The galaxy morphology–density relation in the EAGLE simulation

Joel Pfeffer [ORCID],[1]⋆ Mitchell K. Cavanagh [ORCID],[1] Kenji Bekki,[1] Warrick J. Couch,[2] Michael J. Drinkwater [ORCID],[3] Duncan A. Forbes[2] and Bärbel S. Koribalski [ORCID][4,5]

[1]*International Centre for Radio Astronomy Research (ICRAR), M468, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia*
[2]*Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Hawthorn VIC 3122, Australia*
[3]*School of Mathematics and Physics, University of Queensland, Brisbane, Queensland 4072, Australia*
[4]*Australia Telescope National Facility, CSIRO Astronomy and Space Science, P.O. Box 76, Epping, NSW 1710, Australia*
[5]*School of Science, Western Sydney University, Locked Bag 1797, Penrith, NSW 2751, Australia*

ABSTRACT

The optical morphology of galaxies is strongly related to galactic environment, with the fraction of early-type galaxies increasing with local galaxy density. In this work, we present the first analysis of the galaxy morphology–density relation in a cosmological hydrodynamical simulation. We use a convolutional neural network, trained on observed galaxies, to perform visual morphological classification of galaxies with stellar masses $M_* > 10^{10} \, \mathrm{M_\odot}$ in the EAGLE simulation into elliptical, lenticular and late-type (spiral/irregular) classes. We find that EAGLE reproduces both the galaxy morphology–density and morphology–mass relations. Using the simulations, we find three key processes that result in the observed morphology–density relation: (i) transformation of disc-dominated galaxies from late-type (spiral) to lenticular galaxies through gas stripping in high-density environments, (ii) formation of lenticular galaxies by merger-induced black hole feedback in low-density environments, and (iii) an increasing fraction of high-mass galaxies, which are more often elliptical galaxies, at higher galactic densities.

Key words: methods: numerical – galaxies: evolution – galaxies: formation – galaxies: structure.

1 INTRODUCTION

The optical morphology of galaxies contains information about both their shape and recent star formation, and thus may offer clues into galaxy formation and evolution processes. One of the key findings of such studies is that morphology depends on local galaxy density, with the fraction of elliptical (E) and lenticular (S0) galaxies increasing with local density and a corresponding decrease in the fraction of late-type (spiral) galaxies, i.e. the galaxy morphology–density relation (e.g. Oemler 1974; Davis & Geller 1976; Dressler 1980; Postman & Geller 1984). Alternatively, the morphology–density relation may be viewed as a morphology–radius relation in clusters (Dressler 1980; Whitmore & Gilmore 1991; Whitmore, Gilmore & Jones 1993; Fasano et al. 2015).

Given S0s have stellar discs it is often suggested S0s are formed as 'stripped spirals', either by direct stripping of the interstellar medium or 'starvation' following stripping of the circumgalactic medium (e.g. Spitzer & Baade 1951; Gunn & Gott 1972; Larson, Tinsley & Caldwell 1980; Bekki, Couch & Shioya 2002; Bekki & Couch 2011). Such a mechanism may explain the morphology–density relation for both S0 and spiral galaxies, as stripping mechanisms will naturally be environmentally dependent. However, simple disc fading does not explain the larger bulge fractions of S0s compared to spiral galaxies (Faber & Gallagher 1976; Dressler 1980; Croom et al. 2021). This

morphological transformation could potentially be explained instead by 'galaxy harassment' (i.e. multiple high-speed encounters between galaxies, Moore et al. 1996; Moore, Lake & Katz 1998; Moore et al. 1999), though galaxies observed in intermediate-redshift ($z \sim 0.3$) clusters do not show evidence of distortion that might be expected from such a mechanism (Couch et al. 1998). Alternatively, concentrated star formation in the late stages of gas stripping could result in bulge growth (Bekki & Couch 2011). A stripped-spiral origin (or other environmentally dependent processes) also cannot explain the origin of field or isolated S0 galaxies, which account for ∼20 per cent of the population, though may explain the increase above the field fraction (Postman & Geller 1984; van der Wel et al. 2010).

Alternatively, S0 galaxies may form from galaxy mergers (both major and minor, e.g. Toth & Ostriker 1992; Quinn, Hernquist & Fullagar 1993; Walker, Mihos & Hernquist 1996; Bekki 1998; Diaz et al. 2018; Dolfi et al. 2021). Mergers may explain many properties of S0s, such as their structure, dynamics and scaling relations (Eliche-Moral et al. 2012, 2013; Borlaff et al. 2014; Querejeta et al. 2015; Tapia et al. 2017). In particular, galaxy mergers may provide an origin for field S0s, but mergers become less likely once a galaxy becomes a satellite in a cluster (Ghigna et al. 1998; see also Section 3.4).

These considerations suggest S0 galaxies may not have a single origin. Quiescent galaxies appear to be significantly flatter in denser environments (van der Wel et al. 2010), and recent work has established that S0s in clusters are more rotationally supported than

⋆ E-mail: joel.pfeffer@uwa.edu.au

those in the field (Coccato et al. 2020; Deeley et al. 2020). These observations are consistent with a merger origin for galaxies in low-density environments and an increasing importance of stripped spirals at higher densities, in agreement with the findings of S0 galaxy formation in cosmological hydrodynamical simulations (Deeley et al. 2021).

The main scenario for the formation of elliptical galaxies is morphological transformation of galaxies through mergers (e.g. Toomre & Toomre 1972; Barnes 1992; Naab & Burkert 2003; Khochfar et al. 2011). Thus, formation through mergers may explain why the elliptical fraction rises steeply at the highest galaxy number densities (Dressler 1980), where galaxy mergers would be expected to be most frequent. The morphology–density relation also persists when only considering slow-rotating galaxies, which only account for one third of ellipticals (Cappellari et al. 2011; D'Eugenio et al. 2013; Houghton et al. 2013; van de Sande et al. 2021). However, the centres of galaxy clusters appear to host a constant fraction of elliptical galaxies, regardless of density (Whitmore et al. 1993). This results in a morphology–density relation for ellipticals that rises at higher densities for clusters with higher average density (Houghton 2015). van der Wel et al. (2010) suggest that these trends are in fact driven by mass, with massive galaxies, which are more often ellipticals, occupying the centres of galaxy groups/clusters.

Cosmological hydrodynamical simulations have now increased sufficiently in realism such that they can reproduce many properties of the evolving galaxy population (see reviews by Somerville & Davé 2015; Vogelsberger et al. 2020). Beginning with initial conditions consistent with a Λ Cold Dark Matter universe, such simulations model the growth and evolution of galaxies in a cosmological context context by including the effects of gravity, cooling of gas, formation of stars and energy feedback from star formation and supermassive black holes. By modelling and following the evolution of galaxies through time, simulations provide important tools to interpret the wealth of information provided by observational surveys.

To date, most work on the morphology of galaxies in hydrodynamical simulations focuses on morphological indicators (shape, kinematic or other non-parametric diagnostics) rather than visual morphology (e.g. Correa et al. 2017; Rodriguez-Gomez et al. 2019; Tacchella et al. 2019; Thob et al. 2019; Trayford et al. 2019; Bignone et al. 2020, but see Huertas-Company et al. 2019). This body of work has shown that modern hydrodynamical simulations produce galaxies with morphological diagnostics and correlations between morphology, colour/star-formation rate and galaxy mass broadly in agreement with observed galaxies. Though physically motivated (and less time-demanding or subjective than visual classification), the various morphological indicators are generally not straightforward to compare with observed morphologies (Hubble type).

In this work, we use simulations from the 'Evolution and Assembly of GaLaxies and their Environments' (EAGLE) project (Crain et al. 2015; Schaye et al. 2015), combined with a convolutional neural network (CNN) trained on observed galaxies to perform morphological classification (Cavanagh et al. 2022), to investigate whether modern hydrodynamical simulations can reproduce the galaxy morphology–density relation and offer insights into its origin. Using a CNN to perform 'visual' classification of simulated galaxies means that morphological relations can be compared in a consistent manner with observed galaxies. With this work, we aim to test which physical processes drive the increase in lenticular and elliptical fractions in denser environments, and the origin of field early-type galaxy fractions.

This paper is structured as follows. In Section 2, we describe the EAGLE simulations and the CNN used to perform galaxy classification. Section 3 presents the main results from this work, including the morphology–density relation in EAGLE and its origin in the simulations. Finally, in Section 4 we discuss the connection between the morphology–density relation and S0 galaxy formation and summarize the results of this work.

2 METHODS

2.1 EAGLE simulations

The EAGLE project is a campaign of cosmological hydrodynamical simulations of galaxy formation and evolution (for full details, see Schaye et al. 2015; Crain et al. 2015). The simulations assume cosmological parameters consistent with the Planck Collaboration XVI (2014, namely, $\Omega_\Lambda = 0.693$, $\Omega_m = 0.307$, $\Omega_b = 0.04825$, $h = 0.6777$, $\omega_8 = 0.8288$). The simulations were performed with a highly modified version of the N-body, TreePM, smoothed particle hydrodynamics code GADGET 3 (last described by Springel 2005).

Importantly, the simulations include routines for radiative cooling (Wiersma, Schaye & Smith 2009a), star formation (Schaye & Dalla Vecchia 2008), stellar evolution (Wiersma et al. 2009b), stellar feedback (Dalla Vecchia & Schaye 2012), black holes (Rosas-Guevara et al. 2015), and active galactic nucleus (AGN) feedback (Booth & Schaye 2009). The stellar and AGN feedback prescriptions were calibrated so that the simulations reproduce the galaxy stellar mass function, galaxy sizes, and black hole masses at $z \approx 0$ (Crain et al. 2015).

The simulations have since been shown to reproduce a broad range of galaxy population properties, including the evolution of galaxy masses and sizes (Furlong et al. 2015, 2017), galaxy luminosities and colours (Trayford et al. 2015), cold gas properties (Lagos et al. 2015; Marasco et al. 2016; Crain et al. 2017), and circumgalactic and intergalactic absorption system properties (Rahmati et al. 2015; Oppenheimer et al. 2016; Turner et al. 2016). The simulations are therefore ideal for comparing with observed galaxy populations.

In this work, we analyse the EAGLE reference simulation (Ref-L100N1504 Schaye et al. 2015) of a 100^3 comoving Mpc^3 (cMpc) periodic volume. The volume initially contains 1504^3 gas and dark matter particles with initial masses of 1.81×10^6 M$_\odot$ and 9.7×10^6 M$_\odot$, respectively, and has a maximum gravitational softening length of 0.7 kpc (for $z < 2.8$, scaling as 2.66 comoving kpc at earlier times). In the Appendix, we also use the simulation of a smaller 50^3 cMpc^3 volume at the same resolution (Ref-L050N752) to test the accuracy of the CNN. Galaxies are identified in the simulation by first running a friends-of-friends algorithm (Davis et al. 1985) to identify dark matter structures, and then using the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009) to identify bound structures (galaxies/subhaloes). To follow the evolution of galaxies between successive snapshots, galaxy merger trees were constructed using the D-TREES algorithm (Jiang et al. 2014; Qu et al. 2017). This data are available publically through the EAGLE data base (McAlpine et al. 2016).

2.2 Sample selection and analysis

We limit our analysis to galaxies with $M_*(r < 50\,\mathrm{kpc}) > 10^{10}$ M$_\odot$, such that they are well resolved with $\gtrsim 10^4$ stellar particles. Particles within galaxies are defined as those bound to the subhalo according to SUBFIND (i.e. unbound particles and those bound to other subhaloes are excluded). This gives us a sample of 3607 galaxies at $z = 0$.

We calculate projected number densities ($\Sigma_N = N/(\pi d_N^2)$, where d_N is the distance to the Nth neighbour) for each galaxy by projecting

galaxies along the *z*-axis of the simulation volume (using only galaxies with $M_* > 10^{10}\,\mathrm{M}_\odot$ for the neighbour counts). To mimic observational radial velocity cuts (i.e. $\pm 1000\,\mathrm{km\,s^{-1}}$ in Baldry et al. 2006), we use a maximum line-of-sight (*z* axis) separation between galaxies of $1000\,\mathrm{km\,s^{-1}}/H_0 \approx 15$ Mpc. For each galaxy, we calculate Σ_3 (as in Dressler 1980) and an average density $<\Sigma>$ (the average of Σ_4 and Σ_5, as in Baldry et al. 2006).

Synthetic SDSS *g*-band images to be classified by the CNN were generated for each galaxy using the process described in Cavanagh et al. (2022), which we briefly summarize here. The images with 100×100 pixels of size 0.5 kpc were generated for the galaxies using SPHVIEWER (Benitez-Llambay 2015), with particles distributed according to their local smoothing length. Galaxies were oriented face-on using the spin vector for star particles between 2.5 and 30 kpc from the centre of the galaxy. Rest-frame SDSS *g*-band luminosities were calculated for each stellar particle using the FSPS stellar population model (Conroy, Gunn & White 2009; Conroy & Gunn 2010). Following Trayford et al. (2017), young star particles (< 100 Myr) and star-forming gas particles were resampled at higher resolution ($10^3\,\mathrm{M}_\odot$) based on the star-formation rate of the particles, to ensure that (e.g.) the spiral arms of galaxies are adequately sampled. The new particles adopt the same properties (position, metallicity, smoothing lengths) as the particle from which they were sampled, and new ages were randomly assigned between 0 and 100 Myr. For the image smoothing lengths, young (< 100 Myr) particles adopt the local SPH smoothing, while older stars use a local stellar smoothing length calculated from the nearest 64 neighbours.

2.3 Galaxy classification

Machine learning, and convolutional neural networks (CNNs) in particular, have been shown to be an extremely effective tool to perform image classification (see LeCun, Bengio & Hinton 2015, for a review). In astronomy, CNNs have now been applied to a wide range of topics, such as classifying galaxy morphologies and bars, detecting mergers and identifying gravitational lenses (e.g. Dieleman, Willett & Dambre 2015; Abraham et al. 2018; Ackermann et al. 2018; Schaefer et al. 2018; Huertas-Company et al. 2019; Cavanagh, Bekki & Groves 2021).

To classify galaxy morphologies, we utilized a modified CNN based on the four-class model (i.e. four classification categories) from Cavanagh et al. (2022), originally developed in Cavanagh et al. (2021). The CNN was trained on the Nair & Abraham (2010) morphological catalogue and *g*-band imaging from the Sloan Digital Sky Survey DR7 (York et al. 2000; Abazajian et al. 2009) to classify galaxies into E, S0, spirals and irregulars. This catalogue contains 14 034 galaxies with redshifts $0.01 < z < 0.1$. Though the difference between the training (real) and classification (simulated) images implies a domain shift,[1] the realism of the training set has more importance than that of the images being classified (e.g. as found by Bottrell et al. 2019, in the context of galaxy images). Therefore, the realism of the images being classified (e.g. lack of realistic fore-/backgrounds) is not expected to have a major impact on the results.

The core elements of a CNN are its convolutional layers, used to extract features using trainable filter weights, and pooling layers, used to downsample the input and preserve dominant features (LeCun

[1]This case is reversed compared to the 'reality gap' problem (Tobin et al. 2017) that occurs when training on simulated images.

et al. 2015). The CNN used in this work is adapted from Cavanagh et al. (2022) and features four blocks of alternating convolutional and max pooling layers, followed by two fully-connected dense layers with 512 nodes each, for a total of 5.6 million trainable parameters. Each convolutional layer utilizes a 7×7 kernel with ReLU activation (Glorot, Bordes & Bengio 2011). We use ReLU activation for the dense layers, with softmax activation for the output layer. We also use dropout layers after each pooling layer. Dropout is regularization technique in which a set fraction of inputs are ignored; this helps mitigate overfitting and also improves model robustness (Srivastava et al. 2014). The CNN was retrained from scratch with the new architecture.

All network parameters, including the learning rate for training, were further refined through extensive hyperparameter tuning using Optuna, an optimization framework. These changes further increased the accuracy of this CNN to 84 per cent, compared to 81 per cent in Cavanagh et al. (2022). The network was trained and tested on an Nvidia RTX 3060 desktop GPU. The classification of 35,082 EAGLE galaxies from Cavanagh et al. (2022, with $M_* > 10^{10}\,\mathrm{M}_\odot$ and all snapshots in the redshift range $0 \leq z \leq 1$) took less than 2 minutes, though in this work we focus on the $z = 0$ sample of 3,607 galaxies. Our galaxy sample size is comparable to that from Huertas-Company et al. (2019), who used a CNN to classify the morphologies of $z \approx 0$ galaxies from the IllustrisTNG simulation ($\sim 12,000$ galaxies with $M_* > 10^{9.5}\,\mathrm{M}_\odot$). We show examples of galaxies classified into each class (E, S0, Sp, Irr) in Fig. 1. Each example galaxy has a classification confidence above 90 per cent (indicated in bottom left of each panel).

To demonstrate the accuracy of the CNN when applied to the simulated galaxy images, in the Appendix we compare the classifications of the CNN with human visual classifications for the EAGLE RefL050N752 simulation (i.e. the reference model for a smaller 50^3 cMpc^3 volume). Two views of the galaxies (face- and edge-on) were used in the visual classifications to distinguish spheroidal and discy galaxies. Overall, the CNN is 76–88 per cent accurate when considering early- (E+S0) and late-type (Sp + Irr) galaxies. However, as discussed further in the Appendix and Section 3.1, flattened high-mass galaxies are often identified as ellipticals (analogous to fast-rotating galaxies visually classified as ellipticals, e.g. Cappellari et al. 2007; Emsellem et al. 2007), while all low-mass early-type galaxies are classified by the CNN as S0s leading to underestimated elliptical fractions.

3 RESULTS

3.1 Properties of morphological types

Before investigating the morphology–density relation (Section 3.2), we first verify the dependence of morphology on different galaxy properties. In Fig. 2, we compare the morphology–mass relations from EAGLE (3607 galaxies) with observed relations from Nair & Abraham (2010, 2721 galaxies) and the Sydney-AAO Multi-object Integral field (SAMI) Galaxy Survey (1823 galaxies, $0.004 < z < 0.095$ Bryant et al. 2015; Cortese et al. 2016). The Nair & Abraham (2010) catalogue has an apparent magnitude limit of SDSS $g < 16$ mag. Thus, we limit the catalogue to galaxies with redshifts $z < 0.03$ to be reasonably complete for $M_* > 10^{10}\,\mathrm{M}_\odot$. The predictions from the EAGLE simulation are generally in good agreement with the observed relations: the fraction of elliptical galaxies increases with stellar mass; the fraction of late-type galaxies decreases with stellar mass; while the fraction of lenticular galaxies is relatively constant with mass (decreasing only slightly at $M_* > 10^{11}\,\mathrm{M}_\odot$). The
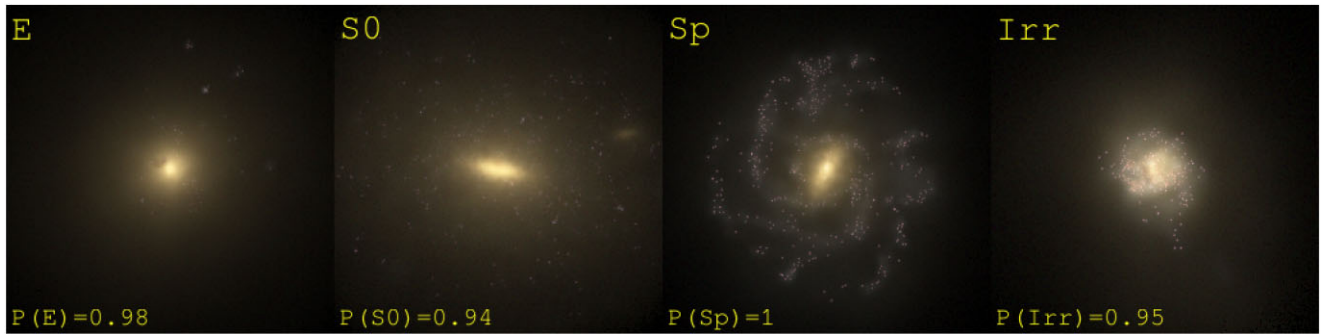
Figure 1. Examples of galaxies classified by the CNN as elliptical (left; GalaxyID = 19054212), S0 (middle left; GalaxyID = 21109760), spiral (middle right; GalaxyID = 16413979), and irregular (right; GalaxyID = 9299365). The images show face-on inclinations for each galaxy. The classification confidences are indicated in the bottom left of each panel. The mock three-colour (*gri*) images were generated for EAGLE galaxies by Trayford et al. (2017), but note that CNN classifications use only *g*-band images.
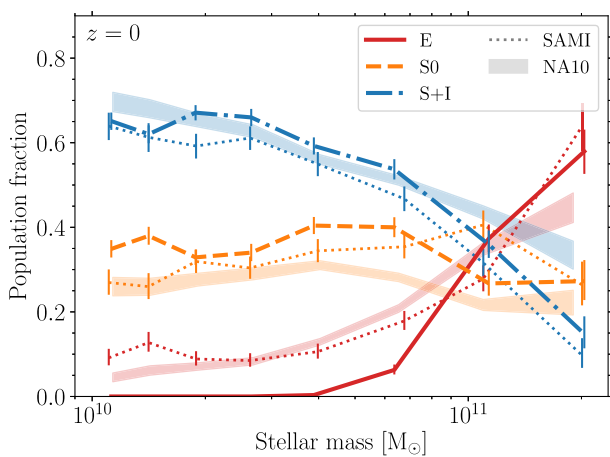

Figure 2. Dependence of morphology on galaxy stellar mass in the EAGLE simulation at $z = 0$. Elliptical (E), lenticular (S0), and late-type (S + I) galaxies are shown as solid red, dashed and dash–dotted lines, respectively. Uncertainties in the figure are calculated using binomial statistics. For comparison, the shaded regions show the confidence intervals for the morphology–mass relations from Nair & Abraham (2010, using their sample for $z < 0.03$), while dotted lines show the SAMI survey (0.004 < z < 0.095; Bryant et al. 2015; Cortese et al. 2016). Other than an absence of low-mass elliptical galaxies (discussed in Section 3.1), the EAGLE simulation shows good agreement with observed morphology fractions as a function of mass.

strong increase in elliptical fraction at $M_* \gtrsim 10^{11}$ may be expected from the increasing contribution of mergers to galaxy stellar mass growth at higher masses (e.g. Robotham et al. 2014; Rodriguez-Gomez et al. 2016; Clauwens et al. 2018; Tacchella et al. 2019; Davison et al. 2020). At masses > $10^{11}\,M_\odot$ the late-type fraction from EAGLE agrees better with the SAMI survey than the Nair & Abraham (2010) sample. The origin of this difference between the observational samples is not clear, but might simply be due to sample variance.

One issue with the classifications is the lack of EAGLE galaxies classified as elliptical at $M_* \lesssim 5 \times 10^{10}\,M_\odot$, with a corresponding fraction of lenticular galaxies that is slightly too high compared to the observed fractions. The likely origin of this issue is that the simulated galaxies are insufficiently concentrated due to numerical effects (i.e. gravitational softening, the ISM cooling floor and/or two-body scattering due to more massive dark matter particles, see Snyder et al. 2015; Bottrell et al. 2017; Ludlow et al. 2019; de Graaff

et al. 2021). Thus, otherwise spheroidal galaxies may be identified by the CNN as lenticular galaxies due to having density profiles closer to exponential discs. We discuss this issue further in the Appendix.

Next, in Fig. 3 we compare the distribution of specific star formation rates (sSFRs; using the instantaneous SFRs of gas particles bound to the subhalo and within 50 kpc of each galaxy) and the spin parameter $\kappa_{\rm co,\,rot}$ (the fraction of stellar kinetic energy in ordered co-rotation). In the left-hand panel of Fig. 3, we also show the median sSFR of different galaxy types from the Nair & Abraham (2010) catalogue for reference (where the SFRs are from Brinchmann et al. 2004). EAGLE galaxies in the different morphological classes generally have properties expected for their given morphology: Late-type galaxies (S + I) are typically star forming (sSFR $\gtrsim 10^{-1.8}\,{\rm Gyr}^{-1}$, in good agreement with the observed galaxies) and disc dominated ($\kappa_{\rm co,\,rot} \gtrsim 0.4$). Elliptical galaxies (E) are non-star forming (sSFR ~ 0) and spheroidal ($\kappa_{\rm co,rot} \sim 0.2$). Lenticular galaxies (S0) are intermediate between late-type and elliptical galaxies in both $\kappa_{\rm co,rot}$ (in qualitative agreement with angular momentum measurements, Cortese et al. 2016; Falcón-Barroso et al. 2019) and sSFR, accounting for the majority of 'green valley' ($10^{-2.8} \lesssim {\rm sSFR}/\,{\rm Gyr}^{-1} \lesssim 10^{-1.8}$) galaxies (in agreement with Bait, Barway & Wadadekar 2017). The sSFRs of E and S0 galaxies from Nair & Abraham (2010) are somewhat higher than EAGLE galaxies. This may be due to the particular method for calculating sSFR (using H α luminosity and 4000 Å break, Brinchmann et al. 2004), as Bait et al. (2017) find lower sSFRs for quenched galaxies using spectral energy distribution fitting. However, qualitatively, the comparison shows that S0 galaxies in both observations and the EAGLE simulations have median sSFRs slightly higher than elliptical galaxies, but significantly lower than late-type galaxies.

As discussed above, with regard to the lack of lower mass elliptical galaxies, many spheroidal galaxies ($\kappa_{\rm co,\,rot} \lesssim 0.2$) are classified as lenticular galaxies. A small fraction ($\sim$9 per cent) of galaxies classified as late-type (S + I) also have sSFR = 0. These galaxies are primarily from galaxy group/cluster regions (median $M_{200} = 10^{13.9}\,M_\odot$, median $\Sigma_3 = 10^{1.2}\,{\rm Mpc}^{-2}$). Some of these objects are galaxies classified as having irregular morphologies ($\approx$35 per cent of such galaxies). Others may be due to misclassification by the CNN (recall the typical accuracy is 84 per cent). From visual inspection, the majority of these cases would be classified by eye as S0 galaxies.

Elliptical galaxies feature a tail in $\kappa_{\rm co,rot}$ towards rotationally supported systems ($\kappa_{\rm co,\,rot} \gtrsim 0.3$). This is somewhat similar to the

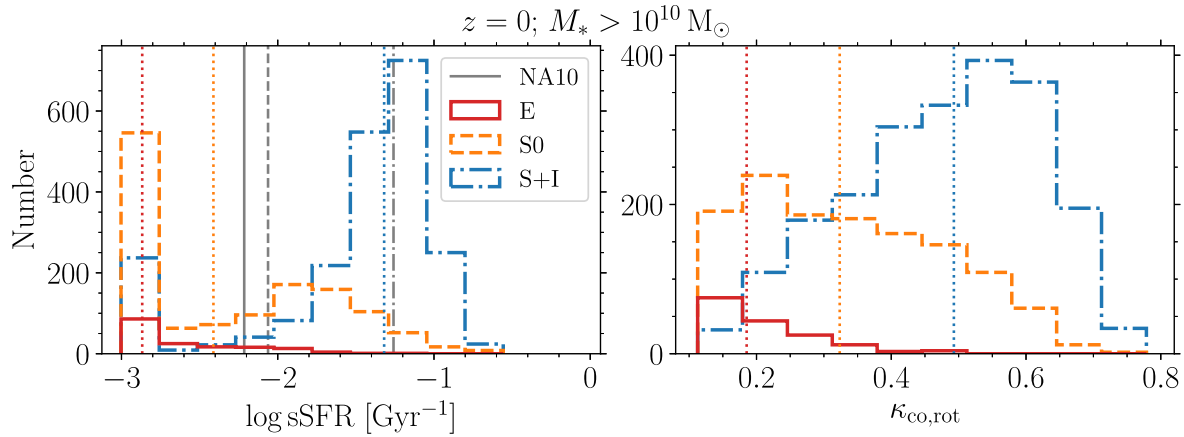$$z = 0;\ M_* > 10^{10}\,\mathrm{M_\odot}$$

Figure 3. Distribution of specific star formation rate (sSFR) and $\kappa_{\mathrm{co,rot}}$ (fraction of stellar kinetic energy in ordered co-rotation) for the different galaxy morphological types in the EAGLE simulation. Galaxies with low (or zero) specific star formation rates have been placed at sSFR $= 10^{-3}\,\mathrm{Gyr^{-1}}$. The median values for each morphological types are shown as vertical dotted lines. For comparison, the median sSFRs from the Nair & Abraham (2010) sample (as in Fig. 2) are shown as grey lines (with E, S0, and S+I galaxies shown as solid, dashed, and dash–dotted lines, respectively). Morphologically classified late-type (S + I) galaxies are star-forming and disc-dominated, while elliptical galaxies are quiescent and dispersion supported. S0 galaxies are typically intermediate between late-type and elliptical galaxies in both sSFR and $\kappa_{\mathrm{co,rot}}$.

distribution in stellar spin found for observed elliptical galaxies, i.e. morphologically classified elliptical galaxies are a mix of 'fast' and 'slow' rotating galaxies (and similarly some S0 galaxies have low stellar spin Cappellari et al. 2007; Emsellem et al. 2007, 2011). The origin of slow rotators in the EAGLE simulations has been investigated in previous works (Lagos et al. 2018b, 2022). In this work, we focus on the visual morphology–density relation, rather than the kinematic morphology–density relation (Cappellari et al. 2011), as most works to date have investigated the former.

Given the reasonable correspondence of the physical properties of different morphological types between the EAGLE simulation and observed galaxies, we now test whether EAGLE reproduces the observed galaxy morphology–density relation (Section 3.2) and what physical processes shape the relation in the simulation (Sections 3.3 and 3.4).

3.2 Morphology–density relation in EAGLE

In Fig. 4, we compare the morphology–density relation from the EAGLE simulation with observed relations from Houghton (2015, a revised catalogue from Dressler 1980, with updated redshifts and densities) and Nair & Abraham (2010, restricted to $z < 0.03$). Given the dependence of morphology on galaxy mass (Fig. 2), we apply mass limits to the simulations to reasonably match observational limits. The catalogue of Dressler (1980) has a galaxy luminosity limit $M_V < -20.4$, or $M_* \gtrsim 10^{10.5}\,\mathrm{M_\odot}$ for a typical $M/L_V \approx 2$–$3\,(\mathrm{M/L})_\odot$. We note that the galaxy sample from Nair & Abraham (2010), with median group halo masses of $\sim 10^{12}\,\mathrm{M_\odot}$, is more comparable to the EAGLE galaxy sample than the Dressler (1980) sample (55 rich galaxy clusters).

The left-hand panel of Fig. 4 shows the comparison of the EAGLE predictions with Houghton (2015, showing the polynomial fits in their table 1), with the galaxy number density calculated using the distance to the third nearest galaxy, Σ_3. Elliptical (E), lenticular (S0), and late-type (S + I) galaxy fractions from EAGLE are shown as solid, dashed, and dash–dotted lines, respectively, while the fits from Houghton (2015) are shown as dotted lines. The volume of the EAGLE simulation ($100^3\,\mathrm{cMpc^3}$) is not large enough to contain very massive ($\sim 10^{15}\,\mathrm{M_\odot}$) galaxy clusters, and thus does not probe the

highest densities of the observed relations. Over the densities covered by the simulations, we find an extremely good match between the EAGLE and observed morphology–density relations.

The right-hand panel of Fig. 4 shows the comparison of EAGLE predictions with the catalogue of Nair & Abraham (2010), using the average galaxy number density (average of Σ_4 and Σ_5) from Baldry et al. (2006). The Nair & Abraham (2010) results are shown as shaded regions (showing the uncertainty range calculated using binomial statistics). As discussed in Section 3.1, there are too few elliptical galaxies identified by the CNN at low galaxy masses, resulting in a morphology–density relation for ellipticals that is too low by a fraction ≈ 0.1. However, the late-type (S + I) fraction from EAGLE is in good agreement with observed morphology–density relation.

Therefore, the EAGLE simulation reproduces the observed $z \approx 0$ galaxy morphology–density relation. We stress here that these are genuine predictions from the EAGLE model. The EAGLE galaxy formation model was calibrated with the $z \approx 0$ galaxy mass function, galaxy sizes, and black hole masses (Schaye et al. 2015), not galaxy morphologies.

Given that the EAGLE model can reproduce the observed morphology–density relation, in the following sections we investigate what physical processes are important for shaping the relation in the simulation.

3.3 Dependence of galaxy properties on environment

We compare the typical physical properties of galaxies as a function of density in Fig. 5. The figure shows how sSFR (upper panel) and $\kappa_{\mathrm{co,rot}}$ (middle panel) correlate with density for each morphological type at $z = 0$ (for galaxies with stellar masses $M_* > 10^{10}\,\mathrm{M_\odot}$ as in the right-hand panel of Fig. 4). As already expected from Fig. 3, S0 galaxies in low-density environments ($\langle \Sigma \rangle \lesssim 3\,\mathrm{Mpc^{-2}}$) are intermediate between late-type (S+I) and elliptical (E) galaxies in both sSFR and $\kappa_{\mathrm{co,rot}}$. However, the figure shows that the properties of late-type and S0 galaxies correlate with environment: in the median, late-type galaxies show lower rotational support (lower $\kappa_{\mathrm{co,rot}}$) at higher densities ($\langle \Sigma \rangle \gtrsim 3\,\mathrm{Mpc^{-2}}$), potentially because large rotationally-supported discs are impacted by dynamical processes in dense environments (e.g. tidal stripping and harassment, Moore et al.
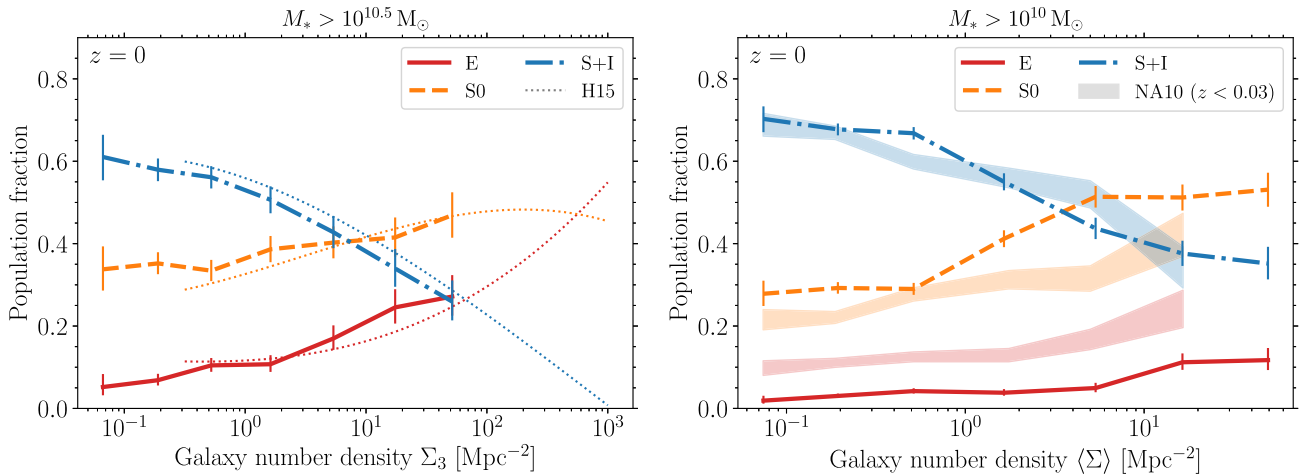
Figure 4. Morphology–density relations in the EAGLE simulation at $z = 0$. Elliptical (E), lenticular (S0) and late-type (S + I) galaxies from EAGLE are shown as solid, dashed and dash–dotted lines, respectively. The left-hand panel shows a comparison with the relations of observed galaxies from Houghton (2015, an updated catalogue from Dressler 1980) as a function of Σ_3 (projected galaxy number density to third nearest neighbour). The right-hand panel shows a comparison with the relations determined from the catalogue of Nair & Abraham (2010, for $z < 0.03$) as a function of $<\Sigma>$ (average of Σ_4 and Σ_5). The title for each panel shows the mass limit applied in each comparison. We only show bins in Σ where there are more than 50 galaxies, such that the results are not influenced by very low number statistics. Uncertainties in the figure are calculated using binomial statistics. The morphology–density relations from EAGLE agree very well with the Dressler (1980) and Houghton (2015) relations for all morphology types. The late-type (S + I) relation from EAGLE also agrees well with the Nair & Abraham (2010) catalogue, but the elliptical fraction is too low by ≈ 0.1 due to issues with identification at low masses (see Section 3.1 and Fig. 2).

1996, 1999). S0 galaxies show the inverse trend, becoming more rotationally supported (higher $\kappa_{\rm co, rot}$) at higher densities (with similar median $\kappa_{\rm co,rot}$ to S + I galaxies in the highest density bin). Correa et al. (2017) similarly showed that the 'red sequence' for EAGLE satellite galaxies has a much larger proportion of disc galaxies than central galaxies. This trend for S0 galaxies is in agreement with the observational results by Coccato et al. (2020) and Deeley et al. (2020), who found that S0 galaxies in groups and clusters have higher rotational support (v/σ) than field S0 galaxies. Similarly, it is consistent with the kinematic morphology-density relation, where fast-rotating galaxies tend to become non-star forming (lenticular) at high densities (Cappellari et al. 2011).

The bottom panel of Fig. 5 shows the fraction of the total galaxy population that is contributed by S0 galaxies with high ($\kappa_{\rm co,rot} > 0.35$) and low rotation ($\kappa_{\rm co, rot} < 0.35$). Low rotation S0 galaxies contain a fixed fraction (≈ 0.2) of the population at all densities, while the fraction of high rotation S0 galaxies increase with density. Therefore, the increased S0 fraction in high-density environments in EAGLE appears to be due to the increase in quiescent disc galaxies, in agreement with the findings of previous studies (e.g. Postman & Geller 1984; van der Wel et al. 2010; Deeley et al. 2020).

At high densities, both S0 and late-type galaxies typically become completely non-star forming (sSFR = $0\,{\rm Gyr}^{-1}$; at $\langle\Sigma\rangle \gtrsim 5\,{\rm Mpc}^{-2}$ for S0 galaxies, at $\langle\Sigma\rangle \gtrsim 10\,{\rm Mpc}^{-2}$ for late-type galaxies). A similar transition to non-star-forming galaxies at $\Sigma \gtrsim 10\,{\rm Mpc}^{-2}$ for observed galaxies was found by Lewis et al. (2002). Possible causes are that the gas in the galaxies has been heated (e.g. via AGN feedback) such that it cannot cool to form stars, 'strangulation' of star formation through prevention of gas accretion following cluster infall (Larson et al. 1980) or that environmental processes such as tidal and ram pressure stripping (Spitzer & Baade 1951; Gunn & Gott 1972) have completely removed the gas in these galaxies (possibly combined with consumption of residual cold gas through star formation, e.g. Larson et al. 1980; Bekki et al. 2002). We discuss this further in Section 3.4.2. Unlike S0 and late-type galaxies, the properties of

elliptical galaxies (low sSFR and low $\kappa_{\rm co,rot}$) are relatively constant across all environments.

3.4 Physical drivers of the morphology–density relation

To further investigate the formation and evolution processes affecting galaxies in different environments, in Figs 6 and 7 we compare the redshift evolution for the progenitors of different morphological types. We divide the sample into 'low' (Fig. 6) and 'high' (Fig. 7) density regions at $\langle\Sigma\rangle = 5\,{\rm Mpc}^{-2}$ (where above this density S0 fraction is highest, Fig. 4, and sSFRs of S0 galaxies are lowest, Fig. 5). In each figure, we compare the specific star formation rates (panel a), gas-to-stellar mass ratio (panel b), stellar mass (panel c), spin parameter $\kappa_{\rm co, rot}$ (panel d), relative major merger rate (panel e), and black hole accretion mass (panel f). The major merger (stellar mass ratio >1:4) rates are shown relative to the average merger rate of all progenitors of galaxies with $M_*(z = 0) > 10^{10}\,{\rm M}_\odot$ in the EAGLE volume, in order to highlight the differences between each galaxy type. The merger times are measured at the snapshot where the galaxies have fully merged according to the SUBFIND algorithm, though we note the merging process may begin a number of snapshots prior to this time. To determine mass ratios, we calculate the ratio at the snapshot where the 'satellite' galaxy was at its maximum stellar mass (excluding snapshots when the main branch and satellite have 'exchanged' mass temporarily, see Qu et al. 2017). In the rest of the section, we discuss low-density environments in Section 3.4.1, high-density environments in Section 3.4.2, and focus on elliptical galaxies in Section 3.4.3.

3.4.1 Low-density environments

In Fig. 6, we compare the evolution of galaxies which are in low-density environments at $z = 0$. At $z \gtrsim 1$, the progenitors of S0 and late-type galaxies show a very similar evolution in both sSFR
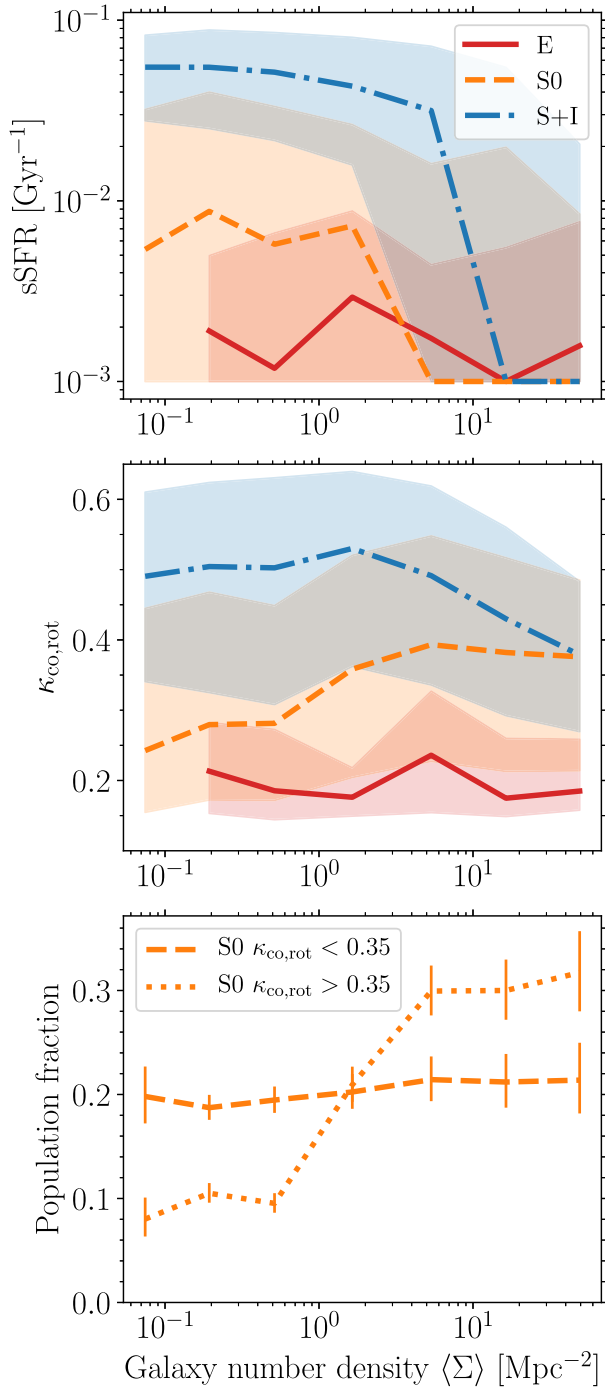
Figure 5. Properties for the different galaxy morphological types in the EAGLE simulation at $z = 0$ as a function of galaxy number density (comparing the average density, as in the right-hand panel of Fig. 4). The panels show specific star formation rate (sSFR, upper panel), $\kappa_{\mathrm{co,rot}}$ (fraction of stellar kinetic energy in ordered co-rotation, middle panel), and the fraction of the total galaxy population contributed by S0 galaxies with low (dashed line) and high (dotted line) rotation (lower panel). The thick lines show the median relation for each type, shaded regions show the 16th–84th percentiles. Solid red lines show elliptical galaxies (E), dashed orange lines show lenticular galaxies (S0), and dash–dotted blue lines show late-type galaxies (S + I). Galaxies with very low (typically zero) sSFRs have been placed at sSFR $= 10^{-3}\,\mathrm{Gyr}^{-1}$. Only bins in density with at least 10 galaxies are shown.

and $\kappa_{\mathrm{co,rot}}$. The galaxy types then diverge at $z \lesssim 1$, with late-type galaxies becoming more disc dominated, and S0 galaxies becoming more dispersion supported and less star forming. We argue this is due to a difference in galaxy merger rates: at $z < 1$, the progenitors of S0 galaxies experience a higher merger rate than late-type progenitors (by a factor of $\sim$2–3 at $z < 0.4$, well above the expected uncertainties in the merger rates from binomial statistics), which may explain the significantly lower $\kappa_{\mathrm{co,rot}}$ for S0 galaxies. This explanation is further supported by considering the progenitors of elliptical galaxies, which have the lowest median $\kappa_{\mathrm{co,rot}}$ at $z < 1$ and a merger rate nearly twice that of S0 progenitors over the same period.[2]

Potentially, the difference in sSFRs between S0 and late-type progenitors at $z < 1$ may also be related to mergers, but modulated by AGN feedback. Previous works have shown that star formation quenching of central galaxies in EAGLE is directly related to black hole growth and feedback (e.g. Bower et al. 2017; Correa, Schaye & Trayford 2019; Davies et al. 2020). In the EAGLE model, the energy released via AGN feedback is directly proportional to the mass accretion rate (Booth & Schaye 2009; Schaye et al. 2015). Galaxies for which black holes have accreted more mass will have received more energy feedback from AGN (and thus are expected to have lower SFRs) than galaxies, where black holes have accreted less mass (e.g. Di Matteo, Springel & Hernquist 2005; Sijacki et al. 2007, 2015; Bower et al. 2017).

We investigate the mass accreted by black holes in panel (f) of Fig. 6. As expected (given the known relationship between black hole mass and bulge mass/velocity dispersion, e.g. Kormendy & Richstone 1995; Magorrian et al. 1998; Ferrarese & Merritt 2000; Savorgnan et al. 2016), despite having similar median stellar masses (panel c), S0 galaxies have median black hole accretion masses over three times as large as late-type galaxies at $z = 0$ ($\approx 2 \times 10^7\,\mathrm{M}_\odot$ compared to $\approx 6 \times 10^6\,\mathrm{M}_\odot$, respectively). The median evolution of black hole accretion mass between late-type and S0 galaxies is largely similar at $z > 1.2$. The black holes in S0 galaxies then typically undergo a 'rapid growth phase' (Bower et al. 2017) at $z \approx 1.2$–0.6, before returning to a low accretion-rate phase ($z \lesssim 0.6$). This rapid growth phase is consistent with the time at which the sSFRs of S0 and late-type galaxies begin to diverge (panel a). In contrast, black holes in late-type galaxies typically have not undergone this rapid growth phase, leading to lower black hole accretion masses and higher SFRs. We also investigated the gas fractions of the progenitor mergers (following Lagos et al. 2018a), but found little difference in the fraction of gas-rich ('wet') mergers between S0 and late-type galaxies.

Previous works have shown that major galaxy mergers can disrupt the rotational motion of gas and drive strong inflows (e.g. Hernquist 1989; Barnes & Hernquist 1996; Mihos & Hernquist 1996), which may then drive rapid black hole growth and strong AGN feedback (Di Matteo et al. 2005; Springel et al. 2005; Hani et al. 2018; McAlpine et al. 2020; Davies et al. 2022). In turn, AGN feedback prevents the replenishment of the interstellar medium through heating and expulsion of the circumgalactic medium (Davies et al. 2019, 2020;

---

[2]We note that the transformation of the progenitors of massive elliptical galaxies from rotation to dispersion supported systems may be driven by frequent minor mergers, rather than major mergers (e.g. Naab, Johansson & Ostriker 2009; Khochfar et al. 2011). However, the relative merger rates in Fig. 6 are quantitatively consistent when also considering minor mergers (>1:10 mass ratio), i.e. major merger rate is correlated with overall merger rate. We focus on major mergers here because of their connections between gas inflows and AGN feedback (e.g. Hernquist 1989; Springel, Di Matteo & Hernquist 2005; Davies, Pontzen & Crain 2022).
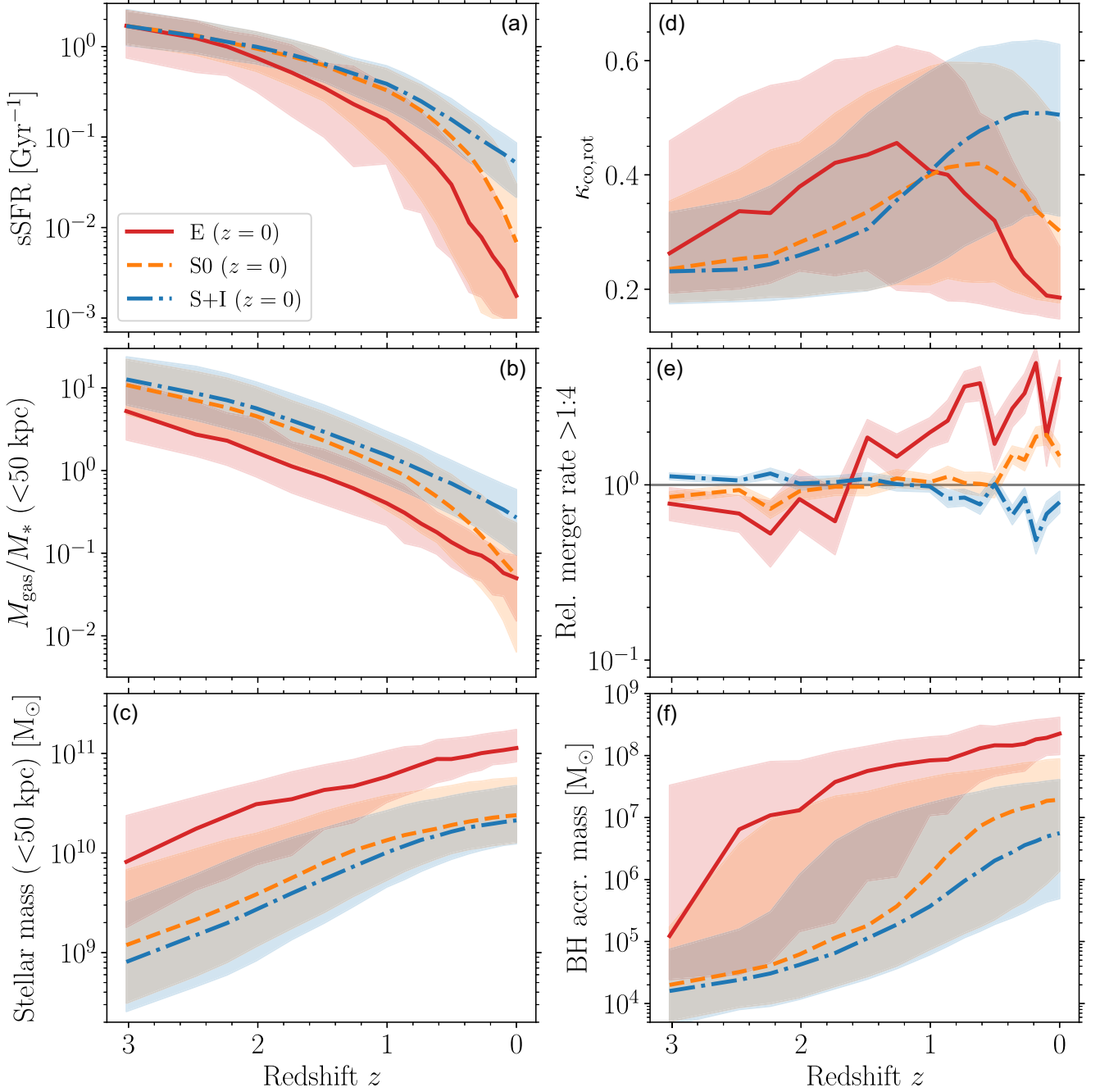
Progenitors: $M_*(z=0) > 10^{10}\,\mathrm{M_\odot}$; $\langle\Sigma\rangle\,(z=0) < 5\,\mathrm{Mpc}^{-2}$

Figure 6. Redshift evolution of properties of main progenitors for galaxies residing in *low*-density environments ($\Sigma_3 < 5\,\mathrm{Mpc}^{-2}$) at $z = 0$ (high-density environments are investigated in Fig. 7). Galaxies are grouped according to their $z = 0$ morphology, with solid red lines showing the median evolution of elliptical galaxies, dashed orange lines showing the median evolution of S0 galaxies and dash–dotted blue lines showing the median evolution of late-type galaxies. Shaded regions show the 16th–84th percentiles, except for panel (e) where the shaded regions show uncertainties in merger rates using binomial statistics. The subpanels of the figure show the evolution of: (a) specific star-formation rate; (b) gas-to-stellar mass ratio; (c) stellar mass within 50 kpc; (d) spin parameter $\kappa_{\mathrm{co,\,rot}}$; (e) relative major merger rate (stellar mass ratio >1:4); (f) black hole accretion mass. The merger rates are calculated relative to the merger rate of all progenitors of galaxies with $M_*(z=0) > 10^{10}\,\mathrm{M_\odot}$ in the EAGLE volume.

Oppenheimer et al. 2020). Mergers alone are unlikely to affect SFRs long term as, without some process to heat or expel gas, gaseous discs may survive or re-form after the merger (Barnes & Hernquist 1996; Springel & Hernquist 2005; Robertson et al. 2006; Governato

et al. 2007; Hopkins et al. 2009) and the merger remnant may remain star forming (Di Matteo et al. 2005; Springel et al. 2005; Davies et al. 2022). Similarly, AGN feedback alone is unlikely to significantly alter the kinematic properties of the stars in the

Progenitors: $M_*(z=0) > 10^{10}\,\mathrm{M_\odot}$; $\langle\Sigma\rangle\,(z=0) > 5\,\mathrm{Mpc^{-2}}$
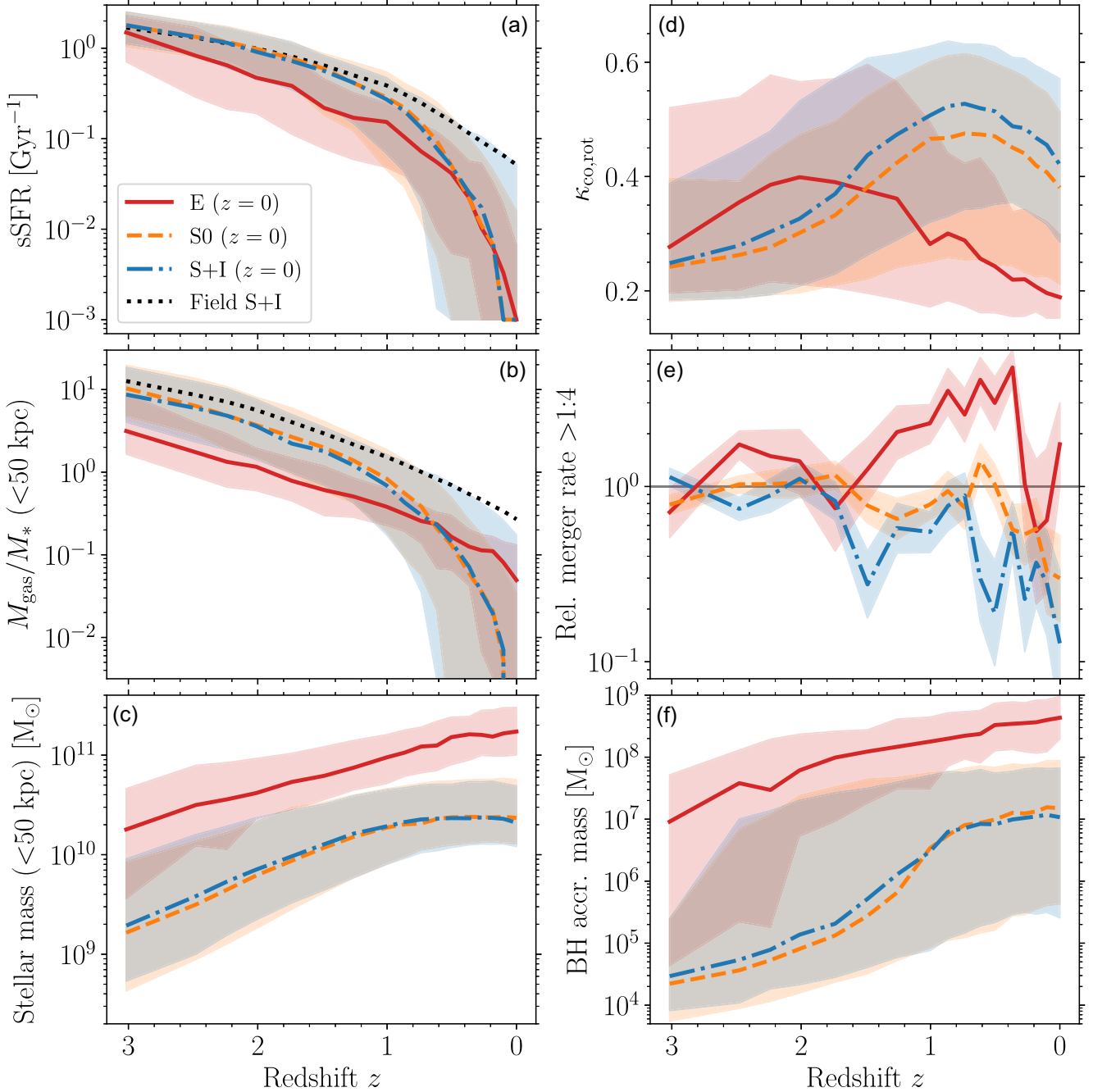

Figure 7. Redshift evolution of properties of main progenitors for galaxies residing in *high*-density environments ($\Sigma_3 > 5\,\mathrm{Mpc^{-2}}$) at $z = 0$. Panels and line styles are as in Fig. 6. For reference, the black-dotted lines show the sSFR and gas-to-stellar mass ratio evolution of low-density S + I galaxies in Fig. 6.

galaxy (i.e. $\kappa_{\mathrm{co,rot}}$ in panel d). Thus, we arrive at a coherent picture in which the higher merger rates of the progenitors of field S0 and elliptical galaxies results in both their lower rotational support (Cortese et al. 2016; Falcón-Barroso et al. 2019) and lower SFRs (Bait et al. 2017) compared to late-type galaxies. We note that the difference in merger rates between S0 and late-type galaxies is most different at $z < 0.4$, while the sSFRs diverge somewhat earlier at $z \sim 0.8$–0.6. This may be evidence of a 'delay' (see Davies et al. 2022) between initial galaxy interactions (that triggers the black hole

growth and AGN feedback) and the final coalescence of the galaxies (the snapshot at which the merger is identified in the galaxy merger tree).

3.4.2 High-density environments

In Fig. 7, we compare the evolution of galaxies in high-density environments at $z = 0$. Contrasting the results with Fig. 6 reveals
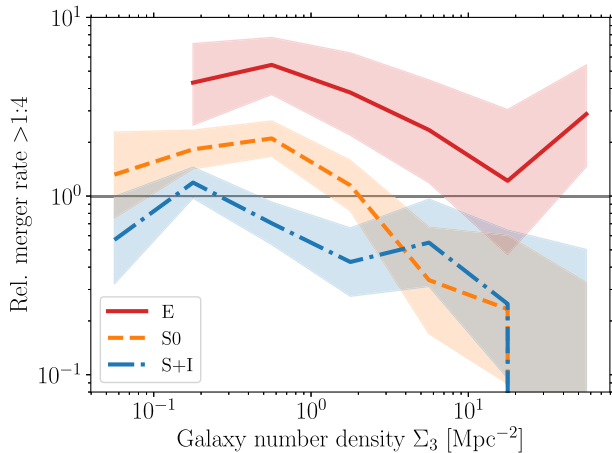
Figure 8. Relative major merger rates (stellar mass ratio >1: 4) of galaxies at $z = 0$ as a function of environment (galaxy number density) and galaxy morphological type. Galaxies are binned in density in an identical manner to Fig. 4, and only show density bins with a least 20 galaxies of a given morphological type. Shaded regions show the uncertainties on the merger rates from binomial statistics. The solid horizontal line shows the unity line for reference. As expected from *N*-body simulations (Ghigna et al. 1998), mergers become less likely for galaxies in higher density environments.

very different histories compared to galaxies in low densities, with the exception of the progenitors of elliptical galaxies which evolve similarly in both environments (the main differences being slightly higher masses, earlier stellar mass and black hole growth and an earlier peak in $\kappa_{co,rot}$ for high-density ellipticals).

The progenitors of late-type and S0 galaxies in high-density environments show very similar evolutionary histories, with the main difference being that late-type galaxies show slightly higher median $\kappa_{co,rot}$ at all redshifts (panel d). This might be explained by the slightly higher meger rates for S0 galaxies (panel e) and the larger tail (16th–84th percentiles) to lower $\kappa_{co, rot}$. However, unlike low-density S0s, the progenitors of high-density S0s show *lower* than average major merger rates (panel e) at $z < 0.5$, indicating infall into a larger group/cluster has resulted in lower merger rates. We demonstrate this further in Fig. 8 by comparing the relative major merger rates at $z = 0$ as a function of galaxy morphology and environment. All galaxy types show a correlation between merger rate and galactic density such that major mergers become less likely in high-density environments, with the largest change in merger rate occurring for S0 galaxies. This is (at least qualitatively) in agreement with expectations from *N*-body simulations that mergers are less frequent in galaxy clusters (Ghigna et al. 1998). A possible exception to the trend is elliptical galaxies at the highest densities in the sample, which could be a result of being the central galaxies in groups/clusters (for ~50 per cent of such galaxies, as compared to <10 per cent for S0 and late-type galaxies at the same density). However, as the binomial uncertainties show, it is not a statistically significant result given the lower number of elliptical galaxies (22) in the highest density range.

In Fig. 7, both late-type and S0 progenitors in high-density environments show a peak $\kappa_{co,rot} \sim 0.5$ (panel d) that is similar to the presenRedshift evolution of properties of main progenitorstday value of low-density late-type galaxies (Fig. 6). The redshift of peak $\kappa_{co,rot}$ also coincides with the typical infall redshift of ≈ 0.74 for both galaxy types. Therefore, given their similar sSFR and $\kappa_{co,rot}$, prior to group/cluster infall the progenitors of high-density S0s may be expected to appear visually similar to late-type galaxies. That the

decline in median $\kappa_{co,rot}$ at $z < 0.7$ is very similar between both late-type and S0 galaxies, despite the lower merger rate of late-type galaxies, suggests that mergers are unlikely to be the origin of the decline. Other possible processes[3] include heating following gas stripping (where gas makes a non-negligible contribution to the potential; e.g. at $z \sim 1$ both galaxy types have typical gas-to-stellar mass ratios ≈ 1 within 50 kpc, panel b), harassment (Moore et al. 1996) or numerical heating (due to interactions between star particles and more massive dark matter particles, Ludlow et al. 2019).

Both late-type and S0 galaxies become (in the median) non-star forming by $z = 0$ (panel a). Unlike low-density environments (where over 70 per cent are central galaxies), ≈ 90 per cent of late-type and S0 galaxies in high-density environments are satellite galaxies, with a typical infall redshift of ≈ 0.75 (i.e. time when a galaxy becomes a satellite in a halo, rather than the central galaxy). The infall redshift coincides with the time at which the median sSFRs of S0 and late-type galaxies (panel a) experience a significant downturn relative to low-density S + I galaxies (black dotted lines in panel). As shown in panel (b), for late-type and S0 galaxies it is not simply that the gas in the galaxies becomes non-star forming (i.e. through AGN feedback, as for elliptical galaxies which still have gas-to-stellar mass ratios of ~5 per cent at $z = 0$ but are largely non-star forming), but that most galaxies become completely gas free. We suggest environmental processes such as tidal and ram pressure stripping (Spitzer & Baade 1951; Gunn & Gott 1972) have removed the gas in most galaxies in high-density environments. As shown by Pfeffer et al. (2022), the gas-stripping process in EAGLE galaxies occurs in an outside-in manner (leading to only galaxy centres remaining star forming; consistent with observations of cluster galaxies, e.g. Koopmann & Kenney 2004a, b; Reynolds et al. 2022), with stronger stripping occurring in higher-mass group/cluster haloes. However, AGN feedback could still play some role in these galaxies, even if stripping is responsible for complete quenching, as we note that high-density late-type and S0 galaxies have higher median accretion masses for black holes (panel f) than low-density late-type galaxies. For example, one could conceive of a scenario in which the black holes experience a rapid growth phase prior to, or about the same time as, the galaxies become satellites and gas stripping takes place, possibly due to ram pressure driving gas inflows (Bekki 2009; Poggianti et al. 2017; Ricarte et al. 2020).

To investigate this further, we have divided the high-density S0 and S + I galaxies into high rotation (Fig. 9) and low rotation (Fig. 10) galaxies at $\kappa_{co,rot}(z = 0) = 0.4$ (i.e. intermediate between S0 and S + I galaxies in Fig. 6). The cut in $\kappa_{co,rot}$ roughly divides the samples into two (144/314 S0 and 121/246 S + I galaxies in the high rotation samples). Following from the discussion in Section 3.4.1, we investigate whether low and high $\kappa_{co,rot}$ S0 and S + I galaxies in high-density environments also show different origins due to the presence or absence of AGN feedback.

In Fig. 9, we show the progenitor evolution for high-density S0 and S + I galaxies with high rotation, $\kappa_{co,rot}(z = 0.4) > 0.4$. These galaxies generally do not experience a rapid black hole growth phase (panel f), and thus AGN feedback plays a limited role in their evolution. Instead, both their sSFR (panel a) and gas-to-stellar mass ratios (panel b), along with $\kappa_{co, rot}$ (panel d), only begin to significantly decrease (relative to low density S + I galaxies) upon group/cluster infall (vertical dotted lines in each panel), indicating

---

[3]Croom et al. (2021) showed that stellar population fading may lower the *observed* stellar spin of galaxies. This process is unlikely to strongly affect $\kappa_{co,rot}$, which depends on the intrinsic angular momentum and not luminosity.

Progenitors: $M_*(z = 0) > 10^{10}\,\mathrm{M_\odot}$; $\langle\Sigma\rangle\,(z = 0) > 5\,\mathrm{Mpc^{-2}}$; $\kappa_{\mathrm{co,rot}}(z = 0) > 0.4$
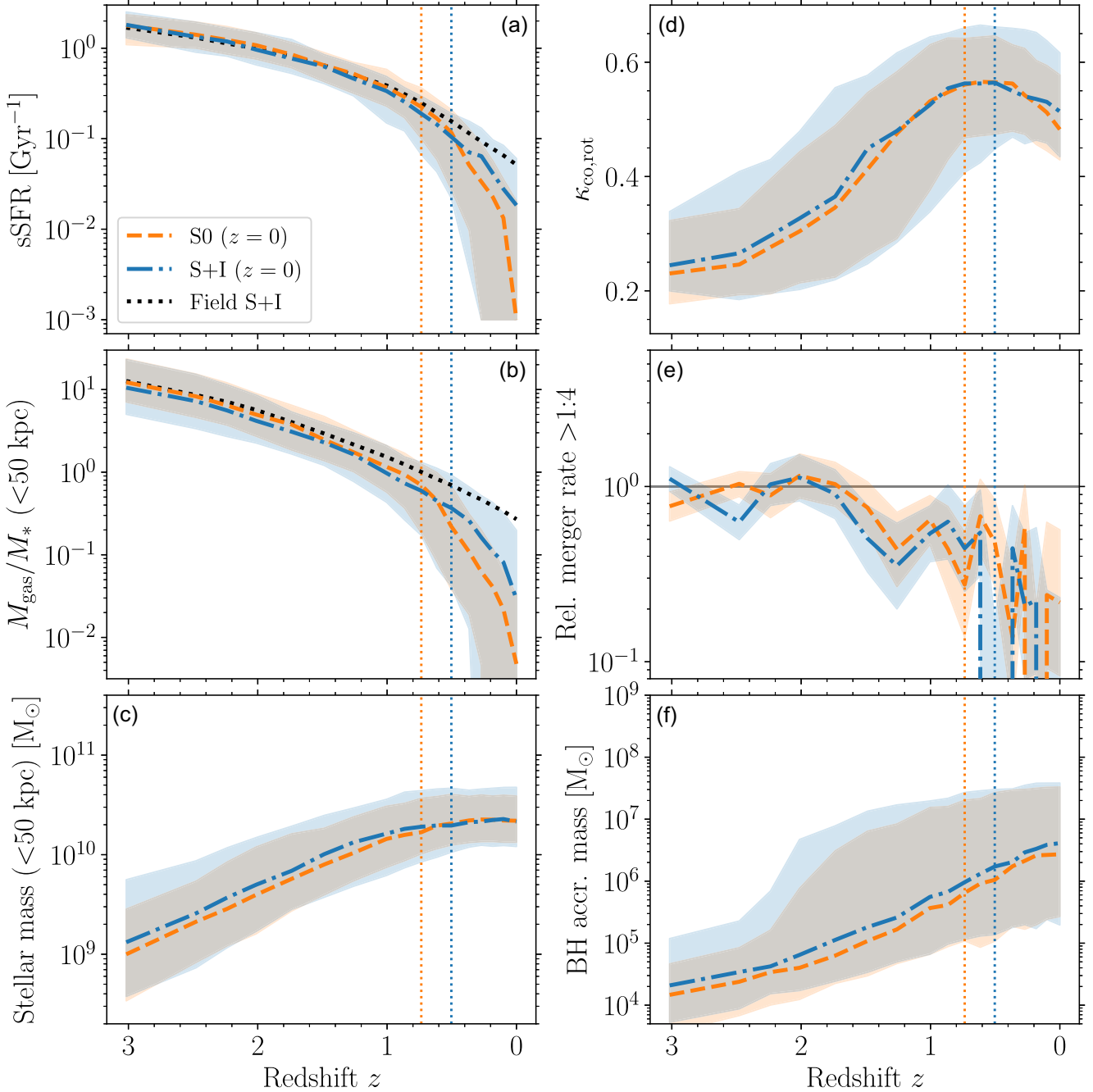

Figure 9. Redshift evolution of properties of main progenitors for galaxies residing in *high*-density environments ($\Sigma_3 > 5\,\mathrm{Mpc^{-2}}$) and with $\kappa_{\mathrm{co,\,rot}} > 0.4$ (disc dominated) at $z = 0$. Panels and line styles are as in Fig. 6. For reference, the black-dotted lines show the sSFR and gas-to-stellar mass ratio evolution of low-density S + I galaxies in Fig. 6. Satellite S0 and S + I galaxies have median infall redshifts of 0.74 and 0.5, respectively, indicated by vertical dotted lines in each panel.

environmental gas stripping (Spitzer & Baade 1951; Gunn & Gott 1972) as the origin of their decline. At $z = 0$, most S0 galaxies are almost completely quenched of star formation, while S+I galaxies have sSFRs $\approx$2.5 times lower than field galaxies. Relative to S + I galaxies, high $\kappa_{\mathrm{co,rot}}$ S0 galaxies have slightly earlier infall times and lower sSFRs by $z = 0$, but otherwise similar evolution in $\kappa_{\mathrm{co,rot}}$, black hole accretion, and merger rates (panels d, e, f).

In Fig. 10, we show the progenitor evolution for high-density S0 and S + I galaxies with low rotation, $\kappa_{\mathrm{co,rot}}(z = 0.4) < 0.4$, which show very different evolution compared to their high rotation counterparts. Like the low-density S0 galaxies (Fig. 6), low-rotation S0 and S + I galaxies in high-density environments have typically undergone a rapid black hole growth phase, though at earlier times from $z \approx 2$–1.5 (panel f). Their sSFRs (panel a) begin decreasing

Progenitors: $M_*(z=0) > 10^{10}\,\mathrm{M_\odot}$; $\langle\Sigma\rangle\,(z=0) > 5\,\mathrm{Mpc^{-2}}$; $\kappa_{\mathrm{co,rot}}(z=0) < 0.4$
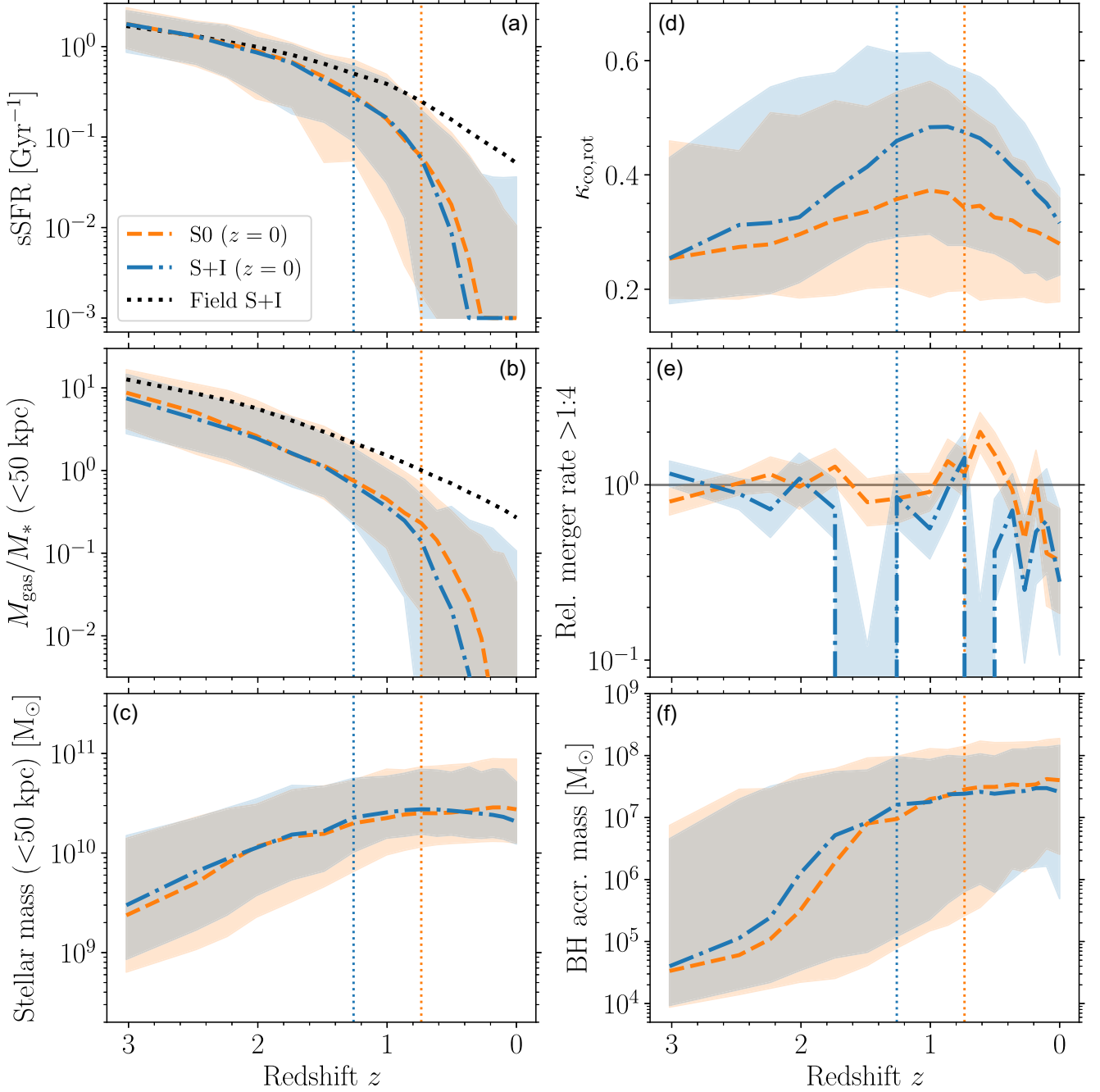

Figure 10. Redshift evolution of properties of main progenitors for galaxies residing in *high*-density environments ($\Sigma_3 > 5\,\mathrm{Mpc^{-2}}$) and with $\kappa_{\mathrm{co,rot}} < 0.4$ (spheroid-dominated) at $z = 0$. Panels and line styles are as in Fig. 6. For reference, the black dotted lines show the sSFR and gas-to-stellar mass ratio evolution of low-density S + I galaxies in Fig. 6. Satellite S0 and S + I galaxies have median infall redshifts of 0.74 and 1.26, respectively, indicated by vertical dotted lines in each panel.

at similar redshifts ($z \approx 1.6$) as the rapid black hole growth, while the gas-to-stellar mass ratios (panel b) do not significantly decrease until after group/cluster infall. Thus, AGN feedback initially begins to quench the galaxies, which is then completed by gas stripping after infall. Interestingly, the evolution of $\kappa_{\mathrm{co,rot}}$ is markedly different between the S0 and S + I galaxies (panel d), with S + I galaxy

progenitors typically achieving much higher peak $\kappa_{\mathrm{co,rot}}$ ($\approx$0.5) compared to S0 galaxies ($\approx$0.35). Panel (c) shows the S + I galaxies experience some mass-loss from $z \approx 0.8$ to $z = 0$, which might indicate that tidal stripping/harassment effects on the discs is responsible for lowering $\kappa_{\mathrm{co,rot}}$ (c.f. Moore et al. 1996, 1999; Bekki & Couch 2011).
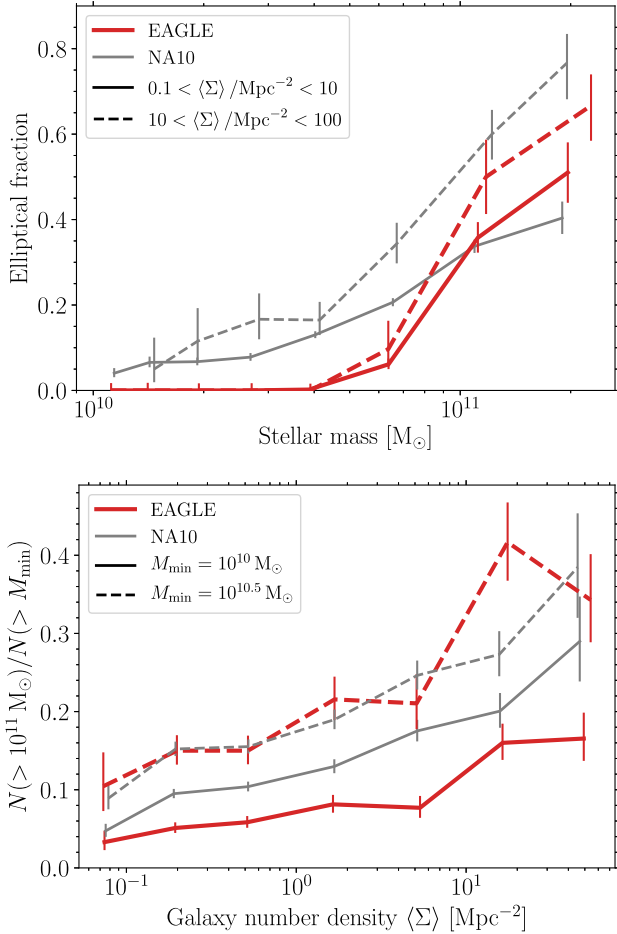
Figure 11. Comparison of elliptical and high-mass ($M_* > 10^{11}\,\mathrm{M_\odot}$) galaxy fractions between EAGLE (thick red lines) and Nair & Abraham (2010) catalogue (thin grey lines). *Top*: Elliptical fraction as a function of mass in low- (solid lines) and high-density environments (dashed lines), divided at $\langle\Sigma\rangle = 10\,\mathrm{Mpc^{-2}}$, where elliptical fraction is highest (Fig. 4). *Bottom*: Fraction of galaxies more massive than $M_* > 10^{11}\,\mathrm{M_\odot}$ as a function of environment. Solid lines show fraction for a minimum stellar mass limit of $10^{10}\,\mathrm{M_\odot}$ and dashed lines show a minimum limit of $10^{10.5}\,\mathrm{M_\odot}$. In both panels, we show bins in mass or number density with more than 20 galaxies.

3.4.3 Elliptical galaxies

In Sections 3.4.1 and 3.4.2, we only briefly discussed the formation of elliptical galaxies relative to S0 and late-type galaxies. As shown in Figs 6 and 7, elliptical galaxies experience early star formation quenching (through AGN feedback) due to rapid black hole growth at early times ($z \gtrsim 2$) and have the lowest rotational support due to their high galaxy merger rates (from $z \lesssim 1.5$). In low-density environments, the majority of ellipticals ($\approx$90 per cent) are central galaxies in their dark matter halo, compared with only $\approx$44 per cent in high-density environments. However, ellipticals which are classed as satellite galaxies in high-density environments have typical infall redshifts of $\approx$0.27, much later than S0 and late-type galaxies, but consistent with the drop in merger rate at $z < 0.3$ in panel (e) of Fig. 7 and the later assembly times of more massive dark matter haloes (e.g. De Lucia et al. 2006; Qu et al. 2017).

In Fig. 11, we investigate the origin of the increasing fraction of elliptical galaxies with density. As the top panel of Fig. 11 shows, for lower-mass galaxies ($M_* \sim 10^{10}\,\mathrm{M_\odot}$), which dominate galaxy numbers overall due to the steep galaxy mass function, there is almost

no difference in elliptical fraction for low ($\langle\Sigma\rangle < 10\,\mathrm{Mpc^{-2}}$) and high ($\langle\Sigma\rangle > 10\,\mathrm{Mpc^{-2}}$) densities in the Nair & Abraham (2010) catalogue. Given the issues identifying low-mass elliptical galaxies in the EAGLE simulation (Section 3.1), if low-mass galaxies drove the morphology–density relation for elliptical galaxies then we would expect a flat relation for EAGLE in Fig. 4, which is not what is found.

Instead, as the bottom panel of Fig. 11 shows, the fraction of high-mass galaxies ($M_* > 10^{11}\,\mathrm{M_\odot}$, compared with all galaxies greater than 10^{10} and $10^{10.5}\,\mathrm{M_\odot}$) increases with galaxy density, likely to be driven by mass segregation (through dynamical friction) and tidal disruption/merging of lower mass galaxies (c.f. Whitmore et al. 1993). This, combined with the elliptical fraction being a strong function of galaxy mass, and a secondary dependence on density for high-mass galaxies (top panel, i.e. in dense environments more high-mass galaxies are elliptical), drives the morphology–density relation for elliptical galaxies.

Whitmore et al. (1993) found that $\sim$55 per cent of galaxies at the centres of clusters are elliptical, regardless of local density, and showed that the morphology–radius relation appears to be the more fundamental correlation. Analogously, Houghton (2015) found the elliptical fraction rises faster with density in lower density clusters (confirming trends found when only considering slow-rotating galaxies Cappellari et al. 2011; D'Eugenio et al. 2013; Houghton et al. 2013; van de Sande et al. 2021). van der Wel et al. (2010) argued both the elliptical morphology–density and morphology–radius correlations are due to cluster centres hosting more massive galaxies, which are more often ellipticals, and indeed we find a similar result for the morphology–density relation with the EAGLE simulation.

4 DISCUSSION AND SUMMARY

4.1 The galaxy morphology–density relation and the origin of S0 galaxies

Much of the discussion on the origin of the morphology–density relation is closely tied with discussion on the origin of S0 galaxies. Recent works have shown that S0 galaxies do not appear to be a homogeneous population, but that their formation may vary with environment. Indeed, the form of the morphology–density relation itself, with a non-zero fraction of S0 galaxies at low densities, suggests environmentally dependent mechanisms (like stripping) cannot be the only origin (Dressler 1980; Postman & Geller 1984; van der Wel et al. 2010). In particular, S0 galaxies in clusters appear more rotationally supported than those in field (Coccato et al. 2020; Deeley et al. 2020) and, similarly, quiescent galaxies are more disc dominated in denser environments (van der Wel et al. 2010; which is also found for the EAGLE simulation, Correa et al. 2017). This appears to be consistent with a transition from (largely) a merger origin for field S0s to a stripped spiral origin for cluster S0s (this work Deeley et al. 2021). Low-mass ($M_* < 10^{10}\,\mathrm{M_\odot}$) field S0 galaxies may have a different origin again to higher mass galaxies (Fraser-McKelvie et al. 2018; Dolfi et al. 2021), which might be expected if they are too low mass for their formation to be affected by black hole feedback (e.g. Crain et al. 2015).

At face value, that S0 galaxies in clusters are more rotationally supported than field galaxies appears at odds with the lower disc-to-bulge ratios of S0 galaxies compared to spiral galaxies (Dressler 1980). Comparing to the disc-to-bulge ratios of S0 galaxies from Dressler (1980) over a similar range in Σ (1-50 $\mathrm{Mpc^{-2}}$) indeed shows that $\kappa_{\mathrm{co, rot}}$ (which correlates with disc-to-bulge ratio; Thob et al.

2019) is relatively constant and lower than that of disc galaxies, while the drop to lower $\kappa_{\mathrm{co,rot}}$, merger-dominated formation occurs at $\lesssim 1 \mathrm{Mpc}^{-2}$ (Fig. 5).

A merger origin (both major and minor) for S0 galaxies in low-density environments has of course been discussed in many works (e.g. Bekki 1998; Eliche-Moral et al. 2012, 2013; Borlaff et al. 2014; Querejeta et al. 2015; Tapia et al. 2017; Diaz et al. 2018; Deeley et al. 2021). One of the key findings of this work (Section 3.4.1) is that, though mergers are responsible for the transformation to more dispersion-dominated systems, it is the black hole feedback induced by mergers that results in their lower star formation rates compared to late-type galaxies (e.g. Bait et al. 2017).

4.2 Summary

In this paper, we present the first analysis of the galaxy morphology–density relation in a cosmological hydrodynamical simulation. We used a CNN trained on observed galaxies (Cavanagh et al. 2022) to perform 'visual' morphological classification of galaxies in the EAGLE simulation (Schaye et al. 2015; Crain et al. 2015). Our main findings from the work are as follows.

We first demonstrated in Section 3.1 that EAGLE reproduces the observed morphology–galaxy mass relations (Fig. 2). Late-type and elliptical fractions are a strong function of galaxy mass, with late-type fraction decreasing and elliptical fraction increasing with mass, while the fraction of lenticular galaxies is relatively constant ($\sim$30–40 per cent). However, almost no galaxies with stellar masses $< 5 \times 10^{10} \mathrm{M}_\odot$ are classified as elliptical, compared to observed fractions of $\sim$5–10 per cent. This issue likely stems from the insufficiently concentrated galaxy profiles due to numerical effects (de Graaff et al. 2021), resulting in slightly elevated S0 fractions.

We also showed in Fig. 3 that the physical properties for simulated galaxies of each morphological type are in line with those expected from observations (e.g. Cortese et al. 2016; Bait et al. 2017; Falcón-Barroso et al. 2019). Late-type galaxies are typically star forming (high sSFR) and disc dominated (high $\kappa_{\mathrm{co,rot}}$), elliptical galaxies are spheroidal (low $\kappa_{\mathrm{co,rot}}$) and non-star forming (low sSFR), while S0 galaxies are intermediate in both properties.

In Section 3.2, we showed, for the first time, that simulations can reproduce the observed galaxy morphology–density relation (Fig. 4). The EAGLE model is therefore ideal to test the origin of the relation. However, the simulation does not reach the highest densities probed by observations (i.e. centres of rich galaxy clusters) due to the limited simulation volume.

In Sections 3.3 and 3.4, we investigated the physical processes that shape the morphology–density relation in the simulations. We found three key drivers of the relation:

(i) Transformation of disc-dominated galaxies from late-type to S0 through gas stripping in higher density environments (c.f. Gunn & Gott 1972; Larson et al. 1980; Bekki et al. 2002), which accounts for the decreasing late-type fraction at higher densities.

(ii) An origin for S0 galaxies that transitions from merger-dominated to stripping-/starvation-dominated from low to high-density environments (c.f. Deeley et al. 2021), which accounts for the increasing S0 fraction with density. In addition, we showed that AGN feedback driven by mergers may play a significant role in the formation of field S0s (c.f. Davies et al. 2022).

(iii) The increasing fraction of high-mass galaxies, that are more likely ellipticals, which drives the rapid increase in elliptical galaxies at higher densities (c.f. van der Wel et al. 2010).

4.3 Future directions

In this work, we have not investigated the redshift dependence of the morphology–density relation. Most high-redshift observations focus on rich galaxy clusters (Dressler et al. 1997; Couch et al. 1998; Fasano et al. 2000; Treu et al. 2003; Postman et al. 2005; Smith et al. 2005, though see van der Wel et al. 2007; Shimakawa et al. 2021), while the EAGLE volume ($100^3 \mathrm{cMpc}^3$) is too small to contain very massive clusters. Therefore, in future work we will extend our analysis to the Cluster-EAGLE/Hydrangea simulations (Bahé et al. 2017; Barnes et al. 2017), a set of 30 zoom-in simulations of galaxy clusters ($M_{200} \sim 10^{14}$–$10^{15.4} \mathrm{M}_\odot$). This will enable us to compare the redshift evolution of the morphology–density relation, the cluster morphology–radius relation (Dressler 1980; Whitmore et al. 1993) and the Butcher-Oemler effect (Butcher & Oemler 1978, 1984; Couch et al. 1994) with observations in similar environments.

We also have not discussed the neutral hydrogen (H I) morphology–density relation (Serra et al. 2012) and its connection to the optical morphology–density relation. In principle, the combination of optical and H I data may aid in determining the origin of individual early-type galaxies, rather than examining population-averaged results as we have done in this work. In future, the Widefield ASKAP L-band Legacy All-sky Blind surveY (WALLABY) survey with the Australian Square Kilometre Array Pathfinder (ASKAP) telescope (Koribalski et al. 2020) is expected to provide H I properties for more than 10^5 galaxies, significantly expanding the capability to study the morphology–density relation in the nearby Universe.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author. All data from the EAGLE simulations (including galaxy catalogues, merger trees and particle data) is publicly available (McAlpine et al. 2016) at http://www.eaglesim.org/database.php.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Abraham S., Aniyan A. K., Kembhavi A. K., Philip N. S., Vaghmare K., 2018, MNRAS, 477, 894
Ackermann S., Schawinski K., Zhang C., Weigel A. K., Turp M. D., 2018, MNRAS, 479, 415
Bahé Y. M. et al., 2017, MNRAS, 470, 4186
Bait O., Barway S., Wadadekar Y., 2017, MNRAS, 471, 2687
Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469
Barnes J. E., 1992, ApJ, 393, 484
Barnes J. E., Hernquist L., 1996, ApJ, 471, 115

Barnes D. J. et al., 2017, MNRAS, 471, 1088
Bekki K., 1998, ApJ, 502, L133
Bekki K., 2009, MNRAS, 399, 2221
Bekki K., Couch W. J., 2011, MNRAS, 415, 1783
Bekki K., Couch W. J., Shioya Y., 2002, ApJ, 577, 651
Benitez-Llambay A., 2015, py-sphviewer: Py-SPHViewer v1.0.0. Zenodo, available at https://doi.org/10.5281/zenodo.21703
Bignone L. A., Pedrosa S. E., Trayford J. W., Tissera P. B., Pellizza L. J., 2020, MNRAS, 491, 3624
Booth C. M., Schaye J., 2009, MNRAS, 398, 53
Borlaff A. et al., 2014, A&A, 570, A103
Bottrell C., Torrey P., Simard L., Ellison S. L., 2017, MNRAS, 467, 2879
Bottrell C. et al., 2019, MNRAS, 490, 5390
Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, MNRAS, 465, 32
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Bryant J. J. et al., 2015, MNRAS, 447, 2857
Butcher H., Oemler A. J., 1978, ApJ, 219, 18
Butcher H., Oemler A. J., 1984, ApJ, 285, 426
Cappellari M. et al., 2007, MNRAS, 379, 418
Cappellari M. et al., 2011, MNRAS, 416, 1680
Cavanagh M. K., Bekki K., Groves B. A., 2021, MNRAS, 506, 659
Cavanagh M. K., Bekki K., Groves B. A., Pfeffer J., 2022, MNRAS, 510, 5164
Clauwens B., Schaye J., Franx M., Bower R. G., 2018, MNRAS, 478, 3994
Coccato L. et al., 2020, MNRAS, 492, 2955
Conroy C., Gunn J. E., 2010, ApJ, 712, 833
Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
Correa C. A., Schaye J., Clauwens B., Bower R. G., Crain R. A., Schaller M., Theuns T., Thob A. C. R., 2017, MNRAS, 472, L45
Correa C. A., Schaye J., Trayford J. W., 2019, MNRAS, 484, 4401
Cortese L. et al., 2016, MNRAS, 463, 170
Couch W. J., Ellis R. S., Sharples R. M., Smail I., 1994, ApJ, 430, 121
Couch W. J., Barger A. J., Smail I., Ellis R. S., Sharples R. M., 1998, ApJ, 497, 188
Crain R. A. et al., 2015, MNRAS, 450, 1937
Crain R. A. et al., 2017, MNRAS, 464, 4204
Croom S. M. et al., 2021, MNRAS, 505, 2247
D'Eugenio F., Houghton R. C. W., Davies R. L., Dalla Bontà E., 2013, MNRAS, 429, 1258
Dalla Vecchia C., Schaye J., 2012, MNRAS, 426, 140
Davies J. J., Crain R. A., McCarthy I. G., Oppenheimer B. D., Schaye J., Schaller M., McAlpine S., 2019, MNRAS, 485, 3783
Davies J. J., Crain R. A., Oppenheimer B. D., Schaye J., 2020, MNRAS, 491, 4462
Davies J. J., Pontzen A., Crain R. A., 2022, MNRAS, 515, 1430
Davis M., Geller M. J., 1976, ApJ, 208, 13
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371
Davison T. A., Norris M. A., Pfeffer J. L., Davies J. J., Crain R. A., 2020, MNRAS, 497, 81
de Graaff A., Trayford J., Franx M., Schaller M., Schaye J., van der Wel A., 2021, MNRAS, 511, 2544
De Lucia G., Springel V., White S. D. M., Croton D., Kauffmann G., 2006, MNRAS, 366, 499
Deeley S. et al., 2020, MNRAS, 498, 2372
Deeley S., Drinkwater M. J., Sweet S. M., Bekki K., Couch W. J., Forbes D. A., Dolfi A., 2021, MNRAS, 508, 895
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Diaz J., Bekki K., Forbes D. A., Couch W. J., Drinkwater M. J., Deeley S., 2018, MNRAS, 477, 2030
Dieleman S., Willett K. W., Dambre J., 2015, MNRAS, 450, 1441
Dolag K., Borgani S., Murante G., Springel V., 2009, MNRAS, 399, 497
Dolfi A., Forbes D. A., Couch W. J., Bekki K., Ferré-Mateu A., Romanowsky A. J., Brodie J. P., 2021, MNRAS, 504, 4923
Dressler A., 1980, ApJ, 236, 351
Dressler A. et al., 1997, ApJ, 490, 577

Eliche-Moral M. C., González-García A. C., Aguerri J. A. L., Gallego J., Zamorano J., Balcells M., Prieto M., 2012, A&A, 547, A48
Eliche-Moral M. C., González-García A. C., Aguerri J. A. L., Gallego J., Zamorano J., Balcells M., Prieto M., 2013, A&A, 552, A67
Emsellem E. et al., 2007, MNRAS, 379, 401
Emsellem E. et al., 2011, MNRAS, 414, 888
Faber S. M., Gallagher J. S., 1976, ApJ, 204, 365
Falcón-Barroso J. et al., 2019, A&A, 632, A59
Fasano G., Poggianti B. M., Couch W. J., Bettoni D., Kjærgaard P., Moles M., 2000, ApJ, 542, 673
Fasano G. et al., 2015, MNRAS, 449, 3927
Ferrarese L., Merritt D., 2000, ApJ, 539, L9
Fraser-McKelvie A., Aragón-Salamanca A., Merrifield M., Tabor M., Bernardi M., Drory N., Parikh T., Argudo-Fernández M., 2018, MNRAS, 481, 5580
Furlong M. et al., 2015, MNRAS, 450, 4486
Furlong M. et al., 2017, MNRAS, 465, 722
Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., 1998, MNRAS, 300, 146
Glorot X., Bordes A., Bengio Y., 2011, in Gordon G., Dunson D., Dudík M., eds, Proc. Machine Learning Research Vol. 15, Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics. PMLR, Fort Lauderdale, FL, USA, p. 315
Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, MNRAS, 374, 1479
Gunn J. E., Gott J. R. I., 1972, ApJ, 176, 1
Hani M. H., Sparre M., Ellison S. L., Torrey P., Vogelsberger M., 2018, MNRAS, 475, 1160
Hernquist L., 1989, Nature, 340, 687
Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, ApJ, 691, 1168
Houghton R. C. W., 2015, MNRAS, 451, 3427
Houghton R. C. W. et al., 2013, MNRAS, 436, 19
Huertas-Company M. et al., 2019, MNRAS, 489, 1859
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Khochfar S. et al., 2011, MNRAS, 417, 845
Koopmann R. A., Kenney J. D. P., 2004a, ApJ, 613, 851
Koopmann R. A., Kenney J. D. P., 2004b, ApJ, 613, 866
Koribalski B. S. et al., 2020, Ap&SS, 365, 118
Kormendy J., Richstone D., 1995, ARA&A, 33, 581
Lagos C. d. P. et al., 2015, MNRAS, 452, 3815
Lagos C. d. P. et al., 2018a, MNRAS, 473, 4956
Lagos C. d. P., Schaye J., Bahé Y., van de Sande J., Kay S. T., Barnes D., Davis T. A., Dalla Vecchia C., 2018b, MNRAS, 476, 4327
Lagos C. d. P., Emsellem E., van de Sande J., Harborne K. E., Cortese L., Davison T., Foster C., Wright R. J., 2022, MNRAS, 509, 4372
Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692
LeCun Y., Bengio Y., Hinton G., 2015, Nature, 521, 436
Lewis I. et al., 2002, MNRAS, 334, 673
Ludlow A. D., Schaye J., Schaller M., Richings J., 2019, MNRAS, 488, L123
Magorrian J. et al., 1998, AJ, 115, 2285
Marasco A., Crain R. A., Schaye J., Bahé Y. M., van der Hulst T., Theuns T., Bower R. G., 2016, MNRAS, 461, 2630
McAlpine S. et al., 2016, Astron. Comput., 15, 72
McAlpine S., Harrison C. M., Rosario D. J., Alexander D. M., Ellison S. L., Johansson P. H., Patton D. R., 2020, MNRAS, 494, 5713
Mihos J. C., Hernquist L., 1996, ApJ, 464, 641
Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613
Moore B., Lake G., Katz N., 1998, ApJ, 495, 139
Moore B., Lake G., Quinn T., Stadel J., 1999, MNRAS, 304, 465
Naab T., Burkert A., 2003, ApJ, 597, 893
Naab T., Johansson P. H., Ostriker J. P., 2009, ApJ, 699, L178
Nair P. B., Abraham R. G., 2010, ApJS, 186, 427
Oemler Augustus J., 1974, ApJ, 194, 1
Oppenheimer B. D. et al., 2016, MNRAS, 460, 2157
Oppenheimer B. D. et al., 2020, MNRAS, 491, 2939
Pfeffer J., Bekki K., Couch W. J., Koribalski B. S., Forbes D. A., 2022, MNRAS, 511, 1072

Planck Collaboration XVI, 2014, A&A, 571, A16

Poggianti B. M. et al., 2017, Nature, 548, 304

Postman M., Geller M. J., 1984, ApJ, 281, 95

Postman M. et al., 2005, ApJ, 623, 721

Qu Y. et al., 2017, MNRAS, 464, 1659

Querejeta M. et al., 2015, A&A, 579, L2

Quinn P. J., Hernquist L., Fullagar D. P., 1993, ApJ, 403, 74

Rahmati A., Schaye J., Bower R. G., Crain R. A., Furlong M., Schaller M., Theuns T., 2015, MNRAS, 452, 2034

Reynolds T. N. et al., 2022, MNRAS, 510, 1716

Ricarte A., Tremmel M., Natarajan P., Quinn T., 2020, ApJ, 895, L8

Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006, ApJ, 645, 986

Robotham A. S. G. et al., 2014, MNRAS, 444, 3986

Rodriguez-Gomez V. et al., 2016, MNRAS, 458, 2371

Rodriguez-Gomez V. et al., 2019, MNRAS, 483, 4140

Rosas-Guevara Y. M. et al., 2015, MNRAS, 454, 1038

Savorgnan G. A. D., Graham A. W., Marconi A., Sani E., 2016, ApJ, 817, 21

Schaefer C., Geiger M., Kuntzer T., Kneib J. P., 2018, A&A, 611, A2

Schaye J., Dalla Vecchia C., 2008, MNRAS, 383, 1210

Schaye J. et al., 2015, MNRAS, 446, 521

Serra P. et al., 2012, MNRAS, 422, 1835

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Shimakawa R., Tanaka T. S., Toshikage S., Tanaka M., 2021, PASJ, 73, 1575

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Smith G. P., Treu T., Ellis R. S., Moran S. M., Dressler A., 2005, ApJ, 620, 78

Snyder G. F. et al., 2015, MNRAS, 454, 1886

Somerville R. S., Davé R., 2015, ARA&A, 53, 51

Spitzer Lyman J., Baade W., 1951, ApJ, 113, 413

Springel V., 2005, MNRAS, 364, 1105

Springel V., Hernquist L., 2005, ApJ, 622, L9

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726

Springel V., Di Matteo T., Hernquist L., 2005, ApJ, 620, L79

Srivastava N., Hinton G., Krizhevsky A., Sutskever I., Salakhutdinov R., 2014, J. Mach. Learn. Res., 15, 1929

Tacchella S. et al., 2019, MNRAS, 487, 5416

Tapia T., Eliche-Moral M. C., Aceves H., Rodríguez-Pérez C., Borlaff A., Querejeta M., 2017, A&A, 604, A105

Thob A. C. R. et al., 2019, MNRAS, 485, 972

Tobin J., Fong R., Ray A., Schneider J., Zaremba W., Abbeel P., 2017, IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). p. 23

Toomre A., Toomre J., 1972, ApJ, 178, 623

Toth G., Ostriker J. P., 1992, ApJ, 389, 5

Trayford J. W. et al., 2015, MNRAS, 452, 2879

Trayford J. W. et al., 2017, MNRAS, 470, 771

Trayford J. W., Frenk C. S., Theuns T., Schaye J., Correa C., 2019, MNRAS, 483, 744

Treu T., Ellis R. S., Kneib J.-P., Dressler A., Smail I., Czoske O., Oemler A., Natarajan P., 2003, ApJ, 591, 53

Turner M. L., Schaye J., Crain R. A., Theuns T., Wendt M., 2016, MNRAS, 462, 2440

van de Sande J. et al., 2021, MNRAS, 508, 2307

van der Wel A. et al., 2007, ApJ, 670, 206

van der Wel A., Bell E. F., Holden B. P., Skibba R. A., Rix H.-W., 2010, ApJ, 714, 1779

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Rev. Phys., 2, 42

Walker I. R., Mihos J. C., Hernquist L., 1996, ApJ, 460, 121

Whitmore B. C., Gilmore D. M., 1991, ApJ, 367, 64

Whitmore B. C., Gilmore D. M., Jones C., 1993, ApJ, 407, 489

Wiersma R. P. C., Schaye J., Smith B. D., 2009a, MNRAS, 393, 99

Wiersma R. P. C., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, MNRAS, 399, 574

York D. G. et al., 2000, AJ, 120, 1579

APPENDIX: CNN ACCURACY TESTING

To test the accuracy of the CNN (Section 2.3) when applied to simulated galaxy images, in this appendix we compare the CNN classifications with human classifications of the same galaxies. As previously, the CNN classifications were performed using a single face-on image for each galaxy. However, both face-on and edge-on images of the galaxies were used for the human visual classifications to help determine the 'true' morphology of each galaxy. In particular, we aim to distinguish between spheroid-dominated (E) and disc-dominated (S0) early-type galaxies to compare with the classifications from the CNN. The issue of distinguishing E and S0 galaxies from single images also exists in classifications of observed galaxies, where galaxy kinematics are necessary to distinguish between rotating and non-rotating early-type galaxies (e.g. see Cappellari et al. 2007; Emsellem et al. 2007, 2011).

Using the process described in Section 2.2, images were generated for galaxies with $M_* > 10^{10}\,\mathrm{M_\odot}$ in the $z = 0$ snapshot of the EAGLE RefL050N752 simulation of a $50^3\,\mathrm{cMpc^3}$ volume (which uses an identical resolution and subgrid model parameters as the largest EAGLE volume, Schaye et al. 2015). This gives us a volume-limited sample of 475 galaxies, which are independent from the simulation used for the main results of the paper (EAGLE RefL100N1504).

Four of the authors (JP, WJC, MJD, DAF) classified each galaxy by visual inspection under an agreed set of criteria. The classification tree can be summarized as follows:

(i) Does the galaxy have prominent spiral or star formation features (i.e. is it a late-type galaxy)?

    (a) If yes, does the galaxy have a regular disc component?

        (1) If yes ⇒ **Spiral (Sp)**.
        (2) If no ⇒ **Irregular (Irr)**.

    (b) If no, does the galaxy have a regular disc component?

        (1) If yes ⇒ **Lenticular (S0)**.
        (2) If no, does the galaxy have an irregular morphology?

            (A) If yes ⇒ **Irregular (Irr)**.
            (B) If no ⇒ **Elliptical (E)**.

A representative set of images (around 20 of each type) were selected and classified by all four authors to serve as a reference. The full set of galaxies were then classified individually by each of the four authors. This classification scheme uses fewer categories than that of Nair & Abraham (2010), who used the Hubble T-Type system. However, it is intended to be broadly consistent with the down-sampled categories used in the CNN analysis (e.g. spirals are defined as all types from Sa to Sm).

Fig. A1 shows the confusion matrices for the CNN compared against the four sets of visual classifications. Generally, galaxies classified as early- (E, S0) or late-type (Sp, Irr) in visual classifications are also classified as such by the CNN. However, visually classified ellipticals tend to be classified as lenticular galaxies by the CNN, while visually classified irregulars tend to be classified as spiral galaxies by the CNN. The classification accuracy for spiral and lenticular galaxies is in the range of 50–70 per cent, depending on set of visual classifications.

We investigate the classification accuracy further in Fig. A2 by comparing the galaxy morphology-mass relations (as in Fig. 2). The
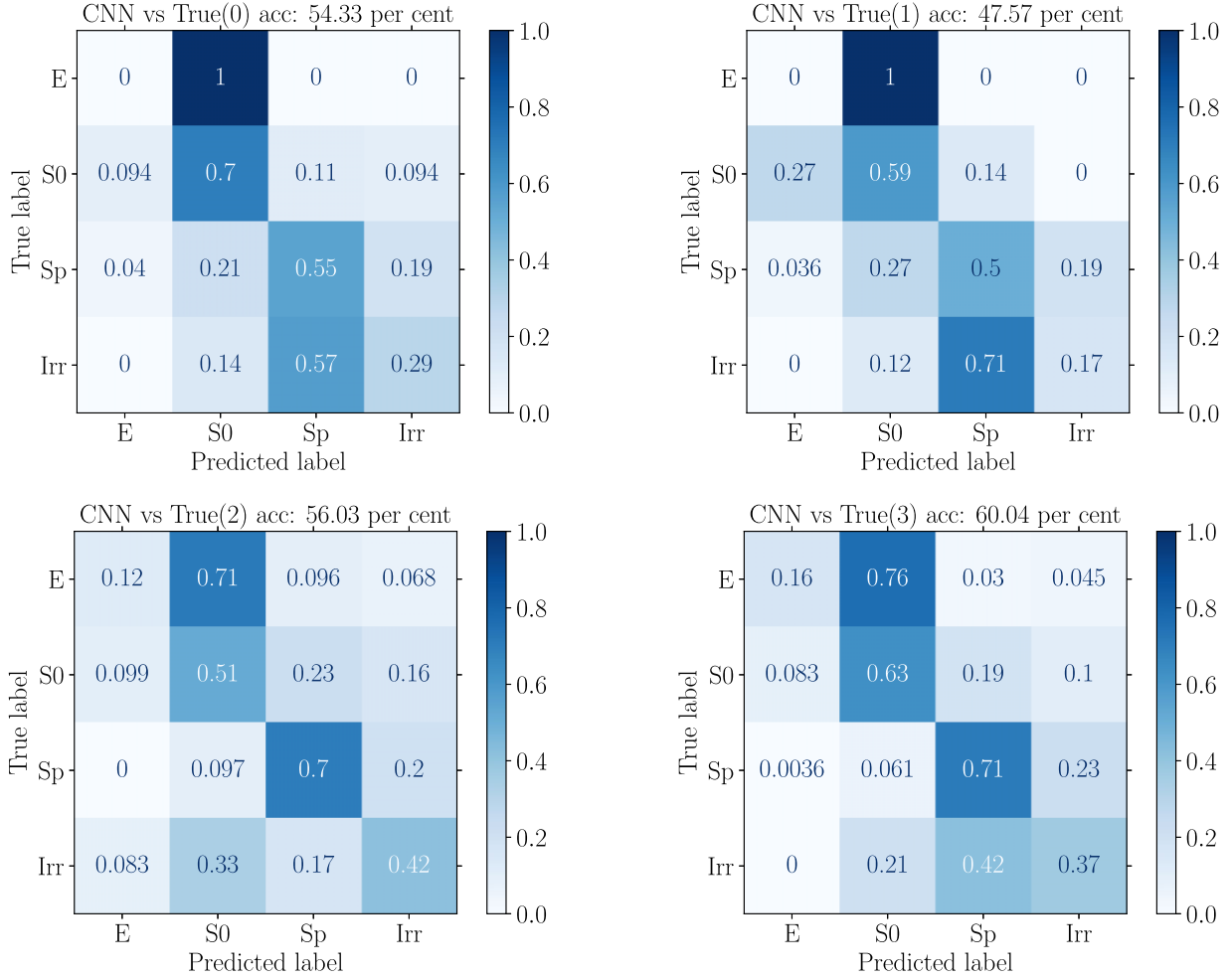
Figure A1. Confusion matrices for the CNN classifications (predicted labels) compared with each set of visual ('true') classifications of the galaxies from EAGLE RefL050N752. The overall accuracies are indicated in the title of each subpanel.
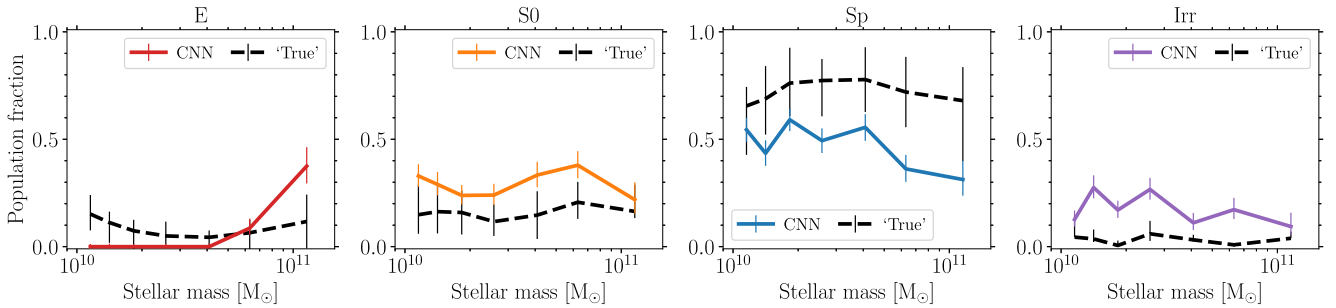
Figure A2. Dependence of morphology on galaxy stellar mass for galaxies in the EAGLE RefL050N752 simulation. Panels from left to right show the fraction of elliptical (E), lenticular (S0), spiral (Sp), and irregular (Irr) galaxies as a function of galaxy mass. Coloured lines show classifications from the CNN, with errorbars showing the uncertainties from binomial statistics. Black dashed lines show the average morphological fractions of four visual classifications of the same galaxies, with errorbars showing the full range of individual classification. Each mass bin contains more than 30 galaxies.

coloured lines show the results for the CNN (with errorbars showing binomial uncertainties), while black dashed lines and errorbars show the mean and range for visual classifications. Between individual authors there exists some systematic offset in the classified fractions of early- and late-type galaxies, which is most evident in the range of spiral fractions (as well as Fig. A1). This reflects the difficulty in classifying galaxies that may be intermediate between S0 and Sp (e.g. Fig. A3).

Below we discuss the key deviations between CNN and visual classifications:

(i) *No lower-mass ($M_* < 5 \times 10^{10} \, \mathrm{M_\odot}$) galaxies are classified by the CNN as elliptical.* Around 9 per cent of such galaxies are classified as elliptical by visual inspection. These galaxies are generally classified by the CNN as S0 galaxies (>80 per cent). In Section 3.1, we suggest that this issue may be due to insufficient
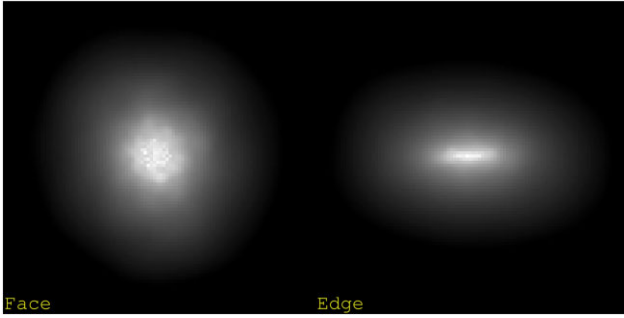
Figure A3. Example of a galaxy where visual classifications were evenly split between S0 and Sp. This galaxy was classified by the CNN as an S0. The left-hand and right-hand panels show the face-on and edge-on projections, respectively.
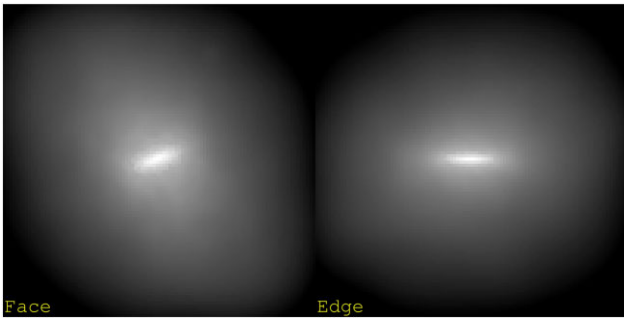

Figure A4. Example of a galaxy classified as an elliptical by the CNN and as a S0 by visual classification. The left-hand and right-hand panels show the face-on and edge-on projections, respectively.

Figure A5. Example of a galaxy classified as an irregular by the CNN and as a spiral by visual classification. The left-hand and right-hand panels show the face-on and edge-on projections, respectively.

Figure A6. Example of a high-mass ($M_* \approx 10^{11}\,\mathrm{M_\odot}$) galaxy classified as a S0 by the CNN and as a late-type by visual classification. The left-hand and right-hand panels show the face-on and edge-on projections, respectively.

concentration of spheroidal galaxies due to numerical effects. To test this, we fit Sérsic (1963) profiles to the g-band profiles of each galaxy (using radially-averaged face-on projections) to compare the Sérsic indices (n_s). For galaxies with $M_* < 5 \times 10^{10}\,\mathrm{M_\odot}$, we find those visually classified (in the majority) as E and S0 have median $n_s \approx 2.2$ and 1.5, respectively. In contrast, high-mass galaxies ($M_* > 5 \times 10^{10}\,\mathrm{M_\odot}$) classified by the CNN as E and S0 have median $n_s \approx 3.8$ and 2.7, respectively. Therefore, lower-mass Es have less concentrated profiles than high mass S0s, which may be leading the CNN to classify the low-mass Es as S0s.

(ii) *The CNN classifies more high-mass galaxies as elliptical than visual classifications.* This difference is a result of the visual classifications making use of edge-on projections to distinguish genuine elliptical galaxies from lenticular galaxies. We show an example of such a galaxy classified as an elliptical by the CNN and a lenticular in visual classifications in Fig. A4. This issue is related (though not identical) to the finding of a lower fraction of slow-rotating (i.e. genuinely spheroidal) galaxies in 'kinematic' morphologies compared to the fraction of elliptical galaxies in visual morphologies (e.g. Emsellem et al. 2007, 2011; Cappellari et al. 2011). However, many of the visually classified elliptical galaxies may still be fast-rotating galaxies. For the visual classifications, 25–40 per cent of early-type galaxies were classified as elliptical. This is much higher than the $\approx$14 per cent of slow-rotating early-type galaxies found for observed galaxies (Emsellem et al. 2011), despite the fraction of slow-rotating galaxies in EAGLE reasonably agreeing with observed fractions (Lagos et al. 2018b).

(iii) *The CNN classifies more galaxies as irregular than visual classifications.* Around 82 per cent of these galaxies are classified
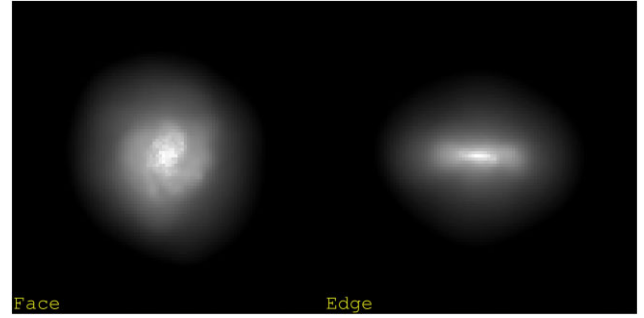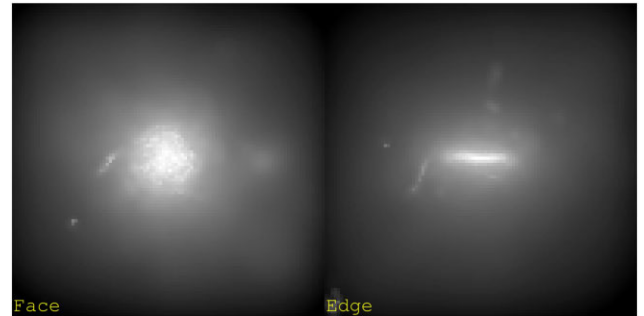
visually as spirals and 5 per cent classified as 'true' irregulars (averaging over the four visual classifications). Taken together, $\approx$92 per cent of galaxies are correctly identified by the CNN as late-type (S + I). The difference in irregular classifications is therefore likely due to difference in definition. The CNN tends to classify only galaxies with strong spiral features as Sp and galaxies with weaker features as Irr. In contrast, visual classifications also labelled galaxies with weaker star-formation features as Sp (e.g. Fig. A5), while the irregular classification is reserved for star-forming galaxies without prominent discs and galaxies with irregular shapes.

(iv) *More high-mass galaxies are classified as late-type (S + I) in visual classifications than by the CNN.* For galaxies with $M_* > 10^{11}\,\mathrm{M_\odot}$ the average late-type fraction in visual classifications is 0.7, compared to 0.4 for the CNN. However, the CNN is consistent within the full range of visual classifications, given the systematic differences between individual sets of classifications (as stated above). For high-mass galaxies in particular, classification is often complicated by compact star-forming discs, and whether more weight should be given to the central region or the lack of features in the majority of the galaxy (e.g. Fig. A6). Such galaxies tend to be labelled by the CNN as S0.

(v) *The CNN classifies more galaxies as lenticular than visual classifications.* This difference is a combination of the factors discussed above. At low masses ($M_* \lesssim 5 \times 10^{10}\,\mathrm{M_\odot}$), spheroidal galaxies are classified by the CNN as S0, accounting for 7–37 per cent of S0s. Given the smooth transition between 'genuine' lenticular and spiral galaxy types, the exact division between the morphological types is difficult to place, leading to the systematic differences in individual visual classifications. The CNN tends to label such intermediate galaxies as lenticular (e.g. Fig. A3). Overall, 50–

70 per cent of galaxies classified as S0 by visual inspection were also classified by the CNN as S0 (Fig. A1). This level of disagreement in the classification of S0 galaxies is relatively similar to that found for observed samples of galaxies (e.g. see fig. 14 in Nair & Abraham 2010).

To summarize, overall the CNN can reasonably accurately distinguish early- and late-type galaxies. Combining the classifications into early- (E + S0) and late-type (Sp + I) galaxies, the CNN is 76–88 per cent accurate over the simulated galaxy sample. The differences between the CNN and visual classifications (for four classes) mainly arise due to differences in how the broad types are subdivided (e.g. late types into Sp and Irr). For early-type galaxies, the differences are largely due to the difficulty in distinguishing spheroid-dominated (E) and disc-dominated (S0) galaxies from a single image, an issue that also occurs when classifying observed galaxies (c.f. Emsellem et al. 2007; Cappellari et al. 2011). For late-type galaxies, the differences appear to be due to how the spiral and irregular classes are defined, in particular whether galaxies with weaker spiral/star-formation features are classified as spiral or irregular.

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