

The maximum accretion rate of hot gas in dark matter halos

Jonathan Stern,¹*† Drummond Fielding,² Claude-André Faucher-Giguère¹ and Eliot Quataert³

¹Department of Physics and Astronomy and CIERA, Northwestern University, Evanston, IL, USA

²Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave, New York, NY 10010, USA

³Astronomy Department and Theoretical Astrophysics Center, University of California Berkeley, Berkeley, CA 94720, USA

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We revisit the question of ‘hot mode’ versus ‘cold mode’ accretion onto galaxies using steady-state cooling flow solutions and idealized 3D hydrodynamic simulations. We demonstrate that for the hot accretion mode to exist, the cooling time is required to be longer than the free-fall time near the radius where the gas is rotationally-supported, R_{circ} , i.e. the existence of the hot mode depends on physical conditions at the galaxy scale rather than on physical conditions at the halo scale. When allowing for the depletion of the halo baryon fraction relative to the cosmic mean, the longer cooling times imply that a virialized gaseous halo may form in halo masses below the threshold of $\sim 10^{12} \text{ M}_\odot$ derived for baryon-complete halos. We show that for any halo mass there is a maximum accretion rate for which the gas is virialized throughout the halo and can accrete via the hot mode of $\dot{M}_{\text{crit}} \approx 0.7(v_c/100 \text{ km s}^{-1})^{5.4}(R_{\text{circ}}/10 \text{ kpc})(Z/Z_\odot)^{-0.9} \text{ M}_\odot \text{ yr}^{-1}$, where Z and v_c are the metallicity and circular velocity measured at R_{circ} . For accretion rates $\gtrsim \dot{M}_{\text{crit}}$ the volume-filling gas phase can in principle be ‘transonic’ – virialized in the outer halo but cool and free-falling near the galaxy. We compare \dot{M}_{crit} to the average star formation rate (SFR) in halos at $0 < z < 10$ implied by the stellar-mass halo-mass relation. For a plausible metallicity evolution with redshift, we find that $\text{SFR} \lesssim \dot{M}_{\text{crit}}$ at most masses and redshifts, suggesting that the SFR of galaxies could be primarily sustained by the hot mode in halo masses well below the classic threshold of $\sim 10^{12} \text{ M}_\odot$.

Key words: –

1 INTRODUCTION

The dynamics of the volume-filling gas phase in dark matter halos, and the nature of its accretion onto the galaxy, crucially depend on whether its cooling time t_{cool} is longer or shorter than the free-fall time t_{ff} . If $t_{\text{cool}} > t_{\text{ff}}$, this phase is expected to be quasi-static, supported against gravity by thermal pressure. Galaxy accretion in this regime is gradual and regulated by energy losses to radiation. In contrast if $t_{\text{cool}} < t_{\text{ff}}$ then the rapid cooling prevents the formation of a pressure-supported gaseous halo, and the halo gas free-falls onto the galaxy. These two distinct regimes for the nature of galaxy accretion, known respectively as ‘hot mode’ and ‘cold mode’ accretion, were originally discussed by White & Rees (1978) who demonstrated that the ratio $t_{\text{cool}}/t_{\text{ff}}$ increases

with halo mass M_{halo} . They identified a threshold mass scale of $M_{\text{thres}} \sim 10^{12} \text{ M}_\odot$ where $t_{\text{cool}} \sim t_{\text{ff}}$, similar to the threshold previously derived for self-gravitating gas clouds (Rees & Ostriker 1977; Silk 1977). Birnboim & Dekel (2003, hereafter BD03) and Dekel & Birnboim (2006) later connected these two regimes to the stability of the virial shock. Using analytic arguments and 1D simulations, they demonstrated that the rapid cooling of postshock gas in low mass halos leads to an unstable shock, so gas accreting from the intergalactic medium (IGM) remains cool ($\sim 10^4 \text{ K}$) and free-falling down to the galaxy scale. Once however M_{halo} surpasses $M_{\text{thres}} \sim 10^{11.5} \text{ M}_\odot$ the conditions for a stable shock are met at the galaxy scale, and a shock forms and expands into the halo heating the volume-filling phase to the virial temperature T_{vir} .

More recently Fielding et al. (2017) used idealized 3D simulations to study how the two regimes for the halo gas are affected by kinetic feedback from stars. They demonstrated

* CIERA Fellow

† E-mail: jonathan.stern@northwestern.edu

that in the $M_{\text{halo}} > M_{\text{thres}}$ regime the outflows are confined by the hot halo gas, and the physics of the volume-filling phase are similar to that suggested by BD03. In low mass halos however the effect of feedback is much more dramatic – galaxy outflows shock against IGM inflows at halo radii, well beyond the radius where the shock initially forms in the BD03 simulations. In this regime the halo gas forms a multi-phase medium dominated by turbulence and bulk inflows/outflows.

A considerable effort has been devoted to detecting these two regimes for galaxy accretion, and the transition between them, in cosmological simulations (e.g., Kereš et al. 2005, 2009, 2012; Birnboim et al. 2007; Ocvirk et al. 2008; Brooks et al. 2009; Oppenheimer et al. 2010; Faucher-Giguère et al. 2011; van de Voort et al. 2011; Nelson et al. 2013). To discriminate between gas which has shocked prior to accretion and gas which has not shocked, most of these studies identified the maximum temperature T_{max} a fluid element had reached before accreting onto the galaxy. In studies where the ‘hot’ and ‘cold’ accretion modes are differentiated by a constant cut in temperature $T_{\text{cut}} \approx 10^{5.5}$ K the gas was found to be entirely cold below $M_{\text{halo}} \approx 10^{11.5} M_{\odot}$, consistent with the conclusion of BD03 (e.g. Kereš et al. 2005). This trend however could be driven by T_{vir} dropping below T_{cut} in low mass halos, in which case even virial-temperature gas would be classified as ‘cold’, as discussed in Nelson et al. (2013) and acknowledged by many of the cited studies. Nelson et al. (2013) scaled T_{cut} with T_{vir} and found that the hot accretion mode is present even in halo masses well below $10^{11.5} M_{\odot}$ (see also figure 8 in van de Voort et al. 2011), in contrast with the 1D simulations of BD03. However, given that even in the free-fall regime inflows potentially shock and reach a temperature $\sim T_{\text{vir}}$ due to the interaction with outflows as seen in the simulations of Fielding et al. (2017), T_{max} may not be a good discriminator between the two regimes. An alternative method to distinguish between gradual, pressure-supported accretion and supersonic free-fall in cosmological simulations would thus be useful.

Another complication arises since t_{cool} depends on the gas density, which implies that the transition between cold and hot mode accretion depends on the gas mass M_{gas} available to form the hot volume-filling phase. The idealized studies mentioned above assumed M_{gas} roughly equals the cosmic halo baryon budget $f_b M_{\text{halo}}$ ($f_b \approx 0.16$ is the cosmic baryon fraction). If however a significant fraction of halo baryons are confined to filaments and subhalos, or, alternatively, if the halo baryons were ejected from the halo by unbound galaxy outflows at earlier times, then M_{gas} will be lower than $f_b M_{\text{halo}}$, t_{cool} would correspondingly be longer, and the transition to pressure support would occur in halos less massive than derived by assuming halos are baryon-complete. Specifically, there is mounting observational evidence for the existence of strong unbound outflows, especially in dwarf galaxies which reside in halos with mass lower than M_{thres} (e.g. Heckman & Thompson 2017; Chisholm et al. 2017). Also, cosmological simulations which model galaxy outflows often predict halo gas masses lower than $f_b M_{\text{halo}}$. In the FIRE zoom-in simulations Hafen et al. (2019) finds a halo gas mass of $\approx 0.2 f_b M_{\text{halo}}$ in $\sim 10^{11} M_{\odot}$ halos at low redshift. In lower mass $\sim 10^{10} M_{\odot}$ halos in FIRE the baryon fraction is even lower, less than 10% of the cosmic baryon budget. A low baryon mass fraction in low mass halos is

found also in the EAGLE cosmological simulations (Davies et al. 2019; Oppenheimer et al. 2019).¹ Thus, both observations and some theoretical studies suggest that M_{gas} could be well below $f_b M_{\text{halo}}$ in halos with mass below the threshold derived assuming $M_{\text{gas}} \approx f_b M_{\text{halo}}$, in which case hot mode accretion could be important also in low mass halos.

In this work (Paper II) we deduce the conditions under which hot mode accretion is possible by analyzing the properties of cooling flow solutions. Cooling flows were originally discussed in the context of gas in the centers of clusters (Mathews & Bregman 1978; Cowie et al. 1980; Fabian et al. 1984; Bertschinger 1989), and adapted to galaxy scale halos in the first paper in this series (Stern et al. 2019, hereafter Paper I). Here, we focus on halo masses which are comparable or below the classic threshold for the formation of a hot halo $M_{\text{thres}} \sim 10^{12} M_{\odot}$. We demonstrate that in low mass halos hot mode accretion depends on the location of the sonic point in the cooling flow that forms – only if the sonic radius is within the galaxy scale is hot accretion possible. This condition was only briefly mentioned in classical studies of the cooling flow solution (Mathews & Bregman 1978) since in cluster-scale halos the expected sonic radius is well within the central galaxy and thus hot mode accretion is always possible. We further show that our formalism for identifying the onset of hot mode accretion yields similar numerical values to the formalism in BD03, though it provides alternative physical intuition for the transition between the two regimes for galaxy accretion. Specifically, the cooling flow formalism suggests that near the threshold for hot mode accretion the halo may assume an ‘inverted’ configuration, in which the volume-filling phase is hot and pressure-supported on large scales but cool and free-falling near the galaxy.

To account for the possibility of a gas mass $< f_b M_{\text{halo}}$ due to e.g. galaxy outflows, we treat the hot gas mass in our analysis as a free parameter. Our derivation thus yields for any M_{halo} and redshift z a maximum gas mass in which hot mode accretion is possible, or equivalently a maximum hot mode accretion rate \dot{M}_{crit} (see below). We then compare the derived \dot{M}_{crit} to the average star formation rate (SFR) in dark matter halos at $0 < z < 10$, which has been constrained via abundance matching techniques and ‘empirical models’ for how galaxies populate dark matter halos (e.g., Moster et al. 2010, 2018; Behroozi et al. 2013, 2019).

This paper is organized as follows. In section 2 we derive the maximum hot mode accretion rate \dot{M}_{crit} using analytic arguments, and corroborate our conclusions with idealized hydrodynamic simulations. In section 3 we explore the dependence of \dot{M}_{crit} on halo and gas parameters, while in section 4 we compare \dot{M}_{crit} with the mean SFR in halos derived by empirical models. We summarize and discuss our results in section 5. In a follow up paper (hereafter Paper III) we compare our results to the properties of halo gas in the FIRE cosmological simulations (Hopkins et al. 2018). Throughout the paper we assume a flat Λ CDM cosmology with Hubble constant $H_0 = 68 \text{ km s}^{-1} \text{ Mpc}^{-1}$ and $\Omega_{\text{m},0} = 0.31$ (Planck Collaboration et al. 2016).

¹ In contrast, the baryon fraction in dwarf halos in the IllustrisTNG simulations appears to be closer to the baryon budget (Nelson et al. 2018).

2 HOT VS. COLD ACCRETION ACCORDING TO COOLING FLOW SOLUTIONS

In this section we use cooling flow solutions to derive a necessary condition for hot mode accretion, and show that this condition can be cast as a maximum hot mode accretion rate \dot{M}_{crit} . In our derivation we assume that the background potential is constant in time, and limit the effects of feedback in our analysis to the possible enrichment and depletion of the halo gas, i.e. ongoing feedback heating is assumed to be negligible. The validity of these assumptions is discussed below and tested in Paper III using cosmological simulations.

We first demonstrate in section 2.1 how \dot{M}_{crit} arises by requiring $T \approx T_{\text{vir}}$ and $t_{\text{cool}} \gtrsim t_{\text{ff}}$ in a steady spherical flow. We then corroborate our derivation using the family of cooling flow solutions to the steady-state flow equations (section 2.2), and using idealized 3D hydrodynamic simulations (section 2.3).

2.1 The $t_{\text{cool}} \gtrsim t_{\text{ff}}$ condition

The energy conservation equation for a steady spherical flow is (appendix A):

$$v_r \frac{d}{dr} \left(\frac{1}{2} v_r^2 + \gamma \epsilon + \Phi \right) = -q, \quad (1)$$

where r is the radius, v_r is the radial velocity, the sum in the brackets is the Bernoulli parameter, ϵ is the specific thermal energy, $\gamma = 5/3$ is the adiabatic index, Φ is the gravitational potential, and q is the cooling rate per unit mass. In a pressure-supported flow the kinetic term is small, while the roughly isothermal potential in dark matter halos implies that the temperature is approximately constant, so the first two terms in the brackets can be neglected. We thus get

$$\frac{d\Phi}{dr} \approx -\frac{q}{v_r} = -\frac{n_{\text{H}}^2 \Lambda}{\rho v_r}, \quad (2)$$

where in the second equality we replaced q with $n_{\text{H}}^2 \Lambda / \rho$ (ρ and n_{H} are the mass and hydrogen density and Λ is the cooling function). Conservation of mass in a steady flow implies an accretion rate $\dot{M} = -4\pi r^2 \rho v_r$, and thus eqn. (2) gives

$$\dot{M} \approx \frac{4\pi r^2 n_{\text{H}}^2 \Lambda}{d\Phi/dr}. \quad (3)$$

The maximum accretion rate for the hot gas can be derived from eqn. (3) by requiring that the density is low enough so $t_{\text{cool}} \gtrsim 0.7 t_{\text{ff}}$. The motivation for the 0.7 prefactor is given in section 2.2. We use

$$t_{\text{ff}} = \frac{\sqrt{2}r}{v_c} \quad (4)$$

where v_c is the circular velocity, and

$$t_{\text{cool}} = \frac{\epsilon}{q} = \frac{\rho \epsilon}{n_{\text{H}}^2 \Lambda}. \quad (5)$$

The maximum gas density is hence

$$n_{\text{H, max}} \approx \frac{m_p v_c \epsilon}{0.7 \cdot \sqrt{2} X \Lambda r} \approx \frac{m_p v_c^3}{X \Lambda r}, \quad (6)$$

where X is the hydrogen mass fraction and m_p is the proton mass. In the second equality we used $\epsilon \approx v_c^2$, which is equivalent to $T \approx (4/3)(v_c/v_{\text{vir}})^2 T_{\text{vir}}$, where $v_{\text{vir}} = v_c(R_{\text{vir}})$ is the virial

velocity and $T_{\text{vir}} = \mu m_p v_{\text{vir}}^2 / 2k$ is the virial temperature. The prefactor in this relation is also justified in section 2.2. Plugging eqn. (6) in eqn. (3) and using $d\Phi/dr = v_c^2/r$ we get a maximum hot gas accretion rate at radius r of

$$\dot{M}_{\text{max}}(r) \approx \frac{4\pi m_p^2 v_c^4 r}{X^2 \Lambda(r)}. \quad (7)$$

In a dark matter halo with an NFW profile (Navarro et al. 1997) v_c is roughly independent of radius, while the roughly constant temperature suggests Λ is also approximately constant, or decreases outwards if metallicity gradients are significant. Eqn. (7) thus suggests that $\dot{M}_{\text{max}}(r)$ increases with radius. For the flow to be pressure-supported at all radii we therefore need to evaluate \dot{M}_{max} at the innermost radius of the flow. For this inner radius we use the circularization radius R_{circ} at which the centrifugal and gravitational forces balance, and thus the gas can be supported by rotation rather than by thermal pressure (see further discussion below). The choice of R_{circ} for the innermost radius of the flow is also motivated by observations which suggest galaxy sizes are $\sim R_{\text{circ}}$ (e.g., Kravtsov 2013; Shibuya et al. 2015). Using R_{circ} in eqn. (7) hence implies a maximum accretion rate for pressure-supported flows of

$$\begin{aligned} \dot{M}_{\text{crit}} &= \dot{M}_{\text{max}}(R_{\text{circ}}) \approx \frac{4\pi m_p^2 v_c^4 R_{\text{circ}}}{X^2 \Lambda} \\ &= 1.7 \left(\frac{v_c(R_{\text{circ}})}{100 \text{ km s}^{-1}} \right)^4 \left(\frac{R_{\text{circ}}}{10 \text{ kpc}} \right) \left(\frac{X^2 \Lambda(R_{\text{circ}})}{10^{-22} \text{ c.g.s}} \right)^{-1} \text{ M}_\odot \text{ yr}^{-1}. \end{aligned} \quad (8)$$

The numerical values of v_c and R_{circ} in eqn. (8) correspond roughly to a halo mass of $\sim 4 \cdot 10^{11} \text{ M}_\odot$ at $z = 0$, though note that the derivation is general and applies to halos of all masses and redshifts. We estimate R_{circ} using the relation

$$v_c(R_{\text{circ}}) R_{\text{circ}} = f_\lambda \frac{J_{\text{halo}}}{M_{\text{halo}}} = \sqrt{2} f_\lambda \lambda v_{\text{vir}} R_{\text{vir}}, \quad (9)$$

where J_{halo} and R_{vir} are the angular momentum and virial radius of the dark matter halo, f_λ accounts for differences between the specific angular momentum of the baryons and the average of the halo, and λ is the halo spin parameter defined in Bullock et al. (2001):

$$\lambda \equiv \frac{J}{\sqrt{2} M_{\text{halo}} v_{\text{vir}} R_{\text{vir}}}. \quad (10)$$

For $f_\lambda \approx 1$, $R_{\text{vir}} \approx 200 \text{ kpc}$, $v_c(R_{\text{circ}}) \approx v_{\text{vir}}$, and $\lambda \approx 0.035$ (e.g., in the Bolshoi-Planck simulation, Rodríguez-Puebla et al. 2016) we get $R_{\text{circ}} \approx 10 \text{ kpc}$. Eqn. (8) can be further elaborated by approximating Λ in the metal-dominated regime as (e.g. Wiersma et al. 2009)

$$\Lambda = 0.5 \cdot 10^{-22} \left(\frac{T}{10^6 \text{ K}} \right)^{-0.7} \left(\frac{Z}{0.3 Z_\odot} \right)^{0.9} \text{ erg cm}^3 \text{ s}^{-1}, \quad (11)$$

where Z is the gas metallicity, and this approximation is valid at $T \sim 10^5 - 10^7 \text{ K}$ and $Z \gtrsim 0.3 Z_\odot$. For $T = 5 \cdot 10^5 (v_c/100 \text{ km s}^{-1})^2 \text{ K}$ implied by $\epsilon = v_c^2$ we get

$$\dot{M}_{\text{crit}} = 0.7 \left(\frac{v_c(R_{\text{circ}})}{100 \text{ km s}^{-1}} \right)^{5.4} \left(\frac{R_{\text{circ}}}{10 \text{ kpc}} \right) \left(\frac{Z(R_{\text{circ}})}{0.3 Z_\odot} \right)^{-0.9} \text{ M}_\odot \text{ yr}^{-1}. \quad (12)$$

2.2 Spherically-symmetric cooling flow solutions

To further demonstrate that hot mode accretion is possible only for accretion rates below the critical value \dot{M}_{crit} derived in the previous section, we utilize the family of cooling flow solutions derived in Paper I. We start by discussing purely radial flows, and then include the effects of angular momentum.

2.2.1 Cooling flows without angular momentum

Cooling flow solutions are derived from the spherical steady-state equations for radiatively cooling gas in a constant gravitational potential:

$$\dot{M} = -4\pi r^2 \rho v_r \quad (13)$$

$$\frac{1}{2} \frac{dv_r^2}{dr} = -\frac{1}{\rho} \frac{dP}{dr} - \frac{v_c^2}{r} \quad (14)$$

$$v_r \frac{d \ln K}{dr} = -\frac{1}{t_{\text{cool}}}, \quad (15)$$

where P is the gas pressure and $\ln K \equiv \ln(kT/n_H^{2/3})$ is the entropy. We integrate these equations as described in Paper I, requiring the solutions to go through a sonic point and to be marginally-bound at large radii (Bernoulli parameter $\mathcal{B} \rightarrow 0^-$ as $r \rightarrow \infty$). The transonic condition is required since non-transonic solutions are either not well-defined at all radii (e.g. Bertschinger 1989), or everywhere supersonic. The exact choice of the outer boundary condition does not affect the conditions near R_{circ} and hence is of no consequence for the discussion here (see Fig. B1 in Paper I). For a given cooling function and gravitational potential the transonic and marginally-bound conditions yield a single-parameter family of solutions. We showed in Paper I that gaseous halos which are initially hydrostatic converge onto these solutions within a cooling time.

For simplicity we assume an isothermal gravitational potential², and address the implications of more realistic potentials below. As instructive examples, we calculate four cooling flow solutions of $Z = Z_\odot/3$ gas in an isothermal potential with $v_c = 100 \text{ km s}^{-1}$, corresponding at $z = 0$ to $M_{\text{halo}} = 4 \cdot 10^{11} M_\odot$. For Λ we use the Wiersma et al. (2009) tables for $z = 0$, which account for photoionization and heating by a Haardt & Madau (2012) UV background. The panels in Figure 1 plot T , n_H , radial Mach number \mathcal{M} , and $t_{\text{cool}}/t_{\text{ff}}$ of the solutions. The solutions differ in their assumed density normalization (second panel), or equivalently in the total gas mass. A higher density normalization corresponds to a higher inflow rate \dot{M} due to the increased cooling (\dot{M} indicated in the top panel) and to a larger sonic radius R_{sonic} (third panel). Specifically for the case of $M_{\text{halo}} = 4 \cdot 10^{11} M_\odot$ at $z = 0$, the purple, blue, green, and yellow solutions have respectively a gas mass within $R_{\text{vir}} = 200 \text{ kpc}$ equal to 0.2, 0.4, 1, and 1.4 times the cosmic baryon budget, and a sonic radius equal to 0.01, 0.04, 0.25, and 1 times R_{vir} .

Figure 1 shows that in the outer subsonic part of the flows the solutions satisfy the conditions for pressure support discussed in the previous section: the gas temperature is roughly equal to $T_{\text{vir}} = 3.6 \cdot 10^5 \text{ K}$ (top panel) and the ratio

² To calculate the Bernoulli parameter in an isothermal potential we assume the potential equals zero at $r = 10 \text{ Mpc}$.

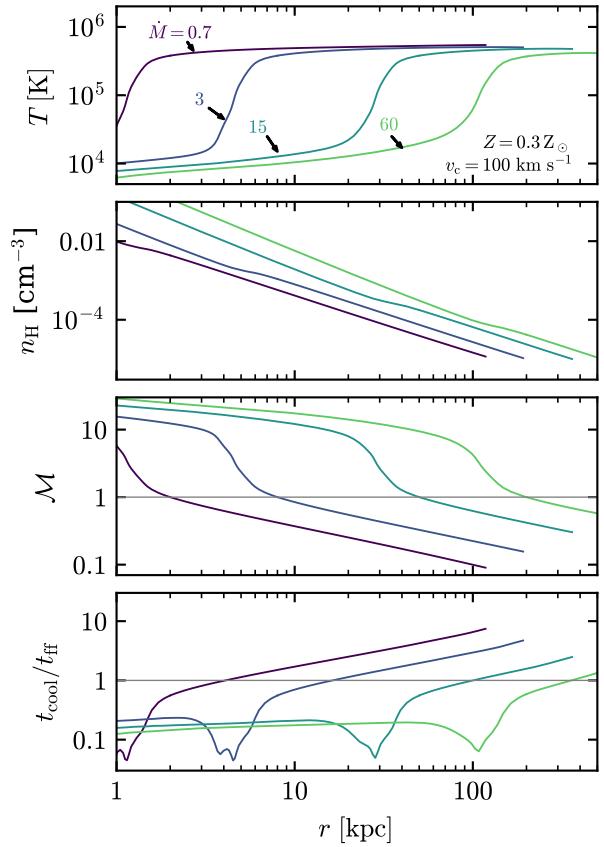


Figure 1. Spherically-symmetric cooling flow solutions for the volume-filling gas phase in an isothermal potential with $v_c = 100 \text{ km s}^{-1}$, corresponding to $M_{\text{halo}} = 4 \cdot 10^{11} M_\odot$ at $z = 0$. The panels show the temperature, density, Mach number, and $t_{\text{cool}}/t_{\text{ff}}$ of the solutions. The four solutions are derived assuming no angular momentum and $Z = 0.3 Z_\odot$, and differ in the assumed density normalization (the green solution is baryon complete within $R_{\text{vir}} = 200 \text{ kpc}$). The solutions are transonic, forming a cool supersonic flow with $t_{\text{cool}} < t_{\text{ff}}$ within the sonic radius. A higher density normalization corresponds to a higher inflow rate \dot{M} (noted in $M_\odot \text{ yr}^{-1}$ in the top panel) and to a larger sonic radius.

$t_{\text{cool}}/t_{\text{ff}}$ is comparable or larger than unity (bottom panel). In this region radiative cooling is balanced via heating by compression as the gas flows inward. This subsonic region can be approximated by the following self-similar solutions³ to the flow equations (13)–(15), which are derived in the subsonic limit ($\mathcal{M}^2 \ll 1$):

$$\epsilon = \frac{9}{10} c_s^2 = v_c^2 \quad (16)$$

$$n_H = \sqrt{\frac{\dot{M} v_c^2}{4\pi \Lambda}} r^{-3/2} \quad (17)$$

$$|v_r| = \frac{r}{t_{\text{cool}}} = \frac{X}{m_p} \sqrt{\frac{\dot{M} \Lambda}{4\pi v_c^2}} r^{-1/2} \quad (18)$$

where $c_s = \sqrt{(10/9)\epsilon}$ is the adiabatic sound speed. From

³ These solutions correspond to the $m = 0$ solutions in Paper I, where m is defined such that $v_c(r) \propto r^m$.

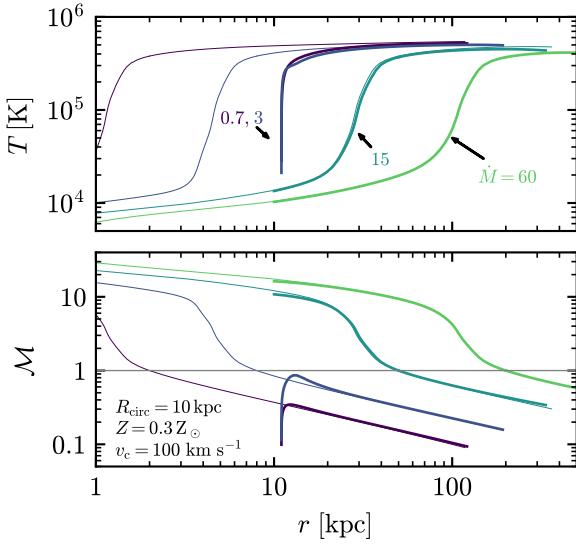


Figure 2. Spherically-symmetric cooling flow solutions for gas in an isothermal potential with $v_c = 100 \text{ km s}^{-1}$. Thick lines plot solutions for a flow with uniform specific angular momentum, corresponding to rotational support at a radius $R_{\text{circ}} = 10 \text{ kpc}$. The values of \dot{M} are the same as in the no angular momentum solutions in Fig. 1 (plotted here as thin lines). In the blue and purple solutions the flow cools just outside R_{circ} and reaches R_{circ} with a vanishing radial velocity. In the green and yellow solutions the flow cools at $R_{\text{sonic}} > R_{\text{circ}}$, and reaches R_{circ} supersonically. Solutions corresponding to hot mode accretion throughout the halo are possible only if $R_{\text{sonic}} < R_{\text{circ}}$ (where R_{sonic} is calculated in the no-angular momentum limit), or equivalently if $\dot{M} < \dot{M}_{\text{crit}}$ (eqn. 8).

eqns. (16) and (18), the Mach number in the self-similar solution is equal to

$$\mathcal{M} \equiv \frac{|v_r|}{c_s} = \frac{X}{m_p} \sqrt{\frac{9\dot{M}\Lambda}{40\pi v_c^4}} r^{-1/2}, \quad (19)$$

i.e. \mathcal{M} increases inwards as in Fig. 1. The flow thus turns supersonic roughly at

$$R_{\text{sonic}} \approx \frac{9\dot{M}X^2\Lambda}{40\pi m_p^2 v_c^4}, \quad (20)$$

where this estimate is approximate due to the inaccuracy of estimating R_{sonic} using a solution in the subsonic limit. The estimate for R_{sonic} in eqn. (20) is also an estimate of the radius where $t_{\text{cool}} \approx t_{\text{ff}}$, since

$$\frac{t_{\text{cool}}}{t_{\text{ff}}} = \sqrt{\frac{9}{20}} \frac{t_{\text{cool}}}{r/c_s} = \sqrt{\frac{9}{20}} \mathcal{M}^{-1} \quad (21)$$

where here we used eqns. (4), (16) and (18). The sonic radius is hence roughly the radius where $t_{\text{cool}}/t_{\text{ff}} = \sqrt{9/20} \approx 0.7$ (see bottom panels of Fig. 1). In section 2.1 we used this prefactor and eqn. (16) to derive eqn. (6), though note that these relations are accurate only in the self-similar case, and should be considered approximate in the general case.

Fig. 1 shows that in the inner supersonic part of the solutions the flow rapidly loses thermal energy until it reaches the equilibrium temperature $T_{\text{eq}} \sim 10^4 \text{ K}$. At the transition t_{cool} is too short to be compensated by heating due to advection as in the subsonic regime, so the temperature decreases,

which due to the shape of the cooling function causes the cooling to accelerate and thus further decrease the temperature. As a result of this rapid cooling process the temperature drops by a factor of $\gtrsim 10$ over merely a factor of ≈ 2 in radius. The sonic radius thus roughly corresponds to the radius where the flow transitions from a flow which is supported against gravity via its thermal pressure, to a free-falling solution which has no support.

2.2.2 Cooling flows with angular momentum

Dark matter halos and the baryons associated with them are expected to have angular momentum, due to tidal torques induced by neighbouring halos. To include the effects of angular momentum in the 1D cooling flow solutions we assume a uniform specific angular momentum equal to $v_c R_{\text{circ}}$, and modify the momentum equation to (see e.g. Cowie et al. 1980; BD03):

$$\frac{1}{2} \frac{dv_r^2}{dr} = -\frac{1}{\rho} \frac{dP}{dr} - \frac{v_c^2}{r} \left[1 - \left(\frac{R_{\text{circ}}}{r} \right)^2 \right]. \quad (22)$$

This equation applies to a flow within the plane defined by the angular momentum vector. To derive solutions relevant for hot mode accretion, we search for solutions to the modified flow equations which satisfy $v_r \rightarrow 0$ as $r \rightarrow R_{\text{circ}}$, i.e. the flow stalls at the circularization radius⁴. These solutions correspond to a radial inflow at $r \gg R_{\text{circ}}$ supported by thermal pressure which connects to a rotating flow at $r \sim R_{\text{circ}}$ supported by angular momentum. We assume $R_{\text{circ}} = 10 \text{ kpc}$ as in eqn. (8). We impose the same marginally-bound outer boundary condition as used for the transonic solutions discussed in the previous section. The thick blue and purple lines in Figure 2 plot two such solutions, for \dot{M} equal to 0.7 and $3 M_\odot \text{ yr}^{-1}$ as in the corresponding non-rotating solutions from Fig. 1 (also plotted in Fig. 2 as thin lines). The rotating and non-rotating solutions differ significantly only at $r \lesssim 1.5 R_{\text{circ}}$, where the rotating solutions stall while the non-rotating solutions continue to accelerate inward.

For the higher values of \dot{M} of 15 and $60 M_\odot \text{ yr}^{-1}$ corresponding to the green and yellow solutions, no solutions which stall at R_{circ} are possible. A transonic solution with a specific \dot{M} is fully-defined, and thus cannot be made to satisfy a specific boundary condition at R_{circ} , as is possible for the blue and purple solutions which are subsonic at all $r > R_{\text{circ}}$. The thick green and yellow lines in Fig. 2 plot the corresponding transonic solution when angular momentum is included in the momentum equations. These solutions are almost identical to the no-angular momentum solutions down to R_{circ} , and indicate that even when angular momentum is included the flow reaches R_{circ} with supersonic speeds (where it would presumably shock in a more realistic calculation).

The conclusion from Fig. 2 is that only if the condition

$$R_{\text{sonic}} \lesssim R_{\text{circ}} \quad (23)$$

is satisfied, where R_{sonic} is calculated in the no-angular momentum limit, then the flow can reach R_{circ} with $T \approx T_{\text{vir}}$

⁴ In practice, we integrate outward from $R_0 = R_{\text{circ}}(1 + \epsilon_1)$ assuming $v_r(R_0) = \epsilon_2 v_c$, with $\epsilon_1 = \epsilon_2 = 0.03$.

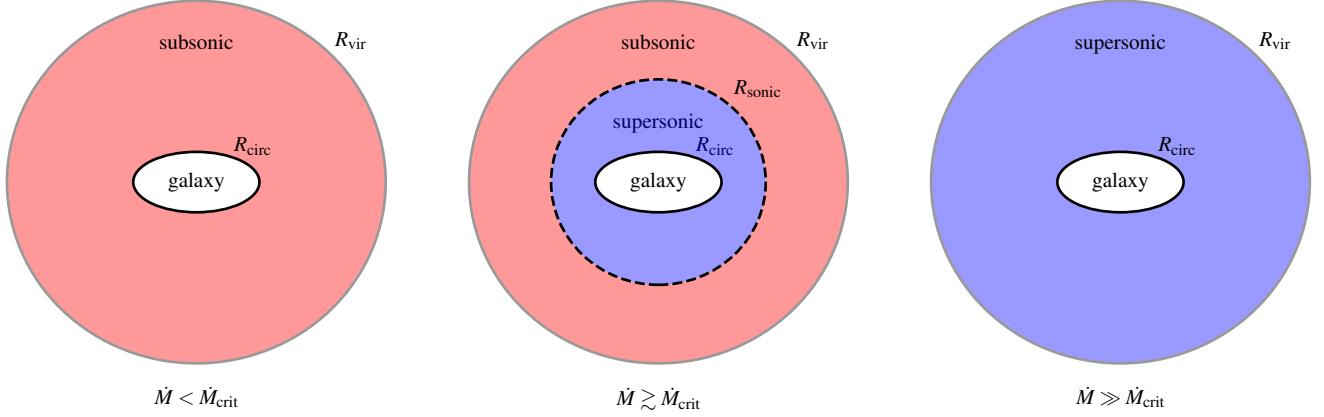


Figure 3. A cartoon picturing the three types of cooling flow solutions discussed in section 2.2 and shown in Fig. 2. (**Left**) When $\dot{M} < \dot{M}_{\text{crit}}$ the flow is subsonic and hot ($T \approx T_{\text{vir}}$) down to the circularization radius. (**Middle**) When $\dot{M} \gtrsim \dot{M}_{\text{crit}}$ the flow goes through a sonic point on the halo scale, and reaches the galaxy as a cool ($T \approx 10^4$ K) supersonic flow. (**Right**) When $\dot{M} \gg \dot{M}_{\text{crit}}$ the flow is supersonic and free falling at all halo radii. We consider only the scenario depicted in the left panel as ‘hot mode accretion’.

and a vanishing radial velocity. If condition (23) is violated as in the green and yellow solutions, than the flow necessarily reaches R_{circ} supersonically.

Figure 3 depicts the three types of solutions discussed in this section. The left panel pictures a cooling flow with $\dot{M} < \dot{M}_{\text{crit}}$, where without angular momentum the sonic radius would be within R_{circ} . In this regime the flow is pressure supported on all halo scales, i.e. the flow is subsonic and has $T \approx T_{\text{vir}}$, down to the radius where the flow is supported by angular momentum. This type of flow corresponds to the classic ‘hot accretion mode’. The right panel plots solutions with \dot{M} sufficiently large such that the sonic radius is beyond the virial radius and hence potentially outside the accretion shock – the outer boundary of the region in which the cooling flow solutions could be valid, since beyond the accretion shock we expect a supersonic flow. This regime corresponds to the classic cold flow regime where gas accreting from the IGM free-falls all the way down to the galaxy. The middle panel plots cooling flow solutions with $R_{\text{circ}} < R_{\text{sonic}} < R_{\text{vir}}$, i.e. the sonic radius is within the range of radii where the cooling flow solutions could be valid. In this regime the gas is hot and pressure-supported in the outer halo, but gas in the inner halo and specifically the gas accreting onto the galaxy is cool and free-falling. In an isothermal potential this scenario applies if $1 < \dot{M}/\dot{M}_{\text{crit}} < R_{\text{vir}}/R_{\text{circ}} \approx 20$, since the sonic radius scales linearly with \dot{M} (eqn. 20). However, since $R_{\text{sonic}} \propto v_{\text{c}}^{-4}$, even a weak decrease of v_{c} with increasing radius would imply that R_{sonic} reaches R_{vir} at inflow rates smaller than $20\dot{M}_{\text{crit}}$, and this intermediate regime would be relevant only over a smaller range of \dot{M} .

Due to the similarity of \mathcal{M}^{-1} and $t_{\text{cool}}/t_{\text{ff}}$ in cooling flows (eqn. 21), the condition (23) is equivalent to the condition $t_{\text{cool}} \gtrsim t_{\text{ff}}$ at R_{circ} used in section 2.1. Using eqn. (23) in eqn. (20) yields the maximum accretion rate of the hot mode \dot{M}_{crit} , which is given by eqn. (8). We note also that Quataert & Narayan (2000) previously discussed the importance of the sonic radius in cooling flows in isothermal potentials,

and the associated mass inflow rate, in the context of the interstellar medium of elliptical galaxies.

2.3 Hydrodynamic simulations

In this section we use an idealized 3D hydrodynamic simulation to support the above analytic results. The simulation is based on the simulations in Fielding et al. (2017), though it does not include galaxy feedback.

We use the grid-based hydrodynamics code ATHENA++ (Stone et al., submitted⁵) in a spherical-polar coordinate system. The computational domain spans $r = 1 \text{ kpc} - 10 \text{ Mpc}$, $\pi/4 \leq \theta \leq 3\pi/4$, and $\pi/4 \leq \phi \leq 3\pi/4$, where θ and ϕ are the polar and azimuthal angles. The grid has 64 cells in each angular direction and 384 logarithmically-spaced cells in r , which give approximately 1:1 cell aspect ratios. We solve the standard hydrodynamics equations with additional source terms to include a static gravitational potential and radiative cooling, using the same cooling function and constant $v_{\text{c}} = 100 \text{ km s}^{-1}$ as used to derive the steady-state solutions in Figures 1 – 2. Self-gravity of the gas is neglected. We adopt periodic boundary conditions in the polar and azimuthal directions, while in the radial direction we adopt outflow boundary conditions.

We run a simulation where \dot{M} increases slowly with time, from $\dot{M} \ll \dot{M}_{\text{crit}}$ to $\dot{M} \gg \dot{M}_{\text{crit}}$, in order to simulate the different cooling flow regimes depicted in Fig. 3. To achieve this goal, we initialize the gas with a hydrostatic pressure profile at all radii, which yields the following dependence of \dot{M} on time (Bertschinger 1989, hereafter B89):

$$\dot{M}_{\text{B89}}(t, r < R_{\text{cool}}) \approx 4\pi R_{\text{cool}}^2 \rho(R_{\text{cool}}) \frac{dR_{\text{cool}}}{dt}. \quad (24)$$

where R_{cool} is the cooling radius at which $t_{\text{cool}} = t$. This relation represents a ‘cooling wave’ expanding in the initially

⁵ <https://princetonuniversity.github.io/athena/index.html>

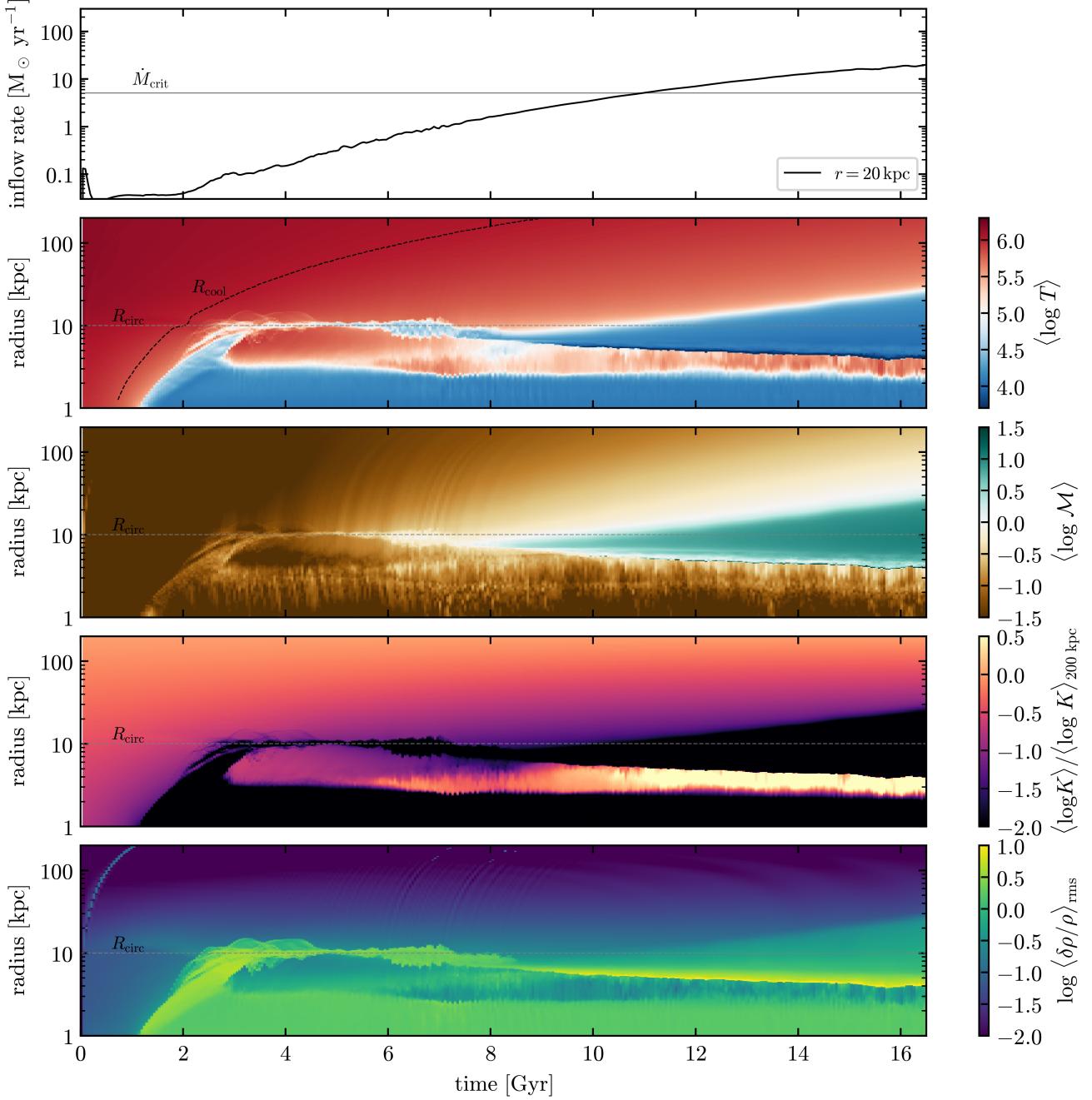


Figure 4. Hydrodynamic 3D simulation of radiatively-cooling gas in an isothermal potential with $v_c = 100 \text{ km s}^{-1}$. Third-solar metallicity is assumed throughout. The gas is initialized at $t = 0$ with a hydrostatic pressure profile, and a density profile chosen to produce the dependence of inflow rate on time shown in the top panel. Initial conditions also include a uniform specific angular momentum corresponding to $R_{\text{circ}} = 10 \text{ kpc}$, and density fluctuations with amplitude $\langle \delta\rho/\rho \rangle_{\text{rms}} = 0.03$. The three middle panels plot mass-weighted shell averages of temperature, Mach number and entropy, as a function of radius and time. Entropy is normalized by the value at 200 kpc at each time. The bottom panel plots the density dispersion in the shells. The critical inflow rate \dot{M}_{crit} and cooling radius R_{cool} are noted in the top two panels. At $3 \lesssim t \lesssim 7 \text{ Gyr}$ when $R_{\text{cool}} > R_{\text{circ}}$ and $\dot{M} \ll \dot{M}_{\text{crit}}$ the halo gas forms a subsonic cooling flow corresponding to hot mode accretion – the inward flow remains hot down to R_{circ} . At $t > 10 \text{ Gyr}$ when $\dot{M} > \dot{M}_{\text{crit}}$ the halo gas forms a transonic flow, with hot subsonic gas overlying cool supersonic gas. Density perturbations develop significantly only in supersonic regions or at $r \lesssim R_{\text{circ}}$.

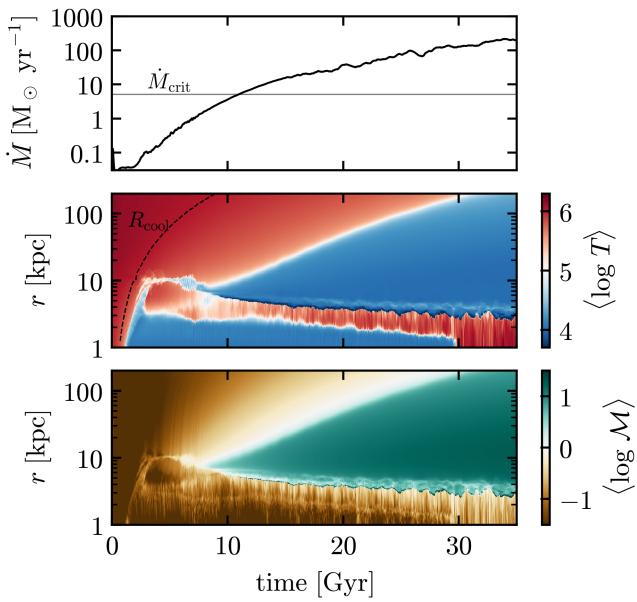


Figure 5. Similar to Figure 4, spanning the entire simulation time. As \dot{M} increases the sonic radius moves outwards. At late times when \dot{M} is sufficiently large, cool gas falls freely throughout the halo, corresponding to the classic ‘cold mode accretion’.

static medium at a velocity dR_{cool}/dt , as gas at increasingly larger radii starts cooling and joins the cooling flow. For a constant initial temperature with sound speed c_s , hydrostatic equilibrium gives $\rho \propto r^{-\alpha}$ with $\alpha = \gamma v_c^2/c_s^2$. The cooling time thus scales as $t_{\text{cool}} \propto r^\alpha$ (eqn. 5), and the cooling radius as $R_{\text{cool}} \propto t^{1/\alpha}$. Equation (24) hence yields

$$\dot{M}_{\text{B89}}(t) \propto t^{3/\alpha-2}, \quad (25)$$

where the constant of the proportionality is determined by the normalization of the initial density profile. As noted by B89, eqns. (24)–(25) are valid only if $dR_{\text{cool}}/dt \ll c_s$, i.e. if the cooling wave is expanding slowly compared to the sound-crossing time. Otherwise, gas at different radii cools out monolithically, and the pressure profile does not have time to adjust to the cooling flow solution. In this latter case the halo gas collapses into a supersonic free-falling solution rather than forming a cooling flow.

We also assume a finite initial velocity in the ϕ direction v_ϕ such that all cells outside R_{circ} have the same specific angular momentum, corresponding to $R_{\text{circ}} = 10$ kpc as in Fig. 2. Within R_{circ} we assume in the initial conditions $v_\phi = v_c$. To avoid the accumulation of $\sim 10^4$ K gas at $\lesssim R_{\text{circ}}$ in the simulation, we implement ‘star formation’ by removing gas that satisfies $T < 3 \cdot 10^4$ K and $n_H > 0.03 \text{ cm}^{-3}$. Testing indicates that the exact parameters of this prescription do not affect the results except where noted below. We also impose small isobaric density perturbations in the initial conditions with an amplitude $\langle \delta \rho / \rho \rangle_{\text{rms}} = 0.03$ and a white-noise spectrum.

The simulation is initialized with a density slope of $\alpha = -0.5$ and a hydrogen particle density $n_H = 10^{-4.5} \text{ cm}^{-3}$ at $r = 100$ kpc, and is run for 35 Gyr. We emphasize that these initial conditions are intended to yield a desired $\dot{M}(t)$ rather than

to describe a realistic halo. The resulting $\dot{M}(t)$ and $R_{\text{cool}}(t)$ are plotted in the top two panels of Figure 4. At $t = 10.3$ Gyr the inflow rate exceeds $\dot{M}_{\text{crit}} = 4.0 \text{ M}_\odot \text{ yr}^{-1}$, where \dot{M}_{crit} is calculated from eqn. (8). The plotted \dot{M} is measured at $r = 20$ kpc in the simulation, though this choice does not affect the time at which \dot{M} crosses \dot{M}_{crit} as long as it is outside R_{circ} and within R_{cool} .

The lower four panels of Fig. 4 plot shell-averaged properties in the simulation as a function of r and t . From top to bottom the panels show the mass-weighted averages of $\log T$, $\log \mathcal{M}$, and $\log K/K(r = 200 \text{ kpc})$, and the density dispersion $\langle \delta \rho / \rho \rangle_{\text{rms}}$. The figure shows that at early times $t \lesssim 2$ Gyr the value of R_{cool} is smaller than R_{circ} and the gas properties remain near the initial conditions. At later times $3 \lesssim t \lesssim 6$ Gyr when $R_{\text{cool}} > R_{\text{circ}}$ and $\dot{M} \ll \dot{M}_{\text{crit}}$, the gas within R_{cool} forms a subsonic cooling flow, where the gas temperature is near virial and the entropy declines inward. Only very close to $R_{\text{circ}} = 10$ kpc the flow cools out, as suggested by the 1D solutions with $\dot{M} = 0.7$ and $3 \text{ M}_\odot \text{ yr}^{-1}$ in Fig. 2. At radii smaller than R_{circ} the flow is hot down to ≈ 3 kpc, and comprises of inflow outside the midplane which overshoots R_{circ} and is then repelled back by the centrifugal force (see Fig. 7 below). The lowest panel in Fig. 4 shows that in this $\dot{M} < \dot{M}_{\text{crit}}$ regime the amplitude of density fluctuations is $\ll 1$ beyond R_{circ} , as expected in subsonic cooling flows (see Paper I and references therein).

Fig. 4 shows that at $t > 10$ Gyr when \dot{M} exceeds \dot{M}_{crit} the flow is transonic with a subsonic region ‘overlying’ a supersonic region, as depicted in the middle panel of Fig. 3. The sonic point is evident as the upper white contour in the \mathcal{M} panel, and it moves outward as \dot{M} increases in the simulation (see Figure 5). We measure R_{sonic} as the outermost shell with $\langle \log \mathcal{M} \rangle = 1$, and plot the relation between R_{sonic} and $\dot{M}(r = 20 \text{ kpc})$ in Figure 6. When $\dot{M} > \dot{M}_{\text{crit}}$, the relation in the simulation is similar to the steady-state, no-angular-momentum calculation (eqn. 20, blue line). When $\dot{M} \lesssim \dot{M}_{\text{crit}}$, i.e. at $7.5 \lesssim t \lesssim 10$ Gyr, the simulation has a sonic point somewhat within R_{circ} , while when $\dot{M} \ll \dot{M}_{\text{crit}}$ the flow in the simulation is entirely subsonic.

Towards the end of the simulation where $\dot{M} \gtrsim 20 \dot{M}_{\text{crit}}$, Fig. 5 shows that the sonic radius exceeds 200 kpc, and the flow is supersonic at all halo scales, i.e. the scenario depicted in the right panel of Fig. 3. We note that in our simulation the flow is transonic even at these late times since there is hot quasi-static gas out to the outer boundary at 10 Mpc, i.e. we effectively assume the accretion shock is at infinity. In a realistic system with a finite accretion shock radius we expect the flow to be purely supersonic if \dot{M} corresponds to R_{sonic} larger than the shock radius.

Figs. 4 and 5 demonstrate that once the flow crosses the sonic radius it cools quickly, as suggested by the steady-state solutions shown in Fig. 1. This rapid cooling is associated with a rapid growth of thermal instabilities (bottom panel in Fig. 4, see also Mathews & Bregman 1978 and Balbus & Soker 1989). The association of the sonic radius with the rapid growth of instabilities occurs since within the sonic radius $t_{\text{cool}} \ll r/v_r$ (see figure 3 in Paper I), so the instabilities grow faster than the rate at which the flow is advected inward, in contrast with the subsonic region where $t_{\text{cool}} \approx r/v_r$ (eqn. 18). The supersonic flow reaches a radius which is substantially smaller than R_{circ} , and is evident as a boundary

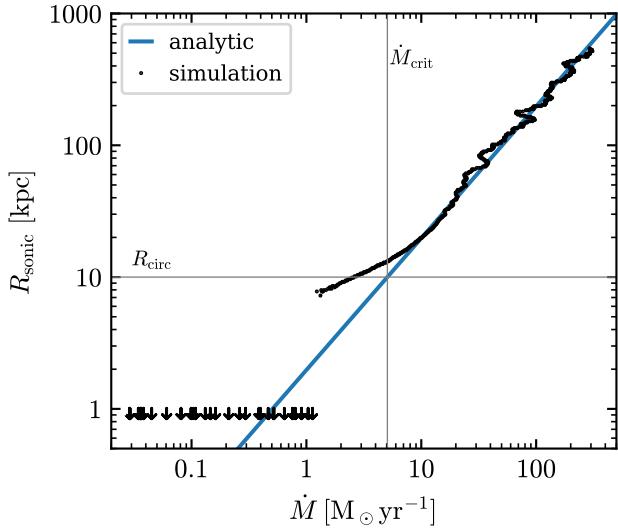


Figure 6. The relation between \dot{M} and $R_{\rm sonic}$. Each dot or arrow corresponds to a snapshot in the simulation shown in Figs. 4–5, while down-pointing arrows at 1 kpc denote snapshots where the flow is entirely subsonic. The cyan line plots the analytic relation (eqn. 20). The simulation and analytic calculations roughly agree when $R_{\rm sonic} > R_{\rm circ}$ and $\dot{M} > \dot{M}_{\rm crit}$.

in all properties plotted in Fig. 4. This minimum radius decreases with increasing \dot{M} .

In Figure 7 we plot streamlines and temperature maps in the meridional plane, mass-weighted over the ϕ -coordinate. The three panels plot snapshots at $t = 4 \text{ Gyr}$ (top), $t = 15 \text{ Gyr}$ (middle) and $t = 30 \text{ Gyr}$ (bottom), corresponding to the hot ($\dot{M} \ll \dot{M}_{\rm crit}$), transonic ($\dot{M} \gtrsim \dot{M}_{\rm crit}$) and purely cold ($\dot{M} \gg \dot{M}_{\rm crit}$) accretion phases. The panels are shaped as wedges similar to half the simulated domain, between $\theta = \pi/2$ (bottom axis) and $\theta = \pi/4$ (top diagonal axis). The streamlines emanate from large radii and are initially evenly spaced in θ , indicating a radial inflow. In the hot accretion mode plotted on top the streamlines converge onto $r = R_{\rm circ}$ and $\theta = \pi/2$, i.e. on the equilibrium position for our assumed specific angular momentum, which corresponds to a ‘ring’ in 3D space. Streamlines initially far away from the midplane first overshoot $R_{\rm circ}$ and reach somewhat smaller radii, and then turn outward as the centrifugal force overcomes gravity. Note though that this latter effect may be artificially enhanced by the boundary of our domain at $\theta = \pi/4$ and hence the lack of streamlines which feed gas and provide pressure support at smaller radii. At the equilibrium position the gas is cool (see also temperature panel of Fig. 4), though the flow cools out just before joining the ring – the region with $T \lesssim 10^5 \text{ K}$ spans $\lesssim 3$ grid cells in the θ direction and $\lesssim 9$ grid cells in the r direction. In the ring our prescription for ‘star formation’ acts as a sink for gas when the density exceeds 0.03 cm^{-3} .

In contrast with the hot accretion mode, in the other two regimes shown in the bottom panels of Fig. 7 the streamlines reach the midplane at a radius of 4–5 kpc, substantially smaller than the equilibrium position at $r = R_{\rm circ}$. This is possible due to the lack of pressure support and high inertia of the flow, which allows a cool flow to ‘overshoot’ the angular

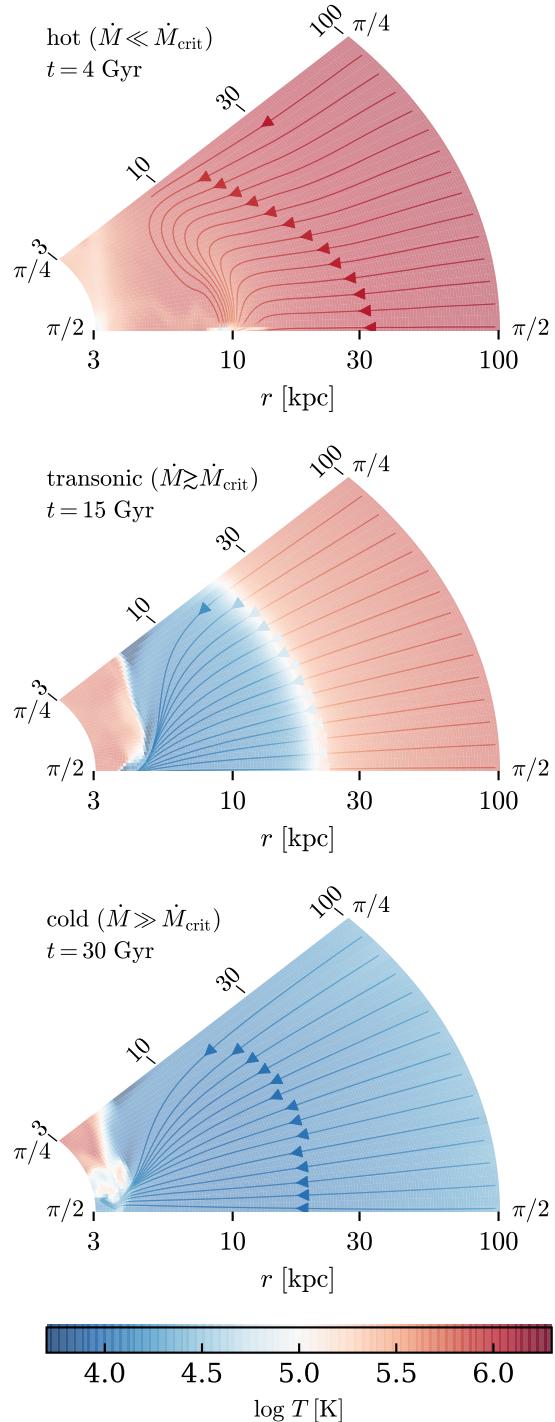


Figure 7. Streamlines and temperature maps in the meridional plane. The panels plot three snapshots of the simulation, corresponding to the hot accretion mode ($t = 4 \text{ Gyr}$, top), the transonic accretion mode ($t = 15 \text{ Gyr}$, middle), and the purely cold accretion mode ($t = 30 \text{ Gyr}$, bottom). The bottom axis of each panel corresponds to the midplane ($\theta = \pi/2$) while the top diagonal axis corresponds to the boundary of the simulated domain at $\theta = \pi/4$. The equilibrium position given the angular momentum of the simulated gas is at $r = R_{\rm circ} = 10 \text{ kpc}$ and $\theta = \pi/2$. In the hot accretion mode the radial flow at large scales converges onto the equilibrium position, at which point the gas cools and is removed from the simulation via our SF prescription. In the other two accretion modes the flow reaches the midplane and is lost to SF at radii smaller than the equilibrium position.

lar momentum barrier. In our simulation the gas is lost to SF at these inner radii, which causes the flowlines to end at a non-equilibrium position. In a similar simulation without the SF prescription the flow circles back to the equilibrium position after crossing the midplane at radii $< R_{\text{circ}}$.

To conclude this section, the flow structure formed in the 3D simulation suggests that the 1D steady-state cooling flow solutions capture the transition between hot and cold mode accretion reasonably well, at least in our idealized setup. If $\dot{M} < \dot{M}_{\text{crit}}$ then the flow ‘quietly’ accretes onto the galaxy disk from a hot ($\approx T_{\text{vir}}$) rotating atmosphere with a vanishing radial velocity, as can be seen in the purple and blue solutions in Fig. 2 and in the top panel of Fig. 7. In contrast if $\dot{M} > \dot{M}_{\text{crit}}$ then $R_{\text{sonic}} > R_{\text{circ}}$ and the gas reaches the galaxy scale as a cool ($\approx T_{\text{eq}}$) supersonic flow, as can be seen in the green and yellow solutions in Fig. 2 and in the two bottom panels of Fig. 7.

2.4 Comparison to the condition for shock stability

The simulation in the previous section and the simulations in Paper I demonstrate that halo gas which is initially hydrostatic converges onto the family of cooling flow solutions (as long as $dR_{\text{cool}}/dt \ll c_s$, see B89). A related question is under which conditions a flow which is initially supersonic⁶ will shock and form a cooling flow. Note that such a transition is non-trivial only if R_{sonic} of the cooling flow that forms is smaller than the outer boundary of the system (e.g. the two left panels in Fig. 3), since otherwise the cooling flow solution is also a supersonic solution (right panel in Fig. 3). This question was addressed by BD03, who argued that supersonic inflows in dark matter halos shock once the conditions for an accretion shock to expand are met at the disk radius $\approx R_{\text{circ}}$. In this section we show that the condition for an expanding accretion shock at a shock radius $R_{\text{sh}} = R_{\text{circ}}$ is equivalent to the condition $R_{\text{sonic}} < R_{\text{circ}}$ derived here for the onset of hot mode accretion. We show this equivalence by utilizing the expectation that the postshock gas forms a cooling flow⁷.

The shock jump condition is

$$v_{\text{sh}} = -\frac{1}{3}(v_0 - 4v_1), \quad (26)$$

where v_{sh} , v_0 , and v_1 are the shock, preshock, and postshock velocities, all measured in the halo frame (inflows have a negative velocity), and for simplicity we assume a strong shock. Note that the postshock velocity in the shock frame $v_1 - v_{\text{sh}}$ must be subsonic, so if v_1 is supersonic v_{sh} must be

⁶ The family of supersonic solutions to eqns. (13)–(15) satisfies $T \approx T_{\text{eq}} \approx 10^4$ K and v_r roughly equal to the free-fall velocity. As ρ is essentially unconstrained, one can find such a supersonic solution for any assumed value of \dot{M} .

⁷ This expectation is not strictly valid, since cooling flows are steady-state solutions while gas immediately within the shock radius is likely not time-steady. The cooling flow solutions for the shocked gas are expected to be accurate only out to radii somewhat smaller than the shock radius (see Paper I). We neglect this complication, and in the next section support our conclusion on the similarity of the two formalisms by comparing the values they yield for M_{thres} .

negative, i.e. the shock is contracting. It hence follows that a necessary condition for an expanding accretion shock is $M(R_{\text{sh}}) < 1$, or equivalently in a cooling flow $R_{\text{sonic}} < R_{\text{sh}}$. To show that the $R_{\text{sonic}} < R_{\text{sh}}$ condition is likely to be also a sufficient condition for an expanding shock, we replace v_1 with $R_{\text{sh}}/t_{\text{cool}}$ (eqn. 18):

$$v_{\text{sh}} = -\frac{1}{3}\left(v_0 + 4\frac{R_{\text{sh}}}{t_{\text{cool}}}\right) \quad (27)$$

Using the definition of t_{ff} (eqn. 4) and extracting v_0 from the parentheses we get

$$v_{\text{sh}} = -\frac{v_0}{3} \cdot \left(1 - \frac{t_{\text{ff}}/t_{\text{cool}}}{|v_0|/(\sqrt{8}v_c)}\right), \quad (28)$$

where all quantities are estimated at R_{sh} . Approximating the inflow velocity as $|v_0| \approx v_c(1 + \sqrt{2 \ln(R_{\text{vir}}/R_{\text{sh}})})$, as expected for an inflow ‘dropped’ from $2R_{\text{vir}}$ in an NFW potential, we get $|v_0|/\sqrt{8}v_c = 1.2$ for $R_{\text{sh}} = 0.05R_{\text{vir}}$. Eqn. (28) thus implies that if $t_{\text{ff}}/t_{\text{cool}} \lesssim 1.2$ at the shock radius, then the term in the brackets is positive and the shock expands outward. Since $t_{\text{ff}}/t_{\text{cool}} \approx 1.5\mathcal{M}$ (eqn. 21), we get that if $M(R_{\text{circ}}) \lesssim 0.8$ then $v_{\text{sh}} > 0$, i.e. if the condition $R_{\text{sonic}} \lesssim R_{\text{circ}}$ is satisfied the shock would expand.

The above derivation suggests that the BD03 condition for shock stability at R_{circ} is similar to the condition for hot mode accretion $R_{\text{sonic}} \lesssim R_{\text{circ}}$ derived in this work. It is important to note though that our derivation does not assume an accretion shock exists, in contrast with the derivation of BD03. Rather, our derivation is based solely on the properties of radiatively cooling gas with $T \approx T_{\text{vir}}$, regardless of whether the gas was heated to this temperature in a single shock, in a series of shocks, or by feedback at earlier epochs. Our analysis thus suggests that the conditions under which hot mode accretion is possible apply more generally.

In the simulations in BD03 the halo mass grows with time, so $\dot{M}/\dot{M}_{\text{crit}}$ decreases since \dot{M}_{crit} increases faster than \dot{M} (see next section). The initially supersonic flows in BD03 though shock directly into subsonic flows, without going through an intermediate transonic cooling flow phase. We have verified this behavior using a setup similar to that in the previous section but with supersonic initial conditions. We set the outer boundary condition so \dot{M} decreases with time from an initial $\dot{M} \gg \dot{M}_{\text{crit}}$, and indeed a shock and subsonic cooling flow form only when $\dot{M} \lesssim \dot{M}_{\text{crit}}$, while when $\dot{M} \gtrsim \dot{M}_{\text{crit}}$ the flow remains purely supersonic rather than forming a transonic flow. A possible limitation of this simulation and the simulation in BD03 is the lack of sufficiently strong shocks beyond R_{circ} . If the flow shocks in the supersonic part of the flow the shock cannot propagate outward, and hence the subsonic part of the cooling flow will not form. However, in the presence of outflows from the galaxy the inflows from the IGM are expected to experience strong shocks at radii $\gg R_{\text{circ}}$ (e.g. Fielding et al. 2017), and thus supersonic flows may shock directly into transonic cooling flows. We leave exploring this possibility to future work.

2.5 The condition for cooling-regulated accretion

White & Frenk (1991) argued that the condition

$$t_{\text{cool}}(R_{\text{vir}}) = t_{\text{H}} \quad (29)$$

separates between ‘cooling-limited’ systems in which accretion is regulated by radiative cooling, and ‘supply-limited’ systems in which accretion is regulated by the inflow rate from the IGM. We now show that eqn. (29) is similar to the condition $t_{\text{cool}} = t_{\text{ff}}$ at R_{circ} derived above for the onset of hot mode accretion. This similarity follows since in cooling flows $t_{\text{cool}}/t_{\text{ff}} \propto r^{1/2}$ (eqn. 21), so in the critical solution $t_{\text{cool}}/t_{\text{ff}} \propto (r/R_{\text{circ}})^{1/2}$ and hence

$$t_{\text{cool}}(R_{\text{vir}}, \dot{M} = \dot{M}_{\text{crit}}) \approx \sqrt{\frac{R_{\text{vir}}}{R_{\text{circ}}}} t_{\text{ff}}(R_{\text{vir}}) \approx t_{\text{H}} \quad (30)$$

where the last approximation follows from $R_{\text{vir}}/R_{\text{circ}} \approx 20$ and $t_{\text{H}} \approx 5t_{\text{ff}}(R_{\text{vir}})$.⁸ Thus, systems with $\dot{M} < \dot{M}_{\text{crit}}$ are cooling-limited according to the condition (29), while systems with $\dot{M} > \dot{M}_{\text{crit}}$ are ‘supply-limited’, even if the halo gas has shocked and forms a transonic cooling flow.

We note in passing that semi-analytic models such as Somerville & Primack (1999) which employ the condition (29) could be improved by considering the halo gas density profile $\rho \propto r^{-1.5}$ suggested by the physical cooling flow solution (eqn. 17), rather than say an isothermal profile with $\rho \propto r^{-2}$.

3 THE CRITICAL COOLING RATE AS A FUNCTION OF HALO AND GAS PARAMETERS

We now use eqn. (8) to evaluate \dot{M}_{crit} as a function of halo parameters. We use the following virial relations:

$$R_{\text{vir}} = \frac{v_c}{\sqrt{\frac{\Delta_c(z)}{2}} H(z)} = 263 M_{12}^{1/3} E^{-2/3}(z) \text{kpc} \quad (31)$$

$$v_{\text{vir}} = \left(\sqrt{\frac{\Delta_c}{2}} H G M_{\text{halo}} \right)^{1/3} = 128 M_{12}^{1/3} E^{1/3}(z) \text{km s}^{-1} \quad (32)$$

where $M_{\text{halo}} \equiv 10^{12} M_{12} M_{\odot}$, Δ_c is the virial overdensity with respect to the critical density from Bryan & Norman (1998), $H(z)$ is the Hubble parameter at redshift z , and we absorbed the redshift-dependent term $\sqrt{\Delta_c(z)} H(z)/\sqrt{\Delta_c(0)} H_0$ into a function $E(z)$, which is equal to

$$E(z) = \sqrt{\frac{\Delta_c(z)}{102}} [1 - \Omega_{\text{m},0} + \Omega_{\text{m},0}(1+z)^3] \approx (1+z)^{0.9} \quad (33)$$

where the approximation is accurate to 25% at $0 < z < 10$. The critical accretion rate is then derived using eqns. (31), (32) and (9) in eqn. (8):

$$\begin{aligned} \dot{M}_{\text{crit}} &= \frac{2^{1/6} \cdot 4\pi G^{5/3} m_p^2}{X^2} f_{v_c}^3 f_{\lambda} \lambda \Delta_c^{1/3} H^{2/3} M_{\text{halo}}^{5/3} \Lambda^{-1} \\ &= 10.6 f_{v_c}^3 f_{\lambda} \lambda_{0.035} \Lambda_{-22}^{-1} M_{12}^{5/3} E^{2/3}(z) M_{\odot} \text{yr}^{-1} \end{aligned} \quad (34)$$

where we defined $f_{v_c} \equiv v_{\text{vir}}/v_c(R_{\text{vir}})$ which depends both on the halo concentration and on the properties of the galaxy. In the numerical evaluation in eqn. (34) we used $X = 0.7$, and defined $\Lambda \equiv 10^{-22} \Lambda_{-22} \text{erg cm}^3 \text{s}^{-1}$ and $\lambda \equiv 0.035 \lambda_{0.035}$.

⁸ The relation $t_{\text{H}} = 5t_{\text{ff}}(R_{\text{vir}})$ can be derived from $t_{\text{ff}}(R_{\text{vir}}) = \sqrt{2} R_{\text{vir}}/v_{\text{vir}} \approx \sqrt{2}/(10H)$ and $t_{\text{H}} = 2/(3H)$, where H is the Hubble parameter.

The value of λ is normalized to the mean value found in the Bolshoi-Planck simulation in halos with mass $10^{10} < M_{\text{halo}} < 10^{15} M_{\odot}$ and redshift $0 < z < 8$ (Rodríguez-Puebla et al. 2016).

To estimate Λ we use the tables of Wiersma et al. (2009), which depend on T , Z , z , and n_{H} . The value of T is calculated from $\epsilon = v_c^2$ (eqn. 16) which gives

$$kT = \frac{2}{3} \mu m_p v_c^2 = \frac{4}{3} f_{v_c}^2 k T_{\text{vir}} . \quad (35)$$

For Z we use as a fiducial estimate the metallicity of gas in the central galaxy, since at R_{circ} which is roughly the size of the galaxy (e.g. Kravtsov 2013; Shibuya et al. 2015) significant mixing of the hot gas and ISM is likely. We calculate this metallicity estimate based on the observed $z = 0$ mass-metallicity relation from Andrews & Martini (2013):

$$\frac{Z_{\text{MZR}}(M_{\text{halo}}, z = 0)}{Z_{\odot}} = \frac{1.28}{1 + \left(\frac{M_{*}(M_{\text{halo}}, z = 0)}{10^{8.901} M_{\odot}} \right)^{-0.64}} , \quad (36)$$

where we converted [O/H] in Andrews & Martini (2013) to Z_{MZR}/Z_{\odot} assuming $12 + [\text{O}/\text{H}]_{\odot} = 8.69$ (Asplund et al. 2009), and we use the stellar-mass halo-mass relation (SMHM) from Behroozi et al. (2019, hereafter B19) to convert between M_{halo} and stellar-mass M_{*} . The dependence of Λ on n_{H} and z is due to heating and ionization by the UV background (UVB) and cooling off the cosmic microwave background, where only the UVB effect is significant in the halo masses of interest. To calculate n_{H} we solve eqn. (6) for $n_{\text{H}, \text{max}}$ including the dependence of Λ on n_{H} , and then rederive \dot{M}_{crit} accordingly. To gauge the importance of the UVB on \dot{M}_{crit} we also calculate Λ assuming no UVB (i.e., in the $n_{\text{H}} \rightarrow \infty$ limit), using the collisional-ionization equilibrium cooling tables from Gnat & Sternberg (2007).

In Figure 8 we explore the dependence of \dot{M}_{crit} on halo mass, gas metallicity, and redshift, assuming $f_{\lambda} \lambda = 0.035$ and an NFW mass profile to calculate f_{v_c} from eqn. (9). NFW concentration parameters are calculated using the fitting formulas of Klypin et al. (2016), which are based on the Bolshoi-Planck dark matter simulation.⁹ The value of f_{v_c} decreases with decreasing concentration, spanning at $z = 0$ between $f_{v_c}^3 \approx 2$ at $M_{\text{halo}} = 10^9 M_{\odot}$ and $f_{v_c}^3 \approx 0.8$ at $10^{13} M_{\odot}$, and spanning $f_{v_c}^3 \sim 0.4 - 0.6$ for all halo masses at redshift two or higher. The dependence of f_{v_c} on the gravity of the central galaxy is discussed later in this section. The four solid lines in the top panel plot \dot{M}_{crit} for $z = 0$ halos, with different assumptions on the calculation of Λ . The top three curves include heating by the UVB and differ in the assumed metallicity, with a metallicity either equal to Z_{MZR} given by eqn. (36), a metallicity equal to half this value, or no metal contribution to the cooling. An increase in \dot{M}_{crit} with decreasing Z is apparent at $M_{\text{halo}} > 10^{10.5} M_{\odot}$, at which $T_{\text{vir}} > 10^5 \text{K}$ and the metals can dominate the cooling. The bottom curve assumes $Z = Z_{\text{MZR}}$ and no UVB. This curve demonstrates that the UVB significantly affects \dot{M}_{crit} only at $M_{\text{halo}} < 10^{10} M_{\odot}$ at which the gas temperature is close to the equilibrium temperature. When the gas temperature equals the equilibrium temperature at $M_{\text{halo}} \approx 10^9 M_{\odot}$ then \dot{M}_{crit} goes to infinity since Λ goes to zero. Below this threshold

⁹ Klypin et al. (2016) published concentration parameters of halos up to $z = 5.4$ so we use these values also at higher redshift.

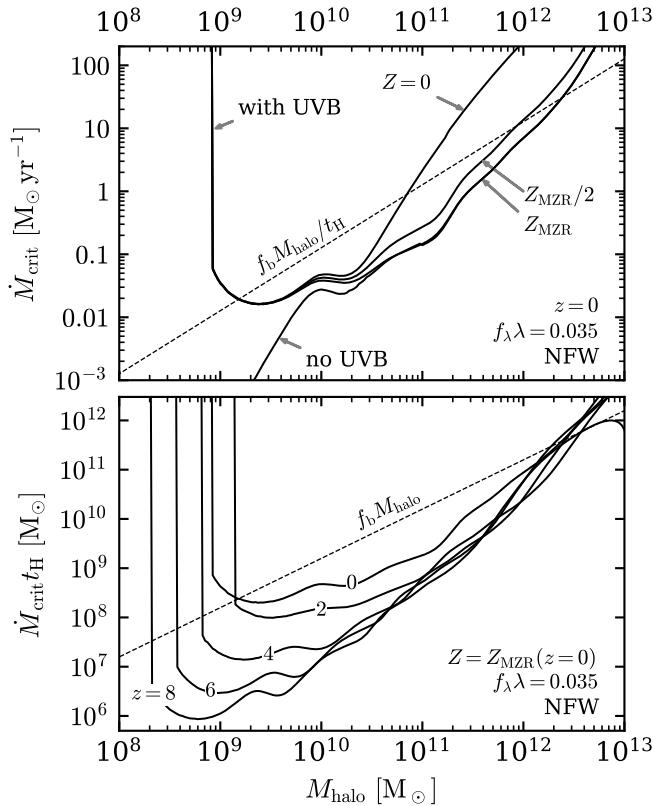


Figure 8. The critical accretion rate (eqn. 34) versus halo and gas parameters, assuming an NFW potential. (**Top**) Halos at $z = 0$. The curve marked ‘no UVB’ neglects heating by the UV background and assumes a hot gas metallicity (at $R_{\text{circ}} \approx 0.05R_{\text{vir}}$) of Z_{MZR} – the gas metallicity in the central galaxy (eqn. 36). The other curves include heating by the UVB and assume different hot gas metallicities as noted. The dashed line marks the cosmic halo baryon budget divided by t_{H} , an estimate of the accretion rate onto the halo if baryons follow the dark matter. (**Bottom**) The dependence of $\dot{M}_{\text{crit}} t_{\text{H}}$ (where t_{H} is the Hubble time) on redshift, assuming no redshift evolution in the MZR and including heating by the UVB. The dashed line marks the cosmic halo baryon budget. At $M_{\text{halo}} > 10^{10.5} M_{\odot}$ the product $\dot{M}_{\text{crit}} t_{\text{H}}$ is roughly independent of redshift.

there is no net cooling and the cooling flow solutions do not apply. Note also that Wiersma et al. (2009) did not account for local ionization sources in the galaxy (e.g. Cantalupo 2010), which may also decrease Λ and increase \dot{M}_{crit} .

In the bottom panel of Fig. 8 we vary the redshift while keeping λ and Z constant. The vertical axis in this panel is $\dot{M}_{\text{crit}} t_{\text{H}}$ where t_{H} is the Hubble time. This product gives a characteristic mass associated with accretion at a rate \dot{M}_{crit} . Note that with increasing redshift the virial temperature increases for a given halo mass (eqn. 35), causing the maximum in Λ and hence minimum in \dot{M}_{crit} to shift toward lower masses. At halo masses above the minima in \dot{M}_{crit} the value of $\dot{M}_{\text{crit}} t_{\text{H}}$ is almost independent of redshift, especially at $z \gtrsim 2$. This independence follows since at high M_{halo} metals dominate the cooling and hence $\dot{M}_{\text{crit}} \propto v_{\text{c}}^{5.4} R_{\text{circ}}$ (eqn. 12). Since $v_{\text{c}} \propto (1+z)^{1/3}$ and $R_{\text{circ}} \propto R_{\text{vir}} \propto (1+z)^{-2/3}$ we get that $\dot{M}_{\text{crit}} t_{\text{H}} \propto (1+z)^{0.13}$, i.e. $\dot{M}_{\text{crit}} t_{\text{H}}$ is roughly independent

of redshift if the metallicity is held constant. The offset of $\dot{M}_{\text{crit}} t_{\text{H}}$ at high M_{halo} and $z = 0$ relative to at $z \geq 2$ is mainly due to the higher concentration of $z = 0$ halos, and hence a higher f_{v_c} as mentioned above.

Also plotted in the top panel of Fig. 8 is $f_b M_{\text{halo}} / t_{\text{H}}$ (dashed line) which provides a rough approximation of the baryon accretion rate onto the halo if baryons follow the dark matter (the average mass accretion rate onto the halo is typically within a factor of two of $M_{\text{halo}} / t_{\text{H}}$, e.g. McBride et al. 2009; Rodríguez-Puebla et al. 2016). Thus, the halo mass where this expected baryon accretion rate intersects with \dot{M}_{crit} corresponds to the threshold for the onset of the hot mode accretion, if baryons follow the dark matter. Specifically, the intersection with the $\dot{M}_{\text{crit}}(Z_{\text{MZR}})$ curve at $M_{\text{halo}} = 2 \cdot 10^{12} M_{\odot}$ is similar to the threshold of $2.5 \cdot 10^{12} M_{\odot}$ derived by Dekel & Birnboim (2006) for the same parameters, i.e. $z = 0$, $Z \approx Z_{\odot}$, and a shock radius $\approx 0.1R_{\text{vir}}$ (see their figure 4).

In halo masses below the (Z -dependent) intersection between \dot{M}_{crit} and $f_b M_{\text{halo}} / t_{\text{H}}$, the value of \dot{M}_{crit} is lower than $f_b M_{\text{halo}} / t_{\text{H}}$, which gives the classic result that there is no hot mode accretion in low mass halos, if the baryons follow the dark matter. However, Fig. 8 shows that hot mode accretion can in principle be relevant also in these lower mass halos, if the gas mass and density are low enough such that $\dot{M} < \dot{M}_{\text{crit}}$.

The dashed line in the bottom panel of Fig. 8 denotes $f_b M_{\text{halo}}$, the cosmic baryon budget of the halo. Due to the weak dependence of $\dot{M}_{\text{crit}} t_{\text{H}}$ on redshift discussed above, $f_b M_{\text{halo}}$ and $\dot{M}_{\text{crit}} t_{\text{H}}$ intersect at a halo mass which is independent of redshift. This again corresponds to previous conclusions that M_{thres} is independent of redshift if the baryons follow the dark matter and the metallicity is held constant.

The galaxy can affect \dot{M}_{crit} via its modification of the gravitational potential and hence of $f_{\text{v}_c} = v_{\text{c}}(R_{\text{circ}})/v_{\text{vir}}$, since $\dot{M}_{\text{crit}} \propto f_{\text{v}_c}^3$ (eqn. 34). To estimate this effect as a function of halo mass, we assume the galaxy mass given by the SMHM from B19, as used above to estimate the gas metallicity. The galaxy is assumed to have an exponential disc distribution with half mass radius $R_{1/2}$ taken from Kravtsov (2013):

$$R_{1/2} = 0.015R_{200c} \approx 0.012R_{\text{vir}} . \quad (37)$$

where R_{200c} is the radius enclosing an overdensity of 200 relative to the critical density, and we remind the reader that R_{vir} is calculated using the overdensity from Bryan & Norman (1998). Since f_{v_c} depends on the mass enclosed within $R_{\text{circ}} \approx 0.05R_{\text{vir}}$, this size estimate is practically equivalent to assuming the galaxy is a point source, so any $R_{1/2}$ up to a factor of ≈ 2 above the estimate in eqn. (37) yields similar results. We then sum the galaxy mass profile with an NFW profile normalized by $1 - M_{*}/M_{\text{halo}}$ and solve eqn. (9) for R_{circ} and f_{v_c} . For completeness, we also calculate a mass profile which accounts for adiabatic contraction of the dark matter due to the galaxy using the CONTRA package (Gnedin et al. 2004). Observations and cosmological simulations though suggest that this contraction may be negated by dark matter expansion induced by clumpy gas accretion or feedback (Dutton et al. 2007; Macciò et al. 2012; Chan et al. 2015).

Figure 9 compares the calculations of \dot{M}_{crit} for the different assumed mass profiles. The vertical axis is normalized by $f_b M_{\text{halo}} / t_{\text{H}}$, so if the baryon accretion rate follows the dark matter accretion rate we expect hot mode accretion at halo

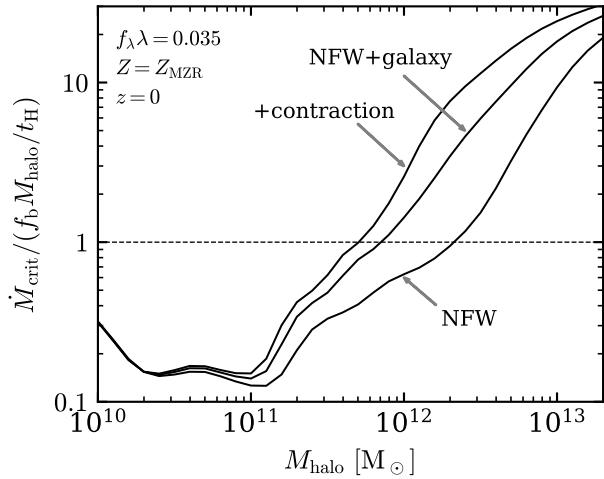


Figure 9. The dependence of \dot{M}_{crit} on the mass profile. The y-axis is normalized by the expected \dot{M} if baryons follow the dark matter. The lower curve assumes an NFW mass profile as in Fig. 8, the middle curve includes the galaxy contribution to the mass profile with M_* based on the SMHM from Behroozi et al. (2019), while the upper curve assumes also that the dark matter adiabatically contracts due to the galaxy. If $\dot{M} = f_b M_{\text{halo}} / t_H$ (i.e., for a gas accretion rate following the dark matter), the change in the mass profile due to the galaxy decreases the threshold halo mass for hot mode accretion from $2 \cdot 10^{12} M_\odot$ to $5 - 7 \cdot 10^{11} M_\odot$.

masses where the ordinate is above unity. The figure shows that the effect of the galaxy is largest at $M_{\text{halo}} = 2.5 \cdot 10^{12} M_\odot$ where the SMHM peaks, in which \dot{M}_{crit} increases by a factor of 4–8. The change in \dot{M}_{crit} is smaller at lower and higher M_{halo} , and almost vanishes at $M_{\text{halo}} < 10^{11} M_\odot$ due to the small galaxy mass. The threshold halo mass where \dot{M}_{crit} crosses $f_b M_{\text{halo}} / t_H$ (corresponding to the case where gas accretion into halos follows the dark matter) decreases by a factor of 3–4, from $M_{\text{thres}} = 2 \cdot 10^{12} M_\odot$ when the galaxy is not included in the mass profile calculation to $M_{\text{thres}} = 7 \cdot 10^{11} M_\odot$ when the galaxy is included and to $5 \cdot 10^{11} M_\odot$ when the dark matter is assumed to contract. Fig. 9 thus demonstrates that the onset of hot mode accretion depends not only on the properties of the halo but also on the properties of the galaxy.

4 COMPARISON OF THE CRITICAL COOLING RATE WITH THE STAR FORMATION RATE

In Figure 10 we compare \dot{M}_{crit} in $z = 0$ dark matter halos (thick black line) with the average star formation rate (SFR, gray stripe). The value of \dot{M}_{crit} is calculated from eqn. (34) using our fiducial parameters: $f_\lambda \lambda = 0.035$, f_{v_c} calculated from an NFW + galaxy profile with $M_*(M_{\text{halo}})$ from B19 (without adiabatic contraction), and the ISM metallicity corresponding to the same M_* (eqn. 36). Thin grey lines plot \dot{M}_{crit} assuming the metallicity is a factor of two lower (top curve) or higher (lower curve) than this fiducial estimate. Including also contraction of the halo in the calculation does not alter our main conclusions. The background

colors emphasize the two regimes for how the volume filling phase accretes onto the galaxy, gradual accretion of hot gas if $\dot{M} < \dot{M}_{\text{crit}}$ and free-fall if $\dot{M} > \dot{M}_{\text{crit}}$. The average SFR is also taken from B19, and is equal to the time derivative of the SMHM. We plot their mean SFRs for central galaxies (i.e. excluding satellites), and use the width of the grey stripe to denote the statistical uncertainty in the B19 model fits. The figure demonstrates that the average SFR derived by B19 is less than or comparable to \dot{M}_{crit} at any halo mass, and is within a factor of a few of \dot{M}_{crit} in the mass range $10^{10.5} \lesssim M_{\text{halo}} \lesssim 10^{11.5} M_\odot$ (which includes most of the stars formed at $z \sim 0$). As \dot{M}_{crit} is the maximum possible accretion rate of the hot mode, this result suggests that hot mode accretion can in principle dominate the gas supply for star formation in low mass halos.

To extend the comparison of \dot{M}_{crit} with the SFR to high redshift, we assume $f_\lambda \lambda = 0.035$ at all redshifts, motivated by the constant median $\lambda = 0.035$ found in the Bolshoi-Planck simulation (Rodríguez-Puebla et al. 2016). The metallicity at high redshift is a major uncertainty. As a fiducial model for the metallicity evolution we utilize the scaling suggested by Dekel & Birnboim (2006) based on semi-analytic models:

$$Z(M_*, z) = 10^{-sz} Z(M_*, z=0) \quad (38)$$

with an enrichment rate $s = 0.17$. This enrichment rate is consistent with the factor of two lower normalization of the mass-metallicity relation at $z \approx 2$ relative to its local value (Erb et al. 2006; Sanders et al. 2015), and is similar to $s = 0.22 \pm 0.03$ deduced for damped Ly α absorbers (DLAs) at $0 < z < 5$ by Rafelski et al. (2012)¹⁰. The value of $Z(M_*, z=0)$ in eqn. (38) is calculated as above using eqn. (36) for the mass-metallicity relation in the local universe, and using the B19 SMHM to derive M_* from M_{halo} and z . The same M_* is also used for the calculation of f_{v_c} , and we assume all the galaxy mass is within $R_{\text{circ}} = 0.05 R_{\text{vir}}$. This latter assumption is consistent with the $R_{1/2} \approx 0.02 R_{\text{vir}}$ found at $z = 0 - 8$ by Shibuya et al. (2015).

Solid lines in Figure 11 plot the implied \dot{M}_{crit} using these parameters. Each panel corresponds to a different redshift as noted at the top of the panels. We normalize the vertical axes in this plot by $f_b M_{\text{halo}} / t_H$ in order to decrease the dynamical range, so in this plot the baryon accretion rate if baryons follow the dark matter is roughly at unity, and the SFR stripes roughly track the ratio of stellar mass to halo baryon budget (since $\text{SFR} \cdot t_H / f_b M_{\text{halo}} \sim M_* / f_b M_{\text{halo}}$). To bracket the range of \dot{M}_{crit} implied by the uncertainty in metallicity we also plot \dot{M}_{crit} assuming the local mass-metallicity relation holds at higher redshift (i.e., $s = 0$), and \dot{M}_{crit} assuming no contribution of metals to the cooling (marked as $Z = 0 Z_\odot$, though in practice any $Z \lesssim 0.01 Z_\odot$ gives identical results). Figure 11 demonstrates that for the metallicity evolution rate in eqn. (38), $\text{SFR} \lesssim \dot{M}_{\text{crit}}$ at most halo masses and redshifts. The hot mode accretion can thus in principle dominate the gas supply for star formation also in low mass halos at high redshift.

¹⁰ The trend of DLA metallicity versus redshift found by Rafelski et al. (2012) does not account for the possible trend of M_* with z in their sample, so their quoted value potentially overestimates s as defined in eqn. (38).

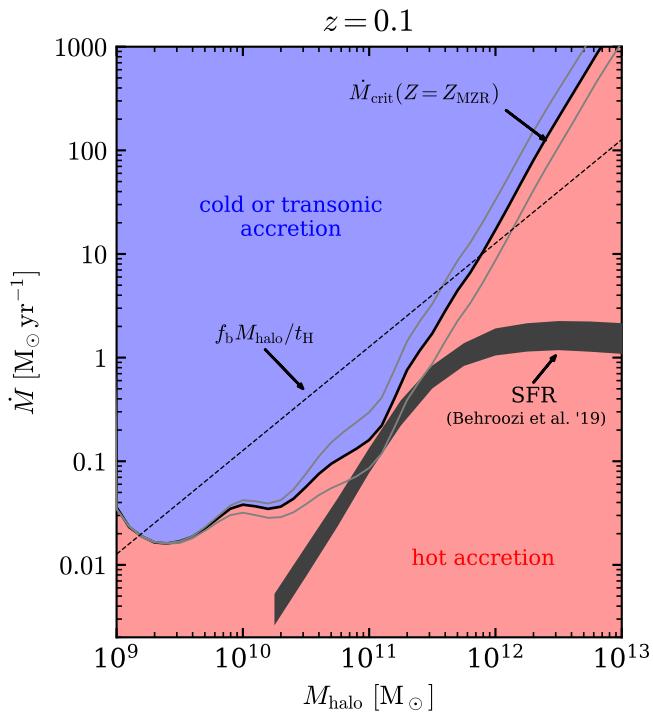


Figure 10. Comparison of the critical cooling rate with the mean star formation rate at $z = 0.1$. The thick solid line plots \dot{M}_{crit} (eqn. 34), assuming $Z = Z_{\text{MZR}}$, $f_A \lambda = 0.035$, and contracted NFW + galaxy mass profiles as in Fig. 9. Thin grey lines plot \dot{M}_{crit} for a factor of two lower or higher Z . For $\dot{M} < \dot{M}_{\text{crit}}$ we expect the volume-filling phase to be pressure-supported down to the galaxy scale (hot accretion mode, red background), while for $\dot{M} > \dot{M}_{\text{crit}}$ we expect the gas to reach the galaxy with supersonic velocities (cold or transonic accretion modes, blue background). The dashed line marks $f_b M_{\text{halo}} / t_H$, an estimate of the accretion rate onto the halo if baryons follow the dark matter. The mean SFR for central galaxies derived by B19 is indicated by a gray stripe, where the stripe width marks the uncertainty in their model fit. The mean SFR is comparable to or lower than \dot{M}_{crit} at all halo masses.

5 DISCUSSION AND CONCLUSIONS

The physical properties of the volume-filling gas phase in dark matter halos are crucial both for understanding the nature of galaxy accretion and for understanding the consequences of feedback (e.g. White & Rees 1978; White & Frenk 1991, BD03, Dekel & Birnboim 2006; Fielding et al. 2017). In this paper we revisit the question of whether this gas phase is predominantly hot and pressure-supported or predominantly cool and free-falling. We limit the effect of feedback in our analysis to the possible enrichment and depletion of the halo gas. Absent dynamical effects of feedback (e.g. heating), hot pressure-supported gas in halos forms a cooling flow. We demonstrate that the family of cooling flow solutions separates the physical states of the halo gas into three regimes, according to whether the cooling flow sonic radius R_{sonic} is on the scale of the galaxy, on the scale of the halo, or beyond the halo (Fig. 3). The first regime corresponds to the classic hot accretion mode where the flow is subsonic (i.e. pressure supported) and smooth from the accretion shock down to the galaxy scale. The third regime cor-

responds to the classic cold accretion regime where clumpy gas falls in supersonically from the IGM down to the galaxy without experiencing a strong shock. In the second intermediate regime the gas forms a hot inflow over some range of radii, and then cools out at R_{sonic} and free-falls onto the galaxy. This ‘transonic’ regime resembles the classic cold mode in terms of the properties of gas when it accretes onto the galaxy, since the gas reaches the galaxy scale as a cold and free-falling flow (Fig. 7). However, in terms of coupling with feedback this intermediate scenario may in some aspects more closely resemble the hot mode, due to the existence of a layer of hot and homogeneous pressure-supported gas situated beyond the cold and clumpy free-falling flow. This will presumably depend on where in the halo relative to R_{sonic} the feedback energy is deposited.

In the simulation shown in Figs. 4–5 the intermediate transonic scenario for the halo gas develops from hydrostatic initial conditions, when the mass inflow rate crosses the critical value of \dot{M}_{crit} . It is less clear if this scenario can be realized if the gas inflow is supersonic at large radii as is often the case in the cosmological context. Indeed, a sonic transition at intermediate radii in the halo is not seen in the idealized simulations of Birnboim & Dekel (2003) where the halo gas is initially inflowing supersonically. We argued in section 2.4 that this difference could be due to the lack of a source of strong shocks in the outer halo in the Birnboim & Dekel simulations. It would thus be interesting to check whether this intermediate regime materializes in setups which include outflows that shock against supersonic inflows at large scales (e.g. Fielding et al. 2017) and in the more realistic conditions in cosmological simulations. The latter can include strong feedback at high redshift that ‘pre-heats’ the gas, stifling later supersonic inflows and more closely resembling the hydrostatic initial conditions used in this work (e.g. Figs. 4–5). We leave addressing this question and deriving the implications of this possible new accretion regime of halo gas to future work.

We demonstrate that hot mode accretion is possible only if $R_{\text{sonic}} < R_{\text{circ}} \approx 0.05 R_{\text{vir}}$, because this condition determines when the gas is virialized and roughly hydrostatic down to galaxy scales. This condition on R_{sonic} is equivalent to the condition $t_{\text{cool}} \gtrsim t_{\text{ff}}$ at $r = R_{\text{circ}}$ (eqn. 21), and can be cast as a maximum accretion rate in the hot mode \dot{M}_{crit} (eqns. 8, 12). We emphasize that ‘hot’ corresponds to the virial temperature, which is relatively low for low-mass halos. We explore the dependence of \dot{M}_{crit} on halo mass, redshift, and gas metallicity in Fig. 8. We find that in halos where metals dominate the cooling the product $\dot{M}_{\text{crit}} t_H$ is roughly independent of redshift if the metallicity is held constant (Fig. 8).

The classic threshold halo mass M_{thres} for the onset of hot mode accretion can be derived by equating \dot{M}_{crit} with the spherically-averaged baryon accretion rate expected from dark matter simulations $\sim f_b M_{\text{halo}} / t_H$ (see Fig. 8). Assuming an NFW mass profile and a gas metallicity at R_{circ} following the $z = 0$ stellar mass-metallicity relation, we find $M_{\text{thres}} \approx 2 \cdot 10^{12} M_\odot$ roughly independent of redshift. This result is consistent with the calculations of BD03 for the formation of a stable accretion shock near R_{circ} for the same parameters (section 3 and Fig. 8). As our derivation does not assume the gas was heated to $\sim T_{\text{vir}}$ in a single shock, our results suggest that the condition for hot mode accretion

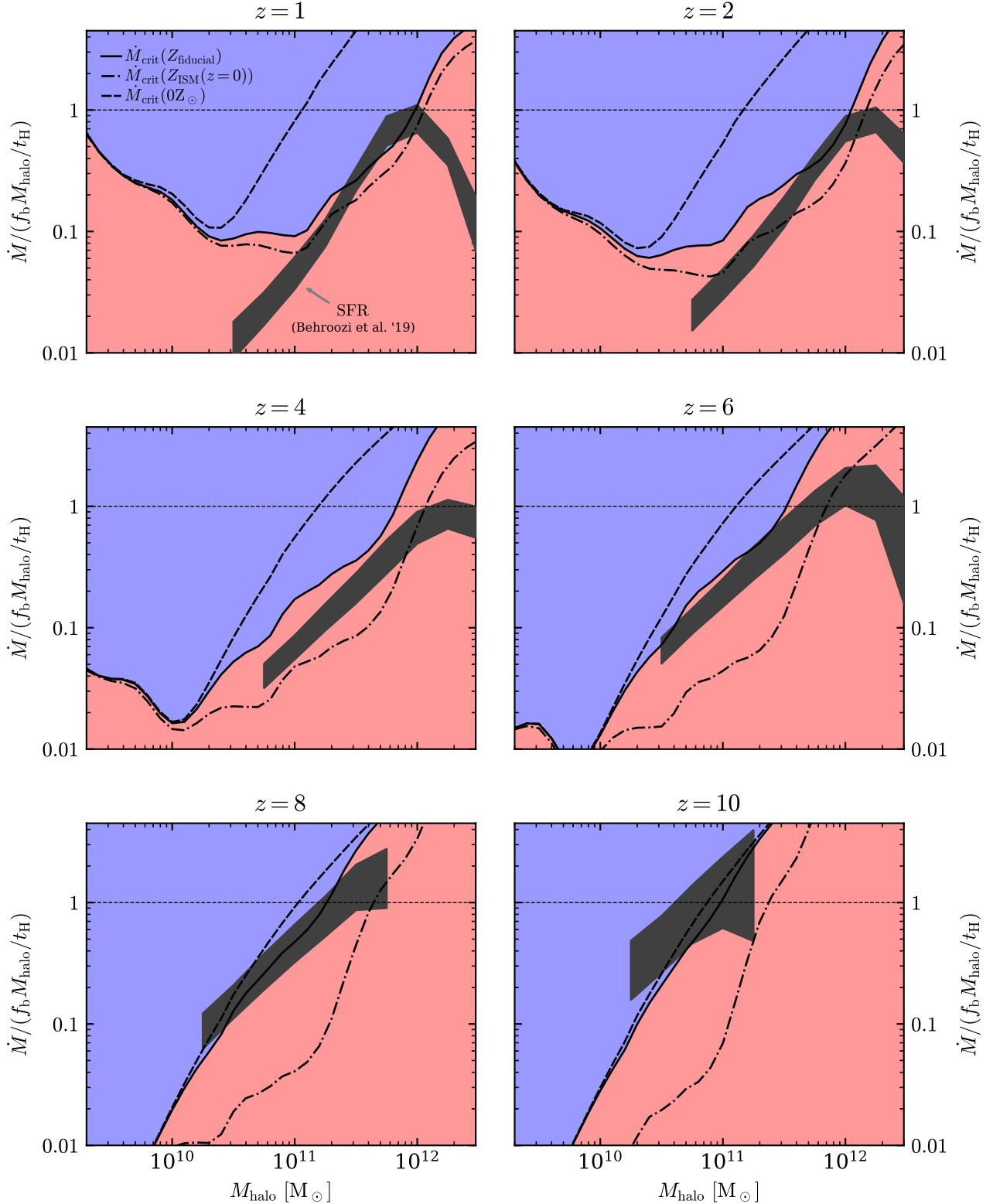


Figure 11. Comparison of the critical cooling rate with the mean SFR at different redshifts. The vertical axis is normalized by $f_b M_{\text{halo}}/t_{\text{H}}$, the estimated accretion rate if baryons follow the dark matter. The solid black line plots \dot{M}_{crit} assuming the redshift evolution of the metallicity from Dekel & Birnboim (2006) given in eqn. (38). Blue and red backgrounds mark respectively hot and cold / transonic accretion modes for the volume-filling phase, for this calculation of \dot{M}_{crit} . The two other black lines bracket the uncertainty in \dot{M}_{crit} due to the uncertain metallicity. The thick dashed lines plot \dot{M}_{crit} assuming no metal contribution to the cooling, while the dash-dotted lines assume no redshift evolution in the metallicity. Gray bands mark the mean SFR of central galaxies from B19. For the fiducial metallicity model $\text{SFR} \lesssim \dot{M}_{\text{crit}}$ at most halo masses and redshifts.

derived by BD03 apply more generally. We also show that when accounting for the gravitational effects of the galaxy M_{thres} decreases by a factor of $3 - 4$ at $z = 0$ (Fig. 9). Moreover, we showed that if the baryon accretion rate or halo gas mass is lower than expected based on dark matter simulations and spherical averaging, such that the cooling and accretion rates are smaller than \dot{M}_{crit} , then hot mode accretion would be relevant also in halos with $M_{\text{halo}} < M_{\text{thres}}$.

The conclusion that hot mode accretion is determined by conditions at the galaxy scale implies that the relevant metallicity for calculating \dot{M}_{crit} is the metallicity of the hot gas just outside the galaxy, which is potentially higher than at larger scales due to more intense enrichment by outflows. Also, while our calculations neglect possible deviations from spherical symmetry induced by cosmological filaments, we expect these to affect our results regarding the nature of accretion from the volume-filling phase only if filaments retain their identity down to the galaxy scale. Cosmological simulations currently differ on whether this is indeed the case, or whether instead the filaments dissolve farther out in the halo (e.g., Kereš et al. 2005; Ceverino et al. 2010; Faucher-Giguère et al. 2011; Nelson et al. 2013; Danovich et al. 2015, see also Mandelker et al. 2016, 2019; Padnos et al. 2018).

Our analysis assumes steady-state conditions, while various physical processes associated with galaxy formation, such as the growth of the background potential, bursty stellar feedback (e.g. Muratov et al. 2015), and clumpy accretion, could drive the system away from steady state. An interesting question is thus what are the relevant timescales on which steady state can be achieved? Our results suggest that the relevant timescale for determining the nature of accretion is the dynamical time of the galaxy. The importance of this timescale emerges from the critical solution, in which $t_{\text{cool}} \approx t_{\text{ff}}$ at $r \approx R_{\text{circ}}$. Since in any cooling flow solution $t_{\text{cool}} \approx t_{\text{flow}} \equiv r/|v_r|$ (eqn. 18), the critical solution satisfies

$$t_{\text{cool}}(R_{\text{circ}}) \approx t_{\text{flow}}(R_{\text{circ}}) \approx t_{\text{ff}}(R_{\text{circ}}) = \frac{2f_{\lambda}\lambda R_{\text{vir}}}{v_c}, \quad (39)$$

where the last equality follows from eqns. (4) and (9). Note that this relation differs from the predictions of several feedback-regulation models in which t_{cool} in the halo is regulated to some factor of t_{ff} (e.g., Sharma et al. 2012; Voit et al. 2017), since in these models $t_{\text{flow}} \gg t_{\text{cool}}$ due to heating by feedback, in contrast with $t_{\text{flow}} \approx t_{\text{cool}}$ in the cooling flow solution. Equation (39) implies that the relevant timescales for the onset of hot mode accretion are a factor of $(\sqrt{2}f_{\lambda}\lambda)^{-1} \sim 20$ shorter than the halo dynamical time, or a factor of ~ 100 shorter than the Hubble time at the corresponding redshift. We thus expect our results to be roughly valid as long as other processes change the relevant physical conditions on timescales longer than this characteristic value. Moreover, the fact that this timescale is relatively short implies that the nature of accretion can be determined by transient processes, if the transient conditions last longer than the galaxy dynamical time. For example, if a burst of feedback depletes gas in the galaxy vicinity such that \dot{M} drops below \dot{M}_{crit} , then the remaining gas may accrete in the hot mode even if the accretion rate averaged over longer timescales is larger than \dot{M}_{crit} .

Figures 10 and 11 plot the average SFR in dark matter halos empirically derived by Behroozi et al. (2019) based on predictions from dark matter-only simulations and ob-

servational constraints. These figures show that the average SFR is consistent with or lower than \dot{M}_{crit} at almost all halo masses and redshifts, for the fiducial metallicity evolution discussed in section 4. It is unclear if this result is a coincidence, or indicates a physical connection between \dot{M}_{crit} and the SFR in low mass halos. However, we have shown that hot mode accretion and \dot{M}_{crit} may be relevant also to low mass halos if they are sufficiently depleted of baryons. It would thus be valuable to explore scenarios in which \dot{M}_{crit} provides a physical upper limit to the SFR in all halos at all times. How could this be the case? Due to the different nature of accretion and different consequences of feedback according to the state of the halo gas, it is plausible that the star formation efficiency SFR/\dot{M} during the hot accretion phase differs significantly from the SF efficiency during the phase where gas reaches the galaxy in free-fall. In a low-mass halo where gas accretes onto the galaxy via the hot mode for some fraction of the time, and the SF efficiency during this hot mode phase is high while it is low in the free-fall phase due to strong winds, the SFR would tend to be $\lesssim \dot{M}_{\text{crit}}$, since during the hot phase $\dot{M} \lesssim \dot{M}_{\text{crit}}$. Simulations of low mass halos which include star formation and feedback could test if such a scenario is realized.

Last, since \dot{M}_{crit} is determined by physical properties at the galaxy scale it can be estimated from observations of galaxy properties, and then compared to the SFR (or other properties) of individual galaxies. This is in contrast with the statistical modelling required to derive average SFR and \dot{M}_{crit} in dark matter halos using techniques such as abundance matching (as in section 4 above). It would be interesting to derive the relation between SFR and \dot{M}_{crit} on a galaxy-by-galaxy basis and for different galaxy subtypes. This may provide new insights into the importance of the hot accretion mode for fuelling and/or quenching star formation, as well as the origin of galaxy scaling relations involving parameters determining \dot{M}_{crit} , such as the Tully & Fisher (1977) relation between v_c and stellar mass.

ACKNOWLEDGEMENTS

JS is supported by the CIERA Postdoctoral Fellowship Program. DF is supported by the Flatiron Institute, which is supported by the Simons Foundation. CAFG is supported by NSF through grants AST-1517491, AST-1715216, and CAREER award AST-1652522, by NASA through grants NNX15AB22G and 17-ATP17-0067, by STScI through grants HST-GO-14681.011, HST-GO-14268.022-A, and HST-AR-14293.001-A, and by a Cottrell Scholar Award from the Research Corporation for Science Advancement. EQ was supported in part by a Simons Investigator Award from the Simons Foundation and by NSF grant AST-1715070.

REFERENCES

- Andrews, B. H., & Martini, P. 2013, ApJ, 765, 140
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
- Balbus, S. A., & Soker, N. 1989, ApJ, 341, 611
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, ApJ, 770, 57

- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143 (B19)
- Bertschinger, E. 1989, ApJ, 340, 666
- Birnboim, Y., Dekel, A., & Neistein, E. 2007, MNRAS, 380, 339
- Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349 (BD03)
- Brooks, A. M., Governato, F., Quinn, T., Brook, C. B., & Wadsley, J. 2009, ApJ, 694, 396
- Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80
- Bullock, J. S., Dekel, A., Kolatt, T. S., et al. 2001, ApJ, 555, 240
- Cantalupo, S. 2010, MNRAS, 403, L16
- Ceverino, D., Dekel, A., & Bournaud, F. 2010, MNRAS, 404, 2151
- Chan, T. K., Kereš, D., Oñorbe, J., et al. 2015, MNRAS, 454, 2981
- Chisholm, J., Tremonti, C. A., Leitherer, C., & Chen, Y. 2017, MNRAS, 469, 4831
- Cowie, L. L., Fabian, A. C., & Nulsen, P. E. J. 1980, MNRAS, 191, 399
- Danovich, M., Dekel, A., Hahn, O., Ceverino, D., & Primack, J. 2015, MNRAS, 449, 2087
- Davies, J. J., Crain, R. A., McCarthy, I. G., et al. 2019, MNRAS, 485, 3783
- Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
- Dutton, A. A., van den Bosch, F. C., Dekel, A., & Courteau, S. 2007, ApJ, 654, 27
- Erb, D. K., Shapley, A. E., Pettini, M., et al. 2006, ApJ, 644, 813
- Fabian, A. C., Nulsen, P. E. J., & Canizares, C. R. 1984, Nature, 310, 733
- Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982
- Fielding, D., Quataert, E., McCourt, M., & Thompson, T. A. 2017, MNRAS, 466, 3810
- Gnat, O., & Sternberg, A. 2007, ApJS, 168, 213
- Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, ApJ, 616, 16
- Haardt, F., & Madau, P. 2012, ApJ, 746, 125
- Hafen, Z., Faucher-Giguère, C.-A., Anglés-Alcázar, D., et al. 2019, MNRAS, 488, 1248
- Heckman, T. M., & Thompson, T. A. 2017, arXiv:1701.09062
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800
- Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
- Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160
- Kereš, D., Vogelsberger, M., Sijacki, D., Springel, V., & Hernquist, L. 2012, MNRAS, 425, 2027
- Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, MNRAS, 457, 4340
- Kravtsov, A. V. 2013, ApJ, 764, L31
- Macciò, A. V., Stinson, G., Brook, C. B., et al. 2012, ApJ, 744, L9
- Mandelker, N., Padnos, D., Dekel, A., et al. 2016, MNRAS, 463, 3921
- Mandelker, N., Nagai, D., Aung, H., et al. 2019, MNRAS, 484, 1100
- Mathews, W. G., & Bregman, J. N. 1978, ApJ, 224, 308
- McBride, J., Fakhouri, O., & Ma, C.-P. 2009, MNRAS, 398, 1858
- Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903
- Moster, B. P., Naab, T., & White, S. D. M. 2018, MNRAS, 477, 1822
- Muratov, A. L., Kereš, D., Faucher-Giguère, C.-A., et al. 2015, MNRAS, 454, 2691
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, ApJ, 490, 493
- Nelson, D., Vogelsberger, M., Genel, S., et al. 2013, MNRAS, 429, 3353
- Nelson, D., Kauffmann, G., Pillepich, A., et al. 2018, MNRAS, 477, 450
- Ocvirk, P., Pichon, C., & Teyssier, R. 2008, MNRAS, 390, 1326
- Oppenheimer, B. D., Davé, R., Kereš, D., et al. 2010, MNRAS, 406, 2325
- Oppenheimer, B. D., Davies, J. J., Crain, R. A., et al. 2019, arXiv:1904.05904
- Padnos, D., Mandelker, N., Birnboim, Y., et al. 2018, MNRAS, 477, 3293
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
- Quataert, E., & Narayan, R. 2000, ApJ, 528, 236
- Rafelski, M., Wolfe, A. M., Prochaska, J. X., Neeleman, M., & Mendez, A. J. 2012, ApJ, 755, 89
- Rees, M. J., & Ostriker, J. P. 1977, MNRAS, 179, 541
- Rodríguez-Puebla, A., Behroozi, P., Primack, J., et al. 2016, MNRAS, 462, 893
- Sanders, R. L., Shapley, A. E., Kriek, M., et al. 2015, ApJ, 799, 138
- Sharma, P., McCourt, M., Parrish, I. J., & Quataert, E. 2012b, MNRAS, 427, 1219
- Shibuya, T., Ouchi, M., & Harikane, Y. 2015, ApJS, 219, 15
- Silk, J. 1977, ApJ, 211, 638
- Somerville, R. S., & Primack, J. R. 1999, MNRAS, 310, 1087
- Stern, J., Fielding, D., Faucher-Giguère, C.-A., & Quataert, E. 2019, MNRAS, 488, 2549 (Paper I)
- Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661
- van de Voort, F., Schaye, J., Booth, C. M., Haas, M. R., & Dalla Vecchia, C. 2011, MNRAS, 414, 2458
- White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
- White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
- Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009, MNRAS, 393, 99
- van de Voort, F., Schaye, J., Booth, C. M., Haas, M. R., & Dalla Vecchia, C. 2011, MNRAS, 414, 2458
- Voit, G. M., Meece, G., Li, Y., et al. 2017, ApJ, 845, 80

APPENDIX A: THE BERNOULLI PARAMETER IN THE PRESENCE OF RADIATIVE LOSSES

When accounting for radiative losses, energy conservation can be stated as

$$\frac{d\epsilon}{dt} = -\frac{P d\rho^{-1}}{dt} - q = \frac{P}{\rho} \frac{d \ln \rho}{dt} - q , \quad (\text{A1})$$

where ϵ and q are the specific thermal energy and specific luminosity, and the other variables have their usual meaning. For a spherical steady-state flow $d/dt = v_r d/dr$ so we get

$$v_r \left(\frac{d\epsilon}{dr} - \frac{P}{\rho} \frac{d \ln \rho}{dr} \right) = -q . \quad (\text{A2})$$

Using $d \ln \rho = d \ln P - d \ln \epsilon$ and $P/\rho = (\gamma - 1)\epsilon$ then gives

$$v_r \left(\gamma \frac{d\epsilon}{dr} - \frac{1}{\rho} \frac{dP}{dr} \right) = -q . \quad (\text{A3})$$

Using the momentum equation (14) we then arrive at eqn. (1):

$$v_r \frac{d}{dr} \left(\frac{1}{2} v_r^2 + \gamma \epsilon + \Phi \right) = -q , \quad (\text{A4})$$