figure 2](#) (see also [Table 1](#)).

2.2 The high-redshift and extended samples

Our high-redshift sample of ETGs is a sample of 127 galaxies in the redshift range $0.8 \lesssim z \lesssim 2.5$, built as follows. We

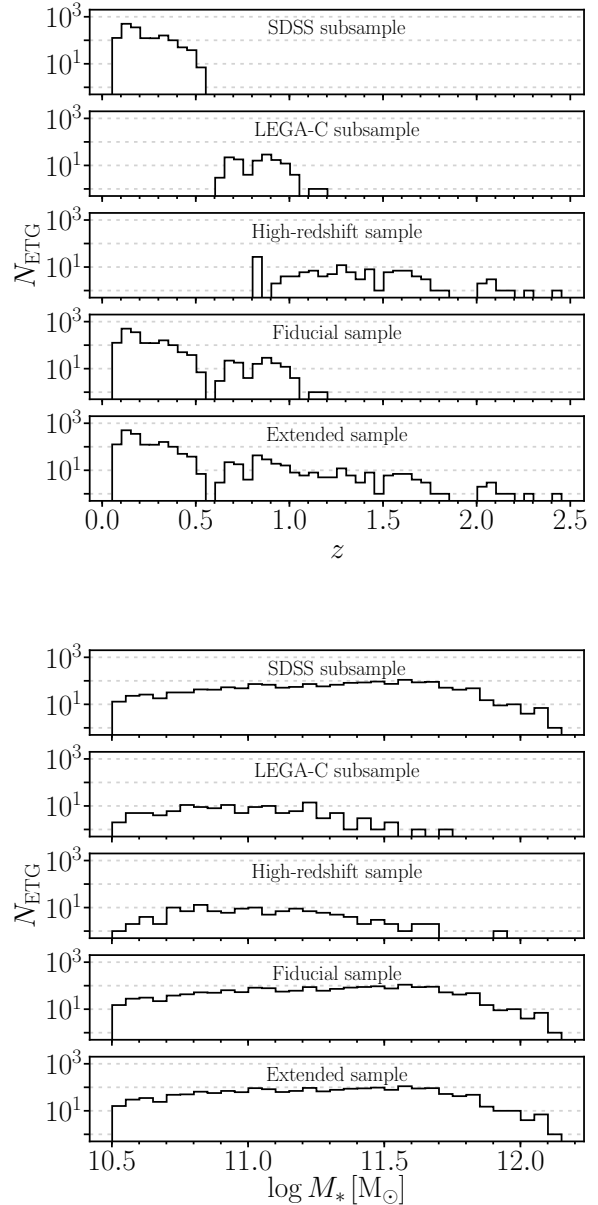


Figure 2. Distributions of the subsamples and samples of ETGs in redshift (top panel) and stellar mass (bottom panel). From the top to the bottom, the SDSS subsample, the LEGA-C subsample, the high-redshift sample (vdS13+B14+G15+B17 subsamples), the fiducial sample (SDSS+LEGA-C subsamples) and the extended sample (fiducial sample+high-redshift sample) distributions are shown.

obtain measurements of the stellar mass and stellar velocity dispersion of quiescent galaxies out to $z \approx 2.5$ from a variety of studies. In order of increasing median redshift, we take 26 galaxies drawn from the LRIS sample presented in [Belli et al. \(2014\)](#), hereafter B14), including only those galaxies for which $\text{EW}_{[\text{OIII}]} \geq -5 \text{ \AA}$ (as done for the fiducial sample; [subsection 2.1](#)), 73 galaxies from [van de Sande et al. \(2013a\)](#), hereafter vdS13), 4 galaxies from [Gargiulo et al. \(2015\)](#), hereafter G15), and 24 galaxies from [Belli et al. \(2017\)](#), hereafter

B17). The main properties of each of these subsamples are summarised in [Table 1](#). Among the 73 galaxies of the vdS13 subsample, only 5 galaxies are presented for the first time in vdS13, while the remaining 68 sources are collected from different studies. All the galaxies in the high-redshift samples are classified as quiescent, based on their *UVJ* colours, morphology and/or spectra. For the vast majority of these galaxies, stellar masses are measured by fitting SPS models to broadband imaging data and by scaling the total flux to match that measured by fitting a Sérsic surface brightness profile to high-resolution images from *Hubble Space Telescope* (HST). The details of the SPS models are very similar to those we adopted in our measurement of the stellar masses of the fiducial sample. In all these subsamples stellar masses are computed assuming Chabrier IMF and central velocity dispersions are given within an aperture of radius R_e . Our extended sample, obtained by combining the fiducial and high-redshift samples, consists of 1837 ETGs with $M_* \geq 10^{10.5} M_\odot$ in the redshift interval $0.05 \lesssim z \lesssim 2.44$. The distributions in redshift and in stellar mass of the high-redshift and extended samples are shown in [Figure 2](#).

In summary, after homogenising the observational data, for all galaxies in our samples we have measurements of spectroscopic redshift z , of stellar mass M_* and of the central stellar velocity dispersion σ_e . In [Table 2](#) we list the identifier, redshift, effective radius R_e , stellar velocity dispersion σ_e and stellar mass M_* for all the galaxies except for those drawn from the SDSS catalogue.

Table 2. Physical properties of the ETGs of the LEGA-C, vdS13, B14, G15 and B17 subsamples used to build our fiducial and high-redshift samples (we omit here galaxies taken from the SDSS subsample). Column 1: name of the subsample (in the case of the subsample vdS13, when a galaxy is taken from previous work, we indicate also the original reference: Be=*Bezanson et al. (2013a)*, vD=*van Dokkum et al. (2009)*, On=*Onodera et al. (2012)*, Ca=*Cappellari et al. (2009)*, Ne=*Newman et al. (2010)*, vdW&Bl=*van der Wel et al. (2008)* & *Blakeslee et al. (2006)*, To=*Toft et al. (2012)*). Column 2: galaxy identifier. Column 3: redshift. Column 4: central stellar velocity dispersion with 1σ uncertainty. Column 5: logarithm of the stellar mass with 1σ uncertainty. For the ETGs of the G15, vds13 and B17 subsamples there are no estimates of the uncertainty on M_* : for these galaxies we assumed uncertainty 0.15 dex in our analysis.

| Sample | Object | z | σ_e [km/s] | $\log M_*$ [M_\odot] | Sample | Object | z | σ_e [km/s] | $\log M_*$ [M_\odot] |
|--------|--------|--------|-------------------|--------------------------|--------|--------|--------|-------------------|--------------------------|
| LEGA-C | 128682 | 0.9281 | 148.5 ± 21.1 | 10.856 ± 0.12 | LEGA-C | 120781 | 0.6279 | 176.1 ± 6.9 | 10.811 ± 0.076 |
| LEGA-C | 129358 | 0.8523 | 172.1 ± 8.4 | 11.224 ± 0.128 | LEGA-C | 123441 | 0.8932 | 193.2 ± 19.8 | 10.664 ± 0.088 |
| LEGA-C | 129957 | 0.6655 | 200.1 ± 11.4 | 10.589 ± 0.079 | LEGA-C | 125017 | 0.6574 | 206.7 ± 16.8 | 10.964 ± 0.094 |
| LEGA-C | 133163 | 0.6966 | 178.3 ± 7.5 | 10.751 ± 0.129 | LEGA-C | 145294 | 0.685 | 205.4 ± 13.4 | 10.665 ± 0.072 |
| LEGA-C | 133240 | 0.7282 | 137.1 ± 9.1 | 10.577 ± 0.079 | LEGA-C | 158315 | 0.89 | 168.5 ± 10.7 | 10.95 ± 0.129 |
| LEGA-C | 134358 | 0.902 | 192.8 ± 12.8 | 10.752 ± 0.107 | LEGA-C | 158934 | 0.837 | 205.1 ± 7.5 | 11.323 ± 0.118 |
| LEGA-C | 140050 | 0.8993 | 249.6 ± 8.9 | 11.324 ± 0.105 | LEGA-C | 217545 | 1.0143 | 271 ± 8.6 | 11.498 ± 0.095 |
| LEGA-C | 210980 | 0.8968 | 188.8 ± 15.5 | 10.835 ± 0.089 | LEGA-C | 219193 | 0.8512 | 221.8 ± 9.5 | 10.992 ± 0.101 |
| LEGA-C | 95941 | 0.926 | 172.4 ± 7.4 | 11.269 ± 0.091 | LEGA-C | 219834 | 0.661 | 219.6 ± 10.2 | 11.01 ± 0.078 |
| LEGA-C | 96621 | 0.7132 | 245.7 ± 8.7 | 11.263 ± 0.089 | LEGA-C | 226945 | 1.1741 | 258.8 ± 34.2 | 10.814 ± 0.245 |
| LEGA-C | 97310 | 0.9428 | 154.9 ± 12.3 | 10.8 ± 0.159 | LEGA-C | 228562 | 0.7256 | 174.9 ± 7.9 | 10.776 ± 0.116 |
| LEGA-C | 98227 | 0.901 | 147.9 ± 34.1 | 10.705 ± 0.111 | LEGA-C | 230835 | 0.9769 | 143.7 ± 12.3 | 10.772 ± 0.155 |
| LEGA-C | 131869 | 0.7261 | 166.5 ± 7.9 | 10.603 ± 0.093 | LEGA-C | 231630 | 0.7249 | 181 ± 4.6 | 10.904 ± 0.155 |
| LEGA-C | 135149 | 0.8813 | 221.6 ± 13.4 | 10.856 ± 0.08 | LEGA-C | 232890 | 0.7657 | 162 ± 9.5 | 10.722 ± 0.095 |
| LEGA-C | 105208 | 0.9345 | 202 ± 26 | 11.23 ± 0.125 | LEGA-C | 203467 | 0.9184 | 172.3 ± 10 | 10.983 ± 0.089 |
| LEGA-C | 117907 | 0.9047 | 167.5 ± 11.2 | 10.836 ± 0.128 | LEGA-C | 212327 | 0.885 | 107.6 ± 14.1 | 10.532 ± 0.121 |
| LEGA-C | 119005 | 0.9243 | 140.1 ± 28.6 | 11.013 ± 0.109 | LEGA-C | 213599 | 0.7366 | 149.1 ± 10.2 | 10.944 ± 0.088 |
| LEGA-C | 119395 | 1.037 | 245.7 ± 19.1 | 11.251 ± 0.082 | LEGA-C | 215711 | 0.6653 | 168.2 ± 9.7 | 10.664 ± 0.128 |
| LEGA-C | 123726 | 0.7254 | 167.2 ± 8.2 | 11.037 ± 0.078 | LEGA-C | 216804 | 0.6661 | 183.4 ± 8.2 | 11.241 ± 0.071 |
| LEGA-C | 86744 | 0.6714 | 282.8 ± 8.4 | 11.081 ± 0.107 | LEGA-C | 225216 | 0.8938 | 187.5 ± 13.2 | 11.037 ± 0.123 |
| LEGA-C | 87479 | 0.7412 | 222.8 ± 9.7 | 11.231 ± 0.097 | LEGA-C | 225383 | 0.6983 | 160.3 ± 10.2 | 10.714 ± 0.103 |
| LEGA-C | 88863 | 0.8124 | 108.6 ± 7.7 | 10.686 ± 0.155 | LEGA-C | 226458 | 0.8902 | 154.9 ± 40.1 | 10.782 ± 0.065 |
| LEGA-C | 106394 | 0.8381 | 143.5 ± 11.4 | 10.907 ± 0.118 | LEGA-C | 228931 | 0.6967 | 219.4 ± 12.2 | 11.133 ± 0.076 |
| LEGA-C | 112593 | 0.8653 | 167 ± 24.8 | 10.569 ± 0.112 | LEGA-C | 230889 | 0.6961 | 68.2 ± 5.7 | 10.503 ± 0.083 |
| LEGA-C | 118972 | 0.8922 | 136.4 ± 12.3 | 10.855 ± 0.123 | LEGA-C | 232884 | 1.005 | 196.1 ± 29.6 | 11.135 ± 0.184 |
| LEGA-C | 119426 | 0.7998 | 258.4 ± 39.5 | 10.848 ± 0.077 | LEGA-C | 247031 | 0.8659 | 128.4 ± 13.4 | 11.244 ± 0.097 |
| LEGA-C | 126603 | 0.9331 | 201 ± 14.6 | 10.883 ± 0.123 | LEGA-C | 249582 | 0.8911 | 235 ± 24.5 | 10.741 ± 0.08 |
| LEGA-C | 127946 | 0.9387 | 209.8 ± 17.1 | 11.312 ± 0.083 | LEGA-C | 258091 | 0.8905 | 186.8 ± 14.6 | 11.362 ± 0.09 |
| LEGA-C | 128478 | 0.7513 | 294.8 ± 23.2 | 11.326 ± 0.064 | LEGA-C | 259363 | 0.6656 | 151.9 ± 7.5 | 10.819 ± 0.108 |
| LEGA-C | 128834 | 0.6985 | 59.9 ± 16.1 | 10.874 ± 0.171 | LEGA-C | 93681 | 0.6068 | 131.1 ± 15.5 | 11.051 ± 0.107 |
| LEGA-C | 129897 | 0.7282 | 189.9 ± 9.7 | 11.213 ± 0.105 | LEGA-C | 95303 | 0.8818 | 211.8 ± 10 | 11.507 ± 0.141 |
| LEGA-C | 129947 | 0.732 | 195.1 ± 12.5 | 10.65 ± 0.101 | LEGA-C | 96221 | 0.8781 | 229.5 ± 15 | 11.217 ± 0.088 |
| LEGA-C | 130982 | 0.852 | 204.4 ± 28.6 | 10.921 ± 0.061 | LEGA-C | 96616 | 0.8374 | 227.8 ± 10.2 | 11.235 ± 0.106 |
| LEGA-C | 131942 | 0.6919 | 246 ± 10.4 | 11.184 ± 0.082 | LEGA-C | 98016 | 0.8246 | 198.3 ± 11.4 | 11.229 ± 0.084 |
| LEGA-C | 138718 | 0.6558 | 230.3 ± 12.8 | 11.211 ± 0.055 | LEGA-C | 98104 | 0.9814 | 234.7 ± 23.4 | 11.127 ± 0.115 |
| LEGA-C | 179040 | 0.8902 | 202 ± 11.8 | 11.07 ± 0.111 | LEGA-C | 101430 | 0.9871 | 206.7 ± 26.5 | 11.139 ± 0.109 |
| LEGA-C | 206616 | 0.7276 | 179 ± 8.7 | 11.051 ± 0.091 | LEGA-C | 102581 | 0.9762 | 245.7 ± 10.4 | 11.712 ± 0.11 |
| LEGA-C | 127745 | 0.7307 | 166.3 ± 14 | 10.918 ± 0.102 | LEGA-C | 182000 | 0.8247 | 205.3 ± 14 | 11.085 ± 0.078 |
| LEGA-C | 128668 | 0.995 | 183.4 ± 19.8 | 11.317 ± 0.106 | LEGA-C | 183906 | 0.891 | 299.9 ± 21.1 | 11.435 ± 0.084 |
| LEGA-C | 129730 | 0.9362 | 132.5 ± 18.2 | 10.951 ± 0.111 | LEGA-C | 185631 | 0.9253 | 100.7 ± 12.5 | 10.858 ± 0.105 |
| LEGA-C | 131716 | 0.942 | 209 ± 27.1 | 10.737 ± 0.11 | LEGA-C | 215209 | 0.7288 | 212.6 ± 11.2 | 11.17 ± 0.064 |
| LEGA-C | 133833 | 0.73 | 179.9 ± 12.7 | 10.565 ± 0.096 | LEGA-C | 227297 | 0.8857 | 205.8 ± 19.8 | 10.757 ± 0.105 |
| LEGA-C | 166634 | 0.8527 | 191.2 ± 47.7 | 10.933 ± 0.135 | LEGA-C | 228250 | 0.998 | 245.6 ± 12 | 11.417 ± 0.126 |
| LEGA-C | 169134 | 0.8548 | 233.1 ± 15.5 | 11.004 ± 0.091 | LEGA-C | 228340 | 0.6968 | 213.6 ± 10.2 | 11.098 ± 0.105 |
| LEGA-C | 169320 | 0.8759 | 155.9 ± 8.2 | 11.17 ± 0.104 | LEGA-C | 231931 | 0.9968 | 270.7 ± 15.3 | 11.549 ± 0.082 |
| LEGA-C | 178560 | 0.7636 | 158.2 ± 18.6 | 10.833 ± 0.118 | LEGA-C | 243246 | 0.6963 | 268.1 ± 13.4 | 10.935 ± 0.078 |
| LEGA-C | 180162 | 0.6784 | 189.8 ± 11.2 | 10.757 ± 0.078 | LEGA-C | 244680 | 0.9576 | 204.5 ± 23.9 | 11.085 ± 0.088 |
| LEGA-C | 208824 | 0.7312 | 144.3 ± 6.4 | 11.261 ± 0.083 | LEGA-C | 245763 | 0.7542 | 146.2 ± 15.3 | 10.625 ± 0.108 |
| LEGA-C | 236382 | 0.8855 | 184.8 ± 13.5 | 10.882 ± 0.12 | LEGA-C | 246345 | 0.9989 | 173.5 ± 6.4 | 11.22 ± 0.168 |
| LEGA-C | 238743 | 0.6997 | 111.8 ± 6.3 | 10.858 ± 0.106 | LEGA-C | 261136 | 0.8843 | 173.4 ± 25.2 | 10.618 ± 0.091 |
| LEGA-C | 119489 | 0.6819 | 219.1 ± 8.9 | 11.102 ± 0.114 | LEGA-C | 262197 | 1.1127 | 275.8 ± 31.3 | 11.409 ± 0.115 |
| LEGA-C | 121033 | 0.938 | 235.9 ± 16.3 | 11.05 ± 0.117 | LEGA-C | 27265 | 0.7329 | 288 ± 15.5 | 11.081 ± 0.077 |
| LEGA-C | 121293 | 1.0501 | 154.2 ± 20.4 | 10.908 ± 0.197 | LEGA-C | 32375 | 0.9164 | 205.5 ± 35 | 10.783 ± 0.079 |
| LEGA-C | 124139 | 0.9249 | 287.9 ± 11 | 11.626 ± 0.088 | LEGA-C | 32382 | 0.9735 | 264.4 ± 18.2 | 11.188 ± 0.128 |
| LEGA-C | 161188 | 0.8916 | 222.4 ± 11 | 11.206 ± 0.076 | LEGA-C | 40014 | 0.9359 | 255.9 ± 15.8 | 11.07 ± 0.107 |
| LEGA-C | 108345 | 0.6773 | 147.5 ± 7.1 | 10.941 ± 0.091 | LEGA-C | 73790 | 0.8411 | 197.2 ± 25.7 | 11.052 ± 0.125 |
| LEGA-C | 109076 | 0.9509 | 199.4 ± 14.5 | 11.07 ± 0.126 | LEGA-C | 76336 | 0.8242 | 216.5 ± 16.9 | 10.786 ± 0.111 |
| LEGA-C | 118184 | 0.6767 | 252.8 ± 7.7 | 11.085 ± 0.105 | LEGA-C | 76593 | 0.8982 | 208.7 ± 18.6 | 10.834 ± 0.112 |
| LEGA-C | 119022 | 0.7344 | 161.4 ± 8.9 | 11.034 ± 0.097 | LEGA-C | 77934 | 0.9889 | 206.2 ± 18.4 | 10.977 ± 0.124 |
| LEGA-C | 120120 | 0.6287 | 245.1 ± 8.9 | 11.237 ± 0.075 | LEGA-C | 77957 | 0.8993 | 298.1 ± 20.4 | 11.454 ± 0.128 |

| Sample | Object | z | σ_e [km/s] | $\log M_*$ [M_\odot] | Sample | Object | z | σ_e [km/s] | $\log M_*$ [M_\odot] |
|--------------|----------|--------|-------------------|--------------------------|--------------|---------|-------|-------------------|--------------------------|
| LEGA-C | 78554 | 0.9752 | 208.1 ± 18.1 | 11.161 ± 0.137 | vdS13-vdW&Bl | 1287 | 0.85 | 345 ± 23 | 11.45 |
| LEGA-C | 79405 | 0.8265 | 106.3 ± 11.2 | 10.635 ± 0.11 | vdS13-vdW&Bl | 1328 | 0.84 | 252 ± 36 | 10.67 |
| LEGA-C | 85318 | 0.824 | 238.7 ± 14 | 10.984 ± 0.088 | vdS13-vdW&Bl | 1559 | 0.94 | 179 ± 13 | 10.79 |
| LEGA-C | 86395 | 0.8663 | 161.1 ± 20.7 | 10.591 ± 0.117 | vdS13-vdW&Bl | 1633 | 0.84 | 333 ± 14 | 10.98 |
| LEGA-C | 87207 | 0.8325 | 140.5 ± 12.8 | 10.753 ± 0.105 | vdS13-vdW&Bl | 1706 | 0.91 | 217 ± 13 | 11.17 |
| LEGA-C | 89072 | 0.837 | 190.3 ± 18.1 | 10.822 ± 0.084 | vdS13-vdW&Bl | 1709 | 0.84 | 220 ± 12 | 10.71 |
| LEGA-C | 90888 | 0.8837 | 199.9 ± 13.5 | 10.789 ± 0.093 | vdS13-vdW&Bl | 1 | 1.09 | 233 ± 17 | 11.17 |
| G15 | S2F1-511 | 1.267 | 281 ± 23 | 11.05 | vdS13-vdW&Bl | 2 | 0.96 | 202 ± 10 | 11.28 |
| G15 | S2F1-633 | 1.297 | 447 ± 27 | 11.48 | vdS13-vdW&Bl | 3 | 1.04 | 302 ± 34 | 10.82 |
| G15 | S2F1-527 | 1.331 | 240 ± 26 | 11.14 | vdS13-vdW&Bl | 4 | 0.96 | 337 ± 19 | 11.63 |
| G15 | S2F1-389 | 1.406 | 234 ± 45 | 11.14 | vdS13-vdW&Bl | 7 | 1.14 | 233 ± 20 | 11.54 |
| vdS13 | 7447 | 1.800 | 287 ± 53.5 | 11.22 | vdS13-vdW&Bl | 12 | 1.12 | 263 ± 23 | 10.94 |
| vdS13 | 18265 | 1.583 | 401 ± 72 | 11.32 | vdS13-vdW&Bl | 13 | 0.98 | 249 ± 11 | 11.24 |
| vdS13 | 7865 | 2.091 | 446 ± 56.5 | 11.64 | vdS13-vdW&Bl | 14 | 0.98 | 199 ± 24 | 10.89 |
| vdS13 | 19627 | 2.036 | 304 ± 41 | 11.20 | vdS13-vdW&Bl | 18 | 1.10 | 327 ± 37 | 11.50 |
| vdS13 | 29410 | 1.456 | 371 ± 102 | 11.24 | vdS13-vdW&Bl | 20 | 1.02 | 201 ± 17 | 11.16 |
| vdS13-Be | 17300 | 1.423 | 295 ± 8 | 11.24 | vdS13-To | 19627 | 2.039 | 326 ± 54 | 11.37 |
| vdS13-Be | 21129 | 1.584 | 301 ± 10 | 11.24 | B14 | 28739 | 1.029 | 238 ± 11 | 10.94 ± 0.08 |
| vdS13-Be | 22260 | 1.240 | 256 ± 16 | 11.41 | B14 | 31377 | 1.085 | 133 ± 18 | 10.83 ± 0.09 |
| vdS13-Be | 21434 | 1.522 | 250 ± 17 | 11.27 | B14 | 16343 | 1.098 | 290 ± 8 | 11.04 ± 0.03 |
| vdS13-Be | 20866 | 1.522 | 306 ± 24 | 11.34 | B14 | 28656 | 1.101 | 251 ± 15 | 11.08 ± 0.08 |
| vdS13-Be | 53937 | 1.621 | 280 ± 21 | 11.90 | B14 | 35291 | 1.110 | 245 ± 10 | 11.22 ± 0.11 |
| vdS13-vD | 1255 | 2.186 | 544 ± 138.5 | 11.26 | B14 | 21657 | 1.125 | 270 ± 13 | 10.97 ± 0.09 |
| vdS13-On | 254025 | 1.823 | 270 ± 106.5 | 11.39 | B14 | 21870 | 1.179 | 230 ± 12 | 11.02 ± 0.07 |
| vdS13-Ca | 2239 | 1.415 | 111 ± 35 | 10.54 | B14 | 1241357 | 1.188 | 207 ± 13 | 10.86 ± 0.04 |
| vdS13-Ca | 2470 | 1.415 | 141 ± 26 | 10.71 | B14 | 33887 | 1.193 | 162 ± 33 | 10.74 ± 0.11 |
| vdS13-Ne | E1 | 1.054 | 204 ± 22 | 11.07 | B14 | 18249 | 1.252 | 286 ± 109 | 10.77 ± 0.04 |
| vdS13-Ne | S1 | 1.110 | 220 ± 13 | 11.29 | B14 | 7310 | 1.255 | 176 ± 16 | 11.13 ± 0.07 |
| vdS13-Ne | E2 | 1.113 | 137 ± 11 | 10.89 | B14 | 13073 | 1.258 | 265 ± 12 | 10.97 ± 0.03 |
| vdS13-Ne | E3 | 1.124 | 241 ± 23 | 11.14 | B14 | 30822 | 1.259 | 271 ± 25 | 10.96 ± 0.07 |
| vdS13-Ne | E4 | 1.179 | 237 ± 14 | 10.95 | B14 | 1244914 | 1.261 | 252 ± 13 | 11.18 ± 0.07 |
| vdS13-Ne | E5 | 1.225 | 129 ± 20 | 10.61 | B14 | 32915 | 1.261 | 264 ± 17 | 10.88 ± 0.05 |
| vdS13-Ne | E6 | 1.243 | 278 ± 22 | 11.07 | B14 | 22760 | 1.262 | 232 ± 17 | 10.83 ± 0.06 |
| vdS13-Ne | GN1a | 1.253 | 268 ± 15 | 11.10 | B14 | 22780 | 1.264 | 88 ± 18 | 10.75 ± 0.07 |
| vdS13-Ne | E7b | 1.262 | 96 ± 19 | 10.71 | B14 | 2823 | 1.316 | 215 ± 21 | 11.01 ± 0.16 |
| vdS13-Ne | E8 | 1.262 | 259 ± 23 | 10.81 | B14 | 37085 | 1.316 | 164 ± 14 | 10.60 ± 0.15 |
| vdS13-Ne | GN2 | 1.266 | 220 ± 17 | 10.93 | B14 | 2337 | 1.327 | 279 ± 20 | 11.04 ± 0.06 |
| vdS13-Ne | GN3 | 1.315 | 263 ± 19 | 11.37 | B14 | 19498 | 1.401 | 250 ± 39 | 10.75 ± 0.07 |
| vdS13-Ne | S2b | 1.315 | 157 ± 23 | 10.70 | B14 | 5020 | 1.415 | 181 ± 54 | 10.83 ± 0.08 |
| vdS13-Ne | S3 | 1.394 | 248 ± 61 | 11.16 | B14 | 4906 | 1.419 | 298 ± 26 | 11.34 ± 0.07 |
| vdS13-Ne | GN4b | 1.395 | 193 ± 38 | 10.72 | B14 | 20275 | 1.442 | 221 ± 70 | 10.80 ± 0.07 |
| vdS13-Ne | E9 | 1.406 | 313 ± 54 | 10.85 | B14 | 34256 | 1.582 | 377 ± 54 | 11.33 ± 0.04 |
| vdS13-Ne | GN5b | 1.598 | 260 ± 27 | 11.92 | B14 | 2653 | 1.598 | 174 ± 27 | 10.82 ± 0.18 |
| vdS13-vdW&Bl | 1649 | 0.84 | 245 ± 29 | 11.03 | B17 | 17364 | 1.526 | 168 ± 84 | 10.83 |
| vdS13-vdW&Bl | 2409 | 0.84 | 290 ± 34 | 10.90 | B17 | 17361 | 1.527 | 169 ± 43 | 10.80 |
| vdS13-vdW&Bl | 3058 | 0.83 | 304 ± 33 | 11.30 | B17 | 17641 | 1.528 | 142 ± 54 | 10.65 |
| vdS13-vdW&Bl | 3768 | 0.82 | 224 ± 25 | 10.97 | B17 | 17089 | 1.528 | 348 ± 57 | 11.56 |
| vdS13-vdW&Bl | 3910 | 0.83 | 298 ± 44 | 10.91 | B17 | 17926 | 1.573 | 231 ± 39 | 11.01 |
| vdS13-vdW&Bl | 4345 | 0.83 | 339 ± 35 | 11.18 | B17 | 22719 | 1.579 | 262 ± 51 | 11.03 |
| vdS13-vdW&Bl | 4520 | 0.83 | 320 ± 29 | 11.68 | B17 | 24891 | 1.604 | 391 ± 71 | 10.85 |
| vdS13-vdW&Bl | 4926 | 0.83 | 313 ± 40 | 10.84 | B17 | 35616 | 1.609 | 198 ± 49 | 11.11 |
| vdS13-vdW&Bl | 5280 | 0.83 | 261 ± 32 | 10.98 | B17 | 30737 | 1.620 | 307 ± 82 | 11.23 |
| vdS13-vdW&Bl | 5298 | 0.83 | 286 ± 41 | 10.60 | B17 | 43367 | 1.624 | 299 ± 74 | 11.07 |
| vdS13-vdW&Bl | 5347 | 0.83 | 255 ± 25 | 10.76 | B17 | 30475 | 1.633 | 296 ± 109 | 10.74 |
| vdS13-vdW&Bl | 5450 | 0.84 | 234 ± 26 | 11.33 | B17 | 32707 | 1.647 | 174 ± 30 | 11.14 |
| vdS13-vdW&Bl | 5529 | 0.82 | 183 ± 24 | 10.76 | B17 | 16629 | 1.657 | 358 ± 76 | 10.61 |
| vdS13-vdW&Bl | 5577 | 0.83 | 308 ± 42 | 10.89 | B17 | 37529 | 1.665 | 232 ± 60 | 11.00 |
| vdS13-vdW&Bl | 5666 | 0.83 | 287 ± 23 | 11.20 | B17 | 22802 | 1.667 | 291 ± 31 | 10.92 |
| vdS13-vdW&Bl | 5756 | 0.83 | 234 ± 28 | 11.02 | B17 | 29352 | 1.690 | 146 ± 31 | 10.84 |
| vdS13-vdW&Bl | 6688 | 0.84 | 276 ± 38 | 10.98 | B17 | 19958 | 1.722 | 169 ± 87 | 10.72 |
| vdS13-vdW&Bl | 237 | 0.85 | 283 ± 22 | 11.30 | B17 | 17255 | 1.739 | 147 ± 40 | 10.84 |
| vdS13-vdW&Bl | 635 | 0.82 | 203 ± 18 | 10.77 | B17 | 25526 | 1.752 | 134 ± 36 | 10.73 |
| vdS13-vdW&Bl | 681 | 0.84 | 344 ± 32 | 10.61 | B17 | 13083 | 2.088 | 197 ± 52 | 11.11 |
| vdS13-vdW&Bl | 761 | 1.01 | 377 ± 40 | 11.53 | B17 | 11494 | 2.093 | 319 ± 26 | 11.67 |
| vdS13-vdW&Bl | 951 | 0.85 | 237 ± 18 | 11.37 | B17 | 7884 | 2.105 | 430 ± 69 | 11.47 |
| vdS13-vdW&Bl | 1236 | 0.85 | 219 ± 13 | 10.85 | B17 | 1769 | 2.300 | 338 ± 46 | 11.17 |
| vdS13-vdW&Bl | 1286 | 0.85 | 249 ± 17 | 11.04 | B17 | 5681 | 2.435 | 452 ± 130 | 10.96 |

3 METHOD

We use a Bayesian hierarchical method to infer the distribution of stellar velocity dispersion as a function of stellar mass and redshift for the ETGs in our samples. This method allows us to properly propagate observational uncertainties, to disentangle intrinsic scatter from observational errors and to correct for Eddington bias (Eddington 1913), which is introduced by our choice of imposing a lower cutoff to the stellar mass distribution. Throughout this section stellar masses are expressed in units of M_{\odot} .

3.1 Bayesian hierarchical formalism

We describe each galaxy in our sample by its redshift, stellar mass and central stellar velocity dispersion. We refer to these parameters collectively as $\Theta = \{M_*, \sigma_e, z\}$. These represent the true values of the three quantities, which are in general different from the corresponding observed values. We assume that the values of Θ are drawn from a probability distribution, described in turn by a set of *hyper-parameters* Φ :

$$P(\Theta) = P(\Theta|\Phi). \quad (5)$$

Our goal is to infer plausible values of the hyper-parameters, which summarise the distribution of our galaxies in the (M_*, σ_e, z) space, given our data. We will describe in detail the functional form of the distribution $P(\Theta|\Phi)$ in [subsection 3.2](#).

Using Bayes' theorem, the posterior probability distribution of the hyper-parameters given the data \mathbf{d} is

$$P(\Phi|\mathbf{d}) \propto P(\Phi)P(\mathbf{d}|\Phi), \quad (6)$$

where $P(\Phi)$ is the prior probability distribution of the model hyper-parameters and $P(\mathbf{d}|\Phi)$ is the likelihood of observing the data given the model.

The data consist of observed stellar masses, stellar velocity dispersions and redshifts,

$$\mathbf{d} \equiv \{M_*^{\text{obs}}, \sigma_e^{\text{obs}}, z^{\text{obs}}\}, \quad (7)$$

and related uncertainties. Since measurements on different galaxies are independent from each other, the likelihood term can be written as

$$P(\mathbf{d}|\Phi) = \prod_i P(\mathbf{d}_i|\Phi), \quad (8)$$

where \mathbf{d}_i is the data relative to the i -th galaxy. For each galaxy in our sample, the likelihood of the data depends only on the true values of the redshift, stellar mass and velocity dispersion, Θ , and not on the hyper-parameters Φ . In order to compute the $P(\mathbf{d}_i|\Phi)$ terms in [equation \(8\)](#), then, we need to marginalise over all possible values of the individual object parameters Θ_i :

$$P(\mathbf{d}_i|\Phi) = \int d\Theta_i P(\mathbf{d}_i|\Theta_i, \Phi) = \int d\Theta_i P(\mathbf{d}_i|\Theta_i)P(\Theta_i|\Phi). \quad (9)$$

This allows us to evaluate the posterior probability distribution, [equation \(6\)](#), provided that a model distribution $P(\Theta|\Phi)$ is specified, priors are defined and the shape of the likelihood is known. The method is hierarchical in the sense that there exists a hierarchy of parameters: individual object parameters Θ_i are drawn from a distribution that is, in turn, described by a set of hyper-parameters.

3.2 The model

The dependent variable of our model is the central velocity dispersion, σ_e , while stellar mass and redshift are independent variables. As such, it is useful to write the probability distribution of individual galaxy parameters as

$$P(\Theta|\Phi) = P(M_*, z|\Phi)P(\sigma_e|M_*, z, \Phi). \quad (10)$$

Here, $P(M_*, z|\Phi)$ describes the prior probability distribution for a galaxy in our sample to have true stellar mass M_* and true redshift z . This probability depends on some hyper-parameters, which may vary between different subsamples. Our galaxies have been selected by applying a lower cut to the observed stellar masses, $\log M_*^{\text{obs}} > 10.5$. We then expect the probability distribution in the true stellar mass to go to zero for low values of M_* . We also expect $P(M_*, z|\Phi)$ to vanish for very large values of M_* , as there are few known galaxies with $M_* > 10^{12}$. For simplicity, we assume that $P(M_*, z|\Phi)$ separates as follows:

$$P(M_*, z|\Phi) = P(M_*|\Phi)P(z|\Phi), \quad (11)$$

where $P(M_*|\Phi)$ is a skew Gaussian distribution in $\log M_*$,

$$P(M_*|\Phi) \propto \frac{1}{\sqrt{2\pi\sigma_*^2}} \exp\left\{-\frac{(\log M_* - \mu_*)^2}{2\sigma_*^2}\right\} \mathcal{E}(\log M_*|\Phi), \quad (12)$$

with

$$\mathcal{E}(\log M_*|\Phi) = 1 + \text{erf}\left(\alpha_* \frac{\log M_* - \mu_*}{\sqrt{2}\sigma_*}\right), \quad (13)$$

where μ_* , σ_* and α_* are three hyper-parameters. Since this is a prior on the stellar mass distribution, and since the typical uncertainty on the stellar mass measurements is much smaller than the width of this distribution (as shown in [section 4](#)), the particular choice of the functional form of $P(M_*|\Phi)$ does not matter in practice, because the likelihood term dominates over the prior. The main role of the prior is downweighting extreme outliers and measurements with very large uncertainties. The term $P(z|\Phi)$ in [equation \(11\)](#) describes the redshift distribution of the galaxies in our sample. As we will show later, this term does not enter the problem, because uncertainties on the observed redshifts can be neglected.

The second term on the right hand side of [equation \(10\)](#) is the core of our model. We assume that the logarithm of the stellar velocity dispersion is normally distributed, with a mean that can scale with redshift and stellar mass and with a variance that can evolve with redshift:

$$P(\log \sigma_e|M_*, z, \Phi) = \mathcal{N}(\mu_{\sigma}(M_*, z), \sigma_{\sigma}^2(z)). \quad (14)$$

We adopt the following functional form for the median of this distribution:

$$\mu_{\sigma} = \mu_0^{\text{SDSS}} + \beta \log\left(\frac{M_*}{M_*^{\text{piv}}}\right) + \zeta \log\left(\frac{1+z}{1+z^{\text{piv}}}\right). \quad (15)$$

In general, the slope β is allowed to depend on z as

$$\beta = \beta_0^{\text{SDSS}} + \eta \log\left(\frac{1+z}{1+z^{\text{piv}}}\right). \quad (16)$$

We perform our analysis considering two different cases: the first is a *constant-slope* case (model $\mathcal{M}_{\text{const}}$), i.e. [equation \(16\)](#) with $\eta = 0$; in the second, which we refer to as

the *evolving-slope* case (model \mathcal{M}_{evo}), η is a free hyper-parameter. For the standard deviation σ_σ in equation (14), namely the intrinsic scatter of our relation, we adopt the form

$$\sigma_\sigma = \psi_0^{\text{SDSS}} + \xi \log \left(\frac{1+z}{1+z^{\text{piv}}} \right). \quad (17)$$

In equations (15-17) $M_*^{\text{piv}} = 10^{11.361}$ and $z^{\text{piv}} = 0.17311$, i.e. the median values of stellar mass and redshift of the SDSS ETGs, respectively, while the quantities μ_0^{SDSS} , β_0^{SDSS} and ψ_0^{SDSS} are the median values of the hyper-parameters μ_0 , β_0 and ψ_0 obtained when fitting equation (14) to the ETGs of the SDSS subsample with

$$\mu_\sigma = \mu_0 + \beta_0 \log \left(\frac{M_*}{M_*^{\text{piv}}} \right) \quad \text{and} \quad \sigma_\sigma = \psi_0, \quad (18)$$

i.e. neglecting any dependence on z . In order to prevent the relation being dominated at lower redshifts by the SDSS subsample, which constitutes more than 85% of the extended sample, we assume the model in equation (18) as the *zero point* at z^{piv} for our redshift-dependent models, because our main interest is to trace the evolution of the relation at higher redshift ($z \gtrsim 0.5$). Hereafter, we will refer to the model in equation (18) as model $\mathcal{M}_{\text{SDSS}}$.

3.3 Sampling the posterior probability distribution functions of the model hyper-parameters

Our goal is to sample the posterior probability distribution function (PDF) of the model hyper-parameters Φ given the data \mathbf{d} , $\text{P}(\Phi|\mathbf{d})$. For this purpose, we use an MCMC approach, using a Python adaptation of the affine-invariant ensemble sampler of Goodman & Weare (2010), *emcee* (Foreman-Mackey et al. 2013). For each set of values of the hyper-parameters, we need to evaluate the likelihood of the data. This is given by the product over the galaxies in our sample of the integrals in equation (9). Using $\log M_*$, $\log \sigma_e$ and z as the integration variables and omitting the subscript i in order to simplify the notation, equation (9) reads

$$\begin{aligned} \text{P}(M_*^{\text{obs}}, \sigma_e^{\text{obs}}, z^{\text{obs}}|\Phi) &= \\ &= \iiint d \log M_* d \log \sigma_e dz \times \\ &\quad \times \text{P}(M_*^{\text{obs}}, \sigma_e^{\text{obs}}, z^{\text{obs}}|M_*, \sigma_e, z) \times \\ &\quad \times \text{P}(M_*, \sigma_e, z|\Phi) = \\ &= \iiint d \log M_* d \log \sigma_e dz \times \\ &\quad \times \text{P}(M_*^{\text{obs}}|M_*) \text{P}(\sigma_e^{\text{obs}}|\sigma_e) \delta(z^{\text{obs}} - z) \times \\ &\quad \times \text{P}(M_*|\Phi) \text{P}(z|\Phi) \text{P}(\sigma_e|M_*, z, \Phi). \end{aligned} \quad (19)$$

In the last line, we have used equations (10) and (11), and we have approximated the likelihood of observing redshift z^{obs} as a delta function, in virtue of the very small uncertainties on the redshift (typical errors are $< 10^{-4}$). As a result, the redshift distribution term $\text{P}(z|\Phi)$ becomes irrelevant, as it contributes to the integral only through a multiplicative constant that we can ignore.

Assuming a Gaussian likelihood in $\log \sigma_e^{\text{obs}}$ for the term

Table 3. Jeffreys' scale (Jeffreys 1961), giving the strength of evidence in the comparison of two models having Bayes factor \mathcal{B} (equation 24).

| $ \ln \mathcal{B} $ | Strength of evidence |
|---------------------|----------------------|
| 0 – 1 | Inconclusive |
| 1 – 2.5 | Weak evidence |
| 2.5 – 5 | Strong evidence |
| > 5 | Decisive evidence |

$\text{P}(\sigma_e^{\text{obs}}|\sigma_e)$, the integral over $d \log \sigma_e$ can be performed analytically, as we show in Appendix A. We also assume a Gaussian likelihood for the measurements of $\log M_*^{\text{obs}}$,

$$\text{P}(M_*^{\text{obs}}|M_*) = \frac{\mathcal{A}(\log M_*)}{\sqrt{2\pi\sigma_{M_*}^2}} \exp \left\{ -\frac{(\log M_* - \log M_*^{\text{obs}})^2}{2\sigma_{M_*}^2} \right\}, \quad (20)$$

with one caveat: since we are only selecting galaxies with $\log M_*^{\text{obs}} > 10.5$, the likelihood must be normalised accordingly,

$$\int_{10.5}^{\infty} d \log M_*^{\text{obs}} \frac{\mathcal{A}(\log M_*)}{\sqrt{2\pi\sigma_{M_*}^2}} \exp \left\{ -\frac{(\log M_* - \log M_*^{\text{obs}})^2}{2\sigma_{M_*}^2} \right\} = 1. \quad (21)$$

In other words, the probability of measuring any value of the stellar mass larger than $\log M_*^{\text{obs}} = 10.5$, given that a galaxy is part of our sample, is one. We perform the final integration over $\log M_*$ numerically with a Monte Carlo method (see Appendix A). We assume flat priors on all model hyper-parameters.

3.4 Bayesian evidence

In our analysis, we consider models with different numbers of free hyper-parameters. To evaluate the performance of a given model in fitting the data, we rely on the *Bayesian evidence* \mathcal{Z} that is the average of the likelihood under priors for a given model \mathcal{M} :

$$\mathcal{Z} = \text{P}(\mathbf{d}|\mathcal{M}) = \int d\Theta \text{P}(\mathbf{d}|\Theta, \mathcal{M}) \text{P}(\Theta|\mathcal{M}). \quad (22)$$

We remark that, in our approach, the parameters Θ are described by a set of global hyper-parameters Φ . When comparing two models, say models \mathcal{M}_1 and \mathcal{M}_2 , we are interested in computing the ratio of the posterior probabilities of the models

$$\frac{\text{P}(\mathcal{M}_1|\mathbf{d})}{\text{P}(\mathcal{M}_2|\mathbf{d})} = \mathcal{B} \frac{\text{P}(\mathcal{M}_1)}{\text{P}(\mathcal{M}_2)}, \quad (23)$$

where

$$\mathcal{B} \equiv \frac{\text{P}(\mathbf{d}|\mathcal{M}_1)}{\text{P}(\mathbf{d}|\mathcal{M}_2)} = \frac{\mathcal{Z}_1}{\mathcal{Z}_2} \quad (24)$$

is the *Bayes factor*. When $\mathcal{B} \gg 1$, \mathcal{M}_1 provides a better description of the data than \mathcal{M}_2 , and vice versa when $\mathcal{B} \ll 1$. The value of the Bayes factor is usually compared with the reference values of the empirical *Jeffreys' scale* (Jeffreys 1961), reported in Table 3. Given two different models, the quantity $|\ln \mathcal{B}|$ is a measure of the strength of evidence that one of the two models is preferable. We compute the

Bayesian evidence \mathcal{Z} of a model exploiting the *nested sampling* technique (Skilling 2004). Briefly, the nested sampling algorithm estimates the Bayesian evidence reducing the n -dimensional evidence integral (where n is the number of the parameters of a given model) into a 1D integral that is less expensive to evaluate numerically. In practice, we evaluate \mathcal{Z} for a model using the MULTINEST algorithm (see Feroz & Hobson 2008; Feroz et al. 2009) included in the Python module PyMultiNest (Buchner et al. 2014). For details about the estimates of the Bayesian evidence and the algorithm exploited to compute them, we refer the interested readers to Feroz & Hobson (2008) and Buchner et al. (2014).

4 RESULTS

In this section we present the results obtained applying the Bayesian method described in section 3 to the fiducial sample and to the extended sample of ETGs, the latter consisting of a combination of the fiducial and high-redshift samples (see section 2).

In subsection 3.2 we have introduced three models: model $\mathcal{M}_{\text{SDSS}}$ (representing the present-day $M_* - \sigma_e$ relation), model $\mathcal{M}_{\text{const}}$ (representing the evolution of the $M_* - \sigma_e$ relation with redshift-independent slope β) and model \mathcal{M}_{evo} (representing the evolution of the $M_* - \sigma_e$ relation with redshift-dependent slope β). In models $\mathcal{M}_{\text{const}}$ and \mathcal{M}_{evo} the intrinsic scatter of the $M_* - \sigma_e$ relation is allowed to vary with redshift. In addition to these models, we also explore simpler models in which the intrinsic scatter is assumed to be independent of redshift. These models are named $\mathcal{M}_{\text{const}}^{\text{NES}}$ and $\mathcal{M}_{\text{evo}}^{\text{NES}}$, where NES stands for *non-evolving scatter*. In summary, we take into account five models: model $\mathcal{M}_{\text{SDSS}}$, represented by equation (18), models $\mathcal{M}_{\text{const}}$ and \mathcal{M}_{evo} , described by equation (15) (the former obtained by assuming $\eta = 0$ in equation (16)), and the models $\mathcal{M}_{\text{const}}^{\text{NES}}$ and $\mathcal{M}_{\text{evo}}^{\text{NES}}$, which are the same as $\mathcal{M}_{\text{const}}$ and \mathcal{M}_{evo} , respectively, but with $\xi = 0$ in equation (17). A description of the hyper-parameters used for each model is provided in Table 4. Model $\mathcal{M}_{\text{SDSS}}$ is applied to the SDSS subsample. The other four models are applied twice, once to the fiducial sample and once to the extended sample (we use the superscripts "fid" and "ext" to indicate that a model is applied, respectively, to the fiducial and extended samples).

The model-data comparison is performed as described in section 3. We validated our method by applying it to a mock dataset similar to the considered observational dataset (see Appendix B). Each MCMC run (see subsection 3.3) uses 50 random walkers running for 800 steps to reach the convergence of the hyper-parameter distribution. The resulting inferences on the hyper-parameters used in model $\mathcal{M}_{\text{SDSS}}$ are shown in Figure 3. The SDSS galaxies are described by $\sigma_e \propto M_*^{\beta_0}$ with $\beta_0 \simeq 0.233$, close to the slope of the classical Faber-Jackson relation $\sigma_0 \propto L^{0.25}$ (Faber & Jackson 1976). The normalisation (μ_0) is such that galaxies with $M_* = 10^{11} M_\odot$ have $\sigma_e \simeq 172 \text{ km s}^{-1}$ and the intrinsic scatter (ψ_0) is $\simeq 0.066$ dex in σ_e at fixed M_* . Our fit to the present-day $M_* - \sigma_e$ relation is broadly consistent with previous analyses (see section 5 for details).

The median values of the hyper-parameters of all models, with the corresponding 1σ errors, are listed in Table 5. In order to compare the models we compute the Bayesian

evidence \mathcal{Z} of each model, using a configuration of 400 live points in the nested sampling algorithm. The resulting \mathcal{Z} and the Bayes factors (see equation 24) are listed in Table 6. The performance of models $\mathcal{M}_{\text{evo}}^{\text{NES}}$ is relatively poor when applied to both the fiducial and the extended samples, so in the following we focus on the other models. In Figures 4–6, we show the inferences of the models $\mathcal{M}_{\text{const}}^{\text{NES}}$, $\mathcal{M}_{\text{const}}$ and \mathcal{M}_{evo} for both the fiducial and the extended samples.

4.1 Fiducial sample ($0 \lesssim z \lesssim 1.2$)

The Bayesian evidences for the four models applied to the fiducial sample (Table 6) indicate that the most representative model is $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$. According to Jeffreys' scale (Table 3), the other three models are significantly worse (with weak evidence against models $\mathcal{M}_{\text{evo}}^{\text{NES, fid}}$ and $\mathcal{M}_{\text{const}}^{\text{fid}}$, and strong evidence against model $\mathcal{M}_{\text{evo}}^{\text{fid}}$). Thus, based on our analysis of the fiducial sample, we conclude that at $z \lesssim 1.2$ the normalisation of the $M_* - \sigma_e$ relation changes with z , while neither the slope nor the intrinsic scatter vary significantly. In this redshift interval the $M_* - \sigma_e$ relation is well described by a power law $\sigma_e \propto M_*^\beta$ with the same slope $\beta = 0.233$ and the same scatter 0.066 dex found for the SDSS subsample. At fixed M_* , the normalisation of the stellar mass–velocity dispersion relation increases back in time as $\sigma_0 \propto (1+z)^\zeta$, with $\zeta \simeq 0.26$, so, at fixed M_* , galaxies tend to have higher σ_e at higher redshift: the median velocity dispersion at fixed M_* is $\simeq 20\%$ higher at $z = 1$ than at $z = 0$.

The best-fitting $M_* - \sigma_e$ relations found for model $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ at $z = 0$, $z = 0.5$ and $z = 1$ are shown in Figure 7. In this diagram the dashed curves are obtained by computing, at given z and $\log M_*$, the median value of $\log \sigma_e$ among all the values sampled by the posterior distribution obtained with the MCMC; similarly, the shaded bands, which we will refer to as 1σ bands, are defined by computing the 16% and the 84% of the distribution of $\log \sigma_e$, at given z and $\log M_*$, for the same sampling. This plot clearly shows that the evolution of the normalisation of the $M_* - \sigma_e$ relation at $0 \lesssim z \lesssim 1$ is significant.

In summary, based on the median values of the hyper-parameters of model $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$, the evolution of the $M_* - \sigma_e$ relation in the redshift range $0 \lesssim z \lesssim 1.2$ can be roughly described by

$$\log \left(\frac{\sigma_e}{\text{km s}^{-1}} \right) \simeq 2.22 + 0.23 \log \left(\frac{M_*}{10^{11} M_\odot} \right) + 0.26 \log(1+z), \quad (25)$$

with redshift-independent intrinsic scatter $\sigma_\sigma \simeq 0.07$ in $\log \sigma_e$ at given M_* .

4.2 Extended sample ($0 \lesssim z \lesssim 2.5$)

We move here to the analysis of the models applied to the extended sample. Based on the Bayesian evidences \mathcal{Z} , we find that, among the models applied to the extended sample, model $\mathcal{M}_{\text{const}}^{\text{ext}}$ has the highest value of \mathcal{Z} , so it is the most representative model to trace the evolution of the $M_* - \sigma_e$ relation. However, $\mathcal{M}_{\text{evo}}^{\text{ext}}$ cannot be rejected: its Bayes factor relative to model $\mathcal{M}_{\text{const}}^{\text{ext}}$ is such that $|\ln \mathcal{B}| < 1$, thus, according to Jeffrey's scale (Table 3), $\mathcal{M}_{\text{evo}}^{\text{ext}}$ is not significantly worse than $\mathcal{M}_{\text{const}}^{\text{ext}}$. We conclude that the considered observational data do not allow us to determine whether

Table 4. Hyper-parameters used in the models. Column 1: name of the model. Column 2: name of the hyper-parameter. Column 3: description of the hyper-parameter. Column 4: priors used in the models ("low" and "up" indicate, respectively, the lower and upper bounds and "guess" is the starting value). M_*^{piv} and z^{piv} are the median values of stellar mass and redshift of the SDSS ETGs (subsection 3.2).

| Model | Hyper-parameter | Description | Prior (low; up; guess) |
|---|---|---|--------------------------|
| $\mathcal{M}_{\text{SDSS}}$ | μ_0 | Median value of $\log \sigma_e$ at M_*^{piv} | Uniform (1; 3; 2.3) |
| | β_0 | Index of the $M_* - \sigma_e$ relation: $\sigma_e \propto M_*^{\beta_0}$ | Uniform (0; 1; 0.2) |
| | ψ_0 | Intrinsic scatter in $\log \sigma_e$ | Uniform (0; 1; 0.1) |
| | μ_* | Mean of Gaussian prior in the stellar mass distribution | Uniform (10; 13; 11) |
| | σ_* | Standard deviation in the Gaussian prior of stellar mass distribution | Uniform (0; 2; 0.15) |
| | α_* | Skewness parameter in the Gaussian prior of stellar mass distribution | Uniform (0.1; 10; 1) |
| \mathcal{M}_{evo} | μ_0^{SDSS} | Median value of $\log \sigma_e$ at $M_* = M_*^{\text{piv}}$ and $z = z^{\text{piv}}$ | ≈ 2.319 |
| | β_0^{SDSS} | Index of the $M_* - \sigma_e$ relation at $z = z^{\text{piv}}$: $\sigma_e \propto M_*^{\beta_0^{\text{SDSS}}}$ | ≈ 0.233 |
| | η | Index of the $\beta - (1+z)$ relation: $\beta \propto (1+z)^\eta$ | Uniform (-2; 2; 0; 0.05) |
| | ζ | Index of the $\sigma_e - (1+z)$ relation: $\sigma_e \propto (1+z)^\zeta$ | Uniform (-2; 2; 0; 0.05) |
| | ψ_0^{SDSS} | Median value of ψ_0 of the intrinsic scatter at $z = z^{\text{piv}}$ | ≈ 0.066 |
| | ξ | Index of the $\sigma_\sigma - (1+z)$ relation: $\sigma_\sigma \propto (1+z)^\xi$ | Uniform (-2; 2; 0; 0.05) |
| | μ_* | Mean of Gaussian prior in the stellar mass distribution | Uniform(10; 13; 11) |
| | σ_* | Standard deviation in the Gaussian prior of stellar mass distribution | Uniform (0; 2; 0.15) |
| α_* | Skewness parameter in the Gaussian prior of stellar mass distribution | Uniform (0.1; 10; 1) | |
| $\mathcal{M}_{\text{const}}$ | | Same as \mathcal{M}_{evo} , but with $\eta = 0$ | |
| $\mathcal{M}_{\text{evo}}^{\text{NES}}$ | | Same as \mathcal{M}_{evo} , but with $\xi = 0$ | |
| $\mathcal{M}_{\text{const}}^{\text{NES}}$ | | Same as \mathcal{M}_{evo} , but with $\eta = \xi = 0$ | |

Table 5. Inferred median and 68% limits on the PDFs of the hyper-parameters of the models.

| | μ_0 | β_0 | ψ_0 | η | ζ | ξ | μ_* | σ_* | α_* |
|--|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|---------------------------|----------------------------|---------------------------|---------------------------|
| $\mathcal{M}_{\text{SDSS}}$ | $2.319^{+0.002}_{-0.002}$ | $0.233^{+0.006}_{-0.006}$ | $0.066^{+0.002}_{-0.002}$ | – | – | – | $11.241^{+0.028}_{-0.035}$ | $0.351^{+0.01}_{-0.008}$ | $0.247^{+0.128}_{-0.098}$ |
| $\mathcal{M}_{\text{const}}^{\text{fid}}$ | – | – | – | – | $0.255^{+0.042}_{-0.044}$ | $0.069^{+0.04}_{-0.034}$ | $10.816^{+0.089}_{-0.091}$ | $0.305^{+0.06}_{-0.042}$ | $0.928^{+0.946}_{-0.53}$ |
| $\mathcal{M}_{\text{evo}}^{\text{fid}}$ | – | – | – | $0.116^{+0.182}_{-0.202}$ | $0.289^{+0.074}_{-0.076}$ | $0.067^{+0.04}_{-0.036}$ | $10.819^{+0.095}_{-0.089}$ | $0.305^{+0.054}_{-0.042}$ | $0.944^{+0.834}_{-0.575}$ |
| $\mathcal{M}_{\text{const}}^{\text{ext}}$ | – | – | – | – | $0.422^{+0.028}_{-0.029}$ | $0.098^{+0.025}_{-0.025}$ | $10.808^{+0.096}_{-0.066}$ | $0.323^{+0.051}_{-0.049}$ | $1.259^{+0.875}_{-0.707}$ |
| $\mathcal{M}_{\text{evo}}^{\text{ext}}$ | – | – | – | $0.222^{+0.120}_{-0.122}$ | $0.492^{+0.047}_{-0.052}$ | $0.088^{+0.26}_{-0.023}$ | $10.819^{+0.097}_{-0.064}$ | $0.313^{+0.047}_{-0.045}$ | $1.175^{+0.79}_{-0.681}$ |
| $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ | – | – | – | – | $0.258^{+0.034}_{-0.038}$ | – | $10.834^{+0.081}_{-0.091}$ | $0.3^{+0.057}_{-0.035}$ | $0.786^{+0.781}_{-0.469}$ |
| $\mathcal{M}_{\text{evo}}^{\text{NES, fid}}$ | – | – | – | $0.153^{+0.174}_{-0.015}$ | $0.31^{+0.069}_{-0.067}$ | – | $10.832^{+0.083}_{-0.086}$ | $0.299^{+0.054}_{-0.036}$ | $0.827^{+0.751}_{-0.51}$ |
| $\mathcal{M}_{\text{const}}^{\text{NES, ext}}$ | – | – | – | – | $0.431^{+0.022}_{-0.023}$ | – | $10.844^{+0.094}_{-0.075}$ | $0.308^{+0.048}_{-0.039}$ | $0.911^{+0.738}_{-0.567}$ |
| $\mathcal{M}_{\text{evo}}^{\text{NES, ext}}$ | – | – | – | $0.29^{+0.108}_{-0.099}$ | $0.517^{+0.04}_{-0.041}$ | – | $10.866^{+0.087}_{-0.076}$ | $0.29^{+0.046}_{-0.032}$ | $0.79^{+0.66}_{-0.506}$ |

the slope of the $M_* - \sigma_0$ relation evolves with redshift in the range $0 \lesssim z \lesssim 2.5$. Instead, there is a strong evidence that both model $\mathcal{M}_{\text{const}}^{\text{NES, ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{NES, ext}}$ must be rejected (having Bayes factors such that $\ln \mathcal{B} < -4$, relative to $\mathcal{M}_{\text{const}}^{\text{ext}}$). Hence, considering the extended sample, the $M_* - \sigma_e$ relation has a significant intrinsic scatter evolution.

The median $M_* - \sigma_e$ relations (with 1σ bands) for the two best models of the extended sample, i.e. $\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$, at six representative redshifts ($z = 0, 0.5, 1, 1.5, 2$ and 2.5) are shown in Figure 8. At given redshift, the central stellar velocity dispersion σ_e increases with stellar mass: the slope β increases from ≈ 0.22 at $z = 0$ up to ≈ 0.32 at $z = 2$ for model $\mathcal{M}_{\text{const}}^{\text{ext}}$ (while it is fixed at $\beta = 0.233$ in model $\mathcal{M}_{\text{const}}^{\text{ext}}$). At fixed M_* , the normalisation of the stellar mass–velocity

dispersion relation varies with redshift as $\sigma_0 \propto (1+z)^\zeta$, with $\zeta \approx 0.42$ for model $\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\zeta \approx 0.49$ for model $\mathcal{M}_{\text{evo}}^{\text{ext}}$. For instance, the median velocity dispersion at $M_* = 10^{11} M_\odot$ varies from $\sigma_e \approx 160 \text{ km s}^{-1}$ at $z = 0$ to $\sigma_0 \approx 250 \text{ km s}^{-1}$ at $z = 2$ for both models $\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$. The intrinsic scatter increases with redshift from ≈ 0.06 dex at $z = 0$ to ≈ 0.11 dex at $z = 2$. A direct comparison of the the two best models of the extended sample ($\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$) can be found in Figure 9, in which the median correlations, each with its 1σ uncertainty band and intrinsic scatter, are plotted at six representative redshifts. At all redshifts, the central stellar velocity dispersions predicted by the two models are essentially indistinguishable around $M_* = 10^{11} M_\odot$ and differ significantly only at the high-mass end.

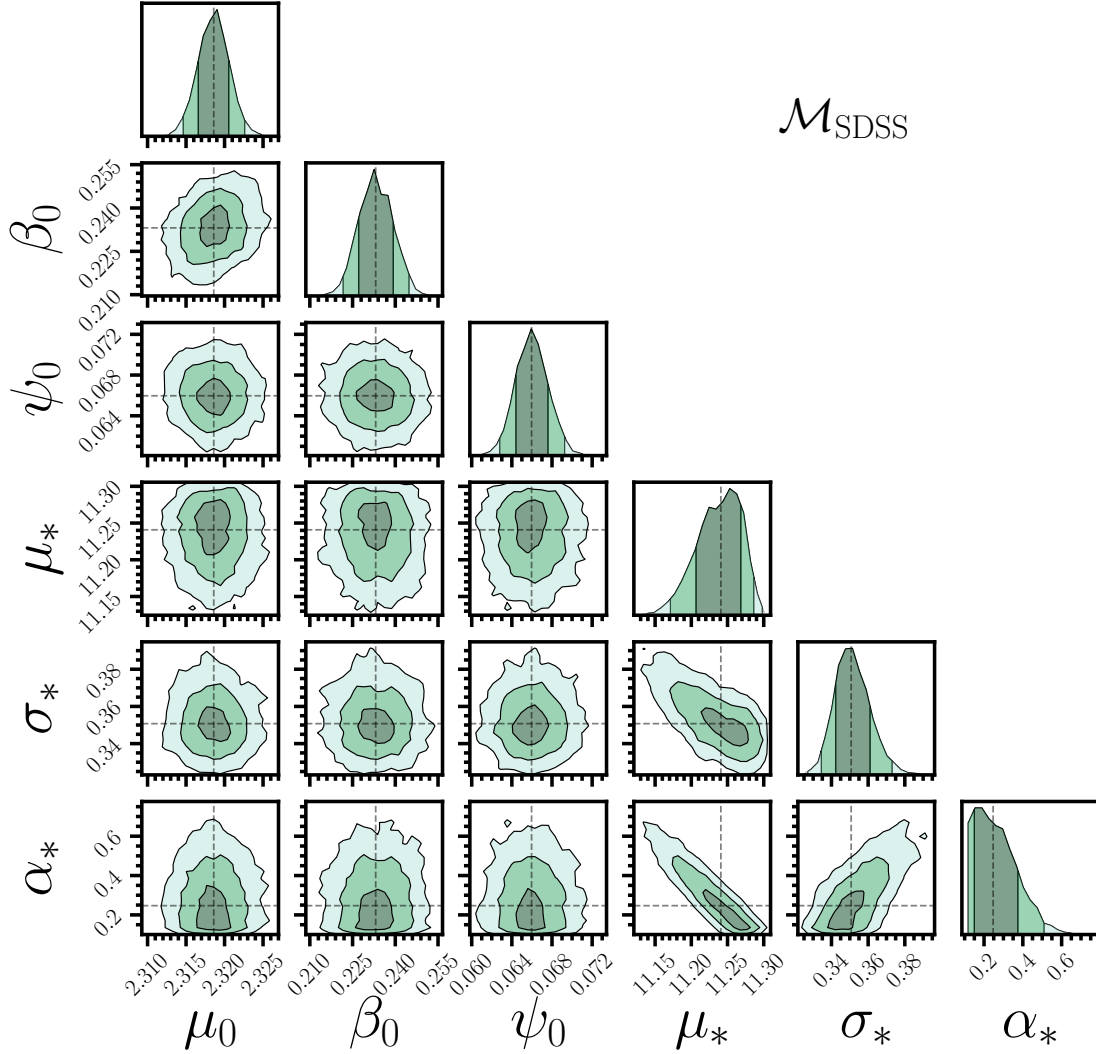


Figure 3. Posterior probability distributions of the hyper-parameters for model $\mathcal{M}_{\text{SDSS}}$ (see Tables 4 and 5). In the 1D distributions (top panel of each column) the vertical solid lines and colours delimit the 1σ , 2σ and 3σ levels. In the 2D distributions (all the other panels) the contours enclose the 68, 95 and 99.7 percent probability regions. The dashed lines indicate the median values of the hyper-parameters.

In summary, based on the median values of the hyper-parameters of model $\mathcal{M}_{\text{const}}^{\text{ext}}$ the evolution of the $M_* - \sigma_0$ relation in the redshift range $0 \lesssim z \lesssim 2.5$ can be roughly described by

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) \approx 2.21 + 0.23 \log\left(\frac{M_*}{10^{11} M_\odot}\right) + 0.42 \log(1+z), \quad (26)$$

with an intrinsic scatter

$$\sigma_\sigma \approx 0.06 + 0.1 \log(1+z). \quad (27)$$

According to model $\mathcal{M}_{\text{ev0}}^{\text{ext}}$, the evolution of the $M_* - \sigma_0$ relation in the redshift range $0 \lesssim z \lesssim 2.5$ can be roughly

described by

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) \approx 2.22 + \beta(z) \log\left(\frac{M_*}{10^{11} M_\odot}\right) + 0.49 \log(1+z), \quad (28)$$

with

$$\beta(z) \approx 0.22 + 0.22 \log(1+z) \quad (29)$$

and an intrinsic scatter

$$\sigma_\sigma \approx 0.06 + 0.09 \log(1+z). \quad (30)$$

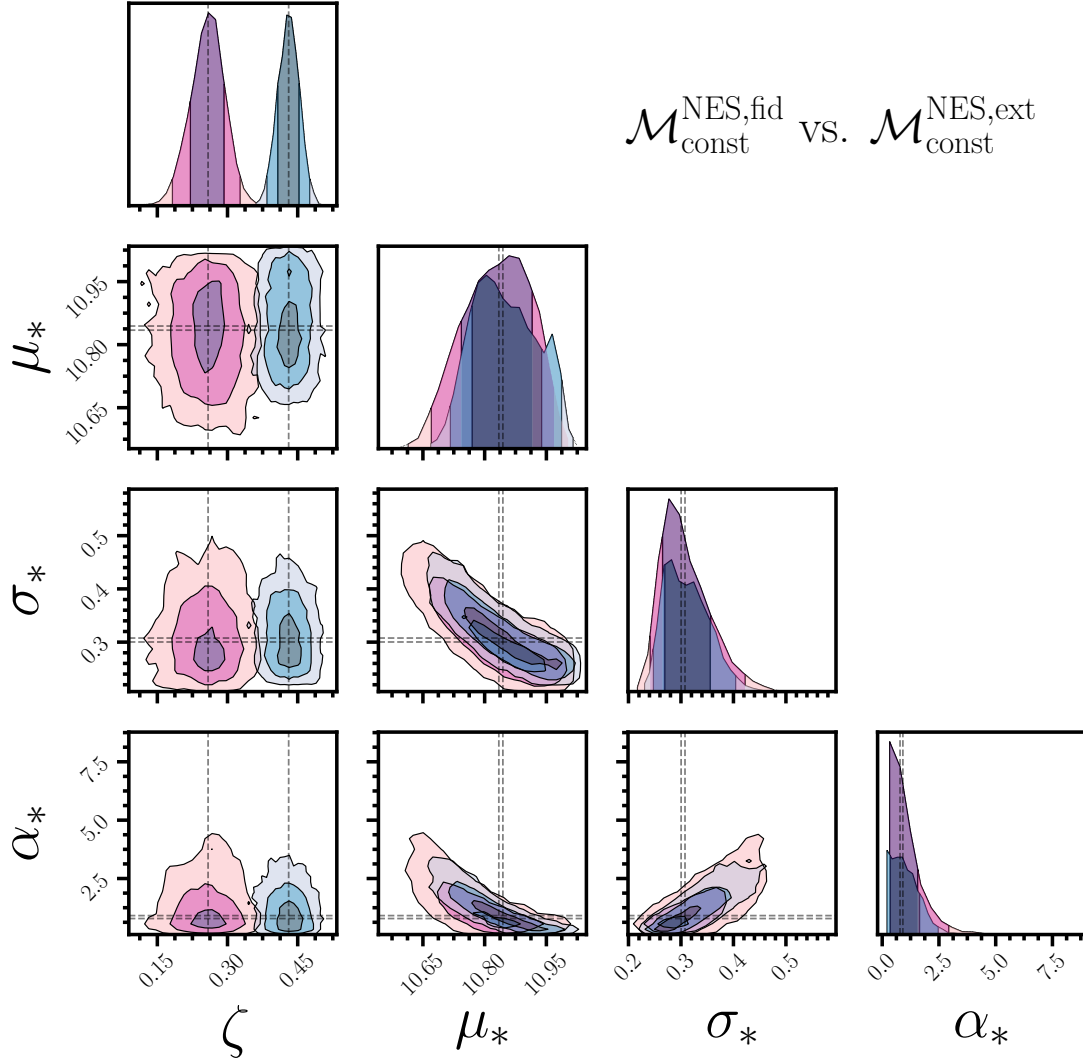


Figure 4. Same as Figure 3, but for the models $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ (pink contours) and $\mathcal{M}_{\text{const}}^{\text{NES, ext}}$ (azure contours; see Tables 4 and 5).

4.3 Comparing the results for the fiducial and extended samples

In Figures 4-6 we compare the posterior PDFs of the hyper-parameters obtained for models $\mathcal{M}_{\text{const}}^{\text{NES}}$, $\mathcal{M}_{\text{const}}$ and \mathcal{M}_{evo} applied to the fiducial and extended samples. In all cases the main difference between the fiducial-sample and extended-sample cases is in the distribution of the hyper-parameter ζ , which quantifies the redshift-evolution of the normalisation of the $M_* - \sigma_e$ relation. The median values are $\zeta = 0.26 - 0.29$ for the fiducial sample and $\zeta = 0.42 - 0.49$ for the extended sample, so the evolution of σ_0 at given M_* is stronger for the extended sample than for the fiducial sample. The differences in the distributions of ζ of the same model applied to the fiducial and to the extended samples are about $2-2.5\sigma$.

The distributions of all the other hyper-parameters (for instance, η , which quantifies the evolution of the slope, or ξ , which quantifies the evolution of the intrinsic scatter) are instead consistent within 1σ when comparing the results for the fiducial and extended samples.

The best model of the fiducial sample (model $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$) is directly compared to the best models of the extended sample ($\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$) in the top panels of Figure 9, in which the correlations are plotted at $z = 0$, $z = 0.5$ and $z = 1$. The aforementioned discrepancy in ζ is apparent in the $z = 1$ panel of this figure, in which the fiducial-sample and extended-sample curves are offset by 0.15-0.2 dex in M_* at given σ_e . The origin of this offset might be due to different factors. A contribution could come from observa-

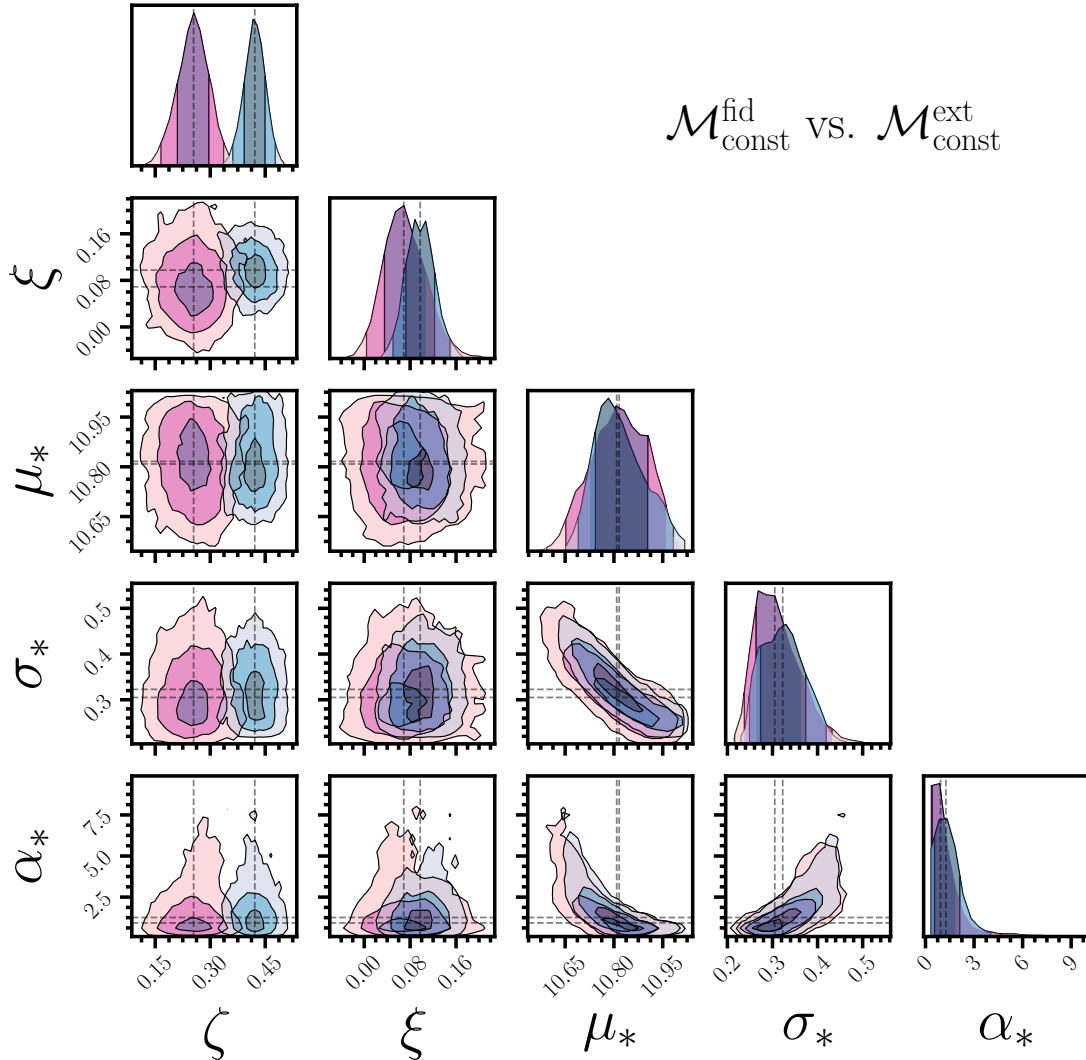


Figure 5. Same as [Figure 3](#), but for the models $\mathcal{M}_{\text{const}}^{\text{fid}}$ (pink contours) and $\mathcal{M}_{\text{const}}^{\text{ext}}$ (azure contours; see [Tables 4 and 5](#)).

tional biases. For instance, while the estimates of the stellar masses of the fiducial sample are consistent with each other, for the extended sample the measurements are derived independently from different studies. The offset in M_* is comparable to the mean uncertainty of the stellar mass estimates in the high-redshift sample, so systematic differences in the stellar mass estimates might well contribute to this offset. As a further test, we applied our model to the high-redshift sample of galaxies, finding at $z = 1$ a $M_* - \sigma_e$ relation offset by $0.2 - 0.25$ dex in M_* with respect to the best model of the fiducial sample. Such offset appears hard to explain entirely with systematic effects in the stellar mass estimates. This suggests that the higher values of ζ found for the extended sample might be at least partly due to the fact that

the redshift-dependence of the normalisation of the $M_* - \sigma_e$ relation is actually stronger at higher redshift.

5 COMPARISON WITH PREVIOUS WORKS

In this section we compare our results on the $M_* - \sigma_0$ relation with previous works in the literature. Specifically, we compare with the studies of [Auger et al. \(2010\)](#), [Hyde & Bernardi \(2009a\)](#), [Zahid et al. \(2016b\)](#), [Belli et al. \(2014\)](#) and [Mason et al. \(2015\)](#), which we briefly describe in the following. All the authors assumed a Chabrier IMF for their estimates of the stellar masses, except for [Mason et al. \(2015\)](#) who assumed a Salpeter IMF ([Salpeter 1955](#)).

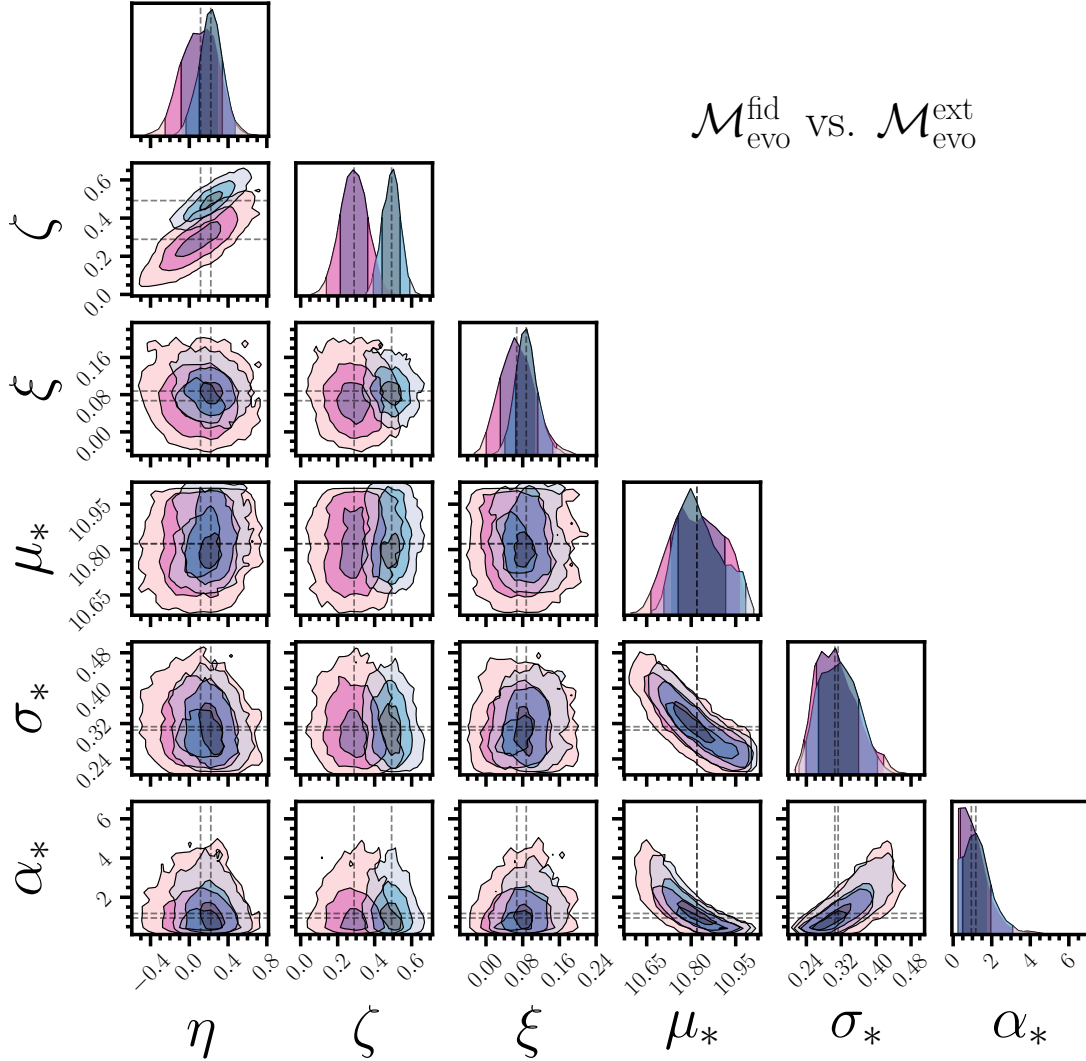


Figure 6. Same as Figure 3, but for the models $\mathcal{M}_{\text{evo}}^{\text{fid}}$ (pink contours) and $\mathcal{M}_{\text{evo}}^{\text{ext}}$ (azure contours; see Tables 4 and 5).

- Auger et al. (2010) take into account 59 ETGs (morphologically classified as ellipticals or S0s) identified as strong gravitational lenses in the Sloan Lens ACS Survey (SLACS) (Bolton et al. 2008; Auger et al. 2009) with a mean redshift $z \approx 0.2$. The stellar masses of these ETGs span the range $11 < \log(M_*/M_\odot) < 12$. The measured velocity dispersion is corrected to $\sigma_{e/2}$, the velocity dispersion within an aperture $R_e/2$, by applying the correction of Jorgensen et al. (1995). Auger et al. (2010) report best fits of the stellar mass-velocity dispersion relation both allowing and not allowing for the presence of intrinsic scatter.

- Hyde & Bernardi (2009a) extract 46410 ETGs from the SDSS DR4 with parameters updated to the DR6 values (Adelman-McCarthy et al. 2008). They include galaxies

with $60 < \sigma_{e/8}/(\text{km s}^{-1}) < 400$, where $\sigma_{e/8}$ is the stellar velocity dispersion measured within an aperture $R_e/8$. Hyde & Bernardi (2009a) consider both linear and quadratic fits to the $\log M_* - \log \sigma_{e/8}$ relation. Here we compare our results with their linear fit for their sample of ETGs in the range $10.5 < \log(M_*/M_\odot) < 11.5$ and $-23 < M_r < -20.5$, where M_r is the absolute magnitude in the r band. This sample spans the redshift range $0.07 < z \leq 0.35$.

- Zahid et al. (2016b) analyse the $M_* - \sigma_0$ relation for massive quiescent galaxies out to $z \approx 0.7$. For our comparison, we use their power-law fit obtained for a subsample drawn from the Smithsonian Hectospec Lensing Survey (SHELS) (Geller et al. 2005) at $0.5 < z < 0.6$, observed with the Hectospec spectrograph whose fibers have an aperture radius of

Table 6. Bayesian evidences \mathcal{Z} and Bayes factors \mathcal{B} of the models. The Bayes factors are relative to the Bayesian evidences of model $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ for the fiducial sample and of model $\mathcal{M}_{\text{const}}^{\text{ext}}$ for the extended sample, i.e. the models with the highest evidences for given sample.

| Model | $\ln \mathcal{Z}$ | $\ln \mathcal{B}$ |
|--|---------------------|--------------------|
| $\mathcal{M}_{\text{const}}^{\text{fid}}$ | 101.201 ± 0.187 | -1.904 ± 0.351 |
| $\mathcal{M}_{\text{evo}}^{\text{fid}}$ | 100.238 ± 0.192 | -2.867 ± 0.356 |
| $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ | 103.105 ± 0.164 | – |
| $\mathcal{M}_{\text{evo}}^{\text{NES, fid}}$ | 101.618 ± 0.176 | -1.487 ± 0.34 |
| $\mathcal{M}_{\text{const}}^{\text{ext}}$ | 168.65 ± 0.191 | – |
| $\mathcal{M}_{\text{evo}}^{\text{ext}}$ | 167.741 ± 0.203 | -0.909 ± 0.394 |
| $\mathcal{M}_{\text{const}}^{\text{NES, ext}}$ | 162.75 ± 0.174 | -5.9 ± 0.365 |
| $\mathcal{M}_{\text{evo}}^{\text{NES, ext}}$ | 164.373 ± 0.187 | -4.277 ± 0.378 |

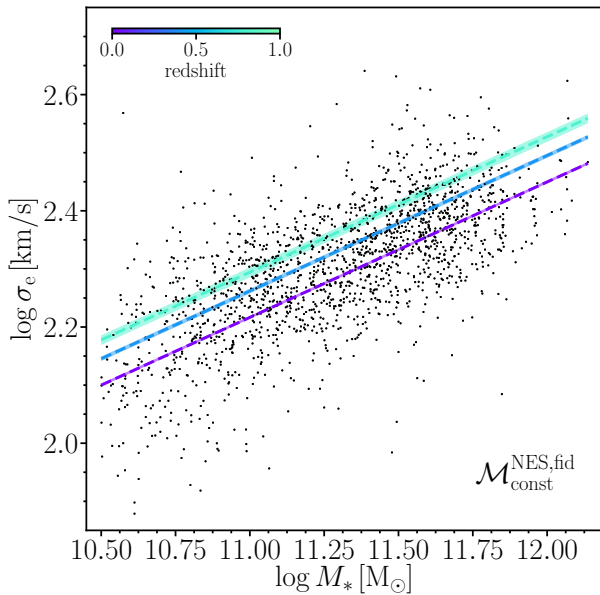


Figure 7. Central stellar velocity dispersion within $1 R_e$ as a function of the stellar mass. The dashed lines represent the median relations at three representative redshifts ($z = 0, 0.5$ and 1) for the model $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$. The shaded areas indicate the 1σ uncertainty ranges. The intrinsic scatter is *not* shown in these plots. The black dots represent the ETGs of the fiducial sample.

$0.75''$. Concerning SHELS galaxies, at $z \approx 0.1$ velocity dispersion is measured in an aperture of around 1.4 kpc . As highlighted by [Damjanov et al. \(2019\)](#), though the observed velocity dispersions are not corrected for the aperture, these measurements are estimated to differ at most by 3–4% from velocity dispersions within R_e .

- [Belli et al. \(2014\)](#) measure σ_e and M_* for a sample of galaxies with median redshift $z \approx 1.23$. We take from [Zahid et al. \(2016b\)](#) the best fit parameters of the sample of [Belli et al. \(2014\)](#).

Table 7. Values of the parameters of [equation \(31\)](#), according to the fits of the considered literature works.

| Reference | μ_0 | β_0 |
|---|---------|-----------|
| Auger et al. (2010) | 2.41 | 0.24 |
| Auger et al. (2010) with scatter (± 0.04) | 2.40 | 0.18 |
| Belli et al. (2014) | 2.49 | 0.30 |
| Hyde & Bernardi (2009a) | 2.39 | 0.286 |
| Mason et al. (2015) | 2.44 | 0.18 |
| Zahid et al. (2016b) SHELS ($0.5 < z < 0.6$) | 2.37 | 0.26 |

- [Mason et al. \(2015\)](#) provide a fitting formula describing the redshift evolution of the $M_* - \sigma_e$ relation, assuming redshift-independent slope determined by the low- z relation measured by [Auger et al. \(2010\)](#). In particular, we compare our fit with the constraints inferred by [Mason et al.](#) at $z = 2$.

In order to make a direct comparison as homogeneous as possible between our work and the others in the literature, we rescale all the fit parameters to the following functional form

$$\log \left(\frac{\sigma_e}{\text{km s}^{-1}} \right) = \mu_0 + \beta_0 \log \left(\frac{M_*}{M_*^{\text{piv}}} \right), \quad (31)$$

where $M_*^{\text{piv}} = 10^{11.361} M_\odot$ (Chabrier IMF). The Salpeter $M_{*, \text{Salp}}$ stellar mass used in [Mason et al. \(2015\)](#) is converted into a Chabrier stellar mass $M_{*, \text{Chab}}$, assuming that a Salpeter IMF implies 0.25 dex higher stellar masses than a Chabrier IMF ([Bernardi et al. 2010](#)). Finally, to directly compare with our fit the fits of [Auger et al. \(2010\)](#) and [Hyde & Bernardi \(2009a\)](#), who report $\sigma_{e/2}$ and $\sigma_{e/8}$, respectively, we estimate $\sigma_{e/2}$ and $\sigma_{e/8}$ for our SDSS galaxies (using [equation 3](#)), we compute the median values of $\log \sigma_e - \log \sigma_{e/8} \approx -0.06$ and $\log \sigma_e - \log \sigma_{e/2} \approx -0.02$, and we correct the fits of [Auger et al. \(2010\)](#) and [Hyde & Bernardi \(2009a\)](#) using $\log \sigma_e = \log \sigma_{e/8} - 0.06$ and $\log \sigma_e = \log \sigma_{e/2} - 0.02$.

In [Figure 10](#), we show the comparison between our models and the previous works at $z = 0.2, 0.55, 1.23$ and 2 . At low redshift ($z = 0.2$; top-left panel of [Figure 10](#)) we compare our results with [Auger et al. \(2010\)](#) (two fits of which one accounting and the other not accounting for the intrinsic scatter) and [Hyde & Bernardi \(2009a\)](#). For clarity reasons, here we show only the $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ fit, because the $\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$ curves are essentially indistinguishable from that of $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ at this redshift, which is close to z^{piv} (see [Figure 9](#)). We find a good agreement with [Hyde & Bernardi \(2009a\)](#), whose fit is only slightly steeper than ours. Also the fits of [Auger et al.](#) have slope similar to that our model, but have normalisation slightly higher than our estimates, for instance by ≈ 0.09 dex in σ_e at $M_* \approx 10^{11} M_\odot$. This is probably related to the fact that the [Auger et al.](#) sample consists of strong lenses: at fixed stellar mass, the probability for a galaxy of being a strong lens increases with central velocity dispersion. Lensing selection effects might tend to bias a sample towards higher values of σ_e .

Moving towards higher redshifts, the fit of SHELS galaxies at $0.5 < z < 0.6$ of [Zahid et al. \(2016b\)](#) almost perfectly overlaps with our models $\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$ evaluated at redshift $z = 0.55$ (top-right panel of [Figure 10](#)). We

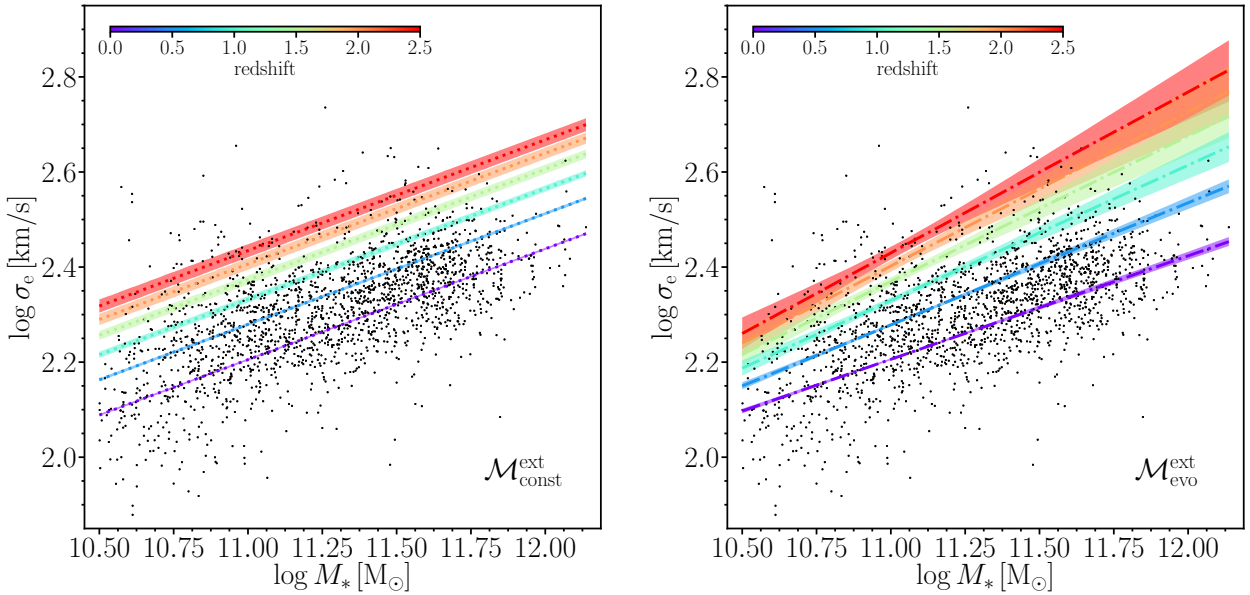


Figure 8. Same as Figure 7, but for model $M_{\text{const}}^{\text{ext}}$ (left-hand panel, dotted lines) and model $M_{\text{evo}}^{\text{ext}}$ (right-hand panel, dash-dotted lines). The lines represent the median relations at six representative redshifts ($z = 0, 0.5, 1, 1.5, 2$ and 2.5). The black dots represent the ETGs of the extended sample.

find also a good agreement with the $M_{\text{const}}^{\text{NES, fid}}$ model, which is only slightly shallower than the SHELS fit.

In the bottom-left panel of Figure 10 we compare our models evaluated at $z = 1.23$ (mean redshift of the sample of Belli et al. 2014) with the linear fit obtained by Belli et al. (2014). While the $M_{\text{const}}^{\text{ext}}$ is shallower than the fit by Belli et al. (2014), the median relation of $M_{\text{evo}}^{\text{ext}}$ model is remarkably similar in slope to Belli et al. (2014) relation, and differs only slightly in normalisation by ≈ 0.1 dex.

Finally, we compare the median relations of our models $M_{\text{const}}^{\text{ext}}$ and $M_{\text{evo}}^{\text{ext}}$ with the estimates done by Mason et al. (2015) at $z = 2$ (bottom-right panel of Figure 10). Both our curves are somewhat steeper than that of Mason et al. (2015), but we find a very good match at $M_* \approx 10^{11} M_{\odot}$.

Overall, we do find a satisfactory agreement among our results and previous works in the literature at all the explored redshifts. Some of the differences highlighted above may be ascribed to differences in the redshift distribution of the galaxy sample, stellar mass ranges, data and models used in the measurements of the stellar masses, selection criteria or fitting methods.

6 CONCLUSIONS

We have studied the evolution of the correlation between central stellar velocity dispersion σ_0 (here measured within R_e) and stellar mass M_* for massive ($M_* \gtrsim 10^{10.5} M_{\odot}$) ETGs observed in the redshift range $0 \lesssim z \lesssim 2.5$. We have modelled the evolution of this scaling law using a Bayesian hierarchical method. This allowed us to optimally exploit the available observational data, without resorting to binning in either redshift or stellar-mass space. The main conclusions of this work are the following.

- The central velocity dispersion of ETGs increases with

stellar mass following a power-law relation $\sigma_0 \propto M_*^{\beta}$ with $\beta \approx 0.23$, similar to the classical Faber-Jackson relation $L \propto \sigma_0^4$ (Faber & Jackson 1976), with stellar mass M_* replacing luminosity L .

- The normalisation of the $M_* - \sigma_0$ relation increases with redshift: at given stellar mass, $\sigma_0 \propto (1+z)^{\zeta}$ with $\zeta \approx 0.26$ in the redshift range $0 \lesssim z \lesssim 1.2$ probed by our fiducial sample. This implies that a typical ETG at $z \approx 0$ has σ_0 lower by about 20% than ETGs of similar mass at $z \approx 1$. Over the wider redshift range $0 \lesssim z \lesssim 2.5$, probed by our extended sample, we find $\zeta = 0.4 - 0.5$, which, modulo systematic effects (for instance in the estimate of the stellar mass), could suggest that the evolution of the normalisation of the $M_* - \sigma_0$ is stronger at higher redshift.

- In the redshift range $0 \lesssim z \lesssim 1.2$ the $M_* - \sigma_0$ relation has redshift-independent intrinsic scatter ≈ 0.07 dex in σ_0 at given M_* . The analysis of the extended sample leads to the conclusion that the intrinsic scatter is higher at higher redshift, up to ≈ 0.11 dex at $z = 2$.

- In the redshift range $0 \lesssim z \lesssim 2.5$, probed by the extended sample, a model in which the slope β of the $M_* - \sigma_0$ relation increases with redshift as $\beta \approx 0.2 + 0.2 \log(1+z)$ describes the data as well as the model with constant $\beta \approx 0.23$. Thus, current data do not allow to determine whether or not the $M_* - \sigma_0$ relation is actually steeper at higher redshift: additional observational data, especially at high redshift, appear necessary to address the question of the redshift-dependence of β .

The results of this work confirm and strengthen previous indications that the $M_* - \sigma_0$ relation of massive ETGs evolves with cosmic time. The theoretical interpretation of the observed evolution is not straightforward. Of course, the stellar mass of an individual galaxy can vary with time: it can increase as a consequence of mergers and star formation

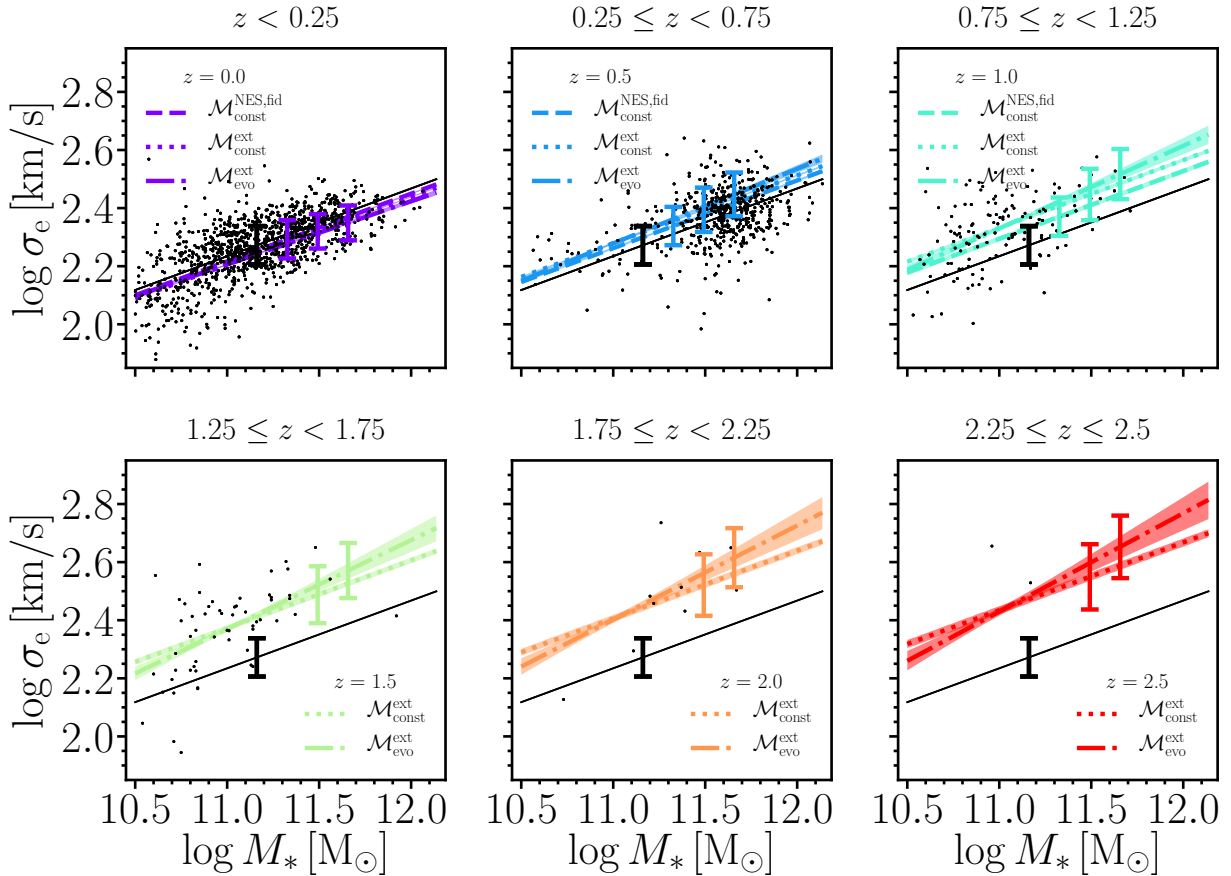


Figure 9. Central stellar velocity dispersion within $1R_e$ as a function of the stellar mass according to models $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ (dashed lines), $\mathcal{M}_{\text{const}}^{\text{ext}}$ (dotted lines) and $\mathcal{M}_{\text{evo}}^{\text{ext}}$ (dash-dotted lines) at six representative redshifts ($z = 0, 0.5, 1, 1.5, 2$ and 2.5). In each panel, the lines and the shaded areas indicate the median relations and 1σ bands, as in Figures 7 and 8, while the error bars indicate the intrinsic scatter at a given redshift. In all panels the black solid lines and the black error bars refer to the relation computed at $z \approx 0.17$ (model $\mathcal{M}_{\text{SDSS}}$). In each panel, the black dots represent the ETGs in the indicated redshift range.

and decrease as a consequence of mass return by ageing stellar populations. In the standard paradigm, the first effect is dominant, so we expect that, as cosmic time goes on, an individual galaxy moves in the $M_* - \sigma_0$ plane in the direction of increasing M_* . As pointed out in section 1, the variation of σ_0 for an individual galaxy is more uncertain: even pure dry mergers can make it increase or decrease depending on the merging orbital parameters and mass ratio. It is then clear that, at least qualitatively, the evolution shown in Figures 7 and 8 could be reproduced by individual galaxies evolving at decreasing σ_0 , but, at least at the low-mass end, even an evolution of individual galaxies at constant or slightly increasing σ_0 is not excluded. Remarkably, our results suggest that, on average, the stellar velocity dispersion of individual galaxies with $M_* \gtrsim 5 \times 10^{11} M_\odot$ at $z \approx 1$ must decrease from $z \approx 1$ to $z \approx 0$ for them to end up on the present-day $M_* - \sigma_0$ relation (see Figure 7).

An additional complication to the theoretical interpretation of the evolution of the scaling laws of ETGs is that it is not guaranteed that the high- z (say $z \approx 2$) ETGs are representative of the progenitors of all present-day ETGs. If the progenitors of some of the present-day ETGs were star-

forming at $z \approx 2$, they would not be included in our sample of $z \approx 2$ ETGs: this is the so-called progenitor bias, which must be accounted for when interpreting the evolution of a population of objects. However, the effect of progenitor bias should be small at least for the most massive ETGs in the redshift range $0 \lesssim z \lesssim 1$, in which the number density of quiescent galaxies shows little evolution (López-Sanjuan et al. 2012).

The theoretical interpretation of the evolution of the scaling relations of ETGs can benefit from the comparison of the observational data with the results of cosmological simulations of galaxy formation. In this approach, the progenitor bias can be taken into account automatically if simulated and observed galaxies are selected with consistent criteria. Moreover, in the simulations we can trace the evolution of individual galaxies, which is a crucial piece of information that we do not have for individual observed galaxies. The method presented in this paper is suitable to be applied to samples of simulated as well as observed galaxies. In the near future we plan to apply this method to compare the observed evolution of the $M_* - \sigma_0$ relation of ETGs with the

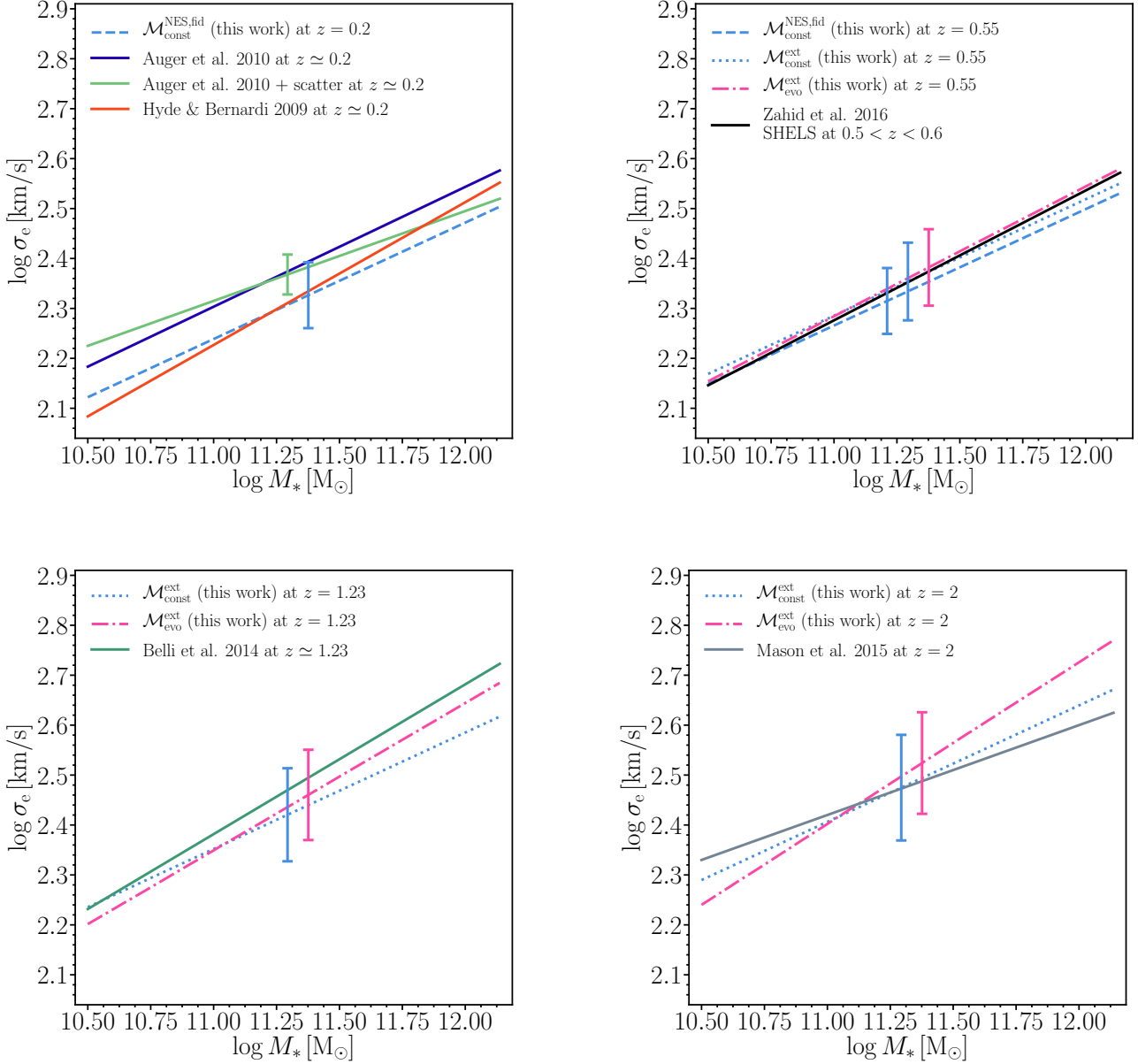


Figure 10. Comparison between the median $M_* - \sigma_e$ relations of our best models ($\mathcal{M}_{\text{const}}^{\text{NES, fid}}$, azure dashed line; $\mathcal{M}_{\text{const}}^{\text{ext}}$, azure dotted line; $\mathcal{M}_{\text{evo}}^{\text{ext}}$, pink dash-dotted line) and of previous works in the literature for the $M_* - \sigma_e$ relation at $z = 0.2$ (top-left panel), $z = 0.55$ (top-right panel), $z = 1.23$ (bottom-left panel) and $z = 2$ (bottom-right panel). Top-left panel: the light green and dark blue solid lines are the linear fits of Auger et al. (2010), including and not including the intrinsic scatter, respectively, while the orange solid line is the linear fit of Hyde & Bernardi (2009a). Here we do not show our fit at $z = 0.2$ for the $\mathcal{M}_{\text{const}}^{\text{ext}}$ and $\mathcal{M}_{\text{evo}}^{\text{ext}}$ models because at this redshift they are indistinguishable from that of the $\mathcal{M}_{\text{const}}^{\text{NES, fid}}$ model. Top-right panel: the black solid line is the linear fit of Zahid et al. (2016b) for the SHELS ($0.5 < z < 0.6$) sample. Bottom-left panel: the emerald green solid line represents the linear fit of Belli et al. (2014). Bottom-right panel: the grey solid line is the linear fit of Mason et al. (2015). The error bars indicate the intrinsic scatters.

results of state-of-the-art cosmological simulations of galaxy formation.

ACKNOWLEDGEMENTS

We are grateful to S. Belli for sharing data and for helpful suggestions, and to A. Cimatti, I. Damjanov and M. Moresco for useful discussions. AS acknowledges funding from the European Union's Horizon 2020 research and innovation programme under grand agreement No. 792916.

References

- Adelman-McCarthy J. K., et al., 2008, *ApJS*, **175**, 297
- Aihara H., et al., 2018, *PASJ*, **70**, S8
- Alam S., et al., 2015, *ApJS*, **219**, 12
- Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Bundy K., Moustakas L. A., 2009, *ApJ*, **705**, 1099
- Auger M. W., Treu T., Bolton A. S., Gavazzi R., Koopmans L. V. E., Marshall P. J., Moustakas L. A., Burles S., 2010, *ApJ*, **724**, 511
- Belli S., Newman A. B., Ellis R. S., 2014, *ApJ*, **783**, 117
- Belli S., Newman A. B., Ellis R. S., 2017, *ApJ*, **834**, 18
- Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, *MNRAS*, **404**, 2087
- Bertin E., Arnouts S., 1996, *A&AS*, **117**, 393
- Bezanson R., van Dokkum P., van de Sande J., Franx M., Kriek M., 2013a, *ApJ*, **764**, L8
- Bezanson R., van Dokkum P. G., van de Sande J., Franx M., Leja J., Kriek M., 2013b, *ApJ*, **779**, L21
- Bezanson R., Franx M., van Dokkum P. G., 2015, *ApJ*, **799**, 148
- Blakeslee J. P., et al., 2006, *ApJ*, **644**, 30
- Bolton A. S., Treu T., Koopmans L. V. E., Gavazzi R., Moustakas L. A., Burles S., Schlegel D. J., Wayth R., 2008, *ApJ*, **684**, 248
- Boylan-Kolchin M., Ma C.-P., Quataert E., 2006, *MNRAS*, **369**, 1081
- Bruzual G., Charlot S., 2003, *MNRAS*, **344**, 1000
- Buchner J., et al., 2014, *A&A*, **564**, A125
- Cappellari M., et al., 2006, *MNRAS*, **366**, 1126
- Cappellari M., et al., 2009, *ApJ*, **704**, L34
- Chabrier G., 2003, *PASP*, **115**, 763
- Cimatti A., Nipoti C., Cassata P., 2012, *MNRAS*, **422**, L62
- Ciotti L., Lanzoni B., Volonteri M., 2007, *ApJ*, **658**, 65
- Damjanov I., Zahid H. J., Geller M. J., Utsumi Y., Sohn J., Souchereau H., 2019, *ApJ*, **872**, 91
- Djorgovski S., Davis M., 1987, *ApJ*, **313**, 59
- Dressler A., Lynden-Bell D., Burstein D., Davies R. L., Faber S. M., Terlevich R., Wegner G., 1987, *ApJ*, **313**, 42
- Eddington A. S., 1913, *MNRAS*, **73**, 359
- Eisenstein D. J., et al., 2011, *AJ*, **142**, 72
- Faber S. M., Jackson R. E., 1976, *ApJ*, **204**, 668
- 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
- Fagotto F., Bressan A., Bertelli G., Chiosi C., 1994c, *A&AS*, **105**, 29
- Ferguson H. C., et al., 2004, *ApJ*, **600**, L107
- Feroz F., Hobson M. P., 2008, *MNRAS*, **384**, 449
- Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, **398**, 1601
- Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, *PASP*, **125**, 306
- Frigo M., Balcells M., 2017, *MNRAS*, **469**, 2184
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, **362**, 41
- Gargiulo A., Saracco P., Longhetti M., Tamburri S., Lonoce I., Ciocca F., 2015, *A&A*, **573**, A110
- Gargiulo A., Saracco P., Tamburri S., Lonoce I., Ciocca F., 2016, *A&A*, **592**, A132
- Geller M. J., Dell'Antonio I. P., Kurtz M. J., Ramella M., Fabricant D. G., Caldwell N., Tyson J. A., Wittman D., 2005, *ApJ*, **635**, L125
- Goodman J., Weare J., 2010, *Communications in Applied Mathematics and Computational Science*, **5**, 65
- Hilz M., Naab T., Ostriker J. P., 2013, *MNRAS*, **429**, 2924
- Hyde J. B., Bernardi M., 2009a, *MNRAS*, **394**, 1978
- Hyde J. B., Bernardi M., 2009b, *MNRAS*, **396**, 1171
- Jeffreys H., 1961, *Theory of Probability*, third edn. Oxford, Oxford, England
- Jorgensen I., Franx M., Kjaergaard P., 1995, *MNRAS*, **276**, 1341
- Kormendy J., 1977, *ApJ*, **218**, 333
- Le Fèvre O., et al., 2003, in Iye M., Moorwood A. F. M., eds, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. pp 1670–1681, doi:10.1117/12.460959
- López-Sanjuan C., et al., 2012, *A&A*, **548**, A7
- Mason C. A., et al., 2015, *ApJ*, **805**, 79
- Miyazaki S., et al., 2018, *PASJ*, **70**, S1
- Muzzin A., et al., 2013, *ApJ*, **777**, 18
- Naab T., Ostriker J. P., 2017, *ARA&A*, **55**, 59
- Naab T., Johansson P. H., Ostriker J. P., 2009, *ApJ*, **699**, L178
- Newman A. B., Ellis R. S., Treu T., Bundy K., 2010, *ApJ*, **717**, L103
- Nipoti C., Londrillo P., Ciotti L., 2003, *MNRAS*, **342**, 501
- Nipoti C., Treu T., Bolton A. S., 2009a, *ApJ*, **703**, 1531
- Nipoti C., Treu T., Auger M. W., Bolton A. S., 2009b, *ApJ*, **706**, L86
- Nipoti C., Treu T., Leauthaud A., Bundy K., Newman A. B., Auger M. W., 2012, *MNRAS*, **422**, 1714
- Onodera M., et al., 2012, *ApJ*, **755**, 26
- Oogi T., Habe A., 2013, *MNRAS*, **428**, 641
- Posti L., Nipoti C., Stiavelli M., Ciotti L., 2014, *MNRAS*, **440**, 610
- Robertson B., Cox T. J., Hernquist L., Franx M., Hopkins P. F., Martini P., Springel V., 2006, *ApJ*, **641**, 21
- Salpeter E. E., 1955, *ApJ*, **121**, 161
- Scoville N., et al., 2007, *ApJS*, **172**, 1
- Sérsic J. L., 1968, *Atlas de galaxies australes*
- Skilling J., 2004, *AIP Conference Proceedings*, **735**, 395
- Somerville R. S., Davé R., 2015, *ARA&A*, **53**, 51
- Sonnenfeld A., Wang W., Bahcall N., 2019, *A&A*, **622**, A30
- Straatman C. M. S., et al., 2018, *ApJS*, **239**, 27
- Tanaka M., et al., 2019, arXiv e-prints, p. arXiv:1909.10721
- Toft S., Gallazzi A., Zirm A., Wold M., Zibetti S., Grillo C., Man A., 2012, *ApJ*, **754**, 3
- Westera P., Lejeune T., Buser R., Cuisinier F., Bruzual G., 2002, *A&A*, **381**, 524
- Zahid H. J., Damjanov I., Geller M. J., Hwang H. S., Fabricant D. G., 2016a, *ApJ*, **821**, 101
- Zahid H. J., Geller M. J., Fabricant D. G., Hwang H. S., 2016b, *ApJ*, **832**, 203
- van Dokkum P. G., Kriek M., Franx M., 2009, *Nature*, **460**, 717
- van de Sande J., et al., 2013a, *ApJ*, **771**, 85
- van de Sande J., et al., 2013b, *ApJ*, **771**, 85
- van der Wel A., Holden B. P., Zirm A. W., Franx M., Rettura A., Illingworth G. D., Ford H. C., 2008, *ApJ*, **688**, 48
- van der Wel A., et al., 2014, *ApJ*, **788**, 28
- van der Wel A., et al., 2016, *ApJS*, **223**, 29

APPENDIX A: DETAILS OF THE CALCULATION OF THE LIKELIHOOD USED IN THE MODEL-DATA COMPARISON

Here we provide some steps of the calculation of the likelihood in [equation \(14\)](#). In this section all masses are in units of solar masses. By writing explicitly each term in [equation \(8\)](#) for our case, we obtain

$$P(\mathbf{d}|\Phi) = \prod_{i=1}^n \int d \log M_{*,i} d \log \sigma_{e,i} dz_i P(\log M_{*,i}^{\text{obs}}, \log \sigma_{e,i}^{\text{obs}}, z_i^{\text{obs}} | \log M_{*,i}, \log \sigma_{e,i}, z_i) P(\log M_{*,i}, \log \sigma_{e,i}, z_i | \Phi). \quad (\text{A1})$$

As explained in [subsection 3.2](#), we neglect the uncertainty on redshift, so that the first term on the right-side of [equation \(A1\)](#) becomes

$$P(\log M_{*,i}^{\text{obs}}, \log \sigma_{e,i}^{\text{obs}}, z_i^{\text{obs}} | \log M_{*,i}, \log \sigma_{e,i}, z_i) = P(\log M_{*,i}^{\text{obs}} | \log M_{*,i}) P(\log \sigma_{e,i}^{\text{obs}} | \log \sigma_{e,i}) \delta(z_i^{\text{obs}} - z_i). \quad (\text{A2})$$

Therefore, we can rewrite [equation \(14\)](#) for the i -th galaxy as follows:

$$\begin{aligned} P(\log M_{*,i}^{\text{obs}}, \log \sigma_{e,i}^{\text{obs}}, z_i^{\text{obs}} | \Phi) &= \int d \log M_{*,i} \frac{\mathcal{A}(\log M_{*,i})}{\sqrt{2\pi\sigma_{M_{*,i}}^2}} \exp \left\{ -\frac{(\log M_{*,i} - \log M_{*,i}^{\text{obs}})^2}{2\sigma_{M_{*,i}}^2} \right\} \frac{1}{\sqrt{2\pi\sigma_*^2}} \exp \left\{ -\frac{(\log M_{*,i} - \mu_*)^2}{2\sigma_*^2} \right\} \mathcal{E}(\log M_{*,i} | \Phi) \times \\ &\times \int d \log \sigma_{e,i} \frac{1}{\sqrt{2\pi\sigma_{\sigma_{e,i}}^2}} \exp \left\{ -\frac{(\log \sigma_{e,i} - \log \sigma_{e,i}^{\text{obs}})^2}{2\sigma_{\sigma_{e,i}}^2} \right\} \frac{1}{\sqrt{2\pi\sigma_{\sigma,i}^2}} \exp \left\{ -\frac{(\log \sigma_{e,i} - \mu_{\sigma,i})^2}{2\sigma_{\sigma,i}^2} \right\}, \end{aligned} \quad (\text{A3})$$

where

$$\mu_{\sigma,i} = \mu_0 + \beta \log \left(\frac{M_{*,i}}{M_*^{\text{piv}}} \right) + \zeta \log \left(\frac{1 + z_i}{1 + z^{\text{piv}}} \right) \quad (\text{A4})$$

and

$$\sigma_{\sigma,i} = \psi_0 + \xi \log \left(\frac{1 + z}{1 + z^{\text{piv}}} \right). \quad (\text{A5})$$

In [equation \(A3\)](#), the term $\mathcal{A}(\log M_*)$ allows to normalise the distribution over all values of the observed stellar mass. Specifically, $\mathcal{A}(\log M_*)$ ensures that the probability of having an ETG with $\log M_*^{\text{obs}}$ between 10.5 (the lower bound of the considered observed stellar mass interval) and $+\infty$ is one:

$$\int_{10.5}^{\infty} d \log M_{*,i}^{\text{obs}} \frac{\mathcal{A}(\log M_{*,i})}{\sqrt{2\pi\sigma_{M_{*,i}}^2}} \exp \left\{ -\frac{(\log M_{*,i} - \log M_{*,i}^{\text{obs}})^2}{2\sigma_{M_{*,i}}^2} \right\} = 1. \quad (\text{A6})$$

Hence, $\mathcal{A}(\log M_{*,i})$ is given by

$$\mathcal{A}(\log M_{*,i}) = \frac{1}{\int_{10.5}^{\infty} dM' \frac{1}{\sqrt{2\pi\sigma_{M_{*,i}}^2}} \exp \left\{ -\frac{(\log M_{*,i} - M')^2}{2\sigma_{M_{*,i}}^2} \right\}} = \left[\frac{\sqrt{\sigma_{M_{*,i}}^2}}{\frac{1}{2}\sigma_{M_{*,i}}^2 \operatorname{erf} \left(\frac{\sqrt{2} \log M_{*,i} - M}{\sigma_{M_{*,i}}^2} \right)} \right]_{10.5}^{+\infty}. \quad (\text{A7})$$

The integral term in $d \log \sigma_{e,i}$ of [equation \(A3\)](#) can be written as

$$\frac{1}{\sqrt{2\pi(\sigma_{\sigma_{e,i}}^2 + \sigma_{\sigma,i}^2)}} \exp \left\{ -\frac{(\log \sigma_{e,i}^{\text{obs}} - \mu_{\sigma,i})^2}{2(\sigma_{\sigma_{e,i}}^2 + \sigma_{\sigma,i}^2)} \right\} \int d \log \sigma_{e,i} \frac{1}{\sqrt{2\pi\tilde{\sigma}_i^2}} \exp \left\{ -\frac{(\log \sigma_{e,i} - \tilde{\mu}_i)^2}{2\tilde{\sigma}_i^2} \right\} = \frac{1}{\sqrt{2\pi(\sigma_{\sigma_{e,i}}^2 + \sigma_{\sigma,i}^2)}} \exp \left\{ -\frac{(\log \sigma_{e,i}^{\text{obs}} - \mu_{\sigma,i})^2}{2(\sigma_{\sigma_{e,i}}^2 + \sigma_{\sigma,i}^2)} \right\}, \quad (\text{A8})$$

where

$$\tilde{\mu}_i = \frac{\log \sigma_{e,i}^{\text{obs}} \sigma_{\sigma,i}^2 + \mu_{\sigma,i} \sigma_{\sigma_{e,i}}^2}{\sigma_{\sigma,i}^2 + \sigma_{\sigma_{e,i}}^2} \quad \text{and} \quad \tilde{\sigma}_i = \sqrt{\frac{\sigma_{\sigma,i}^2 \sigma_{\sigma_{e,i}}^2}{\sigma_{\sigma,i}^2 + \sigma_{\sigma_{e,i}}^2}}. \quad (\text{A9})$$

By writing $\mu_{\sigma,i}$ explicitly, [equation \(A3\)](#) becomes

$$\begin{aligned} P(\log M_{*,i}^{\text{obs}}, \log \sigma_{e,i}^{\text{obs}}, z_i^{\text{obs}} | \Phi) &= \int d \log M_{*,i} \frac{\mathcal{A}(\log M_{*,i})}{\sqrt{2\pi\sigma_{M_{*,i}}^2}} \exp \left\{ -\frac{(\log M_{*,i} - \log M_{*,i}^{\text{obs}})^2}{2\sigma_{M_{*,i}}^2} \right\} \frac{1}{\sqrt{2\pi\sigma_*^2}} \exp \left\{ -\frac{(\log M_{*,i} - \mu_*)^2}{2\sigma_*^2} \right\} \mathcal{E}(\log M_{*,i} | \Phi) \times \\ &\times \frac{1}{\sqrt{2\pi\sigma_{\text{eff},i}|\beta|}} \exp \left\{ -\frac{(\log M_{*,i} - \mu_{\text{eff},i})^2}{2\sigma_{\text{eff},i}^2} \right\} \end{aligned}$$

(A10)

with

$$\mu_{\text{eff},i} = \log M_*^{\text{piv}} + \frac{\log \sigma_{e,i}^{\text{obs}} - \log \sigma_0 - \zeta [\log(1+z_i) - \log(1+z^{\text{piv}})]}{\beta} \quad \text{and} \quad \sigma_{\text{eff},i} = \frac{(\sigma_{\sigma_e,i}^2 + \sigma_{\sigma,i}^2)}{\beta^2}. \quad (\text{A11})$$

By combining the Gaussian terms centred in $\log M_{*,i}^{\text{obs}}$ and $\mu_{\text{eff},i}$, we obtain the final form of the likelihood:

$$\begin{aligned} \text{P}(\log M_{*,i}^{\text{obs}}, \log \sigma_{e,i}^{\text{obs}}, z_i^{\text{obs}} | \Phi) &= \frac{1}{|\beta|} \frac{1}{\sqrt{2\pi(\sigma_{M_{*,i}}^2 + \sigma_{\text{eff},i}^2)}} \exp \left\{ -\frac{(\log M_{*,i}^{\text{obs}} - \mu_{\text{eff},i})^2}{2(\sigma_{M_{*,i}}^2 + \sigma_{\text{eff},i}^2)} \right\} \times \\ &\times \int d \log M_{*,i} \frac{1}{\sqrt{2\pi\sigma_i'^2}} \exp \left\{ -\frac{(\log M_{*,i} - \mu_i')^2}{2\sigma_i'^2} \right\} \mathcal{A}(\log M_{*,i}) \mathcal{S}(\log M_{*,i}), \end{aligned} \quad (\text{A12})$$

with

$$\mu_i' = \frac{\log M_{*,i} \sigma_{\text{eff},i}^2 + \mu_{\text{eff},i} \sigma_{M_{*,i}}^2}{\sigma_{\text{eff},i}^2 + \sigma_{M_{*,i}}^2} \quad \text{and} \quad \sigma_i' = \sqrt{\frac{\sigma_{\text{eff},i}^2 \sigma_{M_{*,i}}^2}{\sigma_{\text{eff},i}^2 + \sigma_{M_{*,i}}^2}}. \quad (\text{A13})$$

We compute the integral term in [equation \(A12\)](#) numerically, using the trapezoidal rule.

APPENDIX B: MOCK CATALOGUE

In order to check the reliability of our method, we performed some tests on mock catalogues. In the following, we provide an example on a mock catalogue of around 1000 ETGs generated as follows (masses are in units of M_\odot and velocity dispersions in units of km s^{-1}):

- the true stellar masses $\log M_*^t$ have been extracted randomly from a normal distribution centred in 10.7 with standard deviation 0.5;
- the true velocity dispersions $\log \sigma_e^t \equiv \mu^t$ have been derived from a normal distribution centred in

$$\mu^t = \mu_0^{\text{mock}} + \beta_0^{\text{mock}} \log \left(\frac{M_*^t}{M_*^{\text{mock}}} \right) \quad (\text{B1})$$

with $M_*^{\text{mock}} = 10^{11.035}$, $\mu_0^{\text{mock}} = 2.20$, $\beta_0^{\text{mock}} = 0.32$ and standard deviation $\psi_0^{\text{mock}} = 0.06$ dex;

- the errors on the stellar masses σ_{M_*} have been arbitrarily assumed 0.15 dex for all galaxies;
- the errors on the velocity dispersions σ_{σ_e} have been picked randomly from the normal distribution $\mathcal{N}(\mu \simeq 0.05, \sigma \simeq 0.02)$;
- the values of $\log M_*^{\text{obs}}$ and $\log \sigma_e^{\text{obs}}$ are extracted from $\mathcal{N}(\mu = \log M_*^t, \sigma = \sigma_M)$ and $\mathcal{N}(\mu = \log \sigma_e^t, \sigma = \sigma_{\sigma_e})$, respectively;
- finally, we apply the same selection in observed stellar mass used in this work, i.e. $\log M_* \geq 10.5$.

In order to sample the PDFs of the model applied to our mock catalogue (hereafter, model $\mathcal{M}_{\text{mock}}$), we perform a MCMC run (see [subsection 3.3](#)), using 50 random walkers running for 800 steps to reach the convergence of the hyper-parameter distribution. In [Figure B1](#), the obtained μ_0 , β_0 and σ_0 hyper-parameters with their 1σ uncertainties for model $\mathcal{M}_{\text{mock}}$ are reported and we show the PDFs of all hyper-parameters computed by our fitting procedure. The input values of the hyper-parameters are all recovered within 1σ .

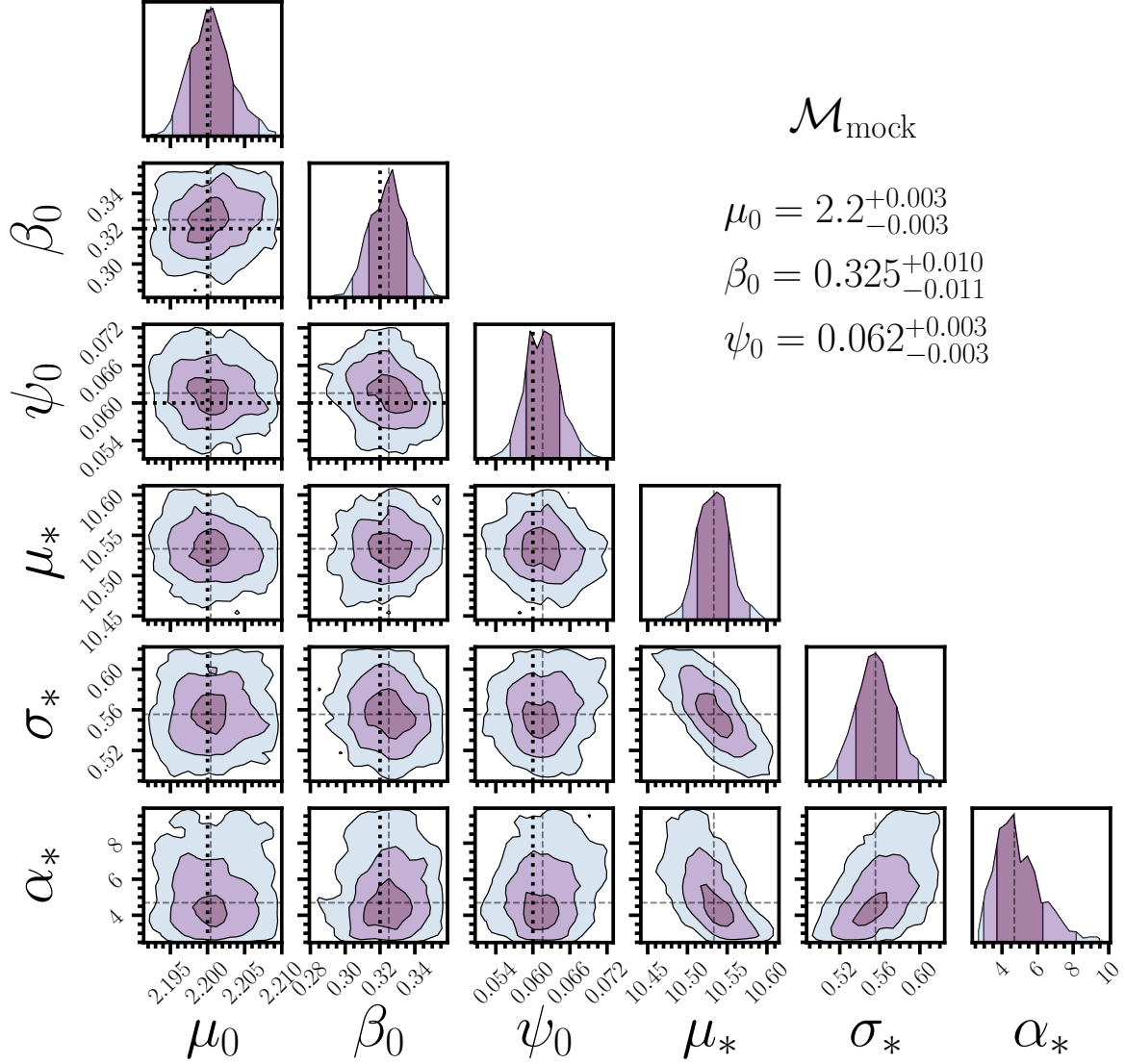


Figure B1. Same as Figure 3, but for the mock catalogue. The black dotted lines represent the true values of the hyper-parameters μ_0^{mock} , β_0^{mock} and ψ_0^{mock} .

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.