Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **510,** 3967–3985 (2022) https://doi.org/10.1093/mnras/stab3625
Advance Access publication 2021 December 15

The galaxy–halo size relation of low-mass galaxies in FIRE

Eric Rohr [1]★ Robert Feldmann [2] James S. Bullock [3] Onur Çatmabacak [2] Michael Boylan-Kolchin [4] Claude-André Faucher-Giguère [5] Dušan Kereš,[6] Lichen Liang [2,7] Jorge Moreno [8] and Andrew Wetzel [9]

[1]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*
[2]*Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, Zurich CH-8057, Switzerland*
[3]*Department of Physics and Astronomy, 4129 Reines Hall, University of California, Irvine, CA 92697, USA*
[4]*Department of Astronomy, The University of Texas at Austin, 2515 Speedway, Stop C1400, Austin, TX 78712, USA*
[5]*Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208, USA*
[6]*Department of Physics, Center for Astrophysics and Space Sciences, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA*
[7]*Canadian Institute for Theoretical Astrophysics, 60 St. George Street, Toronto, ON M5S 3H8, Canada*
[8]*Department of Physics and Astronomy, Pomona College, Claremont, CA 91711, USA*
[9]*Department of Physics and Astronomy, University of California, Davis, CA 95616, USA*

Accepted 2021 December 9. Received 2021 December 9; in original form 2021 July 16

ABSTRACT

Galaxy sizes correlate closely with the sizes of their parent dark matter haloes, suggesting a link between halo formation and galaxy growth. However, the precise nature of this relation and its scatter remains to be understood fully, especially for low-mass galaxies. We analyse the galaxy–halo size relation (GHSR) for low-mass ($M_\star \sim 10^{7-9}\,\mathrm{M}_\odot$) central galaxies over the past 12.5 billion years with the help of cosmological volume simulations (FIREbox) from the Feedback in Realistic Environments (FIRE) project. We find a nearly linear relationship between the half-stellar mass galaxy size $R_{1/2}$ and the parent dark matter halo virial radius $R_{\rm vir}$. This relation evolves only weakly since redshift $z = 5$: $R_{1/2}\,[\mathrm{kpc}] = (0.053 \pm 0.002)(R_{\rm vir}/35\,\mathrm{kpc})^{0.934\pm0.054}$, with a nearly constant scatter $\langle\sigma\rangle = 0.084\,[\mathrm{dex}]$. While this ratio is similar to what is expected from models where galaxy disc sizes are set by halo angular momentum, the low-mass galaxies in our sample are not angular momentum supported, with stellar rotational to circular velocity ratios $v_{\rm rot}/v_{\rm circ} \sim 0.15$. Introducing redshift as another parameter to the GHSR does not decrease the scatter. Furthermore, this scatter does not correlate with any of the halo properties we investigate – including spin and concentration – suggesting that baryonic processes and feedback physics are instead critical in setting the scatter in the GHSR. Given the relatively small scatter and the weak dependence of the GHSR on redshift and halo properties for these low-mass central galaxies, we propose using galaxy sizes as an independent method from stellar masses to infer halo masses.

Key words: methods: numerical – galaxies: dwarf – galaxies: evolution – galaxies: haloes – cosmology: theory.

1 INTRODUCTION

According to the standard picture of galaxy formation, galaxies form at the centres of their parent dark matter haloes (White & Rees 1978; Cole et al. 2000). Haloes tend to form bottom-up, with small haloes collapsing first and subsequently merging into larger and more massive haloes (Blumenthal et al. 1984). This gravity-driven hierarchical picture of halo collapse, however, contrasts with a much more complex picture of galaxy formation, which is set by a variety of baryonic processes (e.g. Somerville & Davé 2014, and references therein). The link between galaxies and their parent haloes is therefore far from trivial, especially given that the latter are more than order of magnitude larger and more massive than the former.

None the less, central galaxies and their parent haloes do appear to be linked tightly, at least in the local Universe (see Wechsler & Tinker 2018, for a recent review). For instance, the stellar-to-halo-mass relation, also called the $M_\star - M_{\rm halo}$ relation, has been widely studied empirically (via abundance matching, Kravtsov & Klypin 1999; Kravtsov et al. 2004; Conroy & Wechsler 2009; Behroozi, Conroy & Wechsler 2010), observationally via – e.g. gravitational lensing (Mandelbaum et al. 2006; Behroozi et al. 2019) or galaxy kinematics (More et al. 2011; Kravtsov, Vikhlinin & Meshcheryakov 2018), and theoretically, with the help of analytic models (e.g. White & Frank 1991; Mo, Mao & White 1998; Wechsler et al. 1998; White et al. 2007) and cosmological simulations (Pearce et al. 2001; Berlind et al. 2003; Simha et al. 2012; Hopkins et al. 2014; Khandai et al. 2015; Chaves-Montero et al. 2016; Feldmann et al. 2016; McAlpine et al. 2016; Pillepich et al. 2018b; Engler et al. 2020). This relation forms a broken power law of increasing galaxy formation efficiency with halo mass until a peak at $M_{\rm halo} \sim 10^{12}\,\mathrm{M}_\odot$ and a sharp decrease towards higher halo masses (Wechsler & Tinker 2018). Interestingly, this relation has a nearly constant scatter over five orders of magnitude in halo mass ($M_{\rm halo} \sim 10^{10-15}\,\mathrm{M}_\odot$), and the baryonic feedback processes – such as stellar winds, supernovae, cosmic rays, and active galactic nuclei – largely determine the shape

★ E-mail: rohr@mpia.de

© 2021 The Author(s)
Published by Oxford University Press on behalf of Royal Astronomical Society

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

of this relation despite the haloes dominating in size and mass (e.g. Dekel & Silk 1986; Silk & Rees 1998; Bullock, Kravtsov & Weinberg 2000; Croton et al. 2006; Hopkins et al. 2012b).

More recently, Kravtsov (2013) discovered a tight connection between the sizes of galaxies and haloes at $z = 0$. Unlike the SHMR, this galaxy–halo size relation (GHSR) forms an approximately linear relation with constant scatter across nearly three orders of magnitude in halo size ($R_{\rm halo} \sim 5$–1500 kpc, corresponding to $M_{\rm halo} \sim 10^{8-15}\,{\rm M}_{\odot}$). Furthermore, this relationship is largely independent of galaxy morphology and nearly identical for centrals and satellites. The observation of the GHSR at $z = 0$ thus raises a variety of intriguing questions such as: When was the GHSR established? How do its slope, normalization, and scatter evolve over cosmic time? Are the GHSR and its scatter linked to the growth histories of structure in the Universe?

However, observing this galaxy–halo connection at high redshift poses a number of challenges. For instance, measuring galaxy and halo sizes with sufficiently high spatial resolution becomes increasingly more challenging at higher redshift. Fortunately, a variety of approaches have been developed to probe the galaxy–halo link at earlier cosmic times. For instance, di Teodoro & Fraternali (2015) show that fitting a 3D-tilted ring models can resolve intrinsic rotation curves and velocity dispersion from low spatial resolution observations. Other recent works have recovered connections between galaxies and their host haloes at high redshifts (Shibuya, Ouchi & Harikane 2015; Huang et al. 2017; Hirtenstein et al. 2019; Zanisi et al. 2020, 2021a). But the current observations are limited to the most massive galaxies, leaving the low-mass regime uncertain.

Another approach is to simulate galaxy formation in a cosmological context to study the galaxy–halo connection across cosmic history (see Somerville & Davé 2014, for a recent review). Over the past decade, advances in computational power and numerical techniques (e.g. Springel 2005, 2010; Hopkins 2015) have made it possible to simulate not only the gravitational collapse of dark matter but also the hydrodynamics of gas, and the complex baryonic processes such as gas cooling, star formation, and stellar feedback (e.g. Vogelsberger et al. 2014; Crain et al. 2015; Schaye et al. 2015; Tremmel et al. 2017; Hopkins et al. 2018; Nelson et al. 2019; Pillepich et al. 2018a). Recent simulations provide a deeper understanding of, and challenge long-standing assumptions about, galaxy formation and evolution. Perhaps contrary to the expectation based on specific angular momentum conservation (Fall & Efstathiou 1980; Mo et al. 1998), Desmond et al. (2017) find in the EAGLE simulation that at a fixed stellar mass, the galaxy size weakly correlates with halo mass, concentration, or spin. On top of that, Somerville et al. (2018) conclude from their sample of Galaxy And Mass Assembly (GAMA) and the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS) mapped to the Bolshoi–Planck dissipationless N-body simulation that the ratio of galaxy-to-halo size decreases slightly with cosmic time for less massive galaxies, while the ratio of galaxy size to halo size times halo spin – $R_{1/2}/(R_{\rm vir}\lambda)$ – is lower for more massive galaxies below $z \lesssim 3$. In the VELA (Ceverino et al. 2014; Zolotov et al. 2015) and NIHAO (Wang et al. 2015) zoom-in simulations, Jiang et al. (2019) find that the halo spin only weakly correlates with that of the galaxy, and this correlation becomes weaker with increasing redshift. Moreover, the gas that builds galaxies in cosmological simulations typically has higher specific angular momentum than that of the dark matter (e.g. Danovich et al. 2015; Stewart et al. 2017; Zjupa & Springel 2017; El-Badry et al. 2018; Kretschmer, Agertz & Teyssier 2020).

In general, these recent results point to differences between halo and galactic properties once thought to be tightly linked. It thus appears that baryonic properties are more significant in setting galaxy sizes. Using zooms of Milky Way mass objects with the same FIRE-2 model, Garrison-Kimmel et al. (2018) find that of their studied parameters, the best predictor of galaxy size and morphology is the gas spin at the time the galaxy formed half of its $z = 0$ stars. Similarly for massive galaxies ($M_{\star} \sim 10^{9-12.5}\,{\rm M}_{\odot}$) in the Sloan Digital Sky Survey Data Release 7, Zanisi et al. (2020) find in their semi-empirical models that the specific stellar angular momentum is the best mediator to the GHSR. Despite the numerous works on potential correlations between halo properties and galaxy sizes (Rodriguez et al. 2021), the GHSR has not been well studied in simulations and observations, in particular, in the regime of low stellar masses and at higher redshifts.

In this paper, we study the GHSR and its scatter for low-mass centrals, as predicted by high-resolution cosmological volume simulations from the FIREbox simulation suite, which is part of the Feedback in Realistic Environments (FIRE)[1] project. FIREbox follows the growth of galaxies and haloes in a $15\,{\rm cMpc}\,h^{-1}$ ($22\,{\rm pMpc}$ at $z = 0$) side-length cosmological box with the help of the FIRE-2 baryonic model (Hopkins et al. 2018) and the Meshless Finite Mass hydrodynamic solver GIZMO[2] (Hopkins 2015). This simulation suite contains a large sample of galaxies ranging from isolated dwarfs to Milky Way (MW) analogues, facilitating an in-depth analysis of the GHSR from $z = 5$ until today.

The layout of this paper is as follows. In Section 2, we outline the methodologies for the numerics of the simulation (Section 2.1), the halo-finding algorithms (Section 2.2), and the sample selection (Section 2.3). Section 3 compares the galaxy/halo pairs to observations and other recent works. Section 4 constructs and details the GHSR from $0 \leq z \leq 5$. Section 5 analyses the scatter in the GHSR and addresses potential halo (Section 5.1), galaxy (Section 5.2), and environment (Section 5.3) properties affecting the $R_{1/2} - R_{\rm vir}$ relation. Specifically, Section 5.1.1 investigates the effects of halo spin and concentration on the GHSR and SHMR in more detail. Lastly, we summarize the major findings in Section 6.

2 METHODOLOGY

2.1 FIREbox simulation suite

The galaxies and haloes analysed in this paper are extracted from the FIREbox suite of $V = (15\,{\rm cMpc}\,h^{-1})^3$ cosmological volume simulations (Feldmann et al. in preparation), which are part of the FIRE project (Hopkins et al. 2014, 2018). Unlike all previous FIRE simulations, FIREbox does not use the zoom-in setup to study galaxy evolution – but instead, it simulates gas, stars, and dark matter in a cubic cosmological volume with periodic boundary conditions. Initial conditions at $z = 120$ were created with MUlti Scale Initial Conditions (MUSIC; Hahn & Abel 2011) using cosmological parameters consistent with Planck 2015 results (Alves et al. 2016): $\Omega_{\rm m} = 0.3089$, $\Omega_{\Lambda} = 1 - \Omega_{\rm m}$, $\Omega_{\rm b} = 0.0486$, $h = 0.6774$, $\sigma_8 = 0.8159$, $n_{\rm s} = 0.9667$, and a transfer function calculated with camb[3] (Lewis, Challinor & Lasenby 2000; Lewis, Challinor & Hanson 2011).

All FIREbox simulations start from the same initial conditions but they differ in particle number, numerical resolution, and whether

[1] https://fire.northwestern.edu/
[2] http://www.tapir.caltech.edu/~phopkins/Site/GIZMO.html
[3] http://camb.info

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

they are run as dark-matter-only (DMO) simulations or include baryonic physics. All simulations are run with GIZMO (Hopkins 2015). Gravitational forces between particles are calculated with a heavily modified version of the parallelization and tree gravity solver of GADGET-3 (Springel 2005) allowing for adaptive force softening, while hydrodynamics is solved with the meshless-finite-mass method introduced in Hopkins (2015). All hydrodynamical FIREbox simulations are run with the FIRE-2 model to account for gas cooling and heating, star formation, and stellar feedback (Hopkins et al. 2018). Feedback from supermassive black holes is not included. Star formation is modelled to occur in dense ($n > 300$ cm^{-3} for the 1024^3 FIREbox simulation; $n > 100$ cm^{-3} and >10 cm^{-3} for the 512^3 and 256^3 simulations, respectively), molecular, self-gravitating gas, and the gas to star conversion takes place on a local free-fall time with a 100 per cent local efficiency. Due to stellar feedback, the realized star formation efficiency is lower, consistent with Kennicutt–Schmidt relations (Schmidt 1959; Kennicutt 1998; Orr et al. 2018). Stellar feedback includes energy, momentum, mass, and metal injections from supernovae (Type II and Type Ia) and stellar winds (OB and AGB stars). Radiative feedback (photo-ionization and photo-electric heating) and radiation pressure from young stars are accounted for in the locally extincted background radiation in optically thin networks (LEBRON) approximation (Hopkins, Quataert & Murray 2012a). The FIRE-2 model has been extensively validated in a number of publications analysing properties of galaxies across a range in stellar masses and numerical resolutions, including simulations at this FIREbox resolution (Wetzel et al. 2016; Hopkins et al. 2018; Ma et al. 2018a, b).

Most of the analysis in this paper is based on the FIREbox pathfinder hydrodynamical simulation ($N_{\rm b} = 1024^3$ and $N_{\rm DM} = 1024^3$). The mass resolution of this run is $m_{\rm b} = 6.3 \times 10^4$ M$_\odot$ for baryonic (gas and star) particles and $m_{\rm DM} = 3.3 \times 10^5$ M$_\odot$ for dark matter particles. The force-softening lengths for star and dark matter particles are $h_\star = 12$ pc (physical) and $h_{\rm DM} = 80$ pc, respectively. The force softening of gas particles is set to their smoothing length down to a minimum of 1.5 pc, which is reached only in the densest parts of the interstellar medium. The force resolution is set such that the highest density we formally resolve is 1000 times the star formation threshold (see Hopkins et al. 2018, Section 2.2 for more details). Mass and force resolution of the $N_{\rm DM} = 512^3$ ($N_{\rm DM} = 256^3$) FIREbox run are correspondingly lower, e.g. $m_{\rm b} \sim 5 \times 10^5$ M$_\odot$ ($m_{\rm b} \sim 4 \times 10^6$ M$_\odot$) and $h_{\rm star} = 32$ pc ($h_{\rm star} = 128$ pc). All FIREbox simulations examined in this paper are evolved to $z = 0$.

2.2 Halo finding and definitions

We employ the AMIGA Halo Finder (AHF)[4] to identify and characterize the properties of dark matter haloes (Knollmann & Knebe 2009). We consider only haloes containing at least 100 particles of any type, which corresponds to a minimum halo mass of $M_{\rm vir} \sim 10^7$ M$_\odot\, h^{-1}$. The halo radius, $R_{\rm vir}$, is defined based on the virial overdensity criterion

$$M_{\rm vir} = (4/3)\pi\, \Delta(z) \rho_m(z) R_{\rm vir}^3, \quad (1)$$

where $\rho_m(z)$ is the matter density at a given redshift z, and $\Delta(z)$ is defined by Bryan & Norman (1998). Halo centres and the centres of their central galaxies are identified as the halo region with the highest total matter density using AHF's maximum-density (MAX) setting.

[4] http://popia.ft.uam.es/AHF/Download.html

Table 1. Selection and convergence criteria of the galaxies from the FIREbox 1024^3 simulation.

z (1)	lg[$R_{\rm vir}$/kpc] (2)	lg[$M_{\rm vir}$/M$_\odot$] (3)	$N_{\rm FB\text{-}S}$ (4)	$N_{\rm small}$ (5)	$N_{\rm FB\text{-}L}$ (6)
0	[1.70, 2.10]	[9.80, 11.00]	826	47	189
1	[1.50, 1.80]	[9.90, 11.25]	1092	88	284
2	[1.30, 1.65]	[9.80, 10.90]	1373	49	210
3	[1.20, 1.55]	[9.85, 10.95]	1160	80	121
4	[1.10, 1.45]	[9.85, 10.90]	801	59	57
5	[1.00, 1.35]	[9.80, 10.85]	488	21	32

Note. (1) Redshift; (2) and (3) ranges of the halo virial radius and mass, respectively, for the 'FB-S' galaxies; (4) number of galaxy/halo pairs included in the analysis; and (5) and (6) numbers of objects excluded because their halo sizes are below and above the virial radius ranges (column 2).

2.3 Sample selection and galaxy definitions

This paper considers only 'central galaxies', galaxies that form in main haloes (i.e. not satellite galaxies nor subhaloes). These galaxies are the most massive within $R_{\rm vir}$ of the host halo and dominate the baryonic processes of the halo. Galaxy sizes and masses are calculated from the `discs` files provided by AHF. We consider the total galaxy radius to be 10 per cent of the halo virial radius (Price et al. 2017). At lower redshifts, Hopkins et al. (2018) and Samuel et al. (2020) find that satellites can exist within $0.1R_{\rm vir}$ for massive galaxies, but this does not affect our sample of low-mass galaxies. When including all stellar material within $0.2R_{\rm vir}$, the same qualitative results hold (see the last paragraph of Section 4 for the results). Additionally, see Appendix A for the results using an iterative $R_{1/2}$ calculation starting from all stars within $0.15R_{\rm vir}$ (Hopkins et al. 2018).

We calculate the galaxy stellar mass $M_\star \equiv M_\star(< 0.1R_{\rm vir})$ by linearly interpolating in log−log space between radii r and the cumulative stellar mass $M_\star(< r)$. Subsequently, the inverse interpolation of $M_\star(< r)$ at $0.5M_\star$ yields the three-dimensional spherical half-stellar mass radius $R_{1/2}$ (see Appendix A for the cumulative stellar mass radial profiles for the 1024^3 'FB-S' galaxies at redshifts $z = 0, 2$). The average stellar particle mass in the 1024^3 simulation is $\sim 3 \times 10^4$ M$_\odot$ at $z = 2$ (roughly half the baryonic mass resolution listed in Section 2.1 due to stellar mass loss). We consider the galaxies resolved when $N_\star(< 0.1R_{\rm vir}) \gtrsim 300$, implying $M_\star > 10^7$ M$_\odot$ in the 1024^3 simulation. We maintain this $N_\star$ criterion for all resolutions, meaning that the $M_\star$ lower limit in lower resolution simulations increases by factors of 8.

Of the resolved galaxies, we study a selected range of halo sizes that depends on redshift and simulation resolution. At each redshift, we define a lower limit on $R_{\rm vir}$ such that the number of haloes in each $R_{\rm vir}$ bin decreases with increasing halo size. This lower limit on the halo size translates to a nearly constant lower limit on the halo mass of $M_{\rm vir} \gtrsim 10^{9.8}$ M$_\odot$ since $z = 5$. This lower limit cuts out a small percentage ($\sim$ 4–8 per cent, see Table 1 column 5) of additional objects at each redshift. When we instead employ a looser criterion of a minimum halo mass of $M_{\rm vir} > 10^9$ M$_\odot$, we exclude only a few additional haloes (<5 at each redshift), and the results remain qualitatively consistent. At large $R_{\rm vir}$, the FIREbox galaxy sizes no longer increase with increasing halo size, which disagrees with observational results at these halo sizes (Kravtsov 2013; Huang et al. 2017; Zanisi et al. 2020). At a fixed halo mass at $z = 0$, these galaxies also have slightly larger stellar masses than is expected from the SHMR (see Feldmann et al., in preparation, for more details). Moreover, these galaxies have more rotational support

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

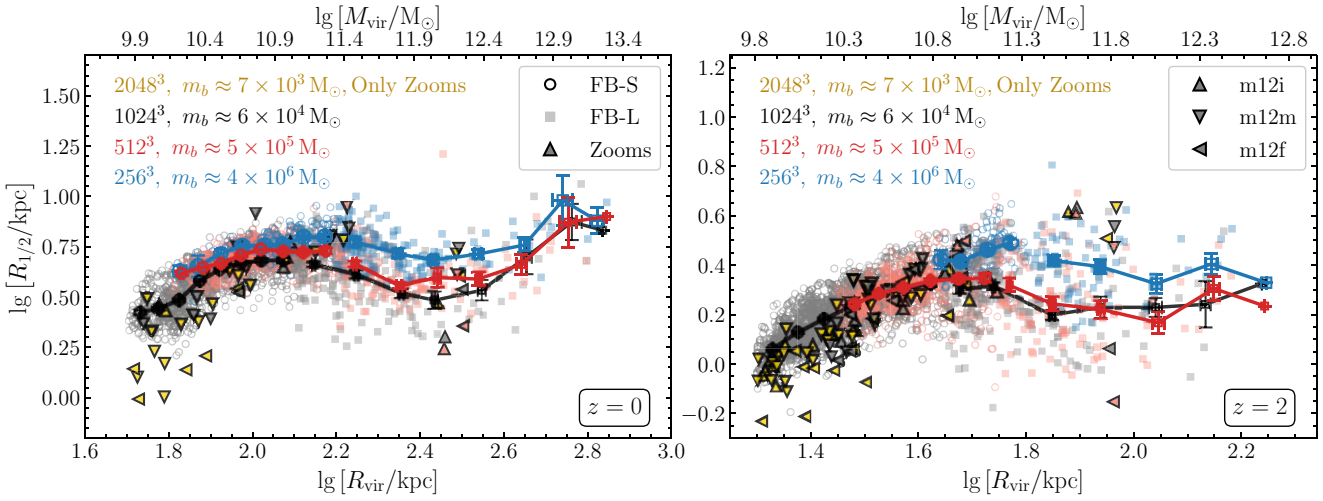

Figure 1. A convergence test of the GHSR at redshift $z = 0$ (left) and $z = 2$ (right), displaying the 256^3 (blue), 512^3 (red), and 1024^3 (black) runs from FIREbox (FB). The open circles represent the galaxies within the selected R_{vir}, denoted 'FB-S,' and the filled squares the galaxies with larger R_{vir}, denoted 'FB-L.' Average sizes within each lg R_{vir} bin are overplotted with standard errors indicated by error bars. We also plot galaxy–halo pairs from FIRE-2 zoom simulations as triangles (Wetzel et al. 2016; Hopkins et al. 2018). These high-resolution (gold triangles, $m_b \approx 7 \times 10^3$ M$_\odot$) zoom-in simulations have an 8 times higher peak resolution than the FIREbox pathfinder (1024^3). Medium resolution zoom-in runs (grey triangles; $m_b \approx 6 \times 10^4$ M$_\odot$) have comparable mass resolution to the 1024^3 run. FIREbox galaxies in moderately massive haloes ($M_{vir} \sim 10^{11-13}$ M$_\odot$ at $z = 0$–2) have approximately constant size with increasing halo size. In this study, we focus on the 'FB-S' galaxies. The figure shows that the $R_{1/2} - R_{vir}$ relation appears converged at both $z = 0$ and $z = 2$ for the FIREbox pathfinder run.

($v_{rot}/v_{circ} \sim 0.4$ at $z = 0$; see Section 5.2 for more details) than the galaxies at smaller R_{vir} ($v_{rot}/v_{circ} \sim 0.1$–$0.2$). Lower-mass galaxies are more dispersion supported both for isolated dwarf galaxies in the Local Group (Wheeler et al. 2017) and in the FIRE-1 zoom simulations (Wheeler et al. 2015, 2017). Hereafter, we denote the galaxies in the large-size regime as 'FB-L', and the fiducial sample within the selected halo size range as 'FB-S'. Due to the differences in the GHSR and amount of dispersion or rotational support between the 'FB-L' and 'FB-S' samples, we focus on the low-mass 'FB-S' galaxies here. We plan to study the rotation-supported, higher-mass 'FB-L' galaxies in more detail in future work. Table 1 summarizes these limits and gives the number of galaxy/halo pairs at each redshift.

Fig. 1 displays the GHSR at redshifts $z = 0$ and $z = 2$ for the 256^3, 512^3, and 1024^3 FIREbox simulations for galaxies residing in both small ('FB-S') and large haloes ('FB-L'). We also compare the predictions by FIREbox with recent FIRE-2 zoom simulations (Wetzel et al. 2016; Hopkins et al. 2018). The latter reach a mass resolution up to ∼8× higher than FIREbox pathfinder run allowing us to check for resolution effects. We also show results from 8 and 64 times lower resolution re-runs of these zoom-ins. The figure shows that the $R_{1/2} - R_{vir}$ relation appears well converged at both $z = 0$ and $z = 2$ for the 1024^3 FIREbox run. However, Hopkins et al. (2018) shows that galaxy properties, including size, may still change when the mass resolution changes from $m_b \approx 6 \times 10^4$ M$_\odot$ to $m_b \approx 7 \times 10^3$ M$_\odot$. Interestingly, the 'FB-L' galaxies show a prominent turnover, i.e. a decrease in galaxy size with increasing halo size for $M_{vir} \sim 10^{11} - 10^{12}$ M$_\odot$ haloes. They also show an increased scatter. A few of the galaxies from the highest resolution zoom simulations have smaller sizes than the 'FB-S' galaxies, which stems from the $M_\star$ criterion. At a fixed R_{vir}, the zoom simulations can resolve galaxies with 8× lower stellar masses than the FIREbox pathfinder simulation. However, the 'FB-S' sample contains many more galaxies larger than the zoom simulations, facilitating statistical analysis across a range of halo sizes.

In summary, the baryonic particles within an AHF-defined dark matter halo must meet these criteria to be included in this analysis:

(i) The host halo must be a main halo, i.e. not a proper subhalo.
(ii) The galaxy must be resolved with $N_\star \gtrsim 300$.
(iii) The halo's virial radius must be in the selected range outlined in Table 1.

3 COMPARISONS TO OBSERVATIONS

In Fig. 2, we compare our galaxy size–stellar mass relation to observations at $z = 2$ from the CANDELS/3D-*HST* (van der Wel et al. 2014) and Cosmic Evolution Survey (COSMOS)/Drift And SHift (DASH) (Mowla et al. 2019b) surveys. We assume that on average for a random projection, the intrinsic three-dimensional half-stellar mass radius $R_{1/2}$ is similar to the projected two-dimensional half-mass radius R_{mass} (van de Ven & van der Wel 2021, see their text for how this approximation varies with intrinsic ellipsoidal axis ratios and for comparisons between $R_{1/2}$ and R_{eff}). We convert R_{mass} to the 2D half-light radius R_{eff} using the empirical fit from Suess et al. (2019a, their table 2 fit using $M_\star$ with a completeness limit of $M_\star > 10^{10}$ M$_\odot$ at $1 < z < 2.5$). Converting between $R_{1/2}$ and R_{eff} typically introduces a smaller scatter than the intrinsic scatter of the galaxy size–stellar mass relation (Price et al. 2017; Genel et al. 2018; van de Ven & van der Wel 2021).

Extrapolating to lower masses where necessary, the 'FB-S' galaxies lie along the extrapolated galaxy size–stellar mass relation for star-forming galaxies, while the 'FB-L' galaxies turn over and lie partly between the relations for star-forming and quiescent galaxies. Recent observational (Lange et al. 2016; Kawinwanichakij et al. 2021; Nedkova et al. 2021; Prole 2021) and numerical (Genel et al. 2018; Sales et al. 2020; Tremmel et al. 2020) studies at $z = 0$ suggest that the relation for star-forming galaxies either continues or becomes flatter in the low-mass regime, while the quiescent relation flattens out for low-mass galaxies. There are no observational studies of

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

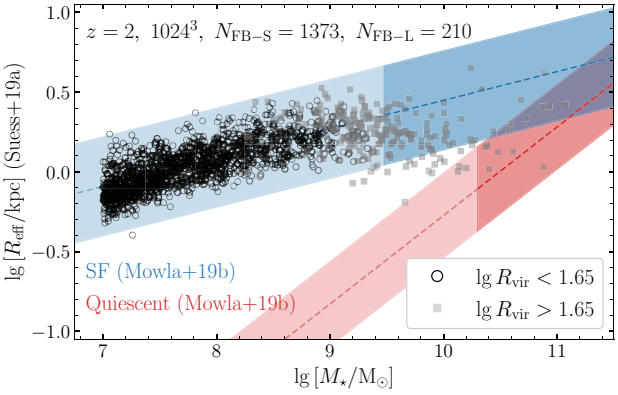

Figure 2. A comparison to the $R_{\rm eff} - M_\star$ relation obtained from CANDELS and COSMOS–DASH (van der Wel et al. 2014; Mowla et al. 2019b). Assuming that our 3D half-stellar mass radius $R_{1/2}$ has on average the same value in 2D projected space (van de Ven & van der Wel 2021), then we convert our $R_{1/2}$ to a projected half-light radius $R_{\rm eff}$ (Suess et al. 2019a, see text for more details). Similarly to Fig. 1, the open circles and filled squares represent the 'FB-S' and 'FB-L' galaxies at $z = 2$, respectively. The dark-shaded areas mark the mass range above the completeness limits of Mowla et al. (2019b), and the lighter regions show the relations extrapolated to lower masses. The completeness limit of Suess et al. (2019a) is $M_\star \gtrsim 10^{10}\,{\rm M}_\odot$. The 'FB-S' galaxies lie along the extrapolated relation for star-forming galaxies.

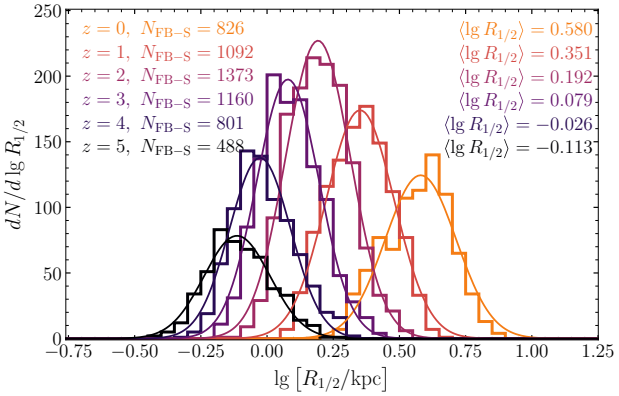

Figure 3. The galaxy size distribution for the 'FB-S' galaxies ($M_\star \sim 10^{7-9}\,{\rm M}_\odot$) at $z = 0 - 5$. The distribution is well described by a lognormal at each redshift. The distribution parameters change with redshift, however. The Gaussian fitted means $\langle \lg[R_{1/2}/{\rm kpc}] \rangle$ are listed in the upper right corner. The average galaxy size increases with cosmic time (decreases with redshift and lookback time), qualitatively agreeing with observations (Shibuya et al. 2015). We find an approximately constant standard deviation $\sigma_{\lg R_{1/2}} = 0.12$ [dex] since $z = 5$, also in qualitative agreement with Shibuya et al. (2015).

dwarf galaxies at $z = 2$. See Appendix B for more details regarding the galaxy size–stellar mass relation, and for the figures at $z = 0$ (Fig. B1). Briefly, at $z = 0$, the FIREbox galaxies follow the galaxy size–stellar mass relation for star-forming galaxies from Nedkova et al. (2021), but they are systematically larger by $\sim 0.3 - 0.5$ [dex] at a fixed stellar mass. The value of the discrepancy depends on the correction factor employed from Suess et al. (2019b). We emphasize that Figs 2 and B1 should be understood as illustrations rather than proper comparisons with observations; we plan to analyse synthetic images of FIREbox galaxies using radiative transfer (e.g. Liang et al. 2019, 2021) in future work.

Shibuya et al. (2015) find that a lognormal distribution well approximates the $R_{\rm eff}$ distribution in their combined legacy data set of 3D-*HST*, CANDELS, Hubble Ultra Deep Fields (HUDF) 09 + 12 and the Hubble Frontier Fields (HFF) surveys. While their sample of star-forming and Lyman break galaxies focuses on more massive galaxies ($M_\star \sim 10^{8.5-11}\,{\rm M}_\odot$ since $z = 6$), our $R_{1/2}$ distribution qualitatively agrees. Fig. 3 shows our best-fitting Gaussians to the histograms of $\lg R_{1/2}$ in our final samples at each redshift. We create Q–Q plots for the fits to each redshift (omitted here), which indicate that a Gaussian well approximates each snapshot's $\lg R_{1/2}$ distribution. However, the Shapiro–Wilks tests and chi-squared p values for normal distributions indicate that the distributions are likely not perfectly lognormal. Further, Shibuya et al. (2015) find that the average galaxy size decreases significantly towards higher redshift while maintaining a roughly constant standard deviation. Fig. 3 demonstrates that 'FB-S' galaxies display the same behaviour.

We find an approximately constant standard deviation $\sigma_{\lg R_{1/2}} = 0.12$ [dex] since $z = 5$ in qualitative agreement with Shibuya et al. (2015). However, their value for the scatter is somewhat higher at $\sigma \sim 0.2-0.3$ [dex]. Their sample of galaxies covers a broad range of luminosities, so it is not surprising that their measured scatter is slightly higher than ours. Additionally, van der Wel et al. (2014) find a scatter $\sigma \sim 0.15 - 0.22$ [dex], in closer agreement to our measured value. We note that our galaxy sample changes across the different redshifts due to the minimum stellar mass criterion. Thus, we are reporting the population average and not the progenitor average, which is more similar to observations.

4 THE GALAXY–HALO SIZE RELATION

At each redshift, we construct the GHSR by fitting a power law of the form

$$R_{1/2}\,[{\rm kpc}] = \beta \left(\frac{R_{\rm vir}}{35\,{\rm kpc}} \right)^{\alpha}, \qquad (2)$$

where 35 kpc is the midpoint of the $R_{\rm vir}$ range between $z = 0-5$ in the 'FB-S' haloes. The first and fiducial power law is to the unbinned data set. The α and β values denote the power-law index and normalization, respectively. Then, we calculate the average scatter $\langle \sigma \rangle$ as the standard deviation of the residuals std$(\Delta \lg R_{1/2})$ from the fiducial power law. Lastly, we fit a power law to this scatter as function of $\lg R_{\rm vir}$.

Fig. 4 displays these results for $z = 2$, where the top panel shows the GHSR and the bottom panel details the scatter (standard deviation of the residuals). We also fit power laws to the binned means and medians and found similar best-fitting functions. The best-fitting power-law index of the GHSR at $z = 2$, $\alpha_{\rm GHSR} = 0.939 \pm 0.025$, is close to unity. The approximately linear GHSR in FIREbox at Cosmic Noon in our target sample is qualitatively consistent with similar findings at $z = 0$ (Kravtsov 2013) and at higher redshifts (Huang et al. 2017; Zanisi et al. 2020). However, we note that our sample consists of smaller and lower-mass galaxies than in those observational studies. The scatter in the GHSR of the 'FB-S' galaxies exhibits a moderate increase with $R_{\rm vir}$, as indicated by the non-vanishing power-law index $\alpha_\sigma = 0.144 \pm 0.023$. It is possible that the decreasing number of galaxies and known turnover effects at larger $R_{\rm vir}$ cause the sub-linearities and increasing scatter.

We show the quality of our power-law fits by binning the residuals and fitting a Gaussian distribution to the resulting histogram, as Fig. 5 shows for $z = 2$. The average value of the residuals $\langle \Delta \lg R_{1/2} \rangle$ is null, and the standard deviation $\sigma_{\Delta \lg R_{1/2}}$ is consistent with the average scatter in the GHSR. A Shapiro–Wilks test and the chi-squared p value for a normal distribution indicate that the GHSR

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

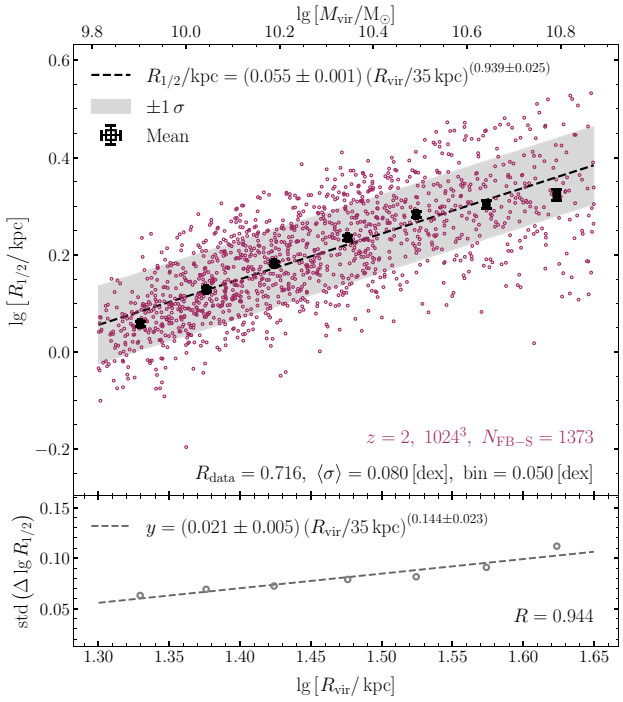

Figure 4. The GHSR relation for the galaxies at $z = 2$ in FIREbox. *Top panel*: The open black squares represent the means ($\langle \lg R_{vir} \rangle$, $\langle \lg R_{1/2} \rangle$) and standard errors within each lg R_{vir} bin of width 0.050 [dex]. The grey region represents $\pm 1\sigma$ scatter above and below the best-fitting power law (black dashed line). The least-squares best-fitting equation to the data is included in the upper left, and the Pearson correlation coefficient R_{data} and average scatter $\langle \sigma \rangle$ are shown in the bottom centre. The slope of the best-fitting power law is approximately linear ($\alpha_{GHSR} = 0.939 \pm 0.025$), agreeing with Kravtsov (2013), Huang et al. (2017), and Somerville et al. (2018). *Bottom panel*: The standard deviation of the residuals for each lg R_{vir} bin from the GHSR in the upper panel. The least-squares best-fitting equation is included in the upper left, and its Pearson correlation coefficient is shown in the lower right, suggesting that the scatter slightly increases with R_{vir}.

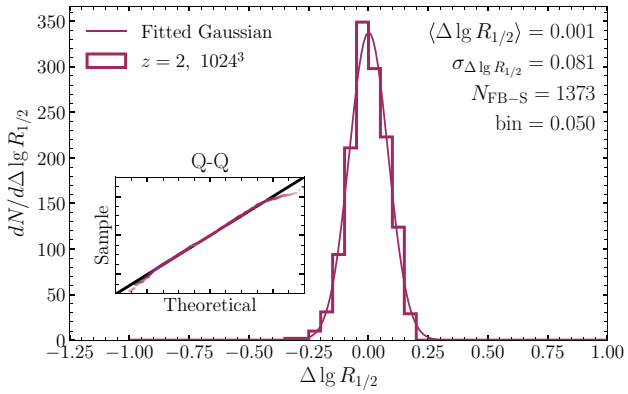

Figure 5. The Gaussian fit to the GHSR residuals $\Delta \lg R_{1/2}$ at $z = 2$. The best-fitting parameters are in the upper right corner, and the binwidth is 0.050 in lg space. The Q–Q plot in the lower left plot indicates that a Gaussian well approximates the $\Delta \lg R_{vir}$ distribution.

residuals $\Delta \lg R_{1/2}$ do not form a perfect normal distribution, but the Gaussian fit is a good approximation.

Table 2 summarizes the main parameters for the GHSR and residuals for each of the analysed redshifts. See Table A1 for the same table using the iterative galaxy definition (Hopkins et al. 2018). Given that the GHSR is consistent with being linear (column 2), the normalization stays roughly constant (column 3), and the average scatter does not vary with redshift (column 4), we conclude that the GHSR in FIREbox is approximately constant since $z = 5$. We calculate the weighted averages of each of the redshifts considered and find that the power-law index $\alpha_{GHSR} = 0.934 \pm 0.054$ is approximately linear and the scatter $\langle \sigma \rangle = 0.084$ is constant with individual snapshots.

Fig. 6 displays the probability distributions of the galaxy-to-halo size *ratio* at each redshift. We fit lognormal distributions to histograms of lg[$R_{1/2}/R_{vir}$], but we do not show the histograms for clarity. There is little redshift evolution in either the mean or the scatter of this size ratio, and we find typical values of $R_{1/2}/R_{vir} \sim 0.05$, which are higher than those of Kravtsov (2013), Shibuya et al. (2015), and Somerville et al. (2018), who find values closer to ∼0.02. However, these studies focus on more massive galaxies than the low-mass centrals analysed here. As Fig. 1 shows, low- and high-mass FIREbox galaxies do not necessarily follow the same GHSR. Somerville et al. (2018) find a weak redshift dependence on this ratio (see their fig. 12). However, this dependence decreases with decreasing halo mass, and our sample is more than an order of magnitude less massive than their least massive bin.

We also combine all galaxy–halo pairs from each redshift and construct a GHSR from this total data set, treating objects from different redshifts equally. This means that the definition of R_{vir} is a function of redshift, because R_{vir} depends on the top hat collapse factor times the background density $\Delta\rho_{back} = \Delta(z)\rho_{back}(z)$ (Bryan & Norman 1998). Fig. 7 displays these results, where the top panel shows the GHSR and the bottom panel the standard deviation of the residuals. The combined sample spans over an order of magnitude in halo size, and there is a 'discretely smooth' transition between the six redshift snapshots. The best-fitting power-law index decreases to $\alpha = 0.894 \pm 0.005$, indicating that this combined sample is slightly sub-linear. This contrasts with the individual and averaged GHSRs, whose power-law index remains consistent with linear at $\alpha = 0.934 \pm 0.054$. At any given instant in time the GHSR is linear, but the size of a given halo typically grows slightly faster than the size of its central galaxy over much of cosmic history. This sub-linearity and the slight decrease of $\langle \lg[R_{1/2}/R_{vir}] \rangle$ with cosmic time (Fig. 6) are qualitatively consistent with expectations of pseudo-evolution of halo sizes (Diemer et al. 2013). However, these low-mass central galaxies remain star forming until $z = 0$, so their galaxy sizes continue growing as well. None the less, it appears that halo sizes grow slightly quicker than galaxy sizes since $z = 5$. Hence, the average $\langle \lg[R_{1/2}/R_{vir}] \rangle$ in Fig. 6 shift leftward with cosmic time (decreasing z), and the power-law fit to the all-z sample is less linear than the fits to each individual redshift in Fig. 7.

As Fig. 7 shows and Table 2 summarizes, the GHSR at each redshift, as well as the combined GHSR, has nearly constant scatter at $\langle \sigma \rangle \approx 0.08 - 0.09$ [dex], suggesting that redshift is *not* a significant factor in explaining the scatter in the GHSR. We emphasize this point by explicitly accounting for a redshift dependence by fitting

$$\lg\left[R_{1/2}/\mathrm{kpc}\right] = \alpha_{R_{vir}} \lg\left[R_{vir}/\mathrm{kpc}\right] + \alpha_{1+z} \lg\left[1+z\right] + \beta. \quad (3)$$

The power-law index for redshift $\alpha_{1+z} = 0.049 \pm 0.013$ is approximately 0, and the index for virial radius increases from $\alpha_{GHSR} = 0.894 \pm 0.005$ (equation 2; Fig. 7) to $\alpha_{R_{vir}} = 0.939 \pm 0.013$. The average scatter $\langle \sigma \rangle = 0.085$ remains exactly the same as that of Fig. 7. Moreover, the log-Gaussian fits to the residuals have the same standard deviation of $\sigma_{\Delta \lg R_{1/2}} = 0.086$ [dex]. Hence, redshift is not

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

Table 2. GHSR at each and all redshifts from the FIREbox 1024^3 'FB-S' galaxies.

z (1)	$\alpha_{\rm GHSR}$ (2)	$\beta_{\rm GHSR}$ (3)	$\langle\sigma\rangle$ (4)	α_σ (5)	β_σ (6)
0	0.929 ± 0.033	0.051 ± 0.002	0.089	0.041 ± 0.035	0.029 ± 0.309
1	1.020 ± 0.032	0.052 ± 0.001	0.082	0.097 ± 0.037	0.024 ± 0.019
2	0.939 ± 0.024	0.055 ± 0.001	0.080	0.144 ± 0.023	0.021 ± 0.005
3	0.909 ± 0.027	0.054 ± 0.001	0.082	0.189 ± 0.029	0.019 ± 0.004
4	0.838 ± 0.036	0.051 ± 0.002	0.089	0.262 ± 0.033	0.017 ± 0.003
5	0.954 ± 0.047	0.055 ± 0.004	0.089	0.189 ± 0.039	0.021 ± 0.007
Mean[a]	0.934 ± 0.054	0.053 ± 0.002	0.084	0.150 ± 0.067	0.022 ± 0.004
All$-z$[b]	0.894 ± 0.005	0.053 ± 0.001	0.085	0.015 ± 0.005	0.033 ± 0.005

[a]The averages weighted by the number of objects at each redshift.
[b]The combined sample of all redshift snapshots, treating objects from different snapshots equally.
(1) Redshift; (2) and (3) power-law index and normalization for the GHSR from equation (2); (4) average scatter in the GHSR; and (5) and (6) power-law index and normalization for the GHSR scatter as a function of lg $R_{\rm vir}$ bin.

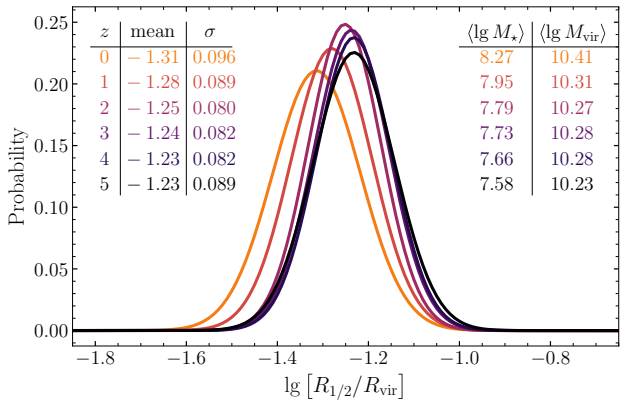

Figure 6. The lognormal probability distributions of the galaxy-to-halo size ratio of the FIREbox galaxies at different redshifts. We omit the histograms for clarity. The best-fitting Gaussian parameters are shown in the upper left corner, and the average galaxy stellar and halo virial masses are included in the upper right corner. There is little redshift evolution in either the scatter ($\sigma \sim 0.08-0.10$) or the mean of this size ratio ($R_{1/2}/R_{\rm vir} \sim 0.049-0.059$) since redshift $z = 5$.

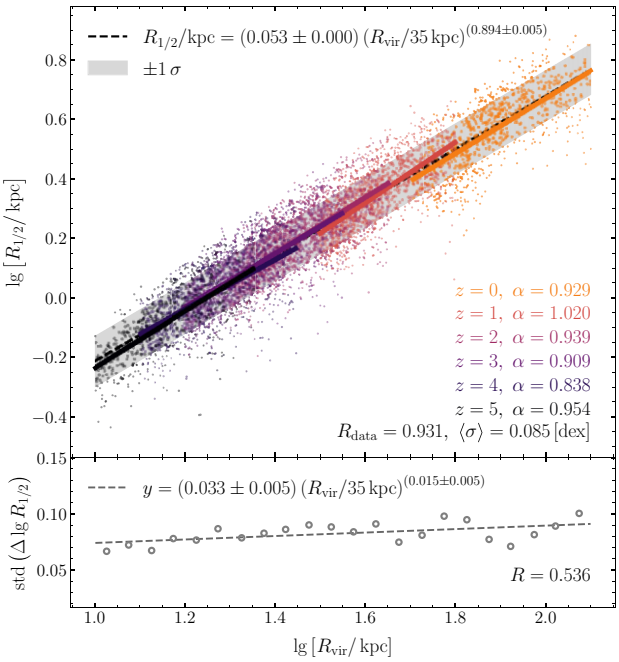

Figure 7. Similar to Fig. 4, but we plot the best-fitting GHSR for each redshift as well as for the combined sample (all$-z$). *Top panel*: the colour of the points represents their redshift (see legend), and the best-fitting power-law index α of the GHSR at that redshift is listed on the bottom right. The least-squares fit to the combined (all$-z$) data is shown at the top of the panel. The grey region represents $\pm 1\sigma$ scatter above and below the best-fitting power law (black dashed line). The Pearson correlation coefficient $R_{\rm data}$ and the average scatter $\langle\sigma\rangle$ are listed at the bottom. The power-law relation of the combined data set is slightly sub-linear. However, the individual power-law relations are much closer to linear (except at $z = 4$) suggesting that the size of a halo grows slightly faster than that of its central galaxy, qualitatively consistent with expectations from pseudo-evolution (Diemer, More & Kravtsov 2013). *Bottom panel*: The standard deviation of the residuals for each lg $R_{\rm vir}$ bin of width 0.050 [dex] from the GHSR in the upper panel. The equation of the best fit and the Pearson correlation coefficient are shown at the top and bottom of the panel. The average scatter is small ($\langle\sigma\rangle < 0.1$ [dex]) and only mildly dependent on halo size.

a significant factor in explaining the scatter in the GHSR in these low-mass galaxies.

We repeat this analysis for R_{25} and R_{80}, i.e. the radii containing 25 per cent and 80 per cent of the stellar mass within $0.1R_{\rm vir}$, respectively (Miller et al. 2019; Mowla et al. 2019a). The scatter changes from $\langle\sigma\rangle \approx 0.08 - 0.09$ [dex] with $R_{1/2}$ to $\approx$0.05 and $\approx$0.11 [dex] with R_{25} and R_{80}, respectively. We also repeat this analysis for $R_{1/2}$ using all stars within $0.2R_{\rm vir}$, and we find that the average scatter of the GHSR increases to $\langle\sigma\rangle \approx 0.12 - 0.13$ [dex]. When including redshift in the two-parameter fit, there is a larger redshift power-law index α_{1+z} to $R_{1/2}$ from equation (3): $\alpha_{R_{\rm vir}} = 0.885 \pm 0.01$, $\alpha_{1+z} = 0.128 \pm 0.018$. However the overall scatter remains constant. Lastly, we repeat this analysis for the iterative definition of galaxy size $R_{1/2}$ described in Hopkins et al. (2018, see Appendix A for more details). Again, the scatter increases to $\langle\sigma\rangle = 0.146$ [dex], and there is a larger redshift power-law index α_{1+z} in the two-parameter fit: $\alpha_{R_{\rm vir}} = 0.925 \pm 0.022$, $\alpha_{1+z} = 0.175 \pm 0.023$. The scatter still does not decrease by including redshift in the two-parameter fit. For each definition of the galaxy size, the GHSRs are approximately linear, and the scatter remains unchanged when including redshift as a second parameter for galaxy size.

5 WHAT SETS THE SCATTER IN THE GHSR?

The scatter in the GHSR is approximately constant since $z = 5$ in the low-mass central galaxies in FIREbox ('FB-S'). Because including

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

redshift as another parameter to the GHSR does not decrease this scatter, we now investigate halo (Section 5.1), galaxy (Section 5.2), and environment (Section 5.3) properties that could determine this scatter. Specifically, we correlate the GHSR residuals with these physical properties. If there is a strong correlation *and* the average scatter decreases, then this property is important in setting the GHSR scatter.

Fig. 8 exemplifies six of these scatter plots for the 'FB-S' galaxies from all redshifts. The vertical axis shows the residuals – i.e. thedifference in log space of the true galaxy size $R_{1/2}$ and that predicted from the GHSR and the halo's $R_{\rm vir}$ – as a function of various properties. We employ a linear regression between the log residuals and functions of the properties. For physical properties – e.g. halo spin λ, stellar mass $M_\star$, and local density $\langle\rho\rangle$ – we fit the residuals $\Delta \lg R_{1/2}$ with the $\lg(x)$ of the physical property x; for fractional properties – e.g. shape parameters c/a and mass fractions $M_\star/M_{\rm bar}$ – we use the logit $(x) \equiv \lg(x/(1-x))$ of the property. These functional choices [$\lg(x)$ for physical properties $x \in (0, \infty)$; logit (x) for fractional properties $x \in (0, 1)$] map their respective properties to all real numbers $(-\infty, \infty)$. For each of these residual–parameter figures, we study the statistical significance $\sigma_{m,0}$ of the slope m from 0 and the per cent reduction in scatter per cent$\Delta\sigma$. We also calculate the vertical offset b of the fit and the Pearson correlation coefficient R. Table 3 details these results for the 'FB-S' galaxies for the halo (top), galaxy (middle), and environment (bottom) properties, respectively. We also repeat this analysis at each redshift and find that all results are consistent with the combined 'all-z' sample.

5.1 Halo properties

The top set of properties correlates the GHSR residuals with the halo properties, employing the AHF values of the properties at the virial radius $R_{\rm vir}$ using all – dark matter and baryonic – interior particles.

The first six halo properties describe the shape using the principal axes of the moment of inertia tensor, such that $a > b > c$. Namely, b/a, c/a, and c/b are the ratios of these axes, and the elongation E, flattening F, and triaxiality T are derived from these ratios. In general, as a halo becomes less spherical/more triaxial, the axis ratios decrease while the elongation, flattening, and triaxiality increase. Each of these six values is between 0 and 1, so we correlate the logit of each parameter with the GHSR residuals. While logit b/a, logit c/a, and logit E have slopes $>4\sigma$ from null, they and the other four properties do not decrease the scatter. Consequently, we conclude that the halo shape is not significant in setting the scatter in the GHSR.

Many theoretical and empirical works suggest that the halo spin and concentration are significant in setting the galaxy sizes, especially in rotationally supported disc galaxies (Fall & Efstathiou 1980; Mo et al. 1998; Somerville et al. 2008; Desmond & Wechsler 2015; Diemer & Kravtsov 2015; Desmond et al. 2017; Somerville et al. 2018). We study the halo spin $\lg\lambda$, $\lg\lambda_e$ (top left-hand panel of Fig. 8) employing the definition of Bullock et al. (2001):

$$\lambda = \frac{J}{\sqrt{2} M_{\rm vir} V_{\rm circ} R_{\rm vir}}, \tag{4}$$

and the classical definition of Peebles (1969):

$$\lambda_e(R_{\rm vir}) = \frac{J|E|^{1/2}}{G M_{\rm vir}^{5/2}}, \tag{5}$$

where $J = J(< r)$ is the angular momentum, $V_{\rm circ}$ is the circular velocity of all particles within the halo, E is the total energy of the system, and G is Newton's gravitational constant. Again, while their slopes are $m \sim 7\sigma$ from null, they do not decrease the scatter.

Assuming a Navarro-Frank-White (NFW) dark matter profile (Navarro, Frenk & White 1997), we calculate the halo concentration using the definition from Prada et al. (2012, see their section 3, equations 9 and 10 for more details). This concentration is also a proxy for the formation history of the halo (Wechsler et al. 2002). The fit to lg cNFW has no significant correlation (top right-hand panel of Fig. 8). Because the concentration is linked to the halo formation history, this lack of correlation suggests that the halo formation time is not critical in the GHSR scatter.

We fit the centre-of-mass offset lg ΔCOM – defined as the distance between the halo's centre of mass and the halo's centre calculated as the densest cell. A larger offset implies some spherical asymmetry in the halo, such as massive satellites. Interestingly, this property has the strongest correlation with the residuals with a slope that is 19σ from null, but the scatter decreases only by ≈ 3 per cent.

Thus, we conclude that none of the studied halo properties – including the halo spin and the concentration – significantly explain the scatter in the GHSR in our FIREbox sample of low-mass central galaxies over the past 12.5 billion years.

In Appendix C, we also correlate the GHSR scatter with the difference between a property's value and the average value of all similar-massed haloes, the value of the cross-matched halo in the DMO simulation, and the difference between the cross-matched halo value and the average value of all similar-massed DMO haloes. There are no significant reductions in the GHSR scatter for any of the studied properties. Thus all null correlations between the halo properties and GHSR scatter are consistent, and we conclude that the halo properties are not important in setting the GHSR scatter in the 'FB-S' galaxies.

5.1.1 *Halo concentration and spin*

In Fig. 9, we construct the stellar–halo mass (top row) and galaxy–halo size (bottom row) relations for the 'FB-S' galaxies at $z = 2$ (the same results hold at the other redshifts), where the sample is divided into four equal quartiles by halo concentration (left column) and spin (right column). For the SHMR, at a fixed halo mass, haloes with a higher concentration have higher stellar masses (top left-hand panel). The four SHMRs follow the same shape, but the high concentration curves have larger normalizations, appearing as vertical shifts. These results for the halo concentration agree with FIRE-1 results from Feldmann, Faucher-Giguère & Kereš (2019). These vertical translations in the SHMR with concentration are equivalent to varying normalizations of the stellar density profiles (Lilly & Carollo 2016). The trend reverses for the halo spin; at a fixed halo mass, haloes with smaller spin have larger stellar masses (top right-hand panel). For the 'FB-S' galaxies, the concentration and the spin are significant properties in setting the SHMR.

However, there is no such distinction in the GHSR relations when separated into quartiles by either concentration or spin. The null result with concentration agrees with the semi-empirical model from Zanisi et al. (2021a) and disagrees with results from zoom-ins from Jiang et al. (2019), although both studies focused on more massive galaxies. The null correlation with halo spin agrees with Jiang et al. (2019). For the largest, most massive haloes in the 'FB-S' galaxies, there may be a slight dependence on the halo spin (bottom right-hand panel). This could be due to the transition from dispersion to rotationally supported galaxies, where the formation of discs becomes increasingly significant (Fall & Efstathiou 1980; Mo et al.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

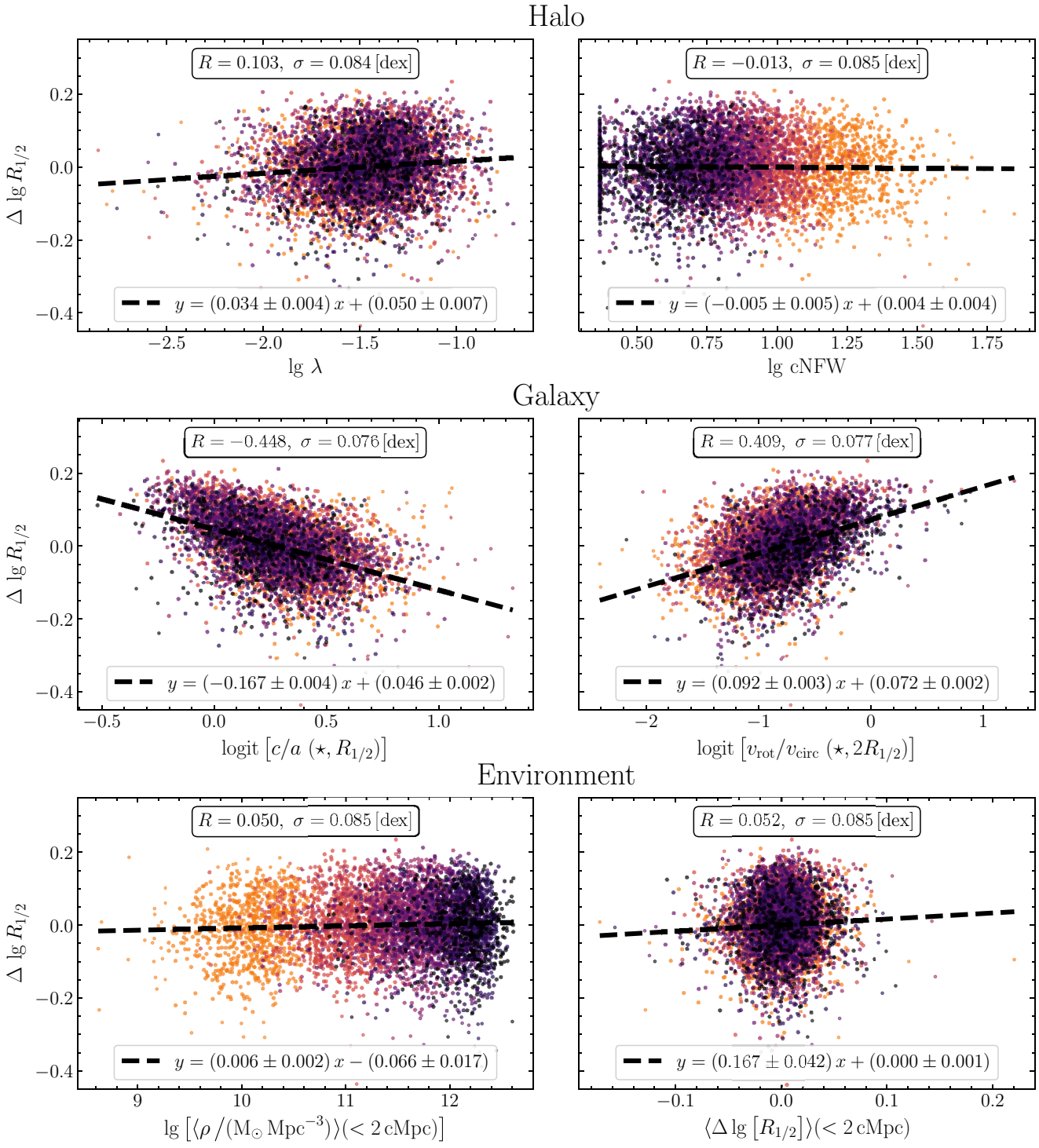

Figure 8. Six example figures graphing the residuals $\Delta \lg R_{1/2}$ (i.e. the scatter) in the galaxy–halo size relation (GHSR) versus two halo (top panels), galaxy (middle), and environment properties (bottom). The best-fitting power laws are shown with black dashed lines, and their respective equations, Pearson correlation coefficients, and new scatter are given in each panel. These values are summarized for all properties in Table 3. The average scatter of the GHSR is $\sigma_{\Delta \lg R_{1/2}} = 0.085$ [dex]. *Top left*: The halo spin. While the best-fitting slope is 8.5σ from 0, the scatter does not decrease. *Top right*: The halo concentration does not correlate with the GHSR scatter. *Middle left*: The shape of the stars correlates with GHSR scatter and deceases the overall scatter. At a fixed halo size, galaxies with stars that are less spherical (lower *c*/*a*) have larger sizes. This property is significant in setting the GHSR scatter. *Middle right*: The ratio of stellar rotational velocity to the circular velocity at the half-mass radius. At a fixed halo size, galaxies with more rotational support (higher $v_{\mathrm{rot}}/v_{\mathrm{circ}}$) have larger sizes. This property is anticorrelated with the stellar shape *c*/*a*. *Bottom left*: The average total density of all central neighbours within 2 cMpc does not correlate with the GHSR scatter. *Bottom right*: The average GHSR residuals of all central neighbours (size conformity) within 2 cMpc. The average size conformity correlates weakly with the residuals, but the GHSR scatter does not decrease.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

Table 3. Correlations of the 'FB-S' GHSR's scatter with halo (top), galaxy (middle), and environment (bottom) properties at all considered redshifts.

Parameter (1)	m (2)	$\sigma_{m,0}$ (3)	b (4)	R (5)	$\langle\sigma\rangle$ (6)	$\%\Delta\sigma$ (7)
logit[b/a (halo, $R_{\rm vir}$)]	0.013 ± 0.003	4.14	−0.012 ± 0.003	0.055	0.085 ± 0.001	0.1
logit[c/a (halo, $R_{\rm vir}$)]	0.022 ± 0.005	4.43	−0.012 ± 0.003	0.058	0.085 ± 0.001	0.2
logit[c/b (halo, $R_{\rm vir}$)]	0.004 ± 0.004	0.99	−0.003 ± 0.004	0.013	0.085 ± 0.001	0.0
logit $\left[E \equiv \sqrt{1-(b/a)^2}\ (\text{halo}, R_{\rm vir})\right]$	−0.018 ± 0.004	4.37	−0.001 ± 0.001	−0.058	0.085 ± 0.001	0.2
logit $\left[F \equiv \sqrt{1-(c/b)^2}\ (\text{halo}, R_{\rm vir})\right]$	−0.005 ± 0.005	0.97	−0.000 ± 0.001	−0.013	0.085 ± 0.001	0.0
logit $\left[T \equiv \left(1-(b/a)^2\right) / \left(1-(c/a)^2\right)\ (\text{halo}, R_{\rm vir})\right]$	−0.007 ± 0.002	2.78	0.001 ± 0.001	−0.037	0.085 ± 0.001	0.1
lg λ	0.034 ± 0.004	7.82	0.050 ± 0.007	0.103	0.084 ± 0.001	0.5
lg λ_e	0.030 ± 0.005	6.58	0.046 ± 0.007	0.086	0.084 ± 0.001	0.4
lg cNFW	−0.005 ± 0.005	0.99	0.004 ± 0.004	−0.013	0.085 ± 0.001	0.0
lg $\left[\sigma_v/\text{km s}^{-1}\right]$	0.021 ± 0.009	2.28	−0.038 ± 0.017	0.030	0.085 ± 0.001	0.0
lg $\left[\Delta\,\text{COM/kpc}\right]$	0.064 ± 0.003	19.32	−0.039 ± 0.002	0.247	0.082 ± 0.001	3.1
logit $\left[b/a\ (\star, R_{1/2})\right]$	**−0.101 ± 0.003**	**31.84**	**0.063 ± 0.002**	**−0.388**	**0.078 ± 0.001**	**7.8**
logit $\left[c/a\ (\star, R_{1/2})\right]$	**−0.167 ± 0.004**	**38.00**	**0.046 ± 0.002**	**−0.448**	**0.076 ± 0.001**	**10.6**
logit[c/b ($\star$, $R_{1/2}$)]	−0.056 ± 0.004	14.54	0.040 ± 0.003	−0.188	0.083 ± 0.001	1.8
logit $\left[E \equiv \sqrt{1-(b/a)^2}\ (\star, R_{1/2})\right]$	**0.120 ± 0.004**	**33.22**	**−0.021 ± 0.001**	**0.402**	**0.078 ± 0.001**	**8.4**
logit $\left[F \equiv \sqrt{1-(c/b)^2}\ (\star, R_{1/2})\right]$	0.069 ± 0.005	15.06	−0.006 ± 0.001	0.195	0.083 ± 0.001	1.9
logit $\left[T \equiv \left(1-(b/a)^2\right) / \left(1-(c/a)^2\right)\ (\star, R_{1/2})\right]$	0.046 ± 0.003	18.40	−0.013 ± 0.001	0.236	0.082 ± 0.001	2.8
lg $\left[M_\star/\text{M}_\odot\right]$	−0.007 ± 0.002	4.01	0.057 ± 0.014	−0.053	0.085 ± 0.001	0.1
logit $\left[M_\star(< R_{1/2})/M_{\rm bar}(< R_{1/2})\right]$	**−0.047 ± 0.002**	**26.27**	**−0.019 ± 0.001**	**−0.336**	**0.079 ± 0.001**	**5.8**
logit $\left[M_\star(< R_{1/2})/M_{\rm tot}(< R_{1/2})\right]$	−0.062 ± 0.003	22.86	−0.074 ± 0.003	−0.289	0.081 ± 0.001	4.3
lg $\left[M_{\rm gas}(< R_{1/2})/\text{M}_\odot\right]$	0.029 ± 0.002	18.94	−0.226 ± 0.012	0.249	0.081 ± 0.001	3.1
logit $\left[M_{\rm gas}(< R_{1/2})/M_{\rm tot}(< R_{1/2})\right]$	0.048 ± 0.002	20.29	0.034 ± 0.002	0.267	0.081 ± 0.001	3.6
lg $\left[M_{\rm bar}(< R_{1/2})\text{M}_\odot\right]$	0.027 ± 0.002	14.13	−0.220 ± 0.016	0.183	0.083 ± 0.001	1.7
logit $\left[M_{\rm bar}(< R_{1/2})/M_{\rm tot}(< R_{1/2})\right]$	0.025 ± 0.003	8.74	0.011 ± 0.002	0.115	0.084 ± 0.001	0.7
lg $\left[M_{\rm dm}(< R_{1/2})/\text{M}_\odot\right]$	0.057 ± 0.004	15.80	−0.491 ± 0.031	0.204	0.083 ± 0.001	2.1
lg $\left[M_{\rm tot}(< R_{1/2})/\text{M}_\odot\right]$	0.043 ± 0.003	13.80	−0.376 ± 0.027	0.179	0.083 ± 0.001	1.6
logit $\left[v_{\rm rot}/v_{\rm circ}\ (\star, R_{1/2})\right]$	**0.092 ± 0.003**	**33.89**	**0.072 ± 0.002**	**0.409**	**0.077 ± 0.001**	**8.7**
lg $\left[\text{min(d)/kpc}\right]$	−0.016 ± 0.003	5.10	0.036 ± 0.007	−0.067	0.085 ± 0.001	0.2
lg $\left[\text{min}(R_{\rm Hill})/\text{kpc}\right]$	−0.013 ± 0.003	4.26	0.027 ± 0.006	−0.056	0.085 ± 0.001	0.2
lg $\left[\langle n_{\rm gal}/\text{Mpc}^{-3}\rangle(< 2\,\text{cMpc})\right]$	0.003 ± 0.002	1.76	−0.002 ± 0.002	0.024	0.084 ± 0.001	0.0
lg $\left[\langle\rho/(\text{M}_\odot\,\text{Mpc}^{-3})\rangle(< 2\,\text{cMpc})\right]$	0.006 ± 0.002	3.80	−0.066 ± 0.017	0.050	0.085 ± 0.001	0.1
Δ lg $\left[R_{1/2,\text{min}(R_{\rm Hill})}\right]$	0.025 ± 0.012	2.13	0.000 ± 0.001	0.028	0.085 ± 0.001	0.0
$\langle\Delta$ lg $\left[R_{1/2}\right]\rangle(< 2\,\text{cMpc})$	0.167 ± 0.042	3.97	0.000 ± 0.001	0.052	0.085 ± 0.001	0.1

Note. (1) Parameter used as the horizontal axis – logit $x \equiv \lg x/(1-x)$ for $x \in (0, 1)$; (2) slope of the fit; (3) σ from the slope being 0 using the statistical error of m; (4) vertical offset in the fit; (5) Pearson correlation coefficient; (6) scatter in regression in the residuals versus parameter; and (7) percentage difference in scatter between the GHSR and the residual–parameter relation. From top to bottom the sections are the halo, galaxy, and environment properties. Section 5 describes what each parameter is and how it is calculated. Parameters that explain at least 5 per cent of the scatter [column (7)] – $\%\Delta\sigma > 5$ – are in boldface.

1998). However, in the MW halo mass regime ($M_{\rm halo} \sim 10^{12}\,\text{M}_\odot$) in FIRE-2 zoom simulations, Garrison-Kimmel et al. (2018) found only weak correlations between the galaxy size and halo spin. Similarly, Zanisi et al. (2020) found in their semi-empirical model using galaxies ($M_\star \sim 10^{9-12}\,\text{M}_\odot$) from the Sloan Digital Sky Survey that the halo spin may not be significant in the GHSR. Including the halo spin in the 'FB-S' GHSR does not reduce its scatter.

For the 'FB-S' galaxies, the SHMR depends on the halo concentration (top left-hand panel) and spin (top right-hand panel); the GHSR does not depend on either of these halo properties (bottom panels). Thus, we suggest that for these low-mass central galaxies, the GHSR is just as or more fundamental than the SHMR, agreeing with semi-empirical results for more massive galaxies (Zanisi et al. 2021a, b).

5.2 Galaxy properties

The middle section of Table 3 summarizes the residual–parameter correlations for the galaxies from all redshifts. Here, the rows in boldface represent parameters that decrease the scatter by > 5 per cent. Each such power-law index also has $>25\sigma$ difference from 0.

We define the galaxy-shape parameters using only the stellar particles within $R_{1/2}$. Many of these properties have significant correlations with the scatter. The middle left-hand panel of Fig. 8 displays the figure for the most significant shape property, logit $[c/a]$. The three axis ratios all have negative correlations, while the elongation, flattening, and triaxiality have positive correlations. Hence, at a fixed halo size, galaxies that are less spherical have larger $R_{1/2}$. This can be expected based on our spherical definition of $R_{1/2}$. For example, let a triaxial galaxy have axes $a > b > c$ and total stellar mass $M_\star(<$

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

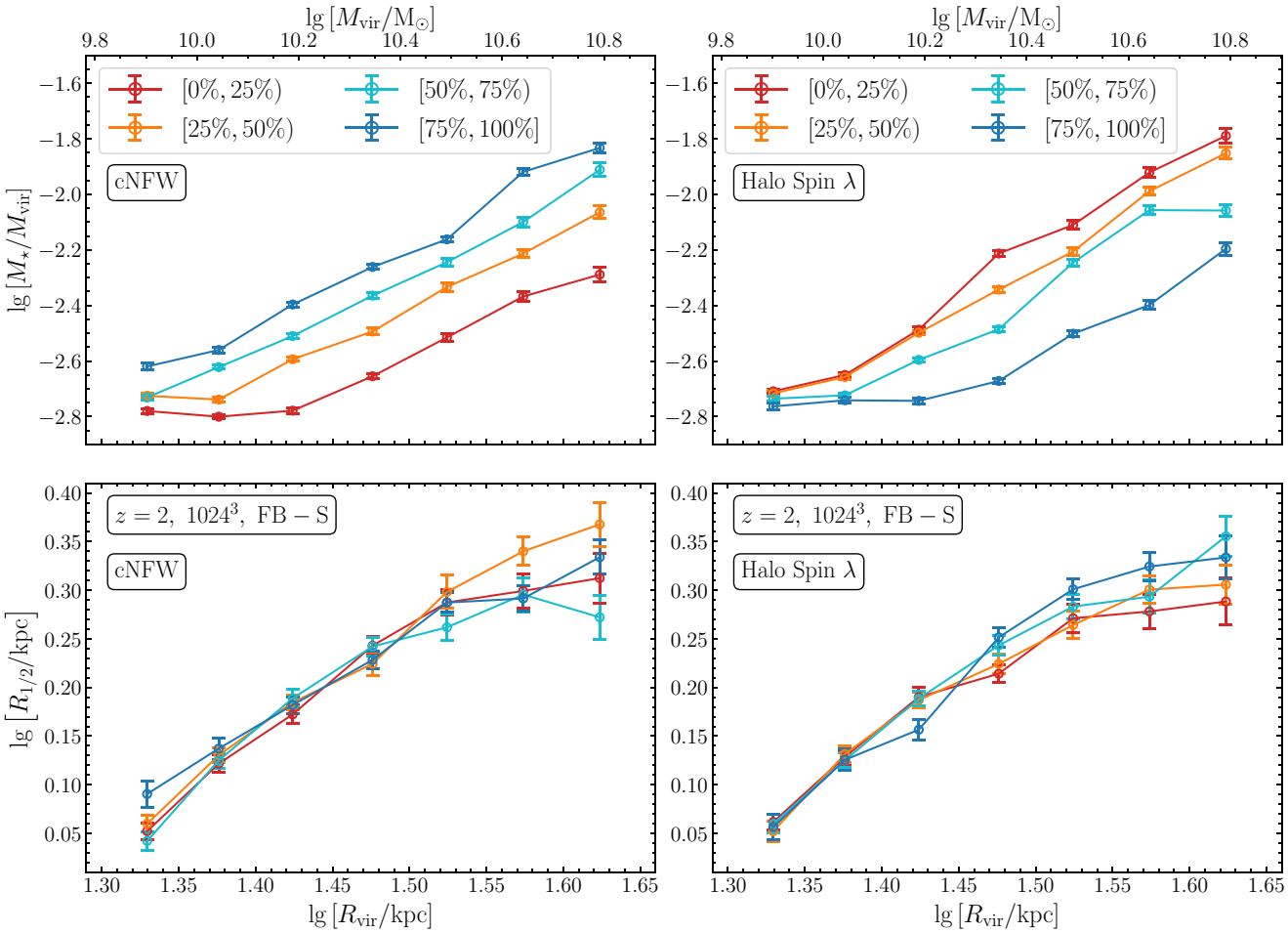

Figure 9. The stellar-halo mass (top row) and galaxy-halo size (bottom row) relations for the FIREbox galaxies at $z = 2$, separated into equal-sized quartiles by halo concentration cNFW (left column) and spin λ (right-hand panel). The $M_\star - M_{vir}$ relation depends on the concentration and halo spin, while the $R_{1/2} - R_{vir}$ relation does not. However, the largest, most massive haloes in the GHSR may have a slight dependence on halo spin (bottom right-hand panel). The same qualitative results hold at other redshifts. Thus, we suggest that for the low-mass, central, dispersion-supported systems, the galaxy–halo size relation is just as or more fundamental than the stellar–halo mass relation.

a). As we integrate the interior stellar mass $M_\star(< r)$ with increasing radius r, we reach the edge of the shortest axis first at $r = c$. If $M_\star(< r = c) < 0.5M_\star(< a)$, then we must continue to further radii to determine $R_{1/2}$. Thus, in the direction of the c axis, we do not accumulate any more stellar material. The same argument holds for the b axis, when $M_\star(< r = b) < 0.5M_\star(< a)$. Therefore, at a fixed R_{vir}, we expect galaxy sizes to be negatively correlated with the axis ratios and positively correlated with the elongation, flattening, and triaxiality.

Additionally, the stellar-to-baryonic mass ratio logit$[M_\star/M_{bar}]$ – or similarly the gas-to-baryonic mass ratio logit$[M_{gas}/M_{bar}]$ = logit$[1 - M_\star/M_{bar}]$ – and the stellar-to-total mass ratio logit$[M_\star/M_{tot}]$ have significant power-law indices and slightly decrease the scatter. These properties are related to the star formation rates and histories of the galaxies, and these properties warrant future study.

We calculate the rotational velocity v_{rot} of the stars,

$$v_{rot} = \frac{L_\star(< R_{1/2})}{M_\star(< R_{1/2})R_{1/2}}, \tag{6}$$

and the circular velocity v_{circ} at the half-mass radius,

$$v_{circ} = \sqrt{\frac{GM_{tot}(< R_{1/2})}{R_{1/2}}}, \tag{7}$$

where $L_\star(R_{1/2})$ is the angular momentum of all stars within $R_{1/2}$. This ratio v_{rot}/v_{circ} determines the amount of rotational support for the galaxy, where $v_{rot}/v_{circ} \sim 1$ is rotationally supported and $\sim$0 is dispersion supported. The middle right-hand panel of Fig. 8 shows this ratio. We find that at fixed halo size, more discy galaxies (higher v_{rot}/v_{circ}) have larger sizes. This amount of rotational support is anticorrelated with c/a, meaning that at a fixed R_{vir}, more discy galaxies (lower c/a) have more rotational support (higher v_{rot}/v_{circ}) and larger sizes. Importantly, most of the low-mass galaxies in our sample are not rotationally supported, with $v_{rot}/v_{circ} \sim 0.0-0.2$. This suggests that our sample is largely dispersion supported, rather than angular momentum supported, which helps explain why there is no observed correlation between halo spin and GHSR scatter. However, it remains to be understood why the sizes of low-mass galaxies track their halo virial radii so closely.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

5.3 Environment properties

Lastly, we examine the environmental properties by constructing a co-moving sphere of radius $R = 2$ cMpc and counting each main (central) halo (galaxy) within this sphere (we varied the sphere's radius and found similar results). For each object's catalogue of neighbours, we calculate the distance $\lg \min(d)$ to the nearest main galaxy; the minimum tidal disruption (Hill) radius $\lg \min(R_{\rm Hill})$; the mean number $\lg\langle n\rangle$; and mass $\lg\langle\rho\rangle$ (bottom left-hand panel of Fig. 8) densities and find no significant trends. Lastly, we test if galaxies that form spatially near each other systematically lie above or below the GHSR. That is, does galaxy size conformity exist? We correlate the average GHSR offset $\langle\Delta \lg R_{1/2}\rangle$ of all neighbours (bottom right-hand panel of Fig. 8), again finding no decrease in scatter for any of the environmental properties.

6 CONCLUSIONS

In this work, we investigate the link between galaxy size and halo size based on a high-resolution, cosmological volume simulation suite from the FIRE collaboration. We focus on low-mass centrals with stellar masses $M_\star \sim 10^{7-9}\,{\rm M}_\odot$. Our main results and conclusions are as follows:

(i) Our galaxy sizes appear consistent with observations, lying along the extrapolated star-forming $M_\star - R_{\rm eff}$ relation at $z = 2$ (Fig. 2, van der Wel et al. 2014; Mowla et al. 2019b; Suess et al. 2019a; Nedkova et al. 2021) and approximating lognormal distributions at $0 \leq z \leq 5$ (Fig. 3; Shibuya et al. 2015). At $z = 0$, the FIREbox galaxies follow the galaxy size–stellar mass relation for star-forming galaxies from Nedkova et al. (2021), but they are systematically larger by $\sim 0.3 - 0.5$ [dex] at a fixed stellar mass (Fig. B1).

(ii) The galaxy–halo size relations (GHSRs) for the low-mass objects at each redshift $0 \leq z \leq 5$ are consistent with being linear, agreeing with previous studies (Kravtsov 2013; Huang et al. 2017; Somerville et al. 2018; Jiang et al. 2019; Zanisi et al. 2020). The power-law index $\alpha = 0.934 \pm 0.054$ and scatter $\langle\sigma\rangle = 0.084$ [dex] from the weighted average from all redshifts are consistent with the individual redshifts, suggesting that the GHSR is constant for low-mass galaxies since $z = 5$. In general, we find $R_{1/2}/R_{\rm vir} \sim 0.05$, which is similar to expectations from spin-based models (Mo et al. 1998).

(iii) While the GHSR at each redshift is roughly linear, the power-law index $\alpha_{\text{all-z}} = 0.894 \pm 0.005$ of the combined sample across redshifts suggests that individual objects may trace out paths that are sub-linear (Fig. 7). This result and the leftward shift of $\lg[R_{1/2}/R_{\rm vir}]$ distribution with cosmic time (shown in Fig. 6) are qualitatively consistent with expectations of pseudo-evolution of haloes (Diemer et al. 2013). The power-law fit to the all-z sample still details a smooth transition since $z = 5$, and the scatter does not decrease when accounting for redshift.

(iv) The halo properties we explore – including spin and concentration – do *not* reduce the scatter in the GHSR in our sample of low-mass, dispersion-supported galaxies in FIREbox. Our weak dependence on spin disagrees with classical theoretical ideas of galaxy formation (Mo et al. 1998) but agrees with recent numerical works of more massive galaxies (Desmond et al. 2017; Garrison-Kimmel et al. 2018; Jiang et al. 2019). At a fixed halo size, the weak correlation of galaxy sizes with halo spin and concentration agrees with FIRE-2 zoom simulations (Garrison-Kimmel et al. 2018) and semi-empirical models (Zanisi et al. 2020, 2021a) of more massive, rotationally supported galaxies.

(v) The galaxy shape, amount of stellar rotational support, and potentially the stellar or gas fractions correlate with GHSR residuals. This suggests that baryonic feedback processes of galaxy evolution as well as observable galactic structure/kinematics may be significant in setting the scatter in the GHSR.

The remarkably tight GHSR with nearly constant scatter and linear power-law indices since $z = 5$ allow for estimating halo masses from the sizes of galaxies. This technique is independent from, and as accurate as, other commonly employed mass-based methods, namely abundance matching, and has the potential to be a better environmental indicator in low-mass galaxy surveys (Yang et al. 2007; Yang, Mo & Van den Bosch 2009). Especially with upcoming low-mass galaxy surveys at high redshifts, with, for example, the James Webb Space Telescope, we suggest inferring halo masses and sizes from the galaxy sizes. Given a measurement of a galaxy's stellar mass $M_\star$, one can estimate the halo mass $M_{\rm vir}$ using the SHMR and associated scatter $\sigma \approx 0.25$ [dex]. That is, a halo mass $\lg[M_{\rm vir}(M_\star)/{\rm M}_\odot]$ inferred from the stellar mass $M_\star$ has an associated error of roughly ± 0.25 [dex]. Now with the galaxy size $R_{1/2}$, one can estimate the halo size $R_{\rm vir}$ via the GHSR using the scatter $\sigma \approx 0.08 - 0.09$ [dex] for the 'FB-S' galaxies. Then, a halo size $\lg[R_{\rm vir}(R_{1/2})/{\rm kpc}]$ inferred from the galaxy size $R_{1/2}$ has an associated error $\pm 0.08 - 0.09$ [dex]. Because $M_{\rm vir}$ scales exactly as $M_{\rm vir} \propto R_{\rm vir}^3$, then the error on estimating $\lg[M_{\rm vir}/{\rm M}_\odot]$ is three times that of estimating the halo size, given the galaxy size. That is, a halo mass $\lg[M_{\rm vir}(R_{1/2})/{\rm M}_\odot]$ estimated from the galaxy size $R_{1/2}$ has an associated error of approximately $\pm 0.24 - 0.27$ [dex], comparable to the error from using $M_\star$. However, using galaxy sizes could be more beneficial than using stellar masses because the GHSR depends less on the unobservable dark matter halo properties, namely spin and concentration. Thus, for dispersion-supported, low-mass, central galaxies, we suggest that the galaxy–halo size relation is just as or even more fundamental than the SHMR.

ACKNOWLEDGEMENTS

We extend our deepest appreciation to the innumerable colleagues, staff, and loved ones at each of our institutes and in each others' lives, especially during the COVID-19 pandemic, for their continued encouragement and necessary assistance in the research and writing of this paper. We thank the anonymous reviewer for his or her insightful comments, which significantly improved the quality and clarity of this paper. We acknowledge PRACE for awarding us access to MareNostrum at the Barcelona Supercomputing Center (BSC), Spain. This research was partly carried out via the Frontera computing project at the Texas Advanced Computing Center. Frontera is made possible by National Science Foundation award OAC-1818253. This work was supported in part by a grant from the Swiss National Supercomputing Centre (CSCS) under project IDs s697 and s698. We acknowledge access to Piz Daint at the Swiss National Supercomputing Centre, Switzerland, under the University of Zurich's share with the project ID uzh18. This work made use of infrastructure services provided by S^{3}IT (www.s3it.uzh.ch), the Service and Support for Science IT team at the University of Zürich.

ECR acknowledges the ThinkSwiss Research Scholarship, funded by the State Secretariat for Education, Research and Innovation (SERI) for the opportunity to spend 3 months at the University of Zürich Institute for Computational Science; travel support from the Alexander Vyssotsky Award from the University of Virginia to present this work; and to many associates – namely Sven De Rejicke, Arjen van der Wel, Charles Steinhardt, Luca Beale, Robin Leichtnam, and Hugues Lascombes – for their conversations regarding this analysis. ECR is a fellow of the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg (IMPRS-HD).

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

RF acknowledges financial support from the Swiss National Science Foundation (grant nos. PP00P2_157591, PP00P2_194814, and 200021_188552).

JSB was supported by National Science Foundation (NSF) grant AST-1910346.

MBK acknowledges support from NSF CAREER award AST-1752913, NSF grant AST-1910346, NASA grant NNX17AG29G, and *HST*-AR-15006, *HST*-AR-15809, *HST*-GO-15658, *HST*-GO-15901, *HST*-GO-15902, *HST*-AR-16159, and *HST*-GO-16226 from the Space Telescope Science Institute (STScI), which is operated by AURA, Inc., under NASA contract NAS5-26555.

CAFG was supported by NSF through grants AST-1715216 and CAREER award AST-1652522; by NASA through grant 17-ATP17-0067; by STScI through grant *HST*-AR-16124.001-A; and by the Research Corporation for Science Advancement through a Cottrell Scholar Award and a Scialog Award.

AW received support from NSF CAREER grant 2045928; NASA ATP grants 80NSSC18K1097 and 80NSSC20K0513; *HST* grants AR-15057, AR-15809, and GO-15902 from STScI; a Scialog Award from the Heising-Simons Foundation; and a Hellman Fellowship.

DK was supported by NSF grant AST-1715101.

DATA AVAILABILITY AND SOFTWARE USED

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

Software used: PYTHON (Van Der Walt, Colbert & Varoquaux 2011); IPYTHON (Pérez & Granger 2007); NUMPY (Van Der Walt et al. 2011; Harris et al. 2020); SCIPY (Virtanen et al. 2020); MATPLOTLIB (Hunter 2007); and JUPYTER (Kluyver et al. 2016).

This work made extensive use of the NASA Astrophysics Data System and arXiv.org preprint server.

REFERENCES

Alves J., Combes F., Ferrara A., Forveille T., Shore S., 2016, A&A, 594, A13
Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, MNRAS, 488, 3143
Berlind A. A. et al., 2003, The ApJ, 593, 1
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, ApJ, 555, 240
Bullock J. S., Kravtsov A. V., Weinberg D. H., 2000, ApJ, 539, 517
Ceverino D., Klypin A., Klimek E. S., Trujillo-Gomez S., Churchill C. W., Primack J., Dekel A., 2014, MNRAS, 442, 1545
Chaves-Montero J., Angulo R. E., Schaye J., Schaller M., Crain R. A., Furlong M., Theuns T., 2016, MNRAS, 460, 3100
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Conroy C., Wechsler R. H., 2009, ApJ, 696, 620
Crain R. A. et al., 2015, MNRAS, 450, 1937
Croton D. J. et al., 2006, MNRAS, 365, 11
Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, MNRAS, 449, 2087
Dekel A., Silk J., 1986, ApJ, 303, 39
Desmond H., Mao Y.-Y., Wechsler R. H., Crain R. A., Schaye J., 2017, MNRAS, 471, L11
Desmond H., Wechsler R. H., 2015, MNRAS, 454, 322
di Teodoro E. M., Fraternali F., 2015, MNRAS, 451, 3021
Diemer B., Kravtsov A. V., 2015, ApJ, 799, 108
Diemer B., More S., Kravtsov A. V., 2013, ApJ, 766, 25
El-Badry K. et al., 2018, MNRAS, 473, 1930
Engler C. et al., 2020, MNRAS, 500, 3957
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Feldmann R., Faucher-Giguère C.-A., Kereš D., 2019, ApJ, 871, L21
Feldmann R., Hopkins P. F., Quataert E., Faucher-Giguère C.-A. A., Kereš D., Keřes D., 2016, MNRAS, 458, L14
Garrison-Kimmel S. et al., 2018, MNRAS, 481, 4133
Genel S. et al., 2018, MNRAS, 474, 3976
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Harris C. R. et al., 2020, Nature, 585, 357
Hirtenstein J. et al., 2019, ApJ, 880, 54
Hopkins P. F. et al., 2018, MNRAS, 480, 800
Hopkins P. F., 2015, MNRAS, 450, 53
Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012b, MNRAS, 427, 968
Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C. A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Hopkins P. F., Quataert E., Murray N., 2012a, MNRAS, 421, 3488
Huang K.-H. et al., 2017, ApJ, 838, 6
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jiang F. et al., 2019, MNRAS, 488, 4801
Kawinwanichakij L. et al., 2021, ApJ, 921, 38
Kennicutt R. C., Jr., 1998, ApJ, 498, 541
Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M. S., 2015, MNRAS, 450, 1349
Kluyver T. et al., 2016, Jupyter Notebooks-a publishing format for reproducible computational workflows, 2016, 87
Knollmann S. R., Knebe A., 2009, ApJS, 182, 608
Kravtsov A. V., 2013, ApJ, 764, L31
Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlober S., Allgood B., Primack J. R., 2004, ApJ, 609, 35
Kravtsov A. V., Klypin A. A., 1999, Evolution, 1, 437
Kravtsov A. V., Vikhlinin A. A., Meshcheryakov A. V., 2018, Astron. Lett., 44, 8
Kretschmer M., Agertz O., Teyssier R., 2020, MNRAS, 497, 4346
Lange R. et al., 2016, MNRAS, 462, 1470
Lewis A., Challinor A., Hanson D., 2011, JCAP, 2011, 018
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Liang L. et al., 2019, MNRAS, 489, 1397
Liang L., Feldmann R., Hayward C. C., Narayanan D., Çatmabacak O., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., 2021, MNRAS, 502, 3210
Lilly S. J., Carollo C. M., 2016, ApJ, 833, 1
Ma X. et al., 2018a, MNRAS, 477, 219
Ma X. et al., 2018b, MNRAS, 478, 1694
Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715
McAlpine S. et al., 2016, Astron. Comput., 15, 72
Miller T. B., Dokkum P. v., Mowla L., Wel A. v. d., 2019, ApJ, 872, L14
Mo H. J., Mao S., White S. D., 1998, MNRAS, 295, 319
More S., Van Den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210
Mowla L. A. et al., 2019b, ApJ, 880, 57
Mowla L., Wel A. v. d., Dokkum P. v., Miller T. B., 2019a, ApJL, 872, L13
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nedkova K. V. et al., 2021, MNRAS, 506, 928
Nelson D. et al., 2019, Computational Astrophysics and Cosmology, 6, 2
Orr M. E. et al., 2018, MNRAS, 478, 3653
Pearce F. R., Jenkins A., Frenk C. S., White S. D., Thomas P. A., Couchman H. M., Peacock J. A., Efstathiou G., 2001, MNRAS, 326, 649
Peebles P. J. E., 1969, ApJ, 155, 393
Pérez F., Granger B. E., 2007, Comput. Sci. Eng., 9, 21
Pillepich A. et al., 2018a, MNRAS, 473, 4077
Pillepich A. et al., 2018b, MNRAS, 475, 648
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Price S. H., Kriek M., Feldmann R., Quataert E., Hopkins P. F., Faucher-Giguère C.-A., Kereš D., Barro G., 2017, ApJ, 844, L6
Prole D. J., 2021, MNRAS, 506, L59
Rodriguez F., Montero-Dorta A. D., Angulo R. E., Artale M. C., Merchán M., 2021, MNRAS, 505, 3192

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

Sales L. V., Navarro J. F., Peñafiel L., Peng E. W., Lim S., Hernquist L., 2020, MNRAS, 494, 1848
Samuel J. et al., 2020, MNRAS, 491, 1471
Schaye J. et al., 2015, MNRAS, 446, 521
Schmidt M., 1959, ApJ, 129, 243
Shibuya T., Ouchi M., Harikane Y., 2015, ApJS, 219, 15
Silk J., Rees M. J., 1998, MNRAS, 324, 128
Simha V., Weinberg D. H., Davé R., Fardal M., Katz N., Oppenheimer B. D., 2012, MNRAS, 423, 3458
Somerville R. S. et al., 2008, ApJ, 672, 776
Somerville R. S. et al., 2018, MNRAS, 473, 2714
Somerville R. S., Davé R., 2014, ARAA, 53, 51
Springel V., 2005, MNRAS, 364, 1105
Springel V., 2010, MNRAS, 401, 791
Stewart K. R. et al., 2017, ApJ, 843, 47
Suess K. A., Kriek M., Price S. H., Barro G., 2019a, ApJ, 877, 103
Suess K. A., Kriek M., Price S. H., Barro G., 2019b, ApJ, 885, L22
Tremmel M., Karcher M., Governato F., Volonteri M., Quinn T. R., Pontzen A., Anderson L., Bellovary J., 2017, MNRAS, 470, 1121
Tremmel M., Wright A. C., Brooks A. M., Munshi F., Nagai D., Quinn T. R., 2020, MNRAS, 497, 2786
van de Ven G., van der Wel A., 2021, ApJ, 914, 45
Van Der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
van der Wel A. et al., 2014, ApJ, 788, 28
Virtanen P. et al., 2020, Nature Methods, 17, 261
Vogelsberger M. et al., 2014, MNRAS, 444, 1518
Wang L., Dutton A. A., Stinson G. S., Macciò A. V., Penzo C., Kang X., Keller B. W., Wadsley J., 2015, MNRAS, 454, 83
Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52
Wechsler R. H., Gross M. A. K., Primack J. R., Blumenthal G. R., Dekel A., 1998, ApJ, 506, 19
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguere C.-A., Keres D., Quataert E., 2016, ApJL, 827, L23
Wheeler C. et al., 2017, MNRAS, 465, 2420
Wheeler C., Oñorbe J., Bullock J. S., Boylan-Kolchin M., Elbert O. D., Shea Garrison-Kimmel M., Hopkins P. F., Kereš D., 2015, MNRAS, 453, 1305
White M., Zheng Z., Brown M. J. I., Dey A., Jannuzi B. T., 2007, ApJ, 655, L69
White S. D. M., Frank C. S., 1991, ApJ, 379, 52
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Yang X., Mo H. J., Van den Bosch F. C., 2009, ApJ, 695, 900
Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153
Zanisi L. et al., 2020, MNRAS, 492, 1671
Zanisi L. et al., 2021a, MNRAS, 505, 4555
Zanisi L., Shankar F., Bernardi M., Mei S., Huertas-Company M., 2021b, MNRAS, 505, L84
Zjupa J., Springel V., 2017, MNRAS, 466, 1625
Zolotov A. et al., 2015, MNRAS, 450, 2327

APPENDIX A: GALAXY-SIZE DEFINITION CHECKS

We check our definition of galaxy size – $M_\star(< R_{1/2}) = 0.5M_\star(< 0.1R_{\rm vir})$ – versus an iterative approach. Hopkins et al. (2018) calculate an initial three-dimensional half-stellar mass radius $R_{1/2,\,0}$ within a large cutoff at $0.15R_{\rm vir}$. Then, they define an intermediate total radius $R_{\rm tot} = 3R_{1/2,\,0}$ and recalculate the half-stellar mass radius: $M_\star(< R_{1/2,\,1}) = 0.5M_\star(< R_{\rm tot}) = 0.5M_\star(< 3R_{1/2,\,0})$. They repeat this process until convergence

$$\frac{|R_{1/2,n} - R_{1/2,n+1}|}{R_{1/2,n+1}} < \epsilon \tag{A1}$$

for some tolerance ϵ, where n is the number of iterations. We choose $\epsilon = 10^{-5}$, which typically takes $n \lesssim 10$ iterations. We require convergence within $n_{\rm max} = 100$ iterations, within which every galaxy-size estimate converges.

Fig. A1 details the similarities between these definitions, where the left-hand panels compare the galaxy sizes $R_{1/2}$ directly and the right-hand panels the stellar masses $M_\star$ at $z = 0$ (top) and $z = 2$ (bottom). In general, there is good agreement between the definitions, except at the smallest and largest galaxy sizes. Hopkins et al. (2018) warn that at lower redshifts our definition fails because satellites can exist within $0.1R_{\rm vir}$, but this typically affects more massive galaxies. We also check our results using all stellar material within $0.2R_{\rm vir}$ and find qualitatively similar results. The top panels at $z = 0$ show good agreement between the two galaxy-size definitions.

We repeat Table 2 and Fig. 7 using the iterative galaxy definition in Table A1 and Fig. A2. The same qualitative results hold, except that there is larger overall scatter $\langle\sigma\rangle = 0.146$ [dex] compared to that in Fig. 7 ($\langle\sigma\rangle = 0.085$ [dex]). At any given instant in time, the GHSR is roughly linear, but the size of a given halo typically grows faster than the size of its central galaxy over much of cosmic history.

As a final check for our definition of $R_{1/2}$, we plot the stellar mass radial profiles in Fig. A3 for the galaxies in the 1024^3 run at redshift $z = 0$ (left) and $z - 2$ (right). The median stellar mass curves flatten to $M_\star \approx M_\star(< R_{\rm vir})$ (horizontal dashed line) by the total galaxy size $R \approx 0.1R_{\rm vir}$ (vertical dashed line), affirming that this is a robust value of the total galaxy size.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

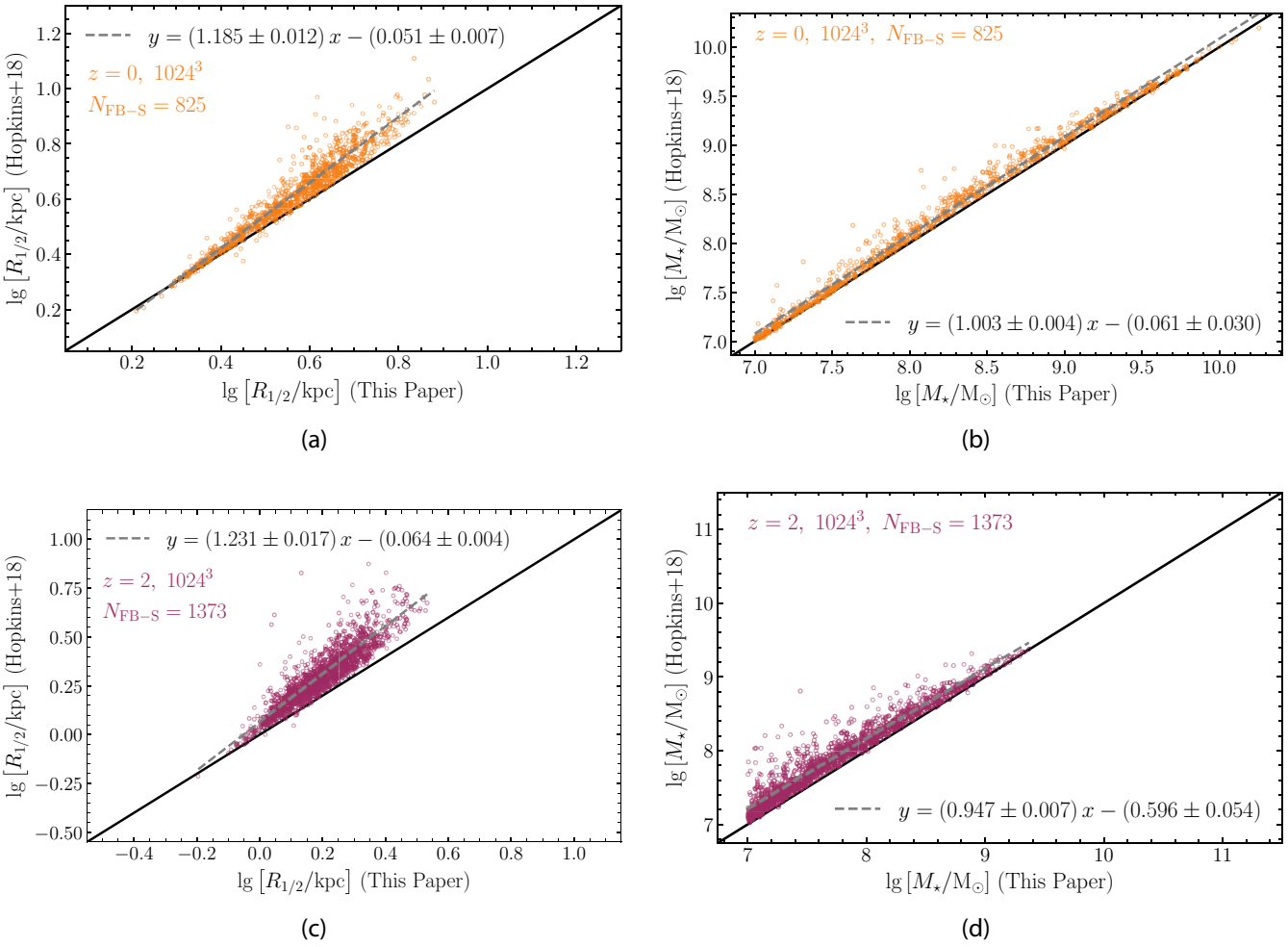

Figure A1. A comparison between galaxy-size definitions in the $z = 0$ (top) and $z = 2$ (bottom) snapshots. The horizontal axes denote the values used in this paper, while the vertical axes the values using the iterative approach (Hopkins et al. 2018). The best-fitting power-law equations are given in each panel. The left-hand panels display the galaxy sizes $R_{1/2}$, and the right-hand panels the stellar masses $M_\star$. In general, the values of the galaxy sizes and stellar masses calculated using the two methods agree.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

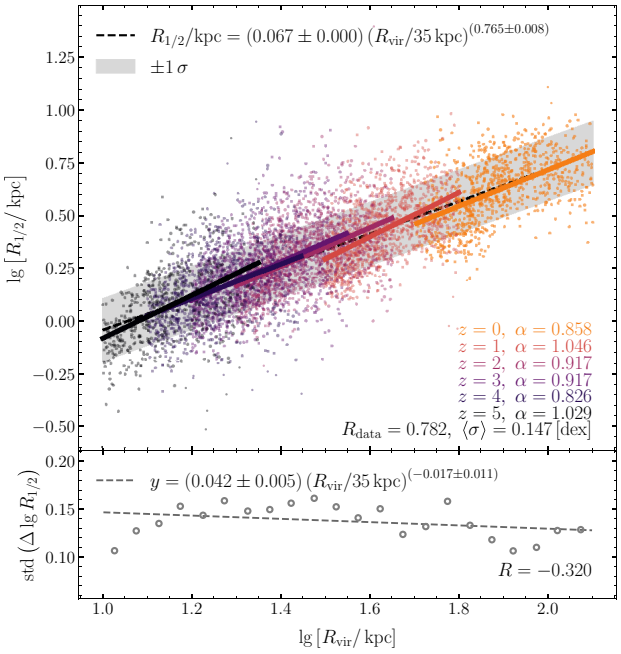

Figure A2. Similar to Fig. 7, but we use the Hopkins et al. (2018) iterative definition of galaxy size. *Top panel*: the colour of the points represents their redshift (see legend), and the best-fitting power-law index α of the GHSR at that redshift is listed on the bottom right. The linear least-squares fit to the combined (all$-z$) data is shown at the top of the panel. The Pearson correlation coefficient and the average scatter are listed at the bottom. The power-law relation of the combined data set is sub-linear. *Bottom panel*: The standard deviation of the residuals for each lg R_{vir} bin of width 0.050 [dex] from the GHSR in the upper panel. The equation of the best fit and the Pearson correlation coefficient are shown at the top and bottom of the panel. The average scatter $\langle\sigma\rangle = 0.147$ [dex] is larger than that in Fig. 7 $\langle\sigma\rangle = 0.085$ [dex].

Table A1. GHSR at each and all redshifts for the iterative definition of galaxy size from Hopkins et al. (2018) for the FIREbox 1024^3 'FB-S' galaxies.

z (1)	α_{GHSR} (2)	β_{GHSR} (3)	$\langle\sigma\rangle$ (4)	α_σ (5)	β_σ (6)
0	0.858 ± 0.049	0.061 ± 0.003	0.131	-0.020 ± 0.055	0.042 ± 0.027
1	1.046 ± 0.051	0.063 ± 0.001	0.134	0.134 ± 0.051	0.023 ± 0.022
2	0.917 ± 0.043	0.070 ± 0.001	0.137	0.259 ± 0.045	0.016 ± 0.004
3	0.917 ± 0.050	0.074 ± 0.002	0.151	0.368 ± 0.023	0.012 ± 0.001
4	0.826 ± 0.068	0.070 ± 0.004	0.166	0.473 ± 0.035	0.010 ± 0.001
5	1.029 ± 0.093	0.086 ± 0.007	0.174	0.461 ± 0.041	0.013 ± 0.002
Mean[a]	0.930 ± 0.076	0.070 ± 0.018	0.146	0.26 ± 0.20	0.013 ± 0.011
All$-z$[b]	0.765 ± 0.008	0.067 ± 0.001	0.147	-0.017 ± 0.011	0.042 ± 0.005

[a]The averages using the number of objects at each redshift as weights.
[b]The combined sample of all redshift snapshots, treating objects from different snapshots equally.
The same as Table 2, here using the iterative definition of galaxy size (Hopkins et al. 2018). (1) The redshift; (2) and (3) the power-law index and normalization for the GHSR from equation (2); (4) the average scatter in the GHSR; and (5) and (6) the power-law index and normalization for the GHSR scatter as a function of lg R_{vir} bin.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

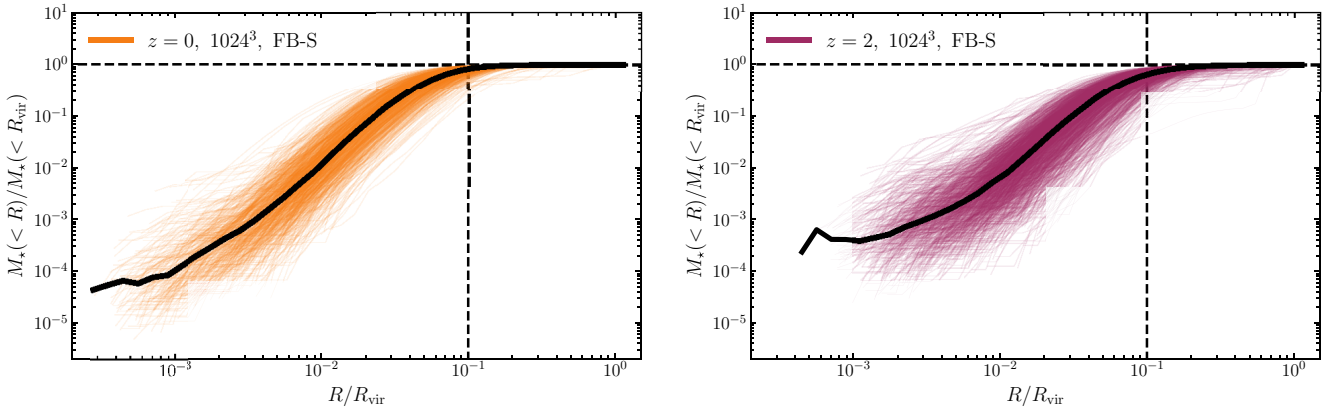

Figure A3. The stellar mass profiles for the galaxies in FIREbox pathfinder at redshift $z = 0$ (left) and $z = 2$ (right). The thick black curves show the median values within $R/R_{\rm vir}$ bins of width 0.1 [dex]. We normalize the stellar masses $M_\star(< R)$ by the total stellar mass within the halo $M_\star(< R_{\rm vir})$, and the radius R by virial radius $R_{\rm vir}$. We mark where $M_\star(< R) = M_\star(< R_{\rm vir})$ and $R = 0.1R_{\rm vir}$ with horizontal and vertical dashed lines, respectively. The median curves flatten to $M_\star \approx M_\star(< R_{\rm vir})$ by $0.1R_{\rm vir}$, affirming that this cutoff radius contains most of the haloes' stellar mass.

APPENDIX B: GALAXY-SIZE STELLAR MASS RELATIONS

In Fig. 2, we use a correction from Suess et al. (2019a) to convert our 3D intrinsic half-mass radii $R_{1/2}$ to 2D projected effective half-light radii $R_{\rm eff}$ at $z = 2$, assuming that the 3D $R_{1/2}$ has on average a similar value to the 2D half-mass radius $R_{\rm mass}$ (van de Ven & van der Wel 2021). The completeness limit for Suess et al. (2019a) is $M_\star \gtrsim 10^{10}\,{\rm M}_\odot$, and it is unclear if this holds in the stellar mass regime of the 'FB-S' galaxies ($M_\star \sim 10^{7-9}\,{\rm M}_\odot$). However, the correction does not significantly affect the results at $z = 2$, where $\langle R_{1/2}/R_{\rm eff}\rangle \sim 0.9$ for both star-forming and quiescent galaxies. Without this correction, the 'FB-S' galaxies are still within the scatter of the star-forming relation.

In Fig. B1, we compare our galaxy size–stellar mass relations to those observed at $z = 0$ (Nedkova et al. 2021) in the Hubble Frontier Fields (HFF) and Cosmic Assembly Near-infrared Deep Extragalactic Survey (CANDELS). In the left-hand panel, we approximate the effective radius $R_{\rm eff}$ as our half-mass radius $R_{1/2}$; in the right-hand panel, we convert $R_{1/2}$ to $R_{\rm eff}$ using the correction factor from Suess et al. (2019b), assuming $R_{1/2} \sim R_{\rm mass}$. Specifically, Suess et al. (2019b) find for both quiescent and star-forming galaxies of stellar mass $M_\star > 10^{10.1}\,{\rm M}_\odot$ at redshift $z \lesssim 1$ that $\langle R_{\rm mass}/R_{\rm eff}\rangle \sim 0.7$. It is unclear if this correction holds at the stellar masses of the 'FB-S' galaxies ($M_\star \sim 10^{7-9.5}\,{\rm M}_\odot$) at $z = 0$. At a fixed stellar mass, the 'FB-S' galaxies systematically have a half-mass radius $R_{1/2}$ larger than the star-forming galaxies from Nedkova et al. (2021). At stellar masses $M_\star \sim 10^{9-10}\,{\rm M}_\odot$, the 'FB-S' and 'FB-L' galaxy sizes flatten out. Then, the 'FB-L' galaxy sizes turnover at $M_\star \sim 10^{10}\,{\rm M}_\odot$, where $R_{1/2}$ starts to decrease with increasing stellar mass. Without the correction from Suess et al. (2019b, left-hand panel), the 'FB-S' galaxy sizes are systematically ~ 0.3 [dex] larger than observations. Then, the 'FB-L' galaxies lie partly between the galaxy size–stellar mass relations for star-forming and quiescent galaxies. With the correction (right-hand panel), the galaxy size offset is ~ 0.5 [dex], and the 'FB-L' galaxies are consistent with the size–mass relation for star-forming galaxies.

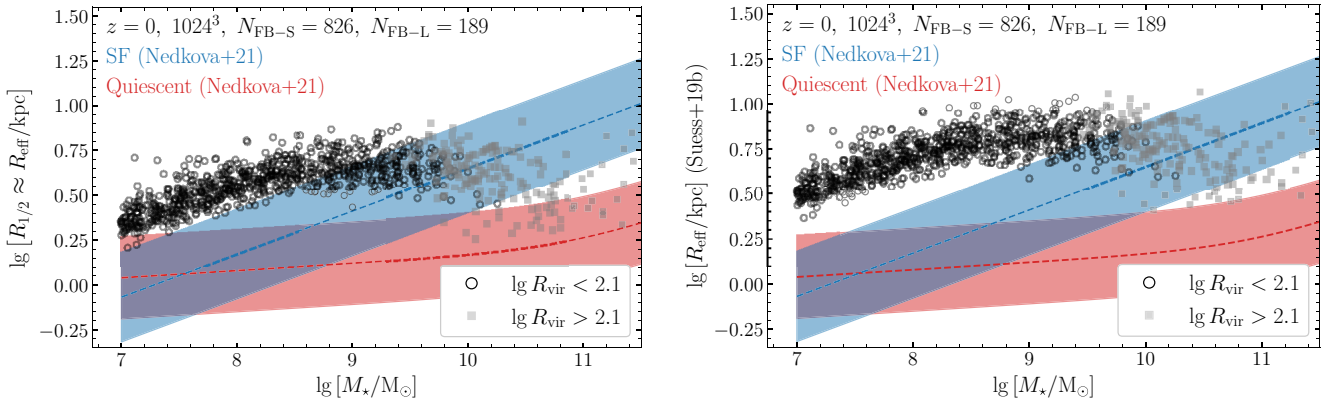

Figure B1. Similar to Fig. 2. A comparison to the galaxy size–stellar mass relation obtained from the Hubble Frontier Fields (HFF) and Cosmic Assembly Near-infrared Deep Extragalactic Survey (CANDELS) at $z = 0$ (Nedkova et al. 2021). In the left-hand panel, we approximate the effective radius $R_{\rm eff}$ from Nedkova et al. (2021) with our half-mass radius $R_{1/2}$; in the right-hand panel, we convert $R_{1/2}$ to $R_{\rm eff}$ using the correction factor from Suess et al. (2019b). Specifically, Suess et al. (2019b) find for both quiescent and star-forming galaxies of stellar mass $M_\star > 10^{10.1}\,{\rm M}_\odot$ at redshift $z \lesssim 1$ that the 2D half-mass radius $R_{\rm mass} \sim 0.7R_{\rm eff}$. The 'FB-S' galaxies follow a similar galaxy size–stellar mass relation to the star-forming galaxies from Nedkova et al. (2021), but the 'FB-S' galaxies are systematically larger by ~ 0.5 [dex].

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

APPENDIX C: ADDITIONAL HALO PROPERTIES AND THE GHSR SCATTER

Similar to Table 3, Table C1 summarizes the correlation fits between the GHSR scatter and additional halo properties at $z = 2$.

The top set of parameters correlates the GHSR residuals with the difference between the property of a given halo and the mean property of all haloes within the lg $R_{\rm vir}$ bin of width 0.050 [dex]. Similar to the direct value of the halo property, only the centre-of-mass offset has a slope significantly different from 0, but no offset parameters decrease the scatter by > 5 per cent.

So far, the halo properties are always from the hydrodynamic simulation, but perhaps the baryonic properties within the halo affect the dark matter properties, specifically the spin and concentration. We run FIREbox with dark matter only (DMO) and with all particles (ALL-hydro), and we create a mapping between haloes in the hydro and DMO simulations.

First, we create two AHF catalogues from the hydro simulation: one using only dark matter particles (DMO-hydro), one using all particles (ALL). We correlate haloes between these catalogues using the spatial coordinates of the halo centres. Specifically, we create a sphere of radius $R_{\rm vir}$ centred on the ALL-halo centre's coordinates in the DMO-hydro catalogue and tabulate every DMO-hydro halo within this sphere. Then, we choose the halo that is most similar in mass to the ALL-halo as the best match. This completes the mapping between the two AHF catalogues generated from the hydro simulation.

Table C1. Correlations of the 'FB-S' GHSR's scatter with additional halo properties at all redshifts.

Parameter (1)	m (2)	$\sigma_{m,0}$ (3)	b (4)	R (5)	$\langle\sigma\rangle$ (6)	$\%\Delta\sigma$ (7)
Δlogit[b/a (halo, $R_{\rm vir}$)]	0.016 ± 0.003	4.76	−0.001 ± 0.001	0.055	0.085 ± 0.001	0.2
Δlogit[c/a (halo, $R_{\rm vir}$)]	0.027 ± 0.005	5.17	−0.001 ± 0.001	0.058	0.085 ± 0.001	0.2
Δlogit[c/b (halo, $R_{\rm vir}$)]	0.004 ± 0.004	1.09	−0.000 ± 0.001	0.013	0.085 ± 0.001	0.0
$\Delta\text{logit}\left[E \equiv \sqrt{1-(b/a)^2}\ (\text{halo}, R_{\rm vir})\right]$	−0.021 ± 0.004	4.99	−0.001 ± 0.001	−0.058	0.085 ± 0.001	0.2
$\Delta\text{logit}\left[F \equiv \sqrt{1-(c/b)^2}\ (\text{halo}, R_{\rm vir})\right]$	−0.005 ± 0.005	1.08	−0.000 ± 0.001	−0.013	0.085 ± 0.001	0.0
$\Delta\text{logit}\left[T \equiv \left(1-(b/a)^2\right) / \left(1-(c/a)^2\right)\ (\text{halo}, R_{\rm vir})\right]$	−0.007 ± 0.002	3.08	−0.000 ± 0.001	−0.037	0.085 ± 0.001	0.1
$\Delta \lg \lambda$	0.030 ± 0.004	6.93	0.002 ± 0.001	0.103	0.084 ± 0.001	0.4
$\Delta \lg \lambda_e$	0.028 ± 0.005	6.11	0.002 ± 0.001	0.086	0.084 ± 0.001	0.3
$\Delta \lg$ cNFW	0.000 ± 0.008	0.06	0.000 ± 0.001	−0.013	0.085 ± 0.001	0.0
$\Delta \lg \left[\sigma_v/(\text{km s}^{-1})\right]$	0.010 ± 0.011	0.99	0.000 ± 0.001	0.030	0.085 ± 0.001	0.0
$\Delta \lg \left[\Delta\text{COM/kpc}\right]$	0.063 ± 0.004	18.02	0.007 ± 0.001	0.247	0.082 ± 0.001	2.7
dmo (logit[b/a (halo, $R_{\rm vir}$)])	0.011 ± 0.003	3.32	−0.009 ± 0.003	0.044	0.084 ± 0.001	0.1
dmo (logit[c/a (halo, $R_{\rm vir}$)])	0.017 ± 0.005	3.54	−0.009 ± 0.003	0.047	0.084 ± 0.001	0.1
dmo (logit[c/b (halo, $R_{\rm vir}$)])	0.003 ± 0.004	0.76	−0.003 ± 0.004	0.010	0.084 ± 0.001	0.0
$\text{dmo}\left(\text{logit}\left[E \equiv \sqrt{1-(b/a)^2}\ (\text{halo}, R_{\rm vir})\right]\right)$	−0.015 ± 0.004	3.58	−0.000 ± 0.001	−0.048	0.084 ± 0.001	0.1
$\text{dmo}\left(\text{logit}\left[F \equiv \sqrt{1-(c/b)^2}\ (\text{halo}, R_{\rm vir})\right]\right)$	−0.004 ± 0.005	0.78	−0.000 ± 0.001	−0.010	0.084 ± 0.001	0.0
$\text{dmo}\left(\text{logit}\left[T \equiv \left(1-(b/a)^2\right) / \left(1-(c/a)^2\right)\ (\text{halo}, R_{\rm vir})\right]\right)$	−0.006 ± 0.002	2.32	0.001 ± 0.001	−0.031	0.084 ± 0.001	0.0
dmo ($\lg \lambda$)	0.022 ± 0.009	2.39	0.029 ± 0.016	0.048	0.084 ± 0.002	0.1
dmo ($\lg \lambda_e$)	0.022 ± 0.008	2.70	0.032 ± 0.014	0.049	0.085 ± 0.002	0.1
dmo ($\lg$ cNFW)	−0.018 ± 0.004	4.02	0.014 ± 0.004	−0.054	0.084 ± 0.001	0.1
$\text{dmo}\left(\lg \left[\sigma_v/(\text{km s}^{-1})\right]\right)$	0.001 ± 0.010	0.09	−0.002 ± 0.018	0.001	0.084 ± 0.001	0.0
$\text{dmo}\left(\lg \left[\Delta\text{COM/kpc}\right]\right)$	0.009 ± 0.003	2.98	−0.005 ± 0.002	0.040	0.084 ± 0.001	0.1
dmo (Δlogit[b/a (halo, $R_{\rm vir}$)])	0.013 ± 0.003	3.85	−0.001 ± 0.001	0.044	0.084 ± 0.001	0.1
dmo (Δlogit[c/a (halo, $R_{\rm vir}$)])	0.023 ± 0.005	4.32	−0.001 ± 0.001	0.047	0.084 ± 0.001	0.2
dmo (Δlogit[c/b (halo, $R_{\rm vir}$)])	0.004 ± 0.004	1.03	−0.000 ± 0.001	0.010	0.084 ± 0.001	0.0
$\text{dmo}\left(\Delta\text{logit}\left[E \equiv \sqrt{1-(b/a)^2}\ (\text{halo}, R_{\rm vir})\right]\right)$	−0.018 ± 0.004	4.12	−0.001 ± 0.001	−0.048	0.084 ± 0.001	0.2
$\text{dmo}\left(\Delta\text{logit}\left[F \equiv \sqrt{1-(c/b)^2}\ (\text{halo}, R_{\rm vir})\right]\right)$	−0.005 ± 0.005	1.05	−0.000 ± 0.001	−0.010	0.084 ± 0.001	0.0
$\text{dmo}\left(\Delta\text{logit}\left[T \equiv \left(1-(b/a)^2\right) / \left(1-(c/a)^2\right)\ (\text{halo}, R_{\rm vir})\right]\right)$	−0.006 ± 0.002	2.49	−0.000 ± 0.001	−0.031	0.084 ± 0.001	0.1
dmo ($\Delta \lg \lambda$)	0.017 ± 0.009	1.89	−0.003 ± 0.003	0.048	0.084 ± 0.002	0.1
dmo ($\Delta \lg \lambda_e$)	0.020 ± 0.008	2.44	−0.001 ± 0.002	0.049	0.085 ± 0.002	0.1
dmo ($\Delta \lg$ cNFW)	−0.030 ± 0.007	4.48	−0.001 ± 0.001	−0.054	0.084 ± 0.001	0.2
$\text{dmo}\left(\Delta \lg \left[\sigma_v/(\text{km s}^{-1})\right]\right)$	−0.015 ± 0.011	1.34	−0.000 ± 0.001	0.001	0.084 ± 0.001	0.0
$\text{dmo}\left(\Delta \lg \left[\Delta\text{COM/kpc}\right]\right)$	0.005 ± 0.003	1.68	0.001 ± 0.001	0.040	0.084 ± 0.001	0.0

Note. The same as Table 3, except only for the alternative halo properties. (1) The parameter used as the horizontal axis; (2) the slope of the fit; (3) the σ from the slope being 0 using the statistical error of m; (4) the vertical offset in the fit; (5) the Pearson correlation coefficient; (6) the scatter in regression in the residuals versus parameter; and (7) the percentage difference in scatter between the GHSR and the residual–parameter relation. From top to bottom, the sections are the difference between the value and the average value of similar-massed haloes, the value of the crossmatched halo in the dark matter only (DMO) simulation, and the difference between the crossmatched halo value and the average value of all similar-massed DMO haloes. Section 5.1 describes what each parameter is.

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022

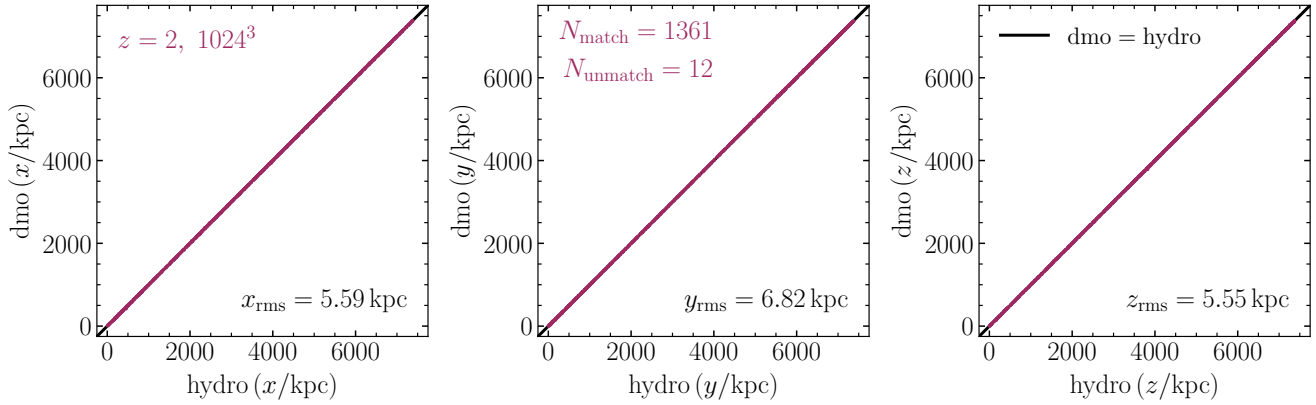

Figure C1. A comparison between the halo centres in the hydrodynamical (horizontal axis) and DMO simulation (vertical axis). There is strong agreement in the halo centres (shown above) and the masses (not shown).

Next, we map the AHF catalogue from the DMO simulation with that of the DMO-hydro using the dark matter particle IDs. We use AHF MERGERTREE to link DMO-hydro haloes to those of the DMO. The likelihood of connection is

$$\max_{\rm j}\left({\rm M_{ij}} = \frac{{\rm N_{i\cap j}^2}}{{\rm N_i N_j}}\right), \tag{C1}$$

where $\rm M_{ij}$ is the merit function, $\rm N_{i\cap j}$ is the number of shared particles between a DMO and its corresponding DMO-hydro halo, while $\rm N_i$ and $\rm N_j$ are the total number of particles in the DMO and DMO-hydro haloes, respectively. This completes the mapping between the DMO-hydro and DMO-halo catalogues.

Lastly, we combine the two maps to correlate the ALL-hydro and DMO haloes. We require that the DMO halo is a main halo, since the ALL-hydro halo is by a main halo by definition. Then, we require the DMO virial mass be within a factor of 2 of that of the ALL-hydro halo. There are no explicit requirements on the halo centres between these catalogues, and we use this as a final check of accurate mapping. Fig. C1 details the positions, and we find that the halo centres are in great agreement. We successfully map 1361/1373 haloes at $z - 2$, where each halo is central and has similar size, mass, and coordinates. The same results hold at the other redshifts.

Thus, we correlate the GHSR residuals with the pristine dark matter properties of the matched haloes in the DMO simulation (middle set of Table C1). Nevertheless, the DMO halo properties do not decrease the GHSR scatter, and the COM offset has a weaker significance. We expect that the lack of satellite galaxies causes the typical COM values to be smaller, causing the weaker correlation.

The bottom set combines the techniques of the second and third sets of correlations by correlating the GHSR residuals with the difference between the DMO property of a given halo and the mean property of all DMO haloes in the same $R_{\rm vir}$ bin. Agreeing with our previous results, we find that no studied halo property significantly explains the scatter in the GHSR in the 'FB-S' galaxies since $z = 5$.

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

Downloaded from https://academic.oup.com/mnras/article/510/3/3967/6462922 by University of Texas Libraries user on 22 January 2022