

The Cosmic Baryon Cycle in IllustrisTNG: flows of mass, energy, and metals

Yossi Oren,^{1*} Viraj Pandya,² Rachel S. Somerville,³ Shy Genel,^{2,3} Osase Omoruyi,⁴ and Amiel Sternberg^{1,3,5}

¹*School of Physics and Astronomy, Tel Aviv University, Ramat Aviv 69978, Israel*

²*Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA*

³*Center for Computational Astrophysics, Flatiron Institute, 162 5th Ave., New York, NY, 10010, USA*

⁴*Center for Astrophysics, Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA*

⁵*Max-Planck-Institut für extraterrestrische Physik (MPE), Giessenbachstr., 85748 Garching, Germany*

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We measure and analyze the inflows and outflows of mass, energy, and metals through the interstellar medium (ISM) and circumgalactic medium (CGM) of galaxies in the IllustrisTNG100 simulations. We identify the dominant feedback mechanism in bins of halo virial mass and redshift by computing the integrated energy input from SNe and the “kinetic” and “thermal” mode of AGN feedback. We measure all quantities in a shell at the virial radius (“halo scale”) and one chosen to be approximately at the interface of the CGM and the interstellar medium (ISM; “ISM scale”). We find that galaxies have strong net positive inflows on halo scales, and weaker but still net positive inflows on ISM scales, at $z \gtrsim 2$. At later times, partially due to the onset of kinetic AGN feedback in massive halos, inflows and outflows nearly balance one another, leading to the familiar effects of the slow-down of galaxy growth and the onset of quenching. Halos dominated by SN feedback show only weak evidence of preventative feedback on halo scales, and we see excess ISM scale accretion indicative of rapid gas recycling. Wind mass loadings decrease with increasing halo mass, and with increasing redshift, while energy loadings are nearly independent of both mass and redshift. The detailed catalogs of these mass, metal, and energy inflow and outflow rates on galaxy and halo scales can be used to guide empirical and semi-analytic models, and provide deeper insight into how galaxy growth and quenching is regulated in the IllustrisTNG simulations.

Key words: keyword1 – keyword2 – keyword3

1 INTRODUCTION

Our current understanding of galaxy evolution is based on the Λ Cold Dark Matter model (Λ CDM) for hierarchical structure formation (White & Rees 1978; Blumenthal et al. 1984; Davis et al. 1985). It posits that gravitational instability drives the collapse of dark matter dominated over-densities, along with the associated baryons, into bound structures. These dark matter halos formed the sites where gas was able to cool and form the first stars. Halos grow continuously over time via accretion of gas and dark matter as well as via mergers, as larger and larger scales become gravitationally unstable and collapse.

From the first theoretical attempts to account for the properties of galaxy populations within the Λ CDM framework, it has been clear that some kind of “feedback” is required across the spectrum of halo masses (White & Rees 1978; White & Frenk 1991; Dekel & Silk 1986). Not all of the accreted gas that is able to cool rapidly can be allowed to form stars efficiently, or both low mass and high mass galaxies would be overproduced relative to observations. Another way to cast this is in terms of the “baryon conversion efficiency”, or the fraction of a halo’s “baryon budget” that is in the form of stars ($m_*/(f_B M_{\text{vir}})$, where m_* is the mass in stars, f_B is the universal

baryon fraction and M_{vir} is the mass of the halo). Observational constraints from abundance matching, lensing, and galaxy clustering show that this baryon conversion efficiency is quite low in the low to intermediate redshift universe ($z \sim 0$ –10), reaching a maximum value of ~ 20 percent at its peak and dropping to only a few percent in the lowest and highest halo masses that host observable galaxies (Moster et al. 2010; Behroozi et al. 2013; Behroozi et al. 2019; Wechsler & Tinker 2018). This implies that a large fraction of the gas that would be expected to cool and accrete into the interstellar medium (ISM) and form stars in the absence of feedback does not do so, either because it is prevented from accreting and/or cooling as rapidly as expected (preventative feedback) or because mass is efficiently ejected from the ISM and/or circumgalactic medium (CGM; ejective feedback).

There is also direct evidence that stars and supermassive black holes deposit copious amounts of energy and momentum into their surroundings (Heckman & Best 2023, and references therein). Decades of observations of ‘galactic winds’ show large scale outflows (e.g. Osterbrock 1960; Burbidge et al. 1964; Veilleux et al. 2005; Rupke 2018). Galactic winds can extend well beyond the ISM and into the CGM (e.g. Yoshida et al. 2016; Falgarone et al. 2017; Guo et al. 2023), and in some cases may even be powerful enough to overcome the binding energy of the halo and reach the intergalactic medium (IGM; Rupke et al. 2019).

* E-mail: orenyossi01@gmail.com

Two primary mechanisms drive these large-scale outflows: supernova explosions (SNe) and active galactic nuclei (AGN). Massive stars and supernovae deposit energy and momentum into the ISM, driving an over-pressured bubble of hot gas that can entrain cold material and vent mass, energy, and metals out of the galaxy (e.g. [Strickland & Heckman 2009](#); [Rupke 2018](#)). AGN can drive outflows via two mechanisms. Rapid accretion onto the BH is thought to lead to efficient production of radiation (associated with classical quasars and AGN), which can drive winds via radiation pressure, magnetic pressure, and line pressure (sometimes called “quasar mode” or “bright mode”). BH accreting at lower rates can drive powerful relativistic jets, associated with radio galaxies (sometimes called “radio mode” or “jet mode”; [Fabian 2012](#); [Heckman & Best 2014](#)). Observationally, it is often difficult to disentangle whether a wind is driven by SNe or an AGN, and many winds may have contributions from both types of sources ([Veilleux et al. 2005](#)). However, it is likely that winds in lower mass galaxies, which tend to be star forming and may have lower mass black holes, are primarily driven by SNe and observed outflows in more massive galaxies (particularly ‘quenched’ galaxies with little ongoing star formation) are driven by AGN. These wind processes occur across a wide range of spatial and temporal time scales. The energy and momentum deposition from SNe and black holes occurs on physical scales of less than a parsec, while galactic winds and BH jets can extend out to hundreds of kpc. Stars that evolve over millions of years release supernova energy in seconds, and black hole accretion rates can vary by orders of magnitude on timescales of months to years.

The processes of intergalactic gas inflow, ejection via feedback, and reaccretion onto the galaxy are collectively termed “the baryon cycle”, and it is viewed as crucial to our understanding of galaxy evolution (Astro2020 Decadal Report¹). The vast range of spatial and temporal scales, and the importance of a wide range of complex, non-linearly interconnected physical processes including gravity, hydrodynamics, thermodynamics, magnetic fields, and radiation, make modeling the baryon cycle in a cosmological context extremely challenging ([Somerville & Davé 2015](#)). Numerical cosmological simulations approach this task by explicitly solving the relevant equations numerically on a (sometimes adaptive or moving) mesh or using Lagrangian, particle-based methods such as Smoothed Particle Hydrodynamics (SPH; [Vogelsberger et al. 2020](#)). Many of the key processes that shape galaxy evolution, including star formation, stellar/SNe feedback, and black hole seeding, accretion, and feedback occur on spatial scales below the resolution that is achievable in large volume cosmological simulations (typically hundreds of parsec to ~ 1 kpc). As a result, these processes are modeled using “sub-grid” recipes, which historically have been implemented in a somewhat heuristic, phenomenological manner (see the discussion in [Somerville & Davé 2015](#); [Naab & Ostriker 2017](#)). These recipes contain free parameters that reflect our ignorance about the details of how the smaller scale phenomenon manifest on the coarse grained scale, which are commonly tuned to reproduce certain quasi-observable global galaxy properties at the population level, such as the galaxy stellar mass function. Different sub-grid implementations can give rise to very different predictions for quantities that have not been explicitly calibrated, particularly properties of difficult-to-observe diffuse gas in the CGM and IGM ([Tillman et al. 2023](#); [Wright et al. 2024](#)). Furthermore, different sub-grid implementations can arrive at similar $z \sim 0$ galaxy stellar properties via very different routes, exhibiting rather

dramatically different baryon cycles ([Pandya et al. 2020](#); [Wright et al. 2024](#)).

In spite of its fundamental importance, many basic questions about feedback and the baryon cycle in these large cosmological simulations remain open. These include: Is feedback primarily ejective or preventative? How far does ejected material travel, and how many times does it get recycled back into the ISM? What is the specific energy (or ratio of energy loading to mass loading) in winds? Which feedback processes (SNe driven winds, radiatively efficient or inefficient AGN) are important at different times, scales, and halo masses? How do metal flows trace baryon flows? What is the relationship between wind launching parameters and emergent properties of outflows on larger scales?

Although numerical cosmological simulations are a powerful tool, they are computationally expensive, making it difficult to simulate large volumes at adequate resolution, and to explore the effects of varying the sub-grid parameters or implementations. A complementary approach to modeling the baryon cycle is a technique called semi-analytic modeling (SAM), in which galaxies are represented by a collection of reservoirs such as the stellar bulge and disk, ISM, CGM, etc, set within “merger trees” that represent the cosmological growth of halos through accretion and merging (e.g. [Kauffmann et al. 1993](#); [Somerville & Primack 1999](#); [Cole et al. 2000](#); [Benson 2010](#)). Simplified models of physical processes such as gas accretion, heating and cooling, star formation, chemical enrichment, and stellar and black hole feedback are used to track the flows of material between these different reservoirs by solving a system of non-linearly coupled ordinary differential equations (ODE). An important recent development is the addition of ODEs describing the *energy* flows between reservoirs, in addition to mass and metals as in traditional SAMs. [Carr et al. \(2023\)](#) and [Pandya et al. \(2023\)](#) showed that energy flows can play a critical role in regulating cooling and accretion from the CGM to the ISM, thereby effectively regulating star formation without requiring large amounts of mass to be ejected from the ISM (see also [Voit et al. 2024a,b](#)).

Traditionally, the scaling relations adopted in SAMs to describe processes such as star formation or stellar driven winds were largely empirical or phenomenological, and were parameterized and tuned to match observational constraints. However, more recently, there have been attempts to implement scalings extracted from hydrodynamic simulations within SAMs ([Pandya et al. 2023](#)) or to use them to motivate the scalings used in closely related “gas regulator” models ([Mitchell & Schaye 2022](#)). It is an interesting open question whether in this way the SAM framework can effectively be used as a physically interpretable ‘emulator’ for more complex and more expensive numerical simulations.

The goal of this work is to measure the detailed dependence of the inflow and outflow rates of mass, metals, and energy, on both CGM and ISM scales, on halo mass and redshift in the IllustrisTNG100 numerical cosmological simulation ([Nelson et al. 2019a](#); [Pillepich et al. 2018](#)). In a companion paper, we use these results to inform and benchmark a new semi-analytic model that attempts to reproduce the predictions of IllustrisTNG not only for the ‘state variables’ (galaxy properties such as stellar mass, gas mass, metallicity, etc) but also *flow rates* between different reservoirs, over a broad range in halo mass and redshift (Omoruyi et al. in prep).

Previous works have described aspects of the baryon cycle in and around simulated galaxies. For example, [Nelson et al. \(2015\)](#) tested the impact of feedback on cosmological gas accretion in the original Illustris simulation; [Nelson et al. \(2019b\)](#) investigated outflows in IllustrisTNG galaxies; [Pandya et al. \(2021\)](#) performed an analysis of mass, momentum, energy, and metal outflows in the FIRE-2 simu-

¹ <https://nap.nationalacademies.org/catalog/26141/pathways-to-discovery-in-astronomy-and-astrophysics-for-the-2020s>

lation suite; and Mitchell et al. (2020a) and Mitchell et al. (2020b) described the baryon cycle in the EAGLE simulation. Wright et al. (2024) tracked the mass flow rates in and out of galaxies in several cosmological simulations, including IllustrisTNG. In this paper, we build on these works by extracting both inflow and outflow rates, on scales chosen to match the traditional reservoirs used in SAMs (ISM and CGM), and tabulate these results over a more finely binned grid in halo mass and redshift than previous works. Another novel aspect of our analysis is that in addition to mass and metallicity, we also measure the energy flows on both ISM and halo scales, which has not been presented previously for the IllustrisTNG simulations. We present our results separately for galaxies that are dominated by SNe feedback and those that are AGN dominated, which has not generally been done in previous studies. In addition to being useful for developing next generation SAMs, our results provide numerous insights into the physics of how both stellar and AGN feedback shape the properties of galaxies and their CGM in the IllustrisTNG universe.

The structure of our paper is as follows: In Section 2 we describe the IllustrisTNG simulation suite and the galaxy sample we use for our analysis; in Section 3 we describe in depth the way feedback is implemented in the TNG simulation, and determine the dominant feedback mechanism in TNG halos as a function of mass and redshift; in Section 4 we describe how we measure the flow of mass, energy and metals, alongside defining loading factors; in Sections 5, 6, and 7 we analyze the flow of mass, energy, and metals in and out of TNG galaxies and halos; in Section 8 we discuss the interpretation of our results, and 9 we present our summary and conclusions.

2 THE ILLUSTRIS TNG SIMULATIONS

The IllustrisTNG (The Next Generation) simulation is an advanced cosmological hydrodynamical simulation designed to study galaxy formation, large-scale structures, and the interaction between dark and baryonic matter. It uses the AREPO code (Springel 2010), which adopts a moving mesh finite-volume method, to solve coupled equations of magneto-hydrodynamics (MHD), and a tree-based approach for gravity (Pakmor et al. 2011; Pakmor & Springel 2013).

To account for physical phenomena unresolved at the simulation scale, IllustrisTNG includes several subgrid models. Star formation is modeled with a density-dependent prescription, where dense, cold gas can form stars; a chemical enrichment model tracks the evolution of elements in stars and their distribution in the surrounding gas; stellar feedback in the form of supernovae injects energy into the surrounding gas, driving galactic winds; radiative mechanisms allow gas heating in the presence of background radiation alongside metal-line cooling; and supermassive black holes are formed, grow, and deposit energy into their surroundings in the form of thermal and kinetic AGN feedback (Pillepich et al. 2018; Weinberger et al. 2017). In Sec. 3 we dive deeper into the implementation of feedback via these subgrid processes.

The IllustrisTNG simulation suite includes three box sizes: 51.7, 110.7, and 302.6 Mpc on a side, each with a different resolution marked 1 (highest) to 3 (lowest). In this work we use Illustris TNG100-1, which is the highest resolution available for the 110.7 Mpc box size. This run includes 1820^3 DM particles with a mass of $7.5 \times 10^6 M_\odot$, and gas particles with a mass of $1.4 \times 10^6 M_\odot$, which form stars in a similar mass resolution. As dark-matter halos and galaxies form in the simulation, their identification is carried out using two algorithms: the Friends-of-Friends (FoF) algorithm is used to identify dark-matter halos, while the SUBFIND algorithm (Springel et al. 2001) identifies substructures in the formed halos,

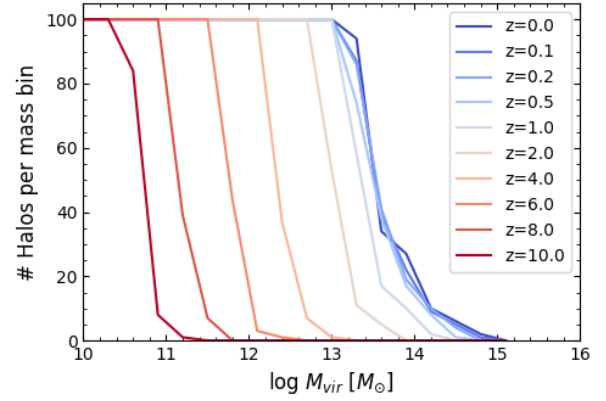


Figure 1. The distribution of TNG100 halos analyzed in this work as a function of halo virial mass, colored by redshift. In total, 9522 TNG100 halos are analyzed across 10 redshifts and in virial mass bins of 0.3 dex, starting from $10^{10} M_\odot$ and up to the most massive halo in each snapshot.

i.e. galaxies. To track the evolution of galaxies over time, the Sublink algorithm (Rodríguez-Gomez et al. 2015) is used to construct merger trees at the subhalo or galaxy level.

The cosmological parameters used in Illustris TNG are based on the Planck Collaboration et al. (2016) results: a Hubble constant $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$, matter and baryon density parameters $\Omega_{m,0} = 0.3089$ and $\Omega_{B,0} = 0.0486$, and a cosmological constant $\Omega_{\Lambda,0} = 0.6911$. The simulation evolves the universe from high redshift ($z = 127$) to the present ($z = 0$), spanning 100 snapshots. At $z = 0$, there are over 6×10^6 identified dark matter halos containing a total of over 4×10^6 galaxies.

2.1 Halo sample

To perform a robust analysis of galactic inflows and outflows and understand their influence over cosmic time, our analysis must span a wide range of redshifts and virial masses. We therefore analyze halos taken from 10 redshifts, each a different snapshot of the simulation: 0, 0.1, 0.2, 0.5, 1, 2, 4, 6, 8, and 10.

At each redshift, we randomly select up to 100 halos in bins of 0.3 dex in virial mass, starting from $M_{\text{vir}} = 10^{10} M_\odot$ and reaching up to the most massive halo in each snapshot. In total, across all redshifts we analyze 9522 halos. In Fig. 1 we present the number of halos per redshift and virial mass bin. Note that because we pick halos randomly at each redshift, we do not track the direct evolution of individual halos along a common merger tree, unlike Pandya et al. (2020), Pandya et al. (2021), and Pandya et al. (2023).

For each halo we collect all gas, dark matter, black hole, and star particles within 1.5 virial radii from the point of the minimum gravitational potential energy inside the halo. While we use the FoF and SUBFIND algorithms to identify halos and their central galaxies, we do not collect particles based on their association with each FoF halo, but pick particles based on their location. We have found that as we reach higher fractions of the virial radius, a significant portion of the particles that can be found around the halo are not associated with its FoF group. To analyze the feedback history in each halo, we additionally track the main progenitor branch of its central galaxy using the Sublink merger trees.

When discussing the virial mass we refer to $M_{\Delta,c}$, using the Bryan & Norman (1998) definition of the overdensity relative to the critical density of the universe. Appropriately, we define the virial radius

as $R_{\Delta,c}$. We define the ISM radius as $0.1R_{\text{vir}}$, in accordance with previous works (e.g., Muratov et al. 2015; Anglés-Alcázar et al. 2017a; Mitchell et al. 2020b; Pandya et al. 2021; Hafen et al. 2022). As for galactic properties, we take the radius of the galaxy to be twice its stellar half-mass radius, inside of which we estimate properties such as stellar mass or star formation rate².

A significant point to consider is whether the simulation results we analyze are well resolved. Unresolved galaxies can be dominated by shot noise that increases with fewer particles. It is common to consider resolved galaxies as those that have more than 100 particles, corresponding to a shot noise of under 10% (e.g. Pillepich et al. 2018; Nelson et al. 2019b; Mitchell et al. 2020a). In TNG100, this limit translates to stellar masses $\sim 10^8 M_\odot$, equivalent to a virial mass of $\sim 10^{10.8} M_\odot$. As the mass, energy, and metal outflow rates we measure are driven by galactic subgrid processes, in our results we use a gray band to highlight all results from halos with $M_{\text{vir}} < 10^{11} M_\odot$, or our first four mass bins, to indicate they may be dominated by shot noise.

An important distinction to make between galaxies in our sample in the context of their star formation rate is whether they are considered “star forming” or “quenched”. In Fig. 2 we plot the total instantaneous star formation rates inside the central galaxy associated with each halo, for 4 of the redshifts we analyze in this work. We mark halos whose central galaxy is star forming in blue, and halos whose central galaxy is quenched in red. We determine which is which using the method described in Karmakar et al. (2023): for each redshift, we estimate the median specific star formation rate ($sSFR$) for galaxies with stellar masses in the ranges $\log(M_*/M_\odot) = 9.25 \pm 0.25$ and $\log(M_*/M_\odot) = 9.75 \pm 0.25$. We draw a line on the $\log(M_*) - \log(sSFR)$ plane between these median $sSFR$ s and the centers of the stellar mass bins to which they correspond, which is then extrapolated to higher and lower masses. Galaxies with $sSFR$ higher than 25% of the value on this line (according to their stellar mass) are considered star forming. We mark the median SFR per virial mass bin with a black dashed line. As described in previous works such as Davies et al. (2020), the quenching of star formation above the virial masses of halos surrounding L^* galaxies in TNG100 is associated with AGN feedback starving galaxies of their star forming fuel by heating and ejecting significant amounts of gas from their ISM. That being said, to better understand the effects of feedback on the galaxy and on the halo in which it resides, we need to further discuss how the feedback is implemented.

3 FEEDBACK IN TNG100

In the TNG100 simulation, winds are driven by SNe and AGN, both implemented as subgrid recipes. The implementation of star formation is described in Springel & Hernquist (2003) and Vogelsberger et al. (2013), and the launching of winds due to SNe is described in Pillepich et al. (2018). The seeding of BH particles in growing halos is also described in Pillepich et al. (2018), and the growth and feedback modes from BH particles are described in Weinberger et al. (2017). The numerical values shown in the following two sections are all taken from the fiducial, publicly available, TNG100 run.

² This radius is included in the public TNG catalog.

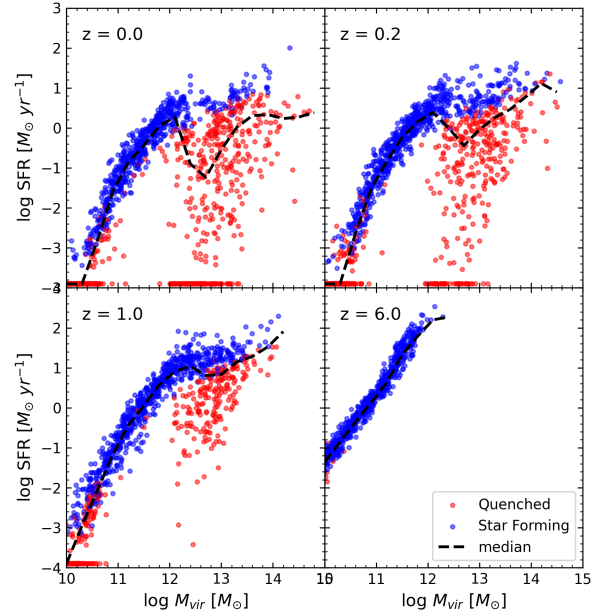


Figure 2. The instantaneous star formation rates in the central galaxies of all halos from four of the redshifts analyzed in this work (top left: $z = 0$; top right: $z = 0.2$; bottom left: $z = 1$; bottom right: $z = 6$), as a function of halo virial mass. Halos whose central galaxies are star forming are marked in blue, and halos whose central galaxies are quenched are marked in red (classification into star forming or quenched is based on the method introduced in Karmakar et al. 2023). At each redshift, the median is plotted using a dashed black line.

3.1 SFR and SNe feedback

Star formation in the TNG simulation is implemented stochastically. Any gas particle above a certain density threshold has a probability to form stars in a rate determined by a characteristic star-formation timescale. Assuming a Chabrier (2003) initial mass function, A fraction of the formed stellar mass then explodes as SNe, which launch wind particles into the surrounding interstellar medium. These wind particles are launched in random directions with an initial velocity that is given by:

$$v_w = \max \left[7.4 \sigma_{\text{DM}} \left(\frac{H_0}{H(z)} \right)^{1/3}, 350 \text{ km/s} \right], \quad (1)$$

where σ_{DM} is the local dark matter velocity dispersion. Notice that the wind velocity is capped at a maximum value of 350 km/s. The rate of wind mass flow per unit star formation rate is determined by a mass loading factor, given by:

$$\eta_{w, \text{TNG}} = \frac{2}{v_w} e_w \times 0.9. \quad (2)$$

The multiplication of the mass loading factor by a factor of 0.9 implies that 90% of the energy of the formed wind particle is kinetic, and the remaining 10% is thermal. The parameter e_w is a specific energy that is attributed to a wind particle depending on the metallicity of the star forming gas cell and on a Type II SN rate per unit mass of stars formed, calculated assuming a Chabrier (2003) IMF:

$$e_w = \left[1 + \frac{3}{1 + (Z/0.002)^2} \right] \times 1.06 \times 10^{49} \text{ erg } M_\odot^{-1}. \quad (3)$$

Thus, the specific energy ranges between $e_w = 4.25 \times 10^{49} \text{ erg } M_\odot^{-1}$

for metallicities far below the $Z = 0.002$ threshold (equivalent to $\sim 16\%$ of $Z_\odot$) and $e_w = 1.06 \times 10^{49} \text{ erg M}_\odot^{-1}$ for metallicities far above it. The newly formed wind particle has the following energy rate:

$$\dot{E}_{\text{SNe}} = e_w \times SFR. \quad (4)$$

The formed wind particles are decoupled from the gas that spawned them, until leaving the ISM and recoupling with the surrounding gas, depositing their momentum, mass, metals, and thermal energy content. The metallicity of the newly formed wind particle is 0.4 relative to its parent gas cell, as described in [Vogelsberger et al. \(2013\)](#) for the Illustris simulation, and is adopted for IllustrisTNG as well. For a more detailed description of the implementation of SNe feedback in IllustrisTNG, see § 2.3.2 of [Pillepich et al. 2018](#). In our analysis of flow rates, we track stellar wind particles as if they were gas particles.

3.2 BH growth and AGN feedback

In the simulation, a BH seed mass of $1.2 \times 10^6 \text{ M}_\odot$ is created and placed in the gravitational potential minimum of a FoF halo that has reached a virial mass threshold of $M_{200,c} = 7.4 \times 10^{10} \text{ M}_\odot$. Each BH particle in TNG100 is allowed to accrete gas at the Bondi accretion rate, which is capped at the Eddington accretion rate. Unlike stellar feedback, AGN feedback in TNG100 comes in one of two states, determined by $\dot{M}_{\text{Bondi}}/\dot{M}_{\text{Eddington}}$, the Bondi accretion rate of the BH particle relative to its Eddington accretion rate. The threshold for delineating between the two feedback states is given by:

$$\chi = \min \left[2 \times 10^{-3} \left(\frac{M_{\text{BH}}}{10^8 \text{ M}_\odot} \right)^2, 0.1 \right]. \quad (5)$$

As the mass of the black hole particle reaches $\sim 7 \times 10^8 \text{ M}_\odot$ (the median BH mass in $\sim 10 \times 10^{13} \text{ M}_\odot$ halos), the value of χ is capped at 0.1.

If $\dot{M}_{\text{Bondi}}/\dot{M}_{\text{Eddington}} \geq \chi$, the BH is assumed to be in a “high-accretion” state, in which pure *thermal* energy is injected in the region around the BH particle. The rate of energy injection is:

$$\dot{E}_{\text{th, AGN}} = 0.02 \dot{M}_{\text{BH}} c^2. \quad (6)$$

If $\dot{M}_{\text{Bondi}}/\dot{M}_{\text{Eddington}} < \chi$, the BH is in a “low-accretion” state. In this state, the injected energy rate is 10 times higher:

$$\dot{E}_{\text{kin, AGN}} = 0.2 \dot{M}_{\text{BH}} c^2. \quad (7)$$

The energy released by black holes in the “low-accretion” mode is accumulated until reaching a threshold of:

$$E_{\text{inj, min}} = 10 \sigma_{\text{DM}}^2 m_{\text{enc}}, \quad (8)$$

where m_{enc} is the mass enclosed in the feedback region. Past this threshold, a feedback “event” occurs where the energy is injected into the surrounding gas in the form of pure *kinetic* energy. Per feedback event, gas particles around the BH are given momentum in a random direction. For a more detailed description of the way AGN feedback is implemented in IllustrisTNG, see Sec. 2 of [Weinberger et al. 2017](#). In this work we interchangeably refer to the AGN “high-accretion” state as “thermal feedback”, and the “low-accretion” state as “kinetic feedback”.

The transition from thermal to kinetic AGN feedback depends on the BH mass such that a newly formed BH particle is in the thermal state, and the likelihood for a BH particle to be in the kinetic state increases with mass. In fact, [Weinberger et al. \(2017\)](#) show that even

before reaching the mass relevant to the maximum value of χ , most black holes have transitioned from thermal to kinetic feedback. As kinetic feedback drives gas away from the BH more effectively, it also limits the gas available for accretion by the BH particle, further reducing its accretion rate and preventing it from getting back to a thermal feedback mode.

3.3 Dominant feedback mechanism

In Fig. 3 we plot the median instantaneous energy injection rates for all three mechanisms (SNe (thermal + kinetic), thermal AGN feedback, and kinetic AGN feedback) as a function of virial mass and for different redshifts. Scattered points are colored by their redshift, and error bars represent the standard deviation for the mass bin, equivalent to the 16-84 percentile range (this is true for all errorbars presented in this work). While SNe feedback occurs for all masses and redshifts, no AGN feedback occurs for $M_{\text{vir}} \lesssim 4 \times 10^{10} \text{ M}_\odot$ due to the lower limit on the halo mass for BH seeding in TNG. Due to the way the AGN feedback sub-grid recipe is implemented, TNG galaxies generally experience either thermal or kinetic feedback, but not both simultaneously. The relative importance of kinetic and thermal AGN feedback depends on both mass and redshift. We find that the criteria for kinetic AGN feedback are only met at $z \lesssim 2$ and for $M_{\text{vir}} \gtrsim 10^{12} \text{ M}_\odot$, while thermal AGN feedback occurs at all redshifts, but only in the mass range of $4 \times 10^{10} \lesssim M_{\text{vir}}/\text{M}_\odot \lesssim 2 \times 10^{13}$.

We wish to determine which feedback mechanism is more significant in driving outflows at a given time. As a proxy to this question, we compare the energies injected into the interstellar medium by each mechanism. To prevent bias due to peaks in the instantaneous energy injection rates, we estimate for each halo in our analysis the total energy injected by each mechanism over time periods that correspond to three typical travel times from the galaxy to the virial radius. We measure the distribution of outflow velocities at the virial radius of each halo at the redshift of interest, and record the 10th percentile, 50th percentile and 90th percentile velocities. Using the virial radius at the redshift of interest and under the simplifying assumption that the gas moves at a constant velocity from its launching point, we estimate the travel times associated with the aforementioned velocity percentiles. We then track the evolution of the central galaxy associated with each halo using the SubLink merger trees created for TNG100, and collect the instantaneous energy injection rates from each feedback mechanism down to the earliest snapshot in TNG100. With the travel times and energy injection rate histories in hand, we can integrate from the three estimated launch times up to the snapshot in which we are measuring the gas properties. We determine the dominant feedback mechanism as the one whose integrated injected energy is highest for two out of the three estimated timescales.

Thermal feedback does not necessarily drive an outflow, as radiative losses have an appreciable effect ([Weinberger et al. 2017](#)). We therefore perform this comparison twice — once only for kinetic feedback, and once for kinetic and thermal feedback together (i.e. using the full feedback model)³. The results are shown in Fig. 4, in which we count the fraction of AGN dominated halos in bins of virial mass and redshift.

We find that for $z \leq 6$, at some point there is a transition from SNe dominated feedback to AGN dominated feedback as halos pass mass thresholds that decrease with cosmic time. Except for one case

³ While a BH particle in TNG100 can only be in either thermal mode or kinetic mode, a galaxy can contain more than one BH particle, and each of them can be in a different feedback mode.

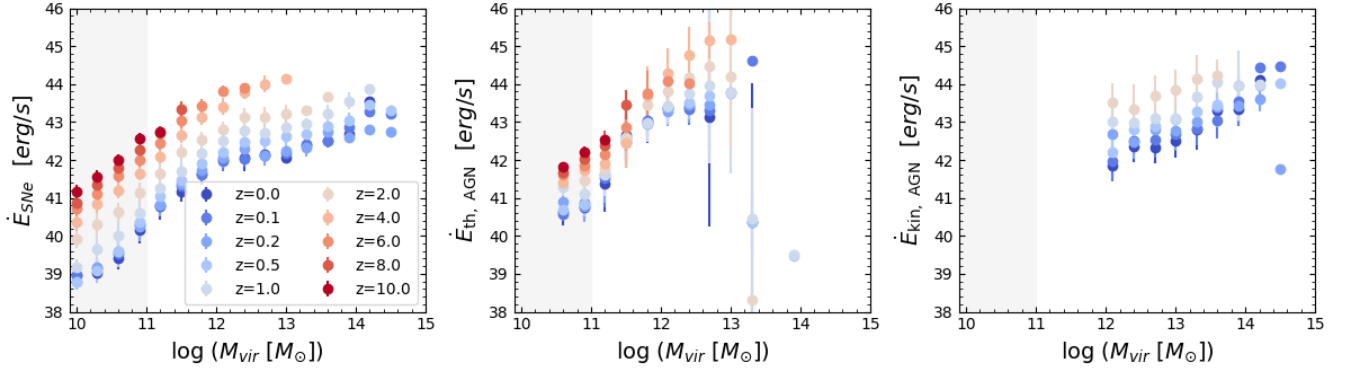


Figure 3. Instantaneous energy injection rates from the different feedback mechanisms in the TNG100 simulation, for all halos and redshifts analyzed in this work. Left panel: SNe feedback, based on the subgrid recipe described in [Pillepich et al. \(2018\)](#). Central panel: thermal / high-accretion mode AGN feedback. Right panel: kinetic / low-accretion mode AGN feedback. Both AGN energy injection rates are based on the subgrid recipe described in [Weinberger et al. \(2017\)](#). In all panels, marker colors range from blue ($z = 0$) to red ($z = 10$). SMBH generally begin their life in the thermal mode, and eventually transition to the kinetic mode.

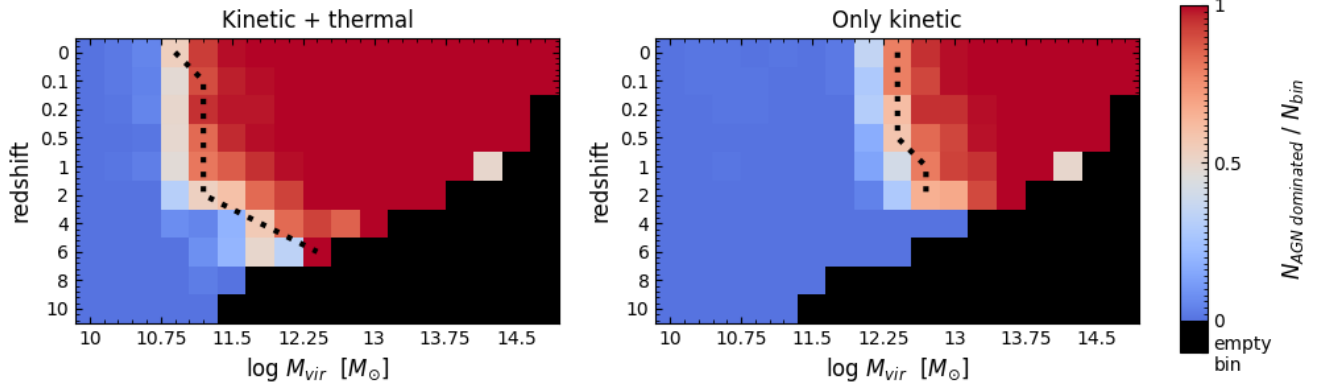


Figure 4. Fraction of halos, in redshift and virial mass bins, for which AGN are the dominant feedback energy injection mechanism. Fractions range from 0 (blue — all halos in bin dominated by SNe feedback) to 1 (red — all halos in bin dominated by AGN feedback). Bins with no halos are colored in black. We use a dotted black line to mark the lowest mass bins, per redshift, where the fraction of AGN dominated halos is greater than 0.5. Left panel: combination of kinetic + thermal feedback. Right panel: only kinetic feedback. In both cases, soon after AGN begin injecting energy into their surroundings, they become the dominant feedback energy injection mechanism.

at $z = 1$, the transition is total, meaning that all halos become AGN dominated as we go to higher and higher masses⁴.

It is evident that soon after AGN feedback kicks in it becomes dominant, in both the kinetic case alone and for the full feedback model. As the energy injection rates plotted in Fig. 3 show that thermal AGN feedback does not occur in halos more massive than $\sim 2 \times 10^{13} M_{\odot}$, the transitions between dominant feedback energy injection mechanisms shown in Fig. 4 are also an output of the “built in” transition from thermal to kinetic AGN feedback in TNG100 halos. We hereafter refer to the virial mass bin in which, for a given

redshift, the dominant feedback energy injection mechanism transitions from SNe feedback to AGN feedback as “the onset of kinetic AGN feedback” when considering kinetic AGN feedback alone, and “the onset of thermal AGN feedback” when considering both thermal and kinetic modes. We use a dotted black line in Fig. 4 to mark the onset of thermal (left panel) and kinetic (right panel) AGN feedback.

4 EXTRACTING THE FLOW QUANTITIES

With an understanding of the dominant feedback energy injection mechanism for each virial mass bin and redshift, we can now analyze the gas flows in and out of galaxies and their host halos and correlate said flows with how the feedback is implemented. As mentioned in Sec. 2.1, for each of our 9522 halos we collect all gas, dark matter, and black hole particles within $1.5R_{\text{vir}}$. To estimate flow rates at the inner ISM scale, we select particles in shells between $0.05R_{\text{vir}}$

⁴ The outlier at $z = 1$ is the most massive halo for this snapshot in TNG100, and is one of only two halos in its mass bin (hence the fraction of AGN dominated halos in the appropriate cell is exactly 0.5). This halo has a very high star formation rate — in fact, it is > 6 times higher than the median SFR of halos in adjacent mass bins and the highest SFR in the entire $z = 1$ snapshot — while its BH accretion rate is close to that of similar mass halos.

and $0.15R_{\text{vir}}$. At the halo scale the flow shells are from $0.95R_{\text{vir}}$ to $1.05R_{\text{vir}}$. To separate inflows from outflows, we consider the radial velocity component of each particle relative to the halo centers. Outflowing particles have positive radial velocities, and inflowing particles have negative radial velocities.

With each particle flagged by its direction and location, we can calculate the instantaneous mass inflow and outflow rates through a radial shell using:

$$\dot{M} = \sum_i \dot{m}_i = \sum_i \frac{m_i v_{r,i}}{\Delta r}, \quad (9)$$

where the sum is over all outflowing or inflowing particles inside the shell, $\dot{m}_i$ is the mass flow rate for a single gas particle, m_i is the particle mass, $v_{r,i}$ is its radial velocity, and $\Delta r = 0.1R_{\text{vir}}$ is the width of the shell.

We consider two components for the energy flows. First is the thermal energy flow rate:

$$\dot{E}_{\text{th}} = \frac{3}{2} \sum_i \dot{m}_i c_{s,i}^2, \quad (10)$$

where $c_{s,i}^2$ is the particle sound speed. Second is the kinetic energy flow rate:

$$\dot{E}_{\text{kin}} = \frac{1}{2} \sum_i \dot{m}_i v_i^2 \quad (11)$$

where here, v_i is the total velocity of the gas particle (not just its radial component.) The total energy flow rate, kinetic plus thermal, is $\dot{E} \equiv \dot{E}_{\text{kin}} + \dot{E}_{\text{th}}$.

Finally is the metal mass flow rate:

$$\dot{M}_Z = \sum_i \frac{m_i Z_i v_{r,i}}{\Delta r}, \quad (12)$$

where Z_i is the heavy element mass fraction of the particle.

Similar to our discussion of resolved galaxies in TNG100, we may ask ourselves whether the flow rates themselves are resolved due to a small number of particles per shell. To inspect this, we tested how the number of particles in the ISM and halo shells scale with the halo virial mass, and whether the flow rates are sensitive to the shell width. We find that the number of gas particles per shell keeps a constant relation with the host halo mass down to our lowest mass bin, and that changing our shell thickness from $0.05R_{\text{vir}}$ to larger fractions of the virial radius does not alter any of our results at all halo masses and shell widths. These scalings alongside the large sample of 100 galaxies per mass bin lead us to conclude that our flow rates should be robust and properly represent the behavior in TNG100, regardless of whether the galaxy that drives the flow is resolved.

4.1 Loading factors

As is standard, we define the loading factors relative to the star formation rates, associated SNe energy injection rates, and estimated metal injection rates. The mass loading factor is (as mentioned in Sec. 3.1):

$$\eta_M = \dot{M}_{\text{out}} / SFR \quad (13)$$

where SFR is the instantaneous galaxy star formation rate as given by the simulation, and $\dot{M}_{\text{out}}$ is our computed mass outflow rate.

The energy loading factor is:

$$\eta_E = \dot{E}_{\text{out}} / \left(\frac{SFR}{100 M_{\odot}} \right) \times E_{\text{SN}}, \quad (14)$$

where $E_{\text{SN}} = 10^{51}$ erg is the energy associated with a SN explosion, and $SFR/(100 M_{\odot})$ is a SN rate of 1 per 100 solar masses formed assuming a Kroupa (2001) IMF.

The metal loading factor is:

$$\eta_Z = \dot{M}_{Z,\text{out}} / (y \times SFR) \quad (15)$$

Here $y = 0.02$ is the approximate metal yield from SNe, obtained using the same assumption for SN rate and a mean ejecta mass of $10 M_{\odot}$, out of which 20% are metals (adopted by e.g. Kim et al. 2020a and Pandya et al. 2021).

Since in this work we define the loading factors relative to rates driven by star formation only (either the star formation rate itself, the energy released in SNe explosions, or the rate of metals deposited into the ISM by SNe), we anticipate a significant increase in all loading factors at the onset of kinetic AGN feedback, or as we reach virial masses of $\sim 3 \times 10^{12} M_{\odot}$, above which galaxies begin to quench (see Figs. 4 and 2) and AGN become the dominant forces in driving gas, energy, and metals out of the ISM. Therefore, in this work we only present the loading factors for halos with virial masses of less than $3 \times 10^{12} M_{\odot}$.

5 MASS FLOW RATES

We now start with analysis of the mass flow rates through the inner galaxy ISM shells and outer halo shells. We begin with a straightforward analysis of inflow and outflow rates, and whether galaxies and halos gain or lose mass. We touch on the phenomena behind these net flows, and how they are correlated with other, more commonly discussed properties of TNG100 halos and galaxies (SFR and baryon fractions). We then focus on the mass inflow rates, first in the context of their deviation from expectations based on DM mass flow rates, and later we track their progression through the halo scale, down to the ISM scale and as they become fuel for star formation. We end with the mass loading factors out of TNG100 galaxies.

5.1 Net inflows and outflows

In Fig. 5 we plot the gas mass inflow and outflow rates for all redshifts at the ISM and halo scales, as functions of the virial mass. We normalize by M_{vir} since to first order we expect the baryonic inflow rates to be proportional to the dark matter inflow rates, which are themselves approximately proportional to the total virial masses (e.g. Genel et al. 2008). For each redshift, we mark mass bins dominated by SNe feedback with open circles, mass bins dominated by thermal AGN feedback with full circles, and mass bins dominated by kinetic AGN feedback with full triangles. In Fig. 6 we plot the ratios of the inflow to outflow rates, at the ISM and halo scales, to determine whether there are net inflows or outflows at these two scales. The colors and markers are as in Fig. 5.

Several trends appear in Figs. 5 and 6. As seen in Fig. 5 at the ISM scale (right panels), the inflow and outflow rates are comparable. For a given redshift, the normalized flow rates at the ISM shell (in both directions) increase with mass and peak near $\sim 2 \times 10^{11} M_{\odot}$. They then decline, and increase again for halo virial masses of $\sim 2 \times 10^{13} M_{\odot}$. As seen in Fig. 6 (right panel) overall the net flows are into the ISM, enabling the galaxies to grow. The net inflow rates decrease with increasing halo mass. At the particular redshifts and masses where kinetic AGN feedback turns on, the outflow rates become nearly equal to the inflow rates, with a slight preference toward net outflows. At redshifts where kinetic AGN feedback is not active, flows remain inwards for all virial masses. The balance between galactic

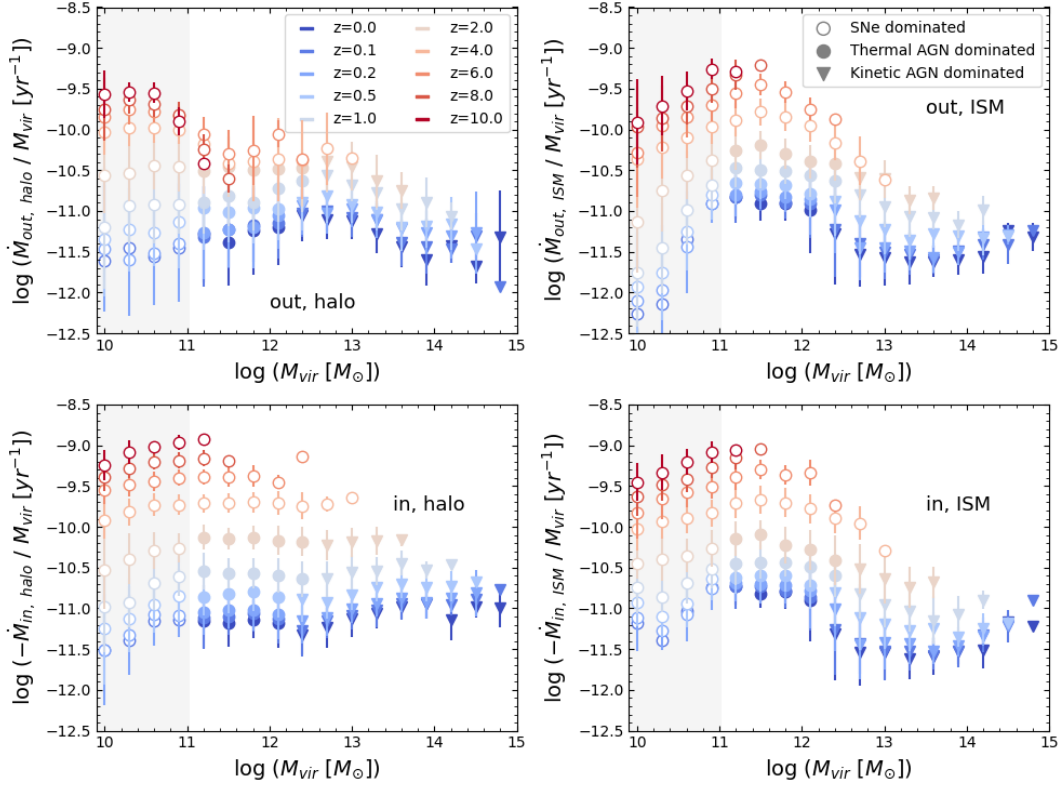


Figure 5. Gas mass outflow rates (top row) and inflow rates (bottom row), estimated at the halo scale (left column) and at the ISM scale (right column), all normalized by the halo virial mass and presented as a function of the halo virial mass. The gas mass flow rates are calculated from $z = 10$ (red markers with error bars) to $z = 0$ (blue markers with error bars). For each redshift, we mark halos dominated by SNe feedback with open circles, halos dominated by thermal AGN feedback with full circles, and halos dominated by kinetic AGN feedback with full triangles.

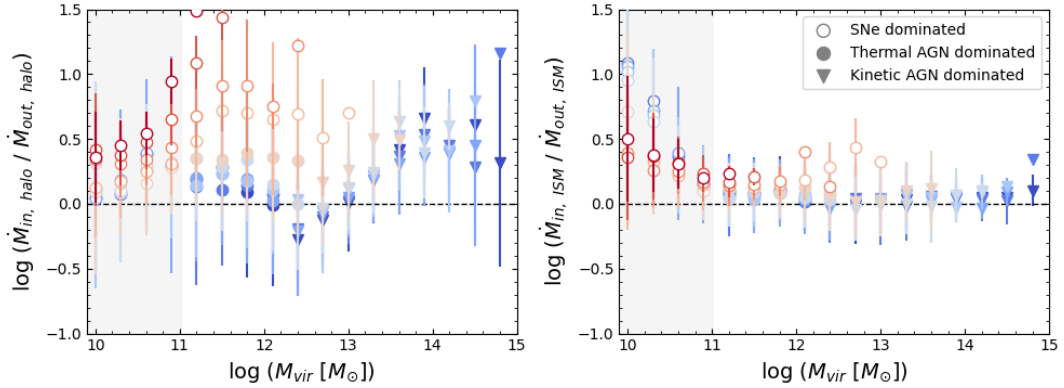


Figure 6. The ratio between the gas mass inflow and outflow rates at the virial scale (left panel) and the ISM scale (right panel). Colors and symbol types are identical to those in Fig. 5. In both panels, and for most redshifts and virial masses, $\dot{M}_{\text{in}} > \dot{M}_{\text{out}}$, suggesting a net inflow into both halos and galaxies. The net inflows, especially on halo scales, are much stronger at high redshift.

outflow and inflow rates persists up to halo virial masses $\sim 10^{14} M_{\odot}$, after which ISM inflow rates again become significantly larger than outflow rates — most likely due to the growing binding energy of the galaxies within. In low mass halos the ratios of the inflow to outflow rates can reach rather large values (> 0.25 dex). Strong net galactic inflows also appear in high mass ($M_{\text{vir}} > 10^{12} M_{\odot}$), high redshift ($z \geq 4$) halos in which there is no recent AGN feedback.

At the halo scale (right panels of Fig. 5) the behavior is more complicated, with different behaviors for the inflows and outflows. For low virial masses and at high redshifts ($z \geq 4$), the normalized *outflow* rates are relatively constant up to halo virial masses of $\sim 10^{11} M_{\odot}$, after which they begin decreasing. At lower redshifts, however, normalized mass outflow rates increase with the halo virial mass. Above $3 \times 10^{12} M_{\odot}$, the normalized mass outflow rates decrease

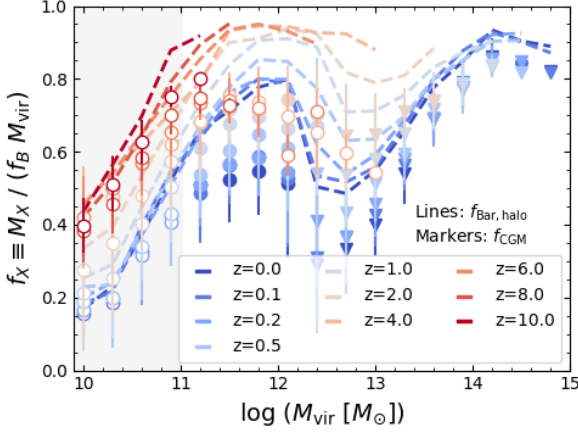


Figure 7. Baryon fractions, defined as the combined mass of all baryons within a certain population normalized by $f_B M_{\text{vir}}$, for the CGM alone (scattered markers with errorbars) and for the entire halo (dashed lines). Colors and marker styles are identical to those in Fig. 5. At redshifts where kinetic AGN feedback is active, both f_{CGM} and $f_{\text{Bar,halo}}$ decrease at the onset of kinetic AGN feedback. At redshifts where there is no kinetic AGN feedback, the drops in f_{CGM} and $f_{\text{Bar,halo}}$ are less significant.

at all redshifts, regardless of the feedback mechanism (the absolute outflow rates are still increasing). The normalized mass inflow rates, on the other hand, slightly increase with the halo virial mass for all redshifts up to $\sim 3 \times 10^{11} M_{\odot}$, after which they decrease until reaching the onset of kinetic AGN feedback, and then increase again. As seen in Fig. 6 the net flows are predominantly inward by large factors. The exceptions are low redshift ($z \leq 0.5$) halos near $\sim 3 \times 10^{12} M_{\odot}$ at the onset of kinetic AGN feedback where substantial net outflows occur. This is consistent with the results shown in e.g. Davies et al. (2020), Wright et al. (2024), and Oren et al. (2024), which show a minimum of the CGM gas mass fraction at this virial mass for $z = 0$. An additional case of net mass outflows at the halo scale is seen for dwarf galaxies (at the centers of $\sim 10^{10} M_{\odot}$ halos) at $z = 0$, where we find slightly higher outflow rates than inflow rates. Unlike for the onset of kinetic AGN feedback, we do not find these net outflows at the ISM shells surrounding dwarf galaxies; in fact, at this mass bin we find ISM inflows to be the highest relative to ISM outflows.

To further investigate the correlation between the onset of kinetic AGN feedback and the decrease in the baryon fractions of TNG100 halos, we turn to Fig. 7. In it, we show the fractions of the virial mass that are taken by the total baryonic mass within the virial radius, and those taken by the CGM mass, both normalized by the cosmological baryon fractions (labeled $f_{\text{Bar,halo}}$ and f_{CGM} respectively). At redshifts where kinetic AGN feedback is active, there is a local minimum of both f_{CGM} and of $f_{\text{Bar,halo}}$ around the onset of kinetic AGN feedback, similar to the case at $z = 0$ and in accordance with the appearance of net mass outflows at the halo scale. The drop in both the CGM and baryon fractions is more significant at lower redshifts, where presumably kinetic AGN feedback has been active longer. On the other hand, at $z = 4$, where there is no kinetic AGN feedback, f_{CGM} decreases past $M_{\text{vir}} \approx 10^{12} M_{\odot}$ while the total baryon fraction remains relatively constant. This suggests that at this redshift, the baryons lost from the CGM fall into the galaxy, unlike at lower redshifts where they are ejected out of the halo entirely due to AGN feedback. This is supported by the right panel of Fig. 6, showing a strong net inflow into the galaxy at high masses at $z = 4$, and the previously discussed increase in star formation rates for this case. At

higher redshifts our sample does not include enough massive halos to investigate this trend.

5.2 Comparison with DM halo growth rates

To first order, we expect the rate that baryons flow into a halo to be proportional to the rate that the halo is accreting dark matter. We examine deviations from this assumption in Fig. 8, in which we present the gas mass inflow rates normalized by $f_B \dot{M}_{\text{in,halo}}^{\text{DM}}$, the DM mass inflow rate at the virial radius multiplied by the cosmological baryon fraction. We show these at the halo scale (left panel) and the ISM scale (right panel).

At the halo scale the normalized flow rates are usually close to but somewhat smaller than unity. We find two interesting cases for which the inflows are significantly reduced: First is for dwarf galaxies at all redshifts, and second is at masses associated with the onset of kinetic AGN feedback. Both reductions are consistent with strong feedback which suppresses inflows into the galaxy; in the case of the low-mass halos, the $z = 0$ point (numerically the lowest) is also a case where there are net outflows from the halo, but not the galaxy as seen in Fig. 6. The case of halos at the onset of kinetic AGN feedback is also consistent with Fig. 6, where we find net outflows at these masses. An additional interesting outlier at the halo scale are high-redshift ($z \geq 2$) inflows into halos with virial masses $\gtrsim 10^{11} M_{\odot}$, which are higher than expected based on DM flow rates alone, suggesting recycling of previously ejected gas at these scales.

At the ISM scale in Fig. 8 we find two notable features. First, at virial masses around $10^{11} M_{\odot}$ the normalized mass inflow rates exceed unity by up to 0.25 dex. This excess may also suggest recycling of galactic outflows that are falling back into the galaxy (also known as a “galactic fountain”), in addition to fresh IGM gas streaming into the halo. Second, in dwarfs and in more massive halos ($M_{\text{vir}} \gtrsim 3 \times 10^{12} M_{\odot}$) the baryonic mass inflow rates are reduced at the ISM scale. This is a signature of preventative feedback: fresh IGM gas which enters the halo but is prevented from reaching the ISM, contributing to growth of the CGM (see e.g. Davé et al. 2012). This preventative feedback is caused by SNe driven winds at the low mass end, and by kinetic AGN feedback at the high mass end. We measure the strongest preventative feedback in halos with $M_{\text{vir}} \approx 2 \times 10^{13} M_{\odot}$, above which galactic mass inflow rates become closer to the expectations from DM flows. Still, galaxies in TNG100 do not reach masses which allow their potential wells to overcome the energy injected by their AGN.

5.3 Progression of inflows

In Fig. 9 we address the question of what fraction of the gas mass that flows into a halo enters the galaxy (left panel), how much of the entering gas is used for star formation (central panel), and with what efficiency the galaxy converts accreted gas into stars (right panel).

For most virial masses and redshifts, the left panel of Fig. 9 is consistent with the comparison between the baryonic inflow rates to the DM inflow rates. For virial masses above $\sim 3 \times 10^{12} M_{\odot}$ in all redshifts, and for virial masses lower than $\sim 10^{11} M_{\odot}$ at $z \geq 2$, the halo inflow rates are higher than the ISM inflow rates, consistent with the markers of preventative feedback. For halos at $z \leq 2$ with virial masses $\sim 10^{11} M_{\odot}$ the inflow rates into the galaxy surpass those that fall into the halo, again supporting an additional contribution to the gas falling in from the IGM, perhaps in the form of previously ejected gas cooling and falling back into the galaxy (a “galactic fountain”).

An interesting outlier, however, is halos in the virial mass bin of $M_{\text{vir}} = 10^{10} M_{\odot}$. At $z \leq 2$, we find that inflows into the ISM surpass

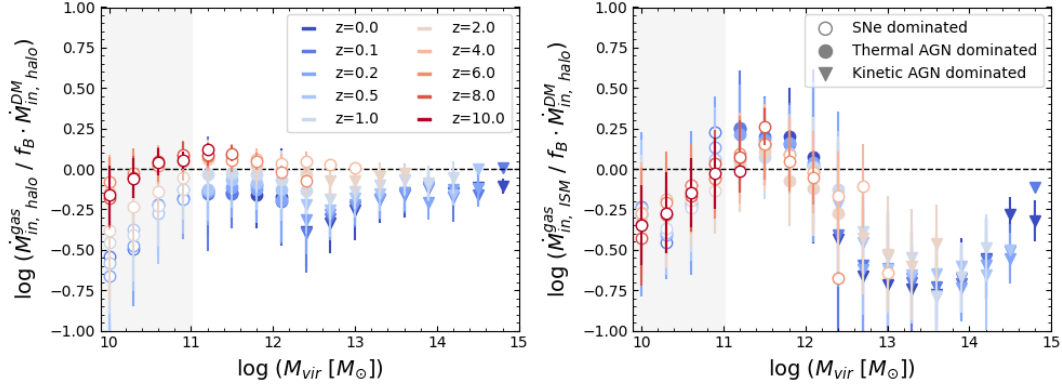


Figure 8. Gas mass inflow rates, estimated at the halo scale (left panel) and at the ISM scale (right panel), normalized by the dark matter inflow rates multiplied by the baryon fraction. Colors and symbol types are identical to those in Fig. 5. At the halo scale, we see inflow rates that slightly exceed the baseline rate $f_B \dot{M}_{in, halo}^{DM}$ at high redshift, and are lower than the baseline at low redshift, indicating preventative feedback. At the ISM scale, at halo masses $\sim 10^{11} M_\odot$ – $10^{12} M_\odot$, we see inflows in excess of the baseline, indicating rapid recycling of ejected gas on this scale in a galactic fountain. At higher halo masses we see a strong suppression of inflows due to preventative feedback, primarily caused by the kinetic mode AGN feedback.

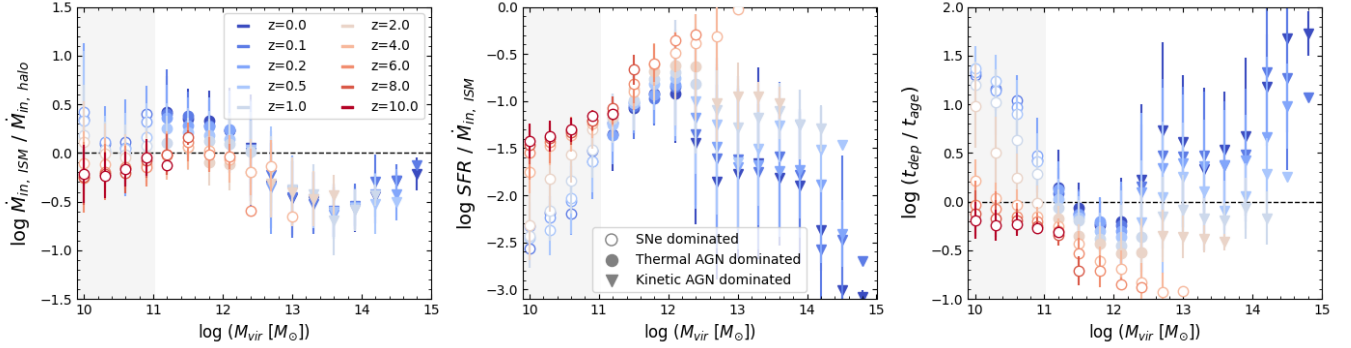


Figure 9. Left panel: The ratio between the mass inflow rates at the ISM scale to those at the halo scale, indicating which fraction of the gas flowing into the halo makes it to the galaxy. At virial masses of $\sim 3 \times 10^{11} M_\odot$ we find that fraction to be larger than 1, indicating rapid cooling and infall of previously ejected ISM gas in addition to fresh IGM gas; in halos with virial masses of $\geq 10^{12} M_\odot$, we find an indication of preventative feedback and build-up of a hot CGM. Central Panel: the ratio between the gas mass flow rates into the galaxy and the instantaneous star formation rate in the central galaxy, indicating which fraction of the galactic inflows was used for star formation. This fraction initially increases with halo mass, only to drop after the onset of kinetic AGN feedback. Right panel: t_{dep} , the ISM depletion time, normalized by the age of the universe (t_{age}). For all redshifts, the depletion time drops below the age of the universe in halos with virial masses of $\sim 10^{11}$ – $10^{12} M_\odot$. The depletion time becomes long compared to t_{age} in more massive halos. In all panels, colors and symbol types match those of Fig. 5.

those into the halo, while a comparison with DM flow rates suggested reduced inflows into both the halo and the ISM at these masses and times. Recall that at $z = 0$ we find net outflows from the halo scale at this mass bin, while simultaneously seeing strong net inflows into the ISM.

The central panel of Fig. 9 shows the ratio between the star formation rates of the central galaxies to the inflow rates into the ISM. For halos with $z \leq 2$, the fraction of inflowing gas that is used for star formation increases with virial mass, only to decrease after the onset of kinetic AGN feedback. For $z \geq 4$, where kinetic AGN feedback has not yet occurred, there is no turnover and the ratio keeps increasing with virial mass. Most notably, for the highest mass bin at $z = 4$, all gas that flowed into the galaxy is used for star formation. While at $z \leq 2$ thermal AGN feedback appears to be in effect, it is not as effective at suppressing star formation as kinetic AGN feedback: the dependence of $SFR/\dot{M}_{in, ISM}$ on the halo virial mass past the onset of

thermal AGN feedback does not appear different from that of halos where SNe feedback is the dominant feedback mechanism.

The right panel of Fig. 9 shows the ratio between the ISM depletion time, defined as $t_{dep} \equiv M_{ISM}/SFR$, normalized by the age of the universe t_{age} . By comparing the two timescales, we can get a sense of the star formation efficiency within the galaxy. In low mass halos, with virial masses $\leq 10^{11} M_\odot$, depletion times are very high at recent times, yet at $z \geq 4$ they converge to $\sim 60\%$ of the age of the universe. This suggests SNe feedback at these masses becomes less effective in suppressing star formation as time progresses. On the other hand, star formation is most effective in galaxies populating halos between $\sim 10^{11} M_\odot$ and $\sim 3 \times 10^{12} M_\odot$, regardless of the dominant feedback mechanism — SNe feedback or thermal AGN feedback.

In halos more massive than $\sim 3 \times 10^{12} M_\odot$, the behavior varies depending on the dominant feedback mechanism. At the onset of kinetic AGN feedback we measure an increase of the depletion time,

which becomes more drastic the lower the redshift. At $z = 2$ and $z = 1$, when kinetic AGN feedback initially kicks in, it slightly increases the depletion time but still does not bring it above the age of the universe. At $z \leq 0.5$, however, star formation rates decrease rapidly as kinetic AGN feedback drive significant amounts of gas out of galaxies. On the other hand, at redshifts where kinetic AGN feedback is not activated, we find that the depletion time keeps decreasing, showing increasingly more intense star formation and insufficient feedback to restrain it.

5.4 Mass loading factors

Before discussing the loading factors out of TNG100 galaxies, we should justify our claim that past the onset of kinetic AGN feedback, loading factors become relatively uninformative — at least in the context in which they are defined in this work. In the left panel of Fig. 10 we plot the ratio between η_M , the mass loading factor we define in Eq. (13), and $\eta_{w, \text{TNG}}$, the input TNG100 mass loading factor defined in Eq. (2). Until the onset of kinetic AGN feedback, the ratio between η_M and $\eta_{w, \text{TNG}}$ is close to unity (ranging from a ratio of ~ 0.3 for our lowest mass bin at $z = 10$ to ~ 5 for $M_{\text{vir}} \approx 3 \times 10^{10} M_\odot$ at $z = 0$). This agreement also continues for masses above our suggested mass cut of $3 \times 10^{12} M_\odot$ — but only at redshifts where AGN feedback is not yet active.

On the other hand, once kinetic AGN feedback becomes the dominant feedback mechanism, the ratio between the measured and the TNG input mass loading factors increases dramatically — up to a ratio larger than 1000 for the highest mass bin at $z = 0$. Since $\eta_{w, \text{TNG}}$ governs the rate at which wind particles are launched in TNG, this suggests that past the onset of kinetic AGN feedback, SNe play a negligible role in driving outflows. Notably, $\eta_M / \eta_{w, \text{TNG}}$ remains within $O(1)$ for halos dominated by thermal AGN feedback, suggesting it is sub-dominant in driving outflows. We also note that for all virial masses, the ratio between η_M and $\eta_{w, \text{TNG}}$ decreases with redshift, such that at high redshifts, less material emerges in the wind than has been launched, suggesting that the winds may have stalled, while at low redshifts, mass flow rates are higher than they were when they were launched. The independence of this result on the halo virial mass leads us to conclude that this phenomenon is related to the higher density of the CGM at early times.

The central panel of Fig. 10 shows the measured mass loading factors at the ISM scale, as described in Eq. (13). There are strong trends with both halo mass and redshifts. For $z \geq 4$, mass loading factors are roughly 10 for all mass bins lower than $10^{11} M_\odot$, and decreases monotonically towards higher masses. The mass loading factors increase with decreasing redshifts but keep a relatively similar functional form, as they remain relatively flat for $M_{\text{vir}} \lesssim 3 \times 10^{10} M_\odot$ and then also monotonically decrease with higher masses. At $z = 0$, we reach $\eta_M \approx 100$ at the lower mass end. In fact, for all halo mass bins the loading factors become higher as redshifts decrease, and the difference between $z = 0$ and $z = 10$ is largest at the $10^{10} M_\odot$ mass bin (this statement is true for both η_M and $\eta_M / \eta_{w, \text{TNG}}$). For $z \leq 2$, the mass loading factors decrease with virial mass. Just before the increase at the $\log(M_{\text{vir}}/M_\odot) = 12.1$ mass bin, η_M drops to ~ 6 for $z = 0$ and ~ 4 for $z = 2$. In redshifts where kinetic AGN feedback does not kick in, the mass loading factor keeps decreasing with redshift until reaching $1 - 2$ in the $\log(M_{\text{vir}}/M_\odot) = 12.5$ mass bin. The onset of thermal AGN feedback does not seem to affect the mass loading factors: for example, the dependence of η_M on the halo virial mass at $z = 0$ does not change with the onset of thermal AGN feedback at the $\log(M_{\text{vir}}/M_\odot) = 11.1$ mass bin, while the

onset of kinetic AGN feedback at the $\log(M_{\text{vir}}/M_\odot) = 12.4$ mass bin increases it significantly.

The right panel of Fig. 10 shows the ratio between the mass loading factor at the halo scale and the mass loading factor at the ISM scale (equivalent to $\dot{M}_{\text{out, halo}} / \dot{M}_{\text{out, ISM}}$ for galaxies with $SFR > 0$). For halos with virial masses $< 3 \times 10^{10} M_\odot$, the mass loading factors increase as they travel through the CGM. For higher virial masses the loading factors decrease by up to ~ 0.5 dex, with no apparent dependence on redshift — except for $z = 8$ and $z = 10$ where the loading factors significantly decrease by ~ 1 dex. The drop is strongest for virial masses of $\sim 3 \times 10^{11} M_\odot$, after which the loading factor ratios start increasing until the onset of kinetic AGN feedback. It is notable that the trend shown by the ratio of loading factors (or outflow rates) is the opposite of the trend shown by the ratio of inflow rates, as seen in the left panel of Fig. 8. There is a correlation between a decrease (increase) in inflows from the halo shell to the ISM shell, and an increase (decrease) in outflows from the ISM shell to the halo shell.

6 ENERGY FLOW RATES

In this section we measure energy inflow and outflow rates and energy loading factors for TNG100 halos and galaxies.

6.1 Net inflows and outflows

In Fig. 11 we plot the energy flow rates for all of our selected halos. In as much as the baryons follow the dark matter, we expect the gaseous energy flow rates to be proportional to the dark matter energy flow rates $\sim \frac{1}{2} \dot{M}_{\text{DM}} V_{\text{vir}}^2$, where V_{vir} is the virial velocity of the parent halo. Hence, we normalize the energy flow rates by $M_{\text{vir}} V_{\text{vir}}^2 = G M_{\text{vir}}^2 / R_{\text{vir}}$. As in Fig. 5 for the mass flows, in Fig. 11 we plot the normalized energy inflow and outflow rates at the halo and ISM scales. In Fig. 12 we plot the ratios of the energy inflow to outflow rates.

The normalized energy flow rates are very similar to the normalized mass flow rates shown in Fig. 5, at both the halo and ISM scales. Most notably, both normalized energy flow rates and normalized mass flow rates occupy the same dynamic range, and show a similar dependence on both halo virial mass and redshift.

One difference arises at the ISM scale (right panels of Fig. 11): past the onset of kinetic AGN feedback, both normalized mass *outflows* and normalized energy outflows decrease. While the normalized energy outflow rates continue to decrease with virial mass until reaching the highest mass bin, the normalized mass outflow rates increase again past halo virial masses of $\sim 2 \times 10^{13} M_\odot$ ⁵. While the normalized galactic energy outflow rates decrease, the normalized galactic energy *inflow* rates begin to increase past the onset of kinetic AGN feedback. As we compare the two in Fig. 12, we find that galactic inflow rates surpass the galactic outflow rates only in halos whose virial masses are higher than $\sim 10^{14} M_\odot$.

As a matter of fact, the right panel of Fig. 12 shows that the net energy flows are outward at most masses and redshifts, whereas the net mass flows are inward. There are net energy inflows into TNG100 galaxies only in 3 regimes: in low redshift dwarfs, where we also find very strong mass inflows into the galaxy alongside net outflows from the halo; at the most massive end of the $z = 4$ sample, where no AGN feedback is yet active and net mass inflow rates are strong; and

⁵ This transitional mass bin is where we measure the strongest preventative feedback caused by AGN (as discussed in Sec. 5.2).

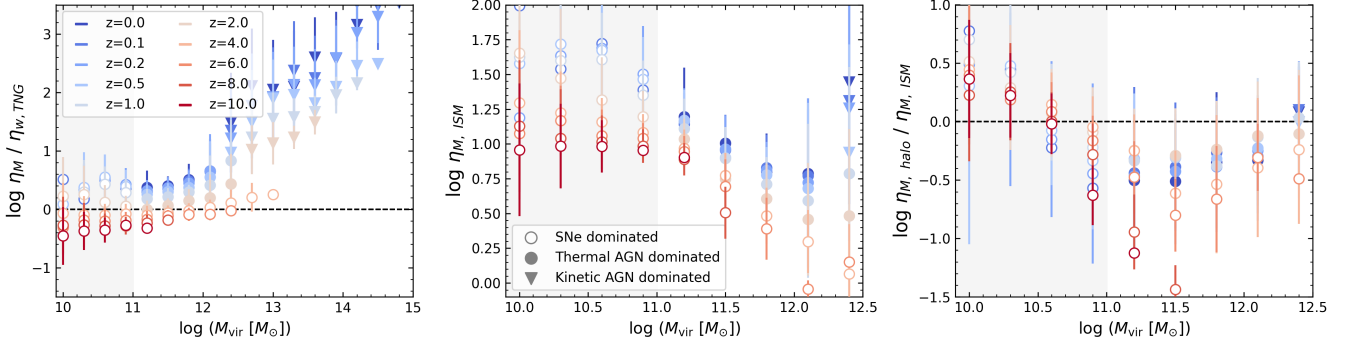


Figure 10. Left panel: the ratio between the mass loading at the ISM scale (estimated as $\dot{M}_{\text{out,ISM}} / SFR$) and the mass loading of injected wind material in the TNG sub-grid recipe described in [Pillepich et al. \(2018\)](#). This ratio is close to unity or slightly below it at high redshift, and increases with decreasing redshift. Central panel: Mass loading factors at the ISM scale for halos where SN are the dominant feedback mechanism. Generally, loading factors at a given halo mass decrease with redshift, to the point where at $z = 10$ they may be 1 dex lower than at the present day. For a given redshift, as virial masses increase and with them the halo gravitational potentials, the loading factors decrease. Right panel: the ratio between the mass loading factors at the halo scale and the mass loading factors at the ISM scale. Mass outflow rates decrease as they travel through the CGM, except for low mass dwarfs and halos at the onset of kinetic AGN feedback.

For all panels, colors and symbol types match those of Fig. 5.

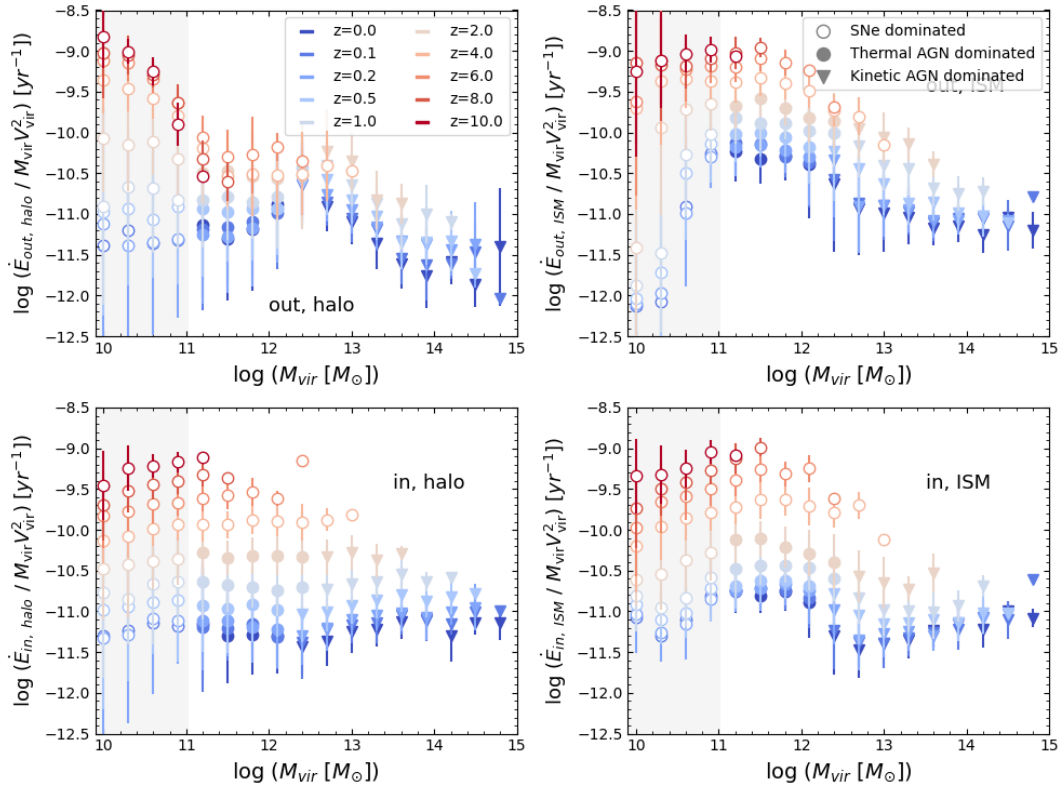


Figure 11. Gas energy outflow rates (top row) and inflow rates (bottom row), estimated at the virial radius (left column) and at the ISM radius (right column), all presented as a function of the halo virial mass and normalized by $M_{\text{vir}} V_{\text{vir}}^2 = GM^2/R$. Colors and symbol types match those of Fig. 5.

at masses above $M_{\text{vir}} \approx 10^{14} M_{\odot}$. For all other masses and redshifts, energy flows outwards from galaxies while they gain mass.

Transitioning to the halo scale (left panels of Fig. 11), we find no significant differences between the normalized energy *inflow* rates and their mass counterparts: for a given redshift, the normalized

inflow rates slightly increase with halo virial mass, except for a dip at the onset of kinetic AGN feedback. There are, however, two notable differences between the normalized energy *outflow* rates and the normalized mass outflow rates. Firstly, high redshift ($z \geq 4$) normalized energy outflow rates increase with decreasing halo virial

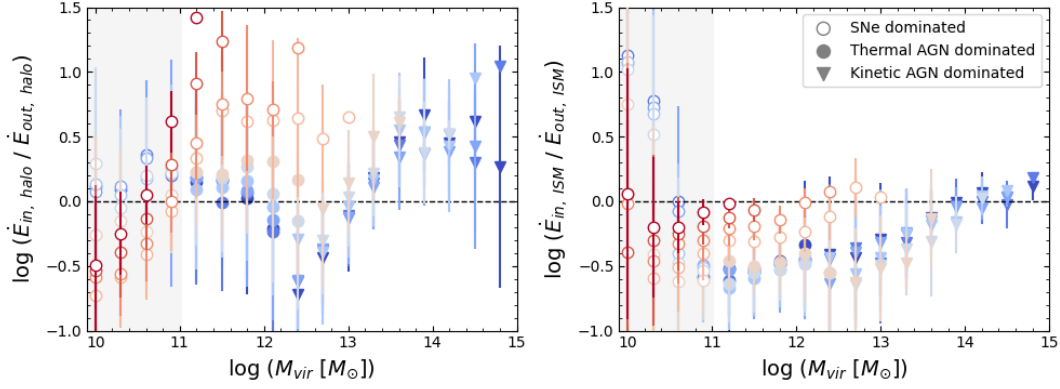


Figure 12. The ratio between the energy inflow and outflow rates at the halo scale (left panel) and the ISM scale (right panel). Colors and symbol types match those of Fig. 5. At the halo scale there is a net outflow of energy only right around the mass scale where kinetic mode AGN feedback becomes active. At the ISM scale, there is a net outflow of energy for almost all virial masses and redshifts.

mass, while the normalized mass outflow rates at the same redshifts flatten for $M_{\text{vir}} < 10^{11} M_{\odot}$. Secondly, the normalized energy outflow rates from halos with virial masses between $10^{12} M_{\odot}$ and $10^{13} M_{\odot}$ at $z \leq 2$ appear to be relatively higher than the normalized mass flow rates at the same mass bin. The normalized energy outflow rates at other masses and redshifts do not differ from their mass counterparts.

The two differences between the mass and energy flow rates at the halo scale also manifest themselves as we observe the ratio between energy inflows and outflows at this scale in Fig. 12. For most redshifts and mass bins, there is a net inflow of energy into the halo. However, at $z \leq 1$ there is a very significant net energy outflow from the halo scale at virial masses between $10^{12} M_{\odot}$ and $10^{13} M_{\odot}$. While all halos that show these energy outflows are dominated by AGN feedback, a significant fraction of those energy outflows occur in halos whose dominant feedback mechanism is thermal AGN feedback. This may either be SNe driven outflows that are also heated by thermal AGN feedback, or kinetic AGN feedback that occurs in some of these halos and is strong enough to raise the median energy outflow rate above the median energy inflow rate (the fraction of kinetic AGN dominated halos at this mass bin is lower than 0.5 — which is why the bin itself is not considered dominated by kinetic AGN feedback — but the fraction is still larger than 0; see right panel of Fig. 4). Another case where we measure a net outflow of energy from the halo scale is for high redshift dwarfs, at $z \geq 4$ and with $M_{\text{vir}} < 10^{11} M_{\odot}$, which becomes more significant at lower masses and earlier times. These outflows must be driven by SNe, which do not lead to mass outflows from the halo (as mass flows inwards) but cause preventative feedback. For all other mass bins and redshifts, energy is flowing into the CGM from both the galaxy and the IGM.

6.2 Energy loading factors

In Fig. 13 we plot the energy loading factors of the central TNG100 galaxies in our sample, calculated using Eq. (14), with markers and colors following our usual convention. The top panel shows the energy loading factors at the ISM shell, and the bottom panel shows the ratio between the energy loading factors at the halo scale and those at the ISM shell (equivalent to plotting $\dot{E}_{\text{out, halo}}/\dot{E}_{\text{out, ISM}}$ for halos with $SFR > 0$).

Before the onset of kinetic AGN feedback, the energy loading factors from the ISM of TNG100 galaxies range between ~ 0.5

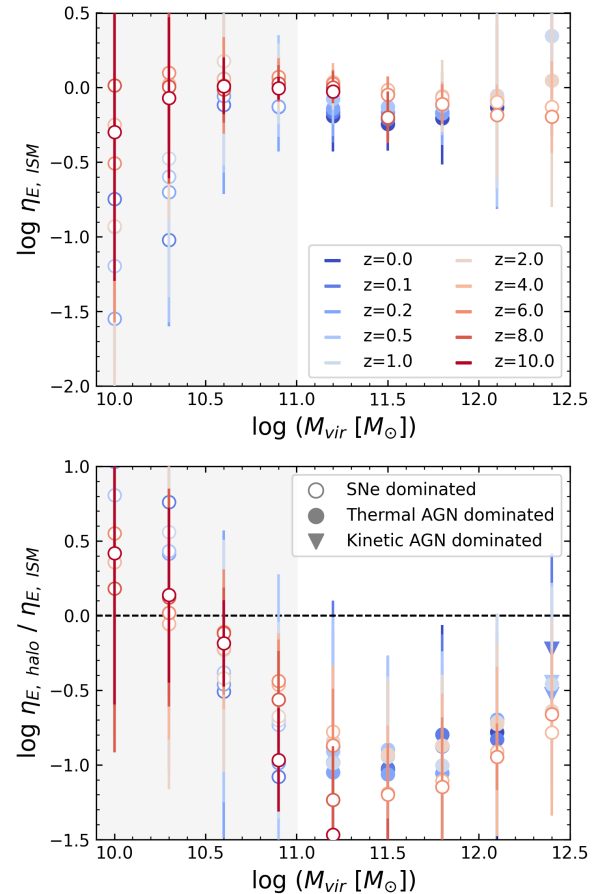


Figure 13. Top: Energy loading factors at the ISM scale, shown for masses and redshift where kinetic AGN feedback is not the dominant feedback mechanism. $\eta_{E, \text{ISM}}$ appears to vary less with the halo virial mass than $\eta_{M, \text{ISM}}$, with values ranging between ~ 0.3 and ~ 1 for most redshifts and halo virial masses. Bottom: Ratio between energy loading factors at the halo scale and energy loading factors at the ISM shell, showing a very similar trend to the ratio between mass loading factors at these scales.

Colors and symbol types match those of Fig. 5.

(for $z = 0$ halos with $M_{\text{vir}} \approx 3 \times 10^{11} M_{\odot}$) and ~ 1.5 (for $M_{\text{vir}} \approx 4 \times 10^{10} M_{\odot}$ halos at $z = 2$). This narrow range of energy loading factors is independent of redshift or halo mass. However, the two lowest mass bins show a pronounced decrease of the energy loading factors at later times. For example, for $z = 0$ η_M increases by ~ 0.9 dex between the second ($2 \times 10^{10} M_{\odot}$) and third ($4 \times 10^{10} M_{\odot}$) mass bin, while for $z = 10$ it only increases by a factor of ~ 0.2 dex. The onset of thermal AGN feedback in $M_{\text{vir}} \approx 10^{11} M_{\odot}$ has no effect on the loading factors.

While the mass loading factors at the ISM shell are very different than the corresponding energy loading factors, the evolution of the two loading factors from the ISM shell to the halo shell is remarkably similar. For dwarfs $\eta_{E, \text{halo}} > \eta_{E, \text{ISM}}$, but as halos masses increases the distance between loading factors quickly narrows and then flips until $\eta_{E, \text{halo}} \approx 0.1 \eta_{E, \text{ISM}}$ for $M_{\text{vir}} \approx 3 \times 10^{11} M_{\odot}$. At higher halo masses the energy loading factor at the halo shell again increases relative to the ISM shell. For all cases there is no redshift dependence.

7 METALLICITY

Lastly, we study the metal contents of the flows. In addition to discussing the metallicity of the gas flows themselves, we examine the relationship between the metallicity of the inflowing or outflowing gas and the reservoir from which it originates (the “metal enrichment factor”), and the ratio between the metallicity of the flow and the estimated metal yield from SNe (metal loading factor), all of which quantify the effectiveness with which gas flows carry and mix metals between the galaxy and the CGM and IGM.

7.1 Metallicity of gas flows

In Fig. 14 we plot the flow metallicities, $Z \equiv \dot{M}_Z / \dot{M}$, where the metal mass flow rates $\dot{M}_Z$ are estimated using Eq. (12), and the total gas mass flow rates $\dot{M}$ are estimated using Eq. (9). We normalize metallicities relative to the solar metal mass fraction $Z_{\odot} = 0.0127$ (adopted from Asplund et al. 2009). The panel ordering, marker types and color coding are as in Fig. 5.

A surprising result, most likely due to resolution issues in low mass TNG100 galaxies, is the dichotomy between halos with virial masses $\lesssim 10^{11} M_{\odot}$ and those above this mass, independent of shell or direction. At the low mass end metallicities increase with redshift (i.e., at $z = 10$ metallicities are *higher* than in $z = 0$), but as we reach one of the mass bins near $10^{11} M_{\odot}$ the variation with redshift decreases and then flips, such that at higher masses, metallicities decrease with redshift. Another expression of the same phenomenon is the dependence of the flow metallicities on the halo virial mass for a given redshift. For $z = 10$, the flow metallicity through a given shell and direction does not vary significantly with the halo virial mass, but as the redshift decreases the flow metallicity increases more sharply with the halo virial mass. This increase continues until the onset of kinetic AGN feedback is reached, after which the mass dependence of the flow metallicity either flattens or drops, depending on the direction of the flow and on the scale in which we measure it. At redshifts where kinetic AGN feedback is not dominant, flow metallicities keep increasing with the halo virial mass.

For most virial masses and redshifts, ISM outflows have the highest metallicities, as they carry the metals formed in the galaxy. Except for dwarfs, ISM outflows at all times have higher metallicities than ISM inflows by roughly 0.2 dex. Outflows at the halo scale are the third highest, as they probably contain metals formed in the galaxy and carried by ejecta, mixed with lower metallicity entrained CGM

gas. Lastly, metallicities of the inflows at the halo scale are the lowest relative to the others, as they are most likely to include only intergalactic, relatively pristine gas (though it is interesting that even at high redshift, gas inflowing into the CGM can be enriched up to more than $0.1 Z_{\odot}$).

The dependence of the gas flow metallicities on the halo virial mass appears similar between all scales and directions, as flow metallicities at both the ISM and halo scales increase with the halo virial mass until peaking at the onset of kinetic AGN feedback. The exception is the metallicity of the inflows at the halo scale, which remain relatively constant past the onset of kinetic AGN feedback. The metallicities of inflows and outflows at the ISM scale and of outflows at the halo scale all decrease (slightly) past the onset of kinetic AGN feedback. This is true for the reservoirs themselves as well, i.e. the ISM and CGM metallicities. The former result can also be seen in Torrey et al. (2019), which observed the drop of the ISM metallicities in TNG100 galaxies more massive than $\sim 10^{10.5} M_{\odot}$, equivalent to the virial mass where we observe the onset of kinetic AGN feedback (see e.g. Fig. 7 of Oren et al. 2024).

7.2 Metal enrichment factors

In addition to describing the metallicity of the flow, we wish to determine the relationship between the gas flow and the environment from which it originated. We therefore define a “metal enrichment factor”:

$$\zeta_{\text{flow}} \equiv \frac{\dot{M}_Z}{\dot{M} \cdot Z_{\text{reservoir}}} = \frac{Z_{\text{flow}}}{Z_{\text{reservoir}}} \quad , \quad (16)$$

where $\dot{M}_Z$ is given by Eq. (12) and $\dot{M}$ is given by Eq. (9), such that Z_{flow} is the parameter we plot in Fig. 14. $Z_{\text{reservoir}}$ describes the average metallicity of the gas reservoir from which the outflows/inflows originate: the ISM metallicity for outflows at the ISM scale, and the CGM metallicity for outflows at the halo scale and for inflows at the ISM scale. Because the IGM metallicity is outside the scope of this work, we do not discuss $\zeta_{\text{in, halo}}$. In most SAMs and gas regulator models, the metallicities of the reservoir and the flow are assumed to be identical, i.e. $\zeta = 1$ (e.g. Finlator & Davé 2008), while in others it is not (e.g. Peebles & Shankar 2011). As we mentioned in Sec. 3.1, the metal enrichment factor for SNe-driven outflows in TNG when they are launched is a constant 0.4^6 .

We start with the outflows at the ISM scale in the central panel of Fig. 15, which shows a dichotomy between two dominant feedback mechanisms. In virial masses where outflows are driven by SNe, $M_{\text{vir}} \lesssim 3 \times 10^{12} M_{\odot}$ (which includes halos dominated by thermal AGN feedback), $\zeta_{\text{out, ISM}}$ ranges from 0.4 at $z = 0$ to 0.7 at $z = 0$, where the TNG metal enrichment factor at injection is 0.4. At high redshifts, where CGM densities are also high, the dense surrounding gas may stall outflows and there is very little mixing of outflows with the surrounding gas, causing little to no change to the metal enrichment at injection. At low redshifts, however, we find evidence of entrainment (e.g. Sec. 5.4) which adds high-metallicity ISM to the winds and contributes to a higher enrichment factor. In virial masses where outflows are driven by AGN, on the other hand, we find that $\zeta_{\text{out, ISM}} = 1$ since kinetic AGN feedback simply gives additional momentum to ISM particles and does not alter their metallicity.

⁶ In the works that describe the Illustris and the IllustrisTNG methods, such as Vogelsberger et al. (2013) and Pillepich et al. (2018), what we call the “metal enrichment factor” is referred to as a “metal loading factor”.

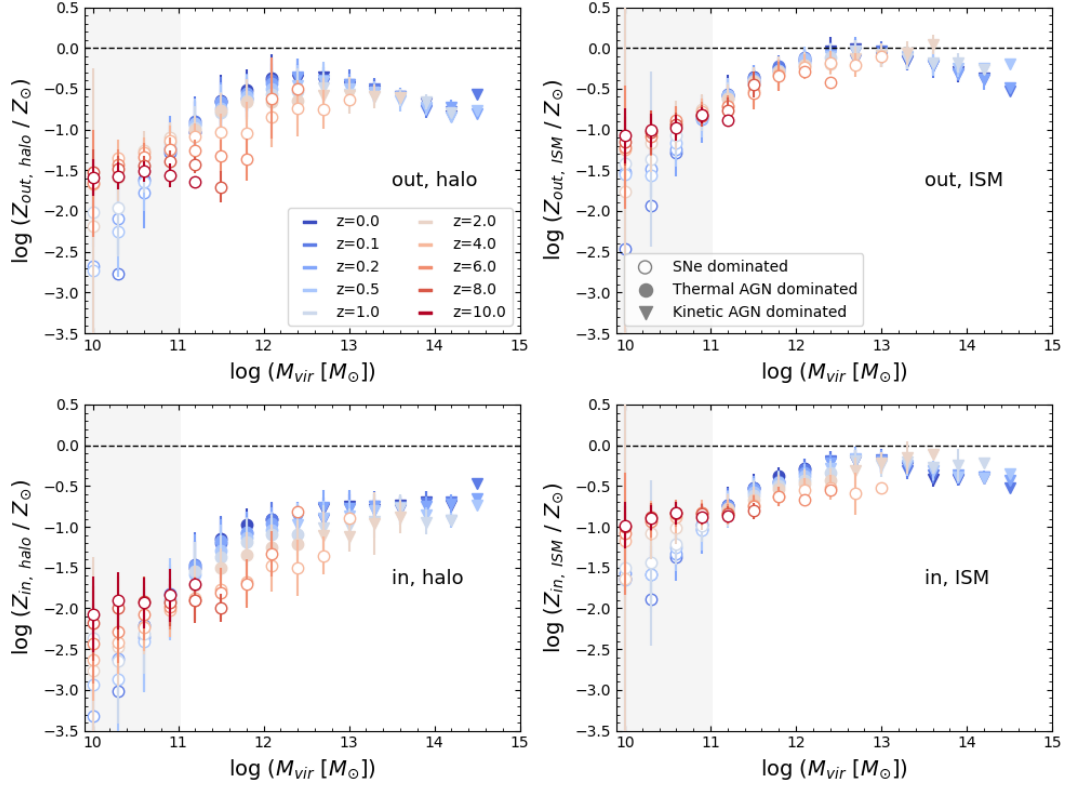


Figure 14. Metallicity of outflows (top row) and inflows (bottom row). The left column shows the halo scale, and the right column shows the ISM scale. Colors and symbol types match those of Fig. 5. The metallicity of flow rates increases with halo virial mass until the onset of kinetic AGN feedback, after which it either stabilizes with virial mass (galactic inflows, lower left panel) or decreases slightly (all other panels).

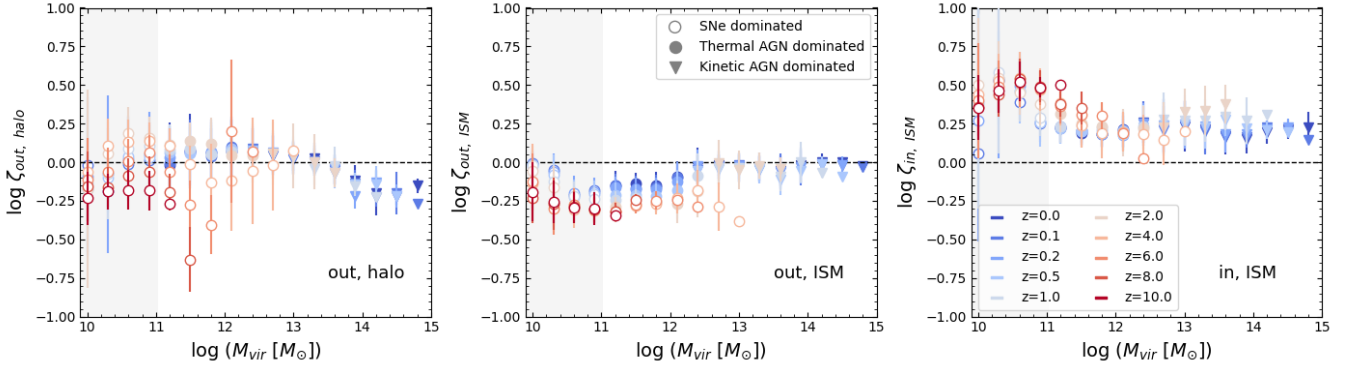


Figure 15. Metal enrichment, as defined in Eq. (16). Colors and symbol types match those of Fig. 5. Left panel: outflows at the halo scale. Generally, outflow metallicity matches the CGM metallicity, except for low- z , high-mass halos where outflows have $\sim 50\%$ fewer metals than the CGM average. Central panel: outflows at the ISM scale. SNe driven outflows appear to have a constant $\zeta = 0.6$ (slightly higher than the TNG metal loading at injection of 0.4), while AGN driven outflows exactly match the metallicity of the ISM. Right panel: inflows into the ISM. The metallicity of the inflows appears to be 1.4 to 3 times higher than the average CGM metallicity, where the highest ratio is at the low-mass end.

The metal enrichment factor for ISM inflows, shown in the right panel of Fig. 15, suggests that ISM inflows have higher metallicities than the average metallicity of the CGM. For halo virial masses $\gtrsim 6 \times 10^{11} M_{\odot}$, there is an almost constant $\zeta_{\text{in, ISM}} \approx 1.5$ for all redshifts. As virial masses decrease, the metal enrichment factor increases slightly until reaching a peak of 3 for $M_{\text{vir}} \approx 3 \times 10^{10}$, and drops back for lower masses. To understand how inflows have

a seemingly higher metallicity than the medium through which they travel, we first consider that the metallicity of the CGM decreases with galactocentric radius (see e.g. Garcia et al. 2023). In Oren et al. (2024) we show that the density of the CGM at $z = 0$ can be very well approximated by a power law of the radius for halos with virial masses higher than $10^{11} M_{\odot}$. We can also fit the metallicity of the CGM to a power law of the radius, with a different power law index than that

of the density. If we use our power laws to estimate the ratio between $Z(0.1 R_{\text{vir}})$ and the average CGM metallicity between $0.1 - 1 R_{\text{vir}}$ ($\langle Z_{\text{CGM}} \rangle$), we find it to be $[5.2, 3.4, 1.7, 2.1]$ for $\log(M_{\text{vir}}/M_{\odot}) = [11, 12, 13, 14]$ — all very close to the numerically calculated $\zeta_{\text{in, ISM}}$ for $z = 0$, supporting our conclusion that the metallicity of ISM inflows matches the metallicity of the CGM at the border between the two⁷.

Last are the metal enrichment of outflows at the halo scale, shown at the left panel of Fig. 15. $\zeta_{\text{out, halo}}$ suggests that the metallicity of outflows from the halo is within a factor of 2 from the average metallicity of the CGM from which they emerge, barring two outliers (one at $z = 6$ and the other at $z = 8$). In low redshifts ($z \leq 2$), $\zeta_{\text{out, halo}}$ appears to start at ~ 0.75 , then increase with the halo virial mass to ~ 1.2 , decrease towards the high-mass end and flatten at ~ 0.6 for $M_{\text{vir}} \gtrsim 10^{14} M_{\odot}$. For $z \leq 1$ the decrease in the metal enrichment starts after the onset of kinetic AGN feedback, and for $z = 2$, $\zeta_{\text{out, halo}}$ starts to decrease at a lower mass, where outflows are dominated by SNe. At high redshifts, however, there is no apparent trend with halo mass and $\zeta_{\text{out, halo}}$ is either constant for all halo virial masses (as is the case for $z = 10$), or fluctuates around 1.

At the highest mass end there are net inflow rates of both mass (Fig. 6) and energy (Fig. 12) plus weakening preventative feedback (Fig. 8), suggesting that much of the outflowing material mixes with the CGM and does not make it out of the halo. We would therefore expect that for halos more massive than $\sim 10^{14} M_{\odot}$ the metallicity of outflows at the halo scale will be similar to that of the CGM at the halo scale, and $\zeta_{\text{out, halo}} = Z(R_{\text{vir}})/\langle Z_{\text{CGM}} \rangle$. Using the same power law distributions described above, we get $\zeta_{\text{out, halo}} = 0.67$, in agreement with the simulation results. While we cannot account for the metal enrichment factors at masses lower than those of galaxy clusters, we still find them to be relatively close to unity. Comparatively, the metallicity of the flows themselves can change by roughly 2 dex in a given redshift. We therefore conclude that the metallicity of halo outflows tracks the metallicity of the CGM to a good approximation.

7.3 Metal loading

We finally describe the efficiency with which metals are ejected from the galaxy relative to a reference yield times the SFR. We observe in Fig. 16 the metal loading factors at the ISM scale, estimated using Eq. (15). Colors and markers follow our standard conventions.

For a given redshift, the metal loading factors of SNe driven outflows do not vary significantly with the halo virial mass above $\sim 3 \times 10^{10} M_{\odot}$. The metal loading factors out of galaxies in the centers of halos with virial masses smaller than $\sim 3 \times 10^{10}$ decrease with halo mass, with a slope that decreases as redshift increases, such that at the low mass end, η_Z for $z = 10$ is comparable to η_Z for $z = 0$. For virial masses higher than the stated cutoff mass, the metal loading factor only increases with decreasing redshift.

The metal loading factor is larger than unity in the region where it flattens with halo virial mass, and for $z \leq 4$. It reaches ~ 3 for $z = 0$, while for $z = 10$ it is ~ 0.7 . This would not necessarily suggest that at late times outflows carry more metals than could have been produced by stars, as $\zeta_{\text{out, ISM}} \leq 1$ for all masses and redshifts, and is entirely calculated using simulation results (whereas for η_Z we assume a

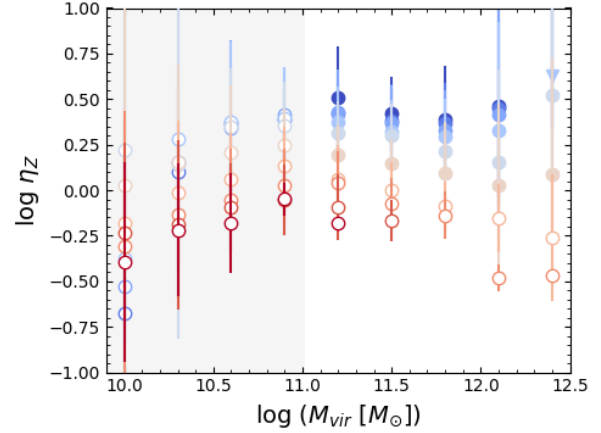


Figure 16. The metal loading factor based on Equation 15, estimated at the ISM scale. Colors and symbol types match those of Fig. 5. For a given redshift, and for $\log(M_{\text{vir}}/M_{\odot}) > 10.5$, η_Z is roughly constant with virial mass. For a given virial mass, the metal loading factor decreases with increasing redshift.

constant metal yield independent of halo mass or redshift). However, this indicates that for most halo masses, outflows carry more metals out of the galaxy (per unit star formation rate) as time progresses. At the low mass end η_Z suggests that outflows carry more metals out of galaxies at earlier times.

8 DISCUSSION

8.1 Baryon flows at different halo mass scales

In our study we find distinct flow behaviors in halos of different total mass, due to the interplay between various feedback mechanisms and the gravitational potential. Within the range of halo masses that is well-resolved and are represented in sufficient numbers in the TNG100 volume, we can identify three main regimes: low mass halos ($10^{11} \lesssim M_{\text{vir}} \lesssim 10^{12.25} M_{\odot}$), intermediate mass halos ($10^{12.25} \lesssim M_{\text{vir}} \lesssim 10^{13.5} M_{\odot}$), and massive halos ($M_{\text{vir}} \gtrsim 10^{14} M_{\odot}$).

8.1.1 Low mass halos

Low mass halos ($10^{11} \lesssim M_{\text{vir}} \lesssim 10^{12.25} M_{\odot}$) are inhabited mostly by galaxies that are star forming at all redshifts (Fig. 2), with outflows that are dominated by stellar feedback. Although most of these galaxies contain black holes, they are mostly in the thermal AGN feedback regime (see Fig. 4). The gas depletion times are less than the Hubble time at all redshifts, leading to efficient star formation (Fig. 9). Low mass halos show *net inflows* of both mass and energy on halo scales at all redshifts, while at the ISM scale they gain mass but simultaneously lose energy. The net energy loss is likely due to cooling of the inflows before they reach the ISM scale. Baryon fractions are strongly sub-universal (Fig. 7), show a weak trend with halo mass over this interval, and have decreased by almost a factor of two over cosmic time since $z \sim 10$. At halo scales, low mass halos do not show signs of preventative feedback at high redshift, and show only weak preventative feedback at $z = 0$ (Fig. 8). On ISM scales, low mass halos show signatures of “galactic fountains”, as inflows into the ISM shells surpass those into the halo shells, suggesting previously ejected material is able to fall back onto the galaxy in addition to fresh gas arriving from the intergalactic medium. This phenomenon

⁷ For $z = 0$, the power law indices fitted to the median TNG halo in each mass bin are: density: $a_n = [1.6, 1.7, 1.2, 1.9]$; metallicity: $a_Z = [1.1, 0.8, 0.3, 0.5]$. The ratio is given by:

$$\frac{Z(0.1 R_{\text{vir}})}{\langle Z_{\text{CGM}} \rangle} = 0.1^{-a_Z} (3 - a_n - a_Z) (1 - 0.1^{3-a_n}) / (3 - a_n) (1 - 0.1^{3-a_n-a_Z}).$$

is considered a crucial source of mass inflow into star forming galaxies, as observed in our own Milky Way (Bregman 1980) and in nearby galactic halos (e.g. Fraternali & Binney 2006, 2008; Tremblay et al. 2018). Fig. 8 suggests that for $z \geq 4$ there is an extra source of inflow at the halo scale, which we interpret as a “halo fountain”.

A very significant factor in the occurrence of a galactic fountain is the ability of the ejecta to cool in order to fall back onto the galaxy. The virial temperature associated with the halos in this mass range, which is also the typical temperature of volume filling phase of the CGM, is in the region of $\sim 10^5$ K (see e.g. Oren et al. 2024). This temperature is associated with the peak of the cooling curve (e.g. Gnat & Sternberg 2007; Wiersma et al. 2009), and therefore CGM gas at this temperature can cool very effectively. Additionally, this may explain why we do not witness a galactic fountain in higher mass halos — higher virial temperatures lead to slower cooling, such that outflows are either ejected entirely from the halo or simply mix with the thermally supported hot CGM.

8.1.2 Intermediate mass halos

We define intermediate mass halos as those with halo virial masses between $10^{12.25} M_\odot - 10^{13.5} M_\odot$, where the onset of both kinetic BH feedback (Fig. 4) and galaxy quenching (Fig. 2) occurs. Before the onset of kinetic AGN feedback, we measure net positive inflow rates on halo scales in both mass and energy (Fig. 12). The ISM depletion times are very short, as low as $\sim 10\%$ of the Hubble time, leading to rapid formation of stars and build-up of metals.

After the onset of kinetic AGN feedback, at $z \leq 2$, the picture changes dramatically. The mass and energy flows become net positive outward at both the ISM and halo scales. The baryon fractions of both the CGM and the entire halo decrease, supporting a picture of AGN driving outflows beyond the virial radius. Star formation rates drop, and the ISM depletion times increase. As this occurs, metallicities begin to drop as fewer new stars are formed and outflows carry the existing metals out of the galaxy. Torrey et al. 2019 also attribute the drop in metallicity of the ISM to the “increased impact of AGN feedback”, and Peebles & Shankar (2011) invoke outflows for their model to reproduce the observed drop in the mass-metallicity relation past MW-like galaxies. The metal enrichment factor of ISM outflows transitions from ~ 0.4 , which is the metal enrichment factor of SNe outflows at injection, to ~ 1 , indicating the outflows now predominantly contain ISM particles kicked by kinetic AGN feedback and not wind particles driven by SNe.

We see the strongest preventative feedback at halo scales in this halo mass range, with the gas inflow rate suppressed relative to $f_B \dot{M}_{\text{in}}^{\text{DM}}$ by about 0.25 dex. The gas inflow rate in ISM scales drops dramatically relative to the DM halo growth rate (Fig. 8). The ISM gas inflow rate $\dot{M}_{\text{in,ISM}}$ also begins to drop below the halo scale gas inflow rate $\dot{M}_{\text{in,halo}}$, indicating the build-up of a hot CGM.

8.1.3 High mass halos

High mass halos include all halos with virial masses higher than $10^{13.5} M_\odot$. These halos are all dominated by kinetic AGN feedback, and at low redshift they are mostly occupied by quenched galaxies. Note that there are very few halos in this mass range at high redshift in the TNG100 volume (Fig. 1), so most of the trends discussed below are measured for redshifts $z \lesssim 1$. There is a strong net inflow of mass into high mass halos on halo scales, while on ISM scales the mass inflow and outflow rates are close to equal. The baryon fractions increase strongly with increasing halo mass, from about half to eighty

percent of the universal value by $M_{\text{vir}} \simeq 10^{14} - 10^{15} M_\odot$. Halo scale preventative feedback is still present, but it is somewhat weaker than in intermediate mass halos, and continues to weaken with increasing halo mass. ISM scale preventative feedback is strongest for halos with mass $3 \times 10^{13} M_\odot$, and from there it weakens as halos and galaxies become more massive. At the highest mass bin we find that $\dot{M}_{\text{in,halo}} \simeq \dot{M}_{\text{in,ISM}} \simeq f_B \dot{M}_{\text{in,DM}}$, suggesting that AGN feedback is beginning to fail to be able to regulate accretion. High mass halos show a net inflow of energy at the halo scale, while at the ISM scale there is net energy loss from $M_{\text{vir}} \sim 10^{13} - 10^{14} M_\odot$, while $\dot{E}_{\text{in}} \simeq \dot{E}_{\text{out}}$ above $10^{14} M_\odot$, with a trend towards net energy inflows towards the highest halo masses. Outflows in these high mass halos are also depleted of metals relative to the overall average metallicity of the CGM (left panel of Fig. 15).

8.2 Interactions between outflows and the surrounding gas — entrainment, mixing, and stalling

The way in which injected momentum or energy couples with the medium through which winds propagate has a large impact on the emergent properties of the outflows. Comparing the mass loading factor, η_M , at the ISM shell to the mass loading factor at injection, η_w , TNG, based on the TNG subgrid recipe (left panel of Fig. 10) shows that the ratio between the two increases with redshift, such that for a given halo mass, the emergent mass flow rates are weaker relative to the amount of injected wind material at earlier times. This result is independent of the halo virial mass, leading us to suggest that this effect is due to higher gas densities in the early universe, where the gas density of the CGM is proportional to $\rho_c \propto (1+z)^3$. The higher densities at high redshift could cause winds to stall, lowering η_M . At low redshift, the winds may entrain additional material, increasing η_M . Evidence of significant amounts of CGM gas entrained by galactic outflows at low z were found in the EAGLE simulation, in agreement with our interpretation, where Mitchell et al. (2020a) show that at $0.1 R_{\text{vir}}$, 50% of outflowing particles are entrained CGM gas.

Additional evidence suggesting that winds stall at higher redshifts (leading to deceleration of outflows) and entrain CGM material at lower redshifts comes from our analysis of the metal enrichment factors of SNe driven outflows at the ISM shell. The central panel of Fig. 15 shows that the metal enrichment factor increases from 0.4 at $z = 10$ — the metal enrichment factor at injection, indicating no mixing with the surrounding CGM — to 0.7 at $z = 0$, suggesting SNe outflows carry surrounding material with them. This enrichment value is the average between 0.4 (ζ of wind particles only) and 1 (ζ of ISM particles only), in agreement with the EAGLE result that 50% of the outflows at the ISM scale consists of entrained gas.

8.3 Comparison with previous works

In this section, we briefly discuss how our results compare and contrast with select results from the literature, including gas regular models, semi-analytic models, and numerical hydrodynamic simulations.

8.3.1 Gas regulator models

Gas regulator models track flow quantities between different reservoirs (such as CGM, ISM, and stars), and generally include the growth of the main progenitor of a dark matter halo based on N-body simulations or fitting functions derived from them. They are similar

to semi-analytic models, but in general are more empirical and do not include the full dark matter merger history (including non-main branches) or the physical processes associated with galaxy mergers.

Most gas regulator models in the literature track only the flows of mass and metals, and do not track energy flows. Carr et al. (2023) was the first to include energy flows in a gas regulator model, and showed that energy flows can play a central role in regulating star formation via regulating cooling of the CGM. In their fiducial model, which was adjusted to approximately reproduce the stellar-mass to halo-mass (SMHM) relation from Behroozi et al. (2019), they adopted mass and energy loading factors that are power law functions of the halo mass. The values of the mass loading range from $\eta_M = 10$ for $10^{10} M_\odot$ halos to $\eta_M = 1$ for the $10^{12} M_\odot$ halos, and the energy loading from $\eta_E \sim 1$ for the low mass end to ~ 0.1 for the high mass end. The adopted mass loading factors in Carr et al. (2023) are much lower than the values that we measure in TNG, and the dependence of energy loading on halo mass is also very different (η_E has only a weak dependence on halo mass in TNG).

Carr et al. (2023) also implements a halo-scale preventative feedback parameter, f_{prevent} , used to moderate baryonic inflows into the halo relative to DM inflows. As the energy flows in Carr et al. (2023) are purely thermal, f_{prevent} is also used to regulated energy inflows into the halo relative to energy outflows. The f_{prevent} obtained from the calibration to the Behroozi et al. (2019) SMHM relation ranges from 0.3 at the low mass end to 0.8 at the high mass end, so while it is by definition capped at 1 this limit is not reached. In comparison, the left panel of Fig. 8 suggests a somewhat different picture in TNG100. $\dot{M}_{\text{in, halo}}^{\text{gas}} / f_B \dot{M}_{\text{in, halo}}^{\text{DM}}$ is close to unity or even slightly above it (indicating no halo scale preventative feedback or halo fountains) at $z \geq 2$, and at $z = 0$ has a nearly constant value of ≈ 0.7 at the lowest well-resolved halo up to about $M_{\text{vir}} \sim 10^{12} M_\odot$, where it begins to decline due to the onset of kinetic AGN feedback.

Another gas regulator model we wish to discuss here was published in a set of companion papers, Voit et al. (2024a) and Voit et al. (2024b). Unlike previous regulator models, the Voit model tracks *all* mass and energy associated with a halo's baryons, even those that may have been pushed beyond the halo's virial radius by feedback, and the energy budget manifestly includes the halo's gravitational potential energy. A key quantity in regulating gas accretion and hence star formation is the specific energy of the outflow $\varepsilon_{\text{fb}} \equiv \eta_E \varepsilon_{\text{SN}} / \eta_M$, where ε_{SN} is a constant specific energy attributed to SNe feedback⁸. They then distinguish between coupled and uncoupled outflows. A coupled outflow can transfer energy to the CGM causing preventative feedback, while uncoupled outflows cannot and therefore cause only ejective feedback. The key idea of the Voit model is that for coupled outflows, the ratio between ε_{fb} and the specific energy of the accreting gas ε_{acc} determines the future of the galaxy. Namely, if $\varepsilon_{\text{fb}} > \varepsilon_{\text{acc}}$ then the halo will expand (leading to reduced cooling and star formation), and if $\varepsilon_{\text{fb}} < \varepsilon_{\text{acc}}$ it will contract (leading to enhanced cooling). If ε_{fb} drops below ε_{acc} , a contraction crisis occurs where the halo keeps contracting and runaway star formation may occur. To prevent this, the halo will require a new source of feedback to counter the collapse — such as AGN feedback.

Voit et al. (2024b) discuss the interpretation of IllustrisTNG in the context of their model (see especially their Fig. 3) and suggest that for $z \gtrsim 4$, $\varepsilon_{\text{fb}} < \varepsilon_{\text{acc}}$ and so halos should be in the contracting regime and therefore strongly cooling and star forming. This is qualitatively

consistent with our finding that in the mass range considered in Voit et al. (2024b) Fig. 3 ($5 \times 10^{11} - 7 \times 10^{12} M_\odot$), there are strong net inflows into halos at $z \gtrsim 4$, and these weaken over time as $\varepsilon_{\text{fb}} / V_{\text{vir}}^2$ increases (due to the assumed scaling in the IllustrisTNG stellar wind sub-grid model), pushing halos in the more strongly 'expanding' regime. However, we note that Voit et al. (2024b) assume that the emergent wind energies are the same as the wind launch properties, while we have shown that emergent wind properties can be quite different due to entrainment and stalling, and these differences can change with the conditions in the CGM (and thus appear different at different cosmic times).

8.3.2 Semi Analytic Models

Traditional semi-analytic models solve ordinary differential equations describing flows of gas mass and metals between the IGM, CGM, and ISM, and the rate at which ISM gas is converted into stars (see Somerville & Davé 2015, for a review). As noted above, SAMs typically include more physically motivated model components than gas regulator models, and generally account for the full dark matter halo formation history, thus including the physical processes associated with galaxy mergers.

Mitchell et al. (2020a) show a useful comparison of the mass loading factors as a function of halo mass for several semi-analytic models from the literature in their Fig. 13, including GALFORM (Mitchell et al. 2018), L-galaxies (Henriques et al. 2015), the Santa Cruz SAM (Somerville et al. 2015), GAEA (Hirschmann et al. 2016), and SHARK (Lagos et al. 2018). Some SAMs do not allow flows directly from the CGM to the IGM (i.e. halo outflows); for those that do include halo outflows, both the halo scale and galaxy scale mass loading is shown. It is also important to note that most traditional SAMs do not include halo-scale preventative feedback arising from internal stellar or AGN feedback, i.e. it is generally assumed that the mass inflow rate into the halo is $f_B \dot{M}_{\text{in, halo}}^{\text{DM}}$ ⁹. Although all of these SAMs are calibrated to reproduce the $z \sim 0$ stellar mass function and stellar mass vs. halo mass relationship, this figure shows that there is an extremely wide range of slopes and normalizations for the mass loading of stellar driven winds in different SAMs. We notice that SAMs that do not include halo outflows (e.g. GALFORM and the Santa Cruz SAM), in which star formation must be regulated solely by ejecting gas from the ISM, unsurprisingly must invoke very large mass loadings, which increase steeply towards lower halo masses. These mass loadings ($\eta \sim 10\text{--}100$ at $M_{\text{vir}} \sim 10^{11} M_\odot$) are a bit higher than the ones we measure in TNG100 over the same halo mass range. The models that include halo-scale outflows (L-galaxies and SHARK) adopt somewhat lower ISM scale mass loadings. The GAEA SAM was one of the first SAMs to include halo-scale preventative feedback, and it adopts the lowest mass loadings — even lower than either our TNG100 measurements or those from EAGLE (presented in Fig. 13 of Mitchell et al. 2020a).

Pandya et al. (2020) presented a detailed comparison of the halo and galaxy scale inflow and outflow rates, along with the ISM, stellar, and CGM masses, between the Santa Cruz SAM and the FIRE-2 simulations (Hopkins et al. 2018). They showed that the mass inflow and outflow rates in the Santa Cruz SAM were much larger than those predicted by FIRE, especially in low-mass halos. This lead to the strong conclusion that SAMs may need to adopt more

⁸ Estimating ε_{fb} for SNe driven outflows in TNG100 using Eqs. 4, 13, and 14 gives a similar relationship, with e_w taking the place of ε_{SN} . Note that e_w depends on the metallicity while ε_{SN} is a constant.

⁹ Many SAMs do include a form of halo-scale preventative feedback arising from heating by the metagalactic UV background after reionization, but this tends to impact only very low mass halos $M_{\text{vir}} \lesssim 10^{10} M_\odot$.

flexible assumptions about galaxy and halo scale inflows and outflows (including the option of halo-scale preventative feedback and halo-scale outflows) in order to faithfully reproduce the baryon cycle in numerical simulations.

This discussion highlights that there are many different baryon cycle configurations that can lead to the same stellar mass vs. halo mass relation in galaxies. Furthermore, several different aspects of the baryon cycle beyond the wind mass loading out of the galaxy are important: also important are the mass outflow rate on the halo scale, the recycling time for mass that has been ejected from the halo, and preventative feedback on both the halo and galaxy scale. As discussed above, energy flows (neglected in traditional SAMs) can also play a crucial role in the baryon cycle. Metal flows (generally assumed to trace mass flows in a simple way in traditional SAMs) can also play a role by affecting cooling rates.

Pandya et al. (2023) presented the first SAM that was designed to reproduce the detailed baryon flow cycle measured from a high-resolution hydrodynamic simulation. Instead of adopting relatively arbitrary choices for many aspects of the baryon cycle, and then calibrating the free parameters to match a limited set of quasi-observables at $z = 0$ as in nearly all previous SAMs, Pandya et al. (2023) used measurements of the mass, metal, and energy inflow and outflow rates on halo (CGM) and ISM scales from the FIRE-2 zoom-in simulations to calibrate their SAM. The model also accounted separately for heating and turbulence driven by supernova winds and cosmic accretion (i.e. thermal and kinetic energy). As in the simpler regulator models of Carr et al. (2023) and Voit et al. (2024a), they found that the specific energy of the supernova driven winds plays a critical role in regulating star formation by over-pressurizing the CGM, leading to halo outflows and halo-scale preventative feedback, and by slowing down cooling from the CGM to the ISM (galaxy scale preventative feedback). Turbulence can also play a significant role in preventing cooling at early times. Importantly, Pandya et al. (2023) showed that, when designed and calibrated in this way, the simplified framework of a SAM could accurately reproduce the main state variables (e.g. stellar mass, ISM mass, CGM mass) and their derivatives (inflow and outflow rates at the galaxy and halo scale) from a numerical hydrodynamic simulation, motivating this work and that of Omoruyi et al. (in prep).

8.3.3 Zoom-in Simulation: FIRE-2

We start our comparison with other simulations with FIRE-2, analyzed by Pandya et al. (2021). FIRE-2 is a zoom-in simulation, in which a region inside a low resolution cosmological simulation is selected and evolved again in a much higher resolution, while boundary conditions are kept from the original simulation¹⁰. The implementation of SNe feedback in FIRE-2 is very different than that of TNG100: in FIRE-2, the probability of a SN explosion is estimated per star-forming particle and time step, according to which a feedback event may then occur. The SN event deposits energy, mass, and metals directly into the surrounding gas particles, and the fraction of thermal to kinetic energy is calculated per neighboring gas particle as well. Additionally, FIRE-2 does not include supermassive black holes nor AGN feedback (but see Anglés-Alcázar et al. 2017b; Hopkins et al. 2023; Wellons et al. 2023). Due to the latter point we will limit our comparison to TNG100 halos in the $10^{10} - 10^{12} M_{\odot}$ region.

Pandya et al. (2021) computed similar mass, energy, and metal loading factors to those we estimate in this work. Unlike this work, which defines outflowing particles as having $v_r > 0$, Pandya et al. (2021) define outflows as particles with $v_B > v_{\text{esc}}$, where v_{esc} is the escape velocity at the ISM shell. Under this definition fewer particles are considered as outflows, which effectively lowers η_M and η_Z by a factor of 2 relative to what they were using our definition¹¹. Secondly, the loading factors are presented as a function of either stellar mass or virial velocity rather than of halo virial mass. We convert the stellar mass to a halo mass using the median ratio presented in Hopkins et al. (2018). In FIRE-2, the equivalent stellar mass range to our analysis is $10^7 - 10^{10.7} M_{\odot}$.

The mass loading factor out of FIRE-2 galaxies reaches ~ 20 at the low mass end (~ 5 times lower than for TNG100), whereas for the high mass end it drops below the TNG100 results and is closer to ~ 0.1 at $z = 0$ (for TNG100, it is ~ 10). Interestingly, the FIRE-2 mass loading factor for a given stellar mass does not appear to vary significantly with redshift, while in TNG100 we find a clear trend of increasing η_M as the redshift drops. If we reorganize the loading factors as a function of the virial velocity rather than of the mass we find that for a given V_{vir} bin the redshift dependence of TNG100 mass loading factors significantly decreases. In FIRE-2, however, η_M increases with redshift for a given V_{vir} bin, but the larger V_{vir} is, the larger the dynamic range covered by η_M . This is because for a given mass bin V_{vir} increases with redshift¹², and thus loading factors from a given mass bin will “skew to the right” as higher virial velocities are attributed to the same mass at higher redshifts.

Energy loading factors in FIRE-2 vary in their dependence on galaxy mass in different redshift bins, but qualitatively they decrease with galactic mass and increase with redshift. For a given mass bin, η_E in FIRE-2 may increase by roughly 1 dex from $z = 0$ to $z = 4$, and more specifically range from ~ 0.1 to ~ 1 at the low mass end, to ~ 0.02 to ~ 0.2 at the high mass end. As a function of the virial velocity, however, the FIRE-2 energy loading factors appear to be constant for $V_{\text{vir}} \lesssim 100$ km/s in both the intermediate and high redshift bins, after which they decrease, while the low redshift η_M appears to increase with V_{vir} . Comparatively, the TNG100 energy loading factors are roughly 0.5 for most masses and redshifts (except for low masses, where they also increase with redshift), and therefore behave similarly as a function of the virial velocity.

The metal loading factors from the FIRE-2 simulation are quite different compared to those in TNG100. In FIRE-2, the metal loading factors remain at a relatively constant ~ 1 until reaching $M_{*} \approx 10^9 M_{\odot}$, after which they decrease until reaching ~ 0.05 for $M_{*} \approx 10^{11} M_{\odot}$. As with the mass loading factors, there is no significant dependence on redshift for a given mass bin. Since η_M is relatively constant for a wide range of stellar masses, as we reorganize it as a function of the virial velocity we find it to have a similar behavior — constant for $V_{\text{vir}} \lesssim 100$ km/s and decreasing afterwards, with no significant variation with redshift. The metal loading factor of TNG100 galaxies, however, remains constant with halo mass (for $M_{\text{vir}} \gtrsim 3 \times 10^{10} M_{\odot}$) until reaching the onset of kinetic AGN feedback, while for a given mass bin it may vary by 1 dex for different redshifts, going from ~ 3 at $z = 0$ down to ~ 0.5 at $z = 10$.

The metallicities of gas flowing into TNG100 halos are very close

¹⁰ FIRE-2 gas particles can reach masses of $\sim 250 - 7100 M_{\odot}$, compared to the TNG100 mass resolution of $\sim 10^6 M_{\odot}$ per gas cell.

¹¹ The comparison between η_E for the two definitions of outflows was not made explicitly by Pandya et al. (2021).

¹² For a given halo virial mass, $V_{\text{vir}} \propto R_{\text{vir}}^{-1/2} \propto (1+z)^{1/2}$.

to those of FIRE-2, but only at the 100 km/s virial velocity bin¹³. In FIRE-2, $Z_{\text{in, halo}}$ keeps decreasing with the virial velocity at a similar pace for all redshifts, while for TNG100, we find that low- z metallicities drop much faster with V_{vir} than high- z metallicities, as is the case in Fig. 14. The preventative feedback parameter estimated for FIRE-2 is also very different from its TNG100 equivalent: on one hand, the trend with redshift seen in TNG100 almost vanishes as we consider f_{prev} as a function of the virial velocity, which is also the case for FIRE-2; on the other hand, this parameter never rises above 1 in FIRE-2, while it does so in $z \geq 2$ in TNG100.

8.3.4 Cosmological simulations

Nelson et al. (2019b) analyzed the outflow rates out of galaxies in TNG50, a higher resolution counterpart of TNG100. Gas particles in TNG50 are roughly 2 dex less massive than those of TNG100. To allow for a higher resolution the box side length was decreased by half, leading to fewer high mass halos appearing in TNG50 compared to TNG100 (e.g. at $z = 0$, only 35 TNG50 halos have virial masses $> 10^{13} M_{\odot}$ compared to 232 in TNG100). On the other hand, lower mass halos are much more abundant and more clearly resolved, allowing to investigate them in better detail. In their work, Nelson et al. (2019b) discuss the mass loading factors out of $z = 2$ galaxies in TNG50, measured at a 10 kpc distance from the galaxy. They also present their results as a function the stellar mass within a 30 kpc aperture, which differs from our definition of stellar mass so far (relative to the stellar half-mass radius). The star formation rates used for the mass loading factors are also taken within the same 30 kpc radius.

In Fig. 17 we compare the mass loading factors from Nelson et al. (2019b) to the TNG100 loading factors at the ISM scale from $z = 2$. For an apples-to-apples comparison we estimate the stellar mass and star formation rates in a 30 kpc aperture as well¹⁴. We mark the newly defined stellar mass whose associated virial radius is approximately 100 kpc with a vertical gray dashed line, and the stellar mass whose associated virial mass is the onset of kinetic AGN feedback for $z = 2$ with a vertical black dashed line. The gray shaded region that marks possibly unresolved galaxies in TNG100 extends to $\sim 10^{8.6} M_{\odot}$ rather than the $10^8 M_{\odot}$ mentioned in Sec. 2.1 since under the new definition the galaxy limits extend farther and therefore include more stellar particles. For the star formation rates, however, there is no significant contribution from the region between twice the stellar half-mass radius and a 30 kpc aperture.

The TNG50 mass loading factors are larger by ~ 0.2 dex than the TNG100 mass loading factors up to the onset of kinetic AGN feedback. We note that for a given stellar mass bin up to the onset of kinetic AGN feedback, the TNG50 star formation rates are comparable to their TNG100 counterparts. This therefore suggests that SNe driven mass outflow rates in TNG50 galaxies are higher relative to their TNG100 counterparts.

Next we compare with the EAGLE cosmological hydrodynamic simulation (Schaye et al. 2015), whose outflow rates have been analyzed in Mitchell et al. (2020a) and inflow rates have been analyzed in Mitchell et al. (2020b). The EAGLE simulation is very similar to Illustris-TNG100 in its method for solving hydrodynamics, its sub-grid physics, its box size, and its resolution. The halos and galaxies

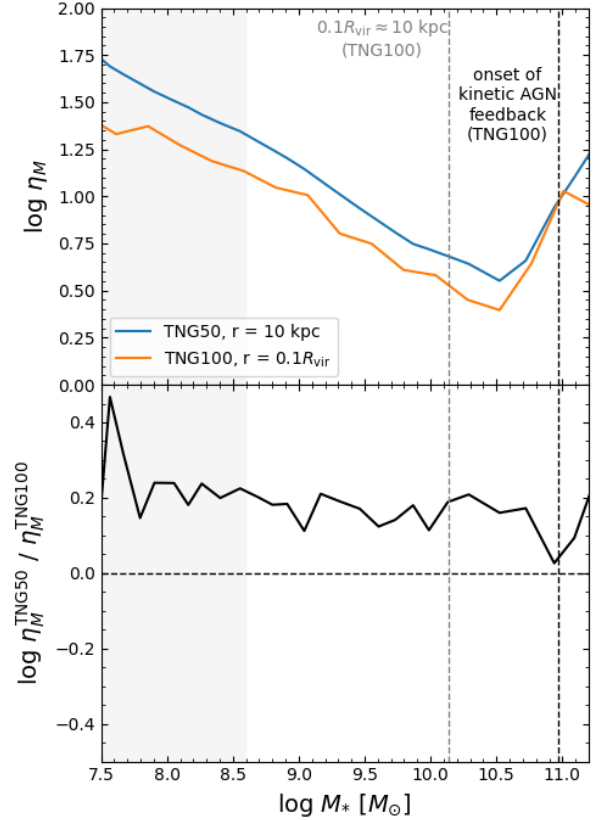


Figure 17. A comparison between the mass loading factors of TNG50 (blue line) and TNG100 (orange line) at $z = 2$, estimated as a function of the stellar mass within a 30 kpc aperture. The TNG50 mass loading factors, taken from Nelson et al. (2019b), are estimated at a distance of 10 kpc from the galactic center while the TNG100 mass loading factors are estimated at $0.1 R_{\text{vir}}$. A gray dashed line marks the stellar mass associated with $R_{\text{vir}} \approx 100$ kpc. A black dashed line marks the stellar mass associated with the onset of kinetic AGN feedback at $z = 2$ in TNG100.

that form within the EAGLE simulation span a similar mass range to those of TNG100. SNe and AGN feedback, however, are implemented differently. SNe feedback is purely thermal, and rather than launching decoupled wind particles, a star forming gas particle adds thermal energy to its neighboring gas particles such that their temperature increases by $\Delta T = 10^{7.5}$ K. Black hole particles are seeded in galaxies in a similar to IllustrisTNG (i.e. planted within halos with $M_{\text{vir}} > 10^{10} M_{\odot}$), and are allowed to accrete material at the Bondi accretion rate, capped by the Eddington accretion rate. SMBH feedback is implemented similarly to SNe feedback, as injection of pure thermal feedback that is applied to neighboring particles. The energy injected into nearby particles is also proportional to $\dot{M}_{\text{BH}} c^2$, and is accumulated until it is large enough to heat a nearby gas particle by $\Delta T = 10^{8.5}$ K, after which a feedback event occurs.

The normalized mass flow rates are relatively close to those of TNG100, and the TNG100 values are higher by at most 1 dex for all redshifts and scales. More can be learned from the mass dependencies. At the ISM scale of EAGLE galaxies, both inflow and outflow rates peak at $M_{\text{vir}} \approx 5 \times 10^{11} M_{\odot}$, similarly to the ISM inflow and outflow rates in TNG100 (in TNG100 this occurs at a slightly lower mass). In EAGLE, the peak in inflow rates is correlated with a relatively high recycling fraction at the ISM scale, similar to our identification of a galactic fountain in TNG100 at this mass scale.

¹³ $V_{\text{vir}} = 100$ km/s is proportional to $M_{\text{vir}} \approx 5 \times 10^{11} M_{\odot}$ at $z = 0$, $M_{\text{vir}} \approx 6 \times 10^{10} M_{\odot}$ at $z = 4$, and $M_{\text{vir}} \approx 2 \times 10^{10}$ at $z = 10$.

¹⁴ We use instantaneous star formation rates whereas Nelson et al. (2019b) takes an average over the last 100 Myr. We find the difference between the two to be negligible for TNG100 galaxies.

However, SNe driven outflows in TNG100 do not leave the halo, while in EAGLE SNe driven outflows travel out to several times the virial radius and act as ejective feedback, draining the halo of its baryons (only a fraction of the EAGLE outflows remain in the halo, where they can cool again and fall onto the galaxy). Comparing AGN dominated halos at the ISM scale, we find that in TNG100 the kinetic AGN feedback is primarily ejective rather than preventative, while the thermal AGN feedback implemented in EAGLE acts in a more preventative manner, suppressing star formation without driving strong ISM outflows.

The mass loading factors of outflows in EAGLE galaxies has a similar mass dependence to those of TNG100, but are lower by a factor of ~ 10 . While in EAGLE the onset of AGN feedback leads to a flattening of η_M at roughly ~ 1 for $M_{\text{vir}} > 10^{12} M_{\odot}$, in TNG the mass loading factor increases rapidly as kinetic AGN feedback kicks in — aligning with the dip in both baryon fractions and mass flow rates near this characteristic mass.

At the halo scale, the difference between the two simulations is more noticeable. In TNG100, the onset of kinetic AGN feedback acts as both preventative and ejective, delaying inflows from entering the halo in addition to ejecting halo baryons, whereas for higher masses we find the kinetic AGN feedback to be mainly preventative. In EAGLE, Mitchell et al. (2020a) report only preventative feedback at the halo scale. For a detailed side-by-side comparison of inflow and outflow rates on galaxy and ISM scales in TNG100, EAGLE, and SIMBA, as well as a comparison of the baryon fractions, please see Wright et al. (2024).

8.4 Caveats in this analysis

8.4.1 Impact of sub-grid implementations

The goal of this work has been simply to characterize the flow rates of mass, energy, and metals on galaxy and halo scales in the widely-used IllustrisTNG simulation. However, as already touched upon, these quantities can be extremely sensitive to the manner in which the sub-grid processes of supernova driven winds and AGN feedback are implemented. This is the reason for the fairly large discrepancies between many aspects of the baryon cycles seen in zoom-in and large volume cosmological simulations, as discussed in more detail in Pandya et al. (2020) and Wright et al. (2024). There are ongoing efforts to improve these sub-grid implementations using insights from smaller-scale, higher resolution, idealized simulations that more directly resolve the launching of large-scale outflows by supernovae explosions (e.g., Walch et al. 2015; Fielding et al. 2017b,a, 2018; Hu 2019; Schneider et al. 2020; Kim et al. 2020a,b; Li & Bryan 2020). An important insight from these simulations is that emergent outflows are *multi-phase* (like observed outflows in real galaxies), with the hot diffuse phase carrying most of the energy (and metals) and the cold dense phase carrying most of the mass. In most past cosmological simulations such as IllustrisTNG and EAGLE, outflows at launch have a single velocity and/or temperature, and typically inject either (mostly) kinetic energy (TNG) or only thermal energy (EAGLE). Thus the predicted baryon cycle emerging from these simulations is almost certainly incomplete.

There are ongoing efforts to develop and implement new multi-phase subgrid wind recipes into cosmological simulations, such as the Arkenstone project (Smith et al. 2024a,b; Bennett et al. 2024). Bennett et al. (2024) pointed out that the required mass loadings at launch adopted by Pillepich et al. (2018) in order to match the stellar mass to halo mass relation and other observables in the TNG wind implementation are much higher than those required in the

Arkenstone model, when it is implemented within the AREPO code with all other physical processes identical to the TNG physics model. This is because Arkenstone’s more accurate treatment of the wind energy coupling with the CGM enables more effective regulation of CGM cooling via preventative feedback, just as predicted by the gas regulator and semi-analytic models of Carr et al. (2023), Voit et al. (2024a), and Pandya et al. (2023).

Similarly, the scales on which AGN feedback operates and the manner in which it regulates star formation and black hole growth are quite sensitive to the way that black hole seeding, accretion, and feedback are implemented (again, see Wright et al. (2024) for a detailed discussion of the baryon cycle in AGN-dominated halos in TNG, SIMBA, and EAGLE). For example, TNG, like most cosmological simulations, only seeds black holes in halos above a critical mass. However, there is mounting evidence for the existence of both black holes and AGN driven outflows in lower mass galaxies (Manzano-King et al. 2019). A second example is the criterion for switching on the powerful kinetic AGN feedback mode, which depends on both accretion rate and black hole mass. This critical black hole mass for the onset of kinetic BH feedback creates a “dip” in the CGM mass fraction at intermediate halo masses, which has been correlated with a minimum in the integrated tSZ signal from the CGMs of galaxies, as shown by Oren et al. (2024); and X-ray brightness of the CGM surrounding galaxies of different masses (see e.g. Lau et al. 2025). This feature is not seen in other simulations that implement AGN feedback differently (e.g. EAGLE).

8.4.2 Determining the dominant feedback mechanism

Our analysis of the dominant feedback mechanism in Sec. 3.3 is key to our interpretation of flow rates in and out of TNG100 galaxies and halos. There are two simplifying assumptions we made in our analysis which may affect our results.

Firstly, in our estimation of the time scales between which we integrate the instantaneous energy injection rates, we assumed outflows do not decelerate as they travel from their source to the halo shell. This work itself has suggested this assumption is incorrect by concluding that as redshifts increase, outflows are more likely to stall as they interact with the dense CGM. Secondly, as we estimate the total energy injected by each feedback mechanism, we integrate over the instantaneous energy injection rates estimated per snapshot, and over time-scales determined by the distance between them. In practice, the time-steps between which TNG100 particles evolve are much smaller than the time difference between snapshots. Additionally, kinetic AGN feedback is not injected per time step like thermal AGN feedback and SNe feedback, but is accumulated until reaching a certain threshold and is then released. In our analysis we treated all three feedback mechanisms as if they are all injected in all available time steps, which may lead to an inaccurate estimate of when kinetic AGN feedback was injected.

Despite that, we opted to calculate the integrated energies as we did due to the fact that we only use them to compare between mechanisms, and not to estimate any of our quantitative results. More importantly, we integrate the energy injection rates over three different time scales and take the dominant feedback mechanism to be the one that injected more energy throughout two out of three, decreasing the chance of skewing our results due to mistiming.

9 CONCLUSIONS

We have analyzed the mass, energy, and metal inflow and outflow rates through 9522 TNG100 galaxies and their surrounding halos spanning five orders of magnitude in halo virial mass and from $z = 10$ to $z = 0$. We separate our sample into halos with the dominant feedback mechanism estimated to be either SNe feedback, thermal AGN feedback, or kinetic AGN feedback. We evaluated the flow rates through an “ISM shell” ($0.05 - 0.15 R_{\text{vir}}$) and a “halo shell” ($0.95 - 1.05 R_{\text{vir}}$). In addition to measuring the direction of the net flows, we identified markers of preventative vs. ejective feedback, and measured the metal enrichment of the flowing gas relative to its source, and the loading factors associated with each type of flow. Our conclusions are as follows:

(i) The dominant feedback mechanism transitions from SNe to kinetic AGN feedback for galaxies in the centers of halos with $M_{\text{vir}} \approx 10^{12} M_{\odot}$ at $z = 2$. While thermal AGN feedback, which kicks in at $M_{\text{vir}} \approx 10^{11} M_{\odot}$, injects more energy into galaxies compared to SNe feedback, it does not appear to significantly affect star formation nor drive any outflows.

(ii) Halo mass normalized mass outflow rates in SN feedback dominated halos decrease by 1-1.5 dex from $z \sim 10$ –0 on halo scales, and by ~ 2 dex on ISM scales. Normalized energy flow rates show a very similar pattern.

(iii) Halos of all masses show strongly net positive mass inflow rates at high redshift ($z \gtrsim 2$) on halo scales; on ISM scales the inflow remains net positive in halos with $M_{\text{vir}} \lesssim 10^{12} M_{\odot}$, but the excess is much smaller.

(iv) At high redshift ($z \gtrsim 2$), there is little evidence for halo scale preventative feedback, with halo accretion rates modestly exceeding the dark matter halo growth rate times the universal baryon fraction. Halo scale preventative feedback remains modest at lower redshift.

(v) On ISM scales, there is evidence for rapid recycling of ejected gas (“galactic fountain”) in intermediate to low mass halos ($M_{\text{vir}} \lesssim 10^{12} M_{\odot}$).

(vi) The mass loading factor in SN dominated halos at $z = 0$ decreases from $\eta_M \approx 20$ –5.6 over the mass range $10^{11} \lesssim M_{\text{vir}} \lesssim 10^{12.5}$. The mass loading at high redshift ($z \sim 2$ –10) has a lower normalization and a steeper decline with increasing mass. In contrast, energy loadings remain nearly constant with both halo mass and redshift. Metal loadings are nearly independent of halo mass, but increase with decreasing redshift.

(vii) At halo masses of $\sim 10^{14} M_{\odot}$ and above, the gravitational potential begins to become strong enough to resist even the powerful kinetic mode AGN feedback, leading to a strong increase in CGM and baryon fractions towards higher halo masses.

(viii) The emergent mass outflow rates do not trace the energy injection rates adopted in the TNG wind subgrid model in a straightforward manner. This may be because outflows are more likely to be stalled by high density CGM at high redshifts, and to mix with and entrain low density CGM at low redshifts.

(ix) The metallicities of outflows on halo scales are similar to those of the CGM, or slightly under-enriched at very high halo masses ($\gtrsim 10^{13.5} M_{\odot}$). The metallicities of ISM-scale outflows are similar to the assumed TNG metal enrichment at launch ($Z_{\text{out}} \sim 0.4 Z_{\text{ISM}}$) for halos where SN feedback is dominant, and increases rapidly to unity for AGN-feedback dominated halos. The metallicity of gas that is flowing into the ISM is enriched relative to the overall average CGM metallicity at all halo masses, consistent with the measured radial metallicity gradients in the CGM of TNG halos.

The analysis presented in this paper can provide greater insight

into *why* stellar and AGN feedback must behave in certain ways as a function of halo mass and cosmic time in order to reproduce basic quantities such as the galaxy stellar mass fraction, CGM baryon fraction, and quenched galaxy fractions. In turn, this will help guide the development of more physically motivated sub-grid recipes and next generation semi-analytic models.

ACKNOWLEDGEMENTS

We wish to thank Jonathan Stern and Amit Nestor-Shachar for their helpful advice and discussions. This work was supported by the German Science Foundation via DFG/DIP grant STE/ 1869-2 GE/ 625 17-1, by the Center for Computational Astrophysics (CCA) of the Flatiron Institute, and by the Mathematical and Physical Sciences (MPS) division of the Simons Foundation, USA. Support for VP was provided by NASA through the NASA Hubble Fellowship grant HST-HF2-51489 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. VP thanks Drummond Fielding, Daniel Anglés-Alcázar and Greg Bryan for getting him interested in gas thermodynamics.

DATA AVAILABILITY

The data presented in this analysis will be made available upon reasonable request.

REFERENCES

- Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Hopkins P. F., Quataert E., Murray N., 2017a, *MNRAS*, **470**, 4698
- Anglés-Alcázar D., Faucher-Giguère C.-A., Quataert E., Hopkins P. F., Feldmann R., Torrey P., Wetzel A., Kereš D., 2017b, *MNRAS*, **472**, L109
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, **47**, 481
- 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
- Bennett J. S., Smith M. C., Fielding D. B., Bryan G. L., Kim C.-G., Springel V., Hernquist L., 2024, *arXiv e-prints*, p. arXiv:2410.12909
- Benson A. J., 2010, *Phys. Rep.*, **495**, 33
- Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, *Nature*, **311**, 517
- Bregman J. N., 1980, *ApJ*, **236**, 577
- Bryan G. L., Norman M. L., 1998, *ApJ*, **495**, 80
- Burbidge E. M., Burbidge G. R., Rubin V. C., 1964, *ApJ*, **140**, 942
- Carr C., Bryan G. L., Fielding D. B., Pandya V., Somerville R. S., 2023, *ApJ*, **949**, 21
- Chabrier G., 2003, *PASP*, **115**, 763
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, **319**, 168
- Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, **421**, 98
- Davies J. J., Crain R. A., Oppenheimer B. D., Schaye J., 2020, *MNRAS*, **491**, 4462
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, **292**, 371
- Dekel A., Silk J., 1986, *ApJ*, **303**, 39
- Fabian A. C., 2012, *ARA&A*, **50**, 455
- Falgarone E., et al., 2017, *Nature*, **548**, 430
- Fielding D., Quataert E., McCourt M., Thompson T. A., 2017a, *MNRAS*, **466**, 3810
- Fielding D., Quataert E., Martizzi D., Faucher-Giguère C.-A., 2017b, *MNRAS*, **470**, L39
- Fielding D., Quataert E., Martizzi D., 2018, *MNRAS*, **481**, 3325
- Finlator K., Davé R., 2008, *MNRAS*, **385**, 2181
- Fraternali F., Binney J. J., 2006, *MNRAS*, **366**, 449

- Fraternali F., Binney J. J., 2008, *MNRAS*, **386**, 935
- Garcia A. M., et al., 2023, *MNRAS*, **519**, 4716
- Genel S., et al., 2008, *ApJ*, **688**, 789
- Gnat O., Sternberg A., 2007, *ApJS*, **168**, 213
- Guo Y., et al., 2023, *Nature*, **624**, 53
- Hafen Z., et al., 2022, *MNRAS*, **514**, 5056
- Heckman T. M., Best P. N., 2014, *ARA&A*, **52**, 589
- Heckman T. M., Best P. N., 2023, *Galaxies*, **11**, 21
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *MNRAS*, **451**, 2663
- Hirschmann M., De Lucia G., Fontanot F., 2016, *MNRAS*, **461**, 1760
- Hopkins P. F., et al., 2018, *MNRAS*, **480**, 800
- Hopkins P. F., et al., 2023, *MNRAS*, **519**, 3154
- Hu C.-Y., 2019, *MNRAS*, **483**, 3363
- Karmakar T., Genel S., Somerville R. S., 2023, *MNRAS*, **520**, 1630
- Kauffmann G., White S. D. M., Guiderdoni B., 1993, *MNRAS*, **264**, 201
- Kim C.-G., et al., 2020a, *ApJ*, **900**, 61
- Kim C.-G., et al., 2020b, *ApJ*, **903**, L34
- Kroupa P., 2001, *MNRAS*, **322**, 231
- Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, *MNRAS*, **481**, 3573
- Lau E. T., et al., 2025, *ApJ*, **984**, 190
- Li M., Bryan G. L., 2020, *ApJ*, **890**, L30
- Manzano-King C. M., Canalizo G., Sales L. V., 2019, *ApJ*, **884**, 54
- Mitchell P. D., Schaye J., 2022, *MNRAS*, **511**, 2948
- Mitchell P. D., et al., 2018, *MNRAS*, **474**, 492
- Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020a, *MNRAS*, **494**, 3971
- Mitchell P. D., Schaye J., Bower R. G., 2020b, *MNRAS*, **497**, 4495
- Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, *ApJ*, **710**, 903
- Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *MNRAS*, **454**, 2691
- Naab T., Ostriker J. P., 2017, *ARA&A*, **55**, 59
- Nelson D., Genel S., Vogelsberger M., Springel V., Sijacki D., Torrey P., Hernquist L., 2015, *MNRAS*, **448**, 59
- Nelson D., et al., 2019a, *Computational Astrophysics and Cosmology*, **6**, 2
- Nelson D., et al., 2019b, *MNRAS*, **490**, 3234
- Oren Y., Sternberg A., McKee C. F., Faerman Y., Genel S., 2024, *ApJ*, **974**, 291
- Osterbrock D. E., 1960, *ApJ*, **132**, 325
- Pakmor R., Springel V., 2013, *MNRAS*, **432**, 176
- Pakmor R., Bauer A., Springel V., 2011, *MNRAS*, **418**, 1392
- Pandya V., et al., 2020, *ApJ*, **905**, 4
- Pandya V., et al., 2021, *MNRAS*, **508**, 2979
- Pandya V., et al., 2023, *ApJ*, **956**, 118
- Peeples M. S., Shankar F., 2011, *MNRAS*, **417**, 2962
- Pillepich A., et al., 2018, *MNRAS*, **473**, 4077
- Planck Collaboration et al., 2016, *Astronomy & Astrophysics*, **594**, A13
- Rodriguez-Gomez V., et al., 2015, *MNRAS*, **449**, 49
- Rupke D. S. N., 2018, *Galaxies*, **6**, 138
- Rupke D. S. N., et al., 2019, *Nature*, **574**, 643
- Schaye J., et al., 2015, *MNRAS*, **446**, 521
- Schneider E. E., Ostriker E. C., Robertson B. E., Thompson T. A., 2020, *ApJ*, **895**, 43
- Smith M. C., et al., 2024a, *MNRAS*, **527**, 1216
- Smith M. C., Fielding D. B., Bryan G. L., Bennett J. S., Kim C.-G., Ostriker E. C., Somerville R. S., 2024b, *MNRAS*, **535**, 3550
- Somerville R. S., Davé R., 2015, *ARA&A*, **53**, 51
- Somerville R. S., Primack J. R., 1999, *MNRAS*, **310**, 1087
- Somerville R. S., Popping G., Trager S. C., 2015, *MNRAS*, **453**, 4337
- Springel V., 2010, *MNRAS*, **401**, 791
- Springel V., Hernquist L., 2003, *MNRAS*, **339**, 289
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, **328**, 726
- Strickland D. K., Heckman T. M., 2009, *ApJ*, **697**, 2030
- Tillman M. T., et al., 2023, *AJ*, **166**, 228
- Torrey P., et al., 2019, *MNRAS*, **484**, 5587
- Tremblay G. R., et al., 2018, *ApJ*, **865**, 13
- Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, **43**, 769
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, **436**, 3031
- Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, *Nature Reviews Physics*, **2**, 42
- Voit G. M., Pandya V., Fielding D. B., Bryan G. L., Carr C., Donahue M., Oppenheimer B. D., Somerville R. S., 2024a, *ApJ*, **976**, 150
- Voit G. M., Carr C., Fielding D. B., Pandya V., Bryan G. L., Donahue M., Oppenheimer B. D., Somerville R. S., 2024b, *ApJ*, **976**, 151
- Walch S., et al., 2015, *MNRAS*, **454**, 238
- Wechsler R. H., Tinker J. L., 2018, *ARA&A*, **56**, 435
- Weinberger R., et al., 2017, *MNRAS*, **465**, 3291
- Wellons S., et al., 2023, *MNRAS*, **520**, 5394
- 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
- Wright R. J., Somerville R. S., Lagos C. d. P., Schaller M., Davé R., Alcazar D., Genel S., 2024, *MNRAS*, **532**, 3417
- Yoshida M., Yagi M., Ohyama Y., Komiyama Y., Kashikawa N., Tanaka H., Okamura S., 2016, *ApJ*, **820**, 48

This paper has been typeset from a $\text{\LaTeX}$ file prepared by the author.