

# The ATLAS<sup>3D</sup> Project - XXX. Star formation histories and stellar population scaling relations of early-type galaxies

Richard M. McDermid<sup>1,2\*</sup>, Katherine Alatalo<sup>3</sup>, Leo Blitz<sup>4</sup>, Frédéric Bournaud<sup>5</sup>, Martin Bureau<sup>6</sup>, Michele Cappellari<sup>6</sup>, Alison F. Crocker<sup>7</sup>, Roger L. Davies<sup>6</sup>, Timothy A. Davis<sup>8</sup>, P. T. de Zeeuw<sup>9,10</sup>, Pierre-Alain Duc<sup>5</sup>, Eric Emsellem<sup>9,11</sup>, Sadegh Khochfar<sup>12</sup>, Davor Krajnović<sup>13</sup>, Harald Kuntschner<sup>9</sup>, Raffaella Morganti<sup>14,15</sup>, Thorsten Naab<sup>16</sup>, Tom Oosterloo<sup>14,15</sup>, Marc Sarzi<sup>16</sup>, Nicholas Scott<sup>17</sup>, Paolo Serra<sup>14,18</sup>, Anne-Marie Weijmans<sup>19</sup> and Lisa M. Young<sup>20,21</sup>

<sup>1</sup>*Department of Physics and Astronomy, Macquarie University, Sydney NSW 2109, Australia*

<sup>2</sup>*Australian Gemini Office, Australian Astronomical Observatory, PO Box 915, Sydney NSW 1670, Australia*

<sup>3</sup>*Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, California 91125, USA*

<sup>4</sup>*Department of Astronomy, Campbell Hall, University of California, Berkeley, CA 94720, USA*

<sup>5</sup>*Laboratoire AIM Paris-Saclay, CEA/IRFU/SAP – CNRS – Université Paris Diderot, 91191 Gif-sur-Yvette Cedex, France*

<sup>6</sup>*Sub-Department of Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, UK*

<sup>7</sup>*Ritter Astrophysical Observatory, University of Toledo, Toledo, OH 43606, USA*

<sup>8</sup>*Centre for Astrophysics Research, University of Hertfordshire, Hatfield, Herts AL1 9AB, UK*

<sup>9</sup>*European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany*

<sup>10</sup>*Sterrewacht Leiden, Leiden University, Postbus 9513, 2300 RA Leiden, the Netherlands*

<sup>11</sup>*Université Lyon 1, Observatoire de Lyon, Centre de Recherche Astrophysique de Lyon and Ecole Normale Supérieure de Lyon, 9 avenue Charles André, F-69230 Saint-Genis Laval, France*

<sup>12</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, UK*

<sup>13</sup>*Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternwarte 16, D-14482 Potsdam, Germany*

<sup>14</sup>*Netherlands Institute for Radio Astronomy (ASTRON), Postbus 2, 7990 AA Dwingeloo, The Netherlands*

<sup>15</sup>*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, 9700 AV Groningen, The Netherlands*

<sup>16</sup>*Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany*

<sup>17</sup>*Sydney Institute for Astronomy (SIfA), School of Physics, The University of Sydney, NSW 2006, Australia*

<sup>18</sup>*CSIRO Astronomy & Space Science, PO Box 76, Epping, NSW 1710, Australia*

<sup>19</sup>*School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews KY16 9SS, UK*

<sup>20</sup>*Physics Department, New Mexico Institute of Mining and Technology, Socorro, NM 87801, USA*

<sup>21</sup>*Academia Sinica Institute of Astronomy & Astrophysics, PO Box 23-141, Taipei 10617, Taiwan, R.O.C.*

\* *richard.mcdermid@mq.edu.au*

Accepted 2015 January 15. Received 2015 January 15; in original form 2014 October 14

arXiv:1501.03723v1 [astro-ph.GA] 15 Jan 2015

**ABSTRACT**

We present the stellar population content of early-type galaxies from the ATLAS<sup>3D</sup> survey. Using spectra integrated within apertures covering up to one effective radius, we apply two methods: one based on measuring line-strength indices and applying single stellar population (SSP) models to derive SSP-equivalent values of stellar age, metallicity, and alpha enhancement; and one based on spectral fitting to derive non-parametric star-formation histories, mass-weighted average values of age, metallicity, and half-mass formation timescales. Using homogeneously derived effective radii and dynamically-determined galaxy masses, we present the distribution of stellar population parameters on the Mass Plane ( $M_{\text{JAM}}$ ,  $\sigma_e$ ,  $R_e^{\text{maj}}$ ), showing that at fixed mass, compact early-type galaxies are on average older, more metal-rich, and more alpha-enhanced than their larger counterparts.

From non-parametric star-formation histories, we find that the duration of star formation is systematically more extended in lower mass objects. Assuming that our sample represents most of the stellar content of today's local Universe, approximately 50% of all stars formed within the first 2 Gyr following the big bang. Most of these stars reside today in the most massive galaxies ( $> 10^{10.5} M_{\odot}$ ), which themselves formed 90% of their stars by  $z \sim 2$ . The lower-mass objects, in contrast, have formed barely half their stars in this time interval. Stellar population properties are independent of environment over two orders of magnitude in local density, varying only with galaxy mass. In the highest-density regions of our volume (dominated by the Virgo cluster), galaxies are older, alpha-enhanced and have shorter star-formation histories with respect to lower density regions.

**Key words:** galaxies: abundances galaxies: elliptical and lenticular, cD galaxies: evolution galaxies: stellar content

**1 INTRODUCTION**

Our understanding of the stellar populations in early-type galaxies (hereafter ETGs) has made significant progress in recent years, through a combination of advancements in both observations and stellar population models. These systems are often cited as ideal laboratories for uncovering the fossil record of galaxy formation processes thanks to the apparent lack of dust or star formation, which obfuscates the stellar continuum in objects of later type. However, the modern observational picture is not quite so straightforward. ETGs are now known to host some degree of young stellar populations (e.g. Trager et al. 2000; Kaviraj et al. 2007; Kuntschner et al. 2010). More surprising is the recent finding that sizable reservoirs of molecular gas can be found residing inside ETGs (Young et al. 2008), coincident with evidence of PAH features in the infrared domain (Shapiro et al. 2010), both of which strongly indicate ongoing star-formation within these galaxies, albeit at a rather low level (Davis et al. 2014).

In addition to a large range in ages, ETGs are also found to show strong correlations between their population parameters and mass-sensitive quantities such as luminosity and velocity dispersion (e.g. Jørgensen 1999; Trager et al. 2000; Thomas et al. 2005). In general, these studies find that high-mass galaxies appear older, metal rich and with enhanced abundance ratios with respect to lower mass objects - a picture emphatically confirmed with very large samples of galaxies from modern surveys (Bernardi et al. 2006; Galazzi et al. 2006; Graves et al. 2009a; Thomas et al. 2010). In addition to this general behaviour, specific co-dependencies between population parameters after controlling for the general mass-dependency have also been proposed (Trager et

al. 2000; Smith et al. 2008; Graves & Faber 2010). Understanding what drives the large intrinsic scatter in the stellar population scaling relations is key to understanding the evolutionary processes involved.

Despite a deepening clarity in our observational view of ETGs, many open questions remain. For example, the role of young stars in these galaxies is still not clear. Although often referred to as ‘rejuvenated’ galaxies, it is currently unclear if these objects have actually rekindled their star formation in recent times, or contrarily, are in the process of recently shutting down star formation for the first time. The role of environment is also somewhat unclear. Previously considered as important for ETGs in general (Thomas et al. 2005; Schawinski et al. 2007), the interplay of mass and environment may also play a role (Thomas et al. 2010).

Here we present the stellar population properties of the ATLAS<sup>3D</sup> sample (Cappellari et al. 2011a) using two approaches: one which uses line strengths to derive single stellar population- (SSP-) equivalent parameters to describe the age, metallicity and alpha-enhancement; and one which uses spectral fitting to obtain a non-parametric star formation history that can be used to derive mass-weighted average age and metallicity, and estimate the duration of star formation. We compare both approaches, which yield complementary information, and check for self-consistency between them. We then examine the trends of stellar populations with various other properties. The ATLAS<sup>3D</sup> project comprises a well-characterized sample with various multi-wavelength data in addition to the SAURON integral field spectroscopy, making it a powerful data set to explore correlations between multiple properties.

This paper proceeds as follows: Section 2 describes the observations and analysis methods; Sections 3 and 4 present our population parameters on the mass-size plane, and against velocity dispersion and mass, respectively; Section 5 presents trends from our empirical star formation histories, and Section 6 concludes.

## 2 OBSERVATIONS, LINE-STRENGTHS AND ANALYSIS

Here we describe the optical spectroscopy used in our analysis. The same spectral data cubes were used for deriving the line strengths and for the spectral fitting, so calibration and removal of emission applies to both techniques.

### 2.1 SAURON observations and basic calibration

The SAURON spectrograph and basic data reduction is described in detail in Bacon et al. (2001). Further details of the ATLAS<sup>3D</sup> SAURON observations are given in Cappellari et al. (2011a). The details of our calibrations and line-strength measurements follow closely the techniques developed and described by Kuntschner et al. (2006). Most of the 260 ATLAS<sup>3D</sup> SAURON spectra were obtained after the spectrograph was upgraded with a volume-phase holographic (VPH) grating in 2004, except for 64 objects observed as part of the SAURON Survey (de Zeeuw et al. 2002) and other related projects. The pre- and post-VPH data are reduced and analysed self-consistently to respect the slightly different dispersion characteristics. The wavelength coverage of the pre- and post-VPH data are very similar, with only a slight difference in resolution. The adopted spectral resolution for pre-VPH data is  $110 \text{ km s}^{-1}$  compared to  $99 \text{ km s}^{-1}$  for VPH data. Observations of the almost featureless white dwarf EG131 were used to verify that the detailed transmission characteristics, including the medium-scale variations described in Kuntschner et al. (2006), were present in the pre- and post-VPH data, agreeing to better than 1%. Corrections for air-mass were applied using the average La Palma extinction curve (King 1985).

A total of 67 different standard stars were observed during both the SAURON and ATLAS<sup>3D</sup> campaigns in order to calibrate our line-strengths onto the standard Lick system (Worthey et al. 1994), as well as to monitor stability and reproducibility in our measurements (160 separate observations were made). Offsets to the Lick system are derived in Appendix A, along with a comparison of pre- and post-VPH values, a comparison with the MILES library (Sánchez-Blázquez et al. 2006), and discussion of the intrinsic uncertainties in our measurements. Appendix B presents a comparison of our line-strength measurements on the Lick system with González (1993) and Trager et al. (1998) for overlapping galaxies, showing good agreement given the uncertainties of matching apertures and differing treatment of emission.

### 2.2 Emission correction

The SAURON spectral range includes several nebular emission lines, all of which lie close to or within the bandpasses of our line-strength indices. Namely: H $\beta$ , present in both emission

and absorption; [OIII] $\lambda\lambda 4959, 5007$ , which strongly affects the Fe5015 index; and [NI] $\lambda\lambda 5198, 5200$ , which falls within the Mg *b* index. As in Kuntschner et al. (2006), we employ the technique developed by Sarzi et al. (2006), which involves simultaneously finding the best linear combination of absorption- and emission-line templates whilst fitting non-linearly for the emission line kinematics. The kinematics of the various lines are tied to that of [OIII], being the more robustly measured line. The resulting emission-line fits are subtracted from the galaxy spectrum if the AoN (ratio of the amplitude of the fitted line to the local noise level in the spectrum, which includes fit residuals) is significant. Following Sarzi et al. (2006), the AoN thresholds used are  $\text{AoN} > 4$  for [OIII], and then if this is detected,  $\text{AoN} > 3$  for H $\beta$  and  $\text{AoN} > 4$  for [NI].

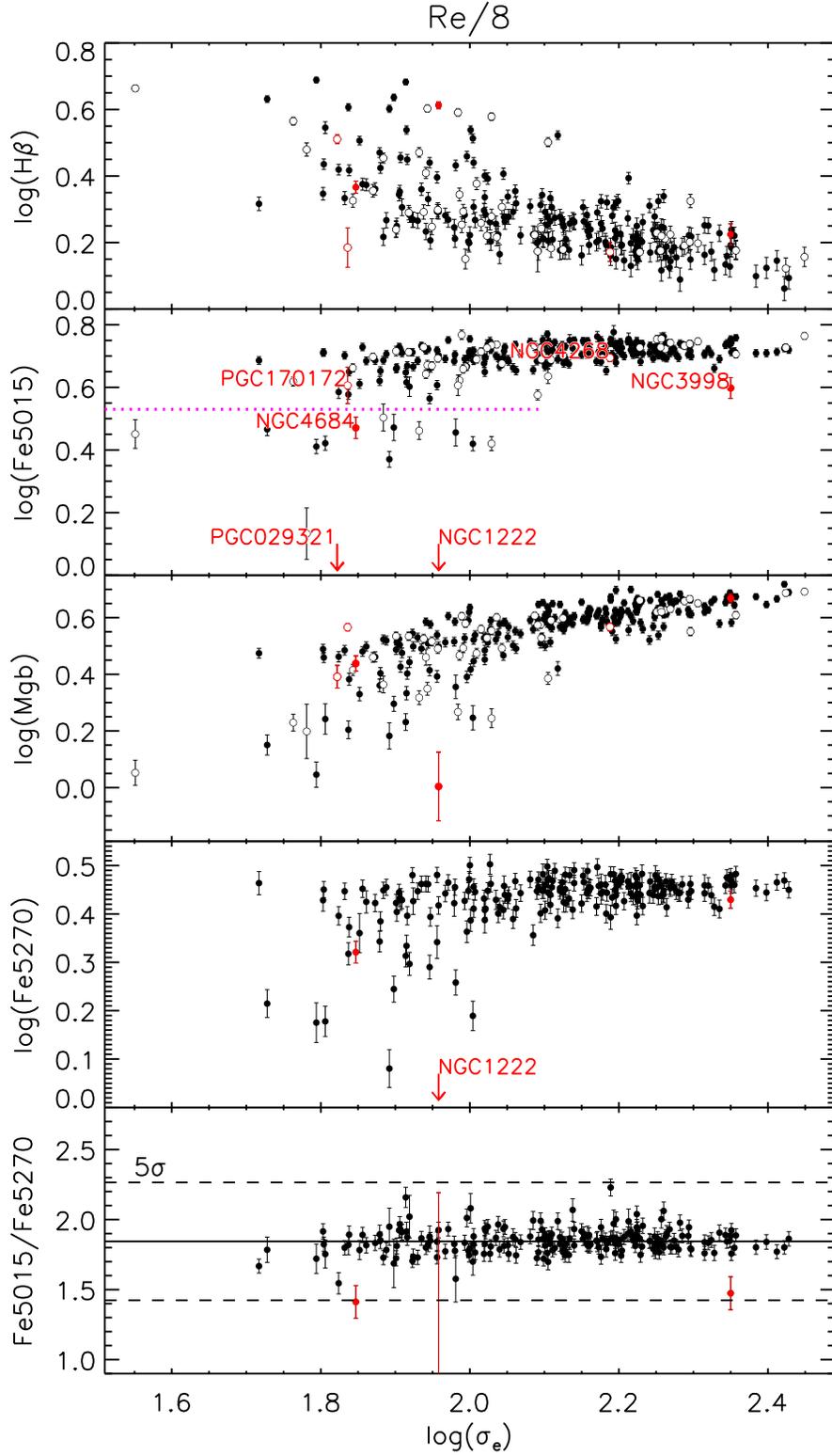
The nebular emission lines are fitted and subtracted on the spatially-resolved data-cubes (after spatially binning to a threshold signal-to-noise ratio of 40 per pixel with the Voronoi binning software of Cappellari & Copin 2003) before generating the aperture spectra. This largely avoids strongly non-Gaussian emission-line profiles (due to the rotation and/or distribution of the gas), which can leave significant residuals.

To estimate how strongly the various indices are affected by residual under- or over-subtracted emission, we use the spectral template fits to provide a good representation of the spectral lines in the absence of any emission. We then compare the observed indices with those measured on the template fits. The differences between the observed and template indices are typically similar to or smaller than the statistical errors on the line measurements. However, a few galaxies show significantly larger differences, due to imperfect emission correction in the presence of very strong emission lines with complex, non-Gaussian line profiles. Rather than use the index measured on the template, which is known to be biased toward solar abundance ratios, we adopt the *difference* between the observed and best-fit template indices as an estimate of the systematic errors. We furthermore take the statistical uncertainties from repeat observations of standard stars as a lower limit on the errors (see Table A2 given in Appendix A).

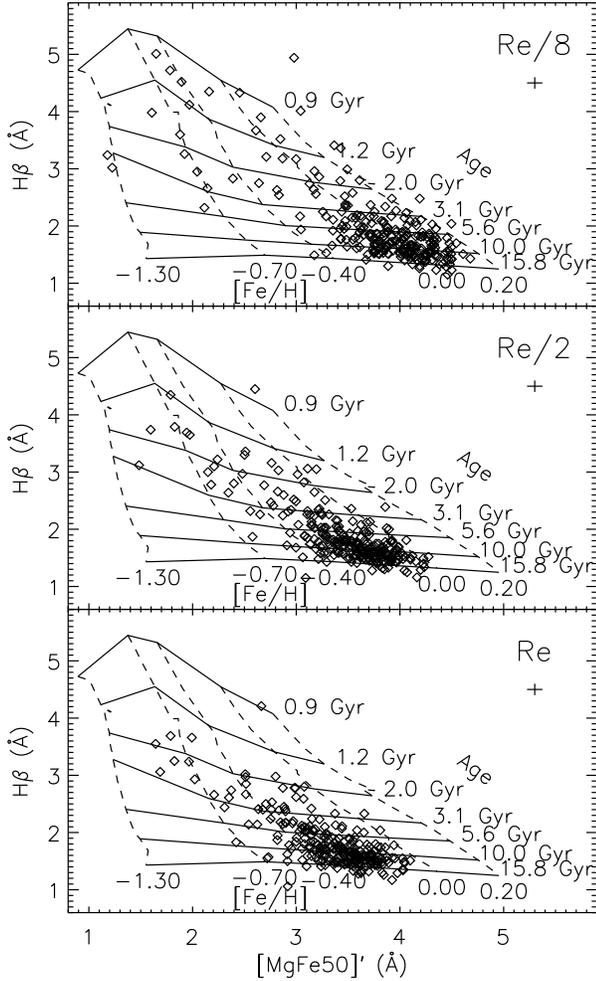
### 2.3 Aperture line strength measurements

Figure 1 of Emsellem et al. (2011) shows the distribution of radii covered by our integral-field spectroscopic data, defined as the maximum circular radius,  $R_{\text{max}}$ , that is at least 85% filled with spectra. Note that this fill factor includes lost spaxels removed due to e.g. foreground stars. The median maximal radius is  $\sim 1.2 R_e$ , with only 10 galaxies having coverage of less than  $R_e/2$ . In terms of aperture coverage, more than half the sample has a filling factor of 85% or more within a circle of  $R_e$ , and all but 11 galaxies have this coverage at  $R_e/2$ .

We measure line strengths in three different circular aperture sizes:  $R_e/8$ ,  $R_e/2$  and  $R_e$ . The smallest of these is comparable to aperture sizes commonly used in conventional long-slit investigations (e.g. Trager et al. 2000; Thomas et al. 2005). The  $R_e/2$  aperture is our largest common aperture that requires minimal aperture corrections in a small number of objects. The  $R_e$  aperture is the most representative ‘global’ value that we can derive based on our data.



**Figure 1.** Absorption line indices measured within  $R_e/8$  against the effective velocity dispersion,  $\sigma_e$  (measured within one  $R_e$ , taken from Cappellari et al. 2013a). Red symbols and text indicate problematic galaxies as described in the main text. The dotted magenta line in the Fe5015 plot indicates the separation of ‘low-Fe’ objects discussed in the text, given here as  $\log(Fe5015) = 0.53$ . The solid line in the bottom panel indicates the biweight mean value, 1.84, with dashed lines showing the  $\pm 5$  times the standard deviation around this value. Closed and open symbols in the upper panels correspond to galaxies with and without Fe5270 measurements respectively.



**Figure 2.** Index-index diagram for our sample using the indices  $H\beta$  versus  $[MgFe50]'$ , overplotted with model predictions from Schiavon (2007) for single stellar populations with constant ages, metallicities and solar abundance ratios. From top to bottom, each panel uses indices measured within different circular apertures, with radii of  $R_e/8$ ,  $R_e/2$  and  $R_e$  as indicated. An error bar in each panel indicates the statistical errors.

To generate the aperture measurements from the data-cubes, we use an analogy of ‘growth curves’, co-adding spectra within increasing circular aperture sizes and measuring the indices at each step. We assign a characteristic radius to each aperture spectrum that represents the radius of a circle having the same area as that of the included SAURON data, computed from the actual bins within the aperture. Thus, where the aperture is well-filled by the SAURON data, this radius is similar to the radius of the desired circular aperture. But as the aperture becomes more poorly filled near the edge of the field, the characteristic radius will tend to a fixed value (the area of the full observed field) regardless of the target aperture size. By using a threshold filling factor of 85%, the difference between the target aperture and delivered area is small.

For the galaxies where the maximum characteristic radius is smaller than the target aperture radius, an aperture correction is applied to the measured indices. Due to the

presence of age gradients that systematically vary in strength with galaxy age, we derive age-dependent aperture corrections. Appendix C gives details of this correction and of the coefficients to apply such corrections. The aperture corrections for our sample are typically 2% or less, with a maximum of 8% for our largest correction, and our overall results do not depend on these corrections.

The line indices themselves are measured on the Lick/IDS system (Worthey et al. 1994) following the standard treatment of broadening the spectra as per the Lick broadening function. We account for the dilution of the line strength due to velocity broadening by measuring the equivalent indices on the broadened and unbroadened optimal template fit, thus accounting for line-strength and LOSVD dependencies of the velocity dispersion corrections (Kuntschner et al. 2006; McDermid et al. 2006). We measure the following indices in all galaxies and aperture sizes:  $H\beta$ , Fe5015 and  $Mgb$ . The SAURON wavelength range has a red cut-off wavelength that changes as a function of position in the field, causing Kuntschner et al. (2006) to define a new index, Fe5270<sub>S</sub>, which linearly traces the Fe5270 Lick index. We use this modified index to estimate the standard Lick Fe5270 index for a subset of objects and apertures where the SAURON wavelength coverage allows.<sup>1</sup> The line strengths measured within the three aperture sizes considered are given in Tables 1–3, which are available in full via the online journal, and also via our project website <http://purl.com/atlas3d>.

Figure 1 presents our line indices measured within one of the three apertures,  $R_e/8$ , plotted as a function of the velocity dispersion measured within one  $R_e$ ,  $\log(\sigma_e)$  taken from Cappellari et al. (2013a). Filled symbols denote the galaxies where Fe5270 could be measured within the given aperture size, with the remainder plotted as open symbols.

As found by numerous other authors (e.g. Trager et al. 2000; Thomas et al. 2005), there are rather tight, positive correlations between the metal-sensitive lines and velocity dispersion, and the age-sensitive  $H\beta$  index shows a negative trend with  $\sigma_e$ . All indices show an increased scatter at lower  $\sigma_e$  values, implying a greater variety of apparent stellar population properties at lower masses.

The Fe5015 index lies on top of the strong [OIII] emission line at 5007Å. For this index, the apparent scatter is driven by a number of galaxies sitting below the main relation. To test whether this is the result of infill from poorly-subtracted [OIII] emission or genuinely weaker iron, we use the fact that some galaxies have a measurable Fe5270 index, which is free from nebular emission contamination. Figure 1 shows several galaxies where Fe5270 also lies below the general relation, but the ratio of the two iron indices (bottom panel of Figure 1) is consistent with the sample average, demonstrating that the weak Fe5015 measurements in these cases are not due to poor emission-line correction. We discuss these ‘low-Fe’ objects further in Section 2.6.

We highlight a few objects in Figure 1 that need special treatment. Firstly, two objects had significantly lower signal-to-noise ratio than the rest of the sample, namely NGC 4268 and PGC 170172. These two objects are not considered further in our analysis. In addition, the following objects show

<sup>1</sup> We derive Fe5270 using the empirical relation  $Fe5270 = 1.28 \times Fe5270_S + 0.03$  from Kuntschner et al. (2006)

large errors in their Fe5015 values due to imperfect subtraction of the very strong and non-Gaussian [OIII] emission line, as indicated by their spuriously low Fe5015/Fe5270 ratio in Figure 1: NGC 1222 (starburst galaxy, Balzano 1983), NGC 4684, NGC 5273, and NGC 3998. These latter two objects are known LINERs with a broad nuclear emission line component (Véron-Cetty & Véron 2010) that only affects the small  $R_e/8$  aperture size, and so we include them when considering the larger aperture measurements. NGC 1222 is excluded for all aperture sizes. PGC 029321 has no Fe5270 measurement, but shows large discrepancies between the Fe5015 index measured on the observed and fitted spectrum due to emission residuals, and so is also excluded in our analysis. For all the problematic cases mentioned here, we still provide measured values in Tables 1–3 where possible, but flag them in the ‘quality’ column.

## 2.4 Deriving SSP parameters

As in McDermid et al. (2006), we make use of a customized version of the Schiavon (2007) models where  $[\text{Ti}/\text{Fe}] = 0.5 \times [\text{Mg}/\text{Fe}]$ . These models were found to give better agreement between abundance ratios computed with Fe5015 and Fe5270 (see Kuntschner et al. 2010, for details), and this relative scaling of  $[\text{Ti}/\text{Fe}]$  and  $[\text{Mg}/\text{Fe}]$  is also approximately consistent with the measured abundances for early-type galaxies from Johansson et al. (2012). The models make predictions of the Lick indices for a grid of various ages, metallicities and alpha-element enhancements. Each position in this grid represents the prediction for a single population of stars sharing these parameters.

We show in Figure 2 the distribution of our full sample in the indices  $\text{H}\beta$  versus  $[\text{MgFe50}]'^2$  for the three aperture sizes, overplotted with these model predictions using  $[\alpha/\text{Fe}] = 0$ . The effects of abundance ratio are largely mitigated with this choice of indices (Thomas et al. 2003). The models have been adapted for zero-point differences between the Lick system and the Schiavon models using Table 1 of Schiavon (2007).

From this figure, it can be seen how the changing distribution of  $\text{H}\beta$  for the different apertures implies that the young stars in general are centrally concentrated. It can also be seen that the inferred metallicity becomes lower for larger apertures, due to the inclusion of the metal-poor outer regions. We also note that several objects have implied ages that are older than the canonical age of the Universe from Planck ( $13.798 \pm 0.037$  Gyr; Planck Collaboration, Ade et al. 2013). In Appendix D we show that these points are consistent with objects close to this upper age limit considering our observational uncertainties.

To derive the best-fitting SSP parameters for our data, we use the three indices that are measured across the full field for all galaxies, namely  $\text{H}\beta$ , Fe5015 and  $\text{Mg } b$ . We employ the straightforward yet robust technique of finding the SSP model that best predicts our observed line indices simultaneously. As per McDermid et al. (2006) and Kuntschner et al. (2010), we do this by means of chi-squared fitting: effectively computing the ‘distance’ from our measured indices to all the predicted values of those indices, and finding the

model with the minimum total distance. To reduce the effects of grid discretization, we oversample the original models using linear interpolation. Measurement errors on the indices are included as weighting factors in the sum. This particular technique has been used by several authors (Proctor et al. 2004; Thomas et al. 2010), although there are other methods for inverting the model grids to derive parameter values (e.g. Graves & Schiavon 2008).

We estimate the parameter errors by taking the range of points at which the difference in chi-square from the minimum is equal to unity, giving an estimate of the projected dispersion on each parameter. This method is computationally efficient, and we verified that it agrees well with a Monte-Carlo approach. All SSP values are given in Tables 1–3, which are available in full via the online journal, and via our project website: <http://purl.com/atlas3d>.

## 2.5 Mass-weighted parameters from spectral fitting

We use the publicly-available<sup>3</sup> penalized pixel fitting code (pPXF, Cappellari & Emsellem 2004) to fit a linear combination of SSP model spectra from the MIUSCAT model library (Vazdekis et al. 2012), spanning a regular grid of  $\log(\text{age})$  (0.1 - 14 Gyr) and metallicity ( $[Z/H] = -1.71 - 0.22$ , equivalent to  $Z = 0.0004 - 0.03$ ), giving 264 templates in total. Due to the degeneracy of inferring the star formation history from the integrated galaxy light, we make use of linear regularization (Press et al. 1992) to impose a smoothness constraint on the solution, using the ‘REGUL’ option of pPXF. The solution is defined by the weights applied to each model in a regularly-sampled grid of age and metallicity, and regularisation imposes a constraint in this parameter-space such that models with neighbouring age and metallicity must have weights that vary smoothly. The relative importance of this smoothness constraint compared to the actual fit quality is controlled via a simple ‘regularisation parameter’, which essentially controls how smooth the solution is. The SSP model spectra are provided to an initial birth cloud mass of 1 solar mass, and so the relative ‘zero-age’ mass-to-light ratios are included in the fitted weights. The distribution of weights from pPXF therefore defines the relative contributions in (zero-age) mass of the different populations: effectively the star formation history (SFH) for that galaxy.

For each galaxy, we fit the spectrum integrated within one effective radius as per the largest aperture considered for the SSP analysis. We derive solutions for a range of regularization parameter values, and select the value that gives a resulting best-fit solution with  $\Delta\chi^2 \sim \sqrt{2N}$ , where  $N$  is the number of pixels in the spectrum being fitted, and  $\Delta\chi^2$  is computed as the difference between the  $\chi^2$  of the current solution with that of the non-regularized case. This criterion is suggested by Press et al. (1992) as a robust indication of when the fit has become unacceptable. In this way, we are deriving the *smoothest* star formation history that still reproduces the SAURON spectrum. The actual star formation histories of galaxies are likely not smooth. Star formation is inherently stochastic when considered with sufficient temporal and spatial resolution, and specific events such as mergers

<sup>2</sup>  $[\text{MgFe50}]' = \frac{0.69 \times \text{Mg}b + \text{Fe5015}}{2}$  (Kuntschner et al. 2010)

<sup>3</sup> <http://purl.org/cappellari/idl>

**Table 1.** Measured Lick index measurements and SSP parameters measured within  $R_e/8$ .

Name	H $\beta$ [ $\text{\AA}$ ] $R_e/8$	Fe5015 [ $\text{\AA}$ ] $R_e/8$	Mg b [ $\text{\AA}$ ] $R_e/8$	Fe5270 [ $\text{\AA}$ ] $R_e/8$	Age <sub>SSP</sub> [Gyr] $R_e/8$	[Z/H] <sub>SSP</sub> $R_e/8$	[ $\alpha$ /Fe] <sub>SSP</sub> $R_e/8$	Quality
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IC0560	2.95 $\pm$ 0.10	4.17 $\pm$ 0.14	2.29 $\pm$ 0.12	2.20 $\pm$ 0.11	1.85 $\pm$ 0.21	-0.11 $\pm$ 0.05	0.09 $\pm$ 0.05	1
IC0598	2.88 $\pm$ 0.10	4.64 $\pm$ 0.14	2.46 $\pm$ 0.12	2.31 $\pm$ 0.12	1.70 $\pm$ 0.22	0.04 $\pm$ 0.05	0.05 $\pm$ 0.04	1
IC0676	3.51 $\pm$ 0.15	2.64 $\pm$ 0.14	1.75 $\pm$ 0.23	1.51 $\pm$ 0.11	2.02 $\pm$ 0.23	-0.64 $\pm$ 0.08	0.37 $\pm$ 0.11	2
IC0719	2.16 $\pm$ 0.10	5.14 $\pm$ 0.14	3.26 $\pm$ 0.12	2.75 $\pm$ 0.11	3.53 $\pm$ 0.68	0.03 $\pm$ 0.05	0.11 $\pm$ 0.04	1
IC0782	2.27 $\pm$ 0.10	4.98 $\pm$ 0.14	2.88 $\pm$ 0.12	- $\pm$ -	3.10 $\pm$ 0.36	-0.03 $\pm$ 0.06	0.06 $\pm$ 0.05	1
IC1024	4.00 $\pm$ 0.10	2.35 $\pm$ 0.14	1.52 $\pm$ 0.17	1.20 $\pm$ 0.11	1.59 $\pm$ 0.18	-0.75 $\pm$ 0.06	0.41 $\pm$ 0.09	2

Columns: (1) Galaxy name. (2)-(5) Lick index measurements within  $R_e/8$ , measured on the Lick/IDS system. Fe5270 is measured indirectly via the Fe5270<sub>S</sub> by applying the relation  $\text{Fe5270} = 1.28 \times \text{Fe5270}_S + 0.03$  (Kuntschner et al. 2006). Due to field coverage issues, this index cannot be measured for all objects. (6)-(8) Estimates of SSP-equivalent age, metallicity and abundance ratio within a circular aperture of radius  $R_e/8$  using the SSP models of Schiavon (2007) and using the three indices H $\beta$ , Fe5015, and Mg b. (9) Quality flag: 1 = Data are of good quality; 2 = Object with weak Fe5015 index in the  $R_e/8$  aperture; 3 = Strong emission line residuals; 4 = low signal-to-noise ratio. The full version of this table with all 260 galaxies is made available via the online journal, and via our project website <http://purl.com/atlas3d>.

**Table 2.** Measured Lick index measurements and SSP parameters measured within  $R_e/2$ .

Name	H $\beta$ [ $\text{\AA}$ ] $R_e/2$	Fe5015 [ $\text{\AA}$ ] $R_e/2$	Mg b [ $\text{\AA}$ ] $R_e/2$	Fe5270 [ $\text{\AA}$ ] $R_e/2$	Age <sub>SSP</sub> [Gyr] $R_e/2$	[Z/H] <sub>SSP</sub> $R_e/2$	[ $\alpha$ /Fe] <sub>SSP</sub> $R_e/2$	Quality
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IC0560	2.53 $\pm$ 0.10	3.86 $\pm$ 0.14	2.32 $\pm$ 0.12	2.15 $\pm$ 0.11	2.67 $\pm$ 0.31	-0.30 $\pm$ 0.05	0.14 $\pm$ 0.06	1
IC0598	2.43 $\pm$ 0.10	4.20 $\pm$ 0.14	2.47 $\pm$ 0.12	2.22 $\pm$ 0.11	2.84 $\pm$ 0.33	-0.22 $\pm$ 0.05	0.11 $\pm$ 0.05	1
IC0676	3.02 $\pm$ 0.10	3.25 $\pm$ 0.14	1.92 $\pm$ 0.12	1.75 $\pm$ 0.11	2.20 $\pm$ 0.25	-0.48 $\pm$ 0.06	0.22 $\pm$ 0.07	2
IC0719	2.43 $\pm$ 0.10	4.16 $\pm$ 0.14	2.69 $\pm$ 0.12	2.46 $\pm$ 0.11	2.84 $\pm$ 0.33	-0.18 $\pm$ 0.04	0.18 $\pm$ 0.05	1
IC0782	2.08 $\pm$ 0.10	4.51 $\pm$ 0.20	2.58 $\pm$ 0.12	- $\pm$ -	4.77 $\pm$ 0.85	-0.25 $\pm$ 0.05	0.03 $\pm$ 0.06	1
IC1024	3.59 $\pm$ 0.15	2.25 $\pm$ 0.14	1.66 $\pm$ 0.20	1.30 $\pm$ 0.11	2.06 $\pm$ 0.24	-0.78 $\pm$ 0.07	0.50 $\pm$ 0.05	2

Columns are as per Table 1 but for an  $R_e/2$  aperture size. The full version of this table with all 260 galaxies is made available via the online journal, and via our project website <http://purl.com/atlas3d>.

**Table 3.** Measured Lick index measurements and SSP parameters measured within  $R_e$ .

Name	H $\beta$ [ $\text{\AA}$ ] $R_e$	Fe5015 [ $\text{\AA}$ ] $R_e$	Mg b [ $\text{\AA}$ ] $R_e$	Fe5270 [ $\text{\AA}$ ] $R_e$	Age <sub>SSP</sub> [Gyr] $R_e$	[Z/H] <sub>SSP</sub> $R_e$	[ $\alpha$ /Fe] <sub>SSP</sub> $R_e$	Quality
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
IC0560	2.37 $\pm$ 0.10	3.75 $\pm$ 0.16	2.31 $\pm$ 0.12	2.10 $\pm$ 0.10	3.53 $\pm$ 0.58	-0.38 $\pm$ 0.06	0.19 $\pm$ 0.07	1
IC0598	2.30 $\pm$ 0.10	4.03 $\pm$ 0.14	2.41 $\pm$ 0.12	2.16 $\pm$ 0.11	3.38 $\pm$ 0.57	-0.30 $\pm$ 0.05	0.12 $\pm$ 0.05	1
IC0676	2.66 $\pm$ 0.10	3.37 $\pm$ 0.14	1.96 $\pm$ 0.13	1.79 $\pm$ 0.11	2.84 $\pm$ 0.38	-0.51 $\pm$ 0.06	0.21 $\pm$ 0.07	2
IC0719	2.35 $\pm$ 0.10	3.84 $\pm$ 0.14	2.56 $\pm$ 0.12	2.33 $\pm$ 0.10	3.38 $\pm$ 0.60	-0.29 $\pm$ 0.05	0.22 $\pm$ 0.06	1
IC0782	2.03 $\pm$ 0.10	4.47 $\pm$ 0.27	2.50 $\pm$ 0.12	- $\pm$ -	5.31 $\pm$ 0.85	-0.30 $\pm$ 0.06	0.01 $\pm$ 0.07	1
IC1024	3.33 $\pm$ 0.18	2.47 $\pm$ 0.14	1.71 $\pm$ 0.16	1.39 $\pm$ 0.12	2.29 $\pm$ 0.42	-0.73 $\pm$ 0.06	0.43 $\pm$ 0.07	2

Columns are as per Table 1 but for an  $R_e$  aperture size. The full version of this table with all 260 galaxies is made available via the online journal, and via our project website <http://purl.com/atlas3d>.

and interactions may trigger or truncate the star formation process on short timescales. In fact, we have also shown that the colors and H $\beta$  indices of the bluest and most H2-rich galaxies can be reproduced with a small amount of current or recent star formation activity superposed on an old stellar population (Young et al. 2014). However, inferring such events from the integrated light is highly degenerate for typical observational data, and the smoothness constraint offers a robust and objective result that can be used to study systematic trends in the entire sample. We also note that non-smooth solutions are in no way excluded by our approach,

but if a smooth solution can give a fit of similar quality (as determined via  $\Delta\chi^2$ ), then the smooth solution will be adopted.

The mass-weighted age and metallicity are given as:

$$\log(\text{Age}_{\text{SFH}}) = \frac{\sum w_i \log(t_{\text{SSP},i})}{\sum w_i} \quad (1)$$

and

$$[\text{Z}/\text{H}]_{\text{SFH}} = \frac{\sum w_i [\text{Z}/\text{H}]_{\text{SSP},i}}{\sum w_i}, \quad (2)$$

**Table 4.** Mass-weighted stellar population properties measured within  $R_e$ . The full version of this table with all 260 galaxies is made available via the online journal, and via our project website <http://purl.com/atlas3d>.

Name	Ages <sub>SFH</sub> [Gyr] $R_e$	[Z/H] <sub>SFH</sub> $R_e$	$t_{50}$ [Gyr] $R_e$
(1)	(2)	(3)	(4)
IC0560	$6.73 \pm 1.07$	$-0.46 \pm 0.06$	$6.55 \pm 1.07$
IC0598	$5.37 \pm 0.99$	$-0.40 \pm 0.05$	$8.60 \pm 0.99$
IC0676	$6.00 \pm 1.92$	$-0.55 \pm 0.18$	$7.22 \pm 1.92$
IC0719	$9.21 \pm 0.86$	$-0.25 \pm 0.08$	$4.23 \pm 0.86$
IC0782	$6.16 \pm 0.83$	$-0.19 \pm 0.11$	$7.09 \pm 0.83$
IC1024	$4.39 \pm 0.63$	$-0.77 \pm 0.15$	$8.65 \pm 0.63$

Columns: (1) Galaxy name. (2) Mass-weighted age. (3) Mass-weighted metallicity. (4) Half-mass formation time. These properties were derived from a regularised spectral fit to a spectrum integrated within  $R_e$ . Errors were computed via Monte-Carlo realisations *without* regularisation constraints. See text for details.

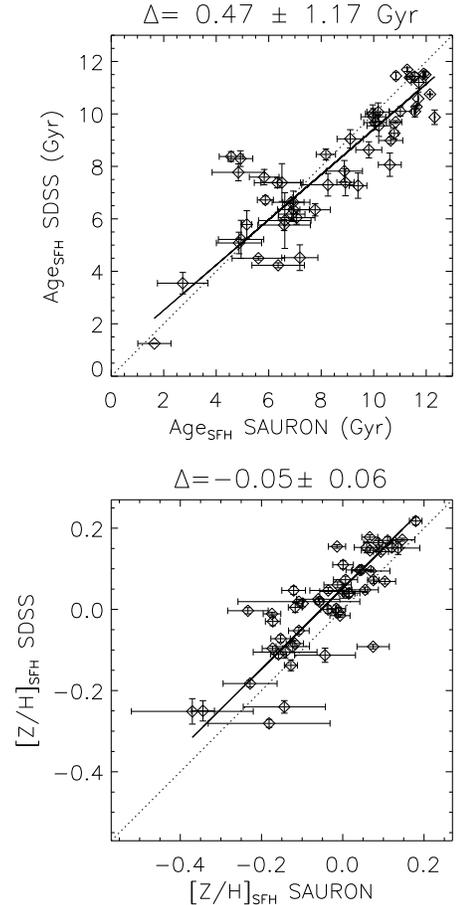
respectively, where  $w_i$  denotes the weight given by pPXF to the  $i$ 'th template, which has age  $t_{\text{SSP},i}$  and metallicity  $[Z/H]_{\text{SSP},i}$ . The grid of model ages is sampled logarithmically, hence the logarithmic form of equation 1. Errors on these quantities were computed via Monte-Carlo simulations *without* regularisation.

The SSP spectral templates we use have fixed (solar) abundance ratios, and so we cannot derive a mass-weighted  $[\alpha/\text{Fe}]$ . We can, however, derive a measure of the star formation timescale directly from the SFH itself. The measure we use is the time,  $t_{50}$ , taken to form 50% of the current day stellar mass within one  $R_e$ , and is derived from the cumulative function of the SFH (see section 5.1). All mass-weighted quantities are given in Table 4.

## 2.6 Comparison of SFHs with SSPs and SDSS

An obvious concern in deriving the SFH from SAURON spectra is the short wavelength range available. To test this, we compare the SFHs derived from SAURON with those derived from fitting SDSS spectroscopy (DR8, York et al. 2000) of the galaxies in common, using the entire SDSS wavelength range to constrain the SFH - a more than ten-fold increase in the wavelength coverage of the spectrum. For this test, we adapt the SAURON aperture to match the  $3''$  diameter SDSS fiber, thus minimizing the effects of population gradients. Figure 3 shows a comparison of the SAURON and SDSS mass-weighted average ages and metallicities derived by applying the above procedure. The mass-weighted parameters compare very well, showing no significant systematic biases across the range of age and metallicity considered despite the dramatically different spectral ranges used. Age shows a slight trend to older ages derived by SAURON, but the oldest case is still consistent with SDSS within the  $1\sigma$  scatter. Metallicity is also very consistent, albeit with a small average offset, but again within the  $1\sigma$  scatter. Overall, the agreement is remarkably good, and suggests that the short wavelength range of SAURON does not significantly bias our derived SFHs.

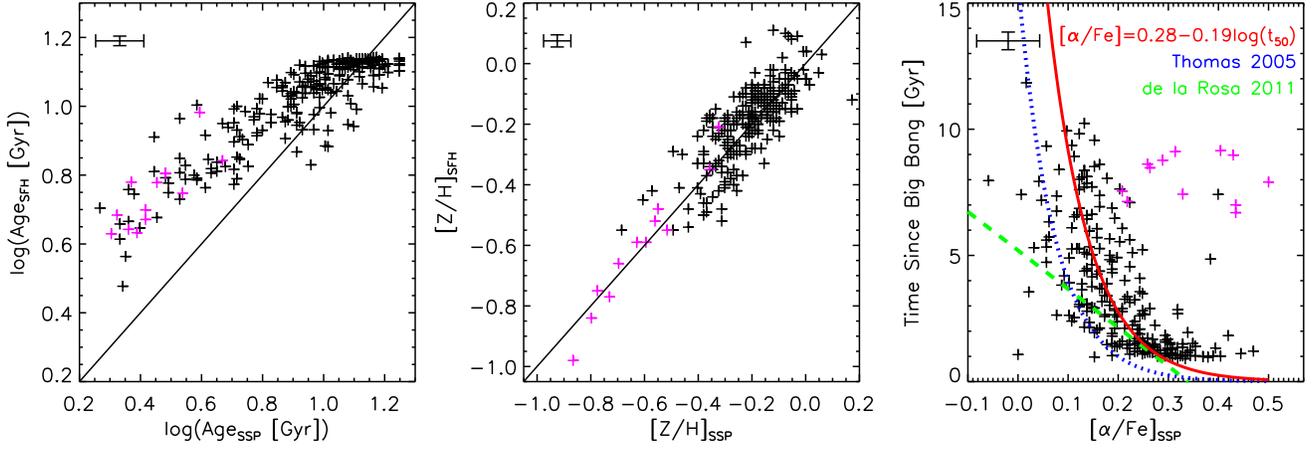
We compare our mass-weighted and SSP-equivalent parameters in Figure 4, this time using a  $1R_e$  aperture. The



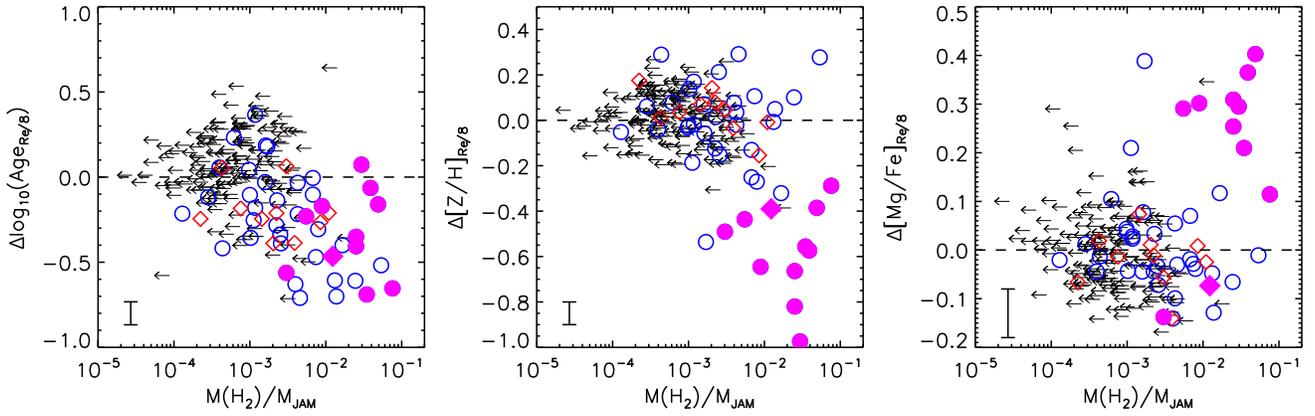
**Figure 3.** Comparison of mass-weighted age (top) and metallicity (bottom) derived from fitting SAURON and SDSS spectra within the same spatial aperture. The mean difference (SAURON-SDSS) and the standard deviation are given in the plot titles. The dotted lines indicate the 1:1 identity line, with a thick line showing the best linear fit. Errors were computed via Monte-Carlo simulations *without* regularisation.

mass-weighted ages are systematically older for all but the oldest ages, where we are biased by the lower maximum age of our spectral template library. Including templates up to 17 Gyr reduces the ‘bunching’ of points at oldest ages, but for clarity of interpreting our SFHs, we use a maximum age of 14 Gyr in our template library, consistent with current cosmological models (Planck Collaboration, Ade et al. 2013). The older mass-weighted ages derived from spectral fitting are a direct result of the well-known bias of SSP-equivalent ages by small mass fractions of young populations, and our findings agree well with the exploration of this issue by Serra & Trager (2007). Metallicities show good agreement, also as expected following the findings of Serra & Trager (2007), who showed that SSP-equivalent metallicities are close to the mass-weighted values, being less biased by the presence of young stars. We also note that there is no obvious bias caused by using models limited to solar abundance ratios.

Finally, we find a clear trend between the SSP-equivalent alpha enhancement and  $t_{50}$ . We over-plot the linear relation given by de La Rosa et al. (2011), who conducted a similar analysis on SDSS spectra using the STARLIGHT



**Figure 4.** Comparison of SSP-equivalent parameters (x-axis) and mass-weighted parameters derived using spectral fitting (y-axis) within the same aperture of one effective radius  $R_e$ . Magenta points highlight the objects with low Fe as discussed in Section 2.3. Identity lines are given for age and metallicity. The right panel shows  $[\alpha/\text{Fe}]$  versus  $t_{50}$ , the time taken to form half the current-day mass. The relation  $t_{50} = -15.3[\alpha/\text{Fe}] + 5.2$  given by eqn. 2 in de La Rosa et al. (2011) is over plotted on our data with a green dashed line. The blue dotted line shows the canonical conversion of abundance ratio to star formation timescale from equation (4) of Thomas et al. (2005) based on chemical evolution models:  $[\alpha/\text{Fe}] \sim \frac{1}{5} - \frac{1}{6} \log \Delta t$ . The red solid line and text indicate our best-fitting function of this same form, giving an empirical relation between  $[\alpha/\text{Fe}]$  and  $t_{50}$  for the ATLAS<sup>3D</sup> sample.



**Figure 5.** Residuals from the SSP-velocity dispersion trends plotted as a function of molecular gas fraction. Median errors are indicated in the bottom left corner of each panel, with  $\text{H}_2$  mass fraction errors within the plotting symbol. Arrows indicate  $3\sigma$  upper-limits for the  $\text{H}_2$  mass. Diamonds denote CO detections in the Virgo cluster. Open circles show detections of non-Virgo objects. Filled magenta symbols correspond to the ‘low-Fe’ objects discussed in Section 2.3. Their weak Fe5015 indices result in low metallicity and high  $[\alpha/\text{Fe}]$ . We speculate that these objects have had recent gas accretions, giving rise to young, metal poor stars that have not yet been enriched with ejecta from Type 1a supernovae. Only one of these low-Fe objects is found in the Virgo cluster (filled magenta diamond).

code of Cid Fernandes et al. (2005). We also show the canonical relation between abundance ratio and star formation timescale from Thomas et al. (2005), equation (4):  $[\alpha/\text{Fe}] \sim \frac{1}{5} - \frac{1}{6} \log \Delta t$ , where  $\Delta t$  is the FWHM of the assumed Gaussian SFH. This relation reproduces the non-linearity of our measured timescale against  $[\alpha/\text{Fe}]$ . We also note that the limited age resolution of the SSP models at the oldest ages creates a ‘floor’ in the  $t_{50}$  values at the shortest timescales. This important region, corresponding to the earliest epochs of massive galaxy formation, is where abundance ratios can yield higher fidelity than our SFH. We therefore use our empirical half-mass formation time to constrain a relation of the same form as that of Thomas et al. (2005) to

relate  $[\alpha/\text{Fe}]$  to  $t_{50}$ . The resulting expression is indicated by the red line in Fig. 4, and is given by:

$$[\alpha/\text{Fe}] = 0.28 - 0.19 \log(t_{50}) \quad (3)$$

The objects plotted with magenta symbols in Figure 4 (the low-Fe objects) are mostly outliers to the  $[\alpha/\text{Fe}]$ - $t_{50}$  trend, having *large* values of  $t_{50}$  - i.e. long formation timescales, contrary to what is expected from their high  $[\alpha/\text{Fe}]$  values. The relatively weak iron indices result in low metallicities and subsequently *high*  $[\alpha/\text{Fe}]$ , but since they also have low velocity dispersions, they are strong outliers in the scaling relations shown in Figure 8. Figure 5 shows the residuals of the SSP parameter - velocity dispersion relations for an  $R_e/8$  aperture, plotted against molecular gas

mass fraction, with the low-Fe objects marked with filled symbols. This shows that, as well as having low metallicity, the low-Fe objects also have high molecular gas fractions and are among the youngest in the sample. This suggests their apparently high SSP-equivalent  $[\alpha/\text{Fe}]$  may indicate the presence of stars that have not yet been appreciably enriched by Type 1a supernovae - the main contributor of iron to the gas involved in star-formation. Their young age causes them to outshine the older, less  $\alpha$ -enhanced stars, and dominate the SSP abundance measurement. They do not show any preferred kinematic misalignment (Davis et al. 2013), and are not Virgo cluster members except for one, NGC 4694, which is a low-mass object in the outskirts of Virgo. We speculate that these low-Fe galaxies have recently accreted a new supply of cold, low-metallicity gas, which is currently in the process of forming stars. These twelve low-Fe objects are shown as magenta symbols in subsequent plots unless stated otherwise, and are indicated in Tables 1-3 via the ‘quality’ flag.

### 3 POPULATION PARAMETERS ON THE MASS PLANE

Galaxy stellar populations have been found to correlate better with stellar surface density (Kauffmann et al. 2003; Franx et al. 2008) or with velocity dispersion (Graves & Faber 2010) than with mass or galaxy size. The same was found to hold for the dynamical M/L, which mainly traces age variations (Cappellari et al. 2006). In Cappellari et al. (2013b) we used accurate masses and velocity dispersions to show that indeed the velocity dispersion is the best simple tracer of multiple galaxy properties. We showed that velocity dispersion is such a good predictor of galaxy properties because it traces the bulge fraction, which appears to be the main driver of the M/L, H $\beta$  and colour, as well as the molecular gas fraction in ETGs. Here we focus on trends in the stellar population parameters.

Figure 6 and 7 show the distribution of population parameters as a function of mass and size (which we hereafter refer to as the Mass Plane) for our sample, using the SSP-equivalent and mass-weighted parameters respectively. The mass  $M_{\text{JAM}}$  is derived applying Equation 3 of Cappellari et al. (2013b), using the mass-to-light ratio  $(M/L)_{\text{JAM}}$  and  $r$ -band luminosity given in Table 1 of Cappellari et al. (2013a), columns 6 and 15 respectively. We use the robust size parameter  $R_e^{\text{maj}}$ , corresponding to the major axis of the half-light isophote (Hopkins et al. 2010) as derived in Cappellari et al. (2013a), and taken from Table 1 there.

The points in Figs. 6 and 7 are coloured according to the population parameters as indicated. The distributions have been adaptively averaged by applying an implementation<sup>4</sup> of the LOESS algorithm (Cleveland & Devlin 1988), using a local linear approximation, and a regularization factor  $f = 0.6$ . The same approach is taken in Cappellari et al. (2013b) and is found to give a good approximation to the underlying intrinsic distribution, mimicking the average behaviour that would be derived from much larger samples (e.g. Gallazzi et al. 2006). We note that, since we are adaptively averaging

the data points, the range of values is reduced in accordance with the reduced effects of errors and intrinsic scatter.

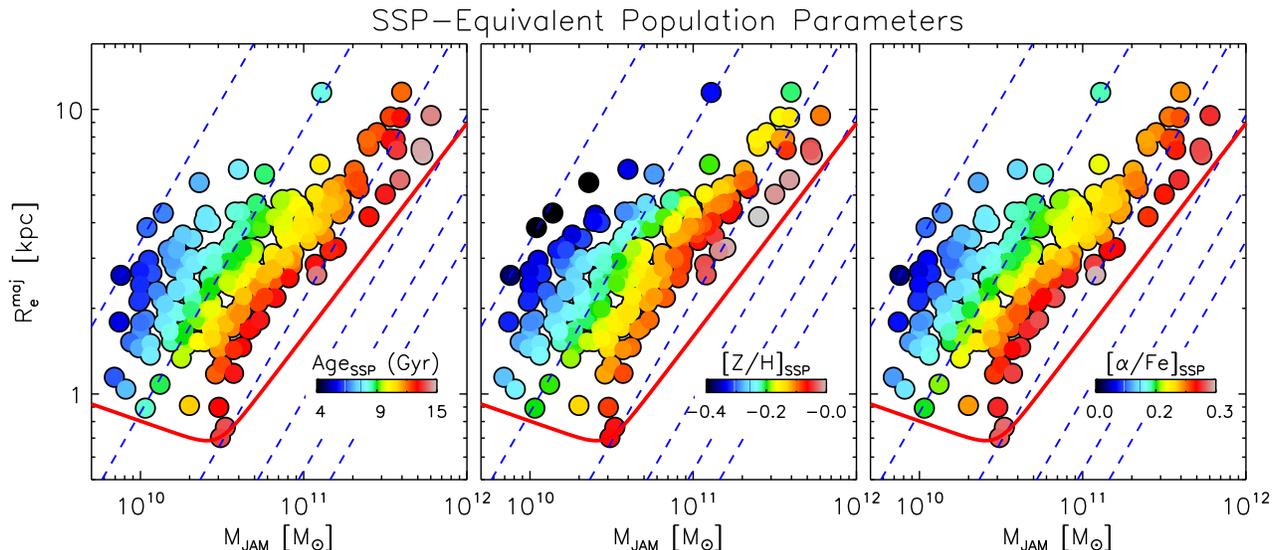
As in Cappellari et al. (2013b), we over-plot lines of constant velocity dispersion, as implied by the virial mass estimator  $M_{\text{JAM}} = 5R_e^{\text{maj}}\sigma^2/G$ . We find that the SSP-equivalent population parameters follow trends of nearly constant velocity dispersion on the Mass Plane, as indicated by the tilt of the iso-colour regions. The mass-weighted age and formation timescale follow even more tightly the lines of constant dispersion, which reflects the lower scatter in the mass-weighted properties as a result of lower formal uncertainties, and less susceptibility to small fractions of young stars being present. The mass-weighted metallicity is also closely linked to the velocity dispersion, though the iso-colour regions show an appreciable dependence on  $M_{\text{JAM}}$  also (i.e. steeper iso-colour regions).

These figures show that, for fixed mass, more compact galaxies are on average older, more metal rich, have higher abundance ratios and shorter formation timescales than larger galaxies. The trends with SSP age and metallicity agree with previous results from averaged properties of stacked SDSS spectra (van der Wel et al. 2009; Shankar et al. 2010). We show here that these trends hold for measurements of individual galaxies. Moreover, the trend with abundance ratio, mass-weighted properties, and formation timescale measured directly from the star-formation history can only be obtained from a spectroscopic analysis, and are reported here for the first time.

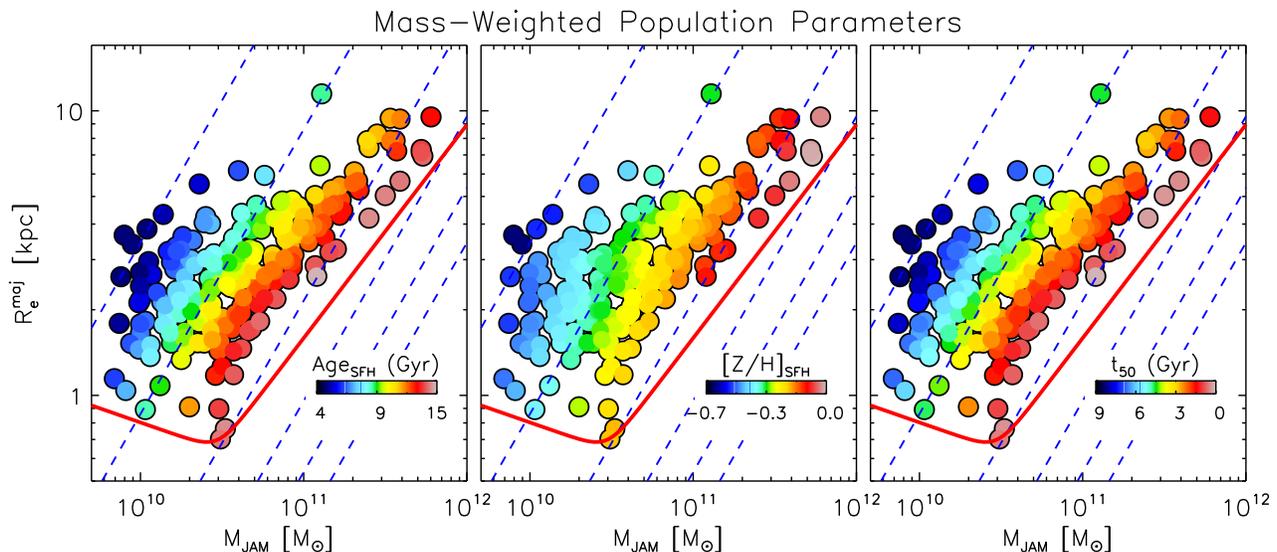
Graves et al. (2009b) found that  $R_e$  does not correlate with either population parameters or their residuals from the population- $\sigma$  relations. This can be understood from the mass-size plane. At each  $R_e^{\text{maj}}$ , there is a similar range of population parameters, thus showing no particular correlation. Likewise, as  $\sigma$  traces the parameters closely, the residuals of the population- $\sigma$  trends show no significant residual correlation with  $R_e^{\text{maj}}$ . Indeed, Figures 6 and 7 show that velocity dispersion is an excellent predictor of all three stellar population parameters, with the smoothed population parameters varying essentially parallel to the lines of constant (virial-predicted) velocity dispersion.

We advocate that the Mass Plane is a novel and useful tool for understanding galaxy properties. Comparing our Figures 6 and 7 to Figures 3–8 of Cappellari et al. (2013b) reveals the close connections between the stellar population parameters and other important quantities, in particular colour, molecular gas mass fraction, and bulge fraction, such that, at a given mass, smaller galaxies are redder in colour, have lower fractions of molecular gas, and have a second moment  $V_{\text{rms}} \equiv \sqrt{V^2 + \sigma^2}$  dominated by a dynamically hot central ‘bulge’ component. There are no particular trends with intrinsic shape other than the fact that the population of round objects above  $M_{\text{JAM}} \gtrsim 2 \times 10^{11} M_{\odot}$  are typically old, metal rich and  $\alpha$ -enhanced. Nor is there a close connection to the specific angular momentum,  $\lambda_{\text{R}}$ , which supports the findings of Naab et al. (2014) where  $\lambda_{\text{R}}$  is shown to trace a galaxy’s evolution, but that there are many evolutionary paths to given current-day value. We explore the connection between the kinematic classes and stellar populations in Section 4.5.

<sup>4</sup> available from <http://purl.org/cappellari/idl>



**Figure 6.** SSP-equivalent population parameters measured within one effective radius, plotted on the plane of effective radius  $R_e^{\text{maj}}$  versus dynamical mass,  $M_{\text{JAM}}$ . Colours indicate the population parameters as indicated in each plot after spatially averaging with the LOESS technique described in the text. For a fixed range in  $M_{\text{JAM}}$ , the smallest galaxies are generally the oldest, most metal rich and alpha-enhanced. Dashed lines show lines of constant velocity dispersion: 50, 100, 200, 300, 400 and 500 km s<sup>-1</sup> from left to right, as implied by the virial mass estimator  $M_{\text{JAM}} = 5R_e^{\text{maj}}\sigma^2/G$ . The red curve shows the ‘zone of exclusion’ (ZOE) defined in Cappellari et al. (2013b).



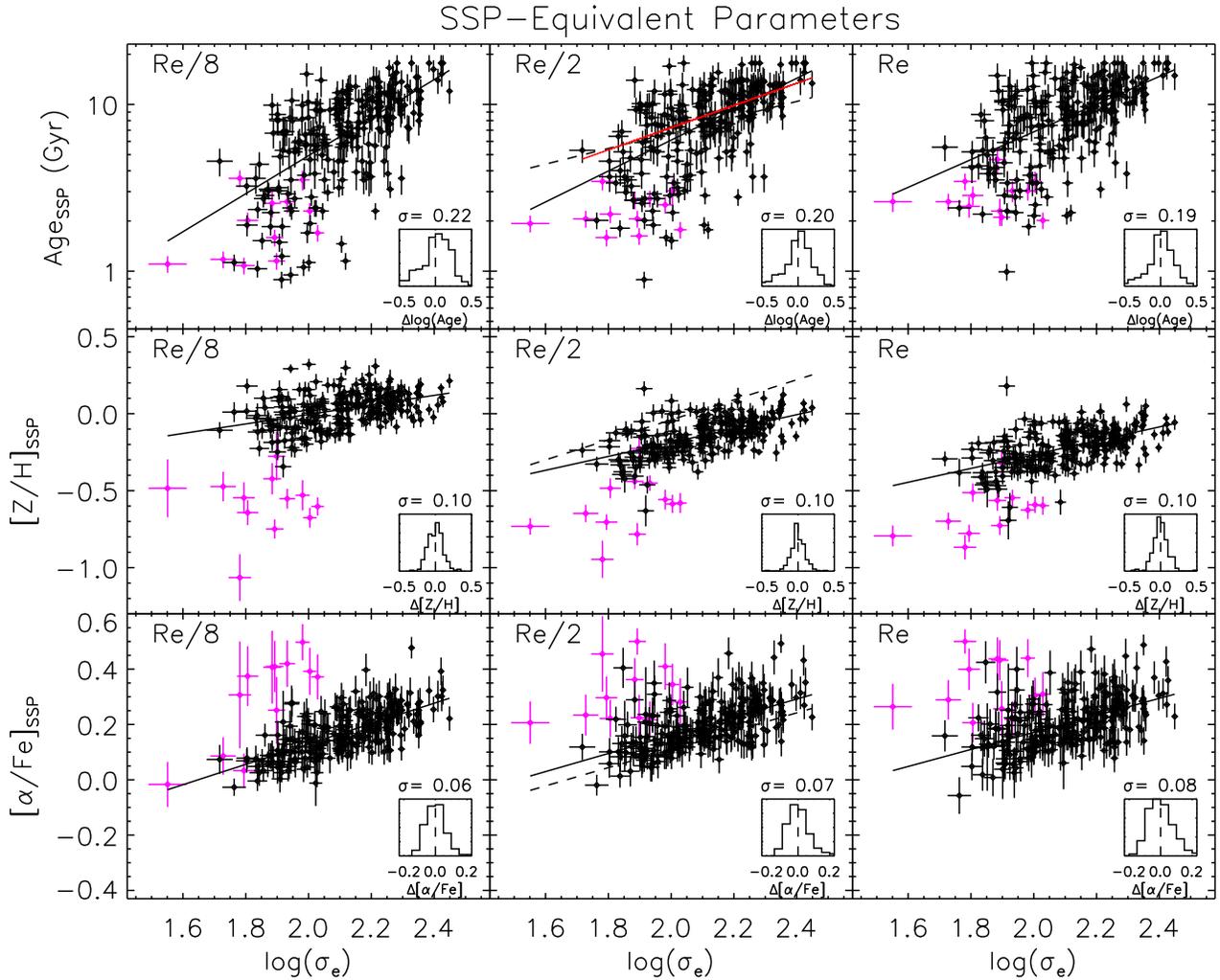
**Figure 7.** Same as Figure 6, but using the mass-weighted age and metallicity, and  $t_{50}$  - the time taken to form half the stars inside  $1R_e$ .

#### 4 POPULATION TRENDS WITH VELOCITY DISPERSION AND MASS

In this section we present a more traditional two-dimensional analysis of the stellar population parameters, examining their scaling properties with velocity dispersion and mass, for SSP-equivalent and mass-weighted properties. This allows a more quantitative study of these relations and their scatter, as well as providing context for our findings on the Mass Plane, allowing direct comparison to previous authors. All linear fit parameters are given in Table 5.

##### 4.1 SSP-equivalent parameters versus velocity dispersion

Figure 8 shows the resulting SSP-equivalent parameters for our three aperture radii, plotted against the effective velocity dispersion,  $\log(\sigma_e)$ , given in column 2 of Table 1 in Cappellari et al. (2013a). As shown in the previous section, the population parameters correlate tightly with velocity dispersion, confirming the trends from previous authors. For comparison, in Figure 8 we overplot with dashed lines the relations from Thomas et al. (2010) based on analysis of 3,360 galax-



**Figure 8.** SSP parameters for the three apertures as a function of  $\log(\sigma_e)$ . Magenta points are the galaxies exhibiting low metallicity as described in Section 2.3. Dashed lines on the  $R_e/2$  aperture plots show the relations from Thomas et al. (2010), which have a comparable aperture size (see text for details). The red line shows the  $R_e/2$  age trend when fitting only galaxies older than 2.5 Gyr, as done by Thomas et al. (2010). Solid lines are robust linear fits. Magenta symbols indicate the low-Fe objects highlighted in section 2.5. Solid lines are robust linear fits to the black points. Parameters of these fits are given in Table 5. Inset histograms show the residuals about the global best fit.

ies observed with a single fibre by the Sloan Digital Sky Survey, selected as part of the ‘MOSES’ project. We make this comparison on the  $R_e/2$  aperture, which should be the most comparable to the effective aperture of that study.

The Thomas et al. (2010) lines show general agreement, but there are some important differences. In particular, our slope with age is significantly steeper, which is trivially explained by the fact that Thomas et al. (2010) exclude objects younger than 2.5 Gyr (at  $z=0.05-0.06$ ) from their fit. We indicate with a red line our fit applying the same selection, and the agreement is excellent. There is also a significant difference in slopes for the metallicity relation. Low velocity dispersion objects are similar, but the Thomas et al. (2010) best-fit linear relation over-estimates the high velocity dispersion objects by 0.3 dex – several times larger than the total scatter. In Appendix B we show that the difference in slope is consistent with the lack of any aperture correction applied to the MOSES data to account for the fixed 3'' diam-

eter aperture used. This steepens the relation by including more of the outer regions in lower-mass (smaller) objects relative to the high-mass (larger) objects, which are dominated by the high metallicity inner regions. Even considering the offset in slope, the MOSES best-fit linear relation is still  $\sim 0.2$  dex higher in  $[Z/H]$ .  $[\alpha/Fe]$  also shows a small, less significant, offset such that our values are  $\sim 0.05$  dex higher than from MOSES. We have verified that these offsets are not caused by differences in the models of Schiavon (2007) and Thomas et al. (2003, 2004), which, when applied consistently to our line-strengths, show average offsets smaller than the  $1\sigma$  scatter, and which would anyway exacerbate the offsets shown in Figure 8. The differences between our findings and Thomas et al. (2010) may arise from the different combination of indices used by the two studies, but it is beyond the scope of this paper to explore this further.

Contrary to Thomas et al. (2010), we do not find strong evidence for bimodality in our distribution of age and veloc-

ity dispersion. Our sample size may be too small to constrain this, though we note that splitting our sample using their age threshold of 2.5 Gyr and our  $R_e/2$  values yields a similar fraction of ‘young’ galaxies (23 out of 260, or 8.8%, for our sample; 10.15% in Thomas et al. 2010).

From Figure 8, the effect of different aperture sizes can be seen. Most notably, as we ‘zoom out’ from the  $R_e/8$  aperture to  $R_e$ , the age- $\sigma$  relation becomes systematically more shallow, showing again that young stars, when present, are preferentially found in the central regions of the low velocity dispersion galaxies. The  $[Z/H]$ - $\sigma$  relation becomes steeper and the average metallicity decreases from around solar for  $R_e/8$  to  $-0.2$  within  $R_e$ , showing that early-type galaxies are only metal rich in their central regions, and are actually rather metal-poor objects when considered on the scales of one effective radius. Enhancement of alpha elements does not show much change considering larger apertures, although we remark that the scatter of the  $[\alpha/Fe]$ - $\sigma$  relation increases by 15% as we go from  $R_e/8$  to  $R_e$ , unlike for age and metallicity, which become 15% and 10% tighter, respectively.

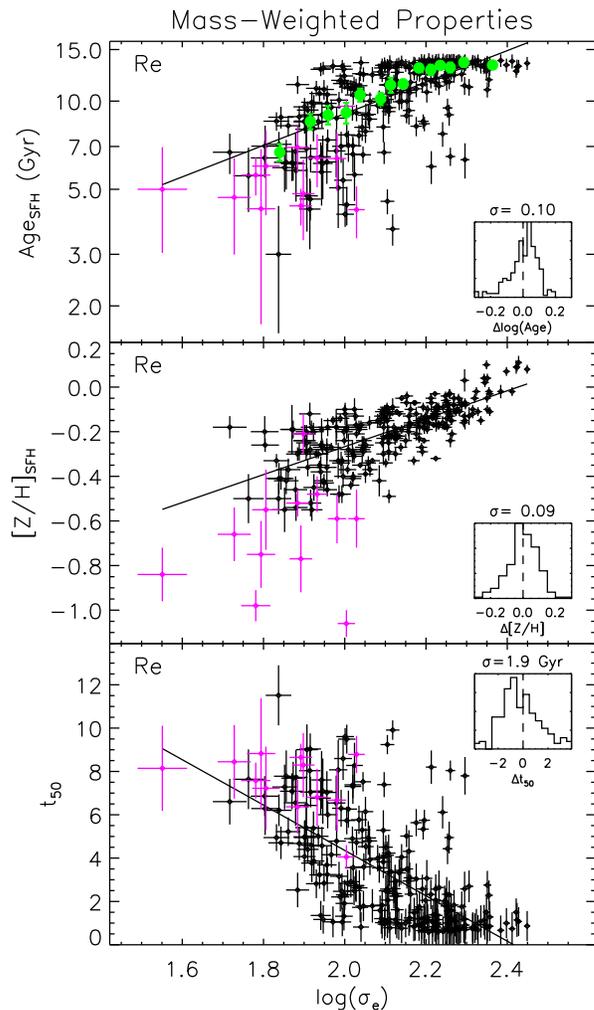
#### 4.2 Mass-weighted parameters versus velocity dispersion

Figure 9 presents the mass-weighted age and total metallicity derived from the spectral fitting technique as described in Section 2.5. The spectrum fitted is integrated within an effective radius, making these values directly comparable to the  $R_e$  aperture of the previous sections. The general trends are similar: age and metallicity increase with velocity dispersion. However, the correlation with mass-weighted age is significantly tighter than with SSP age, showing a factor 2.6 less scatter (though with a somewhat smaller range in age).

We also show in Figure 9 the median-binned mass-weighted ages to illustrate better the general trend, which suggests a possible break in the slope of the relation at velocity dispersions higher than  $\approx 160 \text{ km s}^{-1}$ . We caution, however, that this is largely a result of limiting the maximum age of the stellar population models to 14 Gyr. Including the maximum-aged MIUSCAT templates (17.8 Gyr) in the star formation history removes any significant deviation from a linear relation of mass-weighted age with  $\log(\sigma_e)$ . In addition, by fitting a linear relation to points below the apparent break, the extrapolated age and scatter at the highest velocity dispersion is  $18 \pm 4 \text{ Gyr}$  - only mildly inconsistent with the fiducial age of the Universe. This suggests that a linear relation is indeed an appropriate choice. We return to this issue in Section 4.4.

The metallicity trend is remarkably similar considering the vastly different approaches and models used. These results give direct empirical confirmation of the fact that early-type galaxies are overwhelmingly composed of old stars, but in general have a small (but non-negligible) mass-fraction of younger stars that contribute disproportionately to the integrated light of the galaxy - a fact inferred from early SSP analyses (González 1993; Trager et al. 2000) and shown unambiguously with GALEX UV colours (Kaviraj et al. 2007). We consider this more quantitatively in Section 5.1.

The almost negligible change in the metallicity trend also confirms the insensitivity of SSP-equivalent metallicity to mixed populations as discussed above. The main differences are a slight reduction in scatter and a steepening of



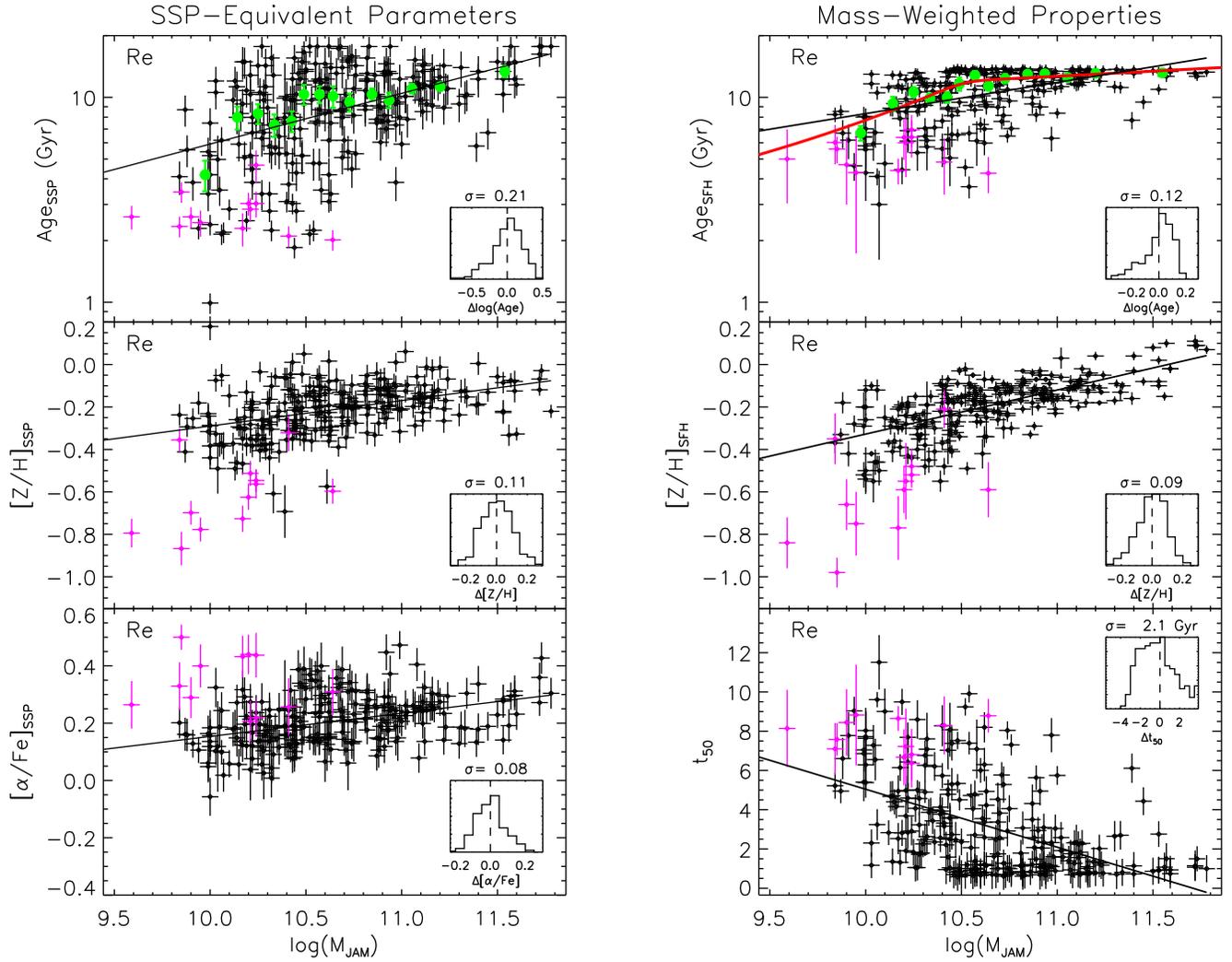
**Figure 9.** Mass-weighted age, metallicity and half-mass formation time of the ATLAS<sup>3D</sup> sample plotted against velocity dispersion. Magenta points are the Fe-poor objects as per Fig. 8. Insert histograms show the residuals about the global best fit. Solid lines are robust linear fits to the black points. Parameters of these fits are given in Table 5.

the relation, likely due to reduced uncertainties by fitting the whole spectrum, and a reduced bias in the contribution of young stars responsible for boosting the metallicity in the SSP relations.

#### 4.3 SSP parameters versus dynamical mass

Figure 10 (*left*) presents the SSP population trends as measured against  $M_{\text{JAM}}$  for the  $1R_e$  aperture (the smaller apertures follow similar relative differences as those mentioned in the previous section). We find that the general trends of population parameters are consistent with the picture derived from velocity dispersion alone, such that lower mass galaxies tend to show younger ages, lower metallicities, and lower abundance ratios.

The relations with mass are generally more shallow compared to the trends with  $\sigma_e$ , especially in age and metallicity, owing to two key differences. Firstly, galaxies with the



**Figure 10.** *Left:* SSP parameters within one effective radius as a function of dynamical mass,  $M_{\text{JAM}}$ . *Right:* Mass-weighted age, metallicity, and half-mass formation time, as a function of  $M_{\text{JAM}}$ . Magenta points indicate the Fe-poor objects as in previous plots. Solid lines are robust linear fits to the black points. Parameters of these fits are given in Table 5. Inset histograms show the residuals about the global best fit. Green points indicate the median values within bins of 20 neighbouring points, with error bars indicating the ‘error’ on the median, computed as  $\sigma/\sqrt{N}$ , where  $\sigma$  is the dispersion within the bin, and  $N = 20$ . The red line indicates a double power law fit to the binned values, as described in Section 4.4.

lowest masses span almost the full range of ages and metallicities, whereas at the lowest velocity dispersion, galaxies are generally the youngest and most metal poor of the sample. Secondly, the high mass regime where galaxies span a narrow range of age and metallicity ( $M_{\text{JAM}} > 10^{11} M_{\odot}$ ) accounts for about one third of the mass range in the sample. For velocity dispersion, this low-scatter region ( $\log(\sigma) > 2.3$ ) is less than a quarter of the sample range. When we move to the mass-based relations, the velocity dispersion distributions are essentially ‘stretched out’ among the older objects, reducing the slope of the relations. Both effects are a direct consequence of the distribution of galaxies on the Mass Plane.

Previous studies have found tighter relations of stellar population parameters with velocity dispersion than with mass (Graves & Faber 2010; Wake et al. 2012), as we also qualitatively highlighted on the Mass Plane. Inset his-

tograms in Figures 8 through 10 show the dispersions around our linear fits and quantify their standard deviations (see Table 5 for the fit parameters and standard deviations). We likewise find that the SSP residuals are somewhat tighter using  $\sigma_e$  than when using mass, although we note that the formal difference is small (0.01 dex in age and 0.02 in  $[Z/H]$ ). We note that the improvement of  $\sigma_e$  over mass as a linear indicator of population parameters is significantly more apparent on the Mass Plane directly than by comparing the residuals of its edge-on projections, where intrinsic scatter makes the interpretation less clear.

#### 4.4 Mass-weighted parameters versus dynamical mass

Figure 10 (*right*) presents the mass-weighted age, metallicity, and half-mass formation time  $t_{50}$ , as a function of  $M_{\text{JAM}}$ .

We see similar trends as before, but there is now a rather clear change in the behaviour of age as a function of mass. Rather than being approximated as a linear increase of age with mass, albeit with a large scatter at lower masses, the trend of mass-weighted age and mass shows a distinct non-linearity. This is made more apparent by considering the median value within mass bins, as indicated by the green points in Figure 10. The binned mass-weighted ages show a distinct departure from the linear fit, whereas the binned SSP values remain reasonably approximated by the best fit line.

To characterise this bimodal trend in more detail, we fit a double power law on the median-binned mass-weighted values of the form:

$$\log(\text{Age}) = 2^{(\beta-\gamma)/\alpha} \log(\text{Age}_b) \left(\frac{M}{M_b}\right)^\gamma \left[1 + \left(\frac{M_b}{M}\right)^\alpha\right]^{\frac{\gamma-\beta}{\alpha}} \quad (4)$$

based on the ‘Nuker’ law for surface brightness profile fitting (Lauer et al. 1995). The best fit curve has a low-mass slope  $\gamma = -0.18$ , a high-mass slope of  $\beta = -0.016$ , and a characteristic mass of  $M_b = 3.1 \times 10^{10} M_\odot$  at a ‘break’ age of  $\text{Age}_b = 11.8$  Gyr that marks the transition of the two slopes, with a ‘sharpness’ parameter for the break in slope given by  $\alpha = 10$ . This bimodal trend is also evident in the other age relations, but to a lesser extent. Using the mass-weighted age, the presence of a narrow sequence of old galaxies can be seen above the transition mass, with a distinct change in the mass-scaling behaviour below this mass. This transition mass is the same as the inflection point of the ‘zone of exclusion’ (ZOE) determined by Cappellari et al. (2013b), which defines the low- $R_e^{\text{maj}}$  envelope of the Mass Plane as shown in Figs. 6 and 7. It also corresponds to the characteristic mass marking the transition between star-forming and passive galaxies found by Kauffmann et al. (2003), as well as a possible change in slope of the relation between supermassive black hole masses and their host galaxy mass (Scott et al. 2013). Considering the population trends on the Mass Plane, the correspondence of this ‘knee’ in the age-mass relation to the inflection of the ZOE on the mass-size plane becomes obvious: below this mass, the low- $R_e^{\text{maj}}$  envelope turns toward larger galaxies, with correspondingly younger ages and longer formation timescales.

We note again that we impose an upper age limit of 14 Gyr on the model templates used, in accordance with the fiducial age of the Universe ( $13.798 \pm 0.037$  Gyr; Planck Collaboration, Ade et al. 2013). This puts a ‘ceiling’ on the mass-weighted age values, as discussed previously in Section 4.2 for the trend of mass-weighted age with  $\log(\sigma_e)$  (see Figure 9). However, unlike the trend with  $\log(\sigma_e)$ , including the full available age range of the MIUSCAT spectral models (up to 17.8 Gyr) in the spectral fit does *not* remove the bimodality. There is a slightly increased spread in mass-weighted age at the oldest values, but the mass at which the ages transition to a narrow range of old values is unchanged, and the clear bimodal shape remains. Moreover, extrapolating the linear trend in Figure 10 from masses below the transition mass results in strongly unphysical ages ( $\approx 30$  Gyr) at our highest masses, suggesting that the departure from a simple linear trend is not artificial.

Mass-weighted metallicity shows a tight relation with  $M_{\text{JAM}}$ , with only slightly higher scatter than with velocity dispersion. This supports the interpretation from the Mass

Plane that iso-surfaces of constant metallicity run steeply with mass, especially for mass-weighted metallicity.

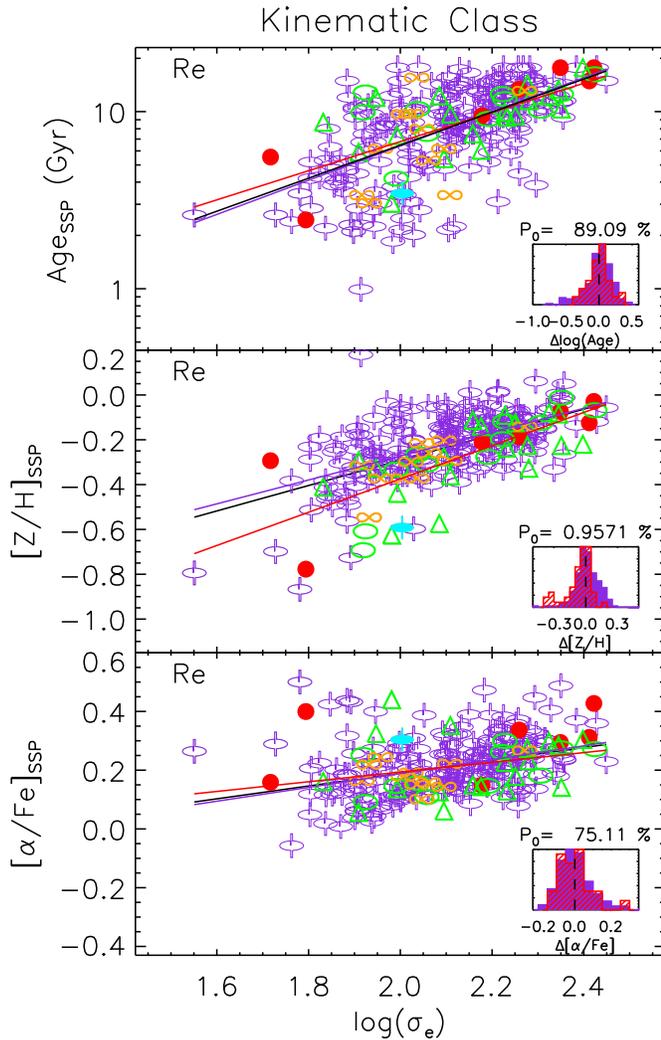
The half-mass formation time is distributed similarly to the mass-weighted age, being strongly related to that property. We include it here for completeness.

#### 4.5 Kinematic class

One of the key findings of the SAURON survey was the discovery that ETGs have essentially two categories of global dynamics, characterised as ‘fast’ and ‘slow’ rotators (Emsellem et al. 2007). Fast rotators have extended regular rotation fields well described as inclined disks, though they are clearly not ‘thin-disk’ systems, having significant velocity dispersion also. Slow rotators do not show regular rotation, and instead show counter-rotating structures, decoupled cores, or no rotation at all. In Emsellem et al. (2011), we expand on this work with the complete ATLAS<sup>3D</sup> sample, using the quantity  $\lambda_R$  as a way of separating these two classes of ETG. In Krajnović et al. (2011), we have shown that, applying the Fourier analysis technique of kinemetry (Krajnović et al. 2006), we can further classify the velocity fields of galaxies, giving ‘regular’ rotators (corresponding very closely to the fast rotators, which show disk-like rotation), and ‘non-regular’ rotators (corresponding closely to the slow rotators, which show various substructures that cannot be described by simple disk-like rotation).

Figure 11 presents the SSP- $\sigma_e$  relations as before, but this time indicating the main kinematic classes described in Krajnović et al. (2011). From this figure it can be seen that the various kinematic classes span broad ranges in velocity dispersion and stellar population properties. Non-regular rotators are on average slightly older than regular rotators, although we note there are several examples of non-regular rotators among the youngest galaxies in our sample. There is little discernible difference in abundance ratios, though we note that the massive non-rotators are among the most enhanced objects in our sample. The most noticeable difference between the regular and non-regular rotators is with metallicity, where the non-regular rotators appear to fall generally on or below the global trend, giving rise to a moderately significant difference in the residual distributions around the global linear fit as quantified by the K-S test probability,  $P_0$ , indicated in the insert histograms. This implies that non-regular rotators have experienced less metallicity enrichment, on average, than their regular-rotating counterparts of the same mass. In Young et al. (2011), we also found that non-regular rotators typically have lower mass fractions and detection rates of molecular gas (though there are notable exceptions at low velocity dispersion values, e.g. NGC1222, NGC3073). Less molecular gas implies less star formation, which is the mechanism for enriching the stellar population with metals.

The offset in metallicity at fixed velocity dispersion can be interpreted conversely as an offset in velocity dispersion for a fixed metallicity, which could arise as a result of a major merger. The metallicity offset is also seen in the relation with  $M_{\text{JAM}}$ . Interpreting the offset as one of increased mass and velocity dispersion at fixed metallicity, we find that the linear-fit mass- ( $\sigma$ -) metallicity relation is offset by 0.3 dex in  $\log(M_{\text{JAM}})$  corresponding to a doubling of the mass, with a corresponding shift of 0.1 dex in  $\log(\sigma)$ . These shifts are



**Figure 11.** As per Fig 8, but with symbols indicating kinematic classes as defined in Krajnović et al. (2011). Red circles show little or no rotation, green ellipses have complex velocity maps with no specific features, green triangles are KDCs, orange lemniscates have off-centre dispersion peaks indicative of superposed counter-rotating disks. These correspond to the respective classes *a*, *b*, *c*, and *d* in Krajnović et al. (2011), referred to collectively as ‘non-regular rotators’. Purple spindles are the class *e* from Krajnović et al. (2011), or ‘regular rotators’. Filled spindles have obvious bars. Red and purple lines show fits to only the non-regular and regular rotators respectively. The solid black lines show fits to all classes together. Parameters of these linear fits are given in Table 5. Inset histograms show the residuals of the regular (purple) and non-regular (red) rotators about the global best fit to the combined population, quoting also  $P_0$ : the probability that the chance deviation between the two samples should be larger than that measured, as derived from a K-S test of the two histograms. Small values of  $P_0$  imply the distributions are different. The only significant difference is in the metallicity relation, where the average offset of the non-regular rotators is consistent with the mass increase expected from a single major (equal mass ratio) merger.

consistent with the predictions of Hilz et al. (2012) for a single 1:1 dry merger, where the velocity dispersion is found to increase by a factor 1.2-1.5 following a 1:1 merger. This is different to the behaviour expected from the idealised virial predictions of Naab et al. (2009), where velocity dispersion remains unchanged. The idealised case, however, ignores the effects of violent relaxation, which acts to unbind a non-negligible fraction of the mass and increase the central density and velocity dispersion. This interpretation implies that, on average, non-regular rotators are consistent with having experienced a single dissipationless major merger that could have shaped their current kinematic characteristics. This picture is consistent with Khochfar et al. (2011), where semi-analytic models show that slow/non-regular rotators have on average experienced more major mergers than fast/regular rotators. It is also consistent with Scott et al. (2013), where we showed that the radial relation of Mg *b* line strength and escape velocity for our sample limits the number of possible major mergers to  $\lesssim 1.5$ , beyond which objects would become inconsistent with the observed trends.

We add a note of caution to this simple interpretation. Firstly, if the offset is in the velocity dispersion, one may reasonably expect to see some sign of a corresponding offset in the other population parameters, which is not observed. This is complicated, however, by the differing response of the parameters to mass- and luminosity-weighting. Metallicity is essentially a mass-weighted property, suggesting that the offset is driven by a process that impacts most of the stellar mass in these galaxies. Age, however, is sensitive to small amounts of recent star formation, making trends difficult to interpret. Figure 4 showed that ages  $\gtrsim 8$  Gyr are close to the mass-weighted values. Considering only galaxies with ages older than this, the offset in metallicity remains apparent in the K-S test, but the relations with age and abundance become sufficiently shallow that any offset in age or velocity dispersion is poorly constrained, so we cannot firmly rule out a shift in the velocity dispersion axis on this basis.

Secondly, as we have shown in Naab et al. (2014), the net angular momentum of a galaxy is a sensitive tracer of its evolution, but with a broad range of possible paths. Some of the non-regular rotators are among the very metal-poor (Fe-poor) objects, which have high mass fractions of cold gas and young ages, consistent with a gas-rich accretion or merger event. On an individual galaxy basis, the specific assembly history can be varied. However, here we highlight the offset in the mass- ( $\sigma$ -) metallicity relation as the only significant *average* difference between the stellar population properties of these broad kinematic classes.

#### 4.6 The role of environment

The ATLAS<sup>3D</sup> sample spans a range of environment, from sparse fields to the central regions of the Virgo cluster. Differences in the galaxy populations of high and low density environments are well documented (e.g. Thomas et al. 2005). However, reported trends and results are surprisingly weak considering the predictions of galaxy formation simulations, where dense regions differ in the star-formation properties and assembly histories (Dekel & Birnboim 2006; Khochfar & Ostriker 2008; Cattaneo et al. 2008).

Figure 12 shows the SSP-equivalent population parameters and velocity dispersion for our  $R_e$  aperture, but this time

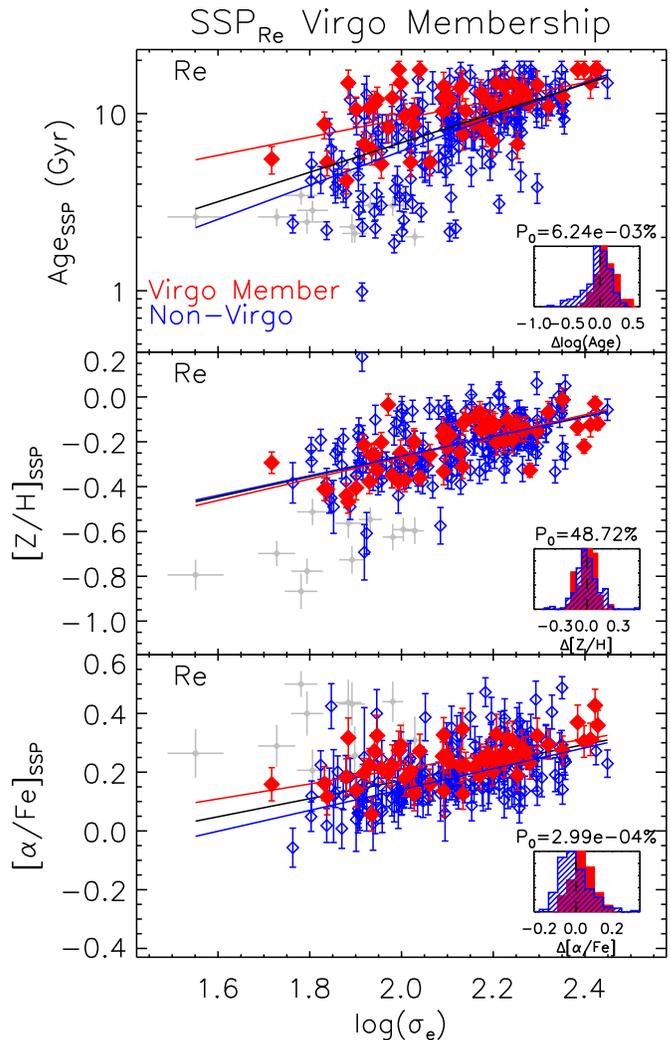
we colour Virgo members red, and non-members blue. Inset histograms show the residuals of these two sub-samples about the global best fit to the combined sample, quoting also  $P_0$ : the probability that the chance deviation between the two samples should be larger than that measured, as derived from a Kolmogorov-Smirnov (K-S) test of the two histograms. Small values of  $P_0$  imply the distributions are different.

Both Virgo and non-Virgo galaxies show similar trends of population properties with velocity dispersion. However Virgo objects are older by approximately 2.5 Gyr on average. The two samples deviate in age most significantly when comparing the largest aperture size. This reflects the fact that when Virgo galaxies have young stars, they tend to be centrally concentrated, whereas outside the cluster environment, recent star formation can be widespread across the galaxy on  $R_e$  scales. Differences in abundance ratios are also statistically most significant within the  $R_e$  aperture, showing a very small probability that the measured deviation of the two distributions occurred by chance. Virgo galaxies have systematically higher abundance ratios on these large spatial scales than their counterparts in the field at the *same* velocity dispersion. At face value, this indicates a truncated star formation history on large scales compared to outside the cluster, indicative of a process acting to suppress star formation preferentially on larger scales, such as ram-pressure stripping (Gunn & Gott 1972). The suppression of spatially extended star formation in Virgo cluster early-type galaxies is also supported by the study of current-day star formation in early-type galaxies by Shapiro et al. (2010), who found that all early-type galaxies with ‘widespread’ star formation are found outside of dense environments. Likewise, in previous papers of the ATLAS<sup>3D</sup> series we have reported on truncated distributions of molecular gas (Davis et al. 2013) and a lack of extended HI disks (Serra et al. 2012) within Virgo compared to lower-density regions, further supporting this picture.

The metallicity distributions are essentially the same between Virgo and non-Virgo objects in all aperture sizes, though we note that the ‘low-Fe’ objects are not present in Virgo, suggesting that the cluster environment specifically prohibits the accretion of low metallicity gas, even though the incidence of molecular gas detections within Virgo is statistically indistinguishable from the field (Young et al. 2011). We return to this point when considering the empirical star formation histories in Section 5.2

Our findings here are consistent with our spatial gradient analysis presented in Scott et al. (2013). Interestingly, although the global properties show differences between Virgo and non-Virgo objects, objects in both environments show the same radial gradient of Mg *b* with local escape velocity (though Virgo is offset to higher Mg *b* values). This implies that the mechanism acting to modify the star formation process in the cluster galaxies maintains the close relationship between the stellar population and the local gravitational potential.

Virgo is the largest bound group of galaxies in our sample. However, there are regions of our volume that have small-scale densities comparable to parts of Virgo. Likewise the outer regions of Virgo have densities comparable to that in the field. In Cappellari et al. (2011b), we showed that using a local estimator of environmental density gives

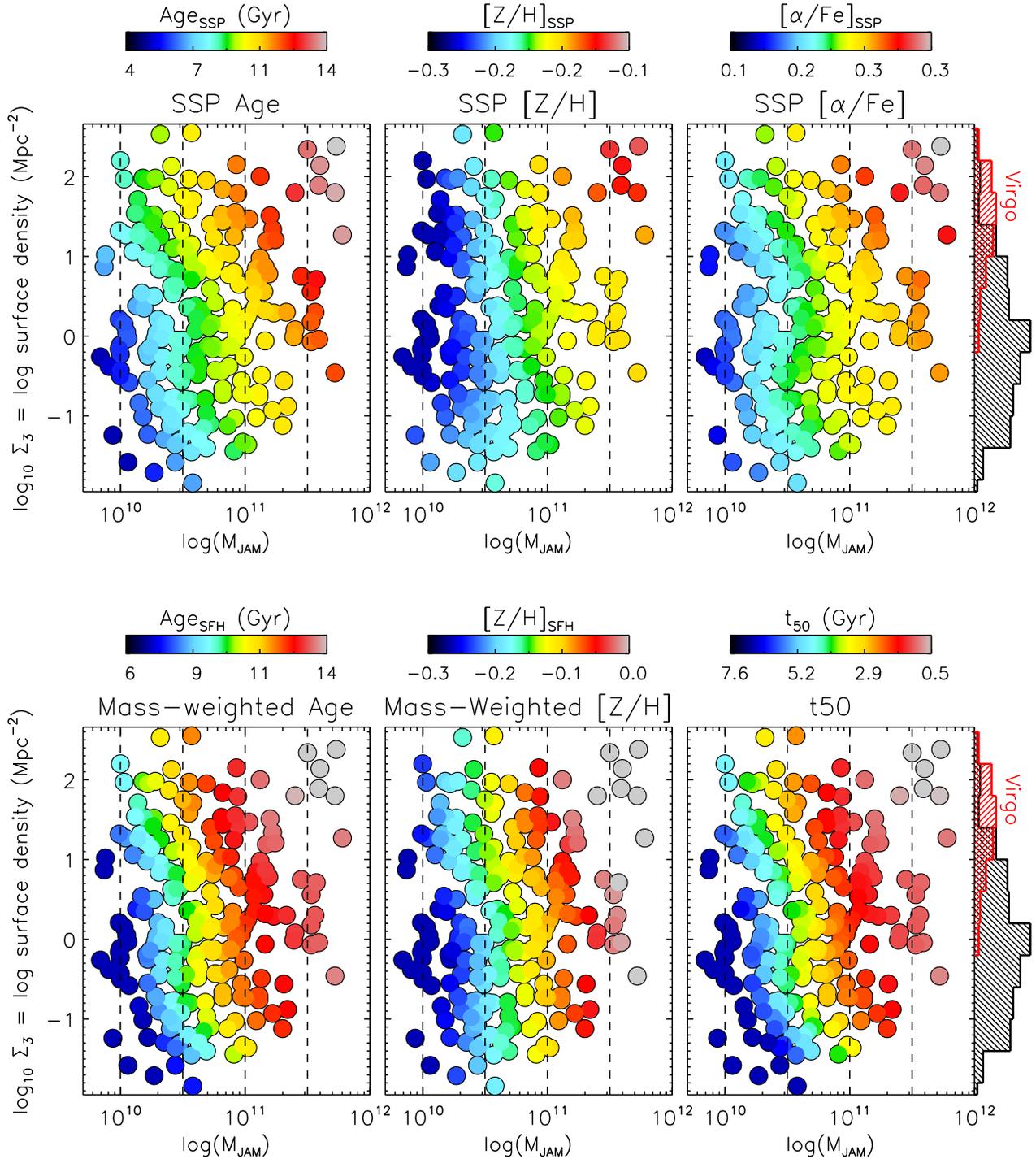


**Figure 12.** Same as Fig. 8, but separated into Virgo member galaxies (red symbols) and non-Virgo (blue). Magenta cross-bars show the outlying Fe-poor galaxies. Black, red and blue lines show the error-weighted best fit straight line to all galaxies (excluding Fe-poor objects), Virgo, and non-Virgo galaxies respectively. Inset histograms show the residuals of the best fit line to all galaxies, separated into Virgo (red) and non-Virgo (blue) members.  $P_0$  gives the probability that the two histograms are drawn from the same populations, as derived from a K-S test. Metallicities are very similar, but age and  $[\alpha/\text{Fe}]$  show a high probability of sampling two different distributions.

a well-defined relation of kinematic morphologies with environment. Here we explore if such similar trends exist for the stellar populations.

Figure 13 shows the LOESS-smoothed distribution of SSP-equivalent and mass-weighted population parameters (both use  $R_e$  apertures) as a function of dynamical mass and environmental density. Environment is quantified by  $\Sigma_3 = N_{\text{gal}}/(\pi R_3^2)$ , which gives the surface density in  $\text{Mpc}^{-2}$  of galaxies inside a cylinder of radius  $R_3$  and height  $h = 600 \text{ km s}^{-1}$  (i.e.  $\Delta V_{\text{hel}} < 300 \text{ km s}^{-1}$ ) centered on the galaxy, which includes the  $N_{\text{gal}} = 3$  nearest neighbours from the parent sample, as described in Cappellari et al. (2011a).

It can be seen from Figure 13 that regions of constant



**Figure 13.** Distribution of dynamical mass  $M_{\text{JAM}}$  versus environment density tracer  $\Sigma_3$ , defined in the main text. In each panel, points are coloured according to the stellar population parameters indicated in the plot title, after locally averaging the measured values using the LOESS technique. The top row provides the SSP-equivalent parameters of age, metallicity and abundance ratio; the bottom row gives the mass-weighted age, metallicity and half-mass formation time. The far-right histograms indicate how Virgo cluster members (red) and non-members (black) are distributed with  $\Sigma_3$ .

stellar population parameters show varying dependency on both mass and environment. The population properties are essentially independent of environment over two orders of magnitude (within the range  $-1 \leq \log(\Sigma_3) \leq 1$ ). The highest-density regions, dominated by Virgo cluster mem-

bers (see vertical histograms for the distribution of members with  $\log(\Sigma_3)$ ), shift markedly to older ages, higher abundance ratios, and earlier half-mass formation timescales for a given mass with respect to the lower density regions. Metallicity shows only a weak difference between high and low

**Table 5.** Line fit parameters.

(i) SSP = $a \log(\sigma_e) + b$ (Figure 8)									
$\log(\text{Age}_{\text{SSP}})$			$[Z/H]_{\text{SSP}}$			$[\alpha/Fe]_{\text{SSP}}$			Notes
a	b	$\sigma$	a	b	$\sigma$	a	b	$\sigma$	
$1.14 \pm 0.09$	$-1.59 \pm 0.13$	0.22	$0.31 \pm 0.04$	$-0.62 \pm 0.06$	0.10	$0.37 \pm 0.03$	$-0.61 \pm 0.04$	0.06	$R_e/8$
$0.93 \pm 0.08$	$-1.07 \pm 0.11$	0.20	$0.45 \pm 0.04$	$-1.09 \pm 0.06$	0.10	$0.33 \pm 0.03$	$-0.49 \pm 0.04$	0.07	$R_e/2$
$0.83 \pm 0.08$	$-0.83 \pm 0.11$	0.19	$0.45 \pm 0.04$	$-1.16 \pm 0.06$	0.10	$0.31 \pm 0.03$	$-0.44 \pm 0.05$	0.08	$R_e$

(ii) SFH = $a \log(\sigma_e) + b$ (Figure 9)									
$\log(\text{Age}_{\text{SFH}})$			$[Z/H]_{\text{SFH}}$			$t_{50}$			Notes
a	b	$\sigma$	a	b	$\sigma$	a	b	$\sigma$	
$0.54 \pm 0.04$	$-0.12 \pm 0.06$	0.10	$0.63 \pm 0.04$	$-1.52 \pm 0.06$	0.09	$-10.47 \pm 0.76$	$25.30 \pm 1.11$	1.93	$R_e$

(iii) SSP = $a \log(M_{\text{JAM}}) + b$ (Figure 10, left)									
$\log(\text{Age}_{\text{SSP}})$			$[Z/H]_{\text{SSP}}$			$[\alpha/Fe]_{\text{SSP}}$			Notes
a	b	$\sigma$	a	b	$\sigma$	a	b	$\sigma$	
$0.25 \pm 0.03$	$-1.70 \pm 0.10$	0.21	$0.12 \pm 0.02$	$-1.49 \pm 0.05$	0.11	$0.08 \pm 0.01$	$-0.66 \pm 0.04$	0.08	$R_e$

(iv) SFH = $a \log(M_{\text{JAM}}) + b$ (Figure 10, right)									
$\log(\text{Age}_{\text{SFH}})$			$[Z/H]_{\text{SFH}}$			$t_{50}$			Notes
a	b	$\sigma$	a	b	$\sigma$	a	b	$\sigma$	
$0.15 \pm 0.02$	$-0.60 \pm 0.05$	0.12	$0.21 \pm 0.01$	$-2.39 \pm 0.05$	0.09	$-2.94 \pm 0.32$	$34.45 \pm 1.04$	2.13	$R_e$

(v) SSP = $a \log(\sigma_e) + b$ (Figure 11, kinematic classes)									
$\log(\text{Age}_{\text{SSP}})$			$[Z/H]_{\text{SSP}}$			$[\alpha/Fe]_{\text{SSP}}$			Notes
a	b	$\sigma$	a	b	$\sigma$	a	b	$\sigma$	
$0.97 \pm 0.09$	$-1.12 \pm 0.12$	0.20	$0.53 \pm 0.05$	$-1.34 \pm 0.07$	0.10	$0.24 \pm 0.04$	$-0.28 \pm 0.06$	0.09	Regular Rotators
$0.82 \pm 0.14$	$-0.81 \pm 0.20$	0.17	$0.74 \pm 0.10$	$-1.86 \pm 0.14$	0.12	$0.17 \pm 0.07$	$-0.14 \pm 0.10$	0.09	Non-Regular Rotators
$0.93 \pm 0.07$	$-1.06 \pm 0.11$	0.20	$0.57 \pm 0.04$	$-1.43 \pm 0.06$	0.12	$0.22 \pm 0.03$	$-0.25 \pm 0.05$	0.09	All (inc. low-Fe)

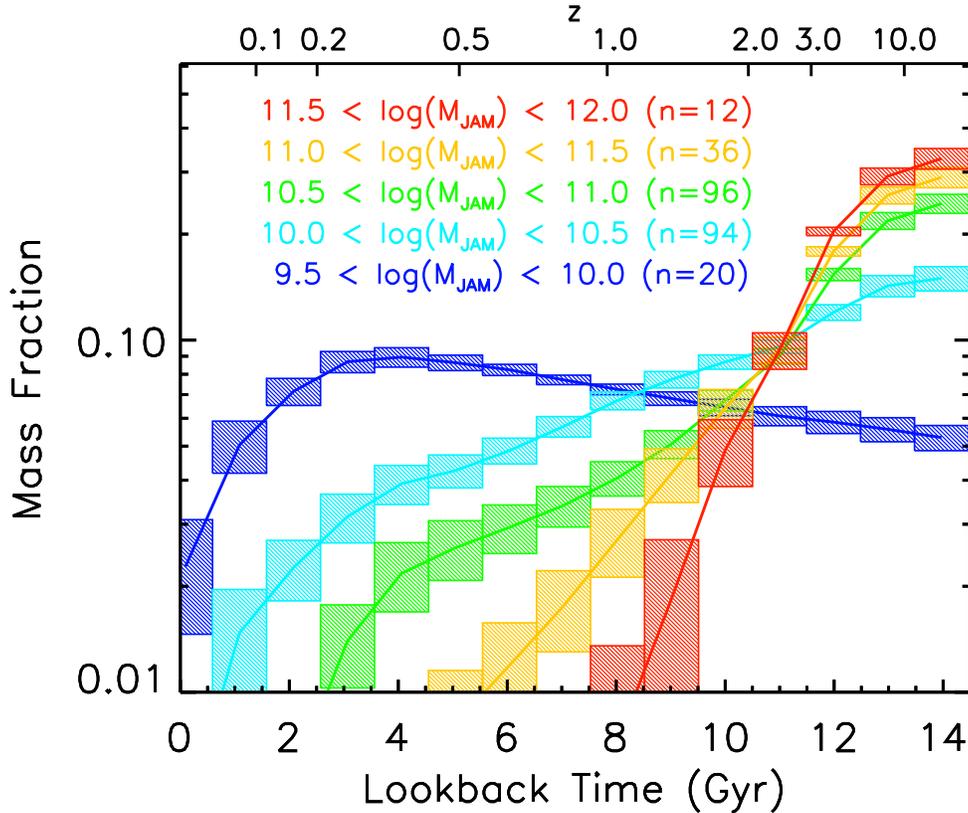
(vi) SSP = $a \log(\sigma_e) + b$ (Figure 12, Virgo members)									
$\log(\text{Age}_{\text{SSP}})$			$[Z/H]_{\text{SSP}}$			$[\alpha/Fe]_{\text{SSP}}$			Notes
a	b	$\sigma$	a	b	$\sigma$	a	b	$\sigma$	
$0.51 \pm 0.11$	$-0.06 \pm 0.16$	0.14	$0.48 \pm 0.07$	$-1.23 \pm 0.10$	0.08	$0.25 \pm 0.05$	$-0.30 \pm 0.07$	0.06	Members
$0.96 \pm 0.09$	$-1.14 \pm 0.13$	0.18	$0.44 \pm 0.05$	$-1.14 \pm 0.07$	0.10	$0.36 \pm 0.04$	$-0.57 \pm 0.06$	0.08	Non-Members

Note: The above table gives the parameters for the linear fits shown in the figures indicated in the table. The fits were made using the outlier-resistant two-variable linear regression method implemented in the IDL routine ‘robust\_linefit.pro’ of the IDL NASA Astronomy Library (Landsman 1993). The  $\sigma$  column gives the standard deviation of the fit to the ATLAS<sup>3D</sup> data. The ‘Notes’ column indicates the aperture size and relevant sub-samples used in the fit. If no aperture size is given,  $R_e$  is assumed.

densities, in the sense that the highest densities show mildly higher metallicities. At the lowest densities ( $\log(\Sigma_3) < -1$ ), there is a weak inflection towards even younger ages, lower abundances, and longer half-mass formation timescales.

Most of our sample (62%) lie within the range  $-1 \leq \log(\Sigma_3) \leq 1$ , within which galaxy mass is the only significant

driver of stellar population differences. As we have shown, mass-weighted properties are less sensitive to small amounts of recent star formation, and therefore trace the build up of stars over the entire lifetime of a galaxy. The consistent trends between the mass-weighted and SSP-equivalent indicators therefore imply that, on average, the influence of



**Figure 14.** The average star formation history of the corresponding bin in dynamical mass,  $M_{\text{JAM}}$ . The star formation history is quantified here as the relative fraction of stellar mass formed at each epoch, derived from the regularised weighted combination of MIUSCAT SSP models fitted to the aperture spectrum. Hatched regions indicate the standard error on the mean at each lookback time, estimated from the standard deviation in each mass-bin of age divided by  $\sqrt{n}$ , where  $n$  is the number of objects in the mass bin as indicated. The upper horizontal scale indicates corresponding redshift, assuming  $\Omega_0 = \Omega_m + \Omega_\Lambda = 1$ , where  $\Omega_m = 0.27$  and  $\Omega_\Lambda = 0.73$ .

environment on the stellar populations of early-type galaxies has a long-term influence on the galaxies' evolution. The suppression of only *recent* star formation in cluster galaxies, for example, is not sufficient to explain the trends we see, which extend across all masses, and in galaxies where significant levels of star formation is long-since over. As we show in Section 5.1, there must also have been a more rapid build of up massive galaxies in high density environments at early times, followed by subsequent preferential suppression of star formation in those regions. Outside of cluster-like over-densities, the mass of individual galaxies (or perhaps more accurately, the parent haloes of today's galaxies) seems to have been the dominant influence governing the star-formation history over the effects of the local galaxy density. This supports a picture of separable effects of mass and environment on the stellar populations of ETGs, consistent with Peng et al. (2010). The weak influence of environment (as compared with mass) in low- and intermediate-density regimes is also similar to the findings of Thomas et al. (2010).

## 5 STAR FORMATION HISTORIES

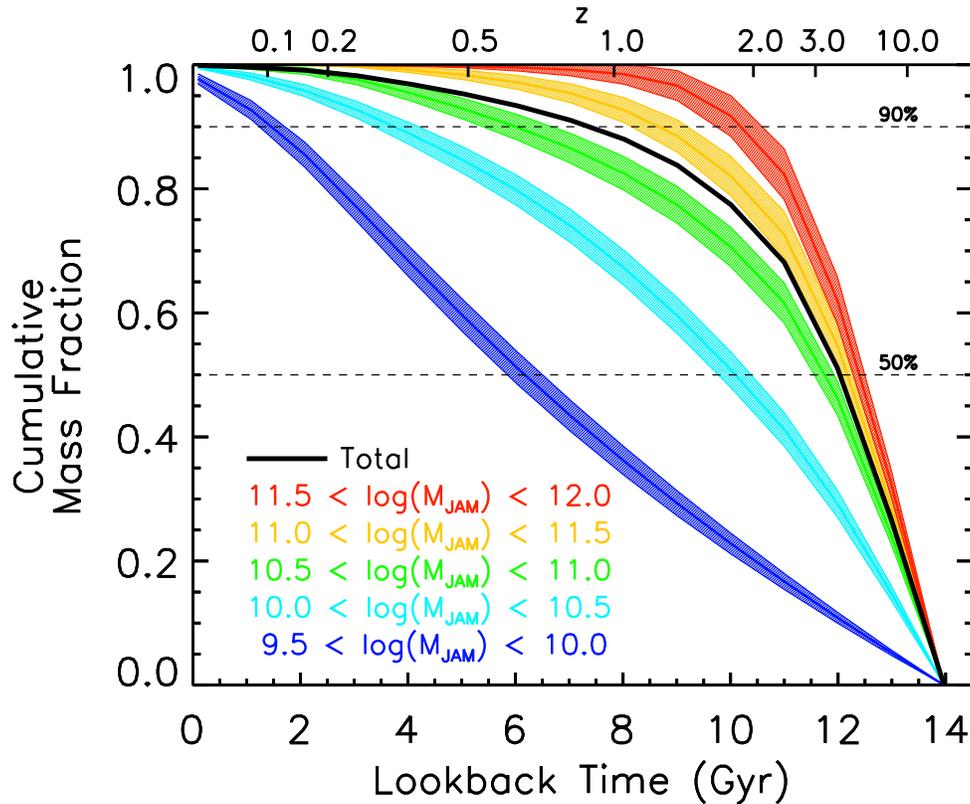
In this section, we use our spectral fits to the integrated  $R_e$  aperture spectra to investigate the history of star formation

within our sample galaxies. The results which follow use the same spectral fits used to determine the mass-weighted population parameters presented above. As described in Section 2.5, we make a regularized linear fit to our aperture spectra of SSP model spectra sampling a grid of age and metallicity. The resulting distribution of weights provides an estimate of the star formation history (SFH) of the galaxy in question, by describing the mass contributions of different populations with given formation epochs. The spectral fit is made using model spectra logarithmically-spaced in age. Here we rebin the resulting logarithmically-spaced weights onto a linear time axis, rigorously preserving the total mass fractions within each linear time step (set as 1 Gyr).

### 5.1 Star formation histories as a function of mass

Figure 14 shows the measured SFHs of our sample averaged within five bins of dynamical mass (as measured at  $z = 0$ ), spanning the range of our sample. The SFH is presented as the mean fraction of stellar mass formed in each 1 Gyr step, down to mass fractions of 1%, for galaxies within the (present-day) mass bin. The variation of mass fraction formed within each 1 Gyr bin is indicated by the hatched regions, quantified as the mean standard deviation within each 1 Gyr bin.

On average, objects with  $M_{\text{JAM}} > 10^{10} M_\odot$  (> 90% of



**Figure 15.** Cumulative functions of the star formation histories shown in Figure 14, showing the integrated mass fraction existing at given time. Colours correspond to mass bins as in Figure 14. Hatched regions indicate the dispersion of values at each time step.

our sample) had their peak in star formation at the earliest epochs, followed by a decline that is systematically more rapid at higher (present-day) masses. In our lowest mass bin, the behaviour is different: on average, these objects build a gradually increasing fraction of their stellar mass as the universe evolves, until a more rapid decline beginning around 4 Gyr ago. This decline is most easily interpreted as a result of our sample selection, which targeted galaxies consistent with an early-type morphology, and thus preferentially selects objects with low present-day star formation rates (though this is not a selection requirement, and many objects show evidence for ongoing star formation). We note also that this lowest mass bin likely suffers from incompleteness, as the sample was selected based on K-band luminosity, which has some variation in mass-to-light ratio with stellar population properties. This may bias against old populations at this low mass extreme, as they will be slightly fainter than their younger counterparts. Nonetheless, the lowest mass bin is consistent with the overall trend of more extended star formation in lower mass early-type galaxies, with a truncation at more recent times than the high mass objects.

The star-formation histories shown in Fig. 14 provide an alternative picture to that of the schematic view inferred from assuming linear trends of SSP-equivalent parameters, as in Thomas et al. (2005, 2010). In that case, galaxies of a given mass have star formation rates that rise and fall symmetrically in time, being described by a Gaussian centred on the mean SSP-equivalent age, with a width given by a conversion of abundance ratio to star-formation timescale.

With that approach, the stellar population scaling relations arise from a systematic delay in the peak of star formation in lower mass galaxies towards later times, combined with a corresponding increase in the star-formation timescale.

By comparison, our non-parametric approach shows that objects of all masses experience a significant epoch of star-formation at the earliest times, and that the subsequent variation in present-day stellar population properties is mainly driven by the differing rate at which star formation declines for galaxies of different masses. The overall picture is qualitatively similar between both approaches. However, by reproducing the observed spectra with a distribution of ages and metallicities directly, our non-parametric star-formation histories emphasise the significance of early star formation in all galaxies, and the subsequent decline in star formation activity that varies systematically with mass, independent of the timescale of the individual star formation events. This reflects the conceptual differences of fitting the full spectrum with a distribution of models to infer the star formation history directly, compared to using line-strength indices to estimate mean properties. Our findings are in good agreement, for example, with the previous studies of Heavens et al. (2004) and Panter et al. (2007), who used comparable spectral-fitting methods on thousands of aperture spectra from SDSS (though without any morphological selection applied).

Figure 15 presents the associated mean cumulative mass fraction as a function of look-back time for the same mass bins as in Figure 14, as well as the total cumulative formed

mass for the entire sample. Assuming that our sample represents most of the stellar content of today’s local Universe, approximately 50% of all stars formed within the first 2 Gyr following the big bang. The vast majority of these stars reside today in the most massive galaxies, which themselves formed 90% of their stars within the first 3 Gyr (i.e. by  $z \sim 2$ ). The low mass objects, in contrast, have formed barely half their stars in this time interval, forming the remaining 50% of their stellar mass in the last 6–10 Gyr. We note that, from this analysis, we can only tell *when* the stars formed, and not whether they formed within the gravitational potential of the current-day galaxy, or instead were formed in separate, smaller systems that later combined to form the galaxy we see today.

Figure 16 shows the fractional formation masses in absolute terms (top panel), which allows comparison to other studies that measure star formation rates. We do this by assuming that our star formation histories apply to the galaxy’s total mass as quantified by  $M_{\text{JAM}}$ , scaling this present-day mass by the weights returned by pPXF. We also derive the history of *specific* star formation rates (middle panel) by normalising the star formation rate by the integrated mass formed by the epoch being considered (though the galaxy may not have assembled yet). In addition, since our sample is volume-limited, we also present the history of star formation rate density for our local volume implied by our analysis, both as a function of mass, and the combined total for our full sample (bottom panel).

The top panel of Figure 16 clearly shows that the most massive current-day galaxies had the highest absolute star formation rates of all such objects for the first 4 Gyr following the big bang. The bottom panel of Figure 16, however, shows that these objects do not dominate the history of integrated star formation rate density, due to their smaller number density.

The middle panel of Figure 16 presents the time-averaged specific star formation rate (sSFR), defined as the mean star formation rate during the 1 Gyr time bin divided by the mass formed up to that epoch (taken as the mass formed up to the centre of the time bin). For comparison, we show direct measurements of the instantaneous sSFR taken from the literature. Our derived sSFR values lie generally well below the direct measurements, which is most easily explained by the inconsistent timescales probed by the two measures of star formation rate. Our ‘archeological’ values probe the mean time-averaged star formation during each  $10^9$  year time step, whereas the direct measurements probe the star formation occurring on timescales of  $10^6 - 10^8$  years. The super-solar abundance ratios of our most massive (and oldest) galaxies require star formation to have occurred on characteristic timescales of order  $10^8$  year (Thomas et al. 2005), which is below the time resolution of our spectral fitting at these early epochs. Such shorter ‘bursts’ of star formation would boost our effective star SFR and sSFR values by an order of magnitude, potentially making them more consistent with direct instantaneous measurements. However, we cannot directly constrain this timescale during the oldest age bins using our current approach.

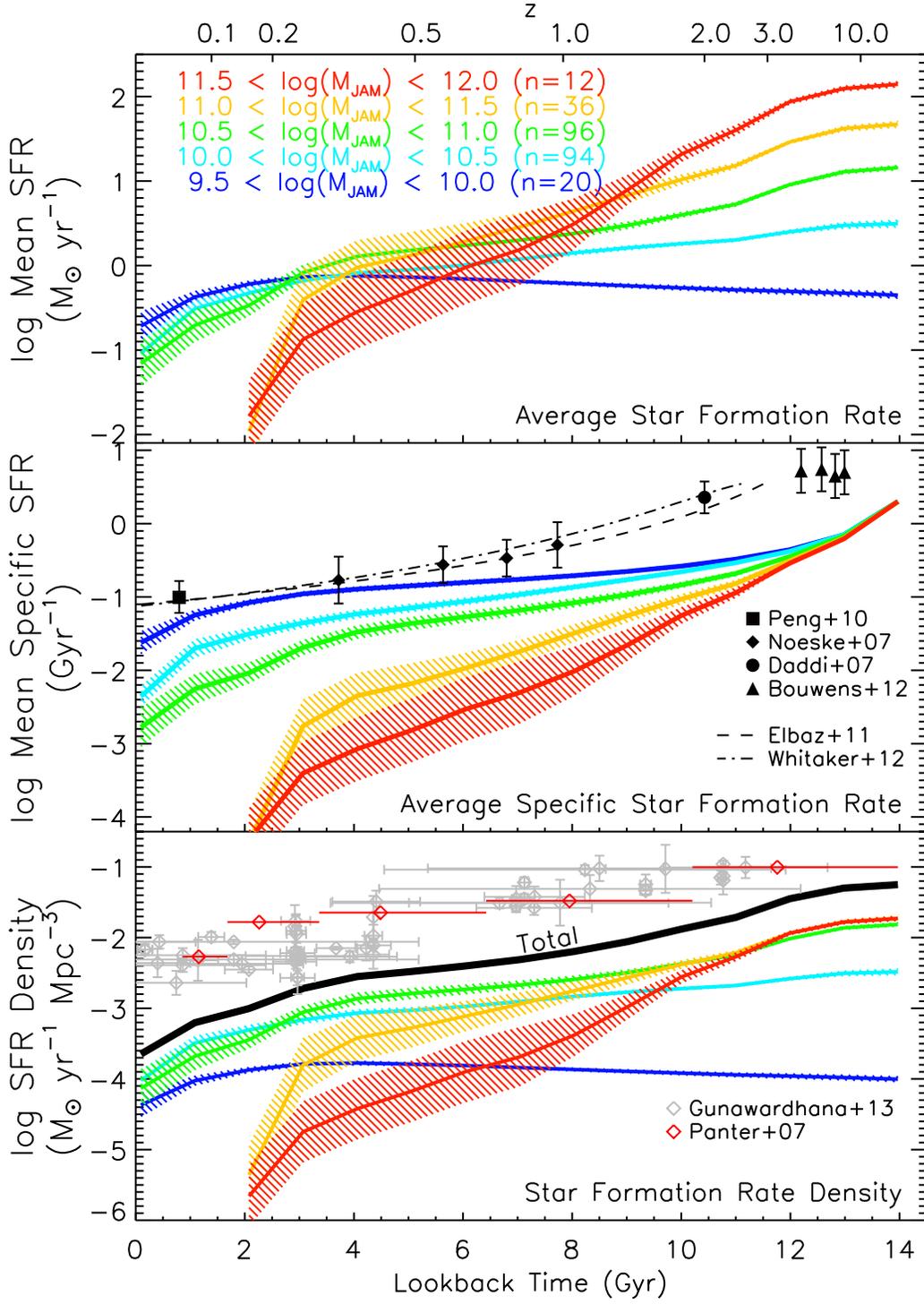
In addition to a redshift dependence, the sSFR histories show a strong mass-dependence of the sSFR at a given redshift, irrespective of the assumed star-formation timescale. This indicates the influence of a mass-dependent feedback

mechanism that is relatively more efficient in higher-mass galaxies, reducing the sSFR more rapidly than at lower masses, and bringing higher-mass galaxies off of the star-formation main sequence at earlier times. Through our ‘archeological’ approach, we can directly trace when galaxies turn off of the star-formation main sequence and become quiescent. Due to the problems of comparing instantaneous and archeological sSFR estimates, however, we cannot easily compare with the measured sSFR main sequence at high redshift. Instead, we adopt the sSFR of  $0.1 \text{ Gyr}^{-1}$  measured in local galaxies from Peng et al. (2010) as a robust threshold between ‘star forming’ and ‘quiescent’ states, and measure the epoch at which our galaxies cross this threshold. Figure 17 presents this ‘quenching redshift’ for the different mass bins. We find a near-linear dependence with  $\log(M_{\text{JAM}})$ , with galaxies of mass above  $10^{11} M_{\odot}$  quenching their star-formation at redshift  $z = 2 - 4$ , whereas galaxies below  $10^{10} M_{\odot}$  quench below redshift 1, with our lowest mass galaxies quenching below  $z \sim 0.2$  (in the last 1 Gyr), on average.

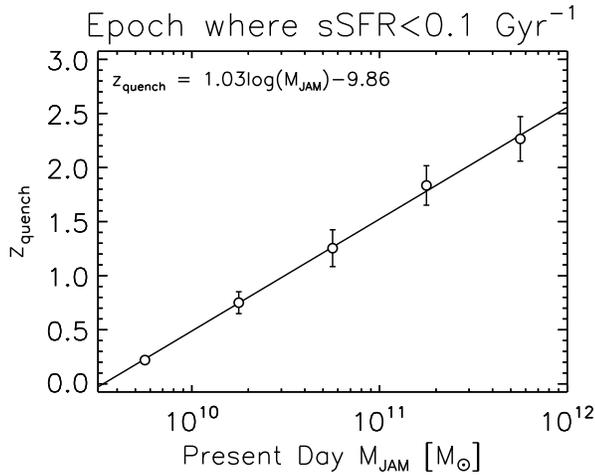
The cosmic SFR density for ATLAS<sup>3D</sup> given in the bottom panel of Figure 16 also shows measurements from Panter et al. (2007) derived using a similar technique to ours using SDSS spectra of  $\sim 3 \times 10^5$  galaxies, and from the compilation of Gunawardhana et al. (2013) using various measures of the (instantaneous) star formation rate. Our estimate of the total star formation rate density is significantly lower than these other studies, differing by an order of magnitude at most recent times, but with better agreement at the earliest epochs. We interpret this as a result of our morphological selection, which effectively excludes galaxies with strong total star formation rates at the current time, and which have likely contributed significantly to the integrated star formation budget of the universe over its lifetime. In contrast, the literature studies either impose no particular mass or morphological selection (as in the case of Panter et al. 2007), or preferentially select galaxies with detectable star formation (as in the data compiled by Gunawardhana et al. 2013), both of which will act to increase the inferred integrated star formation rate. Cosmic variance may also play a role given our modest survey volume, which for example lacks extremely massive clusters like Coma. In general, however, we see that early-type galaxies have experienced a similarly-paced decline in star formation rate as the total population, though with a more rapid drop off and reduction in contributions from galaxies of higher masses.

## 5.2 Star formation histories as a function of environment

In Section 4.6 we presented evidence for differing stellar population properties in the different environments present in our sample. Here we investigate whether this is also recovered directly in the SFHs determined from spectral fitting. Figure 18 shows the mean cumulative fraction of stellar mass formed as a function of time for members of the Virgo cluster and the rest of the sample, binned by present-day mass. The population outside of Virgo (solid lines) show consistently more extended star formation than the objects that are Virgo members (hatched regions), with the exception of the lowest mass bin, where the extended star formation is similar in both populations. This low-mass bin has just three



**Figure 16.** *Top:* Star formation rates averaged within the indicated (present-day) mass bins. These are calculated by multiplying the mass fractions in Figure 14 by the dynamical mass  $M_{\text{JAM}}$ . *Middle:* Mean (time-averaged) specific star formation rate in each  $\Delta t_{\text{SFH}} = 1$  Gyr time step (thick solid lines), obtained by normalising the star formation rate by the mass formed up to the epoch considered. Direct measurements of the instantaneous sSFR from the literature are also shown, either as discrete values (black data points) or empirically-constrained continuous relations (broken lines). *Bottom:* Integrated star formation rate density for our local volume, as a function of present day galaxy mass, and the integrated total for our sample (black line). Grey points show instantaneous star formation rate density measurements from the compilation of Gunawardhana et al. (2013); red points show the measurements of Panter et al. (2007) using a similar approach to our analysis. In all panels, hatched regions show the standard deviation of properties within each age and mass bin.



**Figure 17.** Redshift at which the sSFR in different bins of galaxy mass,  $M_{\text{JAM}}$ , crosses the value of  $0.1 \text{ Gyr}^{-1}$ , corresponding to the local sSFR of main-sequence star-forming galaxies from Peng et al. (2010). The quenching redshift follows an approximately linear relation with mass, as indicated by fit (the function is also given). Errors are computed using the spread within each mass bin, as indicated with hatched regions in Figure 16.

Virgo objects, all of which are located in the outer regions of the cluster, and are presumably entering Virgo for the first time. Overall, these trends directly show the impact of environment on the mass growth and star-formation histories of early-type galaxies, in addition to the mass-dependent effects illustrated in the previous section.

The apparent suppression of extended star formation within Virgo compared to objects of the same mass outside of the cluster environment is qualitatively consistent with an extension of what is found in Virgo late-type galaxies, where the intra-cluster medium appears to truncate extended neutral gas disks with subsequent implications on star-formation activity (Koopmann & Kenney 2004; Crowl & Kenney 2008; Chung et al. 2009). Indeed, even though the detection rate of CO in our sample is not statistically different between Virgo members and non-members (Young et al. 2011), we have shown in Davis et al. (2013) that this present-day reservoir of molecular gas responsible for fuelling current star formation is spatially less extended in the Virgo galaxies, and often shows signs of recent interaction and/or truncation.

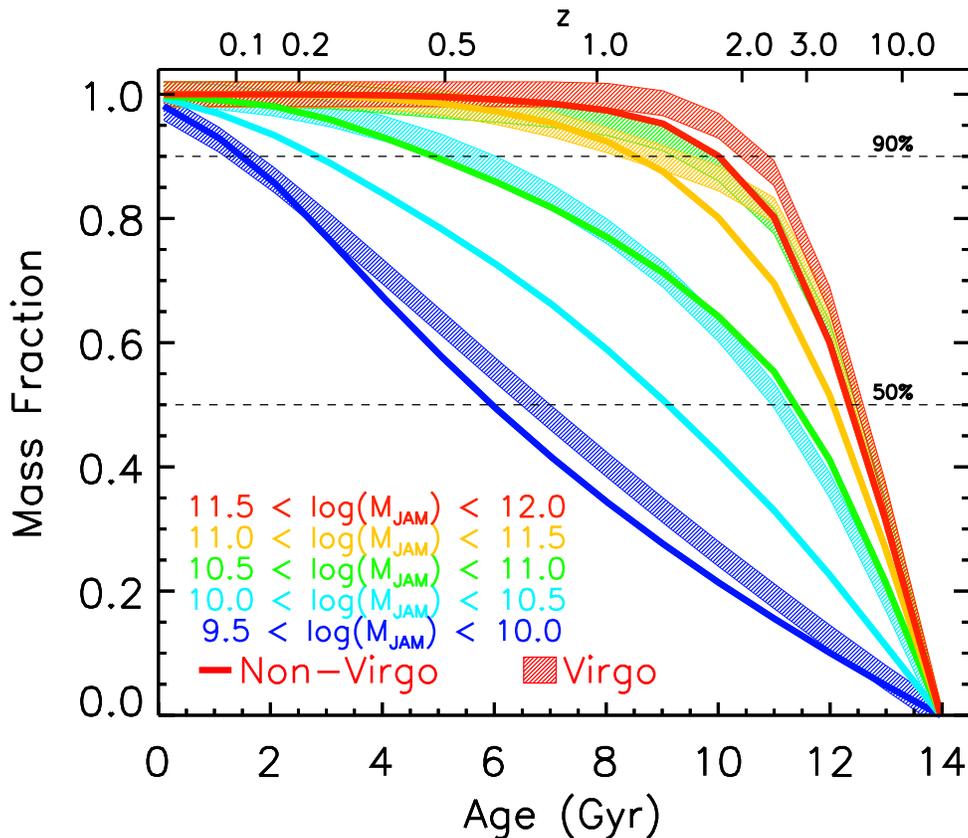
Further evidence comes from the kinematic alignment of gas in our sample. In Davis et al. (2011) we showed that gas in various phases (ionised, molecular or neutral) is almost always aligned with the stars in Virgo galaxies, in contrast to galaxies in low-density environments where more than half show misalignments. This strongly suggests that the gas supply in the cluster is limited to internal recycling. Without any additional external accretion of cold gas to maintain or periodically boost levels of star formation, galaxies in the cluster have less extended star formation histories compared to those of similar mass outside the cluster environment. In less dense environments, accretion of cold gas can still occur, as indicated by the high incidence of gas components misaligned to the stellar motions. Recall also that none of the extremely low metallicity, gas rich objects shown in Figure 5 are found in Virgo.

Our star formation histories suggest that this is not restricted to a recent suppression of star formation. On average, cluster objects have *always* exhibited shorter formation timescales than field galaxies of comparable present-day mass. This finding is in good agreement with the predictions of Khochfar & Ostriker (2008), where a combination of ram pressure stripping and gravitational heating act to reduce the duration of star formation in dense environments. The corollary to this is that, at early times when cluster galaxies were rapidly forming the bulk of their stars, their effective star formation rate would have been higher than that of galaxies in the field that would eventually match their mass in the present day. Such a picture is consistent with direct observations at high redshift, where an apparent ‘reversal’ of the low redshift behaviour (where star formation is lower in high density regions) occurs, and dense environments exhibit enhanced star formation compared to the field (Elbaz et al. 2007; Tran et al. 2010), potentially driven by cold flows (Khochfar & Silk 2009). This phase would be largely completed by redshift  $z \sim 2$ , by which time the most massive galaxies have formed their stars, and galaxies of lower mass and in less dense environments become the primary sources of star formation for the remaining 10 Gyr.

## 6 DISCUSSION AND CONCLUSIONS

Our integral-field spectroscopy provides a unique opportunity to study the mass-scaling properties of stellar populations in early-type galaxies with high accuracy, by virtue of minimising the effects of aperture size and slit losses. For a large range of galaxy masses and apparent sizes, we implicitly include the effects of line-strength gradients in our measurements of population properties by matching scaled apertures with effectively no slit loss effects. Our study also uses  $\sigma_e$ : the velocity dispersion integrated within  $R_e$ , which is empirically close to the true second velocity moment  $\langle v_{\text{los}}^2 \rangle_{\infty}$  that appears in the scalar virial theorem (Binney & Tremaine 2008), and is directly proportional to mass. Moreover, the IFU data allows the dynamical mass to be tightly constrained by fitting general dynamical models that account for non-homology and dark matter. Our sample selection is also based on a classical prescription of determining galaxy type, based on the presence of spiral arms and/or a significant dust disk. In this way, we do not exclude galaxies that have ongoing star-formation on anything other than morphological grounds. Finally, we have a large collection of high quality measurements on individual galaxies, allowing the intrinsic scatter of our relations to be accurately assessed.

From consideration of the classic Fundamental Plane, in Cappellari et al. (2013a) we showed that the intrinsic scatter and tilt of this relation is driven by stellar population variations, including the stellar initial mass function, and that accounting for these effects in the stellar mass-to-light ratio brings galaxy masses into very close agreement with the scalar virial relation  $M_{\text{JAM}} \propto \sigma_e^2 R_e^{\text{maj}}$ . Here we explored the distribution of stellar population parameters on the thin plane of  $(M_{\text{JAM}}, \sigma_e, R_e^{\text{maj}})$  - or the ‘Mass Plane’ - and show that there are indeed smooth variations of age, metallicity and abundance ratio within this plane. By consideration of empirical star formation histories derived from spectral fitting, we confirm that this reflects trends in



**Figure 18.** Cumulative star formation history for Virgo members (thick hatched regions) compared to outside of Virgo (solid lines) for galaxies averaged in bins of dynamical mass. Line colours show the range of dynamical masses considered, as indicated.

the mass-weighted properties and star formation timescales across the Mass Plane. At fixed mass, galaxies with smaller sizes are on average older, more metal rich, and more alpha-enhanced than their larger counterparts, showing that they formed earlier and in a shorter time. These trends tie directly to a higher mass normalisation of the IMF in these more compact objects (Cappellari et al. 2012; McDermid et al. 2014), as well as a larger bulge fraction and lower molecular gas fractions (Cappellari et al. 2013b).

Why are today’s massive galaxies older, more metal rich, and more enhanced in alpha elements compared to their lower-mass counterparts? There seem to be two processes at work. All of today’s early-type galaxies share an early period of intense star formation, making the systems compact, metal rich and alpha enhanced. This is consistent with the picture of Khochfar & Silk (2006a), who argue that high levels of dissipation at high redshift leads to more compact galaxies, with consequently older ages and enhanced abundance ratios. Lower-mass systems can continue to accrete gas from the inter-galactic medium, thereby increasing their apparent size and forming new stars, resulting in younger average ages when viewed today, with lower average metallicities. Dissipation of accreted gas results in an extended and regular stellar rotation field in these objects, which remains in tact in the absence of any major merger events, producing a ‘regular’ or ‘fast’ rotator (Khochfar et al. 2011; Naab et al. 2014). In the case where a major merger does occur, depending on the orbital parameters of the interaction, the

resulting system can remain as a ‘fast’ rotator, or in the case of misaligned spin vectors of the progenitors, give rise to a ‘two-sigma’ galaxy with a significant fraction of stars on counter-rotating orbits (Bois et al. 2011).

Massive systems, in contrast, quench early, and thereafter predominantly accrete stars (Khochfar & Silk 2006; Oser et al. 2010). Most of these stars have also formed early, but in smaller systems with lower metallicity. Via this stellar accretion, these galaxies will also increase in size, but remain old on average. The accreted low mass systems are dispersed at a radius proportional to where the local potential approaches the binding energy of the accreted system (Villumsen 1983), resulting in a gradient of metallicity, decreasing to large radii (Hirschmann et al. 2013; Navarro-González et al. 2013). Moreover, the collisionless accretion also reduces the average angular momentum of the system, resulting in massive ‘non-rotating’ galaxies (Naab et al. 2014). Where neither accretion of gas nor stars is efficient, galaxies remain compact, with characteristic stellar populations that are relatively old, metal-rich and alpha-enhanced for their mass. We note that in both cases (gas accretion at lower mass, and stellar accretion at higher mass), the stellar metallicity must remain closely related to the local potential, as traced by the spatially-resolved relationship between stellar population parameters and the local escape velocity (Scott et al. 2013).

Cappellari et al. (2013b) discuss evolution in the Mass Plane at length, describing the processes by which objects

may evolve within that plane. In summary, the evolutionary picture described there is one of increasing bulge fraction (or decreasing disk fraction) as mass increases, bringing galaxies to higher  $\sigma_e$  and more compact sizes. This process may be driven by dissipative mergers, which encourage central star formation and increase the effective concentration of stellar mass. Above a characteristic mass of  $\log(M_{\text{JAM}})=11.2$ , the range of early-type galaxy sizes is dramatically reduced, together with the scatter in stellar population properties. The process driving mass growth in this regime is effective at increasing galaxy size without significantly changing the stellar population properties or velocity dispersion. Simulations of collisionless (or ‘dry’) mergers show behaviour that is largely consistent with this picture (Khochfar & Burkert 2003; Naab et al. 2009; Hilz et al. 2013), and given the low gas fractions associated with the oldest and/or most massive early-type galaxies (Young et al. 2011), are a suitable mechanism for driving the mass growth along the upper mass envelope of the mass-size plane. Moreover, we present evidence that the mean scaling relations for the metallicity in non-regularly rotating galaxies is offset to higher mass and velocity dispersion for a given metallicity, consistent with a major merger origin for these objects.

We also derived empirical, non-parameteric star formation histories (SFHs) for our sample, in order to derive mass-weighted mean ages and metallicities. As expected, the mass-weighted ages are significantly older than the SSP-equivalent ages, which are strongly biased by the presence of bright young stars in galaxies with recent star formation.

In addition to providing mass-weighted (instead of SSP-equivalent) stellar population properties, our derived star formation histories themselves provide unique insights into the formation of our sample galaxies. We find a very regular dependence of star formation history on present-day mass, consistent with similar results using the SDSS (Heavens et al. 2004; Panter et al. 2007) and at higher redshift (Juneau et al. 2005). Within bins of galaxy mass, the peak of star formation is, on average, within the first billion years following the big bang for all galaxies with  $M_{\text{JAM}} > 10^{10} M_{\odot}$ . This provides an alternative view to the picture of staggered Gaussian star formation that is implied assuming linear relations of SSP-equivalent parameters and abundance-ratios to derive formation timescales, as in Thomas et al. (2005, 2010), though the qualitative picture is similar.

Through consideration of the mean absolute star formation rate, we demonstrate that the classic ‘downsizing’ scenario of Cowie et al. (1996) applies to early-type galaxies, showing that less massive galaxies have more extended star formation histories, and as a result, the characteristic mass of the objects dominating the star formation rate decreases as the Universe evolves. This trend is driven by the increasingly rapid fall off in star formation for galaxies with higher masses.

Our most massive galaxies ( $11.5 < \log(M_{\text{JAM}}) < 12$ ) have formed 50% of their stars no more than 2 Gyr after the big bang ( $z \sim 3$ ). By comparison, the lowest-mass galaxies ( $9.5 < \log(M_{\text{JAM}}) < 10$ ) reach the same fraction after around 8 Gyr ( $z \sim 0.6$ ). This is in reasonable agreement with recent simulation results (e.g. Oser et al. 2010), where models are indeed characterized by early intense star formation fed by cold flows and subsequent assembly of the formed stars, with ongoing ‘in situ’ star formation more prevalent in lower-mass

galaxies. Although we cannot constrain *where* the stars have formed from our integrated measurements, it is certainly the case that the recent (and even ongoing: Shapiro et al. 2010; Davis et al. 2014) star formation is more present in the central regions of our lower-mass galaxies, supporting the view that this is ‘in situ’.

Finally, we explore the environmental dependence of our star formation histories, and find that, on average, galaxies in Virgo have an additional suppression of extended-duration star formation with respect to galaxies outside the cluster environment. As a result, comparing galaxies of similar mass, Virgo members build their stellar mass on a shorter timescale than non-members - an effect that is most pronounced at intermediate masses ( $10^{10} < M_{\text{JAM}} < 10^{11} M_{\odot}$ ). The highest mass ( $M_{\text{JAM}} > 10^{11.5} M_{\odot}$ ) galaxies are very alike in their star formation histories, forming stars on comparably short timescales (90% of stellar mass formed within 3 Gyr). This suggests that the star formation properties of these galaxies are relatively unaffected by their current-day environment. Interestingly, galaxies with the lowest masses ( $M_{\text{JAM}} < 10^{10} M_{\odot}$ ) also share very similar star formation histories, with protracted formation of stars until recent times. These lowest-mass Virgo objects are at large cluster-centric distances, suggesting they have only recently entered the dense cluster environment for the first time, and the effective suppression of star formation by the environment has not yet made an impact.

## ACKNOWLEDGEMENTS

We thank the referee, Daniel Thomas, for a constructive report, which helped improve this paper. RMcD would like to thank Inger Jørgensen for a careful reading of early drafts of this paper, and Ricardo Schiavon for numerous instructive discussions. RMcD and RLD gratefully acknowledge the hospitality of the ESO visitor programme during the preparation of this work. MC acknowledges support from a Royal Society University Research Fellowship. This work was supported by the rolling grants ‘Astrophysics at Oxford’ PP/E001114/1 and ST/H002456/1 and visitor grants PPA/V/S/2002/00553, PP/E001564/1 and ST/H504862/1 from the UK Research Councils. RLD acknowledges travel and computer grants from Christ Church, Oxford and support from the Royal Society in the form of a Wolfson Merit Award 502011.K502/jd. TN acknowledges support from the DFG Cluster of Excellence Origin and Structure of the Universe. MS acknowledges support from a STFC Advanced Fellowship ST/F009186/1. TAD acknowledges the support provided by an ESO fellowship. The research leading to these results has received funding from the European Community’s Seventh Framework Programme (/FP7/2007-2013/) under grant agreement No 229517. The authors acknowledge financial support from ESO. SK acknowledges support from the Royal Society Joint Projects Grant JP0869822. NS acknowledges support of Australian Research Council grant DP110103509.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## REFERENCES

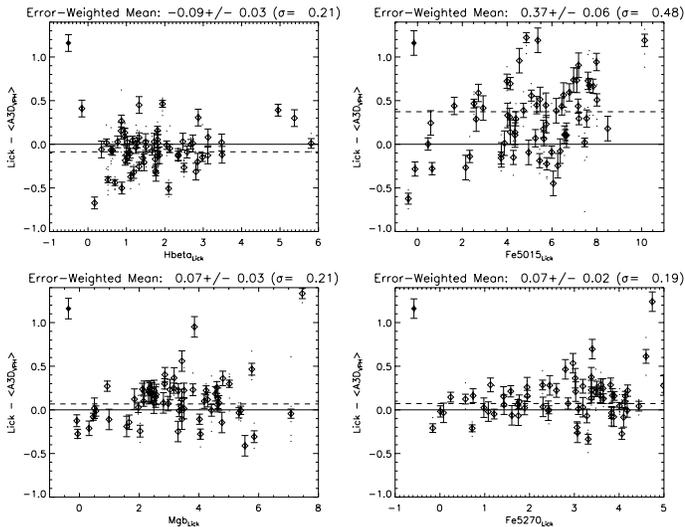
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, arXiv:1303.5062
- Auger, M. W., Treu, T., Bolton, A. S., et al. 2010, *ApJ*, 724, 511
- Bacon R., et al., 2001, *MNRAS*, 326, 23
- Balzano, V. A. 1983, *ApJ*, 268, 602
- Bernardi, M., Nichol, R. C., Sheth, R. K., Miller, C. J., & Brinkmann, J. 2006, *AJ*, 131, 1288
- Binney, J., & Tremaine, S. 2008, *Galactic Dynamics: Second Edition*, Princeton University Press
- Bois, M., Emsellem, E., Bournaud, F., et al. 2011, *MNRAS*, 416, 1654
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, *ApJ*, 754, 83
- Cappellari, M., & Copin, Y. 2003, *MNRAS*, 342, 345
- Cappellari, M., & Emsellem, E. 2004, *PASP*, 116, 138
- Cappellari, M., et al. 2006, *MNRAS*, 366, 1126
- Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, *MNRAS*, 413, 813
- Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011, *MNRAS*, 416, 1680
- Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2012, *Nature*, 484, 485
- Cappellari, M., Scott, N., Alatalo, K., et al. 2013, *MNRAS*, 432, 1709
- Cappellari, M., McDermid, R. M., Alatalo, K., et al. 2013, *MNRAS*, 432, 1862
- Cassisi, S., Castellani, M., & Castellani, V. 1997, *A&A*, 317, 108
- Cattaneo, A., Dekel, A., Faber, S. M., & Guiderdoni, B. 2008, *MNRAS*, 389, 567
- Chung, A., van Gorkom, J. H., Kenney, J. D. P., Crowl, H., & Vollmer, B. 2009, *AJ*, 138, 1741
- Cid Fernandes, R., Mateus, A., Sodré, L., Stasińska, G., & Gomes, J. M. 2005, *MNRAS*, 358, 363
- Cleveland W., & Devlin S., 1988, *Journal of the American Statistical Association*, 596
- Cowie, L. L., Songaila, A., Hu, E. M., & Cohen, J. G. 1996, *AJ*, 112, 839
- Crowl, H. H., & Kenney, J. D. P. 2008, *AJ*, 136, 1623
- Daddi, E., Dickinson, M., Morrison, G., et al. 2007, *ApJ*, 670, 156
- Davis, T. A., Alatalo, K., Sarzi, M., et al. 2011, *MNRAS*, 417, 882
- Davis, T. A., Alatalo, K., Bureau, M., et al. 2013, *MNRAS*, 429, 534
- Davis, T. A., Young, L. M., Crocker, A. F., et al. 2014, *MNRAS*, 444, 3427
- Dekel, A., & Birnboim, Y. 2006, *MNRAS*, 368, 2
- de La Rosa, I. G., La Barbera, F., Ferreras, I., & de Carvalho, R. R. 2011, *MNRAS*, 418, L74
- de Zeeuw, P. T., Bureau, M., Emsellem, E., et al. 2002, *MNRAS*, 329, 513
- Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, *A&A*, 468, 33
- Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, *MNRAS*, 379, 401
- Emsellem, E., et al. 2011, *MNRAS*, 414, 888
- Franx, M., van Dokkum, P. G., Schreiber, N. M. F., et al. 2008, *ApJ*, 688, 770
- Gallazzi, A., Charlot, S., Brinchmann, J., & White, S. D. M. 2006, *MNRAS*, 370, 1106
- Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, *A&AS*, 141, 371
- González, J. J. Ph.D. Thesis, University of California, Santa Cruz, 1993 (Source: Dissertation Abstracts International, Volume: 54-05, Section: B, page: 2551)
- Graves, G. J., & Schiavon, R. P. 2008, *ApJS*, 177, 446
- Graves, G. J., Faber, S. M., & Schiavon, R. P. 2009a, *ApJ*, 693, 486
- Graves, G. J., Faber, S. M., & Schiavon, R. P. 2009b, *ApJ*, 698, 1590
- Graves, G. J., & Faber, S. M. 2010, *ApJ*, 717, 803
- Gunawardhana, M. L. P., Hopkins, A. M., Bland-Hawthorn, J., et al. 2013, *MNRAS*, 433, 2764
- Gunn, J. E., & Gott, J. R., III 1972, *ApJ*, 176, 1
- Heavens, A., Panter, B., Jimenez, R., & Dunlop, J. 2004, *Nature*, 428, 625
- Hilz, M., Naab, T., Ostriker, J. P., et al. 2012, *MNRAS*, 425, 3119
- Hilz, M., Naab, T., & Ostriker, J. P. 2013, *MNRAS*, 429, 2924
- Hirschmann, M., Naab, T., Davé, R., et al. 2013, *MNRAS*, 436, 2929
- Hopkins, P. F., Bundy, K., Hernquist, L., Wuyts, S., & Cox, T. J. 2010, *MNRAS*, 401, 1099
- Johansson, J., Thomas, D., & Maraston, C. 2012, *MNRAS*, 421, 1908
- Jones, L. A., & Worthey, G., 1995, *ApJL*, 446, L31
- Jørgensen, I. 1999, *MNRAS*, 306, 607
- Juneau, S., Glazebrook, K., Crampton, D., et al. 2005, *ApJL*, 619, L135
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 54
- Kaviraj, S., Schawinski, K., Devriendt, J. E. G., et al. 2007, *ApJS*, 173, 619
- Khochfar, S., & Burkert, A. 2003, *ApJL*, 597, L117
- Khochfar, S., & Silk, J. 2006a, *ApJL*, 648, L21
- Khochfar, S., & Silk, J. 2006b, *MNRAS*, 370, 902
- Khochfar, S., & Ostriker, J. P. 2008, *ApJ*, 680, 54
- Khochfar, S., & Silk, J. 2009, *ApJL*, 700, L21
- Khochfar, S., Emsellem, E., Serra, P., et al. 2011, *MNRAS*,

- 417, 845  
 King D. L., 1985, Technical Note 31, Atmospheric Extinction at the Roque de los Muchachos Observatory, La Palma. The Royal Greenwich Observatory (RGO)
- Koopmann, R. A., & Kenney, J. D. P. 2004, *ApJ*, 613, 866
- Krajnović, D., Cappellari, M., de Zeeuw, P. T., & Copin, Y. 2006, *MNRAS*, 366, 787
- Krajnović, D., Emsellem, E., Cappellari, M., et al. 2011, *MNRAS*, 414, 2923
- Kuntschner, H., et al. 2006, *MNRAS*, 369, 497
- Kuntschner, H., Emsellem, E., Bacon, R., et al. 2010, *MNRAS*, 408, 97
- Landsman, W. B. 1993, *Astronomical Data Analysis Software and Systems II*, 52, 246
- Lauer, T. R., Ajhar, E. A., Byun, Y.-I., et al. 1995, *AJ*, 110, 2622
- McDermid, R. M., et al. 2006, *MNRAS*, 373, 906
- McDermid, R. M., Cappellari, M., Alatalo, K., et al. 2014, *ApJL*, 792, L37
- Naab, T., Johansson, P. H., & Ostriker, J. P. 2009, *ApJL*, 699, L178
- Naab, T., Oser, L., Emsellem, E., et al. 2014, *MNRAS*, 444, 3357
- Navarro-González, J., Ricciardelli, E., Quilis, V., & Vazdekis, A. 2013, *MNRAS*, 436, 3507
- Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, *ApJL*, 660, L43
- Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., & Burkert, A. 2010, *ApJ*, 725, 2312
- Panter, B., Jimenez, R., Heavens, A. F., & Charlot, S. 2007, *MNRAS*, 378, 1550
- Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, *ApJ*, 721, 193
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, Cambridge: University Press
- Proctor, R. N., Forbes, D. A., & Beasley, M. A. 2004, *MNRAS*, 355, 1327
- Sánchez-Blázquez, P., et al., 2006, *MNRAS*, 371, 703
- Sarzi M., et al., 2006, *MNRAS*, 366, 1151
- Scott, N., Cappellari, M., Davies, R. L., et al. 2013, *MNRAS*, 432, 1894
- Scott, N., Graham, A. W., & Schombert, J. 2013, *ApJ*, 768, 76
- Schawinski, K., Kaviraj, S., Khochfar, S., et al. 2007, *ApJS*, 173, 512
- Schiavon, R. P. 2007, *ApJS*, 171, 146
- Serra, P., & Trager, S. C. 2007, *MNRAS*, 374, 769
- Serra, P., Oosterloo, T., Morganti, R., et al. 2012, *MNRAS*, 422, 1835
- Shankar, F., Marulli, F., Bernardi, M., Dai, X., Hyde, J. B., & Sheth, R. K. 2010, *MNRAS*, 403, 117
- Shapiro, K. L., et al. 2010, *MNRAS*, 402, 2140
- Smith, R. J., et al. 2008, *MNRAS*, 386, L96
- Thomas, D., Maraston, C., & Bender, R. 2003, *MNRAS*, 339, 897
- Thomas, D., Maraston, C., & Korn, A. 2004, *MNRAS*, 351, L19
- Thomas, D., Maraston, C., Bender, R., & Mendes de Oliveira, C. 2005, *ApJ*, 621, 673
- Thomas, D., Maraston, C., Schawinski, K., Sarzi, M., & Silk, J. 2010, *MNRAS*, 404, 1775
- Thomas, D., Maraston, C., & Johansson, J. 2011, *MNRAS*, 412, 2183
- Trager, S. C., Worthey, G., Faber, S. M., Burstein, D., & Gonzalez, J. J. 1998, *ApJS*, 116, 1
- Trager, S. C., Faber, S. M., Worthey, G., & González, J. J. 2000, *AJ*, 119, 1645
- Tran, K.-V. H., Papovich, C., Saintonge, A., et al. 2010, *ApJL*, 719, L126
- van der Wel, A., Bell, E. F., van den Bosch, F. C., Gallazzi, A., & Rix, H.-W. 2009, *ApJ*, 698, 1232
- van Dokkum, P. G., & Conroy, C. 2010, *Nature*, 468, 940
- Vazdekis, A., Ricciardelli, E., Cenarro, A. J., et al. 2012, *MNRAS*, 424, 157
- Véron-Cetty, M.-P., & Véron, P. 2010, *A&A*, 518, A10
- Villumsen, J. V. 1983, *MNRAS*, 204, 219
- Wake, D. A., van Dokkum, P. G., & Franx, M. 2012, *ApJL*, 751, L44
- Whitaker, K. E., van Dokkum, P. G., Brammer, G., & Franx, M. 2012, *ApJL*, 754, L29
- Worthey G., Faber S. M., Gonzalez J. J., Burstein D., 1994, *ApJS*, 94, 687
- York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, *AJ*, 120, 1579
- Young, L. M., Bureau, M., & Cappellari, M. 2008, *ApJ*, 676, 317
- Young, L. M., Bureau, M., Davis, T. A., et al. 2011, *MNRAS*, 414, 940
- Young, L. M., Scott, N., Serra, P., et al. 2014, *MNRAS*, 444, 3408

## APPENDIX A: LINE-STRENGTH CALIBRATION AND ERRORS FROM STANDARD STARS

For observations made before the installation of the VPH grating in SAURON in 2004, we refer the reader to Kuntschner et al. (2006) for details. An extensive set of line-strength standard star observations were also obtained post-VPH installation to provide a calibration to the Lick system of index measurements, and to monitor run-to-run and inter-run variations. The stars were chosen to overlap with several well-known stellar libraries, including Worthey et al. (1994), Jones & Worthey (1995) and MILES (Sánchez-Blázquez et al. 2006). In total, 67 stars were observed, with several repeat observations, giving a total number of 160 stellar spectra. Wherever possible, velocity measurements for the stars were derived from the matching MILES library spectrum, otherwise the closest spectral type from the MILES library was used. This minimized template mismatch effects from introducing large velocity errors, and subsequent errors in the index measurements. The stars do not suffer from the field coverage issue mentioned previously, so we include it here for reference.

Figure A1 presents the difference between our measurements and the standard reference values of the Lick/IDS system (Worthey et al. 1994). Since some stars have several repeat measurements, and others do not, we use an error-weighted mean to derive the average offsets. Six stars were observed four times or more, upon which we base estimates of our intrinsic measurement errors in Figure A1 and given in Table A2. This characteristic error was adopted for each individual measurement, and repeat measurements are averaged,

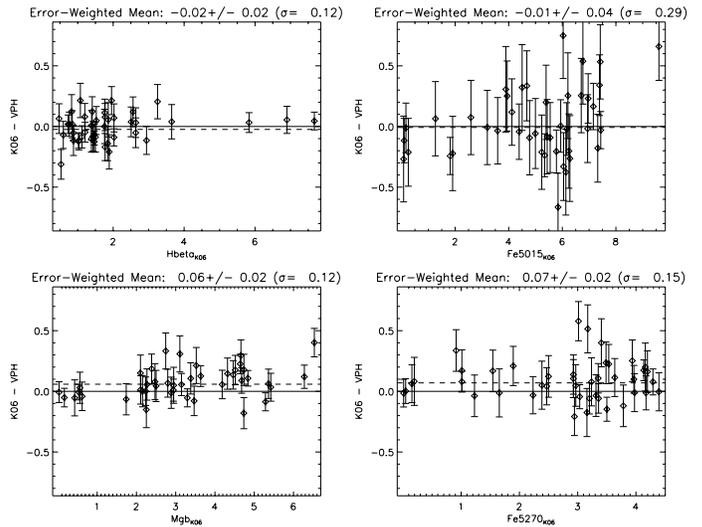


**Figure A1.** Comparison of Lick indices measured from the 67 standard stars overlapping between the ATLAS<sup>3D</sup> survey and those from the Lick/IDS system (Worthey et al. 1994). The derived biweight-mean offsets (shown as a dashed line) and errors on the mean are given in Table A1. Dots show individual measurements; diamonds show individual and (where available) combined measurements. A representative error bar is plotted in the upper left of each panel showing the empirical scatter in repeat measurements from six stars having four or more observations over the four observing runs. This value is also tabulated in Table A1, and shows that the scatter in this comparison is largely driven by scatter in the original Lick measurements, which have an average uncertainty around twice as large.

with the characteristic error scaled by  $(n - 1)^{-1/2}$ , where  $n$  is the number of repeat measurements. We effectively ignore the intrinsic errors of the reference values themselves, which in fact are the dominant source of scatter (compare the errors in Table A2 with the scatter in Figure A1, quantified by  $\sigma$  in the panel titles). As with all figures in this section, we employ outlier-resistant ‘robust’ estimators of the standard deviation.

The average offsets from our post-VPH standard stars are consistent within the  $1\sigma$  uncertainties of the pre-VPH ‘stars only’ values of Kuntschner et al. (2006, column 2 of their table 3), with the exception of  $Mg\ b$ , for which we derive a larger offset, indicating slightly weaker index values with the VPH grating. To further compare the pre- and post-VPH consistency of our SAURON data, we compare measurements of the *same* stars measured with the two different gratings. Figure A2 shows this comparison, where the values have been averaged wherever possible. The mean offsets in this comparison are fully consistent with the differences found in the comparison with the Lick/IDS values, indicating that the small differences between the two gratings are real. We therefore apply *different* offsets for the pre- and post-VPH data. The final applied offsets are given in Table A1. We also derive an estimate of our uncertainties from this differential comparison, assuming that the scatter comes from equally-sized errors from both observations. The inferred values are also given in Table A2, and are in good agreement with the estimates from repeated stars.

Whilst these error estimates are inferred from observa-



**Figure A2.** Comparison of Lick indices measured from the 43 stars observed with SAURON before (from Kuntschner et al. 2006) and after the installation of the VPH grating. Repeat measurements were averaged in both data sets, and corresponding errors combined. The plotted errors assume the errors come equally from both data sets, giving  $\sigma_{\text{Overlap}}$  in Table A2. Dashed line indicates the biweight mean.

**Table A1.** Estimate of Lick offsets derived from all stars.

Index	Pre-VPH Offset	Post-VPH Offset
H $\beta$	$-0.13 \pm 0.05 \text{ \AA}$	$-0.09 \pm 0.03 \text{ \AA}$
Fe5015	$+0.28 \pm 0.05 \text{ \AA}$	$+0.37 \pm 0.06 \text{ \AA}$
Mg <i>b</i>	$+0.00 \pm 0.05 \text{ \AA}$	$+0.07 \pm 0.03 \text{ \AA}$
Fe5270	$+0.00 \pm 0.05 \text{ \AA}$	$+0.07 \pm 0.02 \text{ \AA}$

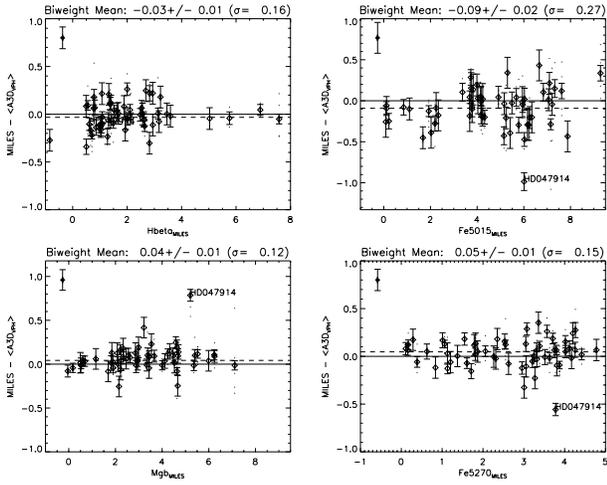
Note: For each index, the biweight mean and dispersion estimate was used to reduce the influence of outlying values. The error on the mean offset derived from the  $N$  observations is given as the dispersion scaled by  $1/\sqrt{N}$ . Pre-VPH offsets are taken from (Kuntschner et al. 2006).

tions spanning several years, they are still ‘internal’ error estimates, coming from the same instrument and similar data reduction etc. As a way to check our ‘external’ errors, i.e. how we may compare with others, we present in Figure A3 a comparison of the Lick indices derived from our measurements (post-VPH) and from the MILES spectra for the same stars. The scatter is consistent with our internal comparison, suggesting that our error estimates are reasonable. Moreover, the mean offset and scatter are comparable to those found between targeted stellar libraries (e.g. see Sánchez-Blázquez et al. 2006).

## APPENDIX B: COMPARISON WITH OTHER AUTHORS

### B1 Comparison of line-strengths

The dominant uncertainty in line strength measurements from high signal-to-noise ratio data such as is available for



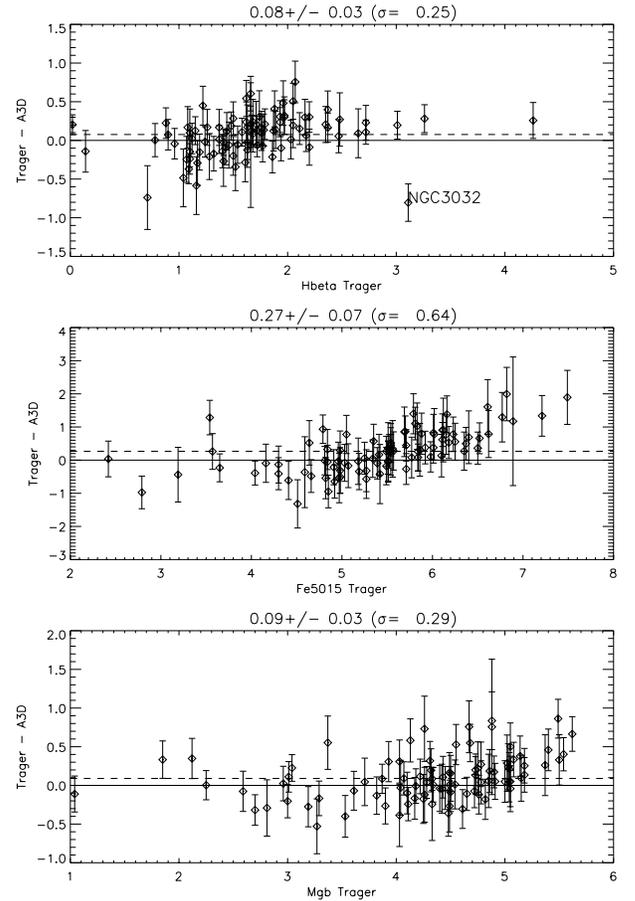
**Figure A3.** Comparison of Lick indices measured from the 57 standard stars in common with the MILES library of Sánchez-Blázquez et al. (2006). Small dots show all 147 individual measurements, and diamonds show individual and averaged values where repeat measurements are available. The mean, error on the mean and dispersion are given in the plot titles, which are computed from the *individual* values to provide a measure of our uncertainties. The representative error bar (filled symbol) assumes the scatter comes equally from both data sets.

**Table A2.** Error estimates derived from the standard star measurements.  $\sigma_{\text{Repeat}}$  is derived from the mean standard deviation of six stars with four or more measurements.  $\sigma_{\text{Overlap}}$  is derived by assuming constant and equal errors between pre- and post-VPH observations of stars in common to both set-ups.  $\sigma_{\text{MILES}}$  is derived by comparing our Lick indices with those from MILES spectra for the same stars, and assuming the scatter comes from both data sets equally. The last column gives the average deviation, which we take as our minimum uncertainty in each index.

Index	$\sigma_{\text{Repeat}}$	$\sigma_{\text{Overlap}}$	$\sigma_{\text{MILES}}$	$\bar{\sigma}$
H $\beta$	0.10 Å	0.08 Å	0.11 Å	0.10 Å
Fe5015	0.14 Å	0.21 Å	0.19 Å	0.18 Å
Mgb	0.12 Å	0.08 Å	0.08 Å	0.09 Å
Fe5270	0.11 Å	0.11 Å	0.11 Å	0.11 Å

nearby galaxies are the *systematic* errors from data calibration, rather than statistical errors from noise. We therefore choose from the literature two studies that have reasonable overlap in sample galaxies, namely the extensive Lick/IDS sample of Trager et al. (1998) (90 galaxies in common), and the higher-quality data from González (1993) (20 galaxies in common).

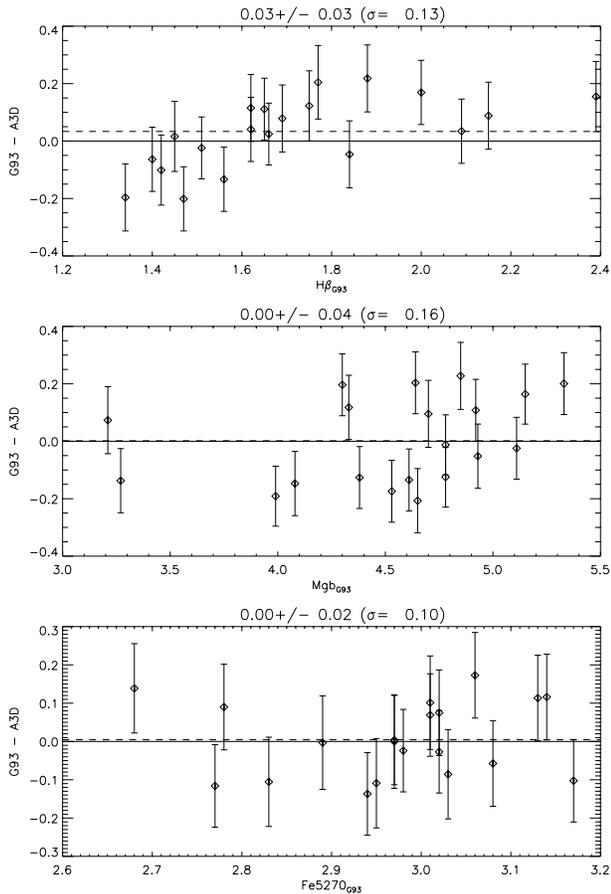
For the Trager et al. (1998) data, as per Kuntschner et al. (2006) we extract a circular aperture with the same area as their  $1.4'' \times 4''$  (i.e. with a radius of  $1.335''$ ), since we do not know the position angle of the slits used. The Trager et al. (1998) data is also not corrected for emission, which is a major difference to our data. We therefore measure the comparison values with no correction for emission. Figure B1 presents the comparison, with robust mean, er-



**Figure B1.** Comparison with the Lick indices of Trager et al. (1998, T98). No correction for emission is made in either data set. The T98 data were measured in  $1.4'' \times 4''$  apertures of unknown orientation. The ATLAS<sup>3D</sup> data (A3D) are measured in circles of equivalent area (i.e. of radius  $1.335''$ ). Errors show the quadrature sum of both data set errors. Plot titles show the robust mean (also shown as a dashed line), error on the mean, and robust dispersion (in parentheses).

ror on the mean, and dispersion (in parenthesis) shown in the title. Observational errors for both data sets have been added in quadrature. The comparison is reasonable, though systematic differences can be seen, in particular an apparent trend with increasing Fe5015. Corrections for velocity dispersion are comparable between the two studies to  $\sim 1\%$ . The outlier in the H $\beta$  comparison, NGC3032, demonstrates the sensitivity to such comparisons to PSF and aperture differences. This particular galaxy has an unresolved blue nucleus where H $\beta$  strengthens considerably (McDermid et al. 2006), making the PSF a dominant uncertainty.

Figure B2 presents the comparison with the  $R_e/8$  aperture values from González (1993), with titles and errors as per Figure B1. In this work, careful consideration of the aperture definitions was made, and the translation between the slit radial length and the corresponding circular aperture size are conveniently tabulated there (his table 4.7). Using this radius, we extract spectra within circles, and compare the measured line-strengths. In this case, a correction for H $\beta$  emission was determined from the measured [OIII] emission strength, as described earlier. We therefore apply the same

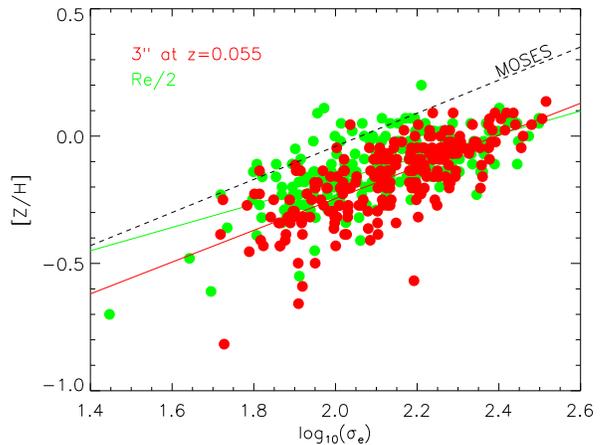


**Figure B2.** Comparison with the indices of González (1993, G93). In both cases,  $H\beta$  is corrected for emission by adding 0.7 times the  $[OIII]$  emission. The G93 values were measured from one or two long-slit position angles. The A3D values were measured on circular apertures using the radii given in table 4.7 of G93, which give the luminosity-weighted circular aperture radius equivalent to the long-slit apertures. Errors, dashed lines and plot titles are as per Figure B1.

correction procedure, adding  $0.7 \times [OIII]$  as measured using the GANDALF code to the measured  $H\beta$  index, with no additional emission correction.

Instead of comparing Fe5015 (which is not in the González table), we compare Fe5270, which has the advantage of being free from emission corrections. To avoid the filter truncation effects described earlier, we convert our measured Fe5270<sub>S</sub> values to the Lick Fe5270 index following Kuntschner et al. (2006):  $Fe5270 = 1.28 \times Fe5270_S + 0.03$ .

The comparison of all three comparable indices is rather good, with the mean difference in all three indices consistent with zero within the scatter. The effective reduced  $\chi^2$  values are also close to unity, implying that the assumed errors are representative. The higher degree of consistency with González (1993) over that with Trager et al. (1998) likely reflects more the more accurate aperture matching with the latter study, rather than some intrinsic difference in the measurement quality (with respect to the errors).



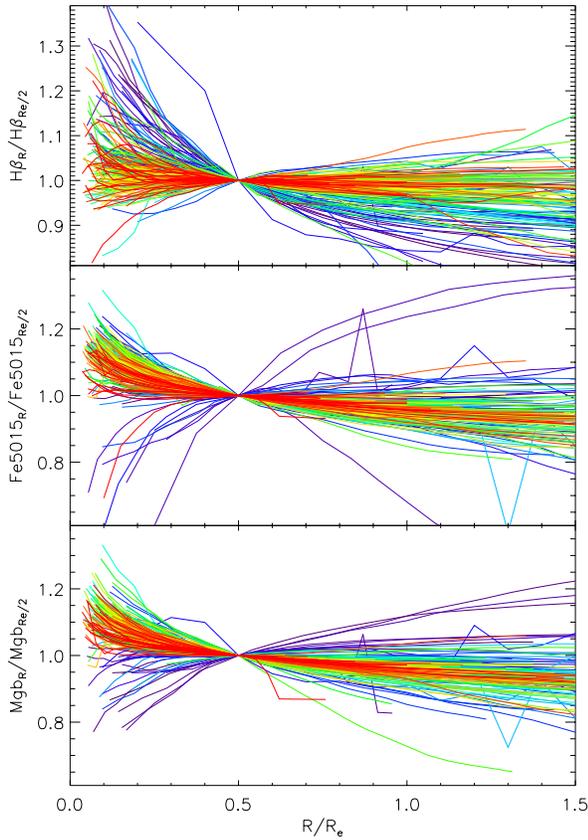
**Figure B3.** Simulation of the effect of a fixed 3'' aperture on the slope of the relation between metallicity and velocity dispersion. Green points show the  $R_e/2$  values presented in Figure 8. The red points were measured assuming a fixed aperture, scaled to mimic the redshift of the MOSES sample Thomas et al. (2010). The effect of the fixed aperture is to include more of the lower metallicity regions in the outskirts of smaller, lower dispersion galaxies, thus steepening the slope compared to our values. Red and green lines show robust fits to the points of corresponding colour. The black dashed line is the relation from Thomas et al. (2010). The slopes match within the uncertainties, but the offsets differ by 0.2 dex.

## B2 Comparison with MOSES sample

In Section 4.1 we compared our relations of SSP parameters and velocity dispersion to those of Thomas et al. (2010), and found that their metallicity relation was significantly steeper than ours. We can understand the difference in the metallicity trend by considering that no corrections were made by Thomas et al. (2010) for the fixed 3'' diameter fiber used. In Figure B3 we simulate the effect of measuring metallicity for our sample with a 3'' fiber at a redshift matching the mean redshift of the MOSES study (0.055). We overplot robust linear fits, excluding objects younger than 3.2 Gyr, which accounts for the redshift difference of our sample. The slopes agree within the errors, but a 0.2 dex offset in the relations remains. As in Kuntschner et al. (2006), we find that applying the models of either Schiavon (2007) or Thomas et al. (2005) yield very similar results, thus cannot explain the large metallicity offset (indeed, the offset becomes even slightly larger). Most likely it arises from the different combination of indices used by the two studies, but it is beyond the scope of this paper to explore this further.

## APPENDIX C: GENERIC APERTURE CORRECTIONS FROM THE ATLAS<sup>3D</sup> SAMPLE

Since we have the privilege of extensive two-dimensional information for a complete sample of galaxies, it is useful to investigate aperture corrections more generally, for use by studies without access to this spatial information. Figure C1 presents index growth curves as a function of  $R_e$  for all objects with maximal radial coverage of  $R_e/2$  or greater. The

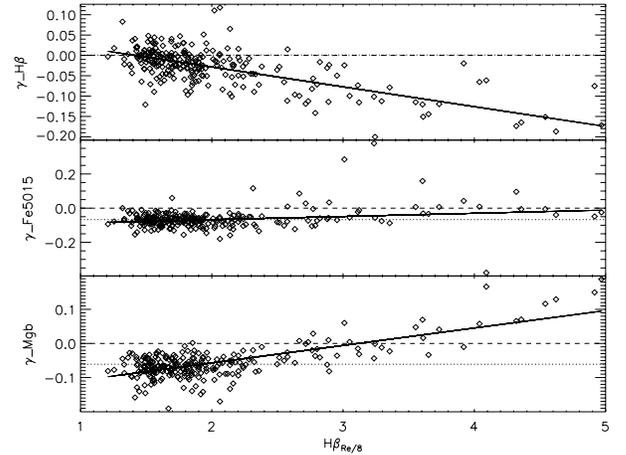


**Figure C1.** Curves of growth for each of our line indices, measured within circular apertures of increasing radius, plotted as a function of the effective radius. The curves are normalized by the aperture value at  $R_e/2$ . Colours indicate the velocity dispersion within the normalization radius, ordered low (blue) to high (red). The central  $1.2''$  are excluded, as are aperture sizes that are less than 85% filled with spectra.

curves are normalized by the value at  $R_e/2$ . The colour of the curves are ordered in velocity dispersion from low (blue) to high (red). We do not include Fe5270 in this analysis due to the limited field coverage issue mentioned above.

From this one can see differences between the red and blue curves, especially in the case of  $H\beta$  and  $Mg b$ . Comparing this figure to the almost equivalent Figure 16 of Kuntschner et al. (2006), we see that there are many more galaxies with rising central  $H\beta$  strengths in our larger sample, and these galaxies tend to have lower velocity dispersions, although for a given velocity dispersion, there is a wide variety of slopes.

We attribute this behaviour to the presence of a centrally-concentrated population of young stars that give rise to a steep gradient in  $H\beta$ . If this is not taken into account in the aperture correction, the strength of  $H\beta$  may be over-estimated. Figure C2 plots the curve-of-growth slopes (measured from robust line fits to the coloured curves in Figure C1) for all galaxies as a function of the  $H\beta$  line-strength measured within  $R_e/8$ , for each of our three indices. There is clearly a relationship between the slopes of  $H\beta$  and  $Mg b$  with the  $R_e/8$   $H\beta$  line-strength, presumably reflecting the known



**Figure C2.** The power-law slope values for fits to all the galaxies in Figure C1, plotted as a function of  $H\beta$  line-strength within an  $R_e/8$  aperture. This illustrates how the size of aperture correction changes as a function of the central  $H\beta$  line-strength, due to the typically centrally-concentrated nature of the young stars. Dotted lines show the fixed corrections from Kuntschner et al. (2006).

**Table C1.** Coefficients of linear fits to aperture correction factors as a function of central  $H\beta$  strength, for use with equation C2.

Index	$a$	$b$
$H\beta$	-0.052	0.069
Fe5015	0.021	-0.105
$Mg b$	0.047	-0.151

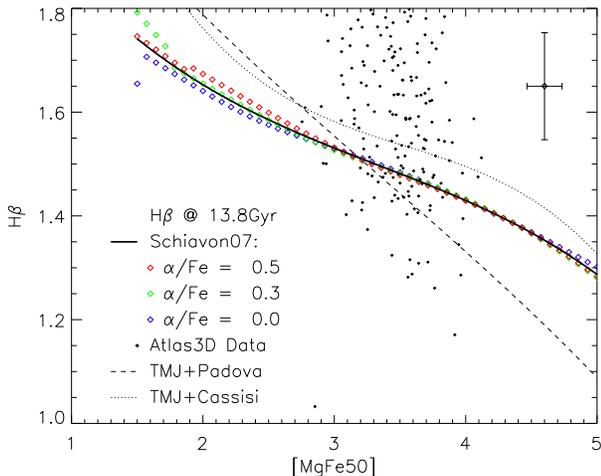
age-sensitivity of these two lines. Thus, instead of providing a single aperture correction value for our three indices, we construct robust linear fits to the data in Figure C2, giving the corrections as a function of the  $H\beta$  lines-strength within an  $R_e/8$  aperture. Table C1 gives the coefficients of these linear fits, which can be used to derive the aperture correction for each index such that:

$$\text{Index}_R / \text{Index}_{norm} = (R/R_{norm})^\gamma \quad (\text{C1})$$

where:

$$\gamma = aH\beta_{Re/8} + b \quad (\text{C2})$$

and  $a$  and  $b$  are given in Table C1. The  $H\beta$  dependence is important for central  $H\beta$  values larger than around  $2 \text{ \AA}$ , and neglecting the effect in this regime will lead to over-estimated Balmer line strengths, and under-estimated  $Mg b$ . Below this threshold, however, the dependence is marginal, and adopting a constant correction (such as that of Kuntschner et al. (2006), shown as dotted lines) gives similar results.



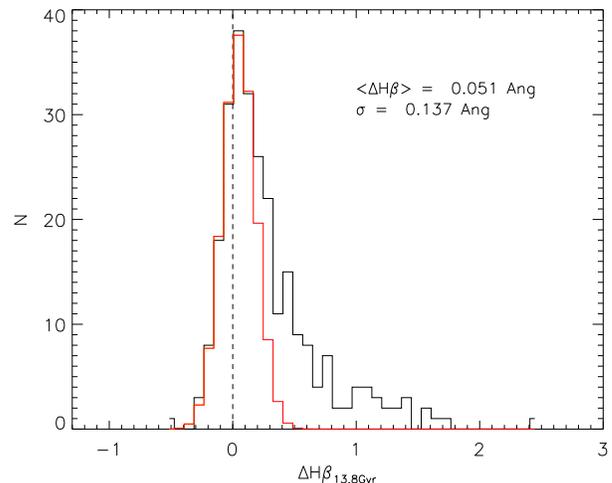
**Figure D1.** Large open diamonds show the Schiavon (2007) SSP model index predictions for a fixed age of 13.8 Gyr, as a function of metallicity (increasing left to right) and abundance ratio (indicated by colour). The solid black line is a cubic fit to these predictions which describes the  $H\beta$  at this age as a function of  $[MgFe50]$ . This line is used to find the offset in  $H\beta$  between the measured points (small diamonds) and the canonical ages of the Universe, as shown in Figure D2. For comparison, we also show the equivalent predictions from the models of Thomas et al. (2011), using the stellar evolutionary tracks of Cassisi et al. (1997, dotted line) and the Padova group (Girardi et al. 2000, dashed line). The mean measurement error is also indicated in the upper-right corner.

#### APPENDIX D: GALAXIES OLDER THAN THE UNIVERSE?

A number of galaxies show  $H\beta$  absorption values that imply ages older than the currently accepted age of the Universe ( $13.798 \pm 0.037$  Gyr; Planck Collaboration, Ade et al. 2013). Here we ask the question: Are these objects consistent with this upper limit given our observational uncertainties? Even though the SSP models make predictions up to 17 Gyr, some data points still lie just outside the model grid. For this reason, we determine instead the differences of the measured  $H\beta$  to the  $H\beta$  predicted by the stellar population models for the age of the universe.

We note that  $H\beta$  and age are not completely orthogonal, and there is a small but significant metallicity dependence. Abundance ratio dependences are small for the Schiavon (2007) models used here. Figure D1 shows the line-strength predictions at the WMAP age for all metallicities and abundance ratios in the model grid (coloured diamonds). The solid line shows a cubic fit to the model points, giving the mean predicted  $H\beta$  value as a function of  $[MgFe50]$ . For comparison, we also show equivalent cubic fits derived from the “TMJ” models of Thomas et al. (2011) using the stellar evolutionary tracks of Cassisi et al. (1997, “TMJ+Cassisi”) and the Padova group (Girardi et al. 2000, “TMJ+Padova”). We shift all model predictions to the Lick system using the offsets from Schiavon (2007). This comparison shows that the data extend below predictions for the fiducial age for all models, though the issue is somewhat less for the TMJ+Padova models, and worse for the TMJ+Cassisi models.

Figure D2 shows the distribution of differences between



**Figure D2.** Histogram showing the distribution of differences between the measured  $H\beta$  values and the model predictions for the canonical age of the Universe (13.8 Gyr). The red line shows a Gaussian fitted to the lower portion of the distribution (fitted region is indicated), with a mean close to zero, and a dispersion similar to the observational error in  $H\beta$ . From this we conclude that the galaxies appearing apparently older than the Universe are consistent with our measurement uncertainties.

the observed  $H\beta$  values and the prediction for 13.8 Gyr from the Schiavon (2007) models used in this paper. Galaxies with negative differences have implied ages older than the Universe. We then fit a Gaussian to the lower portion of this distribution, and find that the tail of negative points can be described by objects with a slightly positive mean offset (i.e. younger than 13.8 Gyr), and a dispersion of  $0.137 \text{ \AA}$ . This is quite consistent with our estimate of the  $H\beta$  measurement error, which is assumed to have a minimum of  $0.1 \text{ \AA}$ . Part of the additional observed dispersion may arise from uncertainties in the emission-line correction, in the form of emission lines below our detection limit, which is typically a few hundredths of an Angstrom at the signal-to-noise ratios considered here (Sarzi et al. 2006). We therefore conclude that the distribution of ages we measure are consistent with the fiducial age of the Universe after taking into account our measurement uncertainties, as well as systematic uncertainties in the models.