We use a simple model of disc-galaxy formation in the ΛCDM cosmology to simultaneously reproduce the slopes, zero-points, scatter and uncorrelated residuals in the observed velocity-luminosity (VL) and radius-luminosity (RL) relations. Observed I-band luminosities are converted to stellar masses using the IMF-dependent relation between stellar mass-to-light ratio, Υ, and color. The model treats halo concentration, spin parameter, and disc mass fraction as independent log-normal random variables. Our main conclusion is that the VL and RL zero-points and the uncorrelated residuals can only be reproduced simultaneously if adiabatic contraction is avoided. One or the other could be fixed by appealing to unrealistic values for Υ, halo concentration c, and spin parameter λ, but not all together. The small VL scatter is naturally determined by the predicted scatter in c and in Υ, quite independent of the large scatter in λ. However, the RL scatter, driven by the scatter in λ, can be as low as observed only if σ_in_λ ≈ 0.25, about half the value predicted for CDM haloes. This may indicate that discs form in a special subset of haloes. The VL slope is reproduced once star formation occurs only above a threshold surface density, and the RL slope implies that the disc mass fraction is increasing with halo mass, consistent with feedback effects. A model that incorporates the above ingredients provides a simultaneous fit to all the observed features. In particular, the elimination of adiabatic contraction allows 80% disc contribution to the observed rotation velocity of bright discs at 2.2 disc scale lengths. The lack of halo contraction may indicate that disc formation is not as smooth as typically envisioned, but instead may involve clumpy, cold streams.
1. Introduction

The correlations between the global structure parameters of disc galaxies, rotation velocity \( V \), exponential radius \( R \), and luminosity \( L \), provide invaluable constraints on the formation process of these galaxies. Although various studies have attempted to explain the origin of the RL or the VL (also known as Tully-Fisher) relation, the true challenge lies in finding a self-consistent model of disc formation that can match both simultaneously. Here we investigate how standard models of disc formation fair in this respect.

We use the observed structural parameters from an \( I \)-band sample of \( \sim 1800 \) bright disc galaxies compiled by Courteau et al. (2003). In order to facilitate a simple comparison with theory, we convert the observed \( I \)-band luminosity into a stellar masses \( M_{\ast} \), using the stellar mass-to-light ratio \( \Upsilon_I \) at a given color as obtained by Bell et al. (2003) and Bell & de Jong (2001). The main uncertainty in this conversion is the normalization of \( \Upsilon_I \), reflecting the unknown stellar initial mass function (IMF). We adopt the IMF advocated by Bell et al. (2003) and note that the most extreme top-heavy IMF realistically considered (the Kroupa-IMF; Kroupa, Tout & Gilmore 1993) has values of \( \log(\Upsilon_I) \) that are only 0.15 dex smaller. We use as joint constraints on our model the observed log slopes and zero-points of the conditional \( V|M_{\ast} \) and \( R|M_{\ast} \) scaling laws, the conditional scatter about each at a given \( M \), and especially the cross-correlation between the corresponding residuals \( \Delta V \) and \( \Delta R \) at \( M_{\ast} \). Throughout we assume \( H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1} \).

2. Disc Formation Models

Our model starts in the spirit of the classical model by Fall & Efstathiou (1980) and Mo, Mao, & White (1998, hereafter MMW). Each model galaxy is specified by four parameters: the total virial mass \( M_{\text{vir}} \), the halo concentration parameter \( c \), the baryon spin parameter \( \lambda \), and the mass fraction that ends up in the disc \( m_d \). The dark-matter is assumed to follow a spherical NFW density profile, with \( c \) that varies systematically with \( M_{\text{vir}} \). The disc, with mass \( M_d = m_d M_{\text{vir}} \), is assumed to have an exponential density profile, whose scale length \( R_d \) follows from \( \lambda, m_d \), and \( M_{\text{vir}} \) based on angular-momentum conservation as outlined in MMW. In this standard model, the halo is assumed to contract adiabatically under the dissipative condensation of the gas following Blumenthal et al. (1986).

A major novel ingredient considered here is the possible avoidance of the adiabatic contraction, hereafter AC (see § 4 for possible physical motivation). Another straightforward novelty adopted here is splitting the disc into a gaseous and a stellar component. Using Toomre’s critical surface density \( \Sigma_{\text{crit}} \), we compute the stellar mass as the disc mass with surface density \( \Sigma \geq \Sigma_{\text{crit}} \), mimicking a star-formation threshold density in agreement with the empirical findings of Kennicutt (1989). Although this is clearly an oversimplification, it helps explaining (a) why the gas mass fraction is higher in lower mass galaxies (Kannappan 2004), (b) why the scale length of the cold gas is larger than that of the stars, and (c) the truncation radii of stellar discs (e.g. van den Bosch 2001).

The parameters \( c, \lambda \) and \( m_d \) are assumed to be independent random variables, each with a log-normal distribution in which the mean could vary systematically with \( M_{\text{vir}} \). In order to pursue the comparison to observations, we construct a large sample of random realizations of model galaxies. For a given \( M_{\text{vir}} \), the halo concentration \( c \) is drawn from a log-normal distribution, with
Table 1: Slopes of observed versus model scaling relations. The upper 2 rows refer to the observations, while the lower 5 rows to the model. $V_{\text{obs}}$ is the observed (inclination corrected) rotation velocity. $V_{3.2}$ is the model circular velocity at $3.2R_s$. $L_I$ is the observed (dust corrected) $I$-band luminosity. $M_s$ and $M_d$ are the model stellar and total disc mass, respectively. $R_s$ and $R_d$ are the corresponding disc scale lengths.

$\bar{\tau} = 12(V_{\text{vir}}/100)^{-0.33}$ and $\sigma_{\ln c} = 0.32$, based on the CDM-simulated haloes (Bullock et al. 2001a). The baryonic spin parameter $\lambda$ is initially drawn from an independent log-normal distribution, with $\bar{\lambda} = 0.035$ and $\sigma_{\ln \lambda} = 0.5$, again based on the simulated properties of CDM haloes (Bullock et al. 2001b), but we are forced to allow lower scatter values (see below). We then draw $m_d$ from an independent log-normal distribution, but consider $m_d(M_{\text{vir}})$ and $\sigma_{\ln m_d}$ as free parameters. Once we compute the disc mass and radius, we derive the stellar surface density profile and evaluate its mass $M_s$, exponential radius $R_s$, and “observed” circular velocity $V_{3.2}$ at $3.2R_s$.

For a fair comparison with the data, we attempt to reproduce similar sampling in the model realizations. We first sample with equal probability in log $M_{\text{vir}}$, and then select a subsample for which the distribution of the derived $V_{3.2}$ resembles that of the data. The model scaling relations and scatter are obtained by linear regressions in log space of $V$ on $M$ and $R$ on $M$.

3. Comparison of Models and Observations

3.1 Slopes

We start our exploration with the simplest MMW model possible. We keep $c$, $\lambda$, and $m_d$ fixed at 12, 0.035 and 0.05 respectively, and compute the (scatter-free) scaling relations VM and RM. The apparent agreement between the predicted slopes and the $I$-band scaling relations, listed in lines 3 and 1 of Table 1 respectively, is misleading. This comparison naively assumes that all the disc mass is transformed into stars and with a constant $\Upsilon_I$. A more realistic comparison of our model is with the relations listed in the second line of Table 1, where $M_s$ replaces $L_I$ in the data. Clearly, the basic MMW model yields steeper slopes than observed. The model VM slope gets smaller when the proper scaling of $c$ with halo mass is incorporated, and it becomes almost as shallow as observed when the star-formation threshold density is imposed. A simultaneous match of the RM slope is obtained when $m_d$ is taken to scale with halo mass, $m_d \propto M_{\text{vir}}^{0.16}$, as qualitatively expected from stellar feedback effects and as inferred from rotation curves (Persic, Salucci & Stel 1996).

3.2 Zero-points

Fig. 1 demonstrates that the model which matches the slopes in §3.1 suffers from severe zero-point offsets compared to the observed relations. This could have been fixed by a $\sim 0.4$ dex de-
Figure 1: Scaling relations for the model which matches the slopes (last row of Table 1, open circles), versus the mean and 2σ scatter of the observed relations (dashed lines). The effects of 2σ variations in $m_d$, $c$ and $\lambda$ are shown, where $\sigma_{\ln m_d} = 0.23$, $\sigma_{\ln c} = 0.32$ and $\sigma_{\ln \lambda} = 0.50$. The arrows indicate the direction of increasing $m_d$, $c$, and $\lambda$.

crease in $\Upsilon_I$, but realistic IMFs limit the correction to 0.15 dex, which we adopt in our models below. Also seen in Fig. 1 is the impact of varying $c$, $\lambda$ and $m_d$ over the 2-σ range of their distributions. Clearly, the VM zero-point is mostly sensitive to $\bar{\tau}$, while the RM zero-point is most sensitive to $\bar{\lambda}$ and to a lesser extent to the other variables (the secondary dependencies are partly due to the effect of AC).

The RM zero-point could have been matched by decreasing $\bar{\lambda}$ or increasing $\bar{m_d}$, but these changes would have left the VM offset unchanged or worsen it. Furthermore, such changes would worsen the conflict with the observed, uncorrelated residuals (see § 3.4 below), and thus fail to provide a simultaneous fit. The VM zero-point could have been matched with the acceptable 0.15 dex decrease in $\Upsilon_I$ plus a factor of 2 decrease in $\bar{\tau}$, 1 but this requires a drastic change of cosmological parameters, and also worsens the problem with the observed, uncorrelated residuals.

The above is a reflection of a well known failure of current semi-analytical models of galaxy formation, being unable to simultaneously fit the luminosity function (and hence the average virial mass-to-light ratio) and the TF zero-point. A VM zero-point match could be obtained if $V_{\text{obs}} \approx V_{\text{vir}}$, but realistic halo concentrations in the standard ΛCDM cosmology with $\sigma_8 = 0.9$ imply $V_{\text{obs}}/V_{\text{vir}} \approx 1.4$, and standard AC actually boosts it up to $V_{\text{obs}}/V_{\text{vir}} \approx 1.6$.

We are led to the conclusion that the only way to obtain a simultaneous match to the slopes and zero-points of both the VM and RM relations, within the realistic range of parameter values in the ΛCDM scenario, and with a hope to match the uncorrelated residuals (§3.4), is by weakening or completely removing the effects of AC on $V_{3.2}$ (see Fig. 3).

3.3 Scatter

The observed scatters about the scaling relations are $\sigma_{\ln}(V|L_I) = 0.14$ and $\sigma_{\ln}(R|L_I) = 0.35$. The measurement errors are estimated to be $\sigma_{\ln} \approx 0.09$ for each of the observables $V$, $R$, and $L_I$. In addition, there is a contribution to the scatter in our evaluated $M_s$ due to scatter in the relation.

---

1Introducing a central density core does not affect $V_{3.2}$ significantly (e.g. Dutton et al. 2003).
between color and $\Upsilon_I$, which we estimate to be $\sigma_{\ln \Upsilon_I} \gtrsim 0.23$ based on Bell et al. (2003). The remainder of the observed scatter is assumed to originate from the intrinsic scatter in $c$, $\lambda$, and $m_d$.

Table 2 illustrates how each of these five sources of scatter contributes to the total scatter in the two relations. Note that although the different sources of scatter are assumed to be uncorrelated, the residuals in the derived quantities ($L_I$, $R_s$, $V_{3.2}$) are partly correlated. This explains why the total scatter is less than the quadratic sum of the scatter in each component.

As evident from Fig. 1 and Table 2, the scatter in $\lambda$ completely dominates the RM scatter. The observed RM scatter places a robust upper limit of $\sigma_{\ln \lambda} \simeq 0.25$, independent of the other properties of the structural correlations. This is roughly half the value found in cosmological simulations, and may indicate that disc formation only occurs in a subset of haloes (with a restricted range in $\lambda$).

In the VL scatter, the contributions of observational errors and scatter in $\Upsilon_I$ and in $c$ are all comparable, while the contribution of $\sigma_{\ln \lambda} = 0.25$ is smaller, and that of $m_d$ is negligible. While it would have been possible for scatter in $\Upsilon_I$ or $c$ alone to account for all the intrinsic VL scatter, each would have resulted in a strong residual correlation, in conflict with the observations (§3.4). We therefore conclude that both $c$ and $\Upsilon_I$ have to contribute comparably to the VL scatter. Note that with a significantly higher $\sigma_{\ln \lambda}$, the residuals about the VL relation would have been correlated with $\lambda$. Given that the radius, and thus the surface brightness, is a strong function of $\lambda$, this would have implied a surface-brightness dependence in the VL relation, contrary to observations (Courteau & Rix 1999; hereafter CR99).

$$\sigma_{\ln \lambda}$$ | Relation | $c$  | $\lambda$ | $m_d$ | $\Upsilon_I$ | err | total | obs |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0.50</td>
<td>$\sigma_{\ln}(R_s</td>
<td>L_I)$</td>
<td>0.09</td>
<td>0.56</td>
<td>0.13</td>
<td>0.07</td>
<td>0.10</td>
<td>0.67</td>
</tr>
<tr>
<td>0.25</td>
<td>$\sigma_{\ln}(R_s</td>
<td>L_I)$</td>
<td>0.32</td>
<td>0.08</td>
<td>0.02</td>
<td>0.07</td>
<td>0.09</td>
<td>0.16</td>
</tr>
<tr>
<td>0.50</td>
<td>$\sigma_{\ln}(V_{3.2}</td>
<td>L_I)$</td>
<td>0.09</td>
<td>0.08</td>
<td>0.02</td>
<td>0.07</td>
<td>0.09</td>
<td>0.16</td>
</tr>
<tr>
<td>0.25</td>
<td>$\sigma_{\ln}(V_{3.2}</td>
<td>L_I)$</td>
<td>0.05</td>
<td>0.08</td>
<td>0.02</td>
<td>0.07</td>
<td>0.09</td>
<td>0.16</td>
</tr>
</tbody>
</table>

Table 2: sources of scatter in the size-luminosity and velocity-luminosity relations. The intrinsic sources of scatter are: $\sigma_{\ln c} = 0.32$; $\sigma_{\ln \lambda} = 0.25$ or 0.50; $\sigma_{\ln m_d} = 0.23$; $\sigma_{\ln \Upsilon_I} = 0.23$, and $\sigma_{\ln \text{err}} = 0.09$.

### 3.4 Residual Correlation

As noticed by CR99, the degree of correlation between the VL and RL residuals at a given $L$ provides a powerful constraint. Courteau et al. (2003) measured a weak anti-correlation between these residuals, fit in log space by a line of slope $S = \Delta \log V / \Delta \log R = -0.07 \pm 0.03$, confirming the original result of CR99. Note that $S = 0$ for completely uncorrelated residuals.

The slope $S$ is connected to the relative contributions of the disc and halo to the potential, and hence to the circular velocity $V_{2.2}$, measured at $2.2R_s$. The disc contributes $V_{\text{disc}}^2 \propto M_d / R_s$, namely $S = -0.5$. A fixed halo, not reacting to the baryonic contraction, contributes $V_{\text{halo}}^2 \propto R_s$, and thus ($S = +0.5$) for an NFW density cusp. One can therefore expect that a weak anti-correlation, as observed, would be obtained with comparable contributions of disc and halo. AC increases the halo contribution to $V_{3.2}$ in a way that anti-correlates with the disc radius, thus lowering $S$ toward small or negative values. In order to maintain a small negative $S$ with AC, the contribution of the halo, therefore, has to be larger. Indeed, assuming AC, CR99 interpreted their uncorrelated residuals as indicating a “sub-maximal” disc with $V_{\text{disc}} / V_{2.2} \simeq 0.55$. Matching the low negative $S$ observed with
our models that include AC, requires very low values of $\bar{\kappa}$, pushing the parameters outside their reasonable ranges. Furthermore, this is at the expense of increasing the zero-point offsets (§3.2), so this model does not provide a simultaneous fit to the residuals and the zero-points.

When the AC is eliminated, our model obeys the zero-point constraints and reproduces the weakly anti-correlated residuals. Fig. 2 illustrates how the various sources of scatter affect the distribution of residuals $\Delta \log V$ and $\Delta \log R$ at a given $M$. When we use the total disc mass and radius, with scatter only in $\lambda$ and $c$, we still find a strong anti-correlation, $S = -0.22$. Adding scatter in $m_d$ reduces it to $S = -0.20$. When considering stellar mass and radius, we obtain a significantly weaker $S = -0.13$; this is because $\Sigma_{\text{crit}}$ links $M_s$ to $\lambda$ (van den Bosch 2000, Firmani & Avila-Reese 2000). When adding the scatter in $\Upsilon_I$, there is another significant reduction to $S = -0.087$, due to the fact that the $\Upsilon_I$ scatter by itself induces a positive correlation of $S = +1$ (bottom-right panel). Finally, when adding the (uncorrelated) measurement errors we obtain the residual distribution shown in Fig. 3, with $S = -0.074$ — in good agreement with the observations.

Fig. 3 shows the joint distribution of structural parameters in our successful model compared to the observed results. The model assumes no adiabatic contraction. It adopts the $c$ distribution of CDM haloes. It assumes $\bar{\lambda} = 0.035$, as for haloes, but $\sigma_{\ln \lambda} = 0.25$, about half the width of the halo distribution. It has $\bar{m}_d = 0.07(M_{\text{vir}}/10^{12}M_\odot)^{0.16}$, with $\sigma_{\ln m_d} = 0.23$. The translation to $M_s$ is done using a $\Upsilon_I$ that corresponds to a Kroupa IMF, with $\sigma_{\Upsilon_I} = 0.23$. A measurement error of $\sigma_{\ln} = 0.09$ is assigned to each of the three observed quantities.

4. Conclusion

We used the slope, zero-point, scatter and residuals of the observed velocity-mass and radius-mass distributions for disc galaxies to place constraints on the basic CDM-based model of galaxy
Figure 3: Distribution of structure parameters for our fiducial model (no AC, Kroupa IMF) versus the observations. The top-right panel refers to the residuals. Colored symbols refer to central surface density $M_{\odot} \text{pc}^{-2}$: $\Sigma_0 < 200$ (black triangles); $200 < \Sigma_0 < 630$ (blue squares); $630 < \Sigma_0 < 2000$ (green circles); and $2000 < \Sigma_0$ (red stars). Dashed lines refer to the mean and 2σ scatter in the data.

formation. By applying these constraints simultaneously, we obtain non-trivial conclusions. Our simple model assumes independent log-normal distributions for 3 random variables at each halo mass: halo concentration $c$, baryonic spin $\lambda$, and disc mass fraction $m_d$. Additional sources of scatter come from the stellar mass-to-light ratio $\Upsilon_I$ and the measurement errors.

Contrary to previous studies, we take account of the fact that discs are not purely stellar, and that the stellar mass-to-light ratio is not constant. For our model-data comparison, we have converted the observed $I$-band luminosities to stellar masses, using the relation between color and $\Upsilon_I$. This immediately demonstrates that even the slopes of the scaling relations are not automatically reproduced from the basic halo virial relation $M_{\text{vir}} \propto V_{\text{vir}}^3 \propto R_{\text{vir}}^3$. We find that for a simultaneous match of the observed slopes, given the $c$ distribution as predicted for CDM haloes, stars have to form only above a threshold density $\Sigma_{\text{crit}}$, and the stellar mass fraction has to increase with halo mass. This is consistent with the expected effects of supernova feedback, known to be more efficient in suppressing star formation in haloes of lower masses (e.g., Dekel & Silk 1986; van den Bosch 2002; Dekel & Woo 2003).

The RM scatter is dominated by the scatter in spin parameter $\lambda$. The relatively small scatter observed requires robustly that $\sigma_{\ln \lambda} \simeq 0.25$, about half the value predicted for typical CDM haloes.
This could be due to a selection effect in the data (e.g. against LSB or S0 galaxies), or may indicate that discs form in a special subset of haloes with a narrower $\lambda$ distribution. Interestingly, haloes with a quiescent history, which are the natural hosts of discs, indeed tend to have a narrower $\lambda$ distribution (D’Onghia & Burkert 2004). Alternatively it may indicate that the baryons somehow developed a narrower $\lambda$ distribution than that of the dark matter haloes. The VM relation is much tighter because it is largely independent of the scatter in $\lambda$. The intrinsic VM scatter is determined by the scatter in both $c$ and $\Upsilon_I$, as each alone violates the constraint from the uncorrelated residuals.

Matching the zero-points of the scaling relations, combined with the weak anti-correlations between the residuals, turns out to be the greatest challenge. It could have been achieved by drastically lowering $c$ relative to the predictions for typical $\Lambda$CDM haloes, or by assuming an unrealistic top-heavy IMF. In both cases the parameters have to be pushed away from their reasonable limits, and the zero-points are matched without reproducing the uncorrelated residuals. One could alternatively achieve weakly anti-correlated residuals by assuming unrealistically low $m_d/\lambda$, but only at the expense of worsening the zero-point offsets.

The only way we found to match simultaneously the zero-points and the uncorrelated residuals is by eliminating the adiabatic response of the dark matter to the condensation of the gas into the central disc. Once the AC is removed, it becomes relatively straightforward to match simultaneously all the constraints (though with a low scatter in $\lambda$). Except for a model with unrealistically low stellar mass-to-light ratios, no other reasonable fix of the model parameters comes even close to simultaneously matching the zero-points and the residual constraints, making our conclusion quite robust. Our successful model with no AC predicts that discs contribute about 50% to $V_{2.2}$ in the smaller, low surface brightness (LSB) galaxies, but as much as 80% in the brighter, high surface brightness (HSB) galaxies. Thus, whereas the LSBs are indeed sub-maximal, with the haloes dominating the inner regions, the HSBs are closer to maximum discs.

In an ongoing work, we attempt to quantify the degree of contraction permitted by the data; yet to be confirmed is our impression that the corrected AC suggested by Gnedin et al. (2004) is not enough. Once the discs are not built by smooth spherical infall, the adiabatic contraction could, in principle, be counter-balanced by a variety of processes, and may even end up as an expansion. One such process is dynamical friction (e.g. El-Zant et al. 2004; Ma & Boylan-Kolchin 2004), which could be important if the discs are built by relatively big clumps (e.g. Maller & Bullock 2004). Indeed, the physics of gas infall in dark haloes suggests that most big discs have formed by clumpy, cold streams rather than smooth infall (Birnboim & Dekel 2003; Dekel & Birnboim 2005).

Two caveats are worth mentioning. First, we should verify using more accurate color information that there is no systematic uncertainty in the conversion of $L_I$ to $M_*$, which could cause an overestimate of $\Upsilon_I$ for a given IMF. Second, our analysis assumed independence of the scatter in the variables at a given $M$. We suspect that $c$ and $\lambda$ may be slightly correlated, e.g. via the halo accretion history (Wechsler et al. 2002), and one could imagine scenarios where $m_d$ is also slightly correlated with the others. To what extent these weak correlations may affect our conclusions remains to be seen.

We thank A. Maller & S.M. Faber for stimulating discussions. AD has been partly supported by ISF 213/02, NASA ATP NAG5-8218, a Miller Professorship at UC Berkeley, and a Blaise Pascal International Chair in Paris. SC recognizes the support of NSERC through a Discovery grant.
References