

# TNG SAM: Bridging Hydrodynamical Complexity and Semi-Analytic Efficiency to Model Galaxy Formation

Osase Omoruyi,<sup>1\*</sup> Bryan Terrazas,<sup>2</sup> Yossi Oren<sup>3</sup> Viraj Pandya<sup>4</sup> Rachel Somerville<sup>5</sup> Lars Hernquist<sup>1</sup>

<sup>1</sup>Center for Astrophysics | Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA

<sup>2</sup>Department of Physics & Astronomy, Oberlin College, Oberlin, OH, 44074, USA

<sup>3</sup>School of Physics and Astronomy, Tel Aviv University, Ramat Aviv 69978, Israel

<sup>4</sup>Columbia University, Department of Astronomy, New York, NY 10027, USA

<sup>5</sup>Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

All cosmological models of galaxy formation weigh physical accuracy against computational efficiency. Computationally expensive hydrodynamical simulations provide detailed insights into the co-evolution of dark matter, gas, stars, and black holes, but still rely on simplified subgrid models for small-scale processes (e.g., star formation). In contrast, semi-analytic models (SAMs) achieve efficiency by relying entirely on subgrid models, limiting their physical fidelity. In this work, we leverage the strengths of the Santa Cruz SAM and the IllustrisTNG 100 Mpc hydrodynamical simulation to develop the TNG SAM. Calibrated to emulate the underlying flow of baryons in stellar feedback-dominated TNG galaxies ( $\sim 10^{10} M_{\odot} < M_{200} < 10^{12} M_{\odot}$ ), the TNG SAM introduces several key updates to the Santa Cruz SAM framework: 1) a prescription for the efficiency of halo gas (re-)accretion, 2) a revised cooling recipe for gas transitioning from the hot halo to the galaxy based on the cooling time, eliminating the traditional “cold mode” vs. “hot mode” dichotomy, 3) the incorporation of both galactic and halo-scale outflows, and 4) metallicity-dependent mass loading factors that describe the circulation of metals between galaxies and their surroundings. These updates not only enable the TNG SAM to reproduce TNG’s underlying flow of gas and metals from the scale of the galaxy to the halo, but also global galaxy (e.g., stellar mass) and halo (e.g. baryon fraction) properties within  $\sim 30\%$  accuracy up to  $z = 6$ . This work demonstrates that, with proper calibration, SAMs can effectively capture the complex physics of galaxy formation seen in hydrodynamical simulations, offering a flexible and physically motivated framework for studying galaxy evolution across the large cosmological volumes targeted by future surveys.

**Key words:** galaxies: evolution – galaxies: formation – galaxies: halos – methods: numerical

## 1 INTRODUCTION

Observations across the electromagnetic spectrum have revealed a Universe rich in dark matter, stars, gas, and dust, sparking numerous questions regarding their formation and evolution. The  $\Lambda$  Cold Dark Matter ( $\Lambda$ CDM) model of cosmological structure formation proposes that the Universe consists of a complex filamentary network molded by gravity and populated with dark matter halos. Within these halos, baryonic matter falls to the center of the gravitational potential well, condenses, and forms galaxies (White & Rees 1978). To study this process from cosmological to stellar scales, numerous studies have used numerical simulations to predict and correlate observable galaxy properties—such as stellar mass, cold gas mass, and rate of star formation—with dark matter halo characteristics (see Somerville & Davé (2015) for a comprehensive review). With the advent of powerful observatories like *JWST* and the upcoming *Nancy Grace Roman Space Telescope*, our ability to observe galaxies in the earli-

est days of the Universe and across larger cosmic volumes is rapidly improving, necessitating the development of more sophisticated and flexible cosmological models of galaxy formation.

Modeling how galaxies form and evolve requires investigating how they interact with their surrounding environment. As dark matter halos grow, galaxies accrete gas from the intergalactic medium (IGM) into the circumgalactic medium (CGM), connecting the large-scale cosmic web with the galaxies at the centers of these halos. Gravity continues to drive flows of gas from the CGM into the interstellar medium (ISM), fueling star formation. As stars form, evolve, and eventually die, the resulting feedback from stellar winds, supernovae (SNe), and black hole activity heats and/or expels material back into the CGM or IGM. This matter can later be re-accreted, fueling new star formation. This cycle of gas accretion, star formation, feedback, and re-accretion—known as the “baryon cycle”—is fundamental to shaping how galaxies evolve.

To model the baryon cycle within the  $\Lambda$ CDM paradigm, two primary techniques have emerged: numerical hydrodynamical simulations and semi-analytic models (SAMs). Improving upon dark-matter

\* E-mail: osase.omoruyi@gmail.com

and gravity-only simulations to include baryonic physics, numerical hydrodynamical simulations explicitly solve equations of gravity, (magneto)hydrodynamics, and thermodynamics. This comprehensive approach enables them to self-consistently model the coupled evolution of dark matter, gas, stars, and black holes, providing detailed information on each component’s spatial distribution, temperature, and kinematics. However, simulating large cosmological volumes at high resolution requires sacrificing capturing the detailed physics of key processes occurring below their resolution limits, such as star formation and black hole feedback. As a result, these simulations rely on phenomenological “subgrid recipes” that contain adjustable parameters calibrated to reproduce a given set of observations. Additionally, the computational expense of hydrodynamical simulations also limits their ability to explore variations in subgrid models across vast volumes.

In contrast, SAMs offer a computationally efficient alternative by treating galaxies as unresolved entities and relying solely on “subgrid” recipes to simulate their formation and evolution. Built from dark matter-only N-body simulations or semi-analytic approaches based on merger trees (e.g., Press-Schechter formalism), SAMs bridge the gap between the hierarchical growth of dark matter halos and observed galaxy populations by employing simplified analytical models of key physical processes relevant to galaxy evolution (White & Frenk 1991; Cole 1991; Lacey & Silk 1991). In a SAM, a galaxy is typically represented as a collection of mass reservoirs, such as the stellar disk, stellar bulge, cold gas disk, hot halo gas, and ejected material. As a result, SAMs essentially function as “flow models,” solving ordinary differential equations to track the flow of baryons between each component, with the flows governed by the physical prescriptions included in the model. The specific processes modeled, such as radiative cooling, chemical evolution, and feedback, lie at the discretion of the model’s creator. Some SAMs incorporate a wide variety of detailed physical prescriptions (e.g., Santa Cruz Somerville et al. 2008b, SAGE Croton et al. 2006, GALFORM Lacey et al. 2016, SHARK Lagos et al. 2018), while others focus on a more limited set of deemed most critical for galaxy formation (e.g., Davé et al. 2012; Bouché et al. 2010). This flexibility allows SAMs to be tailored to specific scientific questions and computational resources, making them a valuable tool for understanding the “big picture” of the complexities shaping galaxy evolution.

Despite their different approaches, both SAMs and hydrodynamical simulations share a reliance on phenomenological prescriptions to simulate fundamental processes in the baryon cycle. This has led to the practice of “fine-tuning” or calibrating these models to reproduce a selection of observations, introducing additional uncertainty when comparing and interpreting results. Early comparisons of SAMs and hydrodynamical simulations (e.g. Benson et al. 2001; Helly et al. 2003; Yoshida et al. 2002; Saro et al. 2010; Stringer et al. 2010) found general agreement in predicting properties such as the overall distribution of galaxies and their cold gas mass fractions, but differences in properties like the efficiency of gas cooling, star formation histories, and the merger history of specific galaxies. More detailed comparisons (e.g. Stringer et al. 2010; Monaco et al. 2014; Guo et al. 2016; Mitchell et al. 2018; Côté et al. 2018) similarly found that SAMs can approximate key galaxy properties such as stellar mass functions and specific star formation rates to a reasonable degree of accuracy when compared to hydrodynamical simulations. However, they also highlighted significant differences in the treatment of gas cooling, star formation efficiencies, galaxy sizes, and the evolution of star formation rate density. These results suggest that discrepancies between the two methods may arise from specific parameterizations and models used, rather than inherent limitations of each approach.

Recognizing SAMs and hydrodynamical simulations as complementary, rather than competing, techniques, can significantly enhance their predictive power. For example, the detailed physical insights from hydrodynamical simulations can refine the subgrid models in SAMs, as demonstrated by Pandya et al. (2020), who compared the FIRE-2 hydrodynamic simulations (Hopkins et al. 2018) with the Santa Cruz SAM (SC SAM) for halos with masses between  $10^{10}$  and  $10^{12} M_{\odot}$ . They found substantial agreement between FIRE-2 and the SC SAM for stellar and cold gas masses, but significant disagreements in the properties of the hot halo gas — a measurement not widely available for calibrating or validating models, particularly for Milky Way and lower mass halos. By adjusting the SAM to account for the suppression of halo gas accretion via stellar winds, particularly in dwarf galaxies, Pandya et al. (2020) reproduced the reduced halo gas accretion efficiencies of the FIRE-2 complementary galaxies remarkably well. Similarly, Côté et al. (2018) refined GAMMA, a SAM aimed at understanding the chemical evolution of low-mass galaxies and the origins of metal-poor stars, by calibrating it against the observable properties of the most massive galaxy in the Wise et al. (2011) high-redshift hydrodynamical simulation. They found that implementing a non-uniform mixing model was key to reproducing the observable quantities in the hydrodynamical simulation. Lastly, in a more extensive effort to incorporate insights from hydrodynamical simulations into simplified models, Pandya et al. 2023 used FIRE-2’s predictions for mass and energy flows in a given set of halos to calibrate their regulator model, which emphasizes the role of the CGM and its energy content in governing galaxy evolution. These calibrations allowed their model to successfully emulate the detailed baryon cycle observed in FIRE-2 halos.

Of course, SAMs can also enhance the predictive power of hydrodynamical simulations. For instance, the IllustrisTNG model for galactic-scale, star formation-driven kinetic winds incorporates several refinements over its predecessor in Illustris, described in detail in Section 2.1. Among these, the TNG model introduced a scaling of wind velocity with the local dark matter velocity dispersion,  $\sigma_{\text{DM}}$ , and a redshift-dependent factor to better match observed stellar mass functions and luminosity functions across redshift (Henriques et al. 2013). These modifications—rooted in SAM-driven findings—enhanced the realism of feedback prescriptions in IllustrisTNG, showcasing the bidirectional synergy between these modeling approaches.

In this work, we focus on the former direction — improving the fidelity of SAMs with insights from hydrodynamical simulations — by comparing the predictions of IllustrisTNG (hereafter TNG) with the Santa Cruz SAM, both of which have demonstrated success in reproducing a wide range of observations. Building on the initial comparisons between the two simulations made by Gabrielpillai et al. (2022), we modify the Santa Cruz SAM’s galaxy formation model to better emulate the detailed physical processes observed in the IllustrisTNG simulations, with a specific focus on mirroring TNG’s treatment of stellar feedback. By calibrating the SAM with measurements of galaxy- and halo-scale inflow and outflow rates for a subset of TNG galaxies (Cohen et al., 2025, in prep), we aim to emulate the underlying baryon cycle modeled in TNG, resulting in the creation of a new “TNG SAM.” This endeavor serves three primary goals: (1) to simplify and distill TNG’s complex feedback and baryon cycling behaviors into a more interpretable semi-analytic framework, (2) to facilitate direct comparisons between SAMs and hydrodynamical simulations, clarifying their respective strengths and limitations, and (3) to ultimately extend the predictive power of SAMs to volumes beyond the reach of hydrodynamical simulations.

This paper is organized as follows: In Section 2, we provide an

	TNG100 Pillepich et al. (2018a); Springel & Hernquist (2003)	SC SAM Somerville et al. (2008a, 2015); Yung et al. (2019)	TNG SAM this paper
<i>Box Size</i>	$75 h^{-1} \text{ Mpc}$	$75 h^{-1} \text{ Mpc}$	$75 h^{-1} \text{ Mpc}$
<i>Particle Resolution</i>	$m_{\text{DM}} = 5.06 \times 10^6 M_{\odot}/h$	$\text{mass}_{\text{min}} = 8.85 \times 10^8 M_{\odot}$	$\text{mass}_{\text{min}} = 8.85 \times 10^8 M_{\odot}$
<i>Halo Gas Accretion</i>	...	NFW	$f_{\text{in}}^{\text{CGM}} \propto f(M_{\text{halo}}, z)$
<i>Hot Halo Gas Re-incorporation</i>	...	$\chi_{\text{re-infall}} = 0.1$	bundled into $f_{\text{in}}^{\text{CGM}}$
<i>Hot Halo Gas Cooling</i>	...	$r_{\text{cool}} < r_{\text{vir}}$ : cold mode accretion $r_{\text{cool}} > r_{\text{vir}}$ : hot mode accretion	$t_{\text{cool}} \propto f(M_{\text{halo}}, z)$
<i>Star Formation</i>	Kennicutt-Schmidt; gas above density threshold forms stars; $\Sigma_{\text{SFR}} \propto \rho^{1.5}$ ; $\tau_{*,0} = 2.1 \text{ Gyr}$ ; Chabrier IMF	multi-phase gas partitioning into neutral, ionized, and molecular components (Gnedin & Kravtsov 2011); H <sub>2</sub> -based star formation recipe (Bigiel et al. 2008); $\tau_{*,0} = 1.0 \text{ Gyr}$ ; Chabrier IMF	Kennicutt-Schmidt; gas above $\Sigma_{\text{crit}} = 4.5 M_{\odot} \text{ pc}^{-2}$ forms stars $\Sigma_{\text{SFR}} \propto \rho^{1.5}$ ; $\tau_{*,0} = 1.0 \text{ Gyr}$ ; Chabrier IMF
<i>Chemical Enrichment</i>	delayed recycling; metals produced by AGB stars ( $1-8 M_{\odot}$ ), SNII ( $8-100 M_{\odot}$ ), and SNIa; yields for each source determined from rescaled yield tables; enrichment events discretized when mass fraction $> 0.0001$ or star age $< 100 \text{ Myr}$	instantaneous recycling; fixed yield $y = 1.2$ ; $f_{\text{recycle}} = 0.43$	instantaneous recycling; fixed yield $y = 2.0$ ; $f_{\text{recycle}} = 0.43$
<i>Stellar Feedback</i>	originates from star-forming gas particles; Injection velocity $\propto [350 \text{ km s}^{-1}, f(\sigma_{\text{DM}}, H(z))]$ ; $\eta_{\text{ISM}} \propto f(v_w, e_w)$	originates from scale of galaxy; SNe ejection efficiency $\epsilon_{\text{SN}} = 1.7$ , ejection scale $V_{\text{eject}} = 110 \text{ km/s}$ ; gas deposited in ejected reservoir	originates from scale of galaxy; $\eta_{\text{ISM}} \propto f(v_w, e_w, M_{\text{halo}}, z)$ ; gas deposited in hot halo; leftover SNe energy drives halo outflows (Henriques et al. 2015); $\eta_{\text{halo}} \propto f(\eta_{\text{ISM}}, M_{\text{halo}}, z)$ ; gas deposited in ejected reservoir
<i>Metal Circulation</i>	metal loading of wind particles: $\gamma_w = 0.4$	metal flows carry the same metallicity as the source reservoir	metal flows carry metallicity $\propto f(M_{\text{halo}}, z)$ of the source reservoir
<i>Calibrations</i>	cosmic SFR density as a function of redshift; galaxy stellar mass vs radius function at $z = 0$ ; stellar-to-halo mass relation, BH mass versus stellar mass relation; hot gas fraction in galaxy clusters	stellar-to-halo mass ratio; stellar mass function; stellar mass–metallicity relation; cold gas fraction vs stellar mass relation for disc-dominated galaxies; black hole mass versus bulge mass relation	<i>TNG100</i> : $f_{\text{in,CGM}}$ ; $t_{\text{cool}}$ ; $R_{\text{disk}}$ vs. $r_{\text{vir}}$ ; $\eta_{\text{ISM}}$ ; $\eta_{\text{halo}}$ ; $\zeta^{\text{in/out}}_{\text{ISM}}$ ; $\zeta^{\text{in/out}}_{\text{halo}}$

**Table 1.** Comparison of the different approaches to modeling galaxy formation taken by TNG100, the Santa Cruz SAM, and the newly developed TNG SAM. Key aspects include simulation box size, resolution, halo gas accretion and cooling, star formation prescriptions, chemical enrichment, stellar feedback, metal circulation, and calibration strategies.

overview of the IllustrisTNG hydrodynamical simulations and the Santa Cruz SAM, while Section 3 details the methods used to compare their respective outputs. Section 4 outlines the modifications made to the SC SAM’s galaxy formation framework that underpin the development of the new “TNG SAM.” The results of these modifications, namely the performance of the TNG SAM in reproducing key global and flow properties of the galaxies modeled in the TNG Universe, are presented in Section 5. In Section 6, we analyze how each modification made contributes to the TNG SAM’s success, and

also discuss the model’s current limitations and potential areas for further refinement. Finally, Section 7 provides a concise summary of the work done and its implications for developing the next generation of SAMs.

## 2 MODEL DESCRIPTIONS

In this section, we describe the two models used: the IllustrisTNG hydrodynamical simulations and the Santa-Cruz semi-analytic model. We summarize their key parameters in Table 1 and outline the methods used to compare the outputs of the Santa Cruz SAM with those of TNG.

### 2.1 The IllustrisTNG Simulations

The IllustrisTNG simulations are a suite of gravitomagnetohydrodynamical simulations that model the physical processes governing the formation and evolution of galaxies across different cosmological volumes (Springel et al. 2018; Weinberger et al. 2017; Pillepich et al. 2018b; Nelson et al. 2018). The suite includes nine simulations, with box sizes of 50, 100, and 300 Mpc<sup>3</sup>, each run at three different resolutions (1 being the highest and 3 the lowest; Marinacci et al. 2018; Naiman et al. 2018; Springel et al. 2018; Nelson et al. 2019; Pillepich et al. 2019). Each simulation has a companion dark matter-only (DMO) and an analogous full-physics (FP) run. The cosmological parameters used in each simulation are taken from Planck Collaboration et al. (2016), with matter density  $\Omega_{M,0} = 0.3089$ , baryon density  $\Omega_{b,0} = 0.0486$ , dark energy density  $\Omega_{\Lambda,0} = 0.6911$ , Hubble constant  $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , power spectrum normalization factor  $\sigma_8 = 0.8159$ , and spectral index  $n_s = 0.9667$ .

TNG builds upon the original Illustris model (Vogelsberger et al. 2013; Torrey et al. 2014), which uses the AREPO code (Springel 2010) to solve coupled equations of self-gravity and magnetohydrodynamics (Pakmor et al. 2011; Pakmor & Springel 2013). To achieve improved consistency with observations, TNG refined the original Illustris models for active galactic nuclei (AGN) feedback, galactic winds, and magnetic fields (Weinberger et al. 2017; Pillepich et al. 2018a).

Given that this paper focuses on low mass halos  $M_{\text{halo}} < 10^{12} M_{\odot}$ , TNG's treatment of star formation and stellar feedback, particularly the treatment of supernova-driven galactic winds (Pillepich et al. 2018a), is most relevant for our study. In both the TNG and Illustris models, stars form in dense gas regions that meet a threshold for star formation, following a subgrid implementation of the Kennicutt-Schmidt law (Kennicutt 1998). Specifically, the TNG model adopts a two-phase ISM framework, where gas above a density threshold is partitioned into cold, star-forming clouds and hot ionized gas, with the star formation rate scaling as  $\propto \rho^{1.5}$  (Springel & Hernquist 2003).

Building on the original Illustris simulation (Vogelsberger et al. 2013), TNG models supernova-driven galactic winds that are launched isotropically from star-forming gas cells in the dense ISM. The winds are modeled as collisionless particles, with the kinetic energy release rate determined by a mass-loading function that depends on the gas-phase metallicity and the dark matter velocity dispersion,  $\sigma_{DM}$ . This velocity dispersion is calculated using a weighted kernel over the nearest 64 DM particles, and the wind velocity is further modulated by  $\sigma_{DM}$  and redshift. These outflows are powerful enough to eject mass from the ISM (Nelson et al. 2019) and, as we discuss in Section 4, can push gas entirely out of the halo.

Several physical processes in TNG, such as star formation, stellar feedback, black hole seeding, black hole accretion, and AGN feedback, cannot be directly simulated due to resolution limitations and are thus implemented using ‘sub-grid’ recipes. The parameters for these sub-grid processes are calibrated to align with various observations, similar to the methodology used in SAMs. However, the specific observations and the precision of calibration differ. For TNG,

calibration targets include the cosmic star formation rate density as a function of redshift, stellar mass functions, the relationship between black hole mass and stellar mass, the hot gas fraction in galaxy clusters, and the galaxy stellar mass versus radius relation (Pillepich et al. 2018a). Additionally, the distribution of galaxy optical colors in stellar mass bins was used to refine AGN feedback parameters, as detailed in Nelson et al. (2018). While these calibration choices overlap with those used for the SC SAM, described in the section below, there are differences in the specific observations and calibration techniques used.

In this paper, we use both the hydrodynamical (full-physics) and dark matter only (N-body) outputs of the intermediate-sized box ( $L_{\text{box}} = 75 \text{ Mpc h}^{-1}$ ) at its highest resolution, TNG100-1.

### 2.2 The Santa Cruz semi-analytic model

The Santa Cruz SAM for galaxy formation, first introduced in Somerville & Primack (1999), and subsequently updated in Somerville et al. (2008a, 2012), Porter et al. (2014), Popping et al. (2014), and Somerville et al. (2015), traces the evolution of baryons within galaxies by partitioning them into four distinct reservoirs: the cold gas in the galaxy disk, the hot gas associated with the dark matter halo, the ejected gas reservoir, and the stellar component, the latter of which is further divided into disk and bulge populations. To model the mass and energy transfer between these reservoirs, the Santa Cruz SAM incorporates physically and observationally motivated prescriptions for a variety of baryonic processes. These include gas cooling, star formation, stellar feedback, gas recycling, chemical enrichment, and the growth of supermassive black holes and their associated feedback. In this work, we use the latest version of the Santa Cruz SAM published in Somerville et al. (2015) and recently referenced in Gabrielpillai et al. (2022) and Yung et al. (2023). Since we limit our focus to central galaxies with  $M_{\text{halo}} < 10^{12} M_{\odot}$ , we turn AGN feedback “off” in the SAM. Below, we briefly summarize the key processes modeled in the SAM, and refer the reader to Somerville et al. (2008a) and Porter et al. (2014) for details.

#### 2.2.1 Gas Cooling and Accretion

For any given halo, the SAM estimates the dark matter accretion rate,  $\dot{m}_{DM}$ , for each halo by using the virial mass provided by the halo merger tree and calculating its rate of change between consecutive snapshots. Prior to reionization, the SAM assumes that gas accretion  $\dot{m}_{\text{acc}}$  closely follows the accretion of dark matter, scaled by the universal baryon fraction  $f_b$  as:

$$\dot{m}_{\text{acc}} = f_b \dot{m}_{DM}, \quad (1)$$

where  $f_b = 0.1573$  (Planck Collaboration et al. 2016).

Following reionization, the photoionizing background suppresses gas accretion into halos, with the fraction of baryons able to collapse determined by  $f_{\text{coll}}$ , a function of halo mass and redshift (Okamoto et al. 2008). The SAM also accounts for the recycling of gas previously expelled by stellar feedback, described in detail in Section 2.2.2. The rate of gas re-accretion is given by:

$$\dot{M}_{\text{CGM,in,recycled}} = \chi_{\text{re}} - \text{infall} \left( \frac{M_{\text{ejected}}}{t_{\text{dyn}}} \right), \quad (2)$$

where  $M_{\text{ejected}}$  represents the total mass of gas in the ejected reservoir, and  $\chi_{\text{re-infall}}$  specifies the fraction of this gas that can return to the

hot halo at each time step. Typically,  $\chi_{\text{re-infall}} = 0.1$ , meaning that the ejected gas is recycled over  $\sim 10$  dynamical times. The dynamical time is defined as  $t_{\text{dyn}} \equiv R_{\text{vir}}/V_{\text{vir}}$ , where  $V_{\text{vir}}$  is the circular velocity of the halo at the virial radius. Additionally, when a halo becomes a subhalo, the SAM assumes its CGM is immediately transferred to the host halo, adding to the host's hot gas reservoir.

Following the framework of White & Frenk (1991), the SAM assumes that the CGM is uniformly at the virial temperature of the halo at each time step. The radiative cooling time, which determines how quickly the gas loses thermal energy via radiation, is calculated using the cooling function  $\Lambda(T_{\text{vir}}, Z_h)$  (Sutherland & Dopita 1993), where  $T_{\text{vir}}$  is the virial temperature and  $Z_h$  is the metallicity of the hot halo gas. The gas density profile is modeled as a singular isothermal sphere, and the cooled gas mass within the radius  $r_{\text{cool}}$ —the radius within which all gas can cool within the cooling time  $t_{\text{cool}}$ —determines the ISM mass accretion rate.

If  $r_{\text{cool}}$  is smaller than the virial radius  $r_{\text{vir}}$ , the SAM applies a standard cooling flow model, with  $t_{\text{cool}}$  calculated following Somerville et al. (2008a) under the assumption of an isothermal and isobaric gas density profile. When  $r_{\text{cool}}$  exceeds  $r_{\text{vir}}$ , the cooling rate matches the gas accretion rate into the halo, representing “cold/fast/filamentary” mode accretion. In these cases, the SAM ignores radiative cooling predictions and sets the ISM accretion rate equal to the halo gas accretion rate. Variations in this cooling model, such as adjustments to the cooling time definition or changes to the gas density profile, can impact the ISM accretion rate by a factor of 2–3 (Somerville et al. 2008b).

### 2.2.2 Star Formation and Stellar Feedback

All star formation recipes available for use in the SAM generally follow the form given by:

$$m_* = \frac{m_{\text{cold}}}{\tau_*}, \quad (3)$$

where  $\tau_*$  is an adjustable efficiency parameter, typically a constant. In the most recently used Santa Cruz SAM, gas accreted into the ISM is partitioned into H I, H<sub>2</sub>, H II, and metals, with their respective mass surface densities tracked in radial disk annuli (Popping et al. 2014; Somerville & Davé 2015). The default prescription for the star formation rate (SFR) surface density is based on the molecular hydrogen gas phase alone, accounting for a higher conversion efficiency above a critical H<sub>2</sub> surface density (Bigiel et al. 2008; Narayanan et al. 2012):

$$\dot{\Sigma}_{\text{SFR}} = A_{\text{SF}} \left( \frac{\Sigma_{H_2}}{10 M_\odot \text{pc}^{-2}} \right) \left( 1 + \frac{\Sigma_{H_2}}{\Sigma_{H_2, \text{crit}}} \right)^N_{\text{SF}}, \quad (4)$$

where  $A_{\text{SF}}$ ,  $N_{\text{SF}}$  and  $\Sigma_{H_2, \text{crit}}$  are free parameters. We set  $A_{\text{SF}} = 5.98 \times 10^{-3} M_\odot \text{yr}^{-1} \text{kpc}^{-2}$ ,  $N_{\text{SF}} = 1.0$  and  $\Sigma_{H_2, \text{crit}} = 70 M_\odot \text{pc}^{-2}$ , following Popping et al. (2014, 2019). In the SAM, the default setting for estimating the molecular hydrogen gas density  $\Sigma_{H_2}$  uses the metallicity-dependent partitioning approach of (Gnedin & Kravtsov 2011).

In the SAM, stellar feedback ejects cold gas from the ISM, with the ejection rate modeled as a power law:

$$\dot{m}_{\text{eject}} = \epsilon_{\text{SN}} \left( \frac{200 \text{km/s}}{V_{\text{disk}}} \right)^{\alpha_{\text{rh}}} \dot{m}_*, \quad (5)$$

where the free parameters  $\epsilon_{\text{SN}}$  and  $\alpha_{\text{rh}}$  modulate the efficiency and dependence on the disk's circular velocity  $V_{\text{disk}}$ , and  $\dot{m}_*$  is the star

formation rate.  $V_{\text{disk}}$  is approximated as the circular velocity of the uncontracted halo at twice the Navarro-Frenk-White (NFW) scale radius  $r_s$  (Navarro et al. 1996). All ejected gas is expelled from the halo entirely and deposited in the “ejected” reservoir. A fraction of this gas is allowed to return to the hot halo on dynamical timescales, as described in Section 2.2.1.

### 2.2.3 Metal Production

The Santa Cruz SAM models metal production using the instantaneous recycling approximation, a commonly employed approach in SAMs. When a mass of stars,  $dm_*$ , forms, it generates a corresponding mass of metals,  $dM_Z = y \cdot dm_*$ , which are instantly mixed with the cold gas in the disk. The yield,  $y$ , is typically calibrated to match observational data, such as the normalization of the stellar metallicity–mass relation from Gallazzi et al. 2005, as shown in Somerville et al. 2015. Unlike most SAMs, which derive yields from stellar evolution models tied to observational constraints, the Santa Cruz SAM adopts the simple calibration approach above given the crude nature of the instantaneous recycling approximation.

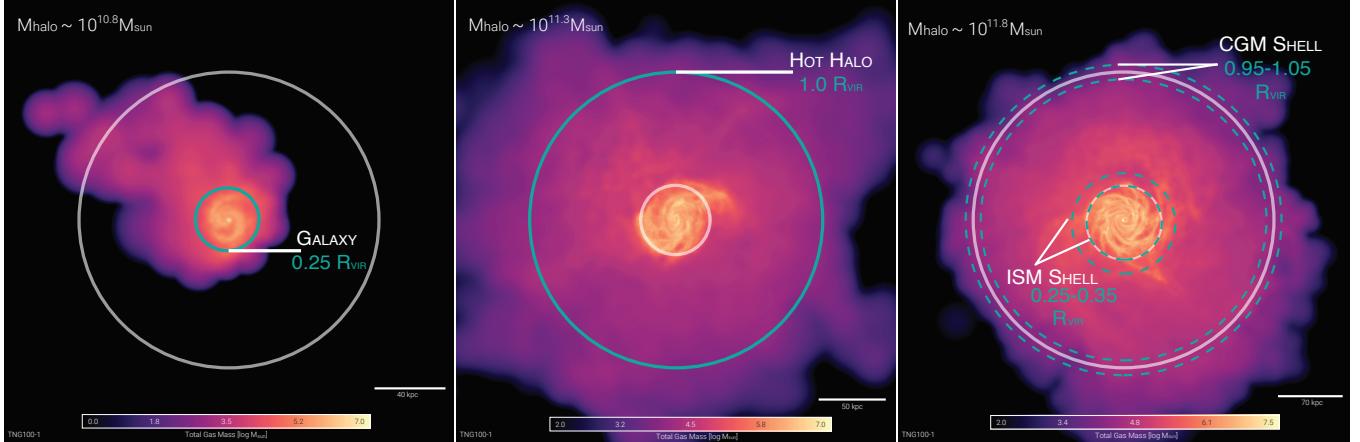
The metallicity of the cold gas,  $Z_{\text{cold}}$ , evolves over time, with newly formed stars inheriting this metallicity. Supernova feedback drives both gas and metals out of the disk, transferring them to the hot halo or ejecting them into an external reservoir at rates proportional to the mass of outflowing gas. These ejected metals are not lost permanently; they can be re-accreted into the halo, contributing to ongoing chemical enrichment over time.

### 2.2.4 Observational Calibration

The SC SAM is calibrated by adjusting free parameters to match key observations at  $z \sim 0$ , such as the stellar-to-halo mass relation, stellar mass function, stellar mass-metallicity relation, cold gas fraction versus stellar mass for disk-dominated galaxies, and the black hole mass vs. bulge mass relation. Observational constraints for these calibrations are sourced from Rodriguez-Puebla et al. (2017), Bernardi et al. (2013), Gallazzi et al. (2005), Peebles et al. (2014), Calette et al. (2018), and McConnell & Ma (2013). Additional cross-checks involve the cold gas phase mass-metallicity relation and the H<sub>2</sub> mass function, using constraints from Obreschkow et al. (2009), Keres et al. (2003), Anderson et al. (2014), Zahid et al. (2013), and Boselli et al. (2014). The calibration considers the sensitivity of gas fraction, stellar metallicity, and star formation efficiency, fine-tuning the parameters to balance these observational constraints while adjusting for the effects of AGN and SNe feedback to match both high-mass and low-mass galaxy populations. For a more detailed description of the observations used for calibration, refer to Appendix B in Yung et al. (2023).

### 2.2.5 New Modifications made to the Santa Cruz SAM

The version of the Santa Cruz SAM presented in this work revises how the SAM computes the rate of gas condensation from the CGM into the ISM. When  $r_{\text{cool}} > r_{\text{vir}}$ , the original Santa Cruz SAM disregards the radiative cooling prediction and equates the ISM accretion rate to the halo gas accretion rate. This approach, however, leads to unrealistically low hot halo gas masses in dwarf galaxies (Pandya et al. 2020). To rectify this issue, we follow Pandya (2021) and adopt a modified approach inspired by Guo et al. (2011): when  $r_{\text{cool}} > r_{\text{vir}}$ ,



**Figure 1.** Three central subhalos (IDs: 715660, 909911, 557236) from TNG100, spanning the  $10^{10.5} - 10^{12} M_{\odot}$  mass range analyzed in this work, are shown to demonstrate how TNG’s outputs are compared to the SAM. The maps are color-coded by total gas mass, with brighter regions indicating higher mass concentrations. The “galaxy” is defined as all material within  $0.25 r_{\text{vir}}$  (left panel, green circle), and the “halo” as all material within  $r_{\text{vir}}$  (center panel, green circle). To analyze the flow of baryons at the galactic and halo scales, we examine inflow and outflow rates within the boundaries highlighted by dashed green circles in the right panel:  $0.2-0.3 r_{\text{vir}}$  for the galaxy and  $0.95-1.05 r_{\text{vir}}$  for the halo.

the rate of cooling is limited solely by the freefall time of the hot halo gas, such that:

$$\dot{M}_{\text{ISM}}^{\text{in}} = \begin{cases} \frac{M_{\text{CGM}}}{t_{\text{dyn}}} \frac{r_{\text{cool}}}{r_{\text{vir}}} & \text{when } r_{\text{cool}} < r_{\text{vir}} \\ \frac{M_{\text{CGM}}}{t_{\text{dyn}}} & \text{when } r_{\text{cool}} \geq r_{\text{vir}}. \end{cases} \quad (6)$$

While this modification no longer predicts extremely low CGM masses in dwarf galaxies (see Figure A1), it has its own limitations. When  $r_{\text{cool}} > r_{\text{vir}}$ , the model assumes that gas accretes onto the galaxy solely at the freefall rate of the hot halo gas. However, this approach may underestimate the total gas inflow, as it ignores contributions from gas external to the halo, such as cold streams or filaments from the IGM. In these cases, the freefall-limited rate represents a lower bound on the actual accretion rate. A more complete model should incorporate contributions from both cold and hot accretion modes.

### 3 COMPARING THE OUTPUT BETWEEN TNG AND THE SAM

#### 3.1 Matching Galaxies Between TNG FP and SC SAM Run on TNG-DMO

The precise identification of halos and construction of merger trees forms the backbone of all SAMs. Thus, to facilitate a direct comparison between the Santa Cruz SAM and TNG, we must use reasonably similar merger trees. The SC SAM employs the ROCKSTAR halo finder and CONSISTENT TREES, a gravitationally consistent merger tree algorithm developed by Behroozi et al. (2012), to extract merger trees from the Bolshoi cosmological N-body simulation (Klypin et al. 2011). TNG, however, utilizes a different halo finder based on the friends-of-friends (FoF) algorithm (Davis et al. 1985) to identify “groups,” and the SUBFIND algorithm (Springel et al. 2001) to identify substructure.

To create bijective matches between the two sets of halo catalogs, Gabrielpillai et al. (2022) ran the ROCKSTAR halo finder and CONSISTENT TREES on TNG100-1-DM. They then matched halos identified by ROCKSTAR and SUBFIND catalogs with SubLink, a software tool developed by Rodriguez-Gomez et al. (2015) that uses a merit function for matching subhalos. This produced successful bijective matches between the ROCKSTAR and FoF/SUBFIND halos

for central galaxies, achieving a 99% match rate at  $z = 0$ . The fraction of bijective matches for non-central subhalos, however, is often substantially lower, with a match rate from 50–70%. As a result, we limit our analysis to central galaxies only.

Although 99% of central galaxies have bijective matches, a direct comparison between the halo mass of the central galaxies between the two catalogs, however, reveals slight offsets up to 20% for low mass halos (see Figure 2 in Gabrielpillai et al. (2022)). These discrepancies are likely primarily influenced by baryonic physics rather than differences in halo-finding algorithms. For a more detailed description of the halo finder, merger tree algorithms, and method used to create these bijective matches, we refer the reader to Section 3 of Gabrielpillai et al. (2022).

#### 3.2 Selecting Comparable Galaxy and Halo Properties

To make reliable comparisons between the outputs of TNG and the SAM, we identify comparable quantities at both the galaxy and halo scales to minimize potential biases in interpretation. Previous comparisons of SAMs and hydrodynamical simulations, discussed in detail in Section 6.3, were not always conducted systematically, which may have led to inconsistencies and misinterpretations of key differences. For instance, comparing the stellar mass within the stellar half-mass radius in a hydrodynamical simulation to the total stellar mass reported by a SAM can lead to misleading conclusions if the SAM does not provide stellar mass measurements within the same radius.

In our approach, illustrated in the left and middle panels of Figure 1, we define the galaxy as all material within  $0.25 r_{\text{vir}}$  and the halo as all material within  $r_{\text{vir}}$ . Properties at the “halo scale” include the virial mass ( $M_{\text{halo}}$ ), the mass of the hot halo ( $M_{\text{hot}}$ ), the metallicity of the gas in the hot halo ( $Z_{\text{hot}}$ ), and the overall baryon fraction ( $f_{\text{baryon}}$ ). Properties at the “galaxy scale” describe the stellar mass ( $M_{\star}$ ), cold gas mass ( $M_{\text{cold}}$ ), their respective metallicities ( $Z_{\star}$  and  $Z_{\text{cold}}$ ), and the star formation rate (SFR) of the central galaxy within the halo. We define the baryon fraction as the sum of  $M_{\text{cold}}$ ,  $M_{\star}$ , and  $M_{\text{hot}}$  divided by the total halo mass. Although the SC SAM stores mass in an ejected reservoir as well, we do not count this mass towards the overall baryon fraction.

In TNG, we define the CGM as all gas within the FoF halo, excluding the gas contribution from each galaxy within the same halo.

The galaxy is defined as all matter within twice the stellar half-mass radius (SHMR). In the SAM, the CGM consists of gas accreted into the halo and ejected from galaxies, making it directly comparable to TNG’s definition. However, the definition of the galaxy, particularly its cold gas component ( $M_{\text{cold}}$ ), is less straightforward. In the SAM, cold gas is explicitly tracked via gas partitioning, whereas in TNG, this mass is not directly provided within twice the SHMR. For consistency, we define  $M_{\text{cold}}$  in both TNG and the SAM as the total gas mass of the galaxy.

To compare these properties effectively, we require that all TNG halos be adequately resolved, given the finite mass resolution of the simulation. The mass resolution in the SAM is determined by the resolution of the N-body simulation used to extract the merger trees. [Gabriellipillai et al. 2022](#) adopt the general guideline that a halo should contain at least 100 particles to be well resolved, and that the root halo should be at least 100 times the mass of the smallest progenitor that can be resolved in order to accurately resolve the halo’s merger history. For TNG100-1-Dark, the mass resolution is  $8.9 \times 10^6 M_{\odot}$ , leading [Gabriellipillai et al. 2022](#) to adopt a conservative minimum root halo mass of approximately  $8.9 \times 10^8 M_{\odot}$ . We adopt the same framework in this paper.

For the hydrodynamical TNG100-1, the target baryon mass is  $1.4 \times 10^6 M_{\odot}$ , meaning that the mass of gas in each cell is maintained within a factor of two of this value, as is the mass of star particles at birth ([Pillepich et al. 2018b](#)). We restrict our analysis to TNG galaxies that have a minimum of approximately 100 stellar particles  $M_{\text{star}} \geq 1.4 \times 10^8 M_{\odot}$ , and 100 gas particles  $M_{\text{gas}} \geq 1.4 \times 10^8 M_{\odot}$ . Our minimum requirement for the gas particles naturally sets a lower limit on the SFR. Based on the analysis by [Donnari et al. \(2019\)](#), we adopt a minimum SFR of  $1.0 \times 10^{-3} M_{\odot} \text{ yr}^{-1}$ . Applying this filter allows us to resolve 19,650 galaxies with halo masses ranging from  $\sim 4 \times 10^{10} - 10^{12} M_{\odot}$  at  $z = 0$ .

### 3.2.1 Flow Cycle Measurements in TNG

To establish a common language for describing the baryon cycle between TNG and the SC SAM, it is essential to carefully analyze how gas flows in TNG. TNG, however, does not directly provide measurements of the rates of gas inflow and outflow within and around the galaxy. To address this, Cohen et al. (in preparation, hereafter CIP), have extracted these measurements for a randomly selected subset of  $\sim 9,522$  central galaxies in TNG100-1 across 10 redshifts ranging from  $z = 0$  to  $z = 10$ . At each redshift, CIP selected halos in bins of 0.3 dex in virial mass, spanning  $M_{\text{vir}} = 10^{10} M_{\odot}$  to the most massive halo at that redshift.

Using an Eulerian approach, CIP calculated instantaneous gas flows by defining volumetric shells of thickness  $\Delta r \sim 0.05 r_{\text{vir}}$  spanning  $0.05r_{\text{vir}}$  to  $2r_{\text{vir}}$  and tracking the mass and velocity of gas crossing these boundaries between consecutive simulation outputs. For further details on the methodology and the specific tools and algorithms used to achieve these measurements, we refer the reader to CIP.

To ensure reliability, we divided the CIP “flow sample” into “resolved” vs. “unresolved” categories. After applying the resolution criteria described above, the resolved subsample comprises  $\sim 400$  galaxies at each redshift. The remaining galaxies are categorized as unresolved. For the galaxies in each subsample, we define the halo boundary as the region between  $0.95 - 1.05 r_{\text{vir}}$ , and the galaxy boundary from  $0.25 - 0.35 r_{\text{vir}}$ . Both boundaries are illustrated as dashed circles in the right panel of Figure 1. We volumetrically sum the inflow and outflow rates within each boundary, providing a basis for comparing gas flows in TNG to those in the SAM.

## 4 TRANSLATING TNG’S DETAILED PHYSICS INTO THE SC SAM’S GALAXY FORMATION MODEL

To translate the complex physics of a hydrodynamical simulation like TNG into the simplified framework of the SC SAM, we focus on modeling the baryon cycle—how gas and metals circulate between the galaxy and its surrounding halo—over time. We achieve this using the CIP sample of TNG galaxies described in Section 3.2.1, which provides inflow and outflow rates of gas and metals for each galaxy. Treating this subset as representative of the entire TNG central galaxy population, we derived physical parameters that govern the evolution of these baryon flows, namely halo (re)-accretion efficiency, cooling time, and mass and metal loading. For galaxy sizes, we use the full sample of  $\sim 20,000$  TNG galaxies. These parameters, shown in Figure 3 as functions of halo mass and redshift, form the foundation for calibrating the SC SAM to capture the baryon cycling processes observed in TNG.

To implement the scaling relations in the SAM, we adopt a two-pronged approach that leverages both the “resolved” and “unresolved” populations of the CIP sample. For halos within the well-resolved mass range of  $10.5 < \log M_{\text{halo}} < 12$ , we directly apply the median scaling relations derived from the resolved sample, as shown by the solid lines in Figure 3. When halos fall below this mass range, we extrapolate the scaling relations to lower-mass halos ( $\log M_{\text{halo}} < 10.5$ ) by analyzing trends in the complete CIP sample, which includes both resolved and unresolved populations down to  $\log M_{\text{halo}} = 10$ . Specifically, we analyze the behavior of halos in the range  $10 < \log M_{\text{halo}} < 10.5$  and use the resulting relations (shown as dashed lines in Figure 3) to model all halos with  $\log M_{\text{halo}} < 10.5$ . This approach ensures that the SAM can capture the baryon cycle for galaxies in unresolved halos, even though these lower-mass halos are not directly provided by the CIP flow sample. This extrapolation, however, introduces significant uncertainties in the predicted evolution of lower-mass halos. We discuss the implications and limitations of this method in greater detail in Section 6.2.

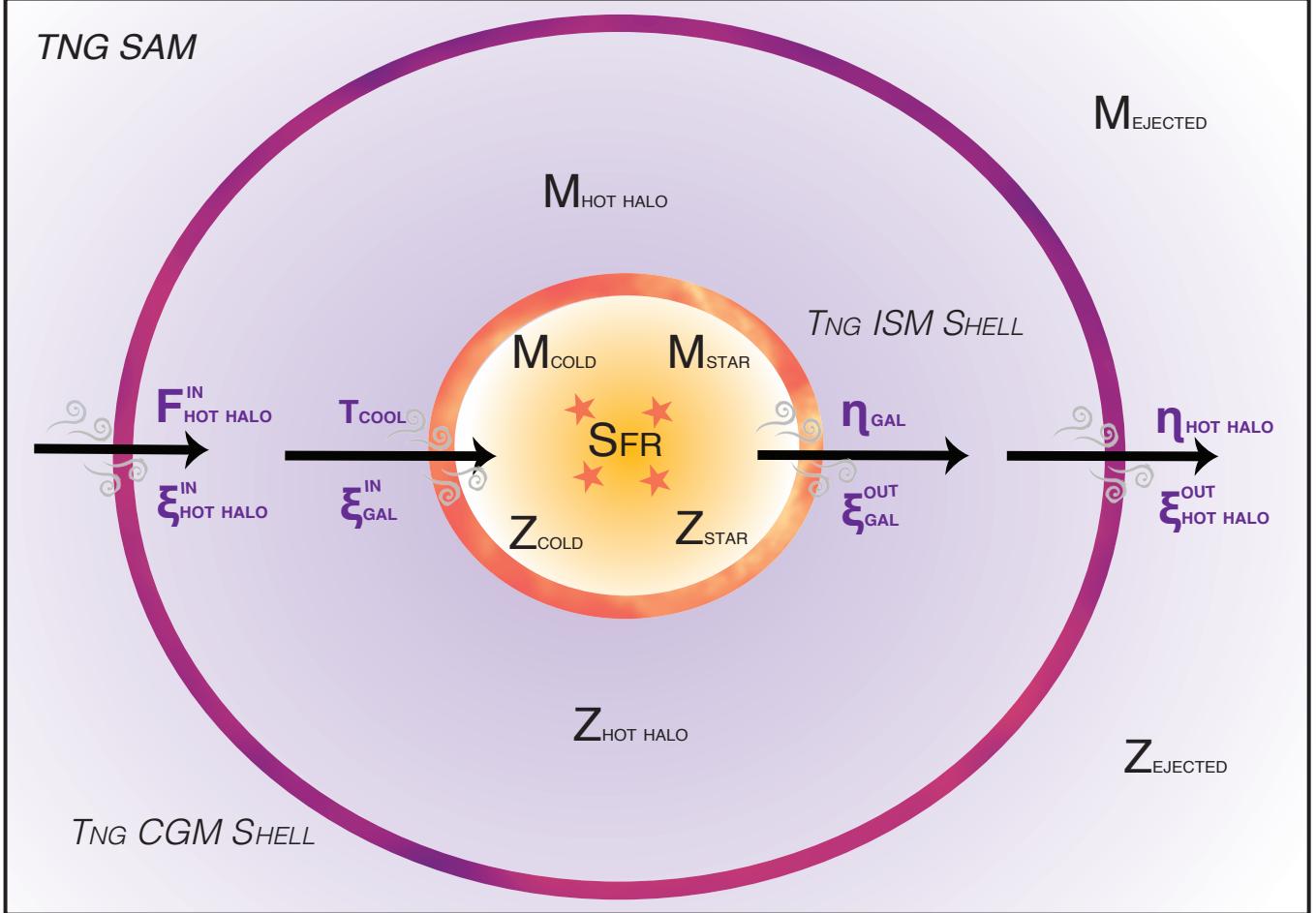
The newly calibrated SAM applies these scaling relations across halo mass and time, performing linear interpolation to determine the relations at intermediary halo masses and redshifts. While the physical processes governing gas flows are not expected to fundamentally change with redshift, TNG shows that the behavior clearly evolves over time, likely driven by underlying physical mechanisms that remain poorly understood. We bundle this ignorance into our interpolation approach, providing a pragmatic approximation of TNG’s redshift-dependent behavior. We also provide analytic functions of halo mass and redshift that capture each relation within 30% accuracy in Table A1. This approach, instead of simply directly inputting the median flow rates from TNG, enables the SAM to compute flow rates in a self-consistent manner according to the provided physical parameters, creating a more physically motivated and flexible SAM. We refer to this newly calibrated model, illustrated in Figure 2, as the ‘TNG SAM.’

### 4.1 Halo Accretion and Cooling

#### 4.1.1 IGM - CGM Accretion and Recycling

In the traditional SAM framework, the rate at which gas accretes onto the CGM is tightly coupled to the dark matter accretion rate, as detailed in Section 2.2.1. In the TNG SAM, we introduce the parameter  $f_{\text{in,CGM}}$  to regulate total gas inflow into the CGM:

$$f_{\text{in,CGM}} \equiv \dot{M}_{\text{in,CGM}} / (f_b \dot{M}_{\text{halo}}), \quad (7)$$



**Figure 2.** Illustration of the TNG SAM, calibrated to reproduce the baryon cycle of  $\sim 400$  central galaxies in the TNG100 simulation. The diagram shows the main baryonic reservoirs—cold gas ( $M_{\text{cold}}$ ), stars ( $M_{\text{star}}$ ), hot halo gas ( $M_{\text{hot halo}}$ ), and ejected gas ( $M_{\text{ejected}}$ )—and the flows of gas and metals between them. Arrows represent baryon flow directions, with governing parameters labeled in purple (e.g.,  $t_{\text{cool}}$  regulates cooling from the hot halo to the ISM, while  $\eta_{\text{ISM}}$  and  $\eta_{\text{halo}}$  control gas outflows). Metallicity enrichment is tracked for each reservoir as  $Z_{\text{cold}}$ ,  $Z_{\text{star}}$ ,  $Z_{\text{hot halo}}$ , and  $Z_{\text{ejected}}$ . Each parameter was calibrated using inflow and outflow rates measured at boundaries defined in Figure 1.

where  $f_{\text{in},\text{CGM}}$  measures the efficiency of gas accretion relative to the expected baryonic inflow, without distinguishing between pristine and recycled gas. Here,  $\dot{M}_{\text{in},\text{CGM}}$  is the gas inflow rate into the CGM,  $f_b$  is the cosmic baryon fraction, and  $\dot{M}_{\text{halo}}$  is the net mass accretion rate of the halo in TNG. We use the *net* accretion rate, rather than the total inflow rate, because it aligns with how the SAM calculates  $\dot{M}_{\text{halo}}$ . Specifically, the SAM estimates halo growth by finite differencing the total halo mass between consecutive timesteps, capturing the *net* change in halo mass, which includes both smooth accretion and mergers.

When  $f_{\text{in},\text{CGM}} = 1$ , the inflow rate matches the expected baryon-to-dark matter ratio. Values below 1 indicate reduced inflow, while values above 1 suggest that gas recycling contributes to enhanced inflow. The top left panel of Figure 3 shows that in TNG,  $f_{\text{in},\text{CGM}}$  fluctuates at or above unity across halo mass and redshift, indicating that the gas inflow generally meets or exceeds the expected baryonic fraction. We further explore the implications of this behavior in Section 6.1.1.

#### 4.1.2 CGM-ISM Cooling

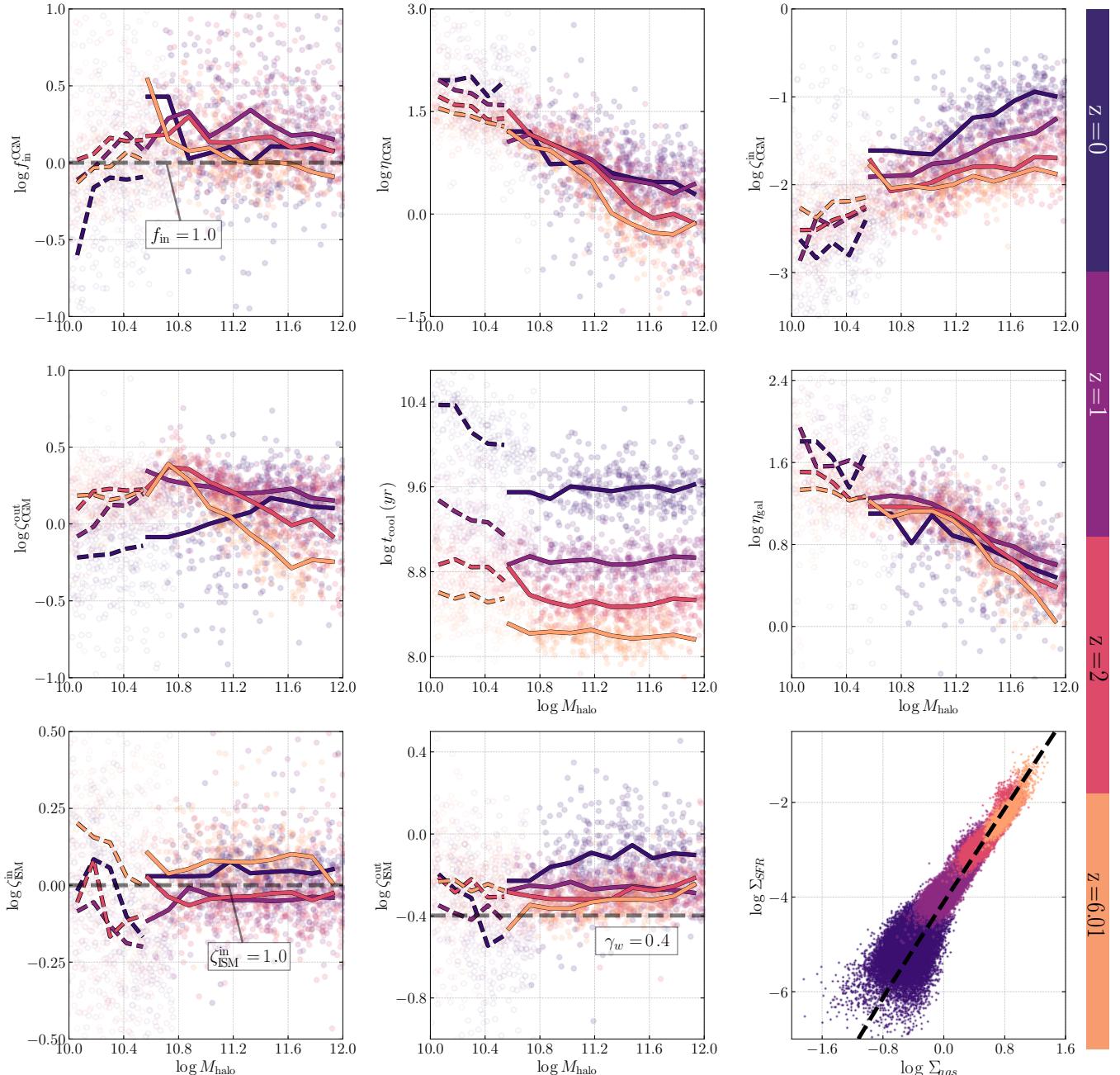
At the scale of the galaxy, the SC SAM, like most conventional SAMs, relies on the cooling radius  $r_{\text{cool}}$  to determine the rate at which gas cools into the ISM, as described in Section 2.2.1. In the

traditional model, the “cooling time”  $t_{\text{cool}}$ , i.e. the time needed for the gas to cool, is calculated assuming an isothermal, isobaric gas density profile. In “cold” mode accretion ( $r_{\text{cool}} > r_{\text{vir}}$ ), gas cools rapidly and falls directly into the galaxy, so the cooling time scales with the dynamical time. Conversely, in “hot” mode accretion ( $r_{\text{cool}} < r_{\text{vir}}$ ), the hot halo gas is shock-heated and the cooling time is significantly longer.

In TNG, we define the cooling time as

$$t_{\text{cool}} \equiv M_{\text{hot}} / \dot{M}_{\text{cool,CGM-ISM}}, \quad (8)$$

where  $\dot{M}_{\text{cool,CGM-ISM}}$  describes the rate that gas cools from the hot halo into the ISM. The left middle panel of Figure 3 shows the resulting cooling times as a function of halo mass and redshift. We find that the cooling times consistently exceed the dynamical time across all halo masses explored, implying that the “cold” vs. “hot” mode dichotomy does not adequately describe the rate of cooling in TNG, discussed further in Section 6.1.2. In the new TNG SAM, we use the cooling times observed in TNG to determine the rate at which gas accretes onto the galaxy by rearranging Equation 8.



**Figure 3.** Overview of the physical parameters measured from the subset of  $\sim 400$  TNG100 galaxies described in Section 3.2.1. Each galaxy is represented as a circle, with colors indicating redshift ( $z = 0, 1, 2, 6$ ); unresolved galaxies are shown with faded colors. Median values are plotted as solid lines for the resolved sample and dashed lines for the unresolved sample, as a function of halo mass and redshift. These median relations are directly input into the TNG SAM. Analytic functions for each parameter, accurate to within  $\sim 30\%$ , are provided in Appendix A3.

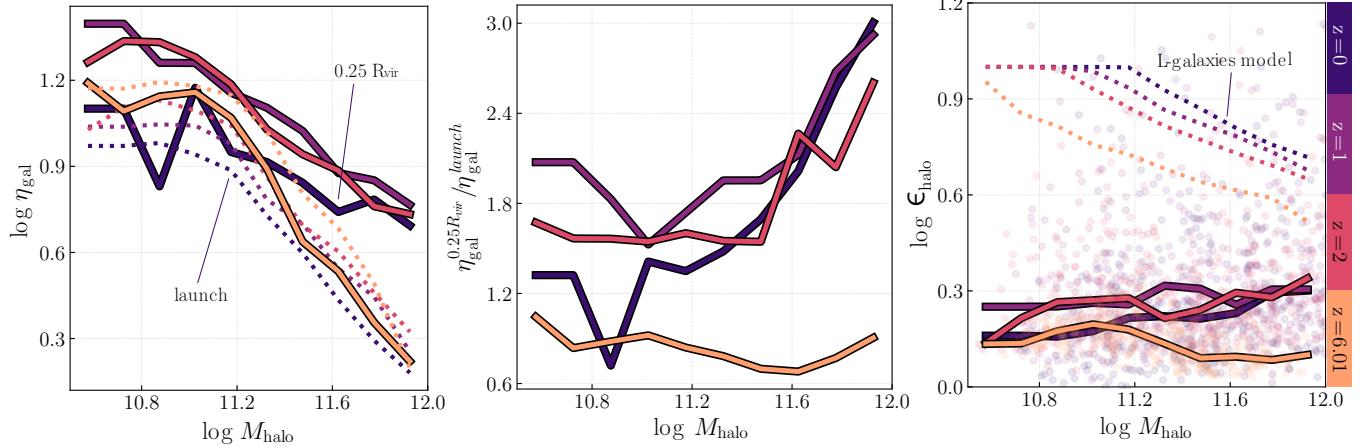
## 4.2 Star Formation

In TNG, star formation occurs below the mass resolution and is therefore treated using a subgrid model. Specifically, the TNG model adopts the two-phase ISM model of Springel & Hernquist (2003), where gas cells above the density threshold  $\Sigma_{\text{gas}}$  are split into cold star-forming clouds and hot ionized gas. Star formation is computed in equilibrium, with the SFR proportional to the cold gas density and scaled by an adjustable free-fall time  $t_{*,0}$ , which is tuned to reproduce the global observed Kennicutt-Schmidt relation, given by Kennicutt (1998):

$$\Sigma_{\text{SFR}} = (2.5 \pm 0.7) \times 10^{-4} \left( \frac{\Sigma_{\text{gas}}}{M_{\odot}, \text{pc}^{-2}} \right)^{1.4 \pm 0.15} M_{\odot}, \text{yr}^{-1}, \text{kpc}^{-2}. \quad (9)$$

Instead of the power law index  $n = 1.4$ , the TNG model applies a scaling of  $\text{SFR} \propto \rho^{n=1.5}$  such that the hydrodynamics roughly reproduce the observed KS relation across the global scale of individual galaxies.

In the new TNG SAM, we also implement the KS relation, but at the scale of the entire galaxy, and with the goal of reproducing the aggregate Kennicutt Schmidt relation of TNG galaxies. Consistent



**Figure 4.** How the TNG SAM models stellar feedback-driven outflows at the galaxy (left panels) and halo (right panel) scales across redshift. The left panel shows that  $\eta_{\text{ISM}}$  (solid lines) almost always exceeds the launch value from star-forming gas cells in TNG (dotted lines), suggesting additional mass entrainment as winds reach the galaxy’s radius. To account for this, the TNG SAM scales the launch value by the ratio of galaxy-scale to launch-scale values, shown in the middle panel. At the halo scale,  $\eta_{\text{halo}}$  is modeled using the dimensionless parameter  $\epsilon_{\text{halo}}$ . The TNG SAM replaces the far stronger  $\epsilon_{\text{halo}}$  values from L-galaxies (dotted lines) with values directly derived from TNG (solid lines) using  $\eta_{\text{halo}}$  and  $\eta_{\text{ISM}}$ .

with previous versions of the Santa Cruz SAM (Somerville et al. 2008a, 2012, 2015), the KS relation is applied to the total surface density of cold neutral gas, including both atomic and molecular components. Star formation only occurs when the gas surface density exceeds a critical threshold, after which the SFR density is calculated using Equation 9.

The KS relation’s dependence on gas surface density, however, necessitates consistent definitions of disk sizes in the SAM and TNG. To address this, we adopt a fixed galaxy radius of  $0.25 r_{\text{vir}}$  in both models, determined by identifying the radius where the change in stellar mass becomes negligible, and the disk gas content is fully captured. In TNG, we find that  $\chi_{\text{gas}}$ , the gas-to-total-mass fraction within this radius, is approximately 1.7, which is typical for SAMs. With disk sizes aligned, we reproduce the median global KS relation in TNG within 10% after setting the slope to a much higher value of  $n = 2.0$ , the normalization parameter  $A_{\text{SF}} = 1.66667 \times 10^{-4}$  and  $\Sigma_0 = 0.01$ , as shown by the dashed line in the lower right panel of Figure 3.

### 4.3 Stellar Feedback

#### 4.3.1 Galactic Winds and Galaxy-scale Outflows

To emulate TNG’s treatment of stellar feedback, the TNG SAM incorporates the galactic wind model from (Pillepich et al. 2018b). Inspired by the results of semi-analytic models, the TNG model launches wind particles with an initial speed that scales with the local dark matter velocity dispersion  $\sigma_{\text{DM}}$ . It also includes a redshift-dependent factor and a minimum wind velocity, such that the wind velocity  $v_w$  scales as:

$$v_w = \max \left[ \kappa_w \sigma_{\text{DM}} \left( \frac{H_0}{H(z)} \right)^{1/3}, v_{w,\min} \right] \quad (10)$$

where  $\kappa_w = 7.4$  and  $v_{w,\min} = 350 \text{ km s}^{-1}$  in the fiducial TNG model. This form ensures a physically plausible scaling of wind velocities with redshift and halo mass, while preventing unphysically large mass loading factors in low-mass halos.

The TNG model further refines wind generation by linking the wind mass loading factor  $\eta_w$  to the specific energy available for wind generation  $e_w$ . This energy is primarily determined by the energy released from Type II supernovae, with a fraction ( $\tau_w$ ) being

thermal. The wind energy itself is a function of the metallicity of the star-forming gas, leading to a metallicity-dependent galactic wind model.

The wind mass loading factor, defined as the ratio of the mass outflow rate  $\dot{m}_{\text{ISM}}$ , out to the star formation rate  $\dot{m}_*$ , quantifies the efficiency of these winds in removing gas from the ISM.  $\eta_w$  is related to  $v_w$  and the wind energy,  $e_w$ , as:

$$\eta_w = \frac{2}{v_w^2} e_w (1 - \tau_w) \quad (11)$$

where  $e_w$  is given by:

$$e_w = \bar{e}_w \left[ f_{w,Z} + \frac{1 - f_{w,Z}}{1 + (Z/Z_{w,\text{ref}})^{\gamma_{w,Z}}} \right] \times N_{\text{SNII}} E_{\text{SNII},51} 10^{51} \text{ erg M}_{\odot}^{-1} \quad (12)$$

Here,  $\bar{e}_w$  is the average energy per unit stellar mass formed,  $f_{w,Z}$  modulates the metallicity dependence,  $Z$  represents the gas metallicity,  $Z_{w,\text{ref}}$  is a reference metallicity,  $\gamma_{w,Z}$  controls the strength of the metallicity dependence,  $N_{\text{SNII}}$  denotes the number of Type II supernovae, and  $E_{\text{SNII},51}$  is the energy released per supernova in units of  $10^{51}$  erg.

In the TNG model,  $\eta_w$  is implemented at the scale of the launched wind particle. To implement this model within the TNG SAM, we instead apply  $\eta_w$  at  $R_{\text{disk}}$ . However, the left panel of Figure 4 shows that the value of  $\eta$  at the scale of the galaxy (solid lines) almost consistently exceeds the value at launch (dashed lines), indicating that the winds entrain additional material as they propagate outwards. To account for this in the SAM, we model the ratio of  $\eta_w$  at the site of the launched wind particle to  $\eta_w$  at the scale of the galaxy in TNG, shown in the middle panel of Figure 3. Since the TNG SAM applies  $\eta_w$  at the scale of the disk, the material ejected from the ISM is always deposited in the CGM.

#### 4.3.2 Halo-scale Outflows

In TNG, galactic winds not only expel gas from the galaxy but also entrain additional material from the CGM, ejecting it into the IGM. In the SC SAM, however, stellar feedback drives gas directly from the ISM into an ejected reservoir, bypassing the CGM. The ejected

gas only reconnects with the CGM indirectly, when a portion of it is re-accreted into the hot halo at later times. In the new TNG SAM, we explicitly account for the CGM-IGM outflow channel with the parameter  $\eta_{\text{halo}}$ , defined as the ratio of the mass outflow rate  $\dot{m}_{\text{CGM,out}}$  to the star formation rate  $\dot{m}_*$ . Complementing the existing  $\eta_w$ ,  $\eta_{\text{halo}}$  quantifies the efficiency with which stellar feedback ejects gas from the halo.

The model for  $\eta_{\text{halo}}$  in the TNG SAM is inspired by the halo-scale outflow model used in the L-galaxies model for the Millennium simulations (Henriques et al. (2015); H15 hereafter). In the H15 model, surplus energy from supernovae, after reheating gas within the ISM and driving it into the CGM, can further expel the hot halo gas into the IGM as:

$$\frac{1}{2}\Delta m_{\text{eject}}V_{\text{vir}}^2 = \Delta E_{\text{SN}} - \Delta E_{\text{reheat}}, \quad (13)$$

where  $\Delta E_{\text{reheat}} = \frac{1}{2}\Delta m_{\text{reheat}}V_{\text{vir}}^2$ .

By assuming the quantity  $\Delta m_{\text{eject}}/\Delta m_*$  is equivalent to our  $\eta_{\text{halo}} \equiv \dot{m}_{\text{CGM-IGM}}/\dot{m}_*$ , we derive the expression for  $\eta_{\text{halo}}$  as:

$$\eta_{\text{halo}} = \frac{2\epsilon_{\text{halo}}V_{\text{SN}}^2}{V_{\text{vir}}^2} - \eta_{\text{ISM}} \quad (14)$$

where  $V_{\text{SN}} = 630$  km/s is a constant and  $\epsilon_{\text{halo}}$  is a dimensionless efficiency parameter capped at unity and is given by:

$$\epsilon_{\text{halo}} = \epsilon_{\text{halo},0} \left[ 0.5 + \left( \frac{V_{\text{max}}}{V_{\text{eject}}} \right)^{-\beta_2} \right]. \quad (15)$$

In H15, the best fit values for  $\epsilon_{\text{halo}}$  are  $\epsilon_{\text{halo},0} = 0.62$ ,  $\beta_2 = 0.80$ , and  $V_{\text{eject}} = 100$  km/s. The right panel of Figure 4, however, shows that applying these values in TNG (dashed lines) leads to significantly stronger  $\epsilon_{\text{halo}}$  values than actually observed (solid lines), overestimating the efficiency by up to a factor of 10. In the TNG SAM, we parameterize  $\epsilon_{\text{halo}}$  using the values observed in TNG across halo mass and redshift.

All gas ejected from the halo enters the ejected reservoir and is not allowed to re-accrete. However, in the TNG SAM implementation, the  $f_{\text{in,CGM}}$  parameter accounts for some of this gas being re-accreted, though the reaccretion process itself is not modeled explicitly due to our inability to distinguish between pristine and recycled gas.

#### 4.4 Metal Enrichment and Circulation

In TNG, metals not only trace star formation and chemical enrichment, but also modulate the strength of galactic winds and cooling rates. The TNG model explicitly tracks the gradual release of metals over time from asymptotic giant branch (AGB) stars, core-collapse supernovae (SNII), and Type Ia supernovae (SNIa). This allows the simulation to trace the redistribution of metals within the ISM and CGM. In contrast, the SC SAM adopts a simplified, instantaneous model for metal production and circulation, detailed in Section 2.2.3.

To account for the detailed metal circulation observed in TNG, we introduce metallicity-weighted mass-loading factors (or ‘‘metal-loadings’’) into the TNG SAM. These factors quantify the fraction of metals transported by gas flows relative to the total gas flow. Following Peeples & Shankar (2011), we compute the metal loadings to quantify the fraction of metals transported out of the ISM/CGM as follows:

$$\zeta_{\text{out}} = \frac{\dot{M}_{Z,\text{out}}}{\dot{M}_{\text{out}} \cdot Z_{\text{gas}}}, \quad (16)$$

where  $\dot{M}_{Z,\text{out}}$  is the rate at which metals are outflowing from the ISM/CGM,  $\dot{M}_{\text{out,ISM/CGM}}$  is the rate at which all gas is outflowing from the ISM/CGM, and  $Z_{\text{gas}}$  describes the metallicity of the gas reservoir from which the outflows stem.

Traditionally, the SC SAM assumed  $\zeta_{\text{out,ISM/CGM}} = 1$ , meaning that the metal component of the outflowing gas was directly proportional to the metallicity of the ISM/CGM. However, the two lower right panels in Figure 3 show that this assumption is not valid in TNG. At the CGM scale,  $\zeta_{\text{out,CGM}}$  often deviates from unity, decreasing significantly at higher redshifts. At the ISM scale,  $\zeta_{\text{out,ISM}}$  consistently falls below unity, indicating that the winds eject fewer metals relative to the average metallicity of the ISM. This result is not too surprising, for the fiducial TNG model sets the wind metal loading factor  $\gamma_w = 0.4$  (Vogelsberger et al. 2013), indicating that the winds carry metals with a fraction of the metallicity of the ISM. However, as the winds reach the scale of the galaxy, they entrain additional metal mass, as evidenced by the fraction of metals outflowing consistently being greater than 0.4 (the grey dashed line in the center right panel).

In addition to calibrating outflow metal loadings, we also adjusted the metal inflow efficiencies to capture the metallicity of gas entering the ISM and CGM from the CGM and IGM, respectively. For the ISM, we define  $\zeta_{\text{ISM}}^{\text{in}}$  as:

$$\zeta_{\text{ISM}}^{\text{in}} = \frac{\dot{M}_{Z,\text{ISM}}^{\text{in}}}{\dot{M}_{\text{ISM}}^{\text{in}} \cdot Z_{\text{CGM}}}, \quad (17)$$

where  $\dot{M}_{Z,\text{in,ISM}}$  is the metal inflow rate into the ISM,  $\dot{M}_{\text{in,ISM}}$  is the total gas inflow rate, and  $Z_{\text{CGM}}$  is the metallicity of the CGM. For the CGM, since the SAM does not track the metallicity of the IGM, we define  $\zeta_{\text{CGM}}^{\text{in}}$  as the fraction of metals entering the CGM compared to the total rate of gas flowing into the CGM:

$$\zeta_{\text{CGM}}^{\text{in}} = \frac{\dot{M}_{Z,\text{CGM}}^{\text{in}}}{\dot{M}_{\text{CGM}}^{\text{in}}}. \quad (18)$$

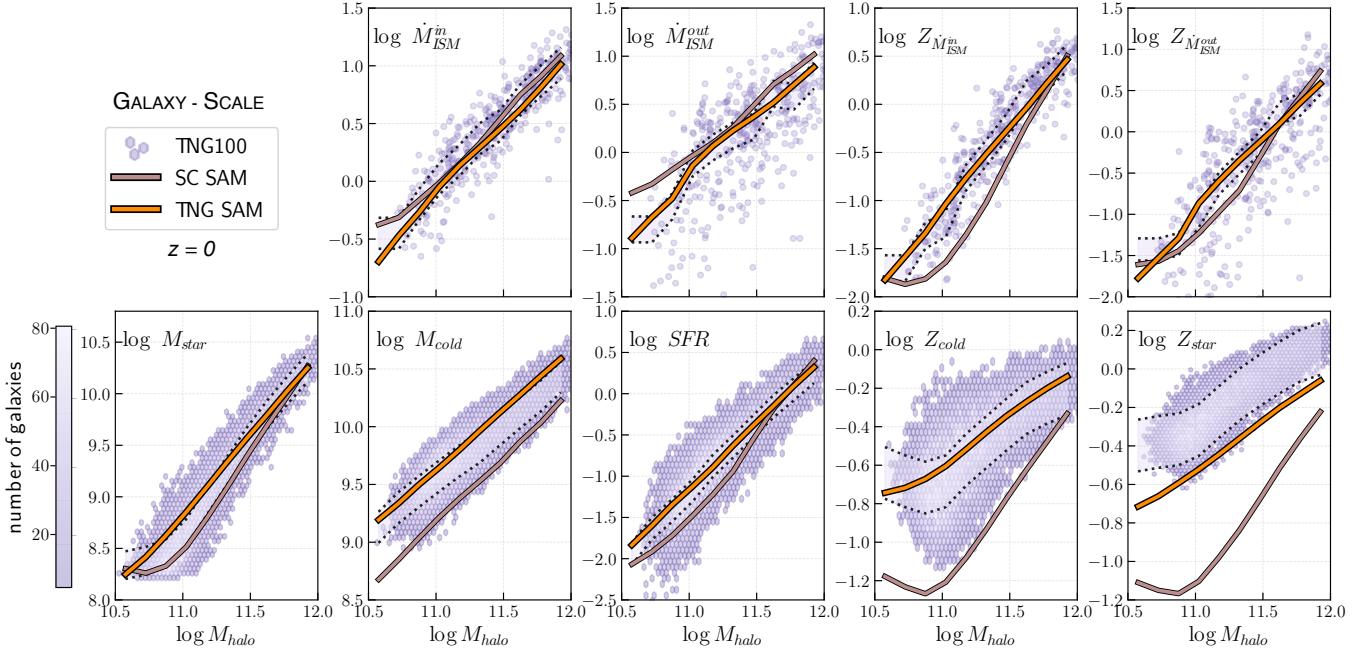
Notably, the top right panel of Figure 3 shows that, in TNG, the proportion of metals returning to the ISM significantly exceeds unity, the original assumption in the SC SAM. This indicates that the gas returning to the ISM is more metal-rich than the CGM, suggesting efficient recycling of expelled metals back into the ISM.

While neither the SAM nor TNG explicitly track the metallicity of the IGM, the non-zero value of  $\zeta_{\text{CGM}}^{\text{in}}$  implies a non-negligible inflow of metals into the CGM. This suggests that a fraction of the metals ejected via feedback can escape the halo and subsequently be re-accreted, highlighting the importance of considering metal recycling on both galactic and halo scales.

Finally, we also increased the stellar yield from the fiducial value of 1.2 in the SC SAM to 2.0 to better match the aggregate stellar metallicity observed in TNG galaxies.

## 5 REPRODUCING TNG100’S RESULTS OVER 12 GYR

As discussed in Section 1, different gas and metal flow cycles, as modeled by hydrodynamical simulations and SAMs, can follow distinct processes yet still converge on similar global predictions for galaxies. The SC SAM and TNG100, for example, represent gas cycling processes differently, but still arrive at similar global outcomes, as shown in Gabrielpillai et al. 2022. The primary goal of the TNG SAM is not only to predict these global properties but also to accurately replicate the underlying gas and metal flow cycles observed in



**Figure 5.** Comparison of baryon flow rates (top panels) and global properties (bottom panels) at the scale of the galaxy for the TNG SAM (orange), SC SAM (brown), and TNG100 (purple). Each panel shows the median (50th percentile) predictions from the SAMs compared to TNG100, with scatter points representing the CIP TNG subsample and histograms representing the full TNG100 dataset. The dotted black lines and light purple shading mark deviations within  $\pm 30\%$  of TNG100’s median. The SC SAM predictions mostly lie outside this range for baryon flows but remain within  $\pm 30\%$  for global properties such as  $M_{\star}$ ,  $M_{\text{cold}}$ , and SFR. The recalibrated TNG SAM, calibrated to both observations and TNG’s baryon cycle, significantly improves upon the SC SAM, with most predictions agreeing within  $\pm 30\%$  of TNG100’s median. The largest improvement is in the metallicities of stars and cold gas, where agreement improves from  $\sim 70\%$  with the SC SAM to 20–40% with the TNG SAM.

TNG, using the recalibrations discussed in Section 4. In this section, we assess how accurately the TNG SAM replicates the global gas and metal flow cycles as well as global properties at the scale of the galaxy and halo in TNG100. We also compare the TNG SAM to the fiducial SC SAM at  $z = 0$ , highlighting how the differences in how each SAM’s galaxy formation model changes its predictions.

### 5.1 The Galaxy Scale

Figure 5 compares the inflow and outflow rates of gas and metals at  $z = 0$  among TNG100, the TNG SAM, and the fiducial SC SAM, focusing on how well each model reproduces galaxy-scale flows. For gas inflow into the ISM (top left panel), the TNG SAM closely follows TNG100’s inflow trends, achieving a median difference of 18% across the halo mass range  $10.5 < \log M_{\text{halo}} < 12$ . While the SC SAM also shows reasonable agreement, it tends to overpredict inflow rates at the lower end of this mass range, likely due to its shorter cooling times.

For gas outflows from the ISM (second panel from the left), the TNG SAM reproduces TNG100’s more moderate outflow rates, with a median difference of 11%. By comparison, the SC SAM predicts significantly higher outflow rates—up to 60% greater than TNG100—particularly for lower-mass halos, reflecting its more aggressive stellar feedback model. For metal flows (top right panels), the TNG SAM continues to perform well, capturing inflow and outflow rates with median differences of 14% and 9%, respectively, across halo masses. In contrast, the SC SAM underestimates both metal inflow and outflow rates by  $\sim 35\text{--}60\%$ , owing to its simplified metal recycling model, discussed further in Section 6.1.5.

The TNG SAM’s ability to reproduce gas flows into and out of galaxies within  $\sim 20\%$  of TNG’s values translates into similarly strong agreement for global galaxy properties. The lower panel of

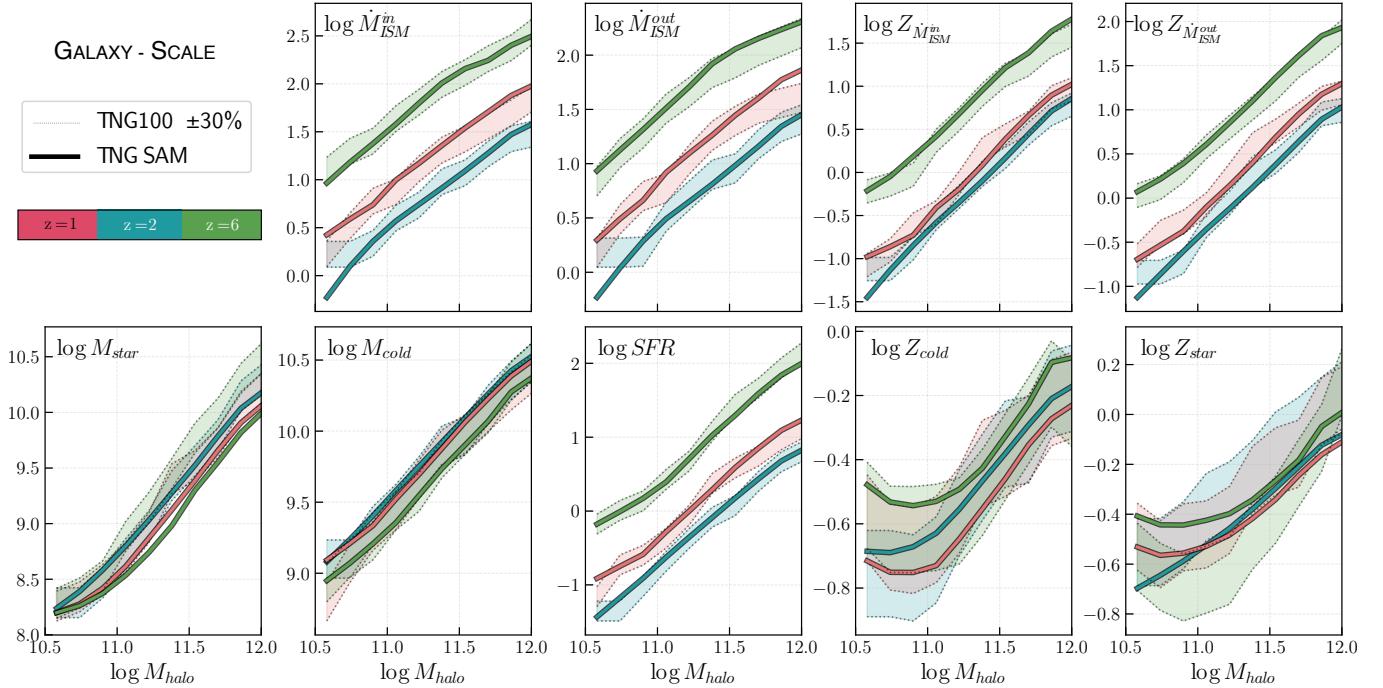
Figure 5 compares key global properties—stellar mass, cold gas mass, star formation rate, and the metallicities of the stellar population and cold gas—between TNG, the TNG SAM, and the SC SAM. For stellar mass, the TNG SAM closely follows TNG, with average discrepancies reduced to 11%. The SC SAM also performs reasonably well, agreeing with TNG within 30% across the mass range, despite differing treatments of gas accretion and star formation. This once again demonstrates that different baryon cycling processes can still produce the “correct” outcome for global galaxy properties, raising the question of which underlying physical processes are truly accurate.

For the cold gas mass (second row, left panel), the TNG SAM performs moderately well, with most deviations within 25%. In contrast, the SC SAM underpredicts cold gas masses by  $\sim 50\%$ . Star formation rates (bottom left panel) show excellent agreement between TNG and the TNG SAM, with differences typically within 6%. However, the SC SAM predicts systematically lower SFRs across the mass range shown.

Metallicities in the TNG SAM show impressive agreement with TNG compared to the fiducial SC SAM model. The SC SAM underpredicts stellar and cold gas metallicities by  $\sim 70\text{--}80\%$ , whereas the TNG SAM matches TNG’s predictions within 8% for cold gas metallicities and 40% for stellar metallicities.

To evaluate how well the TNG SAM’s agreement with TNG at  $z = 0$  holds across cosmic time, Figure 6 presents the same properties at  $z = 1, 2$ , and 6. For gas inflows and outflows in the ISM, the TNG SAM generally remains within 30% of TNG, with deviations increasing to about 50% for lower-mass halos. Metal inflows and outflows show greater consistency, with differences remaining within 20% across the halo mass range and redshifts explored.

For global properties, the TNG SAM generally tracks TNG’s stellar mass, cold gas mass, and star formation rate within 20% across time,



**Figure 6.** Time evolution of the median galaxy-scale flow rates (top panels) and global properties (bottom panels) from Figure 5, comparing the TNG SAM (solid lines) and TNG100 at  $z = 1, 2, 6$ , with red, blue, and green lines representing each redshift. Shaded regions show  $\pm 30\%$  of TNG’s distribution, with dotted gray lines marking the upper and lower bounds. The TNG SAM closely replicates TNG’s gas flow rates, remaining within 30% of the median values. Similar agreement is observed for global galaxy properties, although the agreement for  $M_{\text{star}}$  worsens to  $\sim 60\%$  at  $z = 6$ .

with slightly larger deviations at earlier epochs. For example, the TNG SAM shows a maximum deviation of 60% for stellar mass at early times for higher-mass halos. For the metal populations, the TNG SAM consistently provides good ( $< 30\%$ ) agreement with TNG for the cold gas metallicities. Interestingly, the agreement for the stellar metallicities improves at higher redshifts, with deviations decreasing in the early universe. The largest differences remain observed at  $z = 0$ , where the deviation reaches up to 40%.

## 5.2 The Halo Scale

Now, at the scale of the halo, we evaluate how well the TNG SAM and the fiducial SC SAM emulate the gas and metal flows at  $z = 0$  in TNG100, as shown in the top panel of Figure 7. For gas and metal inflows into the halo (top left and middle right panels), the TNG SAM generally matches TNG’s predictions within 20%. In contrast, the SC SAM, which uses a more simplified model for gas accretion, predicts inflow rates up to 45% higher than TNG across halo masses. This discrepancy is particularly evident for metal inflows, where the SC SAM’s more efficient recycling leads to higher predicted rates, while the TNG SAM aligns more closely with TNG’s gradual metal flows, maintaining accuracy within 30% across the mass range displayed.

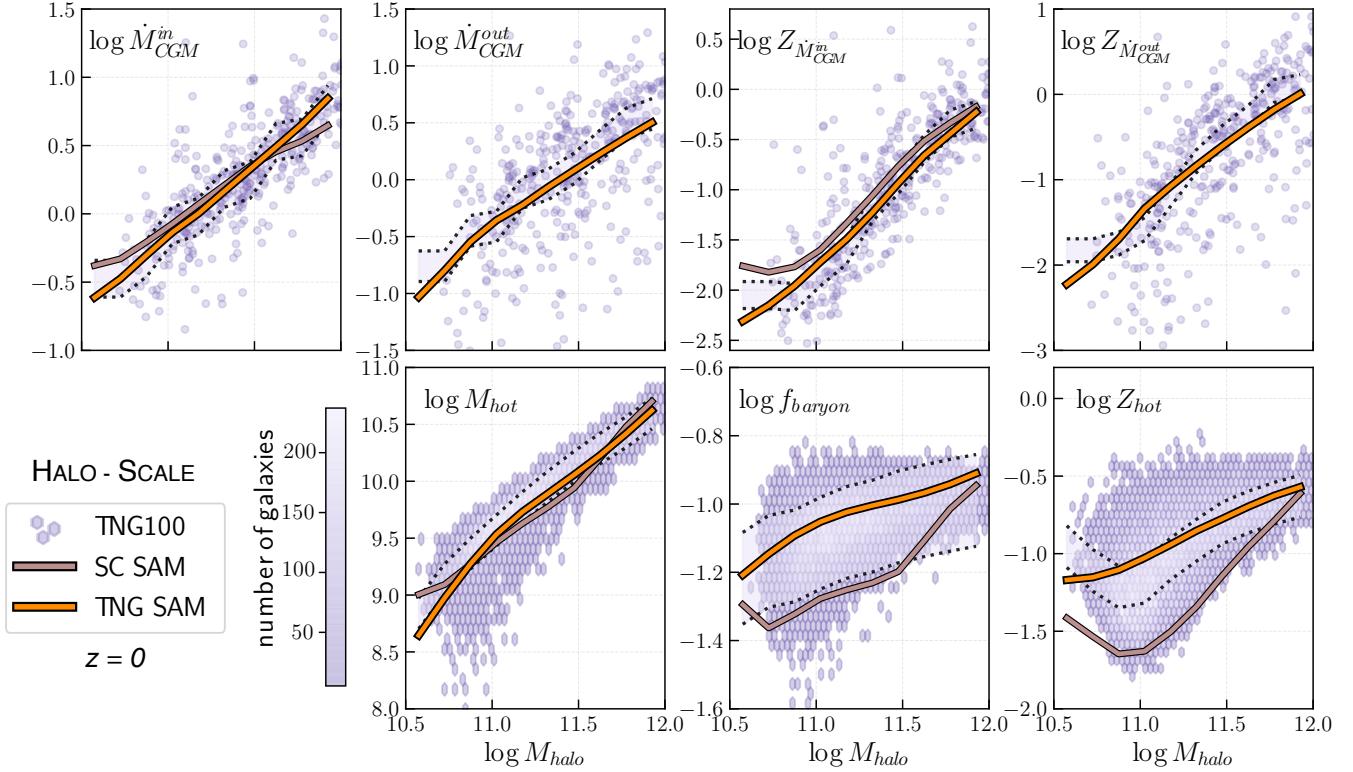
One of the most notable changes in the TNG SAM is the explicit inclusion of gas and metal expulsion from the CGM to the ejected reservoir, a channel not included in the SC SAM. While these outflows are generally an order of magnitude weaker than the inflowing component, they are comparable in strength to gas outflows from the ISM, highlighting their importance in regulating the baryon cycle. The TNG SAM’s gas and metal outflow rates closely match TNG’s predictions, with deviations of  $\sim 20\%$  and  $30\%$ , respectively.

As shown at the galaxy scale in Section 5.1, the TNG SAM’s ability to reproduce gas flows in the CGM within  $\sim 30\%$  of TNG’s values translates into similarly strong agreement for global halo properties.

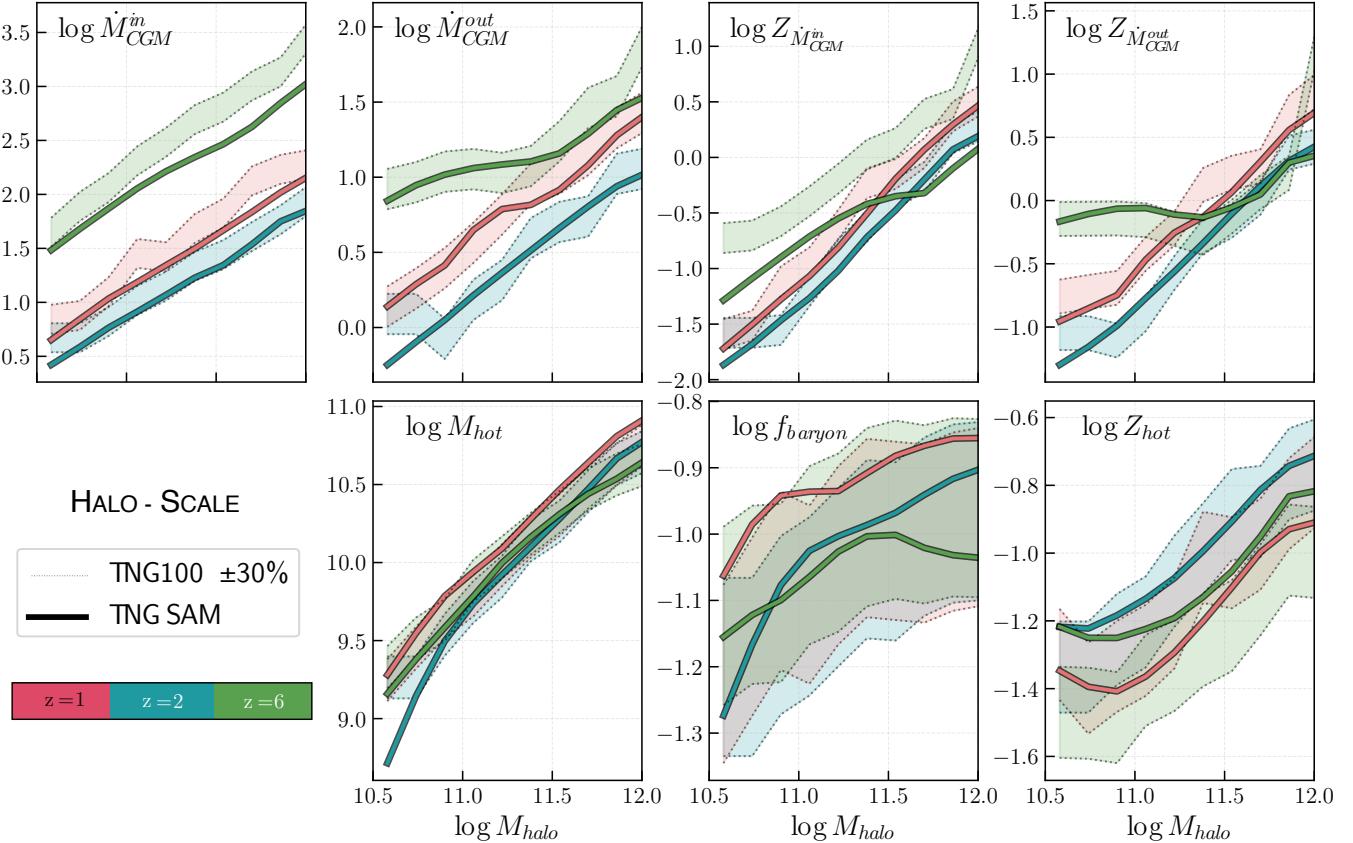
The lower panel of Figure 7 compares key halo-scale quantities—hot gas mass, baryon fraction, and the metallicity of the hot gas reservoir—across TNG100, the SC SAM, and the TNG SAM.

The TNG SAM shows excellent agreement with TNG for the hot halo mass, with deviations generally within 20% across the halo mass range. Despite lacking halo-scale outflows, the SC SAM predicts only slightly lower hot gas masses, again illustrating how variations in baryon cycling processes can still yield similar global outcomes. For the overall baryon fraction, the TNG SAM matches TNG within 10% accuracy, while the SC SAM underpredicts the baryon fraction by  $\sim 35\%$  at low to intermediate halo masses, consistent with its slight underestimation of the hot halo gas mass. Similarly, the TNG SAM closely reproduces TNG’s predictions for hot gas metallicities, with deviations within 12%. In contrast, the SC SAM significantly underpredicts hot gas metallicities by  $\sim 65\%$ .

Figure 8 examines how the TNG SAM’s agreement with TNG evolves over time. For gas and metal inflows and outflows in the CGM, the TNG SAM generally stays within 30% of TNG’s values, except at  $z = 6$ , where inflow rates are underpredicted by 50–60%. This discrepancy arises from the simplified way the SAM calculates dark matter accretion rates, which we discuss further in Section 6.2.2. Nevertheless, for global halo properties, the TNG SAM maintains good agreement with TNG over time, with deviations for the baryon fraction and hot gas mass staying within 30%. For the metallicity of the hot gas reservoir, the TNG SAM consistently achieves strong agreement ( $< 30\%$ ), outperforming the SC SAM across all redshifts explored.



**Figure 7.** The same comparison made in Figure 5, now at the halo scale at  $z = 0$ . The SC SAM fluctuates in and out of the  $\pm 30\%$  range of TNG100’s distribution for flow properties and remains mostly within 35% for global properties such as  $M_{\text{hot}}$  and  $f_{\text{baryon}}$ , while still underpredicting the metallicity of the hot gas. In contrast, the TNG SAM reproduces both global and flow properties within 30% of TNG100’s median values.



**Figure 8.** Time evolution of the median halo-scale flow rates (top panels) and global properties (bottom panels) from Figure 7, comparing the TNG SAM and TNG100. The same color scheme used for galaxy-scale properties in Figure 6 is applied here. The TNG SAM predicts all halo-scale global and flow properties within 30%, except for  $\dot{M}_{\text{CGM}}^{\text{in}}$  at  $z = 6$ , where the agreement falls to 50%.

## 6 DISCUSSION

### 6.1 Why do the SAM and TNG Agree?

The TNG SAM’s ability to reproduce the median galaxy- and halo-scale properties of  $\sim 20,000$  TNG galaxies within  $\sim 10\text{--}30\%$  highlights the effectiveness of recalibrating SAMs using insights from hydrodynamical simulations. Below, we examine the specific physical modifications that facilitated this agreement for the galaxy- and halo-scale properties compared. Figure 9 summarizes the contributions of each modification, showing the TNG SAM’s performance before and after these adjustments at  $z = 0$  and their impact on each of the observed properties.

#### 6.1.1 Efficient Halo (Re-)accretion: Strong Gas Recycling and Weak Preventive Feedback

A key factor behind the TNG SAM’s success in replicating TNG’s baryon cycle is  $f_{\text{in,CGM}}$ , which regulates the efficiency of gas accretion into halos. As shown in Figure 3,  $f_{\text{in,CGM}}$  in TNG often exceeds unity, particularly at lower redshifts, indicating that gas inflow frequently meets or surpasses the expected baryon-to-dark matter ratio. In contrast,  $f_{\text{in,CGM}}$  values in hydrodynamical simulations like FIRE (Pandya et al. 2020; Pandya 2021) and EAGLE (Mitchell et al. 2020; Mitchell & Schaye 2022; Wright et al. 2020) typically fall below unity, reflecting stronger “preventive” feedback, which heats gas in the IGM and limits its accretion onto halos. TNG’s elevated  $f_{\text{in,CGM}}$  suggests that ejected gas efficiently re-enters halos, mitigating the impact of any preventive feedback mechanisms that may be present.

The consistent elevation of  $f_{\text{in,CGM}}$  across the entire halo mass range explored raises the question: what drives the efficient gas recycling in TNG? For low-mass halos ( $10.5 < \log M_{\text{halo}} < 11$ ), stellar feedback plays the primary role in ejecting the gas that gets recycled. As halo mass increases, AGN feedback becomes increasingly significant, and CIP note that for these halos, the number of galaxies dominated by AGN vs supernovae feedback is roughly balanced. Despite this shift,  $f_{\text{in,CGM}}$  remains consistently elevated across all halo masses studied, suggesting that stellar feedback is just as effective as AGN feedback at maintaining the high recycling efficiency in the TNG Universe.

In the TNG SAM  $f_{\text{in,CGM}}$  replaces the fixed gas return fraction ( $f_{\text{return}} = 0.1$ ) used in the SC SAM. Most SAMs also rely on static values for gas recycling (e.g., 0.64 in GALFORM, 1.0 in L-galaxies). However, static return fractions clearly oversimplify the dynamic accretion and recycling processes seen in hydrodynamical simulations like TNG. Without  $f_{\text{in,CGM}}$ , the TNG SAM underpredicts the hot gas mass by 45%, driven by insufficient gas accretion into the halo, which subsequently leads to the SAM underpredicting nearly all other quantities.

Despite the effectiveness of  $f_{\text{in,CGM}}$  in the SAM, uncertainties remain regarding the precise nature of gas recycling in TNG. TNG’s mesh-based approach makes it difficult to track whether accreted gas is pristine or recycled, unlike particle-based simulations like EAGLE, where gas flows can be explicitly traced. In EAGLE, for instance, Mitchell & Schaye (2022) found that the halo recycling efficiency increases monotonically with halo mass and redshift. If such tagging were possible in TNG, it would provide valuable insights into whether preventive feedback still plays a role in first-time gas accretion in low-mass halos, although recycling is clearly the dominant process once it begins.

Looking ahead, the next generation of SAMs would greatly benefit from incorporating more flexible and physically motivated recycling

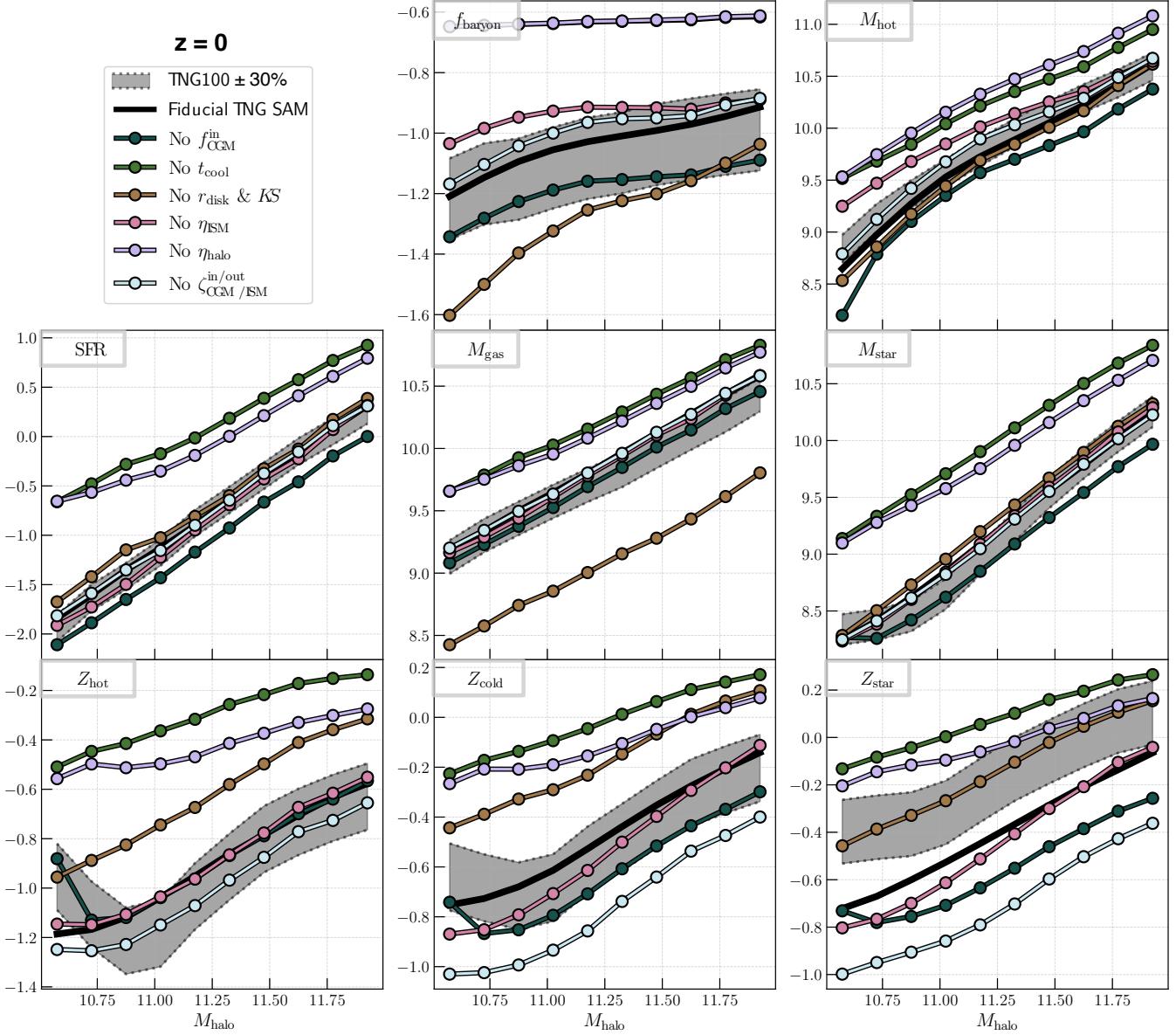
models that are guided by the explicit tracking of gas flows observed in hydrodynamical simulations. While the  $f_{\text{in,CGM}}$  parameter offers a useful step in replacing the ad hoc static recycling fractions used in most SAMs, further refinement—such as directly modeling the recycling efficiency as it evolves with halo mass and redshift could also be helpful. For instance, the latest version of SAGE (Croton et al. 2016) adjusts the reincorporation rate to increase in massive halos, based on fits to multiple simulations, a trend also observed in EAGLE.

#### 6.1.2 Revised Cooling Model: Limitations of the Cold-Mode vs. Hot-Mode Dichotomy

Another significant improvement to the TNG SAM came from revising the cooling model for the hot halo gas onto the central galaxy. The overwhelming majority of SAMs published in the literature (e.g., Galactus Benson 2012, Shark Lagos et al. 2018, GAEA Hirschmann et al. 2016, Morgana Monaco et al. 2007) classify gas accretion as either “cold-mode” or “hot-mode” based on the cooling radius. In these models, radiative cooling is governed by well-established processes such as collisional excitation, ionization, recombination, and bremsstrahlung, with the assumption that the hot gas is shock-heated to the virial temperature of the host halo and that chemical abundances are well-mixed throughout the gas. As cooling depends on the gas density and metallicity, denser gas at the center of the halo cools faster than outer gas, resulting in an inside-out cooling pattern. While this simplified approach captures the basic mechanisms of gas cooling, it has been shown to lead to systematic discrepancies when compared to cosmological hydrodynamic simulations, with SAMs underpredicting gas accretion rates in low-mass halos but overpredicting them in massive halos (Lu et al. 2011).

In TNG, we find the traditional cooling model has the opposite trend for low-mass halos. The left panels of Figure 10 shows that cooling times in TNG often exceed the expected timescales for both cold- and hot-mode accretion. Even in halos where the cooling radius exceeds the virial radius (the cold-mode regime), the cooling time frequently surpasses the dynamical time (dashed line). If cooling were dominated by these rapid, cold-mode flows, as assumed in some SAMs (e.g., GalICS; Cattaneo et al. 2017), we would expect significantly shorter cooling times and higher accretion rates than those observed in TNG. However, the TNG model accounts for additional local factors such as the gas density, temperature, element-based metal line cooling, and radiation field from the UV background and nearby AGN, which modulate the radiative cooling rate. This results in a more gradual cooling of the hot halo gas, contrasting with the rapid collapse predicted by cold-mode flows and the delayed cooling expected from hot-mode accretion. Although identifying the exact refinements in TNG’s radiative cooling model that lead to disagreement with SAMs is beyond the scope of this paper, such an investigation could provide valuable insights for refining SAMs in the future.

Interestingly, this trend persists outside of TNG. The right panel of Figure 10 compares the cooling times in TNG with the cooling times reported in FIRE, adopted from Figure 4.22 of Pandya (2021), with both simulations predicting cooling times longer than the dynamical time at low redshift. In both TNG and FIRE, radiative cooling is implemented using similar physical processes, including collisional excitation, ionization, recombination, and metal-line cooling driven by local gas properties. However, the treatment of feedback differs substantially. The FIRE simulations employ a highly resolved, explicit treatment of stellar feedback, capturing the localized and stochastic nature of supernovae, stellar winds, and radiation pressure. On



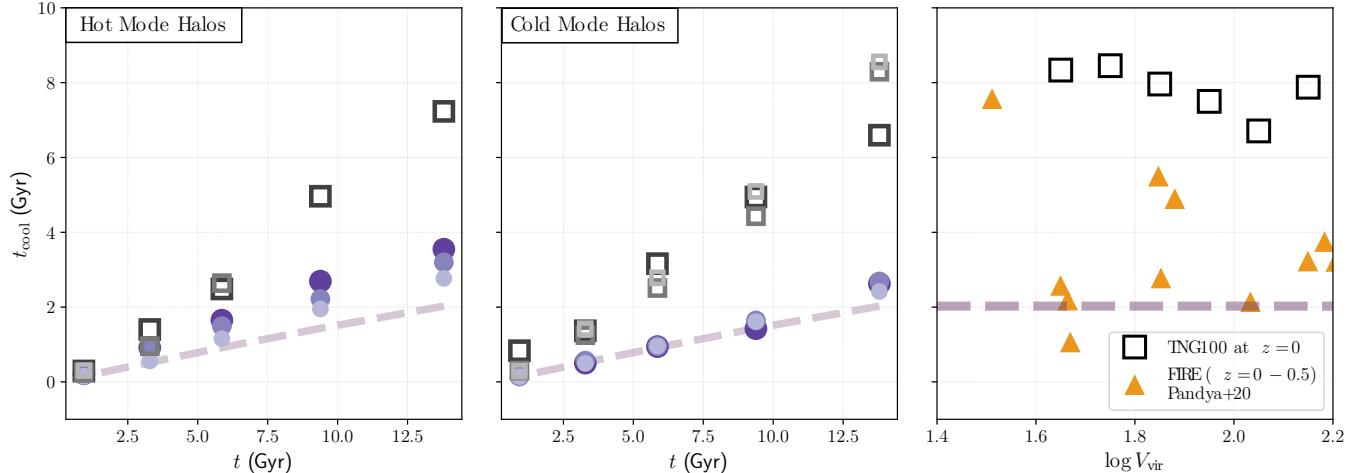
**Figure 9.** To show the importance of each calibration made to the TNG SAM, we evaluate how reverting individual parameters to their original SC SAM formulations impacts the TNG SAM’s ability to reproduce TNG’s global results at the galaxy and halo scales at  $z = 0$ . The solid black line represents the 50th percentile of the fully calibrated TNG SAM, while the shaded gray region marks  $\pm 30\%$  of the TNG100 population, with dotted lines indicating the upper and lower boundaries. Solid lines with circles represent the TNG SAM’s performance when specific calibrations are reverted, with each line colored according to the removed calibration.

the other hand, TNG utilizes a more parameterized approach with sub-grid models that distribute feedback energy and momentum over larger scales. Despite these differences in feedback implementation, both simulations produce similar results in terms of cooling times, even in halos traditionally classified as ‘cold mode.’ This lends further credence to the idea that the traditional cooling model used in SAMs, which simplifies the interaction of radiative cooling, feedback, and turbulence, may miss key aspects of gas dynamics in more realistic galaxy environments.

To address this shortcoming, we revised the TNG SAM’s cooling model to align with the cooling times in TNG. This revision leads to significant improvements in the SAM’s predictions of halo-scale properties compared to the traditional cooling model, as illustrated in the top row of Figure 9. The traditional model, due to its fast cooling times, consistently underpredicts the mass of the hot halo by

$\sim 60\%$ , leading to elevated gas masses in the ISM. The success of this revision underscores the importance of adopting more physically detailed cooling models in SAMs to better replicate the complex gas cooling processes captured by hydrodynamical simulations like TNG and FIRE.

Building on this success, a more physically motivated approach may involve not only explicitly tracking mass flows, but also energy flows within the SAM, further discussed in Section 6.2.4. Pandya et al. 2023; Carr et al. 2023; Voit et al. 2024a,b demonstrated that accounting for the transfer of feedback energy to the CGM using energy-tracking ordinary differential equations naturally slows cooling rates, producing timescales consistent with those observed in FIRE-2 halos. Given that both TNG and FIRE-2 halos exhibit slow cooling times, such an energy-based framework may also achieve cooling times that better align with TNG’s results.



**Figure 10.** Cooling times ( $t_{\text{cool}}$ ) for TNG (squares) and the SC SAM (circles), with marker size indicating different halo mass ranges: smaller markers for  $10.5 < \log M_{\text{halo}} < 11$ , medium for  $11 < \log M_{\text{halo}} < 11.5$ , and larger for  $11.5 < \log M_{\text{halo}} < 12$ . The left and middle panels show the time evolution of  $t_{\text{cool}}$  for hot-mode ( $r_{\text{cool}} < r_{\text{vir}}$ ) and cold-mode ( $r_{\text{cool}} > r_{\text{vir}}$ ) halos, respectively. The dashed purple line represents the dynamical time for a  $10^{11} M_{\odot}$  halo. In TNG,  $t_{\text{cool}}$  remains far above the dynamical time for both modes, highlighting the limitations of the traditional binary accretion classification in SAMs. The right panel compares  $t_{\text{cool}}$  as a function of virial velocity ( $V_{\text{vir}}$ ) in SC SAM (purple circles), TNG100 (black squares), and FIRE (orange triangles,  $z = 0 - 0.5$ ; Pandya+20 2021). Despite differing methodologies, both TNG and FIRE predict  $t_{\text{cool}}$  values significantly longer than the dynamical time (dashed line).

### 6.1.3 Kennicutt-Schmidt Relation and Consistent Disk Sizes Drives ISM Agreement

Despite the fundamental difference in spatial scales at which the KS relation is applied in TNG—locally at the scale of  $10^6 M_{\odot}$  gas cells—and in the TNG SAM—globally across the entire galaxy—the TNG SAM successfully reproduces the global KS relation observed in TNG, leading to good agreement for all mass reservoirs in the ISM, specifically cold gas mass and stellar mass. This result is not particularly surprising, as the star formation law in TNG was explicitly calibrated to match the local, spatially resolved KS relation observed in nearby galaxies (Springel & Hernquist 2003). Although projecting quantities onto surface densities can slightly affect the slope of the KS relation, the assumption of nearly constant scale heights ensures that the relationship between 3D densities and surface densities is preserved, as demonstrated by Diemer et al. (2018), who showed that the KS law also remains consistent globally in TNG. Furthermore, observational studies (e.g., Kumari et al. 2020), have shown that the KS relation holds consistently across a wide range of scales, from macroscopic regions within galaxy disks (hundreds of parsecs) to the entire galaxy. Detailed numerical simulations, too, have demonstrated that the KS law remains robust from global to local scales (e.g., Kravtsov 2003; Li & Nakamura 2006). As a result, the agreement between TNG and the TNG SAM is likely a reflection of the broader robustness of the KS relation.

However, it is important to note that while the global total gas KS relation is reproduced well in both TNG and the TNG SAM, the same cannot necessarily be said for the molecular KS relation. Observational studies (e.g., Bigiel et al. 2008) have shown a linear molecular gas Schmidt relation, in contrast to the super-linear total gas relation seen in entire galaxies. In TNG, Diemer et al. 2018 show that this molecular KS relation is not strictly obeyed, with the simulation results showing a superlinear relation with a slope of 1.5. Since TNG does not explicitly output molecular gas, direct comparisons between molecular gas predictions in TNG and the SAM remain difficult. Given this limitation, the total gas KS relation is the most reliable metric for comparing the two models and is successful.

We stress that the success of the KS relation is highly dependent on consistently modeling the disk sizes in TNG and the TNG SAM.

Most SAMs, including the SC SAM, determine disk sizes through angular momentum conservation under the assumption of an exponential disk. While this method achieves reasonably good agreement with observed radial disk sizes as a function of stellar mass up to  $z \sim 2$  in the SC SAM (Somerville et al. 2008b), it does not align with observations as well as TNG’s predictions (Genel et al. 2018). Particularly, the SAM tends to underestimate disk sizes for lower-mass halos and at higher redshifts (see Karmakar et al. (2023) Figure 2).

Together, the KS relation and consistent definition of disk sizes allow the TNG SAM to predict the star formation rate, cold gas mass, stellar mass, and metallicities within  $\sim 30\%$  accuracy. In contrast, Figure 9 shows that the fiducial SAM’s Bigiel et al. (2008)-inspired star formation recipe and angular momentum-determined disk sizes overestimates the star formation rate and stellar mass by 35%, leading to an 80% depletion of the total cold gas mass. This discrepancy results in higher metallicities in each reservoir and an underprediction of the overall baryon fraction by  $\sim 35\%$ .

These results suggest that SAMs utilizing surface density-based star formation models, like the Kennicutt-Schmidt relation, can still provide robust predictions for ISM properties, even when applied at the global scale of the galaxy rather than locally. However, they also underscore the importance of accurately modeling disk sizes, as mismatches can lead to discrepancies in gas and stellar mass predictions. Therefore, achieving reliable predictions for cold gas and stellar masses in galaxies will require developing more physically motivated models for estimating disk sizes in SAMs.

### 6.1.4 ISM and Halo-Scale Outflows Modulate the Mass of the Hot Halo

While galaxy-scale outflows are well-established as incredibly important regulators of galaxy evolution in numerical simulations, recent studies have underscored the importance of outflows at the halo scale as well (Wright et al. 2024; Pandya et al. 2021; Mitchell et al. 2020). Consistent with these findings, we observe that both galaxy and halo-scale outflows are essential for accurately modeling the gas content in the TNG SAM. Specifically, the mass of the CGM is highly sensitive to the strengths of the outflow efficiency parameters  $\eta_{\text{ISM}}$  and  $\eta_{\text{halo}}$ , as shown in Figure 9.

Incorporating TNG’s galactic winds model into the TNG SAM, which explicitly ties wind strength to the metallicity of the cold gas, significantly improved the model’s ability to reproduce galaxy-scale outflows observed in TNG. Although the metallicity dependence adds complexity, it proved crucial for avoiding the overestimation of galaxy outflows by over 30%, as seen in the SC SAM (Figures 5 and 6). This overestimation in the SC SAM results in excessive gas expulsion, inflating both  $M_{\text{hot}}$  and  $f_{\text{baryon}}$ . By contrast, the TNG SAM’s metallicity-modulated feedback provides a more nuanced treatment of galaxy outflows.

The most critical update to the TNG SAM’s feedback model, however, was the explicit inclusion of halo-scale outflows, a process indirectly present in the SC SAM, and directly included in few other traditional SAMs (e.g. GALFORM SAMs [Bower et al. \(2012\)](#); [Lacey et al. \(2016\)](#)). Unlike the SC SAM, which only allowed gas to be ejected directly from the ISM into the ejected reservoir, the TNG SAM introduces a more realistic channel for these outflows. Gas first leaves the ISM and enters the CGM, where it continues to interact with the hot halo. From there, a significant portion of this gas can be further expelled into the ejected reservoir/IGM, forming the CGM-to-IGM outflow channel. This two-step process provides a more realistic depiction of how gas is cycled and expelled, as opposed to the simplified ISM-to-ejected reservoir flow used in many traditional SAMs. Neglecting halo-scale outflows, as shown in Figure 9, results in substantial overpredictions of  $M_{\text{hot}}$  and  $f_{\text{baryon}}$  by over 200%. These overestimates propagate into the ISM, causing  $M_{\text{cold}}$ , SFR, and  $M_{\star}$  to be overestimated by 100–300%.

#### 6.1.5 Metal Cycling Efficiencies Improve Metallicity Predictions

Incorporating metallicity-weighted mass-loading factors greatly enhanced the TNG SAM’s ability to replicate the evolution of cold gas, stellar, and hot halo metallicities observed in TNG100. Without these refinements, the TNG SAM underpredicts the cold gas and stellar metallicities by up to  $\sim 70\%$ , as shown in Figure 9. While the metallicity of the CGM is less affected, discrepancies up to 20% persist, consistent with the TNG SAM.

Introducing metal cycling efficiencies mostly improves the TNG SAM’s predictions for the cold gas metallicities. The agreement for stellar metallicities, however, shows a more complex pattern. At higher redshifts, the stellar and cold gas metallicities in TNG are nearly equal, consistent with the instantaneous recycling approximation used in the TNG SAM, which assumes that metals produced by stars are immediately mixed into the surrounding gas. This leads to improved agreement for stellar metallicities at early times. However, at lower redshifts, this approximation breaks down in TNG, where time delays in metal transport and mixing become more significant. As a result, the TNG SAM increasingly underpredicts stellar metallicities relative to TNG at later times, reflecting a limitation of the instantaneous recycling approach. Despite these challenges, the introduction of metallicity-weighted mass-loading factors compensates for some of the model’s simplifications, allowing the TNG SAM to better track the movement of metals between the ISM, CGM, and IGM.

Most SAMs in the literature also use similar instantaneous recycling models for metal production, although some models (e.g. GAEA) attempt to modify it by accounting for the lifetimes of different stellar populations. The overwhelming majority of these models, however, like the Santa Cruz SAM, also assume that metals are exchanged between the different baryonic reservoirs in proportion to the exchange of gas mass from each component. Few SAMs, such as SAG ([Collacchioni et al. 2018](#)) and GALFORM ([Lagos et al. 2013](#)),

have experimented with incorporating metal loadings in a similar manner to the TNG SAM. However, these models found little impact on reproducing the observed stellar mass-metallicity relation, likely because their supernova feedback models are not tied to the metallicity of the gas. In contrast, the TNG SAM, mimicking the TNG model, explicitly ties feedback strength to gas metallicity, an approach also found to be critical in high-resolution zoom-in simulations of the ISM like TIGRESS ([Kim et al. 2020b,a](#)) and FIRE ([Hopkins et al. 2018, 2023](#)).

If future generations of SAMs aim to model feedback in a more physically motivated manner, similar to the aforementioned hydrodynamical simulations, incorporating metallicity-weighted feedback models and corresponding metal loadings will likely become necessary. As demonstrated in the TNG SAM, this approach can help address some of the shortcomings of the instantaneous recycling model and lead to more accurate predictions of metal abundances in galaxies and their surrounding gas reservoirs.

## 6.2 Limitations and Future Directions

Although Section 5 shows that the median galaxy population produced by the TNG SAM agrees quite well with TNG given the physical updates made, there are still several limitations to our approach, discussed below.

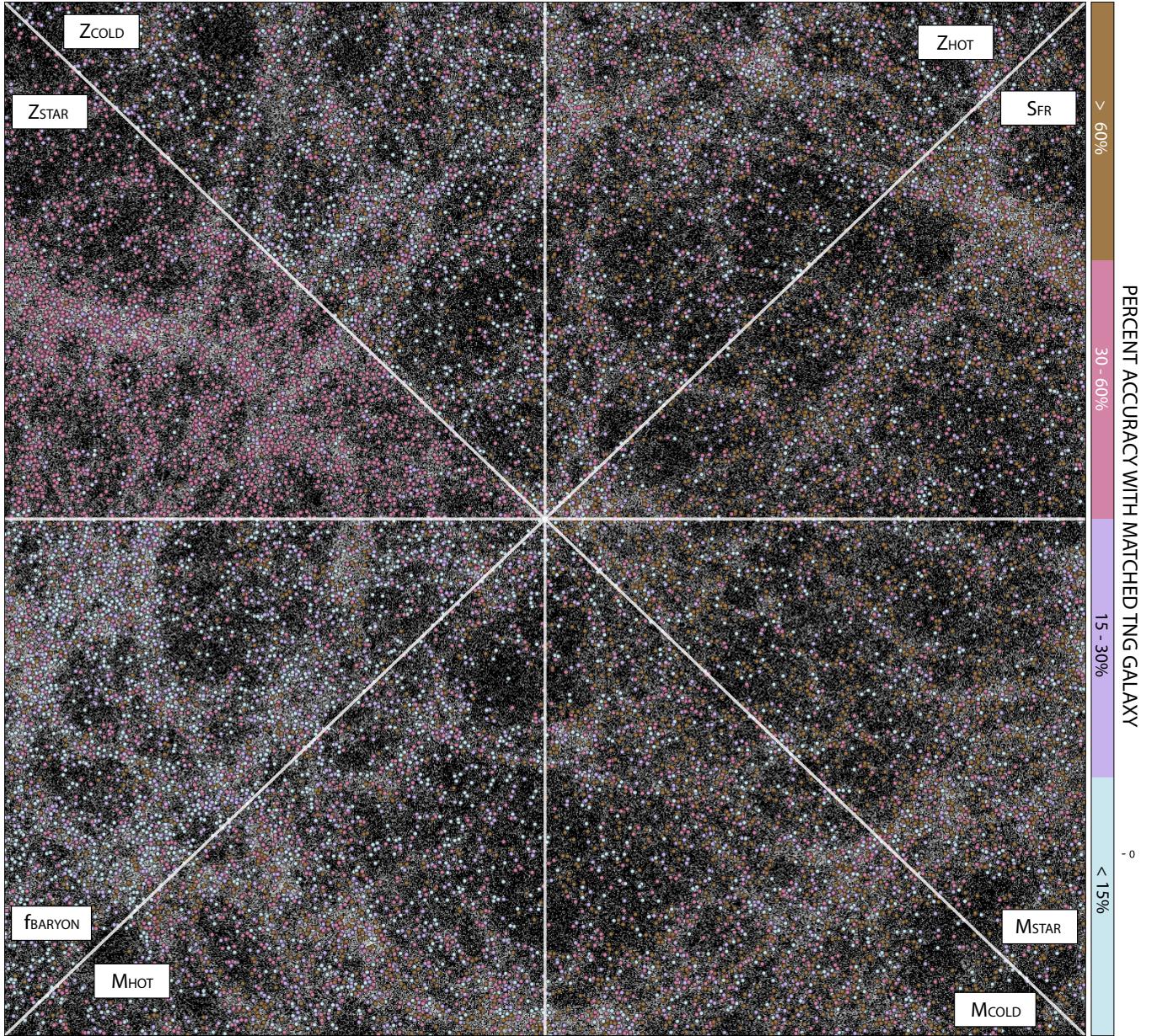
### 6.2.1 Modeling Individual Galaxies with Median-Based Calibrations

The TNG SAM is calibrated using median relations that reflect the overall behavior of galaxies within TNG rather than tailoring the model to fit individual galaxies. To evaluate how well the TNG SAM’s median-based calibrations replicate the detailed growth histories of individual TNG galaxies, we directly matched subhalos between TNG and the SAM to perform an object-by-object comparison.

Although the TNG SAM’s median galaxy population is mostly within 30% of the TNG median, Figure 11 shows that the spread increases significantly when comparing individual galaxies. The TNG SAM still demonstrates generally good agreement for several key properties, although this agreement is primarily observed for properties at the halo scale. For instance, the baryon fraction is accurately predicted within 30% for 66% of galaxies, highlighting the SAM’s strong performance in capturing the overall gas content of halos. Similarly, the hot halo mass is generally well-reproduced, with 54% of galaxies falling within 30%.

The agreement is less consistent for galaxy-scale properties. The TNG SAM predicts stellar masses with only 30% accuracy for 40% of galaxies. Cold gas masses show similar performance, with 40% of galaxies meeting the same threshold. Predictions for star formation rates exhibit the weakest agreement, with only 31% of galaxies being accurate to within 30%.

The disagreement at the galaxy scale largely stems from the SAM’s inability to correctly model the variation in galaxy sizes in TNG. Fixing galaxy sizes to  $0.25 r_{\text{vir}}$  proved effective for reproducing median trends but clearly fails to capture the variability in galaxy sizes, which is critical for modeling the spread in galaxies’ star formation rates and stellar and cold gas masses. Attempts to introduce scatter in disk sizes using the angular momentum model from the fiducial SAM only worsened predictions for individual galaxies. While the angular momentum model introduces variability in galaxy sizes, it does not generate the correct variation observed in TNG. These results further emphasize the need for the SAM community to develop



**Figure 11.** Spatial distribution of matched halos between TNG100 and the TNG SAM at  $z = 0$ , plotted in physical space (x vs. y position). All TNG100 halos with  $10.5 < \log M_{\text{halo}} < 12$  are shown in white. The map is divided into eight octants corresponding to a specific galaxy property: baryon fraction ( $f_{\text{baryon}}$ ), hot gas mass ( $M_{\text{hot}}$ ), cold gas mass ( $M_{\text{cold}}$ ), stellar mass ( $M_{\star}$ ), star formation rate (SFR), and metallicities of hot gas ( $Z_{\text{hot}}$ ), cold gas ( $Z_{\text{cold}}$ ), and stars ( $Z_{\star}$ ). Each matched galaxy is colored by the percent difference between the TNG SAM and TNG100 predictions for the given property in each octant. Blue and purple points represent agreement within 30%, while pink and brown points highlight discrepancies greater than 30%. While the overall agreement is strong across most properties, notable scatter is observed for galaxy-scale properties like  $M_{\star}$  and  $M_{\text{cold}}$ , illustrating the challenges of calibrating to median trends alone.

more sophisticated approaches to modeling galaxy sizes and their variability.

Lastly, the TNG SAM’s metallicity predictions show varying levels of agreement, consistent with the trends observed for the median predictions in Section 5. The cold gas metallicity demonstrates relatively strong performance, with 57% of galaxies within 30% of TNG values. The hot gas metallicity performs slightly worse, with 42% of galaxies meeting this threshold. Stellar metallicity shows the weakest performance, with only 22% of galaxies within 30%, though most ( $\sim 76\%$ ) are within 60%.

Despite these limitations, the results are encouraging. Nearly half of the 20,000 matched galaxies agree with TNG’s predictions within 30%, even though the SAM was calibrated using medians derived

from a much smaller sample of just  $\sim 400$  galaxies. This level of agreement is particularly promising given the computational simplicity of the SAM compared to the detailed hydrodynamical simulation. Refining the work presented here will provide a nice foundation for developing the next generation of flexible, computationally efficient, yet physically detailed SAMs needed to model galaxy formation across wider cosmological volumes.

#### 6.2.2 TNG100 DMO vs Hydrodynamical Simulations

A significant limitation to our approach arises from running the TNG SAM on the TNG100-1-DMO simulation while comparing results to the hydrodynamical TNG100-1 simulations. As noted in Section 3.2,

the halo masses in the TNG100-1-DM and TNG100-1 simulations differ by up to 20% for low halo masses at  $z = 0$ . This discrepancy stems from the inclusion of baryonic processes in the hydrodynamical simulation, and varies inconsistently across redshift.

This variation has significant implications for the TNG SAM, particularly in how the model calculates the rate of gas accretion. The SAM derives the rate of gas accretion by finite differencing the halo masses reported by the DMO simulation. The mismatch between halo masses in the N-body and hydrodynamical simulations affects how well we can model parameters like  $f_{\text{in,CGM}}$ , which governs the growth of the hot halo, which in turn regulates the build up of cold gas and stars in the ISM. Due to this discrepancy, directly tuning  $f_{\text{in,CGM}}$  in the SAM to match the hydrodynamical simulation was not feasible. Instead, we inferred the value of  $f_{\text{in,CGM}}$  by aligning the rate of gas entering the CGM with TNG. This approach performed well at lower redshifts but became problematic at  $z = 6$ , where the difference in the dark matter accretion rates between the SAM and TNG reached nearly 30%. This discrepancy forced the SAM to underpredict the amount of gas entering the CGM, leading to deviations in the modeled baryon cycle at high redshifts.

Furthermore, the above discrepancies in virial mass result in  $\sim 20\%$  differences in the virial radius ( $r_{\text{vir}}$ ), directly affecting the SAM’s ability to consistently define the galaxy radius as  $0.25 r_{\text{vir}}$  to align with TNG. As discussed in Section 6.1.3, accurate disk sizes are crucial for determining surface densities used in the Kennicutt-Schmidt relation, which directly influence predictions for star formation rates, stellar masses, and cold gas masses, introducing inherent uncertainties when disk sizes are incorrectly estimated.

### 6.2.3 Resolution and Sample Size

The TNG SAM’s ability to reproduce TNG100’s results is further constrained by the halo mass and temporal resolution of the calibration sample. We consider halos well-resolved if they contain more than 100 star particles ( $\log M_{\text{halo}} > 10.5$ ); however, the baryon flow measurements used to calibrate TNG SAM are extracted from radial shells  $0.1 r_{\text{vir}}$  thick within each halo. For lower-mass halos, particularly those around  $10^{10} M_{\odot}$ , these shells often contain far fewer resolution elements, producing noisy efficiency estimates and contributing to discrepancies in the SAM’s predictions at the low-mass end, where deviations are most pronounced. Moreover, the  $10^{10} M_{\odot}$  halos modeled at  $z = 0$  evolved from progenitors with halo masses of  $\sim 10^8 M_{\odot}$  at  $z = 6$ , which contain poorly resolved stellar and gas components.

This mass resolution limit precluded us from tracking the baryon cycle’s time evolution in a select group of galaxies selected at  $z = 0$ . Instead, we had to rely on aggregate measurements at each redshift, which introduces additional uncertainties for low-mass systems. More importantly, at redshifts higher than  $z = 6$ , few galaxies exist within the targeted mass range of  $10^{10} - 10^{12} M_{\odot}$ , with resolved galaxies predominantly sampling the  $10^{10} - 10^{11} M_{\odot}$  range. This limits our ability to fully populate the halo mass range at early times and places a practical limit on our analysis at  $z = 6$ , beyond which the data reliability steadily declines.

To account for unresolved galaxies below the well-resolved mass range, we extrapolated their behavior using baryon flow trends observed for halos in the range  $10 < \log M_{\text{halo}} < 10.5$ . While this approach introduces uncertainty—since it is unclear whether these trends hold for  $\log M_{\text{halo}} < 10$ —it plays a critical role in achieving the TNG SAM’s overall 30% agreement with TNG for higher-mass halos. Accurate modeling of low-mass galaxies, even approximately, is essential because their baryon flow efficiencies influence the evolu-

tion of more massive descendants at later times. However, this extrapolation also introduces unquantifiable uncertainty, raising questions about the robustness of the 30% agreement achieved.

This raises two broader questions about the TNG SAM’s calibration process: 1) could its accuracy be improved with a larger calibration sample, and 2) how few galaxies are needed to maintain reliable predictions? Since the TNG SAM achieves an impressive 30% accuracy by only calibrating to a sample of  $\sim 400$  resolved TNG galaxies, increasing the calibration sample size seems like an obvious way to enhance the model’s precision, particularly given the noisy measurements at the low-mass range. However, extracting the baryon flow measurements required for calibration is computationally intensive. Scaling this process to thousands of galaxies would significantly increase the current computational costs (insert comp time), undermining the efficiency that makes semi-analytic models so attractive in the first place.

Interestingly, we find that the TNG SAM can maintain accuracy for galaxy-scale properties within  $\sim 40\%$  even when calibrated with only 100 galaxies per redshift, randomly sampling the  $10^{10} - 10^{12} M_{\odot}$  halo mass range. The reduced sample size does impact the accuracy of the halo-scale properties quite a bit, however, ranging from 20–80% accuracy (see Appendix). These results suggest that a larger sample size could certainly improve the accuracy currently achieved by the TNG SAM, but also that the quality of the calibration sample is more critical than its size.

These results are particularly promising for future efforts, such as those by SMAUG and LTU, to develop next-generation SAMs and hydrodynamical simulations using detailed physical insights from zoom-in simulations and observations. They demonstrate that the quality of the calibration sample—its ability to represent key physical processes—can be as impactful as computationally expensive increases in sample size. This is also promising for multi-wavelength observational surveys, which are often constrained by the difficulty in securing enough observing time across numerous telescope time allocation committees. For example, large surveys like CALIFA comprise around 600 galaxies, while highly detailed surveys such as PHANGS feature  $\sim 90$  galaxies. The TNG SAM’s success with just 100 well-sampled galaxies per snapshot suggests that small but carefully selected observational datasets could also help to refine feedback prescriptions in numerical simulations. While observational data inevitably include greater uncertainties than simulations, high-quality, representative datasets can still aid in improving our understanding of the baryon cycle and galaxy evolution across cosmic time.

### 6.2.4 Future Directions

The TNG SAM, though successful in reproducing TNG100’s baryon cycle and scaling relations for a limited halo mass range ( $10.5 > \log M_{\text{halo}} < 12$ ), has several limitations, as discussed above. Expanding the analysis to galaxies in the TNG50 simulations, which resolves halos down to  $\sim 10^8 M_{\odot}$ , would likely yield even more significant insights into the baryon cycle in lower-mass halos. Although TNG100 was the more appropriate choice for this study given its use in the calibration of the fiducial TNG galaxy formation model and the known resolution-dependent discrepancies in TNG (Pillepich et al. 2018a), running this exercise on TNG50 would offer a more comprehensive test of the TNG SAM’s ability to reproduce galaxy properties across a wider range of stellar feedback dominated halos. Similarly, running this exercise on the TNG300 simulations would also allow us to investigate the lack of convergence between the different TNG resolutions, potentially uncovering which aspects of the baryon cy-

cle are most sensitive to the varying volumes and resolutions in the simulation suite.

The TNG SAM’s focus on replicating the physical processes of galaxy evolution, rather than just the final outcomes, also opens up exciting possibilities for future refinements. Whereas traditional SAMs track mass flows, as we have in this work, Pandya et al. (2023); Carr et al. (2023); Voit et al. (2024a) and Voit et al. (2024b) advocate for the inclusion of explicit tracking of energy flows alongside mass flows. By explicitly tracking the energy budget of the CGM, accounting for both energy sources (e.g., stellar and AGN feedback) and sinks (e.g., radiative cooling and turbulence dissipation), this model enables a more nuanced understanding of the complex interplay of physical processes that shape the CGM and its impact on galaxy evolution.

Integrating such an energy-based CGM model into the TNG SAM would allow for a more self-consistent treatment of feedback processes, where the energy injected by stellar winds and AGN could be directly linked to the thermal and kinetic energy of the CGM. This would provide a more accurate representation of how feedback affects gas cooling, star formation, and other key processes, leading to a more comprehensive and physically motivated picture of galaxy evolution. The strength of this approach lies in its ability to leverage both simulations and observations. The energy-based CGM model can be calibrated against observational constraints on feedback energy and CGM properties at the ISM scale, which are becoming increasingly well-constrained with the advent of high-resolution, multi-wavelength observing facilities (e.g., JWST, MUSE, ALMA). The TNG simulations, in turn, can provide crucial insights into less directly observable aspects, such as the energy budget and dynamics at the CGM scale. This dual calibration, leveraging both simulations and observations, would ensure that the SAM accurately reflects both the theoretical understanding of galaxy evolution and the empirical evidence.

### 6.3 Comparison to Similar Work

Previous attempts to align the predictions of SAMs with hydrodynamical simulations have varied in methodology, scope, and success. Here, we compare our approach and results with those of Stringer et al. (2010), Neistein et al. (2012) and Mitchell & Schaye (2022), highlighting key differences and similarities, and the relative accuracy of each approach.

Early efforts to reconcile SAMs with hydrodynamic simulations (e.g., Helly et al. (2003), Benson et al. (2001), and Yoshida et al. (2002)) demonstrated that the cooled mass predicted by SAMs and hydrodynamical simulations could be consistent when the SAMs are adapted to emulate the assumptions made by the hydrodynamical simulations. Stringer et al. (2010) expanded this approach by attempting to reproduce the entire formation history of a single disk galaxy within the GASOLINE hydrodynamical simulation using a modified version of the GALFORM SAM. To achieve broad agreement with the simulation data, they adjusted parameters related to gas cooling, star formation, and feedback, demonstrating the potential of SAMs to reproduce the detailed evolution of individual galaxies. However, the focus on a single object left uncertainties about the general applicability of their findings.

Neistein et al. (2012) (hereafter N12) expanded on this to a wider range of galaxies by extracting efficiencies describing accretion, cooling, star formation, and feedback from the OverWhelmingly Large Simulations (OWLS) and applying them within the SAM presented in Neistein & Weinmann (2010) to reproduce OWLS’s results. By tuning these efficiencies as functions of halo mass and redshift, they

reproduced various global galaxy and halo properties up to  $z = 3$  within  $\sim 0.1 - 0.2$  dex. In terms of approach and scope, N12 is most similar to our work here with the TNG SAM.

However, there are both similarities and significant differences in our implementations. Similar to the TNG SAM, N12 also utilizes a set of efficiencies to describe the rate at which gas enters the halo, cools from the hot halo to the galaxy, forms stars, and is ejected by stellar feedback. For instance, the halo gas accretion efficiency parameter  $f_a = \frac{\dot{M}_{\text{CGM}}^{\text{in}}}{\dot{M}_{\text{halo}}}$  in N12 is similar to the TNG SAM’s  $f_{\text{in,CGM}}$ , except  $f_{\text{in,CGM}}$  adjusts for the baryon fraction and includes both first-time infall and gas recycling. Despite this nuance, both models similarly use their respective parameters to match the gas accretion rate observed in their hydrodynamical simulations.

In contrast, how we both treat star formation highlights where our methodologies diverge. N12 directly tunes the star formation efficiency as a function of halo mass and redshift to match the star formation rate observed in OWLS. The TNG SAM, however, takes a more physically motivated approach by scaling disk sizes with the virial radius and applying the Kennicutt-Schmidt relation to compute star formation rates based on gas surface density. This method allows the TNG SAM to replicate the underlying star formation processes in TNG without relying on predetermined efficiencies. The inclusion of disk size scaling also captures variations in gas distribution, which are critical for star formation, in a way that static parameterization cannot.

Overall, N12 used less detailed efficiencies to calibrate their SAM against OWLS, reflecting their primary focus on reproducing the hot gas mass, cold gas mass, stellar mass, and total galaxy mass. Their approach yielded very good agreement for these quantities, reporting a standard deviation ranging from 0.1 - 0.2 dex for each mass reservoir. However, they reported higher deviations for the SFR, with a standard deviation of 0.5 dex. The TNG SAM, despite aiming to reproduce a wider range of global galaxy and halo properties and extending to higher redshifts ( $z = 6$ ), achieves a comparable within 0.1-0.3 dex for almost all galaxy and halo scale properties discussed in Section 5, including the SFR.

Here, we must note a broader issue with works comparing the outputs from SAMs and hydrodynamical simulations: the definitions of key galaxy- and halo-scale properties are not always consistent. For instance, in the TNG SAM, consistent definitions of disk sizes with TNG played a critical role in determining accurate star formation rates, cold gas masses, and stellar masses. Although most studies comparing SAMs and hydrodynamical simulations report the stellar mass within twice the stellar half-mass radius, the cold gas mass can refer to a variety of definitions— in SAMs, it often refers to all gas within the ISM above a density threshold, while in hydrodynamical simulations, it typically includes gas below a temperature threshold (e.g.,  $T < 10^4$  K) or gas designated as star-forming. These discrepancies can propagate to related quantities like hot gas mass, where the gas classified as “hot” depends on what is excluded as “cold” (see Appendix for examples of these definitional inconsistencies). To our knowledge, the TNG SAM is the first such comparison exercise to apply consistent definitions across models. This raises questions about whether the reported accuracies in earlier works, including N12, fully reflect the agreement claimed in their results, given the potential biases introduced by inconsistent definitions.

Despite these uncertainties, even comparisons with inconsistent definitions can provide valuable insights into the baryon cycle. Recently, Mitchell & Schaye (2022) (hereafter MS22) expanded on the N12 framework to emulate the results of the EAGLE simulations and investigate the parameters most critical to reproducing the stellar

mass–halo mass (SHM) relation. They found that EAGLE’s SHM relation is primarily shaped by gas ejection of gas via outflows for halos with  $M_{\text{vir}} < 10^{12} M_{\odot}$ , with halo-scale preventative feedback and recycling of ejected gas playing secondary roles. They also found that the redshift evolution of the SHM relation is most sensitive to the efficiencies of first-time gas accretion and ejection by outflows, and is less sensitive to the efficiency of wind recycling, and of gas consumption by star formation. In the TNG SAM, we find similar trends, with Figure 9 showing that the stellar mass is most sensitive to  $\eta_{\text{halo}}$ , and then  $f_{\text{in,CGM}}$ .

MS22 also investigated how star formation and gas flows affect the relationship between halo mass and the masses of both the ISM and CGM. They found that the CGM mass is most sensitive to variations in gas inflows and outflows, particularly those occurring at the halo scale. This finding is echoed in our analysis of the TNG SAM, where we also observed a strong dependence of the CGM mass in TNG on halo-scale gas flows. However, while MS22 found that the mass of the ISM is more sensitive to halo-scale gas flows compared to stellar mass, the TNG SAM finds that  $\eta_{\text{halo}}$  has a comparable impact on  $M_{\text{cold}}$  and  $M_{\star}$ . MS22 found this impact was because the ISM reflects recent accretion, star formation, and outflow activity, while galaxy-scale outflows are more efficient at high redshifts in EAGLE and halo-scale outflows are more efficient at low redshifts. In the TNG SAM, however, the mass of the ISM is directly coupled to the stellar mass through the Kennicutt-Schmidt relation. This implies that any variation in the gas outflow rate, whether originating from the ISM or the CGM, will indirectly affect the stellar mass through its impact on the available gas reservoir for star formation.

A significant difference between the TNG SAM and the above studies is its explicit modeling of metal content and cycling within the hydrodynamical simulation. This step is crucial for incorporating physically motivated TNG-like efficiencies, as many processes in both TNG and EAGLE rely on accurate metallicity predictions. It remains unclear whether the other approaches could replicate metal content as effectively as the TNG SAM. However, given our success using the metal loadings, it is reasonable to assume that incorporating similar metal-tracking mechanisms in those models could yield comparable improvements.

## 7 SUMMARY AND CONCLUSIONS

In this paper, we introduced the TNG SAM, a recalibrated version of the Santa Cruz semi-analytic model, designed to replicate the complex baryon cycle of galaxies in the IllustrisTNG cosmological hydrodynamical simulations. This model bridges the detailed physical processes captured in hydrodynamical simulations with the computational efficiency of SAMs, offering a powerful tool for studying galaxy evolution. By incorporating TNG’s galaxy formation physics into a simplified framework, the TNG SAM allows for deeper insights into the interplay of baryonic processes that shape galaxies.

Focusing on stellar feedback-dominated systems, we aimed to reproduce the baryon cycle in low- to intermediate-mass dwarf galaxies ( $M_{\text{halo}} \sim 10^{10} - 10^{11}$ ) and Milky Way-mass galaxies ( $M_{\text{halo}} \sim 10^{12}$ ) in TNG100. Using measurements of gas flows from a subset of  $\sim 400$  central galaxies as a proxy for the larger TNG100 sample, we updated the SC SAM’s physical prescriptions for halo gas accretion, cooling, stellar feedback, and metal circulation. These updates, implemented as a function of halo mass and redshift, led to more accurate predictions of galaxy-scale properties such as stellar mass, gas content, and metallicity, as well as halo-scale properties like the overall baryon content.

Several key insights about the baryon cycle in TNG emerged from this work, with important implications to keep in mind when building the next generation of SAMs:

(i) *The complexities of gas cooling in TNG necessitate a departure from the traditional “cold mode” vs. “hot mode” dichotomy in SAMs.* In TNG and FIRE, cooling times often exceed dynamical timescales, showing that the simple cold-mode vs. hot-mode framework commonly used in SAMs does not adequately represent the cooling process. The TNG SAM’s revised cooling model, calibrated to the cooling times in TNG, results in improved predictions of gas accretion onto galaxies and the mass of the hot halo. This suggests that SAMs need to move beyond the traditional “cold mode” vs. “hot mode” dichotomy and instead adopt models that better reflect the full range of gas cooling timescales seen in hydrodynamical simulations (Section 6.1.2, Figure 10).

(ii) *Global star formation is relatively insensitive to the detailed spatial distribution of gas, but galaxy sizes must be properly calibrated.* Although the TNG SAM applies the Kennicutt-Schmidt relation at the scale of the entire galaxy, while TNG applies it locally, the TNG SAM successfully reproduces the global Kennicutt-Schmidt relation seen in TNG, leading to accurate stellar masses, star formation rates and cold gas masses. This suggests that global star formation rates are relatively insensitive to the spatial distribution of gas within galaxies. However, accurate predictions still require proper calibration of galaxy sizes, which are not adequately captured by the angular momentum models used in SAMs. (Section 6.1.3, Figure 9).

(iii) *Modeling outflows at the galaxy- and halo-scale is essential for accurately predicting the gas content in galaxies.* While galaxy-scale outflows are a well-known driver of galaxy evolution, the TNG SAM highlights that halo-scale outflows are equally critical for accurately predicting gas content. Incorporating TNG’s galactic winds model, which ties outflow strength to the metallicity of the cold gas, allowed the TNG SAM to capture not just the behavior of galaxy-scale outflows, but also their continuation into the halo—a channel neglected in most traditional SAMs. This two-step outflow process, where gas flows from the ISM to the CGM and from the CGM to the IGM, provides a more realistic depiction of outflows and avoids the overprediction of hot gas mass and baryon fraction. (Section 6.1.4, Figure 9).

(iv) *Non-uniform metal flows between the ISM and CGM leads to more realistic predictions of galaxy metallicities.* The TNG SAM’s use of metallicity-weighted mass-loading factors significantly improves its ability to match the metallicity evolution observed in TNG. This approach compensates for the simplified instantaneous recycling model used in most SAMs, which typically overlook the delayed mixing and transport of metals. By tying feedback strength to gas metallicity, as done in TNG and high-resolution simulations like FIRE and TIGRESS, the TNG SAM demonstrates that future SAMs aiming to adopt more realistic feedback models will need to incorporate metal loadings to properly capture metal cycling processes between the ISM, CGM, and IGM (Section 6.1.5, Figure 9).

## ACKNOWLEDGEMENTS

OO acknowledges support from the National Science Foundation Graduate Research Fellowship.

*Software:* ASTROPY Robitaille et al. (2013); Collaboration et al. (2018, 2022), IPYTHON (Pérez & Granger 2007), MATPLOTLIB (Caswell et al. 2022), NUMPY (van der Walt et al. 2011), SCIPY (Virtanen et al. 2020)

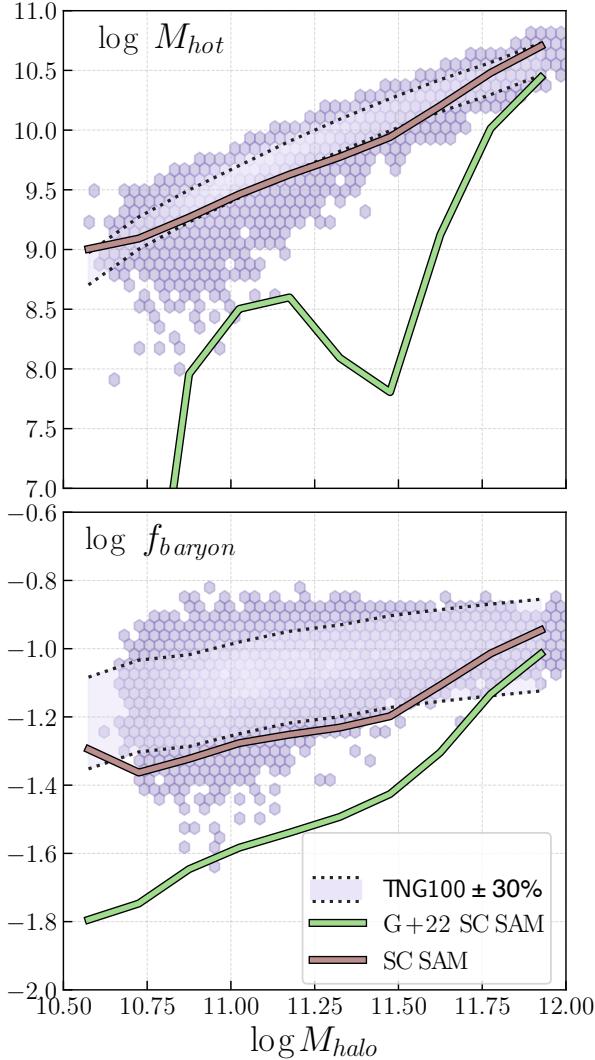
## DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author

## REFERENCES

- Anderson L. D., Bania T. M., Balser D. S., Cunningham V., Wenger T. V., Johnstone B. M., Armentrout W. P., 2014, *The Astrophysical Journal Supplement Series*, 212, 1
- Behroozi P. S., Wechsler R. H., Wu H.-Y., 2012, *The Astrophysical Journal*, 762, 109
- Benson A. J., 2012, *New Astronomy*, 17, 175
- Benson A. J., Pearce F. R., Frenk C. S., Baugh C. M., Jenkins A., 2001, *Monthly Notices of the Royal Astronomical Society*, 320, 261
- Bernardi M., Meert A., Sheth R. K., Vikram V., Huertas-Company M., Mei S., Shankar F., 2013, *Monthly Notices of the Royal Astronomical Society*, 436, 697
- Bigiel F., Leroy A., Walter F., Brinks E., Blok W. J. G. d., Madore B., Thornley M. D., 2008, *The Astronomical Journal*, 136, 2846
- Boselli A., Cortese L., Boquien M., Boissier S., Catinella B., Lagos C., Saintonge A., 2014, *Astronomy & Astrophysics*, 564, A66
- Bouché N., et al., 2010, *The Astrophysical Journal*, 718, 1001
- Bower R. G., Benson A. J., Crain R. A., 2012, *Monthly Notices of the Royal Astronomical Society*, 422, 2816
- Calette A. R., Avila-Reese V., Rodríguez-Puebla A., Hernández-Toledo H., Papastergis E., 2018, *Revista Mexicana de Astronomía y Astrofísica*, 54, 443
- Carr C., Bryan G. L., Fielding D. B., Pandya V., Somerville R. S., 2023, *The Astrophysical Journal*, 949, 21
- Caswell T. A., et al., 2022, matplotlib/matplotlib: REL: v3.5.3, doi:10.5281/zenodo.6982547, <https://zenodo.org/record/6982547>
- Cattaneo A., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, 471, 1401
- Cole S., 1991, *The Astrophysical Journal*, 367, 45
- Collaboration T. A., et al., 2018, *The Astronomical Journal*, 156, 123
- Collaboration T. A., et al., 2022, *The Astrophysical Journal*, 935, 167
- Collacchioni F., Cora S. A., Lagos C. D. P., Vega-Martínez C. A., 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 954
- Croton D. J., et al., 2006, *Monthly Notices of the Royal Astronomical Society*, 365, 11
- Croton D. J., et al., 2016, *The Astrophysical Journal Supplement Series*, 222, 22
- Côté B., Silvia D. W., O’Shea B. W., Smith B., Wise J. H., 2018, *The Astrophysical Journal*, 859, 67
- Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *The Astrophysical Journal*, 292, 371
- Davé R., Finlator K., Oppenheimer B. D., 2012, *Monthly Notices of the Royal Astronomical Society*, 421, 98
- Diemer B., et al., 2018, *The Astrophysical Journal Supplement Series*, 238, 33
- Donnari M., et al., 2019, *Monthly Notices of the Royal Astronomical Society*, 485, 4817
- Gabrielpillai A., Somerville R. S., Genel S., Rodriguez-Gomez V., Pandya V., Yung L. Y. A., Hernquist L., 2022, *Monthly Notices of the Royal Astronomical Society*, 517, 6091
- Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *Monthly Notices of the Royal Astronomical Society*, 362, 41
- Genel S., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 474, 3976
- Gnedin N. Y., Kravtsov A. V., 2011, *The Astrophysical Journal*, 728, 88
- Guo Q., et al., 2011, *Monthly Notices of the Royal Astronomical Society*, 413, 101
- Guo Q., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 3457
- Helly J. C., Cole S., Frenk C. S., Baugh C. M., Benson A., Lacey C., Pearce F. R., 2003, *Monthly Notices of the Royal Astronomical Society*, 338, 913
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, *Monthly Notices of the Royal Astronomical Society*, 431, 3373
- Henriques B., White S., Thomas P., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, *Monthly Notices of the Royal Astronomical Society*, 451, 2663
- Hirschmann M., De Lucia G., Fontanot F., 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 1760
- Hopkins P. F., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 800
- Hopkins P. F., et al., 2023, *Monthly Notices of the Royal Astronomical Society*, 519, 3154
- Karmakar T., Genel S., Somerville R. S., 2023, *Monthly Notices of the Royal Astronomical Society*, 520, 1630
- Kennicutt Jr. R. C., 1998, *Apj*, 498, 541
- Keres D., Yun M. S., Young J. S., 2003, *The Astrophysical Journal*, 582, 659
- Kim C.-G., et al., 2020a, *The Astrophysical Journal*, 900, 61
- Kim C.-G., et al., 2020b, *The Astrophysical Journal Letters*, 903, L34
- Klypin A. A., Trujillo-Gomez S., Primack J., 2011, *The Astrophysical Journal*, 740, 102
- Kravtsov A. V., 2003, *The Astrophysical Journal*, 590, L1
- Kumari N., Irwin M. J., James B. L., 2020, *Astronomy & Astrophysics*, 634, A24
- Lacey C., Silk J., 1991, *The Astrophysical Journal*, 381, 14
- Lacey C. G., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, 462, 3854
- Lagos C. d. P., Lacey C. G., Baugh C. M., 2013, *Monthly Notices of the Royal Astronomical Society*, 436, 1787
- Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, *Monthly Notices of the Royal Astronomical Society*, 481, 3573
- Li Z.-Y., Nakamura F., 2006, *Apj*, 640, L187
- Lu Y., Kereš D., Katz N., Mo H. J., Fardal M., Weinberg M. D., 2011, *Monthly Notices of the Royal Astronomical Society*, 416, 660
- Marinacci F., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 480, 5113
- McConnell N. J., Ma C.-P., 2013, *The Astrophysical Journal*, 764, 184
- Mitchell P. D., Schaye J., 2022, *Monthly Notices of the Royal Astronomical Society*, 511, 2948
- Mitchell P. D., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 474, 492
- Mitchell P. D., Schaye J., Bower R. G., Crain R. A., 2020, *Monthly Notices of the Royal Astronomical Society*, 494, 3971
- Monaco P., Fontanot F., Taffoni G., 2007, *Monthly Notices of the Royal Astronomical Society*, 375, 1189
- Monaco P., Benson A. J., De Lucia G., Fontanot F., Borgani S., Boylan-Kolchin M., 2014, *Monthly Notices of the Royal Astronomical Society*, 441, 2058
- Naiman J. P., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 477, 1206
- Narayanan D., Krumholz M. R., Ostriker E. C., Hernquist L., 2012, *Monthly Notices of the Royal Astronomical Society*, 421, 3127
- Navarro J. F., Frenk C. S., White S. D. M., 1996, *The Astrophysical Journal*, 462, 563
- Neistein E., Weinmann S. M., 2010, *Monthly Notices of the Royal Astronomical Society*, 405, 2717
- Neistein E., Khochfar S., Dalla Vecchia C., Schaye J., 2012, *Monthly Notices of the Royal Astronomical Society*, 421, 3579
- Nelson D., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 624
- Nelson D., et al., 2019, *Monthly Notices of the Royal Astronomical Society*, 490, 3234
- Obreschkow D., Croton D., Lucia G. D., Khochfar S., Rawlings S., 2009, *The Astrophysical Journal*, 698, 1467
- Okamoto T., Gao L., Theuns T., 2008, *Monthly Notices of the Royal Astronomical Society*, 390, 920

- Pakmor R., Springel V., 2013, *Monthly Notices of the Royal Astronomical Society*, 432, 176
- Pakmor R., Bauer A., Springel V., 2011, *Monthly Notices of the Royal Astronomical Society*, 418, 1392
- Pandya V., 2021, PhD thesis, UC Santa Cruz, <https://escholarship.org/uc/item/9xc1v7c9>
- Pandya V., et al., 2020, *The Astrophysical Journal*, 905, 4
- Pandya V., et al., 2021, *Monthly Notices of the Royal Astronomical Society*, 508, 2979
- Pandya V., et al., 2023, *The Astrophysical Journal*, 956, 118
- Peeples M. S., Shankar F., 2011, *Monthly Notices of the Royal Astronomical Society*, 417, 2962
- Peeples M. S., Werk J. K., Tumlinson J., Oppenheimer B. D., Prochaska J. X., Katz N., Weinberg D. H., 2014, *The Astrophysical Journal*, 786, 54
- Pillepich A., et al., 2018a, *Monthly Notices of the Royal Astronomical Society*, 473, 4077
- Pillepich A., et al., 2018b, *Monthly Notices of the Royal Astronomical Society*, 475, 648
- Pillepich A., et al., 2019, *Monthly Notices of the Royal Astronomical Society*, 490, 3196
- Planck Collaboration et al., 2016, *Astronomy & Astrophysics*, 594, A13
- Popping G., Somerville R. S., Trager S. C., 2014, *Monthly Notices of the Royal Astronomical Society*, 442, 2398
- Popping G., et al., 2019, *The Astrophysical Journal*, 882, 137
- Porter L. A., Somerville R. S., Primack J. R., Johansson P. H., 2014, *Monthly Notices of the Royal Astronomical Society*, 444, 942
- Pérez F., Granger B. E., 2007, *Computing in Science & Engineering*, 9, 21
- Robitaille T. P., et al., 2013, *Astronomy & Astrophysics*, 558, A33
- Rodríguez-Gómez V., et al., 2015, *Monthly Notices of the Royal Astronomical Society*, 449, 49
- Rodríguez-Puebla A., Primack J. R., Avila-Reese V., Faber S. M., 2017, *Monthly Notices of the Royal Astronomical Society*, 470, 651
- Saro A., De Lucia G., Borgani S., Dolag K., 2010, *Monthly Notices of the Royal Astronomical Society*, 406, 729
- Somerville R. S., Davé R., 2015, *Annual Review of Astronomy and Astrophysics*, 53, 51
- Somerville R. S., Primack J. R., 1999, *Monthly Notices of the Royal Astronomical Society*, 310, 1087
- Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008a, *Monthly Notices of the Royal Astronomical Society*, 391, 481
- Somerville R. S., et al., 2008b, *The Astrophysical Journal*, 672, 776
- Somerville R. S., Gilmore R. C., Primack J. R., Domínguez A., 2012, *Monthly Notices of the Royal Astronomical Society*, 423, 1992
- Somerville R. S., Popping G., Trager S. C., 2015, *Monthly Notices of the Royal Astronomical Society*, 453, 4337
- Springel V., 2010, *Monthly Notices of the Royal Astronomical Society*, 401, 791
- Springel V., Hernquist L., 2003, *Monthly Notices of the Royal Astronomical Society*, 339, 289
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *Monthly Notices of the Royal Astronomical Society*, 328, 726
- Springel V., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, 475, 676
- Stringer M. J., Brooks A. M., Benson A. J., Governato F., 2010, *Monthly Notices of the Royal Astronomical Society*, 407, 632
- Sutherland R. S., Dopita M. A., 1993, *The Astrophysical Journal Supplement Series*, 88, 253
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *Monthly Notices of the Royal Astronomical Society*, 438, 1985
- Virtanen P., et al., 2020, *Nature Methods*, 17, 261
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *Monthly Notices of the Royal Astronomical Society*, 436, 3031
- Voit G. M., Pandya V., Fielding D. B., Bryan G. L., Carr C., Donahue M., Oppenheimer B. D., Somerville R. S., 2024a, Equilibrium States of Galactic Atmospheres I: The Flip Side of Mass Loading, doi:10.48550/arXiv.2406.07631, <http://arxiv.org/abs/2406.07631>
- Voit G. M., Carr C., Fielding D. B., Pandya V., Bryan G. L., Donahue M., Oppenheimer B. D., Somerville R. S., 2024b, Equilibrium States of Galactic Atmospheres II: Interpretation and Implications, doi:10.48550/arXiv.2406.07632, <http://arxiv.org/abs/2406.07632>
- Weinberger R., et al., 2017, *Monthly Notices of the Royal Astronomical Society*, 465, 3291
- White S. D. M., Frenk C. S., 1991, *The Astrophysical Journal*, 379, 52
- White S. D. M., Rees M. J., 1978, *Monthly Notices of the Royal Astronomical Society*, 183, 341
- Wise J. H., Turk M. J., Norman M. L., Abel T., 2011, *The Astrophysical Journal*, 745, 50
- Wright R. J., Lagos C. d. P., Power C., Mitchell P. D., 2020, *Monthly Notices of the Royal Astronomical Society*, 498, 1668
- Wright R. J., Somerville R. S., Lagos C. d. P., Schaller M., Davé R., Anglés-Alcázar D., Genel S., 2024, The baryon cycle in modern cosmological hydrodynamical simulations, doi:10.48550/arXiv.2402.08408, <http://arxiv.org/abs/2402.08408>
- Yoshida N., Stoehr F., Springel V., White S. D. M., 2002, *Monthly Notices of the Royal Astronomical Society*, 335, 762
- Yung L. Y. A., Somerville R. S., Finkelstein S. L., Popping G., Davé R., 2019, *Monthly Notices of the Royal Astronomical Society*, 483, 2983
- Yung L. Y. A., et al., 2023, *Monthly Notices of the Royal Astronomical Society*, 519, 1578
- Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, *The Astrophysical Journal Letters*, 771, L19
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Computing in Science & Engineering*, 13, 22



**Figure A1.** Comparison of the Santa Cruz SAM used in this work (solid brown line) to TNG100 (purple shaded region) and the SC SAM from [Gabrielpillai et al. \(2022\)](#) (green line) for hot halo gas mass ( $M_{\text{hot}}$ , top) and baryon fraction ( $f_{\text{baryon}}$ , bottom) as a function of halo mass. The modified cooling model described in Section 2.2.5 substantially improves agreement with TNG100 for halo-scale properties.

## A2 The Impact of Consistent Radial Definitions

While creating the TNG SAM, it became clear that the definitions of key mass reservoirs—such as  $M_{\text{hot}}$ ,  $M_{\text{cold}}$ , and  $M_{\star}$ —vary significantly in the literature, particularly when comparing the outputs of SAMs and hydrodynamical simulations. For example, one can define ... Inconsistent definitions can introduce substantial biases, particularly for quantities like gas masses, where the choice of definition directly impacts the results.

## A3 Analytic Scaling relations

We alternatively calibrated the SC SAM using analytic scaling relations as a function of halo mass and redshift using the following sophisticated functional form:

$$\log(f(M_{\text{halo}}, z, b, c, d, \alpha_0, \alpha_z)) = -a(\arctan(b \cdot [\log(M_{\text{halo}}) - c]) + d) \cdot [\alpha_0 + e^{\alpha_z}(1+z)] \quad (\text{A1})$$

Using these scaling relations also provided agreements with TNG within the same range as the linearly interpolated values. All coefficients used in Equation A1 are described in Table A2.

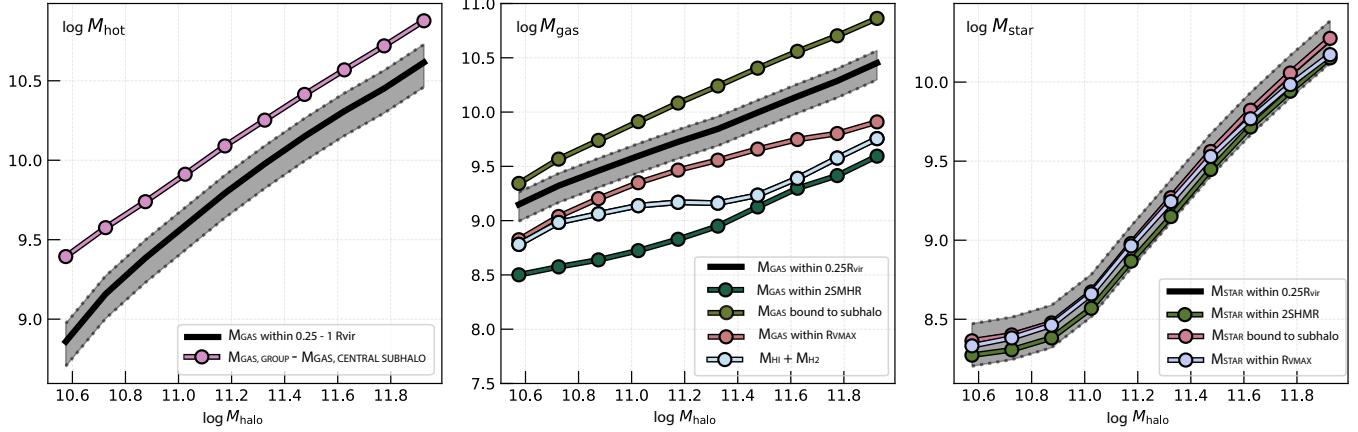
## A4 Results of the TNG SAM with 100 Galaxies

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

## APPENDIX A: CREATION OF THE TNG SAM

### A1 Modifications to the SC SAM Presented

The modifications described in Section 2.2.5 had the most substantial impact on the halo-scale properties of the Santa Cruz SAM. The updated cooling model significantly improved the prediction of the hot halo gas mass, previously underestimated by several orders of magnitude, as shown in Figure A1 and first noted in [Pandya et al. \(2020\)](#). Naturally, this correction also led to a substantial improvement in the SAM’s prediction of the baryon fraction, illustrated in the lower panel of Figure A1. Although these changes drastically affected the halo-scale gas properties, the impact on the ISM-scale gas properties and overall metal content was minimal, with deviations from the Santa Cruz SAM published in [Gabrielpillai et al. \(2022\)](#) (G+22) reaching a maximum of 0.5 dex.



**Figure A2.** Comparison of  $M_{\text{hot}}$ ,  $M_{\text{cold}}$ , and  $M_{\text{star}}$  using different definitions in TNG at  $z = 0$ . The black solid line shows the definitions adopted in this paper ( $M_{\text{hot}}$  within 0.25–1  $r_{\text{vir}}$ ,  $M_{\text{cold}}$  within 0.25  $r_{\text{vir}}$ , and  $M_{\text{star}}$  within 0.25  $r_{\text{vir}}$ ), with shaded regions indicating deviations within 30%. Colored lines represent alternate definitions commonly used in the literature, showing significant variations for  $M_{\text{hot}}$  and  $M_{\text{cold}}$  (up to 50%) and smaller deviations for  $M_{\text{star}}$  (within 30%).

Scale	Parameter	SAM Label	TNG Label	Description
<i>Halo</i>	$M_{\text{halo}}$	GalpropMvir	Group_M_TopHat_200['SubhaloGrNr']	Total virial mass
	$M_{\text{DM}}$	GalpropMvir ( $1 - f_{\text{baryon}}$ )	SubhaloMassType Part 1 (DM)	Total dark matter mass
	$M_{\text{hot}}$	HalopropMhot	$\sum_i^n m_{\text{gas}}^i$ for $0.25r_{\text{vir}} < r < r_{\text{vir}}$	Gas mass of the circumgalactic halo
<i>Galaxy</i>	$M_{*}$	GalpropMstar	$\sum_i^n m_{\text{star}}^i$ for $r < 0.25r_{\text{vir}}$	Stellar mass
	$M_{\text{cold}}$	GalpropMcold	$\sum_i^n m_{\text{gas}}^i$ for $r < 0.25r_{\text{vir}}$	Gas mass of the interstellar medium
	SFR	GalpropSfr	$\sum_i^n m_{\text{star}}^i$ for $r < 0.25r_{\text{vir}}$	Star formation rate
	$< Z_{*} >$	GalpropZstar	$\frac{1}{n} \sum_i^n Z_{\text{star}}^i$ for $r < 0.25r_{\text{vir}}$	Average stellar metallicity
	$< Z_{\text{cold}} >$	GalpropZcold	$\frac{1}{n} \sum_i^n Z_{\text{gas}}^i$ for $r < 0.25r_{\text{vir}}$ ,	Average cold gas metallicity
	$< Z_{\text{hot}} >$	HalopropZhot	$\frac{1}{n} \sum_i^n Z_{\text{gas}}^i$ for $0.25r_{\text{vir}} < r < r_{\text{vir}}$	Average hot halo gas metallicity

**Table A1.** Correspondence between the parameters studied in this paper, their fields in the SAM, and their corresponding fields in the IllustrisTNG simulations. A description of each parameter is also provided.

Model Update	Parameter	Description	Coefficients (a, b, c, d, $\alpha_0$ , $\alpha_z$ )	Comments
Gas Cooling	$f_{\text{in,CGM}}$		1.15, 0.0088, 10.2189, 0.0454, -3.5025, -1.0791	
	$t_{\text{cool}}$		1.15, 9.4401, 11.5681, 359.5587, -0.0275, -7.4258	
Star Formation	$R_{\text{disk,TNG}}/R_{\text{disk,SAM}}$		1.0, 0.0009, 9.4550, -0.0028, 9.2402, 4.2137	
	$\eta_{\text{launch}}/\eta_{\text{ISM}}$		1.2, 6.7252, 11.6654, 4.6427, -0.1362, -3.7849	
Stellar Feedback	$\eta_{\text{halo}}$		0.85, -1.6941, 10.8106, 1.7008, -0.7913, -3.1102	
	$\zeta_{\text{in,halo}}$		1.0, -5.0310, 11.3533, 7.8483, 0.1622, -4.3837	
Metal Circulation	$\zeta_{\text{out,halo}}$		1.0, 4.0473, 11.3201, -0.0693, -0.1187, -2.9203	
	$\zeta_{\text{in,ISM}}$		1.0, 1.1155, 10.0801, -1.4657, 0.0161, -2.9189	
	$\zeta_{\text{out,ISM}}$		1.0, 48.4092, -62.9615, 5.9549, 0.0501, -18.1282	$z \leq 1$
	...		1.0, 7.6844, 13.7629, 1.4179, -4.2491, -0.8384	$z > 1$

**Table A2.** Parameters calibrated to match TNG and the coefficients used in Equation A1. (1) Parameter; (2) SAM Label; (3) TNG Label; (4) comment.