A Mass Threshold for Galactic Gas Discs by Spin Flips

Avishai Dekel\textsuperscript{1,2*}, Omri Ginzburg\textsuperscript{1}, Fangzhou Jiang\textsuperscript{1}, Jonathan Freundlich\textsuperscript{1}, Sharon Lapiner\textsuperscript{1}, Daniel Ceverino\textsuperscript{3}, Joel Primack\textsuperscript{4}

\textsuperscript{1}Racah Institute of Physics, The Hebrew University, Jerusalem 91904 Israel
\textsuperscript{2}SCIPP, University of California, Santa Cruz, CA 95064, USA
\textsuperscript{3}Departamento de Fisica Teorics, Facultad de Ciencias, Universidad Autonoma de Madrid, Cantoblanco, 28049 Madrid, Spain
\textsuperscript{4}Physics Department, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA

19 December 2019

ABSTRACT
We predict, analytically and by simulations, that gas discs tend to survive only in haloes above a threshold mass $\sim 2 \times 10^{11} M_\odot$ (stellar mass $\sim 10^9 M_\odot$), with only a weak redshift dependence. At lower masses, the disc spins typically flip in less than an orbital time due to mergers associated with a change in the pattern of the feeding cosmic-web streams. This threshold arises from the halo merger rate when accounting for the mass dependence of the ratio of galactic baryons and halo mass. Above the threshold, wet compactions lead to massive central nuggets that allow the longevity of extended clumpy gas rings. Supernova feedback has a major role in disrupting discs below the critical mass, by driving the stellar-to-halo mass ratio that affects the merger rate, by stirring up turbulence and suppressing high-angular-momentum gas supply, and by confining major compactions to the critical mass. Our predictions seem consistent with current observed fractions of gas discs, to be explored by future observations that will resolve galaxies below $10^9 M_\odot$ at high redshifts, e.g. by JWST.

Key words: galaxies: discs — galaxies: evolution — galaxies: formation — galaxies: haloes — galaxies: mergers — galaxies: spirals

1 INTRODUCTION

Galaxies tend to appear as discs, triaxial spheroids and combinations of the two, as well as in irregular morphologies. For massive galaxies, the fraction of quenched spheroids is known to be growing with cosmological time and with mass at the expense of discs, consistent with the common wisdom that discs tend to become spheroids due to dry mergers. We do not address this well-studied phenomenon here, but rather focus on the survival versus disruption of gas discs, especially at lower masses.

On intermediate and small mass scales, where the shape and kinematics are harder to resolve, the observational situation is less clear, hopefully to be explored with future telescopes such as JWST. One wishes to find out the range of disc viability in terms of cosmological epoch and mass. In particular, can star-forming gas discs form and survive at high redshifts? Can they form and survive at low masses?

Discs are fragile. They are expected not to survive under intense mass accretion including mergers, especially major mergers, if the orbital angular momentum and spins of the incoming objects are not aligned with the angular momentum of the disc. A naive estimate guided by the merger rate of dark-matter haloes would lead to the apparent expectation that discs of all masses had hard time surviving at high redshifts. This is because, in the Einstein-deSitter cosmology regime (practically $z > 1$), the characteristic time for the occurrence of halo mergers is $t_{\text{mer}} \propto (1+z)^{-5/2}$ (Neistein & Dekel 2008, Dekel et al. 2013), while the galaxy orbital time is proportional to the Hubble time, $t_{\text{orb}} \propto (1+z)^{-3/2}$. The ratio, $t_{\text{mer}}/t_{\text{orb}} \propto (1+z)^{-1}$, is thus smaller at higher redshifts, indicating that there is a critical redshift above which major mergers occur in every disc orbital time and thus do not allow a long-lived disc.

Contrary to this naive expectation, using below zoom-in cosmological simulations to explore the relative abundance of discs in the plane of halo mass $M_\nu$ and redshift $z$, we find instead a threshold mass for disc viability, at $M_\nu \sim 2 \times 10^{11} M_\odot$, roughly independent of redshift.

We show that this is actually expected when recall-
ing that the ratio of baryonic mass of a galaxy (stars plus cold gas) to its total halo mass is increasing systematically with mass and with redshift. While the external inflow (and merger) rate of mass and angular-momentum that could damage a disc is primarily determined by the total halo mass, the angular momentum of the existing central galaxy that it is affecting is increasing with its baryonic mass. Therefore the damaging relative change in angular momentum is expected to be larger for lower-mass galaxies and at lower redshifts. This should introduce a strong mass dependence in the disc survivability and weaken its redshift dependence. Indeed, when evaluating analytically the disc survivability under angular-momentum flips due to mergers, as a function of mass and redshift, we predict a threshold halo mass at $M_* \sim 2 \times 10^{11} M_\odot$ with only little sensitivity to redshift, consistent with the threshold seen in the simulations.

Another phenomenon that points to a characteristic mass of a similar value is the process of wet compaction into a blue nugget, sometimes driven by mergers and in other times by other mechanisms such as counter-rotating streams (Zolotov et al. 2015; Tackella et al. 2016a; Tomassetti et al. 2016; Dekel, Lapiner & Dubois 2019). We address below, and in more detail in Dekel et al. (2020), how the compaction process may have a role in the transition from disc disruption to the survival of an extended gas disc, then ring, at a similar threshold mass. At low masses, pre compaction, a disc disruption is expected to be driven by inward transport of gas and clumps driven by violent disc instability and possibly also by torques exerted by the prolate inner system. At masses above the critical mass, namely post compaction, we show that an extended disc, actually a ring fed by cold streams of high angular momentum, is expected to survive under the influence of the massive, rounder, post-compaction bulge (or by a massive cuspy halo). The ring survival is evaluated through the disruptive torques exerted by a spiral structure (Dekel et al. 2020).

It is interesting to note that the origin of the threshold mass for extended rotating discs or rings by the above mechanisms may result in several ways from the robust mass dependence of supernova feedback, namely the upper limit for the mass of a halo within which supernova feedback can effectively heat up the gas or remove it and thus suppress star formation (Dekel & Silk 1986). Supernova feedback could directly disrupt discs of low masses by driving excessive turbulence or by halting the supply of new high-angular-momentum gas. Indirectly, the increase of baryon-to-halo mass ratio with mass, which determines the threshold due to merger-driven spin flips, is argued to arise from supernova feedback (Dekel & Silk 1986; Dekel & Woo 2003). Furthermore, the characteristic mass for major compaction events is argued to also be due to the supernova feedback being effective in suppressing compaction events in galaxies of lower masses (Dekel, Lapiner & Dubois 2019).

The paper is organized as follows. In §2 we briefly present the VELA simulations, which are described in Appendix A, and present the threshold mass for discs as seen in the simulations. In §3 we show that the disc disruption below the critical mass is consistent with spin flips in an orbital timescale. In §4 we work out an analytic evaluation of the threshold mass for merger-driven spin flips. In §5 we briefly describe the possible effect of a compaction to a blue nugget on the longevity of extended discs and then rings above the threshold mass, and refer to Dekel et al. (2020) for more details. In §6 we discuss the key role of supernova feedback in the disruption versus survival of discs. In §7 we discuss other simulation results and compare our predictions to potentially relevant observations. In §8 we summarize our conclusions.

2 MASS THRESHOLD IN SIMULATIONS

2.1 The VELA simulations

In parallel to our analytic modeling, we utilize a suite of 34 VELA zoom-in hydro-cosmological simulations with high resolution. A few relevant characteristics are summarized here, while more details are provided in Table A1 and in several published papers (e.g. Ceverino et al. 2014; Zolotov et al. 2015), recently in Dekel et al. (2019). The simulations are based on an Adaptive Refinement Tree (ART) code (Kravtsov, Klypin & Khokhlov 1997; Ceverino & Klypin 2009). The suite consists of 34 galaxies that were evolved to $z \sim 1$, with a unique maximum spatial resolution ranging from 17.5 to 35 pc at all times. The dark-matter halo masses at $z = 2$ range from $10^{11}$ to $10^{12} M_\odot$. The galaxies were selected at $z = 1$ such that their dark-matter haloes did not suffer a major merger near $z \sim 1$. This eliminated less than ten percent of the haloes which would have become massive ellipticals at low redshifts. The halo mass selection also eliminated dwarf galaxies at $z < 4$.

Besides gravity and hydrodynamics, the code incorporates physical process relevant for galaxy formation such as gas cooling by atomic hydrogen and helium, metal and molecular hydrogen cooling, photoionization heating by the UV background with partial self-shielding, star formation, stellar mass loss, metal enrichment of the ISM and stellar feedback. Supernovae and stellar winds are implemented by local injection of thermal energy (as described in Ceverino & Klypin 2009; Ceverino, Dekel & Bournaud 2010; Ceverino et al. 2012). Radiation-pressure stellar feedback is implemented at a moderate level, following Dekel & Krumholz (2013), as described in Ceverino et al. (2014). In general, the feedback as implemented in this suite is on the weak side of the range of feedback strengths in common cosmological simulations, and no AGN feedback is incorporated.

In the analysis of the simulations, the disc plane...
Figure 1. Discs versus non-discs in the VELA simulations. The disc measures $R_d/H_d$ and $V_{\text{rot}}/\sigma$ (color) are averaged within bins in the mass-redshift plane. We see a mass threshold for disc viability at $M_v \sim 10^{11} M_\odot$ (highlighted by red horizontal line), with negligible redshift dependence, as predicted in §4. The two measures of disciness are consistent with each other. The black curve refers to the upper limit for effective supernova feedback at a virial velocity $V_v = 100 \text{ km s}^{-1}$ (§6). The cyan curves refer to the Press-Schechter $\nu\sigma$ peaks, for $\nu = 1, 2, 3, 4$ from left to right, respectively.

Figure 2. From non-discs to discs in the VELA simulations. Two measures of disciness, $R_d/H_d$ and $V_{\text{rot}}/\sigma$, as a function of halo mass $M_v$ for all the snapshots of all the 34 evolving galaxies. Each point refers to a snapshot, with the median and 1σ scatter (16% and 84% percentiles) marked in bins of $M_v$. The color refers to the measure of disciness in the other panel. The two measures of disciness are consistent with each other. We see a marked transition from non-discs to discs at a threshold mass of $M_v \sim 10^{11} M_\odot$, as predicted.

and dimensions are determined iteratively, as detailed in Mandelker et al. (2014). In short, the disc axis is defined by the angular momentum of cold gas ($T < 1.5 \times 10^4 \text{ K}$), which on average accounts for $\sim 97\%$ of the total gas mass in the disc. The disc radius $R_d$ is chosen to contain 85% of the cold gas mass in the galactic mid-plane out to $0.15 R_v$, where $R_v$ is the halo virial radius. The disc half-height $H_d$ is defined to encompass 85% of the cold gas mass in a thick cylinder where both the radius and half-height equal $R_d$. The values of the disc dimensions at $z = 2$ are listed in Table A1. The ratio $R_d/H_d$ is used below as one of the measures of gas disciness.

Another measure of disciness is the kinematic ratio of rotation velocity to velocity dispersion $V_{\text{rot}}/\sigma$. The rotation velocity and the velocity dispersion are computed by mass-weighted averaging over cells inside a cylinder whose minor axis is along the angular-momentum direction of the cold gas ($T < 4 \times 10^4 \text{ K}$).
2.2 Mass threshold for discs in the simulations

We use the simulations to explore how discs and non-discs populate the $M_v - z$ plane. Figure 1 shows the two measures of gas disciness, $R_d/H_d$ and $V_{rot}/\sigma$, in this plane. The ratios $R_d/H_d$ and $V_{rot}/\sigma$ (color) are averaged over all the simulation snapshots in bins of $M_v$ and $z$. We see a systematic gradient of “disciness” with mass, and a division between the zones of non-disc and disc dominance at a critical halo mass of $M_v \simeq (1-2) \times 10^{11} M_\odot$, with no significant redshift dependence. This is predicted by simple analytic model below, contrary to the naive expectation mentioned in the Introduction of a redshift threshold based on halo merger rates.

Focusing on the mass dependence, Fig. 2 shows the

within a sphere of radius $0.1 R_e$. The cylinder radius is $0.1 R_e$ and its half-height is $0.25 R_e$, where $R_e$ is the cold-gas half-mass radius (more details in Kretchmer et al., in prep.). The radial velocity dispersion is measured with respect to the mean radial velocity within the cylinder.
distribution of the two measures of gas disciness as a function of halo mass. The symbols correspond to all the $z < 9$ snapshots of all the 34 simulated galaxies as they evolve in time and grow in mass. We again see a clear transition from non-disc to disc dominance at about $M_v \approx 10^{11} M_\odot$, separated at $R_d/H_d \approx 2.5$ and $V_{rot}/\sigma \approx 2$.

To illustrate the distribution of disciness in the $M_v-z$ plane, Fig. 3 shows images of face-on projected gas density in example VELA galaxies that fall in the different zones of the $M_v-z$ plane (see the edge-on analog images in Fig. B1). Indeed, we see that the gas in galaxies of $M_v > 10^{11} M_\odot$ is typically in an extended rotating disc. In contrast, the gas in galaxies of $M_v < 10^{11} M_\odot$ is typically irregular and undergoing mergers. In between, the gas in galaxies of $M_v \sim 10^{11} M_\odot$ is typically in a compact configuration - a blue nugget generated by wet compaction.

3 SPIN FLIPS IN THE SIMULATIONS

3.1 Spin flips below the critical mass

In order to see whether the threshold mass for discs at $M_v \sim 2 \times 10^{11} M_\odot$ indeed correlates with the rate of spin flips in a galaxy orbital time, the left panel of Fig. 4 shows the average change in the spin direction over an orbital time, $\cos \theta$, in bins of $M_v$ and $z$. We indeed see a strong mass dependence with no apparent redshift dependence. For $M_v < 10^{11} M_\odot$, the tilt angle is typically large, $\cos \theta < 0.7$ ($\theta > 45^\circ$), which we can consider “a flip”. In contrast, for $M_v > 2 \times 10^{11} M_\odot$, the tilt is small, $\cos \theta > 0.85$ ($\theta < 30^\circ$), namely the disc orientation is stable over an orbital time.

To explore the association of major mergers with spin flips, the right panel of Fig. 4 shows all major mergers in the $M_v-z$ plane, with the color referring to the tilt $\cos \theta$ over an orbital time about the merger. We see that the mergers of galaxies in haloes of mass below the threshold tend to be associated with flips, while the mergers of more massive haloes tend to have a stable spin direction despite the merger. This is with the exception of three massive outliers, possibly affected by the fact that these mergers occur near the final time of the simulation making it hard to assess flips in a full orbital time.

It is interesting to note in the right panel of Fig. 4 the tendency of high-sigma peaks, $\nu \geq 2$ (near and to the right of the second black curve from the left), to flip in association with major mergers. This is consistent with the tendency of non-discs to populate the high-sigma peaks regime in Fig. 4. It is along the lines of the notion that mergers of high-sigma peaks, at the nodes of the cosmic web, are associated with a change in the pattern of streams that feed the galaxy in the node, including filament and sheet mergers, which is in turn associated with a change in the angular momentum that is brought by these streams to the galaxy (Danovich et al. 2015). Mergers of low-sigma peaks, occurring at high masses and low redshifts, tend to occur instead along cosmic-web filaments. These are less likely to be associated with a change in stream pattern that causes a spin flip.

Figure 5 makes the case for a mass threshold for spin flips even more explicit. It shows the evolution of gas spin of each of six example VELA galaxies. In each output time, the position of the symbol marks the spin direction (with arbitrary zero-points for the spin direction in spherical coordinates) and its size indicates
Figure 5. Evolution tracks on the sky of spin direction and magnitude (symbol size) for six VELA galaxies. The angles $\theta$ and $\phi$ are the standard spherical coordinates in a randomly oriented frame, such that the line intervals fairly represent the angles between the spins in successive snapshots. The color indicates halo mass $M_v$ as the galaxy is growing in time. Blue and red triangles mark the earliest and latest times, typically $z \sim 8$ and $1$, when the galaxies are of the lowest and highest mass, respectively. The white cross marks $\sim 10^{11} M_\odot$. The galaxy orbital time ranges from less than a time step to a couple of time steps from the initial to the final snapshots. We see frequent spin flips as long as $M_v < 10^{11} M_\odot$ (blue), relaxing to a stable spin direction and amplitude at $M_v > 10^{11} M_\odot$ (red), post compaction.

Figure 6. Cold mass fractions with respect to the total baryons within $0.15 R_v$, averaged within bins in the $M_v - z$ plane. The cold mass includes cold gas and young stars, with the latter contributing about a third of the mass. There is a gradient with virial mass, crossing $\sim 0.5$ near the critical mass, indicating that the mergers in low mass galaxies tend to be wet, and therefore more disruptive to the gas discs.

3.2 Wet mergers

The disruption by each merger-driven flip can be more effective when the mergers are more gas rich. Figure 6 shows the cold fraction with respect to the total baryons within the larger disc radius, $0.15 R_v$, averaged in bins within the $M_v - z$ plane. The young stars contribute about a third of the cold mass. We see a strong gradient with virial mass, rising to above 50% below the critical mass for merger-driven disc flips and for major compaction events (see below). We learn that the magnitude of the angular momentum. The starting and ending snapshots, typically at $z \sim 8$ and $z \sim 1$, are marked by a blue and a red triangle. The halo mass is indicated by color, with the snapshot closest to $10^{11} M_\odot$ marked by a white cross. In all six cases we clearly see big flips between output times, which are comparable to the orbital times, as long as $M_v < 10^{11} M_\odot$ (blue to white symbols), and a stable spin when $M_v > 2 \times 10^{11} M_\odot$ (pink to red symbols).

In §4 below we evaluate analytically the threshold mass for merger-driven spin flips, in agreement with what we extracted from the simulations. We learn that merger-driven flips can be the reason for the disruption of discs below the critical mass.
critical mass the major mergers are not only more frequent per orbital time (e.g., Fig. 1), but they are more gas rich, and therefore more disruptive for the gas discs. This emphasizes further the transition from non-discs to discs near the critical mass.

We discuss below how other mechanisms can contribute to the disruption of discs below the critical mass. These include the effects of supernova feedback, and the inward mass transport within violently unstable discs (VDI), which may be effective below a similar critical mass that refers to compaction events, to be discussed next.

4 DERIVATION OF THE THRESHOLD MASS

4.1 Flips in an orbital time

We expect a disc to be disrupted once the direction of its angular-momentum (AM) is changed significantly during a disc orbital time. We assume that such a flip occurs in a major merger of high-sigma nodes of the cosmic web, which is associated with a pattern change of the cosmic-web streams that feed the galaxy with mass and angular momentum (Danovich et al. 2015 see below). The characteristic time $t_{\text{mer}}$ between mergers of baryon mass ratio $r \leq 1$ can be related to the baryon accretion timescale by

$$ t_{\text{mer}} \simeq \beta(r) \frac{M_b}{M_b}, $$

where $M_b$ is the galaxy baryonic mass, $\dot{M}_b$ is the total baryonic accretion rate, and $\beta(r)$ is a factor of order unity for major mergers of $r \sim 1/3$ (more in 4.3).

We write

$$ t_{\text{mer}} \simeq \frac{M_b}{M_b} = \frac{M_b}{M_b} \frac{M_v}{M_v}, $$

where $M_b$ is the stellar mass and $M_v$ is the virial mass. We made here the assumption that the baryonic accretion rate relates to the total accretion rate via

$$ \dot{M}_b = \dot{f}_b \dot{M}_v, $$

where $f_b = 0.16 \ f_{b, 0.16}$ is the universal baryonic fraction. This is supported by simulations (Dekel et al. 2013), and by a match to the observed SFR on the main sequence as a function of redshift (Dekel et al. 2009, Dekel & Mandelker 2014, Rodriguez-Puebla et al. 2016).

In the third term of eq. (2), $M_v/M_v$ is the average timescale for total accretion into the halo. It can be derived analytically in simple terms (Dekel et al. 2013) using the self-invariant Press-Schechter time variable (the inverse of the linear growing mode of density fluctuations), in good agreement with simulations (Dekel et al. 2013). In the EdS regime ($z > 1$), the approximation is

$$ t_{\text{ac}} = \frac{M_v}{M_v} = \frac{1}{\tau \ G\ y (1+z)^{-3/2} M_{v,11}}, $$

where $M_v = 10^{11} M_\odot \ M_{v,11}$ and the weak mass dependence with the small power index $-0.1$ is an estimate for the LCDM power spectrum in the vicinity of $M_v \sim 10^{11} M_\odot$. The value of $\tau$ as derived from the simulations is $\sim 30$, so we denote below $\tau = 30 \ t_{\text{30}}$.

The first term in eq. (2) is the stellar-to-halo mass ratio, which can be estimated for LCDM haloes via abundance matching to observed galaxies. We term $f_{sv} = M_s/M_v$ and denote $f_{sv} = 10^{2.5} f_{sv,2.5}$. The second term in eq. (2) is the inverse of the stellar fraction in the galaxy with respect to the total baryons, which we term $f_{sb} = M_b/M_v$ and denote $f_{sb} = 0.5 f_{sb,0.5}$. We obtain from eq. (1), eq. (2) and eq. (4)

$$ t_{\text{mer}} \simeq 1.1 \ Gyr \ f_{sv,2.5} \ f_{sb,0.5} \ t_{\text{30}} \ f_{sb,0.16}^{-1} \ \beta \ \times (1+z)^{-5/2} M_{v,11}^{-0.1}. $$

Recall that $t_{\text{30}}, f_{sb,0.16}^{-1}$ and $\beta$ are all parameters of order unity, while $f_{sv}$ and $f_{sb}$ depend on $z$ and/or $M_v$.

In comparison, the disc orbital time is

$$ t_{\text{orb}} = 2 \pi \ R_d \ V_d = 2 \pi \ \frac{\lambda R_v}{V_v}, $$

where $V_v$ and $V_d$ are the virial velocity and disc rotation velocity respectively, and where $\lambda = R_d/R_v$ is the contraction factor from the halo to the gas disc. We write $\lambda = 0.08 \lambda_{0.08}$, where $\lambda \sim 0.08$ is derived for the gas from the VELA simulations. The virial crossing time in the EdS regime is (Dekel et al. 2013)

$$ t_v = \frac{R_v}{V_v} = 0.14 \ t_{\text{hubble}} \simeq 2.45 \ Gyr \ (1+z)^{-3/2}. $$

This gives

$$ t_{\text{orb}} \simeq 1.23 \ Gyr \ \lambda_{0.08} \ (1+z)^{-3/2}. $$

The ratio of timescales, from eq. (5) and eq. (8), is

$$ \frac{t_{\text{mer}}}{t_{\text{orb}}} \simeq 0.89 \ f_{sv,2.5} \ f_{sb,0.5} \ t_{\text{30}} \ \lambda_{0.08}^{-1} \ f_{sb,0.16}^{-1} \ \beta \ \times (1+z)^{-1} M_{v,11}^{-0.1}. $$

This may seem to indicate a redshift dependence, emerging from $t_{\text{ac}}/t_v \propto (1+z)^{-1}$, which would have implied that disc disruption is more efficient at higher redshifts and is rather insensitive to mass. This is misleading because $f_{sv}$ has a strong mass dependence that becomes the dominant factor in $t_{\text{mer}}/t_{\text{orb}}$, while $f_{sb}$ has a redshift dependence that roughly cancels the overall redshift dependence of $t_{\text{mer}}/t_{\text{orb}}$, as we see next.

4.2 Baryon-to-halo ratio and the threshold mass

From abundance matching of LCDM haloes to observed galaxies (Behroozi et al. 2018, Fig. 7), we read for $M_v = 10^{10–12} M_\odot$

$$ f_{sv} = \frac{M_s}{M_v} \simeq 3 \times 10^{-3} f_{sv,11} \ M_{v,11}^{-2.5}, $$

This is sometimes identified with the halo spin parameter but it has been shown not to correlate with the spin parameter on a halo-by-halo basis (Jiang et al. 2019).
where $f_{sv,11,-2.5}$ is the value of $f_{sv,-2.5}$ at $M_v = 10^{11} M_\odot$, and $f_{sv,11,-2.5} \approx 1$ with only little redshift dependence in the redshift range $z = 1 - 5$. The ratio at a given mass is indicated to rise with $z$ at higher redshifts, but not so in other attempts to estimate this ratio (e.g. Rodríguez-Puebla et al. 2017), so we ignore here any variations with $z$. An analytic argument based on energetics of supernova feedback predicts a qualitatively similar increase with mass as in eq. (10) but with a slightly flatter mass dependence, $M_s/M_\bullet \propto M_z^{2/3}$ Dekel & Silk 1986 see (6.1). We adopt here the latest empirical abundance-matching estimate from eq. (10) without a strong effect on the results.

We determine $f_{sb}$, the stellar mass fraction with respect to the baryonic mass, from the data of Tacconi et al. (2018). This is a compilation of more than 1400 molecular gas measurements in typical star-forming galaxies at $z = 0 - 4$, including the PHIBSS (Tacconi et al. 2013) and PHIBSS2 (Freundlich et al. 2019) samples. The dependence of $f_{sb}$ on mass and redshift is evaluated through a simultaneous power-laws fit in terms of the variables $(1 + z)$, $M_s$ and $\Delta M S$, the deviation from the star-forming main sequence ridge. This is assuming that the variables can be separated and that the dependence on each can be approximated by a power law. The best fit obtained is

$$f_{sb} = 1.27 (1 + z)^{-1} M_{s,11}^{0.14}, \quad (11)$$

where $M_s = 10^{11} M_\odot M_{s,11}$.

Combined with eq. (10), the required $f_{sb}^{-1}$ can be expressed in terms of $M_v$ as

$$f_{sb}^{-1} = 0.88 f_{sv,11,-2.5}^{0.14} (1 + z) M_{v,11}^{-0.28}. \quad (12)$$

Similar exponents are obtained when carrying instead successive fits, as in Genzel et al. (2015) and Tacconi et al. (2018), although such fits may indicate that the variable separation is not entirely justified, and they hint to a weak redshift dependence of the exponents of $M_s$ and $\Delta M S$.

In fact, the redshift dependence of $f_{sb}$ could be predicted using a bathtub steady-state model, where mass conservation is assumed as a balance between input and output of gas mass and stellar mass. Combining equations 19, 24 and 26 in Dekel & Mandelker (2014), we indeed obtain in the range $z \gg 1$

$$f_{sb}^{-1} \propto t_{orb} / t_{sc} \propto (1 + z), \quad (13)$$

assuming a constant star-formation-rate efficiency in a dynamical time, $\epsilon_B$. A zero-point of $f_{sb}^{-1} \sim 0.8$ at $z = 0$ is obtained for $\epsilon_B \sim 0.01$, and for $\tau \sim 30$ Gyr in the accretion time, ignoring stellar mass in the accretion at very high redshift. One should note, however, that eq. (13) is valid for a galaxy as it evolves, so the redshift dependence at a fixed mass is slightly different.

4 Note that the value of $f_{sb}$ here is properly smaller than unity only for $z$ safely above 0.27. Furthermore, this relation identifies the total gas mass with the molecular gas mass, which is a sensible approximation only at $z$ well above 0.4. Thus, eq. (11) is a valid approximation only at redshifts of order unity and above.

We learn in passing by combining eq. (10) and eq. (12) that at a given $z$

$$M_s \propto M_v^{0.12}. \quad (14)$$

Substituting eq. (10) and eq. (12) in eq. (9) we obtain

$$t_{mer} \simeq 0.79 \tau_{sc} \lambda_0 \lambda_{s,11,11}^{-1.15} f_{sv,11}^{-0.62} M_v^{0.62}. \quad (15)$$

Here all the parameters are assumed to be constants, independent of $M_v$ and $z$. We learn that the mass dependence of $t_{mer}/t_{orb}$ is pronounced, but the redshift dependence has vanished.

If we adopt $t_{mer}/t_{orb} = 1$ as the threshold for disc survival, we finally obtain from eq. (15) a redshift-independent virial mass threshold for disc survival,

$$M_v,11 \simeq 1.36 \tau_{sc}^{-1.62} \lambda_0 \lambda_{s,11,11}^{-1.15} f_{sv,11}^{-0.62} M_{v,11}^{-1.62}. \quad (16)$$

We expect all the parameters to be of order unity and independent of mass and redshift. We thus predict a threshold halo mass for disc survival at $M_\bullet \sim (1 - 2) \times 10^{11} M_\odot$, corresponding to $M_s \sim (3 - 6) \times 10^8 M_\odot f_{sv,11,11,-2.5}$. This is indeed seen very convincingly in the simulations, Fig. 1 and Fig. 2.

In summary, the relation $t_{mer} \propto (1 + z)^{−5/2}$ for haloes turns into $t_{mer} \propto M_v^{0.62}(1 + z)^{−3/2}$ when one incorporates the mass dependence of $M_s/M_\bullet$ and the redshift and mass dependences of $M_s/M_\bullet$. The resultant redshift dependence balances the redshift dependence of $t_{orb}$, yielding a pure mass dependence for $t_{mer}/t_{orb}$, which is translated to a mass threshold for disc survival with no redshift dependence.

We note that a weak redshift dependence in the threshold mass would appear if $f_{sv} = M_s/M_\bullet$ actually evolves with redshift, or if the log slope of the redshift dependence of $f_{sb}$ in eq. (11) is somewhat different from unity, as may be obtained based on a different method of linear-log multi-variable fit of the scaling relation. As for the former possibility, in the abundance-matching analysis of Rodríguez-Puebla et al. (2017), the zero point of $f_{sv}$ at a given halo mass is actually decreasing with increasing redshift in the redshift range $z = 0 - 3$ and it remains constant at higher redshifts. This behavior is different from the findings of Behroozi et al. (2018), thus reflecting a significant uncertainty in the stellar-to-halo mass ratio from abundance matching. A decrease of $f_{sv}$ with redshift would introduce a negative redshift dependence in eq. (15), namely a positive redshift dependence in eq. (10), which would result in a larger mass threshold at higher redshifts. This would imply a lower fraction of discs at higher redshifts at a fixed mass, as indicated by certain observations at $z < 2$ (see § 2 Fig. 11). Such a redshift dependence is not seen in the VELA simulations at higher redshifts, Fig. 1 because $f_{sv}$ does not vary significantly with redshift in these simulations, consistent with Behroozi et al. (2018) at $z < 5$.

4.3 Mergers and spin flips

Back to the merger timescale assumed in eq. (1), it can be obtained by expressing the mass added by a merger
in two ways,
\[ t_{\text{mer}} M_b f_r = \dot{r} M_b , \]  
(17)
where \( f_r \) is the mass fraction in mergers with mass ratio \( > r \) with respect to the total accreted baryonic mass, and \( \dot{r} \gtrsim r \) is the average ratio for such mergers that are of a ratio \( > r \). This gives eq. [1] with \( \beta(r) = \dot{r} / f_r \).

We can crudely estimate \( \beta \sim 1 \) from major mergers of haloes, which can be evaluated using the EPS formalism. On one hand, \( \dot{r} \sim 2r \) in the major-merger regime (based on Neistein & Dekel [2008], Fig. 7). On the other hand, the mass ratio for the baryonic masses, \( r_h \), is smaller than the ratio for the total halo masses, \( r_h \), because the baryon-to-halo mass ratio is an increasing function of mass. We have \( r_h \propto r_h^{1.72} \) in the relevant mass range (from eq. (10) and eq. (12)), so \( r_h \sim 2r_h \) in the major-merger regime. The corresponding values of \( f_r \) differ by a factor of 2, roughly balancing the difference between \( r \) and \( \dot{r} \). For haloes, using EPS, we read from Fig. 7 of Neistein & Dekel [2008] that \( f_r \approx 1/4 \), \( 1/3 \) and \( 1/2 \) for \( r = 1/2 \), \( 1/3 \) and \( 1/10 \), respectively. This gives \( \beta \approx 2 \), 1 and 0.2 for these choices of \( r \). We adopt for major mergers \( r = 1/3 \) and thus \( \beta \approx 1 \) as our fiducial value.

For a high-sigma halo, we expect a major merger to be associated with a significant change in the AM direction of the cosmological inflow. This is because high-sigma haloes form in the nodes of the cosmic web, and a coalescence of such nodes is likely associated with a change of the pattern of streams that feed the galaxy and in the inflowing mass. The incoming AM is therefore expected to be sufficient for causing a flip in the galaxy AM direction for any sensible value of \( r \).

5 POST-COMPATION DISCS AND RINGS

5.1 Mass transport & compaction to blue nuggets

While we learned so far that discs above the threshold mass are not likely to suffer frequent merger-driven spin flips, violently unstable discs are themselves subject to inward mass transport which may in principle disrupt the discs from the inside even in the absence of major mergers. In a disc undergoing violent disc instability (VDI), the non-cylindrically symmetric density perturbations exert torques that induce transport of AM outward. AM conservation implies an associated mass transport inward, e.g., in terms of clump migration and gas inflow through the disc (e.g., [Noguchi 1999], [Dekel, Sari & Ceverino 2009]) estimated the corresponding evacuation time of a disc to be
\[ t_{\text{inflow,disc}} \sim \delta_d^{-2} t_{\text{orb}}, \quad \delta_d = \frac{M_d}{M_{\text{tot}}}, \]  
(19)
where the key variable \( \delta_d \) is the mass ratio of cold disc to total mass within radius \( r \). This has been estimated either via the encounters between giant clumps or based on the shear-driven mass-inflow rate of [Shakura & Sunyaev 1973] with the maximum AM flux density by Gammie [2001]. With \( \delta_d \lesssim 1 \), we thus expect an inward mass transport within a few orbital times. This is contrary to our finding from the simulations that extended rings survive for long term once above the threshold mass (5.2).

An immediate suspect that could have a role in stabilizing discs above the threshold mass is the process of wet compaction to a blue nugget, which tends to occur near a similar characteristic mass (Zolotov et al. 2015). In the simulations, the major compaction event is typically associated with the formation of an extended long-lived disc or ring. We argue that the post-compaction appearance of a massive bulge is responsible for the transition from a disc to a ring and for slowing down its inward mass transport. This is addressed in detail in [Dekel et al. 2020] and summarized here.

Cosmological simulations show that galaxies evolve through a dramatic wet-compaction event, pronounced when the galaxy mass is near or above \( M_* \sim 10^{11.5} M_{\odot} \) (\( M_* \sim 10^{10.5} M_{\odot} \)) and in the redshift range \( z = 1-5 \) when the gas fraction is high (Zolotov et al. 2015). This is gaseous contraction into a compact central star-forming body within the inner \( \sim 1 \) kpc, “a blue nugget” (BN). The gas consumption by star formation and ejection by supernova feedback cause central gas depletion and inside-out quenching of star-formation rate (SFR) (Tacchella et al. 2016a). Figure [3] below illustrates the evolution through such a sequence of events in an example VELA galaxy. The cartoon in Fig. 3 of [Dekel et al. 2020] shows the typical evolution of gas mass, stellar mass and SFR within the inner kpc. Figure 4 there shows evolution tracks of galaxies in the plane
Figure 7. Compaction to a blue nugget and post-compaction disc and ring. Shown are the projected densities of gas (top) and stars (bottom) in four phases (expansion factor $a = (1+z)^{-1}$ is marked) during the evolution of one of the VELA simulated galaxies (V07). The projections are face on with respect to the AM. Top, from left to right. First: during the compaction process. Second: at the blue-nugget phase. Third: post-compaction VDI disc. Fourth: post-compaction, clumpy, long-lived ring, fed by incoming streams. Bottom: the stellar compact red nugget forms during and soon after the compaction and the resulting bulge remains compact and grows massive thereafter.

of specific SFR (sSFR) versus compactness as measured by the stellar surface density within 1 kpc, $\Sigma_1$. A compaction at a roughly constant sSFR turns into quenching at a constant $\Sigma_1$ with the “knee” marking the blue-nugget phase. This universal L-shape evolution track has been confirmed observationally (Barro et al. 2017, Fig. 7). A dissipative “wet compaction” (Dekel & Burkert 2014) is actually required by the compactness of the observed massive passive galaxies at $z \sim 2−3$ (“red nuggets”, e.g. van Dokkum et al. 2015), which implies the presence of “blue nuggets” as their immediate progenitors. Indeed, star-forming blue nuggets are observed with consistent masses, structure, kinematics and abundance (e.g. Barro et al. 2013; van Dokkum et al. 2015; Barro et al. 2017). A deep-learning study (Huertas-Company et al. 2018), trained on “observed” mock images of blue nuggets as identified in the simulations, recognized with high confidence blue nuggets in the CANDELS-HST imaging survey of $z = 1−3$ galaxies.

The compaction requires dissipative AM loss, caused in the simulations either by wet mergers (∼40%), by colliding counter-rotating streams, or other processes. The compaction marks transitions in the galaxy structural, compositional, and kinematic properties, which translate to variations with mass near the critical mass for blue nuggets. For example, the dark-matter dominated center becomes baryon dominated, inducing a transition in global shape from a oblate to prolate (Tomassetti et al. 2016). The growth rate of the central black hole changes from slow to fast (Dekel, Lapiner & Dubois 2019), inducing a transition from supernova feedback to AGN feedback as the main source for quenching.

Repeated episodes of compactions and quenching attempts explain the confinement of SFGs to a narrow Main Sequence (Tacchella et al. 2016). However, the major compaction events that produce massive bulges are predicted to occur near a characteristic halo mass $M_h \sim 10^{11.5} M_\odot$ (Zolotov et al. 2015; Tomassetti et al. 2016). This has been confirmed by the deep-learning study of VELA simulations versus observed CANDELS galaxies (Huertas-Company et al. 2018), which detected a preferred stellar mass for the observed blue nuggets near the golden mass of $M_\star \sim 10^{10.5−10} M_\odot$. We argue (Dekel, Lapiner & Dubois 2019 and (6) below) that the compaction events are especially pronounced near the critical mass because supernova feedback becomes inefficient near and above this mass.

5.2 Rings stabilized by a massive bulge

Figure 7 shows the evolution of one VELA galaxy through the compaction, blue-nugget and post-compaction phases. The face-on images of projected gas density reflect the SFR via the Kennicutt-Schmidt relation. The blue-nugget at the end of the compaction is the central blob of high gas density (second from left). A highly turbulent rotating disc develops thereafter and is growing in extent (third from left), showing a spiral-arm pattern and irregular perturbations including giant clumps. In this violent-disc-instability (VDI) phase the giant clumps and the gas between them migrate inwards (e.g. Dekel, Sari & Ceverino 2009). Then,
the central gas is depleted into star formation and outflows (Zolotov et al. 2015) and an extended clumpy ring forms (fourth panel), continuously fed by incoming cold streams, showing tightly wound spiral arms and giant clumps. The ring is sometimes maintained at its extended form for several Gyr, with only little inward migration. The stellar-density maps show how a compact stellar system forms following the gas compaction process and how it remains massive and compact as it quenches to a passive red nugget. These post-compaction discs and rings can be identified with the observed massive and extended star-forming rotating and highly turbulent discs showing giant clumps (Genzel et al. 2008, 2014; Guo et al. 2015, 2018; Forster Schreiber et al. 2018).

We learn in Dekel et al. (2020) that during the compaction the VELA galaxies evolve from pressure to rotation support, with the median $V_{\text{rot}}/\sigma$ growing from near unity to about 4±1. It is the rotation velocity that is dramatically growing during the compaction, from $V_{\text{rot}}/V_c = 0.4 ± 0.1$ to $V_{\text{rot}}/V_c = 1.4 ± 0.1$, while the velocity dispersion remains at $\sigma/V_c = 0.4 ± 0.1$. The kinematic-ratio transition is thus due to an abrupt increase in the gas AM rather than to a change in turbulence. This implies that the inflowing high-AM gas is prevented from forming a long-lived disc in the pre-compaction phase because of an efficient loss of AM, e.g., due to the merger-driven flips discussed above. In turn, the gas seems to retain its incoming high AM in the post-compaction phase. This motivates our effort to understand the post-compaction survival of extended discs and rings by means of AM exchange.

In Dekel et al. (2020) we show that the longevity of the rings could be understood in the context of a VDI disc if $\delta_d \ll 1$, which could arise from the post-compaction appearance of a massive central bulge. By computing the torques exerted by a tightly wound spiral structure in a thin disc on its outer parts, we find that the pitch angle can be approximated by

$$\tan^2 \alpha \sim 4\delta_d^2$$

and the timescale for mass inflow is roughly

$$t_{\text{inflow, ring}} \sim \delta_d^{-3} t_{\text{orb}}.$$  \hfill (21)

With $\delta_d$ small compared to unity, this implies a very small pitch angle, namely a ring, and a very long inflow time, much longer than the orbital time. This is longer than the analogous inflow time of eq. (19), which is proportional to $\delta_d^{-2}$. The post-compaction massive bulge, that appears above the critical mass for blue nuggets, reduces $\delta_d$, and thus acts to stabilize the ring against evacuation inwards.

The survival of the ring is determined by the interplay between the inflow rate of eq. (21), the star-formation rate (SFR) within the ring that suppresses the dissipative inflow (Dekel & Burkert 2014), and the external accretion rate of high-AM gas in spiraling-in streams that replenish a new outer ring. If we assume an EdS cosmology (approximately valid at $z > 1$), the ring orbital time is $t_{\text{orb}} \sim 0.75 \text{Gyr} (1 + z)^{-3/2}$. The accretion timescale is then (Dekel et al. 2013) $t_{\text{acc}} \sim 40 t_{\text{orb}} (1 + z)^{-1}$. In comparison, the inflow rate computed above is $t_{\text{inflow}} > 40 t_{\text{orb}}$ for $\delta_d < 0.3$. Thus $t_{\text{acc}} < t_{\text{inflow}}$ in the ring at all redshifts. This implies that the gas accumulates in the ring.

Gas dissipation may speed up the contraction of the ring. The gas turns into stars on a timescale $t_{\text{sf}} \sim \epsilon_{\text{ff}}^{-1} t_{\text{ff}} \sim 5t_{\text{orb}}$, where $\epsilon_{\text{ff}} \sim 0.01$ is the efficiency of SFR in a free-fall time and $t_{\text{ff}}$ is the free-fall time in the star-forming regions (Krumholz, Dekel & McKee 2012). We adopted $t_{\text{sf}} \sim 0.3t_{\text{dyn}}$, assuming that stars form in clumps that are denser than the mean density of baryons in the ring by a factor of $\sim 10$ (Ceverino et al. 2012). A similar timescale for SFR is obtained in the VELA simulations. We learn that $t_{\text{sf}} \ll t_{\text{inflow}}$, which implies that the gas in the ring turns into stars before it has a chance to flow in such that a dissipative contraction of the ring is not likely.

6 ON THE ROLE OF SUPERNOVA FEEDBACK AND HOT CGM

Figure 8 emphasizes the relevant physical characteristic masses in the mass-redshift plane. It shows (a) the threshold mass for merger-driven disc flips in an orbital time from $\epsilon_{\text{ff}}^3 t_{\text{ff}} \sim 5t_{\text{orb}}$, (b) the favored mass for compaction-driven blue nuggets, $M_\star \sim 2 \times 10^{11} M_\odot$ (Zolotov et al. 2015; Tomassetti et al. 2016), (c) the upper limit for effective supernova feedback, corresponding to $V_c \sim 120 \text{km s}^{-1}$ (Dekel & Silk 1986), and (d) the threshold for virial shock heating of the CGM and the upper limit for penetrating cold streams that feed the galaxies (Dekel & Birnboim 2006; Dekel et al. 2009). Also shown are half of the VELA snapshots, selected to be the upper and lower quadrants in terms of their disciness as measured by $V_{\text{rot}}/\sigma$ and $R_d/H_d$. One can see that each of the first three lines can serve as a border line distinguishing between non-disc and disc dominance. We argue below that SN feedback has a major role in disrupting discs below the critical mass in several different ways, direct and indirect, by itself and through the effects of disc flipping and compactions to blue nuggets. We then comment on the possible complementary role of a hot CGM in more massive galaxies.

6.1 Supernova feedback

Simple energetic arguments, using the standard theory for supernova bubbles, yield an upper limit for the dark-matter halo mass (actually its virial velocity) within which supernova feedback can be effective in heating or ejecting the gas and thus suppressing the SFR (Dekel & Silk 1986). The energy deposited in the ISM by supernovae that arise from a burst of stellar mass $M_\star$ is estimated to be $E_{\text{SN}} \sim M_\star V_{\text{SN}}^2$, with $V_{\text{SN}} \sim 120 \text{km s}^{-1}$. 

© 2002 RAS, MNRAS 000.
The cyan curves refer to the Press-Schechter mass corresponding to virial shock heating and the diagonal dashed black line is the characteristic threshold for the gas is either clearly discy (blue, from the VELA simulations, selected from snapshots where the z−masses. Haloes of $1 M_{\odot}$ for high-sigma peaks. Extended discs of star-forming gas are expected to survive above $M_{\text{vir}} \sim 2 \times 10^{13} M_{\odot}$ and disrupt at lower masses. Haloes of $1.4 \times 10^{13} M_{\odot}$ are expected at a comoving number density (in $\text{Mpc}^{-3}$) on the order of $10^{-2}$ at $z \sim 1$−3, $10^{-3}$ at $z \sim 4$−6, and $10^{-4}$ at $z \sim 8$ (Fig. 4). The symbols refer to galaxies from the VELA simulations, selected from snapshots where the gas is either clearly discy (blue, $R_{\text{d}}/H_{\text{s}} > 3.1$ and $v_{\text{rot}}/\sigma > 3$), or clearly non-discy (red, $R_{\text{d}}/H_{\text{s}} < 2.8$ and $v_{\text{rot}}/\sigma < 1.3$), respectively, each consisting of a quarter of all the snapshots. Each of the three theoretical curves (blue nuggets, flips and SN) could serve as a border line between non-discs and discs.

This is obtained by comparing the cooling and dynamical timescales for the adiabatic phase of SN bubbles and for the star-formation rate, respectively. For the supernova energy to heat or eject most of the gas of mass $M_{\text{g}}$ that has accreted into the galaxy it should be comparable to the binding energy of this gas in the dark-matter halo potential well, $E_{\text{CGM}} \sim M_{\text{g}} V_{\text{c}}^2$, where $V_{\text{c}}^2 = GM_{\text{halo}}/R_{\text{halo}}$ characterizes the halo potential well. At the peak of star-formation efficiency, if a large fraction of the cumulative accreted gas turned into stars with little ejection, one can approximate $M_{\text{g}} \sim M_{\text{d}}$. Then, comparing $E_{\text{SN}}$ and $E_{\text{CGM}}$, one obtains an upper limit for effective SN feedback at $V_{\text{c}} \sim V_{\text{s,N}} \sim 120 \text{ km s}^{-1}$. The corresponding halo mass in the EdS regime ($z > 1$) is (e.g. Dekel & Birnboim 2006, Appendix A2)

$$M_{\text{vir}} \simeq 6 \times 10^{11} M_{\odot} V_{\text{c,100}}^3 (1 + z)^{-3/2}, \quad (22)$$

where $V_{\text{c}} = 100 \text{ km s}^{-1} V_{\text{c,100}}$. This SN critical mass is plotted in Fig. 8. Despite the redshift dependence of the mass, it is in the ballpark of the mass threshold for transition from non-discs to discs (5) at all redshifts, which is also similar to the critical mass for blue nuggets (5).

We saw in (4) that the characteristic mass for disc flips arises from the relation between stellar and halo mass in the SN zone, below the critical mass, eq. (10). The above simple energetics argument for SN feedback can also be used to explain the slope of this relation. When $V_{\text{c}} < V_{\text{SN}}$, the shallow potential well allows significant gas ejection such that $M_{\text{g}} < M_{\text{d}}$. If the gas that has been accreted into the halo is roughly proportional to the halo mass, $M_{\text{g}} \propto M_{\text{d}}$, by comparing $E_{\text{SN}}$ and $E_{\text{CGM}}$ one obtains (Dekel & Woo 2003)

$$\frac{M_{\text{g}}}{M_{\text{vir}}} \propto V_{\text{c}}^2 \propto V_{\text{SN}}^{2/3} \propto M_{\text{d}}^{2/3}.$$ \quad (23)

This mass dependence is in the ballpark of the relation deduced from observations via abundance matching with LCDM haloes (Moster et al. 2010; Behroozi, Wechsler & Conroy 2013), while the analysis of Behroozi et al. (2018) as summarized in eq. (10) indicates a slightly steeper slope, closer to unity. Thus, SN feedback seems to have an indirect but important role in determining the critical mass for disc flips (5 and 4).

Each of the major mergers that cause spin flips below the critical mass also generates a starburst, which leads to a boosted SN feedback. This SN feedback may drive excessive turbulence in the disc, which puffs it up and possibly disrupt it. Furthermore, the resultant supernova-driven winds suppress the inflow into the galaxy, as has been demonstrated by Tollet et al. (2019, Fig. 5). Since the inflowing gas is typically of a higher AM than the existing galaxy, the SN-driven winds limit the growth of an extended disc.

Another possible mechanism for losing AM involves an “inverse fountain”, where supernovae at the outskirts of the disc preferentially remove high-AM gas (that may return after losing AM in the inner CGM), thus trimming the galaxy from its extended gas disc (Degraff et al., in prep.). The fountain becomes of an opposite, more standard effect on the disc viability after the compaction to a blue nugget, where the SFR becomes concentrated in the central regions such that SN feedback preferentially removes low-AM gas, which increases the net specific AM in the disc.

Finally, it has been argued that SN feedback is responsible for confining the major compaction events to the golden halo mass of a few $10^{11} M_{\odot}$ (Dekel, Lapiner & Dubois 2019). Compaction events do occur also at smaller masses, where they help confining the star-forming galaxies into a narrow main sequence (Tacchella et al. 2016b). However, SN winds in the heart of the SN zone prevent these compactions to proceed deep into the galaxy central kiloparsec. They do not cause significant long-term post-compaction quenching as the gas is rapidly replenished by and the centres remain dark-matter dominated. Only near the upper-limit critical mass for SN feedback the winds become ineffective.

© 2002 RAS, MNRAS 000, [121]
and the compaction process can proceed deep into the galaxy centre, generating a blue nugget that triggers a lasting central quenching. Given that the blue nuggets seem to be a major driver of an extended disc or ring (\cite[5.2]{5.2}), we learn that the SN feedback has an indirect but important role in determining the mass threshold for the formation of long-lived rings.

6.2 Hot CGM

Another characteristic mass of a complementary role here is the upper limit for the halo mass within which cold inflow can supply gas for an extended star-forming disc or ring (originally \citet{1977,1977,1977}). It is obtained by comparing the gas radiative cooling time to the dynamical time for gas inflow. The key is whether the shock that forms at the halo virial radius, behind which the gas heats to the virial temperature, can be supported against gravitational collapse. For this, its cooling time has to be longer than the dynamical time for gas compression behind the shock. A shock-stability analysis reveals a threshold mass for a hot CGM (\cite{2003,2006}, marked in Fig. 8) on the order of

\[ M_c \sim 5 \times 10^{11} M_\odot, \]

roughly independent of redshift. The associated uncertainty and scatter between galaxies are rather large, spanning about a decade, e.g., reflecting uncertainties in metallicity and the location within the halo where the shock stability is evaluated.

At \( z > 2 \), narrow cold streams are expected to penetrate through the otherwise hot CGM and supply gas for discs or rings even in haloes above the critical mass for shock heating \cite[Figure 7]{2006}. These predictions have been confirmed in cosmological simulations \cite{2005,2008,2008,2013,2016}, and observationally based on abundance matching of galaxies to dark-matter haloes may provide clues for what one should expect for the disc spins, but one should be aware of the significant AM gains and losses of the in-streaming baryons as they penetrate from outside the halo to the inner galaxy \cite{2015}, which tends to smear out most of the one-to-one correlation between the disc and halo spins \cite{2019}.

\textbf{7 OTHER SIMULATIONS & OBSERVATIONS}

Here we compare our results to related simulation work and tentatively address existing observational results, looking forward to future observations, e.g., by JWST, at the relevant low masses \( M_\ast < 10^6 M_\odot \) and at high redshifts.

7.1 Other Simulations

As a relevant background, spin flips of dark-matter haloes were studied in cosmological simulations by \cite{2012} and \cite{2016}, who found that they could indeed be associated with either major mergers or with non-major mergers and other changes in the AM feeding pattern and rate. First steps in addressing the effects of mergers of cosmic-web filaments, using excursion-set theory for saddle-point mergers, were made by \cite{2018}. Also related are the correlations of galactic spins with the filaments and walls of the cosmic web as analyzed in cosmological simulations by \cite{2019}. The spins of the dark-matter haloes may provide clues for what one should expect for the disc spins, but one should be aware of the significant AM gains and losses of the in-flowing baryons as they penetrate from outside the halo to the inner galaxy \cite{2015}, which tends to smear out most of the one-to-one correlation between the disc and halo spins \cite{2019}.

\cite{2018} studied the viability of discs as a function of stellar mass at \( z = 0 \) in 24 FIRE zoom-in cosmological simulations, spanning stellar masses of \( M_\ast = 10^{6-11} M_\odot \). They find at \( z = 0 \) a trend with mass similar to our finding at \( z > 0.8 \), with a transition from dominance by non-discs to discs near a similar critical mass, \( M_\ast \sim 10^6 M_\odot \). They also realize that stellar feedback has a major role in reducing the rotational support in low-mass galaxies, owing to the suppression of the accretion of high-AM gas and to driving large non-circular gas motions. We argued that, on top of these direct effects of SN feedback, it affects the transition from non-discs to discs in two other major ways, namely trough frequent merger-driven disc flips in low-mass galaxies, and through the compactions to massive bulges above the same critical mass, which make the extended rings long-lived.

\textbf{El-Badry et al. (2018)} also point out that the stellar systems tend to be even less rotationally supported than the gas, particularly at low masses. This is consistent with or findings in \cite{2015}. Indeed, \textbf{Wheeler et al. (2018)} performed very high resolution simulations of dwarf galaxies, and found stellar kinematics with very low rotational support, \( V_{\text{rot}}/\sigma \sim 0.5 \),
for a few dwarfs in the range $M_s = 10^{5.7} M_\odot$. It will be useful to obtain analogous kinematics for the gas component.

7.2 Tentative comparison to observations

The existing observations of gas kinematics are largely limited either to low redshifts or to masses above the predicted threshold for disc dominance. Furthermore, the selection is typically biased toward disc morphologies, which makes it hard to explore the predicted non-disc zone below $M_v \sim 10^{9} M_\odot$. We summarize here preliminary comparisons to existing observations. More suitable observations will probably have to wait for studies using future tools such as JWST.

7.2.1 Predicted abundance of disc galaxies

Given the threshold mass for discs, the expected abundance of disc galaxies in a given redshift can be estimated by the number density of haloes above the threshold mass. Figure 9 shows the comoving number density of haloes as a function of redshift, as derived by the Press-Schechter formalism for haloes above the threshold mass. Shown are curves for $M_\sigma > 10^{11} M_\odot$ and $> 2 \times 10^{11} M_\odot$, as could be deduced from Fig. 1.

Figure 9. Comoving number density for discs as a function of redshift, as derived by the Press-Schechter formalism for haloes above the threshold mass. Shown are curves for $M_\sigma > 10^{11} M_\odot$ and $> 2 \times 10^{11} M_\odot$, as could be deduced from Fig. 1.

7.2.2 Gas kinematics at low redshift: SAMI

The SAMI survey provides HI gas kinematics for galaxies at low redshift in the mass range $M_v = 10^{8.5-11} M_\odot$ (Cortese et al. 2014, Fig. 2). This survey was supposedly selected blind to morphology, though in practice it seems to be biased at low masses towards late type discs, Sc or later (their Fig. 2, bottom panels). Figure 10 shows the disciness, as measured by $V_{rot}/\sigma$, versus virial mass and versus stellar mass, for the VELA simulations compared to the observational estimates from the SAMI survey. The data provided are $V_{rot}$ and $\sigma$ as a function of stellar mass. The virial halo mass was determined from the stellar mass using our functional fit to the $M_\sigma - M_v$ relation by Rodríguez-Puebla et al. (2017), $M_v = 25 M_\sigma (M_{s,10.6} + M_{s,10.6})$, where $M_{s,10.6} = M_\sigma/10^{10.6} M_\odot$. The stellar-to-halo mass ratio is slightly higher in the VELA3 simulations, which introduces a horizontal uncertainty in the comparison between the simulations and observations. We note that the observations are at low $z$, while the simulations span a range of redshifts, generally monotonic with mass as the galaxies grow in time, such that the low-mass tail largely refers to $z>4$.

We see a qualitative agreement between theory and observations, both showing a similar systematic increase of disciness with mass, with a steep drop below $M_v \sim 2 \times 10^{11} M_\odot$. The simulation disciness seems to be marginally higher than the observational estimates at large masses, though this difference is only at the level of $1\sigma$ or less. This may well be due to the different ways $V_{rot}$ and $\sigma$ are evaluated in 2D (in the observations) versus 3D (in the simulations). In particular, because of the limited resolution, the observations naturally tend to underestimate $V_{rot}$ and overestimate $\sigma$. The observational analyses attempt to correct for this effect, but this may be an under-correction or an over-correction, leaving a significant uncertainty in the measured $V_{rot}/\sigma$.

El-Badry et al. (2018) show the observed Tully-Fisher relation based on the gas rotation velocities from the SAMI survey. Here, too, the data show a steeper slope in the range $M_v \sim 10^{8.5-9} M_\odot$, crudely $V_{rot} \propto M_v^{0.7}$ or steeper, compared to a log slope $\sim 0.3$ in the higher-mass range $\sim 10^{9-10} M_\odot$, just above the critical mass. These are to be compared to the much flatter slope that one would have obtained by assuming rotational support, $V_{rot} \propto V_c$, and $M_v \propto M_\sigma^2$ from eq. (10), which would give $V_{rot} \propto M_\sigma^{1/3} \propto M_v^{1/6}$. The steeper observed slope is qualitatively consistent with our predicted lower rotational support at lower masses, particularly below the critical mass.
7.2.3 Gas kinematics above the critical mass: Simons et al.

Kassin et al. (2012 Fig. 11) and Simons et al. (2017 Fig. 4) show their estimate of the fraction of observed disc galaxies with $V_{\text{rot}}/\sigma > 3$ as a function of stellar mass and redshift. The former reference spans the range $M_8 = 10^{8-10} M_\odot$ at $z = 0.2 - 1.1$ and the latter more comprehensive study covers $M_8 = 10^{9-11} M_\odot$ at $z = 0.2 - 2.4$. Figure 11 shows the fraction of discs ($V_{\text{rot}}/\sigma > 3$) as a function of redshift in stellar mass bins, in the VELA simulations compared to the observational estimates by Simons et al. (2017 Fig. 4b).

The data are for three stellar mass bins, in the limited overlapping redshift range. The observations show a systematic mass dependence at any given redshift, crudely $f_{\text{disc}} \propto M_8^{0.3-0.4}$. This is similar to what we find in the VELA simulations in the same mass and redshift range. Note that the mass range studied by Simons et al. (2017) is limited to near and above the predicted transition mass.

The observations by Simons et al. (2017) also indicate in the given redshift range a redshift dependence of $f_{\text{disc}} \propto (1 + z)^{0.8}$ for $M_8 = 10^{9-10} M_\odot$, which seems to be in apparent conflict with our general prediction of no significant redshift dependence over the larger redshift range $z = 1 - 8$. However, when focusing on the limited redshift range $z = 0.8 - 2.4$, where there are both simulated and observed galaxies, as in Fig. 11 the simulations predict a similar redshift dependence for the two low mass bins. In the most massive bin the simulations predict a higher disc fraction than estimated observationally at $z \sim 2$. On one hand, this difference may reflect a flaw in the simulations. For example, as mentioned in [1] a redshift dependence of the stellar-to-halo mass ratio, e.g., as deduced by Rodríguez-Puebla et al. (2017) contrary to Behroozi et al. (2018), would have introduced a redshift dependence in the disc fraction along the lines of the observations over a large redshift range. On the other hand, the difference in the high-mass bin at $z \sim 2$ could reflect differences in the ways the measurements of $V_{\text{rot}}$ and $\sigma$ are performed in the simulations and observations, and especially the effects of resolution on the kinematic ratio derived from observations.

7.2.4 Gas kinematics in low redshift discs: THINGS

El-Badry et al. (2018) summarize observations of HI gas specific AM. In their Fig. 13, they show the specific AM of HI gas, $j_{\text{HI}}$, versus baryonic mass, as measured in the HI surveys THINGS (Obreschkow & Glazebrook 2014) and LITTLE THINGS (Butler, Obreschkow & Off 2017) as well as from van den Bosch, Burkert & Swaters (2001), and Chowdhury & Chengalur (2017). These data indicate a marginal bend of the $j_{\text{HI}} - M_{\text{bar}}$ relation near $M_{\text{bar}} \sim 10^9 M_\odot$, as predicted. The logarithmic slope of the relation above this mass is $\sim 0.4$ and below this mass it is $\sim 0.75$. These slopes can be compared to what one would have deduced from the halo virial relations. Assuming that the effective gas radius and rotation velocity are constant fractions of $R_e$ and $V_e$, one obtains $j_{\text{gas}} \propto V_e R_e \propto M_e^{2/3}$. Adopting $M_{\text{bar}} \propto M_e^{1/7}$ from eq. (14), the expected relation becomes $j_{\text{gas}} \propto M_e^{0.4}$. This is consistent with the observed slope just above the bend, while the slope below the bend is steeper, indicating less rotational support for
lower masses. When plotted against $M_V$, the change of slope is expected to become more pronounced. This is qualitatively along the lines of our predictions of a gradient of rotational support as a function of mass, with the transition near $M_*=10^{9.0-10.5}$. Unfortunately, the observed galaxies were selected to have a disc morphology, so the sample misses the low-rotation non-discs at low masses, and may thus miss much of the predicted effect.

7.2.5 Stellar kinematics of dwarfs: Wheeler et al.

Wheeler et al. (2017) performed a Bayesian analysis of the observed $V_{\text{rot}}/\sigma$ for stars in 40 dwarf galaxies in the local volume over $M_*=10^{3.5-8} M_\odot$, namely below the predicted critical mass for rotation-supported discs. Their sample includes both dwarf spheroidals and star-forming dwarf irregulars. They find that 80 percent of the stellar systems in these dwarfs are dispersion supported. They see no clear trend between the stellar $V_{\text{rot}}/\sigma$ and $M_*$ in the mass range $M_*=10^{3.5-6} M_\odot$, which we note is well below the threshold mass for disc formation. They also found that in four FIRE simulated isolated galaxies in the range $M_*=10^{9-10} M_\odot$ the values of the stellar $V_{\text{rot}}/\sigma$ are consistent with the observed values. While this is a hint, these results are not straightforwardly comparable to our predictions as extending our analysis of gaseous discs to stellar kinematics is beyond the scope of our current paper.

7.2.6 Rings in massive galaxies at high $z$: Genzel et al.

We predict that above the threshold mass the rotating gas discs are in many cases clumpy rings surrounding a massive body. When observing such rings, one should be careful to correct for dust, which may introduce false rings. More details concerning the expected ring properties are provided in Dekel et al. (2020).

At $z \sim 2$, Genzel et al. (2014) detected a significant fraction of massive galaxies with extended $H_\alpha$ star-forming rings, most of which surrounding a central massive stellar bulge. Genzel et al. (2019, in prep.), following Genzel et al. (2017), who determined rotation curves for 40 massive star-forming galaxies at $z=0.6-2.7$ using data from the 3D-HST/KMOS$^{3D}$/SINFONI/NOEMA surveys, verify the common existence of extended gas rings. Some of the rings surround massive compact bulges, typically with little dark-matter mass within the effective radius, while other rings surround less massive bulges but with higher central dark-matter masses. In both cases this is consistent with the presence of a large central mass.

In an ongoing work (Zhiyuan, Giavalisco et al. 2020, in prep.), we find many star-forming ring galaxies in medium-deep HST images observed in the U-F366W bandpass in the GOODS-S field, complementing them with the corresponding B-F453W and H-F160W images, within the CANDELS survey. Our preliminary inspections indicate that, among the galaxies of $M_*>10^{9.5} M_\odot$ at $z=0.5-2$, a significant fraction show star-forming, clumpy rings, typically surrounding a massive bulge, which is either star forming or quenched.

At low redshifts, Salim, Fang & Rich (2012) found that most “Green-Valley” galaxies, at the early stages of their quenching process, consist of massive quenched bulges surrounded by star-forming rings. These rings may be related to the high-$z$ rings discussed above as the Green Valley may be the low-$z$ analog of the post-compaction phase that triggers the quenching at high $z$ (Tacchella et al. 2016b).

8 CONCLUSION

We predict that galactic gas discs are likely to survive only in dark-matter haloes of mass above a threshold of $\sim 2 \times 10^{11} M_\odot$, corresponding to a stellar mass of $\sim 10^{9.5} M_\odot$, with only little dependence on redshift. This is derived by analytic toy models and confirmed in zoom-in hydrodynamical cosmological simulations. In haloes of lower masses, the gas does not tend to settle into an extended long-lived rotating disc, as several different mechanisms act to reduce the angular momentum and thus disrupt the disc.

Most importantly, the angular momentum is predicted to flip on a timescale shorter than the orbital timescale due to mergers associated with a pattern-change in the cosmic-web streams that feed the galaxy with angular momentum. We show analytically why the mass threshold for flips is due to the dependence of the baryon-to-halo mass ratio on mass. The redshift dependencies of the ratio of accretion-to-orbital timescales...
and of the stellar fraction in the galaxy balance each other, giving rise to a mass threshold that hardly varies with redshift.

Secondly, in this pre-compaction low-mass regime, violent disc instability exerts torques that drive angular-momentum out and mass in, thus making the disc shrink in a few orbital times. Furthermore, in this regime the central dark-matter and stellar system tend to be prolate and thus capable of producing torques that reduce the angular momentum of the incoming new gas.

We argue that supernova feedback has an important role in the disruption of discs below the critical mass. It determines the stellar-to-halo mass ratio that affects the merger rate, it stirs up turbulence that puffs up the disc and it suppresses the supply of new gas with high angular momentum, possibly even ejecting high-AM gas from the disc outskirts. Supernova feedback also confines the major compaction events to near and above the critical mass and thus stabilize the extended rings above this threshold.

Our simulations indicate only a weak redshift dependence of disc survival. This is associated with an insensitivity of the stellar-to-halo mass ratio to redshift, as implied by certain abundance matching analyses (Behroozi et al. 2018). A preferred occurrence of discs of a given mass at late times could arise instead if the stellar-to-halo mass ratio is actually increasing with time (Rodríguez-Puebla et al. 2017). The weak redshift dependence is also a manifestation of the weak redshift dependence of the effectiveness of supernova feedback, which is primarily determined by the depth of the potential well, a function of the halo virial velocity which is primarily a function of halo mass.

Above the critical mass, the disruptive mergers are less frequent and are not necessarily associated with a change in the pattern of the feeding streams, allowing the discs to survive for several orbital times. In parallel, the effects of supernova feedback are reduced due to the depth of the halo potential well.

Potential disc disruption by inward mass transport due to violent disc instability is limited to below a similar critical mass. In the post-compaction regime, when the galaxy mass is typically above the threshold mass, the simulations show that the inflowing high-AM streams settle into long-lived extended discs that evolve into rings. By computing the torques exerted by a tightly wound spiral structure on an extended ring, we show in Dekel et al. (2020) that the timescale for mass inflow is roughly $\delta_d^{-1} t_{\text{orb}}$, where $\delta_d$ is the disc-to-total mass ratio. This implies that the post-compaction massive bulge, that appears above the critical mass for blue nuggets, acts to stabilize the ring against shrinkage by reducing $\delta_d$ to values well below unity. Preliminary observational studies in progress indeed indicate detections of gaseous star-forming rings around massive bulges or dark-matter dominated centers in a large fraction of $z \sim 1$–2 galaxies above the threshold mass, consistent with the predicted role of compaction in the longevity of extended rings.

The threshold of $M_c \sim 1.4 \times 10^{11} M_\odot$ implies a disc comoving number density of $n \sim 10^{-2}, 10^{-3}, 10^{-5} \text{Mpc}^{-3}$ at $z = 2, 6, 10$. The results from existing kinematical observations are consistent with the predicted mass dependence and threshold mass for long-lived gas discs. These include the studies based on the SAMI survey (Cortese et al. 2014) and the analysis of observations by Simons et al. (2017). However, the current observations hardly explore the mass range below $M_c \sim 10^9 M_\odot$ at moderate and high redshifts, where our simulations and analysis indicate that disc disruption is expected to dominate. In several cases (e.g. THINGS) the selection is biased toward discs, thus disabling a study of non-discs at low masses. Future observations, e.g. with JWST, may allow an exploration of this regime.

ACKNOWLEDGMENTS

We are grateful for a stimulating interaction with Doug Lin on torques from spiral structure, and for stimulating discussions with Sandy Faber, Reinhard Genzel and David Koo. This work was partly supported by the grants Germany-Israel GIF 1-1341-303.7/2016, Germany-Israel DIP STE1869/2-1 GE625/17-1, I-CORE Program of the PBC/ISF 1829/12, ISF 857/14, US-Israel BSF 2014-273, and NSF AST-1405962. The cosmological VELA simulations were performed at the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory, and at NASA Advanced Supercomputing (NAS) at NASA Ames Research Center. Development and analysis have been performed in the astro cluster at HU.

REFERENCES

Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Barro G. et al., 2017, ApJ, 840, 47
Barro G. et al., 2013, ApJ, 765, 104
Behroozi P., Wechsler R., Hearin A., Conroy C., 2018, arXiv e-prints
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 762, L31
Bett P. E., Frenk C. S., 2012, MNRAS, 420, 3324
Bett P. E., Frenk C. S., 2016, MNRAS, 461, 1338
Binney J., 1977, ApJ, 215, 483
Binboin Y., Dekel A., 2003, MNRAS, 345, 349
Binboin Y., Dekel A., Neistein E., 2007, MNRAS, 380, 339
Bryan G. L., Norman M. L., 1998, ApJ, 495, 80
Butler K. M., Obreschkow D., Oh S.-H., 2017, ApJ, 834, L4
Cattaneo A., Teyssier R., 2007, MNRAS, 376, 1547
Ceverino D., Dekel A., Bournaud F., 2010, MNRAS, 404, 2151
Ceverino D., Dekel A., Mandelker N., Bournaud F., Burkert A., Genzel R., Primack J., 2012, MNRAS, 424, 1284
APPENDIX A: THE VELA COSMOLOGICAL SIMULATIONS

The VELA suite consists of hydro-cosmological simulations of 34 moderately massive galaxies. Full details are presented in Ceverino et al. (2014); Zolotov et al. (2015). This suite has been used to study central issues in the evolution of galaxies at high redshifts, including compaction to blue nuggets and the trigger of quenching (Zolotov et al. 2015); Tacchella et al. (2016b), evolution of global shape (Ceverino, Primack & Dekel 2015; Tomassetti et al. 2016), violent disc instability (Mandelker et al. 2014, 2017), OVI in the CGM (Roca-Fàbrega et al. 2018), and galaxy size and angular momentum (Jiang et al. 2019). Additional analysis of the same suite of simulations are discussed in Moody et al. (2014). Snyder et al. (2015). In this appendix we give an overview of the key aspects of the simulations and their limitations.

A1 The Cosmological Simulations

The VELA simulations make use of the Adaptive Refinement Tree (ART) code (Kravtsov, Klypin & Khokhlov 1997; Kravtsov 2003; Ceverino & Klypin 2009), which accurately follows the evolution of a gravitating N-body system and the Eulerian gas dynamics using an adaptive mesh refinement approach. The adaptive mesh refinement maximum resolution is 17 − 35 pc at all times, which is achieved at densities of ∼ 10^{-4} − 10^{-3} cm^{-3}. Beside gravity and hydrodynamics, the code incorporates physical process relevant for galaxy formation such as gas cooling by atomic hydrogen and helium, metal and molecular hydrogen cooling, photoionization heating by the UV background with partial self-shielding, star formation, stellar mass loss, metal enrichment of the ISM and stellar feedback. Supernovae and stellar winds are implemented by local injection of thermal energy as described in Ceverino & Klypin (2009); Ceverino, Dekel & Bournaud (2010) and Ceverino et al. (2012). Radiation-pressure stellar feedback is implemented at a moderate level, following Dekel et al. (2013), as described in Ceverino et al. (2014).

Cooling and heating rates are tabulated for a given gas density, temperature, metallicity and UV background based on the CLOUDY code (Ferland et al. 1998), assuming a slab of thickness 1 kpc. A uniform UV background based on the redshift-dependent Haardt & Madau (1996) model is assumed, except at gas densities higher than 0.1 cm^{-3}, where a substantially suppressed UV background is used (5.9 × 10^{6} erg s^{-1} cm^{-2} Hz^{-1}) in order to mimic the partial self-shielding of dense gas, allowing dense gas to cool down to temperatures of ∼ 300K. The assumed equation of state is that of an ideal mono-atomic gas. Artificial fragmentation on the cell size is prevented by introducing a pressure floor, which ensures that the Jeans scale is resolved by at least 7 cells (see Ceverino, Dekel & Bournaud 2010).

Star particles form in timesteps of 5 Myr in cells where the gas density exceeds a threshold of 1 cm^{-3} and the temperatures is below 10^{4}K. Most stars (> 90%) end up forming at temperatures well below 10^{3}K, and more than half of the stars form near 300K in cells where the gas density is higher than 10^{-3}. The code implements a stochastic star-formation where a star particle with a mass of 42% of the gas mass forms with a probability \( P = \left(\frac{\rho_{\text{gas}}}{10^{3} \text{cm}^{-3}}\right)^{1/2} \) but not higher than 0.2. This corresponds to a local SFR that crudely mimics \( \rho_{\text{gas}} \epsilon_{\text{ff}} / \epsilon_{\text{ff}} \text{with } \epsilon_{\text{ff}} \sim 0.02 \). A stellar initial mass function of Chabrier (2003) is assumed.

Thermal feedback that mimics the energy release from stellar winds and supernova explosions is incorporated as a constant heating rate over the 40 Myr following star formation. A velocity kick of ∼ 10 km s^{-1} is applied to 30% of the newly formed stellar particles – this enables SN explosions in lower density regions where the cooling may not overcome the heating without implementing an artificial shutdown of cooling (Ceverino & Klypin 2009). The code also incorporates the later effects of Type Ia supernova and stellar mass loss, and it follows the metal enrichment of the ISM.

Radiation pressure is incorporated through the addition of a non-thermal pressure term to the total gas pressure in regions where ionizing photons from massive stars are produced and may be trapped. This ionizing radiation injects momentum in the cells neighbouring massive star particles younger than 5 Myr, and whose column density exceeds 10^{21} cm^{-2}, isotropically pressurizing the star-forming regions (see more details in Agertz et al. 2013; Ceverino et al. 2014).

The initial conditions for the simulations are based on dark-matter haloes that were drawn from dissipationless N-body simulations at lower resolution in cosmological boxes of 15 − 60 Mpc. The LCDM cosmological model was assumed with the WMAP5 values of the cosmological parameters, \( \Omega_{\text{m}} = 0.27, \Omega_{\Lambda} = 0.73, \Omega_{\text{b}} = 0.045, h = 0.7, \sigma_{8} = 0.82 \) (Komatsu et al. 2009). Each halo was selected to have a given virial mass at \( z = 1 \) and no ongoing major merger at \( z = 1 \). This latter criterion eliminated less than 10% of the haloes, those that tend to be in a dense, proto-cluster environment at \( z \sim 1 \). The virial masses at \( z = 1 \) were chosen to be in the range \( M_{\text{vir}} = 2 \times 10^{11} − 2 \times 10^{12} M_{\odot} \), about a median of 4.6 × 10^{11} M_{\odot}. If left in isolation, the median mass at \( z = 0 \) was intended to be \( \sim 10^{12} M_{\odot} \).
ΩΛ(z) and Ωm(z) are the cosmological parameters at z (Bryan & Norman 1998; Dekel & Birnboim 2006). The stellar mass M⋆ is the instantaneous mass in stars within a radius of 0.2Rc, accounting for past stellar mass loss.

We start the analysis at the cosmological time corresponding to expansion factor a = 0.125 (redshift z = 7). As can be seen in Table A1 most galaxies reach a = 0.50 (z = 1). Each galaxy is analyzed at output times separated by a constant interval in a, Δa = 0.01, corresponding at z = 2 to ∼100 Myr (roughly half an orbital time at the disc edge). The sample consists of totally ∼1000 snapshots in the redshift range z = 7 − 0.8 from 35 galaxies that at z = 2 span the stellar mass range (0.2 − 6.4) × 10^{11}M⊙. The half-mass sizes Rc are determined from the M⋆ that are measured within a radius of 0.2Rc and they range R∞ ≃ 0.4 − 3.2 kpc at z = 2.

The SFR for a simulated galaxy is obtained by SFR = ⟨M⋆(tage < tmax)/tmax⟩, where M⋆(tage < tmax) is the mass at birth in stars younger than tmax. The average ⟨⟩tmax is obtained by averaging over all tmax in the interval [40, 80] Myr in steps of 0.2 Myr. The tmax in this range are long enough to ensure good statistics. The SFR ranges from ∼1 to 33M⊙ yr^{-1} at z ∼ 2.

The instantaneous mass of each star particle is derived from its initial mass at birth and its age using a fitting formula for the mass loss from the stellar population represented by the star particle, according to which 10%, 20% and 30% of the mass is lost after 30 Myr, 260 Myr, and 2 Gyr from birth, respectively. We consistently use here the instantaneous stellar mass, Ms, and define the specific SFR by sSFR = SFR/Ms.

The determination of the centre of the galaxy is outlined in detail in Appendix B of Mandelker et al. (2014). Briefly, starting from the most bound star, the centre is refined iteratively by calculating the centre of mass of stellar particles in spheres of decreasing radii, updating the centre and decreasing the radius at each iteration. We begin with an initial radius of 600 pc, and decrease the radius by a factor of 1.1 at each iteration. The iteration terminates when the radius reaches 130 pc or when the number of stellar particles in the sphere drops below 20.

The disc plane and dimensions are determined iteratively, as detailed in Mandelker et al. (2014). The disc axis is defined by the angular momentum of cold gas (T < 1.5 × 10^4K), which on average accounts for ∼97% of the total gas mass in the disc. The radius Rd is chosen to contain 85% of the cold gas mass in the galactic mid-plane out to 0.15Rc, and the half-height Hd is defined to encompass 85% of the cold gas mass in a thick cylinder where both the radius and half-height equal Rd.

Relevant global properties of the VELA 3 galaxies at z = 2 are listed in Table A1 and explained in the caption. It includes the global masses and sizes of the different components, the shape and kinematic properties.

We attempt to identify the major event of compaction to a blue nugget for each galaxy. This is the one that leads to a significant central gas depletion and SFR quenching, and marks the transition from dark-matter to baryon dominance within R∞. Following Zolotov et al. (2015) and Tacchella et al. (2016a), the most physical way to identify the compaction and blue nugget is by the steep rise of gas density (and SFR) within the inner 1 kpc to the highest peak, as long as it is followed by a significant, long-term decline in central gas mass density (and SFR). The onset of compaction can also be identified as the start of the steep rise of central gas density prior to the blue-nugget peak. An alternative identification is using the shoulder of the stellar mass density within 1 kpc where its rise due to starburst associated with the compaction turns into a plateau of maximum long-term compactness slightly after the blue-nugget peak of gas density. This is a more practical way to identify blue nuggets in observations (e.g. Barro et al. 2017).

Major mergers in the history of each galaxy, for the limited purpose they are used here, are identified in a simplified way by following sudden increases in the stellar mass.

APPENDIX B: COMPLEMENTARY FIGURES

Here we show complementary relevant images from the VELA simulations.

Figure B1 is the edge-on analog of Fig. 3. The galaxies above the critical mass are perturbed, thick, turbulent, rotating discs, with axial ratios ∼4. At and below the critical mass there is no evidence for discs, and ongoing mergers can be seen.

Figure B2 displays six examples of gas in galaxies in the pre-compaction and compaction phases, showing ongoing wet mergers in progress.
Figure B1. Projected gas density in example VELA galaxies in different zones in the $M_v - z$ plane, the edge-on counterpart of Fig. 8. The rows from top to bottom represent galaxies above, near and below $M_v = 10^{11} M_\odot$, while the columns from left to right represent low, intermediate and high redshifts. The projections are edge-on with respect to the angular momentum.
Figure B2. Several more examples of pre-compaction non-discs, complementary to Fig. 3. Shown are face-on projected images of gas density in six VELA galaxies at the pre-compaction phase or during the compaction process itself. They display a typical non-disc perturbed morphology and sometimes showing wet mergers in progress.