



The High Fraction of Thin Disk Galaxies Continues to Challenge Λ CDM Cosmology

Moritz Haslbauer^{1,2} , Indranil Banik^{1,3} , Pavel Kroupa^{1,4} , Nils Wittenburg¹ , and Behnam Javanmardi⁵ ¹ Helmholtz-Institut für Strahlen- und Kernphysik (HISKP), University of Bonn, Nussallee 14–16, D-53115 Bonn, Germany; mhaslbauer@astro.uni-bonn.de² Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany³ Scottish Universities Physics Alliance, University of Saint Andrews, North Haugh, Saint Andrews, Fife, KY16 9SS, UK; ib45@st-andrews.ac.uk⁴ Charles University in Prague, Faculty of Mathematics and Physics, Astronomical Institute, V Holešovičkách 2, CZ-180 00 Praha 8, Czech Republic⁵ Institute for Genomic Statistics and Bioinformatics, University of Bonn, Venusberg-Campus 1, D-53127 Bonn, Germany

Received 2021 June 23; revised 2021 December 22; accepted 2021 December 24; published 2022 February 4

Abstract

Any viable cosmological framework has to match the observed proportion of early- and late-type galaxies. In this contribution, we focus on the distribution of galaxy morphological types in the standard model of cosmology (Lambda cold dark matter, Λ CDM). Using the latest state-of-the-art cosmological Λ CDM simulations known as Illustris, IllustrisTNG, and EAGLE, we calculate the intrinsic and sky-projected aspect ratio distribution of the stars in subhalos with stellar mass $M_* > 10^{10} M_\odot$ at redshift $z = 0$. There is a significant deficit of intrinsically thin disk galaxies, which however comprise most of the locally observed galaxy population. Consequently, the sky-projected aspect ratio distribution produced by these Λ CDM simulations disagrees with the Galaxy And Mass Assembly (GAMA) survey and Sloan Digital Sky Survey at $\geq 12.52\sigma$ (TNG50-1) and $\geq 14.82\sigma$ (EAGLE50) confidence. The deficit of intrinsically thin galaxies could be due to a much less hierarchical merger-driven build-up of observed galaxies than is given by the Λ CDM framework. It might also arise from the implemented sub-grid models, or from the limited resolution of the above-mentioned hydrodynamical simulations. We estimate that an 8^5 times better mass resolution realization than TNG50-1 would reduce the tension with GAMA to the 5.58σ level. Finally, we show that galaxies with fewer major mergers have a somewhat thinner aspect ratio distribution. Given also the high expected frequency of minor mergers in Λ CDM, the problem may be due to minor mergers. In this case, the angular momentum problem could be alleviated in Milgromian dynamics because of a reduced merger frequency arising from the absence of dynamical friction between extended dark matter halos.

Unified Astronomy Thesaurus concepts: [Galaxies \(573\)](#); [Galaxy properties \(615\)](#); [Galaxy structure \(622\)](#); [Galaxy evolution \(594\)](#); [Galaxy mergers \(608\)](#); [Disk galaxies \(391\)](#); [Elliptical galaxies \(456\)](#); [Cosmology \(343\)](#); [Cold dark matter \(265\)](#); [Modified Newtonian dynamics \(1069\)](#)

1. Introduction

Observed galaxies show a wide spectrum of structural and dynamical properties. According to the morphological classification scheme, early-type galaxies typically have a smooth ellipsoidal shape whereas late-type galaxies have a flattened disk that often contains spiral features. A dynamical characterization divides galaxies into dispersion- and rotation-dominated systems. These classifications are not identical, e.g., most early-type galaxies in the ATLAS^{3D} sample are rotation-supported (Emsellem et al. 2011).

The vast majority of nearby galaxies with stellar mass $M_* > 10^{10} M_\odot$ are of late type (e.g., Kautsch et al. 2006; Delgado-Serrano et al. 2010; Kormendy et al. 2010). In particular, Delgado-Serrano et al. (2010) analyzed local galaxies with an absolute magnitude $J < -20.3$ ($M_* \gtrsim 1.5 \times 10^{10} M_\odot$) from the Sloan Digital Sky Survey (SDSS; SDSS Collaboration 2000) and found that only $3\% \pm 1\%$ of galaxies are elliptical, $15\% \pm 4\%$ are lenticular, $72\% \pm 8\%$ are spiral, and $10\% \pm 3\%$ are peculiar. Interestingly, they also showed that the relative fraction of ellipticals and lenticulars has hardly evolved over the last 6 Gyr (see Table 3 and Figure 5 of Delgado-Serrano et al. 2010), which is consistent with their early and rapid formation (Kroupa et al. 2020, and references therein). Moreover, the relative fraction of

non-spheroidals (70%–80%; i.e., spirals, irregulars, and interacting galaxies) and early types (20%–30%; i.e., ellipticals and transition E/S0 galaxies) with an apparent K_s magnitude brighter than 22 selected from the GOODS-MUSIC catalog (Grazian et al. 2006; Santini et al. 2009) of the Great Observatory Origins Deep Survey (Giavalisco et al. 2004) remains constant over the redshift range $0.6 \leq z \leq 2.5$ (see Figure 8 of Tamburri et al. 2014).

The morphology of a galaxy is closely related to internal physical processes (e.g., rapid monolithic collapse of post-Big Bang gas clouds, star formation, feedback from supernovae and active galactic nuclei), its dynamical history (interactions and mergers with other galaxies), and its environment (e.g., tidal and ram pressure effects). Thus, the observed distribution of galaxy morphological types constrains models of galaxy formation and evolution. Indeed, Disney et al. (2008) emphasized that the observed population of galaxies shows a significantly smaller variation of individual properties than expected in a hierarchical formation model where galaxies undergo mergers stochastically. Several simulations in the standard cosmological model known as Lambda cold dark matter (Λ CDM; Efstathiou et al. 1990; Ostriker & Steinhardt 1995) show an excessive loss of angular momentum (e.g., Katz & Gunn 1991; Navarro & Benz 1991; Navarro & White 1994; Navarro & Steinmetz 2000; van den Bosch 2001; Piontek & Steinmetz 2011; Scannapieco et al. 2012). This hampers the formation of bulgeless disk galaxies. Due to dynamical friction on the extended dark matter halos (Kroupa 2015), galactic mergers are common in Λ CDM. Indeed, N -body simulations yield that $\approx 95\%$ of galaxies with dark matter halo

mass $M_{\text{halo}} \approx 10^{12} h^{-1} M_{\odot}$ accreted at least one galaxy with $M_{\text{halo}} > 5 \times 10^{10} h^{-1} M_{\odot}$ within the last 10 Gyr, where h is the present Hubble constant H_0 in units of $100 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (Stewart et al. 2008). In the Millennium-II simulation (Springel et al. 2005), 69% of galaxies with a similar halo mass have had a major merger since $z = 3$ (Fakhouri et al. 2010). Galaxy mergers thicken the stellar disk and grow the bulge component, making it difficult for Λ CDM to account for the observed large population of bulgeless disk galaxies (Graham & Worley 2008; Kormendy et al. 2010). Trayford et al. (2017) showed that disk galaxies in the Evolution and Assembly of GaLaxies and their Environments (EAGLE) simulation are thicker than those observed (see their Figure 3), but concluded that this is because of the underlying sub-grid physics (we discuss this further in Section 7).

Although the formation of such galaxies is generally known to be a challenge for the Λ CDM paradigm, some works claimed that the angular momentum problem has been resolved in the latest self-consistent Λ CDM simulations. For example, Vogelsberger et al. (2014) argued that the loss of angular momentum was caused by numerical and physical modeling limitations rather than a failure of the Λ CDM paradigm, because the Illustris-1 simulation produces a mix of disk galaxies and ellipticals (but see our Section 6.2, which comes to a different conclusion). Indeed, it is possible to form a Milky Way–like galaxy with a small bulge in the Λ CDM framework, but only under very special conditions of a quiescent merger history and rapid star formation, which removes low angular momentum gas from the inner part of the galaxy (Guedes et al. 2011). However, in the observed universe, such late-type galaxies are frequent, with $\approx 50\%$ of them having no classical merger-built bulge (Kormendy et al. 2010). The same problem was identified by Graham & Worley (2008) two years earlier, who found that most real lenticular galaxies have bulge/total fractions $\lesssim 1/3$. A recent attempt to quantify the tension (Rodríguez-Gomez et al. 2019) showed that the median of the sky-projected ellipticity distribution of galaxies in the “Illustris The Next Generation” simulation (IllustrisTNG; Pillepich et al. 2018; Nelson et al. 2019) lies within the 16th–84th percentile range of the Panoramic Survey Telescope And Rapid Response System (Pan-STARRS) observational sample (Chambers et al. 2016). However, this is only a very crude test. In order to rigorously assess the angular momentum problem in a cosmological framework, one has to consider the overall distribution of galaxy morphology, as addressed by this contribution.

The present-day morphological distribution of galaxies has already been studied in the Illustris-1 simulation using mock photometry. Based on dust-free synthetic images of simulated galaxies with $M_* > 10^{10} M_{\odot}$ at $z = 0$, Bottrell et al. (2017a) derived photometric bulge/total fractions $(B/T)_{\text{phot}}$ by performing a 2D parametric surface brightness decomposition with a fixed Sérsic index of $n_d = 1$ for the disk component and $n_b = 4$ for the bulge. This was done in the SDSS g and r bands at four different camera angles. In a subsequent study (Bottrell et al. 2017b), those authors applied the same decomposition analysis to observed galaxies from the SDSS. By comparing the simulated and observed galaxy samples in the space of M_* and $(B/T)_{\text{phot}}$, they surprisingly found a significant deficit of bulge-dominated subhalos in the Illustris-1 simulation at $10^{10} \leq M_*/M_{\odot} \leq 10^{11}$ (see also their Figures 4 and 6). This would imply that the angular momentum problem has been resolved in Λ CDM cosmology despite mergers being very common. However, we will argue in Section 7.2 that their derived $(B/T)_{\text{phot}}$ is not an appropriate measure to quantify the

morphology of a galaxy in the Illustris-1 simulation. Instead of a deficit of bulge-dominated subhalos, the simulation in fact overproduces these and lacks disk-dominated galaxies, contrary to the claims of Bottrell et al. (2017b).

In this contribution, we statistically compare the observed sky-projected aspect ratio (q_{sky}) distribution with that provided by the Λ CDM framework based on the latest state-of-the-art hydrodynamical cosmological Λ CDM simulations from the projects known as EAGLE (Schaye et al. 2015; McAlpine et al. 2016), Illustris (Vogelsberger et al. 2014; Nelson et al. 2015), and IllustrisTNG (Pillepich et al. 2018; Nelson et al. 2019). Our analysis focuses on the stellar distribution in a galaxy without regards to individual structural components like its thin or thick disk. The main aim of our work is to test whether state-of-the-art Λ CDM simulations form galaxies with a realistic distribution of morphologies.

The layout of this paper is as follows: Section 2 describes the here assessed hydrodynamical cosmological Λ CDM simulations. The methods to calculate the intrinsic and sky-projected aspect ratio of a galaxy are given in Section 3. In Section 4, we introduce the observational galaxy samples from which we extract the aspect ratio distribution. The statistical method to quantify the tension between the simulated and observed galaxy populations is explained in Section 5. The present-day intrinsic and sky-projected aspect ratio distributions yielded by the Λ CDM framework are presented and the latter are compared with observations in Section 6. In Section 7, we test the numerical convergence of the TNG50 and EAGLE runs, seek to understand the mismatch between the photometric parameters (Bottrell et al. 2017a, 2017b) and intrinsic aspect ratios in the Illustris-1 simulation, and investigate the effect of different merger histories on galaxy shapes. We also qualitatively compare the q_{sky} distribution of SDSS spirals with that of disk galaxies formed in hydrodynamical Milgromian dynamics (MOND; Milgrom 1983) simulations. Our conclusions are given in Section 8.

2. Cosmological Λ CDM Simulations

We investigate the aspect ratio distribution provided by different simulation runs of the projects known as EAGLE (Schaye et al. 2015; McAlpine et al. 2016), Illustris (Vogelsberger et al. 2014; Nelson et al. 2015), and TNG (Pillepich et al. 2018; Nelson et al. 2019).⁶ These state-of-the-art simulations self-consistently evolve dark matter and baryons from shortly after the Big Bang up to the present time in a Λ CDM cosmology consistent with the WMAP-9 (Hinshaw et al. 2013), Planck 2015 (Planck Collaboration XIII 2016), and Planck 2013 (Planck Collaboration I 2014) measurements of the cosmic microwave background for Illustris, TNG, and EAGLE, respectively. These simulations differ in the implemented baryonic feedback models and underlying grid solvers. Both the Illustris and TNG simulations were performed with the moving-mesh code AREPO (Springel 2010), whereas the EAGLE simulations employed the GADGET-3 (Springel 2005) smoothed particle hydrodynamics (SPH) code. Here, we use the TNG50-1, TNG100-1, Illustris-1, EAGLE Ref-L050N0752 (hereafter EAGLE50), and EAGLE Ref-L100N1504 (EAGLE100) simulations, with TNG50-1 having the highest resolution. In

⁶ IllustrisTNG (abbreviated as TNG hereafter) is a further development of the Illustris project with an improved galaxy formation and evolutionary model (Nelson et al. 2019).

Table 1
Numerical and Cosmological Parameters of the Here Analyzed Λ CDM Simulations

Simulation	L (cMpc)	N ...	H_0 (km s ⁻¹ Mpc ⁻¹)	$\Omega_{b,0}$...	$\Omega_{m,0}$ —	$\Omega_{\Lambda,0}$...	m_b (M_\odot)	m_{dm} (M_\odot)
EAGLE100	100	2×1504^3	67.77	0.04825	0.307	0.693	1.81×10^6	9.70×10^6
EAGLE50	50	2×752^3	67.77	0.04825	0.307	0.693	1.81×10^6	9.70×10^6
EAGLE25	25	2×752^3	67.77	0.04825	0.307	0.693	2.26×10^5	1.21×10^6
Illustris-1	106.5	2×1820^3	70.4	0.0456	0.2726	0.7274	1.3×10^6	6.3×10^6
TNG50-1	51.7	2×2160^3	67.74	0.0486	0.3089	0.6911	8.5×10^4	4.5×10^5
TNG50-2	51.7	2×1080^3	67.74	0.0486	0.3089	0.6911	6.8×10^5	3.6×10^6
TNG50-3	51.7	2×540^3	67.74	0.0486	0.3089	0.6911	5.4×10^6	2.9×10^7
TNG50-4	51.7	2×270^3	67.74	0.0486	0.3089	0.6911	4.3×10^7	2.3×10^8
TNG100-1	110.7	2×1820^3	67.74	0.0486	0.3089	0.6911	1.4×10^6	7.5×10^6

Note. From left to right: simulation name; co-moving box size (cubic side length); number of dark matter particles plus the initial number of gas cells/particles (hence the factor of two); present-day Hubble constant; present baryonic density in units of the cosmic critical density; same for the total matter density; the dark energy density; the baryonic mass resolution; and the dark matter mass resolution. The highest-resolution realization TNG50-1 has a Plummer-equivalent gravitational softening length for the collisionless component of 288 pc and a minimum adaptive gas gravitational softening length of 72 pc at redshift $z = 0$ (for more information, see Table 1 of Pillepich et al. 2019). EAGLE25 refers to the EAGLE Recal-L0025N0752 simulation. Additional parameters for the EAGLE set can be found in Table 1 of McAlpine et al. (2016), for Illustris-1 in Table 1 of Nelson et al. (2015), and for TNG100-1 in Table 1 of Nelson et al. (2018).

Section 7.1, we also employ the lower-resolution realizations TNG50-2, TNG50-3, and TNG50-4 of the TNG50 sets and the higher-resolution realization EAGLE Recal-L0025N0752 (hereafter EAGLE25) in order to study the effect of resolution on the shapes of simulated galaxies. The numerical and cosmological parameters of these simulations are listed in Table 1.⁷

3. Quantifying the Shape of a Galaxy

The shape of a galaxy can be quantified by the 3D intrinsic aspect ratio of its mass distribution as defined by $q_{\text{int}} \equiv \lambda_1 / \sqrt{\lambda_2 \lambda_3}$, where λ_1 , λ_2 , and λ_3 (sorted so $\lambda_1 < \lambda_2 < \lambda_3$) are the square roots of the eigenvalues of the mass distribution tensor (MDT, sometimes also called the moment of inertia tensor) divided by the total mass (see, e.g., Equation D.39 in Appendix D of Binney & Tremaine 2008). In this contribution, we consider only the stellar MDT because we are interested in the appearance of a galaxy in optical images. This analysis does not distinguish different structural components of the galaxy like its thin or thick disk. A completely intrinsically thin disk galaxy has $q_{\text{int}} = 0$, whereas a perfectly spherical galaxy has $q_{\text{int}} = 1$. Spiral galaxies account for the bulk of galaxies in the locally observed universe (Delgado-Serrano et al. 2010), with typical $q_{\text{int}} \approx 0.2$ (see, e.g., Figure 1 of Mosenkov et al. 2010 and Figure 1 of Hoffmann et al. 2020). About 57% (79%) of all galaxies in the the Sydney-AAO Multi-object Integral field spectrograph (SAMI) Galaxy Survey (Croom et al. 2012; Bryant et al. 2015) have $q_{\text{int}} < 0.4$ (< 0.6) (see Figure 15 of Oh et al. 2020). The Galactic thin disk has an exponential scale length of $l = 2.6 \pm 0.5$ kpc and an exponential scale height of $h = 220\text{--}450$ pc (Bland-Hawthorn & Gerhard 2016), which results in an aspect ratio of $h/l \approx 0.07\text{--}0.21$. The Andromeda galaxy (M31) has a thin disk with $1 - \epsilon \equiv b/a \equiv q_{\text{sky}} = 0.27 \pm 0.03$, where ϵ and b/a are the sky-projected ellipticity and axis ratio, respectively (Courteau et al. 2011). Because M31 is not viewed exactly edge-on, it must be intrinsically thinner than it appears on the sky (Section 2.1 of Banik & Zhao 2018).

⁷ In the Illustris and TNG simulations, the initial speed of a baryonic wind particle has an unphysical dependence on the local 1D dark matter velocity dispersion (see Equation (1) in Pillepich et al. 2018).

We extract the eigenvalues of the stellar MDT from supplementary data catalogs provided by the Illustris, TNG, and EAGLE teams (Thob et al. 2019).⁸ The MDTs of subhalos in the Illustris and TNG simulations are calculated from the stellar particles within twice the stellar half-mass-radius $r_{0.5,*}$ (Genel et al. 2015). This slightly differs from the EAGLE simulations in which an iterative form of the reduced MDT is used, with the initial selection being all stellar particles inside a spherical aperture of physical radius 30 kpc (Section 2.3 of Thob et al. 2019). As demonstrated in Section 7.5, this method provides a secure division into spirals and ellipticals, whereas other bulge-disk decomposition methods face problems when applied to Λ CDM simulations (Section 7.2).

3.1. Projecting an Ellipsoid onto the Sky

In order to compare simulations with observations, we have to determine q_{sky} for a galaxy with λ_i , where $i = 1\text{--}3$ (Section 3). Only the ratios of the λ_i are relevant for our analysis, but it will be helpful to think of them as actual lengths. We approximate that the galaxy is an ellipsoid and find what this looks like when viewed by a distant observer. We work in Cartesian coordinates in a reference frame centered on the galaxy and aligned with the eigenvectors of its inertia tensor. Thus, the “edge” of the galaxy is given by

$$\frac{x^2}{\lambda_1^2} + \frac{y^2}{\lambda_2^2} + \frac{z^2}{\lambda_3^2} = 1. \quad (1)$$

Suppose a distant observer is located toward the direction \hat{n} , where we use hats to denote unit vectors. Our approach is to find the extent of the image along the direction \hat{n}_3 , which lies entirely within the sky plane (i.e., $\hat{n} \cdot \hat{n}_3 = 0$). Our main goal is to find the position vector \mathbf{r} of the point corresponding to the edge of the galaxy image along the direction \hat{n}_3 . For this purpose, it will be useful to define the unit vector within the sky plane in the orthogonal direction, which we call \hat{n}_2 (orthogonal to both \hat{n} and \hat{n}_3).

We know that $\mathbf{r} \cdot \hat{n}_2 = 0$, or else the point would not appear to be in the direction \hat{n}_3 from the galaxy’s center. To get an

⁸ The eigenvalues of the stellar MDT for Illustris and TNG subhalos can be downloaded from <https://www.tng-project.org/data/docs/specifications/#sec5c> [21.07.2020].

additional constraint, we note that the tangent plane to the galaxy at the point \mathbf{r} must not contain $\hat{\mathbf{n}}_3$, or else it would be possible to move along the galaxy’s boundary and reach a larger apparent separation along $\hat{\mathbf{n}}_3$. The plane normal must therefore be $\pm\hat{\mathbf{n}}_3$, but for our analysis, it is sufficient to know that the plane normal is orthogonal to $\hat{\mathbf{n}}$.

These two constraints are sufficient to determine the direction of \mathbf{r} . Its magnitude is found through Equation (1). Thus, the extent of the image along $\hat{\mathbf{n}}_3$ is d , which we find through the following procedure involving the intermediate vectors \mathbf{q} and \mathbf{v} :

$$\mathbf{q}_i \equiv \frac{\hat{\mathbf{n}}_i}{\lambda_i^2}, \quad (2)$$

$$\mathbf{v} \equiv \hat{\mathbf{n}}_2 \times \mathbf{q}, \quad (3)$$

$$d = \frac{\mathbf{v} \cdot \hat{\mathbf{n}}_3}{\sqrt{\sum_{i=1}^3 \left(\frac{v_i}{\lambda_i}\right)^2}}. \quad (4)$$

We repeat this for a low-resolution grid of $\hat{\mathbf{n}}_3$, whose direction we parameterize using the so-called position angle. Starting from the $\hat{\mathbf{n}}_3$ in this grid which gives the lowest d , we apply the gradient descent algorithm (Fletcher & Powell 1963) to find the minimum value of d to high precision. We then rotate $\hat{\mathbf{n}}_3$ through a right angle, and start a gradient ascent stage to search for the maximum extent of the image. The ratio of these d values is the sky-projected aspect ratio $q_{\text{sky}} \leq 1$, which forms the heart of our analysis.

To build up the q_{sky} distribution, we repeat this procedure for a 2D grid of viewing angles $\hat{\mathbf{n}}$, with each result weighted according to the solid angle it represents. The observer is thus assumed to be in a random direction relative to the galaxy.

4. Observational Galaxy Samples

The sky-projected aspect ratio distributions of the following observational galaxy samples are statistically compared with the simulations discussed in Section 2.

4.1. GAMA Survey

The Galaxy And Mass Assembly survey (GAMA; Driver et al. 2009, 2011) is a multiwavelength photometric and spectroscopic survey. An overview of the survey regions and their corresponding magnitude limits is given in Table 1 of Baldry et al. (2018). Here, we download the stellar masses, redshifts, and ellipticities by submitting an SQL query to the GAMA DR3 database (Baldry et al. 2018).⁹ In detail, we extract the GALFIT (Peng et al. 2002, 2010) r -band ellipticity $(1 - b/a)$ from the SérsicPhotometry (v09) - SérsicCatSDSS catalog (Kelvin et al. 2012). We use the stellar mass labeled as “logmstar” from the StellarMasses (v20) catalog (Taylor et al. 2011), which is derived from matched aperture photometry in the r band, missing therewith flux beyond the AUTO aperture.¹⁰ Consequently, the stellar masses $M_{*,\text{AUTO}}$ within the photometric aperture (“logmstar”) are corrected by

$$\log_{10}\left(\frac{M_*}{M_\odot}\right) = \log_{10}\left(\frac{M_{*,\text{AUTO}}}{M_\odot}\right) + \log_{10}\left(\frac{f_{\text{Sérsic}}}{f_{\text{AUTO}}}\right),$$

where $f_{\text{Sérsic}}/f_{\text{AUTO}}$ is the so-called “fluxscale” parameter, i.e., the linear ratio between the total r -band flux from a Sérsic profile fit cut at 10 effective radii and the r -band AUTO aperture flux (see also, e.g., Taylor et al. 2011; Kelvin et al. 2014; Lange et al. 2016; Vázquez-Mata et al. 2020). The stellar masses are calculated by assuming concordance cosmology with the cosmological parameters being $\Omega_{\text{m},0} = 0.3$, $\Omega_{\Lambda,0} = 0.7$, and $h=0.7$ (Taylor et al. 2011).

For our analysis, we only consider galaxies with a fluxscale correction of $0.5 < \text{fluxscale} < 2$ and a spectral energy distribution (SED) fit with a posterior predictive P -value $PPP > 0$, which removes failed SED fits. In addition, we exclude objects with a heliocentric redshift $z < 0.005$ to remove stars.¹¹ Also requiring $M_* > 10^{10} M_\odot$ and $z < 0.1$ yields a final sample of 5304 galaxies that pass the above quality cuts.

4.2. GAMA and Cluster Input Catalogs for the SAMI Galaxy Survey

The SAMI Galaxy Survey Data Release 3 (DR3) includes 3068 galaxies in the redshift range $0.004 < z < 0.095$. The ellipticity distribution of a subsample of the SAMI Galaxy Survey consisting of 826 galaxies is shown in the third panel of Figure 6 in Oh et al. (2020).

In order to increase the sample size, we analyze the aspect ratio distribution of galaxies listed in the input catalogs of the three equatorial regions of the GAMA survey (i.e., G09, G12, and G15; Bryant et al. 2015) and the eight cluster regions (i.e., APMCC0917, A168, A4038, EDCC442, A3880, A2399, A119, and A85; Owers et al. 2017) for the SAMI DR3 Galaxy Survey. For this, we download the stellar masses, redshifts, and projected ellipticities by using an SQL/Astronomical Data Query Language (ADQL) query to the Data Central database servers¹² of the InputCatGAMADR3 and InputCatClustersDR3 catalogs (for more details, see also Table 1 of Croom et al. 2021). The ellipticities are derived from Sérsic fits in the r band (Bryant et al. 2015; Owers et al. 2019).

Requiring $M_* > 10^{10} M_\odot$ gives a sample of 4252 galaxies, of which 3238 are from GAMA and 1014 are from the cluster regions. As we show in Section 5, the aspect ratio distributions of the GAMA survey and the here described sample disagree only at the 0.013σ confidence level. Although the ellipticity values of the same galaxy listed in the InputCatGAMADR3 and the GAMA survey are slightly different, these catalogs are not fully independent of each other. Because of these reasons and since the GAMA survey contains more galaxies (5304), we do not use the input catalogs of the GAMA and cluster regions for SAMI DR3 in our statistical comparison with simulations (Section 5).

4.3. SDSS

The SDSS is a flux-limited galaxy survey including galaxies with a Petrosian r magnitude brighter than 17.77 (SDSS Collaboration 2000). Here, we download the stellar masses, redshifts, and photometric parameters from its DR16¹³ (Ahumada et al. 2020) using an SQL query to the SDSS database server.¹⁴

⁹ <http://www.gama-survey.org/dr3/query/> [11.11.2021].

¹⁰ For further information on the stellar masses, see also <http://www.gama-survey.org/dr3/schema/dmu.php?id=9> [11.11.2021].

¹¹ Private communication with Edward Taylor.

¹² <https://datacentral.org.au/services/query/> [04.11.2021].

¹³ <https://www.sdss.org/dr16/>

¹⁴ <http://skyserver.sdss.org/dr16/en/home.aspx> [11.11.2021].

We selected galaxies listed in the PhotoPrimary catalog, which also contains the aspect ratios of galaxies derived from an exponential and a de Vaucouleurs profile (de Vaucouleurs 1948) for the SDSS r , i , u , z , and g filters. Throughout this analysis, we use the aspect ratio parameters based on the SDSS r -band magnitude. In order to distinguish between spiral and elliptical galaxies, we access the fracDeV parameter, the weight of the de Vaucouleurs component in a linear combination of the exponential and de Vaucouleurs profiles (for a more detailed description, see Section 3.1 of Abazajian et al. 2004). Following Padilla & Strauss (2008), we use the aspect ratio derived from the exponential fit if fracDeV < 0.8 and from the de Vaucouleurs fit if fracDeV \geq 0.8.

The here used stellar masses from the galSpecExtra table correspond to the median of the estimated logarithmic stellar mass probability density function using model photometry (Kauffmann et al. 2003; Salim et al. 2007). The redshifts are extracted from the SpecObj table.¹⁵

Selecting galaxies with $z < 0.1$ and $M_* > 10^{10} M_\odot$ gives a sample of 232,315 galaxies. The so-obtained and here analyzed q_{sky} distribution is consistent with Padilla & Strauss (2008), as shown in Section 5. They obtained a volume-limited sample by weighting each galaxy in the SDSS DR6 (Adelman-McCarthy et al. 2008) by $1/V_{\text{max}}$, where V_{max} is the volume corresponding to the maximum distance at which we can observe a galaxy with its absolute magnitude (see Section 3.1 of Padilla & Strauss 2008). In total, their sample contains 303,290 “spirals” (fracDeV < 0.8) and 282,203 “ellipticals” (fracDeV \geq 0.8). Here, we extract the $1/V_{\text{max}}$ weighted q_{sky} distributions of the spiral and elliptical samples from their Figure 1 (bottom panels), and combine these in order to obtain the q_{sky} distribution of the total galaxy sample.

4.4. The Catalog of Neighboring Galaxies

As a consistency check, we also investigate the q_{sky} distribution of galaxies within the Local Volume (LV), a sphere of radius 11 Mpc centered on the Sun. The stellar masses are calculated from K -band luminosities using a mass-to-light ratio of 0.6 (e.g., McGaugh & Schombert 2014). We select 52 galaxies with $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$ from the Updated Nearby Galaxy Catalog (Karachentsev et al. 2013), a renewed version of The Catalog of Neighboring Galaxies¹⁶ (Karachentsev et al. 2004). These galaxies cover a wide range of q_{sky} (0.13–0.94), as shown in the right panel of Figure 1. Due to the small sample size and high resulting Poisson uncertainties, we do not calculate its tension with the simulated q_{sky} distributions.

These are nearby galaxies with well-established morphological types, allowing direct well-resolved images to inform us of which galaxies are thin disks. Therefore, we show the distribution of the morphological T -types of our LV galaxy sample in the left panel of Figure 1. The T -types of these galaxies are listed in the Updated Nearby Galaxy Catalog and discussed and analyzed in detail in Karachentsev et al. (2018). Of the 52 galaxies, 50% are early-type spirals (Sa, Sab, Sb,

Sbc, Sc; morphological type $T = 1$ –5), 23% are late-type spirals (Scd, Sd, Sdm, Sm; $T = 6$ –8), and the remaining 27% are E, S0, dSph, or S0a galaxies ($T < 1$). Thus, the morphological classification scheme used by Karachentsev et al. (2013, 2018) is slightly different to the de Vaucouleurs system.

It is not in general possible to disentangle the contributions to q_{sky} from inclination and intrinsic thickness. The inclination i between disk and sky planes listed in the Updated Nearby Galaxy Catalog is derived with Equation (5) of Karachentsev et al. (2013):

$$\sin^2 i = \frac{1 - q_{\text{sky}}^2}{1 - \tilde{q}_{\text{int}}^2}, \quad (5)$$

where the assumed intrinsic aspect ratio \tilde{q}_{int} is assigned a value depending on the morphological T -type (T) of the considered galaxy as given by their Equation (6):¹⁷

$$\log_{10} \tilde{q}_{\text{int}} = \begin{cases} -0.43 - 0.053T & (T \leq 8), \\ -0.38 & (T = 9, 10), \end{cases} \quad (6)$$

where according to Karachentsev et al. (2018) $T = 9$ and $T = 10$ refer to Im/BCD and Ir galaxies, respectively. Thus, the so-obtained intrinsic aspect ratios are a best guess based on the observed morphology. As an example, the M31 galaxy with $T = 3$ has with this approach $\tilde{q}_{\text{int}} = 0.26$ ($q_{\text{sky}} = 0.33$).

Applying Equation (6) to the 52 selected LV galaxies shows that their intrinsic aspect ratio distribution covers the range $\tilde{q}_{\text{int}} = 0.07$ –0.55 with a global peak at $\tilde{q}_{\text{int}} \approx 0.23$, as presented in the right panel of Figure 1. Converting q_{sky} to q_{int} depends on the relation between \tilde{q}_{int} and the morphological T -type, so the intrinsic aspect ratio distribution of our LV sample is mainly considered for illustrative purposes. However, it helps to demonstrate that the LV galaxies seem to be mostly thin disks. According to this, if the LV were to be representative of the universe, then $81\% \pm 13\%$ (42/52) of all galaxies heavier than $M_* = 10^{10} M_\odot$ are thin disk galaxies with $\tilde{q}_{\text{int}} < 0.4$. For completeness, we quantify the tension between the here obtained intrinsic aspect ratio distribution of LV galaxies and the Λ CDM models in Section 6.1.

In order to avoid any bias on the shapes of galaxies introduced by model assumptions when converting q_{sky} to q_{int} , we only consider the former when statistically comparing observations to simulations in our main analysis. As explained in Section 3.1, this requires us to generate the q_{sky} distribution given the intrinsic shapes of simulated galaxies. We can then directly compare the resulting distribution with the observed q_{sky} distribution (Section 6.2).

5. Quantifying the Tension between Simulations and Observations

The stellar mass distributions of the GAMA survey and SDSS are compared with that of the TNG50-1 run in the left panel of Figure 2. Galaxies with $10.30 \lesssim \log_{10}(M_*/M_\odot) \lesssim 11.05$ are more abundant in the

¹⁵ Although the SpecObj catalog does not contain duplicate observations, there is the rare situation that the catalog lists more than one redshift measurement for the same object. In this case, the R.A. and decl. values of the fibers are different during the spectroscopic observation. We take this into account by checking if our downloaded catalog lists more than one redshift measurement for the same object ID. If so, we remove the measurement with the lower signal-to-noise ratio. Of our 232,315 selected galaxies, a duplicate happened only once for the object with ID 1237663782590021834.

¹⁶ <https://www.sao.ru/lv/lvqdb/tables.php> [22.11.2021].

¹⁷ There seems to be a typo in Equation (6) of Karachentsev et al. (2013). The relation between the assumed intrinsic aspect ratio and the T -type parameter for galaxies with $T \leq 8$ should be $\log_{10}(a/b)_0 = -\log_{10} \tilde{q}_{\text{int}} = 0.43 - 0.053 T$ (see also Equation (14) of Paturel et al. 1997) rather than $0.43 + 0.53 T$, which would make late-type galaxies far too thin.

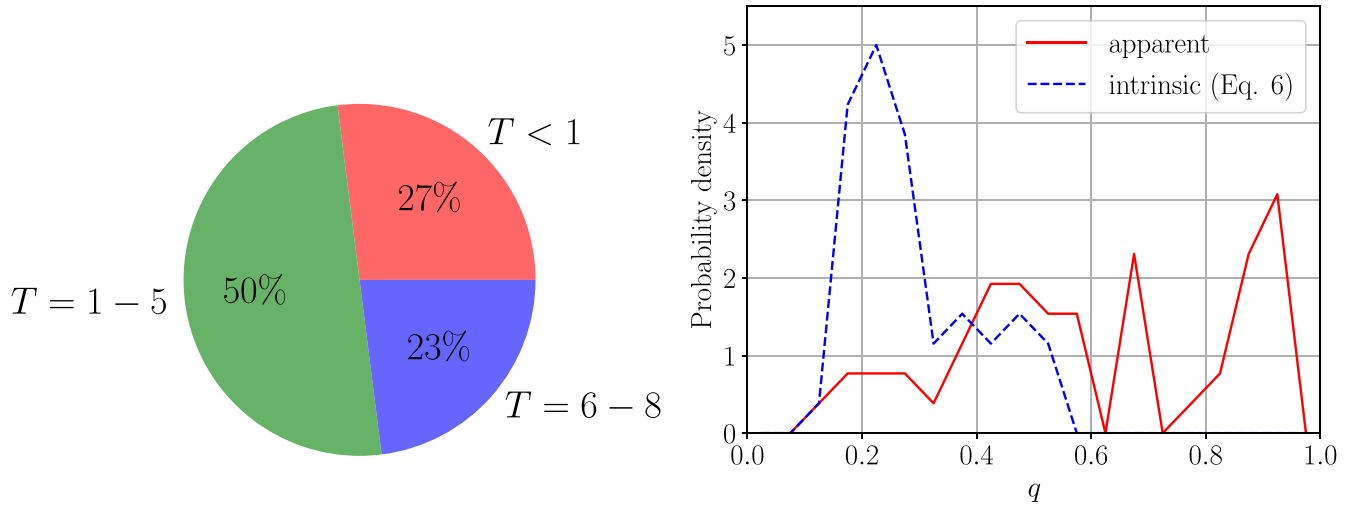


Figure 1. Left: distribution of the morphological T -types of galaxies with $M_* > 10^{10} M_\odot$ in the Local Volume (LV). The red slice refers to E, S0, dSph, and S0a galaxies ($T < 1$), the green slice to early-type spirals (Sa, Sab, Sb, Sbc, Sc; $T = 1-5$), and the blue slice to late-type spirals (Scd, Sd, Sdm, Sm; $T = 6-8$). Right: the sky-projected (solid red) and intrinsic (dashed blue) aspect ratio distribution of galaxies with $M_* > 10^{10} M_\odot$ in the LV. The bin width is $\Delta q_{\text{sky}} = \Delta \tilde{q}_{\text{int}} = 0.05$. The intrinsic aspect ratios are derived using Equation (6), which is a best guess based on the morphological T -type (see the text). It is statistically compared with the TNG50-1 simulation run in Section 6.1, but the model dependence of the \tilde{q}_{int} values means this is not our main result.

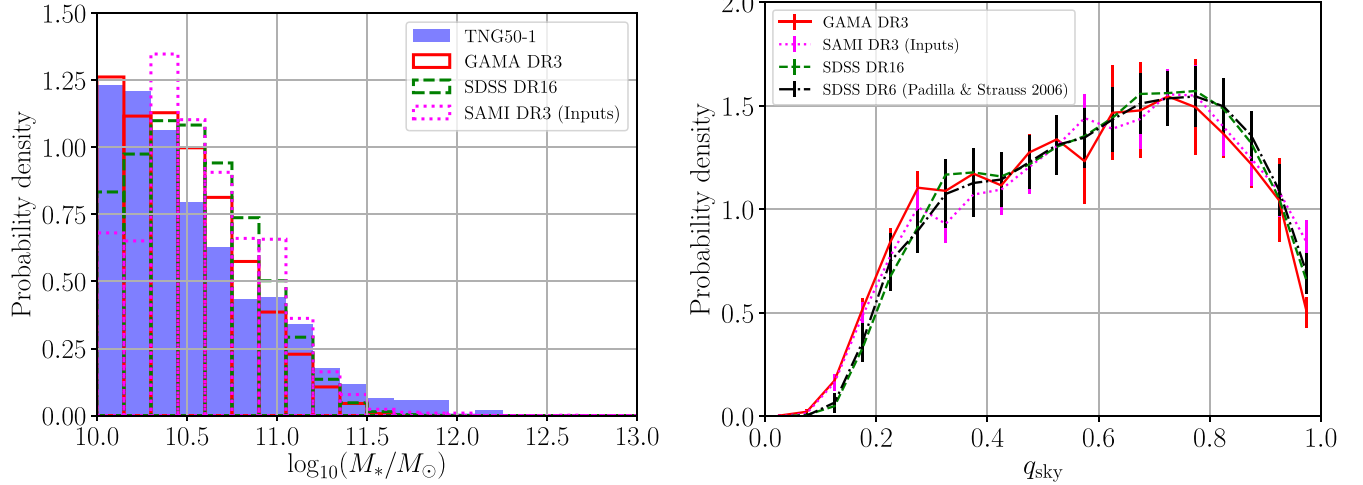


Figure 2. Left: stellar mass distribution of the TNG50-1 simulation run (filled blue) and on the observational side GAMA DR3 (open solid red, Section 4.1), the input catalogs of the GAMA and cluster regions for SAMI DR3 (open dotted magenta, Section 4.2), and SDSS DR16 (open dashed green, Section 4.3). We use a bin width of $\Delta \log_{10}(M_*/M_\odot) = 0.15$. Right: sky-projected aspect ratio distribution of GAMA DR3 (solid red), the input catalogs of the GAMA and cluster regions for the SAMI DR3 (dotted magenta), and SDSS DR16 (dashed green) after weighting the galaxies in order to match the stellar mass distribution of the TNG50-1 run. Invisible error bars (Equation (8)) are smaller than the data point symbol. The black dotted–dashed line refers to the volume-weighted SDSS DR6 sample (Padilla & Strauss 2008) with $N_{\text{spirals}} = 303,290$ and $N_{\text{ellipticals}} = 282,203$ (see the bottom panels of their Figure 1). The error bars of the spiral and elliptical aspect ratio distributions have been obtained from the jackknife technique and were here weighted by the above-mentioned proportion of spirals and ellipticals. The bin width is $\Delta q_{\text{sky}} = 0.05$.

observational samples than in the simulation, while the opposite is true at higher mass. This is most likely because the used stellar masses in the Subfind catalog of the TNG50-1 simulation refer to the mass of all stellar particles bound to each considered subhalo, not to the stellar mass within a certain aperture size.

The different M_* distributions would bias the comparison between the observed and simulated q_{sky} distributions. Thus, we apply a M_* -weighting scheme to each observed galaxy in order to match the simulated M_* distribution, as explained in the following. The total χ^2 between the simulated and observed sky-projected aspect ratio distributions is

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \frac{\left(\frac{w_{\text{tot},i}}{w_{\text{tot}}} - \frac{N_{\text{model},i}}{N_{\text{model,tot}}} \right)^2}{\sigma_{\text{obs},i}^2 + \sigma_{\text{model},i}^2}, \quad (7)$$

$$\sigma_{\text{obs},i} = \begin{cases} \frac{w_{\text{max},i}}{w_{\text{tot}}} \sqrt{\frac{w_{\text{tot},i}}{w_{\text{max},i}} + 1} & (w_{\text{obs},i} \neq 0), \\ \frac{w_{\text{max}}}{w_{\text{tot}}} & (w_{\text{obs},i} = 0), \end{cases} \quad (8)$$

$$\sigma_{\text{model},i} = \frac{\sqrt{N_{\text{model},i} + 1}}{N_{\text{model,tot}}}, \quad (9)$$

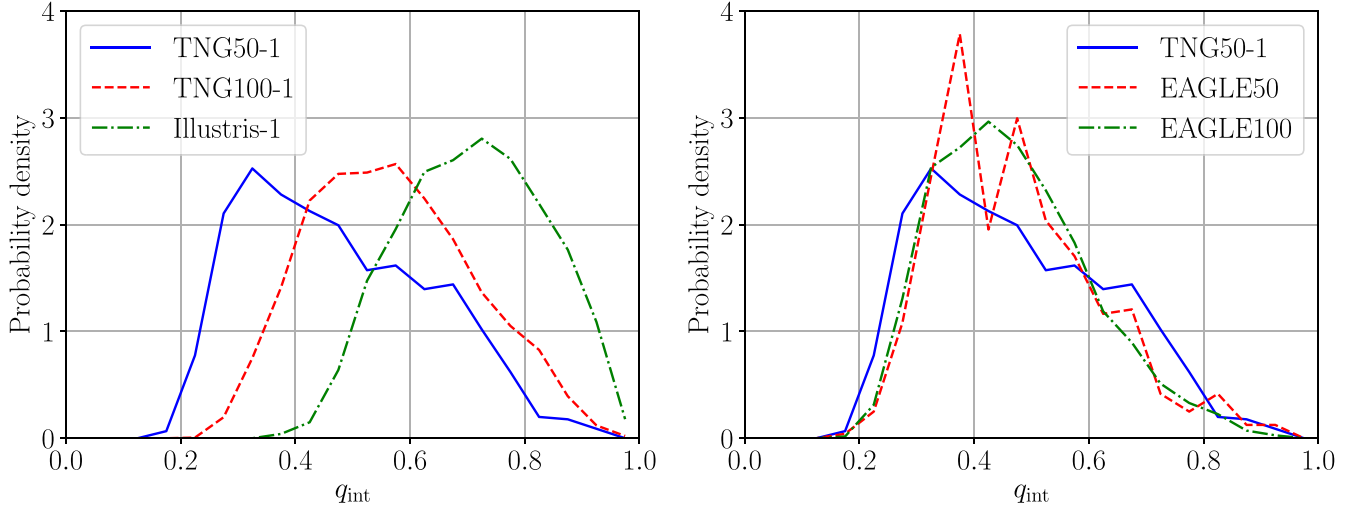


Figure 3. Distribution of the intrinsic aspect ratio q_{int} for galaxies with $M_* > 10^{10} M_{\odot}$ in the Illustris and TNG (left) and EAGLE (right) simulations, with the TNG50-1 result shown in both panels as a solid blue line for clarity. The bin width is $\Delta q_{\text{int}} = 0.05$ throughout this work.

where $N_{\text{model},i}$ is the number of simulated galaxies in bin i , $N_{\text{bins}} = 20$ is the number of bins in q_{sky} , and $N_{\text{model,tot}}$ is the total number of simulated galaxies. In order to avoid the bias caused by the different M_* distributions of the simulations and observations, we group simulated and observed galaxies in M_* bins of width 0.15 dex over the range $10.0 < \log_{10}(M_*/M_{\odot}) \leq 11.65$. Each observed galaxy is weighted by $w_{\text{obs}} = N_{\text{sim}}/N_{\text{obs}}$, where N_{sim} (N_{obs}) is the number of simulated (observed) galaxies in the M_* bin of the considered galaxy. The lower limit on M_* is set to guarantee that only well-resolved simulated galaxies are analyzed. The maximum limit is applied because the GAMA DR3 sample runs out of galaxies at higher stellar mass, leading to an undefined w_{obs} . These criteria give final GAMA, SAMI DR3 inputs, and SDSS sample sizes of 5304, 4229, and 232,128, respectively. $w_{\text{tot},i}$ is then the weighted number of galaxies in q_{sky} bin i , while $w_{\text{max},i}$ is the weight of the galaxy in this q_{sky} bin with the maximum weight, w_{max} is the maximum weight of all considered observed galaxies (regardless of q_{sky}), and w_{tot} is the total weight of all observed galaxies, which by definition must match the number of simulated galaxies used for the comparison. This approach is invariant to a uniform scaling of the weights. $\sigma_{\text{obs},i}$ and $\sigma_{\text{model},i}$ are Poisson uncertainties. The use of Poisson statistics is valid because we choose a bin width of $\Delta q_{\text{sky}} = 0.05$, so each bin contains only a small fraction of the full sample.

This M_* -weighting scheme is not applied when the aspect ratio distributions are being compared between simulations. The total χ^2 between any two models is

$$\chi^2 = \sum_{i=1}^{N_{\text{bins}}} \frac{\left(\frac{N_{\text{model } 1,i}}{N_{\text{model } 1,\text{tot}}} - \frac{N_{\text{model } 2,i}}{N_{\text{model } 2,\text{tot}}} \right)^2}{\sigma_{\text{model } 1,i}^2 + \sigma_{\text{model } 2,i}^2}. \quad (10)$$

The uncertainties $\sigma_{\text{model } 1}$ and $\sigma_{\text{model } 2}$ can be thought of as given by Equation (8) with all galaxies equally weighted.

The total χ^2 between any model and observations (Equation (7)) or between two models (Equation (10)) is converted to a probability or P -value of a more extreme outlier using the χ^2 distribution for the appropriate number of degrees of freedom, which we call n . For convenience, we then express

the P -value as an equivalent number of standard deviations for a single Gaussian variable. We denote this x , which we find by iteratively solving

$$1 - \frac{1}{\sqrt{2\pi}} \int_{-x}^x \exp\left(-\frac{x^2}{2}\right) dx \equiv P. \quad (11)$$

A protocol to convert particularly large χ^2 values is given in the Appendix, which provides a way to solve this equation despite numerical difficulties that arise for very high χ^2 .

The right panel of Figure 2 shows the q_{sky} distributions of the GAMA Galaxy Survey and SDSS, weighted based on the TNG50-1 run for galaxies with $10.0 < \log_{10}(M_*/M_{\odot}) \leq 11.65$. The GAMA DR3 and SDSS DR16 results disagree at 3.25 σ confidence, indicating good agreement considering the large sample sizes. The aspect ratio distributions of GAMA DR3 and the input catalogs of the GAMA and cluster regions for the SAMI DR3 differ only at 0.013 σ confidence. Because of this and the larger sample size of the GAMA survey, we only use the GAMA survey and SDSS for the following analysis.

6. Results

In this section, we present both the intrinsic and the sky-projected aspect ratio distribution at $z = 0$ as produced by the Λ CDM paradigm. We then statistically compare the simulated q_{sky} distribution with local observations from the GAMA survey and SDSS.

6.1. Intrinsic Aspect Ratio Distribution

Figure 3 shows the q_{int} distribution of galaxies in central and noncentral subhalos with $M_* > 10^{10} M_{\odot}$ in the TNG50-1, TNG100-1, Illustris-1, EAGLE50, and EAGLE100 simulations. These all show a unimodal distribution with different peak positions. We show later that increasing the resolution broadens the peak and shifts it to smaller q_{int} (Section 7.1). Galaxies in the EAGLE and TNG50-1 runs are typically thinner than in the Illustris-1 and TNG100-1 runs. The TNG50-1 run contains 133 galaxies with $q_{\text{int}} < 0.3$ and $M_* > 10^{10} M_{\odot}$, with the thinnest galaxy of this sample having $q_{\text{int}} \approx 0.19$. Clearly, the resolution and the adopted temperature floor of about 10^4 K (10 kK; see Trayford et al. 2017 for EAGLE and

Nelson et al. 2019 for TNG) allow for the formation of massive thin disk galaxies. This is consistent with the work of Sellwood et al. (2019), who managed to maintain a thin disk in a hydrodynamical CDM-based simulation of M33 with a slightly higher temperature floor of 12 kK.

The EAGLE runs have a very similar q_{int} distribution—the peak position is similar to TNG50, and EAGLE forms galaxies as thin as $q_{\text{int}} = 0.18$ (the thinnest galaxies in each run are $q_{\text{int}} = 0.20$ for EAGLE25, $q_{\text{int}} = 0.19$ for EAGLE50, and $q_{\text{int}} = 0.18$ for EAGLE100). This is in agreement with TNG50-1 (right panel of Figure 3). By using different resolution realizations of the TNG50 and EAGLE simulations, we test the numerical convergence of the aspect ratio distribution in Section 7.1. While there is no evidence that the TNG50 simulation has numerically converged, the aspect ratio distributions of different EAGLE runs closely agree with each other, possibly because they have the same resolution. The similarity between EAGLE runs is consistent with Lagos et al. (2018), who showed that higher-resolution realizations of the EAGLE project yield galaxies with a similar ellipticity distribution (see their Figures A2 and A3). Moreover, the q_{int} distributions of EAGLE and TNG50-1 are very similar despite differences in resolution and other details of the models.

In the LV, $81\% \pm 13\%$ of galaxies with $M_* > 10^{10} M_{\odot}$ have $\tilde{q}_{\text{int}} < 0.4$ (Section 4.4). In contrast, the fractions of galaxies with such masses that have $q_{\text{int}} < 0.4$ in the highest-resolution simulations analyzed here (TNG50-1, EAGLE25) are only $39\% \pm 2\%$ and $46\% \pm 8\%$, respectively. The stated Poisson uncertainties are estimated as $\sigma = \sqrt{N_{<0.4} + 1}/N_{\text{tot}}$, where $N_{<0.4}$ and N_{tot} are the number of galaxies with $q_{\text{int}} < 0.4$ and the total number of galaxies, respectively. The overall intrinsic aspect ratio distribution of the LV galaxies (blue dashed line in the left panel of Figure 1) is in 5.42σ tension with the TNG50-1 run. We emphasize again that the intrinsic aspect ratios of LV galaxies are based on the morphological T -type (Equation (6)). Therefore, we provide in the following section another test of the Λ CDM framework in which we compare the sky-projected aspect ratio distributions of the Λ CDM simulations with the GAMA survey and SDSS.

6.2. Sky-projected Aspect Ratio Distribution and Comparison with Observations

In order to compare these simulation results with observations from the GAMA survey and SDSS (Section 4), the simulated galaxies are projected onto the sky from a grid of viewing angles to find the distribution of q_{sky} (Section 3.1). The resulting q_{sky} distribution of each simulation is compared with the GAMA survey and SDSS in Figure 4. This reveals that the simulations significantly underproduce the fraction of galaxies with a low q_{sky} . In particular, only about 10% of simulated galaxies with $10.0 < \log_{10}(M_*/M_{\odot}) \leq 11.65$ have $q_{\text{sky}} < 0.4$, while these galaxies make up about 22%–25% of locally observed galaxies (Table 2).

The tension between the simulated and observed q_{sky} distributions is quantified using a standard χ^2 statistic (Section 5). All simulations significantly disagree with local observations. The smallest tension is 12.52σ , which arises from comparing the GAMA survey with TNG50-1. The smallest tension for the SDSS is 14.82σ , corresponding to a comparison with EAGLE50 (we do not consider results from the EAGLE25 run as it only has 81 galaxies within the analyzed M_* range, but see Section 7.1). The total χ^2 values and the corresponding levels of tension are listed in Table 3 for different simulations.

7. Discussion

Although state-of-the-art cosmological Λ CDM simulations produce a variety of galaxy types (Vogelsberger et al. 2014; Schaye et al. 2015), we showed that the overall morphological distribution produced by the Λ CDM framework significantly disagrees with local observations. The q_{sky} distribution is similar between different observational samples, with the GAMA survey and SDSS being in 3.25σ tension with each other despite the large sample size of both (Figure 2). Here the important result has been documented that the different Λ CDM simulations disagree with the observed galaxy population. Galaxies formed in cosmological Λ CDM simulations are typically intrinsically thick rather than intrinsically thin (Section 6.2), making it challenging to explain the observed common formation of galaxies like the Milky Way or M31. This contrasts with the conclusion of Vogelsberger et al. (2014) based on the Illustris-1 simulation, who claimed that the angular momentum problem of Λ CDM has been resolved. Although the simulations can produce *some* disk galaxies, the fraction of such galaxies is far too small compared with observations (Table 2). The lack of such thin disk galaxies in simulations underlies the significant discrepancy between the observed and simulated distributions of galaxy shapes.

Our results broadly agree with those of van de Sande et al. (2019), who showed that the intrinsic (see their Figure 10) and sky-projected aspect ratio distributions (their Figures 4 and 8) of the EAGLE (Crain et al. 2015; Schaye et al. 2015; McAlpine et al. 2016), HYDRANGEA (Bahé et al. 2017; Barnes et al. 2017), HORIZON-AGN (Dubois et al. 2014), and MAGNETICUM (Hirschmann et al. 2014; Dolag et al. 2016, 2017) simulations all disagree with observational data from the ATLAS^{3D}, Calar Alto Legacy Integral Field Area Survey (CALIFA), and MASSIVE surveys, and also with the SAMI Galaxy Survey (MAGNETICUM produced too few very round galaxies). Similar results were obtained by Peebles (2020), who recently showed that stellar particles in Λ CDM subhalos have kinematics different to stars in local galaxies. However, the very small sample size made a statistical comparison difficult.

7.1. The Effect of Numerical Resolution

The fact that zoomed-in Λ CDM simulations can produce thin disk galaxies (e.g., Wetzel et al. 2016) suggests that further increasing the numerical resolution may solve the discrepancy reported in Section 6.2. For example, Ludlow et al. (2021) recently argued that the coarse-grained implementation of dark matter halos causes an artificial heating of the stellar particles, which increases the vertical velocity dispersion of the galaxies. This yields thicker disks than better resolved galaxies, which could be the reason for the here reported tension. According to their Table 3, the vertical velocity dispersion σ_z of TNG50-1 galaxies embedded in a dark matter halo of $M_{200} < 7.9 \times 10^{11} M_{\odot}$ is numerically increased by $\Delta\sigma_z > 0.1 V_{200}$, where V_{200} is the virial velocity of the halo, and M_{200} is its virial mass.

In this section, we analyze the effect of numerical resolution on the distribution of galaxy shapes using the TNG50 simulation. The flagship of the TNG project is called TNG50-1, which has the highest resolution among the here analyzed simulation runs. In addition, the TNG50 simulation suite includes the runs TNG50-2, TNG50-3, and TNG50-4, where a higher suffixed number indicates a lower-resolution realization as summarized in Table 1. The particle mass differs by a factor of eight between any TNG50 run and the next higher-resolution realization.

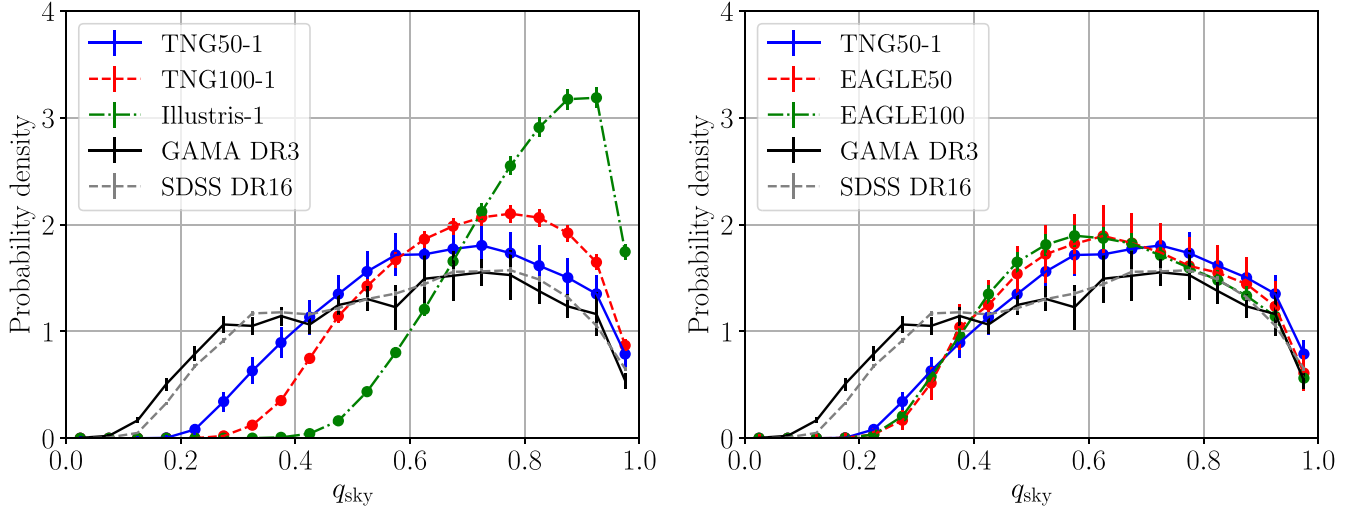


Figure 4. Comparison between the observed distribution of q_{sky} and that produced by different cosmological Λ CDM simulations for galaxies with $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$. The observed q_{sky} distributions (black and gray points with error bars) have been weighted based on the stellar mass distribution of the TNG50-1 run (shown in Figure 2). Invisible error bars (Equation (8)) are smaller than the data point symbol. The total χ^2 values between the observed and simulated distributions (Equation (7)) and the corresponding levels of tension are reported in Table 3. The TNG50-1, TNG100-1, Illustris-1, EAGLE50, and EAGLE100 computations shown here use 882, 6424, 6842, 480, and 3613 subhalos, respectively.

Table 2

Fraction of Galaxies with $q_{\text{sky}} < 0.4$ in the Stellar Mass Range $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$ in Different Simulation Runs (Left Columns) and Observational Surveys (Right Columns)

Simulations		Observations	
EAGLE100	0.089 ± 0.005	Local 11 Mpc ^a	0.212 ± 0.067
EAGLE50	0.088 ± 0.014	GAMA DR3 ^a	0.246 ± 0.007
EAGLE25	0.106 ± 0.038	GAMA DR3 ^b	0.238 ± 0.009
Illustris-1	0.0005 ± 0.0003	SDSS DR16 ^a	0.220 ± 0.001
TNG100-1	0.025 ± 0.002	SDSS DR16 ^b	0.216 ± 0.002
TNG50-1	0.098 ± 0.011		

Notes. The stated Poisson uncertainties are estimated as $\sigma = \sqrt{N_{<0.4} + 1}/N_{\text{tot}}$, where $N_{<0.4}$ and N_{tot} are the number of galaxies with $q_{\text{sky}} < 0.4$ and the total number of galaxies, respectively. The Poisson uncertainties for M_* -weighted distributions are given by Equation (8), where $w_{\text{tot},i}$ is the sum of weights of galaxies with $q_{\text{sky}} < 0.4$, and $w_{\text{max},i}$ is the maximum weight of such galaxies.

^a No M_* -weighting applied.

^b M_* -weighting applied based on the TNG50-1 stellar mass distribution.

Table 3

Statistical Comparison of the Observed Sky-projected Aspect Ratio Distribution from GAMA DR3 and SDSS DR16 with the Results of Different Cosmological Λ CDM Simulations

Simulation	GAMA DR3	SDSS DR16
EAGLE100	551.20 (21.68 σ)	2679.88 (50.68 σ)
EAGLE50	252.89 (13.65 σ)	288.51 (14.82 σ)
Illustris-1	1768.16 (40.80 σ)	14689.73 (120.61 σ)
TNG100-1	964.54 (29.54 σ)	7421.62 (85.39 σ)
TNG50-1	220.69 (12.52 σ)	358.36 (16.89 σ)

Note. The numbers show the total χ^2 calculated from 20 bins (Equation (7)). The bracketed numbers correspond to the level of tension for 20 degrees of freedom. The method for calculating the statistical significance of such extreme events is presented in the Appendix.

The left panel of Figure 5 shows the intrinsic aspect ratio distribution for different resolution realizations of the TNG50 simulation. The formation of thin galaxies and the peak position of the q_{int} distribution strongly depend on the

resolution: the mode shifts from $q_{\text{int}} = 0.88$ for TNG50-4 to $q_{\text{int}} = 0.33$ for TNG50-1.

The convergence of different simulation runs is statistically quantified in Table 4. The q_{int} distributions of the TNG50-1 and TNG50-2 runs disagree with each other at 7.60σ . Although this is still a high tension, it is much lower compared to that between TNG50-2(-3) and TNG50-3(-4), which differ at the 13.75σ (14.17σ) confidence level. Based on the q_{sky} distribution, the TNG50-1 and TNG50-2 runs disagree at only 0.83σ .

The right panel of Figure 5 compares the q_{int} distribution of TNG50-1 with EAGLE100, EAGLE50, and the higher-resolution realization EAGLE25 (Recal-L0025N0752). The similarity of the EAGLE results (Table 4) implies that the EAGLE simulations have numerically converged, so the deficit of intrinsically thin galaxies cannot be explained by resolution effects—as also concluded by Lagos et al. (2018) and apparent in their Figures A2 and A3. Although the q_{int} distribution of the EAGLE25 run is very similar to EAGLE50 or EAGLE100, the tension between the sky-projected aspect ratio distribution of EAGLE25 and GAMA DR3 (SDSS) is only 1.44σ (0.37σ). This is because EAGLE25 only has 81 galaxies in the stellar mass range $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$, which results in large Poisson uncertainties that decrease therewith the total χ^2 .

Importantly, the q_{sky} distributions of galaxies in TNG50-1 and EAGLE50 (EAGLE100) are consistent with each other at $1.4 \times 10^{-3}\sigma$ (0.58σ) confidence (Table 4). Thus, we demonstrate for the first time that Λ CDM simulations with SPH-based (EAGLE) and adaptive grid (Illustris/TNG) methods produce the same sky-projected aspect ratio distribution. The baryonic algorithms are also quite different between these simulations.

In order to quantify the thickening of simulated galaxies due to limited numerical resolution (e.g., Ludlow et al. 2021), we apply as an ansatz a parametric correction to the intrinsic shape distribution. The short axis λ_1 of each galaxy is scaled down by the factor x such that its q_{int} is corrected by

$$q_{\text{int, corr}} = xq_{\text{int}}, \quad x = \alpha + (1 - \alpha)q_{\text{int}}. \quad (12)$$

α is a variable ranging from 0–1 in steps of $\Delta\alpha = 0.01$. Applying this correction to the q_{int} of simulated galaxies keeps

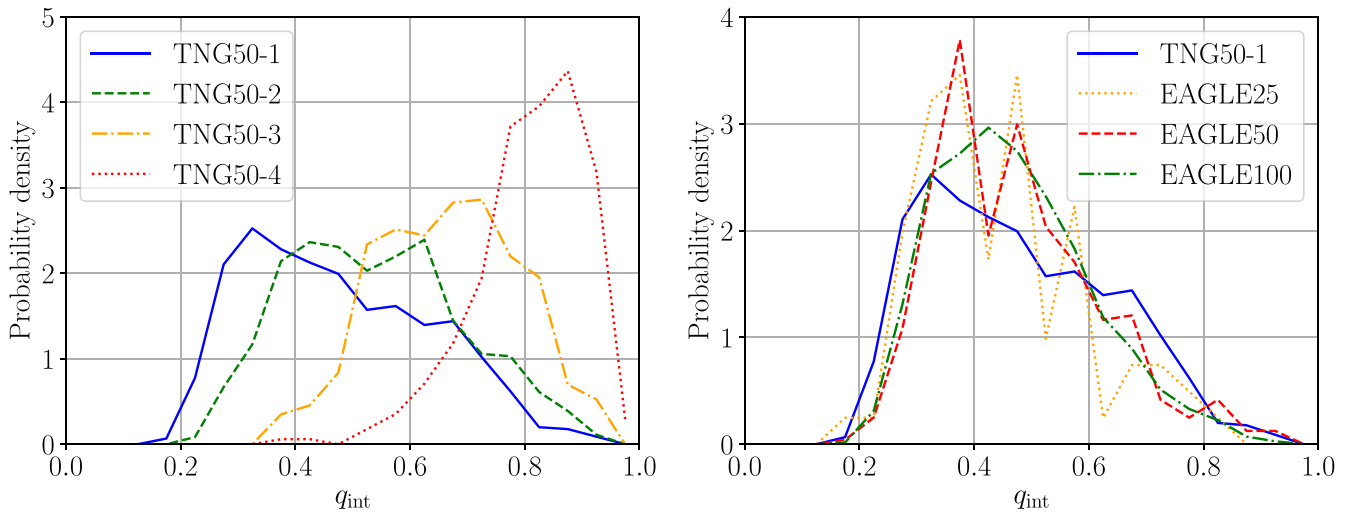


Figure 5. Intrinsic aspect ratio distribution of subhalos with $M_* > 10^{10} M_\odot$ for different resolution realizations of the simulation TNG50 (left) and EAGLE (right). A higher suffixed number for the TNG50 runs indicates a lower-resolution realization (i.e., TNG50-1 has the highest resolution). The simulated distributions are statistically compared in Table 4. EAGLE100 and EAGLE50 have the same resolution, while EAGLE25 has a higher resolution (Table 1).

Table 4

Testing the Numerical Convergence of Different Simulation Runs (Column 1) for Subhalos with $M_* > 10^{10} M_\odot$ by Showing the Total χ^2 (Equation (10)) between Their Distributions of q_{int} (Column 2) and q_{sky} (Column 3), along with the Corresponding Level of Tension (Bracketed Numbers)

Simulations	χ^2_{int} between Them (Tension)	χ^2_{sky} between Them (Tension)
TNG50-3 versus TNG50-4	268.26 (14.17 σ)	91.80 (6.69 σ)
TNG50-2 versus TNG50-3	255.73 (13.75 σ)	105.10 (7.45 σ)
TNG50-1 versus TNG50-2	107.85 (7.60 σ)	20.82 (0.83 σ)
TNG50-1 versus EAGLE100	87.83 (6.35 σ)	18.35 (0.58 σ)
TNG50-1 versus EAGLE50	56.78 (4.24 σ)	6.04 (0.0014 σ)
TNG50-1 versus EAGLE25	19.70 (0.71 σ)	2.98 (4.83 $\times 10^{-6}$ σ)
EAGLE100 versus EAGLE25	19.23 (0.66 σ)	0.50 (2.80 $\times 10^{-13}$ σ)
EAGLE50 versus EAGLE25	12.87 (0.15 σ)	0.95 (1.26 $\times 10^{-10}$ σ)

Note. We use 20 bins in q_{int} and q_{sky} , giving 20 degrees of freedom. EAGLE50 and EAGLE100 use the same resolution, so the similarity between their results (χ^2 not shown) is not a strong test of numerical convergence. Note that the higher-resolution realization EAGLE25 only has 81 galaxies with $M_* > 10^{10} M_\odot$, which reduces the total χ^2 because of the higher Poisson uncertainties. The intrinsic aspect ratio distributions underlying these comparisons are plotted in Figure 5.

a galaxy with a high q_{int} round, but makes a thin disk galaxy even thinner. $\alpha = 1$ does not change the q_{int} of a galaxy, while $\alpha = 0$ yields a corrected $q_{\text{int, corr}} = q_{\text{int}}^2$, therewith causing the most substantial thinning of the galaxy population in our parameterization.

The tension between the observed and rescaled simulated q_{sky} distributions is shown in Figure 6 for different resolution realizations of the TNG50 simulation. As expected from our previous analysis, the tension systematically decreases with higher resolution. The tension between TNG50-1 and the GAMA survey (SDSS) reaches the 5σ confidence level if galaxies in the TNG50-1 run are corrected by $\alpha = 0.668$ ($\alpha = 0.738$). The tension between TNG50-1 and GAMA (SDSS) is minimized for a correction factor of $\alpha = 0.32$ ($\alpha = 0.47$), the tension in this case being only $4.6 \times 10^{-3}\sigma$ ($1.2 \times 10^{-2}\sigma$). This demonstrates that our simple parametric correction with a single parameter α (Equation (12)) can make the simulations agree very well with observations.

We can use this procedure to quantify how improving the resolution affects the simulated aspect ratio distribution, which then allows us to quantify if the simulations are numerically converged. To do this, we first find which α value must be applied to a TNG50 run to minimize the tension with the uncorrected q_{int} distribution of the next higher-resolution realization with an

eight-times-higher mass resolution. We find that the tensions between the q_{int} distributions of the rescaled TNG50-4 and TNG50-3, rescaled TNG50-3 and TNG50-2, and rescaled TNG50-2 and TNG50-1 runs become minimal for α values of 0.03 (0.85 σ), 0.44 ($5.9 \times 10^{-3}\sigma$), and 0.69 (0.80 σ) applied to the TNG50-4, TNG50-3, and TNG50-2 runs, respectively. The true q_{int} distribution of each TNG50 run and the so-corrected distribution of the next lower-resolution TNG50 run is compared in the left panel of Figure 7, demonstrating that our parametric correction to, e.g., TNG50-2 galaxy shapes reproduces quite well the actual q_{int} distribution of TNG50-1. The required α value increases with higher resolution, but we cannot confirm that the TNG50 simulations have numerically converged, as that would be achieved if $\alpha = 1$.

In a second step, we extrapolate the α values by plotting $\ln(1 - \alpha)$ against the logarithmic dark matter particle mass of the corresponding TNG50 run (right panel of Figure 7). The α value required for galaxies in the TNG50-1 run to mimic the result of an eight-times-lower dark matter particle mass simulation is estimated using a parabolic and a linear fit in the $\log_{10}(m_{\text{dm}}/M_\odot)$ versus $\ln(1 - \alpha)$ diagram of Figure 7. This predicts $\alpha = 0.824$ (0.834) for a linear (parabolic) fit. Being more conservative, we apply the linearly extrapolated result

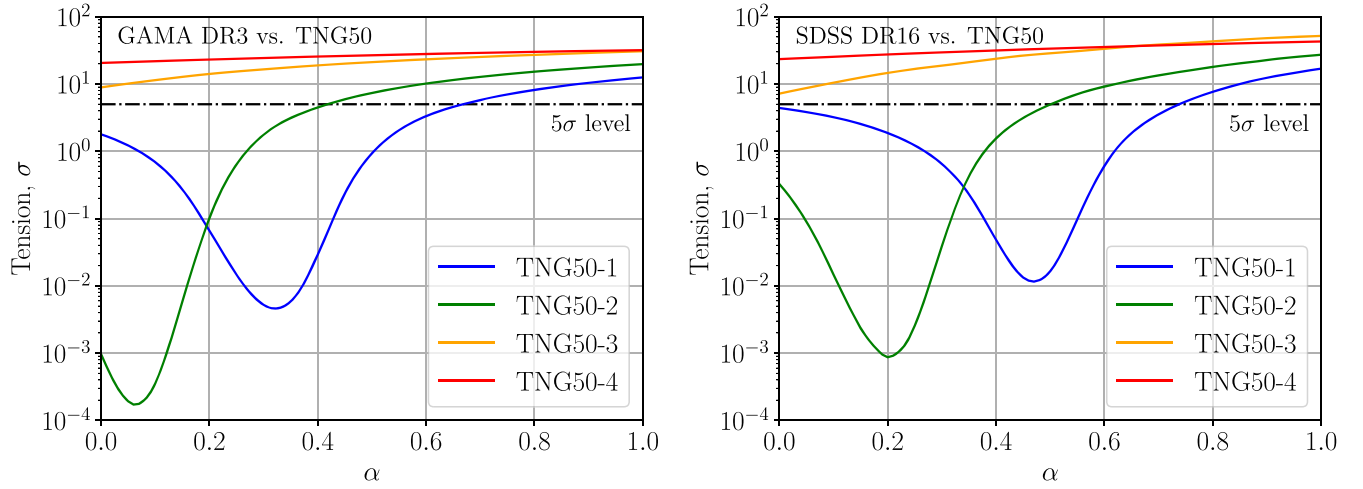


Figure 6. Level of tension between the observed and simulated q_{sky} distribution of galaxies with $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$ in dependence of α (Equation (12)) for different resolution realizations of the TNG50 simulation, which we show with different colored lines (see the legend). The left panel considers GAMA as the observational sample, while the right panel uses SDSS. The tension between TNG50-1 and GAMA DR3 (SDSS DR16) reaches the 5σ confidence level (dashed horizontal line) if $\alpha = 0.668$ ($\alpha = 0.738$), with the tension becoming minimal if instead $\alpha = 0.32$ ($\alpha = 0.47$), the tension in this case being only $4.6 \times 10^{-3}\sigma$ ($1.2 \times 10^{-2}\sigma$).

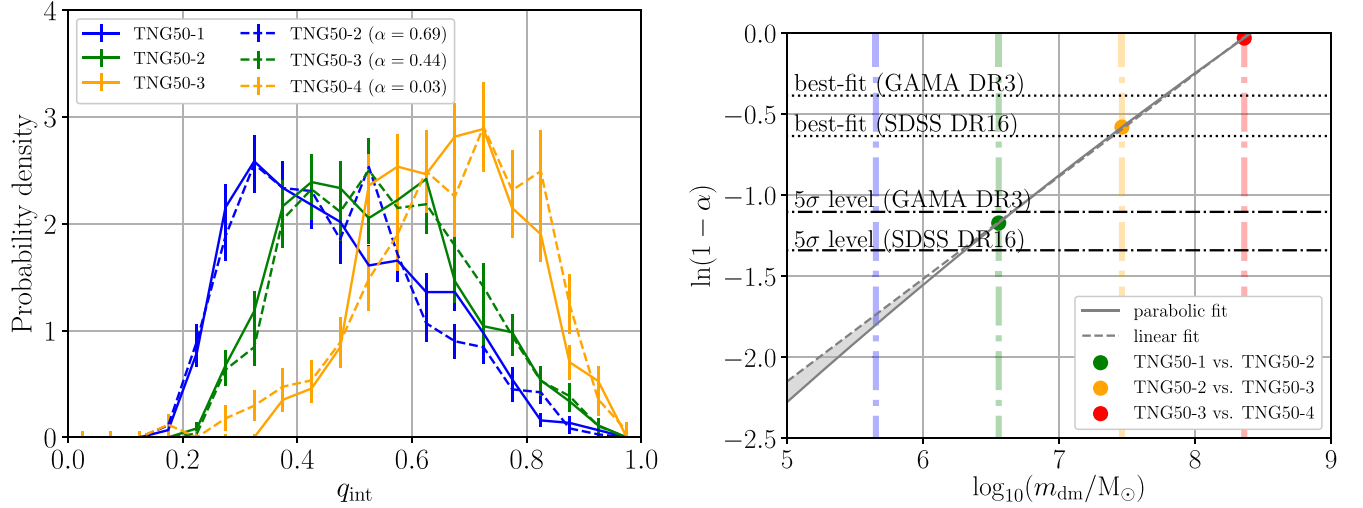


Figure 7. Left: the q_{int} distribution of subhalos with $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$ in each TNG50 simulation (solid lines) is compared with the corrected distribution for subhalos in the next lower-resolution simulation, which we show using a dashed line of the same color. The applied correction factors and resulting levels of tension are $\alpha = 0.69$ (0.80σ), $\alpha = 0.44$ ($5.9 \times 10^{-3}\sigma$), and $\alpha = 0.03$ (0.85σ) for subhalos in TNG50-2, TNG50-3, and TNG50-4, respectively. These corrections minimize the tension with the uncorrected q_{int} distribution of the next higher-resolution simulation. Right: checking for convergence of the TNG50 q_{int} distribution with respect to the dark matter mass resolution m_{dm} for galaxies with $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$. The above-mentioned α values are plotted in terms of $\ln(1 - \alpha)$ against $\log_{10}(m_{\text{dm}}/M_\odot)$. The vertical lines mark the mass of a dark matter particle in the TNG50 runs as listed in Table 1. The dotted–dashed horizontal lines mark $\alpha = 0.668$ and $\alpha = 0.738$, which are required for the q_{sky} distribution of the TNG50-1 run to match that of GAMA DR3 and SDSS DR16, respectively, at the 5σ confidence level. The dotted lines represent $\alpha = 0.32$ and 0.47 because these values would minimize the tension of TNG50-1 with GAMA and SDSS, respectively. The dashed (solid) gray line is a linear (parabolic) fit. The gray shaded region indicates how much the q_{int} distribution could differ in an even higher-resolution run than TNG50-1. We estimate that using an eight-times lower dark matter particle mass than TNG50-1 (dotted–dashed blue vertical line) is equivalent to scaling galaxies in it by α in the range between 0.834 (parabolic fit) and 0.824 (linear fit).

that $\alpha = 0.824$ to the TNG50-1 run, which yields an 8.68σ (8.71σ) tension with GAMA (SDSS). Thus, an eight-times higher-resolution realization of the TNG50-1 run would almost certainly not reduce the here reported tension below the 5σ confidence level.

In the third and final step, we estimate the tension with observations by extrapolating the data further to five more refinement levels than in the TNG50-1 run. Therefore, we extract the α values not only for an eight-times-lower m_{dm} but also for an $8^2\times$, $8^3\times$, $8^4\times$, and $8^5\times$ lower m_{dm} than in the TNG50-1 run. These so-obtained α values for the five different resolution levels are listed in Table 5 for the parabolic and

linear extrapolations. The α values are then successively applied to estimate the intrinsic aspect ratio of each subhalo in an $8^5\times$ higher-resolution realization than TNG50-1. In other words, the aspect ratio distribution for an $8^2\times$ higher-resolution run than TNG50-1 is obtained by correcting the aspect ratios of subhalos in the eight-times higher-resolution run by applying Equation (12) with $\alpha = 0.915$ (parabolic fit) or $\alpha = 0.900$ (linear fit). This procedure is repeated until we reach an $8^5\times$ higher refinement level, which on the last step requires applying $\alpha = 0.991$ (parabolic fit) or $\alpha = 0.982$ (linear fit) to the q_{int} distribution of the $8^4\times$ higher-resolution realization than TNG50-1. Since α is now close to unity, this almost

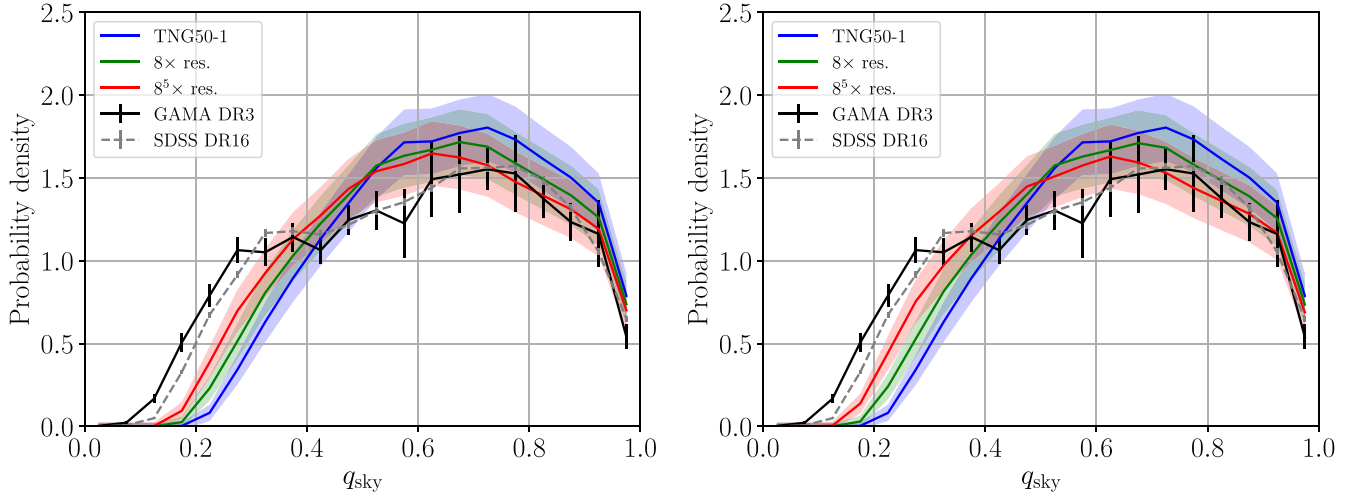


Figure 8. Similar to Figure 4, but for an eight-times-higher (green line) and $8^5 \times$ higher (red line) resolution run than TNG50-1 (blue line). The colored shaded regions highlight the uncertainties given by Equation (8). The simulated q_{sky} distributions for a higher-resolution run than TNG50-1 have been estimated with a parabolic (left panel) or linear (right panel) extrapolation in the $\log_{10}(m_{\text{dm}}/M_{\odot})$ vs. $\ln(1 - \alpha)$ diagram as shown in the right panel of Figure 7 (see the text). The total χ^2 (Equation (7)) between the observed and simulated distribution for different refinement levels is reported in Table 5 along with the corresponding level of tension.

Table 5

Statistical Comparison of the Observed Sky-projected Aspect Ratio Distributions from GAMA DR3 (Fourth and Fifth Columns) and SDSS DR16 (Sixth and Seventh Columns) with the Results of Five More Refinement Levels (First Column) Than in the TNG50-1 Run

Resolution	α (Equation (12))		χ^2 and Tension with GAMA DR3		χ^2 and Tension with SDSS DR16	
	Parabolic Fit	Linear Fit	Parabolic Fit	Linear Fit	Parabolic Fit	Linear Fit
TNG50-1	220.69 (12.52 σ)	220.69 (12.52 σ)	358.36 (16.89 σ)	358.36 (16.89 σ)
$8 \times$	0.834	0.824	134.17 (8.94 σ)	128.96 (8.68 σ)	140.34 (9.23 σ)	129.60 (8.71 σ)
$8^2 \times$	0.915	0.900	99.29 (7.12 σ)	90.08 (6.59 σ)	77.38 (5.72 σ)	63.95 (4.78 σ)
$8^3 \times$	0.958	0.944	84.94 (6.18 σ)	71.83 (5.34 σ)	57.06 (4.26 σ)	41.14 (2.91 σ)
$8^4 \times$	0.980	0.968	78.27 (5.77 σ)	63.12 (4.72 σ)	48.57 (3.57 σ)	32.31 (2.05 σ)
$8^5 \times$	0.991	0.982	75.37 (5.58 σ)	58.36 (4.36 σ)	45.12 (3.27 σ)	28.06 (1.61 σ)

Note. The second and third columns list the α values obtained by extrapolating the TNG50 runs to an $8 \times$, $8^2 \times$, $8^3 \times$, $8^4 \times$, and $8^5 \times$ lower dark matter particle mass than in TNG50-1 using a parabolic or a linear fit in the $\log_{10}(m_{\text{dm}}/M_{\odot})$ versus $\ln(1 - \alpha)$ diagram (right panel of Figure 7). These so-obtained α values have been successively applied to the intrinsic aspect ratios of subhalos in TNG50-1 (see the text).

reaches full convergence in q_{int} . Further improvements to the resolution can be expected to have sub-percent level effects on the intrinsic aspect ratios.

The q_{sky} distributions for an $8 \times$ and $8^5 \times$ higher dark matter mass resolution than the TNG50-1 run derived from a parabolic (linear) extrapolation are presented in the left (right) panel of Figure 8. Raising the resolution of the dark matter mass in TNG50-1 by a factor of 8^5 yields an estimated tension of 5.58σ (parabolic fit) and 4.36σ (linear fit) with GAMA. In contrast, the discrepancy is significantly alleviated with respect to the older SDSS data set—we estimate a tension of 3.27σ with a parabolic fit and only a 1.61σ tension with a linear fit. The difference between the two observational samples is mainly because the fraction of galaxies with $q_{\text{sky}} \lesssim 0.3$ is higher in the GAMA survey compared to SDSS (Figure 8). The total χ^2 values and corresponding levels of tension with the GAMA and SDSS surveys for different resolution realizations and extrapolation methods are summarized in Table 5.

Since the parabolic fit considers more information than the linear fit and exactly matches the data, we consider it as the nominal case in our main analysis. The linear fit is still shown in order to illustrate the effect of uncertainties in the extrapolation procedure. Therefore, our estimate is that increasing the mass resolution of TNG50-1 by $8^5 \times$ still leaves a significant tension

of 5.58σ with GAMA, exceeding the 5σ plausibility threshold. The SDSS data set is then in 3.27σ tension, which still points to an underestimated fraction of very thin galaxies in the Λ CDM simulations (see Figure 8).

Finally, we note that our approach is very conservative with respect to the Λ CDM framework. First of all, the extrapolation to five more refinement levels than TNG50-1 is based only on the four resolution realizations of the TNG50 simulation. For example, the TNG50-4 simulation has $m_{\text{dm}} = 2.3 \times 10^8 M_{\odot}$, which is a factor of $8^8 = 1.7 \times 10^7$ more massive than dark matter particles in an $8^5 \times$ higher-resolution realization than the TNG50-1 run. The parabolic and linear fits have been derived from simulations with $4.5 \times 10^5 (TNG50 - 1) \leq m_{\text{dm}}/M_{\odot} \leq 2.3 \times 10^8 (TNG50 - 4)$. This introduces additional uncertainties because the shape of the relation between $\log_{10}(m_{\text{dm}}/M_{\odot})$ and $\ln(1 - \alpha)$ may deviate significantly from the derived polynomial fits at lower m_{dm} . Second, we performed the extrapolation to very high-resolution realizations without assuming that numerical convergence is reached between two different refinement levels. It is not impossible that numerical convergence ($\alpha = 1$) is already attained before reaching the $8^5 \times$ higher-resolution level. Consequently, higher-resolution runs of self-consistent Λ CDM simulations are still required to test when numerical convergence is reached and if the tension with observations is then alleviated.

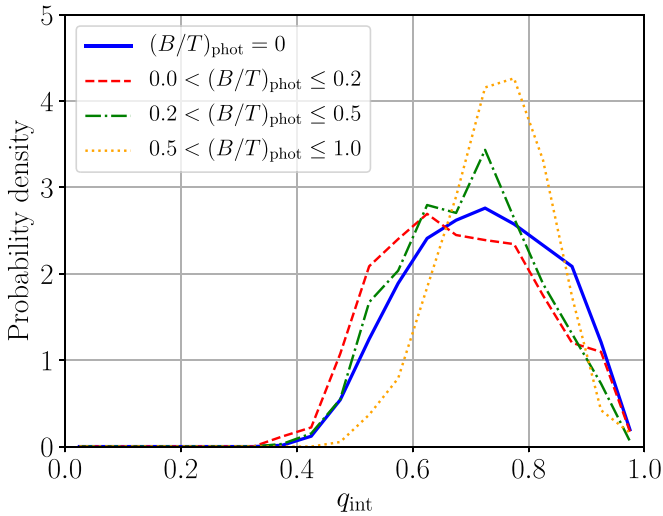


Figure 9. The q_{int} distribution of subhalos with $M_* > 10^{10} M_\odot$ in different photometric $(B/T)_{\text{phot}}$ bins in the Illustris-1 simulation as derived by Bottrell et al. (2017a) for camera angle 0. The solid blue, dashed red, dotted-dashed green, and dotted yellow lines refer to galaxy subsamples with $(B/T)_{\text{phot}} = 0$ (3999 subhalos), $0 < (B/T)_{\text{phot}} \leq 0.2$ (1715 subhalos), $0.2 < (B/T)_{\text{phot}} \leq 0.5$ (658 subhalos), and $0.5 < (B/T)_{\text{phot}} \leq 1.0$ (380 subhalos), respectively.

7.2. Photometric Bulge/Total Ratios of Illustris-1 Subhalos

In contrast to our findings, Bottrell et al. (2017b) found a significant deficit of bulge-dominated galaxies based on the photometric bulge/total ratios of subhalos in the Illustris-1 simulation compared to SDSS. In particular, Bottrell et al. (2017a) photometrically derived $(B/T)_{\text{phot}}$ ratios of subhalos with $M_* > 10^{10} M_\odot$ at $z=0$ in the Illustris-1 simulation by performing a bulge-disk decomposition of the surface brightness profile with fixed Sérsic indices $n_b = 4$ for the bulge and $n_d = 1$ for the disk. By applying the same decomposition analysis to observed SDSS galaxies, they found a significant deficit of bulge-dominated galaxies with $M_*/M_\odot = 10^{10} - 10^{11}$ in the Illustris-1 simulation (see Figures 4 and 6 in Bottrell et al. 2017b).

If the Illustris-1 simulation indeed lacks bulge-dominated galaxies, this deficit should also be evident in the q_{int} parameter (Section 6.1). To test if $(B/T)_{\text{phot}}$ reflects the shapes of simulated galaxies as quantified by q_{int} , we show in Figure 9 its distribution for four simulated galaxy subsamples of Bottrell et al. (2017a) with different photometric morphologies, i.e., $(B/T)_{\text{phot}}: 0, (0, 0.2], (0.2, 0.5],$ and $(0.5, 1]$.¹⁸ Throughout this work, we use the $(B/T)_{\text{phot}}$ ratio derived for the r band from camera angle 0. Remarkably, all four galaxy subsamples have a very similar aspect ratio distribution, with the peak between $q_{\text{int}} = 0.6 - 0.8$. The distributions have very few galaxies with $q_{\text{int}} < 0.5$: the proportions for the above-mentioned $(B/T)_{\text{phot}}$ bins are only 3.4%, 7.1%, 3.6%, and 0.26%, respectively, implying that the sample of Bottrell et al. (2017a) is actually bulge-dominated.

In order to understand this mismatch between $(B/T)_{\text{phot}}$ and q_{int} , we begin by presenting color images¹⁹ generated from the Friends-of-Friends (FoF) halo finder. Subhalos with low

$0 < (B/T)_{\text{phot}} < 0.2$ and $q_{\text{int}} < 0.4$ indeed have spiral structures and a disk (Figure 10). The κ_{rot} morphological parameter (Equation (2) in Rodríguez-Gomez et al. 2015) indicates that these subhalos are rotation-dominated ($\kappa_{\text{rot}} > 0.5$).

The mismatch between the morphological parameters becomes evident in Figure 11, which shows the significant fraction of subhalos (13.1%) with low $0 < (B/T)_{\text{phot}} < 0.2$ but high $q_{\text{int}} > 0.8$ for their stellar component.²⁰ These are featureless-looking dispersion-dominated ($\kappa_{\text{rot}} < 0.5$) galaxies without a pronounced disk or spiral structures. Based on these images and given that the intrinsic aspect ratio is a very robust parameter to quantify the actual shape of a galaxy, we conclude that the 2D parametric surface brightness decomposition applied by Bottrell et al. (2017a) is inadequate if applied to simulated galaxies.

In addition to a bulge-disk decomposition, Bottrell et al. (2017a) also applied a pure Sérsic model to the galaxies in subhalos by varying the Sérsic index between 0.5 and 8. Figure 12 shows a significant correlation between $(B/T)_{\text{phot}}$ and the Sérsic index, highlighting the inevitable confusion between an elliptical galaxy with $n_s \approx 1$ and a thin exponential disk if only considering the surface brightness profile. Choosing a different camera angle leads to the same results. In combination with Figure 9, this shows that neither n_s nor $(B/T)_{\text{phot}}$ correlates with q_{int} .

This might be due to the resolution of the Illustris-1 simulation not being sufficient to apply a photometric decomposition. Importantly, we have shown in this contribution that the observed and simulated galaxy morphology distributions differ significantly (Table 3). The vast majority of observed galaxies are spirals (e.g., Loveday 1996; Delgado-Serrano et al. 2010), while the Λ CDM simulations form a far-too-large fraction of bulge-dominated galaxies (Figures 3 and 4). Thus, assuming that all galaxies with low $(B/T)_{\text{phot}}$ are intrinsically thin is quite accurate in the real universe where spirals are quite common, but not in a Λ CDM universe where they are rare (Section 6.2). In the Illustris-1 simulation, there are so few disk galaxies that galaxies with low $(B/T)_{\text{phot}}$ are nearly always ellipticals with an exponential-like surface brightness profile (Sérsic index close to 1).

These problems with the $(B/T)_{\text{phot}}$ parameter are avoided by using the sky-projected aspect ratio (Figure 4), which is closely linked to the intrinsic aspect ratio and thus yields a more robust measurement of the galaxy morphology. Since q_{sky} is observable, it allows for a much more direct test of the model, provided the sample size is sufficient to statistically sample over projection effects. This is true in our case because the TNG50-1, GAMA DR3, and SDSS DR16 samples contain 882, 5304, and 232,128 galaxies, respectively, in the range $10.0 < \log_{10}(M_*/M_\odot) \leq 11.65$ used for our comparisons.

7.3. Impact of the Merger History

The Λ CDM theory strictly implies a hierarchical merger-driven build-up of the galaxy population. Galaxy mergers are driven by dynamical friction on the extended dark matter halos of interacting galaxies (e.g., Kroupa 2015). Mergers grow the bulge component and thicken the stellar disk. By studying galaxy-galaxy interactions in N -body/hydrodynamical simulations, Hwang et al. (2021) showed that the disk angular

¹⁸ We use their DISTINCT catalog, which contains the galaxy sample analyzed by Torrey et al. (2015).

¹⁹ Downloaded [26.06.2020] from the Illustris-1 webpage using the Illustris Galaxy Observatory tool: https://www.illustris-project.org/galaxy_obs.

²⁰ Subhalos with $(B/T)_{\text{phot}} = 0$ are probably erroneous (Bottrell et al. 2017a) and are excluded from our analysis except in Figure 9.

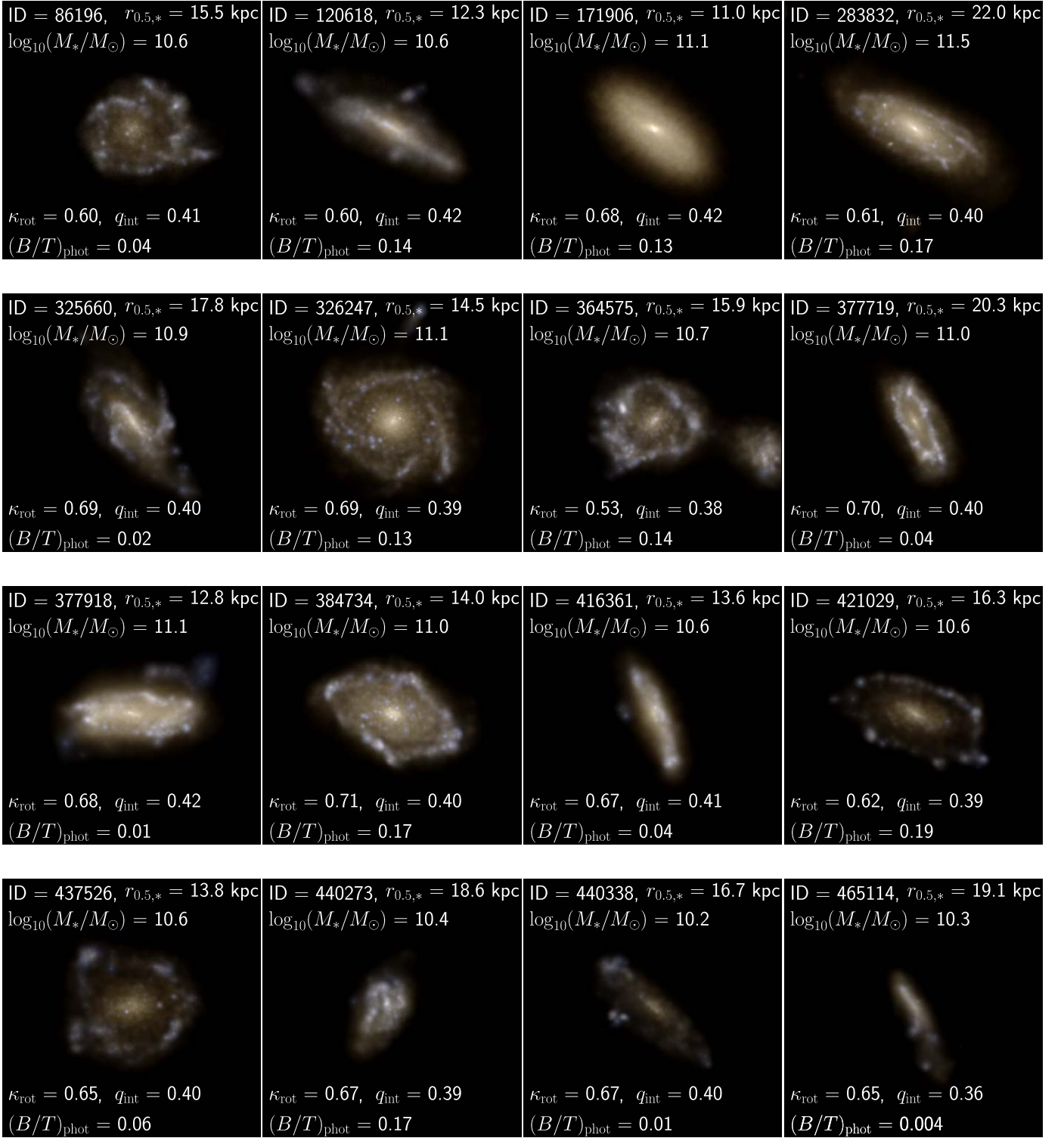


Figure 10. Images of Illustris-1 galaxies in subhalos with $0 < (B/T)_{\text{phot}} < 0.2$ from camera angle 0 that have the lowest q_{int} values in the sample of Bottrell et al. (2017a). These galaxies are rotation-dominated ($\kappa_{\text{rot}} > 0.5$) and have spiral features, with $q_{\text{int}} \approx 0.4$. The FoF images were downloaded from the Illustris Galaxy Observatory tool: https://www.illustris-project.org/galaxy_obs. The field of view is ten stellar half-mass radii of the shown subhalo. These are genuinely disk-dominated galaxies.

momentum of the late-type galaxy decreases by about 15%–20% after a prograde collision. Thus, galaxies with a quiescent merger history are expected to have lower bulge fractions than galaxies that have undergone a major merger (see also, e.g., Bournaud et al. 2005; D’Onghia et al. 2006). Since most

observed galaxies are late types (e.g., Delgado-Serrano et al. 2010) and $\approx 50\%$ of these have no classical bulge (Graham & Worley 2008; Kormendy et al. 2010), we might expect that galactic mergers are less frequent in the universe than in Λ CDM simulations (Disney et al. 2008; Stewart et al. 2008;

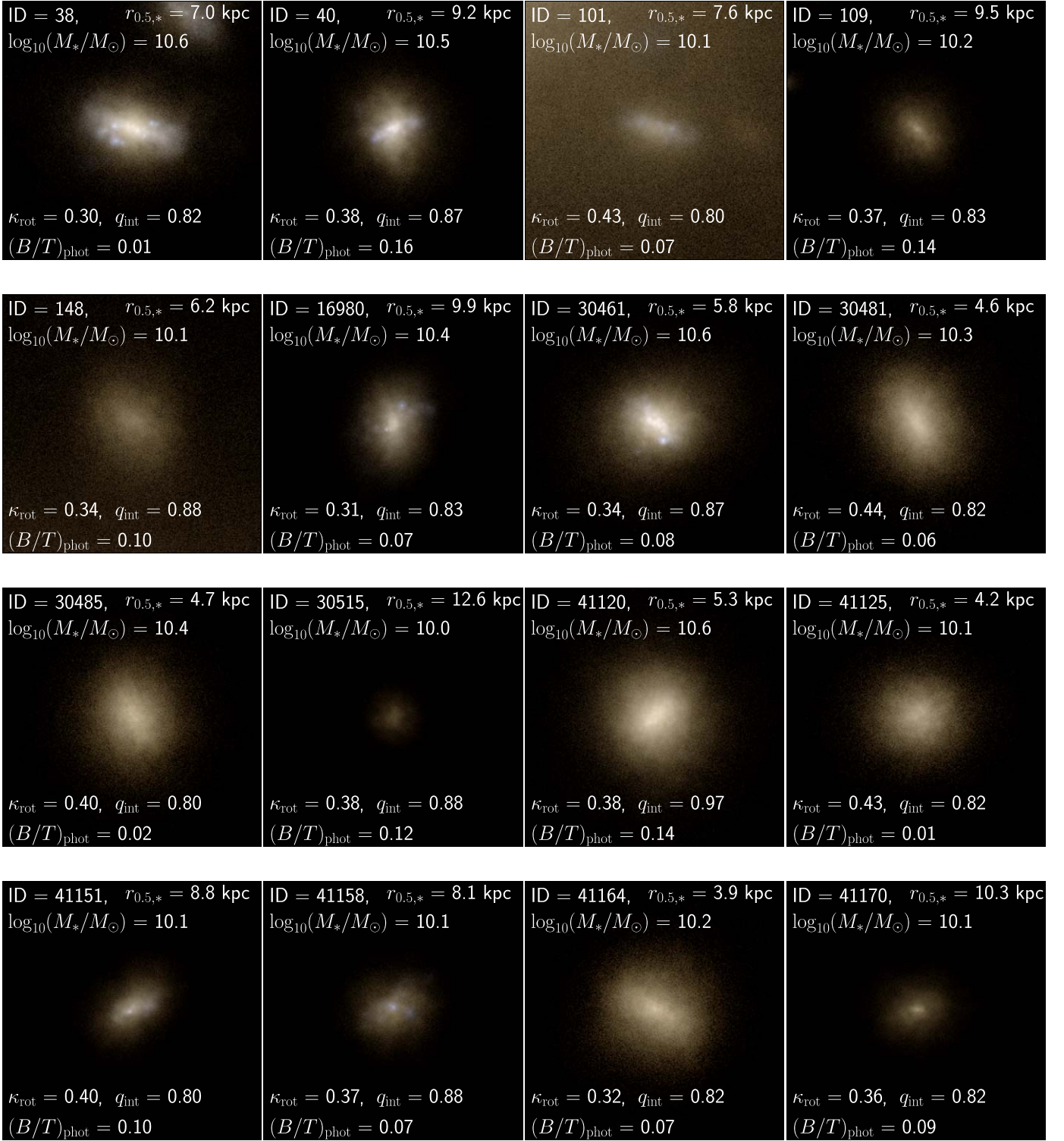


Figure 11. Similar to Figure 10, but for galaxies with high $q_{\text{int}} > 0.8$ despite a low $(B/T)_{\text{phot}} < 0.2$. These are dispersion-dominated ($\kappa_{\text{rot}} < 0.5$) featureless-looking galaxies. They are best understood as early-type galaxies where the photometric classification failed by assigning a low $(B/T)_{\text{phot}}$.

Fakhouri et al. 2010; Kroupa 2015; Wu & Kroupa 2015). This may underlie the tension between the observed and simulated sky-projected aspect ratio distributions (Figure 4).

We therefore investigate if galaxies with a quiescent merger history in the Λ CDM framework are typically much thinner. We focus on the TNG50-1 run, as it has the highest resolution (Table 1) and yields the lowest tension in the here analyzed

simulations (Table 3). The merger trees of the Illustris and TNG projects can be downloaded from their webpages,²¹ with a detailed description available in Rodriguez-Gomez et al. (2015). We quantify the merger history of a galaxy by

²¹ <https://www.illustris-project.org> [21.07.2020] and <https://www.tng-project.org> [21.07.2020].

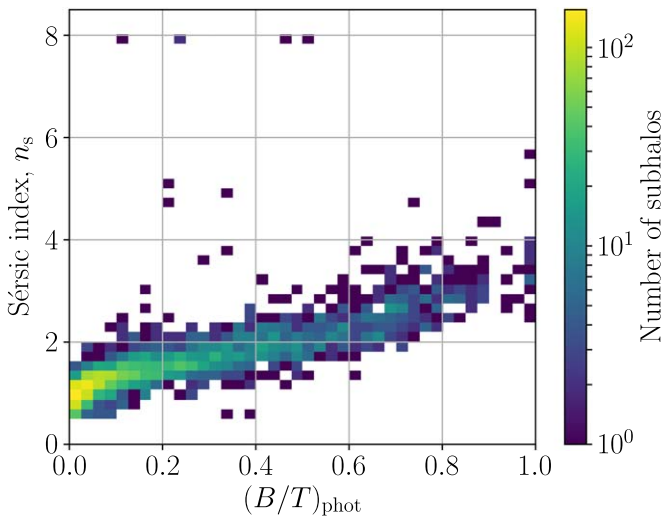


Figure 12. Correlation between the photometric $(B/T)_{\text{phot}}$ ratio and the Sérsic index n_s for the Illustris-1 simulated subhalo sample analyzed by Bottrell et al. (2017a). The parameters shown here are from catalogs provided in their Tables A1 and A2 for camera angle 0, with different camera orientations leading to the same results (not shown). Notice the tight correlation between $(B/T)_{\text{phot}}$ and n_s , but a lack of correlation between $(B/T)_{\text{phot}}$ and q_{int} (see Figure 9).

considering the total mass ratio $\mu \leq 1$ of the two progenitors and the maximum lookback time t_{max} during which we consider mergers. If there are multiple mergers in this timeframe, we use the most major merger, i.e., that with the highest μ . This μ value is used to construct two galaxy samples that differ according to whether the galaxy had at least one major merger with $\mu \geq 1/12$ within a lookback time of $t_{\text{max}} = 12$ Gyr ($z = 3.7$). We restrict ourselves to galaxies with $M_* > 10^{10} M_\odot$ and $M_{\text{dm}}/M_* > 1$ at $z = 0$. The constraint on the M_{dm}/M_* ratio is applied to exclude dark matter-poor galaxies, which cannot be traced back accurately. This excludes 26 galaxies from the original sample. About 89% of all subhalos at $z = 0$ have undergone at least one major merger, which is broadly consistent with other Λ CDM simulations (Stewart et al. 2008; Fakhouri et al. 2010). The remaining 11% of all subhalos had at most only minor merger(s) with $\mu \leq 1/12$ during the past 12 Gyr.

As expected from zoomed-in simulations of galaxies in underdense environments, intrinsically thin galaxies are more frequent in the sample of galaxies with a quiescent merger history compared to that with an active merger history (see the left panel of Figure 13). In particular, $39\% \pm 2\%$ of galaxies with an active merger history have $q_{\text{int}} < 0.4$, but this rises to $50\% \pm 7\%$ for galaxies with a quiescent merger history. The aspect ratio of the thinnest such galaxy ($q_{\text{int}} = 0.22$) is similar to that in the sample with an active merger history ($q_{\text{int}} = 0.19$). Additionally, $9.5\% \pm 1.1\%$ of galaxies with an active merger history have $q_{\text{sky}} < 0.4$, but this rises to $12.5\% \pm 3.7\%$ for the galaxies with a quiescent merger history (right panel of Figure 13).

Thus, the lack of intrinsically thin galaxies in Λ CDM could be partly due to major mergers. More likely, it is due to a combination of major and minor mergers, which are unavoidable in Λ CDM, as galaxies grow their mass through mergers. We emphasize that minor mergers are expected to be much less frequent in an alternative framework where galaxies lack extended dark matter halos, as occurs in MOND (Milgrom 1983). In addition, secular processes like disk-halo angular

momentum exchange could also drive the formation of significant bars and bulges in Λ CDM (Athanasoula 2002; Sellwood et al. 2019), but perhaps not in the real universe (Banik et al. 2020; Roshan et al. 2021a) where the fraction of bars differs substantially from Λ CDM expectations (Reddish et al. 2021). There is also a highly significant discrepancy between the pattern speeds of bars in observations and in Λ CDM simulations, where the results seem to have converged with respect to the ratio of bar length to corotation radius (Roshan et al. 2021b). The tension is caused by the fact that bars are expected to be slowed down by dynamical friction with the dark matter halo, but observed bars are fast. This is another indication against dynamical friction from massive dark matter halos.

7.4. Feedback

In the previous section, we discussed that the disagreement between the observed and Λ CDM simulated galaxy shapes could be partly due to the frequency of mergers being too high in this framework. Another possibility is that the tension is caused by the feedback description used in the simulations. In particular, Lagos et al. (2018) discussed the link between loss/gain of angular momentum and dry/wet mergers. Using the EAGLE and HYDRANGEA (Bahé et al. 2017; Barnes et al. 2017) hydrodynamical simulations, they showed that dry mergers typically decrease the stellar spin parameter while wet mergers increase it. Therefore, galaxies that further accrete cold gas are able to reform their disks. This process is sensitive to the implemented feedback description.

As shown in Sections 6.1 and 7.1, the TNG50-1 and EAGLE simulations produce very similar aspect ratio distributions despite relying on different sub-grid models. This is a strong indication that the tension is likely not caused by the implemented sub-grid feedback models. Even so, improved models might yet alleviate or resolve the tension. For example, the feedback recipe in both the TNG and EAGLE projects could be too strong to allow the (re)formation of disks. In this case, the shape of the galaxies would be set by the merger rate independently of whether the mergers are dry or wet. However, a weaker feedback description could potentially increase the efficiency of disk formation. We note that strong feedback is required in Λ CDM simulations for them to explain why the Newtonian dynamical mass of a galaxy or galaxy group often greatly exceeds $6.4\times$ its baryonic mass (e.g., Müller et al. 2021), even though Λ CDM needs this to be the cosmic ratio between baryonic and total mass (Planck Collaboration VI 2020). A strong feedback prescription is also needed to solve the missing satellite galaxy problem (e.g., Brooks et al. 2013).

7.5. Disk Galaxies in MOND

The difficulties faced by Λ CDM with regards to the high fraction of thin disk galaxies motivate us to consider MOND one of the main alternative frameworks that is generally considered to perform better on galaxy scales (for reviews, see, e.g., Famaey & McGaugh 2012; Kroupa 2015; Banik & Zhao 2022). For illustrative purposes, we show in the left panel of Figure 14 the present-day q_{int} distribution of six disk galaxies with $10^{10} < M_*/M_\odot \leq 9.56 \times 10^{10}$ formed in hydrodynamical MOND simulations of collapsing gas clouds conducted by Wittenburg et al. (2020) using the PHANTOM

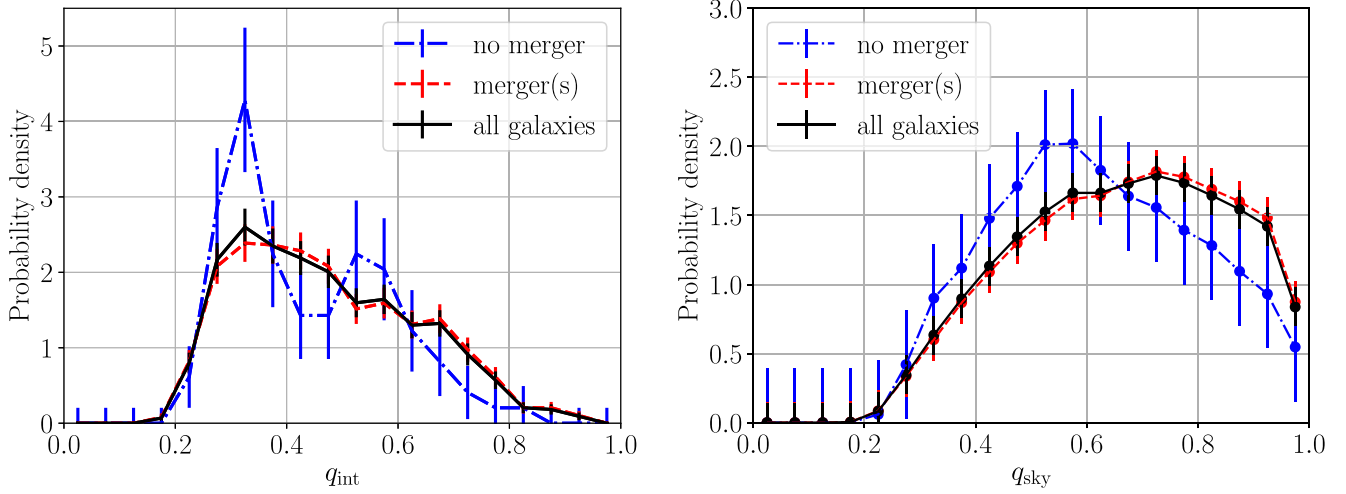


Figure 13. Intrinsic (left) and sky-projected (right) aspect ratio distribution of subhalos in the TNG50-1 simulation with $M_* > 10^{10} M_\odot$ and $M_{\text{dm}}/M_* > 1$ with an active (dashed red) and quiescent (dotted–dashed blue) merger history, defined according to whether the galaxy had at least one major merger with total mass ratio $\mu \geq 1/12$ within the last 12 Gyr. The intrinsic aspect ratio distribution of the active (quiescent) sample, which contains 779 (98) galaxies, has a median of $q_{\text{int}} = 0.40$ (0.45). Results for all 877 galaxies (regardless of merger history) are shown by the solid black lines. The proportion of galaxies with $q_{\text{int}} < 0.4$ is $39\% \pm 2\%$ and $50\% \pm 7\%$ in the sample with an active and a quiescent merger history, respectively. If instead we require $q_{\text{sky}} < 0.4$, these proportions become $9.5\% \pm 1.1\%$ (active) and $12.5\% \pm 3.7\%$ (quiescent).

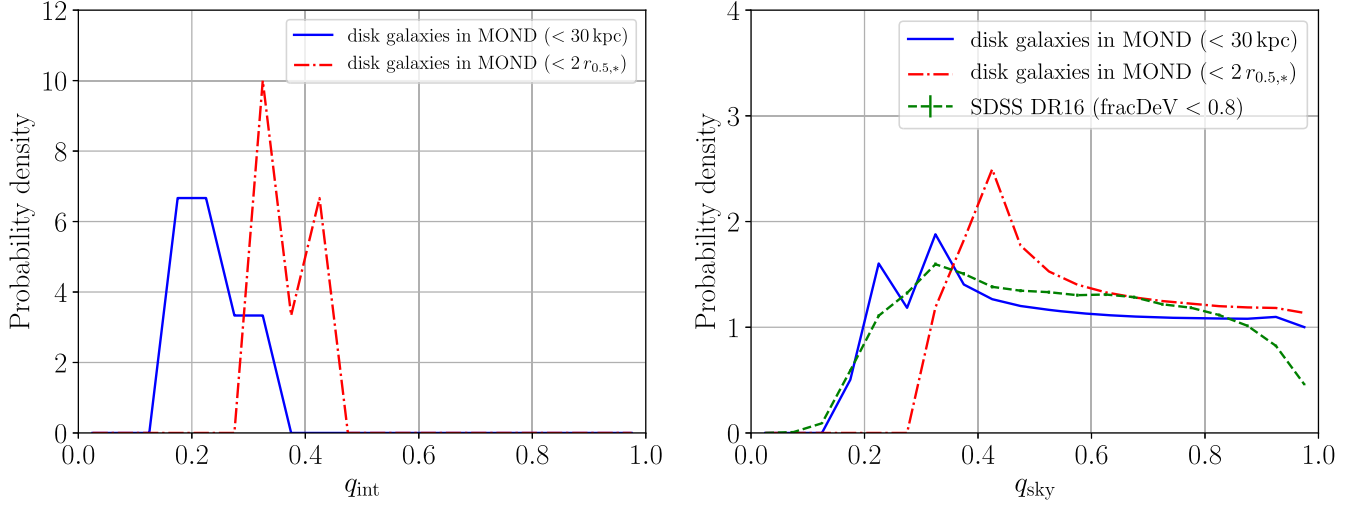


Figure 14. Distribution of the intrinsic (left) and sky-projected (right) aspect ratio for six disk galaxies with $10^{10} < M_*/M_\odot \leq 9.56 \times 10^{10}$ formed in MOND simulations conducted by Wittenburg et al. (2020). The solid blue (dotted–dashed red) line refers to the aspect ratio derived from the mass tensor of particles within a sphere of radius 30 kpc ($2 r_{0.5,*}$). These MOND results are shown for illustrative purposes only—they are not directly comparable to the Λ CDM simulations because the Milgromian galaxies are not formed in a self-consistent cosmological simulation. The dashed green line in the right panel shows the q_{sky} distribution derived from an exponential fit to SDSS galaxies with $\text{fracDeV} < 0.8$ and $10^{10} < M_*/M_\odot \leq 9.56 \times 10^{10}$. We do not apply an M_* -weighting to the observed sample as we only have a small number of simulated MONDian galaxies.

OF RAMSES code (Lüghausen et al. 2015), an adaptation of the N -body and hydrodynamics solver RAMSES (Teyssier 2002) for MOND gravity (for a user guide, see Nagesh et al. 2021). These non-cosmological simulations have a box size of 960 kpc per side, a minimum stellar mass of $M_* \approx 3 \times 10^4 M_\odot$, and depending on the run a maximum grid cell resolution of 117.19 pc, 234.38 pc, or 468.75 pc. The temperature floor for the gas is set to 10 kK (for further information on the initial conditions and numerical parameters of the individual simulations, see Table 1 of Wittenburg et al. 2020). The Milgromian disk galaxies are systematically thinner than in the here analyzed Λ CDM simulations, as evidenced by their $q_{\text{int}} = 0.18\text{--}0.30$ (0.31–0.54) for $r < 30$ kpc ($r < 2 r_{0.5,*}$), with the global peak at $q_{\text{int}} \approx 0.2$ (0.3). While these isolated MOND results cannot yet be directly compared with self-consistent

cosmological Λ CDM simulations that allow for galaxy interactions and mergers, we note that neglecting mergers may be a good approximation in MOND as mergers are expected to be rare due to the absence of dynamical friction with the dark matter halo (Renaud et al. 2016).

Interestingly, the right panel of Figure 14 shows that the sky-projected aspect ratio distribution of these MONDian disk galaxies is very consistent with that of SDSS spiral galaxies ($\text{fracDeV} < 0.8$; see Section 4.3). This also demonstrates that the distinction of spiral from elliptical galaxies based on the linear combination of the exponential and de Vaucouleurs models (Abazajian et al. 2004) succeeds in SDSS, in contrast to the bulge-disk decomposition applied to the Illustris-1 simulation (possible reasons were discussed in Section 7.2). In the future, whether a self-consistent cosmological MOND simulation would

be able to reproduce the observed fraction of spiral and elliptical galaxies needs to be explicitly shown. Such MOND simulations are not available at the moment, but are underway in the Bonn-Prague group based on the promising cosmological MOND framework detailed in Haslbauer et al. (2020). Given the problems in reproducing the observed distribution of galactic morphologies with Λ CDM simulations, it is often argued that these simulations depend sensitively on the implemented feedback model and that the problem is likely to be alleviated once the correct feedback implementation has been found. Hydrodynamical MOND simulations of galaxy formation out of post-Big Bang gas clouds, on the other hand, naturally lead to realistic galaxies similar to the observed ones with the available feedback implementations (Wittenburg et al. 2020, Eappen et al. 2022, submitted).

8. Conclusions

In this contribution, we considered the distribution of galaxy morphologies in state-of-the-art cosmological Λ CDM simulations. The present-day sky-projected aspect ratio distribution of galaxies in the TNG50-1 (EAGLE50) simulation disagrees with the GAMA survey and SDSS at $\geq 12.52\sigma$ ($\geq 14.82\sigma$) confidence (Section 6.2). The lowest tension is obtained when comparing GAMA with the TNG50-1 simulation run, which has the highest resolution of the TNG project. The main reason for this mismatch is that the Λ CDM simulations significantly underproduce galaxies with $q_{\text{sky}} < 0.4$ (Table 2), making it difficult for the latest Λ CDM simulations to form thin disk galaxies like the Milky Way (with a ratio of scale height to scale length of $h/l \approx 0.07\text{--}0.21$; Bland-Hawthorn & Gerhard 2016) or M31 (with a sky-projected aspect ratio of 0.27 ± 0.03 ; Courteau et al. 2011). The intrinsic aspect ratio distribution of the highest-resolution models conflicts with the best-observed local galaxy sample at 5.42σ significance if we use the recently released TNG50-1 run (Section 6.1), confirming this discrepancy independently of the GAMA and SDSS data sets. The advantage of the LV sample is the higher spatial resolution of the galaxy images compared to SAMI, GAMA, and SDSS.

Our results agree with other recent studies (e.g., Lagos et al. 2018; van de Sande et al. 2019; Peebles 2020). We therefore disagree that the angular momentum problem has been resolved in the Illustris-1 simulation as concluded by Vogelsberger et al. (2014). The loss of angular momentum remains a significant problem in the latest cosmological Λ CDM simulations, as quantified in the present work.

The aspect ratio distribution has numerically converged in the EAGLE simulations (Section 7.1; see also Lagos et al. 2018). However, convergence cannot be confirmed for the TNG50 simulation, where the tension reported here could be related to numerical heating of the stellar particles by the coarse-grained implementation of dark matter halos (Ludlow et al. 2021). Therefore, we apply a parametric correction to the q_{int} of each simulated galaxy in TNG50 (Equation (12)). Depending on the extrapolation, we estimate that galaxies in an eight-times-higher particle resolution realization compared to the TNG50-1 run would be thinned by a factor in the range $0.824 \leq \alpha \leq 0.834$, where $\alpha = 1.0$ would imply numerical convergence in q_{int} . Applying the lower limit and being therewith more conservative with respect to the Λ CDM framework, the here reported tension would decrease to the 8.68σ (8.71σ) confidence level for GAMA DR3 (SDSS DR16).

Extrapolating the TNG50 results to $8^5 \times$ better dark matter mass resolution than TNG50-1, the tension with GAMA DR3 (SDSS DR16) becomes 5.58σ (3.27σ) for the parabolic extrapolation, which better fits the available data (the tension is slightly lower with a linear extrapolation; see Table 5). However, such an extrapolation to much higher resolution than the TNG50 runs introduces additional uncertainties, so it is not yet clear if an arbitrarily higher-resolution realization than TNG50-1 can indeed resolve the here reported tension.

Thus, additional self-consistent cosmological Λ CDM simulations are useful to test if higher-resolution realizations and further improved sub-grid models for the interstellar medium can resolve the tension (e.g., Trayford et al. 2017; Lagos et al. 2018; van de Sande et al. 2019). This long-standing problem is likely not caused by limitations of the sub-grid model because the latest EAGLE and TNG simulations that are based on very different computational algorithms and baryonic feedback coding lead to galaxy populations whose sky-projected aspect ratio distributions agree with each other (Table 4). Moreover, isolated CDM simulations with a similar temperature floor of about 10 kK in the gas are able to produce thin disk galaxies (Sellwood et al. 2019)—as indeed are EAGLE and TNG50.

An uncertainty that has not been elaborated on in our work is how observed aspect ratios could be affected by the difference in stellar ages between the typically older bulges in spiral galaxies and their typically younger disks, which thus end up with a lower mass-to-light ratio. This could cause a difference between the mass-weighted aspect ratios obtained from simulations and the luminosity-weighted aspect ratios obtained from observations. However, the high fraction of bulgeless disk galaxies locally (Kormendy et al. 2010) suggests that this issue is not by itself sufficient to reconcile Λ CDM with the observed galaxy population. Observational systematics can be expected to differ between, e.g., GAMA and the high-resolution LV observations (Section 4.4), but even here there is still a significant 5.42σ tension with TNG50.

It appears to be impossible to form as many thin disk galaxies as observed. Consequently, the angular momentum problem persists in the hierarchical cosmological Λ CDM framework and is unlikely to be solved by improving the resolution. This conclusion can also be reached independently through observed galaxy bars being fast with no sign of slowdown through Chandrasekhar dynamical friction on the hypothetical dark matter halo (Roshan et al. 2021b). Λ CDM also faces many problems on other scales (Kroupa 2012, 2015; Pawlowski 2021; Banik & Zhao 2022). Almost 50 yr after dark matter halos were first postulated to surround galaxies (Ostriker & Peebles 1973), numerical implementations of this model still cannot explain the observed fraction of early- and late-type galaxies, so the observed galaxy population continues to pose a severe challenge to this framework.

If better resolved and/or improved sub-grid models of the interstellar medium in self-consistent Λ CDM cosmological simulations cannot resolve this tension, the loss of angular momentum would question the hierarchical merger-driven build-up of galaxies in this paradigm. We showed that intrinsically thin galaxies are more frequent in a Λ CDM galaxy sample selected to have a very quiescent merger history compared to one with an active merger history. However, a sample with a quiescent merger history cannot resolve the tension.

This leaves open the possibility that the tension we reported here is due to minor mergers and/or disk-halo angular momentum exchange (see also Sellwood et al. 2019; Banik et al. 2020; Roshan et al. 2021a, 2021b). If so, the angular momentum problem might be alleviated in a cosmological MOND framework (Milgrom 1983) due to the absence of dynamical friction on extended dark matter halos reducing the merger rate (Kroupa 2015; Renaud et al. 2016). In MOND, cluster-scale and early-universe observables can be explained within the neutrino hot dark matter (ν HDM) model (Angus 2009; Haslbauer et al. 2020; Asencio et al. 2021). In particular, we emphasize that the statistically significant Hubble tension does not appear in this model (Haslbauer et al. 2020) but does exist in the standard Λ CDM model, independently showing Λ CDM to be invalid (Riess et al. 2021). Thus, the Hubble tension is naturally solved by using the standard MOND cosmological model without tuning any theoretical parameter. The sky-projected aspect ratio distribution of disk galaxies formed in hydrodynamical MOND simulations of collapsing post-Big Bang gas clouds (Wittenburg et al. 2020) is very consistent with observed SDSS spiral galaxies (Section 7.5). Self-consistent cosmological MOND simulations underway in Bonn and in Prague will allow us to determine the entire galactic morphological distribution for comparison with observations, enabling the same tests as documented here for Λ CDM.

I.B. is supported by Science and Technology Facilities Council grant ST/V000861/1. He acknowledges support from an Alexander von Humboldt Foundation postdoctoral research fellowship (2018–2020) and the University of Bonn “Pathways to Research” program. The authors are grateful to Veselina Kalinova for valuable comments, Karl Menten for his support, Sylvia Plöckinger for assistance with the EAGLE database, Sree Oh for explanations on the Data Central platform (<https://datacentral.org.au/>), and Edward Taylor for clarifications on the stellar masses of GAMA DR3 galaxies. The authors are very grateful to the anonymous referee for helpful suggestions that significantly improved this publication. They also thank the GAMA and SDSS teams for providing their data and useful related discussions.

GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalog is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT, and ASKAP, providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is <http://www.gama-survey.org/>.

The SAMI Galaxy Survey is based on observations made at the Anglo-Australian Telescope. The Sydney-AAO Multi-object Integral field spectrograph (SAMI) was developed jointly by the University of Sydney and the Australian Astronomical Observatory. The SAMI input catalog is based on data taken from the Sloan Digital Sky Survey, the GAMA Survey, and the VST ATLAS Survey. The SAMI Galaxy Survey is funded by the Australian Research Council Centre of Excellence for All-sky Astrophysics (CAASTRO), through

project No. CE110001020, and other participating institutions. The SAMI Galaxy Survey website is <http://sami-survey.org/>.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington. The SDSS website is <http://www.sdss.org/>.

Appendix

Statistical Significance of Extreme Events

If the χ^2 value is particularly large, the P -value is too low for a finite element computer to handle. We therefore make an analytic approximation for the integrals of the χ^2 and Gaussian distributions. In both cases, as the integrand declines very rapidly, we locally approximate it as declining exponentially. This allows us to approximate the integral out to infinity. Using this approach, we need to solve for x based on the known value of χ^2 using

$$\sqrt{\frac{2}{\pi}} \frac{\exp(-x^2)}{x} = \frac{(\chi^2)^{\frac{n}{2}-1} \exp\left(-\frac{\chi^2}{2}\right)}{2^{\frac{n}{2}} \left(\frac{n}{2}-1\right)! \left(\frac{1}{2} - \frac{\frac{n}{2}-1}{\chi^2}\right)}. \quad (\text{A1})$$

Factorials of non-integer numbers are defined using the Γ function.

To minimize numerical errors, we set up Equation (A1) as an equality between the logarithms of both sides. We then solve for x using the Newton-Raphson algorithm. The statistical significance of the result is approximately x standard deviations, with the approximation becoming very accurate for $x > 7$. This is fortunate, as lower values of x allow Equation (11) to be solved directly without the approximation of Equation (A1). We checked that both methods give similar results for $x = 5-7$, but numerical difficulties mean that Equation (A1) is necessary for higher x .

ORCID iDs

Moritz Haslbauer  <https://orcid.org/0000-0002-5101-6366>
 Indranil Banik  <https://orcid.org/0000-0002-4123-7325>
 Pavel Kroupa  <https://orcid.org/0000-0002-7301-3377>
 Nils Wittenburg  <https://orcid.org/0000-0001-9332-0000>
 Behnam Javanmardi  <https://orcid.org/0000-0002-9317-6114>

References

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2004, *AJ*, **128**, 502
- Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2008, *ApJS*, **175**, 297
- Ahumada, R., Prieto, C. A., Almeida, A., et al. 2020, *ApJS*, **249**, 3
- Angus, G. W. 2009, *MNRAS*, **394**, 527
- Asencio, E., Banik, I., & Kroupa, P. 2021, *MNRAS*, **500**, 5249
- Athanassoula, E. 2002, *ApJL*, **569**, L83
- Bahé, Y. M., Barnes, D. J., Dalla Vecchia, C., et al. 2017, *MNRAS*, **470**, 4186
- Baldry, I. K., Liske, J., Brown, M. J. I., et al. 2018, *MNRAS*, **474**, 3875
- Banik, I., Thies, I., Candlish, G., et al. 2020, *ApJ*, **905**, 135
- Banik, I., & Zhao, H. 2018, *MNRAS*, **473**, 4033
- Banik, I., & Zhao, H. 2022, arXiv:2110.06936
- Barnes, D. J., Kay, S. T., Bahé, Y. M., et al. 2017, *MNRAS*, **471**, 1088
- Binney, J., & Tremaine, S. 2008, *Galactic Dynamics* (2nd ed.; Princeton, NJ: Princeton Univ. Press)
- Bland-Hawthorn, J., & Gerhard, O. 2016, *ARA&A*, **54**, 529
- Bottrell, C., Torrey, P., Simard, L., & Ellison, S. L. 2017a, *MNRAS*, **467**, 1033
- Bottrell, C., Torrey, P., Simard, L., & Ellison, S. L. 2017b, *MNRAS*, **467**, 2879
- Bournaud, F., Jog, C. J., & Combes, F. 2005, *A&A*, **437**, 69
- Brooks, A. M., Kuhlen, M., Zolotov, A., & Hooper, D. 2013, *ApJ*, **765**, 22
- Bryant, J. J., Owers, M. S., Robotham, A. S. G., et al. 2015, *MNRAS*, **447**, 2857
- Chambers, K. C., Magnier, E. A., Metcalfe, N., et al. 2016, arXiv:1612.05560
- Courteau, S., Widrow, L. M., McDonald, M., et al. 2011, *ApJ*, **739**, 20
- Crain, R. A., Schaye, J., Bower, R. G., et al. 2015, *MNRAS*, **450**, 1937
- Croom, S. M., Lawrence, J. S., Bland-Hawthorn, J., et al. 2012, *MNRAS*, **421**, 872
- Croom, S. M., Owers, M. S., Scott, N., et al. 2021, *MNRAS*, **505**, 991
- de Vaucouleurs, G. 1948, *AnAp*, **11**, 247
- Delgado-Serrano, R., Hammer, F., Yang, Y. B., et al. 2010, *A&A*, **509**, A78
- Disney, M. J., Romano, J. D., Garcia-Appadoo, D. A., et al. 2008, *Natur*, **455**, 1082
- Dolag, K., Komatsu, E., & Sunyaev, R. 2016, *MNRAS*, **463**, 1797
- Dolag, K., Mevius, E., & Remus, R.-S. 2017, *Galax*, **5**, 35
- D’Onghia, E., Burkert, A., Murante, G., & Khochfar, S. 2006, *MNRAS*, **372**, 1525
- Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, *MNRAS*, **413**, 971
- Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, *A&G*, **50**, 5.12
- Dubois, Y., Pichon, C., Welker, C., et al. 2014, *MNRAS*, **444**, 1453
- Efstathiou, G., Sutherland, W. J., & Maddox, S. J. 1990, *Natur*, **348**, 705
- Emsellem, E., Cappellari, M., Krajnović, D., et al. 2011, *MNRAS*, **414**, 888
- Fakhouri, O., Ma, C.-P., & Boylan-Kolchin, M. 2010, *MNRAS*, **406**, 2267
- Famaey, B., & McGaugh, S. S. 2012, *LRR*, **15**, 10
- Fletcher, R., & Powell, M. J. D. 1963, *CompJ*, **6**, 163
- Genel, S., Fall, S. M., Hernquist, L., et al. 2015, *ApJL*, **804**, L40
- Giavalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, **600**, L93
- Graham, A. W., & Worley, C. C. 2008, *MNRAS*, **388**, 1708
- Grazian, A., Fontana, A., de Santis, C., et al. 2006, *A&A*, **449**, 951
- Guedes, J., Callegari, S., Madau, P., & Mayer, L. 2011, *ApJ*, **742**, 76
- Haslbauer, M., Banik, I., & Kroupa, P. 2020, *MNRAS*, **499**, 2845
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *ApJS*, **208**, 19
- Hirschmann, M., Dolag, K., Saro, A., et al. 2014, *MNRAS*, **442**, 2304
- Hoffmann, K., Laigle, C., Chisari, N. E., et al. 2020, arXiv:2010.13845
- Hwang, J.-S., Park, C., Nam, S.-h., & Chung, H. 2021, *JKAS*, **54**, 71
- Karachentsev, I. D., Kaisina, E. I., & Makarov, D. I. 2018, *MNRAS*, **479**, 4136
- Karachentsev, I. D., Karachentseva, V. E., Huchtmeier, W. K., & Makarov, D. I. 2004, *AJ*, **127**, 2031
- Karachentsev, I. D., Makarov, D. I., & Kaisina, E. I. 2013, *AJ*, **145**, 101
- Katz, N., & Gunn, J. E. 1991, *ApJ*, **377**, 365
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, **341**, 33
- Kautsch, S. J., Grebel, E. K., Barazza, F. D., & Gallagher, J. S. I. 2006, *A&A*, **445**, 765
- Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2012, *MNRAS*, **421**, 1007
- Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2014, *MNRAS*, **444**, 1647
- Kormendy, J., Drory, N., Bender, R., & Cornell, M. E. 2010, *ApJ*, **723**, 54
- Kroupa, P. 2012, *PASA*, **29**, 395
- Kroupa, P. 2015, *CaJPh*, **93**, 169
- Kroupa, P., Subr, L., Jerabkova, T., & Wang, L. 2020, *MNRAS*, **498**, 5652
- Lagos, C. D. P., Schaye, J., Bahé, Y., et al. 2018, *MNRAS*, **476**, 4327
- Lange, R., Moffett, A. J., Driver, S. P., et al. 2016, *MNRAS*, **462**, 1470
- Loveday, J. 1996, *MNRAS*, **278**, 1025
- Ludlow, A. D., Fall, S. M., Schaye, J., & Obreschkow, D. 2021, *MNRAS*, **508**, 5114
- Lüghausen, F., Famaey, B., & Kroupa, P. 2015, *CaJPh*, **93**, 232
- McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, *A&C*, **15**, 72
- McGaugh, S. S., & Schombert, J. M. 2014, *AJ*, **148**, 77
- Milgrom, M. 1983, *ApJ*, **270**, 365
- Mosenkov, A. V., Sotnikova, N. Y., & Reshetnikov, V. P. 2010, *MNRAS*, **401**, 559
- Müller, O., Lelli, F., Famaey, B., et al. 2021, *A&A*, in press (arXiv:2111.10306)
- Nagesh, S. T., Banik, I., Thies, I., et al. 2021, *CaJPh*, **99**, 607
- Navarro, J. F., & Benz, W. 1991, *ApJ*, **380**, 320
- Navarro, J. F., & Steinmetz, M. 2000, *ApJ*, **538**, 477
- Navarro, J. F., & White, S. D. M. 1994, *MNRAS*, **267**, 401
- Nelson, D., Pillepich, A., Genel, S., et al. 2015, *A&C*, **13**, 12
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *ComAC*, **6**, 2
- Oh, S., Colless, M., Barsanti, S., et al. 2020, *MNRAS*, **495**, 4638
- Ostriker, J. P., & Peebles, P. J. E. 1973, *ApJ*, **186**, 467
- Ostriker, J. P., & Steinhardt, P. J. 1995, *Natur*, **377**, 600
- Owers, M. S., Allen, J. T., Baldry, I., et al. 2017, *MNRAS*, **468**, 1824
- Owers, M. S., Hudson, M. J., Oman, K. A., et al. 2019, *ApJ*, **873**, 52
- Padilla, N. D., & Strauss, M. A. 2008, *MNRAS*, **388**, 1321
- Patuel, G., Andernach, H., Bottinelli, L., et al. 1997, *A&AS*, **124**, 109
- Pawlowski, M. S. 2021, *NatAs*, **5**, 1185
- Peebles, P. J. E. 2020, *MNRAS*, **498**, 4386
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, **124**, 266
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2010, *AJ*, **139**, 2097
- Pillepich, A., Nelson, D., Springel, V., et al. 2019, *MNRAS*, **490**, 3196
- Pillepich, A., Springel, V., Nelson, D., et al. 2018, *MNRAS*, **473**, 4077
- Piontek, F., & Steinmetz, M. 2011, *MNRAS*, **410**, 2625
- Planck Collaboration I 2014, *A&A*, **571**, A1
- Planck Collaboration VI 2020, *A&A*, **641**, A6
- Planck Collaboration XIII 2016, *A&A*, **594**, A13
- Reddish, J., Kraljic, K., Petersen, M. S., et al. 2021, arXiv:2106.02622
- Renaud, F., Ho, L. C., Impey, C. D., & Rix, H.-W. 2016, *MNRAS*, **463**, 3637
- Riess, A. G., Yuan, W., Macri, L. M., et al. 2021, arXiv:2112.04510
- Rodríguez-Gomez, V., Genel, S., Vogelsberger, M., et al. 2015, *MNRAS*, **449**, 49
- Rodríguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, *MNRAS*, **483**, 4140
- Roshan, M., Banik, I., Ghafourian, N., et al. 2021a, *MNRAS*, **503**, 2833
- Roshan, M., Ghafourian, N., Kashfi, T., et al. 2021b, *MNRAS*, **508**, 926
- Salim, S., Rich, R. M., Charlot, S., et al. 2007, *ApJS*, **173**, 267
- Santini, P., Fontana, A., Grazian, A., et al. 2009, *A&A*, **504**, 751
- Scannapieco, C., Wadepuhl, M., Parry, O. H., et al. 2012, *MNRAS*, **423**, 1726
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, **446**, 521
- SDSS Collaboration 2000, *AJ*, **120**, 1579
- Sellwood, J. A., Shen, J., & Li, Z. 2019, *MNRAS*, **486**, 4710
- Springel, V. 2005, *MNRAS*, **364**, 1105
- Springel, V. 2010, *MNRAS*, **401**, 791
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, *Natur*, **435**, 629
- Stewart, K. R., Bullock, J. S., Wechsler, R. H., Maller, A. H., & Zentner, A. R. 2008, *ApJ*, **683**, 597
- Tamburri, S., Saracco, P., Longhetti, M., et al. 2014, *A&A*, **570**, A102
- Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, *MNRAS*, **418**, 1587
- Teyssier, R. 2002, *A&A*, **385**, 337
- Thob, A. C. R., Crain, R. A., McCarthy, I. G., et al. 2019, *MNRAS*, **485**, 972
- Torrey, P., Snyder, G. F., Vogelsberger, M., et al. 2015, *MNRAS*, **447**, 2753
- Trayford, J. W., Camps, P., Theuns, T., et al. 2017, *MNRAS*, **470**, 771
- van de Sande, J., Lagos, C. D. P., Welker, C., et al. 2019, *MNRAS*, **484**, 869
- van den Bosch, F. C. 2001, *MNRAS*, **327**, 1334
- Vázquez-Mata, J. A., Loveday, J., Riggs, S. D., et al. 2020, *MNRAS*, **499**, 631
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *Natur*, **509**, 177
- Wetzel, A. R., Hopkins, P. F., Kim, J.-H., et al. 2016, *ApJL*, **827**, L23
- Wittenburg, N., Kroupa, P., & Famaey, B. 2020, *ApJ*, **890**, 173
- Wu, X., & Kroupa, P. 2015, *MNRAS*, **446**, 330