

# The morphology of star-forming gas and its alignment with galaxies and dark matter haloes in the EAGLE simulations

Alexander D. Hill <sup>\*</sup>, Robert A. Crain , Juliana Kwan and Ian G. McCarthy 

*Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK*

Accepted 2021 April 30. Received 2021 April 27; in original form 2021 February 26

## ABSTRACT

We present measurements of the morphology of star-forming gas in galaxies from the EAGLE simulations, and its alignment relative to stars and dark matter (DM). Imaging of such gas in the radio continuum enables weak lensing experiments that complement traditional optical approaches. Star-forming gas is typically more flattened than the stars and DM within halo centres, particularly for present-day structures of total mass  $\sim 10^{12-12.5} M_{\odot}$ , which preferentially host star-forming galaxies with rotationally supported stellar discs. Such systems have oblate, spheroidal star-forming gas distributions, but in both less- and more-massive subhaloes the distributions tend to be prolate, and its morphology correlates positively and significantly with that of its host galaxy's stars, both in terms of sphericity and triaxiality. The minor axis of star-forming gas most commonly aligns with the minor axis of its host subhalo's central DM distribution, but this alignment is often poor in subhaloes with a prolate DM distribution. Star-forming gas aligns with the DM at the centre of its parent subhalo less strongly than is the case for stars, but its morphological minor axis aligns closely with its kinematic axis, affording a route to observational identification of the unsheared morphological axis. The projected ellipticities of star-forming gas in EAGLE are consistent with shapes inferred from high-fidelity radio continuum images, and they exhibit greater shape noise than is the case for images of the stars, owing to the greater characteristic flattening of star-forming gas with respect to stars.

**Key words:** gravitational lensing: weak – methods: numerical – large-scale structure of Universe – radio continuum: ISM.

## 1 INTRODUCTION

The currently preferred  $\Lambda$ -cold dark matter ( $\Lambda$ CDM) cosmogony posits that the large-scale cosmic matter distribution (spatial scales  $\gtrsim 1$  Mpc) is best described as a highly non-uniform system of voids, sheets, filaments, and haloes, colloquially termed the ‘cosmic web’. This structure forms in response to the gravitational growth of small instabilities in the matter distribution of the early Universe (e.g. Bond, Kofman & Pogosyan 1996; Faucher-Giguère, Lidz & Hernquist 2008; Shandarin, Habib & Heitmann 2010). Spectroscopic redshift surveys have revealed that galaxies are themselves distributed in a cosmic web, as expected if they broadly trace the underlying matter distribution. The non-uniform distribution of galaxies was apparent in early redshift surveys (see e.g. de Lapparent, Geller & Huchra 1986; Geller & Huchra 1989), but was demonstrated spectacularly by those exploiting the advent of highly multiplexed spectrographs, notably the 2dF Galaxy Redshift Survey (2dFGRS; Colless et al. 2003) and the Sloan Digital Sky Survey (SDSS; Tegmark et al. 2004).

A fundamental tenet of galaxy formation models is that galaxies form within the dark matter (DM) haloes that permeate the cosmic web (e.g. White & Rees 1978). The broad correspondence between the clustering of galaxies inferred from observational surveys on one hand, and on the other that of the galaxies that form in semi-analytic models of galaxy formation (e.g. Kauffmann et al. 1999;

Springel et al. 2005b; Wechsler et al. 2006; Guo et al. 2011) and, more recently, hydrodynamical simulations of large cosmic volumes (e.g. Crain et al. 2017; McCarthy et al. 2017; Springel et al. 2018), can be considered a remarkable corroboration of the  $\Lambda$ CDM paradigm. However, being subject to the rich array of dissipative physical processes that govern their growth, galaxies inevitably represent imperfect tracers of their local environment (e.g. Kaiser 1984; White et al. 1987), such that their baryonic components do not necessarily trace the shape and orientation of their DM haloes in a simple fashion.

Besides their potential use as a means to place constraints on the ill-understood microphysics of galaxy formation, and to reveal the nature of the environment of galaxies (e.g. Codis et al. 2015; Zhang et al. 2015), differences in the shape and orientation of baryonic components of galaxies with respect to those of their DM haloes are of particular interest because they represent sources of uncertainty in observational inferences of the morphology of DM haloes, and of their orientation with respect to the large-scale matter distribution (e.g. Troxel & Ishak 2015). This is of consequence for efforts to constrain cosmological parameters via the shape correlation function of galaxies, a key aim of ongoing optical/near-infrared weak lensing surveys such as the Canada–France–Hawaii Telescope Lensing Survey (CFHTLenS; Erben et al. 2013), Kilo-Degree Survey (KiDS; de Jong et al. 2015), the Hyper Suprime Cam Subaru Strategic Program (HSC; Aihara et al. 2018), and the Dark Energy Survey (DES; The Dark Energy Survey Collaboration 2005), and ambitious forthcoming surveys with the Vera Rubin Observatory (LSST Science Collaboration 2009), the *Euclid* spacecraft (Laureijs et al. 2012), and the *Nancy Grace Roman Space Telescope* (e.g.

\* E-mail: a.d.hill@2017.ljmu.ac.uk

Spergel et al. 2015). Moreover, the severity of differences between the shape and alignment of haloes and those of the observable structures used to infer them has a strong bearing on the accuracy of weak gravitational lensing predictions derived from dark matter-only simulations. At present, such simulations are the only means of modelling the evolution of cosmic volumes comparable to those mapped out by lensing surveys.

Simplified techniques such as halo occupation distribution (HOD) modelling, subhalo abundance matching (SHAM), and semi-analytic models have, in order of increasing sophistication, proven valuable means of understanding the connection between galaxies and the matter distribution (see e.g. Schneider & Bridle 2010; Joachimi et al. 2013). However, such methods have been shown to exhibit significant systematic differences with respect to the predictions of cosmological hydrodynamical simulations on small-to-intermediate spatial scales (e.g. Chaves-Montero et al. 2016; Springel et al. 2018), in large part because they (by design) do not self-consistently capture the back-reaction of baryon evolution on the structure of DM haloes (Bett et al. 2010; Guo et al. 2016). A comprehensive understanding of the influence of systematic uncertainties stemming from the differences in the shape and orientation of galaxies and their host haloes therefore requires self-consistent and realistic physical models of galaxy formation in a fully cosmological framework.

There is a rich history of the use of numerical simulations to establish the correspondence between the morphology, angular momentum, and orientation of galaxies, their satellite systems and their host DM haloes, with particular emphases on the roles played by gas accretion (e.g. Chen, Jing & Yoshikawa 2003; Sharma & Steinmetz 2005; Sales et al. 2012), mergers (e.g. Dubinski 1998; Boylan-Kolchin, Ma & Quataert 2006; Naab, Khochfar & Burkert 2006) and environment (e.g. Croft et al. 2009; Hahn, Teyssier & Carollo 2010; Shao et al. 2016). However, prior studies have tended to suffer from one or more significant shortcomings, namely relatively poor spatial and mass resolution, relatively small sample sizes, and a poor correspondence between the properties of simulated galaxies with observed counterparts. These shortcomings are significantly ameliorated by the current generation of state-of-the-art hydrodynamical simulations, such as EAGLE (Crain et al. 2015; Schaye et al. 2015), HorizonAGN (Dubois et al. 2014), Illustris/IllustrisTNG (e.g. Vogelsberger et al. 2014; Pillepich et al. 2018) and MassiveBlack-II (Khandai et al. 2015). Each of these simulations broadly reproduces key observed properties of the present-day galaxy population, thus engendering confidence that they capture (albeit with varying degrees of accuracy) the complexity of the interaction between the baryonic components of galaxies and their DM haloes. The simulations each follow a cosmological volume sufficient to yield representative galaxy population ( $\sim 100^3 \text{ cMpc}^3$ ), and do so with a mass resolution ( $\sim 10^6 M_\odot$ ) and spatial resolution ( $\sim 1 \text{ pkpc}$ ) that enables examination of the properties and evolution of even sub- $L^*$  galaxies. Moreover, they capture important second-order effects such as the back-reaction of baryons on the structure and clustering of DM haloes.

The emergence of optical weak lensing surveys as a promising means of constraining the nature of DM and dark energy has intensified the need to assess the severity of systematic uncertainties afflicting cosmic shear measurements (specifically, the galaxy shape correlation function). Cosmological hydrodynamical simulations have proven a valuable tool for this purpose, highlighting that galaxies can be significantly misaligned with respect to their DM haloes (e.g. Bett et al. 2010; Hahn et al. 2010; Bett 2012; Tenneti et al. 2014; Velliscig et al. 2015a; Shao et al. 2016; Chisari et al. 2017) and that the shapes and alignments of galaxies and their

haloes are correlated over large distances via tidal forces (Tenneti et al. 2014; Chisari et al. 2015; Codis et al. 2015; Velliscig et al. 2015b). The simulations have also been exploited to examine the morphological and kinematic alignment of galaxies with the cosmic large-scale structure (see e.g. Cuesta et al. 2008; Codis et al. 2018). The current generation of state-of-the-art simulations remains reliant on the use of subgrid treatments of many of the key physical processes governing galaxy evolution and, as noted by Joachimi et al. (2015) the details of their particular implementation can in principle influence the alignment of cosmic structures (see also Velliscig et al. 2015a). However, in key respects the simulations appear to be quantitatively compatible with extant observational constraints, e.g. the  $w_{g+}$  correlation function of luminous red galaxies in SDSS and their analogues in the MassiveBlack-II simulation (Tenneti et al. 2015, their fig. 21).

A complementary approach to optical/near-IR weak lensing surveys is to measure shear at radio frequencies. The concept has been demonstrated both by exploiting very large area, low source density radio data (Chang, Refregier & Helfand 2004), and deep, pointed observations with greater source density (Patel et al. 2010). Ambitious future radio continuum surveys such as those envisaged for the Square Kilometre Array (SKA) may prove to be competitive with the largest optical surveys. An SKA Phase-1 continuum survey of  $5000 \text{ deg}^2$  is predicted to observe a source density of resolved star-forming galaxies of  $2.7 \text{ arcmin}^{-2}$  (Square Kilometre Array Cosmology Science Working Group et al. 2020). Brown et al. (2015) argue that, in the most optimistic case, a full Phase-2 SKA survey over  $3\pi$  steradians would yield twice the areal coverage of the *Euclid* ‘wide survey’, with a similar source density of  $\simeq 30$  galaxies  $\text{arcmin}^{-2}$ .

The characteristic redshift of sensitive radio continuum surveys may also prove to be significantly greater than that of optical counterparts. By bridging the gap between traditional shear measurements and those derived from maps of the cosmic microwave background (CMB) radiation, radio weak lensing surveys offer the promise of tomographic mapping of cosmic structure evolution in both the quasi-linear and strongly non-linear regimes. Shear mapping in the radio regime offers advantageous complementarity with optical surveys, in particular to suppress key systematic uncertainties. For example, the use of kinematic and/or polarization information may enable improved characterization of the intrinsic (unsheared) ellipticity, and suppress the influence of intrinsic alignment, the deviation from random of the observed ellipticity of a sample (Blain 2002; Morales 2006; de Burgh-Day et al. 2015; Whittaker, Brown & Battye 2015).

Shear measurements in the radio regime are derived from images of the extended radio continuum emission from galaxies, which effectively traces the star-forming component of the interstellar medium (ISM). The morphology and kinematics of this component, and their relationship with those of the underlying DM distribution, can in principle differ markedly from the analogous quantities traced by the stellar component imaged by conventional lensing surveys. However, by design, leading models of the radio continuum sky (e.g. Wilman et al. 2008; Bonaldi et al. 2019) do not account for such differences. This therefore motivates an extension of prior examinations of the relationship between galaxies and the overall matter distribution, and correlation of shapes and alignments of galaxies separated over cosmic distances, focusing on the use of the star-forming ISM to characterize the morphology and orientation of galaxies. The current generation of state-of-the-art cosmological hydrodynamical simulations are well suited to this application since, as for the stellar component, they self-consistently model the evolution of star-forming gas within galaxies, including cosmological accretion

from the intergalactic and circumgalactic media (IGM and CGM, respectively), expulsion by feedback processes, and its interaction with a dynamically ‘live’ DM halo.

In this study, we use the cosmological hydrodynamical simulations of the EAGLE project (Crain et al. 2015; Schaye et al. 2015) to examine the correspondence between the morphology and orientation of the star-forming ISM of galaxies and those of their parent DM haloes. EAGLE is well suited to this application: although the simulations do not explicitly model the balance between molecular, atomic, and ionized hydrogen, the use of empirical or theoretical models to partition gas into these phases indicates that the simulations broadly reproduce key properties of the atomic and molecular reservoirs of galaxies (see e.g. Lagos et al. 2015; Bahé et al. 2016; Crain et al. 2017; Davé et al. 2020) including, crucially, the ‘Fundamental Plane of star formation’ that relates their stellar mass, star formation rate, and neutral hydrogen fraction (Lagos et al. 2016). This study complements prior examinations of the morphology of stars, hot gas, and DM in the EAGLE simulations (e.g. Velliscig et al. 2015a, b; Shao et al. 2016). The morphology of the star-forming ISM of galaxies in the IllustrisTNG-50 simulation (hereafter TNG50) was also examined by Pillepich et al. (2019); whilst the motivation for that study was quite different to that of ours, their findings are of direct relevance and offer an opportunity to assess the degree of consensus between different simulations.

This paper is structured as follows. We discuss our numerical methods in Section 2, as well as summarizing briefly details of the EAGLE simulation and galaxy finding algorithms, and our sample selection criteria. In Section 3, we examine the morphology of star-forming gas and its dependence on subhalo mass and redshift. In Section 4, we examine the internal alignment of star-forming gas with DM and stars, and its mutual alignment with its kinematic axis, again as a function of subhalo mass and redshift. In Section 5, we investigate the shapes and alignments of the various matter components in 2D. In Section 6, we discuss and summarize our findings. In a series of appendices, we examine the influence of a series of numerical and modelling factors on our findings.

## 2 METHODS

In this section, we briefly introduce the EAGLE simulation (Section 2.1) and key numerical techniques for identifying haloes and subhaloes (Section 2.2), and for characterizing their morphology with shape parameters (Section 2.3). Our sample selection criteria are discussed in Section 2.4. Detailed descriptions of the simulations are provided by many other studies using them, so we present only a concise summary of the most relevant aspects and refer the interested reader to the project’s reference articles (Crain et al. 2015; Schaye et al. 2015).

### 2.1 Simulations

The EAGLE project (the Evolution and Assembly of GaLaxies and their Environments) comprises a suite of hydrodynamical simulations that model the formation and evolution of galaxies and the cosmic large-scale structure in a  $\Lambda$ CDM cosmogony (Crain et al. 2015; Schaye et al. 2015). Particle data, and derived data products, from the simulations have been released to the community as detailed by McAlpine et al. (2016). The simulations were evolved with a modified version of the Tree-Particle-Mesh (TreePM) smoothed particle hydrodynamics (SPH) solver GADGET-3 (last described by Springel 2005). The main modifications include the implementation of the pressure-entropy formulation of SPH introduced by Hopkins

(2013), a time-step limiter as proposed by Durier & Dalla Vecchia (2012), switches for artificial viscosity and artificial conduction, as per Cullen & Dehnen (2010) and Price (2008), respectively, and the use of the Wendland (1995)  $C^2$  smoothing kernel. The influence of these developments on the properties of the galaxy population yielded by the simulations is explored by Schaller et al. (2015).

EAGLE includes subgrid treatments of several physical processes that are unresolved by the simulations. These include element-by-element radiative heating and cooling of 11 species (Wiersma, Schaye & Smith 2009a) in the presence of a spatially uniform, temporally evolving UV/X-ray background radiation field (Haardt & Madau 2001) and the cosmic microwave background (CMB); a model for the treatment of the multiphase ISM as a single-phase fluid with a polytropic pressure floor (Schaye & Dalla Vecchia 2008); a metallicity-dependent density threshold above which gas becomes eligible for star formation (Schaye 2004), with a probability of conversion dependent on the gas pressure (Schaye & Dalla Vecchia 2008); stellar evolution and mass-loss (Wiersma et al. 2009b); the seeding of BHs and their growth via gas accretion and mergers (Springel, Di Matteo & Hernquist 2005a; Booth & Schaye 2009; Rosas-Guevara et al. 2015); and feedback associated with the formation of stars (Dalla Vecchia & Schaye 2012) and the growth of BHs (Booth & Schaye 2009; Schaye et al. 2015). The simulations adopt the stellar initial mass function (IMF) of Chabrier (2003). The efficiency of stellar feedback was calibrated to reproduce the stellar mass function of the low-redshift galaxy population and, broadly, the sizes of local disc galaxies. The efficiency of AGN feedback was calibrated to reproduce the present-day scaling relation between the stellar mass and central black hole mass of galaxies. The gaseous properties of galaxies and their haloes were not considered during the calibration.

EAGLE adopts values of the cosmological parameters derived from the initial Planck data release (Planck Collaboration XVI 2014), namely  $\Omega_0 = 0.307$ ,  $\Omega_b = 0.04825$ ,  $\Omega_\Lambda = 0.693$ ,  $\sigma_8 = 0.8288$ ,  $n_s = 0.9611$ ,  $h = 0.6777$ ,  $Y = 0.248$ . Our analyses focus primarily on the EAGLE simulation of the largest cosmic volume, Ref-L100N1504, which follows a cubic periodic volume of side  $L = 100$  cMpc, realized with  $N = 1504^3$  collision-less DM particles of mass  $m_{\text{DM}} = 9.7 \times 10^6 M_\odot$ , and an initially equal number of baryonic particles of mass  $m_g = 1.81 \times 10^6 M_\odot$ . The Plummer-equivalent gravitational softening length is  $1/25$  of the mean inter-particle separation ( $\epsilon_{\text{com}} = 2.66$  kpc), limited to a maximum proper length of  $\epsilon_{\text{com}} = 0.7$  pkpc. We explore the numerical convergence of the morphology and orientation of the star-forming gas component of galaxies in Appendix A, using the pair of high-resolution L025N0752 EAGLE simulations introduced by Schaye et al. (2015). These follow a cosmic volume of  $L = 25$  cMpc realized with  $N = 752^3$  particles of each species, with masses  $m_{\text{DM}} = 1.21 \times 10^6 M_\odot$  and  $m_g = 2.26 \times 10^5 M_\odot$ . For these simulations, the Plummer-equivalent gravitational softening length is  $\epsilon_{\text{com}} = 1.33$  kpc, limited to a maximum proper length of  $\epsilon_{\text{com}} = 0.35$  pkpc. The first of these simulations, Ref-L025N0752, uses the same ‘Reference’ subgrid model parameters as the Ref-L100N1504, whilst the second, Recal-L025N0752, uses a model whose parameters were recalibrated to achieve a better match to the calibration diagnostics. A summary of the simulations used in this paper are given in Table 1.

The standard-resolution simulations marginally resolve the Jeans scales at the density threshold for star formation in the warm and diffuse photoionized ISM. They hence lack the resolution to model the cold, dense phase of the ISM explicitly, and so impose a temperature floor to inhibit the unphysical fragmentation of star-forming gas. This floor takes the form  $T_{\text{eos}}(\rho)$ , corresponding to the



**Table 1.** The box sizes and resolution details of the EAGLE simulations used in this study. The columns are: comoving box side length,  $L$ ; number of DM particles (there is initially an equal number of baryon particles); the initial baryon particle mass; the Plummer-equivalent gravitational softening length in comoving units; the maximum proper softening length.

Identifier	$L$ (cMpc)	$N$	$m_g$ ( $M_\odot$ )	$\epsilon_{\text{com}}$ (ckpc)	$\epsilon_{\text{phys}}$ (pkpc)
L025N0376	25	376 <sup>3</sup>	$1.81 \times 10^6$	2.66	0.70
L025N0752	25	752 <sup>3</sup>	$2.26 \times 10^5$	1.33	0.35
L100N1504	100	1504 <sup>3</sup>	$1.81 \times 10^6$	2.66	0.70

equation of state  $P_{\text{eos}} \propto \rho_g^{4/3}$  normalized to  $T_{\text{eos}} = 8 \times 10^3 \text{ K}$  at  $n_{\text{H}} = 10^{-1} \text{ cm}^{-3}$ . The temperature of star-forming gas thus reflects the effective pressure of the ISM, rather than its actual temperature. A drawback of the use of this floor is the suppression of the formation gas discs with scale heights much less than Jeans length of the gas on the temperature floor ( $\sim 1 \text{ pkpc}$ ). In Appendix B, we explore the sensitivity of the star-forming gas morphology to the slope of the ISM equation of state, and the normalization of the star formation law.

In a recent study, Ludlow et al. (2019) demonstrated that the scale height of discs can be artificially increased by 2-body scattering of particles with unequal mass, as is the case here since we use (initially) equal numbers of baryon and DM particles, meaning that  $m_{\text{dm}}/m_b \equiv (\Omega_0 - \Omega_b)/\Omega_b \simeq 5.4$ . The vertical support of the disc may also have physical causes, such as turbulence stemming from gas accretion and energy injection from feedback (Benítez-Llambay et al. 2018), although it is likely that these influences are artificially strong in the simulations. Therefore, we caution that both the gas and stellar discs of galaxies in EAGLE are generally thicker than their counterparts in nature (see also Trayford et al. 2017). We note however that these effects are unlikely to influence significantly the mutual alignment of the stellar and gaseous discs, nor their alignment with their parent DM halo.

## 2.2 Identifying and characterizing haloes, subhaloes, and galaxies

We define galaxies as the cold baryonic component of gravitationally self-bound structures, identified by the application of the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009) to DM haloes first identified with the friends-of-friends (FoF) algorithm (with a linking length of 0.2 times the mean interparticle separation). Subhaloes are identified as overdense regions in the FoF halo bounded by saddle points in the density distribution. Within a given FoF halo, the subhalo comprising the particle (of any type) with the lowest gravitational potential energy is defined as the central subhalo, others are then satellites.

The position of galaxies is defined as the location of the particle in their subhalo with the lowest gravitational potential energy. The position of the central galaxy is used as a centre about which to compute the spherical overdensity mass (see Lacey & Cole 1993),  $M_{200}$ , for the adopted enclosed density contrast of 200 times the critical density,  $\rho_c$ . In general, the properties of galaxies are computed by aggregating the properties of the appropriate particles located within 30 pkpc of the galaxy centre, as this yields stellar masses comparable to those recovered within a projected circular aperture of the Petrosian radius (see Schaye et al. 2015).

## 2.3 Characterizing the morphology and orientation of galaxy components

Following Thob et al. (2019), we obtain quantitative descriptions of the morphology of galaxies and their subhaloes by modelling the spatial distribution of their constituent particles as ellipsoids, characterized by their sphericity,<sup>1</sup>  $S = c/a$ , and triaxiality,  $T = (a^2 - b^2)/(a^2 - c^2)$ , parameters, where  $a$ ,  $b$ , and  $c$  are, respectively, the moduli of the major, intermediate, and minor axes of the ellipsoid.<sup>2</sup> Therefore  $S = 0$  corresponds to a perfectly flattened (but potentially elongated) disc, and  $S = 1$  corresponds to a perfect sphere, whilst low and high values of  $T$  correspond, respectively, to oblate and prolate ellipsoids.

Axis lengths are given by the square root of the eigenvalues of a matrix describing the 3D mass distribution of the particles in question. The simplest choice is the mass distribution tensor (e.g. Davis et al. 1985; Cole & Lacey 1996), defined as

$$M_{ij} = \frac{\sum_p m_p r_{p,i} r_{p,j}}{\sum_p m_p}, \quad (1)$$

where the sum runs over all particles,  $p$ , comprising the structure,  $r_{p,i}$  denotes the  $i$ -th component ( $i, j = 0, 1, 2$ ) of each particle's coordinate vector with respect to the galaxy centre, and  $m_p$  is the particle's mass. As has been widely noted elsewhere, the mass distribution tensor is often referred to as the moment of inertia tensor, as the two share common eigenvectors.

There are several well-motivated alternative choices to the mass distribution tensor and, as per Thob et al. (2019), we elect here to use an iterative form of the reduced inertia tensor (see also Dubinski & Carlberg 1991; Bett 2012; Schneider, Frenk & Cole 2012). The reduced form is advantageous because it suppresses a potentially strong influence on the tensor of structural features in the outskirts of galaxies, by down-weighting the contribution of particles at a large (ellipsoidal) radius. The use of an iterative scheme is further advantageous because it enables the scheme to adapt to particle distributions that deviate significantly from the initial particle selection. Since the latter is usually (quasi-)spherical, this is particularly relevant for strongly flattened or triaxial systems. This form of the tensor is thus

$$M_{ij}^R = \frac{\sum_p \frac{m_p}{\tilde{r}_p^2} r_{p,i} r_{p,j}}{\sum_p \frac{m_p}{\tilde{r}_p^2}}, \quad (2)$$

where  $\tilde{r}_p$  is the ellipsoidal radius, and the superscript  $R$  denotes that this is the reduced form of the tensor. In the first iteration, all particles of the relevant species within a spherical aperture of a prescribed radius,  $r_{\text{sph}}$ , are considered. This yields a initial estimate of the axis lengths ( $a, b, c$ ). In the next iteration, particles satisfying the following condition relating to the ellipsoidal distance are considered:

$$\tilde{r}_p^2 \equiv \frac{r_{p,a}^2}{\tilde{a}^2} + \frac{r_{p,b}^2}{\tilde{b}^2} + \frac{r_{p,c}^2}{\tilde{c}^2} \leq 1, \quad (3)$$

where  $r_{p,a}$ ,  $r_{p,b}$ , and  $r_{p,c}$  are the particle radii projected along the eigenvectors of the previous iteration,  $\tilde{a}$ ,  $\tilde{b}$  and  $\tilde{c}$  are the re-scaled axis lengths calculated as  $\tilde{a} = a \times r_{\text{sph}}/(abc)^{1/3}$ . This ensures the ellipsoid maintains a constant volume; in this respect, we differ from the scheme used by Thob et al. (2019), who maintained a constant

<sup>1</sup>Thob et al. (2019) used the flattening,  $\epsilon = 1 - S$ , rather than the sphericity but, as is clear from their definitions, the two are interchangeable.

<sup>2</sup>Thob et al. (2019) present publicly available Python routines for this procedure at <https://github.com/athob/morphokinematics>.

major axis length between iterations. We opt for this scheme to avoid artificial suppression of the major axis in cases of highly flattened geometry, which is more common when examining star-forming gas than is the case for stellar distributions. We note that our definition of  $\hat{r}_p^2$  differs with respect to that of Thob et al. (2019) by a factor  $\hat{a}^2$ , and that often the normalization factor  $\sum_p m_p / \hat{r}_p^2$  is not explicitly adopted in the definition of this tensor (see e.g. Bett 2012, their equation 6). The axis lengths (and by extension, the shape parameters) recovered from the use of either form of the tensor are identical.

Iterations continue until the fractional change in the axis ratios  $c/a$  and  $b/a$  falls below 1 per cent. If this criterion is not satisfied after 100 iterations, or if the number of particles enclosed by the ellipsoid falls below 10, the algorithm is deemed to have failed and the object's morphology is declared unclassified. We find a failure to converge only in cases of low particle number (e.g. subhaloes with very few gas or star particles) and, crucially, our selection criteria (Section 2.4) ensure that no subhaloes with unclassified morphologies are included in our sample.

For consistency with the aperture generally used when computing galaxy properties by aggregating particle properties (see e.g. Section 5.1.1. of Schaye et al. 2015), we adopt a radius of  $r = 30$  pkpc for the initial spherical aperture. We use this aperture for all three matter types, star-forming gas, stars, and DM, and note that for the latter, this focuses our morphology measurements towards halo centres, since haloes are in general much more extended than their cold baryons (see Section 2.5). We retain the use of this aperture for the DM component in order to focus on the DM structure local to star-forming gas discs, and note that the global morphology of DM haloes in EAGLE was presented by Velliscig et al. (2015a). In Section 5, we examine the 2D projected morphology and alignment of galaxies. When performing these measurements for star-forming gas and stars, we use an initial circular aperture of  $r = \max(30 \text{ pkpc}, 2r_{1/2, \text{SF}})$ , where  $r_{1/2, \text{SF}}$  is the half-mass radius of star-forming gas bound to the subhalo. This ensures a robust morphological characterization of the image projected by the most extended gas discs when viewed close to a face-on orientation.

Equation (2) can be generalized to be weighted by any particle variable, rather than its mass. To crudely mimic the morphology of continuum-luminous regions, when computing the tensor for star-forming gas, we weight by their star formation rate (SFR) rather than their mass, since it is well-established that the relationship between SFR and radio continuum luminosity is broadly linear (see e.g. Condon 1992; Schober, Schleicher & Klessen 2017). We do not consider radio continuum emission due to AGN, since this is not extended. Pillepich et al. (2019) recently employed a similar approach to assess the morphology and alignment of  $\text{H}\alpha$ -luminous regions of star-forming galaxies in the TNG50 simulation, via the use of the SFR as a proxy for the  $\text{H}\alpha$  luminosity. The recovered shape parameters and orientation are little changed with respect to the use of particle mass as the weighting variable, or indeed a uniform weighting, largely because the SFR of particles scales as  $\dot{m}_* \propto P^{1/5}$  for a Kennicutt–Schmidt law with index  $n_s = 1.4$  (see Schaye & Dalla Vecchia 2008) and the pressure distribution of star-forming gas particles is relatively narrow: at  $z = 0$ , the 10th and 90th percentiles of the pressure of star-forming particles in the Ref-L100N1504 volume spans less than two decades in dynamic range.

We define the orientation of galaxies and subhaloes as the unit vector parallel to the minor axis of the best-fitting ellipsoid, and hence measure the relative alignment of structures as the angle between these unit vectors. We note that it is more typical in the literature to use the unit vector parallel to the major axis; this is arguably the

best-motivated choice for describing the alignment of systems that are in general prolate (e.g. DM haloes), since in such systems the major axis is the most ‘distinct’. In contrast, it is the minor axis that is the most distinct in systems that are preferentially oblate, as is the case for a flattened disc. In Section 4.2, we examine the correspondence between the morphological and kinematic axes of the star-forming gas distribution; we define the latter as the unit vector parallel to the angular momentum vector of all star-forming gas particles located within 30 pkpc of the galaxy centre.

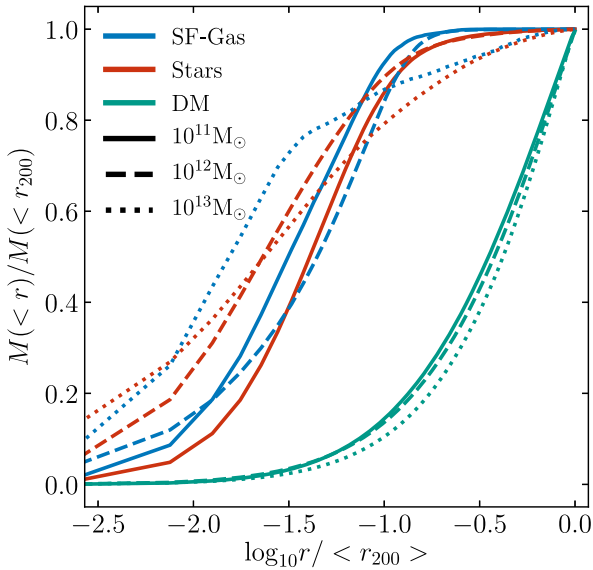
## 2.4 Sample selection

We identify subhaloes comprising a minimum of 100 each of star-forming gas particles, stellar particles, and DM particles. This numerical threshold is motivated by tests, presented in Appendix C, that assess the fractional error on shape parameters induced when performing the measurement on sub-samples, randomly selected and of decreasing size, of the particles comprising exemplar subhaloes. These tests indicate that a minimum of 100 particles are needed to recover a measurement error of the flattening of star-forming gas discs of less than 10 per cent, when using the iterative reduced inertia tensor. As noted by Thob et al. (2019), the sphericity and triaxiality shape parameters are poor descriptors of systems that deviate strongly from axisymmetry, so we excise subhaloes with strongly non-axisymmetric star-forming gas distributions. We quantify this characteristic by adapting the method of Trayford et al. (2019), binning the mass of star-forming gas into pixels of solid angle about the galaxy centre using HEALPIX (Górski et al. 2005). The asymmetry of the star-forming gas distribution,  $A_{3D}$ , is then computed by summing the (absolute) mass difference between diametrically opposed pixels and normalizing by the total star-forming gas mass. As per Trayford et al. (2019), we use coarse maps of 12 pixels, and exclude systems with  $A_{3D}^{\text{SFG}} > 0.6$ . This criterion excises 534 subhaloes, mostly of low mass, and leaves us with a sample of 6764 subhaloes at  $z = 0$ .

Our selection criteria, in particular the requirement for subhaloes to be comprised of at least 100 particles each of stars and star-forming gas, impose a strong selection bias at low halo masses. In practice, for simulations at the resolution of the EAGLE Ref-L100N1504 simulation, the criteria dictate that subhaloes host a galaxy with a minimum stellar mass of  $\sim 10^8 M_\odot$  and a minimum SFR of  $\simeq 6 \times 10^{-2} M_\odot \text{ yr}^{-1}$ , where the latter assumes the star-forming particles have a density of  $0.1 \text{ cm}^{-3}$  and pressure corresponding to a temperature of 8000 K. This corresponds to a specific star formation rate (sSFR) of  $6 \times 10^{-10} \text{ yr}^{-1}$  for the lowest (stellar) mass galaxies, a value that is above the canonical threshold separating the blue cloud of star-forming galaxies and the red sequence of quenched counterparts (e.g. Schawinski et al. 2014). Our selection criteria result in the selection of approximately (0.1, 10, 80) per cent of all subhaloes of mass  $\log_{10}(M_{\text{sub}}/M_\odot) \sim (10, 11, 12)$ , respectively, corresponding to approximately (16, 65, 60) per cent of all subhaloes of stellar mass  $\log_{10}(M_*/M_\odot) \sim (9, 10, 11)$ .

## 2.5 Mass distribution profiles

Prior studies have demonstrated that the shape and orientation of stars and DM in haloes can vary significantly as a function of radius (see e.g. Velliscig et al. 2015a). Fig. 1 shows the mean, spherically averaged, cumulative radial mass distribution profiles of the star-forming gas (blue curve), stars (red), and DM (green) comprising present-day central subhaloes with halo mass in ranges  $M_{200} \sim 10^{11} M_\odot$  (solid curves),  $10^{12} M_\odot$  (dashed) and  $10^{13} M_\odot$



**Figure 1.** Mean, spherically averaged cumulative radial mass distribution profiles of the star-forming gas (blue curve), stars (red), and DM (green) of central subhaloes within our sample that have a halo mass  $M_{200} \sim 10^{11} M_{\odot}$  (solid curve),  $10^{12} M_{\odot}$  (dashed), and  $10^{13} M_{\odot}$  (dotted). The distributions are normalized relative to the total mass of each component within  $r_{200}$ . Star-forming gas is much more centrally concentrated than dark matter at all masses.

(dotted). As might be naïvely expected, the baryonic components are much more centrally concentrated than the DM, in each of the subhalo mass bins: the median half-mass radius of star-forming gas is (3, 4.5, 1.5) per cent of  $r_{200}$  for the low, middle, and high mass bins, respectively, compared with (35, 37, 42) per cent of  $r_{200}$  for the DM.<sup>3</sup> Owing to this central concentration of the star-forming gas, we do not consider here how the shape parameters of the star-forming gas distribution change in response to the use of an initial aperture that envelops an ever-greater fraction of the virial radius.

### 3 THE MORPHOLOGY OF STAR-FORMING GAS

We begin with an examination of the morphology of star-forming gas associated with subhaloes. To illustrate visually how the method described in Section 2.3 yields shape and orientation diagnostics for the simulated galaxies, we show in Fig. 2 the star formation rate surface density,  $\Sigma_{\text{SFR}}$ , of star-forming gas (upper row), in face-on and edge-on views, and the mass surface density of stars ( $\Sigma_{\star}$ , bottom left-hand panel) and DM ( $\Sigma_{\text{DM}}$ , bottom right-hand panel) of a present-day star-forming galaxy from Recal-L025N0752. The galaxy is taken from the high-resolution Recal-L025N0752 run, and its stellar mass is  $M_{\star} = 10^{10.5} M_{\odot}$ , with a subhalo mass of  $M_{\text{sub}} = 10^{12.4} M_{\odot}$ . The galaxy’s sSFR is  $\dot{M}_{\star}/M_{\star} = 10^{-10.5} \text{ yr}^{-1}$ , and it exhibits reasonably strong rotational support: stars residing within 30 pkpc of its centre of potential have a significant fraction of their kinetic energy invested in corotation ( $\kappa_{\text{co}}^{\star} = 0.44$ ). For reference, Correa et al. (2017) argue

<sup>3</sup>The figures for the low subhalo mass bin are significantly influenced by our sample selection criteria: removal of the minimum particle number criterion results in the inclusion of systems with less-extended star-forming gas distributions, and further reduces the characteristic half-mass radius of the star-forming gas.

that  $\kappa_{\text{co}}^{\star} > 0.4$  is a useful and simple criterion for identifying star-forming disc galaxies in EAGLE. The star-forming gas exhibits a very high degree of rotation support,  $\kappa_{\text{co}}^{\text{SF}} = 0.97$ .

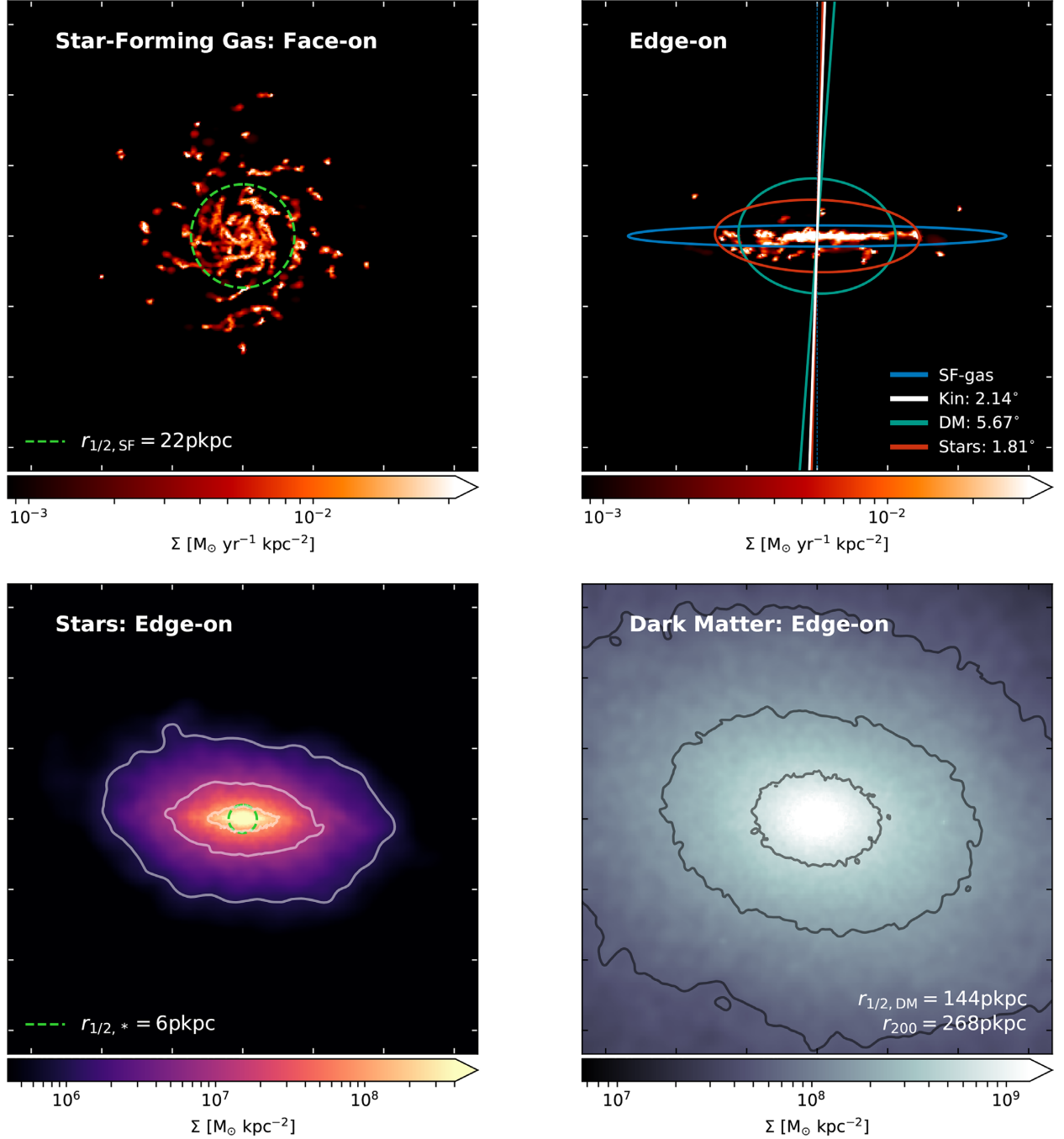
The field of view of each panel is 200 pkpc, and overlaid dashed green circles denote the half-mass radius of the matter type in question. Edge-on images are aligned such that horizontal and vertical image axes are parallel to the major and minor axes, respectively, of the star-forming gas distribution. In the upper right-hand panel, coloured ellipses correspond to projections of the best-fitting ellipsoids describing the respective matter components, whilst the solid coloured lines show the (projected) minor axes of the stellar and DM distributions, and the white line shows the projected rotation axis of the star-forming gas. Contours overlaid on the stellar and DM surface density images correspond to surface densities of  $\log_{10}(\Sigma_{\star}) = 6, 7, 8 M_{\odot} \text{ kpc}^{-2}$  and  $\log_{10}(\Sigma_{\text{DM}}) = 7.75, 8.25, 8.75 M_{\odot} \text{ kpc}^{-2}$ , respectively.

As expected for a galaxy whose gas disc has strong rotational support, the star-forming gas distribution is much more flattened than the corresponding distributions of stars and DM. In this example, the distributions of the three matter components are well aligned: the minor axis of the star-forming gas is misaligned with respect to that of the stars by  $\simeq 2$  deg and the DM by  $\simeq 6$  deg. As shown in Appendix C, these offsets are comparable to the measurement uncertainty for well-resolved and well-sampled structures. The rotational axis of the star-forming gas is also closely aligned with the minor axis in this example, as naïvely expected for an extended, rotationally supported disc.

#### 3.1 Shape parameters as a function of subhalo mass

Fig. 3 shows probability distribution functions (PDFs) of the shape parameters of the star-forming gas (blue curves), stellar (red), and DM (green) distributions of the subhaloes comprising our sample from Ref-L100N1504 at  $z = 0$ . We reiterate that measurements of the stellar and DM distributions are included here, despite being previously presented for EAGLE subhaloes by Velliscig et al. (2015a), because we use an alternative form of the mass distribution tensor. Thick and thin lines represent the sphericity and triaxiality parameters, respectively. Each panel shows subhaloes split by total mass in bins of 0.5 dex, spanning  $M_{\text{sub}} = 10^{10} - 10^{14} M_{\odot}$ . For clarity, the PDFs of triaxiality have been artificially elevated in the vertical axis by an increment of 0.4. Down arrows denote the median value of each distribution. The bottom right-hand panel shows the volumetric subhalo mass function, split into central and satellite subhalo populations, highlighting that the sample is dominated by central galaxies at all subhalo masses except for the lowest mass bin. For clarity, we also show the median values of the shape parameters for star-forming gas as a function of subhalo mass in Fig. 4. The solid and dashed curves of that plot correspond to the samples, identified as discussed in Section 2.4, at  $z = 0$  and  $z = 1$ , respectively. The lower panel of the figure shows the subhalo volumetric mass function at the two epochs.

These figures show that the distribution of sphericities of star-forming gas distributions is peaked at relatively low values for all subhalo masses, but with a long tail towards high  $S$  (i.e. quasi-spherical systems). The median value of the distributions, which is qualitatively similar to the peak value of the distribution, declines from  $\tilde{S} \simeq 0.25$  for subhaloes of  $M_{\text{sub}} \sim 10^{10} M_{\odot}$ , to a minimum of  $\tilde{S} \simeq 0.1$  at  $M_{\text{sub}} \sim 10^{12.5} M_{\odot}$ . The sphericity of the star-forming gas is therefore systematically lower than is the case for that of the stars, and much more so than is the case for the DM, consistent with the naïve expectation that this dissipational component is found

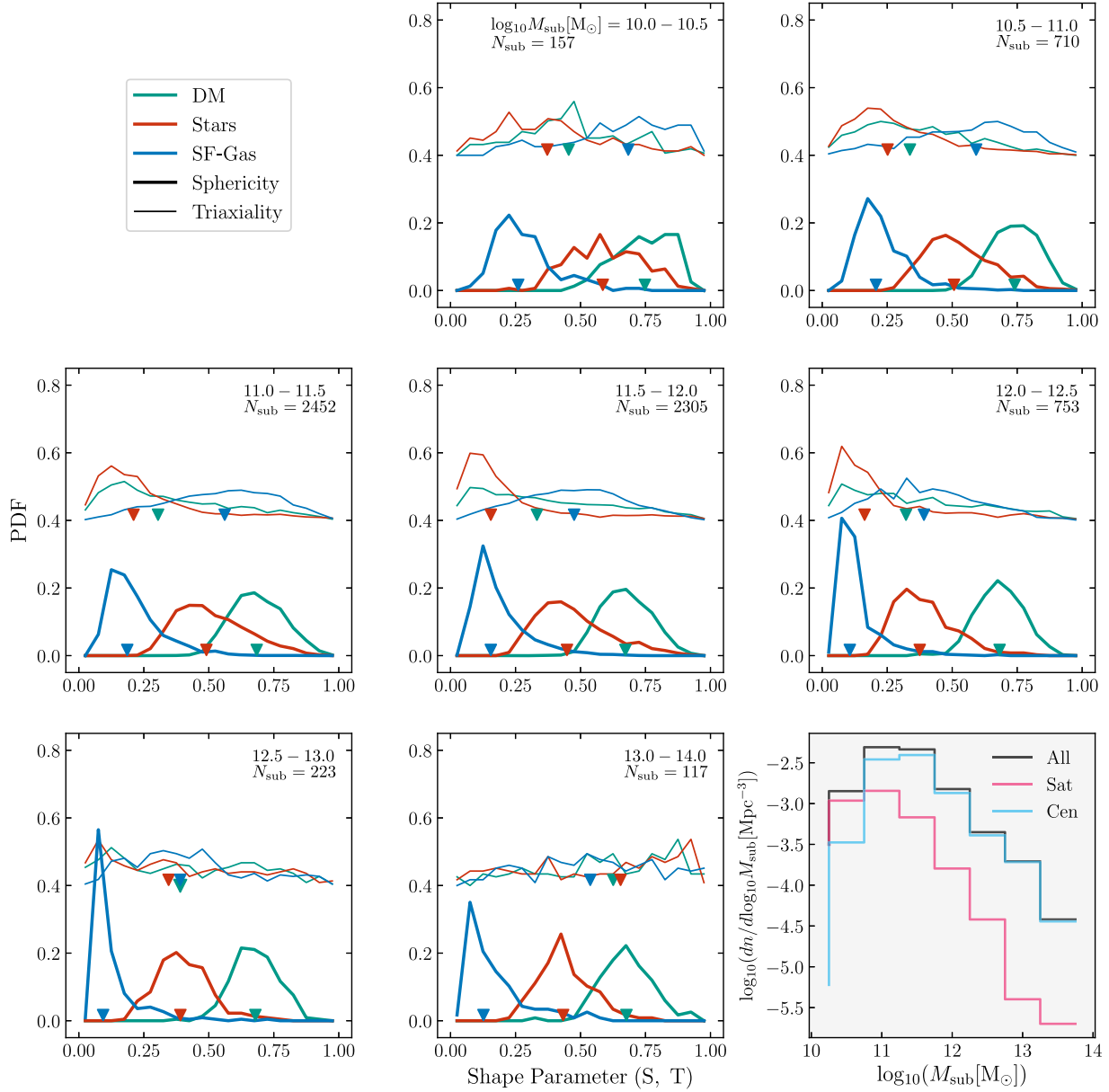


**Figure 2.** The star-forming gas, stars, and dark matter (DM) comprising a star-forming central galaxy drawn from Recal-L025N0752, with stellar mass  $M_{\star} = 10^{10.5} M_{\odot}$ . Each panel is 200 pkpc on a side. The galaxy's subhalo mass is  $M_{\text{sub}} = 10^{12.4} M_{\odot}$ , and its sSFR is  $\dot{M}_{\star}/M_{\star} = 10^{-10.5} \text{ yr}^{-1}$ . The upper panels show the star formation rate surface density, a simple proxy for the radio continuum surface brightness, viewed face-on and edge-on. The green circle in the upper-left-hand panel denotes the spherical half-mass radius of star-forming gas within 30 pkpc. Ellipsoids in the upper-right-hand panel show projections of the best-fitting ellipsoids of the three matter components recovered by the iterative reduced inertia tensor. Overlaid solid lines show the minor axis of the stars (red) and DM (green), whilst the white line corresponds to the rotation axis. The SF-gas is much flatter ( $S = 0.06$ ) than the stars ( $S = 0.35$ ) and DM ( $S = 0.73$ ). The lower panels show the surface densities of stars (left-hand panel) and DM (right-hand panel). Overlaid contours denote surface densities of  $\log_{10}(\Sigma_{\star}) = 6, 7, 8 M_{\odot} \text{ kpc}^{-2}$  and  $\log_{10}(\Sigma_{\text{DM}}) = 7.75, 8.25, 8.75 M_{\odot} \text{ kpc}^{-2}$  for the stars and DM, respectively. The spherical half-mass radii for the stars and dark matter are 6 and 144 pkpc, respectively. These images have been made using the publicly available code PY-SPHVIEWER (Benitez-Llambay 2015).

primarily in flattened discs. Broadly, the peaks of the sphericity PDFs of stars and DM are found at  $S \simeq 0.3\text{--}0.5$  and  $S \simeq 0.7\text{--}0.75$ , respectively, irrespective of subhalo mass. Thob et al. (2019) noted that present-day galaxies whose stellar component exhibit a sphericity of  $S \lesssim 0.6$  generally exhibit stellar corotation kinetic

energy fractions of  $\kappa_{\text{co}}^{\star} > 0.4$  and so correspond broadly to blue, star-forming disc galaxies (Correa et al. 2017). Despite our use of an initial 30 pkpc aperture for the mass tensor, the median values of the sphericity of the stars and DM are broadly consistent with those recovered by Velliscig et al. (2015a) when applying





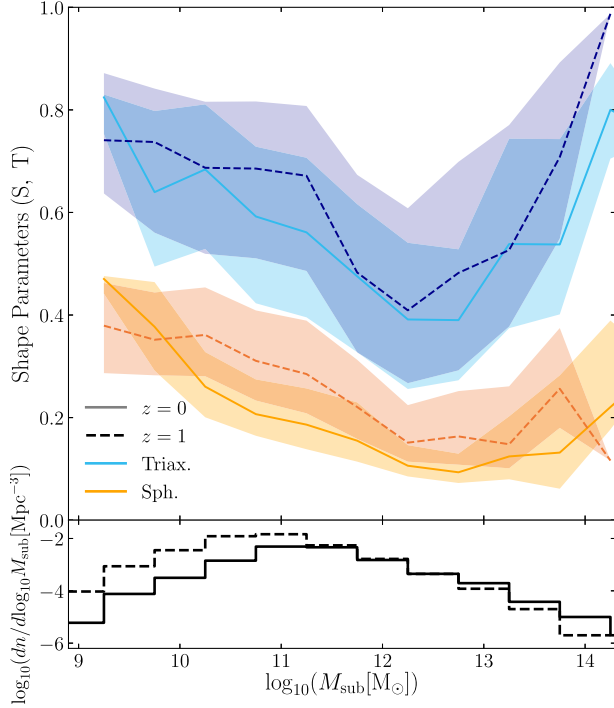
**Figure 3.** Probability distribution functions (PDFs) for the sphericity and triaxiality of the star-forming gas (blue), stars (red), and DM (green) comprising sampled subhaloes in Ref-L100N1504 at  $z = 0$ . Subhaloes are binned by their total mass. The triaxiality PDFs have been raised artificially by 0.4 for clarity. The arrows represent the location of the median for each distribution. The bottom right-hand panel shows volume density function of the satellite and central galaxies. For all subhalo masses, the star-forming gas is significantly more flattened than the stars and DM, and the distribution of the sphericity parameter for star-forming gas is particularly narrow in  $\sim L^*$  galaxies. Star-forming gas does not exhibit a characteristic triaxiality, spanning a wide range of  $T$  at all subhalo masses.

the standard mass distribution tensor to the entirety of EAGLE subhaloes, and those recovered by Tenneti et al. (2014) for subhaloes in the MassiveBlack-II simulation in the mass range for which our respective selection criteria recover broadly similar samples of galaxies ( $M_{\text{sub}} \gtrsim 10^{11} M_{\odot}$ ). Similarly, the distribution of sphericities of star-forming gas discs are consistent with those recovered by Pillepich et al. (2019) when applying the standard mass distribution tensor to galaxies in the TNG50 simulation. We remark that we have also computed the morphology of star-forming gas structures using an iterative form of the simple mass tensor (equation 1), and do not find a significant systematic change.

The sphericity of star-forming gas is most uniform in subhaloes of intermediate mass,  $M_{\text{sub}} \sim 10^{11.5-13} M_{\odot}$ . In such structures, the

distribution of  $S$  is strongly peaked at low values corresponding to flattened discs, albeit with a long tail to more spherical configurations. Owing to this asymmetry, which is most prominent for the star-forming gas, we quantify the diversity of the shape parameter distributions via the interquartile range (IQR) rather than their variance (see Table 2). The IQR of the star-forming gas sphericity decreases from 0.13 for subhaloes of  $\log_{10}(M_{\text{sub}}/M_{\odot}) = 10 - 10.5$  to a minimum of 0.06 for subhaloes of  $\log_{10}(M_{\text{sub}}/M_{\odot}) = 12 - 13$ , before increasing again to 0.14 for the most massive haloes in our sample. The greater diversity in low-mass subhaloes is driven largely by stochasticity in the structure of star-forming gas, with star formation in many low-mass galaxies being confined to a small number of gas clumps rather than being distributed throughout a well-defined





**Figure 4.** The median sphericity and triaxiality of the star-forming gas distribution of subhaloes, as a function of subhalo mass, at  $z = 0$  and  $z = 1$  in bins of 0.5 dex. Shaded regions denote the interquartile range. The orange and blue curves correspond to the sphericity and triaxiality, whilst solid and dashed curves correspond to  $z = 0$  and  $z = 1$ , respectively. The lower panel shows the volumetric mass function of the sampled subhaloes.

**Table 2.** Interquartile ranges of the distributions of the sphericity ( $S$ ) and triaxiality ( $T$ ) shape parameters of the star-forming gas (SF-gas), stars and dark matter (DM) comprising subhaloes in our sample, as a function of subhalo mass.

$\log_{10} M_{\text{sub}}$ ( $M_{\odot}$ )	Sphericity, $S$			Triaxiality, $T$		
	SF-gas	Stars	DM	SF-gas	Stars	DM
10.0–10.5	0.13	0.21	0.16	0.28	0.27	0.26
10.5–11.0	0.11	0.18	0.13	0.31	0.24	0.31
11.0–11.5	0.12	0.19	0.14	0.31	0.23	0.36
11.5–12.0	0.12	0.18	0.13	0.29	0.2	0.41
12.0–12.5	0.06	0.15	0.12	0.29	0.23	0.4
12.5–13.0	0.06	0.14	0.11	0.26	0.48	0.48
13.0–14.0	0.14	0.13	0.13	0.37	0.51	0.38

disc. In massive subhaloes, cold gas discs are readily disturbed by outflows driven by efficient AGN feedback (see e.g. Bower et al. 2017; Oppenheimer et al. 2020), and are less readily replenished with high-angular momentum gas from coherent circumgalactic inflows (see e.g. Davies et al. 2020; Davies, Crain & Pontzen 2021).

A potentially surprising finding highlighted by Figs 3 and 4 is that the characteristic morphology of present-day star-forming gas discs can deviate significantly from that of a disc. The characteristic triaxiality of star-forming gas in subhaloes of mass  $M_{\text{sub}} \sim 10^{12-12.5} M_{\odot}$  is  $T < 0.5$ , consistent with a flattened, oblate spheroid. Subhaloes of all masses exhibit a broad distribution of  $T$ , in marked contrast with that of  $S$ , and for subhaloes in the lower and higher mass bins, the median value is  $T > 0.5$ , signifying that the characteristic morphology is prolate, such that even though the structures are flattened, their isodensity contours when viewed face-on deviate significantly from circular. A

similar finding from the TNG50 simulation was recently reported by Pillepich et al. (2019). Inspection of face-on projections of the star-forming gas surface density highlights that this behaviour again stems primarily from the stochasticity of star-forming gas structure in low-mass subhaloes. In more massive subhaloes, stochasticity is also relevant, owing to the efficient disruption of well-sampled cold gas discs by AGN feedback. However we note that the stellar and DM components tend towards more prolate configurations in more massive subhaloes (as has been widely reported elsewhere, e.g. Tenneti et al. 2014; Velliscig et al. 2015a), suggesting that the morphology of the gravitational potential may influence that of the cold gas. We examine this further in Section 3.3.

As previously noted by Velliscig et al. (2015a), the triaxiality of the stars and DM in EAGLE subhaloes increases as a function of the subhalo mass, such that these components in the most-massive structures are strongly prolate. We note that our quantitative measures are however slightly lower than those reported by Velliscig et al. (2015a), owing to our use of an initial 30 kpc aperture and the reduced inertia tensor, which ascribes less weight to morphology of these structures at large (elliptical) radius. It is well established from prior studies that the condensation of baryons in halo centres drives the morphology towards a more spherical configuration than is realized in dark matter-only simulations (see e.g. Dubinski 1994; Katz et al. 1994; Kazantzidis et al. 2004; Springel, White & Hernquist 2004; Zemp et al. 2012).

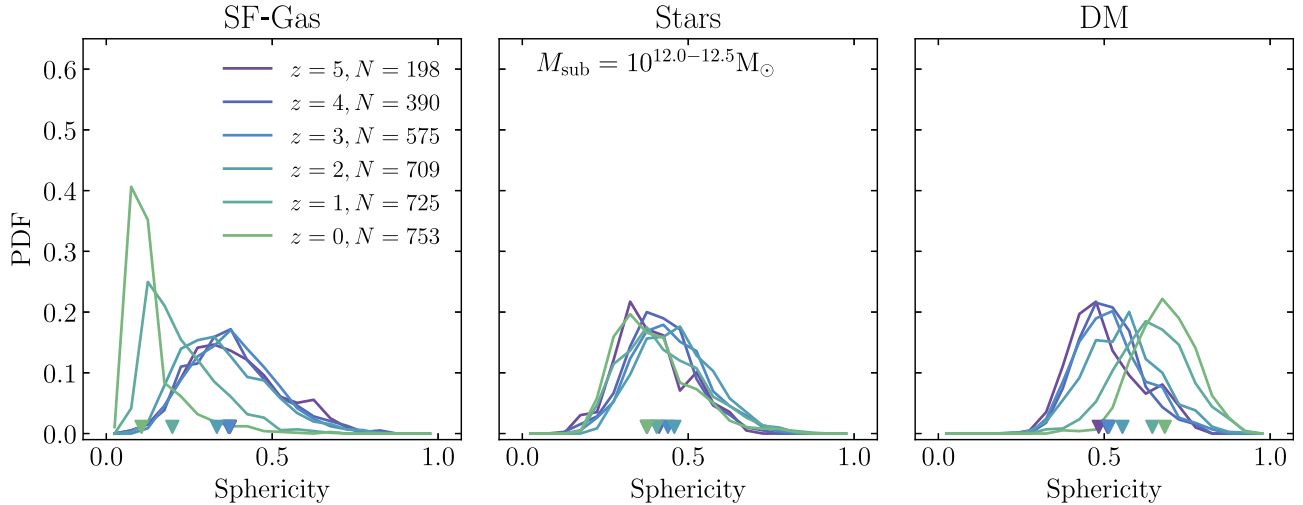
### 3.2 Shape parameters at $z > 0$

We now turn to the morphology of subhaloes at  $z > 0$ , for which we take two approaches. First, we identify subhaloes at  $z = 1$  (which, for our adopted cosmogony, corresponds to a lookback time of 8.1 Gyr) that satisfy the selection criteria specified in Section 2.4, and compare the shape parameters of the samples at these epochs. We subsequently explore the evolution of the shape parameters of the main progenitors of subhaloes that satisfy the selection criteria at  $z = 0$ . Clearly, these approaches require the examination of increasingly dissimilar subhalo samples as one advances to higher redshift.

The evolution of the characteristic morphology of the star-forming gas for identically selected samples at  $z = 0$  and  $z = 1$ , respectively, can be assessed from comparison of the solid and dashed curves of Fig. 4. These curves denote the median values of the shape parameters (sphericity in orange, triaxiality in blue) as a function of subhalo mass, whilst the shaded regions correspond to the interquartile range. The darker (lighter) shaded areas for each parameter correspond to  $z = 1$  ( $z = 0$ ).

It reveals that cold gas structures of fixed subhalo mass, for  $\log_{10}(M_{\text{sub}}/M_{\odot}) \gtrsim 10$ , are slightly more spherical (i.e. less flattened) at  $z = 1$  than the present day. However this difference ( $< 0.1$  for all subhalo masses) is smaller than, or comparable to, the interquartile range of  $S$  at either epoch, which varies between 0.03 and 0.22 at  $z = 0$  and 0.11 and 0.19 at  $z = 1$ , over the subhalo mass range from  $\log_{10}(M_{\text{sub}}/M_{\odot}) = 9 - 14$ . Similarly, the star-forming gas in subhaloes of the same mass tends to be less oblate / more prolate at  $z = 1$  than the present day, but again the difference is small in comparison to the scatter at fixed subhalo mass. We note that the trend for sphericity is in marked contrast with the qualitative behaviour of the DM, for which there is a consensus that structures become more spherical with advancing cosmic time (see e.g. Bryan et al. 2013; Tenneti et al. 2014; Velliscig et al. 2015a).

Fig. 5 shows the sphericity PDFs of the three matter components (star-forming gas, stars, and DM from left to right, respectively) of the main progenitor subhalo, at  $z = (0, 1, 2, 3, 4, 5)$ , of present-day



**Figure 5.** The sphericities of the three matter components (star-forming gas, stars, and DM, from left to right, respectively) comprising subhaloes of present-day mass  $\log_{10}(M_{\text{sub}}/M_{\odot}) = 12 - 12.5$ , and their main progenitor subhaloes at  $z = (1, 2, 3, 4, 5)$ . The latter are included only whilst still satisfying the  $z = 0$  selection criteria, so the sample size  $N$ , shown in the legend of the left-hand panel, is a monotonically declining function of redshift. The arrows represent the location of the median for each distribution. The characterization of the matter distributions with the iterative reduced inertia tensor indicates that DM becomes more spherical at late cosmic epochs, the stellar distribution evolves mildly towards a more flattened configuration, and the star-forming gas typically evolves strongly towards a very flattened configuration by the present day.

central subhaloes with mass  $\log_{10}(M_{\text{sub}}/M_{\odot}) = 12.0 - 12.5$ . Such subhaloes broadly correspond to those that host present-day  $\sim L^*$  galaxies. The progenitors are identified using the D-TREES algorithm (Jiang et al. 2014); a full description of its application to the EAGLE simulations is provided by Qu et al. (2017). The standard 30 pkpc aperture is used at all redshifts.<sup>4</sup> Progenitor subhaloes are included in the  $z > 0$  samples only while they still satisfy the selection criteria concerning particle number and asymmetry, to ensure that a robust measurement of their shape parameters can be made. As such, the sample size,  $N$ , is a monotonically declining function of redshift, as denoted in the legend of the left-hand panel of the figure.

We saw from Fig. 3 that present-day galaxies hosted by subhaloes in this mass range typically exhibit strongly flattened ( $S \simeq 0.1$ ) star-forming gas discs. The left-hand panel of Fig. 5 highlights that, although star-forming gas discs are predominantly flattened<sup>5</sup> even at early epochs, the median sphericity at  $z = 5$  is  $\tilde{S}_{\text{SF-gas}} \simeq 0.37$ . The star-forming gas of the main progenitor becomes increasingly flattened with advancing cosmic time, but the emergence of strongly flattened discs ( $S \lesssim 0.2$ ) is generally limited to  $z < 2$ : the median sphericity evolves from  $\tilde{S}_{\text{SF-gas}} \simeq 0.33$  at  $z = 2$  to  $\tilde{S}_{\text{SF-gas}} \simeq 0.1$  at  $z = 0$ . The strong evolution of the star-forming gas sphericity of these progenitors is broadly coincident with the growth of the gas disc’s median scale length, which grows only from  $\simeq 2$  pkpc to  $\simeq 3$  pkpc from  $z = 5$  to  $z = 2$ , but by  $z = 0$  reaches  $\simeq 9$  pkpc. The decrease in the accretion rate (of all matter types) on to the galaxy + halo ecosystem at later epochs (see e.g. Fakhouri, Ma & Boylan-Kolchin 2010; van de Voort et al. 2011) likely also results in a steady decline of the scale height of the gas disc (e.g. Benítez-Llambay et al. 2018), further contributing to decrease in sphericity.

<sup>4</sup>We have assessed the impact of using an adaptive aperture of initial spherical radius  $r = 0.15r_{200}(z)$ , to account for the decreasing physical size of progenitors at early times, and do not recover significant differences.

<sup>5</sup>For context, we reiterate that, as noted in Section 3.1, present-day galaxies with a stellar component sphericity of  $S \lesssim 0.6$  are broadly equivalent to star-forming disc galaxies.

Strong evolution of the structural parameters of star-forming gas was similarly reported by Pillepich et al. (2019) based on analysis of the TNG50 simulation. Those authors noted that the evolution of the flattening (‘disciness’ in their terminology, since they also examined kinematic descriptions) of both the star-forming gas and stars increases over time, but that the evolution for the former is much more pronounced than the latter. The same behaviour is evident in EAGLE, as is clear from inspection of the centre panel of Fig. 5, which shows that the sphericity of the stellar component of the progenitors of our present-day  $\sim L^*$  galaxy sample is largely insensitive to redshift. The majority of the galaxies comprising our sample remain actively star-forming at  $z = 0$ , and are characterized by flattened discs ( $\tilde{S}_{\star} \simeq 0.4 - 0.45$  at all redshifts examined). Such galaxies will therefore have assembled primarily via *in-situ* star formation (Qu et al. 2017) and will not have experienced the strong morphological evolution that typically follows internal quenching (see e.g. Davies et al. 2020, 2021). Furlong et al. (2017) showed that the half-mass radius of the stellar component of present-day star-forming  $\sim L^*$  galaxies grows only from  $\simeq 10^{0.5}$  pkpc to  $\simeq 10^{0.85}$  pkpc between  $z = 2$  and  $z = 0$ .

As noted above, it has been shown elsewhere that DM haloes, even in the absence of dissipative baryon physics, tend to become more spherical with advancing cosmic time (e.g. Bryan et al. 2013; Tenneti et al. 2014; Velliscig et al. 2015a), in marked contrast to the behaviour seen for the star-forming gas. The right-hand panel of Fig. 5 shows that this effect is clearly seen for the host subhaloes of present-day  $\sim L^*$  galaxies, even when focusing primarily on the halo centre by defining the shape parameters via the use of the iterative reduced mass tensor. The resolved progenitors exhibit a median sphericity of  $\tilde{S}_{\text{DM}} \simeq 0.5$  at  $z = 5$ , and this median increases monotonically to  $\tilde{S}_{\text{DM}} \simeq 0.7$  at  $z = 0$ .

Besides the evolution of the median sphericity of the matter components, it is interesting to consider the evolution of their diversity. Since the PDFs can exhibit significant asymmetry, we characterize this diversity using the interquartile range. Whilst the IQR of the star-forming gas sphericity decreases markedly at later cosmic epochs (c.f. 0.18 at  $z = 5$  to  $\simeq 0.06$  at  $z = 0$ ), that of the stars

and the DM remain components remain largely unchanged from  $z = 5$  to  $z = 0$ , with values of 0.13 to 0.14 for the stars and 0.14 to 0.12 for the DM.

### 3.3 Correspondence of star-forming gas and stellar structure

We noted in the Section 3.1 that the star-forming gas configuration in massive subhaloes is often well described by a flattened prolate spheroid, similar to the characteristic morphology of the stars and DM in such structures. Thob et al. (2019) previously demonstrated that EAGLE galaxies with flattened stellar distributions are preferentially hosted by flattened DM haloes, motivating a closer examination here of the degree to which the morphology of the star-forming gas correlates with that of the other matter components. Since the density of stars typically dominates over the density of dark matter within the region traced by the star-forming gas, we focus on the correspondence between the morphology of the star-forming gas and the stars.

The main panels of Fig. 6 show, as a function of subhalo mass, the sphericity (left-hand panel) and triaxiality (right-hand panel) of the star-forming gas distributions of the subhaloes comprising our sample. The distribution is shown as a 2D histogram, and black curves denote the running median of the star-forming gas shape parameters, computed via the locally weighted scatterplot smoothing method (LOWESS; e.g. Cleveland 1979). The LOWESS curves are plotted within the interval for which there are at least 10 measurements at both lower and higher  $M_{\text{sub}}$ . The colour of each hexbin denotes to the median value of the corresponding shape parameter of the stellar component: subhaloes in bins denoted by red (blue) colours typically have a stellar component with a high (low) value of the shape parameter in question.

The shape parameters of the star-forming gas and stellar distributions are strongly and positively correlated at effectively all subhalo masses: flattened star-forming gas distributions are generally found in subhaloes with flattened stellar components, and more prolate star-forming gas distributions are found in subhaloes with more prolate stellar components. We quantify the strength and significance of these correlations by computing a ‘running’ Spearman rank correlation coefficient,  $\rho(M_{\text{sub}})$ , for the  $\Delta S_{\text{SF-gas}} - \Delta S_*$  and  $\Delta T_{\text{SF-gas}} - \Delta T_*$  relations, where  $\Delta X_m$  represents the residual of shape parameter  $X$  for matter distribution  $m$  about the LOWESS median. Hence, in the case of sphericity,  $\Delta S_{*,i} = S_{*,i} - \tilde{S}_*(M_{\text{sub},i})$  for the  $i$ -th subhalo. The running Spearman rank correlation coefficient is computed in subhalo mass-ordered subsamples: for bins with a median subhalo mass  $M_{\text{sub}} < 10^{12.5} M_{\odot}$ , we use samples of 200 subhaloes with starting ranks separated by 50 subhaloes (e.g. subhaloes 1–200, 51–250, and 101–300). For bins with median  $M_{\text{sub}} > 10^{12.5} M_{\odot}$ , we use samples of 50 subhaloes with starting ranks separated by 25 subhaloes, to ameliorate the effect of the relative paucity of massive subhaloes. This running  $\rho(M_{\text{sub}})$  is plotted in the lower subpanel. Regions shaded in grey denote a Spearman rank  $p$ -value is  $> 0.05$ , and thus indicate where the recovered correlation cannot be considered significant.

The relatively high correlation coefficient ( $\rho \gtrsim 0.4$ ) for the sphericity over a wide range in subhalo mass indicates that the degree of flattening of the two components is indeed strongly and positively correlated. The correlation is weaker ( $\rho \gtrsim 0.3$ ) for the triaxiality parameter, but remains positive and significant over a wide range of subhalo masses. We have also examined the correlation of the shape parameters of star-forming gas with those of their host subhalo’s DM, and we find that the correlation is not formally significant at any subhalo mass.

Our results suggest that examination of a large sample of galaxies with high-fidelity radio imaging is likely to reveal significant correlations between the radio continuum and optical morphologies of galaxies. There is not currently a firm consensus amongst observational studies, which are necessarily limited to comparisons of projected ellipticities, in regard to correlations between the morphologies of the radio continuum and optical components of galaxies. Battye & Browne (2009) report a strong, positive correlation of the two in late-type galaxies, and a weak negative correlation for early-type galaxies, whilst complementary studies using a smaller sample (Patel et al. 2010), or a sample of fainter, more-distant galaxies (Tunbridge, Harrison & Brown 2016), recovered no significant correlations. More recently, Hillier et al. (2019) examined the correlation of optical and radio continuum measurements of shape and orientation for galaxies in the COSMOS field, and recovered a significant correlation of position angles (projected orientation) between matched 3 GHz radio (VLA) and optical (*HST*-ACS) images (seen in their figs 5 and 6).

## 4 THE ALIGNMENT OF STAR-FORMING GAS WITH GALAXIES AND THEIR DM HALOES

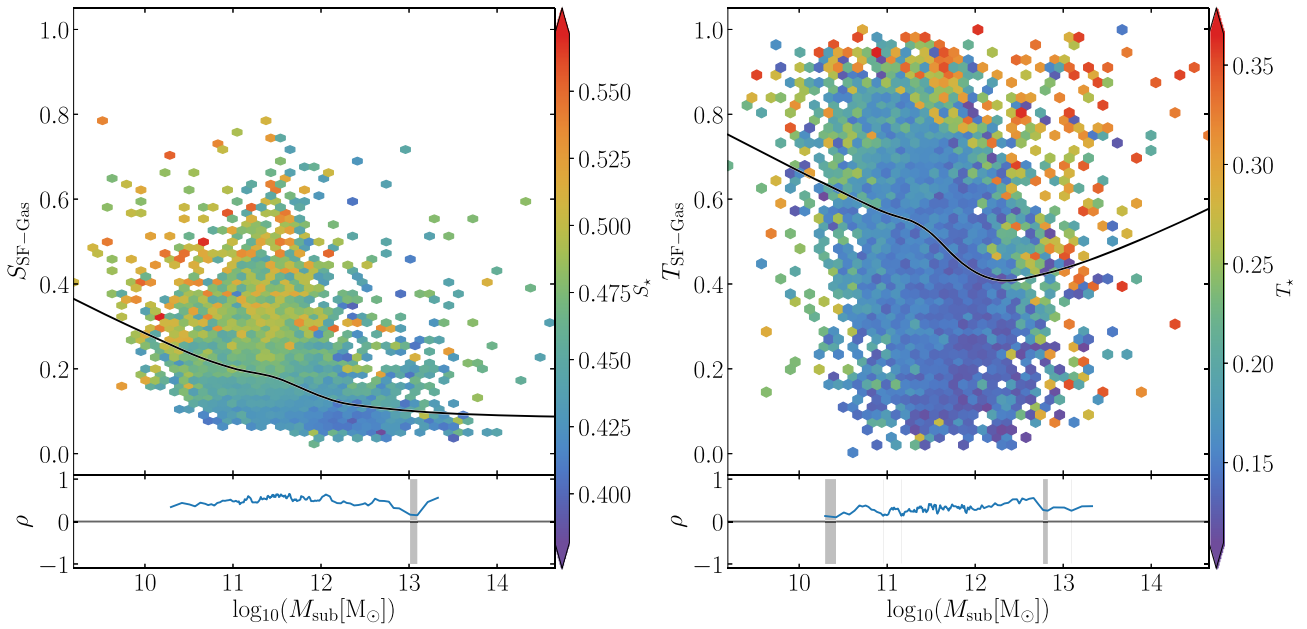
In this section, we examine the orientations of the 3D distribution of star-forming gas in galaxies with respect to the stellar and DM components of their host subhaloes. We begin in Section 4.1 with an examination of the morphological alignment of subhalo components as a function of subhalo mass, triaxiality, and cosmic epoch. In Section 4.2, we consider the alignment of the morphological minor axis of the star-forming gas with its kinematic axis.

### 4.1 Morphological alignment of subhalo matter components

We quantify the morphological alignment of the various components via the angle,  $\theta$ , between the minor axes of the ellipsoids describing each matter distribution, such that  $\theta = 0^\circ$  indicates perfect alignment and  $\theta = 90^\circ$  indicates orthogonality. As noted in Section 2.3, we consider the minor axis to be the natural choice when focusing on discs, as the minor axis is the most distinct axis for oblate discs (though we reiterate the finding from Section 3.1 that many flattened star-forming structures are mildly prolate). Moreover, as seen in Section 3.1, the central regions of the stellar and DM distributions (to which the iterative reduced mass distribution tensor is more strongly weighted) also tend to be mildly oblate.

Fig. 7 shows the alignment between the star-forming gas distribution and that of the DM. In the left-hand panel the alignment is shown as a function of subhalo mass ( $M_{\text{sub}}$ ) and in the right-hand panel it is shown as a function of the triaxiality of the DM. The thick orange curve and associated shading denotes the median alignment angle, and the 10th–90th percentiles of the distribution, when considering the minor axes of the two components. In general, the alignment is strong, with the median alignment angle typically  $\simeq 30^\circ$  for  $M_{\text{sub}} = 10^{10} M_{\odot}$ , declining to  $\simeq 15^\circ$  for  $M_{\text{sub}} = 10^{11} M_{\odot}$  and  $\simeq 10^\circ$  for  $M_{\text{sub}} = 10^{12-12.5} M_{\odot}$ . In more massive subhaloes, the characteristic alignment is typically (marginally) poorer, rising to  $\simeq 20^\circ$  for  $M_{\text{sub}} \gtrsim 10^{13} M_{\odot}$ .

Examination of the right-hand panel shows that the alignment of the minor axes of the star-forming gas and the DM of its host subhalo is a strong function of the latter’s triaxiality, with oblate subhaloes exhibiting close alignment of the two components ( $\theta < 10^\circ$  for  $T_{\text{DM}} \lesssim 0.4$ ) but prolate subhaloes exhibiting much poorer alignment ( $\theta > 50^\circ$  for  $T_{\text{DM}} \gtrsim 0.8$ ). As is clear from the scatter about the median relation, in prolate systems the minor axes of the two components can become



**Figure 6.** The sphericity (left-hand panel) and triaxiality (right-hand panel) of the star-forming gas as a function of subhalo mass. The black curves denote the running medians of the shape parameters, computed via the locally weighted scatterplot smoothing method (LOWESS). The colours represent the median sphericity/triaxiality of the stellar mass distributions of the subhaloes represented by each hexbin. The lower panels display the running Spearman rank correlation coefficient,  $\rho$ , between the residual shape parameters about  $M_{\text{sub}}$  for the two matter distributions, i.e.  $\Delta S_{\text{SF-gas}} - \Delta S_{\star}$  and  $\Delta T_{\text{SF-gas}} - \Delta T_{\star}$ , where for instance  $\Delta S_{\star,i} = S_{\star,i} - \tilde{S}_{\star}(M_{\text{sub},i})$ . The grey shaded regions indicate mass ranges for which the correlation is recovered at low significance ( $p > 0.05$ ). The subpanels corroborate, quantitatively, the impression given by visual inspection, that the morphology of the star-forming gas component correlates positively with that of the stellar component.

effectively orthogonal. If the shape parameters of the two components are dissimilar, as is the case for the common configuration of an oblate disc within a prolate subhalo, alignment of the minor axes might not be the most likely scenario, since in such cases the minor and intermediate DM axes are not distinct. Indeed, the axes that should be ‘expected’ to align are likely to be those most closely aligned with the angular momenta of the respective components (as we discuss in Section 4.2). We therefore examine whether this is a genuine misalignment, or is rather a consequence of the minor axis of the star-forming gas exhibiting a preference to align with one of the other principal axes of the DM.

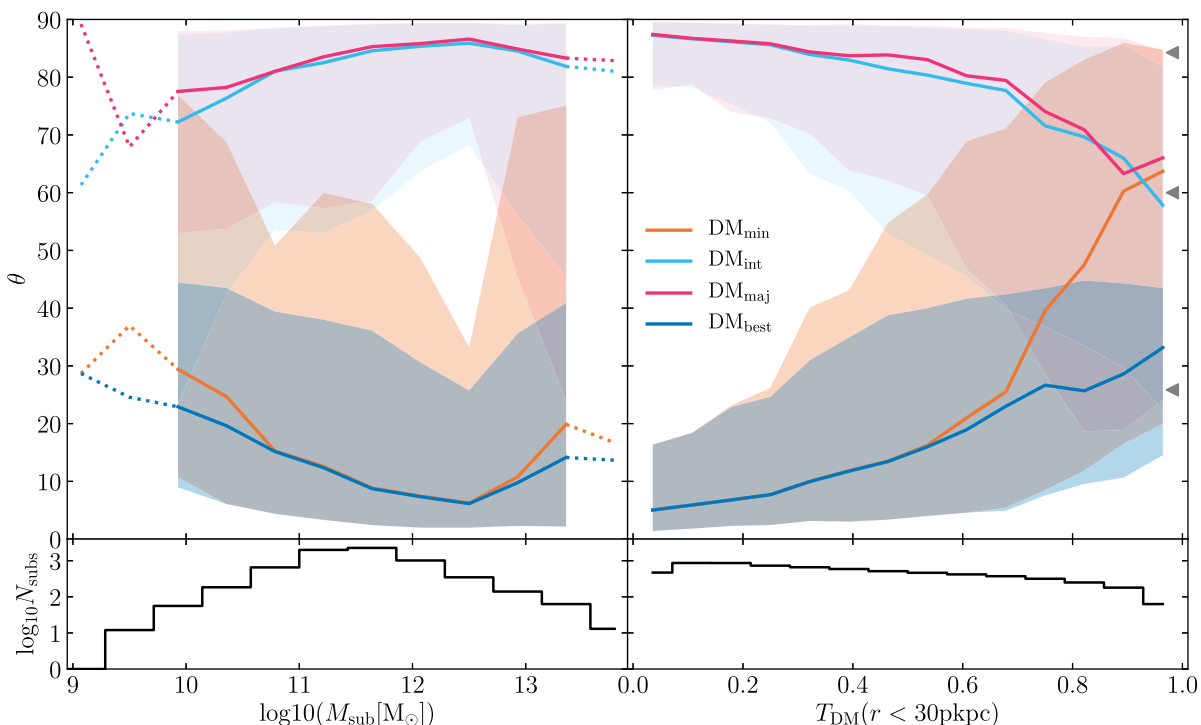
In Fig 7, the intermediate and major axes are denoted by the light blue and red curves, respectively. The dark blue curve denotes the angle between the minor axis of the star-forming gas and that of the DM axis with which it best aligns. In prolate subhaloes, for which the alignment quantified by the standard measure is poor, one can often find good alignment between the star-forming gas minor axis with one of the other principle axes of the DM. However, for  $T_{\text{DM}}(r < 30 \text{ pkpc}) \rightarrow 1.0$  subhaloes, the characteristic alignments of the star-forming gas minor axis with all of the DM morphological axes converge towards  $\simeq 60^\circ$ , the expectation value for the alignment angle of unit vectors randomly oriented in 3 dimensions. This implies that poor alignment between the minor axes of the two components within high-triaxiality subhaloes is not primarily due to a preference for the star-forming gas minor axis to align with a non-minor DM morphological axis. Therefore in what follows, we focus exclusively on the misalignment between the minor axes of the two matter components.

Fig. 8 shows the cumulative distribution function of the alignment angle  $\theta$  for the three pairs of matter components, namely star-forming gas and DM (pink), star-forming gas and stars (blue), and stars and

DM (green). We plot the distribution as a function of  $\log_{10}(1 + \theta)$  because the bulk of the misalignments (for all component pairs) are small, but there are long tails to severe misalignments. Thick lines denote our fiducial measurement, whilst the thin lines show the alignments inferred when the initial characterization of the mass distribution considers all particles of the relevant matter component bound to the subhalo, rather than only those within 30 pkpc of the subhalo’s centre. We show the latter in order to highlight the influence of the initial aperture, since an influence is to be expected: for example, Velliscig et al. (2015a) showed that the alignment of the stellar and DM components is stronger closer to the subhalo centre, i.e. that galaxies are best aligned with the local, rather than global, distribution of matter in the subhalo. For reference, the dotted black line shows the distribution function of alignment angles between randomly oriented vectors.

For our fiducial measurements, half of the sampled subhaloes have star-forming gas distributions misaligned with their stellar components by more than  $5^\circ$ , and half have star-forming gas distributions misaligned with their DM component by more than  $9.5^\circ$ . Half of the subhaloes have stellar components misaligned with their DM component by more than  $6^\circ$ . Assessing the alignments recovered when considering all the particles of a given type associated with subhaloes, we find that half of the subhaloes have stellar-DM misalignments greater than  $17^\circ$ . The poorer star-forming gas – DM alignment with respect to the stars – DM alignment might be expected; since the stars and DM are collision-less components, their relevant evolutionary time-scale is the gravitational dynamical time,  $t_{\text{dyn}} = 1/\sqrt{G\rho} \sim 10^9 \text{ yr}$ , such that their morphologies and orientation effectively ‘encode’ their formation and assembly history over an appreciable fraction of a Hubble time. In contrast, the phase-space structure of the collisional, dissipative gas is not preserved as





**Figure 7.** The misalignment angle,  $\theta$ , of the minor axis of star-forming gas of our sample of subhaloes with respect to each of the principal axes of their DM distribution. The solid lines indicate the binned median values of  $\theta$ , whilst the shading denotes the 10th–90th percentiles. The values are shown as a function of subhalo mass ( $M_{\text{sub}}$ , left-hand panel) and DM triaxiality ( $T_{\text{DM}}$ , right-hand panel). The orange, cyan, and magenta curves correspond to the alignment of the star-forming gas minor axis with respect to the minor, intermediate, and major axis of the DM, respectively, whilst the blue curve is with respect to the principle DM axis with which the star-forming gas minor axis is most closely aligned. The dotted lines indicate where the sampling drops below 30 subhaloes per bin. Sub-panels show the number of subhaloes per bin. The grey arrows denote the expected 10th, 50th, and 90th percentile values of alignment angles between vectors randomly oriented in 3 dimensions. In general the minor axis of star-forming gas is well aligned with the minor axis of its corresponding DM, but in strongly triaxial subhaloes the former often aligns more closely with the intermediate or major axis of the DM distribution.

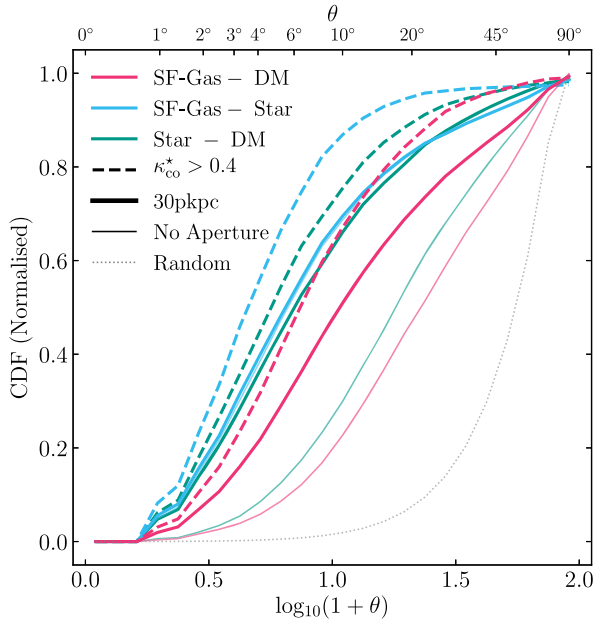
it accretes on to galaxies and condenses into star-forming clouds. Its morphology and orientation therefore reflects a more instantaneous snapshot of the evolution of the subhalo than is the case for the collision-less components.

We note that the stellar–DM alignment shown in Fig. 8 (thick green curve) is significantly better than that inferred by Velliscig et al. (2015a), who found that half of all the subhaloes they examined had misalignments worse than the  $40^\circ$ . This follows primarily from our use of an initial particle selection within a 30 pkpc sphere and the iterative reduced inertia tensor (which weights more strongly towards the halo centre), and also in part due to their measurement of the misalignment angle relative to the major axes of the mass distribution, and the slightly different sample selections. The influence of the initial particle selection can be assessed by comparison of the thick and thin solid curves: as expected, when one considers all matter bound to the subhalo (as opposed to only that within a 30 pkpc sphere) when initializing the iterative characterization of the mass distribution, the misalignments with respect to the DM become significantly more pronounced. As is clear from the thinner curves of Fig. 8, in this case half of the sampled subhaloes have star-forming gas distributions misaligned with their DM components by more than  $\approx 20^\circ$ , and half have stellar components misaligned with their DM component by more than  $15^\circ$ . The misalignment of star-forming gas and the stars is however largely unaffected, since the bulk of both components is typically found within the central 30 pkpc.

Having noted that misalignments are typically most severe in massive, prolate subhaloes, which tend to host quenched elliptical

galaxies (see e.g. Thob et al. 2019), it is reasonable to hypothesize that subhaloes hosting star-forming disc galaxies (i.e. those with  $\kappa_{\text{co}}^* > 0.4$ ) will exhibit significantly better alignment than the broader sample. The misalignment angles for this subset of subhaloes are shown by the dashed lines in Fig. 8, and indeed we find that the primary consequence of restricting our focus to these systems is the exclusion of galaxies with severe misalignments. For this subsample, only 20 percent of galaxies exhibit star-forming gas distributions misaligned with their DM by more than  $\approx 10^\circ$ .

Fig. 9 shows the temporal evolution of the misalignment angle,  $\theta$ , of the minor axes of star-forming gas and DM mass distributions (left-hand panel) and the star-forming gas and stars (right-hand panel). Here, as was the case for Fig. 5, we consider at all epochs subhaloes that satisfy the selection criteria specified in Section 2.4, however we do not here focus solely on main branch progenitors of  $L_*$  subhaloes. It is immediately apparent that the orientation of the star-forming gas is a much poorer tracer of the orientation of both the DM and the stars at early cosmic epochs than at the present day (though the characteristic alignment is always much better than random). As noted above, at  $z = 0$  half of the sampled subhaloes have star-forming gas distributions misaligned with their DM components by more than  $10^\circ$ , but at  $z = 1$  half are misaligned by more than  $25^\circ$  and at  $z = 2$  the figure is  $29^\circ$ . Similarly, at  $z = 0$  half of the sampled subhaloes have star-forming gas distributions misaligned with their stellar components by more than  $6^\circ$ , but at  $z = 1$  half are misaligned by more than  $14^\circ$  and at  $z = 2$  half are misaligned by at least  $19^\circ$ . The deterioration of the alignment of the star-forming



**Figure 8.** The misalignment of the star-forming gas, DM, and stellar distributions within the subhaloes of our sample. The figure shows the cumulative distribution function of the misalignment angle,  $\theta$ , between the minor axes of the matter distribution pairs in the legend. Thick curves correspond to fiducial measurements, thin curves denote alignments inferred when the initial characterization of the mass distribution considers all particles of the relevant type bound to the subhalo, rather than only those within 30 pkpc of the subhalo centre. Thick dashed lines correspond to the subset of galaxies with  $\kappa_{\text{co}}^* > 0.4$ , which broadly identifies star-forming disc galaxies. The dotted black line indicates the distribution of angles between randomly orientated vectors in 3D. Star-forming gas is a poorer tracer of the orientation of the subhalo DM distribution than are the stars.

gas distribution with both the DM and the stars at earlier times is to be expected, since all three components tend to be more spherical (less flattened) at higher redshift. Although in principle even highly spherical distributions can exhibit perfect alignment, as  $S \rightarrow 1$  the minor axis becomes less well defined.

In Appendix D, we provide analytical fits to probability distribution functions of the misalignment angle,  $\theta$ , of the star-forming gas distribution with those of DM and stars, enabling subhaloes in dark matter-only simulations to be populated with galaxies whose star-forming gas has a realistic misalignment distribution.

#### 4.2 Alignment of the kinematic and morphological axes

A novel aspect of radio continuum lensing surveys is that complementary observations of the 21 cm hyperfine transition emission line from atomic hydrogen can, in principle, be obtained simultaneously with little or no extra observing time. The Doppler shift of the 21 cm line is widely used to infer the kinematics of the atomic phase of the ISM (e.g. Bosma 1978; Swaters 1999) and hence affords an independent means of assessing galaxy orientation. As noted by Blain (2002), Morales (2006), and de Burgh-Day et al. (2015), the kinematic axis can be used as a proxy for the unsheared morphological axis, and hence affords a means to suppress the influence of galaxy shape noise and intrinsic alignments.

Clearly, the naïve application of this method assumes perfect alignment of the kinematic and minor morphological axes. To assess the accuracy of this assumption, we define the morphokinematic misalignment angle,  $\beta$ , as the angle between the minor axis of

the star-forming gas distribution, and the unit vector of its angular momentum. Fig. 10 shows the cumulative distribution function of  $\log_{10}(1 + \beta)$ , with solid curves denoting present-day measurements and dashed lines denoting measurements at  $z = 1$ . The blue curves correspond to the fiducial sample, whilst red curves correspond to the subset of galaxies with  $\kappa_{\text{co}}^* > 0.4$ . For reference, the dotted black line again shows the distribution function of alignment angles between randomly oriented vectors.

As naïvely expected, the star-forming gas minor axis and angular momentum vector of star-forming gas are well aligned for present-day subhaloes: 80 per cent of systems exhibit morphokinematic misalignments of less than  $10^\circ$ . However, similar to the internal component alignments, the distribution function exhibits a long tail to severe, but rare, misalignments. The morphokinematic alignment improves if one restricts the analysis to the  $\kappa_{\text{co}}^* > 0.4$  subsample, for which eighty per cent of the systems are misaligned by less than  $6^\circ$ , and the tail to severe misalignments is strongly diminished. As might be expected when considering the reduced prevalence of strongly flattened star-forming discs at  $z = 1$ , the morphokinematic alignment is poorer at this earlier epoch, with 80 per cent of subhaloes aligned to better than  $30^\circ$ , and  $12^\circ$  when restricting to the  $\kappa_{\text{co}}^* > 0.4$  subsample.

To establish the characteristics of the subhaloes that typically suffer from poor morphokinematic alignment, we separate the primary sample of present-day subhaloes into quartiles of  $\beta$ , and quote in Table 3 the median values of key characteristics of subhaloes in each quartile, namely the star-forming gas sphericity, subhalo mass, star formation rate, stellar mass, the star-forming gas corotation parameter and the half-mass radius of the star-forming gas. This exercise illustrates that poor alignment of the minor axis of the star-forming gas with its angular momentum vector is more typical in subhaloes hosting a spheroidal central galaxy, with a low star-formation rate and a less flattened and less extended star-forming gas distribution. In principle, such systems can be readily identified from either optical or radio continuum imaging.

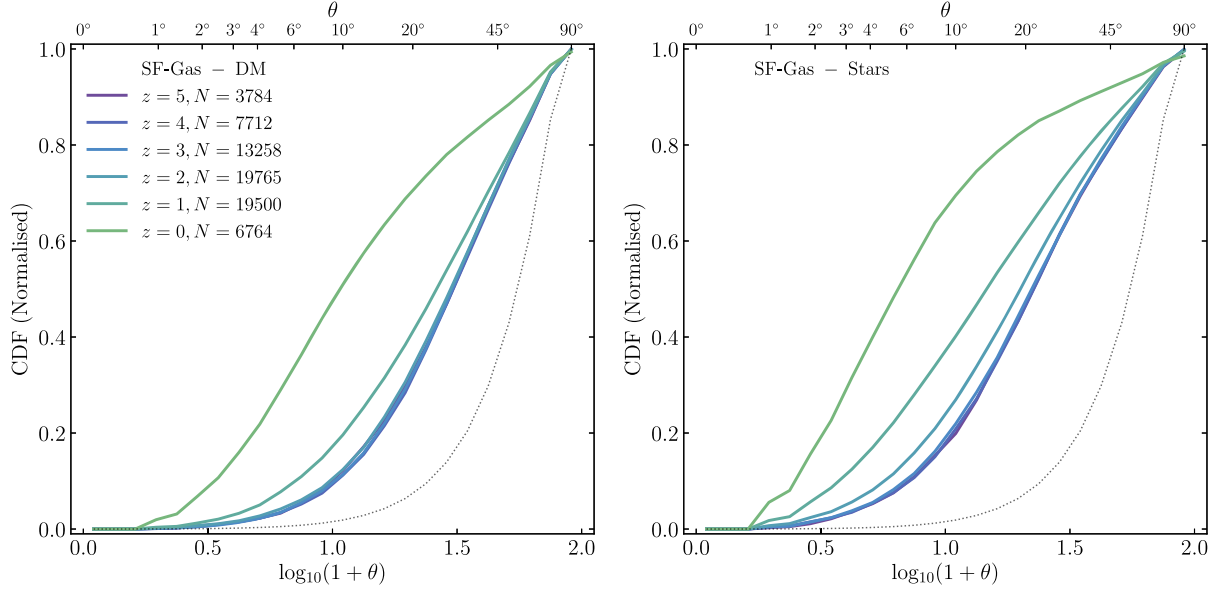
## 5 THE MORPHOLOGY AND ALIGNMENT OF PROJECTED STAR-FORMING GAS DISTRIBUTIONS

In this section, we examine the morphologies, alignments, and orientations of star-forming gas and DM when projected ‘on the sky’ in 2 dimensions, affording a direct connection with observational tests. In Section 5.1, we consider the ellipticity of the matter components, i.e. their projected morphology. In Section 5.2, we consider the projected alignments of galaxies.

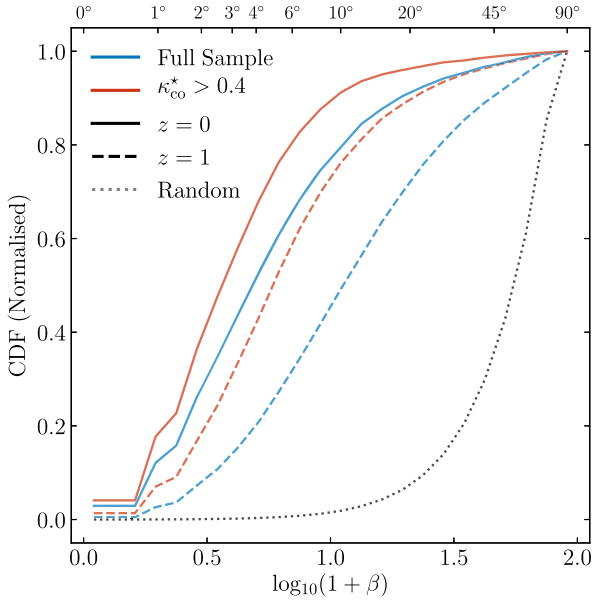
### 5.1 Projected ellipticities

It is via measurement of the morphology of galaxies in projection, i.e. their ellipticity, that the weak gravitational shear is estimated. Since galaxies are intrinsically ellipsoidal (i.e. non-circular), the observed ellipticity is due to both the intrinsic ellipticity of the galaxy, and the lensing shear. The former can therefore be considered as a noise term when measuring the shear, and is often referred to as ‘shape noise’. Since the variance of the observed ellipticity,  $\epsilon_{\text{obs}}$  is the sum of the variances of the intrinsic ellipticity and the (reduced) shear, i.e.  $\sigma^2(\epsilon_{\text{obs}}) = \sigma^2(\epsilon_{\text{int}}) + \sigma^2(\epsilon_{\text{sh}})$ , the signal-to-noise ratio of shear measurements is markedly sensitive to the diversity of the intrinsic ellipticity of the galaxy population being surveyed.

To measure the intrinsic ellipticity of matter distributions, we adapt the iterative reduced inertia tensor algorithm presented in Section 2.3 to consider only two spatial coordinates and so recover the best-fitting



**Figure 9.** The temporal evolution of the misalignment of the minor axes of the star-forming gas and DM distributions (left-hand panel) and of the star-forming gas and stellar distributions (right-hand panel) comprising the subhaloes of our sample, as computed while adopting a 30 pkpc aperture. The figure shows the cumulative distribution function of the misalignment angle,  $\theta$ . Colour indicates the redshift, and the thin black dotted line shows the distribution of angles between randomly orientated 3D vectors.



**Figure 10.** A histogram of the alignment between the morphological and kinematic axes of the star-forming gas within our sample. Alignment angles are given in terms of  $\log_{10}(1 + \beta)$ , as the majority of the subhaloes have small  $\beta$ . The solid (dashed) lines correspond to subhaloes at redshift  $z = 0(1)$ . The black dotted line shows the distribution of angles between randomly orientated 3D vectors. The red lines correspond to the subset of galaxies with  $\kappa_{\text{co}}^*$ , which broadly identifies star-forming disc galaxies, while blue lines correspond to the full sample.

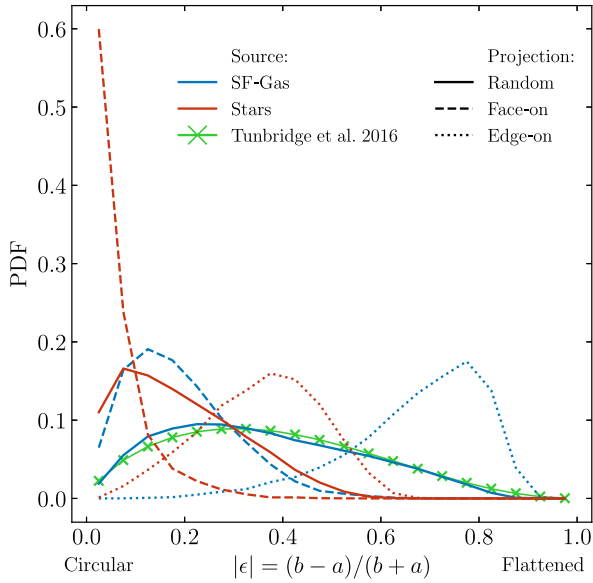
ellipse. The intrinsic ellipticity is then  $|\epsilon_{\text{int}}| = (a - b)/(a + b)$ , where  $a, b$  are the major and minor axis lengths of this ellipse, respectively, such that low ellipticity corresponds to near-circular morphology, and high ellipticity corresponds to a strongly flattened configuration. Hereafter we omit the subscript for brevity, such that  $\epsilon \equiv \epsilon_{\text{int}}$ . As noted in Section 2.3, the first iteration of the algorithm considers

**Table 3.** The median and standard deviation in various subhalo properties for the systems binned into quartiles based on the alignment angle between the star-forming gas kinematic and morphological minor axes. In degrees, the 0th, 25th, 50th, 75th, and 100th percentile values are  $0.0^\circ, 2.14^\circ, 4.36^\circ, 9.15^\circ$ , and  $89.75^\circ$ , respectively. The values from top to bottom are: the star-forming gas sphericity; total subhalo mass (as  $\log_{10} M_{\text{sub}} [M_\odot]$ ); star-formation rate within 30 pkpc (in  $M_\odot \text{yr}^{-1}$ ); stellar mass within 30 pkpc (as  $\log_{10} M_* [M_\odot]$ ); the fraction of kinetic energy in the star-forming gas invested in corotation; and the star-forming gas half-mass radius (in pkpc).

Quartile	1st	2nd	3rd	4th
$S_{\text{SF-gas}}$	$0.11 \pm 0.05$	$0.14 \pm 0.06$	$0.19 \pm 0.07$	$0.29 \pm 0.12$
$M_{\text{sub}}$	$11.67 \pm 0.5$	$11.53 \pm 0.52$	$11.41 \pm 0.51$	$11.44 \pm 0.67$
SFR	$0.59 \pm 1.05$	$0.41 \pm 0.93$	$0.32 \pm 1.31$	$0.31 \pm 0.94$
$M_*$	$9.97 \pm 0.54$	$9.75 \pm 0.56$	$9.6 \pm 0.52$	$9.56 \pm 0.61$
$\kappa_{\text{co}}^{\text{SF}}$	$0.93 \pm 0.06$	$0.9 \pm 0.08$	$0.86 \pm 0.1$	$0.76 \pm 0.18$
$r_{\text{SF-gas}}$	$8.29 \pm 5.37$	$5.88 \pm 6.74$	$3.96 \pm 9.47$	$2.58 \pm 29.75$

all particles of the relevant type within a circular aperture of radius  $r_{\text{ap}} = \max(30 \text{ pkpc}, 2r_{\text{sfg}})$ , where  $r_{\text{sfg}}$  is the 2D half-mass radius of star-forming gas within a circular aperture of 30 pkpc. The use of this additional criterion ensures a robust morphological characterization of the image projected by the most extended gas discs when viewed close to a face-on orientation. At each iteration, the elliptical aperture adapts to maintain a constant area.

Fig. 11 shows the probability distribution function of the projected ellipticity of star-forming gas (blue curves) and the stars (red curves) associated with the subhaloes of our sample. The solid curves denote the distribution of aggregated ellipticities recovered from projection of the 3D mass distributions along the line of sight of 100 ‘observers’ randomly positioned on a unit sphere, thus crudely mimicking a real light cone (albeit without noise or degradation from instrumental limitations). The dashed and dotted curves show the ellipticity distributions recovered when the subhaloes are first oriented such that the projection axis is parallel to, respectively, the minor and



**Figure 11.** Probability distribution functions of the projected 2D ellipticities of the present-day mass distributions of star-forming gas (blue curves) and stars (red curves) bound to the subhaloes of our sample. The solid curves denote the aggregated ellipticities recovered from projection of the 3D mass distributions along 100 random axes of projection. For reference, the dashed and dotted curves show the distributions recovered when the galaxies are oriented face-on and edge-on, respectively, to the axis of projection. An observational comparison (green curve with crosses) is sourced from Tunbridge et al. (2016), who provide a best-fitting model to the distribution of observed ellipticities of galaxies in the radio VLA COSMOS data set.

major principal axes of the respective 3D mass distribution, in order to show the ellipticities when viewed face-on and edge-on.

The distribution of ellipticities when projected along random lines of sight is significantly broader for the star-forming gas than is the case for the stars: the IQRs of two distributions are 0.30 and 0.19, respectively. The origin of this difference is revealed by inspection of the ellipticity distributions for the face-on and edge-on reference cases: as might be inferred from the distribution of 3D shape parameters, star-forming gas is more commonly found in flattened configurations (corresponding to large values of the projected ellipticity) than is the case for the stars. The characteristic ellipticity of the flattened structures is greater for the star-forming gas, as can be quantified via the median ellipticities,  $\bar{e}_{\text{edge}}^{\text{sf}} = 0.70$  and  $\bar{e}_{\text{edge}}^{\text{stars}} = 0.37$ . Consequently, when projected along random lines of sight, there is a mild but significant paucity of low-ellipticity star-forming structures, since observing such a configuration requires that the galaxy is oriented close to face-on. Similarly, there are few high-ellipticity stellar structures, but this deficit is greater: not only does observing such a configuration require that the galaxy is oriented close to edge-on but, crucially, stellar structures that are strongly flattened (in 3D) are rare. We note that our sample selection criteria act to minimize these differences, since galaxies with significant star-forming gas reservoirs preferentially exhibit flattened stellar discs, i.e. elliptical and spheroidal galaxies are underrepresented by our sample.

The solid green curve of Fig. 11 denotes the best-fitting functional form of the galaxy ellipticity distribution recovered from the application of the IM3SHAPE algorithm (Zuntz et al. 2013) to Very Large Array (VLA)  $L$ -band observations of galaxies in the COSMOS field (Tunbridge et al. 2016, see their equation 8). The iterative algorithm

finds the best-fitting two-component Sèrsic (disc and bulge) model, yielding two-component ellipticities  $\epsilon = (e_1, e_2)$ , and is similar in concept, if not in detail, to the approach used here to characterize the simulated galaxies. There is a remarkable correspondence between the observed ellipticity distribution and that recovered from EAGLE. The qualitative similarity is a reassuring indication that the ellipticity distribution of star-forming gas yielded by EAGLE is realistic, however we caution that the degree of agreement is likely to be, in part, coincidental: besides the differences in shape measurement algorithms and the absence of noise or smearing by a point spread function in the simulated shape measurements, the observed sample also spans a wide range of redshifts.

Tunbridge et al. (2016) noted that the dissimilar diversity of the projected ellipticities of the star-forming gas and stellar mass distributions is of practical relevance, because it governs the shape noise. This difference is analogous to the difference in shape noise in the optical regime expected for samples of early- and late-type galaxies: Joachimi et al. (2013) estimate that the former exhibit up to a factor of two less shape noise than the latter at fixed number. We assess the magnitude of this effect in EAGLE, by defining the shape noise of a sample of galaxies,  $\sigma_e$ , as

$$\sigma_e^2 = \frac{1}{N} \sum_i |e_i|^2, \quad (4)$$

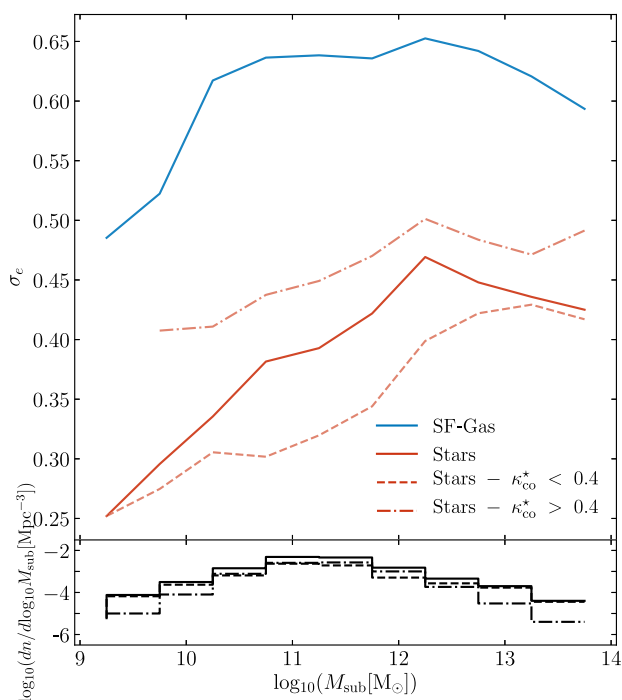
where  $N$  is the total number subhaloes in the sample. The quantity in the summation is often referred to as the polarization (see e.g. Blandford et al. 1991) and is defined as  $|e| = (a^2 - b^2)/(a^2 + b^2)$ . It is thus related to the ellipticity via  $e = 2\epsilon/(1 + |\epsilon|^2)$ .

We compute  $\sigma_e$  for the star-forming gas and stellar distributions of subhaloes as a function of subhalo mass. These measurements are shown in Fig. 12. The solid curves denote measurements for the star-forming gas (blue) and stars (red) considering all subhaloes comprising our sample. To place the difference in shape noise between the two matter types into context, we also show the shape noise of the stellar component when splitting the main sample into two subsamples separated about  $\kappa_{\text{co}}^* = 0.4$ , thus broadly separating the main sample into late- and early-type galaxies. The shape noise of the star-forming gas associated with subhaloes of all masses probed by our sample is systematically greater than is the case for their stars, by  $\Delta\sigma_e \simeq 0.19\text{--}0.25$ , an offset comparable to the difference between the shape noise (at fixed subhalo mass) of the stellar component of subhaloes comprising our crudely defined early- and late-type subsamples. Tunbridge et al. (2016) report a qualitatively similar offset of the shape noise of radio continuum sources relative to their optical images (see their table 3).

## 5.2 Projected alignment

In practice, it is only the misalignment angle of the various matter types in projection that can be measured observationally. We therefore extend the exploration of 3D misalignments presented in Section 4, to examine misalignments in projection. Fig. 13 shows the cumulative distribution function of  $\theta_{2D}$ , the alignment angle of the three pairs of matter components when viewed in projection. As with Fig. 8, we plot the distribution as a function of  $\log_{10}(1 + \theta_{2D})$  since the bulk of the misalignments are small, but show long tails to severe misalignments. Thick lines denote our fiducial measurement, whilst thin lines show the alignments inferred when the initial characterization of the projected mass distribution considers all particles of the relevant matter component bound to the subhalo. Thick dashed lines repeat the fiducial measurement for the subsample of subhaloes hosting late-type galaxies, i.e. those with  $\kappa_{\text{co}}^* > 0.4$ . For



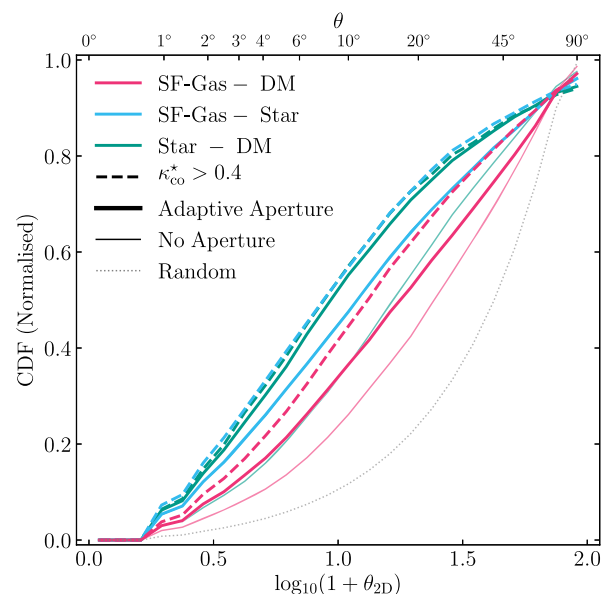


**Figure 12.** The scatter in the projected ellipticities of stars and star-forming gas, calculated as the standard deviation of the minor–major axis ratios within a given mass bin. The blue lines relate to the star-forming gas, and red lines to the stars. The solid lines correspond to the shape error in the projected galaxy shapes and the dotted line corresponds to the edge-on star-forming gas particle coordinate projection for the star-forming gas. The red dashed and dot-dashed lines correspond to the standard deviations for  $\kappa_{\text{co}}^* < 0.4$  and  $\kappa_{\text{co}}^* > 0.4$  stellar systems, respectively.  $\kappa_{\text{co}}^*$  is here the fraction of kinetic energy invested in corotation for the stars, as opposed to the star-forming gas, within 30 kpc, as outlined by Thob et al. (2019). The lower panel displays the volume density of subhaloes in a given mass bin, given as  $\log_{10}(dn/d\log_{10}M_{\text{sub}}[M_{\odot}^{-3}])$ . At all masses, the shape noise is systematically greater for galaxy populations that are intrinsically flatter.

reference, the dotted black line shows the distribution function of alignment angles between randomly oriented vectors.

The plot reveals that the projected alignments are qualitatively similar to those recovered in 3D, in so far that the star-forming gas and DM are most weakly aligned (half of all subhaloes are aligned to better than  $16.9^\circ$ ), whilst the star-forming gas – stars and stars – DM alignments are aligned significantly more closely (half of all subhaloes aligned to better than  $10.9^\circ$  and  $8.1^\circ$ , respectively). Discarding the initial aperture weakens the alignment between the more centrally concentrated baryons and the DM but, in a similar fashion to the 3D case, has little impact on the alignment between star-forming gas and stars. Restricting the sample to late-type galaxies improves the alignment of all component pairs, with half of all subhaloes being aligned to better than  $12.2^\circ$ ,  $7.7^\circ$ , and  $7.8^\circ$  for, respectively, the star-forming gas – DM, star-forming gas – stars, and stars – DM pairs.

We note that, in contrast to Tenneti et al. (2014, their fig. 10) and Velliscig et al. (2015a, their fig. 13), we find that the projected alignments are in general weaker in projection than in 3D. For example, the median alignment angle of star-forming gas and DM using our fiducial aperture choices are  $9.5^\circ$  in 3D and  $16.9^\circ$  in projection. This is a consequence of our choice, motivated in Section 2.3, to measure misalignments relative to the minor axis rather than the major axis; whilst the projected misalignment is



**Figure 13.** The projected 2D internal alignment between the stars, DM, and star-forming gas within the subhaloes of our sample. The figure displays a normalized cumulative distribution function of the angle  $\theta_{2D}$  between the minor axes of various matter distributions within subhaloes. The line colour indicates the two matter types assessed, thick dashed, and thin solid lines correspond to the aperture used in the computation of the iterative reduced inertia tensor. The black dotted line indicates the distribution of angles between randomly orientated vectors in 2D. Star-forming gas is a poorer tracer of the underlying DM distribution than the stars in terms of orientation.

insensitive to this choice, the choice has a significant bearing on the alignments in 3D. We have explicitly confirmed that switching from the use of the minor axis to the major axis to define the misalignment angle results in smaller misalignments when projecting from 3D, consistent with the findings of Tenneti et al. (2014) and Velliscig et al. (2015a). Although not shown in the figure, we have further examined the misalignment angles of all matter component pairs at  $z = 1$ , and find more severe misalignments at the earlier cosmic epoch. This result is largely insensitive to the use of the axisymmetry criterion.

## 6 SUMMARY AND DISCUSSION

We have investigated the morphology of, and mutual alignments between, the star-forming gas, stars and dark matter bound to subhaloes that form in the EAGLE suite of simulations (Crain et al. 2015; Schaye et al. 2015; McAlpine et al. 2016). Our study is motivated by the complementarity of weak lensing experiments conducted using radio continuum surveys with traditional optical surveys. While recent radio weak lensing studies were limited by low source densities (see e.g. Tunbridge et al. 2016; Hillier et al. 2019; Harrison et al. 2020), the next-generation Square Kilometer Array (SKA) radio telescope will be competitive with optical surveys at a higher characteristic redshift. In simulations like EAGLE, gas that has a non-zero star formation rate is a good proxy for gas that is bright in the radio continuum. EAGLE represents a judicious test-bed for an assessment of this kind, as the simulations were calibrated to ensure a good reproduction of the galaxy stellar mass function and the size–mass relation of late-type galaxies. We focus primarily on

present-day subhaloes, but also examine the simulations at earlier times to explore evolutionary trends.

A summary of our results is as follows:

(i) The star-forming gas distribution of present-day subhaloes is typically flattened (i.e. low sphericity) along its minor axis. Flattening is most pronounced in subhaloes of  $M_{\text{sub}} \sim 10^{12.5} M_{\odot}$ , for which the median sphericity is  $\tilde{S}_{\text{SF-gas}} = 0.1$ . The distribution of star-forming gas sphericities is significantly narrower than that of stars and dark matter at all subhalo masses, but particularly for those of  $M_{\text{sub}} = 10^{12-12.5} M_{\odot}$ , for which the interquartile ranges of star-forming gas, stars, and DM are 0.06, 0.15, and 0.12, respectively (Fig. 3).

(ii) Star-forming gas exhibits a diverse range of triaxiality parameters. Subhaloes of mass  $M_{\text{sub}} \sim 10^{12-12.5} M_{\odot}$  typically host oblate distributions consistent with classical gas discs, but in both low and high mass subhaloes, the distributions are more often prolate (Fig. 3).

(iii) Star-forming gas is less flattened at earlier epochs, for all subhalo masses examined, irrespective of whether one considers a sample selected in a similar fashion to the present-day sample, or considers the progenitors of the latter. Strongly flattened star-forming gas structures ( $S \lesssim 0.2$ ) emerge only at  $z \lesssim 2$ , broadly coincident with the growth of the disc's scale length (Figs 4 and 5).

(iv) The shape parameters describing the morphology of star-forming gas are strongly and positively correlated with those describing the stellar morphology of the host galaxy, such that e.g. flattened gas structures are associated with flattened stellar structures (Fig. 6).

(v) The minor axis of the star-forming distribution preferentially aligns most closely with the minor axis of the (inner) DM halo. However, in prolate subhaloes  $T_{\text{DM}}(r < 30 \text{ pkpc}) \gtrsim 0.7$ , a significant fraction of galaxies have star-forming gas distributions whose minor axis most closely aligns with one of the other principal axes of the DM (Fig. 7).

(vi) Characterized by the angle between the minor axes of the respective components of subhaloes, star-forming gas tends to align with the DM (i.e. the alignment is stronger than random), but the alignment is weaker than is the case for stars and the DM. This is the case for both the 3D matter distributions (Fig. 8) and their projections on the sky (Fig. 13). The alignments are strongest when considering the inner DM halo, and in general the alignments are stronger for late-type galaxies.

(vii) The alignment of the star-forming gas distribution with those of both the stars and the DM bound to its parent subhalo is typically weaker at early cosmic epochs (Fig. 9).

(viii) The kinematic axis of star-forming gas aligns closely with its minor morphological axis, with most galaxies being aligned to better than  $10^\circ$  at the present-day, and better than  $6^\circ$  if only late-type galaxies are considered. The alignment is poorer at  $z = 1$ , with these characteristic misalignment angles doubling (Fig. 10).

(ix) The more pronounced flattening of star-forming gas structures leads to them exhibiting a broader distribution of projected ellipticities than is the case for stellar structures, analogous to the differing ellipticity distributions of optical images of late-type and early-type galaxies. The ellipticity distribution of star-forming gas in EAGLE corresponds closely to that recovered from high-fidelity VLA radio continuum images of galaxies in the COSMOS field (Fig. 11). For a fixed subhalo sample, the 'shape noise' of its star-forming gas is therefore systematically greater than that of its stars (12).

Our analyses reveal that the morphology of star-forming gas distributions, and their orientation with respect to the DM of their parent subhalo, are more complex than might be naively assumed. This complexity is particularly relevant in the context of using

extended star-forming gas distributions, which can be imaged in the radio continuum, to conduct weak lensing experiments.

Forecasts for the outcomes of the next generation of the 'megasureys' require that very large cosmic volumes are modelled. The associated expense of including the baryonic component forces the use of empirical, analytic, or semi-analytic models grafted on to treatments of the evolving cosmic dark matter distribution. By construction, such techniques do not capture the full complexity of the evolution of the baryonic component resulting from the diverse range of physical processes that influence galaxies, nor do they capture the 'back reaction' of the baryons on the DM, and so can mask the importance of key systematic uncertainties.

In the specific case of modelling the radio continuum sky, the most popular approach has been to couple observed source populations with either a Press–Schechter or  $N$ -body treatment of the evolving cosmic DM distribution (see e.g. Wilman et al. 2008; Bonaldi et al. 2019). By construction, such models invoke no explicit connection between the properties of star-forming gas structures and their parent DM haloes, and often relate (or equate) the properties of the former to those of the host galaxy's stellar component. Our analyses highlight shortcomings of these approximations: the characteristic morphology of star-forming gas is a strong function of the mass of its host subhalo and, although the simulations indicate that it correlates strongly with the morphology of its associated stellar component, we find that the respective morphologies can differ significantly.

We also find that star-forming gas structures are imperfectly aligned with both the stellar and DM components of their host subhalo. Although the misalignment angle is generally small (particularly with respect to the stellar component), there is a long-tail to severe misalignments, and we find that the misalignment is most pronounced in early-type galaxies. We also find that the misalignment of the star-forming gas with the DM of its host subhalo becomes more pronounced if the outer halo is considered (for instance, if disabling the use of the 30 pkpc spherical aperture). Therefore, when constructing semi-empirical radio sky models based on  $N$ -body simulations, we caution against naively orienting star-forming discs with the principle axes of the DM distribution.

Our analyses also highlight that the shape noise of images of a fixed sample of galaxies seen in the radio continuum should be significantly greater than when seen in the optical. This follows naturally from the lower characteristic sphericity (or, alternatively, the greater flattening) of star-forming gas structures than their stellar counterparts. A systematic offset in shape noise was previously reported by Tunbridge et al. (2016) following the examination of a relatively small sample of galaxies with high-fidelity radio and optical imaging. The corollary of this finding is that radio continuum weak lensing experiments will require a greater source density in order to obtain a signal-to-noise ratio equal to optical experiments. However, our analyses also corroborate the hypothesis that the use of the kinematic axis (revealed by ancillary 21 cm observations) affords an effective means of estimating the unsheared orientation of the minor axis, and thus mitigating the systematic uncertainty in radio weak lensing experiments.

An interesting consequence of the poorer alignment of star-forming gas structures with the DM of their host subhaloes than is the case for the stars – DM alignment, is that it implies that the intrinsic alignment signal may be less severe in radio weak lensing surveys than is the case for optical counterparts. In a follow-up paper, Hill et al. (in preparation), we examine the two key 'intrinsic alignment' signals recoverable from radio continuum imaging, namely the orientation of star-forming gas distributions

with respect to the directions to, and orientations of, the star-forming gas structures of its neighbouring galaxies.

## ACKNOWLEDGEMENTS

We gratefully thank Adrien Thob and Jonathan Davies for fruitful discussions and technical help. We thank the anonymous reviewer whose constructive suggestions improved both the content and presentation of this study. ADH is supported by an STFC doctoral studentship within the Liverpool Big Data Science Centre for Doctoral Training, hosted by Liverpool John Moores University and the University of Liverpool [ST/P006752/1]. RAC is a Royal Society University Research Fellow. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 769130). The study made use of high performance computing facilities at Liverpool John Moores University, partly funded by the Royal Society and LJMU's Faculty of Engineering and Technology, and the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility ([www.dirac.ac.uk](http://www.dirac.ac.uk)). This equipment was funded by BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grants ST/H008519/1 and ST/K00087X/1, STFC DiRAC Operations grant ST/K003267/1, and Durham University. DiRAC is part of the National E-Infrastructure.

## DATA AVAILABILITY

Particle data, and derived data products from the simulations have been released to the community as detailed by McAlpine et al. (2016). Further derived data used in this article will be shared on reasonable request to the corresponding author.

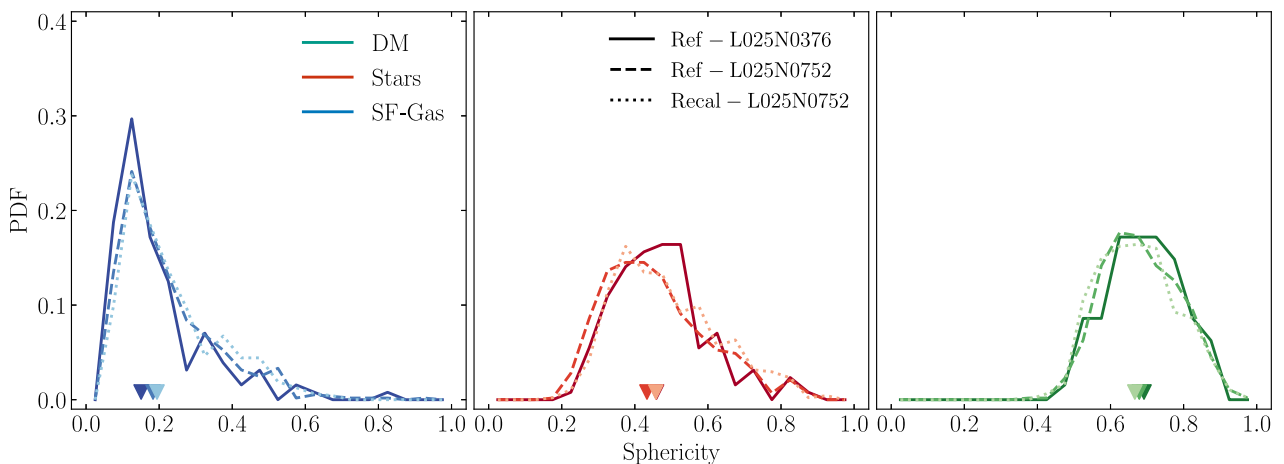
## REFERENCES

- Aihara H. et al., 2018, *PASJ*, 70, S8  
 Bahé Y. M. et al., 2016, *MNRAS*, 456, 1115  
 Battye R. A., Browne I. W. A., 2009, *MNRAS*, 399, 1888  
 Benítez-Llambay A., 2015, py-sphviewer: Py-sphviewer v1.0.0. Available at: <http://dx.doi.org/10.5281/zenodo.21703>  
 Benítez-Llambay A., Navarro J. F., Frenk C. S., Ludlow A. D., 2018, *MNRAS*, 473, 1019  
 Bett P., 2012, *MNRAS*, 420, 3303  
 Bett P., Eke V., Frenk C. S., Jenkins A., Okamoto T., 2010, *MNRAS*, 404, 1137  
 Blain A. W., 2002, *ApJ*, 570, L51  
 Blandford R. D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, *MNRAS*, 251, 600  
 Bonaldi A., Bonato M., Galluzzi V., Harrison I., Massardi M., Kay S., De Zotti G., Brown M. L., 2019, *MNRAS*, 482, 2  
 Bond J. R., Kofman L., Pogosyan D., 1996, *Nature*, 380, 603  
 Booth C. M., Schaye J., 2009, *MNRAS*, 398, 53  
 Bosma A., 1978, PhD thesis, Groningen Univ.  
 Bower R. G., Schaye J., Frenk C. S., Theuns T., Schaller M., Crain R. A., McAlpine S., 2017, *MNRAS*, 465, 32  
 Boylan-Kolchin M., Ma C.-P., Quataert E., 2006, *MNRAS*, 369, 1081  
 Brown M. et al., 2015, Proc. Sci., Weak gravitational lensing with the Square Kilometre Array, SISSA, Trieste, PoS(AASKA14)023  
 Bryan S. E., Kay S. T., Duffy A. R., Schaye J., Dalla Vecchia C., Booth C. M., 2013, *MNRAS*, 429, 3316  
 Chabrier G., 2003, *PASP*, 115, 763  
 Chang T.-C., Refregier A., Helfand D. J., 2004, *ApJ*, 617, 794  
 Chaves-Montero J., Angulo R. E., Schaye J., Schaller M., Crain R. A., Furlong M., Theuns T., 2016, *MNRAS*, 460, 3100  
 Chen D. N., Jing Y. P., Yoshikawa K., 2003, *ApJ*, 597, 35  
 Chisari N. et al., 2015, *MNRAS*, 454, 2736  
 Chisari N. E. et al., 2017, *MNRAS*, 472, 1163  
 Cleveland W. S., 1979, J. Am. Stat. Assoc., 74, 829  
 Codis S. et al., 2015, *MNRAS*, 448, 3391  
 Codis S., Jindal A., Chisari N. E., Vibert D., Dubois Y., Pichon C., Devriendt J., 2018, *MNRAS*, 481, 4753  
 Cole S., Lacey C., 1996, *MNRAS*, 281, 716  
 Colless M. et al., 2003, preprint (astro-ph/0306581)  
 Condon J. J., 1992, *ARA&A*, 30, 575  
 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  
 Crain R. A. et al., 2015, *MNRAS*, 450, 1937  
 Crain R. A. et al., 2017, *MNRAS*, 464, 4204  
 Croft R. A. C., Di Matteo T., Springel V., Hernquist L., 2009, *MNRAS*, 400, 43  
 Cuesta A. J., Betancort-Rijo J. E., Gottlöber S., Patiri S. G., Yepes G., Prada F., 2008, *MNRAS*, 385, 867  
 Cullen L., Dehnen W., 2010, *MNRAS*, 408, 669  
 Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 426, 140  
 Davé R., Crain R. A., Stevens A. R. H., Narayanan D., Saintonge A., Catinella B., Cortese L., 2020, *MNRAS*, 497, 146  
 Davies J. J., Crain R. A., Oppenheimer B. D., Schaye J., 2020, *MNRAS*, 491, 4462  
 Davies J. J., Crain R. A., Pontzen A., 2021, *MNRAS*, 501, 236  
 Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, 292, 371  
 de Burgh-Day C. O., Taylor E. N., Webster R. L., Hopkins A. M., 2015, *MNRAS*, 451, 2161  
 de Jong J. T. A. et al., 2015, *A&A*, 582, A62  
 de Lapparent V., Geller M. J., Huchra J. P., 1986, *ApJ*, 302, L1  
 Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497  
 Dubinski J., 1994, *ApJ*, 431, 617  
 Dubinski J., 1998, *ApJ*, 502, 141  
 Dubinski J., Carlberg R. G., 1991, *ApJ*, 378, 496  
 Dubois Y. et al., 2014, *MNRAS*, 444, 1453  
 Durier F., Dalla Vecchia C., 2012, *MNRAS*, 419, 465  
 Erben T. et al., 2013, *MNRAS*, 433, 2545  
 Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, *MNRAS*, 406, 2267  
 Faucher-Giguère C.-A., Lidz A., Hernquist L., 2008, *Science*, 319, 52  
 Furlong M. et al., 2017, *MNRAS*, 465, 722  
 Geller M. J., Huchra J. P., 1989, *Science*, 246, 897  
 Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *ApJ*, 622, 759  
 Guo Q. et al., 2011, *MNRAS*, 413, 101  
 Guo Q. et al., 2016, *MNRAS*, 461, 3457  
 Haardt F., Madau P., 2001, in Neumann D. M., Tran J. T. V., eds, Clusters of Galaxies and the High Redshift Universe Observed in X-rays, CEA, 2002, Les Arcs, Savoie, France. p. 64  
 Hahn O., Teyssier R., Carollo C. M., 2010, *MNRAS*, 405, 274  
 Harrison I. et al., 2020, *MNRAS*, 495, 1737  
 Hillier T., Brown M. L., Harrison I., Whittaker L., 2019, *MNRAS*, 488, 5420  
 Hopkins P. F., 2013, *MNRAS*, 428, 2840  
 Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, *MNRAS*, 440, 2115  
 Joachimi B., Semboloni E., Bett P. E., Hartlap J., Hilbert S., Hoekstra H., Schneider P., Schrabback T., 2013, *MNRAS*, 431, 477  
 Joachimi B. et al., 2015, Space Sci. Rev., 193, 1  
 Kaiser N., 1984, *ApJ*, 284, L9  
 Katz N., Quinn T., Bertschinger E., Gelb J. M., 1994, *MNRAS*, 270, L71  
 Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, *MNRAS*, 303, 188  
 Kazantzidis S., Kravtsov A. V., Zentner A. R., Allgood B., Nagai D., Moore B., 2004, *ApJ*, 611, L73  
 Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, *MNRAS*, 450, 1349  
 Lacey C., Cole S., 1993, *MNRAS*, 262, 627  
 Lagos C. d. P. et al., 2015, *MNRAS*, 452, 3815  
 Lagos C. d. P. et al., 2016, *MNRAS*, 459, 2632

- Laureijs R. et al., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann Jr. J. M., eds, Proc. SPIE Conf. Ser. Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 84420T
- LSST Science Collaboration et al., 2009, preprint ([arXiv:0912.0201](https://arxiv.org/abs/0912.0201))
- Ludlow A. D., Schaye J., Schaller M., Richings J., 2019, *MNRAS*, 488, L123
- McAlpine S. et al., 2016, *Astron. Comput.*, 15, 72
- McCarthy I. G., Schaye J., Bird S., Le Brun A. M. C., 2017, *MNRAS*, 465, 2936
- Morales M. F., 2006, *ApJ*, 650, L21
- Naab T., Khochfar S., Burkert A., 2006, *ApJ*, 636, L81
- Oppenheimer B. D. et al., 2020, *MNRAS*, 491, 2939
- Patel P., Bacon D. J., Beswick R. J., Muxlow T. W. B., Hoyle B., 2010, *MNRAS*, 401, 2572
- Pillepich A. et al., 2018, *MNRAS*, 475, 648
- Pillepich A. et al., 2019, *MNRAS*, 490, 3196
- Planck Collaboration XVI et al., 2014, *A&A*, 571, A16
- Price D. J., 2008, *J. Comput. Phys.*, 227, 10040
- Qu Y. et al., 2017, *MNRAS*, 464, 1659
- Rosas-Guevara Y. M. et al., 2015, *MNRAS*, 454, 1038
- Sales L. V., Navarro J. F., Theuns T., Schaye J., White S. D. M., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012, *MNRAS*, 423, 1544
- Schaller M., Dalla Vecchia C., Schaye J., Bower R. G., Theuns T., Crain R. A., Furlong M., McCarthy I. G., 2015, *MNRAS*, 454, 2277
- Schawinski K. et al., 2014, *MNRAS*, 440, 889
- Schaye J., 2004, *ApJ*, 609, 667
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schneider M. D., Bridle S., 2010, *MNRAS*, 402, 2127
- Schneider M. D., Frenk C. S., Cole S., 2012, *J. Cosmol. Astropart. Phys.*, 5, 030
- Schober J., Schleicher D. R. G., Klessen R. S., 2017, *MNRAS*, 468, 946
- Shandarin S., Habib S., Heitmann K., 2010, *Phys. Rev. D*, 81, 103006
- Shao S., Cautun M., Frenk C. S., Gao L., Crain R. A., Schaller M., Schaye J., Theuns T., 2016, *MNRAS*, 460, 3772
- Sharma S., Steinmetz M., 2005, *ApJ*, 628, 21
- Spergel D. et al., 2015, preprint ([arXiv:1503.03757](https://arxiv.org/abs/1503.03757))
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springel V., White S. D. M., Hernquist L., 2004, in Ryder S., Pisano D., Walker M., Freeman K., eds, Proc. IAU Symp. 220, Dark Matter in Galaxies. Astron. Soc. Pac., San Francisco, CA, p. 421
- Springel V., Di Matteo T., Hernquist L., 2005a, *MNRAS*, 361, 776
- Springel V. et al., 2005b, *Nature*, 435, 629
- Springel V. et al., 2018, *MNRAS*, 475, 676
- Square Kilometre Array Cosmology Science Working Group, 2020, *Publ. Astron. Soc. Austr.*, 37, e007
- Swaters R. A., 1999, PhD thesis, Rijksuniversiteit Groningen
- Tegmark M. et al., 2004, *ApJ*, 606, 702
- Tenneti A., Mandelbaum R., Di Matteo T., Feng Y., Khandai N., 2014, *MNRAS*, 441, 470
- Tenneti A., Singh S., Mandelbaum R., di Matteo T., Feng Y., Khandai N., 2015, *MNRAS*, 448, 3522
- The Dark Energy Survey Collaboration, 2005, preprint (astro-ph/0510346)
- Thob A. C. R. et al., 2019, *MNRAS*, 485, 972
- 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
- Troxel M. A., Ishak M., 2015, *Phys. Rep.*, 558, 1
- Tunbridge B., Harrison I., Brown M. L., 2016, *MNRAS*, 463, 3339
- van de Voort F., Schaye J., Booth C. M., Haas M. R., Dalla Vecchia C., 2011, *MNRAS*, 414, 2458
- Velliscig M. et al., 2015a, *MNRAS*, 453, 721
- Velliscig M. et al., 2015b, *MNRAS*, 454, 3328
- Vogelsberger M. et al., 2014, *Nature*, 509, 177
- Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, *ApJ*, 652, 71
- Wendland H., 1995, *Adv. Comput. Math.*, 04, 389
- White S. D. M., Rees M. J., 1978, *MNRAS*, 183, 341
- White S. D. M., Davis M., Efstathiou G., Frenk C. S., 1987, *Nature*, 330, 451
- Whittaker L., Brown M. L., Battye R. A., 2015, *MNRAS*, 451, 383
- 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
- Wilman R. J. et al., 2008, *MNRAS*, 388, 1335
- Zemp M., Gnedin O. Y., Gnedin N. Y., Kravtsov A. V., 2012, *ApJ*, 748, 54
- Zhang Y., Yang X., Wang H., Wang L., Luo W., Mo H. J., van den Bosch F. C., 2015, *ApJ*, 798, 17
- Zuntz J., Kacprzak T., Voigt L., Hirsch M., Rowe B., Bridle S., 2013, *MNRAS*, 434, 1604

## APPENDIX A: NUMERICAL CONVERGENCE

In this section, we examine the influence of the numerical resolution of the EAGLE simulations on the recovered sphericity of the star-forming gas, stars, and DM comprising subhaloes. We follow Schaye



**Figure A1.** Probability distribution function of the sphericity parameter of the star-forming gas (left-hand panel), stars (centre), and DM (right-hand panel) of present-day subhaloes drawn from the Ref-L025N0376 (solid dark-coloured curve), Ref-L025N0752 (dashed medium), and Recal-L025N0752 (dotted light) simulations. Down arrows denote the median sphericity of the distribution of each simulation. Comparison of Ref-L025N0376 with Ref-L025N0752 and Recal-L025N0752 affords simple tests of, respectively, the strong and weak convergence behaviour of the star-forming gas sphericity.



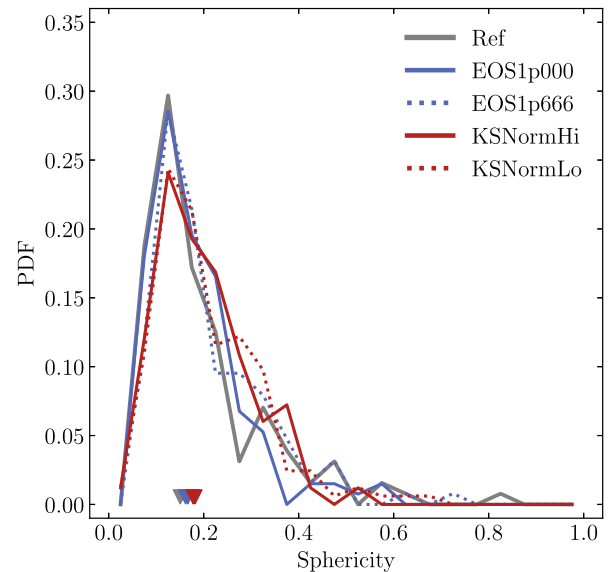
et al. (2015) and adopt the terms ‘strong convergence’ and ‘weak convergence’, where the former denotes a comparison at different resolutions of a fixed physical model, and the latter denotes a comparison at different resolutions of two models calibrated to recover the same observables. We use three  $L = 25$  cMpc simulations introduced by Schaye et al. (2015): Ref-L025N0376, which is identical to the flagship Ref-L100N1504 simulation with the exception of the boxsize; Ref-L025N0752, which adopts the same reference physical model but has a factor of 8 more particles each of both baryons and DM; and Recal-L025N0752 which also adopts values for subgrid parameters governing stellar and AGN feedback that have been recalibrated to improve the match to the galaxy stellar mass function and galaxy sizes at this higher resolution. Comparison of Ref-L025N0376 with Ref-L025N0752 and Recal-L025N0752 thus affords simple tests of, respectively, the strong and weak convergence behaviour.

Fig. A1 shows the probability distribution functions of the sphericities of the star-forming gas (left-hand panel), stars (centre), and DM (right-hand panel) for each of the three  $L = 25$  cMpc simulations. The subhaloes shown are selected according to the standard sampling criteria outlined in the Section 2.4, irrespective of the resolution of the simulation. Down arrows denote the median sphericity of the distribution of each simulation. Inspection shows that the distributions are not strongly influenced by the change in resolution. The median values of the sphericity of the three matter components in the Ref-L025N0376 simulation are 0.15, 0.50, and 0.69 for the star-forming gas, stars and DM, respectively. As can be clearly seen from the figure, when moving to the high resolution simulations, the shift in median values is much smaller than the associated interquartile ranges of the Ref-L025N0376 simulation (IQR = 0.14, 0.15, 0.14 for the three components, respectively). Although not shown here, we recover similar behaviour when focusing on the triaxiality parameter.

## APPENDIX B: INFLUENCE OF SUBGRID ISM TREATMENTS

In this section, we examine the sensitivity of star-forming gas morphologies to aspects of EAGLE’s subgrid models that in principle influence the structure of interstellar gas directly, namely the form of the temperature floor equation of state and the star formation law. To achieve this, we compare the Ref-L025N0376 simulation with two pairs of complementary L025N0376 simulations. The first pair, introduced by Crain et al. (2015), varies the slope of the equations of state from the reference value of  $\gamma_{\text{eos}} = 4/3$  with different slopes, to adopt isothermal ( $\gamma_{\text{eos}} = 1$ ) and adiabatic ( $\gamma_{\text{eos}} = 5/3$ ) equations of state. Schaye & Dalla Vecchia (2008) used simulations of idealized discs to show that a stiffer equation of state generally leads to smoother star-forming gas distributions with a larger scale height. Crain et al. (2015) showed that in EAGLE, a stiffer equation of state also suppresses accretion on to the central BH in massive galaxies. The second pair, introduced by Crain et al. (2017), varies the normalization of the Kennicutt–Schmidt law (the variable  $A$  in equation 1 of Schaye et al. 2015) from its fiducial value of  $1.515 \times 10^{-4} \text{ M}_{\odot} \text{ yr}^{-1} \text{ kpc}^{-2}$  by  $\pm 0.5$  dex. Crain et al. (2017) showed that increasing (decreasing) this parameter tends to decrease (increase) the mass of cold gas associated with galaxies, since it governs the mass of gas that is required to maintain a balance between the gas infall rate and the outflow rate due to effective feedback.

Fig. B1 shows probability distribution function of the sphericity star-forming gas for the reference model (solid black curve) and the simulations with differing equations of state ( $\gamma_{\text{eos}} = 1$ , solid blue;  $\gamma_{\text{eos}} = 5/3$ , dotted blue), and with higher (solid red) and lower (dotted

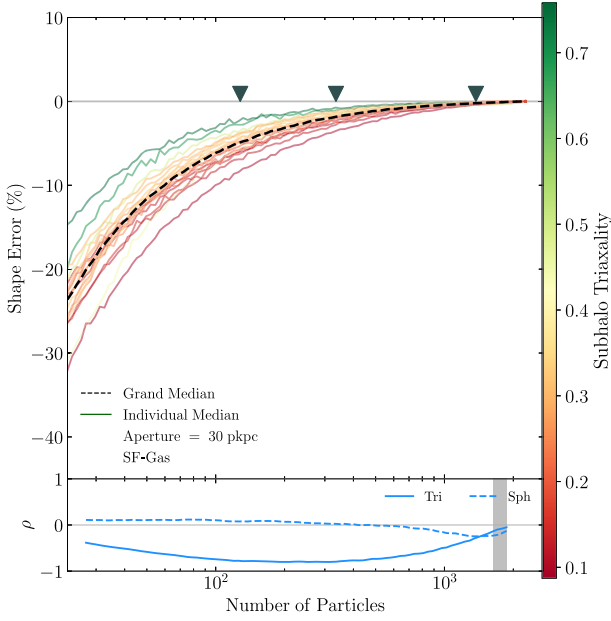


**Figure B1.** Probability distribution function of the sphericity parameter of the star-forming gas of present-day subhaloes drawn from the Ref-L025N0376 simulation (black curve) and two pairs of simulations that incorporate variations of the reference model: pair with alternative equation of state slopes (EOS1p00, solid blue; and EOS1p666, dotted blue) and with normalizations of the star formation law adjusted by  $\pm 0.5$  dex (KSNormHi, solid red; KSNormLo, dotted red). Comparison of these runs with the reference model indicates the influence of the subgrid ISM model on the sphericity of star-forming gas distributions.

red) normalizations of the star formation law with respect to the reference model. The subhaloes shown are selected according to the standard sampling criteria outlined in the Section 2.4. Down arrows denote the median sphericity of the distribution of each simulation. Inspection reveals that the distributions are not strongly influenced by changes to the subgrid modelling of the ISM. The median value of the sphericity of the star-forming gas in the Ref-L025N0376 simulation is 0.15. As can be clearly seen from the figure, the median sphericity in the three variation simulations shifts by  $< 0.05$  with respect to the reference simulation, a value that is much smaller than the interquartile range of the reference case. Although not shown here, we recover similar behaviour when focusing on the triaxiality parameter.

## APPENDIX C: THE INFLUENCE OF PARTICLE SAMPLING ON SHAPE CHARACTERIZATION

The morphological characterization of structures defined by particle distributions is unavoidably influenced by sampling error. It is therefore crucial to establish the reliability of such characterizations as a function of particle number. A common methodology is to realize a mass distribution of a known analytic form with a particle distribution, and assess the deviation of the recovered shape from the input shape as the distribution is progressively subsampled (see e.g. Appendix A2 of Velliscig et al. 2015a). We adopt a similar approach here but, since star-forming gas distributions are not readily characterized by a simple analytic form, we instead draw 20 central subhaloes from the sample described in Section 2.4, with dynamical mass comparable to that of the Milky Way ( $M_{\text{sub}} \simeq 10^{12.0-12.5} \text{ M}_{\odot}$ ). We compute their ‘true’ shape parameters by applying the algorithm defined in Section 2.3 using all star-forming gas particles ( $N_{\text{part}} \simeq 2000$ ). We then progressively subsample the particle distribution to



**Figure C1.** The sphericity shape error recovered as a function of the degree of particle subsampling for star-forming gas in 20 present-day subhaloes with dynamical mass similar to that of the Milky Way. Down arrows correspond to the 10th, 50th, and 90th percentile values of the number of star-forming gas particles within the subhalo sample. Curves show the median shape error recovered from  $10^5$  random subsamplings of the true star-forming gas particle distribution, and are coloured by the latter’s true triaxiality. The dashed curve shows the median recovered by aggregating measurements from all 20 subhaloes. The subpanel shows the running Spearman rank correlation coefficients relating the shape error to the true value of the triaxiality (solid curve) and sphericity (dashed curve) of the star-forming gas.

lower  $N_{\text{part}}$ , generating  $10^5$  realizations at each value of  $N_{\text{part}}$ , and recompute the shape parameters.

Fig. C1 shows the median of the relative error on the sphericity of the star-forming gas distribution recovered from the  $10^5$  subsamplings of the particle distribution as a function of  $N_{\text{part}}$ . The curves are coloured by the ‘true’ value of the triaxiality parameter of the subhalo’s star-forming gas. The dashed black curve shows the

‘grand median’ recovered by aggregating the measurements from all 20 subhaloes. Down arrows show the 10th, 50th, and 90th percentile values of the number of star-forming gas particles within the subhalo sample. Velliscig et al. (2015a) noted that poor particle sampling leads to a systematic underestimate of the sphericity parameter of the DM; we find this is also the case for the star-forming gas. A shape error of less than 10 per cent typically requires at least  $N_{\text{part}} = 100$ , hence we adopt this threshold as our lower limit for our sample selection.

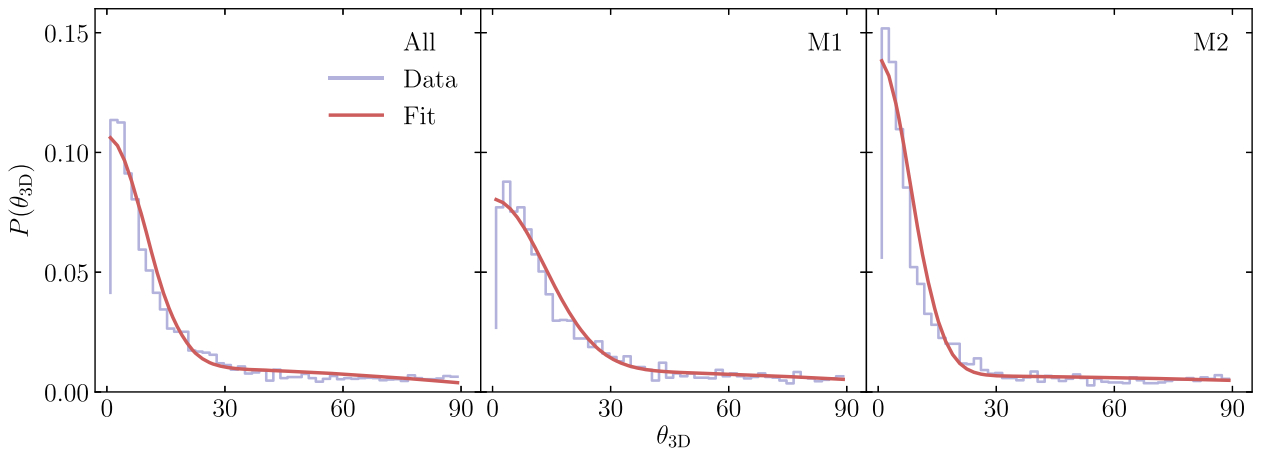
Inspection of the curves for the individual subhaloes indicates that this value is sensitive to the triaxiality of the structure, with accurate recovery of the sphericity requiring fewer particles in prolate ( $T > 0.5$ ) distributions. As shown in Fig. 4, the star-forming gas of low-mass subhaloes is preferentially prolate, hence a minimum of  $N_{\text{part}} = 100$  can be considered a conservative choice. For completeness, the subpanel shows the ‘running’ value of the Spearman rank coefficient recovered from  $N_{\text{part}}$ -ordered subsamples, of the correlation between the absolute shape error and the true shape parameters,  $T$  (solid curve) and  $S$  (dashed curve). The solid curve highlights that a negative correlation between the shape error on sphericity and the true triaxiality persists to over 1000 particles. The dashed curve indicates that there is a very mild positive correlation of the relative shape error on sphericity with the true input sphericity.

#### APPENDIX D: ANALYTIC FITS TO THE MISALIGNMENT ANGLE DISTRIBUTIONS

We provide fitting functions to the distribution of internal misalignment angles between star-forming gas and DM for present-day subhaloes in three mass bins from the EAGLE Ref-L100N1504 simulation, in both 2D and 3D. The fits enable users of  $N$ -body simulations to populate subhaloes with galaxies oriented with respect to the minor axis of the subhalo in a realistic fashion. We fit to  $P(\theta)$  using the following functional form:

$$\mathcal{M}(\theta) = C \exp\left(-\frac{\theta^2}{2\sigma_1^2}\right) + D \exp\left(-\frac{\theta^2}{2\sigma_2^2}\right) + E, \quad (\text{D1})$$

where  $C, D, \sigma_1^2, \sigma_2^2, E$  are the free parameters, and  $\theta$  is the misalignment angle. The same form was used by Velliscig et al. (2015a) to fit to the misalignment angle of stars and DM in projection.



**Figure D1.** Probability distribution functions  $P(\theta_{3\text{D}})$ , where  $\theta_{3\text{D}}$  is the angle between the morphological minor axes of stars and DM within the sample of subhaloes. A fiducial aperture of 30 pkpc is imposed. The faded step functions show the raw histograms, while the smooth lines are their respective analytic fits described by equation (D1). Panels correspond to different mass bins: the full sample (left-hand panel),  $M_{\text{sub}} \leq 10^{11.5} M_{\odot}$  (middle), and  $10^{11.5} M_{\odot} \leq M_{\text{sub}} < 10^{13} M_{\odot}$  (right-hand panel). During fitting, errors in the y-axis were taken to be the  $1\sigma$  Poisson errors.

**Table D1.** Best-fitting parameters for equation (D1), used to fit the probability distribution functions of the intrinsic 3D and projected 2D misalignment angle between star-forming gas and DM within present-day subhaloes of three mass bins (denoted by *italics*). Parameters are provided for the angles recovered using our fiducial initial aperture for the iterative reduced inertia tensor, and for no aperture (denoted by font weight).

Aperture and Mass-Bin	2D					3D				
	<i>C</i>	<i>D</i>	<i>E</i>	$\sigma_1$	$\sigma_2$	<i>C</i>	<i>D</i>	<i>E</i>	$\sigma_1$	$\sigma_2$
<b>Fiducial</b>										
<i>All</i>	24.0	8.33	0.0199	0.0507	0.00717	5130.0	9.62	43.2	0.0962	−43.2
<i>M1</i>	−27.0	9.21	0.0205	0.042	0.00688	−13.3	3050.0	0.0714	9.06	−9.05
<i>M2</i>	9.32	−9.32	−218.0	218.0	0.0107	−8.13	1440.0	0.132	1.03	−1.03
<b>No Aperture</b>										
<i>All</i>	−17.4	17.4	−87.3	87.3	0.0122	29.3	−29.3	−157.0	157.0	0.0125
<i>M1</i>	−17.7	17.7	−71.7	71.7	0.0124	31.2	−31.2	−181.0	181.0	0.013
<i>M2</i>	−17.3	17.3	−79.8	79.8	0.0124	26.9	−26.9	−114.0	114.0	0.0121

We calculate the best-fitting parameters with the Python package `SCIPY.OPTIMIZE.CURVE_FIT`, using  $1\sigma$  Poisson errors.

The best-fitting parameters are quoted in Table D1. Parameters are recovered for the misalignment angles in both the cases of (i) applying our fiducial aperture to the initial step of the iterative algorithm, and (ii) applying no initial aperture, i.e. considering all particles bound to the subhalo. In addition to presenting best-fitting parameters for all subhaloes in our sample (‘All’), we provide fits to subsamples ‘M1’ and ‘M2’, which are subject to the additional criteria  $\log_{10}M_{\text{sub}}[M_{\odot}] \leq 11.5$  (M1) and  $11.5 < \log_{10}M_{\text{sub}}[M_{\odot}] \leq 13$  (M2). This is motivated by two factors. First, below  $M_{\text{sub}} = 10^{11.5} M_{\odot}$  our selection criteria result in significant incompleteness. Secondly, the misalignment of the minor axes of the star-forming gas and the DM components becomes large for  $M_{\text{sub}} > 10^{13} M_{\odot}$  (see discussion in Section 4.1), severely degrading the value of the fits. The best fits for the 3D fiducial aperture case are shown in Fig. D1.

We find that the fitting is able to recover the profile of the input distribution fairly successfully. As an example we find the

percentage difference in the retrieved median as compared with the input distribution to be (1.7, 0.14, 0.51) per cent for the three cases

displayed in the panels of Fig. D1, while for the standard deviation this becomes (0.72, 1.04, 1.16) per cent. For the no aperture version of these cases we find errors of (3.1, 2.2, 2.6) per cent for the median and (1.6, 1.8, 1.7) per cent for the standard deviation. When no aperture is applied, we find that the errors in the 2D fittings are comparable to the 3D case. However, with the fiducial aperture the errors are noticeably larger for the 2D case, the largest being  $\sim 10$  per cent for the median and standard deviation of the M2 bin. Twelve figures comprising all variations displayed in Table D1 (two dimensions  $\times$  two apertures  $\times$  three mass bins) in the style of Fig. D1 may be found at the author’s website.<sup>6</sup>

<sup>6</sup>[www.astro.ljmu.ac.uk/~ariahill/](http://www.astro.ljmu.ac.uk/~ariahill/)

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