

*This article has been accepted for publication in Monthly Notices of the Royal Astronomical Society ©: 2020 The Authors. Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.*

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

Carlo Cannarozzo <sup>1,2★</sup>, Alessandro Sonnenfeld <sup>3</sup> and Carlo Nipoti <sup>1</sup>

<sup>1</sup>Dipartimento di Fisica e Astronomia, Alma Mater Studiorum Università di Bologna, Via Piero Gobetti 93/2, I-40129 Bologna, Italy

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

<sup>3</sup>Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, the Netherlands

Accepted 2020 July 16. Received 2020 July 6; in original form 2019 October 12

## ABSTRACT

We study the evolution of the observed correlation between central stellar velocity dispersion  $\sigma_e$  and stellar mass  $M_*$  of massive ( $M_* \gtrsim 3 \times 10^{10} M_\odot$ ) early-type galaxies (ETGs) out to redshift  $z \approx 2.5$ , taking advantage of a Bayesian hierarchical inference formalism. Collecting ETGs from state-of-the-art literature samples, we build a *fiducial* sample ( $0 \lesssim z \lesssim 1$ ), which is obtained with homogeneous selection criteria, but also a less homogeneous *extended* sample ( $0 \lesssim z \lesssim 2.5$ ). Based on the fiducial sample, we find that at  $z \lesssim 1$  the  $M_*$ – $\sigma_e$  relation is well represented by  $\sigma_e \propto M_*^\beta (1+z)^\zeta$ , with  $\beta \simeq 0.18$  independent of redshift and  $\zeta \simeq 0.4$  (at a given  $M_*$ ,  $\sigma_e$  decreases for decreasing  $z$ , for instance by a factor of  $\approx 1.3$  from  $z = 1$  to  $z = 0$ ). When the slope  $\beta$  is allowed to evolve, we find it increasing with redshift:  $\beta(z) \simeq 0.16 + 0.26 \log(1+z)$  describes the data as well as constant  $\beta \simeq 0.18$ . The intrinsic scatter of the  $M_*$ – $\sigma_e$  relation is  $\simeq 0.08$  dex in  $\sigma_e$  at given  $M_*$ , independent of redshift. Our results suggest that, on average, the velocity dispersion of *individual* massive ( $M_* \gtrsim 3 \times 10^{11} M_\odot$ ) ETGs decreases with time while they evolve from  $z \approx 1$  to  $z \approx 0$ . The analysis of the extended sample, over the wider redshift range  $0 \lesssim z \lesssim 2.5$ , leads to results similar to that of the fiducial sample, with slightly stronger redshift dependence of the normalization ( $\zeta \simeq 0.5$ ) and weaker redshift dependence of the slope ( $d\beta/d\log(1+z) \simeq 0.18$ ) when  $\beta$  varies with time. At  $z = 2$  ETGs with  $M_* \approx 10^{11} M_\odot$  have, on average,  $\approx 1.7$  higher  $\sigma_e$  than ETGs of similar stellar mass at  $z = 0$ .

**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  $\sigma_0$ , 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  $\sigma_0$ ,  $L$ , 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 successful 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. 2013; Belli, Newman & Ellis 2014a; Gargiulo et al. 2016; Belli, Newman & Ellis 2017; Tanaka et al. 2019). Interestingly, the stellar mass Fundamental Plane, relating  $M_*$ ,  $\sigma_0$ , and  $R_e$  appears to change little with redshift (Bezanson et al. 2013b; Bezanson, Franx & van Dokkum 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 formation in the Universe occurs as a consequence of the collapse and virialization of the dark matter haloes, 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.

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

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, Londrillo & Ciotti 2003; Naab, Johansson & Ostriker 2009; Hilz, Naab & Ostriker 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; Oogi & Habe 2013; Posti et al. 2014; 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  $\sigma_0$  is a very interesting quantity to consider. Even for purely dry mergers of spheroids,  $\sigma_0$  can increase, decrease or stay constant following a merger, depending on the merger mass ratio and orbital parameters (Boylan-Kolchin, Ma & Quataert 2006; Naab et al. 2009; Nipoti, Treu & Bolton 2009a; Nipoti et al. 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, Lanzoni & Volonteri 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 characterization of the evolution of the observed scaling relations of the ETGs is thus crucial to use them as test beds for theoretical models. On 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, Nipoti & Cassata 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 (SNR), the study of the redshift evolution of correlations involving  $\sigma_0$ , 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 characterize 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 characterization 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 homogenizing 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  $\sigma_e$ , 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  $\beta$  and the case in which  $\beta$  is free to vary with redshift.

The paper is organized 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 discussed in Section 5. Section 6 concludes. Throughout this work, we adopt a standard Lambda 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 definition of what constitutes an ETG is based mainly on morphology, with the addition of cuts on emission line equivalent width of [O II] aimed at removing star-forming galaxies (as explained in the rest of this section). 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 (i) select galaxies and measure their stellar mass and velocity dispersion in a homogeneous way, and (ii) ensure that, at any given redshift and stellar mass, our selection criteria do not depend, either directly or indirectly, on velocity dispersion. With our main focus on accuracy, we first define a *fiducial sample* of galaxies, for which conditions (i) and (ii) above are satisfied. We drew our fiducial sample from the Sloan Digital Sky Survey (SDSS; Eisenstein et al. 2011) and the Large Early Galaxy Astrophysics Census (LEGA-C; van der Wel et al. 2016). For the galaxies in this sample, we strictly apply consistent selection criteria and 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 two surveys cover the redshift range  $0 \lesssim z \lesssim 1$  and, most importantly, have well-defined selection functions, which is critical to meet condition (ii).

We then define a second *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. We also define an *extended* sample, obtained by combining the fiducial and high-redshift samples. In building our samples, we include only galaxies with stellar mass higher than a minimum mass  $M_{*,\text{min}}$ , which in general depends both on the survey and on  $z$  (see Sections 2.1 and 2.2): in all cases  $M_{*,\text{min}} \geq 10^{10.5} M_\odot$ , which we adopt as absolute lower limit in stellar mass.

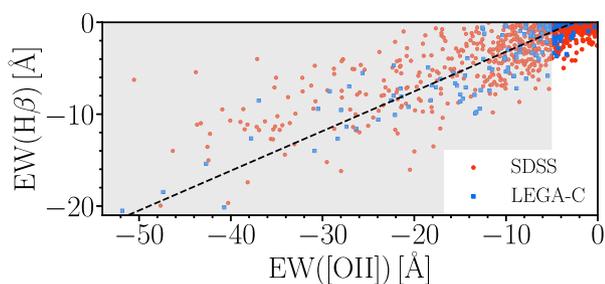
Our strategy is to carry out our inference on both the fiducial and the extended samples. Given the way the samples are built, we expect our results at  $z < 1$  to be more robust (i.e. less prone to observational biases), but it is nevertheless very interesting to examine trends out to  $z \approx 2.5$ , as probed by our extended sample. In the following two 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 (Strauss et al. 2002). 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; Muzzini et al. 2013).

#### 2.1.1 ETG selection

As anticipated, our definition of ETG is based mostly on morphology. For the morphological classification, we opted for visual inspection



**Figure 1.** Equivalent width of  $H\beta$ ,  $EW(H\beta)$ , as a function of equivalent width of  $[O II]$ ,  $EW([O II])$ , for galaxies drawn from the original catalogues of SDSS (circles) and LEGA-C (squares). For LEGA-C galaxies, we show only objects with  $SNR > 10$ . The black dashed line represents a linear fit to the data. Galaxies in the shaded region of the diagram ( $EW([O II]) < -5 \text{ \AA}$ ) are excluded from our sample of ETGs.

because the number of galaxies of our sample is relatively small. Valid alternatives, which are necessarily preferable for larger data sets, are automated morphological classification algorithms (e.g. Domínguez Sánchez et al. 2018). Before the visual inspection, we applied a pre-selection based on star formation activity: we removed star-forming galaxies from our sample, under the assumption that they are mostly associated with a late-type or irregular morphology. We relied on the presence of emission lines in the spectra of our galaxies as an indicator of star formation activity. In particular, we applied a selection based on the equivalent width of the forbidden emission line doublet of  $[O II]$ ,  $EW([O II]) \lambda\lambda 3726, 3729$ . We included only those galaxies that have  $EW([O II]) \geq -5 \text{ \AA}$ , where  $EW([O II])$  of SDSS and LEGA-C galaxies are obtained from the respective data release catalogues. Although  $[O II]$  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 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 used  $[O II]$  as a first step towards obtaining a sample of ETGs. Nevertheless, we found a good correlation between  $EW([O II])$  and  $EW(H\beta)$  for those galaxies drawn from the original catalogues of SDSS and LEGA-C for which both measurements are available (see Fig. 1). Although half of the LEGA-C galaxies do not have values of  $EW([O II])$  in the DR2 catalogue, these are for the most part objects at the low end of the redshift range,  $z < 0.8$ .

The second step in our selection is to include only galaxies with an early-type morphology, according to visual inspection. We used imaging data from the Wide layer of the HSC DR1, for this purpose. The Wide layer of HSC covers approximately 108 square degrees. The number of SDSS main sample galaxies present in this data set is  $\approx 2000$ , 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  $\simeq 1.3 \text{ deg}^2$  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 ( $\approx 1700$ ) 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 arcsec). 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 obtained cutouts in the  $g$ ,  $r$ ,  $i$ ,  $z$ , and  $y$  filters, and then visually inspected colour-composite RGB images made using the  $g$ -,  $r$ - and  $i$ -band data. We removed objects showing any presence of discs, spiral arms, 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). Such objects account for roughly 50 per cent of the inspected galaxies. Additionally, a few per cent of the objects were removed because of contamination from stars, and an even smaller fraction was eliminated because of the presence of close neighbours that make it difficult to carry out accurate photometric measurements. Although this last step could in principle introduce a bias in the inferred  $M_*$ – $\sigma_e$  relation in case this varies as a function of environment, given the small fraction of objects with close neighbours removed, any such bias will in any case be very small.

In Figs 2 and 3, we show colour-composite images of example sets of SDSS galaxies included and excluded from our sample on the basis of our morphological classification.

### 2.1.2 Photometric measurements

Our procedure for measuring stellar masses of the galaxies in the fiducial sample consisted in fitting stellar population synthesis (SPS) models to broad-band photometric data. Although photometric measurements for these galaxies are available from the literature, we chose to carry out new measurements using photometric data from the HSC survey. The data from the HSC survey are much deeper and have a much higher image quality compared to the SDSS data. This is important, because it allows for a cleaner detection and masking of foreground contaminants, and allows for a better characterization of the faint extended envelope of massive galaxies (see e.g. Huang et al. 2018). Moreover, by using the same data and procedure to estimate the stellar masses of the galaxies in the SDSS and LEGA-C samples, our inference on the evolution of the  $M_*$ – $\sigma_e$  relation is less prone to possible systematic effects related to the photometric measurements.

We estimated the  $g$ ,  $r$ ,  $i$ ,  $z$ , and  $y$  magnitudes of each galaxy by fitting a Sérsic surface brightness distribution to the data in these five bands simultaneously. In particular, we obtained  $201 \times 201$  pixel ( $\approx 34 \times 34$  arcsec) sky-subtracted cutouts of each galaxy in each band, we fitted the five-band data simultaneously with a seeing-convolved Sérsic surface brightness profile with elliptical isophotes and spatially uniform colours, while iteratively masking out foreground or background objects using the software SEXTRACTOR (Bertin & Arnouts 1996).

Saturated pixels were also masked, using the masks provided by HSC DR1. We added in quadrature a 0.05 mag systematic uncertainty to the observed flux in each band, to account for zero-point calibration errors in the HSC DR1 photometry, which have been shown to be on this order of magnitude or smaller (see Aihara et al. 2018).

An important data-reduction step on which our measurements rely is the sky subtraction. We checked the robustness of the sky subtraction by repeating the analysis on a subset of galaxies, using the more recent data from the HSC data release 2<sup>1</sup> (DR2; Aihara et al. 2019). The HSC DR2 used a substantially different sky subtraction method, compared to the DR1 (see section 4.1 in Aihara et al. 2019). The corresponding difference in flux leads to an average difference of 0.03 dex on the stellar masses, with a 0.07 dex scatter. While the scatter is well within the observational uncertainty on the stellar mass,

<sup>1</sup>The HSC DR2 was released when the bulk of our analysis was complete.



**Figure 2.** Colour-composite HSC images of a set of SDSS main sample ETGs that passed our selection in  $EW([O II])$  and our visual inspection.



**Figure 3.** Colour-composite HSC images of a set of SDSS main sample galaxies that passed our selection in  $EW([O II])$ , but were rejected in our visual inspection step, due to the presence of discs and/or spiral arms.

this bias is a potential systematic effect that is difficult to correct for and should in principle be taken into account in our global error budget. However, it does not affect the conclusions of our study: our main goal is to measure the slope and evolution of the  $M_*-\sigma_e$  relation, which are robust to overall shifts in the stellar mass measurements of the sample.

### 2.1.3 Stellar mass measurements

To infer stellar masses, we fitted the observed  $g$ ,  $r$ ,  $i$ ,  $z$ , and  $y$  fluxes with composite stellar population models. These were obtained using the BC03 SPS 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 considered star formation histories with an exponentially declining star formation rate and we applied a prior on metallicity based on the mass–metallicity relation measured by Gallazzi et al. (2005). We sampled the posterior probability

distribution of stellar mass, age (time since the initial burst of star formation), star formation rate decline time-scale, metallicity, and dust attenuation with a Markov Chain Monte Carlo (MCMC), following the method introduced by Auger et al. (2009). We then considered the posterior probability distribution in log-stellar mass, marginalized over the other parameters, and approximated 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, Wang & Bahcall (2019) for more details. In Appendix A (online version only), we compare our estimates of stellar mass with those of Mendel et al. (2014; hereafter M14) for the SDSS galaxies of our sample.

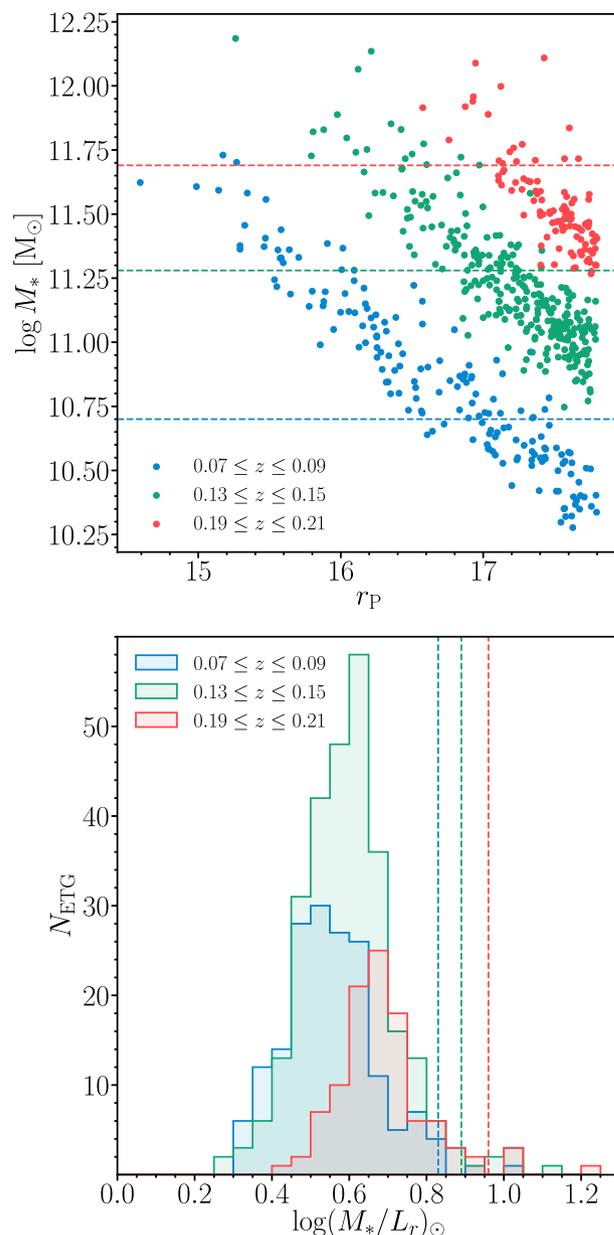
### 2.1.4 A complete sample

In order to accurately infer the  $M_*-\sigma_e$  relation, it is necessary that the selection criteria used to define our sample do not introduce spurious correlations between these two variables. A sufficient condition to achieve this is working with a sample that, at any given redshift, is highly complete in stellar mass, or is randomly drawn from a complete sample. For the SDSS sample, we achieved this condition by first estimating, at each redshift  $z$ , the minimum stellar mass above which our sample is 99 per cent complete,  $M_{*,\min}(z)$ , and then removing from the sample all galaxies with stellar mass below this value. To estimate  $M_{*,\min}(z)$  of the SDSS sample, we proceeded as follows. The SDSS main sample, from which our galaxies are drawn, is complete down to an  $r$ -band Petrosian magnitude  $r_p$  of 17.77 (Strauss et al. 2002). At any redshift, this value of  $r_p$  corresponds to a range of values of the stellar mass, with a spread that is due to scatter in the stellar mass-to-light ratio and to a mismatch between the definition of Petrosian and Sérsic magnitudes. We can nevertheless define the ratio between the observed stellar mass and the observed-frame SDSS  $r$ -band Petrosian luminosity  $L_r$  and consider its distribution  $P(M_*/L_r)$ . We then made narrow redshift bins and, approximating  $P(M_*/L_r)$  as a Gaussian, used the mean and standard deviation of the sample of  $M_*/L_r$  values in each bin to find the 99th percentile of this distribution,  $M_*/L_r|_{99}$ . Finally, we obtained  $M_{*,\min}(z)$  by multiplying  $M_*/L_r|_{99}$  by the Petrosian luminosity corresponding to the limiting value  $r_p = 17.77$ .

In Fig. 4, we illustrate an application of this procedure on three redshift bins: in the upper panel, we show values of stellar mass as a function of  $r_p$ , while in the lower panel, we show the corresponding distributions in  $M_*/L_r$ . The 99th percentile of the  $P(M_*/L_r)$  distribution and the corresponding value of  $M_{*,\min}(z)$  are shown as dashed lines in the two panels.

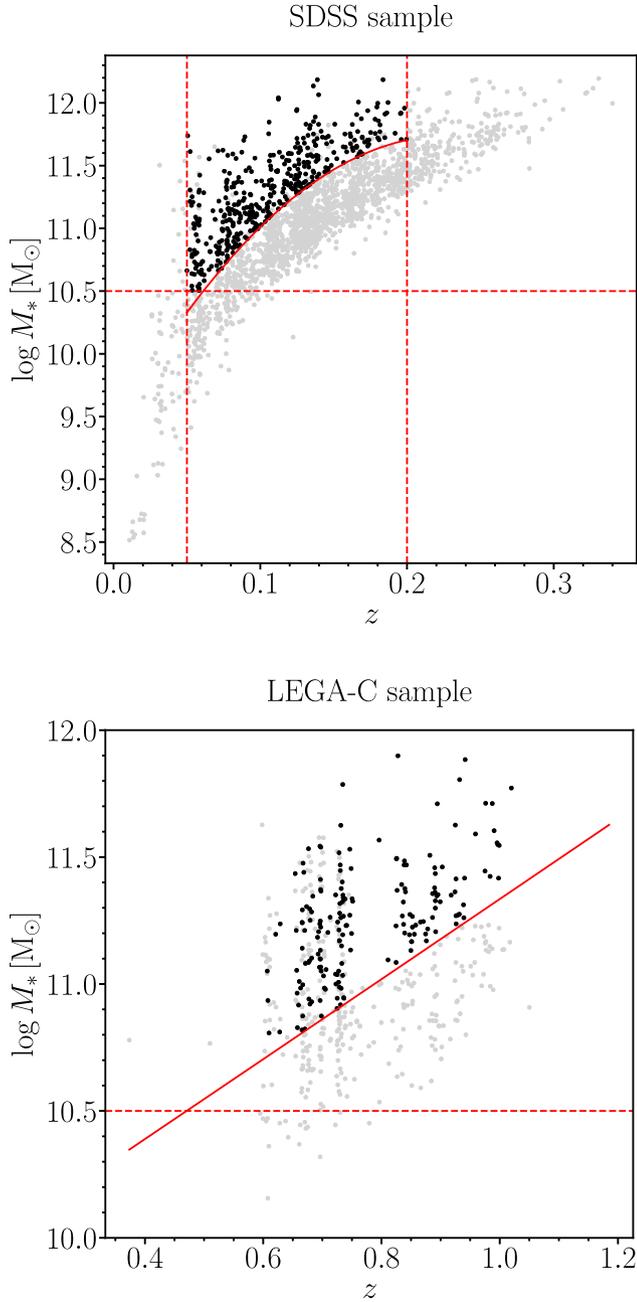
We estimated  $M_{*,\min}(z)$  in a series of bins in the redshift range  $0.05 < z < 0.20$ . Outside this interval, the number of galaxies per redshift bin becomes small, and it is more difficult to obtain an accurate estimate of  $M_{*,\min}$ . We therefore, only included SDSS galaxies in this redshift range, with a stellar mass higher than the value of  $M_{*,\min}$  at the corresponding redshift. We approximated the function  $M_{*,\min}(z)$  as a quadratic polynomial for this purpose. In the upper panel of Fig. 5, we show the stellar mass as a function of redshift for the initial sample of SDSS ETGs composed by 2127 sources (grey dots), as well as for the final sample (black dots), which consists of 413 objects, obtained after applying the cut in stellar mass. The solid curve shows  $M_{*,\min}(z)$ : our SDSS sample is more than 99 per cent complete above this stellar mass.

LEGA-C primary targets have been selected on the basis of their photometric redshift and  $K_s$ -band magnitude, as obtained from the UltraVISTA survey photometric data (Muzzin et al. 2013). Specifically, according to Straatman et al. (2018), primary targets have been selected in the photometric redshift range  $0.6 < z_{\text{photo}} < 1$  and applying a redshift-dependent  $K_s$ -magnitude selection  $K_s < K_{s,\max}(z_{\text{photo}})$ , with  $K_{s,\max}(z_{\text{photo}}) = 20.7 - 7.5 \log[(1 + z_{\text{photo}})/1.8]$ . For the sake of robustness, in order to avoid contamination from objects with incorrect photo- $z$ , we apply a more conservative selection adopting a constant  $K_s$  limit,  $K_s < 20.36 = K_{s,\max}(1)$ . We then obtained  $M_{*,\min}(z)$  for the LEGA-C sample using the method described above for the SDSS sample, simply replacing  $r_p$  with the UltraVISTA  $K_s$ -band magnitude. The resulting distribution in redshift and stellar mass is shown in the lower panel of Fig. 5. From a sample of 492 galaxies selected in morphology,  $K_s$ -band magnitude, and EW([O II]) (grey dots), after selecting only galaxies with stellar mass above  $M_{*,\min}(z)$ , our LEGA-C sample of ETGs reduces to 178 objects (black dots).



**Figure 4.** Stellar mass as a function of SDSS  $r$ -band Petrosian magnitude, for SDSS main sample ETGs in three narrow redshift bins (upper panel). Horizontal dashed lines mark, in each redshift bin, the stellar mass above which an ETG drawn from the SDSS main sample has more than 99 per cent probability of entering our sample. Distribution in the ratio between stellar mass and observed-frame  $r$ -band Petrosian luminosity of the galaxies in the three redshift bins shown in the upper panel (lower panel). The 99th percentile of each distribution is marked by a vertical dashed line. This value, multiplied by the Petrosian luminosity corresponding to the limiting  $r$ -band magnitude of the SDSS main sample,  $r_p = 17.77$ , gives the 99 per cent completeness limit shown in the upper panel.

The LEGA-C DR2 sample, however, does not include all galaxies brighter than the stated magnitude limit, as the survey was not finished at the time of that data release. Instead, the targets included in DR2 were selected according to a  $K_s$ -dependent probability,  $P(K_s)$ . The resulting sample is therefore incomplete, but the incompleteness rate  $P(K_s)$  is a known quantity, provided in the LEGA-C DR2. In



**Figure 5.** Stellar mass as a function of redshift for SDSS (upper panel) and LEGA-C (lower panel) galaxies. The solid curves represent the empirical 99 per cent mass-completeness limits. The horizontal dashed lines represent the absolute lower stellar mass limit  $10^{10.5} M_{\odot}$ , while the vertical dashed lines are the lowest ( $z = 0.05$ ) and highest ( $z = 0.20$ ) redshift limits imposed in the SDSS galaxy selection. In the upper panel, grey dots represent all the 2127 SDSS main sample galaxies morphologically selected and with  $\text{EW}([\text{O II}]) \geq -5 \text{ \AA}$ , while black dots are the 413 objects above the mass-completeness limit, with  $\log(M_*/M_{\odot}) > 10.5$  in the redshift range of  $0.05 < z < 0.2$ . In the lower panel, grey dots are the 492 LEGA-C galaxies selected in morphology,  $\text{EW}([\text{O II}])$  and  $K_s$ -band magnitude, while the black dots represent the final LEGA-C sample of 178 ETGs above the mass-completeness limit.

order to obtain an unbiased inference of the  $M_*$ – $\sigma_e$  relation, it is then sufficient to re-weight each measurement by the inverse of  $P(K_s)$ . In Table 1, we summarize the selection steps used to obtain the final SDSS and LEGA-C samples.

**Table 1.** Summary table of the selection steps adopted to build the final SDSS and LEGA-C samples.

Selection step	$N_{\text{ETG}}$
SDSS sample	
SDSS main sample galaxies selected on morphology and $\text{EW}([\text{O II}])$	2127
SDSS galaxies at $0.05 \leq z \leq 0.2$ with $M_* > 10^{10.5} M_{\odot}$ and $M_* > M_{*,\text{min}}(z)$	413
LEGA-C sample	
LEGA-C galaxies selected on morphology, $\text{EW}([\text{O II}])$ and $K_s$ -band magnitude	492
LEGA-C galaxies with $M_* > 10^{10.5} M_{\odot}$ and $M_* > M_{*,\text{min}}(z)$	178

### 2.1.5 Velocity dispersion measurements

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 arcsec radius fibre of the SDSS spectrograph, which we label  $\sigma_{\text{ap}}$ . We convert this measurement into an estimate of the central velocity dispersion integrated within an aperture equal to the half-light radius,  $\sigma_e$ , 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).

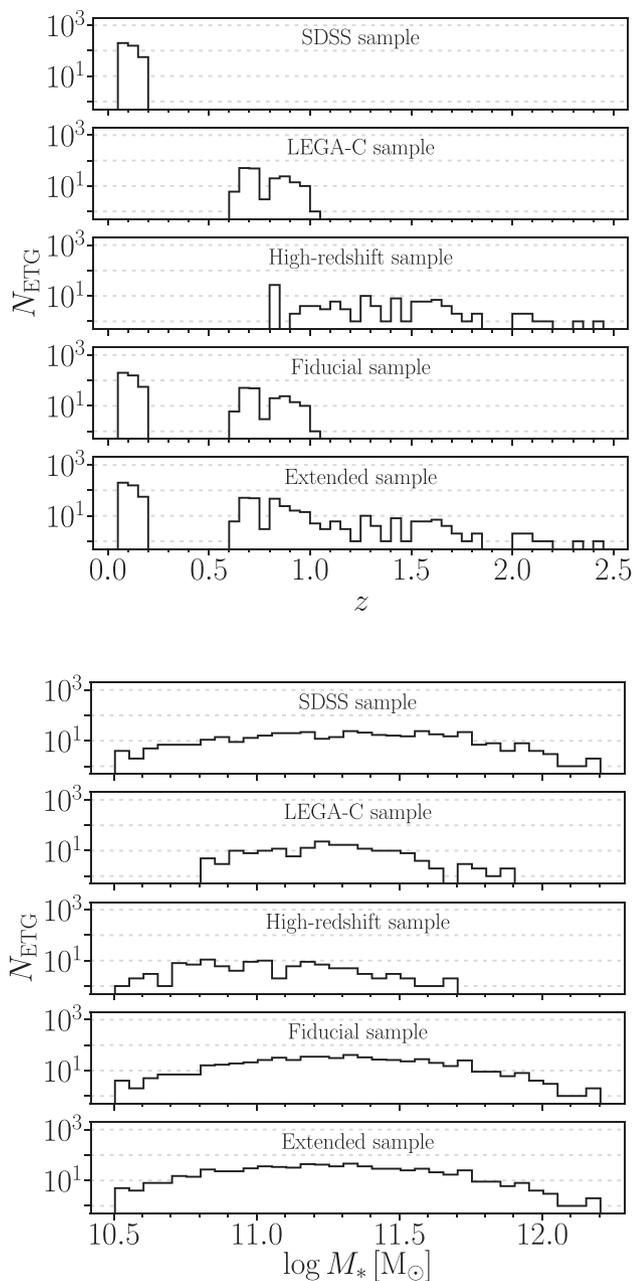
Velocity dispersion measurements provided in the LEGA-C DR2 are converted to values of the central velocity dispersion  $\sigma_e$  applying the aperture correction

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

which is a good approximation for galaxies in the redshift range of the LEGA-C sample (van de Sande et al. 2013; Belli et al. 2014a). The distributions in redshift and in stellar mass of the SDSS and LEGA-C subsamples and of the fiducial sample are shown in Fig. 6 (see also Table 2).

## 2.2 The high-redshift and extended samples

Our high-redshift sample of ETGs is a sample of 110 galaxies with  $\log(M_*/M_{\odot}) > 10.5$  in the redshift range  $0.8 \lesssim z \lesssim 2.5$ , built as follows. We obtain measurements of the stellar mass and stellar velocity dispersion of ETGs 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. (2014a; hereafter B14), including only those galaxies for which  $\text{EW}([\text{O II}]) \geq -5 \text{ \AA}$  (as done for the fiducial sample; Section 2.1), 56 galaxies from van de Sande et al. (2013; 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 summarized in Table 2. Among the original sample of 73 galaxies of vdS13, only 5 galaxies are presented for the first time, while the remaining 68 sources are collected from different studies. We removed 17 of these 73 ETGs because they are already included as part of either B14's or B17's samples. All the galaxies in the high-redshift samples are classified as ETGs, based on their  $UVJ$  colours, morphology and/or spectra. Of course, given the more heterogeneous selection, our extended sample is not as self-consistent as our fiducial sample, and, due to the known correlations between  $\sigma_e$  and some



**Figure 6.** Distributions of the subsamples and samples of ETGs in redshift (upper panel) and stellar mass (lower 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.

structural or spectral properties of ETGs (Zahid & Geller 2017), we cannot exclude that selection biases have non-negligible effects when the high-redshift sample is considered. However, for the vast majority of these galaxies, stellar masses are measured by fitting SPS models to broad-band 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

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

Sample	$z$	$\log(M_*/M_\odot)$	$N_{\text{ETG}}$
SDSS	(0.05; 0.20)	(10.50; 12.19)	413
LEGA-C	(0.60; 1.02)	(10.80; 11.90)	178
vdS13	(0.81; 2.19)	(10.53; 11.69)	56
B14	(1.02; 1.60)	(10.59; 11.35)	26
G15	(1.26; 1.41)	(11.04; 11.49)	4
B17	(1.52; 2.44)	(10.60; 11.68)	24

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 701 ETGs with  $M_* \gtrsim 10^{10.5} M_\odot$  in the redshift interval  $0 \lesssim z \lesssim 2.5$ . The distributions in redshift and in stellar mass of the high-redshift and extended samples are shown in Fig. 6.

### 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 when 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  $\Theta = \{\log M_*, \log \sigma_e, z\}$ . 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  $\Theta$  are drawn from a probability distribution, described in turn by a set of hyper-parameters  $\Phi$ :

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

Our goal is to infer plausible values of the hyper-parameters, which summarize the distribution of our galaxies in the  $(\log M_*, \log \sigma_e, z)$  space, given our data. We will describe in detail the functional form of the distribution  $P(\Theta|\Phi)$  in Section 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,

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

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

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

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,  $\Theta$ , and not on the hyper-parameters  $\Phi$ . In order to compute the  $P(\mathbf{d}_i|\Phi)$  terms in equation (8), then, we need to marginalize over all possible values of the individual object parameters  $\Theta_i$ :

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

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  $\Theta_i$  are drawn from a distribution that is, in turn, described by a set of hyper-parameters.

As explained in Section 2.1, the LEGA-C sample is not representative of a complete sample, but each galaxy was included with a  $K_s$  magnitude-dependent probability  $P(K_s)$ , so that brighter galaxies are overrepresented (see fig. 2 of Straatman et al. 2018). To correct for this, we re-weight the contribution of each LEGA-C measurement to the likelihood by a factor proportional to  $1/P(K_s)$ : we transform equation (8) to

$$P(\mathbf{d}|\Phi) = \prod_i P(\mathbf{d}_i|\Phi)^{w_i}, \quad (10)$$

where  $w_i$  is given by

$$w_i = \frac{1/P(K_{s,i})}{(1/P(K_s))} \quad (11)$$

for LEGA-C galaxies and  $w_i = 1$  otherwise. The normalization of the weights given in the equation above ensures that the effective number of LEGA-C data points equals the number of LEGA-C galaxies.

### 3.2 The model

The purpose of our model is to summarize the distribution in stellar mass and velocity dispersion of our samples of ETGs with a handful of parameters,  $\Phi$ , that can provide an intuitive description of the  $M_*$ - $\sigma_e$  relation. In the absence of a well-established theoretically motivated model, we opt for an empirical one that we describe in this section.

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

$$P(\Theta|\Phi) = P(\log M_*, z|\Phi)P(\log \sigma_e|\log M_*, z, \Phi). \quad (12)$$

Here,  $P(\log M_*, z|\Phi)$  describes the prior probability distribution for a galaxy in our sample to have logarithm of the true stellar mass  $\log M_*$  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,  $M_*^{\text{obs}} > M_{*,\text{min}}$ . We then expect the probability distribution in the true stellar mass to go to zero for low values of  $M_*$ . We also expect  $P(\log M_*, z|\Phi)$  to vanish for very large values of  $M_*$ , as there are few known galaxies with  $M_* > 10^{12}$ . For simplicity, we assume that  $P(\log M_*, z|\Phi)$  separates as follows:

$$P(\log M_*, z|\Phi) = P(\log M_*|\Phi)P(z|\Phi), \quad (13)$$

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

$$P(\log M_*|\Phi) \propto \frac{1}{\sqrt{2\pi\sigma_*^2}} \exp\left\{-\frac{(\log M_* - \mu_*)^2}{2\sigma_*^2}\right\} \mathcal{E}(\log M_*|\Phi), \quad (14)$$

with

$$\mathcal{E}(\log M_*|\Phi) = 1 + \text{erf}\left(\alpha_* \frac{\log M_* - \mu_*}{\sqrt{2\sigma_*^2}}\right), \quad (15)$$

where the three hyper-parameters  $\mu_*$ ,  $\sigma_*$ , and  $\alpha_*$  are modelled as

$$\mu_* = \mu_{*,0} + \mu_{*,s} \log\left(\frac{1+z}{1+z^{\text{piv}}}\right), \quad (16)$$

$$\sigma_* = \sigma_{*,0} + \sigma_{*,s} \log\left(\frac{1+z}{1+z^{\text{piv}}}\right) \quad (17)$$

and

$$\alpha_* = \text{const.} \quad (19)$$

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  $P(\log M_*|\Phi)$  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 (13) describes the redshift distribution of the galaxies in our sample. As we show below, 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 (12) is the core of our model. With it, we wish to capture the following features of the  $M_*$ - $\sigma_e$  relation: its normalization (i.e. the amplitude of the stellar velocity dispersion at a given value of the stellar mass) and its redshift evolution, the correlation between velocity dispersion and stellar mass, and the amplitude of the intrinsic scatter in  $\sigma_e$  at fixed  $M_*$  and redshift. With these requirements in mind, 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:

$$P(\log \sigma_e|\log M_*, z, \Phi) = \frac{1}{\sqrt{2\pi\sigma_\sigma^2}} \exp\left\{-\frac{(\log \sigma_e - \mu_\sigma)^2}{2\sigma_\sigma^2}\right\}. \quad (20)$$

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

$$\mu_\sigma = \mu_0^{\text{SDSS}} + \beta \log\left(\frac{M_*}{M_*^{\text{piv}}}\right) + \zeta \log\left(\frac{1+z}{1+z^{\text{piv}}}\right). \quad (21)$$

In general, the slope  $\beta$  is allowed to depend on  $z$  as

$$\beta = \beta_0^{\text{SDSS}} + \eta \log\left(\frac{1+z}{1+z^{\text{piv}}}\right). \quad (22)$$

We perform our analysis considering two different cases: the first is a *constant-slope* case (model  $\mathcal{M}_{\text{const}}$ ), i.e. equation (22) with  $\eta = 0$ ; in the second, which we refer to as the *evolving-slope* case (model  $\mathcal{M}_{\text{evo}}$ ),  $\eta$  is a free hyper-parameter. For the standard deviation  $\sigma_\sigma$  in equation (20), namely the intrinsic scatter of our relation, we adopt the form

$$\sigma_\sigma = \psi_0^{\text{SDSS}} + \xi \log\left(\frac{1+z}{1+z^{\text{piv}}}\right). \quad (23)$$

In equations (21-23)  $M_*^{\text{piv}} = 10^{11.321}$  and  $z^{\text{piv}} = 0.10436$ , i.e. the median values of stellar mass and redshift of the SDSS ETGs,

respectively, while the quantities  $\mu_0^{\text{SDSS}}$ ,  $\beta_0^{\text{SDSS}}$ , and  $\psi_0^{\text{SDSS}}$  are the median values of the hyper-parameters  $\mu_0$ ,  $\beta_0$ , and  $\psi_0$  obtained when fitting equation (20) to the ETGs of the SDSS subsample with

$$\mu_\sigma = \mu_0 + \beta_0 \log \left( \frac{M_*}{M_*^{\text{piv}}} \right) \quad \text{and} \quad \sigma_\sigma = \psi_0, \quad (24)$$

i.e. neglecting any dependence on  $z$ . In order to prevent the redshift dependence of the relation from being influenced by any redshift dependence within the SDSS sample, which constitutes  $\approx 60$  per cent of the extended sample, we assume the model in equation (24) as the *zero-point* at  $z^{\text{piv}}$  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 (24) applied to the SDSS subsample as model  $\mathcal{M}^{\text{SDSS}}$ .

Allowing for intrinsic scatter is an important feature of our model. Neglecting it leads typically to underestimating the slope of the  $M_* - \sigma_e$  relation (see e.g. Auger et al. 2010). Our choice for the functional form of the distribution in velocity dispersion introduced above is somewhat arbitrary. Although there could exist alternative distributions that fit the data equally well as our model or better, however, exploring such distributions is beyond the scope of this work.

### 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  $\Phi$  given the data  $\mathbf{d}$ ,  $P(\Phi|\mathbf{d})$ . 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

$$\begin{aligned} & P(\log M_*^{\text{obs}}, \log \sigma_e^{\text{obs}}, z^{\text{obs}} | \Phi) \\ &= \iiint d \log M_* d \log \sigma_e dz \\ & \quad \times P(\log M_*^{\text{obs}}, \log \sigma_e^{\text{obs}}, z^{\text{obs}} | \log M_*, \log \sigma_e, z) \\ & \quad \times P(\log M_*, \log \sigma_e, z | \Phi) \\ &= \iiint d \log M_* d \log \sigma_e dz \\ & \quad \times P(\log M_*^{\text{obs}} | \log M_*) P(\log \sigma_e^{\text{obs}} | \log \sigma_e) \delta(z^{\text{obs}} - z) \\ & \quad \times P(\log M_* | \Phi) P(z | \Phi) P(\log \sigma_e | \log M_*, z, \Phi). \end{aligned} \quad (25)$$

In the last line, we have used equations (12) and (13), and we have approximated the likelihood of observing redshift  $z^{\text{obs}}$  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  $P(z|\Phi)$  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  $P(\log \sigma_e^{\text{obs}} | \log \sigma_e)$ , the integral over  $d \log \sigma_e$  can be performed analytically, as we show in Appendix B (online version only). We also

**Table 3.** Jeffreys' scale (Jeffreys 1961), giving the strength of evidence in the comparison of two models having Bayes factor  $\mathcal{B}$  (equation 30).

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

assume a Gaussian likelihood for the measurements of  $\log M_*^{\text{obs}}$ ,

$$\begin{aligned} & P(\log M_*^{\text{obs}} | \log M_*) \\ &= \frac{\mathcal{A}(\log M_*)}{\sqrt{2\pi\sigma_{M_*}^2}} \exp \left\{ -\frac{(\log M_* - \log M_*^{\text{obs}})^2}{2\sigma_{M_*}^2} \right\}, \end{aligned} \quad (26)$$

with one caveat: we are only selecting galaxies with  $\log M_*^{\text{obs}} > \log M_{*,\text{min}}$ , where  $\log M_{*,\text{min}}$  is derived from the mass-completeness limits at a given redshift for SDSS and LEGA-C galaxies (see Section 2.1.4) and it is assumed to be constant and equal to 10.5 for all the ETGs of the high-redshift sample. The likelihood must be normalized accordingly:

$$\int_{\log M_{*,\text{min}}}^{\infty} d \log M_* \frac{\mathcal{A}(\log M_*)}{\sqrt{2\pi\sigma_{M_*}^2}} \exp \left\{ -\frac{(\log M_* - \log M_*^{\text{obs}})^2}{2\sigma_{M_*}^2} \right\} = 1. \quad (27)$$

In other words, the probability of measuring any value of the stellar mass larger than  $M_{*,\text{min}}$ , 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 B, online version only). 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  $\mathcal{M}$ :

$$\mathcal{Z} = P(\mathbf{d}|\mathcal{M}) = \int d\Theta P(\mathbf{d}|\Theta, \mathcal{M}) P(\Theta|\mathcal{M}). \quad (28)$$

We remark that, in our approach, the parameters  $\Theta$  are described by a set of global hyper-parameters  $\Phi$ . 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

$$\frac{P(\mathcal{M}_1|\mathbf{d})}{P(\mathcal{M}_2|\mathbf{d})} = \mathcal{B} \frac{P(\mathcal{M}_1)}{P(\mathcal{M}_2)}, \quad (29)$$

where

$$\mathcal{B} \equiv \frac{P(\mathbf{d}|\mathcal{M}_1)}{P(\mathbf{d}|\mathcal{M}_2)} = \frac{\mathcal{Z}_1}{\mathcal{Z}_2} \quad (30)$$

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

**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: uniform priors used in the models ('low' and 'up' indicate, respectively, the lower and upper bounds). For those hyper-parameters showing two ranges for prior assumptions, the first refers to the fiducial sample and the second to the extended sample.  $M_*^{\text{piv}}$  and  $z^{\text{piv}}$  are the median values of stellar mass and redshift of the SDSS ETGs (Section 3.2).

Model	Hyper-parameter	Description	Prior (low; up)
$\mathcal{M}^{\text{SDSS}}$	$\mu_0$	Median value of $\log \sigma_e$ at $M_*^{\text{piv}}$	(1; 3)
	$\beta_0$	Index of the $M_*-\sigma_e$ relation: $\sigma_e \propto M_*^{\beta_0}$	(0; 1)
	$\psi_0$	Intrinsic scatter in $\log \sigma_e$	(0; 1)
	$\mu_{*,0}$	Normalization of the mean of Gaussian prior of $\log M_*$	(10; 13)
	$\mu_{*,s}$	Slope of the mean of Gaussian prior of $\log M_*$	(15; 30)
	$\sigma_{*,0}$	Normalization of the standard deviation in the Gaussian prior of $\log M_*$	(0; 2)
	$\sigma_{*,s}$	Slope of the standard deviation in the Gaussian prior of $\log M_*$	(-10; 10)
	$\alpha_*$	Skewness parameter in the Gaussian prior of $\log M_*$	(-5; 15)
$\mathcal{M}_{\text{evo}}$	$\mu_0^{\text{SDSS}}$	Median value of $\log \sigma_e$ at $M_* = M_*^{\text{piv}}$ and $z = z^{\text{piv}}$	$\simeq 2.287$
	$\beta_0^{\text{SDSS}}$	Index of the $M_*-\sigma_e$ relation at $z = z^{\text{piv}}$ : $\sigma_e \propto M_*^{\beta_0^{\text{SDSS}}}$	$\simeq 0.176$
	$\eta$	Index of the $\beta-(1+z)$ relation: $\beta \propto (1+z)^\eta$	(-2; 2)
	$\zeta$	Index of the $\sigma_e-(1+z)$ relation: $\sigma_e \propto (1+z)^\zeta$	(-2; 2)
	$\psi_0^{\text{SDSS}}$	Median value of $\psi_0$ of the intrinsic scatter at $z = z^{\text{piv}}$	$\simeq 0.075$
	$\xi$	Index of the $\sigma_\sigma-(1+z)$ relation: $\sigma_\sigma \propto (1+z)^\xi$	(-2; 2)
	$\mu_{*,0}$	Normalization of the mean of Gaussian prior of $\log M_*$	(7; 13)
	$\mu_{*,s}$	Slope of the mean of Gaussian prior of $\log M_*$	(0; 10)    (-5; 5)
	$\sigma_{*,0}$	Normalization of the standard deviation in the Gaussian prior of $\log M_*$	(-5; 5)
	$\sigma_{*,s}$	Slope of the standard deviation in the Gaussian prior of $\log M_*$	(-10; 10)
	$\alpha_*$	Skewness parameter in the Gaussian prior of $\log M_*$	(-15; 15)
$\mathcal{M}_{\text{const}}$		Same as $\mathcal{M}_{\text{evo}}$ , but with $\eta = 0$	
	$\alpha_*$	Skewness parameter in the Gaussian prior of $\log M_*$	(-10; 10)    (-15; 15)
$\mathcal{M}_{\text{evo,NES}}$		Same as $\mathcal{M}_{\text{evo}}$ , but with $\xi = 0$	
$\mathcal{M}_{\text{const,NES}}$		Same as $\mathcal{M}_{\text{evo}}$ , but with $\eta = \xi = 0$	
	$\sigma_{*,s}$	Slope of the standard deviation in the Gaussian prior of $\log M_*$	(-10; 10)    (-5; 5)
	$\alpha_*$	Skewness parameter in the Gaussian prior of $\log M_*$	(-10; 10)    (-15; 15)

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, Hobson & Bridges 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

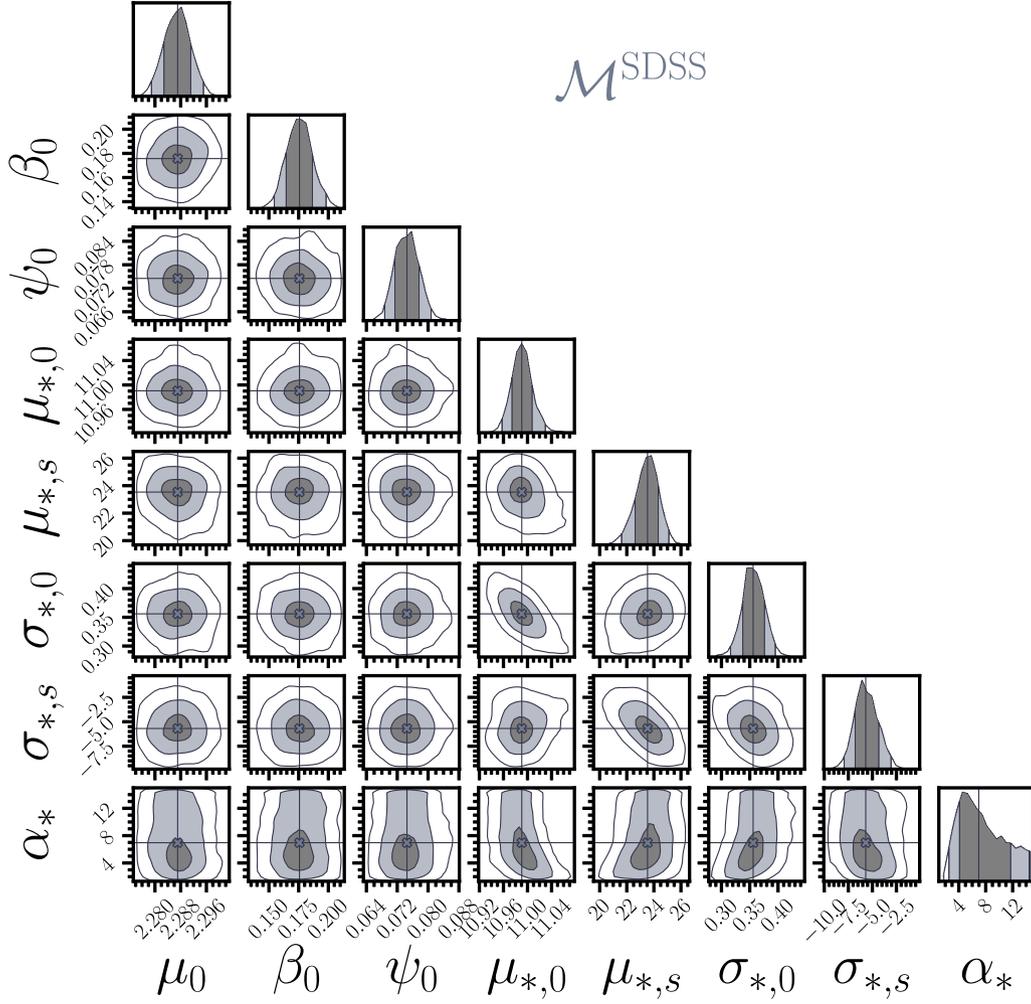
In this section we present the results obtained applying the Bayesian method described in Section 3 to our fiducial and extended samples of ETGs (see Section 2).

In Section 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  $\beta$ ), and model  $\mathcal{M}_{\text{evo}}$  (representing the evolution of the  $M_*-\sigma_e$  relation with redshift-dependent slope  $\beta$ ). In models  $\mathcal{M}_{\text{const}}$  and  $\mathcal{M}_{\text{evo}}$ , 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  $\mathcal{M}_{\text{const,NES}}$  and  $\mathcal{M}_{\text{evo,NES}}$ , where NES stands for *non-evolving scatter*. In summary, we take into account five models: model  $\mathcal{M}^{\text{SDSS}}$ , represented by equation (24), models  $\mathcal{M}_{\text{const}}$  and  $\mathcal{M}_{\text{evo}}$ , described by equation (21)

(the former obtained by assuming  $\eta = 0$  in equation 22), and the models  $\mathcal{M}_{\text{const,NES}}$  and  $\mathcal{M}_{\text{evo,NES}}$ , which are the same as  $\mathcal{M}_{\text{const}}$  and  $\mathcal{M}_{\text{evo}}$ , respectively, but with  $\xi = 0$  in equation (23). 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 data set similar to the our SDSS data set (see Appendix C, online version only). Each MCMC run (see Section 3.3) uses 50 random walkers running for 1000 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 Fig. 7. The SDSS galaxies are described by  $\sigma_e \propto M_*^{\beta_0}$  with  $\beta_0 \simeq 0.176$ . The normalization  $\mu_0 \simeq 2.287$  is such that galaxies with  $M_* = 10^{11} M_\odot$  have  $\sigma_e \simeq 170 \text{ km s}^{-1}$  and the intrinsic scatter is  $\simeq 0.075$  dex in  $\sigma_e$  at fixed  $M_*$ . The posterior distributions of the hyper-parameters  $\mu_0$ ,  $\beta_0$ , and  $\psi_0$  are relatively narrow (Fig. 7), with  $1\sigma$  scatter of at most few per cent (Table 5), so our SDSS sample of ETG is sufficiently numerous for our purposes, even if it contains only a small fraction of the massive ETGs of the entire SDSS sample. Our results on the present-day  $M_*-\sigma_e$  relation are broadly consistent with previous analyses (see Section 5.2 for details).

The median values of the hyper-parameters of all models, with the corresponding  $1\sigma$  uncertainties, are listed in Table 5. In order



**Figure 7.** Posterior probability distributions of the hyper-parameters for model  $\mathcal{M}^{\text{SDSS}}$  (see Tables 4 and 5). In the 1D distributions (upper panel of each column) the vertical solid lines and colours delimit the 68, 95, and 99.7th quantile based posterior credible interval. In the 2D distributions (all the other panels) the contours enclose the 68, 95, and 99.7 per cent posterior credible regions. The dashed lines indicate the median values of the hyper-parameters.

**Table 5.** Inferred median and 68 per cent posterior credible intervals of the hyper-parameters of the models.

Model	$\mu_0$	$\beta_0$	$\psi_0$	$\mu_{*,0}$	$\mu_{*,s}$	$\sigma_{*,0}$	$\sigma_{*,s}$	$\alpha_*$
$\mathcal{M}^{\text{SDSS}}$	$2.287^{+0.004}_{-0.004}$	$0.176^{+0.011}_{-0.011}$	$0.075^{+0.003}_{-0.003}$	$10.990^{+0.017}_{-0.016}$	$23.533^{+0.792}_{-0.910}$	$0.356^{+0.020}_{-0.018}$	$-5.690^{+1.358}_{-1.095}$	$6.969^{+4.725}_{-2.868}$
Model	$\eta$	$\zeta$	$\xi$	$\mu_{*,0}$	$\mu_{*,s}$	$\sigma_{*,0}$	$\sigma_{*,s}$	$\alpha_*$
$\mathcal{M}_{\text{const}}^{\text{fid}}$	–	$0.390^{+0.031}_{-0.031}$	$0.033^{+0.026}_{-0.023}$	$9.827^{+0.259}_{-0.213}$	$5.616^{+0.955}_{-1.054}$	$0.457^{+0.199}_{-0.193}$	$-0.841^{+0.937}_{-0.908}$	$3.993^{+4.089}_{-2.579}$
$\mathcal{M}_{\text{evo}}^{\text{fid}}$	$0.248^{+0.150}_{-0.156}$	$0.415^{+0.034}_{-0.035}$	$0.028^{+0.027}_{-0.024}$	$9.855^{+0.226}_{-0.220}$	$5.455^{+0.995}_{-0.913}$	$0.428^{+0.211}_{-0.191}$	$-0.684^{+0.920}_{-0.950}$	$5.727^{+5.763}_{-4.083}$
$\mathcal{M}_{\text{const}}^{\text{ext}}$	–	$0.474^{+0.024}_{-0.023}$	$0.025^{+0.019}_{-0.019}$	$11.147^{+0.168}_{-0.158}$	$-0.172^{+0.322}_{-0.348}$	$0.249^{+0.088}_{-0.075}$	$0.210^{+0.301}_{-0.271}$	$0.062^{+1.421}_{-1.126}$
$\mathcal{M}_{\text{evo}}^{\text{ext}}$	$0.179^{+0.101}_{-0.106}$	$0.505^{+0.028}_{-0.029}$	$0.019^{+0.021}_{-0.022}$	$11.156^{+0.140}_{-0.121}$	$-0.197^{+0.300}_{-0.319}$	$0.240^{+0.062}_{-0.079}$	$0.209^{+0.286}_{-0.254}$	$0.061^{+1.031}_{-0.952}$
$\mathcal{M}_{\text{const,NES}}^{\text{fid}}$	–	$0.398^{+0.028}_{-0.031}$	–	$9.853^{+0.282}_{-0.220}$	$5.538^{+0.951}_{-1.029}$	$0.415^{+0.202}_{-0.181}$	$-0.657^{+0.846}_{-0.868}$	$3.267^{+3.894}_{-2.494}$
$\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$	$0.264^{+0.146}_{-0.142}$	$0.417^{+0.035}_{-0.034}$	–	$9.867^{+0.211}_{-0.204}$	$5.408^{+0.934}_{-0.870}$	$0.408^{+0.211}_{-0.179}$	$-0.625^{+0.900}_{-0.921}$	$5.195^{+5.947}_{-3.571}$
$\mathcal{M}_{\text{const,NES}}^{\text{ext}}$	–	$0.478^{+0.021}_{-0.021}$	–	$11.142^{+0.174}_{-0.171}$	$-0.148^{+0.272}_{-0.302}$	$0.246^{+0.084}_{-0.076}$	$0.203^{+0.325}_{-0.274}$	$0.093^{+1.289}_{-1.036}$
$\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$	$0.180^{+0.103}_{-0.095}$	$0.506^{+0.029}_{-0.027}$	–	$11.192^{+0.138}_{-0.166}$	$-0.120^{+0.271}_{-0.285}$	$0.238^{+0.081}_{-0.071}$	$0.177^{+0.292}_{-0.246}$	$-0.281^{+1.103}_{-0.704}$

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

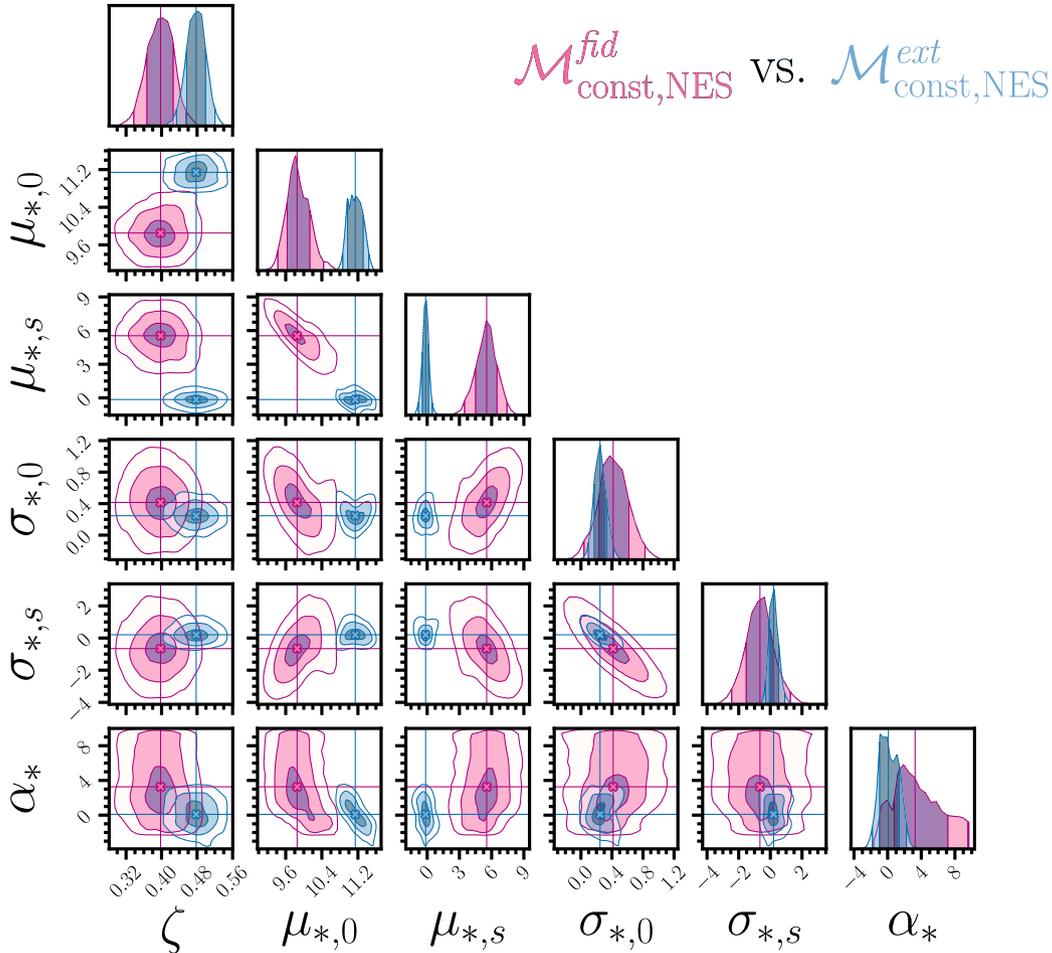
Model	$\ln \mathcal{Z}$	$\ln \mathcal{B}$
$\mathcal{M}_{\text{const}}^{\text{fid}}$	$241.791 \pm 0.204$	$-2.269 \pm 0.391$
$\mathcal{M}_{\text{evo}}^{\text{fid}}$	$240.645 \pm 0.216$	$-3.415 \pm 0.403$
$\mathcal{M}_{\text{const,NES}}^{\text{fid}}$	$244.060 \pm 0.187$	–
$\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$	$243.710 \pm 0.197$	$-0.350 \pm 0.384$
$\mathcal{M}_{\text{const}}^{\text{ext}}$	$284.705 \pm 0.225$	$-2.653 \pm 0.431$
$\mathcal{M}_{\text{evo}}^{\text{ext}}$	$283.908 \pm 0.235$	$-3.450 \pm 0.441$
$\mathcal{M}_{\text{const,NES}}^{\text{ext}}$	$287.358 \pm 0.206$	–
$\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$	$286.584 \pm 0.222$	$-0.774 \pm 0.428$

to compare the models we compute the Bayesian evidence  $\mathcal{Z}$  of each model (Section 3.4), using a configuration of 400 live points in the nested sampling algorithm. The resulting  $\mathcal{Z}$  and the Bayes factors are listed in Table 6. The performances of models  $\mathcal{M}_{\text{const}}$  and  $\mathcal{M}_{\text{evo}}$  are relatively poor when applied to both the fiducial and the extended samples, so in the following we focus on model  $\mathcal{M}_{\text{const,NES}}$  and  $\mathcal{M}_{\text{evo,NES}}$ : in Figs 8 and 9, we show the inferences of these two models applied to both the fiducial and the extended samples.

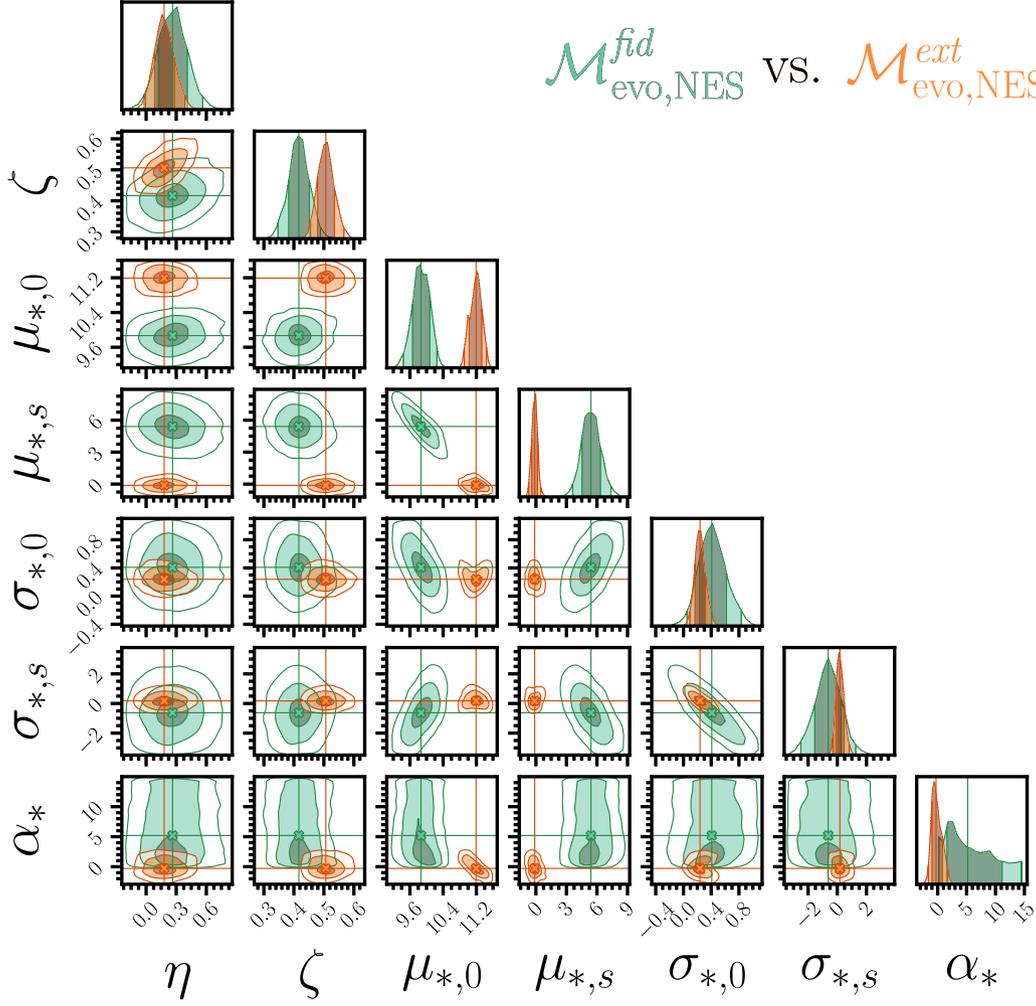
#### 4.1 Fiducial sample ( $0 \lesssim z \lesssim 1$ )

The model with the highest Bayesian evidence, among those applied to the fiducial sample, is  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  (see Table 6). Model  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$ , though with slightly lower evidence, describes the data as well as model  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$ , according to Jeffreys’ scale (Table 3), while models  $\mathcal{M}_{\text{const}}^{\text{fid}}$  and  $\mathcal{M}_{\text{evo}}^{\text{fid}}$  are rejected with strong evidence. Thus, based on our analysis of the fiducial sample, we conclude that at  $z \lesssim 1$  the normalization of the  $M_*$ – $\sigma_e$  relation changes with  $z$ , while the intrinsic scatter is independent of redshift; the slope  $\beta$  is either constant or increasing with redshift (see Fig. 10).

The median  $M_*$ – $\sigma_e$  relations found for models  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  at three representative redshifts are shown in the left-hand panels of Fig. 11. Quantitatively, according to model  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$ , in the redshift interval  $0 \lesssim z \lesssim 1$ , the  $M_*$ – $\sigma_e$  relation is well described by a power law  $\sigma_e \propto M_*^\beta$ , with redshift-independent slope  $\beta \simeq 0.18$  and intrinsic scatter  $\simeq 0.08$  dex in  $\sigma_e$  at given  $M_*$ . At fixed  $M_*$ ,  $\sigma_e \propto (1+z)^\zeta$ , with  $\zeta \simeq 0.40$ , so galaxies of given  $M_*$  tend to have higher  $\sigma_e$  at higher redshift: the median velocity dispersion at fixed  $M_*$  is a factor  $\approx 1.3$  higher at  $z = 1$  than at  $z = 0$ . According to model  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$ ,  $\sigma_e$  varies with  $M_*$  and  $z$  as  $\sigma_e \propto M_*^{\beta(z)}(1+z)^\zeta$ , with  $\zeta \simeq 0.42$  and  $\beta(z) \simeq 0.16 + 0.26 \log(1+z)$ . For instance, at  $z = 1$ , the slope of the  $M_*$ – $\sigma_e$  relation is  $\beta \simeq 0.24$ . The time variation of  $\sigma_e$  at a given  $M_*$  depends on  $M_*$ : at  $M_* = 10^{11} M_\odot$ , it is similar to that inferred according to model  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  (Fig. 10, upper panel).



**Figure 8.** Same as Fig. 7, but for models  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  (pink contours) and  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  (azure contours); see Tables 4 and 5.



**Figure 9.** Same as Fig. 7, but for models  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  (green contours) and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  (orange contours); see Tables 4 and 5.

In summary, the evolution of the  $M_*-\sigma_e$  relation in the redshift range  $0 \lesssim z \lesssim 1$  can be roughly described by

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) \simeq 2.21 + 0.18 \log\left(\frac{M_*}{10^{11}M_\odot}\right) + 0.40 \log(1+z), \quad (31)$$

based on the median values of the hyper-parameters of model  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$ , or

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) \simeq [0.16 + 0.26 \log(1+z)] \log\left(\frac{M_*}{10^{11}M_\odot}\right) + 0.42 \log(1+z) + 2.22, \quad (32)$$

based on the median values of the hyper-parameters of model  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$ , with redshift-independent intrinsic scatter  $\simeq 0.08$  dex in  $\sigma_e$  at a given  $M_*$ .

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

The results obtained for the extended sample are very similar to those obtained for the fiducial sample. The model with the highest Bayesian evidence is model  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  (Table 6), but the performance of model  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  is comparable on the basis of Jeffreys' scale (see Table 3). Models  $\mathcal{M}_{\text{const}}^{\text{ext}}$  and  $\mathcal{M}_{\text{evo}}^{\text{ext}}$  are rejected with strong evidence. Thus, on the basis of our data, over the redshift range  $0 \lesssim z \lesssim 2.5$  the  $M_*-\sigma_e$

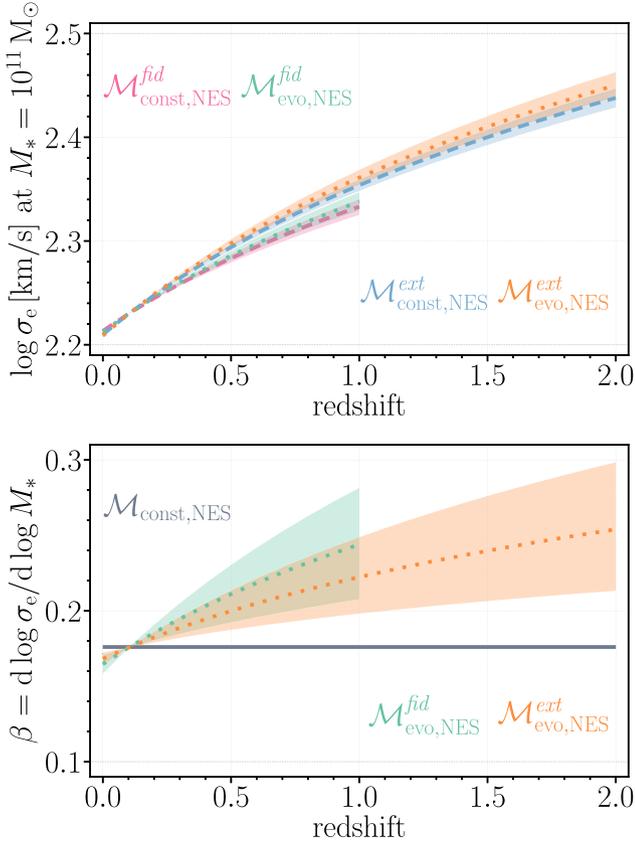
$\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  VS.  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$

relation of ETGs evolves in time by changing its normalization, with redshift-independent intrinsic scatter, and with slope either constant or increasing with redshift (see Fig. 10).

The median  $M_*-\sigma_e$  relations of models  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  at five representative redshifts are shown in the right-hand panels of Fig. 11. Quantitatively, according to model  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$ , in the redshift interval  $0 \lesssim z \lesssim 2.5$ , the  $M_*-\sigma_e$  relation is well described by a power law  $\sigma_e \propto M_*^\beta$ , with slope  $\beta \simeq 0.18$  and intrinsic scatter  $\simeq 0.08$  dex in  $\sigma_e$  at given  $M_*$ : at fixed  $M_*$ ,  $\sigma_e \propto (1+z)^\zeta$ , with  $\zeta \simeq 0.48$ , so galaxies of given  $M_*$  tend to have higher  $\sigma_e$  at higher redshift. For instance, the median velocity dispersion at fixed  $M_*$  is a factor  $\approx 1.7$  higher at  $z = 2$  than at  $z = 0$ . According to model  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$ ,  $\sigma_e$  varies with  $M_*$  and  $z$  as  $\sigma_e \propto M_*^{\beta(z)}(1+z)^\zeta$ , with  $\zeta \simeq 0.51$  and  $\beta(z) \simeq 0.17 + 0.18 \log(1+z)$  (so  $\beta \simeq 0.26$  at  $z = 2$ ; Fig. 10, lower panel). The time variation of  $\sigma_e$  at a given  $M_*$  depends on  $M_*$ , but at  $M_* = 10^{11} M_\odot$  is similar to that obtained with model  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  (Fig. 10, upper panel).

In summary, the evolution of the  $M_*-\sigma_e$  relation in the redshift range  $0 \lesssim z \lesssim 2.5$  can be roughly described by

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) \simeq 2.21 + 0.18 \log\left(\frac{M_*}{10^{11}M_\odot}\right) + 0.48 \log(1+z), \quad (33)$$



**Figure 10.** Median central stellar velocity dispersion  $\sigma_e$  at  $M_* = 10^{11} M_\odot$  (upper panel) and slope  $\beta$  of the  $M_*$ – $\sigma_e$  relation (lower panel) as functions of redshift for models  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  (pink dashed curve),  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  (green dotted curves),  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  (azure dashed curve), and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  (orange dotted curves). In the lower panel, the curves for models  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  and  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  are identical, and are represented by the grey solid line. The curves are obtained by computing, at given  $x$ , the median value of  $y$  (where  $x$  and  $y$  are the quantities in abscissa and ordinate, respectively) among all the values sampled by the posterior distribution obtained with the MCMC; similarly, the shaded bands, which we will refer to as  $1\sigma$  uncertainty bands, are defined by computing the 16 per cent and the 84 per cent of the distribution of  $y$ , at given  $x$ , for the same sampling.

based on the median values of the hyper-parameters of model  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$ , or

$$\log \left( \frac{\sigma_e}{\text{km s}^{-1}} \right) \simeq [0.17 + 0.18 \log(1+z)] \log \left( \frac{M_*}{10^{11} M_\odot} \right) + 0.51 \log(1+z) + 2.21, \quad (34)$$

based on the median values of the hyper-parameters of model  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$ , with redshift-independent intrinsic scatter  $\simeq 0.08$  dex in  $\sigma_e$  at a given  $M_*$ .

### 4.3 Comparing the results for the fiducial and extended samples

Among the hyper-parameters of model  $\mathcal{M}_{\text{const,NES}}$ , only  $\zeta$ , which quantifies the redshift dependence of  $\sigma_e$  at a given  $M_*$ , contains physical information on the  $M_*$ – $\sigma_e$  relation. The other five hyper-parameters describe properties of the galaxy sample. Thus, when comparing the inferences obtained for model  $\mathcal{M}_{\text{const,NES}}$  applied to the fiducial and extended samples (Fig. 8), we must focus on the inference on  $\zeta$ . For model  $\mathcal{M}_{\text{evo,NES}}$ , instead, the physical information is contained in the hyper-parameters  $\eta$  and  $\zeta$ , which must

be considered when comparing the inferences for models  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  (Fig. 9). While the differences in the distributions of  $\eta$  between  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  are well within  $1\sigma$ , the differences in the distributions of  $\zeta$  are between  $1\sigma$  and  $2\sigma$  for both pairs of models ( $\mathcal{M}_{\text{const,NES}}^{\text{fid}} - \mathcal{M}_{\text{const,NES}}^{\text{ext}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}} - \mathcal{M}_{\text{evo,NES}}^{\text{ext}}$ ). Thus, while we find no significant differences in  $\eta$ , the extended-sample data prefer a somewhat higher value of  $\zeta$  than the fiducial-sample data, suggesting that the evolution of  $\sigma_e$  at a given  $M_*$  might be stronger at higher redshift.

However, we recall that the extended sample is not as homogeneous and complete as the fiducial sample, so the aforementioned difference in  $\zeta$  could be produced by some observational bias. For instance, while for the fiducial sample we selected ETGs on the basis of morphology and strength of emission lines, in some of the subsamples of the high-redshift sample (Belli et al. 2014b, 2017), ETGs were selected using also the so-called *UVJ* colour–colour diagram, which is a useful tool to separate passive and star-forming galaxies (e.g. Moresco et al. 2013). To quantify the effect of these different selection criteria, we performed the following test. Using *UVJ* colour measures from the UltraVISTA survey (Muzzin et al. 2013), we placed the LEGA-C galaxies of our fiducial sample in the *UVJ* diagram, finding that 90 per cent of them lie in the locus of passive galaxies (see Cannarozzo et al. 2020). We then modified our fiducial sample by excluding the remaining 10 per cent of the LEGA-C galaxies and applied model  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  to this modified fiducial sample, finding inferences on the hyper-parameters (in particular,  $\zeta = 0.408^{+0.032}_{-0.031}$ ) in agreement within  $1\sigma$  with those of  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  shown in Fig. 8. This test indicates that the results obtained for the extended sample should be independent of whether the *UVJ*-colour selection is used as additional criterion to define the sample of ETGs. Of course, the selection of the extended sample is heterogeneous in other respects also, so we cannot exclude that there are other non-negligible biases.

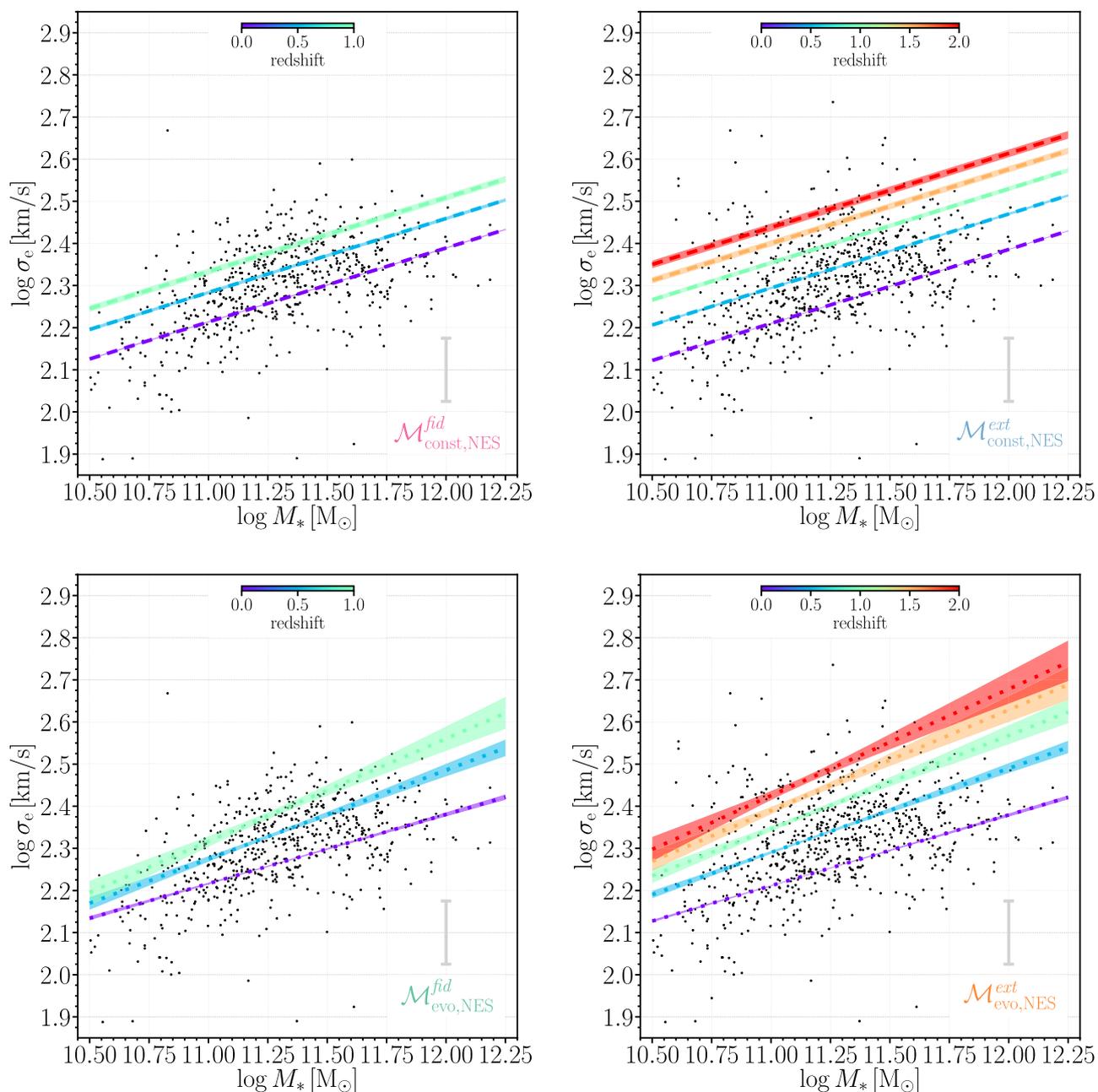
As a more general comment, we note that, for both the fiducial and the extended samples, the results on the evolution of the  $M_*$ – $\sigma_e$  relation hold within the assumption that the slope and the normalization vary as power laws of  $1+z$ . Our inferences on the redshift intervals in which we have no galaxies ( $0.2 < z < 0.6$ ) or very few galaxies ( $z > 1.75$ ) in our samples (Fig. 6) clearly rely on this assumption and are thus driven by the properties of galaxies in other redshift intervals.

## 5 DISCUSSION

### 5.1 Potential systematics

Our inference relies on measurements of the stellar mass and central velocity dispersion of galaxies. Both these quantities are subject to systematic effects that could in principle affect our results. The biggest systematics are those affecting the stellar mass measurements, which we discuss here.

Stellar mass measurements are the results of fits of Sérsic profiles to broad-band photometric data, from which luminosities and colours are derived and subsequently fitted with SPS models. One possible source of systematics is a deviation of the true stellar density profile of a galaxy from a Sérsic profile. For instance, as shown by Sonnenfeld et al. (2019) in their study of a sample of massive ellipticals at  $z \sim 0.6$ , it is difficult to distinguish between a pure Sérsic model and a model consisting of the sum of a Sérsic and an exponential component, even with relatively deep data from the HSC survey. Differences between the two models only arise at large radii and can



**Figure 11.** Central stellar velocity dispersion as a function of stellar mass. In the left-hand panels, the curves represent the median relations of the fiducial-sample models  $\mathcal{M}_{\text{const,NES}}^{\text{fid}}$  (upper panel) and  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$  (lower panel), at  $z = 0$ ,  $z = 0.5$ , and  $z = 1$ . In the right-hand panels, the curves represent the median relations of the extended-sample models  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  (upper panel) and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  (lower panel), at  $z = 0$ ,  $z = 0.5$ ,  $z = 1.5$ , and  $z = 2$ . The bands indicate the  $1\sigma$  uncertainty ranges. The median relations and the  $1\sigma$  bands are computed as described in the caption of Fig. 10. The grey error bars represent the redshift-independent intrinsic scatter. The dots indicate the positions in these diagrams of the ETGs of the fiducial (left-hand panels) and extended (right-hand panels) samples.

lead to variations in the estimated luminosity on the order of 0.1 dex. Secondly, our models assume implicitly that the stellar population parameters of a galaxy are spatially constant. However, if these vary as a function of radius, a bias on the inferred stellar masses could be introduced. More generally, the SPS models on which our  $M_*$  measurements are based are known to be subject to systematics (see e.g. Conroy 2013). Most importantly, uncertainties on the stellar IMF can lead to a global shift of the stellar mass distribution, affecting the

inference on the normalization of the  $M_*-\sigma_e$  relation  $\mu_0$ , and/or the slope of the relation  $\beta_0$ , in case the IMF varies as a function of mass. Additionally, gradients in the IMF can also introduce biases: along with  $M_*/L$  gradients at fixed IMF, these are particularly relevant if our measurements are used to quantify the stellar contribution to the dynamical mass of a galaxy (see e.g. Li et al. 2017; Bernardi et al. 2018; Sonnenfeld et al. 2018; Domínguez Sánchez et al. 2019, and related discussions).

All of these systematic effects are common to virtually all estimates of the  $M_*-\sigma_0$  relation in the literature and are difficult to address, given our current knowledge on the accuracy of our models of galaxy stellar profiles and stellar populations. Nevertheless, they should be taken into consideration when comparing our observations with theoretical models.

## 5.2 Comparison with previous works

In this section, we compare our results on the  $M_*-\sigma_e$  relation with previous works in the literature, which we briefly describe in the following.

(i) Auger et al. (2010) study a sample of 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  $11 < \log(M_*/M_\odot) < 12$ . Auger et al. (2010) report fits of the  $M_*-\sigma_{e/2}$  relation both allowing and not allowing for the presence of intrinsic scatter ( $\sigma_{e/2}$  is the velocity dispersion within an aperture  $R_e/2$ ).

(ii) Hyde & Bernardi (2009a) extract 46 410 ETGs from the SDSS DR4 with parameters updated to the DR6 values (Adelman-McCarthy et al. 2008), selecting galaxies with  $60 < \sigma_{e/8}/(\text{km s}^{-1}) < 400$ , where  $\sigma_{e/8}$  is the stellar velocity dispersion measured within an aperture  $R_e/8$ , in the redshift range  $0.07 < z \leq 0.35$ . Hyde & Bernardi (2009a) fit the distribution of  $\log \sigma_{e/8}$  as a function of  $\log M_*$  both with a linear function, over the stellar mass range  $10.5 < \log(M_*/M_\odot) < 11.5$ , and with a quadratic function, in the range  $9.5 < \log(M_*/M_\odot) < 12$ .

(iii) Damjanov et al. (2018) estimate the  $M_*-\sigma_0$  relation of 565 quiescent galaxies of the hCOS20.6 sample, with  $10.5 < \log(M_*/M_\odot) < 11.4$ , in the redshift range  $0.2 < z < 0.5$ . The velocity dispersions, corrected to an aperture of 3 kpc, can be taken as good approximations (to within 3–4 per cent; I. Damjanov, private communication) of measurements of  $\sigma_e$ .

(iv) 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 of 1316 galaxies drawn from the Smithsonian Hectospec Lensing Survey (SHELS; Geller et al. 2005) at  $0.3 < z < 0.4$ . Also in this case the velocity dispersions, which are corrected to an aperture of 3 kpc, can be taken as measurements of  $\sigma_e$ .

(v) Belli et al. (2014a) measure  $\sigma_e$  and  $M_*$  for a sample of galaxies with median redshift  $z \simeq 1.23$ . We take from Zahid et al. (2016b) the best-fitting parameters of the  $M_*-\sigma_e$  relation for the sample of Belli et al. (2014a).

(vi) Mason et al. (2015) study 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), finding that  $\sigma_e$  at fixed  $M_*$  increases with redshift as  $(1+z)^{0.2}$ . In particular, we consider here the fit of Mason et al. (2015) evaluated at  $z = 0.35, 1.23, \text{ and } 2$ , taking as reference the fit of Auger et al. (2010) with non-zero intrinsic scatter.

In order to compare the results of the different works, we express all the fits in the form

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) = \mu + \beta \log\left(\frac{M_*}{M_*^{\text{piv}}}\right) + \gamma \left[\log\left(\frac{M_*}{M_*^{\text{piv}}}\right)\right]^2, \quad (35)$$

where  $M_*^{\text{piv}} = 10^{11.321} M_\odot$  (Chabrier IMF). We correct for aperture the fits of Auger et al. (2010) and Hyde & Bernardi (2009a) using equation (3), so  $\log \sigma_e = \log \sigma_{e/8} - 0.06$  and  $\log \sigma_e = \log \sigma_{e/2} -$

**Table 7.** Values of the parameters of equation (35), according to the fits of the literature works that we compared with our model.  $\gamma = 0$  in all cases, but in the case of the quadratic fit of Hyde & Bernardi (2009a), for which  $\gamma = -0.044$ .

Redshift	Reference	$\mu$	$\beta$
$z \simeq 0.2$	Auger et al. (2010)	2.38	0.24
	Auger et al. (2010) with intrinsic scatter	2.38	0.18
	Hyde & Bernardi (2009a) Linear fit	2.32	0.29
	Hyde & Bernardi (2009a) Quadratic fit	2.32	0.24
$z \simeq 0.35$	Damjanov et al. (2018)	2.37	0.25
	hCOS20.6 ( $0.2 < z < 0.5$ ) Zahid et al. (2016b)	2.38	0.31
	SHELS ( $0.3 < z < 0.4$ )		
$z = 0.35$	Mason et al. (2015)	2.39	0.18
$z \simeq 1.23$	Belli et al. (2014a)	2.48	0.30
$z = 1.23$	Mason et al. (2015)	2.43	0.18
$z = 2$	Mason et al. (2015)	2.46	0.18

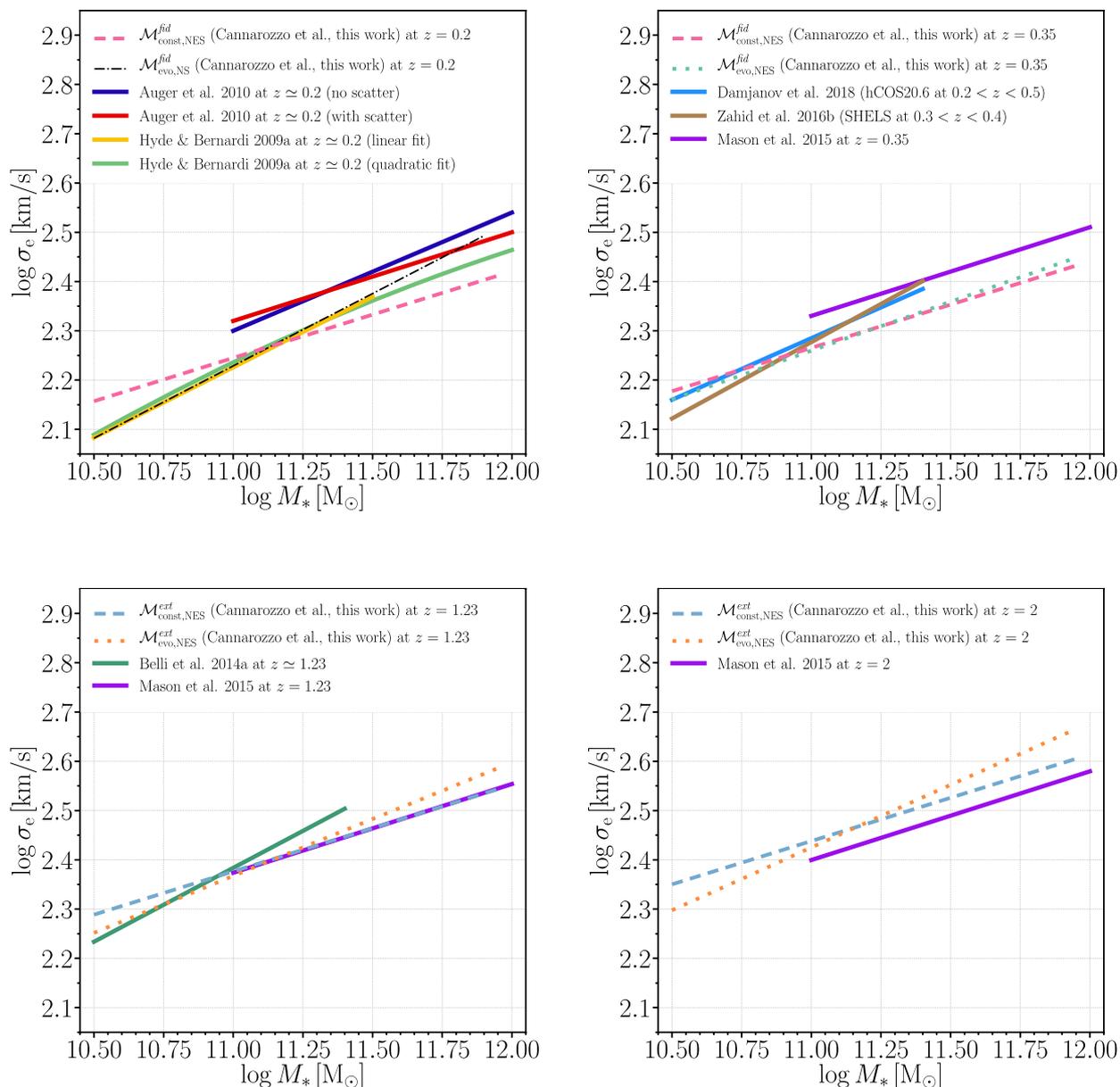
0.02. Except for the quadratic fit of Hyde & Bernardi (2009a), all the other fits assume  $\gamma = 0$  in equation (35). The values of the parameters of equation (35) for the considered literature works are reported in Table 7. In Fig. 12, we show the comparison between our models and the previous works at  $z = 0.2, 0.35, 1.23, \text{ and } 2$ .

At  $z = 0.2$  (upper left-hand panel of Fig. 12) we compare our results with Auger et al. (2010) and Hyde & Bernardi (2009a). The curve of model  $\mathcal{M}_{\text{const, NES}}^{\text{fid}}$  at  $z = 0.2$  intersects the linear fit of Hyde & Bernardi (2009a), but has shallower slope, similar to the fits of Auger et al. (2010), which however have  $\approx 20$  per cent higher normalization. At  $z = 0.2$ , the median relation of model  $\mathcal{M}_{\text{evo, NES}}^{\text{fid}}$ , not shown in the plot, is very similar to that of model  $\mathcal{M}_{\text{const, NES}}^{\text{fid}}$ . The steeper slope  $\beta$  of the fits of Hyde & Bernardi (2009a) can be ascribed to two main reasons: they exclude the highest mass galaxies and they do not allow for intrinsic scatter in their model. The fact that the correlation is shallower at higher  $M_*$  is apparent from the shape of the quadratic fit of Hyde & Bernardi (2009a). The difference between the two fits of Auger et al. (2010) illustrates the effect on  $\beta$  of allowing for intrinsic scatter. As a further test of the importance of considering the intrinsic scatter in the model, we applied to the fiducial sample the analysis described in Section 3, but assuming zero intrinsic scatter ( $\sigma_\sigma = 0$  in equation 23) for models  $\mathcal{M}_{\text{const}}$  and  $\mathcal{M}_{\text{evo}}$ . Based on the Bayesian evidence, in this case the best model is  $\mathcal{M}_{\text{evo, NS}}^{\text{fid}}$ , i.e. an evolving-slope model with a null scatter (NS) that can be approximately described by

$$\log\left(\frac{\sigma_e}{\text{km s}^{-1}}\right) \simeq 2.21 + \beta(z) \left(\frac{M_*}{10^{11} M_\odot}\right) + 0.46 \log(1+z), \quad (36)$$

with  $\beta(z) = 0.22 + 0.89 \log(1+z)$ . This model evaluated at  $z = 0.2$  (dash-dotted curve in upper left-hand panel of Fig. 12) has slope  $\beta \simeq 0.29$  and overlaps almost perfectly with the linear fit of Hyde & Bernardi (2009a).

In the upper right-hand panel of Fig. 12, we compare our models  $\mathcal{M}_{\text{const, NES}}^{\text{fid}}$  and  $\mathcal{M}_{\text{evo, NES}}^{\text{fid}}$  at  $z = 0.35$  with the fits obtained by Mason et al. (2015) at the same redshift, by Damjanov et al. (2018) at  $0.2 < z < 0.5$  and by Zahid et al. (2016b) for SHELS galaxies at  $0.3 < z < 0.4$ . Taking into account the differences in the stellar mass range, and that Damjanov et al. (2018) and Zahid et al. (2016b) do not allow



**Figure 12.** Comparison between the median  $M_*$ – $\sigma_e$  relations of our models  $\mathcal{M}_{\text{const,NES}}$  (dashed curves) and  $\mathcal{M}_{\text{evo,NES}}$  (dotted curves), and fits from the literature (solid curves) at  $z = 0.2$  (upper left-hand panel),  $z = 0.35$  (upper right-hand panel),  $z = 1.23$  (lower left-hand panel), and  $z = 2$  (lower right-hand panel). Upper left-hand panel: the red and blue solid curves are the linear fits of Auger et al. (2010), including and not including the intrinsic scatter, respectively, while the gold and light green solid curves are the linear and the quadratic fits of Hyde & Bernardi (2009a), respectively. The thin dash–dotted curve is the median relation of model  $\mathcal{M}_{\text{evo,NES}}^{\text{fid}}$ , in which we assume zero intrinsic scatter (see text). Upper right-hand panel: the azure curve is the linear fit for the hCOS20.6 galaxies ( $0.2 < z < 0.5$ ) by Damjanov et al. (2018), the brown curve is the linear fit of Zahid et al. (2016b) for the SHELS sample ( $0.3 < z < 0.4$ ), and the violet curve is the fit of Mason et al. (2015) evaluated at  $z = 0.35$ . Lower left-hand panel: the green curve is the linear fit of Belli et al. (2014a) and the violet curve is the fit of Mason et al. (2015) evaluated at  $z = 1.23$ . Lower right-hand panel: the violet curve is the fit of Mason et al. (2015) evaluated at  $z = 2$ . Each curve is shown over the stellar mass range spanned by the considered data set.

for the presence of intrinsic scatter, there is reasonable agreement among the five curves.

In the lower left-hand panel of Fig. 12, we compare our models  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  at  $z = 1.23$  (mean redshift of the sample of Belli et al. 2014a) with the linear fit obtained by Belli et al. (2014a) and that of Mason et al. (2015) at the same redshift. Considering that Belli et al. (2014a) do not allow for the presence of intrinsic scatter, the four curves are broadly consistent.

In the lower right-hand panel of Fig. 12, we compare the median relations of our models  $\mathcal{M}_{\text{const,NES}}^{\text{ext}}$  and  $\mathcal{M}_{\text{evo,NES}}^{\text{ext}}$  with the estimate of Mason et al. (2015) at  $z = 2$ , finding that our relations predict a higher velocity dispersion at the same stellar mass, which is a consequence of the fact that the Mason et al. (2015) find a weaker redshift dependence of the normalization than our models.

Overall, we do find a satisfactory agreement among our results and previous works in the literature. Some of the differences pointed

out above may be ascribed to different redshift distributions of the galaxy sample, stellar mass ranges, data and models used in the measurements of the stellar masses, and selection criteria or fitting methods. For instance, it is apparent from Fig. 12 that different studies consider different stellar mass intervals. Studies focusing on lower stellar masses tend to find steeper slopes than those focusing on higher stellar masses. Furthermore, allowing for the presence of intrinsic scatter when modelling the data leads to shallower slopes. Models allowing for the presence of intrinsic scatter, such as those presented in this work, are expected to provide a more correct description of the  $M_*$ – $\sigma_e$  correlation.

### 5.3 Connection with the size evolution of ETGs

It is useful to discuss the results here obtained for the evolution of the  $M_*$ – $\sigma_e$  relation of ETGs in light of the well-known evolution of the  $M_*$ – $R_e$  relation: the redshift dependence of the median effective radius at fixed stellar mass can be parametrized as  $R_e \propto (1+z)^{a_R}$ . The value of  $a_R$  for ETGs appears to depend somewhat on the considered sample, mass, and redshift intervals, ranging from  $a_R \approx -1.5$  (van der Wel et al. 2014;  $0 \lesssim z \lesssim 3$ ) to  $a_R \approx -0.6$  (Cimatti et al. 2012;  $0 \lesssim z \lesssim 2$ ). If all the ETGs in the considered redshift range were structurally and kinematically homologous (see section 5.4.1 of Cimatti, Fraternali & Nipoti 2019), we would have  $\sigma_e^2 \propto M_*/R_e$  and thus, at fixed stellar mass  $M_*$ ,  $\sigma_e \propto R_e^{-1/2} \propto (1+z)^{\zeta_{\text{hom}}}$ , with  $\zeta_{\text{hom}} = -a_R/2$ . For  $-1.5 \lesssim a_R \lesssim -0.6$ , we get  $0.3 \lesssim \zeta_{\text{hom}} \lesssim 0.75$ . This toy model is consistent with our observational finding  $\sigma_e \propto (1+z)^\zeta$  with  $0.4 \lesssim \zeta \lesssim 0.5$ .

It must be stressed that the observed value of  $\zeta$  must not be necessarily equal to  $\zeta_{\text{hom}}$ . From a theoretical point of view, an observed evolution in  $\sigma_e$  different than predicted by the above toy model can be expected if ETGs do not evolve maintaining homology. For instance, dry merging, which is one of the processes believed to be responsible for the size and velocity dispersion evolution of ETGs (see Section 1), is known to produce non-homology, because it varies the shape and the kinematics of the stellar distribution, and the mutual density distributions of luminous and dark matter (Nipoti et al. 2003; Hilz et al. 2013; Frigo & Balcells 2017). We can quantify the effect of non-homology by defining the dimensionless parameter

$$k_* \equiv \frac{GM_*}{\sigma_e^2 R_e}, \quad (37)$$

such that galaxies that are structurally and kinematically homologous have the same value of  $k_*$ . If, at fixed  $M_*$ ,  $R_e \propto (1+z)^{a_R}$  and  $\sigma_e \propto (1+z)^\zeta$ , the average value of  $k_*$  must vary with redshift as  $k_* \propto (1+z)^{a_k}$  with  $a_k = -(2\zeta + a_R)$ . Thus, we have  $\zeta \neq -a_R/2$  if  $a_k \neq 0$ , i.e. if, on average, galaxies at different redshift have different  $k_*$ . However, from an observational point of view, a significant evolution of  $k_*$  seems to be excluded. Defining the dynamical mass  $M_{\text{dyn}} \equiv 5\sigma_e^2 R_e/G$ , the average ratio  $M_*/M_{\text{dyn}} \propto k_*$  is found to increase mildly with redshift, or even remain constant (van de Sande et al. 2013; Belli et al. 2014a), and the zero-point of the stellar mass Fundamental Plane, which can also be seen as a measure of the average  $k_*$ , varies only little with redshift (Bezanson et al. 2013a; Zahid et al. 2016a).

## 6 CONCLUSIONS

We have studied the evolution of the correlation between central stellar velocity dispersion  $\sigma_e$  (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:

(i) On average, the central velocity dispersion of massive ( $M_* \gtrsim 10^{10.5} M_\odot$ ) ETGs increases with stellar mass following a power-law relation  $\sigma_e \propto M_*^\beta$  with either  $\beta \simeq 0.18$ , independent of redshift, or  $\beta$  increasing with redshift as  $\beta \simeq 0.16 + 0.26 \log(1+z)$  in the redshift range  $0 \lesssim z \lesssim 1$  probed by our fiducial sample.

(ii) The normalization of the  $M_*$ – $\sigma_e$  relation increases with redshift: for instance, when  $\beta \simeq 0.18$  independent of redshift, at fixed stellar mass  $\sigma_e \propto (1+z)^\zeta$  with  $\zeta \simeq 0.4$  out to  $z \approx 1$ . In other words, a typical ETG of  $M_* \approx 10^{11} M_\odot$  at  $z \approx 0$  has  $\sigma_e$  lower by a factor of  $\approx 1.3$  than ETGs of similar stellar mass at  $z \approx 1$ .

(iii) The intrinsic scatter of the  $M_*$ – $\sigma_e$  relation is  $\simeq 0.08$  dex in  $\sigma_e$  at given  $M_*$ , independent of redshift.

(iv) Over the wider redshift range  $0 \lesssim z \lesssim 2.5$ , probed by our extended sample, we find results similar to those found for the  $z \lesssim 1$  fiducial sample, with slightly stronger redshift dependence of the normalization ( $\zeta \simeq 0.5$ ) and weaker redshift dependence of the slope ( $d\beta/d\log(1+z) \simeq 0.18$ ) when  $\beta$  varies with time. On average, the velocity dispersion of ETGs of  $M_* \approx 10^{11} M_\odot$  at  $z = 2$  is a factor of  $\approx 1.7$  higher than that of  $z = 0$  ETGs of similar stellar mass.

The results of this work confirm and strengthen previous indications that the  $M_*$ – $\sigma_e$  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, 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  $M_*$ – $\sigma_e$  plane in the direction of increasing  $M_*$ . As pointed out in Section 1, the variation of  $\sigma_e$  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 Fig. 11 could be reproduced by individual galaxies evolving at decreasing  $\sigma_e$ , but, at least at the low-mass end, even an evolution of individual galaxies at constant or slightly increasing  $\sigma_e$  is not excluded. Remarkably, our results suggest that, on average, the stellar velocity dispersion of individual galaxies with  $M_* \gtrsim 3 \times 10^{11} M_\odot$  at  $z \approx 1$  must decrease from  $z \approx 1$  to  $z \approx 0$  for them to end up on the median present-day  $M_*$ – $\sigma_e$  relation.

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  $M_*-\sigma_e$  relation of ETGs with the results of state-of-the-art cosmological simulations of galaxy formation.

## ACKNOWLEDGEMENTS

We are grateful to S. Belli and M. Maseda for sharing data and for helpful suggestions, and to A. Cimatti, I. Damjanov, S. Faber, M. Moresco and B. Nipoti for useful discussions. We thank the anonymous referees for their advice and comments that considerably contributed to improve the quality of the manuscript. AS acknowledges funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 792916.

## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Adelman-McCarthy J. K. et al., 2008, *ApJS*, 175, 297
- Aihara H. et al., 2018, *PASJ*, 70, S8
- Aihara H. et al., 2019, *PASJ*, 71, 114
- 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., 2014a, *ApJ*, 783, 117 (B14)
- Belli S., Newman A. B., Ellis R. S., Konidaris N. P., 2014b, *ApJ*, 788, L29
- Belli S., Newman A. B., Ellis R. S., 2017, *ApJ*, 834, 18 (B17)
- Bernardi M., Sheth R. K., Domínguez-Sánchez H., Fischer J. L., Chae K. H., Huertas-Company M., Shankar F., 2018, *MNRAS*, 477, 2560
- 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
- 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
- Cannarozzo C., Nipoti C., Sonnenfeld A., Leauthaud A., Huang S., Diemer B., Oyarzún G., 2020, preprint ([arXiv:2006.05427](https://arxiv.org/abs/2006.05427))
- Cappellari M. et al., 2006, *MNRAS*, 366, 1126
- Chabrier G., 2003, *PASP*, 115, 763
- Cimatti A., Nipoti C., Cassata P., 2012, *MNRAS*, 422, L62
- Cimatti A., Fraternali F., Nipoti C., 2019, *Introduction to Galaxy Formation and Evolution: From Primordial Gas to Present-Day Galaxies*. Cambridge Univ. Press, Cambridge
- Ciotti L., Lanzoni B., Volonteri M., 2007, *ApJ*, 658, 65
- Conroy C., 2013, *ARA&A*, 51, 393
- Damjanov I., Zahid H. J., Geller M. J., Fabricant D. G., Hwang H. S., 2018, *ApJS*, 234, 21
- 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
- Domínguez Sánchez H., Huertas-Company M., Bernardi M., Tuccillo D., Fischer J. L., 2018, *MNRAS*, 476, 3661
- Domínguez Sánchez H., Bernardi M., Brownstein J. R., Drory N., Sheth R. K., 2019, *MNRAS*, 489, 5612
- 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, 39
- 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 (G15)
- 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, *Commun. Appl. Math. Comput. Sci.*, 5, 65
- Hilz M., Naab T., Ostriker J. P., 2013, *MNRAS*, 429, 2924
- Huang S., Leauthaud A., Greene J. E., Bundy K., Lin Y. T., Tanaka M., Miyazaki S., Komiyama Y., 2018, *MNRAS*, 475, 3348
- Hyde J. B., Bernardi M., 2009a, *MNRAS*, 394, 1978
- Hyde J. B., Bernardi M., 2009b, *MNRAS*, 396, 1171
- Jeffreys H., 1961, *Theory of Probability*, 3rd edn. Oxford Univ. Press, Oxford
- Kormendy J., 1977, *ApJ*, 218, 333
- Le Fèvre O. et al., 2003, in Iye M., Moorwood A. F. M., eds, *Proc. SPIE Conf. Ser. Vol. 4841, Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*. SPIE, Bellingham, p. 1670
- Li H. et al., 2017, *ApJ*, 838, 77
- López-Sanjuán C. et al., 2012, *A&A*, 548, A7
- Mason C. A. et al., 2015, *ApJ*, 805, 79
- Mendel J. T., Simard L., Palmer M., Ellison S. L., Patton D. R., 2014, *ApJS*, 210, 3 (M14)
- Miyazaki S. et al., 2018, *PASJ*, 70, S1
- Moresco M. et al., 2013, *A&A*, 558, A61
- 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
- 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
- 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
- Scoville N. et al., 2007, *ApJS*, 172, 1
- Sérsic J. L., 1968, *Atlas de Galaxias Australes*, Observatorio Astronomico, Cordoba, Argentina
- Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, *AIP Conf. Proc. Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Am. Inst. Phys., New York, p. 395
- Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
- Sonnenfeld A., Leauthaud A., Auger M. W., Gavazzi R., Treu T., More S., Komiyama Y., 2018, *MNRAS*, 481, 164

- Sonnenfeld A., Wang W., Bahcall N., 2019, *A&A*, 622, A30  
Straatman C. M. S. et al., 2018, *ApJS*, 239, 27  
Strauss M. A. et al., 2002, *AJ*, 124, 1810  
Tanaka M. et al., 2019, *ApJ*, 885, L34  
van de Sande J. et al., 2013, *ApJ*, 771, 85 (vdS13)  
van der Wel A. et al., 2014, *ApJ*, 788, 28  
van der Wel A. et al., 2016, *ApJS*, 223, 29  
Westera P., Lejeune T., Buser R., Cuisinier F., Bruzual G., 2002, *A&A*, 381, 524  
Zahid H. J., Geller M. J., 2017, *ApJ*, 841, 32  
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

## SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](#) online.

### **The\_cosmic\_evolution\_of\_the\_stellar\_mass\_velocity\_dispersion\_relation\_of\_early\_type\_galaxies\_APPENDIX.PDF**

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

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