

The evolution of turbulent galactic discs: gravitational instability, feedback, and accretion

Omri Ginzburg,¹★ Avishal Dekel,^{1,2} Nir Mandelker¹ and Mark R. Krumholz^{3,4}

¹*Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel*

²*SCIPP, University of California, Santa Cruz, CA 95064, USA*

³*Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia*

⁴*Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*

Accepted 2022 May 5. Received 2022 April 20; in original form 2022 February 6

ABSTRACT

We study the driving of turbulence in star-forming disc galaxies of different masses at different epochs, using an analytic ‘bathtub’ model. The disc of gas and stars is assumed to be in marginal Toomre instability. Turbulence is assumed to be sustained via an energy balance between its dissipation and three simultaneous energy sources. These are stellar feedback, inward transport due to disc instability and clumpy accretion via streams. The transport rate is computed with two different formalisms, with similar results. To achieve the energy balance, the disc self-regulates either the mass fraction in clumps or the turbulent viscous torque parameter. In this version of the model, the efficiency by which the stream kinetic energy is converted into turbulence is a free parameter, ξ_a . We find that the contributions of the three energy sources are in the same ball park, within a factor of ~ 2 in all discs at all times. In haloes that evolve to a mass $\leq 10^{12} M_\odot$ by $z = 0$ ($\leq 10^{11.5} M_\odot$ at $z \sim 2$), feedback is the main driver throughout their lifetimes. Above this mass, the main driver is either transport or accretion for very low or very high values of ξ_a , respectively. For an assumed $\xi_a(t)$ that declines in time, galaxies in haloes with present-day mass $> 10^{12} M_\odot$ make a transition from accretion to transport dominance at intermediate redshifts, $z \sim 3$, when their mass was $\geq 10^{11.5} M_\odot$. The predicted relation between star formation rate and gas velocity dispersion is consistent with observations.

Key words: stars: formation – ISM: kinematics and dynamics – galaxies: disc – galaxies: formation – galaxies: star formation.

1 INTRODUCTION

Most of the star formation in galaxies tends to occur in galactic discs (Wuyts et al. 2011; Huertas-Company et al. 2015). This observational result makes the evolution of gaseous discs a key for our understanding of galaxies.

According to our current knowledge, very thin discs, with rotation-to-dispersion velocity ratios $V_d/\sigma_g > 5$, are expected to emerge at or after $z \sim 1$ (Kassin et al. 2012). At higher redshifts, as a result of intense accretion through streams and associated dynamical instabilities, discs tend to be more perturbed, with $V_d/\sigma_g \sim 3$ (Genzel et al. 2006, 2008; Simons et al. 2017; Förster Schreiber et al. 2018), although some observations suggest that massive cold discs can form as early as redshift $z \approx 4$, with $V_d/\sigma_g \sim 10$, at least for short periods (Neeleman et al. 2020; Rizzo et al. 2020; Lelli et al. 2021). Theoretically, Dekel et al. (2020a) found, using analytical estimates as well as cosmological simulations, that frequent spin flips, driven by mergers, tend to destroy discs within an orbital time when they reside in haloes with mass $M_h \lesssim 2 \times 10^{11} M_\odot$, roughly independent of redshift. Kretschmer, Dekel & Teyssier (2022) found cold discs with $V_d/\sigma_g \sim 5$ in cosmological simulations at redshift $z \sim 3.5$, with an even colder molecular disc, of $V_d/\sigma_{g, \text{mol}} \sim 8$.

Self-gravitating, rotating discs tend to undergo various instabilities, collectively termed violent disc instabilities (VDI). The most common of these instabilities is the Toomre instability (Toomre 1964; Noguchi 1999; Dekel, Sari & Ceverino 2009b), governed by the Toomre- Q parameter (see Section 2.2), which expresses the balance between self-gravity on one hand, and pressure and rotation on the other hand. A Q parameter of around unity indicates marginal equilibrium between these forces. Observed galaxies, both local and at high redshifts, indeed seem to have Q parameter of order unity (Genzel et al. 2011; Romeo & Falstad 2013; Genzel et al. 2014; Obreschkow et al. 2015).

In order to maintain a constant level of Q near unity, for a given rotation velocity, the turbulence at the disc must remain at the level required to counteract self-gravity without overstabilizing the disc. However, turbulence, especially at supersonic levels as observed over a wide range of redshifts (Kassin et al. 2012; Ianjamasimanana et al. 2015; Wisnioski et al. 2015; Stott et al. 2016), tends to decay rapidly in a dynamical time. Hence, in order to sustain the turbulence levels required for marginal Toomre instability, a continuous supply of energy to turbulence is required. It is therefore important to understand the different energy inputs from various mechanisms that drive turbulence in marginally Toomre-unstable discs.

There are several physical processes that occur regularly in discs that can provide a continuous supply of energy to turbulence. First, star formation results in supernova explosions and various forms of stellar feedback. These release large amounts of energy

* E-mail: omry.ginzburg@mail.huji.ac.il

and momentum to their surroundings (Draine 2011). A fraction of this energy turns into heat, while the rest can turn into kinetic energy, partially driving outflows and partially driving turbulence (Hayward & Hopkins 2017).

A second energy source for turbulence is radial mass transport. When mass is transported inwards through the disc and down the overall potential well, there is a gravitational energy gain that could be converted into turbulence (Wada, Meurer & Norman 2002). A natural driver of inward mass transport is the disc instability itself, where the deviations from circular symmetry exert torques that tend to drive angular momentum out, and therefore transport mass in by conservation of angular momentum. As previously mentioned, the instability tends to self-regulate and maintain $Q \sim 1$. The self-regulation mechanism can be understood as follows: If Q falls below unity, the torques get stronger, the inward transport rate becomes higher, the gravitational energy gain by this transport down the potential well enhances the turbulence, which in turn increases Q . If Q becomes larger than unity, the torques weaken, the transport rate is reduced, the energy gain becomes smaller, the turbulence is reduced and Q becomes smaller. While the above picture gives an idea of how disc-instability can give rise to mass transport and turbulence in a self-regulated way, the precise details of how VDI generates inwards mass transport are unclear. We here consider two possibilities, detailed below.

The first mechanism is encounters between giant clumps in the disc. Galaxies, in the redshift range $z \approx 1-3$, are typically observed to host multiple massive, star-forming giant clumps (Elmegreen & Elmegreen 2005; Genzel et al. 2011; Shibuya et al. 2016; Guo et al. 2018; Huertas-Company et al. 2020; Ginzburg et al. 2021). These clumps are consistently found in UV, H α , and CO observations (Förster Schreiber et al. 2009, 2011; Dessauges-Zavadsky et al. 2019), which serve as tracers of star formation, as well as in the optical regime (Huertas-Company et al. 2020). Giant star-forming clumps are also observed, though less frequently, in local, gas-rich systems (Fisher et al. 2017; Lenkić et al. 2021). The emergence of these clumps is a robust theoretical prediction, found in isolated as well as cosmological simulations (Noguchi 1999; Agertz, Teyssier & Moore 2009; Dekel et al. 2009b; Genel et al. 2012b; Bournaud et al. 2014; Mandelker et al. 2014; Mandelker et al. 2017; Oklopčič et al. 2017). The rate at which clumps interact with each other and their surroundings, along with the resulting torques, can be computed in the Toomre framework to get an estimate for the mass transport inwards (Dekel et al. 2009b). This transport of mass is partially due to clump migration and partially to interclump inflow.

The second mechanism that we consider below for inducing mass transport is turbulent viscosity. Turbulence can be modelled as an effective viscosity that is added to the gas, well beyond the molecular viscosity. We use the standard theory for viscous accretion discs by Shakura & Sunyaev (1973), followed by Balbus & Papaloizou (1999) and simulations by Gammie (2001). Krumholz & Burkert (2010) introduced a hydrodynamical model for turbulent viscous discs that includes the effect of star formation. They showed that the systems usually reach a steady state, which allows one to compute the torques induced by this effective viscosity, and from that the resulting mass transport inwards.

Finally, the third energy source we consider for driving turbulence in the disc is the intense external accretion on to it. The primary mechanism by which high-redshift galaxies obtain fresh gas is via accretion along cold streams (Kereš et al. 2005; Dekel & Birnboim 2006; Dekel et al. 2009a). Typically, the disc is fed by three streams which are confined to a plane with most of the inflowing mass co-rotating with the disc (Danovich et al. 2012, 2015). As the streams

flow into the disc, they are subject to various instabilities, including gravitational instabilities, Kelvin–Helmholtz instability and thermal instabilities (Mandelker et al. 2016; Padnos et al. 2018; Aung et al. 2019; Mandelker et al. 2018, 2019, 2020a), which cause the streams to fragment and form bound clumps within them (Mandelker et al. 2018, 2020a). Similar instabilities also occur in ISM filaments inside giant molecular clouds (GMCs; Clarke & Whitworth 2015; Clarke, Whitworth & Hubber 2016; Clarke et al. 2017). Such clumpy accretion can be very efficient at transferring the kinetic energy associated with the accretion into turbulent energy within the disc, when the clumps collide with the disc (Klessen & Hennebelle 2010).

The nature of turbulent support in discs undergoing VDI has been studied with various levels of complexity. Dekel et al. (2009b) and Cacciato, Dekel & Genel (2012) studied the evolution of discs in which clump encounters are the main drivers of turbulence and mass transport, without the input of supernova, stellar feedback or accretion. Krumholz & Burkert (2010), Forbes, Krumholz & Burkert (2012), and Forbes et al. (2014) studied the effects of disc instability on the turbulence within the disc, using the turbulent viscosity formalism. Faucher-Giguère, Quataert & Hopkins (2013) and Hayward & Hopkins (2017) argued that supernova feedback-driven turbulence is enough to explain the high levels of turbulence in galaxies. However, this requires a very high star formation rate (SFR), with a star formation law $\dot{\Sigma}_{\text{SF}} \propto \Sigma_g^2$, where $\dot{\Sigma}_{\text{SF}}$ and Σ_g are the SFR and gas densities, steeper than observed (Kennicutt 1998). Elmegreen & Burkert (2010) studied the driving of turbulence by accretion in young discs, finding that accretion alone cannot sustain turbulence at high levels for long periods of time. However, their considered efficiency with which the accretion drives turbulence was low, and does not reflect the possible enhancement of turbulence driving by the clumps in the streams. Genel, Dekel & Cacciato (2012a) considered the energy input by accretion and disc instability, but they did not include the input by supernova or stellar feedback, and their treatment of mass transport due to disc instability did not take into account the turbulent state of the disc.

Krumholz et al. (2018) developed a comprehensive model for the evolution of disc galaxies, accounting for turbulence driven by both supernova feedback and transport, extending the framework of turbulent viscous torques developed in Krumholz & Burkert (2010), but without including the independent contribution of accretion to the driving of turbulence, but only as a source of mass to the disc. They found that at high redshifts, turbulence in galaxies tends to be driven primarily by transport, while supernova feedback is sufficient only in low-mass galaxies at low redshifts.

In this paper, inspired by the recent studies of Mandelker et al. (2018), Aung et al. (2019), and Mandelker et al. (2020b) regarding the clumpiness of cosmological streams, we consider simultaneously the three energy sources for turbulence. In particular, we extend the earlier studies to address the role of accretion-driven turbulence in sustaining turbulence in discs as a function of halo mass and redshift. We wish to ascertain how different prescriptions for mass transport and for the efficiency of converting accretion energy into turbulence affect the result. We defer a study of how exactly the different mechanisms drive turbulence to future work, and focus here only on their energy input to turbulence.

The paper is organized as follows. In Section 2, we present our model. To connect to earlier studies, we model several basic properties of the galaxies using an approach similar to that of Krumholz et al. (2018) while adding the necessary new prescriptions for transport and accretion. In Section 2.3, we describe the different mass sources and sinks, and in Section 2.4, we address the different drivers of turbulence and their corresponding energy input. In Section 3, we

show the results of numerically integrating the mass conservation equation, and study the relative roles of the different energy sources in driving disc turbulence. In Section 4, we discuss caveats of the model. We present our conclusions in Section 5.

2 THE MODEL

We model the galaxy as a disc embedded in a dark matter halo. The discs undergo accretion of fresh gas from the cosmic web, and can lose gas by either star formation or mass transport from the disc to a central bulge, which we consider as separate from the disc. Our model is based on two main assumptions:

- (i) The disc maintains marginal Toomre instability, with a two component Toomre- Q around unity (Section 2.2),
- (ii) The disc maintains turbulent energy balance, where turbulent dissipation is balanced by driving via the three sources discussed above (Section 2.5).

In the following sections, we lay out in detail each component of our model.

2.1 Galaxy–halo connection

Galaxies are assumed to be embedded in dark matter haloes with virial radii and velocities determined by cosmology, given approximately by (Dekel et al. 2013)

$$\frac{R_v}{100 \text{ kpc}} \approx 0.54 M_{h,12}^{1/3} \left(\frac{1+z}{3} \right)^{-1}, \quad (1)$$

$$\frac{V_v}{200 \text{ km} \cdot \text{s}^{-1}} \approx M_{h,12}^{1/3} \left(\frac{1+z}{3} \right)^{1/2}, \quad (2)$$

where $M_{h,12} = M_h/10^{12} M_\odot$ is the halo virial mass. The halo is assumed to follow an NFW profile (Navarro, Frenk & White 1997), characterized by a concentration c .

We follow Krumholz et al. (2018) and assume that the radius of the galaxy is proportional to the virial radius, with a constant proportionality factor

$$R_d \approx 0.035 R_v. \quad (3)$$

This resembles the relation $R_d = \lambda R_v$, with $\lambda \approx 0.04$ the halo spin parameter (Bullock et al. 2001), which follows an assumption of angular momentum conservation during disc formation¹ (Fall & Efstathiou 1980; Mo, Mao & White 1998). We assume the galaxy rotation velocity to be proportional to the halo maximum circular velocity, $V_d \approx \gamma V_{\max}$, which for an NFW halo is given by

$$V_{\max} = 0.465 \sqrt{\frac{c}{\ln(1+c) - c/(1+c)}} V_v. \quad (4)$$

We choose as a fiducial value $c = 10$, valid on average for a halo with $M_{h,12} = 1$ at $z = 1$ (Zhao et al. 2009), and note that values of V_{\max} for the range $c = 1-30$ deviate from each other by less than a factor of 2, and this deviation can be absorbed into the parameter γ . Krumholz et al. (2018) chose $\gamma = 1.4$ quite arbitrarily in order to achieve the correct rotation speed of a Milky Way-like galaxy. It turns out that this choice is in good agreement with cosmological simulations (Dekel et al. 2020b, fig. 7), which find $\gamma \approx 1-1.5$ in

¹We note that it has been shown in simulations that the size of a given galaxy is not correlated with the spin of its host halo (Jiang et al. 2019), because the gas angular momentum is not conserved during its contraction (Danovich et al. 2015).

the relevant halo-mass range for discs (see discussion in Section 3). This small scatter introduces small changes to the dynamical time that will not affect our qualitative results. We therefore also adopt $\gamma = 1.4$ as our fiducial value.

With the disc radius and rotation velocity in hand, we can compute the dynamical time of the galaxy, given by

$$t_d = \frac{1}{\Omega} = \frac{R_d}{V_d} \sim 50-500 \text{ Myr}. \quad (5)$$

2.2 Disc instability

The first main assumption in our model is that self-gravitating, rotating discs undergo gravitational instability and self-regulate to marginal instability. The linear stability of a thin, gaseous, self-gravitating, rotating disc, following Toomre (1964), is governed by the Toomre- Q parameter

$$Q_g = \frac{\kappa \sigma_g}{\pi G \Sigma_g}, \quad (6)$$

where $\Sigma_g = M_g/\pi R_d^2$ is the disc gas mass surface density, σ_g is the gas radial velocity dispersion, and κ is the disc epicyclic frequency. Assuming an isotropic velocity dispersion, σ_g is equivalent to the one-dimensional velocity dispersion. For a disc with a rotation curve $v \propto r^\beta$, the epicyclic frequency is given by $\kappa = \sqrt{2(1+\beta)}\Omega$. We will focus on discs with flat rotation curves, namely $\beta = 0$. The disc is stable to axisymmetric perturbations when $Q_g > 1$. When $Q_g < 1$,² the disc is unstable to axisymmetric perturbations, and will break up into rings that will eventually fragment into bound clumps.

Following Dekel et al. (2009b), we write the Toomre- Q parameter in the form

$$Q_g = \sqrt{2(1+\beta)} \delta^{-1} \frac{\sigma_g}{V_d}, \quad (7)$$

where $\delta = M_g/M_{\text{tot}}$, with M_g being the gas mass within the disc,³ and M_{tot} is the total mass within a sphere of a radius R_d , including the bulge and dark matter component.

The picture of gravitational instability slightly changes when one takes into account both cold material, including cold gas and young stars, together with hotter, older stars as the drivers of instability. When considering such a two component disc, the instability is governed by an effective Toomre- Q parameter, that can be approximated by Rafikov (2001) and Romeo & Falstad (2013):

$$Q = \left(Q_g^{-1} + \frac{2\sigma_g\sigma_s}{\sigma_g^2 + \sigma_s^2} Q_s^{-1} \right)^{-1}, \quad (8)$$

where Q_g is given by equation (6), and Q_s is the analogous Toomre- Q parameter for the stellar component.⁴ The disc of gas and stars is stable whenever $Q > 1$. Following Krumholz et al. (2018), we write the two component Toomre- Q parameter as

$$Q = f_{g,Q} Q_g, \quad (9)$$

²If the disc has a finite thickness, the critical Toomre- Q for instability is modified, depending on the anisotropy of the velocity dispersion. For an isotropic velocity dispersion, the critical Toomre- Q for instability is 0.67 rather than unity (Goldreich & Lynden-Bell 1965).

³The mass that goes into δ in principle consists of all of the cold mass that participates in the instability, including young stars.

⁴When performing the linear analysis for a collisionless stellar disc, the π in equation (6) is replaced with 3.36 (Toomre 1964), a negligible difference that we ignore here.

by defining

$$f_{g,Q} = \frac{M_g}{M_g + 2 \left[\sigma_g^2 / (\sigma_g^2 + \sigma_s^2) \right] M_s}, \quad (10)$$

which is an effective gas fraction that coincides with the common gas fraction when $\sigma_g \approx \sigma_s$. We note that the two component disc can be unstable even when each component is stable had it been a one component disc.

We assume that the discs remain in a state of two-component marginal instability with $Q = 1$. The value of $f_{g,Q}$ is near 0.7 (see Section 2.6) for a gas-rich disc, appropriate for high redshift galaxies (Genzel et al. 2014), while it is near a value of $f_{g,Q} = 0.5$ in the Solar neighbourhood of the Milky Way, as well as star forming galaxies in the xCOLD GASS survey (Krumholz et al. 2018; Yu et al. 2021). In Section 2.6, we explain how we parametrize the transition from high to low redshifts as a continuous function of redshift.

2.3 Mass budget

The gas mass in our model follows a mass conservation equation of the form

$$\dot{M}_g = \dot{M}_{\text{source}} - \dot{M}_{\text{sink}}, \quad (11)$$

where \dot{M}_{source} represents the rate at which mass is added to the gaseous disc, while \dot{M}_{sink} represents the rate at which the disc loses gas mass, for example by star formation or by the transport of mass into a bulge. In the following sections, we lay out the different sources and sinks assumed in the model.

2.3.1 Gas accretion

Galaxies are subject to accretion of baryons and dark matter from the cosmic web. The halo accretion rate is robustly estimated from Extended Press–Schechter (EPS) theory. We adopt an approximation calibrated to match cosmological simulations, and is given by (Neistein & Dekel 2008)

$$\frac{\dot{M}_h}{M_h} = -a M_{h,12}^b \dot{\omega}. \quad (12)$$

Here, $b = 0.14$, $a = 0.628$, and ω is the EPS self-similar time variable, whose time derivative is given by (Neistein & Dekel 2008)

$$\dot{\omega} = -0.0476 (1+z + 0.093(1+z)^{-1.22})^{2.5} \text{Gyr}^{-1}. \quad (13)$$

A good approximation to equation (12) in the EdS cosmological regime, valid for roughly $z > 1$, is given by

$$\frac{\dot{M}_h}{M_h} \approx s(1+z)^\mu \text{Gyr}^{-1}. \quad (14)$$

where $\mu = 5/2$. The power $\mu = 5/2$ can be simply understood from a scaling argument, based on the Press–Schechter theory (Press & Schechter 1974). A key element in the Press–Schechter formalism is the self-invariant time variable, $\omega \propto D(a)^{-1}$, where $D(a)$ is the growth rate of linear perturbations. Self-invariance implies that the growth rate of haloes with respect to ω must be independent of ω , namely $dM/d\omega = \text{const}$. This can be written as $\dot{M} \propto \dot{\omega}$. In the EdS cosmological regime, $\omega \propto D(a)^{-1} \propto a^{-1}$. With $a \propto t^{2/3}$, this implies that $\dot{M} \propto a^{-5/2} = (1+z)^{5/2}$. At lower redshifts, the power is slightly smaller (Neistein & Dekel 2008).

Equation (14) can be integrated analytically, resulting in (Dekel et al. 2013)

$$M_h = M_{h,z_0} e^{-\tilde{s}(z-z_0)}, \quad \tilde{s} \approx 0.79. \quad (15)$$

In practice, we numerically integrate equation (12) to generate mass histories of haloes that will host our galaxies.

We assume that the amount of accreted baryons in the overall accretion is $f_b \dot{M}_h$, where $f_b \approx 0.17$ is the universal baryonic fraction (Dekel et al. 2013). However, not all of the accreted baryons eventually reach the galaxy, as the penetration of cold material is suppressed when the galaxy supports a virial shock, and flows mainly through cold narrow streams that penetrate the hot halo. This virial shock suppression is expected above halo masses $M_{h,12} \gtrsim 1$ (Birnbom & Dekel 2003; Dekel & Birnbom 2006). To take this into account, we use the approximation used by Krumholz et al. (2018), based on fits to numerical simulations by (Faucher-Giguère, Kereš & Ma 2011):

$$\dot{M}_{g,\text{acc}} = \epsilon_{\text{in}} f_b \dot{M}_h, \quad (16)$$

where ϵ_{in} is a penetration parameter, parametrized as

$$\epsilon_{\text{in}} = \min(\epsilon_0 M_{h,12}^{\alpha_1} (1+z)^{\alpha_2}, 1). \quad (17)$$

The best-fitting parameters for $z \geq 2$ are $\epsilon_0 = 0.31$, $\alpha_1 = -0.25$, and $\alpha_2 = 0.38$. In the mass and redshift range of interest for us here, this fit is good to within a factor of 2. At $z \leq 1$ and $M_h \geq 10^{13} M_\odot$, the fit may overestimate the true penetration found by Faucher-Giguère et al. (2011) by up to a factor of 3, but haloes of such masses are not expected to host discs (see below), and are therefore not relevant for us here.

2.3.2 Star formation

Motivated by the Kennicutt–Schmidt law (Schmidt 1959; Kennicutt 1998), we assume that the instantaneous SFR is proportional to the mass in gas,

$$\dot{M}_{\text{SF}} = \epsilon_{\text{ff}} \frac{M_g}{t_{\text{ff}}}, \quad (18)$$

where t_{ff} is the free fall time estimated at the mid-plane of the disc, and ϵ_{ff} is the efficiency of star formation per free-fall time. The latter has been extensively studied, both observationally and theoretically (Krumholz & McKee 2005; Vutisalchavakul, Evans & Heyer 2016). The values are usually in the range $\epsilon_{\text{ff}} \approx 0.01–0.1$, and depend on the relevant scale at which star formation occurs (Dekel et al. 2009b), the time-scale over which it is measured and the virialization of the molecular clouds (Kim, Ostriker & Filippova 2021). Here, we adopt a median value of $\epsilon_{\text{ff}} = 0.015$, similar to other studies (Bouché et al. 2010; Lilly et al. 2013; Dekel & Mandelker 2014). We write the freefall time as $t_{\text{ff}} = \chi t_d$, and we follow Krumholz et al. (2018) to find χ , assuming a vertical force balance between self-gravity, thermal, and non-thermal pressure sources. The resulting proportionality is

$$\chi = \frac{\pi Q_g}{4} \sqrt{\frac{3 f_{g,P} \phi_{\text{mp}}}{2(1+\beta)}}, \quad (19)$$

where $f_{g,P}$ is the fraction of mid-plane pressure which is due to the self-gravity of the gas (Ostriker, McKee & Leroy 2010), and ϕ_{mp} is the excess of pressure due to non-thermal and non-turbulent sources (e.g. magnetic fields and cosmic rays, etc.). The former is about one half for local galaxies (Krumholz et al. 2018; Yu et al. 2021), assuming that the scale height of the gaseous disc is smaller than that of the stars, while it is expected to be larger at high redshifts, where the gas dominates the self-gravity of the disc. Similarly to our treatment of $f_{g,Q}$, we adopt a redshift parameterization that will smoothly change the value of $f_{g,P}$ from 0.7 at high redshifts to 0.5 at low redshifts, as explained in Section 2.6. The value of ϕ_{mp} is in

the range $\phi_{\text{mp}} = 1-2$, and we adopt $\phi_{\text{mp}} = 1.4$ as in Krumholz et al. (2018).

By writing the SFR as in equation (18), we have neglected two corrections. First, is the fact that at low redshifts, star formation typically occurs locally in giant molecular clouds (GMCs), as opposed to high redshifts, where star formation occurs in the giant clumps which are governed by the Toomre instability of the entire disc. The former is called the ‘GMC regime’, and the latter the ‘Toomre regime’ (Krumholz, Dekel & McKee 2012). This transition occurs roughly when $t_{\text{ff}} \epsilon_{\text{ff}}^{-1} \sim 2$ Gyr, above which (namely, in the GMC regime at low redshift where the gas depletion time is long) the relation between the local free fall time to the dynamical in equation (19) is no longer valid. Secondly, we have neglected the fact that at low redshifts, not all of the gas is in a star-forming molecular phase, but rather in a warm atomic phase (Krumholz 2013). As we show below when comparing the results of our model to that of Krumholz et al. (2018), who did take these two effects into account, the resulting difference is less than a factor of 2.

2.3.3 Mass transport

It is well established that gravitationally unstable, turbulent discs exert torques that transport angular momentum outwards, towards larger radii, which by conservation of angular momentum causes mass transport through the disc from the outskirts to the centre. This gradually drains the disc of its gas and partly its stars as well (Gammie 2001; Dekel et al. 2009b; Krumholz & Burkert 2010; Forbes et al. 2012; Goldbaum, Krumholz & Forbes 2015). There are several ways to estimate the transport rate, corresponding to different mechanisms that drive the transport. Here, in order to allow for uncertainties in estimating the transport rate, we focus on two alternative mechanisms: (i) mass transport due to non-axisymmetric torques induced by clump encounters (Dekel et al. 2009b); (ii) mass transport due to turbulent viscosity (Krumholz & Burkert 2010).

The mass transport due to torques induced by clumps was calculated by Dekel et al. (2009b) by estimating the timescale for encounters between clumps and for interactions between clumps and the interclump medium. The transport rate is estimated by

$$\dot{M}_{\text{trans,c}} = \frac{M_g}{t_{\text{evac}}}, \quad (20)$$

where t_{evac} is the disc evacuation time, given by

$$t_{\text{evac}} = \alpha_c^{-1} \frac{24}{\sqrt{3}\pi^2} (1 + \beta) Q_g^2 \delta^{-2} t_d = \alpha_c^{-1} \frac{12}{\sqrt{3}\pi^2} Q_g^4 \left(\frac{V_d}{\sigma_g} \right)^2 t_d. \quad (21)$$

The parameter α_c in equation (21) is the fraction of cold mass in clumps, which will play a key role in regulating the disc in a marginally unstable, turbulent state.

The mass transport in turbulent viscous discs is derived by solving the hydrodynamical equations when taking into account momentum transfer due to turbulent motions and a constant Toomre- Q parameter (Krumholz & Burkert 2010). The resulting equation is an energy conservation equation, which expresses the balance between turbulence dissipation and turbulence driving. We describe in detail below these different mechanisms, and write here the equation with its full generality. Krumholz & Burkert (2010) found that systems reach a steady state, in which the governing equation becomes

$$-\frac{1-\beta}{2} \Omega \mathcal{T} = \mathcal{L} - \mathcal{G}, \quad (22)$$

where \mathcal{T} is the viscous torque due to turbulent viscosity, \mathcal{L} is the turbulence dissipation rate and \mathcal{G} is the turbulence driving term. The

mass transport rate is found by noting that the viscous torque in a steady state is given by $\mathcal{T} = -\dot{M}_{\text{trans,v}} V_d R_d$. The resulting transport rate is

$$\dot{M}_{\text{trans,v}} = \frac{2}{V_d^2 (1 - \beta)} (\mathcal{L} - \mathcal{G}). \quad (23)$$

It is useful to define a dimensionless viscous torque via the equation $\dot{M}_{\text{trans,v}} = \tau M_g / t_d$. Equations (22) and (23) show that the disc self-regulates the turbulent viscous torque in order to sustain energy equilibrium, as discussed in more detail in Section 2.5.

2.4 Turbulence budget

2.4.1 Turbulence dissipation

Due to its dissipative nature, gas tends to lose energy by shocks and radiation. Turbulence in gas, and in particular supersonic turbulence, as observed at both high and low redshifts (Ianjamasimanana et al. 2015; Übler et al. 2019), quickly cascades to smaller scales until reaching the viscous scale at which the turbulence energy dissipates. The turbulence dissipation time-scale is of the order of the crossing time of the largest driving scale of the turbulence (Mac Low et al. 1998). To be explained below, we parametrize the dissipation time-scale as

$$t_{\text{diss}} = \frac{\gamma_{\text{diss}}}{\sqrt{2(1+\beta)}} Q_g^n t_d. \quad (24)$$

If the driving scale of turbulence is the scale height of the disc, then $n = 1$, and $\gamma_{\text{diss}} = 1/2$ at low redshifts, while $\gamma_{\text{diss}} \approx 1$ at high redshifts. To see this, we first adopt the approximation of Forbes et al. (2012) for the scale height

$$H_g = \frac{\sigma_g^2}{\pi G (\Sigma_g + (\sigma_g/\sigma_s) \Sigma_s)} = \frac{\sigma_g^2}{\pi G \Sigma_g \phi_Q}, \quad (25)$$

where $\phi_Q = 1 + Q_g/Q_s$. This parametrization of the gas scale height interpolates between the limit of the scale height being dominated by the self-gravity of the gas ($\sigma_g/\sigma_s \ll 1$) and the limit of the scale height being determined by the gravity of both stars and gas ($\sigma_g \approx \sigma_s$). At low redshifts, $\phi_Q = 2$ (Krumholz et al. 2018; Yu et al. 2021), while it should be smaller for high redshift galaxies, as gas dominates the instability. This implies that $\gamma_{\text{diss}} = 1/2 - 1$.

Alternatively, if the typical clump scale is the relevant scale for the turbulence dissipation time, then $\gamma_{\text{diss}} = 3\pi/8 \sim 1$, and $n = -1$ (Dekel et al. 2009b). For $Q \sim 1$, this is a factor of 2 difference with respect to the disc scale height, however for thick discs this can be larger by a factor of a few.

In this study, we focus on galaxies with $Q \approx 1$, and in order to compare with Krumholz et al. (2018), we focus on dissipation on the disc scale height, remembering that the two dissipation time-scales do not differ by much for $Q = 1$. Thus, we adopt $n = 1$, while evolution of γ_{diss} with redshift is outlined in Section 2.6. We note that an equilibrium value of $Q \sim 0.67$ is possible for thick discs (Goldreich & Lynden-Bell 1965), appropriate at high redshifts. In Section 3.2.3, we discuss how a choice of $Q = 0.67$ will affect our results.

2.4.2 Gravitational driven turbulence

The time-scale for encounters between clumps to generate turbulence in the disc of level σ_g^2 was computed by Dekel et al. (2009b), by considering the gravitational cross-section for two body interactions

between bound clumps. The timescale was found to be

$$t_{\text{enc,turb}} = \frac{36}{\sqrt{3}\pi^2} \alpha_c^{-1} Q_g^4 t_d. \quad (26)$$

Alternatively, in the picture of gravitational turbulence driving in a turbulent viscous disc, the relevant time-scale for transport to generate a turbulence of level σ_g^2 is simply

$$t_{\text{trans,turb}} = 3 \frac{M_g \sigma_g^2}{\dot{M}_{\text{trans,v}} V_d^2 (1 - \beta)}. \quad (27)$$

Once we impose the energy equilibrium, we will be able to find $\dot{M}_{\text{trans,v}}$ and estimate the time-scale for this type of transport to generate turbulence (see Section 2.5).

2.4.3 Supernova feedback turbulence

In order to estimate the amount of turbulence generated by supernovae, we follow several authors (Dekel & Silk 1986; Matzner 2002; Faucher-Giguère et al. 2013; Hayward & Hopkins 2017; Krumholz et al. 2018). During the final phase of the supernova, after the adiabatic phase is over, a thin shell expands while conserving momentum. The shell then fades into with the surrounding medium when its velocity approaches the velocity dispersion of the surrounding medium (for a summary, see Draine 2011; Dekel et al. 2019). If the initial momentum is p , then when the shell mixes with the medium, it adds energy of order $\sim p \sigma_g$ (Matzner 2002). For an SFR of \dot{M}_{SF} and mean momentum injected per stellar mass formed, appropriately averaged over an initial mass function, of $(p/m)_*$ (Ostriker & Shetty 2011), the energy injection rate is

$$\dot{E}_{\text{SN,turb}} = \dot{M}_{\text{SF}} \left(\frac{p}{m} \right)_* \sigma_g. \quad (28)$$

From this expression, we can estimate the time-scale for supernova feedback to drive turbulence of level σ_g^2 , plugging the SFR from equation (18) and χ from equation (19),

$$t_{\text{turb,SN}} = \frac{3}{2} \epsilon_{\text{ff}}^{-1} \chi \left(\frac{p}{m} \right)_* \sigma_g t_d. \quad (29)$$

Simulations of non-clustered supernova suggest that $(p/m)_* \approx 3000 \text{ km s}^{-1}$ (Ostriker et al. 2010; Hayward & Hopkins 2017), which we adopt as our fiducial value. The clustering of supernovae may have a significant effect on $(p/m)_*$, though it is fairly uncertain, ranging from a factor of 2 above our fiducial value to slightly below it (Walch & Naab 2015; Gentry et al. 2017, 2019; Kim, Ostriker & Raileanu 2017). A factor of 2 increase in the adopted $(p/m)_*$ will double the contribution of feedback to the driving of turbulence. This will affect our qualitative results only for discs in haloes of mass $\sim 10^{12} M_\odot$ at $z \sim 2-3$.

2.4.4 Accretion driven turbulence

The main form in which fresh gas is supplied to massive galaxies at high redshifts, $M_v > 10^{12} M_\odot$ and $z > 1$, is in narrow cold streams (Dekel & Birnboim 2006; Dekel et al. 2009a; Danovich et al. 2012). Long-lived discs are expected to exist in this mass range at all redshifts (Dekel et al. 2022). When the streams hit the disc, some of the kinetic energy they carry is converted into heat in an accretion shock, while the rest remains in kinetic form, eventually turning into turbulence. The fraction of the initial kinetic energy that contributes to turbulence depends strongly on the density contrast between the colliding flows of streams and discs (Klessen & Hennebelle 2010).

A series of papers studying simulations of idealized, cosmologically motivated streams (Aung et al. 2019; Mandelker et al. 2020a) as well as analytic work (Mandelker et al. 2018, 2020b) suggest that the streams should fragment and form dense clumps before reaching the galaxy. Upon impact, such dense clumps have a high efficiency in converting their bulk kinetic energy into turbulent kinetic energy within the disc (Klessen & Hennebelle 2010).

Streams are expected to be more efficient at driving turbulence than the more spherical accretion typical at low redshifts. Even if the streams were smooth (i.e. not clumpy), they are much denser and concentrated than a spherical accretion, for a given accretion rate. We therefore expect turbulence driven by clumpy accretion to be most relevant at high redshifts.

In this paper, we want to determine the mass and redshift range within which this energy source can be important for driving turbulence, and leave the detailed modelling of how clumpy accretion drives turbulence in rotating discs to future work. We therefore, similarly to Genel et al. (2012a), parametrize the rate at which streams provide energy to turbulence as some fraction ξ_a of the total kinetic energy carried by the streams, namely

$$\dot{E}_{\text{acc,turb}} = \frac{1}{2} \xi_a \dot{M}_{\text{g,acc}} V_d^2, \quad (30)$$

where we have approximated the instreaming velocity as the virial velocity, which is comparable to the rotation velocity of the disc (see Section 2.1). According to Klessen & Hennebelle (2010), $\xi_a \sim \Delta$, where Δ is the density contrast between the accreting material, either smooth stream material or clumps within the streams, and the disc. Using equations (30) and (6), we derive a time-scale for accretion to drive turbulence of level σ_g^2

$$t_{\text{turb,acc}} = \frac{3\sqrt{2}(1+\beta)}{\xi_a G \dot{M}_{\text{g,acc}} Q_g} \sigma_g^3 t_d. \quad (31)$$

Below, we explore the effect of three constant values for the conversion efficiency, representing low level ($\xi = 0.3$), moderate level ($\xi_a = 0.6$) and maximal ($\xi_a = 1$) efficiencies, and we will see that they result in qualitatively different outcomes. The chosen values for ξ_a are motivated by crude earlier studies of clumpiness in streams (Dekel et al. 2009b; Aung et al. 2019), to be properly evaluated as a function of halo mass and redshift in future work. We will also explore the effects of a time-varying conversion efficiency parameter on the evolution of the turbulence in discs.

2.5 Turbulent energy equilibrium

We now apply the second main assumption of the model, namely that turbulence is in an equilibrium state, in which the dissipation of turbulence is balanced by the three drivers. This condition can be written as

$$\dot{E}_{\text{turb,trans}} + \dot{E}_{\text{turb,SN}} + \dot{E}_{\text{turb,acc}} = \dot{E}_{\text{diss}}. \quad (32)$$

Equation (32) can be written for both pictures of transport, resulting in

$$\begin{aligned} \frac{\sqrt{3}\pi^2}{36} \alpha_c Q_g^{-4} + \frac{2}{3} \epsilon_{\text{ff}} \left(\frac{p}{m} \right)_* \chi^{-1} \sigma_g^{-1} + \frac{1}{3} \xi_a \frac{G \dot{M}_{\text{g,acc}} Q_g}{\sqrt{2(1+\beta)}} \sigma_g^{-3} \\ = \sqrt{2(1+\beta)} \gamma_{\text{diss}}^{-1} Q_g^{-n}, \end{aligned} \quad (33a)$$

$$\begin{aligned} \tau \frac{1-\beta}{2} V_d^2 + \left(\frac{p}{m} \right)_* \epsilon_{\text{ff}} \chi^{-1} \sigma_g + \frac{1}{2} \xi_a \frac{\dot{M}_{\text{g,acc}} t_d}{M_g} V_d^2 \\ = \frac{3}{2} \sqrt{2(1+\beta)} \gamma_{\text{diss}}^{-1} Q_g^{-n} \sigma_g^2, \end{aligned} \quad (33b)$$

Equation (33a) is the energy balance equation when the transport is governed by clump encounters, while equation (33b) is the energy balance equation when the transport is governed by viscous torques.

By imposing the equilibrium turbulent state in equations (33a) and (33b), the disc self-regulates the gravitational-driven turbulence by adjusting the relevant parameter in each picture. If we assume a fixed $Q = f_{g,Q} Q_g$, and that ϵ_{ff} and $(p/m)_*$ are constant in equations (33a) and (33b), then in the picture of clump encounters, the disc self-regulates α_c , the instantaneous fraction of mass in clumps (or equivalently, the number of clumps) in the galaxy, and in the picture of turbulent viscosity, the disc self-regulates the viscous torque τ . By isolating α_c from equation (33a) and τ from equation (33b), we can write

$$\alpha_c = \frac{36 Q_g^4}{\sqrt{3\pi}^2} F(\sigma_g), \quad (34a)$$

$$\tau = \frac{3}{1-\beta} \left(\frac{\sigma_g}{V_d} \right)^2 F(\sigma_g), \quad (34b)$$

where

$$F(\sigma_g) = \sqrt{2(1+\beta)} \gamma_{diss}^{-1} \left(\frac{Q}{f_{g,Q}} \right)^{-n} \left(1 - \frac{\sigma_{SN}}{\sigma_g} - \left(\frac{\sigma_{acc}}{\sigma_g} \right)^3 \right). \quad (35)$$

By writing equations (34a) and (34b) in this form, we have defined two velocities, σ_{SN} and σ_{acc} , which are given by

$$\sigma_{SN} = \frac{2}{3\sqrt{2(1+\beta)}} \left(\frac{p}{m} \right)_* \epsilon_{ff} \chi^{-1} \gamma_{diss} \left(\frac{Q}{f_{g,Q}} \right)^n, \quad (36)$$

which is σ_{sf} as defined in Krumholz et al. (2018), and

$$\sigma_{acc} = \left(\frac{\xi_a G \dot{M}_{g,acc} Q^{1+n} \gamma_{diss}}{6(1+\beta) f_{g,Q}^{1+n}} \right)^{1/3}. \quad (37)$$

The ratio σ_{SN}/σ_g represents the fraction of turbulent energy that is sustained by supernova feedback, while the ratio $(\sigma_{acc}/\sigma_g)^3$ represents the fraction of turbulent energy that is sustained by accretion.⁵ The velocities σ_{SN} and σ_{acc} depend on the redshift-dependent parameters $f_{g,Q}$, $f_{g,P}$, and γ_{diss} . We discuss the adopted redshift parametrization of these parameters in Section 2.6, and quote typical values for σ_{SN} and σ_{acc} there.

Equation (35) suggests that there is a critical gas velocity dispersion, $\sigma_c(M_h, z)$, below which equations (34a) and (34b) no longer hold. Below this critical velocity dispersion, neither of the sources of mass transport discussed here produce turbulence. This means that at this velocity dispersion, mass transport is no longer needed by the disc to sustain turbulence, and the transport therefore shuts-off. At this point, in order to maintain energy equilibrium, as in equations (33a) and (33b), the disc would have to self-regulate a different parameter. As discussed in Section 3.3, this can be either Q or ϵ_{ff} . σ_c can be calculated analytically, and this is done in Appendix A. It is in the range $\sigma_c \approx 20\text{--}100 \text{ km s}^{-1}$, depending on ξ_a , mass and redshift.

Using the self-regulated parameters, we can write expressions for the mass transport rates in the steady turbulent state. In the picture of clump encounters, the mass transport rate is

$$\dot{M}_{trans,c} = 3F(\sigma_g) \left(\frac{\sigma_g}{V_d} \right)^2 \frac{M_g}{t_d}, \quad (38)$$

⁵One might expect these fractions to be scale as σ_g^2 ; however, under the assumption of constant Q , supernova driving scales as σ_g^2 , while accretion driving does not scale with σ_g . Since the turbulent energy scales as σ_g^3 under constant Q , we get the linear and cubic scalings as mentioned in the text.

and in the picture of turbulence viscosity,

$$\dot{M}_{trans,v} = \frac{3}{1-\beta} F(\sigma_g) \left(\frac{\sigma_g}{V_d} \right)^2 \frac{M_g}{t_d}. \quad (39)$$

We can see that up to a factor that depends on the rotation curve, which can introduce a factor of at most two, the resulting mass transport rates are identical. Using this surprising result, in the remainder of the paper we assume a flat rotation curve, $\beta = 0$, ignore the different origins of mass transport, and define

$$\dot{M}_{trans} = 3F(\sigma_g) \left(\frac{\sigma_g}{V_d} \right)^2 \frac{M_g}{t_d} \quad (40)$$

as the sink term in equation (43), which is related to the mass transport within the disc. While the expressions for the transport rates are the same, one must remember that their physical origin is different and that the physical meaning of the corresponding self-regulated parameter, α_c and τ , is different in each picture.

2.6 Redshift evolution of $f_{g,Q}$, $f_{g,P}$, and γ_{diss}

The parameters $f_{g,Q}$, $f_{g,P}$, and γ_{diss} may vary with time. At high redshifts, discs are gas rich and the dynamics is more influenced by the gaseous component. At low redshifts, the discs are gas-poor, in which both the cold gas and the hotter stars contribute the dynamics. The value of ϕ_Q , which sets the value of γ_{diss} (see equation 25), also varies with time, as the height of the gaseous disc is thicker at high redshifts than low redshifts.

According to Genzel et al. (2015), for star-forming galaxies on the main sequence in the redshift range $z = 0\text{--}4$, the typical gas fraction varies as

$$f_g(z) \approx \frac{0.06(1+z)^{2.7}}{1 + 0.06(1+z)^{2.7}}. \quad (41)$$

In order to inspect the qualitative effects of the variation in time of the aforementioned parameters, we make the crude assumption that $f_{g,Q}$, $f_{g,P}$, and ϕ_Q evolve as a function of the gas fraction,

$$f_{g,Q} = f_{g,P} = \phi_Q^{-1} = a \cdot f_g(z) + b, \quad (42)$$

where $a = 0.39$ and $b = 0.38$ are chosen so that the values will vary between 0.7 at $z = 4\text{--}0.5$ at $z = 0$. These values are chosen based on the gas fraction of observed galaxies (Genzel et al. 2015; Tacconi et al. 2018) and velocity dispersions of gas and stars from simulations (Pillepich et al. 2019). We defer to future work a more detailed study, using the redshift dependence as derived from simulations.

Under this parametrization, σ_{SN} evolves from $\sigma_{SN} \approx 13 \text{ km s}^{-1}$ at $z = 2$ to $\sigma_{SN} \approx 11 \text{ km s}^{-1}$ at $z = 0$. For $\xi_a = 0.6$, $\sigma_{acc} \approx 20, 36, 61 \text{ km s}^{-1}$ for galaxies in haloes of mass $\log(M_h/M_\odot) = 11.5, 12, 13$ at $z = 2$, and $\sigma_{acc} \approx 8, 12, 24 \text{ km s}^{-1}$ for galaxies in haloes of the same masses at $z = 0$.

3 RESULTS

Equipped with all of the relevant mass sources and sinks, we can write a conservation equation for the gas mass in the disc

$$\dot{M}_g = \dot{M}_{g,acc} - \dot{M}_{trans} - \dot{M}_{SF}. \quad (43)$$

We integrate equation (43) numerically,⁶ using equations (16), (18), and (40) for the relevant quantities. The mass transported down the

⁶We use the *solve_ivp* method in the *SciPy* package (Virtanen et al. 2020), using the BDF method.

Table 1. Summary of typical values of model parameters.

Parameter	Value	Definition	Meaning
c	10	Section 2.1	Halo concentration parameter
γ	1.4	Section 2.1	V_d in units of V_{\max}
β	0	Section 2.2	Rotation curve log slope
Q	1	Equation (8)	Two component Toomre- Q
$f_{g,Q}$	0.5–0.7	Equation (10)	Effective gas fraction for Toomre- Q
ϵ_{ff}	0.015	Equation (18)	Star formation efficiency
χ	1–1.5	Equation (19)	Free-fall to dynamical time ratio
$f_{g,P}$	0.5–0.7	Section 2.3.2	Fraction of self-gravity due to gas
ϕ_{mp}	1.4	Section 2.3.2	Total-to-turbulent pressure ratio
α_c^a	0–1	Section 2.3.3	Mass fraction in clumps
τ^a	0– 10^{-2}	Section 2.3.3	Dimensionless viscous torque
γ_{diss}	0.5	Equation (24)	Turbulence dissipation parameter
ϕ_Q	1.5–2	Equation (25)	Stellar to gaseous Toomre- Q ratio
$(\frac{p}{m})_*$	3000 km s $^{-1}$	Equation (28)	SN momentum per stellar mass
ξ_a	0.3–1	Equation (30)	Accretion-driven turbulence efficiency
σ_{SN}	11–14 km s $^{-1}$	Equation (36)	Dispersion sustained by SN
σ_{acc}	8–54 km s $^{-1}$	Equation (37)	Dispersion sustained by accretion

Note. ^aThe values of α_c and τ depend on the gas velocity dispersion, halo mass, redshift, and all of the model parameters (see Section 2.5).

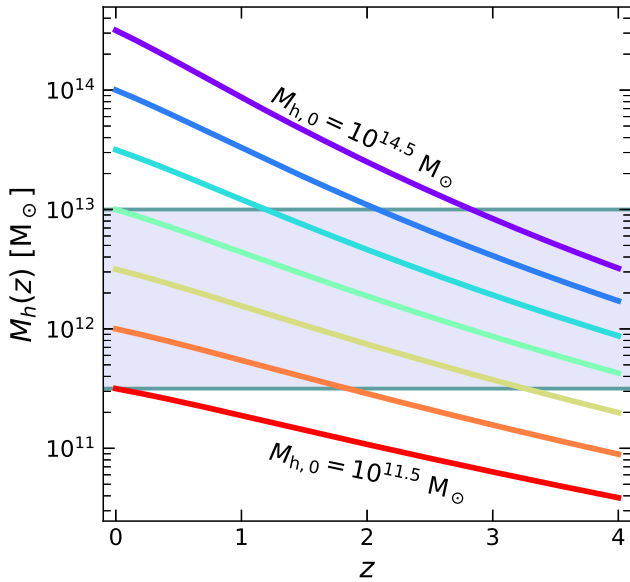


Figure 1. The evolution of the mass of haloes hosting discs, as a result of integrating the average accretion rate (equation 12). Shown are the evolution tracks for haloes with mass at redshift $z = 0$ between $10^{11.5} M_\odot$ and $10^{14.5} M_\odot$, in increments of 0.5 dex. The shaded band indicates the time the galaxy in each halo is expected to be a disc (see the text).

potential well ends up in a bulge, which is assumed to be a distinct component of the galaxy, and we neglect its dynamical implications (see Section 4 for discussion).

Since equation (43) is driven by the external accretion, which evolves in time slower than other relevant timescales, the solution will eventually converge to a unique solution that depends only on the model parameters and the halo mass at redshift zero (due to the model cosmological accretion rate), and not on the particular choice of initial conditions. This characteristic of the bathtub model has already been investigated in other bathtub studies (Bouché et al. 2010; Lilly et al. 2013; Cacciato et al. 2012; Genel et al. 2012a; Dekel & Mandelker 2014).

Once we obtain the evolution of the gas mass, we can use the fact that $\sigma_g \propto Q_g M_g$, along with the marginal Toomre instability condition, to solve for $\sigma_g(t)$. This will allow us to closely inspect how the turbulence is supported against dissipation. The model parameters and their corresponding fiducial values are summarized in Table 1.

We follow the evolution of galaxies inside haloes throughout their lifetimes. The evolution of the discs' host halo masses are shown in Fig. 1. However, the galaxies are not expected to be discs at all times. Dekel et al. (2020a) found that discs tend not to survive for more than an orbital time if they reside in haloes of mass $M_h < 10^{11.5} M_\odot$, roughly independent of redshift. Also, central galaxies inside clusters, i.e. in haloes with $\log M_h \gtrsim 13$, are expected to be ellipticals rather than discs (Dressler 1980). Thus, when presenting our results, we highlight the regions of the parameter space in which we expect to have the galaxies containing extended discs, namely the mass range $11.5 \leq \log M_h(z) \leq 13$ at the given redshift, where our model is applicable. Due to the nature of bathtub models, the assumption that our model is applicable only from a certain point in time does not introduce a significant error, as the solution is not very sensitive to the initial conditions. We find that it takes between one to two disc orbital times, for all initial redshifts, for the solutions to converge to a unique solution, regardless of the initial conditions. We analyse the convergence in Appendix B.

Throughout this section, unless stated otherwise, the quantity $M_{h,0}$ refers to the galaxy's host halo mass at $z = 0$. This mass uniquely determines the entire evolution of the halo mass, assuming the average accretion history given by equation (12).

3.1 Feedback and transport only

First, we examine the main driver of turbulence in discs with no accretion, namely $\xi_a = 0$, which means $\sigma_{\text{acc}} = 0$. In this case, the function $F(\sigma_g)$ that goes into equation (43) is

$$F(\sigma_g) \propto 1 - \frac{\sigma_{\text{SN}}}{\sigma_g}. \quad (44)$$

Fig. 2 shows the evolution of the velocity dispersion in the disc, normalized by σ_{SN} . As discussed in Section 2.5, the ratio $f_{\text{turb, SN}} = \sigma_{\text{SN}}/\sigma_g$ is the fraction of turbulence dissipation that is supported by supernova feedback. So, when $f_{\text{turb, SN}} > 1/2$, most of the turbulence is supported by supernova feedback, while when $f_{\text{turb, SN}} < 1/2$, most of the turbulence is supported by transport.

The region of the parameter space in which galaxies are expected to be discs is bounded by the two solid curves. We can see from Fig. 2 that at high redshifts, roughly $z > 1.5$, the turbulence in all discs is expected to be supported by transport. At lower redshifts, turbulence in discs that reside in haloes of mass $\log M_{h,0} \lesssim 12$ becomes mainly supported by supernova feedback.

Fig. 2 also shows the evolution of the velocity dispersion from Krumholz et al. (2018) for a halo with a mass of $10^{12} M_\odot$ at redshift zero. Note that this result assumes $f_{g,Q} = f_{g,P} = \phi_Q^{-1} = 1/2$ with no redshift evolution. In that case $\sigma_{\text{SN}} = 11 \text{ km s}^{-1}$ at all redshifts.⁷ The results of Krumholz et al. (2018) suggest that discs that reside in haloes of mass $\log M_{h,0} = 12$ have turbulence that is primarily supported by transport at all redshifts until $z \sim 0.5$. However, such galaxies are not expected to be discs until $z \sim 2$, and our model predicts that these galaxies, when they become discs, have turbulence that is supported by supernova feedback and transport with comparable power, rather than predominantly supported by transport. We note that changes to the redshift parametrization of $f_{g,Q}$, $f_{g,P}$ and ϕ_Q as discussed in Section 2.6 can change the conclusions only quantitatively, and we still expect that qualitatively, discs with

⁷In Krumholz et al. (2018), σ_{SN} has a weak redshift dependence at $z \lesssim 0.5$, that comes from the evolution of the fraction of gas that is in a molecular phase rather than an atomic phase.

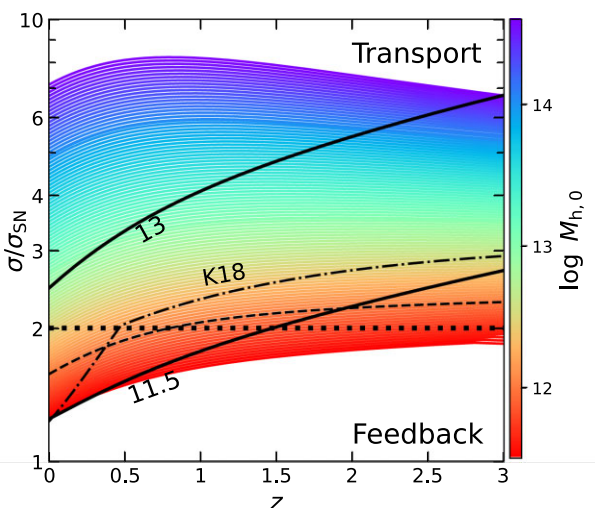


Figure 2. The evolution of gas-disc velocity dispersion, with respect to the contribution of feedback, as predicted by our model, when $\xi_a = 0$, i.e. when accretion is assumed to provide no energy to the turbulence. Here, we consider feedback and transport as the drivers of turbulence, and turn-off the contribution of accretion. The figure shows the evolution for different galaxies, distinguished by the host halo mass at $z = 0$, indicated by the different colours. The dashed line is a sample evolution of a galaxy that resides in a halo with $\log M_{\text{h},0} = 12$ today. The dashed dotted line shows the evolution of a galaxy inside a halo of mass $\log M_{\text{h},0} = 12$ from Krumholz et al. (2018). The thick solid black lines bound the region in which our model predicts the galaxies to be discs, namely when the instantaneous halo mass is in the range $11.5 < \log M_{\text{h},0} < 13$. Note that this region is applicable only to the results of our model and not to that of Krumholz et al. (2018). Above the thick dotted line, the turbulence in galaxies is mainly supported by transport rather than feedback. We learn that turbulence in all discs is sustained mainly by transport at $z \gtrsim 1.5$. At lower redshifts, $z \lesssim 1$ turbulence in discs that reside in haloes with $\log M_{\text{h},0} \lesssim 12$ becomes mainly supported by feedback.

$\log M_{\text{h},0} \lesssim 12$ will have turbulence that is primarily supported by supernova feedback.

We note two differences between our approach to the solution and the one used by Krumholz et al. (2018), beyond the parameterization of $f_{\text{g},\text{Q}}$, $f_{\text{g},\text{P}}$ and ϕ_{Q} . First, Krumholz et al. (2018) produced the cosmological evolution of the velocity dispersion by solving the equation $\dot{M}_{\text{g},\text{acc}} = \dot{M}_{\text{trans}}$ for σ_{g} . The left-hand side introduces the cosmological dependence. The reasoning behind this approach is the fact that the bathtub models for galaxy evolution, driven by an external accretion, reach a steady-state solution that is dictated by this external source, as discussed above. This steady-state solution can be approximated analytically by imposing $\dot{M}_{\text{g}} = 0$ in equation (43); however, Krumholz et al. (2018) neglected the SFR as a contribution to the sink term.

Secondly, we have not included either the GMC regime of star formation nor the fact that, at low redshifts, the fraction of gas in the star forming phase is lower. Both of these effects change the mass transport rate, and hence affect the evolution of σ_{g} . These two corrections introduce a small difference of less than a factor of two, and do not introduce any qualitative difference.

3.2 Adding accretion

We now investigate the dominant source of turbulence when adding accretion as a third driver. We will first assume a constant value for ξ_a , and investigate the evolution assuming three values for the efficiency of converting accreted kinetic energy into turbulence (as defined in

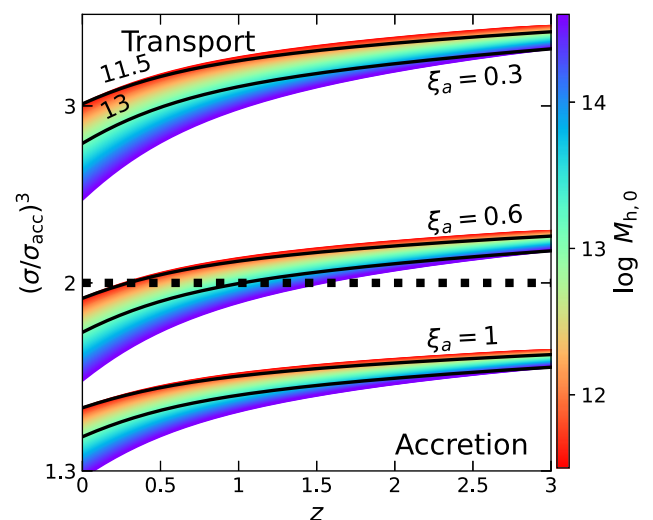


Figure 3. The evolution of $(\sigma_{\text{g}}/\sigma_{\text{acc}})^3$, which corresponds to the fraction of turbulent dissipation balanced by accretion-driven turbulence (see Section 2.5). This figure is analogous to Fig. 2. When $(\sigma_{\text{g}}/\sigma_{\text{acc}})^3 > 2$, mass transport is the primary driver of turbulence. Shown are three sets of evolutionary curves, resulting from a different accretion conversion efficiency: low ($\xi_a = 0.3$; top), moderate ($\xi_a = 0.6$; middle), and maximal ($\xi_a = 1$; bottom). The solid black curves bound the regions in each set where the galaxies are expected to be discs. We learn that for a low accretion conversion efficiency, turbulence in all discs at all times is supported mainly by transport. For a moderate efficiency, turbulence in discs is supported by accretion and transport with comparable power. For a maximal efficiency, all discs are supported mainly by accretion, at all times.

Section 2.4.4): low efficiency, $\xi_a = 0.3$, namely a third of the kinetic energy provided by streams is converted into turbulence, moderate, $\xi_a = 0.6$, and maximal efficiency, $\xi_a = 1$. In Section 3.2.3, we will investigate very crudely the evolution assuming a redshift-dependent conversion efficiency.

3.2.1 Accretion and transport only

First, we consider models with only transport and accretion, namely no feedback ($\sigma_{\text{SN}} = 0$). The function $F(\sigma_{\text{g}})$ that goes into the mass transport rate in equation (43) is in this case

$$F(\sigma_{\text{g}}) \propto 1 - \left(\frac{\sigma_{\text{acc}}}{\sigma_{\text{g}}} \right)^3. \quad (45)$$

Analogously to the previous section, we will focus on the ratio $f_{\text{turb,acc}} = (\sigma_{\text{acc}}/\sigma_{\text{g}})^3$, which is the fraction of turbulence dissipation that is balanced by accretion driven turbulence.

Fig. 3 shows the evolution of the cube of σ_{g} normalized by σ_{acc} , for three levels of accretion efficiency, namely $\xi_a = 0.3$, $\xi_a = 0.6$, and $\xi_a = 1$. Unlike the case with feedback and transport only (Section 3.1), now the normalization, namely σ_{acc} , is mass dependent. We see from Fig. 3 a qualitative difference between the three levels of conversion efficiency. For $\xi_a = 0.3$, we see that all discs have $(\sigma_{\text{g}}/\sigma_{\text{acc}})^3 \sim 3$, namely more than 60 per cent of their turbulence is supported by transport, for all masses and redshifts. For $\xi_a = 0.6$, all discs have $(\sigma_{\text{g}}/\sigma_{\text{acc}})^3 \sim 2$, namely comparable share between transport and accretion. Lastly, we see from Fig. 3 that for $\xi_a = 1$, namely maximal conversion efficiency of accretion kinetic energy to turbulence, accretion is the main driver of turbulence in discs of all masses at all times.

We notice weak variations of $(\sigma_g/\sigma_{\text{acc}})^3$ with halo mass, with the ratio decreasing with increasing $\log M_{h,0}$. This results from the dependence of the gas accretion rate on halo mass. In our model, this dependency comes from the cosmological accretion rate, as defined in equation (12), and the penetration parameter (equation 17). A constant penetration, as well as a weaker mass dependence of the accretion rate, might further weaken and even invert this gradient. For most reasonable choices, the dependence of $f_{\text{turb, acc}}$ on M_h is very weak, and the parametrizations of \dot{M}_h as a function of M_h will not qualitatively affect our conclusions.

The results of this section indicate that turbulence driven by accretion can have a significant role in balancing turbulence dissipation. We do note, however, that while different conversion efficiencies ξ_a give qualitatively different results, the fraction of turbulence dissipation that is balanced by accretion is always in the range of 0.4–0.7, namely its contribution to the total turbulent energy budget is comparable to that of transport. This is unlike the model with feedback and transport, ignoring accretion (Section 3.1), where massive haloes were predominantly supported by transport.

3.2.2 Accretion, feedback, and transport – constant ξ_a

Finally, we consider the case where all three drivers of turbulence are active. First, we consider three constant efficiency parameters for converting accretion kinetic energy into turbulence. In this case, the function $F(\sigma_g)$ that goes into equation (43) is given by equation (35). In order to find out which is the primary driver of turbulence as a function of halo mass and redshift, we compare the energy injection rates of each driver, as derived in Section 2.4.

Fig. 4 presents our main result. It shows evolutionary tracks of galaxies that reside in haloes of different masses at $z = 0$, from $\log M_{h,0} = 11.5$ to $\log M_{h,0} = 14.5$, in increments of 0.5 dex. The tracks are plotted in the plane of $\log(T/A)$ versus $\log(A/F)$, where T , A , and F are the turbulent energy injection rates of transport, accretion, and feedback, respectively. The plane has been divided into several regions of two types – the thick lines separate regions of different primary drivers, while the thin lines separate regions of a given primary driver that have different secondary drivers. The shaded region marks the mass range where galaxies are expected to be discs at each redshift, as explained above. We learn from Fig. 4 the following:

(i) For a low accretion conversion efficiency, $\xi_a = 0.3$ (top panel), discs in haloes with $\log M_{h,0} \lesssim 12$ today, are predominantly supported by feedback at all times, while haloes with $\log M_{h,0} \gtrsim 13$ are mainly supported by transport at all times. Turbulence support in intermediate-mass galaxies transitions from being dominated by transport at high redshifts to feedback at low redshifts. This result is qualitatively similar to the result of the model with feedback and transport only (Section 3.1). Furthermore, for $\log M_{h,0} < 12$, the secondary driver is transport, with the accretion less important. For massive discs, with $\log M_{h,0} \gtrsim 13.5$, corresponding to $M_h(z=2) \approx 10^{12.7} M_\odot$, accretion becomes the secondary driver, with a higher energy injection rate than feedback.

(ii) For a moderate accretion conversion efficiency, $\xi_a = 0.6$ (middle panel), low-mass discs, residing in haloes with $\log M_{h,0} \lesssim 12$, have their turbulence primarily driven by feedback throughout their lifetimes, similar to the $\xi_a = 0.3$ case and the feedback and transport only model. The secondary driver in these discs is transport until $z \sim 0.5$, with accretion becoming the secondary driver at later times. However, this difference is small (see below).

Discs that reside in haloes with $\log M_{h,0} \gtrsim 12$ at $z = 0$ have their turbulence driven primarily by transport at high redshifts. However, we note that the transport energy injection rate and accretion energy injection rate for all of the galaxies with $\log M_{h,0} \gtrsim 12.5$ differ by no more than ~ 40 per cent at all times, meaning that both of these mechanisms contribute comparable power to sustaining turbulence. In particular, at redshift $z \sim 2$, the difference between the energy injection rates from transport and accretion for galaxies in haloes with $12.5 \lesssim \log M_{h,0} \lesssim 14$ is less than ~ 20 per cent. We conclude that, for a moderate level of accretion conversion efficiency, galaxies residing in haloes within the mass range of $12.5 \lesssim \log M_{h,0} \lesssim 14$, corresponding to $\sim 10^{12-13} M_\odot$ at $z = 2$, have their turbulence supported by transport and accretion with comparable power.

We note that our model predicts that, at lower redshifts, the energy injection rates from accretion becomes larger in some of the galaxies. However, as discussed below, a large value of $\xi_a = 0.6$ is not expected at these late times.

(iii) For a maximal accretion conversion efficiency, $\xi_a = 1$ (bottom panel), discs in haloes with $\log M_{h,0} \lesssim 12$ still have their turbulence primarily driven by feedback throughout their lifetimes. In these discs, accretion is the secondary driver of turbulence, being more important than transport. More massive discs, residing in haloes with $\log M_{h,0} \gtrsim 12$, have their turbulence primarily driven by accretion throughout their lifetimes. In particular, at $z \sim 4$, all discs have their turbulence primarily supported by accretion, with energy injection ~ 40 per cent larger than from transport. Feedback is the secondary driver in discs that reside in haloes with $12.5 \lesssim \log M_{h,0} \lesssim 13$, while transport is the secondary driver for discs in haloes with $\log M_{h,0} \gtrsim 13$. As discussed below, such a high value of $\xi_a = 1$ is expected at very high redshifts, but is not expected to be valid at low redshifts.

For all of the cases studied above, we learn that when including all three drivers together, accretion can have a significant role in supporting the turbulence in massive discs, either as a primary driver, when $\xi_a = 1$, as an equal power contributor, when $\xi_a = 0.6$, or as a secondary driver, when $\xi_a = 0.3$.

3.2.3 Accretion, feedback, and transport – time varying ξ_a

The biggest uncertainty in our model is the conversion efficiency from accretion kinetic energy to turbulence in the disc, namely ξ_a . The value of this parameter represents the density contrast between the accreting stream and the disc, which is governed both by the ambient density and cross-section of the stream, and by stream clumpiness. Since the value of this parameter is yet to be studied in a cosmological context, we treat it here in an ad hoc way for a qualitative study.

In particular, in the previous section we have treated it as fixed in time, but it is expected to decline in time following the stream clumpiness. For example, Mandelker et al. (2018) have studied the fragmentation of cosmological streams, and noted that the mass fraction in very high density clumps that may serve as globular-cluster progenitors is a strong function of the halo mass and redshift, with streams being more clumpy at higher redshifts. We defer to future work a detailed analysis of the conversion efficiency of stream energy to disc turbulence as a function of halo mass and redshift. Here, for the purpose of a qualitative understanding, we limit ourselves to a potentially representative example of the time dependence of ξ_a , where we assume ad hoc

$$\xi_a(z) = 0.2(1+z), \quad z \leq 4. \quad (46)$$

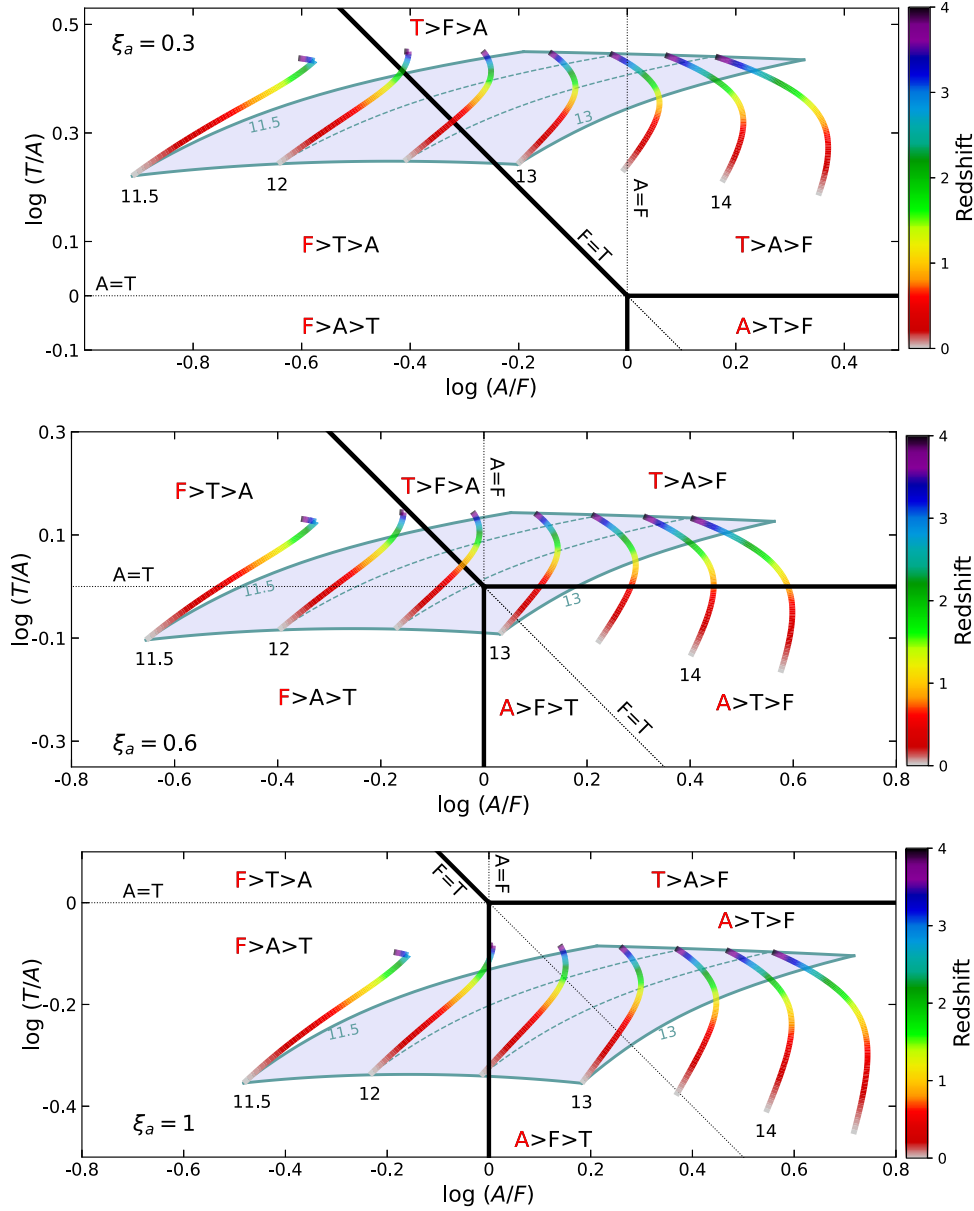


Figure 4. The relative roles of the three processes that drive disc turbulence as a function of halo mass and redshift. A , F , and T refer to accretion, feedback, and transport. The axes refer to the ratios of energy injection rates, T/A and A/F . Shown are evolutionary tracks of discs that reside in haloes of different masses at $z = 0$, $\log M_{h,0} = 11.5 - 14.5$ in increments of 0.5 dex. The colour indicates the redshift along the track at the particular position in the plane. Thick solid black lines divide the plane to three regions with different dominant drivers, marked by a red letter. The thin lines separate regions with a given primary driver based on their secondary driver. The grey region bounded by the turquoise frame marks the mass range, where the galaxies are expected to be discs at the given redshift, $\log M_h(z) = 11.5 - 13$. The dashed lines are lines of constant halo mass of $10^{12} M_\odot$ and $10^{12.5} M_\odot$. The different panels show the evolution of discs for different values of the accretion conversion efficiency – $\xi_a = 0.3$ (top), $\xi_a = 0.6$ (middle), and $\xi_a = 1$ (bottom). We can see that discs in haloes that have $\log M_{h,0} \lesssim 12$ today are mainly supported by feedback throughout their lifetimes, regardless of the accretion conversion efficiency. More massive discs are primarily supported by transport for a low conversion efficiency, while for a moderate conversion efficiency, we can see a transition from transport-dominated turbulence to accretion-dominated turbulence near $z \sim 2$. For a high accretion conversion efficiency, massive discs are always accretion dominated.

This parametrization interpolates between a high value of ξ_a at high redshifts, namely $\xi_a = 1$ at $z = 4$, to a low value at low redshifts, namely $\xi_a = 0.2$ at $z = 0$. The results of our model with this parametrization for ξ_a are shown in Fig. 5. While very crude, the figure shows qualitatively how turbulence support would evolve with an accretion efficiency that is declining with time. We read from the figure that at high redshifts, $z \gtrsim 3$, turbulence in discs is mainly supported by accretion, at all halo masses at which the central galaxy is considered a disc (shaded region, see Section 3).

The turbulence in discs becomes dominated by transport at $z \sim 2.5$. Low-mass discs, residing in haloes with $\log M_{h,0} \lesssim 12$ today, have their turbulence primarily supported by feedback from the moment they become discs, and for the remainder their lifetimes. More massive discs remain in the regime where transport is the dominant driver of turbulence until $z = 0$. From this crude approximation, we expect accretion to be a significant driver at high redshifts, where the streams are more likely to fragment into bound clumps. At low redshifts, the role of accretion weakens, and the evolution of

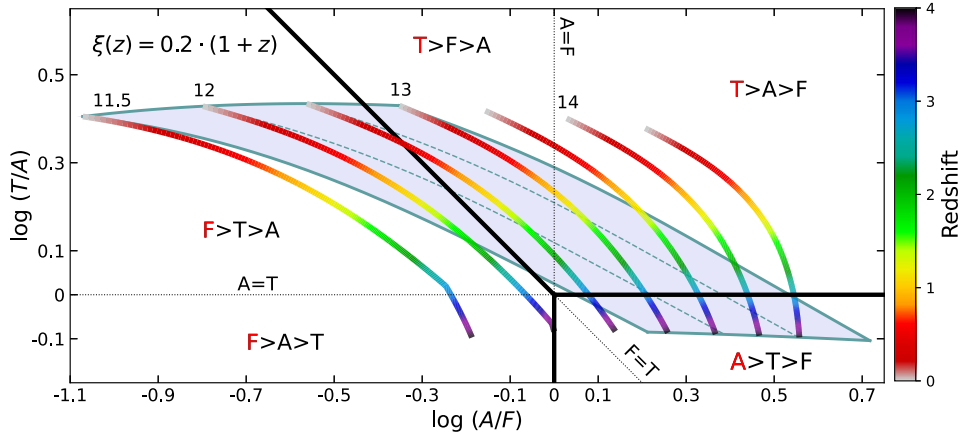


Figure 5. Same as Fig. 4, but with time varying ξ_a (see the text). We see that turbulence in massive discs transition from being primarily supported by accretion to transport. Low-mass discs are primarily supported by feedback at low redshifts.

disc turbulence is practically dominated by feedback and transport (Section 3.1).

It is worth evaluating the qualitative effects on the transport rate of values of Q below unity, as expected for thick discs in marginal instability, $Q \sim 0.67$ (Goldreich & Lynden-Bell 1965). For our fiducial $n = 1$, σ_{SN} is independent of Q (equations 19 and 36), while $\sigma_{\text{acc}} \propto Q^{2/3}$ (equation 37) and $F \propto Q^{-1}$ (equation 35). Combined, a lower value of Q yields a higher transport rate, which results in a lower disc mass, and therefore a smaller σ_g . A smaller velocity dispersion increases the contribution by accretion and feedback at the expense of transport. For the constant $\xi_a = 0.3$, this makes no qualitative difference; however, for $\xi_a = 0.6$, accretion becomes dominant by $z \sim 3$ (and below), compared to our fiducial results. For the time-varying ξ_a model, the transport becomes dominant over the accretion only at $z \sim 1$ and below, compared to our fiducial results.

3.3 Shutoff of transport

As mentioned in Section 2.5 and in Appendix A, there is a critical velocity dispersion $\sigma_c(z, M_h)$ below which instability driven inflow shuts off in order to maintain the energy equilibrium (equations 33a and 33b). While we have seen that in all models considered in Section 3 the transport always has some contribution, it is interesting to see how the addition of accretion-driven turbulence changes the evolution of turbulence in galaxies when transport is absent.

In such cases, $\alpha_c = 0$ or $\tau = 0$ for the two different methods of computing the transport rate, respectively, namely the disc no longer hosts clumps and turbulent viscosity is no longer exerting torques. Equations (33a) and (33b) reduce to the same equation

$$\frac{2}{3} \epsilon_{\text{ff}} \left(\frac{p}{m} \right)_* \chi^{-1} \sigma_g^{-1} + \frac{1}{3} \xi_a \frac{G \dot{M}_{\text{g,acc}} Q}{f_{\text{g,Q}} \sqrt{2(1+\beta)}} \sigma_g^{-3} = \sqrt{2(1+\beta)} \gamma_{\text{diss}}^{-1} \left(\frac{Q}{f_{\text{g,Q}}} \right)^{-n}. \quad (47)$$

Since ξ_a and $\dot{M}_{\text{g,acc}}$ refer to external drivers, they are not self-regulated by the disc. The momentum per unit of stellar mass formed, $(p/m)_*$, is determined by the internal physics of supernova explosions, and hence it is also not self-regulated by the disc. This leaves χ , σ_g , ϵ_{ff} , and Q as possible disc properties that can be self-regulated. The parameter χ depends on Q and on other parameters that are regulated by the assumed vertical force balance, and are unrelated to

the unstable state of the disc (see equation 19). The variations in these parameters are not expected to introduce values beyond the range of values already considered in this work. Hence, the only independent variables that can be self-regulated, in the case of no transport, are Q , ϵ_{ff} , σ_g .

When no transport is in effect, the solution to equation (43) approaches the approximate solution

$$\dot{M}_{\text{g,acc}} = \dot{M}_{\text{SF}} = \epsilon_{\text{ff}} \frac{M_g}{t_{\text{ff}}}. \quad (48)$$

Plugging this expression in equation (47), we get

$$\frac{2}{3} \epsilon_{\text{ff}} \left(\frac{p}{m} \right)_* \chi^{-1} \sigma_g^{-1} + \frac{1}{3} \xi_a \epsilon_{\text{ff}} \frac{G M_g Q}{f_{\text{g,Q}} \chi t_{\text{d}} \sqrt{2(1+\beta)}} \sigma_g^{-3} = \sqrt{2(1+\beta)} \gamma_{\text{diss}}^{-1} \left(\frac{Q}{f_{\text{g,Q}}} \right)^{-n}. \quad (49)$$

From the definition of the Toomre- Q parameter (equation 6), we have that $G M_g Q_g = \sqrt{2(1+\beta)} V_d R_d \sigma_g$. Plugging this into the above equation, we get

$$\frac{2}{3} \epsilon_{\text{ff}} \left(\frac{p}{m} \right)_* \chi^{-1} \sigma_g^{-1} + \frac{1}{3} \xi_a \epsilon_{\text{ff}} \chi^{-1} \left(\frac{\sigma_g}{V_d} \right)^{-2} = \sqrt{2(1+\beta)} \gamma_{\text{diss}}^{-1} \left(\frac{Q}{f_{\text{g,Q}}} \right)^{-n}. \quad (50)$$

Since $\chi \propto Q$ (equation 19), for our fiducial value of $n = 1$, this equation is independent of the value of Q , meaning that even if Q is free to vary, it will not be set by the energy conservation.

We first let ϵ_{ff} be fixed. Under the steady-state approximation, equation (48), the solution of equation (50) is σ_c of equation (A2) for any value of Q , i.e. in the case of a fixed ϵ_{ff} , σ_g remains at the level of σ_c . When only feedback is considered and the contribution of accretion to generating turbulence is ignored, we have $\sigma_c = \sigma_{\text{SN}}$, independent of mass. However, when accretion is included, the critical value of turbulence sustained by feedback and accretion in the absence of transport depends on both halo mass and redshift, $\sigma_c(M_h, z)$ from A2.

On the other hand, if ϵ_{ff} can be self-regulated by the disc while Q remains fixed, we can solve for ϵ_{ff} from equation (50), to get

$$\epsilon_{\text{ff}} = \chi \sqrt{2(1+\beta)} \gamma_{\text{diss}}^{-1} \left(\frac{Q}{f_{\text{g,Q}}} \right)^{-n} \frac{\sigma_g}{\frac{2}{3} \left(\frac{p}{m} \right)_* + \frac{1}{3} \xi_a V_d^2 / \sigma_g}. \quad (51)$$

For a given σ_g , this is smaller than the prediction of equation 54 in Krumholz et al. (2018), as we would expect: since we have added an

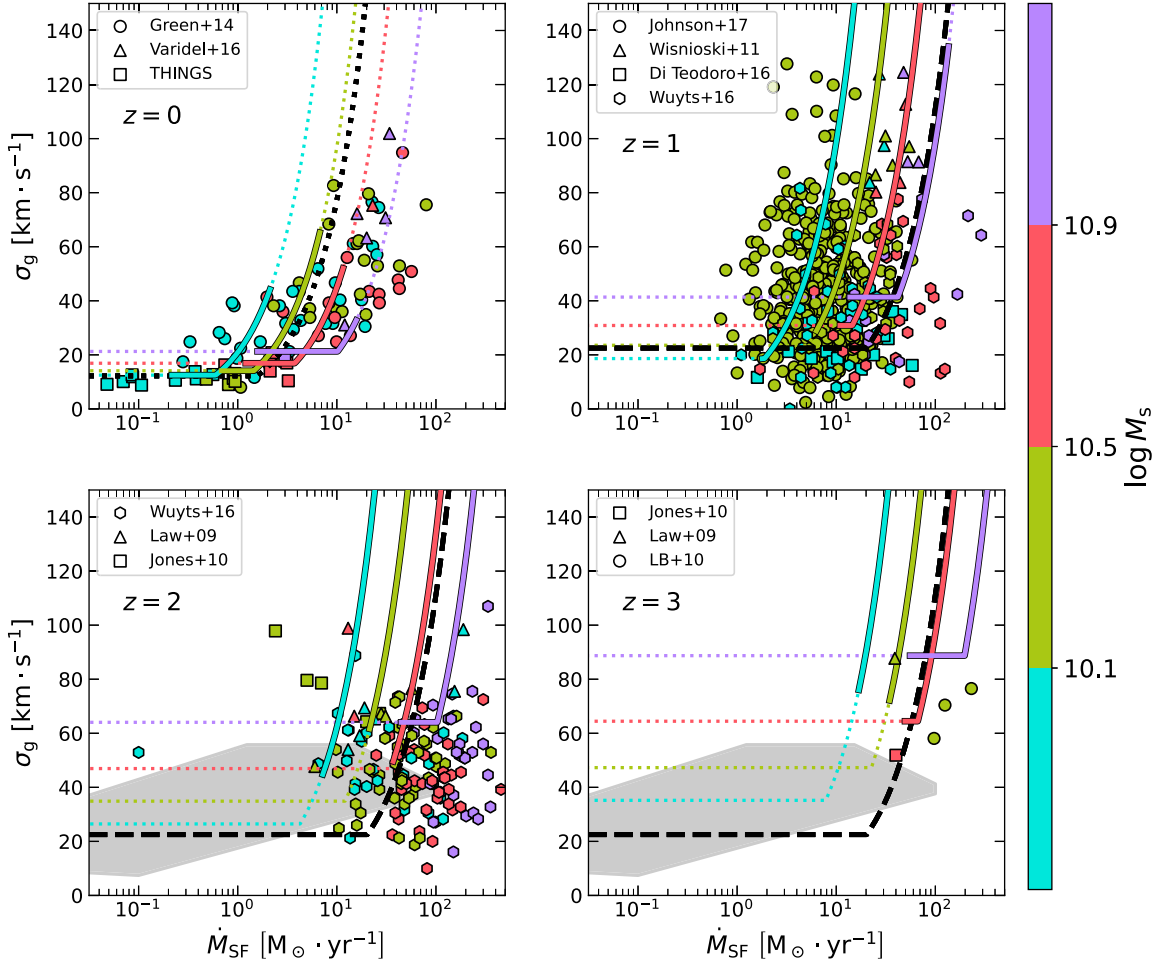


Figure 6. The relation between the gas velocity dispersion, σ_g , and the SFR, \dot{M}_{SF} . Each panel refers to a different redshift. Colours represent different stellar mass bins at each redshift. The coloured curves in each panel represent the relation as predicted by our model, for fixed $M_h(z) = 10^{11.5}, 10^{12}, 10^{12.5}, 10^{13} M_\odot$ at each redshift, converted into stellar mass using the stellar-to-halo mass relation by Behroozi, Wechsler & Conroy (2013). The solid section of each curve indicates the range of SFRs within 0.5 dex of the value at the main sequence, taken from Whitaker et al. (2012). The dashed and dotted lines show the results of the high- z and local spiral model of Krumholz et al. (2018), respectively. The shaded regions are the results from isolated simulations (with feedback) by Ejdetj  m et al. (2021), high gas fraction runs. Symbols represent observational results of H α and H I at low redshifts (Leroy et al. 2008; Walter et al. 2008; Ianjamasimanana et al. 2012; Green et al. 2014; Varidel et al. 2016) and high redshifts (Law et al. 2009; Jones et al. 2010; Lemoine-Busserolle et al. 2010; Wisnioski et al. 2011; Di Teodoro, Fraternali & Miller 2016; Wuyts et al. 2016; Johnson et al. 2018). The masses assigned by colour to the symbols of Jones et al. (2010), Di Teodoro et al. (2016), and Johnson et al. (2018) are the median masses of each sample. We see a general agreement between our model and the observations, both in terms of the level of turbulence and the mass trends.

extra source of energy, in the form of accretion, we require less energy input from star formation, and thus require less extreme values of ϵ_{ff} . Note that for our fiducial $n = 1$, ϵ_{ff} is independent of the value of Q , since $\chi \propto Q$. Equation (51) suggests that, when transport is absent and the star formation efficiency is allowed to vary, more turbulent discs have a higher star formation efficiency.

3.4 $\dot{M}_{\text{SF}}-\sigma_g$ relation

We can use our model to calculate the $\dot{M}_{\text{SF}}-\sigma_g$ relation from equations (6) and (18). According to our model, $\dot{M}_{\text{SF}} \propto \sigma_g$, whenever $\sigma_g \geq \sigma_c(M_h, z)$. In Fig. 6, we show the $\dot{M}_{\text{SF}}-\sigma_g$ in our model⁸

⁸We note that the curves are not produced as a result of integration over the history of a galaxy, but rather represent the functional relation between \dot{M}_{SF} and σ_g .

(assuming the time-varying parametrization for $\xi_a(t)$,⁹ equation 46) at four different redshifts and for four different masses. Each curve in each redshift is for a fixed halo mass at that particular redshift, converted into stellar mass using the stellar-to-halo mass relation, as parametrized in equation 3 of Behroozi et al. (2013). The solid section of each curve indicates the range of SFRs within 0.5 dex of the value at the star-forming main sequence, fitted by Whitaker et al. (2012) to galaxies at $0 \lesssim z \lesssim 2.5$.

When $\sigma_g = \sigma_c$, transport shuts off, and the turbulence remains at the level of σ_c (see discussion in Section 3.3). Our model predicts variations in σ_g as a function of mass at a fixed \dot{M}_{SF} , due to the

⁹While the current tentative choice of the time evolution of ξ_a is ad hoc, the dependence of σ_c on ξ_a is rather weak, a power law with a power of 1/3 (equations 37 and A2), as seen in Fig. A1. The redshift evolution of the $\dot{M}_{\text{SF}}-\sigma_g$ relation is therefore not very sensitive to the choice of $\xi_a(z)$.

dependence of σ_c on the halo mass. This is because galaxies that reside in more massive haloes undergo stronger accretion (due to our chosen accretion model, equation 16) and therefore have a larger σ_c (see Appendix A).¹⁰ The results of analytical models by Krumholz et al. (2018) are also shown in Fig. 6 (their ‘local spirals’ and ‘high- z ’ models). The general behaviour is similar between our model and theirs, with the floor in σ_g in the models by Krumholz et al. (2018) generally smaller, due to the absence of accretion in their model.

In Fig. 6, we compare the relation predicted by our model to H α observations in high redshifts and H α and H I observations at low redshifts (see the figure caption). We use a subset of the observational results compiled in appendix B of Krumholz et al. (2018). We can see a general agreement between our model and the observations,¹¹ with the mass trend in the same ballpark (mass bins indicated by the colour), as well as the varying levels of σ_g at a given SFR. We note, however, that the high levels of σ_g at high SFRs are usually observed for H α rather than atomic or molecular gas, a distinction we have not made in our model.

Ejdetjärn et al. (2021) have studied this relation in simulations of isolated galaxies, and found that when looking at the gas as a whole, the high levels of σ_g observed at high redshifts are not reproduced. On the other hand, the H α -emitting gas alone does seem reproduce the high values for σ_g . Ejdetjärn et al. (2021) have further argued that the high levels of turbulence could be an observational artefact, as a result of beam smearing and inclination effects. While their fiducial sample was not able to reproduce these high levels of turbulence (see shaded regions in Fig. 6), when considering these aforementioned effects, they were able to achieve high values. We note, however, that these simulations, not being in a cosmological setting, do not take into account the possible driving of turbulence by accretion, which can potentially strengthen the turbulence for the total gas.

4 DISCUSSION

It is worth discussing the simplifying assumptions of our model, potential caveats and possible improvements in future work.

4.1 Spatial variations

While we have approximated the galaxy properties (e.g. star formation and mass transport) as spatially independent and only evolving with time, this is clearly an oversimplification. While some of the processes discussed here do approach a spatially independent solution in a steady state (Krumholz & Burkert 2010; Forbes et al. 2012), star formation, feedback, and accretion may depend on the position inside the galaxy.

It has been shown that the cold streams that feed the disc are usually confined to a plane, which is not necessarily parallel to the disc plane (Danovich et al. 2012). While the streams tend to co-rotate with

the disc, about a third of the mass is in-streaming counter rotating to the disc (Danovich et al. 2015). However, the detailed profile of how streams feed the galaxy at different radii has not been studied thoroughly. At $z = 0$, Trapp et al. (2022) used the FIRE simulations (Hopkins et al. 2014) to show that in-streaming gas tends to pile up at the outskirts of the galaxy, and does not directly contribute much flux in inner radii. Since high redshift galaxies tend to reside in more messy regions, and to be fed by more intense inflowing cold streams (Dekel & Birnboim 2006; Dekel et al. 2009a), it is not obvious at all that this picture remains valid at high redshifts.

In future work, we plan to study in detail the different profiles of the different sources and sinks, starting from the feeding of mass and angular momentum by streams at different radii of the disc. These will be used as boundary conditions for a more detailed, spatially dependent model.

4.2 Outflows

Besides stirring up turbulence, feedback is also responsible for generating outflows from the galaxy. This can be due to two main mechanisms. One is supernova explosions and stellar winds, which are related to the SFR of the galaxy, and are effective in low-mass galaxies, below the critical mass of $\log M_h \sim 12$ (Dekel & Silk 1986). The second is due to AGN feedback, which becomes effective after the galaxies go through a wet compaction, above the same critical mass (Lapiner, Dekel & Dubois 2021).

The strength of outflows is usually characterized by the mass loading factor, $\eta = \dot{M}_{g,out}/\dot{M}_{SF}$, where $\dot{M}_{g,out}$ is the mass loss rate due to outflows. Observationally, η varies in the range $\eta \sim 0.1-1$ for star-forming galaxies in the redshift range $0.6 < z < 2.7$ (Davies et al. 2019; Förster Schreiber et al. 2019), while simulations show a wider range of values, $\eta \sim 1-100$ (Muratov et al. 2015; Nelson et al. 2019), with or without AGN (Mitchell et al. 2020a), for galaxies in the same redshift range. The discrepancy between the values of the mass loading factors in different simulations lies in the differences between the assumed feedback models. These feedback models are usually fine tuned to match the observed stellar-to-halo mass relation. However, this practice often disregards other properties of the galaxies that are also affected by feedback. For example, the VELA simulations (Ceverino et al. 2014) produce long lived giant clumps with abundances and physical properties consistent with observations (Mandelker et al. 2017; Guo et al. 2018; Dekel et al. 2022; Ginzburg et al. 2021). However, these simulations often produce galaxies that are more massive than observed, due to the relatively weak feedback model they assume. On the other hand, simulations that implement models with stronger feedback, in order to better agree with the observed stellar-to-halo mass relation, often fail to produce the abundance of long-lived clumps (Genel et al. 2012b; Oklopčić et al. 2017). This conflict between the role of feedback on galactic and clump scales needs to be addressed in future work, which calibrates a feedback model that obeys both large- and small-scale constraints.

Outflows should introduce another sink term in equation (43) of the form $\eta \dot{M}_{SF}$. While it has been shown that the mass outflows from stellar feedback can have an important effect on the evolution of the galaxy (Dekel & Mandelker 2014), this has been studied disregarding other potential forms the energy released by stellar feedback can take, e.g. the turbulent form. Hence, the mass loading factor needs to be modeled consistently, taking into account that not all of the energy from the stellar feedback goes into outflows. Hayward & Hopkins (2017) have developed a turbulent model for the mass loading factor; however, their model takes into account only turbulence driven by

¹⁰We can think of other sources of variations in σ_c , which we do not address in this paper. One source is variations about the average cosmological accretion rate which are translated to variations in σ_c , and thus cause variations in σ_g at a fixed mass. Another source is variations in the strength of feedback at a fixed SFR. Such variations could arise, for example, from variations in $(p/m)_*$ or ϵ_{ff} or from variations in the clustering of supernova (Gentry et al. 2017).

¹¹The galaxies observed by Green et al. (2014) and Varidel et al. (2016) usually lie above the main sequence, being selected intentionally as local analogues of high redshift star-forming galaxies. Even though adopting our model well above the main sequence at $z = 0$ is in agreement with these samples, we note that adopting the average cosmological values for our model parameters for comparing with these galaxies may be incorrect.

supernova feedback while disregarding the other potential drivers of turbulence considered here. Their model needs to be modified to allow for multiple drivers of turbulence, in order to consistently estimate the mass loading factor, which is beyond the scope of this paper.

In order to crudely evaluate the possible qualitative effect of outflows on our results, at the risk of not being self-consistent, we added to equation (43) a sink term of the form $\eta \dot{M}_{\text{SF}}$. The direct effect of a constant mass loading factor is an overall reduction of the mass of the disc. This in turn reduces the velocity dispersion of the disc, thus bringing σ_g closer to σ_c , which means that the role of transport in balancing turbulence dissipation becomes less important. For the fixed ξ_a models, a value of $\eta = 1$ keeps the transport dominant in high mass discs for $\xi_a = 0.3$, but for $\xi_a = 0.6$, the outflows make the accretion dominant already at $z \sim 3$ and below, compared to $z \sim 1$ with no outflows. For the time-varying ξ_a model, the transport overtakes the accretion only at $z \sim 1$ and below, compared to $z \sim 3$ and below with no outflows. A self-consistent analysis incorporating outflows is deferred to future work.

4.3 Recycling

While the external gas accretion is the main source of gas to the galaxy, galaxies that undergo outflows are also subject to recycling of this previously outflowing gas that was unable to escape from the halo. Dekel & Mandelker (2014) have proposed that recycling may be important to explain discrepancies between theoretical models of galaxy formation and observations at around $z \approx 2$, after substantial star formation has occurred. Also, numerical simulations have shown that recycled winds can significantly contribute to the overall accretion of gas, and can sometimes dominate the supply of gas to the galaxy (Anglés-Alcázar et al. 2017). However, there is still disagreement between numerical simulations about the relevant time-scale for recycling, and it can vary between 0.2–1 Gyr (Oppenheimer et al. 2010; Anglés-Alcázar et al. 2017; Tollet et al. 2019), and some simulations find the recycling time-scale to be longer than the Hubble time (Mitchell, Schaye & Bower 2020b).

Modelling recycling theoretically is a challenging task. Dekel & Mandelker (2014) modelled recycling simply as a lower value of the mass loading factor for outflows; however, this treatment assumes that the recycling is instantaneous, namely that the time-scale for recycling is much shorter than other dynamical time-scales in the problem. How recycling drives turbulence is also unknown, in particular its effect on ξ_a and $\dot{M}_{\text{g,acc}}$, and requires more numerical work to inspire analytical models. Recycling is therefore an open question that we do not attempt to address here.

4.4 Modelling of accretion-driven turbulence

Our implementation of accretion-driven turbulence due to the incoming streams has been extremely crude. First, the parametrization of the energy deposited in turbulence by accretion is an oversimplification. The accretion occurs along narrow streams, and the location of the collision between the streams and the disc is not known. Secondly, the conversion efficiency of kinetic energy associated with gas collisions to turbulence was either assumed constant throughout the evolution of the galaxy, or was parametrized ad hoc to have a qualitative idea of how its evolution with time affects the disc as a whole. Klessen & Hennebelle (2010) suggest that the conversion efficiency is $\xi_a \sim \Delta$, where Δ is the density contrast between colliding flows. Mandelker et al. (2018) have estimated analytically that the clumpiness of the streams is a function of both the stream's

host halo mass and redshift, so one would generally expect ξ_a to be a function of those two parameters. In future work, we plan to use the models proposed by Mandelker et al. (2018, 2020b) to estimate analytically the evolution of the conversion parameter. Finally, due to the rotation of the disc, the orientation at which the streams hit the disc can affect the conversion of kinetic to turbulent energy, e.g. streams co-rotating with the disc will stir less turbulence than counter rotating streams. In future work, we plan to study analytically and using simulations how accretion in general, and clumpy accretion in particular, drives turbulence in the disc.

4.5 Mass transport

In this paper, we computed the rate of instability-driven mass transport within the disc in two alternative ways – encounters between clumps and turbulent viscous torques. We found that, when imposing an energy equilibrium, these two processes produce the same mass transport rate, with a non-trivial dependence on the velocity dispersion. However, these two processes do not necessarily cover the whole range of processes involved. Dekel et al. (2020b) have found that once the galaxy exhibits a massive bulge, it tends to generate a long-lived extended ring with suppressed mass transport rate. The transport time-scale in this case was found to be much longer than the time-scale for mass transport in a VDI disc. The time-scale for clump migration as estimated by dynamical friction within the disc has been found to be comparable to and slightly shorter than the estimate based on encounters (Dekel et al. 2013, appendix A). It is still not clear what is the dominant mass transport mechanism, and there is more theoretical work to be done, both analytical and using simulations, to better understand the nature of mass transport in the disc and the implied rate. We plan to study these in more detail in future work.

Furthermore, mass that is being transported down the potential well eventually ends up in the bulge. This induces a stabilizing effect, as described in Dekel et al. (2020b). Furthermore, the gas that flows into the bulge feeds the supermassive black hole (Lapiner et al. 2021), induces AGN feedback, hence another potential source of turbulence we have not considered here.

5 CONCLUSIONS

In this paper, we have addressed the evolution of turbulent support in disc galaxies, as a function of mass and redshift, assuming that they are self-regulated in a marginal Toomre disc instability, with the appropriate level of turbulence maintained for many dynamical times simultaneously by three energy sources that balance the dissipative losses. These energy sources are stellar feedback, mass transport within the disc, and cosmological accretion on to the disc. This is a generalization of previous models which dealt with only two of the driving sources at a time. We used an analytical bathtub model for the evolution of the discs, based on the following assumptions:

- (i) Galaxies are modeled as discs of gas and stars in a turbulent, marginally Toomre unstable state, embedded in dark matter haloes.
- (ii) The discs smoothly transition from gas-dominated discs at high redshifts to two component discs of gas and stars at low redshifts
- (iii) Discs gain fresh gas from external accretion at the average cosmological rate.
- (iv) Discs lose gas by either star formation or mass transport through the disc to the central bulge.

(v) The transport rate has been computed in two alternative ways, via clump encounters and via turbulent viscous torques, with very similar results.

(vi) Turbulence in the discs is driven in concert by three energy sources, namely supernova feedback, instability-driven torques and the associated mass transport, and the impact of clumpy accretion on to the disc.

(vii) The intensity of accretion-driven turbulence is characterized here by an energy conversion efficiency parameter, and we considered low, moderate and maximal efficiencies.

(viii) Turbulence dissipation is balanced by the three energy sources, resulting in an energy balance in the disc.

We then numerically solve a conservation equation for the gas mass and turbulence energy, each time considering a different set of turbulence drivers and eventually considering them all in concert. Under the assumptions of turbulent energy balance and marginal Toomre instability, this directly allows us to compute the turbulent state of the disc, and to determine which turbulence driver is the dominant one, for the given halo mass and redshift. Our results are as follows:

(i) When including only feedback and transport as drivers, turbulence is mainly supported by transport at $z > 1.5$, for all disc masses. At lower redshifts, discs that reside in haloes with masses that evolve to $M_h(z=0) = M_{h,0} \lesssim 10^{12} M_\odot$ today become mainly supported by supernova feedback.

(ii) When including only accretion and transport as drivers, we find that different levels of stream energy conversion efficiencies produce qualitatively different results. For a low conversion efficiency, turbulence in all discs, at all times and masses, is primarily supported by transport, at the level of ~ 70 per cent. For a moderate conversion efficiency, transport and accretion contribute comparable power to sustaining turbulence. Lastly, for a maximal conversion efficiency, 70 per cent of the turbulence dissipation is supported by accretion.

(iii) When considering all three drivers together, we find that for all levels of conversion efficiency, discs that reside in haloes that evolve into $M_{h,0} \lesssim 10^{12} M_\odot$ at $z=0$ have their turbulence primarily supported by supernova feedback throughout their lifetimes. For discs in haloes more massive than $10^{12.5} M_\odot$ today, the result depends on the accretion conversion efficiency as follows:

(a) For a low conversion efficiency, below ~ 40 per cent, the turbulence is primarily supported by transport, throughout the galaxy lifetime.

(b) For a moderate accretion efficiency, at 40–70 per cent, the turbulence is driven primarily by transport at high redshifts; however, the contribution of transport is at most 40 per cent larger than that of accretion. In particular, at $z \sim 2$, the difference between the contributions of transport and accretion to sustaining turbulence is at most 20 per cent. We thus conclude that accretion and transport contribute to sustaining turbulence with comparable power for these galaxies.

(c) For a maximal accretion conversion efficiency, close to unity, the primary driver of turbulence is accretion at all times in the galaxy's history. In particular, at $z \sim 4$, accretion dominates by more than 40 per cent over transport.

(d) Finally, in order to study how an ξ_a that declines with time qualitatively affects the evolution, we make a very crude approximation for its evolution as a function of redshift. We find that discs within haloes that end up with $M_{h,0} \lesssim 10^{12} M_\odot$ have their turbulence supported by feedback at all times,

while turbulence in discs within haloes of $M_{h,0} \gtrsim 10^{12.5} M_\odot$ transition between a marginal accretion dominance to transport dominance at around $z \sim 3$, when their host halo's masses were $M_h(z=3) \gtrsim 10^{11.5} M_\odot$.

(e) The main lesson is that, even though different drivers dominate the turbulence at different stages of the disc's evolution, the relative roles of the three energy sources are comparable to within a factor of 2–3 in all disc galaxies at all times.

(iv) Our model predictions for the $\dot{M}_{\text{SF}}-\sigma_g$ relation are largely consistent with the observations both at low and high redshifts. They are also crudely consistent with simulations of isolated galaxies for low values of SFR, though they may predict for the whole gas higher values of σ_g for high SFR at high redshifts.

Our simplified model suggests that accretion-driven turbulence plays an important and sometimes the dominant role in sustaining turbulence in galactic discs, in parallel with transport and supernova feedback. In the current version of the model, the three ingredients were based on simplified parametric expressions. This should be improved next, based on more detailed physical modelling of the driving of turbulence by accretion, the instability-driven transport rate, and the way feedback drives turbulence. Another important generalization would be to consider radial variations of the physical quantities within the disc, and demanding local equilibrium.

ACKNOWLEDGEMENTS

This work has been supported by the grants.

DATA AVAILABILITY

Data and results underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

- Agertz O., Teyssier R., Moore B., 2009, *MNRAS*, 397, L64
 Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017, *MNRAS*, 470, 4698
 Aung H., Mandelker N., Nagai D., Dekel A., Birnboim Y., 2019, *MNRAS*, 490, 181
 Balbus S. A., Papaloizou J. C. B., 1999, *ApJ*, 521, 650
 Behroozi P. S., Wechsler R. H., Conroy C., 2013, *ApJ*, 770, 57
 Birnboim Y., Dekel A., 2003, *MNRAS*, 345, 349
 Bouché N. et al., 2010, *ApJ*, 718, 1001
 Bournaud F. et al., 2014, *ApJ*, 780, 57
 Bullock J. S., Dekel A., Kolatt T. S., Kravtsov A. V., Klypin A. A., Porciani C., Primack J. R., 2001, *ApJ*, 555, 240
 Cacciato M., Dekel A., Genel S., 2012, *MNRAS*, 421, 818
 Ceverino D., Klypin A., Klimek E. S., Trujillo-Gomez S., Churchill C. W., Primack J., Dekel A., 2014, *MNRAS*, 442, 1545
 Clarke S. D., Whitworth A. P., 2015, *MNRAS*, 449, 1819
 Clarke S. D., Whitworth A. P., Hubber D. A., 2016, *MNRAS*, 458, 319
 Clarke S. D., Whitworth A. P., Duarte-Cabral A., Hubber D. A., 2017, *MNRAS*, 468, 2489
 Danovich M., Dekel A., Hahn O., Teyssier R., 2012, *MNRAS*, 422, 1732
 Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, *MNRAS*, 449, 2087
 Davies R. L. et al., 2019, *ApJ*, 873, 122
 Dekel A., Birnboim Y., 2006, *MNRAS*, 368, 2
 Dekel A., Mandelker N., 2014, *MNRAS*, 444, 2071
 Dekel A., Silk J., 1986, *ApJ*, 303, 39
 Dekel A. et al., 2009a, *Nature*, 457, 451

- Dekel A., Sari R., Ceverino D., 2009b, *ApJ*, 703, 785
- Dekel A., Zolotov A., Tweed D., Cacciato M., Ceverino D., Primack J. R., 2013, *MNRAS*, 435, 999
- Dekel A., Sarkar K. C., Jiang F., Bournaud F., Krumholz M. R., Ceverino D., Primack J. R., 2019, *MNRAS*, 488, 4753
- Dekel A., Ginzburg O., Jiang F., Freundlich J., Lapiner S., Ceverino D., Primack J., 2020a, *MNRAS*, 493, 4126
- Dekel A. et al., 2020b, *MNRAS*, 496, 5372
- Dekel A., Mandelker N., Bournaud F., Ceverino D., Guo Y., primack J., 2022, *MNRAS*, 511, 316
- Dessauges-Zavadsky M. et al., 2019, *Nature Astron.*, 3, 1115
- Di Teodoro E. M., Fraternali F., Miller S. H., 2016, *A&A*, 594, A77
- Draine B. T., 2011, *Physics of the Interstellar and Intergalactic Medium*. Princeton University Press, Princeton
- Dressler A., 1980, *ApJ*, 236, 351
- Ejdetjärn T., Agertz O., Östlin G., Renaud F., Romeo A. B., 2021, *MNRAS*, preprint (arXiv:2111.09322)
- Elmegreen B. G., Burkert A., 2010, *ApJ*, 712, 294
- Elmegreen B. G., Elmegreen D. M., 2005, *ApJ*, 627, 632
- Fall S. M., Efstathiou G., 1980, *MNRAS*, 193, 189
- Faucher-Giguère C.-A., Kereš D., Ma C.-P., 2011, *MNRAS*, 417, 2982
- Faucher-Giguère C.-A., Quataert E., Hopkins P. F., 2013, *MNRAS*, 433, 1970
- Fisher D. B. et al., 2017, *MNRAS*, 464, 491
- Forbes J., Krumholz M., Burkert A., 2012, *ApJ*, 754, 48
- Forbes J. C., Krumholz M. R., Burkert A., Dekel A., 2014, *MNRAS*, 438, 1552
- Förster Schreiber N. M. et al., 2009, *ApJ*, 706, 1364
- Förster Schreiber N. M. et al., 2011, *ApJ*, 739, 45
- Förster Schreiber N. M. et al., 2018, *ApJS*, 238, 21
- Förster Schreiber N. M. et al., 2019, *ApJ*, 875, 21
- Gammie C. F., 2001, *ApJ*, 553, 174
- Genel S., Dekel A., Cacciato M., 2012a, *MNRAS*, 425, 788
- Genel S. et al., 2012b, *ApJ*, 745, 11
- Gentry E. S., Krumholz M. R., Dekel A., Madau P., 2017, *MNRAS*, 465, 2471
- Gentry E. S., Krumholz M. R., Madau P., Lupi A., 2019, *MNRAS*, 483, 3647
- Genzel R. et al., 2006, *Nature*, 442, 786
- Genzel R. et al., 2008, *ApJ*, 687, 59
- Genzel R. et al., 2011, *ApJ*, 733, 101
- Genzel R. et al., 2014, *ApJ*, 785, 75
- Genzel R. et al., 2015, *ApJ*, 800, 20
- Ginzburg O., Huertas-Company M., Dekel A., Mandelker N., Snyder G., Ceverino D., Primack J., 2021, *MNRAS*, 501, 730
- Goldbaum N. J., Krumholz M. R., Forbes J. C., 2015, *ApJ*, 814, 131
- Goldreich P., Lynden-Bell D., 1965, *MNRAS*, 130, 97
- Green A. W. et al., 2014, *MNRAS*, 437, 1070
- Guo Y. et al., 2018, *ApJ*, 853, 108
- Hayward C. C., Hopkins P. F., 2017, *MNRAS*, 465, 1682
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, 445, 581
- Huertas-Company M. et al., 2015, *ApJ*, 809, 95
- Huertas-Company M. et al., 2020, *MNRAS*, 499, 814
- Ianjamasimanana R., de Blok W. J. G., Walter F., Heald G. H., 2012, *AJ*, 144, 96
- Ianjamasimanana R., de Blok W. J. G., Walter F., Heald G. H., Caldu-Primo A., Jarrett T. H., 2015, *AJ*, 150, 47
- Jiang F. et al., 2019, *MNRAS*, 488, 4801
- Johnson H. L. et al., 2018, *MNRAS*, 474, 5076
- Jones T. A., Swinbank A. M., Ellis R. S., Richard J., Stark D. P., 2010, *MNRAS*, 404, 1247
- Kassin S. A. et al., 2012, *ApJ*, 758, 106
- Kennicutt R. C., Jr, 1998, *ARA&A*, 36, 189
- Kereš D., Katz N., Weinberg D. H., Davé R., 2005, *MNRAS*, 363, 2
- Kim C.-G., Ostriker E. C., Raileanu R., 2017, *ApJ*, 834, 25
- Kim J.-G., Ostriker E. C., Filippova N., 2021, *ApJ*, 911, 128
- Klessen R. S., Hennebelle P., 2010, *A&A*, 520, A17
- Kretschmer M., Dekel A., Teyssier R., 2022, *MNRAS*, 510, 3266
- Krumholz M. R., 2013, *MNRAS*, 436, 2747
- Krumholz M., Burkert A., 2010, *ApJ*, 724, 895
- Krumholz M. R., McKee C. F., 2005, *ApJ*, 630, 250
- Krumholz M. R., Dekel A., McKee C. F., 2012, *ApJ*, 745, 69
- Krumholz M. R., Burkert B., Forbes J. C., Crocker R. M., 2018, *MNRAS*, 477, 2716
- Lapiner S., Dekel A., Dubois Y., 2021, *MNRAS*, 505, 172
- Law D. R., Steidel C. C., Erb D. K., Larkin J. E., Pettini M., Shapley A. E., Wright S. A., 2009, *ApJ*, 697, 2057
- Lelli F., Di Teodoro E. M., Fraternali F., Man A. W. S., Zhang Z.-Y., De Breuck C., Davis T. A., Maiolino R., 2021, *Science*, 371, 713
- Lemoine-Busserolle M., Bunker A., Lamareille F., Kissler-Patig M., 2010, *MNRAS*, 401, 1657
- Lenić L., Bolatto A. D., Fisher D. B., Glazebrook K., Obreschkow D., Abraham R., Ambachew L., 2021, *MNRAS*, 506, 3916
- Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2782
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119
- Mac Low M.-M., Klessen R. S., Burkert A., Smith M. D., 1998, *Phys. Rev. Lett.*, 80, 2754
- Mandelker N., Dekel A., Ceverino D., Tweed D., Moody C. E., Primack J., 2014, *MNRAS*, 443, 3675
- Mandelker N., Padnos D., Dekel A., Birnboim Y., Burkert A., Krumholz M. R., Steinberg E., 2016, *MNRAS*, 463, 3921
- Mandelker N., Dekel A., Ceverino D., DeGraf C., Guo Y., Primack J., 2017, *MNRAS*, 464, 635
- Mandelker N., van Dokkum P. G., Brodie J. P., van den Bosch F. C., Ceverino D., 2018, *ApJ*, 861, 148
- Mandelker N., Nagai D., Aung H., Dekel A., Padnos D., Birnboim Y., 2019, *MNRAS*, 484, 1100
- Mandelker N., Nagai D., Aung H., Dekel A., Birnboim Y., van den Bosch F. C., 2020a, *MNRAS*, 494, 2641
- Mandelker N., van den Bosch F. C., Nagai D., Dekel A., Birnboim Y., Aung H., 2020b, *MNRAS*, 498, 2415
- Matzner C. D., 2002, *ApJ*, 566, 302
- Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020a, *MNRAS*, 494, 3971
- Mitchell P. D., Schaye J., Bower R. G., 2020b, *MNRAS*, 497, 4495
- Mo H. J., Mao S., White S. D. M., 1998, *MNRAS*, 295, 319
- Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *MNRAS*, 454, 2691
- Navarro J. F., Frenk C. S., White S. D. M., 1997, *ApJ*, 490, 493
- Neeleman M., Prochaska J. X., Kanekar N., Rafelski M., 2020, *Nature*, 581, 269
- Neistein E., Dekel A., 2008, *MNRAS*, 383, 615
- Nelson D. et al., 2019, *MNRAS*, 490, 3234
- Noguchi M., 1999, *ApJ*, 514, 77
- Obreschkow D. et al., 2015, *ApJ*, 815, 97
- Oklopčić A., Hopkins P. F., Feldmann R., Kereš D., Faucher-Giguère C.-A., Murray N., 2017, *MNRAS*, 465, 952
- Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, *MNRAS*, 406, 2325
- Ostriker E. C., Shetty R., 2011, *ApJ*, 731, 41
- Ostriker E. C., McKee C. F., Leroy A. K., 2010, *ApJ*, 721, 975
- Padnos D., Mandelker N., Birnboim Y., Dekel A., Krumholz M. R., Steinberg E., 2018, *MNRAS*, 477, 3293
- Pillepich A. et al., 2019, *MNRAS*, 490, 3196
- Press W. H., Schechter P., 1974, *ApJ*, 187, 425
- Rafikov R. R., 2001, *MNRAS*, 323, 445
- Rizzo F., Vegetti S., Powell D., Fraternali F., McKee J. P., Stacey H. R., White S. D. M., 2020, *Nature*, 584, 201
- Romeo A. B., Falstad N., 2013, *MNRAS*, 433, 1389
- Schmidt M., 1959, *ApJ*, 129, 243
- Shakura N. I., Sunyaev R. A., 1973, *A&A*, 500, 33
- Shibuya T., Ouchi M., Kubo M., Harikane Y., 2016, *ApJ*, 821, 72
- Simons R. C. et al., 2017, *ApJ*, 843, 46
- Stott J. P. et al., 2016, *MNRAS*, 457, 1888
- Tacconi L. J. et al., 2018, *ApJ*, 853, 179

- Tollet É., Cattaneo A., Macciò A. V., Dutton A. A., Kang X., 2019, *MNRAS*, 485, 2511
- Toomre A., 1964, *ApJ*, 139, 1217
- Trapp C. et al., 2022, *MNRAS*, 509, 4149
- Übler H. et al., 2019, *ApJ*, 880, 48
- Varidel M., Pracy M., Croom S., Owers M. S., Sadler E., 2016, *PASA*, 33, e006
- Virtanen P. et al., 2020, *Nature Methods*, 17, 261
- Vutisalchavakul N., Evans Neal J. I., Heyer M., 2016, *ApJ*, 831, 73
- Wada K., Meurer G., Norman C. A., 2002, *ApJ*, 577, 197
- Walch S., Naab T., 2015, *MNRAS*, 451, 2757
- Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C., Jr, Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563
- Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29
- Wisnioski E. et al., 2011, *MNRAS*, 417, 2601
- Wisnioski E. et al., 2015, *ApJ*, 799, 209
- Wuyts S. et al., 2011, *ApJ*, 742, 96
- Wuyts S. et al., 2016, *ApJ*, 831, 149
- Yu X., Bian F., Krumholz M. R., Shi Y., Li S., Chen J., 2021, *MNRAS*, 505, 5075
- Zhao D. H., Jing Y. P., Mo H. J., Börner G., 2009, *ApJ*, 707, 354

APPENDIX A: CALCULATING σ_c

The critical velocity below which gravitationally driven mass transport halts is calculated by solving a cubic equation of the form

$$1 - \frac{A}{x} - \frac{B^3}{x^3} = 0. \quad (\text{A1})$$

In our case, $x = \sigma_c$, $A = \sigma_{\text{SN}}$ and $B = \sigma_{\text{acc}}$ (see equation 34). Since σ_{acc} depends on $\dot{M}_{\text{g,acc}}$, which in turn depends on M_{h} , σ_c will be a function of halo mass as well as redshift. The cubic in equation (A1)

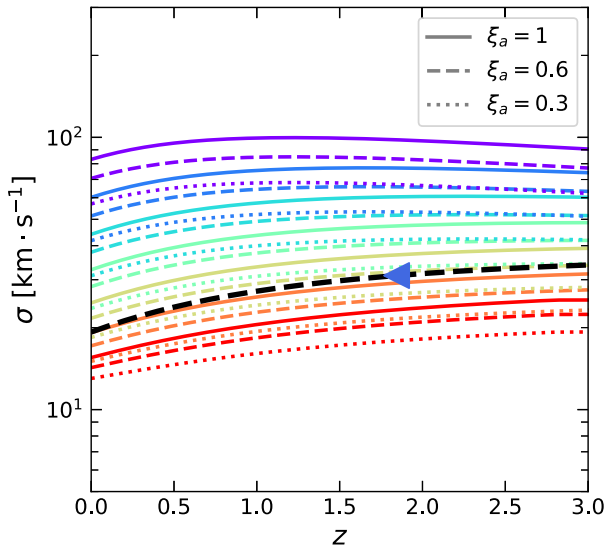


Figure A1. The evolution of the critical velocity dispersion, σ_c , below which mass transport shuts off. Coloured lines show the evolution of σ_c for haloes with $\log M_{\text{h},0} = 11.5$ – 14.5 in intervals of 0.5 dex. Solid, dashed, and dotted lines show the evolution for $\xi = 1, 0.6, 0.3$, respectively. The black dashed line shows the evolution of the velocity dispersion in a disc with $\log M_{\text{h},0} = 12$, evaluated using $\xi_a = 0.6$. The relevant σ_c curve for this galaxy is the dashed orange curve. The blue arrow indicates the point in time in which this galaxy becomes a disc. We can see that the galaxy is always above this critical value, indicating the mass transport is always present, though its contribution decreases with time.

is monotonically increasing, and therefore has a unique real root. It is given by

$$\sigma_c = \frac{1}{3} \left(\sigma_{\text{SN}} + \frac{\sigma_{\text{SN}}^2}{C^{1/3}} + C^{1/3} \right), \quad (\text{A2})$$

where

$$C = \frac{1}{2} \left(2\sigma_{\text{SN}}^3 + 27\sigma_{\text{acc}}^3 + 3\sqrt{12\sigma_{\text{SN}}^3\sigma_{\text{acc}}^3 + 81\sigma_{\text{acc}}^6} \right). \quad (\text{A3})$$

Fig. A1 shows the evolution of σ_c for different $\log M_{\text{h},0}$, for the three levels of conversion efficiencies considered. As the conversion efficiency increases, the values of σ_c increase, indicating that higher levels of turbulence can be sustained by feedback and accretion, without the need for mass transport (see discussion in Section 3.3).

As discussed in the text, discs in our model never reach this threshold of σ_c . As an illustration, the black line in Fig. A1 shows the gas velocity dispersion of a disc with $\log M_{\text{h},0} = 12$, evaluated using a model with $\xi_a = 0.6$. The value for σ_c which corresponds to this halo is the orange dashed line. We see that the velocity dispersion of this disc (dashed black curve) is always larger than σ_c (dashed orange curve), meaning transport always occurs in this disc.

APPENDIX B: CONVERGENCE TO A STEADY-STATE SOLUTION

In Section 3, we rely on the assumption that the numerical solutions to equation (43) converge towards a unique steady-state solution, independent of the initial conditions, which can be obtained by solving the algebraic equation $\dot{M}_{\text{g,acc}} = \dot{M}_{\text{trans}} + \dot{M}_{\text{SF}}$. In this section, we justify this assumption numerically. We define the quantity Δt_{SS} to be the time elapsed since the beginning of the integration until the deviation between the steady-state solution to the true solution of

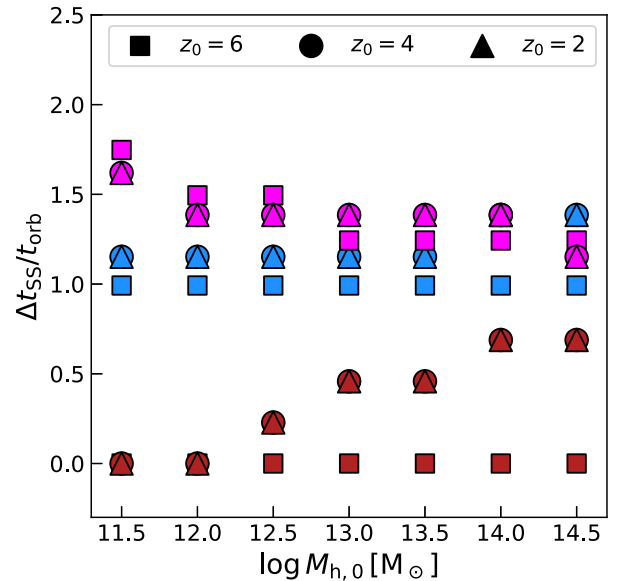


Figure B1. The convergence time of the numerical solution to the steady-state unique solution. Shown are the number of disc orbital times until the solution converges to the steady-state solution, for different values of $\log M_{\text{h},0}$, for different initial conditions and initial redshifts. The different symbols indicate the different initial redshifts. The blue, red, and magenta markers indicate the convergence time for initial conditions $M_{\text{g}} = 0.1, 0.01, 0.005 M_{\text{h}}(z = z_0)$, respectively, at the initial redshift indicated by the symbol type. We see that the convergence time is at most two disc orbital times.

equation (43) is less than 5 per cent. In Fig. B1, we show the number of disc orbital times it takes for the solution to converge to the steady-state solution, for different choices of initial conditions and different choice of initial redshifts, for different halo masses. The reason we test the convergence for different initial redshifts is that since, when presenting our results, we assume that our model is applicable only when the galaxies' host halo mass enters a certain mass range, but we still integrate the evolution of the galaxy using our model from beforehand.

We see from Fig. B1 that the convergence time is less than two disc orbital times, for any choice of initial conditions and initial redshift. This means that from the moment the galaxy is considered disc by

our model, it takes a couple of hundreds of megayears for our model to be applicable to that particular galaxy. For example, for our model with $\xi_a = 0.6$, a disc with $\log M_{h,0} = 12$ today became a disc at $z \sim 2$. As a result of our integration prior to $z = 2$, this galaxy had $M_g \sim 0.006 \cdot M_h(z = 2)$. From our convergence test, we learn that the results of our model are applicable to this galaxy after $\sim t_{\text{orb}}$ (magenta triangle), namely from $z \sim 1.8$.

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