Toy Models for Galaxy Formation versus Simulations

A. Dekel\textsuperscript{1}, A. Zolotov\textsuperscript{1}, D. Tweed\textsuperscript{1}, M. Cacciato\textsuperscript{1,2}, D. Ceverino\textsuperscript{1,3}, J.R. Primack\textsuperscript{4}

\textsuperscript{1}Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel; (avishai.dekel@mail.huji.ac.il)
\textsuperscript{2}Leiden Observatory, Leiden University, Niels Bohrweg 2, NL-2333 CA Leiden, The Netherlands
\textsuperscript{3}Departamento de Fisica Teorica, Universidad Autonoma de Madrid, Madrid E-28049, Spain
\textsuperscript{4}Department of Physics, University of California, Santa Cruz, CA, 95064, USA

6 February 2014

ABSTRACT
We describe simple useful toy models for key processes of galaxy formation in its most active phase, at $z > 1$, and test the approximate expressions against the typical behaviour in a suite of high-resolution hydro-cosmological simulations of massive galaxies at $z = 4 - 1$. We address in particular the evolution of (a) the total mass inflow rate from the cosmic web into galactic haloes based on the EPS approximation, (b) the penetration of baryonic streams into the inner galaxy, (c) the disc size, (d) the implied steady-state gas content and star-formation rate (SFR) in the galaxy subject to mass conservation and a universal star-formation law, (e) the inflow rate within the disc to a central bulge and black hole as derived using energy conservation and self-regulated $Q \sim 1$ violent disc instability (VDI), and (f) the implied steady state in the disc and bulge. The toy models provide useful approximations for the behaviour of the simulated galaxies. We find that (a) the inflow rate is proportional to mass and to $(1 + z)^{5/2}$, (b) the penetration to the inner halo is $\sim 50\%$ at $z = 4 - 2$, (c) the disc radius is $\sim 5\%$ of the virial radius, (d) the galaxies reach a steady state with the SFR following the accretion rate into the galaxy, (e) there is an intense gas inflow through the disc, comparable to the SFR, following the predictions of VDI, and (f) the galaxies approach a steady state with the bulge mass comparable to the disc mass, where the draining of gas by SFR, outflows and disc inflows is replenished by fresh accretion. Given the agreement with simulations, these toy models are useful for understanding the complex phenomena in simple terms and for back-of-the-envelope predictions.

Key words: cosmology — dark matter — galaxies: evolution — galaxies: formation — galaxies: haloes

1 INTRODUCTION
A scientific understanding of a physical phenomenon is often materialized through a simple, idealized, analytic model, which we call a toy model. A toy model is derived from first principles in an attempt to capture the key physical elements of the process and to provide a useful approximation to the complex behaviour seen in numerical simulations or in observations. The simple scaling relations from such a toy model allow back-of-the-envelope calculations that involve several elements or a sequence of events, which may help to develop an understanding for a whole scenario, and hopefully lead to explanations of observed phenomena or to new theoretical predictions.

A different tool in the study of galaxy formation is semi-analytic modeling (SAM, e.g., Somerville et al. 2008, Guo et al. 2011, Benson 2012). In SAMs, the many complex physical processes associated with gas, stars, black holes and radiation are modeled in parallel by recipes with free parameters, incorporated in retrospect in dark-matter halo merger trees that were obtained from dissipationless cosmological N-body simulations or the Extended Press-Schechter approximation (EPS, Bond et al. 1991). The toy models addressed here are much simpler than the sophisticated SAMs. The toy models are typically based on simple analytic arguments using crude approximations, and they tend to focus on one process at a time. These toy models could serve as a basis for physical recipes to be incorporated in SAMs. SAM recipes have been confronted with the more elaborate hydro-cosmological simulations, where gas processes are explicitly simulated (e.g., Benson et al. 2001, Cattaneo et al. 2007, Neistein et al. 2012, Hirschmann et al. 2012), but this is commonly done in the context of a full SAM. Here we compare
each toy model on its own to the corresponding feature of the simulation results.

We address toy models for the formation and evolution of galaxies in their most active phase, at redshifts $z > 1$, where the Einstein-deSitter (EdS) cosmological model serves as a useful approximation. We focus on central galaxies in their haloes, the main progenitors of central galaxies at $z \sim 1$. In particular, we address several key processes as follows. (a) We verify the inflow rate of dark matter and baryons from the cosmic web into distinct haloes based on the EPS approximation. (b) We evaluate the penetration of baryonic streams through the hot gas in the haloes into the inner disc galaxies. (c) We crudely estimate the disc radius as a function of the virial radius compared to the standard model where angular momentum is conserved. (d) We study the implied gas content and star-formation rate (SFR) in the galaxy subject to mass conservation (the “bathtub” model) and a universal SFR law. (e) We address the inflow rate within the disc to a central bulge as derived using energy conservation and Toomre Violent Disc Instability (VDI). (f) We study the implied steady state in the disc as it is fed by cold streams and is drained by star formation, outflows, and inflow to the bulge.

To test the validity of the toy models, we compare their predictions with the results obtained from a suite of 27 Adaptive Mesh Refinement (AMR) hydro-cosmological simulations that zoom in on individual galaxies with a resolution of $\sim 50$ pc. These are state-of-the-art simulations that resolve the key processes addressed. The feedback in these simulations is limited to the thermal effects of stellar winds and supernovae, not addressing radiative feedback and AGN feedback, thus not allowing an immediate study of the effects of very strong outflows and the associated severe suppression of SFR.

The outline of this paper is as follows. In §2 we describe the simulations. In §3 we address the mass inflow rate into haloes. In §4 we study the penetration of cold streams into the central galaxy at the inner halo. In §5 we refer to the disc size as derived from conservation of specific angular momentum. In §6 we investigate the steady-state of the gas and stellar content in which the SFR follows the accretion rate. In §7 we address the VDI-driven inflow within the disc into the bulge. In §8 we refer to the resultant steady state in the disc. In §9 we summarize our conclusions.

2 SIMULATIONS

We use zoom-in hydro cosmological simulations of 27 moderately massive galaxies with an AMR maximum resolution $35 - 70$ pc, all evolved from high redshift to $z = 2$, many to $z \sim 1.5$ and some reaching $z = 1$, as marked by $a_{\text{in}}$ in Table 1. They utilize the Adaptive Refinement Tree (ART) code (Kravtsov, Klypin & Khokhlov 1997; Ceverino, Dekel & Bournaud 2010), which accurately follows the evolution of a gravitating N-body system and the Eulerian gas dynamics using an adaptive mesh. Beyond gravity and hydrodynamics, the code incorporates at the subgrid level many of the physical processes relevant for galaxy formation. They include gas cooling by atomic hydrogen and helium as well as by metals and molecular hydrogen, photoionization heating by the UV background with partial self-shielding, star formation, stellar mass loss, metal enrichment of the ISM, and feedback from stellar winds and supernovae, implemented as local injection of thermal energy (Ceverino & Klypin 2009; Ceverino, Dekel & Bournaud 2010; Ceverino et al. 2012).

2.1 Subgrid Physics

A few relevant details concerning the subgrid physics are as follows. 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 ($5.9 \times 10^{-26}$ erg cm$^{-2}$ Hz$^{-1}$) in order to mimic the partial self-shielding of dense gas. This allows the dense gas to cool down to temperatures of $\sim 300$ K. 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 formation is assumed to occur at cell densities above a threshold of $1$ cm$^{-3}$ and at temperatures below $10^4$ K. More than 90% of the stars form at temperatures well below $10^4$ K, and more than half the stars form at 300 K in cells where the gas density is higher than $10$ cm$^{-3}$. The code implements a stochastic star-formation model that yields a star-formation efficiency per free-fall time of 5%. At the given resolution, this efficiency roughly mimics by construction the empirical Kennicutt-Schmidt law (Schmidt 1959; Kennicutt 1998). When extracting the resultant approximate SFR, we consider the mass in stars born in a given timestep divided by the elapsed time. The code incorporates a thermal stellar feedback model, in which the combined energy from stellar winds and supernova explosions is released as a constant heating rate over 40 Myr following star formation, the typical age of the lightest star that explodes as a type-II supernova. The heating rate due to feedback may or may not overcome the cooling rate, depending on the gas conditions in the star-forming regions (Dekel & Silk 1986; Ceverino & Klypin 2009). We note that no shutdown of cooling or any other artificial mechanism for boosting the feedback effects is implemented in these simulations. We also include the effect of runaway stars by assigning a velocity kick of $\sim 10$ km s$^{-1}$ to 30% of the newly formed stellar particles. The code also includes the later effects of type-Ia...
Table 1. The suite of simulated galaxies. Quoted at $a = 0.33$ ($z = 2$) are the virial radius $R_v$ and within it the total mass $M_v$, stellar mass $M_{\text{star}}$, and gas mass $M_g$. The snapshots used are from $a = 0.2$ ($z = 4$) to $a = a_{\text{fin}}$ ($z_{\text{fin}} = a_{\text{fin}} - 1 = 2 - 1$), and the virial mass at $a_{\text{fin}}$ is quoted.

| Galaxy | $R_v$ (kpc) | $M_v$ ($10^{12}M_\odot$) | $M_{\text{star}}$ ($10^{11}M_\odot$) | $M_g$ ($10^{11}M_\odot$) | $a_{\text{fin}}$ | $M_v(a_{\text{fin}})$ ($10^{12}M_\odot$) |
|--------|-------------|--------------------------|----------------------------|--------------------------|----------------|---------------------------------|
| MW01   | 102         | 0.81                     | 0.72                       | 0.57                     | 0.42           | 1.39                           |
| MW02   | 105         | 0.89                     | 2.56                       | 1.12                     | 0.34           | 0.94                           |
| MW03   | 099         | 0.73                     | 0.60                       | 0.51                     | 0.42           | 1.34                           |
| MW04   | 123         | 1.42                     | 1.41                       | 0.89                     | 0.38           | 1.93                           |
| MW07   | 073         | 0.30                     | 0.30                       | 0.22                     | 0.40           | 0.49                           |
| MW08   | 071         | 0.28                     | 0.28                       | 0.15                     | 0.45           | 0.59                           |
| MW09   | 059         | 0.16                     | 0.19                       | 0.08                     | 0.50           | 0.64                           |
| MW10   | 102         | 0.82                     | 0.72                       | 0.44                     | 0.50           | 1.63                           |
| MW11   | 088         | 0.53                     | 0.51                       | 0.28                     | 0.40           | 0.61                           |
| MW12   | 130         | 1.70                     | 2.06                       | 1.01                     | 0.39           | 3.49                           |
| VL01   | 117         | 1.23                     | 1.54                       | 0.75                     | 0.37           | 1.89                           |
| VL02   | 101         | 0.81                     | 0.89                       | 0.46                     | 0.50           | 2.04                           |
| VL03   | 117         | 1.22                     | 1.44                       | 0.76                     | 0.33           | 1.17                           |
| VL04   | 109         | 1.01                     | 1.33                       | 0.51                     | 0.42           | 1.17                           |
| VL05   | 118         | 1.28                     | 1.29                       | 0.75                     | 0.41           | 2.09                           |
| VL06   | 099         | 0.75                     | 0.94                       | 0.32                     | 0.50           | 1.97                           |
| VL07   | 129         | 1.66                     | 2.15                       | 0.82                     | 0.34           | 1.62                           |
| VL08   | 112         | 1.09                     | 1.35                       | 0.46                     | 0.46           | 1.19                           |
| VL09   | 086         | 0.49                     | 0.61                       | 0.24                     | 0.34           | 0.54                           |
| VL10   | 102         | 0.81                     | 0.95                       | 0.44                     | 0.50           | 3.55                           |
| VL11   | 130         | 1.73                     | 2.02                       | 0.81                     | 0.50           | 2.23                           |
| VL12   | 105         | 0.90                     | 0.96                       | 0.51                     | 0.50           | 2.68                           |
| SFG1   | 129         | 1.66                     | 2.10                       | 0.87                     | 0.38           | 1.98                           |
| SFG4   | 112         | 1.09                     | 1.16                       | 0.66                     | 0.38           | 1.31                           |
| SFG5   | 123         | 1.38                     | 1.52                       | 0.78                     | 0.40           | 1.73                           |
| SFG8   | 121         | 1.38                     | 1.70                       | 0.72                     | 0.35           | 1.42                           |
| SFG9   | 135         | 1.89                     | 2.44                       | 1.22                     | 0.48           | 5.34                           |

Figure 1. Growth of virial mass $M_v$ as a function of expansion factor $a = (1 + z)^{-1}$ for the individual simulated galaxies. Each galaxy has been evolved to a final redshift as specified in Table 1. The simulations span about a decade in mass, from $M_v = 0.16 \times 10^{12}$ to $1.9 \times 10^{12}M_\odot$ at $z = 2$. The solid black curves refer to the toy-model prediction, eq. (9), normalized to $M_v = 10^{12}M_\odot$ at $z = 2$. The overall growth pattern is reproduced by the toy model quite well, though the individual galaxies can grow in different rates at different times in their histories.

supernova and stellar mass loss, and it follows the metal enrichment of the ISM.

2.2 Selected Haloes

The initial conditions for the high-resolution, zoom-in, hydrodynamical simulations that are used in this paper were based on dark-matter haloes that were drawn from dissipationless N-body simulations at lower resolution in three large comoving cosmological boxes. The assumed cosmology is the standard ΛCDM model with the WMAP5 values of the cosmological parameters, namely $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$, $\Omega_b = 0.045$, $h = 0.7$ and $\sigma_8 = 0.82$ (Komatsu et al. 2009). Distinct haloes were...
identified about density peaks, and their virial radius and mass, \( R_v \) and \( M_v \), were measured from the radial mass profile such that the mean overdensity within the sphere of radius \( R_v \) equals \( \Delta(a) \) (eq. 3.2). We do not explicitly address the evolution of subhaloes in this work, but rather consider them as part of the accretion onto the central galaxy of their host distinct halo. The haloes for re-simulation were selected to have a given virial mass at \( z = 1 \) (or \( z = 0 \) in a few cases), as specified in Table 1. The only other selection criterion was that the haloes show no ongoing major merger at \( z = 1 \). This eliminates from the sample less than 10% of the haloes which tend to be in a dense environment at \( z \approx 1 \), and it induces only a minor selection effect at higher redshifts.

2.3 Zoom in

The initial conditions corresponding to each of the selected haloes were filled with gas and refined to a much higher resolution on an adaptive mesh within a zoom-in Lagrangian volume that encompasses the mass within twice the virial radius at \( z = 1 \), which is roughly a sphere of comoving radius \( 1 \) Mpc. This was embedded in a comoving cosmological box of side 20, 40 or \( 80 \) h\(^{-1}\)Mpc. Each galaxy has been evolved with the full hydro ART and subgrid physics on an adaptive comoving mesh refined in the dense regions to cells of minimum size between 35-70 pc in physical units at all times. This maximum resolution is valid in particular throughout the cold discs and dense clumps, allowing cooling to \( \sim 300 \)K and gas densities of \( \sim 10^3 \) cm\(^{-3}\). The dark-matter particle mass is \( 6.7 \times 10^5 M_\odot \), and the particles representing groups of stars have a minimum mass of \( 10^8 M_\odot \).

The virial properties of all 27 galaxies in our sample are listed in Table 1. This includes the virial radius and total virial mass at \( z = 2 \), the stellar mass and gas mass within the virial radius at that time, the latest time of analysis for each galaxy in terms of the expansion factor, \( a_{\text{fin}} \), and the virial mass at that time. The galaxies MW01-03 were the basis for the study of VDI in [Ceverino, Dekel & Bournaud (2010)]. The galaxies MW04 and SFG1 were added to the study of giant clumps in VDI discs by [Ceverino et al. (2012)].

2.4 Analysis

We start the analysis presented below at the cosmological time corresponding to expansion factor \( a = 0.2 \) (redshift \( z = 4 \)). At earlier times, the fixed resolution scale typically corresponds to a larger fraction of the galaxy size, which may bias some of the quantities that we wish to study here. All 27 galaxies reach \( a = 0.33 \) (\( z = 2 \)), 17 galaxies reach \( a = 0.4 \) (\( z = 1.5 \)), and only 7 galaxies have been run all the way to \( a = 0.5 \) (\( z = 1 \)). This gradual degradation of the sample after \( z = 2 \) has been taken into account in our analysis.

The output of each simulation is provided at output times separated by a constant interval in \( a \), \( \Delta a \), which for most galaxies is \( \Delta a = 0.01 \). For two galaxies (SFG-8-9) the timestep is twice as small, \( \Delta a = 0.005 \), and for four galaxies (MW01-4) the timestep is twice as large, \( \Delta a = 0.02 \). For every galaxy we analyze the data averaged over timesteps of \( \Delta a = 0.02 \).

The quantities addressed, for example, are mass \( M \), inflow rate \( \dot{M} \), and specific inflow rate \( \dot{M}/M \). For comparison with a toy model at a given time step, we stack the data from all the available galaxies, sometimes after proper scaling motivated by the model in order to minimize possible mass dependence. For each quantity, we compute and show the median and the linear average. Because of the lognormal nature of the distribution (see below, §3.3.2, Fig. 5), we also compute the average of the log within the 90% percentiles. The tails are eliminated here because \( M \) could be negative in a few rare cases due, for example, to a fly-by satellite caught while crossing the virial radius outwards. The exclusion of the outliers also helps reducing fluctuations between time steps. The scatter is typically marked in our figures by shaded areas between the upper and lower 68% percentiles and 90% percentiles. We also show as error-bars an estimate for the error of the mean, \( \sigma/\sqrt{N} \), where \( \sigma \) is the standard deviation over the \( N \) galaxies averaged over at the given time step.

The statistical analysis considers all the galaxies of the sample in the time range \( a = 0.2 - 0.33 \) (\( z = 4 - 2 \)), but only the gradually degrading subsample between \( a = 0.33 \) (\( z = 2 \)) and \( a = 0.5 \) (\( z = 1 \)) according to the value of \( a_{\text{fin}} \) in Table 1. This translates to larger uncertainties in the results after \( z = 2 \). We therefore focus most of our attention on the simulations at \( z \geq 2 \). When addressing detailed disc properties (e.g., in §3.2 and §5), we limit the main analysis to the range \( z = 3 - 2 \).

2.5 The Sample of Galaxies

Figure 1 shows the curves of virial mass growth in time for the different galaxies listed in Table 1. At \( z = 2 \), the sample spans roughly an order of magnitude in halo mass, from \( M_v = 0.16 \times 10^{11} \) to \( 1.9 \times 10^{12} M_\odot \), with virial radii from 70 to 135 kpc. The simulated haloes typically contain within the virial radius a baryon fraction of 0.14, slightly smaller than the cosmic value, of which the stellar fraction is typically 0.10. This suite of galaxies spans the mass range of typical observed massive star-forming galaxies at \( z \sim 2 \) [Förster Schreiber et al. (2009)], and otherwise no major selection criteria was imposed on their properties at \( z \sim 2 - 4 \). As shown in [Ceverino, Dekel & Bournaud (2010)], our first simulated galaxies, MW01-03, are consistent with the observed scattered scaling relations of \( z \sim 2 \) galaxies, including the relation between SFR and stellar mass and the Tully-Fisher relation [Förster Schreiber et al. (2009)].
and redshift range, excluding galaxies that will be approximately a fair sample of galaxies in the relevant mass below). We can therefore assume that this is approx-
Cresci et al. 2009, but see a discussion of deviations below). The mass growth curves shown in Fig. [1] are to be visually compared to the toy model prediction discussed later in eq. [2], which is arbitrarily normalized in this figure to \( M_f = 10^{12} M_\odot \) at \( z = 2 \). The overall growth pattern is similar to the toy model prediction, though the individual galaxies can grow in different rates at different times in their histories. We note that in some cases galaxies tend to keep their rank order with respect to the other galaxies in terms of virial mass, but in other cases galaxies change their rank order drastically during the evolution (e.g., MW03, left panel, solid green line), commonly following major mergers in their histories.

The simulated growth rates are compared to the toy model in more detail in Figs. [3] to [4].

2.6 Limitations of the Current Sample

These simulations are state-of-the-art in terms of the high-resolution AMR hydrodynamics and the treatment of key physical processes at the subgrid level. In particular, they properly trace the cosmological streams that feed galaxies at high redshift, including mergers and smooth flows, and they resolve the violent disc instability that governs the high-
Ceverino, Dekel & Bournaud 2010. Ceverino et al. 2012. When tracing the important small-scale processes involved in galaxy evolution, AMR codes offer a distinct advantage over SPH codes, e.g., because SPH does not accurately model processes such as shocks and hydrodynamical instabilities (e.g. Agertz et al. 2007, Scannapieco et al. 2012, Bauer & Springel 2012). The AMR codes seem to be comparable in their capabilities to new codes using a moving unstructured grid (Bauer & Springel 2012), but the latter are currently limited to significantly lower resolution, on the order of \( \sim 1 \) kpc (e.g., Nelson et al. 2013), which is not sufficient for tracing the key processes of disc instabilities or the evolution of streams.

However, like other simulations, the current simulations do not yet treat the star formation and feedback processes with sufficient accuracy. For example, the code assumes a somewhat high SFR efficiency per free-fall time, it does not follow in detail the formation of molecules and the effect of metallicity on SFR (Krumholz & Dekel 2012), and it does not explicitly include radiative stellar feedback (Murray, Quataert & Thompson 2010, Krumholz & Dekel 2010, Hopkins et al. 2012, Dekel & Krumholz 2013), or AGN feedback (Silk & Rees 1998, Hopkins et al. 2006, Booth & Schaye 2004, Cattaneo et al. 2009). Therefore, the early SFR is overestimated, while the suppression of SFR in small galaxies is underestimated, resulting in excessive early star formation prior to \( z \sim 3 \), by a factor of order 2. As a result, the typical gas fraction and SFR at \( z \sim 2 \) are lower by a factor of \( \sim 2 \)

the average observed values in star-forming galaxies (Ceverino, Dekel & Bournaud 2010, Daddi et al. 2010, Tacconi et al. 2011).

Furthermore, the simulated galactic mass outflow rate is only a fraction of the SFR, where the mass loading factor ranges from zero to unity with an average \( \eta \sim 0.3 \) at \( 0.5 H_\odot \), not reproducing some of the observed strong outflows with mass loading factors of order unity and above (Steidel et al. 2010, Genzel et al. 2011, Dekel & Krumholz 2013). This leads to a stellar fraction of \( \sim 0.1 \) within the virial radius, a factor of \( \sim 2 - 3 \) higher than the observationally indicated value (e.g., Pérez-González et al. 2008, Behroozi, Wechsler & Conroy 2013). These inaccuracies in the SFR, feedback and outflows introduce a limitation on the generality of our testing. Nevertheless, this situation has the advantage that one can test the toy models for cosmological accretion and the response of the galaxy to it without the extra complication of very strong feedback effects. More accurate recipes for star formation and feedback are being incorporated into simulations that we and others are now running, and they will be improved further in the future. This will enable the next generation comparison to toy models.

One should note that the sample of simulated galaxies is not a fair sample of halo or galaxy mass. The halo masses are limited by selection to a rather narrow mass range at the target redshift \( z = 1 \), and therefore to limited mass ranges at earlier redshifts, and the mass distribution at any redshift does not follow the ΛCDM halo mass function. We do not attempt to follow the properties of a population of galaxies as it evolves, partly because of the mass variation at \( z = 1 \), and partly because of the degrading of our sample at late times. Instead, we treat each outputed snapshot independently of its history, assuming that it represents an (almost) arbitrary galaxy of an instantaneous mass \( M \) at \( z \). Since the galaxies are of different masses at any given redshift, we evaluate “average” properties for the simulated galaxies focusing on quantities that are only weakly dependent on mass within the spanned mass range, such as the specific accretion rate. For quantities that do have a significant mass dependence, such as the accretion rate and the cumulative mass growth, we attempt to scale out the mass dependence using the toy model itself prior to stacking the galaxies of different masses.

It should be emphasized that we only follow the growth of the central galaxies in the main-progenitor haloes of the final haloes selected at \( z = 1 \). The other progenitors, satellite and merging haloes (or galaxies), are all considered here as part of the mass inflow onto the main progenitor. The growth of the non-main-progenitor galaxies in the zoom-in simulations is not studied here, partly because they represent only 10-20% of the population of galaxies of the same mass, which mostly consists of central galaxies, not satellites of more massive galaxies. We do not attempt here an extension of the zoom-in simulations to smaller central galaxies because the accuracy at the given resolution becomes limited, and because the proper treatment of feedback
Cosmological Accretion Rate

3 Halos in the Einstein-deSitter Regime

3.1 Halos in the Einstein-deSitter Regime

We consider the \( \Lambda \)CDM cosmology in the Einstein-deSitter (EdS) regime. This is a useful approximation at \( z > 1 \), and it becomes more and more accurate at higher redshifts. In our toy modeling, in the EdS regime, we adopt the following simple relation between the expansion factor \( a \) and the Hubble time \( t \),

\[
a = (1 + z)^{-1} \simeq \left( \frac{t}{t_1} \right)^{2/3},
\]

(1)

\[
t_1 = \frac{2}{3} \Omega_{\text{m}}^{-1/2} H_0^{-1} \simeq 17.5 \text{ Gyr}.
\]

(2)

A comparison to the accurate expressions in the appendix, \( \boxwedge \), reveals that this approximation overestimates \( t \) at \( z = 1 \) by only 5\%, and by 1.6\% at \( z = 2 \) (but by 28\% at \( z = 0 \)). The mean mass density in the Universe is

\[
\rho_0 = \rho_0 a^{-3}, \quad \rho_0 \simeq 2.5 \times 10^{-30} \text{ g cm}^{-3}.
\]

(3)

Inspired by the spherical-collapse model and the virial theorem, a halo is defined as a sphere about a density peak that in the EdS regime encompasses a mean overdensity of \( \Delta \simeq 200 \) above the cosmological background (see eq. \( \boxwedge \)) for a more accurate expression, where \( \Delta \simeq 207, 187 \) and 178 at \( z = 1, 2, 3 \) respectively. The relations between virial mass \( M_v \), radius \( R_v \) and velocity \( V_v \) are thus

\[
V_v^2 = \frac{G M_v}{R_v}, \quad \frac{M_v}{(4\pi/3) R_v^3} \Delta \rho_0,
\]

(4)

which lead to the approximate virial relations

\[
V_{200} \simeq M_{12}^{1/3} (1 + z)^{1/2}, \quad R_{100} \simeq M_{12}^{1/2} (1 + z)^{-1},
\]

(5)

where \( M_{12} = M_v/10^{12} M_\odot, \) \( V_{200} = V_v/200 \text{ km s}^{-1} \), \( R_{100} = R_v/100 \text{ kpc} \), and \( (1 + z)^3 \simeq (1 + z)/3 \). As can be verified based on eq. \( \boxwedge \) in \( \boxwedge \), eq. \( \boxwedge \) is accurate to a few percent at \( z > 1 \).

Based on eq. \( \boxwedge \) and eq. \( \boxwedge \), the halo crossing time \( t_v \) scales with the cosmological time \( t \),

\[
t_v = \frac{R_v}{V_v} \simeq 0.14 t.
\]

(6)

3.2 Toy Model: Accretion

We use the general term “accretion” to refer to the total inflow into a distinct halo or its central galaxy, including all the mass in dark matter, gas and stars and not distinguishing between smooth and clumpy components, i.e., including all mergers.

As reported in the Appendix, especially \( \boxwedge \), the average specific accretion rate of mass into haloes of mass \( M \) at \( z \) can be approximated by an expression of the form

\[
\frac{\dot{M}}{M} \simeq s M_{12}^\beta (1 + z)^\alpha.
\]

(7)

In the EdS regime \( \mu \rightarrow 5/2 \). With \( \beta = 0.14 \) and the appropriate value of the normalization factor \( s \), this approximation for the average of \( M(z) \) was found to be a good match to cosmological N-body simulations to better than 5\% for \( z > 1 \) \( \text{[Neistein & Dekel 2008b]} \), while it becomes an underestimate of \( \sim 20\% \) at \( z = 0 \). Similar fitting formulae with slight variations in the values of the parameters, valid in different ranges of mass and redshift, were proposed by others \( \text{[Fakhouri & Ma 2008]} \), \( \text{[Genel et al. 2008, 2010]} \).

The power of \( \mu = 5/2 \) can be simply understood from the following scaling argument based on the Press-Schechter (PS) and Extended-PS (EPS) approximations of gravitational structure formation in cosmology \( \text{[Press & Schechter 1974, Bond et al. 1991]} \). A key element in the PS formalism is a self-invariant time variable, \( \omega \propto D(a)^{-1} \), where \( D(a) \) is the growth rate of linear density perturbations (eq. \( \boxwedge \)). The self-invariance means that the growth rate of halo mass with respect to \( \omega \) is independent of \( \omega \), namely \( dM/d\omega = \text{const.} \), which implies \( M \propto \omega \). In the EdS regime, where \( D(a) \propto a \) and \( a \propto t^{3/2} \), this gives \( M \propto a^{-5/2} \). We therefore use \( \mu = 5/2 \) for our toy model at \( z > 1 \).

At lower redshifts, a better fit can be obtained with \( \mu \approx 2.4 \rightarrow 2.2 \) \( \text{[Neistein, van den Bosch & Dekel 2006, Neistein & Dekel 2008b]} \).

The small power \( \beta \) reflects the log-slope of the fluctuation power spectrum. The value \( \beta \approx 0.14 \) fits well the Millennium simulation merger trees for \( M_v \) in the range \( 10^{11} \rightarrow 10^{14} M_\odot \) \( \text{[Neistein & Dekel 2008b]} \). According to EPS, it should be \( \beta = (n + 3)/6 \), where \( n = -2 \) is the power index, \( P(k) \propto k^n \), on the corresponding scales. This implies that \( \beta \) should be even smaller for smaller haloes. Since \( \beta \) is rather small, we approximate \( \beta = 0 \) in the toy model addressed in this paper.

The normalization factor \( s \) is the specific accretion rate into a halo of \( M_v = 10^{12} M_\odot \) at \( z = 0 \), or the inverse of the corresponding accretion timescale, \( s = \tau_{\text{in},0}^{-1} \) \( \text{[}\boxwedge\text{]} \). As described in \( \boxwedge \) the value of \( s \) for haloes of \( M_v = 10^{12} M_\odot \) has been estimated earlier to be \( s \sim 0.030 \text{ Gyr}^{-1} \), by a fit to the systematic mass growth \( M(a) \) in the Millennium cosmological simulation, scaled to the \( \Lambda \)CDM cosmological parameters assumed here \( \text{[Neistein & Dekel 2008b]} \). In particular, the quoted value of \( s \) is \( \sigma_s = 0.82 \), and it scales as \( s \propto \sigma_s^{-1} \).

According to eq. \( \boxwedge \), with \( s \approx 0.030 \text{ Gyr}^{-1} \), a halo of \( 10^{12} M_\odot \) at \( z = 2 \) accretes baryons at a rate \( \dot{M}_\text{bac} \simeq 80 f_{b0,17} M_\odot \text{ yr}^{-1} \), where \( f_{b0,17} \) is the universal baryon fraction in units of 0.17.

In a simple toy model that we address here, valid for massive galaxies at \( z > 1 \), we ignore the weak mass dependence in the average specific accretion rate, \( \beta \rightarrow 0 \), and simplify eq. \( \boxwedge \) to

\[
\frac{\dot{M}}{M} \simeq s (1 + z)^{5/2}, \quad s \approx 0.030 \text{ Gyr}^{-1}.
\]

(8)
Galaxy-Formation Toy Models

Figure 2. Cosmological accretion of total mass: specific mass inflow rate $\dot{M}/M$ through the virial radius as a function of time (expansion factor $a$ and redshift $z$). Shown at each time step are the median (thick black), the average (magenta), and the average of the log within the 90% percentiles (green) over the sample of simulated galaxies. The error bars estimate the error of the mean, and the shaded area marks the 68% percentiles. Shown in comparison is the toy-model prediction, eq. (8) with $s = 0.030$ Gyr$^{-1}$ (thick smooth red). The toy model provides a reasonable fit over the whole redshift range $z = 4 - 1$.

Figure 3. Cosmological accretion of total mass: growth of virial mass. Shown are the median (thick black), the average (magenta), and the average of the log within the 90% percentiles (green) over the simulated galaxies. The mass of each galaxy $M(z)$ has been scaled before stacking by median $[M(z = 2)]/M(z = 2)$ (see text). Shown in comparison is the toy model prediction, eq. (9) with $\alpha = 0.79$ and $s = 0.030$ Gyr$^{-1}$ (dashed red). The toy model is a reasonable approximation over the whole redshift range.

Figure 4. Cosmological accretion of total mass: mass inflow rate $\dot{M}$. Shown are the median (thick black), the average (magenta), and the average of the log within the 90% percentiles (green) over the simulated galaxies. $\dot{M}(z)$ of each galaxy has been scaled before stacking by median $[M(z = 2)]/M(z = 2)$ (see text). Shown in comparison is the toy model prediction, eq. (10) with $\alpha = 0.79$ and $s = 0.030$ Gyr$^{-1}$, normalized like the simulations at $z = 2$ (dashed red). The model reproduces the stacked simulation results to $0.1 - 0.2$ dex.

This can be simply integrated to a growth of halo mass as a function of $z$, where the mass at some fiducial redshift $z_0$ is given to be $M_0$,

$$M(z) = M_0 e^{-\alpha(z-z_0)}, \quad \alpha = (3/2) s t_1 \sim 0.79. \quad \text{(9)}$$

This functional form was indeed found to be a good fit to the halo growth in earlier cosmological N-body simulations of lower resolution (Wechsler et al. 2002; Neistein & Dekel 2008b).

Note from eq. (8) and eq. (9) that the accretion rate into a given halo as it grows is

$$\dot{M}(z) \simeq s M_0 e^{-\alpha(z-z_0)} (1 + z)^{5/2}. \quad \text{(10)}$$

This average accretion rate into a given halo as it evolves does not vary much in time over an extended cosmological period. For $\alpha = 0.79$, this rate has a maximum at $z = 2.5/\alpha - 1 \simeq 2.2$, and it varies by less than a factor of 2 in the range $z \sim 0.3 - 5$. The maximum average baryon accretion rate in the history of a halo of mass $M_0 = 2 \times 10^{12} M_\odot$ today, similar to the Milky Way, is $\dot{M}_{\text{max}} \simeq 33 f_{b0.17} M_0 \text{ yr}^{-1}$.

While the expressions derived above are for the total accretion rate dominated by dark matter, we suspect that the same expressions for the specific accretion rates could be valid for the specific accretion rate of baryons into the virial radius. This should be true when the baryons follow the total-mass inflow with a constant baryonic fraction, and as long as we tentatively ignore the baryonic mass loss from the haloes. Note that strong outflows can in principle make the net baryonic accretion rate smaller than the total accretion rate (e.g. Faucher-Giguère, Keres & Mäk. 2011).
3.3 Simulations: Accretion

3.3.1 Total accretion

In Figs. 2 to 4, we address the total mass inflow rate, dominated by the dark matter component and including the baryons, gas plus stars. The accretion rate $M$ through a spherical boundary of radius $R$ during a timestep $\Delta t$ is taken to be the difference between the masses encompassed by the sphere of radius $R$ at the two snapshots defining the beginning and the end of the timestep, divided by $\Delta t$. In most cases $R$ is either $R_c$ or $0.1 R_c$, and in each of the snapshots we use the actual value of $R_c$ at that time. We verify in Figs. 2, 3, and 4 that using a fixed radius does not make a significant difference at $z > 1$. An alternative way to compute the instantaneous $M$ in a given snapshot would have been to sum over the cells or particles of mass $m_i$ and radial velocity $v_{r,i}$ within a shell of thickness $D$ about $R$, namely $\dot{M} = \sum m_i v_{r,i} / D$. As a sanity check, we have verified that when averaged over a timestep this gives similar results to the former method.

We highlight in each figure the median over the galaxies within timesteps of $\Delta \alpha = 0.02$, and also show the linear average, as well as the average of the log within the 90 percentiles. Because of the lognormal nature of the distribution, the average of the log tends to be similar to the median, while the linear average tends to be larger. These stacked quantities are compared to the corresponding toy-model prediction.

Figure 2 shows the specific total mass inflow rate $\dot{M}/M$ through the virial radius. The simulation results are compared to the toy-model prediction, eq. (8), with $s = 0.030$ Gyr$^{-1}$. There is no other free parameter in the toy model for $\dot{M}/M$. This comparison directly tests the predicted asymptotic systematic time dependence, $\mu = 5/2$, with the approximation that the specific inflow rate is independent of mass within the relatively narrow mass range spanned by the current sample. This mass independence allows a straightforward stacking of the galaxies of different masses without any scaling. This makes the specific accretion rate the most robust quantity for the comparison of the model with the simulations, and for determining the best-fit normalization, the parameter $s$. We note, however, that a success in predicting the systematic evolution of $\dot{M}/M$ does not guarantee an accurate match to the evolution of $\dot{M}$ and of $M$, which could be both off by the same multiplicative factor. Another mismatch may occur because the average of $\dot{M}/M$ is not necessarily the ratio of the averages of $\dot{M}$ and $M$.

One can see in Fig. 2 that, in general, the model provides a good fit to the systematic redshift dependence in the simulations, and that $s = 0.030$ Gyr$^{-1}$ gives a proper normalization. The model matches the median (and log average) of the simulations well in the redshift range $z = 4 - 2$, where our sample is complete and where we focus our analysis. The model is a slight overestimate of the median in the range $z = 1.8 - 1.3$, with deviations at the $2\sigma$ level. As expected, the linear average is somewhat higher than the median at every time, and it is approximately matched by the model throughout the whole redshift range $z = 4 - 1$, though the average is slightly above the model at $z = 2.6 - 2.0$.

Figure 3 shows the corresponding growth of total virial mass in time. In order to stack all the galaxies, the mass of each galaxy at $z$ has been scaled by the inverse of the ratio of its mass at $z = 2$ to the median mass at $z = 2$, namely $M(z)$ is multiplied by median[$M(z = 2) / M(z = 2)$]. For the stacking at a given $z$, the average $\langle M(z = 2) \rangle$ is computed for the sample of galaxies that is averaged over at $z$, namely the whole sample of 27 galaxies at $z \geq 2$, and the gradually diminishing subsample from $z = 2$ to 1. The simulation results are compared to the toy model prediction, eq. (9), with $\alpha = 0.79$, corresponding to $s = 0.030$ Gyr$^{-1}$. The model is normalized at $z_0 = 2$ to match the simulations’ median, $M_0 = \text{median}[M(z = 2)]$, as defined above. This way of stacking yields by construction no scatter and perfect fit of the model at $z = 2$; the comparison thus tests the success of the model at redshifts away from $z = 2$. We learn that this model is a reasonable approximation for the total mass growth over the redshift range studied. The model overestimates the median by $\sim 0.2$ dex at $z = 4 - 3$, but it is less than $0.1$ dex above the linear average. The figure demonstrates that a higher value of $\alpha$ (and $s$) can fit better the average or the median in the range $z = 4 - 2$, where the simulation results are more reliable.

Figure 4 shows the corresponding absolute total mass inflow rate $\dot{M}$ at $R_c$. Motivated by eq. (10) where $\dot{M} \propto M$, the value of $M$ for each galaxy at every $z$ has also been scaled before stacking by median[$M(z = 2) / M(z = 2)$] (similar to the scaling applied to $M(z)$). The simulation results are compared to the toy model prediction, eq. (11), with $s = 0.03$ Gyr$^{-1}$ and $\alpha = 0.79$. The toy model is again normalized at $z_0 = 2$ to $M_0 = \text{median}[M(z = 2)]$. Indeed, the value of $M$ does not vary much in time during the evolution of a halo within this redshift range, because the growth of halo mass compensates for the cosmological decline of specific inflow rate at a given mass. The gradual diminishing of the sample between $z = 2$ and $z = 1$ is reflected in the shape of the model curve in that region. The model reproduces the general trend with time seen in the stacked simulations. It overestimates the median by $\sim 0.1 - 0.2$ dex at $z = 4 - 2.7$, and even by more at $z \sim 1.4 - 1.3$, but it matches well the linear average over the whole range (except at $z < 1.3$ where the sample is very small).

Our results for the total accretion at $R_c$ confirm that the toy model in its simplest form provides a very useful approximation for the characteristic values of the three quantities, $\dot{M}/M$, $M$, and $\dot{M}$, over the whole redshift range. We see that the simplest toy model enjoys different levels of success for the different quantities at different redshifts. Somewhat different values of $s$ and $\alpha$ may provide the best fit to the median and to the average, to each of the three quantities, and at differ-
Galaxy Formation Toy Models

Figure 5. Baryon accretion: Distribution of specific accretion rate of baryons, $\dot{M}/M$, relative to the model prediction, as measured at $R_v$ (left linear, middle log) and at $0.1R_v$ (right). All snapshots of all galaxies are included, averaged within time steps of $\Delta t = 0.02$. The distributions are close to lognormal, as can be seen from the lognormal functional fits, with the associated log10 mean ($\mu_{10}$) and standard deviation ($\sigma_{10}$) quoted. The dashed Gaussian in the right panel is the fit at $R_v$ from the middle panel. The scatter represents variations among the galaxies and along the history of each individual galaxy (see Fig. 7). Marked by vertical bars are the median and linear average. The average of the log within the 90% percentiles almost coincides with the median. The gray and yellow shaded areas denote the 68 and 90 percentiles about the median. The negligible tail of negative values (excluded from the log plot) corresponds to two rare cases. The few percent tail of snapshots with high accretion rates corresponds to mergers.

Figure 6. Baryon accretion: specific inflow rate and penetration. Shown are the median (thick solid) and the average (thin dashed) of $\dot{M}/M$ through $R_v$ (black) and through $0.1R_v$ (green). At $R_v$ it is analogous to the total rate in Fig. 2. The model prediction for $R_v$, eq. (8) with $s = 0.030 \text{ Gyr}^{-1}$ (smooth solid red), is indeed a good approximation at $R_v$, but it also provides a good approximation for the specific accretion rate at $0.1R_v$.

Figure 7. Baryon accretion: specific inflow rate at $R_v$ for four individual galaxies (that enter the average in Fig. 6). The black curve second from bottom is normalized properly, and the other curves are shifted by 2 dex relative to each other. The simulations are compared to the toy model prediction, eq. (8) with $s = 0.030 \text{ Gyr}^{-1}$ (dotted). This figure illustrates the scatter due to variations among the galaxies and due to variations along the history of each individual galaxy.

3.3.2 Accretion of baryons

The baryonic inflow through the virial radius in the simulations consists of stars and gas in comparable fractions, with the stellar fraction increasing with time. Most of the stars are in small merging galaxies, associated with mini-minor mergers of mass ratio smaller than 1:10 (Dekel et al. 2009). These mergers do not have a substantial global dynamical effect on the disc — they tend to join the disc as part of the cold streams and grow the cold disc component that develops the vio-
Baryon accretion: growth of mass within $R_v$, analogous to the total mass in Fig. 3. The toy model with the adopted value of $\alpha$ eq. 8 is a reasonable approximation over the whole redshift range. Also shown is the toy model with $\alpha = 1.1$ (dashed red), which fits the median in the range $z = 4 - 2$.

Figure 8. Baryon accretion: growth of mass within $R_v$, analogous to the total mass in Fig. 3. The toy model with the adopted value of $\alpha$ eq. 8 is a reasonable approximation over the whole redshift range. Also shown is the toy model with $\alpha = 1.1$ (dashed red), which fits the median in the range $z = 4 - 2$.

A comparison of Figs. 6 and 8 to Figs. 2 and 3 shows that, similar to the case of total accretion, the toy model of eq. 8 provides a good fit to the stacked simulation results at $R_v$. With $s = 0.03$ Gyr$^{-1}$, it is an excellent match to the median at $z = 4 - 2$. The model is a slight overestimate of the median at $z = 1.8 - 1.0$, at the level of $0.1 - 0.2$ dex, or 2-sigma in terms of the error of the mean in a few points. With this normalization, the average is slightly higher than the model at $z = 4 - 2$, and is matched well at $z = 1.8 - 1.2$.

Figure 7 shows the specific baryon accretion rate at $R_v$ for four individual galaxies. It shows separately the scatter due to variations among the galaxies and the scatter due to variations along the history of each individual galaxy. Although some galaxies seem to deviate systematically from the model throughout their histories (e.g. the second from bottom case tending to be lower than the model at most times), most galaxies fluctuate about the model prediction. The typical scatter along the histories of individual galaxies is comparable to the variations from galaxy to galaxy.

Figure 6 addresses the baryon mass growth within $R_v$. In analogy to Fig. 3 the galaxies are scaled before stacking to the same value of the median at $z = 2$. The model overestimates the median by 0.2 dex at $z \sim 4$, but it overestimates the average only by 0.1 dex at that redshift. A higher value of $\alpha$ (and $s$) matches the growth curve better in the range $z = 4 - 2$.

**4 PENETRATION TO THE INNER HALO**

**4.1 Toy model: Penetration**

We expect, based on the reasons outlined below, that at $z > 1$ a very large fraction of the baryons that enter the halo at the virial radius efficiently penetrate into the galaxy at the halo centre. Below a critical halo mass of $\sim 10^{12} M_{\odot}$, a stable virial shock encompassing an extended hot gas medium cannot be supported because the radiative cooling rate is too high [Birnboim & Dekel 2003; Keres et al. 2003; Dekel & Birnboim 2006], so the instreaming of cold gas is unperturbed. A stable virial shock is likely to develop in more massive haloes, and it may shut down the cold accretion in sufficiently massive haloes at low redshifts, but at $z > 1$, most of the accreted gas is expected to be in cold streams that penetrate through the hot halo medium deep into the inner halo [Ocvirk, Pichon & Teyssier 2008; Dekel et al. 2009; Keres et al. 2009], as illustrated in Fig. 8 (based...
Galaxy-Formation Toy Models

Figure 9. Predicted penetration of cold gas streams into the halo centre as a function of halo mass and redshift. This schematic diagram (reproduced from Dekel & Birnboim 2006, Dekel et al. 2009) is based on analytic spherically symmetric calculations (Dekel & Birnboim 2006, Dekel et al. 2009). While a virial shock is expected to be present at and above \( M_v \sim 10^{12} M_\odot \), at high redshift the gas is expected to penetrate through it in cold streams along the filaments of the cosmic web.

on Dekel & Birnboim (2006). At high redshift, in massive galaxies that represent high-sigma peaks in the density fluctuation field, the gas density in these streams is high compared to the mean gas density in the halo, as they follow the narrow dark-matter filaments of the cosmic web. The high density enhances the radiative cooling rate in the streams and prevents the development of pressure that could support a stable virial shock in the streams. As a result, the streams penetrate through the halo with a mass inflow rate that is rather constant with radius (Dekel et al. 2009). They flow in at a constant inflow velocity that is comparable to the virial velocity, and they keep cold at \( T > 10^4 \) K while dissipating the gained gravitational energy into cooling Lyman-alpha radiation (Dijkstra & Loeb 2009, Goerdt et al. 2010, Fumagalli et al. 2011).

4.2 Simulations: Penetration

Figure 6, which has been discussed above concerning the specific accretion rate of baryons at \( R_v \), also shows the specific accretion rate of baryons into the inner halo at \( 0.1 R_v \). The stacked simulation results are compared to the model prediction for the accretion rate at \( R_v \), eq. (8). The right panel of Fig. 5 shows the distribution of specific accretion rate of baryons at \( 0.1 R_v \), over all the snapshots of all the galaxies, normalized to the model prediction. The distribution at \( 0.1 R_v \)

should be compared to that shown in the left panels for \( R_v \), and to the Gaussian fit that characterizes it. We learn that the specific baryon accretion rate at \( R_v \) and at \( 0.1 R_v \) are rather similar, and well approximated by the toy model. In fact, the model is a good match to the median of the simulations at \( 0.1 R_v \) over the whole redshift range.
The similarity between the specific accretion rate at 0.1RV and at RV does not necessarily imply 100% penetration through the halo, as both M and M for the baryons could decrease between RV and 0.1RV in a similar way. We conclude that eq. (8), with s = 0.033 Gyr−1, is also a useful approximation for the specific accretion rate of baryons into the galaxy itself. This is for the distribution among galaxies of halo masses that range from below 10¹¹M⊙ to above 10¹²M⊙ in the redshift range z = 4 − 1.

Figure 10 shows the absolute inflow rate ˙M for the baryons at RV and at 0.1RV, and the corresponding toy model prediction at RV, eq. (10). The scaling to the average at z = 2 before stacking the simulated galaxies and the corresponding normalization of the toy model is the same as described in § 3.3.1 We first note in Fig. 10 that the simplified toy model, eq. (10), is a reasonable fit to the inflow rate of the simulations at RV in the range z = 4 − 1. The model overestimates the median by 0.2 dex or less, and it fits the average to within 0.1 dex (except at z < 1.3 where the sample is very small). Indeed, the value of M does not vary much in time during the evolution of a galaxy within this redshift range, because the mass growth compensates for the decline of specific inflow rate, as for Fig. 4 This fit is complementary to the good fits provided by the model to the specific inflow rate and to the mass growth, Figs. 6 and 8.

Figure 10 provides a new insight into the penetration of baryons through the halo into the inner galaxy. One reads that, in the range z = 4 − 2, the baryonic inflow rate at 0.1RV is typically ∼ 50% of the baryonic inflow rate at RV. In the range z = 2 − 1 the penetration seems to be more efficient, though the uncertainty is larger because the sample is smaller. This massive penetration is in general agreement with earlier estimates based on the MareNostrum simulation with lower resolution (Dekel et al. 2009; Danovich et al. 2012) and with the high SFR observed in massive galaxies at z ∼ 2 (Genzel et al. 2008; 2011). If the SFR follows the gas accretion rate into the disc (see § 6), the SFR can in principle be a fixed significant fraction of the instantaneous overall baryon accretion rate into the virial radius. Of course, internal processes within the disc may suppress the SFR at early times (z > 3, say) and less massive haloes (M_v < 10¹¹M⊙, say) (e.g. Krumholz & Dekel 2012), and thus accumulate gas for higher SFR in the more massive galaxies at later times (e.g., z ∼ 2).

One should recall the caveat associated with the significant fraction of stellar mass in the inflow through the halo into the central galaxy. This may be partly responsible for the more efficient penetration seen at z < 1.8, and it reminds us not to assume accuracy of better than a factor of two when addressing the incoming gas and the resultant star formation rate.

The radii of spheres through which the inflow rate has been considered so far were the virial radius for the outer halo, and a fixed fraction of it (0.1RV) for the galaxy at its centre. This is in order to allow us to consider self-similar radii that follow the halo and the galaxy as they grow. However, the growth of RV in time (which can be deduced from the virial relations, eq. (5), and the toy model for halo mass growth, eq. (11)) is responsible for part of the mass growth within RV, which may not be associated with actual inflow (Diemer, More & Kravtsov 2013). In order to quantify this effect, Fig. 11 shows the inflow rate through spheres of fixed radii, 100 kpc and 10 kpc, replacing RV and 0.1RV of Fig. 10. A comparison between the two figures indicates no significant differences between the absolute values of ˙M at z = 4 − 2, and similar conclusions regarding the penetration efficiency. Similar results are obtained for the specific accretion rate. We conclude that, at least at z ≥ 2, the vast majority of the inflow measured through RV or a fixed fraction of it is true inflow, and therefore proceed with the analysis using the radii that grow self-similarly in time, RV and 0.1RV.

5 DISC SIZE
5.1 Toy Model: Disc Size versus Virial Radius

We assume that the characteristic disc radius scales with the halo virial radius via a contraction factor λ, RV = λRV. (11)

We also assume that the characteristic circular velocity of the disc scales with the virial velocity, V_d = vVR. (12)

Given that the halo is roughly an isothermal sphere, the value of v is expected to be of order unity, and we sometimes adopt v ≃ √2, which is a good approximation for the Milky Way. The value of λ can be estimated in the common model where the galactic disc is assumed to form by dissipative gas contraction within the dark-matter halo (Fall & Efstathiou 1980; Mo, Mao & White 1998; Bullock et al. 2001). If one assumes that the original specific angular momentum of the gas, j, is similar to that of the dark matter in the virialized halo, and if j is conserved during the gas contraction, then

λ ≃ √2v−1′λ′, (13)

where λ′ ≡ j/(√2RV v) is the halo spin parameter as defined by Bullock et al. (2001). Based on N-body simulations and tidal-torque theory, the halo spin parameter is assumed to have a constant average value independent of mass or time, λ′ ∼ 0.035 (e.g. Bullock et al. 2001), implying that λ is constant and of a similar value.

The disc dynamical time is then
t_d = R_d/V_d ≃ v−1′λRV ≃ 0.0071v−1λ0.05 t, (14)

where λ0.05 ≃ λ/0.05, and where we have used the relation of the virial crossing time to the cosmological time t, eq. (9).

Note that the distribution of spin parameter among haloes is lognormal with a standard deviation of half a dex (Bullock et al. 2001). This would translate to a large scatter about the average R_d of eq. (11).
Using eq. (11) and eq. (5), one can derive a useful expression for the baryonic surface density in the disc, \[ \Sigma = \frac{M_d}{\pi R_d^2} \]
\[ \simeq 1.3 \times 10^9 M_\odot \text{kpc}^{-2} m_{d,0.1} \tilde{\lambda}_{0.05}^2 M_{1/2}^{1/3} (1 + z)^2, \] (15)
where \( M_d \) is the cold mass in the disc, and \( m_{d,0.1} \equiv m_d/0.1 \) refers to the baryonic mass in the disc relative to \( M_{\odot} \).

5.2 Simulations: Disc Size versus Virial Radius

The disc radius \( R_d \) is determined in the simulations as described in [Mandelker et al. 2013]. In short, the disc is modeled as a cylinder whose \( z \) axis coincides with the spin axis of the cold gas within it (\( T < 1.5 \times 10^4 \text{K} \), which is typically 97% of the gas). We start with a large cylinder of radius 0.15\( R_v \) and half-height 1 kpc, and iteratively converge on the cylinder of radius \( R_d \) that contains 85% of the gas mass within the large cylinder. This determination of the disc radius is commonly in good qualitative agreement with the radius one would have estimated by visual inspection of the face-on surface density of the cold gas. Note, however, that this is only a crude approximation for the radius of the whole disc including its stellar component, which could be dominant.

The disc radii as determined from the simulations at \( z > 3 \) are rather uncertain and sometimes ill-defined. This is because the discs in these early times tend to be small and highly perturbed as the timescale for morphologically damaging mergers is comparable to and shorter than the disc dynamical time. We therefore prefer to de-emphasize the results that explicitly depend on the disc radius at \( z > 3 \), e.g., in Fig. 16 and Fig. 17 below.

Figure 12 shows the average and 68% scatter of the ratio \( R_d/R_v \) over the simulated galaxies. In eq. (11), a value of \( \tilde{\lambda} \simeq 0.05 \) crudely fits the 1-sigma range of values from the simulations over the whole redshift range. However, there seems to be a systematic trend with redshift, from a median of \( \sim 0.06 \) at \( z = 3 - 2 \), through 0.05 at \( z = 2 - 1.5 \) down to 0.04 at \( z = 1 \). The indicated values of \( \tilde{\lambda} \) are somewhat larger than what is implied from the spin parameter of dark matter haloes as estimated from cosmological N-body simulations, especially if \( v \) is larger than unity. This is consistent with observational indications at \( z \sim 2 \) (Genzel et al. 2006).

Indeed, the specific angular momentum of the baryons in the high-redshift discs is likely to be higher than that of the dark-matter haloes (a) because of the way angular momentum is transported into the disc by cold streams from the cosmic web (Kimm et al. 2011; Pichon et al. 2011; Danovich et al. 2012; Stewart et al. 2013), and (b) because outflows tend to preferentially carry away low-angular-momentum material (Maller & Dekel 2002; Brook et al. 2011). We take the toy model for the disc radius and the associated estimate of dynamical time, eq. (11) and eq. (14), to be accurate to within a factor of two.

Using abundance matching to determine halo virial radii, [Kravtsov 2013] finds that the half-mass radii of today’s galaxies of all morphological types, \( R_{\text{half}} \), are related to their today’s halo radius \( R_{200} \) as \( R_{\text{half}} \simeq 0.015 R_{200} \). If the galaxy has formed roughly in its present size at \( z \sim 2 \), when the halo was a few times smaller in radius, the implied ratio then is consistent with our finding for \( \tilde{\lambda} \) in Fig. 12.

6 MASS CONSERVATION: STEADY-STATE OF SFR AND GAS MASS

6.1 Toy Model: Steady State

Assuming mass conservation, the gas mass in the galaxy varies subject to a source term and a sink term (analogous to a “bathtub” [Bouché et al. 2010], e.g.,
\[ \dot{M}_g = \dot{M}_{g,\text{ac}} - \dot{M}_{\text{sf}}, \] (16)
where \( \dot{M}_{g,\text{ac}} \) is the accretion rate of gas, and \( \dot{M}_{\text{sf}} \) is the star formation rate. We temporarily ignore potential gas outflow from the galaxy (to be incorporated later). We assume that the SFR density obeys a universal local volumetric law, \( \dot{\rho}_g = \epsilon_{\text{sf}} \dot{M}_g/\tau_{\text{sf}} \) with \( \epsilon_{\text{sf}} \sim 0.02 \) a constant efficiency and \( \dot{\rho}_g \) the density of molecular gas in the star-forming region (Krumholz, Dekel & McKee 2012).

Then, under certain circumstances, the overall SFR in the galaxy can be crudely assumed to be proportional to the total gas mass,
\[ \dot{M}_{\text{sf}} = \frac{\dot{M}_g}{\tau_{\text{sf}}}, \] (17)
where \( \tau_{\text{sf}} = t_{\text{ff}}/\epsilon_{\text{sf}} \).
If $M_{g,ac}$ and $\tau_{sf}^{-1}$ vary on a timescale longer than $\tau_{sf}$, the simple solution of eq. (16) with eq. (17) is

$$\dot{M}_g = \dot{M}_{g,ac} e^{-t/\tau_{sf}}, \quad M_g = M_{g,ac} \tau_{sf} (1 - e^{-t/\tau_{sf}}).$$  \hspace{1cm} (18)

After a transition period of order $\tau_{sf}$, the solution relaxes to a steady state solution

$$\dot{M}_g \simeq 0, \quad M_g \simeq M_{g,ac} \tau_{sf}, \quad M_{sf} \simeq M_{g,ac}. \hspace{1cm} (19)$$

The SFR sink term adjusts itself to match the external source term $\dot{M}_{g,ac}$.

The value of $\tau_{sf}$ can be related to the global disc crossing time $t_d$ by $t_{lf} = f_{sl} t_d$, with $f_{sl} \sim 0.5$, assuming that star formation occurs in regions (partly bound clumps) where the overdensity is a few with respect to the mean density in the disc. Using the estimate for $t_d$ from eq. (14), we obtain, for $f_{sl} \sim 0.5$ and $\epsilon_{sf} \sim 0.02$,

$$\tau_{sf} = \epsilon_{sf} f_{sl} t_d \sim 0.17 t. \hspace{1cm} (20)$$

At $z = 2$, where the Hubble time is $t \sim 3.25$ Gyr, we have $\tau_{sf} \sim 500 - 600$ Gyr.

The validity of the solution eq. (18) to eq. (16) depends on the timescales for variation of $M_{g,ac}$ and of $\tau_{sf}^{-1}$ compared to $\tau_{sf}$. Using eq. (8) and eq. (1) we obtain

$$t_{var}(M_{g,ac}) = \frac{M_{g,ac}}{|dM_{g,ac}/dt|} \sim \frac{t}{|(2/3) \alpha a^{-1} - 5/3|}. \hspace{1cm} (21)$$

With $\alpha = 0.79$ in eq. (9), we get $t_{var}(M_{g,ac})/t \simeq 1.0, 2.3, 12, 1.6$ for $z = 4, 3, 2, 1$ respectively. This means that $t_{var}(M_{g,ac}) \gg t$ throughout the redshift range $z = 4 - 1$, so, based on eq. (20), $t_{var}(M_{g,ac}) \gg \tau_{sf}$, as required. Similarly, using eq. (21), the timescale for variation of $\tau_{sf}^{-1}$ is

$$t_{var}(\tau_{sf}^{-1}) = \frac{\tau_{sf}^{-1}}{|d\tau_{sf}^{-1}/dt|} = t, \hspace{1cm} (22)$$

namely $|t_{var}(\tau_{sf}^{-1})| \gg \tau_{sf}$, as required. We thus expect the steady-state solution to be valid through the whole period of interest here, and to be approximately valid out to $z \sim 10$ and earlier.

A continuity equation of a similar nature, and its steady-state solution, were used with a variety of sink terms in several recent studies (Dekel, Sari & Ceverino 2009, Boucle et al. 2010, Krumholz & Dekel 2012, Cacciato, Dekel & Genel 2012, Genel, Dekel & Cacciato 2012, Feldmann 2010).

The sink term in eq. (16) can be generalized to include other sink terms. For example, if there is an outflow at a rate $\dot{M}_{out} = f_{out} \dot{M}_g$, the second term of eq. (16) is simply multiplied by the factor $(1 + f_{out})$. Similarly, if eq. (16) is applied only to the disc component, instability-driven mass inflow within the disc adds another sink term of a similar form, see [24]. In these cases, $\tau_{sf}$ in eq. (16) is replaced by a smaller timescale $\tau$, which makes the steady-state solution an even better approximation.

---

**Figure 13.** Overall steady state: Shown are SFR and $M_{gas}$ within the sphere of radius $0.1 R_c$, as well as the gas accretion rate $M_{gas,acc}$ through that radius, which is the sum of the former two. Shown are the medians (solid thick) and the average (dashed thin) over the simulated galaxies. The 68 percentiles (shaded area) and the errors of the mean (error bars) are shown for $M_{gas}$ and SFR. The simulations evolve about the steady-state solution, eq. (19), where SFR= $M_{gas,acc}$ and $M_{gas} = 0$, with a slow increase in the gas mass at the level of ~10% of the SFR and the accretion rate.

---

**6.2 Simulations: Steady State**

Figures 13 to 15 verify the validity of the steady state solution, eq. (19), in the simulations. Figure 13 shows the SFR and the rate of change of gas mass, $M_{gas}$, within the inner sphere of radius $0.1 R_c$. As in §3.3.1, these quantities were scaled before stacking to match the median mass at $z = 2$. The gas accretion rate through that radius, $M_{gas,acc}$, is the sum of the two. We see that the median and average of $M_{gas}$ is indeed small, at the level of ~10% of the SFR or the accretion rate. This corresponds to the SFR being close to the accretion rate, typically only ~10% smaller.

Figure 14 shows the ratio of SFR and $M_{gas,acc}$ at $0.1 R_c$ for four individual galaxies, compared to the steady-state prediction of unity, illustrating the variation along the history of each galaxy and among the different galaxies. As in Fig. 7 we see that the typical galaxy does not tend to show a systematic deviation of the SFR from the gas accretion rate into the disc along its history. The scatter along the history of each individual galaxy is similar in the different galaxies and comparable to the variations between different galaxies.

Figure 15 shows the distribution of SFR/$M_{gas,acc}$ at $0.1 R_c$ over all the snapshots. It is close to lognormal, with a log10 mean $\mu_{10} = -0.04$ and standard deviation $\sigma_{10} = 0.12$. The linear average is only 2% off unity. We conclude that the SFR closely follows the gas accretion rate into the $0.1 R_c$ sphere, as predicted by the steady-state solution, eq. (19). While the scatter in this ratio is only at the level of ~30%, the much larger scatter in SFR seen in Fig. 13 indicates large variations in both.
Galaxy-Formation Toy Models

15

Figure 14. Overall steady state: SFR versus gas accretion rate at 0.1\(R_d\) for four individual galaxies. The black curve second from bottom is normalized properly, and the other curves are shifted by 2 dex relative to each other. The simulations are compared to the toy-model steady-state solution, eq. (19), where SFR = \(\dot{M}_{\text{gas acc}}\) (dotted). This figure illustrates the scatter due to variations among the galaxies and due to variations along the history of each individual galaxy.

Figure 15. Overall steady state: distribution of the ratio of SFR and gas accretion rate at 0.1\(R_d\). All snapshots of all galaxies are included, averaged within time steps of \(\Delta a = 0.02\). The distribution is close to lognormal, with the associated lognormal mean \(\langle \mu_{10} \rangle\) and standard deviation \(\sigma_{10}\) quoted. The scatter represents variations among the galaxies and along the history of each individual galaxy. The radial velocity dispersion \(\sigma\) and the surface density \(\Sigma\) refer to the cold component of the disc that participates in the disc instability (see the practical definition in the simulations in §7.2 below). We may refer to it as “gas”, but it also includes the cold young stars, and one should be careful to specify whether the latter are included (see a multicomponent analysis in Cacciato, Dekel & Genel 2012). The angular velocity \(\Omega\) and circular velocity \(V\) refer to the characteristic disc radius \(R_d\), \(V = \Omega R_d\). The constant \(\sqrt{2}\) refers to a flat rotation curve (and it stands for \(\sqrt{2}(1 + \theta)\) where \(V \propto r^\theta\)). The mass fraction \(\delta\) is the ratio of mass in the cold component \(M_d\) to the total mass within \(R_d\), \(M_{\text{tot}}(R_d)\), the latter including gas, stars and dark matter,

\[
\delta = \frac{M_d}{M_{\text{tot}}(R_d)}.
\] (24)

The instability in high-redshift galaxies is driven by the high cold surface density that reflects the high gas accretion rate and the high mean cosmological density at earlier times. We term this phase Violent Disk Instability (VDI), as the associated dynamical processes occur on a galactic orbital timescales, as opposed to the “secular” processes associated with disc instabilities at low redshifts. The unstable disc tends to self-regulate itself in marginal instability with \(Q \approx 1\), where \(\sigma\) represents supersonic turbulence that provides the pressure while thermal pressure is negligible. The turbulence tends to decay on a timescale comparable to the disc dynamical time (Mac Low 1999), so it should be continuously powered by an energy source that could stir up turbulence and maintain \(\sigma\) at the level required for \(Q \approx 1\). In the perturbed disc, which consists of extended transient features and massive compact clumps (Mandelker et al. 2013), gravitational torques drive angular momentum out and cause mass inflow towards the centre, partly in terms of clump migration (Noguchi 1999; Bournaud, Elmegreen & Elmegreen 2007; Dekel, Sari & Cererind 2009), and partly in terms of off-clump inflow (Gammie 2001; Dekel, Sari & Cererind 2009; Bournaud et al. 2011). This inflow down the potential gradient from the disc outskirts to its centre in turn provides the required energy for \(Q \approx 1\). (Krumholz & Burkert 2010; Cacciato, Dekel & Genel 2012; Forbes, Krumholz & Burkert 2012; Forbes et al.

7 INFLOW WITHIN THE DISC

7.1 Toy Model: Disk Inflow

Galactic discs at high redshift are expected to develop a gravitational disc instability with a Toomre instability parameter \(Q\) smaller than unity (Toomre 1964). The Toomre parameter can be expressed as (e.g. Dekel, Sari & Ceverino 2009)

\[
Q \approx \frac{\sqrt{2}\Omega\sigma}{\pi G \Sigma} \approx \frac{\sqrt{2}}{\delta V}.
\] (23)

The radial velocity dispersion \(\sigma\) and the surface density \(\Sigma\) refer to the cold component of the disc that participates in the disc instability. The Toomre parameter can be expressed as (e.g. Dekel, Sari & Ceverino 2009)

\[
Q \approx \frac{\sqrt{2}\Omega\sigma}{\pi G \Sigma} \approx \frac{\sqrt{2}}{\delta V}.
\] (23)

The instability in high-redshift galaxies is driven by the high cold surface density that reflects the high gas accretion rate and the high mean cosmological density at earlier times. We term this phase Violent Disk Instability (VDI), as the associated dynamical processes occur on a galactic orbital timescales, as opposed to the “secular” processes associated with disc instabilities at low redshifts. The unstable disc tends to self-regulate itself in marginal instability with \(Q \approx 1\), where \(\sigma\) represents supersonic turbulence that provides the pressure while thermal pressure is negligible. The turbulence tends to decay on a timescale comparable to the disc dynamical time (Mac Low 1999), so it should be continuously powered by an energy source that could stir up turbulence and maintain \(\sigma\) at the level required for \(Q \approx 1\). In the perturbed disc, which consists of extended transient features and massive compact clumps (Mandelker et al. 2013), gravitational torques drive angular momentum out and cause mass inflow towards the centre, partly in terms of clump migration (Noguchi 1999; Bournaud, Elmegreen & Elmegreen 2007; Dekel, Sari & Ceverind 2009), and partly in terms of off-clump inflow (Gammie 2001; Dekel, Sari & Ceverind 2009; Bournaud et al. 2011). This inflow down the potential gradient from the disc outskirts to its centre in turn provides the required energy for \(Q \approx 1\). (Krumholz & Burkert 2010; Cacciato, Dekel & Genel 2012; Forbes, Krumholz & Burkert 2012; Forbes et al.
The gas inflow rate $\dot{M}_{g,\text{in}}$ can be estimated by equating this energy gain and the dissipative losses of the turbulence,

$$\dot{M}_{g,\text{in}} V^2 \simeq \frac{M_g \sigma^2}{\gamma t_d}, \tag{25}$$

Here $\gamma = (2/3) \gamma_g^{-1} \gamma_{\text{dis}}$, where $\gamma_g$ is the fraction of gas in the inflowing mass, $\gamma_{\text{dis}} V^2$ is the energy gain per unit mass between the disc radius and the centre, and $\gamma_{\text{dis}} t_d$ is the turbulence decay timescale. The value of $\gamma$ is of order unity, and it could be as large as a few. We thus obtain

$$\frac{\dot{M}_{g,\text{in}}}{M_g} \simeq \frac{1}{\gamma t_d} \sigma^2 \simeq \frac{1}{2 \gamma t_d} \delta^2, \tag{26}$$

assuming $Q \sim 1$ in eq. (26). Note that $\delta$ on the right-hand side refers to the cold component, including young stars, while the left-hand side refers to the gas only.

Independent estimates based on the mechanics of driving the mass inflow by torques yield similar results for the inflow rate, to within a factor of two. Examples of such calculations are (a) an estimate of the rate of energy exchange by gravitational clump encounters (Dekel, Sari & Ceverino 2009, eqs. 7 and 21), and (b) an estimate of the angular-momentum exchange among the transient perturbations in a viscous disc (Gammie 2001; Genzel et al. 2008; Dekel, Sari & Ceverino 2009, eq. 24). In Appendix A5 we estimate that the disc evacuation rate by clump migration due to the dynamical friction exerted by the disc on the clumps is in the same ballpark as the above estimates of the inflow rate as long as $\delta \sim 0.2 - 0.3$ as in the steady-state solution for VDI discs (Dekel, Sari & Ceverino 2009) and in our current simulations, Fig. 13. One should emphasize that the inflow in the disc is a robust feature of the instability, not limited to clump migration. Inflow at a rate comparable to eq. (26) is expected even if bound clumps were to disrupt by stellar feedback in less than a migration time, as sometimes assumed (Murray, Quataert & Thompson 2010; Genel et al. 2012; Hopkins et al. 2012), though this scenario is not realistic based on theoretical and observational grounds (Krumholz & Dekel 2010; Dekel & Krumholz 2013).

We comment in parentheses that the cosmological streams that penetrate through the halo and feed the disc are also potential drivers of turbulence (Dekel, Sari & Ceverino 2003; Khochar & Silk 2005; Genel, Dekel & Cacciato 2012), but this requires strong coupling of the streams with the higher-density disc, which demands that the streams largely consist of dense clumps. Being driven by an external source, this mechanism is not naturally self-regulating.

Stellar feedback may or may not be a major direct driver of the turbulence. There is potentially enough energy and momentum in supernovae, radiative feedback and winds (Dekel, Sari & Ceverino 2009; Genel et al. 2010; Krumholz & Thompson 2012, 2013; Dekel & Krumholz 2013), but it is not clear whether they are properly deposited within the disc. Evidence against stellar feedback as a direct source
of turbulence is in the observation that the velocity dispersion is not tightly correlated with proximity to star-forming clumps [Forster Schreiber et al. 2009; Genzel et al. 2011]. However, outflows can help boosting up $\sigma$ by simply lowering the surface density of that gas that is supposed to be stirred up by a given energy source (Genel, Dekel & Cacciato 2012). Recall that the stellar feedback in the simulations used here is underestimated.

According to eq. (11), the steady-state solution implies $M_g \simeq M_{g,ac} \tau_{sf}$. Combined with eq. (26) and $\tau_{sf} = f_{sf} \epsilon_{sf}^{-1} t_d$ from eq. (24), this gives in steady state

$$\frac{\dot{M}_{g, in}}{M_{g,ac}} \simeq \frac{f_{sf}}{\epsilon_{sf}^2} \frac{\sigma^2}{V^2} \simeq \frac{f_{sf}}{2\epsilon_{sf} \gamma} \delta^2. \quad (27)$$

Note that this expression is independent of $t_d$. The right-hand side is of order unity, so the equation implies that $\dot{M}_{g, in}$ is also comparable to the gas accretion rate, we conclude for the growth of mass in the bulge during the period $\Delta t$. The ratio of rates in eq. (27) can be interpreted as an approximation for the bulge-to-total baryonic mass ratio in the galaxy.

The gas mass fraction within the disc radius, $\delta_{g}$, can be crudely estimated from the toy models as follows. The gas mass is given by the steady-state solution, eq. (11), $M_g \simeq M_{g,ac} \tau_{sf}$. Expressing $\tau_{sf} \simeq f_{sf} \epsilon_{sf}^{-1} t_d$ and $t_d \simeq 0.007 t$ from eq. (11), we obtained $\tau_{sf}/t \simeq 0.17$. We can then write $M_g \simeq (\tau_{sf}/t) f_{g,0} M_{ac,t}$, where $f_{g,0}$ is the gas fraction in the baryon accretion, $f_{b} \simeq 0.17$ is the universal baryon fraction, and $M_{ac}$ is the total accretion rate, such that $M_{ac} \simeq M_V$. The total mass within the disc radius can be crudely estimated as the sum of the baryonic mass and the dark mass, $M_{tot} \simeq (f_b + \lambda) M_V$, assuming $M(r) \propto r$ within the halo. We thus obtain for the gas fraction

$$\delta_g \simeq \frac{(\tau_{sf}/t) f_{g,0} f_b}{f_b + \lambda} \lesssim 0.1. \quad (28)$$

If only a fraction of the baryons in the halo are within the disc radius, the value of $\delta_g$ would be larger accordingly. When including the cold stars in the disc, the value of $\delta$ could be doubled or higher by a factor of a few, so we expect $\delta \sim 0.2 - 0.3$.

### 7.2 Simulations: Disk Inflow

Figure 13 and Fig. 17 test the validity of the toy model estimate for the inflow within the disc, eq. (26) and eq. (27) respectively. For each equation, we compare the left-hand side as measured in the simulations and the right-hand side as deduced from the toy model but also using quantities that are measured from the simulations.

We adopt as our fiducial value for the dissipation timescale $\gamma = 1$, and the simulations can actually be used to determine the best-fit effective value of $\gamma$.

The cold mass fraction within the disc radius, $\delta$, entering the right-hand sides of eq. (26) and eq. (27), is computed for each snapshot inside a sphere of radius $R_d$. “Cold” disk stars are defined by $j_{z}/j_{max} > 0.7$, where $j_z$ is the specific angular momentum of the star particle along the spin axis of the disc, and $j_{max} = R V_{tot}$ is the maximum specific angular momentum at the given energy, where $R$ is the radius at the particle position and $V_{tot}$ is the magnitude of the total particle velocity. Figure 18 shows the average and 68% scatter of $\delta$ in the simulations. The median is $\delta \simeq 0.2$ across the whole redshift range, with a 68% scatter between 0.12 and 0.25. The value of $\delta$ due to gas only, $\delta_{g,0}$, is about half that value. These measured values are consistent with the expectations based on eq. (28).

The dynamical time $t_d$ entering the right-hand-side of eq. (26) is computed for each snapshot from the simulations via eq. (11). Based on eq. (11), the parameter $\bar{\lambda}$ is derived from the radii as determined from the simulations, and $\nu = \sqrt{2}$ is assumed. The Hubble time $t$ is approximated by eq. (1).

For the SFR efficiency in eq. (27) we use $\epsilon_{sf} = 0.05$. This is a high value, estimated to be the effective value of $\epsilon_{sf}$ in our current simulations, dictated by the requirement that the SFR in the simulations at the given resolution roughly matches the Kennicutt relation. This effective value of $\epsilon_{sf}$ is consistent with eq. (17) for the SFR in the simulations with $t_d = 4$ Myr and where the cold gas available for star formation is defined to be the mass with temperature lower than $10^4 K$ and density higher than $1 cm^{-3}$.

For the analysis here, we have used only snapshots where there is a net inflow within the disc, as a signature of VDI. Snapshots where the disc radius is smaller than 3 kpc are excluded; this minimizes problems related to the spatial resolution of the simulations. Including these small disks did not make a significant difference at $z < 3$, but did make some difference at $z = 3 - 4$. Since the disc radii at $z > 3$ are uncertain due to wild perturbations, we de-emphasize the results in this regime even when the disc radius is larger than 3 kpc.

Figure 18 compares the specific gas inflow rate in the disc according to the two sides of eq. (26). The left-hand side is measured directly from the simulations, while in the right-hand side we make use of the $\delta$ and $t_d$ as computed from the simulations. Overall, the toy model is quite successful, at the expected level of uncertainty. The average of the left-hand-side is systematically higher than the right-hand side by a factor of order two in the range $z = 3 - 2$. This is at the level of accuracy expected from the crude toy model and the uncertainties in measuring the simulation results. The apparent discrepancy is somewhat larger at $z = 4 - 3$ where the uncertainty is larger, and somewhat smaller
at $z = 2 - 1$. A better fit could be obtained with either a smaller value of $\gamma$ or a larger value of $v$.

Figure 17 compares the ratio between the gas inflow rate within the disc and the gas accretion rate onto the disc as computed by the two sides of eq. (27). The values of the parameters used in the right hand side are $\gamma = 1$, $\epsilon_{sf} = 0.05$, and $f_{sd} = 0.5$ The agreement is good except in the redshift range $z = 3 - 4$ where the uncertainty in the disc radius is large.

8 STEADY STATE IN THE DISK

8.1 Toy Model: steady state in disc

In analogy to the SFR, $\dot{M}_{sf} = M_g / \tau_{sf}$, based on eq. (25) we can express the inflow rate within the disc as $\dot{M}_{g, in} = M_g / \tau_{in}$, where $\tau_{in} = \gamma (\sigma / V)^2 t_d$. The gas-mass conservation equation, as in eq. (18), could then be applied to the disc alone, and include draining by three processes, namely star formation, outflows, and inflow within the disc. Based on eq. (17) and eq. (25), each of the corresponding sink terms ($i = 1, 3$) could be expressed in the form $\dot{M}_{sink, i} = M_g / \tau_i$, where $\tau_i = \epsilon_{i}^{-1} t_d$, so the continuity equation for the disc is

$$\dot{M}_g = \dot{M}_{g, ac} - \dot{M}_{sink}, \quad \dot{M}_{sink} = \frac{M_g}{\tau},$$

(29)

where

$$\tau = \tau_{sf}^{-1} + \tau_{out}^{-1} + \tau_{in}^{-1},$$

(30)

and its steady state solution is analogous to eq. (19) with $\tau_{sd}$ replaced by the smaller timescale $\tau$.

If a steady-state solution exists, we can evaluate the steady-state value of $\delta$, the mass fraction of cold disc inside $R_d$, following Dekel, Sari & Ceverino (2009). We write

$$\delta = \frac{\beta M_d}{M_b}.$$

(31)

where $\beta = M_b / M_{tot}(R_d)$ is the fraction of baryons within the disc radius, estimated to be $\approx 0.6$. Assuming that $\beta$ is constant, a time derivative of eq. (31), combined with eq. (24), yields

$$\dot{\delta} = \beta (f_g - \beta^{-2} \delta) \tau_{ac}^{-1} - \delta \tau_{sink}^{-1} \dot{\delta}.$$

(32)

where $f_g$ is the mass fraction of gas in the baryonic accretion, and where $\tau_{ac} = M_b / M_b$ and $\tau_{sink} = M_d / M_{sink}$. Here $M_d$ stands for the cold component in the disc, gas and cold young stars. If a steady-state solution exists, it will be provided by setting $\dot{\delta} = 0$.

If the sink timescale scales with $\delta^{-2}$ (see below), we write $\tau_{sink} / \tau_{ac} = \delta^{-2}$, and obtain from eq. (32) that the steady-state solution is the solution of the depressed cubic polynomial,

$$\delta^3 + b \delta - c = 0,$$

(33)

where $c = b \beta f_g$. Its solution is

$$\delta = u - \frac{b}{3u}, \quad u = \left(\frac{c}{2}\right)^{1/3} \left[1 + \left(1 + \frac{4b^3}{27c^2}\right)^{1/2}\right]^{1/3}.$$

(34)

If $b \ll 1$, then $c \ll 1$, and there is no steady-state solution with $\delta \sim 1$. If $\delta \ll 1$, then there is a solution, $\delta \approx c / b = \beta f_g$, which is indeed likely to vary rather slowly with time. If all three terms of eq. (33) are comparable, with $\delta \sim 0.3$, then we need to appeal to the full solution, eq. (34). If the gas dominates the cold disc, we can write $\tau_{sink} = \tau$. Then the long-term time dependence is via $b \propto a^{-1}$. When inserted in eq. (34), we find that the value of $\delta$ is expected to vary rather slowly with time, consistent with a steady-state solution.

The scaling of $\tau_n$ with $\delta^{-2}$ is apparent from eq. (26). The dependence of $\tau_{sf}$ (and therefore $\tau_{out}$) on $\delta$ could be similar if star-formation is limited to the disc clumps. Then the SFR is proportional to the mass in clumps, namely the product of clump mass and the number of clumps, $N_c M_c$. The Toomre clump mass is $M_c / M_d \approx (1/4) \delta^2$ (Dekel, Sari & Ceverino 2009). Constant values of $N_c M_c$, $N_c$ and $\delta$ are consistent with the slow variation of $\delta$ implied by eq. (34).

8.2 Simulations: steady state in disc

Figure 18 shows that throughout the range $z = 3 - 1$, the average value of $\delta$ is consistent with being constant, as implied by the steady-state solution eq. (34). A slight decline may be marginally noticeable toward $z \sim 1$. The average value is $\delta \approx 0.2$ and the 68% range is roughly from 0.12 to 0.24. It is consistent with the crude estimate in eq. (28), once the mass in cold stars is comparable to the gas mass.

Figure 18 shows the evolution of bulge-to-total ratio. The disc stars are distinguished from bulge stars by $j_z / j_{max} > 0.7$. The average is rather constant at $B/T \approx 0.73$, with the 68% range from 0.63 to 0.87. This is consistent with the steady-state bulge growth by VDI, eq. (27), plus the constancy of $\delta$ as implied by eq. (34).
Figure 19. Steady state in disc: Bulge-to-total ratio within $R_d$. Shown are the median (thick black), the average (magenta), and the average of the log within the 90% percentiles (green) over the simulated galaxies. Shown in comparison is the toy model the average of the log within the 90% percentiles (green) over the 90% percentiles (green) of the value that would yield a good fit in Fig. 17, with $\delta \simeq 0.2$, $\epsilon_{sf} \simeq 0.04$ and $f_{sf} \simeq 0.5$. The value of $B/T$ and its constancy in time is consistent with the steady-state bulge growth by VDI, eq. (27), plus the constancy of $\delta$ as implied by eq. (53).

With $\delta \simeq 0.2$, $\epsilon_{sf} \simeq 0.04$ and $f_{sf} \simeq 0.5$, a value of $B/T \simeq 0.7$ is predicted for $\gamma \simeq 0.55$. This is in the ballpark of the value that would yield a good fit in Fig. 17, testing the validity of eq. (20) for the inflow rate within the disc.

9 CONCLUSION

We have verified that key processes of galaxy evolution in its most active phase, at $z = 1 - 4$, as captured by hydrodynamical simulations in a cosmological context, can be approximated by expressions derived from simple toy models. The systematic evolution of the stacked simulated galaxies is recovered in the best cases at the level of 10% accuracy, and in the worse cases only to within a factor of two. Even the latter can be useful for qualitative understanding and back-of-the-envelope estimates. We summarize our main findings below.

Both the average and the median total specific accretion rate into massive haloes are only weakly dependent on mass and is well approximated as a function of redshift by $\propto (1 + z)^{3/2}$, eq. (5). This includes mergers and smoother accretion into the virial radius. The systematic growth of virial mass in an individual halo is well approximated by a simple exponent in $-z$, eq. (4). This implies a slow variation in time of the absolute accretion rate into a given halo as it grows in the range $z = 5-0.3$, with a maximum near $z \sim 2.2$, eq. (10). The distribution of specific accretion rate is lognormal, with a standard deviation of 0.33 dex. The same expressions describe the accretion of baryons in our simulations.

The penetration of the inflowing baryonic mass ($M$) through the halo into the galaxy inside $0.1R_v$ is $\sim 50\%$ at $z = 4 - 2$, and somewhat larger at $z = 2 - 1$, partly reflecting the higher stellar fraction in the accretion at later times. On the other hand, the specific inflow rate $\dot{M}/M$ at $0.1R_v$ remains very similar to that at $R_v$, reflecting the fact that the baryonic mass and its rate of change vary in a similar way between the two radii. This implies that the toy-model expressions for $\dot{M}/M$ at $R_v$ can serve as a rather accurate estimator of the actual characteristic baryon input into the central galaxy.

The gas in the galaxy obeys a simple mass conservation equation, eq. (10), where the accretion is the source and the star formation (and the associated outflows) is the sink. Combined with a universal star-formation law (e.g. Krumholz, Dekel & McKee [2012]), the galaxy is predicted to converge to a cosmological quasi-steady state, with the SFR following the gas accretion rate into the central galaxy and the gas mass constant, eq. (11). We note that this may not be true for less massive galaxies and at higher redshifts, where feedback is likely to be more effective in suppressing star formation and accumulating a gas reservoir (e.g. Zolotov et al. [2012] Krumholz & Dekel [2012]).

The intense inflows into the dense early galaxies trigger violent disc instability (VDI), which drives mass inflow within the disc, building a compact bulge and feeding the central black hole. Toy model estimates of the inflow rate, performed in several independent ways, provide good approximations for the actual inflow rate in the simulations, both relative to the gas mass and relative to the gas accretion rate, eq. (20) and eq. (27). A mass conservation equation that is applied to the disc gas alone, eq. (29), predicts a steady state, in which the gas in the disc drains by star formation, outflows and inflow within the disc to the bulge, and is replenished by cosmological accretion. In this steady state, the mass fraction of gas and cold disc stars within the disc radius is rather constant in time, and the bulge mass is comparable to the disc mass or even higher. In the current simulations $\delta \sim 0.2$ and $B/T \sim 0.6-0.8$, while if the gas fraction at $z \sim 2$ is higher, as observed, then $\delta$ could be somewhat larger and $B/T$ somewhat smaller (as in Dekel, Sari & Ceverino [2009]).

We conclude that these toy models can be very useful tools in the study of key processes, as seen observationally or in simulations, and can thus help us gain understanding of the complex process of galaxy formation. In turn, the toy models allow us to perform simple analytic calculations to predict the expected evolution of galaxies. A more detailed treatment of star formation, and the effects of very strong outflows, are being and will be incorporated into the simulations, and are yet to be reflected in further testing of toy models.

ACKNOWLEDGMENTS

We acknowledge stimulating discussions with F. Bournaud, M. Krumholz, N. Mandelker, E. Neistein, and R. Sari. The simulations were performed at NASA Advanced Supercomputing (NAS) at NASA Ames Research Center, at the National Energy Research Scien-
REFERENCES

Agertz O., Moore B., Stadel J., Potter D., Miniati F., Read J., Mayer L., et al., 2007, MNRAS, 380, 963
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bauer A., Springel V., 2012, MNRAS, 423, 2558
Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 762, L31
Benson A. J., 2012, Nature, 17, 175
Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, MNRAS, 320, 261
Binney J., Tremaine S., 2008, Galactic Dynamics. Princeton, NJ, Princeton Univ. Press
Birnboim Y., Dekel A., 2003, MNRAS, 345, 349
Bond J. R., Cole S., Efstathiou G., Quinn T., et al., 2011, MNRAS, 415, 1051
Brooks A. M., Wadsley J., Quinn T., et al., 2011, MNRAS, 415, 1088
Bouche N. et al., 2010, ApJ, 718, 1001
Bournaud F., Dekel A., Teyssier R., Cacciato M., Daddi E., Juneau S., Shankar F., 2011, ApJ, 741, L33
Bournaud F., Elmegreen B. G., Elmegreen D. M., 2007, ApJ, 670, 237
Bournaud F., Elmegreen B. G., Teyssier R., Block D. L., Puera I., 2010, MNRAS, 409, 1088
Brook C. B., Governato F., Roškar R., Stinson G., Brooks A. M., Wadsley J., Quinn T., et al., 2011, MNRAS, 415, 1051
Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. I., Porciani C., Primack J. R., 2001, ApJ, 555, 240
Cacciato M., Dekel A., Genel S., 2012, MNRAS, 421, 818
Carroll S. M., Press W. H., Turner E. L., 1992, ARA&A, 30, 499
Cattaneo A. et al., 2007, MNRAS, 377, 63
Cattaneo A., Faber S. M., Binney J., Dekel A., Kormendy J., Mushotzky R., Babul A., et al., 2009, Nature, 460, 213
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, 420, 2390
Ceverino D., Klypin A., 2009, ApJ, 695, 292
Cresci G. et al., 2009, ApJ, 697, 115
Daddi E. et al., 2010, ApJ, 713, 686
Danovich M., Dekel A., Hahn O., Teyssier R., 2012, MNRAS, 422, 1732
Dekel A., Birnboim Y., 2006, MNRAS, 368, 2
Dekel A. et al., 2009, Nature, 457, 451
Dekel A., Krumholz M. R., 2013, MNRAS, in press; arXiv:1302.4457
Dekel A., Sari R., Ceverino D., 2009, ApJ, 703, 785
Dekel A., Silk J., 1986, ApJ, 303, 39
Diemer B., More S., Kravtsov A. V., 2013, ApJ, 766, 25
Dijkstra M., Loeb A., 2009, MNRAS, 400, 1109
Fakhouri O., Ma C.-P., 2008, MNRAS, 386, 577
Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189
Faucher-Giguère C.-A., Kereš D., Dijkstra M., Hernquist L., Zaldarriaga M., 2010, ApJ, 725, 633
Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, MNRAS, 417, 2982
Feldmann R., 2010, arXiv:1212.4969
Ferland G. J., Korista K. T., Werner D. A., Ferguson J. W., Kingdon J. B., Turner E. M., 1998, PASP, 110, 761
Forbes J., Krumholz M., Burkert A., 2012, ApJ, 754, 48
Forbes J., Krumholz M., Burkert A., Dekel A., 2013, arXiv:1305.0000
Förster Schreiber N. M. et al., 2009, ApJ, 706, 1364
Fumagalli M., Prochaska J. X., Kasen D., Dekel A., Ceverino D., Primack J. R., 2011, MNRAS, 418, 1796
Gammie C. F., 2001, ApJ, 553, 174
Genel S., Bouché N., Naab T., Sternberg A., Genzel R., 2010, ApJ, 719, 229
Genel S., Dekel A., Cacciato M., 2012, MNRAS, 425, 788
Genel S. et al., 2008, ApJ, 688, 789
Genel S. et al., 2012, ApJ, 745, 11
Genzel R. et al., 2008, ApJ, 687, 59
Genzel R. et al., 2011, ApJ, 733, 101
Genzel R., Tacconi L. J., Eisenhauer F., Förster Schreiber N. M., Cimatti A., Daddi E., Bouché N., et al., 2006, Nature, 442, 786
Goerdt T., Dekel A., Sternberg A., Ceverino D., Teyssier R., Primack J. R., 2010, MNRAS, 407, 613
Guo Q. et al., 2011, MNRAS, 413, 101
Haardt F., Madau P., 1996, ApJ, 461, 20
Hirschmann M., Naab T., Somerville R. S., Burkert A., Oser L., 2012, MNRAS, 419, 3200
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1
Hopkins P. F., Kereš D., Murray N., Quataert E., Hernquist L., 2012, MNRAS, 427, 968
Kennicutt, Jr. R. C., 1998, ApJ, 498, 541
Kereš D., Katz N., Fardal M., Davé R., Weinberg D. H., 2009, MNRAS, 395, 160
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Khochfar S., Silk J., 2009, ApJ, 700, L21
Kim T., Devriendt J., Slyz A., Pichon C., Kassin S. A., Dubois Y., 2011, arXiv:1106.0538
Klypin A. A., Trujillo-Gomez S., Primack J., 2009, ApJ, 740, 102
Kodama T., Katsuda S., Couch S., Hiratsuka T., 2013, ApJ, 764, 73
Komatsu E., Dunkley J., Nolta M. R., Bennett C. L., Hinshaw G., Jarosik N., et al., 2009, ApJS, 180, 330
Kravtsov A. V., 2013, ApJ, 764, L31
Kravtsov A. V., Klypin A. A., Khokhlov A. M., 1997, ApJS, 111, 73
Krumholz M., Burkert A., 2010, ApJ, 724, 895
Krumholz M. R., Dekel A., 2010, MNRAS, 406, 112
Krumholz M. R., Dekel A., 2012, ApJ, 753, 16
Galaxy-Formation Toy Models

APPENDIX A: USEFUL RELATIONS

We summarize here the more accurate cosmological relations, valid in the standard ΛCDM cosmology at all redshifts, which the toy-model expressions of this paper approximate in the EdS regime, z > 1. This is largely basic material, based for example on Lahav, Rees, Lilje, & Primack [1991], Carroll, Press, & Turner [1992], and Mo & White [2002]; Birnboim & Dekel [2003]; Dekel & Birnboim [2006], with the addition of accretion rates from Neistein, van den Bosch, & Dekel (2006); Neistein & Dekel (2008b).

A1 Cosmology

The basic parameters characterizing a flat cosmological model in the matter era are the current values of the mean mass density parameter Ω0, and the Hubble constant H0. At the time associated with expansion factor a = 1/(1 + z), the vacuum-energy density parameter is

\[ \Omega_{\Lambda}(a) = 1 - \Omega_m(a) \]

and the age of the universe is

\[ t(a) = \frac{2}{3} H(a)^{-1} \sinh^{-1}\left(\frac{1 - \Omega_m(a)}{\Omega_m(a)}\right)^{1/2} \]

The mean mass density is

\[ \rho_a \simeq 2.76 \times 10^{-30} \Omega_{m,0.3} h_{0.7}^2 a^{-3} \]

where \( \Omega_{m,0.3} \equiv \Omega_m/0.3, h \equiv H_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1} \), and h_{0.7} ≡ h/0.7.

A2 Virial relations

The virial relations between halo mass, velocity and radius,

\[ V_{\text{c}}^2 = \frac{G M_{\text{c}}}{R_{\text{c}}} - \frac{M_{\text{c}}}{(4\pi/3) R_{\text{c}}^3} = \Delta \rho_a \]

become

\[ V_{200} \simeq 1.02 M_{12}^{1/3} A_{1/3}^{-1/3}, \quad R_{200} \simeq 1.03 M_{12}^{1/3} A_{1/3}, \]

where \( M_{12} \equiv M_c/10^{12} M_\odot \), \( V_{200} \equiv V_c/200 \text{ km s}^{-1} \), \( R_{200} \equiv R_c/100 \text{ kpc} \), and

\[ A \equiv (\Delta_{200} \Omega_{m,0.3} h_{0.7}^2)^{-1/3} a. \]

An approximation for \( \Delta(a) \) in a flat universe (Bryan & Norman 1998) is:

\[ \Delta(a) \simeq (18\pi^2 - 82\Omega_{\Lambda} - 399\Omega_{\Lambda}(a)^2)/\Omega_m(a). \]

In the EdS regime, z > 1, or when referring to \( R_{200} \), instead of \( R_c \) at all redshifts, \( A \approx (1 + z)^{-1} \) to an accuracy of a few percent. The virial temperature, defined by \( kT_c/m = (1/2) V_c^2 \) for an isotropic, isothermal sphere,
Based on the improved formalism of Sheth & Tormen (1999), Carroll, Press & Turner (1992), Mo & White (2002)

\[ D(a) = \frac{g(a)}{g(1)} a, \quad (A10) \]

where

\[ g(a) \approx \frac{5}{2} \Omega_m(a) \times \left[ \Omega_m(a)^{4/7} - \frac{1}{(1 + \Omega_m(a)/2)} \right]^{-1}. \quad (A11) \]

The CDM power spectrum is crudely approximated by Bardeen et al. (1986)

\[ P(k) \propto k^2 T^2(k), \quad (A12) \]

with

\[ T(k) = \frac{\ln(1 + 2.34 q)}{2.34 q} \times \left[ 1 + 3.89 q + (16.1q)^2 + (5.46q)^3 + (6.71q)^4 \right]^{-1/4}, \quad (A13) \]

where

\[ q = k/(\Omega_m h^2 \text{Mpc}^{-1}). \quad (A14) \]

It is normalized by \( \sigma_8 \) at \( R = 8 \text{ h}^{-1} \text{Mpc} \), where

\[ \sigma_8^2(R) = \frac{1}{2\pi} \int_0^\infty dk k^2 P(k) \tilde{W}^2(kR), \quad (A15) \]

and with the Fourier transform of the top-hat window function

\[ \tilde{W}(x) = 3(\sin x - x \cos x)/x^3. \quad (A16) \]

In the Press Schechter (PS) approximation, the characteristic halo mass \( M_{\text{ps}}(a) \) is defined as the mass of the 1-\( \sigma \) fluctuation,

\[ 1 = \nu(M, a) = \frac{\delta_c}{D(a) \sigma(M)}, \quad \delta_c \approx 1.68, \quad (A17) \]

where \( M \) and the comoving radius \( R \) are related via the universal density today: \( M = \frac{4\pi}{3} \rho_0 R^3 \). The mass of 2-\( \sigma \) fluctuations is obtained by setting \( \nu(M, a) = 2 \), etc. Based on the improved formalism of Sheth & Tormen (2002), the fraction of total mass in haloes of masses exceeding \( M \)

\[ F(>M, a) \approx 0.4 \left( 1 + \frac{0.4}{\rho_{0.1}} \right) \text{erfc} \left( \frac{0.85 \nu}{\sqrt{2}} \right). \quad (A18) \]

This fraction for 1-\( \sigma \), 2-\( \sigma \), and 3-\( \sigma \) fluctuations is 22%, 4.7%, and 0.54% respectively.

Figure A1 shows the PS mass \( M_{\text{ps}} \) as a function of redshift (blue). Also shown are the approximations \( \log M_{\text{ps}} = 12.7 - 1.29z \) (green), and \( \log M_{\text{ps}} = 12.78 - 1.46z^{0.88} \) (red).

The average growth of a halo, following its main progenitor, has been derived by a fit to merger trees from the Millennium cosmological simulation (Springel & Hernquist 2003), using a functional form that is motivated by the EPS approximation (Neistein & Dekel 2008). In terms of the self-similar dimensionless time variable \( \omega = \delta_c[D(a)^{-1} - 1] \), the growth rate of halo of mass \( M \) at \( \omega \)

\[ \frac{dM}{d\omega} = -\alpha M^{1+\beta}, \quad (A19) \]

and the mass at \( \omega \), given that the mass at \( z = 0 \) is \( M_0 \), is

\[ M(\omega|M_0) = (M_0^{-\beta} + \alpha \beta \omega)^{-1/\beta}, \quad (A20) \]

Note that the Sheth-Tormen approximation becomes quite inaccurate at very high redshifts (Klypin, Trujillo-Gomez & Primack 2011).

\[ 2 \text{ Note that in the initial conditions for our simulations we use the more accurate CAMB numerical solutions based on CAM}^\text{fast} \text{ (Seljak and Zaldarriaga, http://lambda.gsfc.nasa.gov/toolbox/).} \]

\[ 3 \text{ Note that the Sheth-Tormen approximation becomes quite inaccurate at very high redshifts (Klypin, Trujillo-Gomez & Primack 2011).} \]
where $M$ is in $10^{12} M_\odot$. The value $\beta = 0.14$ fits well the Millennium trees for $M$ in the range $10^{11} - 10^{14} M_\odot$, despite the fact that according to EPS it should vary as $\beta = (n + 3)/6$, where $n$ is the local power-spectrum index. At lower masses, $\beta$ should be smaller, so the mass dependence of the accretion rate becomes even closer to linear. For $\sigma_8 = 0.81$, the best fit is with $\alpha = 0.623$ (it scales as $\propto \sigma_8^{-1}$). The accuracy of this fit to the Millennium trees is better than 5% for $z < 5$.

The following approximations (better than 0.5% accuracy at all $z$ for $\Omega_m = 0.27$, $h = 0.7$, $\sigma_8 = 0.81$) relate $\omega$ to $z$,

$$\omega = 1.28[(1+z) + 0.086(1+z)^{-1} + 0.22e^{-1.2z}], \quad (A21)$$

and provide the standard accretion rate $\dot{M} = \omega dM/d\omega$

$$\omega = -0.0476[(1+z) + 0.093(1+z)^{-1.22}]^{2.5} \text{Gyr}^{-1}. \quad (A22)$$

The latter scales in proportion to $\propto h_0.7$. Note that in the EdS regime, valid at high $z$, the asymptotic dependence is $\dot{\omega} \propto (1+z)^{-2.5}$, because $D(t) = (1+z)^{-1} \propto t^{2/3}$, so $\dot{M} \propto t^{-5/3} \propto (1+z)^{5/2}$.

Based on these expressions, a practical approximation for the typical baryonic accretion rate for a halo of mass $M$ at $z$ is

$$\dot{M}_b = 80 M_\odot^{12/5} (1+z) f_{0.17} M_\odot \text{yr}^{-1}, \quad (A23)$$

where $f_{0.17}$ is the baryonic fraction in units of 0.17. With the EdS asymptotic value $\mu = 2.5$, the accuracy is $< 5\%$ for $z > 1$, and it becomes an underestimate of $\sim 20\%$ at $z = 0$. With $\mu = 2.4$ the accuracy is $< 5\%$ for $0.2 < z < 5$, and is an underestimate of $\sim 10\%$ at $z = 0$ and $z = 10$. For most practical purposes we adopt $\beta = 0.14$ and $\mu = 2.4$.

Figure A2 is a useful summary for appreciating the average baryon accretion rate onto haloes as they grow during cosmological evolution. Each of the blue curves refers to a given value of $M_b$ (as marked). They are obtained using the inverse of eq. (A19) and eq. (A22). Each of the red curves shows the average mass growth of the main progenitor of a halo that ends up with a given mass $M$ at $z = 0$ (for log $M_\odot = 11.5, 12.5, 13.5$), using eq. (A20) and eq. (A21). For a given halo of $M_c$ at $z$, one can read the baryon accretion rate at the corresponding point in the $M_c, z$ plane by its position relative to the nearest reference blue curve. For example, a halo of $M_c = 10^{12} M_\odot$ at $z = 2$ accretes baryons at about $80 M_\odot \text{yr}^{-1}$, as eq. [A23].

By following the red curves one can read the history of accretion rate onto a given halo as it grows. Take for example the middle red curve, referring to a halo of $M_0 = 3 \times 10^{12} M_\odot$ at $z = 0$. It grew between $z = 4$ and 0.5 from $M_c \simeq 2 \times 10^{11} M_\odot$ to $2 \times 10^{12} M_\odot$, while maintaining an accretion rate of $40 - 60 M_\odot \text{yr}^{-1}$.

**Figure A2.** Baryonic accretion rate, halo mass and redshift. The solid blue curves refer to constant values of baryon accretion rate (an upper limit to the SFR) as marked. The long-dashed red curves show the average mass growth of the main progenitor of a halo of a given mass $M_0$ at $z = 0$, for log $M_\odot = 11.5, 12.5, 13.5$ from bottom to top. The short-dash green curve is the Press-Schechter mass.

$$t_{\text{mig}} \sim \delta^{-2} t_d, \quad (A24)$$

where $\delta$ is the cold mass fraction defined by $M_d = \delta M_{\text{tot}}$, with $M_d$ the cold mass in the disc and $M_{\text{tot}}$ the total mass encompassed by the disc radius $R_d$. The timescale for evacuating a disc mass $M_d$ is then

$$t_{\text{evac}} = m_c^{-1} t_{\text{mig}}, \quad (A25)$$

where $m_c \sim 0.2$ is the instantaneous fraction of the disc mass in clumps. This estimate was based on the energy exchange in the gravitational encounters between clumps.

Here we provide an alternative crude estimate of the clump migration time due to the dynamical friction that is exerted on the clump by the off-clump disc mass. Consider a clump of mass $M_c$ in a circular orbit with velocity $V \sim (G M_{\text{tot}} / R_d)^{1/2}$. The deceleration due to dynamical friction is approximated by the Chandrasekhar formula (Binney & Tremaine 2008, sec. 8.1) to be

$$\dot{V} \sim -4\pi G^2 M_c \rho_\odot (\ln \Lambda) V^{-2}. \quad (A26)$$

In a $Q \sim 1$ marginally unstable disc, the clump mass is related to the disc mass by $M_c \simeq \delta^2 M_d = \delta^3 M_{\text{tot}}$, (A27) and the disc half thickness $h$ is related to its radius by $h/R_d \sim \delta$. (A28)

(e.g. Dekel, Sari & Ceverino 2003). Thus, the mean density in the disc is

$$\rho_\odot \simeq \frac{M_d}{2\pi R_d^2 h} \simeq \frac{M_{\text{tot}}}{2\pi R_d^2}. \quad (A29)$$

In the Coulomb logarithm $\Lambda \sim M_d/M_c \sim \delta^{-2}$ so $\ln \Lambda$ is assumed to be $\sim 2$. We finally obtain

**A5 Clump Migration by Dynamical Friction**

The timescale for inward clump migration in a marginally unstable disc, $Q \sim 1$, has been estimated by Dekel, Sari & Ceverino (2003) to roughly be
$t_{\text{mig}} = \frac{V}{\dot{V}} \sim \frac{1}{4} \delta^{-3} t_d$. \hfill (A30)

This is indeed comparable to the estimate in eq. (A24)
for $\delta \sim 0.2 - 0.3$, the typical value in the steady-state
solution of high-z disks (Dekel, Sari & Ceverino 2009;
Cacciato, Dekel & Genel 2012) and in our current sim-
ulations, Fig. 18. The disc evacuation time by clump
migration due to dynamical friction is thus

$t_{\text{evac}} \sim \delta^{-3} t_d$. \hfill (A31)
