Desmond, H, Mao, Y-Y, Wechsler, R, Crain, RA and Schaye, J

On the galaxy-halo connection in the EAGLE simulation

http://researchonline.ljmu.ac.uk/6851/

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)


LJMU has developed LJMU Research Online for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

http://researchonline.ljmu.ac.uk/
On the galaxy–halo connection in the EAGLE simulation

Harry Desmond,1,2,⋆ Yao-Yuan Mao,3,4 Risa H. Wechsler,1,2 Robert A. Crain5 and Joop Schaye6

1Physics Department, Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Stanford, CA 94305, USA
2SLAC National Accelerator Laboratory, Menlo Park, CA 94025, USA
3Department of Physics and Astronomy, University of Pittsburgh, Pittsburgh, PA 15260, USA
4Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), Pittsburgh, PA 15260, USA
5Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK
6Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands

Accepted 2017 June 9. Received 2017 May 22; in original form 2016 December 3

ABSTRACT

Empirical models of galaxy formation require assumptions about the correlations between galaxy and halo properties. These may be calibrated against observations or inferred from physical models such as hydrodynamical simulations. In this Letter, we use the EAGLE simulation to investigate the correlation of galaxy size with halo properties. We motivate this analysis by noting that the common assumption of angular momentum partition between baryons and dark matter in rotationally supported galaxies overpredicts both the spread in the stellar mass–size relation and the anticorrelation of size and velocity residuals, indicating a problem with the galaxy–halo connection it implies. We find the EAGLE galaxy population to perform significantly better on both statistics, and trace this success to the weakness of the correlations of galaxy size with halo mass, concentration and spin at fixed stellar mass. Using these correlations in empirical models will enable fine-grained aspects of galaxy scalings to be matched.

Key words: galaxies: formation – galaxies: fundamental parameters – galaxies: haloes – galaxies: kinematics and dynamics – galaxies: statistics – dark matter.

1 INTRODUCTION

Accurate semi-analytic and empirical modelling of galaxy formation is challenging, in part because the correlations of key galaxy and halo variables remain unknown. Observational manifestations of these correlations include galaxy scaling relations, and through detailed investigations of these relations we may hope to build knowledge of the galaxy–halo connection.

Over the past decades, two models that have proven useful for capturing aspects of the galaxy–halo connection are subhalo abundance matching (SHAM; Kravtsov et al. 2004; Behroozi, Conroy & Wechsler 2010) and the angular momentum model of Mo, Mao & White (1998, hereafter MMW). SHAM asserts a nearly monotonic relationship between stellar mass and a halo proxy, establishing the dependence of galaxy mass on halo mass and concentration required to reproduce galaxy clustering (e.g. Conroy, Wechsler & Kravtsov 2006; Reddick et al. 2013). The MMW model sets galaxy and halo specific angular momentum proportional and assumes galaxies’ velocities to be entirely rotational, making galaxy size a function of galaxy mass and halo mass, concentration and spin. This agrees well with observed average galaxy sizes over a wide range of mass (Kravtsov 2013; Desmond & Wechsler 2015, hereafter DW15).

Despite these successes, however, the conjunction of these models (hereafter ‘SHAM+MMW’) is known to make incorrect predictions for two properties of the galaxy population. The first is the scatter sMSR in the stellar mass–size relation (MSR; de Jong & Lacey 2000; Gnedin et al. 2007). SHAM+MMW sets galaxy size proportional to halo spin, λ, and hence requires the scatter in size at fixed mass to be at least as large as that in λ. In fact, these scatter are ∼0.2 and ∼0.25 dex in observed galaxies and simulations, respectively (DW15). The second is the correlation of residuals of the mass–size and mass–velocity relations (ρΔR–ΔV), which is negligible in observations but predicted to be negative (McGaugh 2005; Dutton et al. 2007; DW15). These discrepancies indicate that the galaxy–halo correlations on which sMSR and ρΔR–ΔV depend are inadequately captured by the model.

This issue is relevant also for semi-analytic models. Many such models set galaxy size proportional to halo virial radius (e.g. Croton et al. 2006; Somerville et al. 2008; Lu et al. 2011), sometimes with a single value for all halo spins. Others that use additional physical assumptions find important correlations of size with variables beyond halo mass and spin, but neglect the scatter in sizes (e.g. Lu, Mo & Wechsler 2015). The empirical identification of the aspects of the
EAGLE is a recently completed set of cosmological hydrodynamical simulations, run with a modified version of GADGET-3 (Springel 2005) and including hydrodynamics, radiative cooling, star formation, stellar feedback and black hole dynamics. The subgrid models were calibrated against the present-day stellar mass function and the galaxy catalogue (Springel et al. 2005), which is matched by our procedure (Section 3.2), but to the correlations of halo variables with galaxy size (Section 3.3). In EAGLE, the sizes of low-redshift galaxies are only weakly correlated at fixed stellar mass with the mass, concentration and spin of their haloes, violating the assumption of angular momentum partition. Section 4 discusses the broader implications of these results, and summarizes.

2 SIMULATIONS AND METHODS

2.1 The EAGLE simulation

EAGLE is a recently completed set of cosmological hydrodynamical simulations, run with a modified version of GADGET-3 (Springel 2005) and including hydrodynamics, radiative cooling, star formation, stellar feedback and black hole dynamics. The subgrid models were calibrated against the present-day stellar mass function and the normalization of the MSR. The simulations used a flat Λ cold dark matter cosmology with Ω_m = 0.307, Ω_b = 0.04825, h = 0.6777, σ_8 = 0.8288 and n_s = 0.9611. We analyse the z = 0 snapshot of simulation Ref-L100N1504, which tracks 1504^3 dark matter and gas particles from z = 127 to the present day in a box with comoving side length 100 Mpc, in addition to the corresponding dark matter only (DMO) run in which baryonic effects were switched off. We refer the reader to Schaye et al. (2015) and Crain et al. (2015) for further information about the simulation.

2.2 Finding and matching haloes

To enable direct comparison with the results of DW15, we perform halo finding on both the DMO and hydrodynamical (hereafter ‘hydro’) runs of the EAGLE simulation using ROCKSTAR (Behroozi, Wechsler & Wu 2013). We define spin as \( \lambda \equiv J/E^{1/2}a^{−1/2}M^{−3/2} \), where \( J \) is a halo’s angular momentum and \( E \) its total energy (Peebles 1969), and calculate concentration \( c \) using \( r_s/klypin \) (derived from \( V_{\text{max}}/V_{\text{vir}} \); Klypin et al. 2001) rather than fitting an NFW profile. We include only dark matter when calculating \( c \) and \( \lambda \). We multiply the DMO haloes’ virial masses by \( 1 − \Omega_b/\Omega_m \) to compare to the hydro haloes, where again we include DMO (\( M_{\text{DM}} \)).

Next, we match the DMO ROCKSTAR catalogue to both the hydro ROCKSTAR catalogue and the SUBFIND catalogue (Springel et al. 2001; Dolag et al. 2009) made by the EAGLE pipeline, as follows. Both halo finders produce a list of particles associated with each halo that they identify. Since the two runs share the same dark matter particle IDs, we can match the haloes by finding common particles. In practice, given a halo in the DMO run (halo \( A \)), we first find the halo (halo \( B \)) in the hydro run that contains the most particles of halo \( A \). If halo \( A \) also contains the most particles of halo \( B \), we identify a ‘match’ between them. Since the SUBFIND catalogue of the hydro run also provides the connection between the haloes and galaxies, this method establishes a link between the haloes in the DMO and hydro ROCKSTAR catalogues, and the galaxies in the SUBFIND catalogue. The fraction of haloes in the hydro run hosting galaxies with \( M_*>10^{10} M_{\odot} \) which are matched by our procedure is 91 per cent; these haloes are not significantly biased in \( M_{\text{DM}} \), \( c \) or \( \lambda \).

2.3 Data, models and statistics

We compare our models with the observations of Pizagno et al. (2007, hereafter P07) for compatibility with DW15. Although larger samples with well-measured sizes now exist (e.g. Huang et al. 2017; Somerville et al. 2017), they produce similar MSRs. P07 require an apparent axis ratio \( b/a \leq 0.6 \) and usable H α rotation curve, which they find not to significantly bias the admitted galaxy population in colour or concentration. We therefore do not make a morphology cut on the EAGLE galaxies in our fiducial analysis, although we have checked that our results change at no more than the ∼1σ level – and our qualitative conclusions remain unchanged – when only including galaxies with a substantial fraction of their kinetic energy in ordered corotation (κ_{\alpha} ≥ 0.4; Correa et al. 2017). We compare the EAGLE results with two semi-empirical models, denoted ‘SHAM’ and ‘SHAM+MMW’ as in Section 1.

For both data and models, we take \( s_{\text{MSR}} \) to be the Gaussian scatter in radius of the best-fitting power-law relation

\[
\Delta R = \log(R_s) - \log(R_s)_{\text{fit}}(M_s)
\]

\[\lambda \equiv J/E^{1/2}a^{−1/2}M^{−3/2} \]

\( \Delta R_{\text{eff}} \) decreases with increasing \( x \) and becomes similar to the predictions of \( \lambda \) and \( s_{\text{MSR}} \).

Note that the use of a power law in this definition means that \( s_{\text{MSR}} \) is increased by curvature in the MSR; thus, \( s_{\text{MSR}} \) for the EAGLE relation (see Fig. 1a) may be considered an upper bound on the ‘true’ intrinsic scatter.

1 http://eagle.strw.leidenuniv.nl
We quantify the dependence of $R_{\text{eff}}$ on halo variables with the Spearman correlation coefficients $\rho_{\Delta R-\Delta V}$, where $X \in \{M_{\text{DM}}, c, \lambda\}$. We record in Table 1 the median and 1σ spread of the statistics over 100 Monte Carlo mock data sets of galaxies with $M_*$ values within 0.01 dex of those of the observational sample (see Section 3.1).

### Table 1. Comparison of statistics of the galaxy–halo connection in observations (P07), the EAGLE simulation, an abundance matching model with sizes chosen to match the stellar MSR by construction (‘SHAM’), and an analogous model with sizes set by angular momentum partition (‘SHAM+MMW’; DW15). $s_{\text{MSR}}$ is the scatter in size of the stellar MSR, $\rho$ denotes Spearman rank correlation coefficient and $\Delta$ is defined in equation (1). Entries in italics are by construction. The SHAM+MMW model overpredicts both $s_{\text{MSR}}$ and $\rho_{\Delta R-\Delta V}$ due to the strong correlations it implies between $R_{\text{eff}}$ and $M_{\text{DM}}$, $c$ and $\lambda$ at fixed $M_*$. The EAGLE galaxy–halo connection, in which these variables are only weakly correlated, performs significantly better on both statistics.

<table>
<thead>
<tr>
<th>P07</th>
<th>EAGLE</th>
<th>SHAM</th>
<th>SHAM+MMW</th>
</tr>
</thead>
<tbody>
<tr>
<td>$s_{\text{MSR}}$</td>
<td>0.18</td>
<td>0.15 ± 0.01</td>
<td>0.18</td>
</tr>
<tr>
<td>$\rho_{\Delta R-\Delta V}$</td>
<td>–0.07</td>
<td>–0.13 ± 0.07</td>
<td>–0.22 ± 0.04</td>
</tr>
<tr>
<td>$\rho_{\Delta R-\Delta M_{\text{DM}}}$</td>
<td>–0.18 ± 0.07</td>
<td>0</td>
<td>0.75 ± 0.04</td>
</tr>
<tr>
<td>$\rho_{\Delta R-\Delta c}$</td>
<td>–0.19 ± 0.07</td>
<td>0</td>
<td>–0.76 ± 0.04</td>
</tr>
<tr>
<td>$\rho_{\Delta R-\Delta \lambda}$</td>
<td>–0.17 ± 0.08</td>
<td>0</td>
<td>0.80 ± 0.04</td>
</tr>
</tbody>
</table>

3 RESULTS

#### 3.1 The EAGLE mass–size and $\Delta R-\Delta V$ relations

Fig. 1(a) shows the MSR of the EAGLE galaxies, and Fig. 1(b) the correlation of their size and velocity residuals. That both EAGLE relations are in approximate agreement with the P07 observations is verified quantitatively in the first two rows of Table 1, which list the $s_{\text{MSR}}$ and $\rho_{\Delta R-\Delta V}$ values.

The third and fourth columns of Table 1 show analogous results for two alternative models. In ‘SHAM’, $M_*$ is set by SHAM using the $V_{\text{peak}}$ proxy and 0.2 dex scatter (Reddick et al. 2013; varying the SHAM parameters within the bounds set by clustering has a negligible effect on our results), and galaxy sizes are chosen randomly from a normal distribution at given $M_*$. To match the P07 MSR by construction. In ‘SHAM+MMW’, sizes are set by the MMW model after SHAM has been performed, using the procedure and best-fitting parameter values of DW15.

As mentioned in Section 1 (and discussed in detail in DW15), the SHAM+MMW model compares poorly with observations in both $s_{\text{MSR}}$ and $\rho_{\Delta R-\Delta V}$. This appears to be in conflict with Somerville et al. (2017), who claim the model generates an $s_{\text{MSR}}$ in agreement with that of a compilation of GAMA and CANDELS data. However, they include only the contribution to $s_{\text{MSR}}$ from scatter in $\lambda$ (~0.25 dex) and neglect the contributions from scatter in $M_{\text{DM}}$ and $c$ at fixed $M_*$, $\rho_{\Delta R-\Delta V}$ has contributions both from baryonic mass (higher surface density means larger rotation velocity), and from the dark matter, since in the MMW model more concentrated haloes, which generate larger rotation velocities, host smaller galaxies at fixed angular momentum. The SHAM model, which includes only the first contribution, predicts a $\Delta R-\Delta V$ anticorrelation that is weaker but still stronger than the data’s. It is important to note, however, that these models assume negligible velocity dispersion $\sigma$. A decrease of $\sigma/\sqrt{\text{rot}}$ with $\lambda$ – as produced in some hydrodynamical simulations (e.g. Rodriguez-Gomez et al. 2017) – could reduce the predicted $\rho_{\Delta R-\Delta V}$ and $s_{\text{MSR}}$. Only with the assumption that $\sigma/\sqrt{\text{rot}}$ does not vary systematically with $\lambda$ does the MMW model follow uniquely from proportionality of galaxy and halo specific angular momentum.

We now investigate the origin of the difference between the EAGLE and SHAM+MMW results.

#### 3.2 Comparison of the haloes in the DMO and hydrodynamical runs of the EAGLE simulation

A possible reason for the apparent failure of the SHAM+MMW model is its application in DW15 to haloes from an $N$-body simulation in which baryonic effects were neglected. In Fig. 2, we show the fractional differences in $M_{\text{DM}}$, $c$ and $\lambda$ of all matched haloes in the EAGLE DMO and hydro runs, and compare in the insets their overall distributions. We find the haloes to be a few per cent less massive on average in the hydro run, and their $c$ and $\lambda$ values to be similar. (Schaller et al. 2015 reported larger differences in halo mass because they included baryons as well as dark matter in the mass definition.) The spin distribution is slightly wider in the hydro.

![Figure 1](image-url)  
**Figure 1.** The $M_*-R_{\text{eff}}$ relation and correlation of $R_{\text{eff}}$ and $V_{\text{max}}$ residuals in the EAGLE simulation, compared to the observations of P07. As in DW15, stellar masses for the latter were taken from the NASA Sloan Atlas. The red lines in the left-hand panel show the best-fitting power law to the data, and its scatter. In this plot and those that follow, points indicate medians and error bars 16th and 84th percentiles. We stack the 100 mock data sets (Section 2.3) to make contour plots, and the levels enclose 90, 80, 60, 40 and 20 per cent of galaxies. The spread in the sizes of EAGLE galaxies is as low as is observed, and they correctly exhibit no significant $\Delta R_{\text{eff}}-\Delta V_{\text{max}}$ correlation. In Fig. 1(b), the observations have $\rho_{\Delta R-\Delta V} = –0.07$, and the EAGLE galaxies have $\rho_{\Delta R-\Delta V} = –0.13$. 

run, which goes in the wrong direction to account for the lower $s_{\text{MSR}}$ in EAGLE than in the SHAM+MMW model.

In Fig. 3, we compare the correlations of $\lambda$ with $M_{\text{DM}}$ and $c$ in the two runs, finding them to be very similar. If spin was more positively correlated with $M_{\text{DM}}$ or $c$ with baryonic effects included, then the corresponding increase in rotational velocity caused by dark matter for larger galaxies would compensate for the reduction in the rotation velocity caused by baryons, which could allow the SHAM+MMW model to agree with the measured $\rho_{{\Delta R,\Delta V}}$. That we do not find such an increased correlation leads us to conclude that the differences between the EAGLE and SHAM+MMW models in their predictions for $s_{\text{MSR}}$ and $\rho_{{\Delta R,\Delta V}}$ are not due to changes to the haloes caused by baryons. They must therefore be due to different galaxy–halo correlations.

3.3 The galaxy–halo connection

Rows 3–5 of Table 1 record the Spearman rank coefficients of the correlations between size residual ($\Delta R_{\text{eff}}$) and $M_{\text{DM}}$, $c$, and $\lambda$ residual in the EAGLE, SHAM, and SHAM+MMW models. As halo properties cannot be observed, there are no corresponding entries in the first column. By construction, the SHAM model does not correlate galaxy size with any halo property at fixed stellar mass. As described in Section 1, however, the SHAM+MMW model implies a strong correlation of $\Delta R_{\text{eff}}$ with $\Delta \lambda$ and a strong anticorrelation with $\Delta c$, and the latter in particular is responsible for the strongly negative value of $\rho_{{\Delta R,\Delta V}}$. In the EAGLE simulation, galaxy size correlates only weakly with each halo variable, with the result that the predicted $\rho_{{\Delta R,\Delta V}}$ is similar to the SHAM case. In addition, this prevents $s_{\text{MSR}}$ from receiving the full contributions from the scatter in halo variables at fixed $M_*$, allowing it to remain below the P07 value. These correlations are shown explicitly in Fig. 4.

4 DISCUSSION AND CONCLUSIONS

While the principle component of the galaxy–halo connection – the relation between galaxy mass and halo mass and concentration – is becoming well constrained by abundance matching studies, secondary components, such as the dependence of galaxy size on halo properties, remain uncertain. A leading model for galaxy size (MMW) assumes $\sigma = 0$ and proportionality of galaxy and halo specific angular momenta, making galaxy size a specific function of stellar mass and halo mass, concentration and spin. Despite success in matching the normalization of the stellar MSR when combined with SHAM, this model overpredicts both the spread in sizes ($s_{\text{MSR}}$) and the strength of the correlation of size and velocity residuals ($\rho_{{\Delta R,\Delta V}}$). This indicates a problem with the galaxy–halo connection it implies.

In this Letter, we investigated this discrepancy in the context of the EAGLE hydrodynamical simulation. We found the galaxy population in this simulation to exhibit near-agreement with measurements of both $s_{\text{MSR}}$ and $\rho_{{\Delta R,\Delta V}}$. We showed that this difference with the SHAM+MMW prediction is due not to modifications to the haloes themselves by baryons, but rather to the weakness of the correlations of galaxy size with $M_{\text{DM}}$, $c$, and $\lambda$. While the MMW model strongly correlates $R_{\text{eff}}$ with $M_{\text{DM}}$ ($\rho = 0.75$), $c$ ($\rho = -0.76$), and $\lambda$ ($\rho = 0.80$) at fixed $M_*$, the Spearman rank coefficients for the corresponding EAGLE correlations are only 0.18, 0.19, and 0.17, respectively. These values are consistent with 0 within 3$\sigma$.

Our results have implications for both galaxy formation theory and semi-analytic and empirical modelling. On one hand, the breakdown of the MMW model requires explanation. Galaxy properties may become weakly correlated with halo spin due to stochastic transfer of angular momentum between baryons and dark matter, or a significant loss or redistribution through feedback or cooling processes (Brook et al. 2011; Zjupa & Springel 2017). On the other hand, the EAGLE galaxy–halo correlations may be used to inform...
empirical models where galaxy sizes are added by hand. To match \(s_{\text{MSR}}\) and \(\rho_{\Delta R-\Delta V}\), at least in a SHAM framework, \(R_{\text{eff}}\) should correlate at most weakly with \(M_{\text{DM}}\), \(c\) and \(\lambda\) at fixed \(M_*\). This is tacitly assumed by several existing models (e.g. Dutton et al. 2011, 2013; Desmond & Wechsler 2017), and implied also by aspects of the mass discrepancy–acceleration relation (Desmond 2017). We suggest such correlations be used by default from now on. Finally, our results facilitate the testing of galaxy formation theories: if a theory’s effective galaxy–halo connection exhibits correlations compatible with those of EAGLE, its success in matching the fine-grained statistics that we investigate here is assured.

ACKNOWLEDGEMENTS

We thank Simon Foreman, Yu Lu and Matthieu Schaller for comments on the manuscript, and Matthieu Schaller for guidance with the EAGLE data. This work used the DiRAC Data Centric system at Durham University, operated by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). This equipment was funded by a BIS National E-infrastructure capital grant ST/K00042X/1, STFC capital grant ST/K00087X/1, STFC capital grants ST/K003267/1 and Durham University. DiRAC is part of the National E-Infrastructure.

REFERENCES

Pizagno J. et al., 2007, AJ, 134, 945

This paper has been typeset from a \TeX/\LaTeX file prepared by the author.