

# How DREAMS are made: Emulating Satellite Galaxy and Subhalo Populations with Diffusion Models and Point Clouds

TRI NGUYEN 1, 2, 3, 4 FRANCISCO VILLAESCUSA-NAVARRO 5, 6 SIDDHARTH MISHRA-SHARMA 1, 3, 7, 8, \*  
 CAROLINA CUESTA-LAZARO 1, 3, 9 PAUL TORREY 10 ARYA FARAHI 11 ALEX M. GARCIA 10 JONAH C. ROSE 5, 12  
 STEPHANIE O'NEIL 1, 2 MARK VOGELSSINGER 1, 2, 3 XUEJIAN SHEN 1, 2 CIAN ROCHE 1, 2 DANIEL ANGLÉS-ALCÁZAR 13, 5  
 NITYA KALLIVAYALIL 10 JULIAN B. MUÑOZ 14 FRANCIS-YAN CYR-RACINE 15 SANDIP ROY 16 LINA NECIB 1, 2, 3 AND  
 KASSIDY E. KOLLMANN 16

<sup>1</sup>Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>2</sup>Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>3</sup>The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, Cambridge, MA 02139, USA

<sup>4</sup>Department of Physics and Astronomy and CIERA, Northwestern University, 2145 Sheridan Road, Evanston, IL 60208

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

<sup>6</sup>Department of Astrophysical Sciences, Princeton University, Peyton Hall, Princeton, NJ 08544, USA

<sup>7</sup>Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>8</sup>Department of Physics, Harvard University, Cambridge, MA 02138, USA

<sup>9</sup>Center for Astrophysics, Harvard & Smithsonian, Cambridge, MA 02138, USA

<sup>10</sup>Department of Astronomy, University of Virginia, 530 McCormick Road, Charlottesville, VA 22904 USA

<sup>11</sup>Departments of Statistics and Data Science, University of Texas at Austin, Austin, TX 78757, USA

<sup>12</sup>Department of Astronomy, University of Florida, Gainesville, FL 32611, USA

<sup>13</sup>Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA

<sup>14</sup>Department of Astronomy, University of Texas at Austin, Austin, TX 78757, USA

<sup>15</sup>Department of Physics and Astronomy, University of New Mexico, 210 Yale Blvd NE, Albuquerque, NM 87106, USA

<sup>16</sup>Department of Physics, Princeton University, Princeton, NJ 08544, USA

## ABSTRACT

The connection between galaxies and their host dark matter (DM) halos is critical to our understanding of cosmology, galaxy formation, and DM physics. To maximize the return of upcoming cosmological surveys, we need an accurate way to model this complex relationship. Many techniques have been developed to model this connection, from Halo Occupation Distribution (HOD) to empirical and semi-analytic models to hydrodynamic. Hydrodynamic simulations can incorporate more detailed astrophysical processes but are computationally expensive; HODs, on the other hand, are computationally cheap but have limited accuracy. In this work, we present NeHOD, a generative framework based on variational diffusion model and Transformer, for painting galaxies/subhalos on top of DM with an accuracy of hydrodynamic simulations but at a computational cost similar to HOD. By modeling galaxies/subhalos as point clouds, instead of binning or voxelization, we can resolve small spatial scales down to the resolution of the simulations. For each halo, NeHOD predicts the positions, velocities, masses, and concentrations of its central and satellite galaxies. We train NeHOD on the TNG-Warm DM suite of the DREAMS project, which consists of 1024 high-resolution zoom-in hydrodynamic simulations of Milky Way-mass halos with varying warm DM mass and astrophysical parameters. We show that our model captures the complex relationships between subhalo properties as a function of the simulation parameters, including the mass functions, stellar-halo mass relations, concentration-mass relations, and spatial clustering. Our method can be used for a large variety of downstream applications, from galaxy clustering to strong lensing studies.

**Keywords:** Galactic and extragalactic astronomy (563) — Warm dark matter (1787)

## 1. INTRODUCTION

Corresponding author: Tri Nguyen  
 tnguy@mit.edu

\* Currently at Anthropic; worked performed while at MIT/IAIFI.

The abundance and spatial clustering of galaxies and dark matter (DM) halos is sensitive to assumptions about the underlying galaxy formation physics and cosmological model, including the nature of DM. Warm DM (WDM) models, for ex-

ample, will predict fewer massive subhalos compared to cold DM due to free-streaming in the early Universe, thereby affecting observable properties such as the satellite abundance, their luminosity functions, and spatial distribution (e.g., [Bode et al. 2001](#); [Lovell et al. 2014](#); [Bose et al. 2016](#)). Likewise, baryonic processes, including stellar winds, feedback from supernovae (SN) and active galactic nuclei (AGN), can disrupt the formation of less massive subhalos and redistribute matter within halos, thus altering galaxy population properties in similar ways (e.g., [Governato et al. 2010](#); [Teyssier et al. 2011](#); [Peirani et al. 2017, 2019](#); [Engler et al. 2021](#); [Gebhardt et al. 2023](#)). The properties of galaxies and their distributions thus serve as powerful probes for DM models and galaxy formation physics.

To leverage these properties, several current surveys are designed to measure these cosmic structures, ranging from large-scale structures to deep imaging of nearby galaxies. Surveys such as the Sloan Digital Sky Survey (SDSS; [York et al. 2000](#)), Dark Energy Survey (DES; [The Dark Energy Survey Collaboration 2005](#); [Abbott et al. 2018, 2021](#)), Dark Energy Spectroscopic Instrument (DESI; [DESI Collaboration et al. 2016](#)), and Euclid ([Euclid Collaboration et al. 2022](#)) map the 3-dimensional distributions and clustering of billions of galaxies. Surveys such as Satellites Around Galactic Analogs (SAGA; [Geha et al. 2017](#); [Mao et al. 2021, 2024](#); [Geha et al. 2024](#); [Wang et al. 2024](#)) and Exploring the Local Volume in the Extended Solar neighborhood (ELVES; [Carlsten et al. 2022](#)) have characterized hundreds of satellite dwarf galaxies around the Milky Way and nearby galaxies. In addition, upcoming surveys such as the Rubin Observatory’s Legacy Survey of Space and Time (LSST; [Ivezić et al. 2019](#)) and Roman Space Telescope ([Akeson et al. 2019](#)) will have the capacity to look deeper and wider, detecting fainter galaxies and providing a more complete sample of galaxies and their distributions. To maximize the information we can extract from these surveys (through, e.g., galaxy clustering or strong lensing studies) and obtain robust constraints on cosmology, astrophysics, and DM properties, we must develop accurate theoretical predictions for galaxies/subhalos as a function of such properties.

Hydrodynamic simulations are the primary method in modern astrophysics for generating theoretical predictions about the phase-space distribution of galaxies and their intrinsic properties (e.g., stellar mass, metallicity, neutral hydrogen mass, etc.). These simulations model the evolution of galaxies and DM halos by solving fluid dynamics and gravitational equations within an expanding universe, starting from the initial density fluctuations derived from measurements of the Cosmic Microwave Background (for a review, see [Vogelsberger et al. 2020](#)). Hydrodynamic simulations incorporate detailed prescriptions of baryonic processes (e.g., gas cooling, star formation, stellar/AGN feedback, etc.) and thus have

been instrumental in understanding galaxies and their DM halos. Key predictions of these simulations range from the large-scale distribution of DM and galaxies (e.g., [Springel et al. 2006](#); [Schaye et al. 2015](#); [Springel et al. 2018](#); [Pillepich et al. 2018a](#); [Marinacci et al. 2018](#)), to the properties of the intergalactic and interstellar medium (e.g., [Hernquist et al. 1996](#); [Gnedin & Kaurov 2014](#); [Ocvirk et al. 2016](#); [Rosdahl et al. 2018](#); [Naiman et al. 2018](#); [Kannan et al. 2022](#); [Smith et al. 2022](#)), down to the internal structure of DM halos (e.g., [Navarro et al. 1996b,a](#)), the properties of individual galaxies (e.g., [Vogelsberger et al. 2014a](#); [Wang et al. 2015](#); [Kaviraj et al. 2017](#); [Grand et al. 2017](#); [Hopkins et al. 2018](#); [Nelson et al. 2018](#)), and the populations of satellite galaxies (e.g., [Wetzel et al. 2016](#); [Garrison-Kimmel et al. 2017](#); [Fitts et al. 2017](#); [Kim et al. 2018](#); [Simpson et al. 2018](#); [Buck et al. 2019](#); [Sales et al. 2022](#)).

To deepen our understanding of the physics governing DM and galaxy formation, we want to thoroughly explore how observed galaxy properties (e.g., abundance and spatial distribution) vary across different DM and astrophysical parameters. Ideally, this would involve conducting a large set of large-volume, high-resolution hydrodynamic simulations that cover this extensive parameter space. However, hydrodynamic simulations are computationally demanding, which limits the simulated volume and/or resolution. Large-volume simulations such as IllustrisTNG ([Pillepich et al. 2018b](#); [Weinberger et al. 2017](#)) and MillenniumTNG ([Springel et al. 2005](#); [Hernández-Aguayo et al. 2023](#)) are constrained by mass resolution and cannot adequately capture the properties of low-mass and/or high-redshift galaxies. Conversely, zoom-in simulations such as FIRE ([Hopkins et al. 2014](#)), NIHAO ([Wang et al. 2015](#)), APOSTLE ([Sawala et al. 2016](#)), and Auriga ([Grand et al. 2017](#)) provide higher resolution but have smaller volumes, resulting in a more limited number of simulated galaxies. In addition, zoom-in simulations have difficulty modeling very massive halos given the smaller box sizes.

These challenges demonstrate the need for developing more efficient methods to study the relationship between galaxies and their DM halos. A common way is to map galaxy properties onto DM halos simulated in N-body (i.e., gravity-only) simulations, which have a much lower computational cost than their full hydrodynamic counterparts. For example, semi-analytic models, which use simplified baryonic physics prescriptions, and empirical models, which use empirical relationships from observational data (e.g., stellar-to-halo mass relations), can be used to assign galaxy properties to DM halos (see [Wechsler & Tinker 2018](#), for a review).

In this paper, we focus on Halo Occupation Distribution (HOD) methods, in which galaxies are assigned to DM halos in N-body simulations based on a statistical framework that describes the probability of a halo hosting a certain number

of galaxies (Neyman & Scott 1952). This process, known as “painting” galaxies onto DM halos, follows a set of recipes that aim to capture the main physics describing the halo-galaxy connection. HODs have been applied in various contexts, from studying galaxy clustering (Benson et al. 2000; Peacock & Smith 2000; Hearin et al. 2016) and the halo-galaxy connection (Yang et al. 2003; Conroy & Wechsler 2009; Croton 2009; Wechsler & Tinker 2018), to generating mock galaxy catalogs for recent surveys such as DESI (Smith et al. 2024) and Euclid (Hamaus et al. 2022). Despite these successes, since HOD methods rely on simplified assumptions about the halo-galaxy connection and are designed to match specific summary statistics (e.g., galaxy abundance and 2-point correlation function), they may fail for other summary statistics or probes of the entire observed distribution (e.g., field-level analyses) (e.g., Schmittfull et al. 2021; Kokron et al. 2021; Wu et al. 2022b; Hadzhiyska et al. 2023).

Ultimately, we want the accuracy of full hydrodynamic simulations at the computational cost of HODs. In principle, machine learning models can reproduce and extend the work of traditional HODs, by not only learning the connection between halos and their constituent galaxies but also linking halo occupation to the underlying assumed physics. Generative models, also called emulators, are particularly well-suited for this task as they learn directly from the outputs of simulations and generate new data samples, bypassing the need for summary statistics. Normalizing flows (Jimenez Rezende & Mohamed 2015; Papamakarios et al. 2017; Papamakarios et al. 2019), for example, have been used for HOD modeling by either painting galaxy properties on top of subhalos or modeling the position and velocity offsets between subhalos and galaxies (see e.g., Lovell et al. 2023a; Kwon & Hahn 2024). Recent developments in diffusion models (Ho et al. 2020; Song et al. 2020; Kingma et al. 2021) open up promising avenues for exploration, as these models are typically more expressive than flow-based models. In recent work, Cuesta-Lazaro & Mishra-Sharma (2023), hereafter CM23, employs diffusion and point clouds to generate large-scale distributions of DM halos as a function of cosmological parameters. Another recent work, Bourdin et al. (2024), combines convolutional neural networks and diffusion to predict the distributions of galaxy counts from DM fields and cosmological parameters, and demonstrates that such an approach outperforms HOD modeling.

In this paper, we explore how diffusion models and point clouds can be used to generate galaxies/subhalos within a single DM halo at field level, across different DM properties and baryonic physics scenarios. We introduce the NeHOD framework, short for Neural Halo Occupation Distribution. Distinct from most HODs and previous emulators, NeHOD employs a point-cloud approach to represent satellite galaxies (instead of relying on binning or voxelization), which allows it to resolve

small spatial scales down to the resolution of the simulations. In addition, NeHOD can take into account intra-galaxy (i.e., properties of a single galaxy) as well as inter-galaxy (i.e., between galaxies in a catalog) properties, since each point cloud/satellite galaxy population is generated at once using a Variational Diffusion Model with a Transformer-based noise prediction model (Vaswani et al. 2023). The primary goal of NeHOD is to generate full galaxy catalogs, including the positions, velocities, and internal properties (such as mass and concentration) of halos, central galaxies, and satellite galaxies, given a set of simulation parameters. Thus, to complement the diffusion model, NeHOD uses normalizing flows (Jimenez Rezende & Mohamed 2015; Papamakarios et al. 2017; Papamakarios et al. 2019) to characterize halos and central galaxies.

We train our model using data from the DREAMS project (Rose et al. 2024a), which contains high-resolution, zoom-in simulations of Milky Way-mass halos. Specifically, we utilize the TNG-Warm Dark Matter (WDM) zoom-in suite of simulations, in which each simulation varies over the astrophysical parameters (the strength of SN and AGN feedback) and the mass of the WDM particles. When WDM decouples at early times, its free-streaming velocity is relativistic. These high speeds allow WDM particles to escape small potential wells, thus suppressing small-scale power (Hogan & Dalcanton 2000; Sommer-Larsen & Dolgov 2001). Consequently, WDM can influence properties of both halos and galaxies, such as the number of satellite galaxies, the satellite mass function, etc. (Bode et al. 2001; Lovell et al. 2014; Bose et al. 2016; Lovell et al. 2016, 2020). Additionally, WDM can affect the internal properties of galaxies (e.g., Macciò & Fontanot 2010; Lovell et al. 2012), but only at sub-kpc scales for allowed WDM particle masses in the TNG-WDM suite.

By employing the TNG-WDM suite, we aim to capture variations due to both astrophysical processes and DM physics, as well as halo-to-halo variance, and demonstrate NeHOD’s capability to model the complex relationships between halos, galaxies, and these parameters. We highlight that emulating halos and galaxies under alternative DM models is, by itself, novel and interesting for understanding the impact of DM on galaxy formation at small scales and placing constraints on its properties. However, we will explore these implications in future work, as this paper focuses primarily on the methodology and validation of our model.

Once trained, NeHOD can generate a catalog that represents galaxies residing in a halo, given some values of the astrophysical parameters and the WDM mass. Our model returns a galaxy catalog with the position, velocity, and internal properties of each galaxy inside the halo. We use multiple summary statistics to quantify the accuracy and precision of the generated galaxy catalogs and compare them with the ones from the simulations. We find NeHOD exhibits

the accuracy of the full hydrodynamic simulations while being orders of magnitude faster. This approach is particularly valuable in any context where populating halos with galaxies is necessary, ranging from galaxy redshift and lensing surveys (Berlind et al. 2003; Zehavi et al. 2011; Hamaus et al. 2022; Smith et al. 2024; Contreras et al. 2024), to near-field cosmology (Jiang et al. 2021; Dropulic et al. in prep), to high-redshift galaxies (Bhowmick et al. 2018; Kayo & Oguri 2012; Reyes-Peraza et al. 2024).

This paper is structured as follows. In §2, we briefly summarize the DREAMS project and the WDM zoom-in suite of simulations used as the training dataset. In §3, we describe the training dataset (in §3.2), the machine learning models (in §3.3), and the training procedure (in §3.4). In §4, we show examples of generated halos and galaxies and quantify the fidelity of the trained emulator model. In §5, we compare NeHOD to traditional modeling techniques and other emulator frameworks, and discuss its current limitations and future prospects. Finally, we state our conclusions in §6.

## 2. SIMULATIONS

### 2.1. The DREAMS project

Introduced in Rose et al. (2024a), the DREAMS project<sup>1</sup> seeks to understand how cosmology, including DM models and baryonic physics, specifically feedback from SN and AGN, influence galaxy formation and evolution, ranging from galaxy clusters to ultra-faint satellites. Prescriptions for SN and AGN feedback are currently among the key sources of systematic uncertainty in cosmological simulations. These feedback mechanisms present significant challenges in distinguishing the effects of baryonic processes from the inherent properties of DM, especially at the Galactic scales, e.g., by altering the inner density profiles in dwarf galaxies (Navarro et al. 1996a; Spekkens et al. 2005; see also Bullock & Boylan-Kolchin 2017 for a review). Ultimately, DREAMS will help understand and disentangle these complex interactions.

The DREAMS suite comprises thousands of state-of-the-art hydrodynamic simulations and is the largest suite of hydrodynamic simulations that vary the DM physics. Similar to the CAMELS project (Villaescusa-Navarro et al. 2021), these simulations vary both cosmological and astrophysical parameters. Notably, these simulations also explore variations in DM particle properties, such as warm DM with different assumed particle masses. The DREAMS suite includes both uniform box cosmological simulations, which provide extensive samples of halos and galaxies, as well as zoom-in cosmological simulations of Milky Way-mass hosts for higher-resolution studies. Among the simulations currently completed are the WDM Uniform Box and Milky Way

suites. For the full specifications of the simulations, we refer readers to Table 1 of Rose et al. (2024a).

Here, we briefly describe the TNG-WDM Milky Way suite used for the machine learning study in this work. The TNG-WDM Milky Way suite contains 1024 simulations of Milky Way-mass halos with varying WDM mass and astrophysical parameters. These simulations are run with the moving-mesh code AREPO (Springel 2010; Weinberger et al. 2020) which accounts for gravity and magneto-hydrodynamics, as well as a range of galaxy formation physics as included in the IllustrisTNG galaxy formation model. The gravitational softening length at  $z = 0$  is  $305 \text{ pc } h^{-1}$ . Cosmological parameters are fixed at  $\Omega_m = 0.301$ ,  $\Omega_\Lambda = 0.698$ ,  $\Omega_b = 0.046$ ,  $\sigma_8 = 0.839$ ,  $h = 0.691$ , which is consistent with the Planck 2016 cosmology (Planck Collaboration et al. 2016). The DM and baryon mass resolution are  $1.2 \times 10^6 \text{ M}_\odot h^{-1}$  and  $1.9 \times 10^5 \text{ M}_\odot h^{-1}$  respectively, which is comparable to the mass resolution of the TNG50 simulation (Nelson et al. 2019; Pillepich et al. 2019), which has a DM and baryon mass resolution of  $6.5 \times 10^5 \text{ M}_\odot h^{-1}$  and  $1.2 \times 10^5 \text{ M}_\odot h^{-1}$ .

The simulations adopt the IllustrisTNG galaxy formation model, as detailed in Weinberger et al. (2017); Pillepich et al. (2018b), which employs a tree and particle mesh (tree-PM) algorithm using periodic boundary conditions to simulate gravity. This model also incorporates baryonic physics, such as AGN feedback, self-consistent magnetohydrodynamics, and an updated prescription of stellar formation and evolution, galactic winds, and outflows from the previous Illustris model (Vogelsberger et al. 2014b; Torrey et al. 2014). The TNG-WDM suite adopts all of the fiducial TNG model parameters, except three parameters that control (i) SN feedback strength, (ii) SN-driven wind velocity, and (iii) AGN feedback strength. For a more detailed discussion of the baryonic prescription in DREAMS, we refer readers to §2.1 of Rose et al. (2024a). Here, we provide a brief description of the varied simulation parameters, including the WDM mass  $m_{\text{WDM}}$  and the three feedback parameters  $\{A_{\text{SN}1}, A_{\text{SN}2}, A_{\text{AGN}}\}$ , and the prior ranged in Table 1. The parameters are distributed uniformly in inverse WDM mass and log uniformly for the feedback parameters. These parameters are varied according to a Sobol sequence (Sobol' 1967), which ensures a quasi-random, well-distributed sampling across the parameter space that is suitable for machine learning applications.

### 2.2. Initial conditions and Milky Way selection criteria

Each simulation zooms in on a high-resolution region that contains a Milky Way-mass halo. The zoom-in procedure includes 3 steps: (1) simulate a low-resolution, DM-only uniform box, (2) randomly select a Milky Way-mass halo and perform a zoom-in DM-only simulation at an intermediate resolution, (3) follow with a final zoom-in with baryons at the fiducial resolution. More details of the zoom-in procedure

<sup>1</sup> <https://www.dreams-project.org/>

Parameter	Description	Range	Prior	TNG Fiducial
$m_{\text{WDM}}$	Mass of WDM particles in keV	[1.8, 30]	Inverse Uniform	$\infty$
$A_{\text{SN}1}$	Specific wind energy of SN wind	[0.9, 14.4]	Log Uniform	3.6
$A_{\text{SN}2}$	Dimensionless scaling factor of SN wind velocity	[3.7, 14.8]	Log Uniform	7.4
$A_{\text{AGN}}$	Fraction of energy transferred to nearby gas due to AGN accretion	[0.025, 0.4]	Log Uniform	0.1

**Table 1.** Descriptions, prior ranges, and their TNG fiducial values of parameters of the 1024 simulations in TNG-WDM suite used in this work.

are outlined in Appendix A of Rose et al. (2024a). Below, we highlight a few key points.

We select the target halo randomly from the uniform box, with virial mass in the range of  $(1.09 - 1.11) \times 10^{12} M_{\odot} h^{-1}$ .<sup>2</sup> This is toward the upper end of the observed Milky Way-mass estimates in Bland-Hawthorn & Gerhard (2016). Additionally, the halo is selected such that it is not in the vicinity of another massive halo (with virial mass greater than  $7.2 \times 10^{11} M_{\odot} h^{-1}$ ) to avoid the high computational costs associated with simulating complex interactions in crowded regions. As a result, the halos used in this work are not direct Local Group-analogs.

Initial conditions are generated by the Multi-Scale Initial Conditions (MUSIC; Hahn & Abel 2011) code at  $z = 127$  using the second-order Lagrangian perturbation theory (2LPT). The initial conditions for the final hydrodynamic zoom-in simulations are derived from intermediate-resolution, DM-only simulations. It is important to note that for each set of simulation parameters (e.g., the WDM mass and astrophysical parameters), a different uniform box was used, and a random target halo was selected from within this box. *This approach ensures a diversity of initial conditions, so each halo in the training dataset represents a unique realization of a Milky Way-mass galaxy.*

Lastly, the halos and subhalos are identified using the friends-of-friends (FoF) and SUBFIND algorithms (Springel et al. 2001) with a linking length of  $b = 0.2$ . The FoF algorithm identifies halos by linking nearby DM particles into groups based on their spatial proximity. SUBFIND then refines these groups by identifying gravitationally bound substructures (subhalos) within the FoF halos. In this work, halo properties are derived from FoF, while subhalo and galaxy properties (whether central or satellite) are from SUBFIND. We will explore the impact of the choice of the halo finder on the emulated halo and subhalo properties in future work. We only include subhalos with at least 100 DM particles, which corresponds to a minimum mass of  $1.2 \times 10^8 M_{\odot} h^{-1}$ ; below this threshold, SUBFIND does not reliably identify subhalos. Additionally, we only consider subhalos that host galaxies,

with stellar mass greater than  $2 \times 10^5 M_{\odot} h^{-1}$ , thus excluding dark subhalos.

### 3. METHODOLOGY

We begin with a brief overview of the NeHOD framework in §3.1. In §3.2, we describe the procedure for creating and preprocessing the dataset as well as the halo properties used as input and output features. We then describe the NeHOD architecture and training strategy in §3.3 and 3.4, respectively.

#### 3.1. Overview of the NeHOD framework

The NeHOD (Neural Halo Occupation Distribution) framework consists of two main components:

1. A normalizing flow that models the conditional likelihood of halos and the galaxies residing at their center (i.e., the central galaxies), including properties such as total mass, stellar mass, and DM concentrations, given the simulation parameters:

$$p(\theta_{\text{halo}}, \theta_{\text{cent}} | \theta_{\text{sim}}), \quad (1)$$

where  $\theta_{\text{halo}}$ ,  $\theta_{\text{cent}}$ , and  $\theta_{\text{sim}}$  are the halo, central galaxy, and simulation parameters respectively.

2. A Variational Diffusion Model (VDM) that models the conditional likelihood of satellite galaxies (e.g., positions, velocities, and internal properties), given the halo, central galaxy, and simulation parameters:

$$p(\theta_{\text{sat}} | \theta_{\text{halo}}, \theta_{\text{cent}}, \theta_{\text{sim}}), \quad (2)$$

where  $\theta_{\text{sat}}$  are the properties of *all* satellites within the halo.

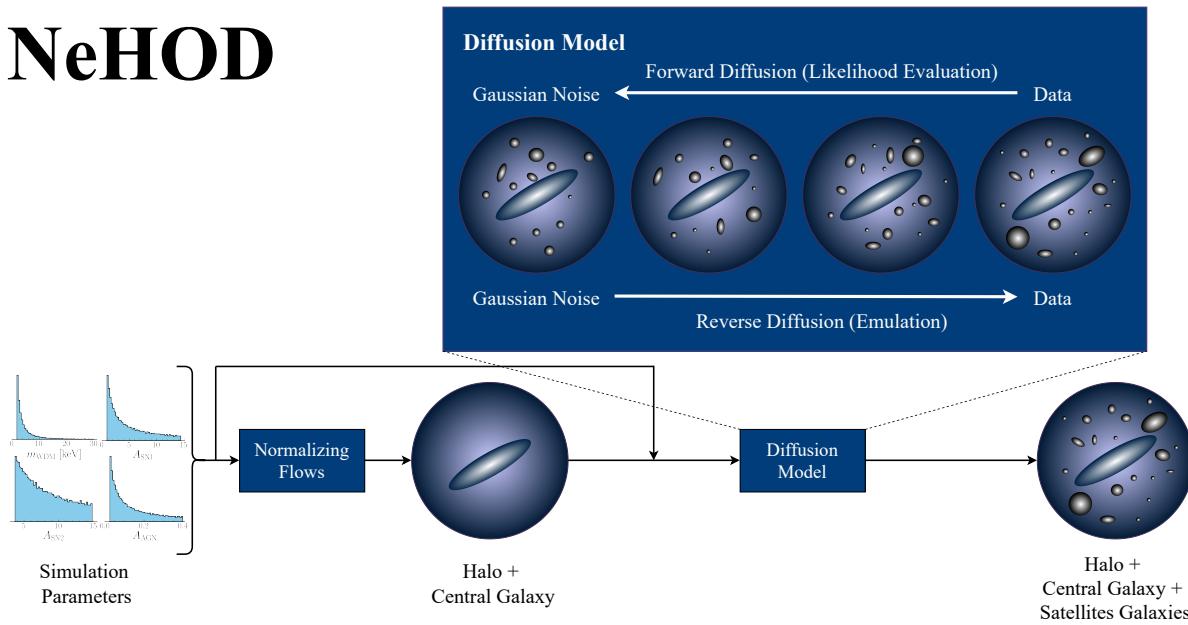
The models operate hierarchically, where the VDM is conditioned on the output from the flows (i.e., the halo and central galaxy properties). In essence, the NeHOD framework takes in the simulation parameters  $\theta_{\text{sim}}$  as inputs and outputs properties of the halo, central galaxy, and satellites within the halo  $\{\theta_{\text{halo}}, \theta_{\text{cent}}, \theta_{\text{sat}}\}$ . The target distribution of NeHOD,  $p(\theta_{\text{sat}}, \theta_{\text{cent}}, \theta_{\text{halo}} | \theta_{\text{sim}})$ , can be written as:

$$p(\theta_{\text{sat}} | \theta_{\text{halo}}, \theta_{\text{cent}}, \theta_{\text{sim}}) p(\theta_{\text{halo}}, \theta_{\text{cent}} | \theta_{\text{sim}}). \quad (3)$$

We note that the flows and diffusion models are trained independently, with more details outlined in §3.4. A visual representation of NeHOD is shown in Figure 1.

<sup>2</sup>This corresponds to  $(1.58 - 1.61) \times 10^{12} M_{\odot}$ .

# NeHOD



**Figure 1.** Flow chart of the NeHOD framework. Black arrows indicate the flow of information into and out of the models. During inference, simulation parameters (i.e., WDM mass and astrophysical parameters) are first input as conditioning (context) features into the normalizing flows, which then generate a halo and a central galaxy. The properties of the halo and central galaxy, along with the simulation parameters, are subsequently passed as conditioning features into the VDM to generate satellite galaxies. During training, the normalizing flows and the VDM are optimized independently (see Appendix A and B for details on the optimization objectives).

We note that this hierarchical approach is a deliberate design choice. One could, for example, model the joint likelihood of central and satellite galaxies, i.e.,  $p(\theta_{\text{sat}}, \theta_{\text{cent}} | \theta_{\text{halo}}, \theta_{\text{sim}})$ , using the VDM. We choose to separate the central and satellite galaxies for the following reasons:

- It is common to study the relationships between central galaxies and their satellites. By modeling the conditional likelihood of satellite galaxies given the central galaxy (instead of the joint likelihood of satellites and the central galaxy), we can more easily adjust the properties of the central galaxies to observe how these changes influence the satellites.
- This is especially important for the “isolated Milky Way” sample in the TNG-WDM suite of simulations used in this paper. The central galaxy dominates the mass within the halo and is significantly more massive than its satellites, typically by a few orders of magnitude. We thus expect it to play a significant role in the dynamics of satellites through processes such as tidal stripping, dynamical friction, and mergers, which can alter the mass distribution within the halo and affect the formation and evolution of satellites (Springel et al. 2005; Binney & Tremaine 2008; Mo et al. 2010).

In addition, the normalizing flows could be further separated into two distinct components, working hierarchically:

one flow that models halos, i.e.,  $p(\theta_{\text{halo}} | \theta_{\text{sim}})$ , and another that models central galaxies given their halos, i.e.,  $p(\theta_{\text{cent}} | \theta_{\text{halo}}, \theta_{\text{sim}})$ . As with satellite and central galaxies, this hierarchical structure could potentially allow for a more modular and flexible approach to modeling the different components within the halo. However, we note that this separation is not critical, as the VDM ultimately takes both  $\theta_{\text{cent}}$  and  $\theta_{\text{halo}}$  as inputs. Our emphasis in this paper is on the VDM approach, which we believe is the core component of our HOD modeling framework. We will further discuss applications of the NeHOD framework, along with potential modifications to the models and their features, in §5.

Lastly, we note that traditional HOD applications typically assign galaxies onto *pre-existing* DM halos from N-body simulations, making the modeling  $p(\theta_{\text{halo}} | \theta_{\text{sim}})$  using the flows unnecessary. However, given our limited number of simulations, this modeling approach facilitates more robust testing against the simulations, as it allows to generate new halos by drawing from the prior  $p(\theta_{\text{sim}})$ . We briefly discuss this in §5.

### 3.2. Dataset and Features

For each simulation, we extract the high-resolution Milky Way-mass halo together with its subhalos from the group catalogs as follows. First, we identify halos with a total mass within  $4.84 \times 10^{11} M_{\odot} h^{-1} < M_{\text{halo}} < 2.07 \times 10^{12} M_{\odot} h^{-1}$ . We require that the total mass of the low-resolution DM particles

does not exceed 2% of the total mass within the virial radius in all of our halos. From the 1,024 simulations in the TNG-WDM suite, 1,018 halos satisfy the contamination criteria.

As mentioned previously, NeHOD consists of two components. Given a set of simulation parameters  $\theta_{\text{sim}}$ , we first predict some properties of the halo  $\theta_{\text{halo}}$  and the central galaxy  $\theta_{\text{cent}}$ . Then, given the halo, central galaxy, and simulation parameters, we predict some properties of all satellite galaxies within the halo  $\theta_{\text{sat}}$ . Below, we provide a detailed description of the set of properties that go into  $\{\theta_{\text{sim}}, \theta_{\text{halo}}, \theta_{\text{cent}}, \theta_{\text{sat}}\}$ . Table 2 provides a summary of the halo and galaxy features and the corresponding FoF/SUBFIND field in the public data release. Descriptions of the simulation parameters can be found in Table 1.

### 3.2.1. Simulation parameters

The simulation parameters  $\theta_{\text{sim}}$  are the input features of the entire NeHOD framework, i.e. both the normalizing flows and the VDM. Here, we let  $\theta_{\text{sim}}$  be the WDM mass and three astrophysical parameters that control the strength of baryonic feedback:

$$\theta_{\text{sim}} = \{m_{\text{WDM}}, A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}. \quad (4)$$

In practice, we use the inverse of the WDM mass and the logarithm of the feedback parameters, since they are sampled uniformly (according to a Sobol sequence) in this space (see Table 1).

We note that the properties of halos and central galaxies can depend on factors beyond these four parameters, such as environmental (e.g., local density field) and assembly information (e.g., star formation history, formation timescales) (e.g., Nusser & Sheth 1999; van den Bosch 2002; Wechsler et al. 2006; Correa et al. 2015; Hadzhiyska et al. 2021; Gabrielpillai et al. 2022). As discussed in §2.2, each halo in our training simulations represents a unique realization of a Milky Way-mass halo, with the only environmental constraint being the isolation criterion. We expect that including dependencies on environmental and assembly factors could further improve the model’s accuracy, which we plan to explore in future work.

### 3.2.2. Halos and frame of reference

The normalizing flows take the simulation parameters  $\theta_{\text{sim}}$  as inputs and outputs the properties of the halo  $\theta_{\text{halo}}$  and central galaxy  $\theta_{\text{cent}}$ . For the halo properties,  $\theta_{\text{halo}}$ , we consider the halo virial mass  $M_{\text{halo}}$ , total stellar mass  $M_{\star}$ , and the number of satellite subhalos  $N_{\text{sat}}$ , i.e.

$$\theta_{\text{halo}} = \{M_{\text{halo}}, M_{\star}, N_{\text{sat}}\}. \quad (5)$$

The virial mass  $M_{\text{halo}}$  is estimated as the total mass enclosed within  $R_{200}$ , the radius where the average density of the halo is 200 times the critical density of the Universe. For each

halo, we use a reference system where its center and peculiar velocity within the cosmological box are both at  $(0, 0, 0)$ . The halo center is the position of the particle with the minimum gravitational potential energy, while the velocity is computed as the sum of the mass-weighted velocities of all particles in the FoF group.

Normalizing flows are designed to model distributions of continuous variables and thus are ill-suited to predict discrete variables such as the number of satellites  $N_{\text{sat}}$ . Thus, in practice, we instead train on and predict  $\log N_{\text{sat}}$ . During generation, we compute  $N_{\text{sat}}$  from the predicted  $\log N_{\text{sat}}$  and then round to the nearest integer. We discuss the limitation of this approach in §5.

### 3.2.3. Central galaxies

For properties of the central galaxy,  $\theta_{\text{cent}}$ , we are interested in modeling the total mass, stellar mass, DM concentration, and the position and velocity offsets relative to the halo. For the position offset, we find that central galaxies in our simulations are at  $(0, 0, 0)$  using the halo reference frame described above. Thus, we do not include the position offset into our features. For the velocity offset, we find a small offset of order 5-10 km/s between the velocities of each halo and its central galaxy. Assuming this velocity offset is isotropic<sup>3</sup>, we only need to model its magnitude, i.e.  $|\mathbf{v}_{\text{cent}}|$ .

In addition to the total mass  $M_{\text{cent}}$ , stellar mass  $M_{\text{cent},\star}$ , and the velocity offset  $|\mathbf{v}_{\text{cent}}|$ , we are interested in modeling structural parameters, such as the DM concentration, which has been found to strongly correlate with the halo formation time and environment in many assembly bias studies (e.g., Nusser & Sheth 1999; van den Bosch 2002; Wechsler et al. 2006; Correa et al. 2015). Because the DM concentration is not available in SUBFIND, we instead use the quantity:

$$\tilde{V}_{\max,i} \equiv \frac{V_{\max,i}}{H_0 r_{\max,i}}, \quad (6)$$

where  $H_0$  is Hubble constant,  $V_{\max,i}$  is the maximum circular velocity of the galaxy and  $r_{\max,i}$  is the radius at this velocity. This follows from Bose et al. (2019), which demonstrates that  $\tilde{V}_{\max}$  is a good proxy for the DM concentration. We also highlight that the relation between  $\tilde{V}_{\max}$  and the DM concentration  $c$  of a Navarro-Frenk-White (NFW, Navarro et al. 1997) is one-to-one (see Equation 9 of Springel et al. 2008), so both  $\tilde{V}_{\max}$  and  $c$  can be used interchangeably as training features. Our post-processing catalog will include both the concentration proxy  $\tilde{V}_{\max}$  and the concentration  $c$ .

<sup>3</sup>This is reasonable because the gravitational forces acting upon the central galaxy from various directions tend to average out due to the symmetric distribution of matter within the halo, resulting in an overall isotropic velocity distribution.

To summarize, the properties of central galaxies used as training features in this work are:

$$\boldsymbol{\theta}_{\text{cent}} = \{M_{\text{cent}}, M_{\text{cent},\star}, |\mathbf{v}_{\text{cent}}|, \tilde{V}_{\text{max,cent}}\}. \quad (7)$$

These features, along with the simulation parameters  $\boldsymbol{\theta}_{\text{sim}}$  and the halo properties  $\boldsymbol{\theta}_{\text{halo}}$ , are used as the context of the VDM for generating satellite galaxies. As before, we emphasize that environmental information and assembly history are not explicitly included as training features. Although the DM concentration proxy  $\tilde{V}_{\text{max}}$  is likely correlated with these factors, it would be beneficial to explicitly include environmental and assembly information in future work.

### 3.2.4. Satellite galaxies

For each satellite galaxy, our model predicts the 3-dimensional positions  $\mathbf{x}_i$  and velocities  $\mathbf{v}_i$  relative to the halo, the total mass  $M_{\text{sub},i}$ , the stellar mass  $M_{\text{sub},\star,i}$ , and the concentration proxy  $\tilde{V}_{\text{max},i}$ . The features of each subhalo, denoted as  $\boldsymbol{\theta}_i$ , are thus:

$$\boldsymbol{\theta}_i = \{\mathbf{x}_i, \mathbf{v}_i, M_{\text{sub},i}, M_{\text{sub},\star,i}, \tilde{V}_{\text{max},i}\}, \quad (8)$$

forming a 9-dimensional output vector for each galaxy. Additionally, it is interesting to include other internal properties of the satellite galaxies (e.g., spin, halo shape, galaxy morphology, etc.) into the framework. We leave this for future work.

Our framework outputs a vector  $\boldsymbol{\theta}_{\text{sat}} = \{\boldsymbol{\theta}_i | i = 1, \dots, N_{\text{sat}}\}$  of shape  $N_{\text{sat}} \times 9$  containing the properties of  $N_{\text{sat}}$  satellites. The number of satellites is a function of each of the properties of the context  $\boldsymbol{\theta}_{\text{sim}}$  (e.g., a larger value of  $m_{\text{WDM}}$  will result in more satellites per halo). However, we note that due to differences in formation history and environment,  $N_{\text{sat}}$  can vary between halos for the same  $\boldsymbol{\theta}_{\text{sim}}$ . This makes the use of permutation equivariant architectures such as Transformers advantageous, as will be discussed in more detail in §3.3.

## 3.3. Machine learning framework

In this section, we describe the normalizing flows and VDM used in NeHOD, including their motivations and neural network architectures. The mathematical formulations of normalizing flows and diffusion models are comprehensive and thoroughly documented in the machine learning literature. Thus, we omit their detailed derivations from this section. We instead provide the mathematical formalism of both models, including their optimization objectives, in Appendix A for normalizing flows and Appendix B for diffusion models.

### 3.3.1. Modeling halo and central galaxies

To model the properties of halos and central galaxies, we employ normalizing flows (Jimenez Rezende & Mohamed 2015; Papamakarios et al. 2017; Papamakarios et al. 2019)

that are conditioned on simulation parameters  $\boldsymbol{\theta}_{\text{sim}}$ . Normalizing flows employ non-linear, invertible transformations (with tractable Jacobians and its inverse) from a standard, base Gaussian distribution to a complex target distribution, allowing for efficient density estimation and sampling. They have been employed in various astrophysical applications including generative modeling (e.g., Lovell et al. 2023a; Nguyen et al. 2023b; Kwon & Hahn 2024) and simulation-based inference (see Cranmer et al. 2020, for a review).

In our framework, our flow model consists of 8 neural spline flows (Durkan et al. 2019) layers, each using a rational quadratic spline transformation with 8 bins (9 knots). To read in the conditioning context, each flow also has a multi-layer perceptron (MLP) with 2 fully-connected layers of hidden size 16 and ReLU activation function. Each fully-connected layer is followed by a batch normalization layer (Ioffe & Szegedy 2015) and a dropout layer (Srivastava et al. 2014) with a rate of 0.2. We find that, due to the limited number of simulations available in our application, batch normalization and dropout are especially important to prevent overfitting in the flows. During training, we also employ early stopping to prevent overfitting (see §3.4).

### 3.3.2. Modeling satellite galaxies

To generate satellites within a halo, we adopt an approach presented in CM23. We represent each satellite population as a point cloud—a set of points in 3-dimensional space with attributes such as velocities, masses, and concentrations. Point clouds eliminate the need for binning and voxelization, which are not viable/effective for HOD applications due to the sparse spatial distribution of satellites. More importantly, this allows us to retain the most information from the data and resolve arbitrary small scales limited only by the spatial resolution of the simulations.

Similarly, we also employ a Variational Diffusion Model (VDM; Vahdat et al. 2021; Song et al. 2021; Kingma et al. 2021). VDMs are a class of generative models that progressively learn to generate complex data distributions through a process of gradually denoising samples (referred to as the “reverse diffusion”), starting from a random noise distribution. They offer a few advantages over other generative models. VDMs are more expressive than variational autoencoders, (VAEs; Kingma & Welling 2013), allowing for more flexible network architecture than normalizing flows. Unlike generative adversarial networks, (GANs; Goodfellow et al. 2014), they are generally much easier to train and less prone to experiencing mode collapse (Salimans et al. 2016). Additionally, they facilitate the tractable evaluation of likelihood, which allows them to be used for inference and anomaly detection.

The VDM consists of a *noise schedule*,  $\gamma(t)$ , which determines how much Gaussian noise is added to the data  $\mathbf{y}$

	Parameters	Description	Relevant FoF/SubFind Field
Halo Properties	$M_{\text{halo}}$	Halo virial mass	Group_M_Crit200
	$\theta_{\text{halo}}$	Halo stellar mass	GroupMassType PartType4
	$N_{\text{sat}}$	Number of satellites	
Central Galaxy Properties $\theta_{\text{cent}}$	$ \mathbf{v}_{\text{cent}} $	Velocity offset with respect to the halo	SubhaloVel
	$M_{\text{cent}}$	Total mass of central galaxy	SubhaloMass
	$M_{\text{cent},\star}$	Stellar mass of central galaxy	SubhaloMassType PartType4
	$\tilde{V}_{\text{max,cent}}$	Proxy for DM concentration of the central galaxy	
Satellite Galaxy Properties $\theta_{\text{sat}}$	$\mathbf{x}_i$	3-dimensional position vector of satellite $i$	SubhaloPos
	$\mathbf{v}_i$	3-dimensional velocity vector of satellite $i$	SubhaloVel
	$M_{\text{sub},i}$	Total mass of satellite $i$	SubhaloMass
	$M_{\text{sub},\star,i}$	Stellar mass of satellite $i$	SubhaloMassType PartType4
	$\tilde{V}_{\text{max},i}$	Proxy for DM concentration of satellite $i$	

**Table 2.** List of parameters used in this work and their corresponding FoF/SUBFIND fields, if available, in the public data release. Note that the properties of central galaxies are extracted from the same FoF/SUBFIND fields as satellite galaxies, but with different indices. The indices of central galaxies are found at GROUPFIRSTSUB in the group catalog.

at each step  $t$  of the forward diffusion process, and a *noise prediction model*,  $\hat{\epsilon}(\mathbf{z}_t, t)$ , which estimates the added noise given the time step  $t$  and the noisy data  $\mathbf{z}_t$  (Ho et al. 2020; Song et al. 2020; Kingma et al. 2021). In our setup, we use a linear noise schedule with trainable parameters  $\eta_{\min}$  and  $\eta_{\max}$  that determine the minimum and maximum amount of noise added, respectively. Further details on the diffusion process, including the noise schedule, and forward and reverse diffusion process, can be found in Appendix B.

For the noise prediction model, we use a Transformer-based architecture. The Transformer, first introduced in Vaswani et al. (2023), leverages the attention mechanism to efficiently process data. Attention allows Transformer-based models to dynamically assess and prioritizes the relevance of each part of the input, making them powerful models for many applications ranging from large language modeling (Radford et al. 2019; Devlin et al. 2018) to more recent astrophysics applications (see e.g. Hwang et al. 2023; He et al. 2023; Lanusse et al. 2023; Cuesta-Lazaro & Mishra-Sharma 2023; Leung & Bovy 2024).

We implement a Transformer with additional MLPs for processing conditioning features (i.e.,  $\{\theta_{\text{sim}}, \theta_{\text{halo}}, \theta_{\text{cent}}\}$ ), the diffusion noise time step  $t$ , and the input data (i.e., the noisy version of  $\theta_{\text{sat}}$  at  $t$ ). During forward pass, the time step  $t$  is first passed through a sinusoidal positional encoding layer similar to that used in Vaswani et al. (2023). The encoded time step, the conditioning features, and the input data are then independently projected onto a 128-dimensional space by three 1-layer MLPs. The projected conditioning features and time step are further concatenated and passed through a 2-layer MLP. This output is then added to the projected input data before being passed into the Transformer. The Transformer itself consists of 6 blocks, each equipped with a Multi-head Attention layer with 128 hidden channels and 4 attention heads, followed by a 2-layer MLP with 256 hid-

den channels and a GELU activation function (Hendrycks & Gimpel 2016). The output is then mapped back into the same dimension of the input data using a 1-layer MLP.

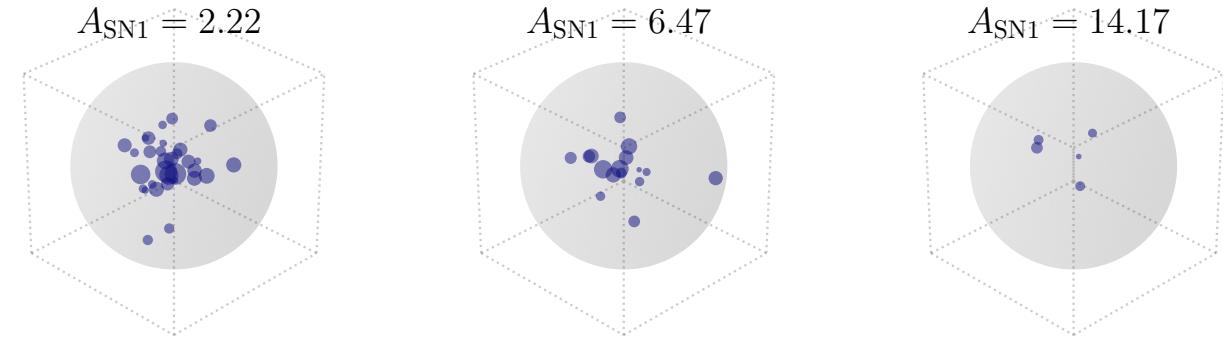
Because each satellite population is modeled as a set of points with no inherent ordering, we do not include any position encoding or causal masking layer, making the model permutation-equivariant. It is also important to highlight that our neural network architecture does not explicitly incorporate geometrical symmetries, such as rotational and translational invariance. Instead, we apply data augmentation techniques during the training process, as detailed in §3.4.

Lastly, it is worth noting that graph-based neural networks are also permutation-equivariant and capable of efficiently modeling point cloud cosmological data, as have been shown in recent studies (e.g., Villanueva-Domingo & Villaescusa-Navarro 2022; Nguyen et al. 2023a; de Santi et al. 2023; Rose et al. 2024b,a). CM23 also demonstrates that using a graph convolutional neural network and  $k$ -nearest neighbors graphs can yield comparable results to Transformers. An advantage of using graph neural networks is that the computational cost scales more slowly with the number of point cloud data (i.e. the number of satellites) increase, as self-attention scales quadratically  $O(N_{\text{sat}}^2)$ . However, in our application, where the number of points is of order  $N_{\text{sat}} \sim O(100)$ , the performance difference becomes negligible. In addition, constructing graphs from point clouds introduces additional hyperparameters, e.g., the number of neighbors  $k$ . For these reasons, we thus choose a Transformer-based architecture for our noise prediction model.

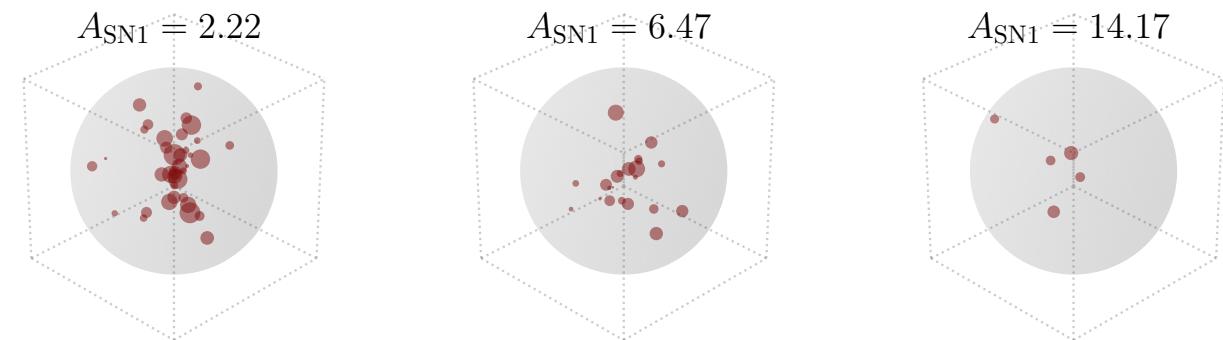
### 3.4. Training and Optimization

We split the dataset into 814 training and 204 validation simulations, following an 80-20 split. The flows and the VDM are trained independently by performing gradient descent using the AdamW optimizer (Loshchilov & Hutter

## NeHOD



## DREAMS



**Figure 2.** Example satellite galaxies generated by NeHOD. The bottom row displays the satellite galaxies of three halos from the DREAMS simulations, with each column corresponding to a different value of  $A_{\text{SN}1}$ . The top row shows corresponding realizations of satellite galaxies for the same halos, as generated by NeHOD. The marker size scales logarithmically with the stellar mass of each satellite. For visual clarity, central galaxies, located at the center of the shaded sphere, are omitted. In each panel, the box size and the diameter of the shaded sphere are set to  $600 \text{ kpc } h^{-1}$ .

2019; Kingma & Ba 2014), with a peak learning rate of 0.0005 and a weight decay coefficient of 0.01. Details on the training objectives can be found in Appendix A for the flows and Appendix B for the VDM. In addition, we use a cosine decay learning rate scheduler (Loshchilov & Hutter 2016). Training parameters, including the linear warm-up and cosine decay steps of the scheduler and the training batch size, are  $\{5000, 10000, 128\}$  and  $\{100, 1000, 64\}$  for the VDM and the flows, respectively. The VDM and flows converge after about 50,000 (400) and 10,000 (150) iterations (epochs), which take approximately 30 and 20 minutes for each model on an NVIDIA Tesla V100 GPU. We implement early stopping with the patience of 1000 iterations for both models to prevent overfitting and select the checkpoints with the lowest validation loss.

As mentioned in §3.3, we apply data augmentation to help the model learn the relevant geometrical symmetries, such as translational and rotational invariance. Data augmentation also mitigates the limitations of our relatively small training dataset by effectively increasing the diversity of the training

examples. We consider only rotational invariance because the coordinates are always in the frame of the halo by construction (as discussed in §3.2).

The data augmentation is applied *during training*. In other words, a random 3-dimensional rotational matrix is applied to the coordinates of the satellites on the fly before being passed through the VDM. *Thus, it is unlikely that the model ever encounters the same sample twice, which improves generalization to unseen data.* The same approach is applied during validation, ensuring that the model is consistently tested on varied data configurations. We do not consider planar symmetry (e.g., the disk plane of the central galaxy) in this work. Depending on specific applications, future work may incorporate planar symmetry into the data augmentation process.

## 4. RESULTS

We present our results by generating new samples using NeHOD and validating them by comparing key summary statistics of the generated samples with those from the DREAMS simulations. To generate the test dataset, we randomly sam-

ple the WDM mass and astrophysical parameters from the prior distributions in Table 1. We first generate the halos and central galaxies using the conditional flows, and then use the VDM to generate satellites given the generated halos and central galaxies. In total, we generate 100,000 samples, which take approximately 50 minutes on a GPU using 1,000 diffusion steps for each sample. Note that the generated sample size is 100 times larger than the sample size of the training simulations (about 1000 samples), which will help identify potential discrepancies and biases between the two datasets more effectively.

Figure 2 presents an example output of NeHOD. In the bottom row, we display the satellite galaxies of three halos randomly selected from the DREAMS simulations, with increasing  $A_{\text{SN}1}$  parameter from left to right. For each halo, we generate a corresponding realization of its satellite galaxies using the same halo properties (including the number of satellites), central galaxy properties, and simulation parameters, and plot them in the top panel of the corresponding column. For visual clarity, we do not show the central galaxies, which are at the center of the shaded sphere. Figure 2 effectively highlights that the point-cloud output of NeHOD, with more detailed validation tests presented in the rest of the section.

#### 4.1. Properties of halos and central galaxies

We evaluate how well the conditional flows can recover the properties of the halos and central galaxies. We divide the samples into 10 bins of the simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN}1}, A_{\text{SN}2}, A_{\text{AGN}}\}$ . For each property of halos and central galaxies (i.e.,  $\theta_{\text{halo}}$ ,  $\theta_{\text{cent}}$ ), we calculate the median, 16th, and 84th percentiles of its distribution within each bin. The bins are divided inverse-uniformly for the WDM mass and log-uniformly for the astrophysical parameters (see Table 1) and thus have the same sample size. In particular, each bin contains about 1,000 and 100 halos for the NeHOD dataset and the simulations, respectively.

Figure 3 shows variations of the number of satellites  $N_{\text{sat}}$  (top) and central galaxy stellar mass  $M_{\text{cent},\star}$  (bottom) with respect to the simulation parameters. We show the full set of output features in Appendix C.1. The halo virial mass  $M_{\text{halo}}$  and central galaxy total mass  $M_{\text{cent}}$  does not vary significantly with these parameters, which is likely because the halos are selected based on the Milky Way-mass criterion (see §3.2). Thus, we do not show the results for  $M_{\text{halo}}$  and  $M_{\text{cent}}$ , and briefly note that the distributions of these properties predicted by NeHOD match well with those from the simulations. Additionally, since the stellar masses of the halo and central galaxy are strongly correlated, we exclude the halo stellar mass  $M_{\star}$  here and instead show their 2-dimensional contours in Figure 4. Lastly, the concentration proxy  $\tilde{V}_{\text{max}}$  follows a bimodal distribution that strongly correlates with  $M_{\text{cent},\star}$ , rendering the above statistics less informative. Thus,

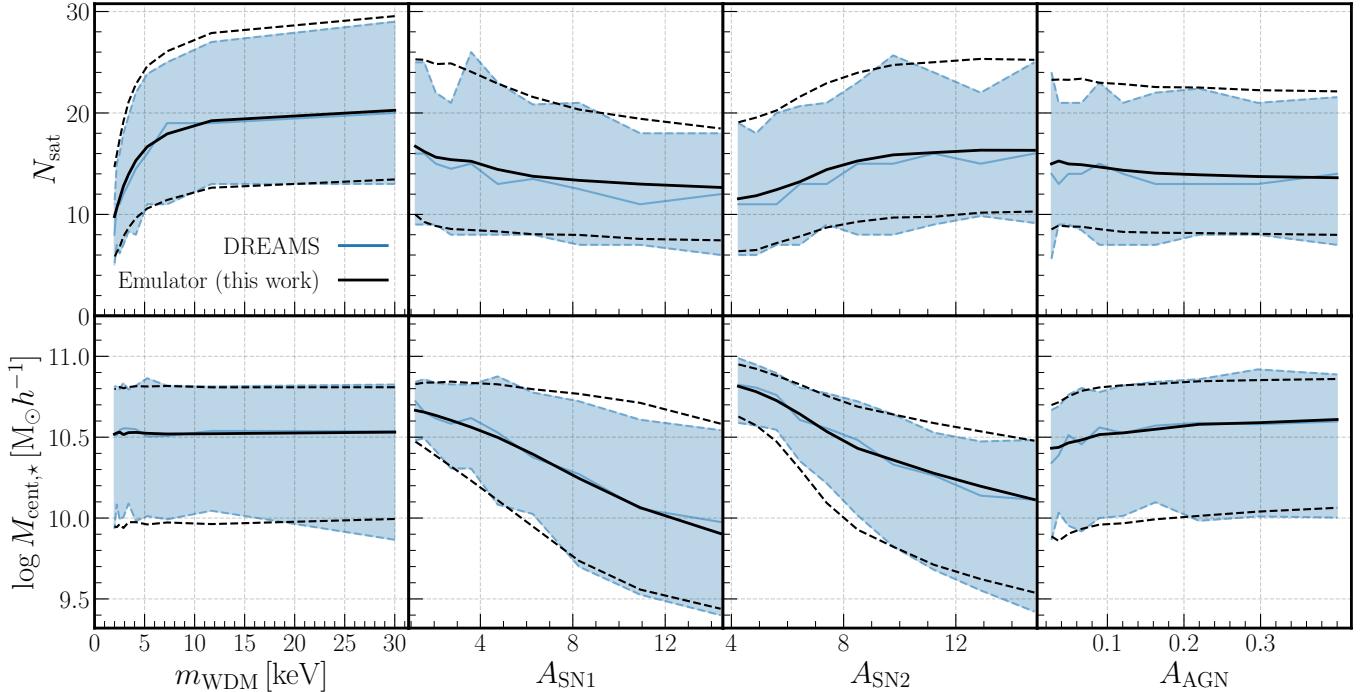
as with the halo stellar mass, we show the 2-dimensional contours of  $\tilde{V}_{\text{max}}$  and  $M_{\text{cent},\star}$  in Figure 4.

In Figure 3, we observe strong variations in the number of satellites  $N_{\text{sat}}$  and the central galaxy stellar mass  $M_{\text{cent},\star}$  with respect to the simulation parameters, except for  $A_{\text{AGN}}$ . The number of satellites  $N_{\text{sat}}$  is sensitive to both the WDM mass and the SN feedback parameters, while  $M_{\text{cent},\star}$  is primarily influenced by the latter. Overall, there is good agreement between the distributions predicted by NeHOD and those from the simulations. However, the flows tend to slightly overestimate  $N_{\text{sat}}$  by 1-2 satellites, though this discrepancy is typically much smaller than the scatter in the  $N_{\text{sat}}$  distribution due to halo-to-halo variations. On the other hand, the distributions of  $M_{\text{cent},\star}$  show good agreement for  $m_{\text{WDM}}$ ,  $A_{\text{SN}1}$ , and  $A_{\text{AGN}}$ , with the largest discrepancy occurring in the 16th percentile for  $A_{\text{SN}2}$ . We note that we do not expect the flows to agree perfectly with the simulations, as the number of training samples is relatively small ( $\approx 1000$  simulations). This is especially true towards the edges of the target distributions, where the models require more training samples to learn (as compared to the median). We expect that with more simulated data, the accuracy of the distributions will further improve.

Figure 4 shows the 2-dimensional distributions of  $M_{\star}$  and  $M_{\text{cent},\star}$  (top), and  $M_{\text{cent},\star}$  and  $\tilde{V}_{\text{max}}$  (bottom). The NeHOD samples and simulations are each divided into 3 bins of the simulation parameters (denoted by different colors), each containing approximately 33,000 and 330 samples respectively. The left and right panels show the 68% and 95% contours of the distributions for the  $A_{\text{SN}1}$  and  $A_{\text{SN}2}$  bins, respectively. We do not show bins of  $m_{\text{WDM}}$  and  $A_{\text{AGN}}$  as we do not find strong variations of  $M_{\text{cent},\star}$ ,  $M_{\star}$ , and  $\tilde{V}_{\text{max}}$  with respect to these parameters. Note that due to the differences in the sample size, the simulation contours are noticeably noisier than the NeHOD contours.

As briefly mentioned previously, there are strong correlations between  $M_{\star}$  and  $M_{\text{cent},\star}$ , and between  $M_{\text{cent},\star}$  and  $\tilde{V}_{\text{max}}$ . The distribution of the concentration proxy  $\tilde{V}_{\text{max}}$  is bimodal, which could be attributed to the bimodality in the formation history of DM halos (Shi et al. 2018; Hearn et al. 2021, e.g.). Since the concentration is highly correlated with formation history (e.g., Nusser & Sheth 1999; van den Bosch 2002; Wechsler et al. 2006; Correa et al. 2015), the observed bimodality in  $\tilde{V}_{\text{max}}$  can arise from differences in the formation histories of these halos. A more detailed investigation of this bimodality is beyond the scope of this work, so we leave this for future work.

From Figure 4, we observe small differences in the case of the central galaxy stellar mass  $M_{\text{cent},\star}$  and concentration proxy  $\tilde{V}_{\text{max}}$ , e.g. the 68% contours in the  $A_{\text{SN}1}$  (lower left) panel. As previously noted, we do not expect the distributions to match perfectly, due to the limited number of training samples. Bimodal distributions are also much more challenging



**Figure 3.** The properties of halos and central galaxies as a function of the simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN}1}, A_{\text{SN}2}, A_{\text{AGN}}\}$  (left to right). The top and bottom panels show distributions of the number of satellites  $N_{\text{sat}}$  and the stellar mass of the central galaxies  $M_{\text{cent},*}$ . The median (solid lines), 16th, and 84th percentiles (dashed lines) of each distribution are shown. The black lines denote the samples generated by conditional flows in this work, while the blue lines and shaded regions denote the simulations.

to fully capture, as both peaks need to be well-sampled to accurately represent the target distribution. Overall, however, the 68% and 95% contours predicted by NeHOD generally align well with the simulations.

Lastly, we note that §3 of Rose et al. 2024a presented a machine learning-based emulator for the WDM satellite count. Given the four simulation parameters, the authors assume a Gaussian distribution of satellite count and employ MLPs to predict the mean and standard deviation. This MLP approach also allows both efficient sampling and likelihood evaluation and achieves somewhat similar results to the conditional flows. However, we opt to use the flows in our framework as they require fewer assumptions on the target distributions and thus can model more complex distributions. As seen in Figure 4, properties such as the concentration proxy  $\tilde{V}_{\max}$  display non-Gaussian features, i.e. bimodal distributions that strongly correlate with the stellar mass of the central subhalos. In such cases, we expect the flows to outperform the MLP approach employed in Rose et al. 2024a.

#### 4.2. Properties of satellite galaxies

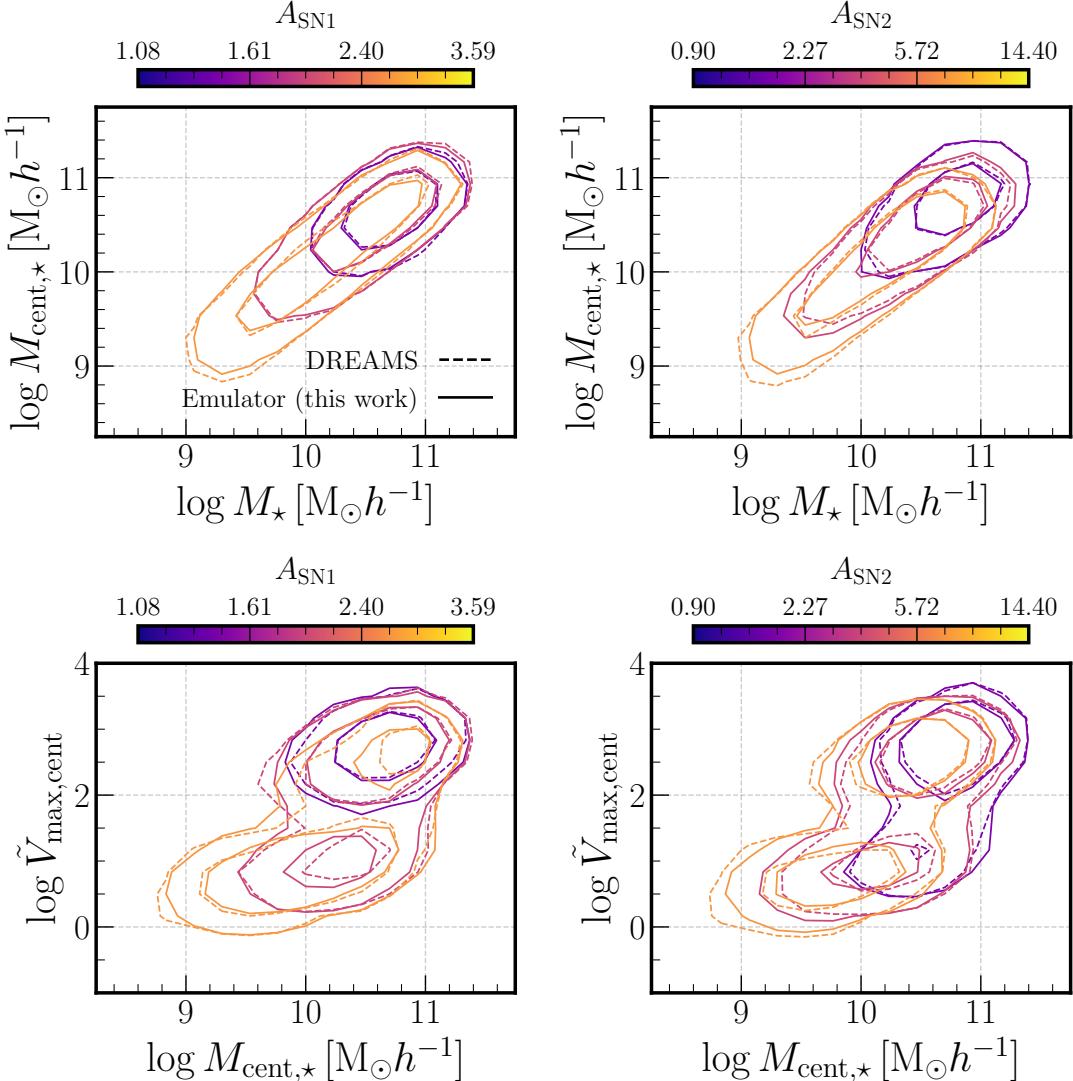
We now evaluate the properties of the satellite populations generated by the VDM, by comparing their key summary statistics with those derived from the simulations. We divide each parameter into three bins such that each bin contains approximately the same number of samples. Due to the prior distributions in the simulation parameters (as shown in Ta-

ble 1), the bins are distributed uniformly in inverse WDM mass (i.e.,  $m_{\text{WDM}}^{-1}$ ) and log-uniformly for the astrophysical parameters. Each bin contains approximately 33,000 NeHOD-generated and 330 simulated satellite populations. This large generated sample size provides a more precise estimate of the distribution modeled by the VDM,

For each bin, we calculate and compare the satellite stellar mass functions (§4.2.1), stellar-to-halo mass relations (§4.2.2), concentration-mass relations (§4.2.3), and position and velocity clustering (§4.2.4). This provides validation tests of the VDM for modeling both intra-galaxy properties (i.e., properties within a single galaxy), such as the stellar-to-halo mass relations and concentration-mass relations, and intergalaxy properties (i.e., between galaxies in a catalog), such as the satellite stellar mass functions and clustering.

##### 4.2.1. Satellite stellar mass functions

The satellite stellar mass function (SSMF) is among the key results of many Milky Way and dwarf galaxy surveys (e.g., Carlsten et al. 2022; Danieli et al. 2023; Mao et al. 2024) and has been used to place constraints on the mass of WDM (e.g., Lovell et al. 2014; Jethwa et al. 2018; Nadler et al. 2021b). Lower WDM particle masses result in fewer low-mass satellites in the SSMF due to free-streaming effects (Bode et al. 2001; Lovell et al. 2014). Baryonic feedback can also disrupt the formation of low-mass satellites and produce similar effects in the mass functions (e.g., Sales et al. 2022; Brown et al.



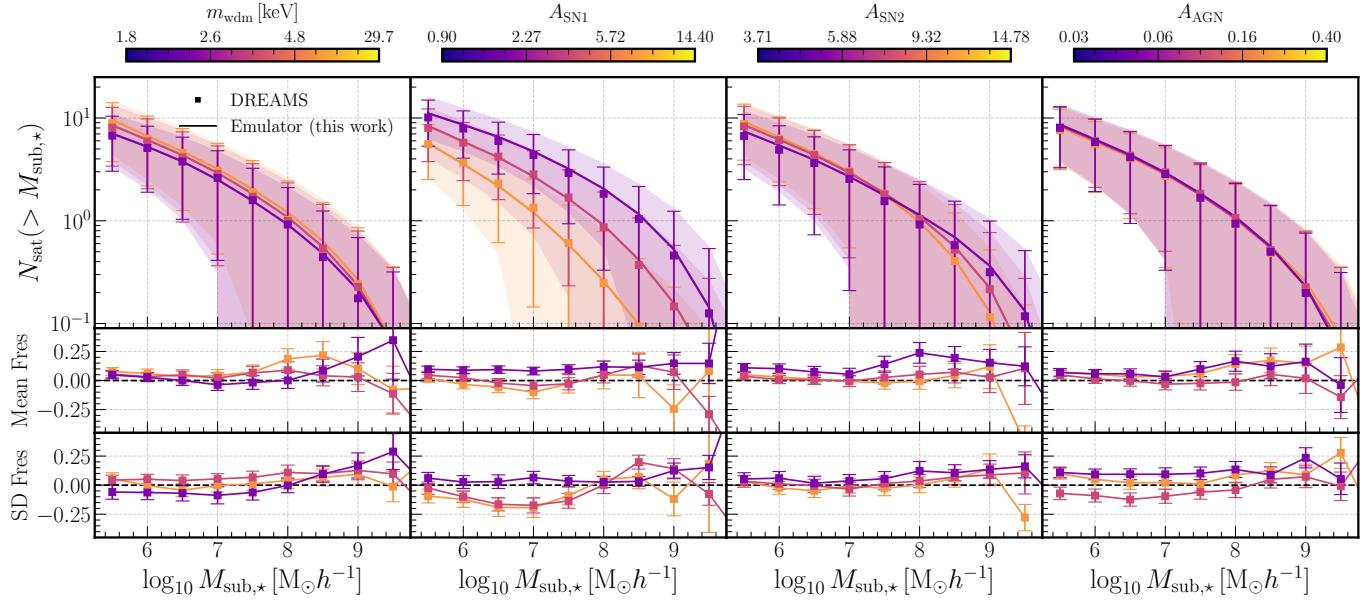
**Figure 4.** *Top:* distributions of the stellar masses of the halos  $M_\star$  and central galaxies  $M_{\mathrm{cent},\star}$ . *Bottom:* distributions of the central galaxy stellar mass  $M_{\mathrm{cent},\star}$  and the concentration proxy  $\tilde{V}_{\mathrm{max}}$ . In both cases, the 68% and 95% contours are shown for 3 bins of  $A_{\mathrm{SN1}}$  (left) and  $A_{\mathrm{SN2}}$  (right), with each bin denoted by a different color. The solid and dashed lines denote the NeHOD samples and the simulations, respectively.

2024). Thus, accurately modeling the SSMFs is important for differentiating the effects of DM from those of baryons and interpreting the results of these surveys.

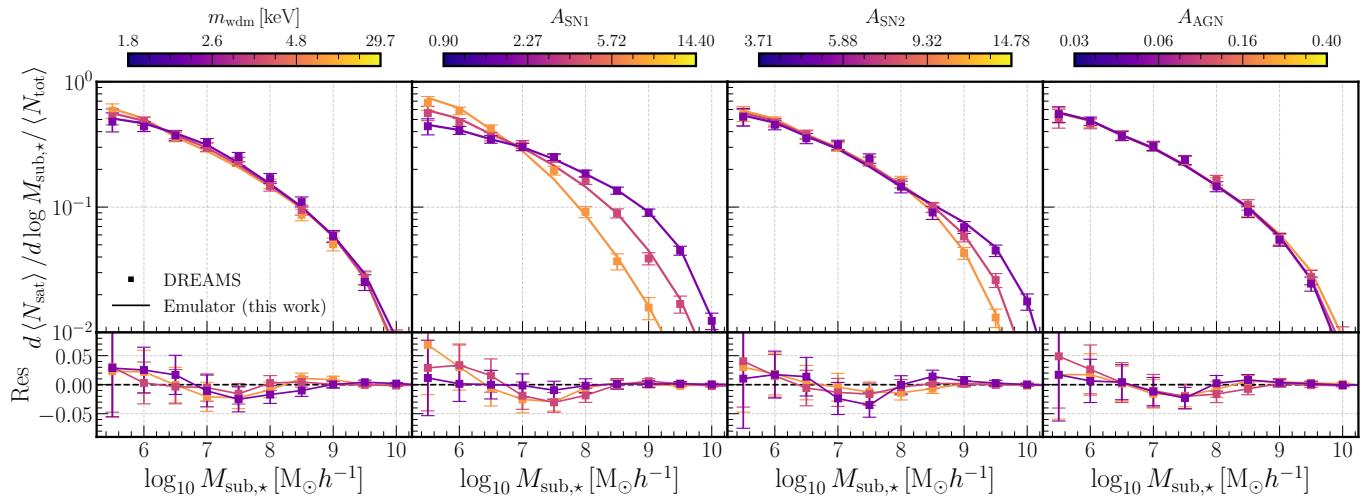
We first calculate the SSMF of each halo by dividing its satellites into  $M_{\mathrm{sub}}$  bins with equal widths of 0.5 dex and counting the cumulative number of satellites  $N_{\mathrm{sat}}(> M_{\mathrm{sub},\star})$  in each bin. Then, for each bin of the WDM mass and astrophysical parameters, we calculate the average and standard deviation of the SSMF over all halos in the bin. These statistics, along with their errors, are estimated using bootstrapping with 1,000 bootstrap samples, where each sample is generated by random sampling with replacement from the halos in each bin. Note that due to the large number of generated halos ( $\sim 33,000$  halos per bin), the errors of the generated SSMFs are a factor of  $\sim 10$  smaller than that of the simulations and

thus negligible. In addition, for each parameter bin, we calculate the differential form of the SSMF, by taking the slope of the average SSMF  $d \langle N_{\mathrm{sat}} \rangle / d \log M_{\mathrm{sub},\star}$ , which is another key summary statistic commonly reported by surveys (Carlsten et al. 2022; Danieli et al. 2023; Mao et al. 2024). It provides an important test of the VDM’s ability to capture the shape of SSMF. To further highlight the shape, we normalize by differential SSMF by the average total number of satellites in each parameter bin. We show additional results on the satellite halo mass functions in Appendix C.2.

Figure 5 shows the SSMFs for each bin of the WDM mass  $m_{\mathrm{WDM}}$  and the three feedback parameters  $\{A_{\mathrm{SN1}}, A_{\mathrm{SN2}}, A_{\mathrm{AGN}}\}$  (left to right column). In each column, the top panel shows the average of SSMFs for the bins of the corresponding parameter, with each bin represented by



**Figure 5.** Top: The satellite stellar mass functions (SSMFs) generated by NeHOD and extracted from the simulations. The columns show the variations of the SSMFs over the WDM mass  $m_{\text{WDM}}$  and astrophysical parameters  $\{A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . In each column, the color denotes the SSMFs of a bin of the corresponding parameter. The average SSMFs, along with their standard deviations, are shown as solid lines and shaded regions for NeHOD and error bars for the simulations. Middle: Fractional residuals of the average SSMFs, calculated as the difference between the NeHOD SSMFs and the simulation SSMFs divided by the simulation SSMFs. The error bars represent the propagated errors of the fractional residuals, derived from the bootstrapped errors of the average SSMFs. Bottom: Fractional residuals of the standard deviation SSMFs, along with their errors, calculated using the same procedure as for the average SSMFs.



**Figure 6.** Top: The differential SSMFs  $d \langle N_{\text{sat}} \rangle / d \log M_{\text{sub},*}$  normalized by the total number of satellites  $\langle N_{\text{tot}} \rangle$ . The columns show the variations of the SSMFs over the WDM mass  $m_{\text{WDM}}$  and astrophysical parameters  $\{A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . The NeHOD and simulation differential SSMFs are shown as solid lines and squares, respectively. Bottom: The residuals between NeHOD and simulation differential SSMFs. The error bars in both rows denote the errors of the simulations.

a different color. The standard deviations of the SSMFs are shown as shaded regions and error bars for NeHOD and simulation samples, respectively. The middle and bottom panels show the fractional residuals of the averages and standard deviations of the SSMFs, respectively. We show the fractional residuals instead of the residuals to highlight differences between NeHOD and the simulations for high  $M_{\text{sub},\star}$  bins with low numbers of satellites. The error bars in the middle and bottom panels denote the errors of the fractional residuals, estimated using error propagation.

As expected, both the WDM mass and the feedback parameters (except for  $A_{\text{AGN}}$ ) significantly influence the SSMFs. There is a noticeable degeneracy between the effects of WDM mass and baryonic feedback and between the feedback parameters  $A_{\text{SN}1}$  and  $A_{\text{SN}2}$ . Additionally, the SSMFs are much more sensitive to  $A_{\text{SN}1}$  than to  $m_{\text{WDM}}$ , making it challenging to constrain  $m_{\text{WDM}}$  observationally with the SSMFs due to the inherent uncertainty in  $A_{\text{SN}1}$ . This further highlights the need for precise modeling to disentangle these influences and place accurate constraints on WDM.

Overall, we see good agreements in both the average and standard deviation of the SSMFs between the simulations and generated samples, with the VDM accurately capturing the general trends across multiple bins of both the WDM mass  $m_{\text{WDM}}$  and the feedback parameters  $\{A_{\text{SN}1}, A_{\text{SN}2}, A_{\text{AGN}}\}$ . However, we observe a 10-15% offset between the generated SSMFs and the simulations, which we attribute mainly to the differences in the total number of satellites  $N_{\text{sat}}$ . As illustrated in Figure 3, the normalizing flows tend to overestimate the total  $N_{\text{sat}}$ , leading to an overestimation at low  $M_{\text{sub},\star}$ . Additionally, the fractional residuals of the average SSMFs increase slightly at high  $M_{\text{sub},\star}$  bins. These bins have a lower number of satellites, and thus the average SSMFs fluctuate more noticeably. The generated SSMFs remain within 1-2 error bars of the simulations, as the errors also increase at high  $M_{\text{sub},\star}$ . We observe similar trends with the standard deviations.

Figure 6 shows the normalized differential SSMFs (top panels), and the residuals between the generated and simulated SSMFs (bottom panels) for the same parameter bins. As previously noted, the normalized differential SSMFs are estimated by dividing the slope of the average SSMFs shown in Figure 5 by the average total number of satellites, i.e.  $d \langle N_{\text{sat}} \rangle / d \log M_{\text{sub},\star} / \langle N_{\text{tot}} \rangle$ . The error bars denote the errors in estimating normalized differential SSMFs of the simulations. The errors of the NeHOD SSMFs are not shown, as they are negligible due to the large number of generated samples.

The differential SSMFs highlight the shape of the mass distributions and provide an important test of the VDM's ability to capture these distributions. Overall, the generated and simulated SSMFs align well, with the residuals mostly

within 1 error bar. We note, however, that we observe a similar offset at low  $M_{\text{sub},\star}$ , due to the overall number of satellites being overestimated by the flows. Additionally, unlike in Figure 5, the errors are higher for lower values of  $M_{\text{sub},\star}$ . This is because the errors scale with the number of satellites in the mass bins, which is higher for lower mass subhalos.

#### 4.2.2. Stellar-to-halo mass relation

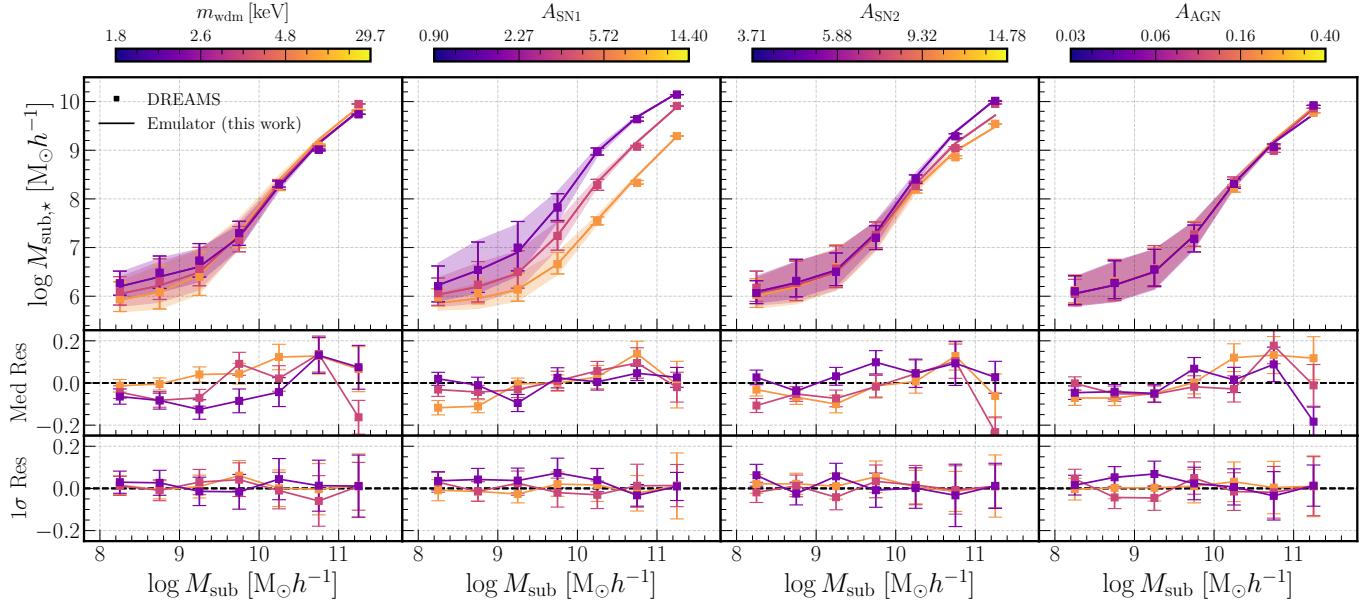
The SSMFs provide an important test for capturing relationships between stellar masses of galaxies within a halo, i.e. the inter-galaxy properties. It is as important for the VDM to accurately model relationships between properties of a single galaxy, i.e. the intra-galaxy properties. To evaluate this aspect, we first compute the stellar-to-halo mass relation (SHMR).

The SHMR is a key summary statistic used in many galaxy formation studies (e.g., Rodríguez-Puebla et al. 2017; Moster et al. 2018; Behroozi et al. 2019; Gabrielpillai et al. 2022). The SHMR provides insights into the efficiency of star formation and the impact of feedback processes and thus is sensitive to baryonic prescriptions of the simulations (e.g., Efstathiou et al. 2000; Hopkins et al. 2014; Somerville & Davé 2015; Marinacci et al. 2019). We note, however, that the SHMR has been employed mainly at the halo mass scales ( $\gtrsim 10^{10} M_{\odot} h^{-1}$ ) and not for satellite galaxies. Nevertheless, it provides a good summary statistic for validating the relationships between stellar and halo mass predicted by NeHOD.

To compute the SHMR of the satellites within each halo, we divide the satellites into  $M_{\text{sub}}$  bins of 0.5 dex width. For each mass bin, we calculate the median, 16th, and 84th percentiles of the stellar mass  $M_{\text{sub},\star}$  of the satellites. For each bin of simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN}1}, A_{\text{SN}2}, A_{\text{AGN}}\}$ , we then determine the average and standard error of these percentiles across all halos. Additionally, we also estimate the  $1-\sigma$  confidence interval of the SHMRs in each parameter bin as the difference between the average 84th and 16 percentiles.

In Figure 7, the top rows show the average of the median, 16th, and 84th percentiles of the generated and simulated SHMRs in each parameter bin. The solid lines and shaded regions represent the SHMRs of the NeHOD samples, while the error bars correspond to those of the simulations. The middle and bottom panels present the residuals between the NeHOD and the simulations of the median and the scatter, respectively. In both cases, the error bars denote the errors of the simulations.

As with the mass functions, we see that both the WDM mass and astrophysical parameters, except for  $A_{\text{AGN}}$ , affect the SHMRs, with the most dominant influence coming from  $A_{\text{SN}1}$ . Interestingly, for low-mass satellites ( $M_{\text{sub}} \lesssim 10^9 M_{\odot} h^{-1}$ ), increasing the WDM mass  $m_{\text{WDM}}$  increases the subhalo stellar mass, though this is not a strong effect since all the lines are consistent to within  $1\sigma$ . Similar effects have



**Figure 7.** Top: The stellar-to-halo mass relations (SHMR) of satellites generated by NeHOD and extracted from the DREAMS simulations. The columns show the variations of the SHMRs over the WDM mass  $m_{\text{WDM}}$  and astrophysical parameters  $\{A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . The average, 16th, and 84th percentiles of the SHMRs are shown as solid lines and shaded regions for NeHOD, and error bars for the simulations. Middle: The residuals of the median SHMRs between NeHOD and the simulations, with the error bars denoting the errors of the simulations. Bottom: The residuals of the SHMR scatters between NeHOD and the simulations, with the error bars denoting the errors of the simulations.

been observed in previous studies (e.g., Read et al. 2017; Lovell et al. 2020). A potential explanation could be that WDM halos form later and thus have delayed star formation, resulting in a more extended period of gas availability for star formation. However, investigating these trends further is beyond the scope of this paper and is left for future work.

The SHMRs of the generated satellites align well with those of the simulations. We see that NeHOD can capture the general trends in both the medians and 16–84th percentiles regions of the SHMRs. Discrepancies from the median SHMRs of the simulations tend to increase slightly at high  $M_{\text{sub}}$  and typically about 0.1 – 0.2 dex. However, the uncertainties of the simulated median SHMRs also increase noticeably due to the limited number of high-mass satellites. For example, there are approximately 30 simulated and 3000 generated samples that have satellites in the last mass bin of the SHMR. Thus, even for this mass bin, the residuals of the median SHMRs between NeHOD and the simulations remain within 1–2 error bars of the simulations.

Next, we compare the average scatters of the SHMRs, estimated as the differences between the average 84th and 16th percentiles. This is particularly important since second-order statistics, such as the SHMR scatters, contain a wealth of information about formation history of galaxies and cosmological (e.g., Truong et al. 2018; Farahi et al. 2018, 2020, 2022; Anbajagane et al. 2020; Hadzhiyska et al. 2021; Gabrielpillai et al. 2022). From the bottom panels of Figure 7, we see a remarkable agreement between the NeHOD and the sim-

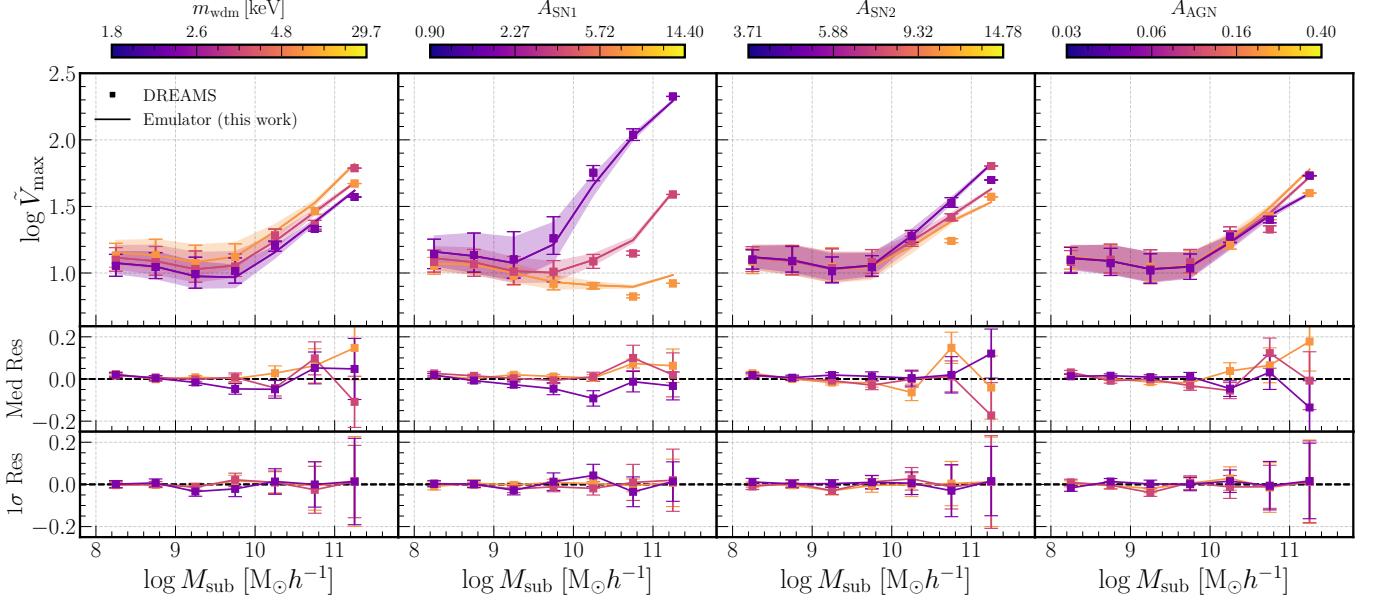
ulations. This suggests that NeHOD can accurately capture not only the central trends but also the intrinsic scatter in the SHMR, which is crucial for understanding the underlying physics governing galaxy formation.

#### 4.2.3. Concentration-mass relation

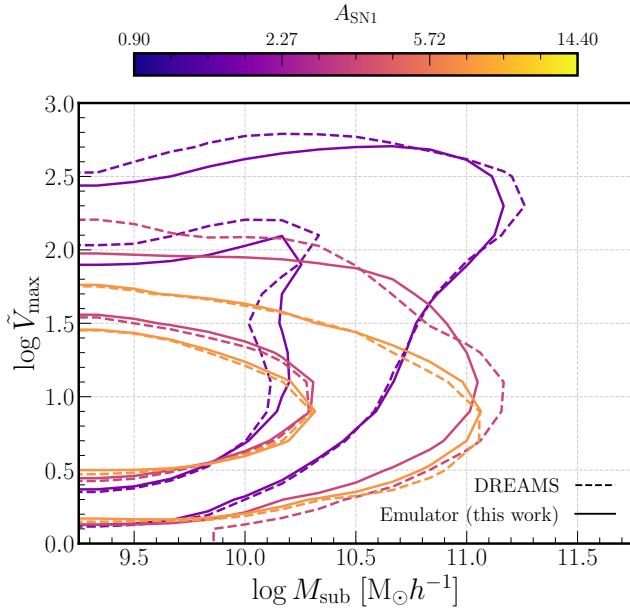
The DM concentration describes the internal structure of a halo and is known to correlate strongly with both the formation time and the environment of the halo (e.g., Nusser & Sheth 1999; van den Bosch 2002; Wechsler et al. 2006; Correa et al. 2015). A key relation in galaxy formation studies is the concentration-mass relation, which links the DM concentration to the halo mass. This will also provide an important validation test for predicting inter-galaxy properties with NeHOD.

We employ a similar procedure with the SHMR to calculate the concentration-mass relations. To summarize, for each halo, the median, 16th, and 84th percentiles of the *concentration proxy*,  $\tilde{V}_{\text{max}}$ , are calculated for  $M_{\text{sub}}$  bins of 0.5 dex width. Then, for each bin of the simulation parameters, we estimate the average and standard error of these percentiles across all halos. Note that as with central galaxies, the distribution of  $\tilde{V}_{\text{max}}$  is bimodal and strongly correlates with the satellite stellar mass and SN feedback parameters, especially  $A_{\text{SN1}}$ . However, the bimodality is much less prominent for the satellite subhalos, making the average percentiles acceptable summaries for the concentration-mass relations.

Figure 8 presents the concentration-mass relations. As with the SHMR, we show the average medians, 16th and



**Figure 8.** The concentration-mass relations of satellites generated by NeHOD and extracted from the DREAMS simulations. Panels are the same as in Figure 7.



**Figure 9.** The 2-dimensional distributions of total mass and concentration proxy of satellites generated by NeHOD (solid) and extracted from the DREAMS simulations (dashed). The 68% and 95% contours are shown for the three bins of  $A_{SN1}$ .

84th percentiles in the top rows, the residuals of the medians in the middle rows, and the residuals of the scatters in the bottom rows. Overall, the relations vary most significantly with  $A_{SN1}$ . Additionally, the concentration proxy decreases with increasing WDM mass, suggesting that WDM halos are less concentrated than CDM halos. This further supports the notion that WDM halos form later, thus affecting the

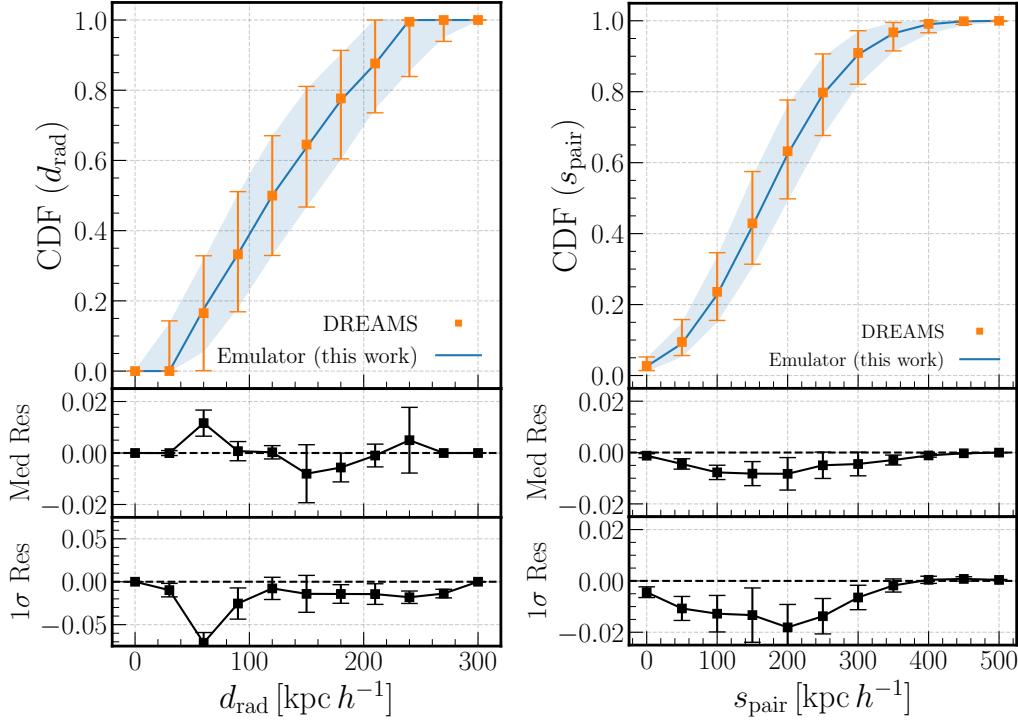
stellar mass as seen in Figure 7. We leave a more detailed investigation to future work.

The concentration-mass relations between NeHOD and the simulations agree remarkably well, with both the average medians and scatters agreeing within 1-2 error bars. As with the SHMRs, the errors in estimating both the average medians and scatters of the simulations increase significantly at high  $M_{sub}$  due to the limited number of satellites at these masses.

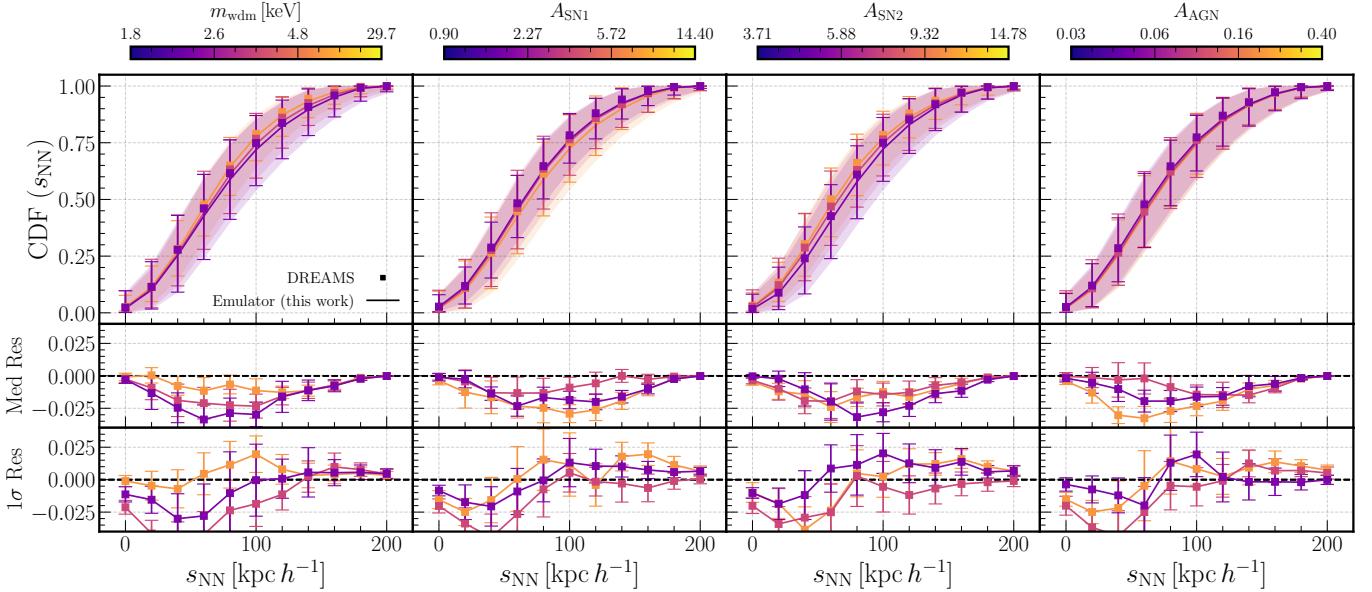
To demonstrate that NeHOD can capture the bimodality in the DM concentration proxy  $\tilde{V}_{max}$  of satellites, we compare the 2-dimensional distributions of  $\tilde{V}_{max}$  and  $M_{sub}$ . Figure 9 shows the 68% and 95% contours of the  $\tilde{V}_{max}$  and  $M_{sub}$  distributions as a function of  $A_{SN1}$ . We only show  $A_{SN1}$  since it has the most significant impact on the concentration, as can be seen in Figure 8. The distributions of  $\tilde{V}_{max}$  and  $M_{sub}$  exhibit a clear bimodal trend, although this bimodality is less pronounced than what we observe in the central galaxies (Figure 4). Overall, the contours predicted by NeHOD align well with those extracted from the simulations. However, the contours in the lowest  $A_{SN1}$  bin show less agreement, likely because the bimodality is more pronounced in this bin and thus requires more training samples to capture fully.

#### 4.2.4. Spatial distribution and clustering

As discussed in §3, the NeHOD framework is unique in its point-cloud approach for modeling the phase-space of satellites. Point-cloud can resolve arbitrarily small spatial scales, down to the spatial resolution of the simulations, and thus are better than binning/voxelization for representing sparse data, including satellite distributions. Here, we provide the



**Figure 10.** The cumulative distribution functions (CDFs) of radial distances (left) and pairwise separations (right). The top panels show the medians, 16th, and 84th percentiles of the CDFs for the NeHOD samples (blue band) and the simulations (error bars). The middle and bottom panels show the residuals between the NeHOD samples and the simulations of the medians and the  $1\sigma$  intervals, respectively.



**Figure 11.** The cumulative distributions of the nearest-neighbor separations of satellites generated by NeHOD and extracted from the DREAMS simulations. Panels are the same as Figure 7.

validation tests for the spatial distribution and clustering of the satellites.

We first compute the radial distribution and the pairwise separation distribution of satellites generated by NeHOD and compare them to those of the simulations. The radial distribution is important for probing the underlying subhalo distribution (e.g., Nagai & Kravtsov 2005), the disruption of subhalo and satellites due to the central galaxy (e.g., Samuel et al. 2020), and the DM model (e.g., Nadler et al. 2021a; Lovell et al. 2021). It is also a common statistic reported by satellite galaxy surveys (e.g., Carlsten et al. 2020; Wu et al. 2022a; Mao et al. 2024), which provides a basis for comparing NeHOD predictions with the results of these surveys. However, before any direct comparison can be made, we note that only *projected* radial distributions are reported in these surveys. Additionally, sample selection effects need to be properly accounted for. On the other hand, the pairwise separation distribution is a measure of the 1-halo clustering and is important for applications from creating mock galaxies catalogs for strong lensing (e.g., Wagner-Carena et al. 2023; Gilman et al. 2024) and galaxy surveys (e.g., Smith et al. 2024; Hamaus et al. 2022) to semi-analytic modeling of satellite distributions (e.g., Jiang et al. 2021; Green et al. 2022; Folsom et al. 2023).

Figure 10 shows the normalized, cumulative distribution of the radial distances  $d_{\text{rad}}$ (left) and the pairwise separations  $s_{\text{pair}}$  (right). We compute and display the median, 16th, and 84th percentiles of each distribution across *all* halos in the NeHOD samples and the simulations. As with the mass functions, we estimate the percentiles and their errors using bootstrapping. The top panels show the percentiles of the distributions, while the middle and bottom panels show the residuals between NeHOD and the simulations of the medians and  $1\sigma$  intervals respectively.

We do not show results for individual bins of simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN}1}, A_{\text{SN}2}, A_{\text{AGN}}\}$ , since we do not observe any significant variations in both the normalized radial distribution and pairwise separation distribution with respect to these parameters. This is despite differences in the stellar mass  $M_{\text{cent},\star}$  and concentration proxy  $\tilde{V}_{\text{max}}$  of the central galaxies between bins of  $A_{\text{SN}1}$  and  $A_{\text{SN}2}$  (see Figure 3). Note that the radial distribution is slightly more concentrated at higher values of  $A_{\text{SN}2}$ , though this effect is minor compared to the considerable scatter in the distribution, likely due to halo-to-halo variations. Additionally, while the normalized distributions are relatively invariant to the simulation parameters, the *un-normalized* distributions are not because the simulation parameters still influence the total number of satellites.

In general, we see a good agreement between the median distributions of the NeHOD samples and the simulations, with the residuals remaining with 1-2 error bars. We find

that NeHOD underestimates the  $1\sigma$  intervals of the radial distributions. Except for 1 radial bin at around  $d_{\text{rad}} \approx 50$  kpc, the biases typically remain with 1-2 error bars and are small compared to the values of the  $1\sigma$  intervals.

On the other hand, we observe a clear bias in the pairwise separation distributions at lower values of  $s_{\text{pair}}$ . Specifically, NeHOD underestimates both the median and the  $1\sigma$  interval of the cumulative distribution up to 1% and 2% respectively, with the largest discrepancy occurring at around  $s_{\text{pair}} \approx 200$  kpc  $h^{-1}$ . This indicates that while satellites generated by NeHOD have similar radial extensions to those in the simulations, they are typically less clustered. We quantify this discrepancy by computing the median and maximum  $s_{\text{pair}}$  between satellites in the same halo for the NeHOD samples and the simulations. On average, the median  $s_{\text{pair}}$  are  $167.94 \pm 34.34$  kpc  $h^{-1}$  and  $165.97 \pm 36.15$  kpc  $h^{-1}$  for the NeHOD and the simulations, respectively. This corresponds to an overestimation by approximately 1%, although this effect is much smaller compared to the spread. Likewise, the maximum  $s_{\text{pair}}$  are  $372.35 \pm 78.02$  kpc  $h^{-1}$  and  $361.32 \pm 83.98$  kpc  $h^{-1}$  for the NeHOD and the simulations, respectively, indicating an overestimation by approximately 3%.

To further investigate the spatial clustering, we identify the nearest neighbor (including the central galaxy) for each satellite in our samples and show the cumulative distribution of nearest-neighbor separations  $s_{\text{NN}}$  in Figure 11. Similar to Figure 10, we show the medians, 16th and 84th percentiles in the top row, the median residuals in the middle row, and the  $1\sigma$  interval residuals in the bottom row. The percentiles and their errors are estimated using the same bootstrapping procedure. We briefly note that, in contrast to the radial distances and pairwise separations, the distributions of  $s_{\text{NN}}$  vary strongly with respect to  $m_{\text{WDM}}$ ,  $A_{\text{SN}1}$ , and  $A_{\text{SN}2}$ . Thus, in Figure 11, we show results for multiple bins of simulation parameters. Additionally, this implies that the simulation parameters may influence the small-scale, local clustering of satellites more strongly than they do for the overall radial structure.

While NeHOD can capture the general trends in the distributions of  $s_{\text{NN}}$  with respect to the simulation parameters, we observe similar biases at small distances. Across all parameter bins, NeHOD underestimates both the medians and  $1\sigma$  intervals up to approximately 2-3%, as evident from the middle and bottom rows. The largest biases occur at around  $s_{\text{NN}} \approx 50$  kpc  $h^{-1}$ . The biases in the  $s_{\text{NN}}$  distribution likely propagate into larger scales, resulting in biases in the pairwise separation  $s_{\text{pair}}$  distributions, as seen in Figure 10.

As with the pairwise distances, we compute the median and maximum  $s_{\text{NN}}$  of satellites within the same halo in the NeHOD samples and the simulations. The median values of  $s_{\text{NN}}$  are  $93.40 \pm 29.6$  kpc  $h^{-1}$  and  $90.72 \pm 31.43$  kpc  $h^{-1}$  for NeHOD and the simulations, respectively. This corresponds

to an overestimation by about 3%, similar to the values obtained from the pairwise distances. However, we find that the NeHOD overestimates the maximum  $s_{\text{NN}}$  by about 13%, with NeHOD and simulation values of  $230.95 \pm 61.91 \text{ kpc } h^{-1}$  and  $203.48 \pm 49.56 \text{ kpc } h^{-1}$ , respectively. This further supports that while NeHOD can capture the overall radial extensions of satellites (as seen in the left panels of Figure 10), it is unable to capture the local clustering between satellites. Lastly, we note that we also observe a similar bias in the velocity space, in which the relative velocities of NeHOD satellites are larger than seen in the simulations. We show this result in Appendix C.3.

To summarize, we examine the radial distribution, pairwise separation distribution, and nearest-neighbor distance distribution of satellites in the NeHOD halos and the simulations. Overall, the shapes of distributions predicted by NeHOD generally match those in the simulations, including variations with respect to the simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . However, we find that while satellites generated by NeHOD match the radial extensions of those in the simulations, they are typically less clustered. The median and maximum pairwise separations between satellites in the NeHOD halos are larger than the simulation values by about 1-3%. However, while the median nearest-neighbor separations between satellites in the NeHOD halos are larger by about 3%, the maximum nearest-neighbor separations can differ by 13%. We thus conclude that NeHOD cannot yet capture the local clustering between satellites and caution readers against applying NeHOD in scenarios where accurate modeling of local clustering is important. We further discuss the current limitations of our model and potential ways to address them in future work in §5.

## 5. DISCUSSION

### 5.1. Comparison with traditional methods

Generative models such as NeHOD can be considered non-parametric techniques for modeling the galaxy-halo connection. Compared to the traditional techniques such as HOD, subhalo abundance matching (SHAM), and semi-analytic models (SAM), NeHOD does not make explicit assumptions about the distributions of halo and galaxy properties. Instead, it learns these distributions directly from simulations, where the underlying assumptions are embedded implicitly. In contrast, HOD assumes a specific form for the galaxy-halo relationship, typically relying on a parametrized model for the number of galaxies as a function of halo mass (Neyman & Scott 1952). SHAM assumes a monotonic relationship between a galaxy property and a corresponding halo property (e.g., Kravtsov et al. 2004; Vale & Ostriker 2004, 2006; Conroy et al. 2006; Behroozi et al. 2010; Moster et al. 2010; Trujillo-Gomez et al. 2011; Reddick et al. 2013; Chaves-Montero et al. 2016). SAMs rely on detailed analytic pre-

scriptions for various physical processes, such as gas cooling, star formation, and feedback, making assumptions about how these processes operate and interact over cosmic time within the framework of hierarchical structure formation (e.g., Cole et al. 1994, 2000; Somerville et al. 2015; Yung et al. 2019; Hutter et al. 2021; Elliott et al. 2021). Below we discuss the key advantages of NeHOD:

- The use of neural networks for non-parametric modeling allows NeHOD to learn more complex distributions. This is especially important for modeling non-Gaussian features in target distributions that are otherwise challenging to capture. For example, both the flows and the VDM can capture the bimodality in the DM concentration proxy  $\tilde{V}_{\text{max}}$  of the central and satellite galaxies, as shown in Figure 4 and Figure 9 respectively.

Additionally, we show that NeHOD can capture second-order statistics, such as the scatters in the target distributions and relations. Second-order statistics are important because they (1) are observationally quantifiable (Farahi et al. 2019), (2) contain information in (e.g., Truong et al. 2018; Farahi et al. 2018, 2020, 2022; Anbajagane et al. 2020; Hadzhiyska et al. 2021; Gabrielpillai et al. 2022), and (3) are a source of systematics in cosmological inference (Zhang et al. 2024). Although we have primarily validated the scatters in the distributions and relations in this work, future studies will extend this analysis to include additional important statistics, such as correlations between the scatters of different properties (e.g., Farahi et al. 2020; Anbajagane et al. 2020). As NeHOD does not make explicit assumptions about the underlying distributions, extending this analysis is both natural and straightforward.

- The VDM and the normalizing flows employed in NeHOD can model high-dimensional distributions, as they learn directly from the data without requiring explicit assumptions for each parameter. Parametric methods often become less efficient and more complex as the number of parameters increases. Additionally, each new parameter typically necessitates additional modeling assumptions. The traditional HOD method does not incorporate information other than the halo mass. SHAMs assume a monotonic relationship between a galaxy property and a halo property, which can oversimplify the underlying physics, particularly in high-dimensional spaces where additional factors like subhalo disruption and galaxy mergers play a significant role. SAMs, while more flexible, require recalibration whenever new parameters or physical processes are introduced. Additionally, modeling subhalos in both SAMs and SHAMs require full merger trees, which can be expensive to generate. In contrast, Ne-

HOD scales naturally with dimensionality, efficiently handling larger parameter spaces by modeling complex, data-driven relationships.

- NeHOD can be readily extended to different DM scenarios and baryonic physics, provided that corresponding simulations are available. Unlike traditional parametric techniques, which require specific assumptions and models tailored to each scenario, NeHOD learns directly from the simulations, allowing it to adapt to a wide range of physical conditions without the need for reparametrization. For example, HODs and SHAMs would require adjustments to their underlying models. SAMs need to be re-calibrated, which involves re-tuning a large number of parameters that govern various physical processes. This can be time-consuming, as it requires extensive comparisons with observational data to ensure that the model predictions remain consistent across different scales and epochs. In comparison, re-training NeHOD simply requires feeding it new data corresponding to the updated simulation scenarios and potentially a new set of output/input features. This flexibility makes NeHOD particularly advantageous for exploring complex or new models that are challenging to capture with traditional methods.
- SHAMs and SAMs assign galaxies onto DM halos in N-body simulations, which can introduce systematic biases when comparing results with those from hydrodynamic simulations. For example, halo finders may lose track of DM subhalos in N-body simulations due to tidal stripping or numerical resolutions, even though the galaxies residing within these DM subhalos may still survive. These galaxies, known as "orphan galaxies," can introduce significant systematics into SHAMs and SAMs, which often struggle to accurately track them (Kitzbichler & White 2008; Guo & White 2014; Pujol et al. 2017). Although the effects of orphan satellites can be quantified and somewhat accounted for in SHAMs and SAMs (e.g., De Lucia & Blaizot 2007; Benson 2012; Gonzalez-Perez et al. 2014; Delfino et al. 2022; Harrold et al. 2024), this requires additional modeling and assumptions. Therefore, it is advantageous to learn directly from hydrodynamic simulations. NeHOD follows this approach by generating galaxies directly from hydrodynamic simulations, with their properties conditioned on halo characteristics that are less susceptible to systematics related to tidal stripping and numerical resolution.

Lastly, we note that machine learning emulators are differentiable, i.e. their outputs change smoothly and predictably with respect to their inputs. In the context of simulations, this allows for more efficient optimization and integration

with gradient-based techniques and can significantly improve downstream tasks such as parameter inference. Building an end-to-end differentiable simulation, whether neural network-based or otherwise, is a growing area of research in astrophysics and cosmology (e.g., Hearn et al. 2021; Modi et al. 2021a,b; Li et al. 2024; Kern 2024). Both the flows and the VDM in NeHOD can individually be considered differentiable. However, NeHOD as a whole is not end-to-end differentiable, since the computation of the number of satellites  $N_{\text{sat}}$  involves a rounding operation (as the flows predict  $\log N_{\text{sat}}$  in practice, see §3.2). In future work, we will experiment with alternative approaches to achieve full end-to-end differentiability.

### 5.2. Comparison with other generative models

Similar non-parametric, machine learning-based techniques have also been employed in various aspects of galaxy formation and cosmology (e.g., Lovell et al. 2023b; Nguyen et al. 2023b; Brown et al. 2024; Lee et al. 2024). Here, we highlight the key advantages of NeHOD:

- Existing emulators have employed techniques such as the Gaussian Process for emulating summary statistics such as satellite counts, and mass functions. In contrast, NeHOD uses a Transformed-based VDM that can model individual satellites, from which we can extract summary statistics as shown in §4. This approach provides access to more detailed information, enabling a wider range of applications, such as generating mock subhalo catalogs, which summary statistics alone cannot offer.
- As discussed in §3, we represent satellite subhalos as point clouds. In contrast to binning and voxelization, point clouds carry more information about the data and can resolve smaller spatial scales, down to the resolution of the simulations. While convolutional neural networks (CNNs) have been successfully used for HOD applications, as shown in Bourdin et al. 2024, it is unclear if they can effectively model the small spatial scales and sparse spatial distributions addressed in this work. Accurately modeling small scales with CNNs would require decreasing the grid size, which significantly increases computational cost and can be challenging to handle with sparse data.
- As also discussed in §3.3, using VDMs for learning point-cloud data has notable advantages compared to other generative machine learning models. Compared to GANs, VDMs are more stable and less prone to mode collapse. They also provide a tractable way to evaluate the likelihood, making them suitable for inference and anomaly detection tasks. Additionally, VDM is more expressive than VAE and normalizing

flows (e.g., Huang et al. 2021; Luo 2022). Compared to normalizing flows, VDMs also scale better with high-dimensional data.

### 5.3. Current limitations and future outlook

We now discuss the current limitations and future prospects of the NeHOD framework. One significant limitation of NeHOD is its inability to fully capture the local clustering of satellites. While the radial distributions of satellites generally agree, satellites generated by NeHOD are typically less clustered than those in the simulations, as shown in §4.2.4. Below we provide a few potential explanations and solutions.

- A potential explanation is a phenomenon in machine learning literature known as the “spectral bias”, where coordinate-based neural networks have difficulty learning high-frequency functions. During training, these networks tend to fit the low-frequency components of a target before the high-frequency components (Rahaman et al. 2018; Basri et al. 2019, 2020). Recent studies demonstrate that using Fourier features can assist networks in learning high-frequency components (Rahaman et al. 2018; Tancik et al. 2020), including 1-dimensional time or position coordinates (Xu et al. 2019; Mehran Kazemi et al. 2019; Vaswani et al. 2023), and more recently, 3-dimensional coordinates (Zhong et al. 2019; Mildenhall et al. 2020). We will explore this approach by employing encoding layers that map the input positions and velocities of satellites onto the Fourier space. Note that this is similar to the sinusoidal encoding layer used for the 1-dimensional diffusion time step in our Transformer (described in §3.3). However, extending this to 3-dimensional positions and velocities is non-trivial, so we leave this for future work.
- In our setup, we have conditioned the halo and central galaxy flows only on the WDM mass and astrophysical parameters, without incorporating other factors. We expect that environmental factors will also play a significant role. Similarly, the VDM is conditioned on the properties of the halos and central galaxies but does not currently include aspects such as the morphology of the central galaxies or environmental information. Although the DM concentration proxy  $\tilde{V}_{\max}$  of the central galaxy is expected to correlate with the environment and formation histories, explicitly including these factors in future work could improve the model’s ability to capture local clustering more accurately.
- Additionally, satellites with a nearest neighbor closer than  $50 \text{ kpc } h^{-1}$  might have been recently or are currently being tidally disrupted. This disruption can affect SUBFIND’s ability to accurately identify and calculate their properties (e.g., Knebe et al. 2011; Onions

et al. 2012; Mansfield et al. 2024), resulting in noisier features that impact the training process.

- Lastly, even with data augmentation, a training dataset of 1,000 simulations might be too small for the VDM to fully model the 9-dimensional feature space for all satellites. We expect more training samples from future simulations will reduce this bias, although the extent of this improvement is currently unclear. This is potentially another major caveat of the current model – we do not know *a priori* how many simulations are necessary for the distributions of modeled properties to be satisfactory. Studying the scaling behavior of downstream properties (e.g., loss function, summary statistics) with the number of training simulations can help answer this in future work.

In addition, we outline the general limitations of NeHOD:

- Our model is conditioned on the properties of the halo and central galaxy from the hydrodynamic simulation. Ideally, we would like to condition our model on the properties of the halo and central subhalo of the corresponding N-body simulation. We note that in this case, the positions of the halos in the N-body and hydrodynamic simulations may be different, so our model should also predict the displacement vector. We leave this for future work.
- As discussed in §2.2, the halos in the TNG-WDM suite are constrained to a relatively narrow mass range, similar to that of the Milky Way. In addition, to manage the computational costs of simulating dense environments, these halos are selected based on an isolation criterion (i.e., they are not close to halos with virial mass greater than  $7.2 \times 10^{11} \text{ M}_\odot h^{-1}$ ). For other applications, we will train the models on halos with much wider variations in mass and environment to fully exploit our method’s capabilities.
- As discussed previously, although both the flows and the VDM are differentiable, the end-to-end NeHOD framework is not, due to the rounding operator when calculating  $N_{\text{sat}}$ . Future work will explore alternative approaches to achieve full end-to-end differentiability.
- Although we apply data augmentation techniques to incorporate rotational equivariance, our model does not explicitly impose this symmetry. Building such symmetry into our model may further improve its accuracy, performance, and simulation efficiency. Additionally, although not necessary for this work, in future work, we will explore incorporating E(3) symmetry, i.e. translation and rotational symmetry. This is important when

there is a positional offset between the halo and central galaxy, which is expected if we condition our model using halos from N-body simulations (instead of hydrodynamic simulations), as discussed in the previous point.

- As discussed in §3.1, the goal of this paper is to highlight the use of VDM and point clouds for modeling satellite galaxies. To simplify the analysis, we combine the properties of halos and central galaxies into the same normalizing flows, effectively modeling the *joint distribution* of halos and central galaxies. In future work, it is better to separate them and instead model the *conditional distribution* of central galaxies given the halos. An interesting avenue for exploration is using a hierarchical normalizing flows approach, similar to what is described in Lovell et al. (2023a). Specifically, we can employ one flow to model the conditional distribution of halos given the simulation parameters, and another to model the conditional distribution of central galaxies given the halo and simulation parameters. This can be helpful for applications in which we already have access to the halos but not the central galaxies, e.g., painting galaxies onto existing halos in N-body simulations.

## 6. CONCLUSION

In this work, we develop an emulator for Milky Way-mass halos, and their central and satellite galaxies by using normalizing flows and variational diffusion models (VDM). Our framework, NeHOD, is hierarchical:

1. The normalizing flows capture the conditional distribution of halo and central galaxies given the simulation parameters.
2. The VDM then learns the properties of satellite galaxies given the halo, central galaxy, and simulation parameters.

This approach allows us to capture the complex dependencies between different levels of structure in the simulations.

We represent satellite galaxies in each halo as a point cloud centered around the halo center. This eliminates the need for binning/voxelization and allows simulation details at fine spatial scales (up to the resolution often simulations) to be captured. The VDM employs a Transformer-based score model, which leverages self-attention mechanisms, to effectively model this point cloud.

We train the NeHOD framework on the TNG-WDM simulation suite of the DREAMS project, which consists of 1024 simulations with varying WDM mass and astrophysical parameters, particularly the strength of baryonic feedback from supernova and active galactic nuclei. We generate a test

dataset by sampling halos and central halos using the flows, and then satellite galaxies using the VDM. Given the WDM mass and astrophysical parameters, the flows sample key properties such as the virial mass and stellar mass of the halos, as well as the total mass, stellar mass, DM concentration proxy, and velocity offset for the central galaxies. The VDM then takes the generated halos and central galaxies, along with the simulation parameters, to generate satellite galaxies, including the positions, velocities, halo and stellar mass, and DM concentration of each satellite. We summarize our main findings below:

- In §4.1 and Appendix C.1, we compare the halos and central galaxies generated by NeHOD with those in the simulations. Overall, the flows effectively capture the distributions of halo and galaxy properties, along with their complex dependencies on the simulation parameters. The edges of the distributions tend to show less agreement, which is expected given the limited number of training samples. Most notably, both the DM concentration proxy and the velocity offset display strong non-Gaussian features; the concentration proxy is bimodal and strongly correlates with the halo stellar mass, while the velocity offset has a long tail extending toward high velocities. Despite the complexities of these distributions, the flows accurately capture them.
- In §4.2, we compare key summary statistics between NeHOD satellites and the simulations. We demonstrate that NeHOD can capture the satellite stellar mass functions, stellar-to-halo mass relations, and concentration-mass relations accurately. Notably, NeHOD can capture both the medians and the scatters of the target distributions and relations. Additionally, we show that NeHOD can model complex relations such as the bi-modal distributions of the concentration proxy  $\tilde{V}_{\max}$  of satellite galaxies (Figure 9). However, similar to the flows, the model tends to struggle towards the extreme values (e.g., high mass range), where there are fewer training samples to learn from. Nonetheless, we note discrepancies between the NeHOD samples and simulations are generally within 1-2 error bars and are typically much smaller than the intrinsic spreads from halo-to-halo variations.
- In §4.2.4, we examine the spatial distribution of satellites, including the radial distribution, pairwise separations, and nearest-neighbor separations. The overall trends predicted by NeHOD are consistent with the simulations. However, we find that although NeHOD satellites have similar radial distributions to those from the simulations, they are typically less clustered than those in the simulations. We observe similar discrepancies in the velocity space (Appendix C.3), where

NeHOD satellites tend to move faster relatively to each other compared to those in the simulations. We discuss the limitations of the current iteration of NeHOD and potential ways to overcome them in §5. For now, we caution readers against over-interpreting results when using NeHOD for applications in which accurate modeling of small-scale clustering and velocity distributions is crucial.

- Although this work is primarily on the methodology and validation of NeHOD, emulating halos and galaxies under alternative DM scenarios and baryonic feedback models has significant implications for constraining DM properties and understanding the broader impact of these factors on galaxy formation and evolution. NeHOD will enable studies of the intricate interaction between the effects of new physics and astrophysics at both *the field as well as summary levels*. During our analysis, we incidentally uncovered several interesting and potentially unexpected correlations between astrophysical and DM parameters while evaluating summary statistics.

In §4.1, we observe that the DM concentration proxy  $\tilde{V}_{\max}$  of central galaxies follows a bimodal distribution, which may be linked to the bimodality in the formation histories of their halos. We see a similar, albeit less prominent, trend in  $\tilde{V}_{\max}$  of satellite galaxies in §4.2.3. In §4.2.2, when examining the stellar-to-halo mass relations, we find that below approximately  $10^{10} M_{\odot} h^{-1}$  total mass, WDM satellites generally have higher stellar masses than their CDM counterparts for the same DM mass, possibly due to WDM halos forming later, leading to delayed star formation. In §4.2.3, when examining the concentration-mass relations, we find that WDM satellites typically have lower concentrations at the same total mass, likely due to the later formation times of WDM halos. As we continue to improve NeHOD and further explore the complex interplay between DM and astrophysical parameters, addressing these correlations and understanding their implications—while also assessing whether they are numerical artifacts—will be a key priority of the DREAMS project.

Overall, the NeHOD framework offers a novel approach for non-parametric modeling of the galaxy-halo connection at the field level, all while operating at a computational cost that is several orders of magnitude lower than that of hydrodynamic simulations. We further emphasize that NeHOD can model individual satellites and complex relationships between halos, galaxies, and the simulation parameters, while not making any explicit assumptions about the underlying distributions. Instead, NeHOD directly learns from the simulations, where

the physical assumptions are embedded implicitly. This allows NeHOD to capture more complex and high-dimensional distributions, as well as offering more flexibility to extend to different DM and baryonic physics scenarios, as discussed in §5.

We plan to utilize NeHOD for generating mock galaxy catalogs for applications such as galaxy clustering, weak and strong lensing, and potentially intensity mapping. For example, in a follow-up project, we will use NeHOD to generate strong lensing catalogs and compare our model against traditional methods such as SAM, SHAM, and HOD, particularly in the context of galaxy clustering at the field level. We envision NeHOD to be especially valuable in applications where populating halos with galaxies is crucial. Additionally, since both the flows and VDM components of NeHOD allow efficient evaluation and sampling of the likelihood function, they hold potential for use in cosmological parameter inference and even outlier detection, though further tests are needed.

In conclusion, NeHOD offers a powerful and flexible framework for exploring various DM models and astrophysical scenarios at the field level. Its ability to model complex, high-dimensional distributions and flexibility makes it a promising tool for a wide range of astrophysical and cosmological applications. As we continue to develop and improve NeHOD, we anticipate its application in even more areas, from generating mock catalogs to conducting cosmological parameter inference.

**Software:** corner (Foreman-Mackey 2016), IPython (Perez & Granger 2007), jax (Bradbury et al. 2018), Jupyter (Kluyver et al. 2016), Matplotlib (Hunter 2007), NumPy (Harris et al. 2020), SciPy (Virtanen et al. 2020), Pandas (pandas development team 2020), PyTorch (Paszke et al. 2019), PyTorch Lightning (Falcon et al. 2020), zuko (Rozet et al. 2022)

## DATA AVAILABILITY

The code used for this article is available at [https://github.com/trivnguyen/nehod\\_torch](https://github.com/trivnguyen/nehod_torch), which includes example tutorials and trained models (both the flows and VDM). Only a subset of the simulation data, in the form of the FoF/SUBFIND catalog, is included. For more information on the DREAMS simulation suites and full data access, please visit at <https://dreams-project.readthedocs.io/en/latest/>.

## ACKNOWLEDGMENTS

We thank Mariangela Lisanti, Justin Read, and Xiaowei Ou for many meaningful discussions and suggestions, and Lucy Reading-Ikkanda for assisting with Figure 1.

The Center for Computational Astrophysics at the Flatiron Institute is supported by the Simons Foundation. The computations in this work were, in part, run at facilities supported by the Scientific Computing Core at the Flatiron In-

stitute, a division of the Simons Foundation. The data used in this work were, in part, hosted on equipment supported by the Scientific Computing Core at the Flatiron Institute, a division of the Simons Foundation. TN, SM, CC, and LN are supported by the National Science Foundation under Cooperative Agreement PHY-2019786 (The NSF AI Institute for Artificial Intelligence and Fundamental Interactions, <http://iaifi.org/>). SM is partly supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics of U.S. Department of Energy under grant Contract Number DE-SC0012567. This work was performed in part at the Aspen Center for Physics, which is supported by National Science Foundation grant PHY-2210452. PT acknowledges

support from NSF-AST 2346977. JBM acknowledges support from NSF AST-2307354 and AST-2408637. LN is also supported by the Sloan Fellowship, the NSF CAREER award 2337864, and NSF award 2307788. FYCR acknowledges the support of NSF award 2327192. KEK is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-2039656. DAA acknowledges support by NSF grant AST-2108944, NASA grant ATP23-0156, STScI grants JWST-GO-01712.009-A and JWST-AR-04357.001-A, Simons Foundation Award CCA-1018464, and Cottrell Scholar Award CS-CSA-2023-028 by the Research Corporation for Science Advancement.

## REFERENCES

- Abbott, T. M. C., Abdalla, F. B., Allam, S., et al. 2018, ApJS, 239, 18, doi: [10.3847/1538-4365/aae9f0](https://doi.org/10.3847/1538-4365/aae9f0)
- Abbott, T. M. C., Adamów, M., Aguena, M., et al. 2021, ApJS, 255, 20, doi: [10.3847/1538-4365/ac00b3](https://doi.org/10.3847/1538-4365/ac00b3)
- Akeson, R., Armus, L., Bachelet, E., et al. 2019, arXiv e-prints, arXiv:1902.05569, doi: [10.48550/arXiv.1902.05569](https://doi.org/10.48550/arXiv.1902.05569)
- Anbajagane, D., Evrard, A. E., Farahi, A., et al. 2020, MNRAS, 495, 686, doi: [10.1093/mnras/staa1147](https://doi.org/10.1093/mnras/staa1147)
- Basri, R., Galun, M., Geifman, A., et al. 2020, arXiv e-prints, arXiv:2003.04560, doi: [10.48550/arXiv.2003.04560](https://doi.org/10.48550/arXiv.2003.04560)
- Basri, R., Jacobs, D., Kasten, Y., & Krritchman, S. 2019, arXiv e-prints, arXiv:1906.00425, doi: [10.48550/arXiv.1906.00425](https://doi.org/10.48550/arXiv.1906.00425)
- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, MNRAS, 488, 3143, doi: [10.1093/mnras/stz1182](https://doi.org/10.1093/mnras/stz1182)
- Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, ApJ, 717, 379, doi: [10.1088/0004-637X/717/1/379](https://doi.org/10.1088/0004-637X/717/1/379)
- Benson, A. J. 2012, NewA, 17, 175, doi: [10.1016/j.newast.2011.07.004](https://doi.org/10.1016/j.newast.2011.07.004)
- Benson, A. J., Cole, S., Frenk, C. S., Baugh, C. M., & Lacey, C. G. 2000, MNRAS, 311, 793, doi: [10.1046/j.1365-8711.2000.03101.x](https://doi.org/10.1046/j.1365-8711.2000.03101.x)
- Berlind, A. A., Weinberg, D. H., Benson, A. J., et al. 2003, ApJ, 593, 1, doi: [10.1086/376517](https://doi.org/10.1086/376517)
- Bhowmick, A. K., Campbell, D., Di Matteo, T., & Feng, Y. 2018, MNRAS, 480, 3177, doi: [10.1093/mnras/sty2128](https://doi.org/10.1093/mnras/sty2128)
- Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition
- Bland-Hawthorn, J., & Gerhard, O. 2016, ARA&A, 54, 529, doi: [10.1146/annurev-astro-081915-023441](https://doi.org/10.1146/annurev-astro-081915-023441)
- Bode, P., Ostriker, J. P., & Turok, N. 2001, ApJ, 556, 93, doi: [10.1086/321541](https://doi.org/10.1086/321541)
- Bose, S., Eisenstein, D. J., Hernquist, L., et al. 2019, MNRAS, 490, 5693, doi: [10.1093/mnras/stz2546](https://doi.org/10.1093/mnras/stz2546)
- Bose, S., Hellwing, W. A., Frenk, C. S., et al. 2016, MNRAS, 455, 318, doi: [10.1093/mnras/stv2294](https://doi.org/10.1093/mnras/stv2294)
- Bourdin, A., Legin, R., Ho, M., et al. 2024, arXiv e-prints, arXiv:2408.00839, doi: [10.48550/arXiv.2408.00839](https://doi.org/10.48550/arXiv.2408.00839)
- Bradbury, J., Frostig, R., Hawkins, P., et al. 2018, JAX: composable transformations of Python+NumPy programs, 0.3.13. <http://github.com/google/jax>
- Brown, S. T., Fattah, A., McCarthy, I. G., et al. 2024, arXiv e-prints, arXiv:2403.11692, doi: [10.48550/arXiv.2403.11692](https://doi.org/10.48550/arXiv.2403.11692)
- Buck, T., Macciò, A. V., Dutton, A. A., Obreja, A., & Frings, J. 2019, MNRAS, 483, 1314, doi: [10.1093/mnras/sty2913](https://doi.org/10.1093/mnras/sty2913)
- Bullock, J. S., & Boylan-Kolchin, M. 2017, ARA&A, 55, 343, doi: [10.1146/annurev-astro-091916-055313](https://doi.org/10.1146/annurev-astro-091916-055313)
- Carlsten, S. G., Greene, J. E., Beaton, R. L., Danieli, S., & Greco, J. P. 2022, ApJ, 933, 47, doi: [10.3847/1538-4357/ac6fd7](https://doi.org/10.3847/1538-4357/ac6fd7)
- Carlsten, S. G., Greene, J. E., Peter, A. H. G., Greco, J. P., & Beaton, R. L. 2020, ApJ, 902, 124, doi: [10.3847/1538-4357/abb60b](https://doi.org/10.3847/1538-4357/abb60b)
- Chaves-Montero, J., Angulo, R. E., Schaye, J., et al. 2016, MNRAS, 460, 3100, doi: [10.1093/mnras/stw1225](https://doi.org/10.1093/mnras/stw1225)
- Chen, T. 2023, arXiv e-prints, arXiv:2301.10972, doi: [10.48550/arXiv.2301.10972](https://doi.org/10.48550/arXiv.2301.10972)
- Cole, S., Aragon-Salamanca, A., Frenk, C. S., Navarro, J. F., & Zepf, S. E. 1994, MNRAS, 271, 781, doi: [10.1093/mnras/271.4.781](https://doi.org/10.1093/mnras/271.4.781)
- Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, MNRAS, 319, 168, doi: [10.1046/j.1365-8711.2000.03879.x](https://doi.org/10.1046/j.1365-8711.2000.03879.x)
- Conroy, C., & Wechsler, R. H. 2009, ApJ, 696, 620, doi: [10.1088/0004-637X/696/1/620](https://doi.org/10.1088/0004-637X/696/1/620)
- Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201, doi: [10.1086/503602](https://doi.org/10.1086/503602)
- Contreras, S., Angulo, R. E., Chaves-Montero, J., et al. 2024, arXiv e-prints, arXiv:2407.18912, doi: [10.48550/arXiv.2407.18912](https://doi.org/10.48550/arXiv.2407.18912)
- Correa, C. A., Wyithe, J. S. B., Schaye, J., & Duffy, A. R. 2015, MNRAS, 450, 1521, doi: [10.1093/mnras/stv697](https://doi.org/10.1093/mnras/stv697)

- Cranmer, K., Brehmer, J., & Louppe, G. 2020, Proceedings of the National Academy of Science, 117, 30055, doi: [10.1073/pnas.1912789117](https://doi.org/10.1073/pnas.1912789117)
- Croton, D. J. 2009, MNRAS, 394, 1109, doi: [10.1111/j.1365-2966.2009.14429.x](https://doi.org/10.1111/j.1365-2966.2009.14429.x)
- Cuesta-Lazaro, C., & Mishra-Sharma, S. 2023, arXiv e-prints, arXiv:2311.17141, doi: [10.48550/arXiv.2311.17141](https://doi.org/10.48550/arXiv.2311.17141)
- Danieli, S., Greene, J. E., Carlsten, S., et al. 2023, ApJ, 956, 6, doi: [10.3847/1538-4357/acefb0](https://doi.org/10.3847/1538-4357/acefb0)
- De Lucia, G., & Blaizot, J. 2007, MNRAS, 375, 2, doi: [10.1111/j.1365-2966.2006.11287.x](https://doi.org/10.1111/j.1365-2966.2006.11287.x)
- de Santi, N. S. M., Shao, H., Villaescusa-Navarro, F., et al. 2023, ApJ, 952, 69, doi: [10.3847/1538-4357/acd1e2](https://doi.org/10.3847/1538-4357/acd1e2)
- Delfino, F. M., Scóccola, C. G., Cora, S. A., Vega-Martínez, C. A., & Gargiulo, I. D. 2022, MNRAS, 510, 2900, doi: [10.1093/mnras/stab3494](https://doi.org/10.1093/mnras/stab3494)
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv e-prints, arXiv:1611.00036, doi: [10.48550/arXiv.1611.00036](https://doi.org/10.48550/arXiv.1611.00036)
- Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. 2018, arXiv e-prints, arXiv:1810.04805, doi: [10.48550/arXiv.1810.04805](https://doi.org/10.48550/arXiv.1810.04805)
- Dropulic, A., Shipp, N., Kim, S., et al. in prep, StreamGen: Connecting Semi-analytic Tidal Debris to Dark Matter Halos
- Durkan, C., Bekasov, A., Murray, I., & Papamakarios, G. 2019, arXiv e-prints, arXiv:1906.04032, doi: [10.48550/arXiv.1906.04032](https://doi.org/10.48550/arXiv.1906.04032)
- Efstathiou, A., Rowan-Robinson, M., & Siebenmorgen, R. 2000, MNRAS, 313, 734, doi: [10.1046/j.1365-8711.2000.03269.x](https://doi.org/10.1046/j.1365-8711.2000.03269.x)
- Elliott, E. J., Baugh, C. M., & Lacey, C. G. 2021, MNRAS, 506, 4011, doi: [10.1093/mnras/stab1837](https://doi.org/10.1093/mnras/stab1837)
- Engler, C., Pillepich, A., Pasquali, A., et al. 2021, MNRAS, 507, 4211, doi: [10.1093/mnras/stab2437](https://doi.org/10.1093/mnras/stab2437)
- Euclid Collaboration, Scaramella, R., Amiaux, J., et al. 2022, A&A, 662, A112, doi: [10.1051/0004-6361/202141938](https://doi.org/10.1051/0004-6361/202141938)
- Falcon, W., et al. 2020, PyTorchLightning/pytorch-lightning: 0.7.6 release, 0.7.6, Zenodo, doi: [10.5281/zenodo.3828935](https://doi.org/10.5281/zenodo.3828935)
- Farahi, A., Evrard, A. E., McCarthy, I., Barnes, D. J., & Kay, S. T. 2018, MNRAS, 478, 2618, doi: [10.1093/mnras/sty1179](https://doi.org/10.1093/mnras/sty1179)
- Farahi, A., Ho, M., & Trac, H. 2020, MNRAS, 493, 1361, doi: [10.1093/mnras/staa291](https://doi.org/10.1093/mnras/staa291)
- Farahi, A., Nagai, D., & Anbjagane, D. 2022, ApJ, 933, 48, doi: [10.3847/1538-4357/ac721e](https://doi.org/10.3847/1538-4357/ac721e)
- Farahi, A., Mulroy, S. L., Evrard, A. E., et al. 2019, Nature Communications, 10, 2504, doi: [10.1038/s41467-019-10471-y](https://doi.org/10.1038/s41467-019-10471-y)
- Fitts, A., Boylan-Kolchin, M., Elbert, O. D., et al. 2017, MNRAS, 471, 3547, doi: [10.1093/mnras/stx1757](https://doi.org/10.1093/mnras/stx1757)
- Folsom, D., Slone, O., Lisanti, M., Jiang, F., & Kaplinghat, M. 2023, arXiv e-prints, arXiv:2311.05676, doi: [10.48550/arXiv.2311.05676](https://doi.org/10.48550/arXiv.2311.05676)
- Foreman-Mackey, D. 2016, The Journal of Open Source Software, 1, 24, doi: [10.21105/joss.00024](https://doi.org/10.21105/joss.00024)
- Gabrielpillai, A., Somerville, R. S., Genel, S., et al. 2022, Monthly Notices of the Royal Astronomical Society, 517, 6091, doi: [10.1093/mnras/stac2297](https://doi.org/10.1093/mnras/stac2297)
- Garrison-Kimmel, S., Wetzel, A., Bullock, J. S., et al. 2017, MNRAS, 471, 1709, doi: [10.1093/mnras/stx1710](https://doi.org/10.1093/mnras/stx1710)
- Gebhardt, M., Anglés-Alcázar, D., Borrow, J., et al. 2023, arXiv e-prints, arXiv:2307.11832, doi: [10.48550/arXiv.2307.11832](https://doi.org/10.48550/arXiv.2307.11832)
- Geha, M., Wechsler, R. H., Mao, Y.-Y., et al. 2017, ApJ, 847, 4, doi: [10.3847/1538-4357/aa8626](https://doi.org/10.3847/1538-4357/aa8626)
- Geha, M., Mao, Y.-Y., Wechsler, R. H., et al. 2024, arXiv e-prints, arXiv:2404.14499, doi: [10.48550/arXiv.2404.14499](https://doi.org/10.48550/arXiv.2404.14499)
- Germain, M., Gregor, K., Murray, I., & Larochelle, H. 2015, in JMLR Workshop and Conference Proceedings, Vol. 37, Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6-11 July 2015, ed. F. R. Bach & D. M. Blei, 881–889. <http://proceedings.mlr.press/v37/germain15.html>
- Gilman, D., Birrer, S., Nierenberg, A., & Oh, M. S. H. 2024, arXiv e-prints, arXiv:2403.03253, doi: [10.48550/arXiv.2403.03253](https://doi.org/10.48550/arXiv.2403.03253)
- Gnedin, N. Y., & Kaurov, A. A. 2014, ApJ, 793, 30, doi: [10.1088/0004-637X/793/1/30](https://doi.org/10.1088/0004-637X/793/1/30)
- Gonzalez-Perez, V., Lacey, C. G., Baugh, C. M., et al. 2014, MNRAS, 439, 264, doi: [10.1093/mnras/stt2410](https://doi.org/10.1093/mnras/stt2410)
- Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., et al. 2014, arXiv e-prints, arXiv:1406.2661, doi: [10.48550/arXiv.1406.2661](https://doi.org/10.48550/arXiv.1406.2661)
- Governato, F., Brook, C., Mayer, L., et al. 2010, Nature, 463, 203, doi: [10.1038/nature08640](https://doi.org/10.1038/nature08640)
- Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179, doi: [10.1093/mnras/stx071](https://doi.org/10.1093/mnras/stx071)
- Green, S. B., van den Bosch, F. C., & Jiang, F. 2022, MNRAS, 509, 2624, doi: [10.1093/mnras/stab3130](https://doi.org/10.1093/mnras/stab3130)
- Guo, Q., & White, S. 2014, MNRAS, 437, 3228, doi: [10.1093/mnras/stt2116](https://doi.org/10.1093/mnras/stt2116)
- Hadzhiyska, B., Liu, S., Somerville, R. S., et al. 2021, MNRAS, 508, 698, doi: [10.1093/mnras/stab2564](https://doi.org/10.1093/mnras/stab2564)
- Hadzhiyska, B., Eisenstein, D., Hernquist, L., et al. 2023, MNRAS, 524, 2507, doi: [10.1093/mnras/stad731](https://doi.org/10.1093/mnras/stad731)
- Hahn, O., & Abel, T. 2011, MNRAS, 415, 2101, doi: [10.1111/j.1365-2966.2011.18820.x](https://doi.org/10.1111/j.1365-2966.2011.18820.x)
- Hamaus, N., Aubert, M., Pisani, A., et al. 2022, A&A, 658, A20, doi: [10.1051/0004-6361/202142073](https://doi.org/10.1051/0004-6361/202142073)
- Harris, C. R., Millman, K. J., Van Der Walt, S. J., et al. 2020, Nature, 585, 357
- Harrold, J. E., Almaini, O., Pearce, F. R., & Yates, R. M. 2024, MNRAS, 532, L61, doi: [10.1093/mnrasl/slae043](https://doi.org/10.1093/mnrasl/slae043)
- He, Y., Wu, J., Wang, W., Jiang, B., & Zhang, Y. 2023, PASJ, 75, 1311, doi: [10.1093/pasj/pssd071](https://doi.org/10.1093/pasj/pssd071)
- Hearin, A. P., Chaves-Montero, J., Becker, M. R., & Alarcon, A. 2021, The Open Journal of Astrophysics, 4, 7, doi: [10.21105/astro.2105.05859](https://doi.org/10.21105/astro.2105.05859)

- Hearin, A. P., Zentner, A. R., van den Bosch, F. C., Campbell, D., & Tollerud, E. 2016, MNRAS, 460, 2552, doi: [10.1093/mnras/stw840](https://doi.org/10.1093/mnras/stw840)
- Hendrycks, D., & Gimpel, K. 2016, arXiv e-prints, arXiv:1606.08415, doi: [10.48550/arXiv.1606.08415](https://doi.org/10.48550/arXiv.1606.08415)
- Hernández-Aguayo, C., Springel, V., Pakmor, R., et al. 2023, MNRAS, 524, 2556, doi: [10.1093/mnras/stad1657](https://doi.org/10.1093/mnras/stad1657)
- Hernquist, L., Katz, N., Weinberg, D. H., & Miralda-Escudé, J. 1996, ApJL, 457, L51, doi: [10.1086/309899](https://doi.org/10.1086/309899)
- Ho, J., Jain, A., & Abbeel, P. 2020, arXiv e-prints, arXiv:2006.11239, doi: [10.48550/arXiv.2006.11239](https://doi.org/10.48550/arXiv.2006.11239)
- Hogan, C. J., & Dalcanton, J. J. 2000, PhRvD, 62, 063511, doi: [10.1103/PhysRevD.62.063511](https://doi.org/10.1103/PhysRevD.62.063511)
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, MNRAS, 445, 581, doi: [10.1093/mnras/stu1738](https://doi.org/10.1093/mnras/stu1738)
- Hopkins, P. F., Wetzel, A., Kereš, D., et al. 2018, MNRAS, 480, 800, doi: [10.1093/mnras/sty1690](https://doi.org/10.1093/mnras/sty1690)
- Huang, C.-W., Lim, J. H., & Courville, A. 2021, arXiv e-prints, arXiv:2106.02808, doi: [10.48550/arXiv.2106.02808](https://doi.org/10.48550/arXiv.2106.02808)
- Hunter, J. D. 2007, Computing In Science & Engineering, 9, 90
- Hutter, A., Dayal, P., Yepes, G., et al. 2021, MNRAS, 503, 3698, doi: [10.1093/mnras/stab602](https://doi.org/10.1093/mnras/stab602)
- Hwang, S. Y., Sabiu, C. G., Park, I., & Hong, S. E. 2023, JCAP, 2023, 075, doi: [10.1088/1475-7516/2023/11/075](https://doi.org/10.1088/1475-7516/2023/11/075)
- Ioffe, S., & Szegedy, C. 2015, arXiv e-prints, arXiv:1502.03167, doi: [10.48550/arXiv.1502.03167](https://doi.org/10.48550/arXiv.1502.03167)
- Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111, doi: [10.3847/1538-4357/ab042c](https://doi.org/10.3847/1538-4357/ab042c)
- Jethwa, P., Erkal, D., & Belokurov, V. 2018, MNRAS, 473, 2060, doi: [10.1093/mnras/stx2330](https://doi.org/10.1093/mnras/stx2330)
- Jiang, F., Dekel, A., Freundlich, J., et al. 2021, MNRAS, 502, 621, doi: [10.1093/mnras/staa4034](https://doi.org/10.1093/mnras/staa4034)
- Jimenez Rezende, D., & Mohamed, S. 2015, arXiv e-prints, arXiv:1505.05770, doi: [10.48550/arXiv.1505.05770](https://doi.org/10.48550/arXiv.1505.05770)
- Kaae Sønderby, C., Raiko, T., Maaløe, L., Kaae Sønderby, S., & Winther, O. 2016, arXiv e-prints, arXiv:1602.02282, doi: [10.48550/arXiv.1602.02282](https://doi.org/10.48550/arXiv.1602.02282)
- Kannan, R., Garaldi, E., Smith, A., et al. 2022, MNRAS, 511, 4005, doi: [10.1093/mnras/stab3710](https://doi.org/10.1093/mnras/stab3710)
- Kaviraj, S., Laigle, C., Kimm, T., et al. 2017, MNRAS, 467, 4739, doi: [10.1093/mnras/stx126](https://doi.org/10.1093/mnras/stx126)
- Kayo, I., & Oguri, M. 2012, MNRAS, 424, 1363, doi: [10.1111/j.1365-2966.2012.21321.x](https://doi.org/10.1111/j.1365-2966.2012.21321.x)
- Kern, N. 2024, in American Astronomical Society Meeting Abstracts, Vol. 243, American Astronomical Society Meeting Abstracts, 354.01
- Kim, S. Y., Peter, A. H. G., & Hargis, J. R. 2018, PhRvL, 121, 211302, doi: [10.1103/PhysRevLett.121.211302](https://doi.org/10.1103/PhysRevLett.121.211302)
- Kingma, D. P., & Ba, J. 2014, arXiv preprint arXiv:1412.6980
- Kingma, D. P., Salimans, T., Jozefowicz, R., et al. 2016, arXiv e-prints, arXiv:1606.04934, doi: [10.48550/arXiv.1606.04934](https://doi.org/10.48550/arXiv.1606.04934)
- Kingma, D. P., Salimans, T., Poole, B., & Ho, J. 2021, arXiv e-prints, arXiv:2107.00630, doi: [10.48550/arXiv.2107.00630](https://doi.org/10.48550/arXiv.2107.00630)
- Kingma, D. P., & Welling, M. 2013, arXiv e-prints, arXiv:1312.6114, doi: [10.48550/arXiv.1312.6114](https://doi.org/10.48550/arXiv.1312.6114)
- Kitzbichler, M. G., & White, S. D. M. 2008, MNRAS, 391, 1489, doi: [10.1111/j.1365-2966.2008.13873.x](https://doi.org/10.1111/j.1365-2966.2008.13873.x)
- Kluyver, T., et al. 2016, in ELPUB
- Knebe, A., Knollmann, S. R., Muldrew, S. I., et al. 2011, MNRAS, 415, 2293, doi: [10.1111/j.1365-2966.2011.18858.x](https://doi.org/10.1111/j.1365-2966.2011.18858.x)
- Kokron, N., DeRose, J., Chen, S.-F., White, M., & Wechsler, R. H. 2021, MNRAS, 505, 1422, doi: [10.1093/mnras/stab1358](https://doi.org/10.1093/mnras/stab1358)
- Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, ApJ, 609, 35, doi: [10.1086/420959](https://doi.org/10.1086/420959)
- Kwon, K. J., & Hahn, C. 2024, arXiv e-prints, arXiv:2401.12318, doi: [10.48550/arXiv.2401.12318](https://doi.org/10.48550/arXiv.2401.12318)
- Lanusse, F., Parker, L., Golkar, S., et al. 2023, arXiv e-prints, arXiv:2310.03024, doi: [10.48550/arXiv.2310.03024](https://doi.org/10.48550/arXiv.2310.03024)
- Lee, M. E., Genel, S., Wandelt, B. D., et al. 2024, arXiv e-prints, arXiv:2403.10609, doi: [10.48550/arXiv.2403.10609](https://doi.org/10.48550/arXiv.2403.10609)
- Leung, H. W., & Bovy, J. 2024, MNRAS, 527, 1494, doi: [10.1093/mnras/stad3015](https://doi.org/10.1093/mnras/stad3015)
- Li, Y., Modi, C., Jamieson, D., et al. 2024, ApJS, 270, 36, doi: [10.3847/1538-4365/ad0ce7](https://doi.org/10.3847/1538-4365/ad0ce7)
- Loshchilov, I., & Hutter, F. 2016, arXiv e-prints, arXiv:1608.03983, doi: [10.48550/arXiv.1608.03983](https://doi.org/10.48550/arXiv.1608.03983)
- Loshchilov, I., & Hutter, F. 2019, in International Conference on Learning Representations, <https://openreview.net/forum?id=Bkg6RiCqY7>
- Lovell, C. C., Hassan, S., Villaescusa-Navarro, F., et al. 2023a, in Machine Learning for Astrophysics, 21, doi: [10.48550/arXiv.2307.06967](https://doi.org/10.48550/arXiv.2307.06967)
- Lovell, C. C., Hassan, S., Villaescusa-Navarro, F., et al. 2023b, in Machine Learning for Astrophysics, 21, doi: [10.48550/arXiv.2307.06967](https://doi.org/10.48550/arXiv.2307.06967)
- Lovell, M. R., Cautun, M., Frenk, C. S., Hellwing, W. A., & Newton, O. 2021, MNRAS, 507, 4826, doi: [10.1093/mnras/stab2452](https://doi.org/10.1093/mnras/stab2452)
- Lovell, M. R., Frenk, C. S., Eke, V. R., et al. 2014, MNRAS, 439, 300, doi: [10.1093/mnras/stt2431](https://doi.org/10.1093/mnras/stt2431)
- Lovell, M. R., Hellwing, W., Ludlow, A., et al. 2020, MNRAS, 498, 702, doi: [10.1093/mnras/staa2525](https://doi.org/10.1093/mnras/staa2525)
- Lovell, M. R., Eke, V., Frenk, C. S., et al. 2012, MNRAS, 420, 2318, doi: [10.1111/j.1365-2966.2011.20200.x](https://doi.org/10.1111/j.1365-2966.2011.20200.x)
- Lovell, M. R., Bose, S., Boyarsky, A., et al. 2016, MNRAS, 461, 60, doi: [10.1093/mnras/stw1317](https://doi.org/10.1093/mnras/stw1317)
- Luo, C. 2022, arXiv e-prints, arXiv:2208.11970, doi: [10.48550/arXiv.2208.11970](https://doi.org/10.48550/arXiv.2208.11970)

- Macciò, A. V., & Fontanot, F. 2010, MNRAS, 404, L16, doi: [10.1111/j.1745-3933.2010.00825.x](https://doi.org/10.1111/j.1745-3933.2010.00825.x)
- Mansfield, P., Darragh-Ford, E., Wang, Y., et al. 2024, ApJ, 970, 178, doi: [10.3847/1538-4357/ad4e33](https://doi.org/10.3847/1538-4357/ad4e33)
- Mao, Y.-Y., Geha, M., Wechsler, R. H., et al. 2021, ApJ, 907, 85, doi: [10.3847/1538-4357/abce58](https://doi.org/10.3847/1538-4357/abce58)
- . 2024, arXiv e-prints, arXiv:2404.14498, doi: [10.48550/arXiv.2404.14498](https://doi.org/10.48550/arXiv.2404.14498)
- Marinacci, F., Sales, L. V., Vogelsberger, M., Torrey, P., & Springel, V. 2019, MNRAS, 489, 4233, doi: [10.1093/mnras/stz2391](https://doi.org/10.1093/mnras/stz2391)
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, MNRAS, 480, 5113, doi: [10.1093/mnras/sty2206](https://doi.org/10.1093/mnras/sty2206)
- Mehran Kazemi, S., Goel, R., Eghbali, S., et al. 2019, arXiv e-prints, arXiv:1907.05321, doi: [10.48550/arXiv.1907.05321](https://doi.org/10.48550/arXiv.1907.05321)
- Mildenhall, B., Srinivasan, P. P., Tancik, M., et al. 2020, arXiv e-prints, arXiv:2003.08934, doi: [10.48550/arXiv.2003.08934](https://doi.org/10.48550/arXiv.2003.08934)
- Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution
- Modi, C., Lanusse, F., & Seljak, U. 2021a, Astronomy and Computing, 37, 100505, doi: [10.1016/j.ascom.2021.100505](https://doi.org/10.1016/j.ascom.2021.100505)
- Modi, C., Lanusse, F., Seljak, U., Spergel, D. N., & Perreault-Levasseur, L. 2021b, arXiv e-prints, arXiv:2104.12864, doi: [10.48550/arXiv.2104.12864](https://doi.org/10.48550/arXiv.2104.12864)
- Moster, B. P., Naab, T., & White, S. D. M. 2018, MNRAS, 477, 1822, doi: [10.1093/mnras/sty655](https://doi.org/10.1093/mnras/sty655)
- Moster, B. P., Somerville, R. S., Maulbetsch, C., et al. 2010, ApJ, 710, 903, doi: [10.1088/0004-637X/710/2/903](https://doi.org/10.1088/0004-637X/710/2/903)
- Nadler, E. O., Banerjee, A., Adhikari, S., Mao, Y.-Y., & Wechsler, R. H. 2021a, ApJL, 920, L11, doi: [10.3847/2041-8213/ac29c1](https://doi.org/10.3847/2041-8213/ac29c1)
- Nadler, E. O., Birrer, S., Gilman, D., et al. 2021b, ApJ, 917, 7, doi: [10.3847/1538-4357/abf9a3](https://doi.org/10.3847/1538-4357/abf9a3)
- Nagai, D., & Kravtsov, A. V. 2005, ApJ, 618, 557, doi: [10.1086/426016](https://doi.org/10.1086/426016)
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206, doi: [10.1093/mnras/sty618](https://doi.org/10.1093/mnras/sty618)
- Navarro, J. F., Eke, V. R., & Frenk, C. S. 1996a, MNRAS, 283, L72, doi: [10.1093/mnras/283.3.L72](https://doi.org/10.1093/mnras/283.3.L72)
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1996b, ApJ, 462, 563, doi: [10.1086/177173](https://doi.org/10.1086/177173)
- . 1997, ApJ, 490, 493, doi: [10.1086/304888](https://doi.org/10.1086/304888)
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, MNRAS, 475, 624, doi: [10.1093/mnras/stx3040](https://doi.org/10.1093/mnras/stx3040)
- . 2019, MNRAS, 490, 3234, doi: [10.1093/mnras/stz2306](https://doi.org/10.1093/mnras/stz2306)
- Neyman, J., & Scott, E. L. 1952, ApJ, 116, 144, doi: [10.1086/145599](https://doi.org/10.1086/145599)
- Nguyen, T., Mishra-Sharma, S., Williams, R., & Necib, L. 2023a, PhRvD, 107, 043015, doi: [10.1103/PhysRevD.107.043015](https://doi.org/10.1103/PhysRevD.107.043015)
- Nguyen, T., Modi, C., Yung, L. Y. A., & Somerville, R. S. 2023b, arXiv e-prints, arXiv:2308.05145, doi: [10.48550/arXiv.2308.05145](https://doi.org/10.48550/arXiv.2308.05145)
- Nusser, A., & Sheth, R. K. 1999, MNRAS, 303, 685, doi: [10.1046/j.1365-8711.1999.02197.x](https://doi.org/10.1046/j.1365-8711.1999.02197.x)
- Ocvirk, P., Gillet, N., Shapiro, P. R., et al. 2016, MNRAS, 463, 1462, doi: [10.1093/mnras/stw2036](https://doi.org/10.1093/mnras/stw2036)
- Onions, J., Knebe, A., Pearce, F. R., et al. 2012, MNRAS, 423, 1200, doi: [10.1111/j.1365-2966.2012.20947.x](https://doi.org/10.1111/j.1365-2966.2012.20947.x)
- pandas development team, T. 2020, pandas-dev/pandas: Pandas, latest, Zenodo, doi: [10.5281/zenodo.3509134](https://doi.org/10.5281/zenodo.3509134)
- Papamakarios, G., Nalisnick, E., Jimenez Rezende, D., Mohamed, S., & Lakshminarayanan, B. 2019, arXiv e-prints, arXiv:1912.02762, doi: [10.48550/arXiv.1912.02762](https://doi.org/10.48550/arXiv.1912.02762)
- Papamakarios, G., Pavlakou, T., & Murray, I. 2017, in Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17 (Red Hook, NY, USA: Curran Associates Inc.), 2335–2344. <https://papers.nips.cc/paper/2017/hash/6c1da886822c67822bcf3679d04369fa-Abstract.html>
- Paszke, A., et al. 2019, in Advances in Neural Information Processing Systems 32, ed. H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Curran Associates, Inc.), 8024–8035
- Peacock, J. A., & Smith, R. E. 2000, MNRAS, 318, 1144, doi: [10.1046/j.1365-8711.2000.03779.x](https://doi.org/10.1046/j.1365-8711.2000.03779.x)
- Peirani, S., Dubois, Y., Volonteri, M., et al. 2017, MNRAS, 472, 2153, doi: [10.1093/mnras/stx2099](https://doi.org/10.1093/mnras/stx2099)
- Peirani, S., Sonnenfeld, A., Gavazzi, R., et al. 2019, MNRAS, 483, 4615, doi: [10.1093/mnras/sty3475](https://doi.org/10.1093/mnras/sty3475)
- Perez, F., & Granger, B. E. 2007, Computing in Science and Engineering, 9, 21, doi: [10.1109/MCSE.2007.53](https://doi.org/10.1109/MCSE.2007.53)
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, MNRAS, 475, 648, doi: [10.1093/mnras/stx3112](https://doi.org/10.1093/mnras/stx3112)
- Pillepich, A., Springel, V., Nelson, D., et al. 2018b, MNRAS, 473, 4077, doi: [10.1093/mnras/stx2656](https://doi.org/10.1093/mnras/stx2656)
- Pillepich, A., Nelson, D., Springel, V., et al. 2019, MNRAS, 490, 3196, doi: [10.1093/mnras/stz2338](https://doi.org/10.1093/mnras/stz2338)
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13, doi: [10.1051/0004-6361/201525830](https://doi.org/10.1051/0004-6361/201525830)
- Pujol, A., Skibba, R. A., Gaztañaga, E., et al. 2017, MNRAS, 469, 749, doi: [10.1093/mnras/stx913](https://doi.org/10.1093/mnras/stx913)
- Radford, A., Wu, J., Child, R., et al. 2019, OpenAI blog, 1, 9
- Rahaman, N., Baratin, A., Arpit, D., et al. 2018, arXiv e-prints, arXiv:1806.08734, doi: [10.48550/arXiv.1806.08734](https://doi.org/10.48550/arXiv.1806.08734)
- Read, J. I., Iorio, G., Agertz, O., & Fraternali, F. 2017, MNRAS, 467, 2019, doi: [10.1093/mnras/stx147](https://doi.org/10.1093/mnras/stx147)
- Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2013, ApJ, 771, 30, doi: [10.1088/0004-637X/771/1/30](https://doi.org/10.1088/0004-637X/771/1/30)
- Reyes-Peraza, G., Avila, S., Gonzalez-Perez, V., et al. 2024, MNRAS, 529, 3877, doi: [10.1093/mnras/stae623](https://doi.org/10.1093/mnras/stae623)
- Rodríguez-Puebla, A., Primack, J. R., Avila-Reese, V., & Faber, S. M. 2017, MNRAS, 470, 651, doi: [10.1093/mnras/stx1172](https://doi.org/10.1093/mnras/stx1172)

- Rosdahl, J., Katz, H., Blaizot, J., et al. 2018, MNRAS, 479, 994, doi: [10.1093/mnras/sty1655](https://doi.org/10.1093/mnras/sty1655)
- Rose, J. C., Torrey, P., Villaescusa-Navarro, F., et al. 2024a, arXiv e-prints, arXiv:2405.00766. <https://arxiv.org/abs/2405.00766>
- . 2024b, MNRAS, 527, 739, doi: [10.1093/mnras/stad3260](https://doi.org/10.1093/mnras/stad3260)
- Rozet, F., et al. 2022, Zukō: Normalizing flows in PyTorch, doi: [10.5281/zenodo.7625672](https://doi.org/10.5281/zenodo.7625672)
- Sales, L. V., Wetzel, A., & Fattah, A. 2022, Nature Astronomy, 6, 897, doi: [10.1038/s41550-022-01689-w](https://doi.org/10.1038/s41550-022-01689-w)
- Salimans, T., Goodfellow, I., Zaremba, W., et al. 2016, arXiv e-prints, arXiv:1606.03498, doi: [10.48550/arXiv.1606.03498](https://doi.org/10.48550/arXiv.1606.03498)
- Samuel, J., Wetzel, A., Tollerud, E., et al. 2020, MNRAS, 491, 1471, doi: [10.1093/mnras/stz3054](https://doi.org/10.1093/mnras/stz3054)
- Sawala, T., Frenk, C. S., Fattah, A., et al. 2016, MNRAS, 457, 1931, doi: [10.1093/mnras/stw145](https://doi.org/10.1093/mnras/stw145)
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521, doi: [10.1093/mnras/stu2058](https://doi.org/10.1093/mnras/stu2058)
- Schmittfull, M., Simonović, M., Ivanov, M. M., Philcox, O. H. E., & Zaldarriaga, M. 2021, JCAP, 2021, 059, doi: [10.1088/1475-7516/2021/05/059](https://doi.org/10.1088/1475-7516/2021/05/059)
- Shi, J., Wang, H., Mo, H. J., et al. 2018, ApJ, 857, 127, doi: [10.3847/1538-4357/aab775](https://doi.org/10.3847/1538-4357/aab775)
- Simpson, C. M., Grand, R. J. J., Gómez, F. A., et al. 2018, MNRAS, 478, 548, doi: [10.1093/mnras/sty774](https://doi.org/10.1093/mnras/sty774)
- Smith, A., Kannan, R., Garaldi, E., et al. 2022, MNRAS, 512, 3243, doi: [10.1093/mnras/stac713](https://doi.org/10.1093/mnras/stac713)
- Smith, A., Grove, C., Cole, S., et al. 2024, MNRAS, 532, 903, doi: [10.1093/mnras/stae1503](https://doi.org/10.1093/mnras/stae1503)
- Sobol', I. M. 1967, Zhurnal Vychislitel'noi Matematiki i Matematicheskoi Fiziki, 7, 784
- Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., & Ganguli, S. 2015, arXiv e-prints, arXiv:1503.03585, doi: [10.48550/arXiv.1503.03585](https://doi.org/10.48550/arXiv.1503.03585)
- Somerville, R. S., & Davé, R. 2015, ARA&A, 53, 51, doi: [10.1146/annurev-astro-082812-140951](https://doi.org/10.1146/annurev-astro-082812-140951)
- Somerville, R. S., Popping, G., & Trager, S. C. 2015, Monthly Notices of the Royal Astronomical Society, 453, 4337, doi: [10.1093/mnras/stv1877](https://doi.org/10.1093/mnras/stv1877)
- Sommer-Larsen, J., & Dolgov, A. 2001, ApJ, 551, 608, doi: [10.1086/320211](https://doi.org/10.1086/320211)
- Song, Y., Durkan, C., Murray, I., & Ermon, S. 2021, arXiv e-prints, arXiv:2101.09258, doi: [10.48550/arXiv.2101.09258](https://doi.org/10.48550/arXiv.2101.09258)
- Song, Y., Sohl-Dickstein, J., Kingma, D. P., et al. 2020, arXiv e-prints, arXiv:2011.13456, doi: [10.48550/arXiv.2011.13456](https://doi.org/10.48550/arXiv.2011.13456)
- Spekkens, K., Giovanelli, R., & Haynes, M. P. 2005, AJ, 129, 2119, doi: [10.1086/429592](https://doi.org/10.1086/429592)
- Springel, V. 2010, MNRAS, 401, 791, doi: [10.1111/j.1365-2966.2009.15715.x](https://doi.org/10.1111/j.1365-2966.2009.15715.x)
- Springel, V., Frenk, C. S., & White, S. D. M. 2006, Nature, 440, 1137, doi: [10.1038/nature04805](https://doi.org/10.1038/nature04805)
- Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, MNRAS, 328, 726, doi: [10.1046/j.1365-8711.2001.04912.x](https://doi.org/10.1046/j.1365-8711.2001.04912.x)
- Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629, doi: [10.1038/nature03597](https://doi.org/10.1038/nature03597)
- Springel, V., Wang, J., Vogelsberger, M., et al. 2008, MNRAS, 391, 1685, doi: [10.1111/j.1365-2966.2008.14066.x](https://doi.org/10.1111/j.1365-2966.2008.14066.x)
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676, doi: [10.1093/mnras/stx3304](https://doi.org/10.1093/mnras/stx3304)
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. 2014, Journal of Machine Learning Research, 15, 1929. <http://jmlr.org/papers/v15/srivastava14a.html>
- Tancik, M., Srinivasan, P. P., Mildenhall, B., et al. 2020, arXiv e-prints, arXiv:2006.10739, doi: [10.48550/arXiv.2006.10739](https://doi.org/10.48550/arXiv.2006.10739)
- Teyssier, R., Moore, B., Martizzi, D., Dubois, Y., & Mayer, L. 2011, MNRAS, 414, 195, doi: [10.1111/j.1365-2966.2011.18399.x](https://doi.org/10.1111/j.1365-2966.2011.18399.x)
- The Dark Energy Survey Collaboration. 2005, arXiv e-prints, astro, doi: [10.48550/arXiv.astro-ph/0510346](https://doi.org/10.48550/arXiv.astro-ph/0510346)
- Torrey, P., Vogelsberger, M., Genel, S., et al. 2014, MNRAS, 438, 1985, doi: [10.1093/mnras/stt2295](https://doi.org/10.1093/mnras/stt2295)
- Trujillo-Gomez, S., Klypin, A., Primack, J., & Romanowsky, A. J. 2011, ApJ, 742, 16, doi: [10.1088/0004-637X/742/1/16](https://doi.org/10.1088/0004-637X/742/1/16)
- Truong, N., Rasia, E., Mazzotta, P., et al. 2018, MNRAS, 474, 4089, doi: [10.1093/mnras/stx2927](https://doi.org/10.1093/mnras/stx2927)
- Vahdat, A., Kreis, K., & Kautz, J. 2021, arXiv e-prints, arXiv:2106.05931, doi: [10.48550/arXiv.2106.05931](https://doi.org/10.48550/arXiv.2106.05931)
- Vale, A., & Ostriker, J. P. 2004, MNRAS, 353, 189, doi: [10.1111/j.1365-2966.2004.08059.x](https://doi.org/10.1111/j.1365-2966.2004.08059.x)
- . 2006, MNRAS, 371, 1173, doi: [10.1111/j.1365-2966.2006.10605.x](https://doi.org/10.1111/j.1365-2966.2006.10605.x)
- van den Bosch, F. C. 2002, MNRAS, 331, 98, doi: [10.1046/j.1365-8711.2002.05171.x](https://doi.org/10.1046/j.1365-8711.2002.05171.x)
- Vaswani, A., Shazeer, N., Parmar, N., et al. 2023, Attention Is All You Need. <https://arxiv.org/abs/1706.03762>
- Villaescusa-Navarro, F., Anglés-Alcázar, D., Genel, S., et al. 2021, ApJ, 915, 71, doi: [10.3847/1538-4357/abf7ba](https://doi.org/10.3847/1538-4357/abf7ba)
- Villanueva-Domingo, P., & Villaescusa-Navarro, F. 2022, ApJ, 937, 115, doi: [10.3847/1538-4357/ac8930](https://doi.org/10.3847/1538-4357/ac8930)
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2)
- Vogelsberger, M., Marinacci, F., Torrey, P., & Puchwein, E. 2020, Nature Reviews Physics, 2, 42, doi: [10.1038/s42254-019-0127-2](https://doi.org/10.1038/s42254-019-0127-2)
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014a, Nature, 509, 177, doi: [10.1038/nature13316](https://doi.org/10.1038/nature13316)
- . 2014b, MNRAS, 444, 1518, doi: [10.1093/mnras/stu1536](https://doi.org/10.1093/mnras/stu1536)
- Wagner-Carena, S., Aalbers, J., Birrer, S., et al. 2023, ApJ, 942, 75, doi: [10.3847/1538-4357/aca525](https://doi.org/10.3847/1538-4357/aca525)
- Wang, L., Dutton, A. A., Stinson, G. S., et al. 2015, MNRAS, 454, 83, doi: [10.1093/mnras/stv1937](https://doi.org/10.1093/mnras/stv1937)

- Wang, Y., Nadler, E. O., Mao, Y.-Y., et al. 2024, arXiv e-prints, arXiv:2404.14500, doi: [10.48550/arXiv.2404.14500](https://doi.org/10.48550/arXiv.2404.14500)
- Wechsler, R. H., & Tinker, J. L. 2018, ARA&A, 56, 435, doi: [10.1146/annurev-astro-081817-051756](https://doi.org/10.1146/annurev-astro-081817-051756)
- Wechsler, R. H., Zentner, A. R., Bullock, J. S., Kravtsov, A. V., & Allgood, B. 2006, ApJ, 652, 71, doi: [10.1086/507120](https://doi.org/10.1086/507120)
- Weinberger, R., Springel, V., & Pakmor, R. 2020, ApJS, 248, 32, doi: [10.3847/1538-4365/ab908c](https://doi.org/10.3847/1538-4365/ab908c)
- Weinberger, R., Springel, V., Hernquist, L., et al. 2017, MNRAS, 465, 3291, doi: [10.1093/mnras/stw2944](https://doi.org/10.1093/mnras/stw2944)
- Wetzel, A. R., Hopkins, P. F., Kim, J.-h., et al. 2016, ApJL, 827, L23, doi: [10.3847/2041-8205/827/2/L23](https://doi.org/10.3847/2041-8205/827/2/L23)
- Wu, J. F., Peek, J. E. G., Tollerud, E. J., et al. 2022a, ApJ, 927, 121, doi: [10.3847/1538-4357/ac4eea](https://doi.org/10.3847/1538-4357/ac4eea)
- Wu, X., Muñoz, J. B., & Eisenstein, D. 2022b, JCAP, 2022, 002, doi: [10.1088/1475-7516/2022/02/002](https://doi.org/10.1088/1475-7516/2022/02/002)
- Xu, D., Ruan, C., Kumar, S., Korpeoglu, E., & Achan, K. 2019, arXiv e-prints, arXiv:1911.12864, doi: [10.48550/arXiv.1911.12864](https://doi.org/10.48550/arXiv.1911.12864)
- Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, MNRAS, 339, 1057, doi: [10.1046/j.1365-8711.2003.06254.x](https://doi.org/10.1046/j.1365-8711.2003.06254.x)
- York, D. G., Adelman, J., Anderson, John E., J., et al. 2000, AJ, 120, 1579, doi: [10.1086/301513](https://doi.org/10.1086/301513)
- Yung, L. Y. A., Somerville, R., Finkelstein, S., Popping, G., & Davé, R. 2019, MNRAS, 483, 2983, doi: [10.1093/mnras/sty3241](https://doi.org/10.1093/mnras/sty3241)
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, ApJ, 736, 59, doi: [10.1088/0004-637X/736/1/59](https://doi.org/10.1088/0004-637X/736/1/59)
- Zhang, Z., Farahi, A., Nagai, D., et al. 2024, MNRAS, 530, 3127, doi: [10.1093/mnras/stae999](https://doi.org/10.1093/mnras/stae999)
- Zhong, E. D., Bepler, T., Davis, J. H., & Berger, B. 2019, arXiv e-prints, arXiv:1909.05215, doi: [10.48550/arXiv.1909.05215](https://doi.org/10.48550/arXiv.1909.05215)

## APPENDIX

### A. NORMALIZING FLOWS

Normalizing flows model the distribution  $p(\mathbf{y})$  of the data  $\mathbf{y}$  by transforming it from a base distribution  $p(\mathbf{u})$  via an invertible transformation. Let  $f$  be an invertible transformation that maps a random variable  $\mathbf{u}$  to  $\mathbf{y}$ , i.e.  $\mathbf{y} = f(\mathbf{u})$ . Using the change of variables formula, we can write the  $p(\mathbf{y})$  as:

$$p(\mathbf{y}) = p(f^{-1}(\mathbf{y})) \left| \det \left( \frac{\partial f^{-1}}{\partial \mathbf{y}} \right) \right| \quad (\text{A1})$$

where the last term is the absolute value of the determinant of the Jacobian matrix of  $f^{-1}$  at  $\mathbf{y}$  and accounts for the change in volume under the transformation. The base distribution  $p(\mathbf{u})$  is often set to a simple, well-understood distribution such as a standard Gaussian. The transformation  $f$  is a neural network with trainable parameters but with a few restrictions. First, we emphasize the invertibility of the transformation, which allows for both sampling (forward transformation from  $p(\mathbf{u})$  to  $p(\mathbf{y})$ ) and density estimation (inverse transformation from  $p(\mathbf{y})$  to  $p(\mathbf{u})$ ). In addition, the Jacobian determinant must be tractable to ensure the model can be efficiently trained and applied to high-dimensional data. These requirements restrict the type of transformations available, with common flows being the masked autoregressive flows (Germain et al. 2015; Papamakarios et al. 2017) and neural spline flows (Durkan et al. 2019).

In practice, the normalizing flows are constructed from a sequence of  $T$  discrete transformations  $f = f_1 \circ f_2 \circ \dots \circ f_T$ , where each transformation  $f_i$  is also invertible and has tractable Jacobian determinant. The data likelihood  $p(\mathbf{y})$  can be written as:

$$p_T(\mathbf{y}) = p_0(\mathbf{u}_0) \prod_{i=1}^T \left| \det \left( \frac{\partial f_i^{-1}}{\partial \mathbf{u}_{i-1}} \right) \right| \quad (\text{A2})$$

where  $\mathbf{y} = \mathbf{u}_T$ . The optimization objective is thus simply the negative log-likelihood  $\log p_T(\mathbf{y})$ , i.e.

$$\mathcal{L}_{\text{flows}}(\mathbf{y}) = -\log p_0(f^{-1}(\mathbf{y})) - \sum_{i=1}^T \log \left| \det \left( \frac{\partial f_i^{-1}}{\partial \mathbf{u}_{i-1}} \right) \right|. \quad (\text{A3})$$

Each transformation  $f_i$  is parameterized by a neural network with parameters  $\phi_i$ . During training, the full set of parameters  $\phi = \{\phi_1, \phi_2, \dots, \phi_T\}$  is optimized simultaneously. In the case of a conditional generative model, with conditioning features  $\theta$ , we simply include  $\theta$  into each transformation, i.e.  $f_i(\mathbf{y}) \rightarrow f_i(\mathbf{y}, \theta)$ .

### B. DIFFUSION MODEL

In this Appendix, we present a summary of the inner workings of variational diffusion models (VDM), following the

discussion in Kingma et al. (2021). For more detailed mathematical formalism and derivation, we also referred readers to a review in Luo (2022).

#### B.1. Forward diffusion

In the context of a Gaussian diffusion model, forward diffusion refers to the process of gradually corrupting an original data  $\mathbf{y}$  by incrementally adding Gaussian noise to it over  $T$  discrete timesteps. Let the noisy version of  $\mathbf{y}$  at some time step  $t \in [0, 1]$  be the *latent variables*  $\mathbf{z}_t$ , we may define the *variance-preserving* mapping from  $\mathbf{y}$  to  $\mathbf{z}_t$  (see Sohl-Dickstein et al. (2015); Ho et al. (2020)) as:

$$q(\mathbf{z}_t | \mathbf{y}) = \mathcal{N} \left( \mathbf{z}_t | \alpha_t \mathbf{y}, \sigma_t^2 \mathbf{I} \right), \quad (\text{B4})$$

where  $\sigma_t^2$  is a strictly positive scalar-valued function of  $t$  that represents the variance of the noise added at each step, and  $\alpha_t = \sqrt{1 - \sigma_t^2}$ . Note that the distribution  $q(\mathbf{z}_t | \mathbf{y})$  is conditioned on the data  $\mathbf{y}$ .

One can define a *signal-to-noise ratio*  $\text{SNR}(t) \equiv \alpha_t^2 / \sigma_t^2$ . As  $t$  progresses from 0 (least noisy) to 1 (most noisy), we require that  $\sigma_t^2$  increases, indicating that more noise is being added, which in turn reduces the signal-to-noise ratio  $\text{SNR}(t)$ . With a sufficiently small  $\text{SNR}(1)$ , the variance-preserving diffusion transformation ensures that the maximally noise-corrupted data at  $t = 1$  follow a standard Gaussian distribution, i.e.

$$q(\mathbf{z}_1 | \mathbf{y}) \approx \mathcal{N}(\mathbf{z}_1 | 0, \mathbf{I}). \quad (\text{B5})$$

The condition offers key benefits, including analytical tractability, consistency in variance, and ease of sampling, which collectively enhance model stability and efficiency.

As mentioned in Section 3.3, the noise schedule, which controls how  $\sigma_t^2$  depends on  $t$ , is the most critical part for the performance of diffusion models (Chen 2023). In our framework, we employ a linear noise schedule with the following functional form:

$$\sigma_t^2 = \text{Sigmoid}(\gamma_\eta(t)), \quad \text{where } \gamma_\eta(t) = \gamma_{\max} - (\gamma_{\max} - \gamma_{\min})t, \quad (\text{B6})$$

where  $\eta = \{\gamma_{\max}, \gamma_{\min}\}$  is the trainable parameter. It is straightforward to see that  $\text{SNR}(t) = \exp(-\gamma(t))$ , so  $\gamma_{\max}$  and  $\gamma_{\min}$  also controls the minimum and maximum signal-to-noise ratio of the forward diffusion process. During our training, we initialize  $\{\gamma_{\min}, \gamma_{\max}\}$  to be  $\{-8, 14\}$ , though there is no restriction of what values they can take. Other possibilities for the noise schedule include, for example, a linear or cosine noise schedule with fixed hyperparameters, or using a monotonic neural network (in which case  $\eta$  will be the parameters of the network).

## B.2. Reverse diffusion

To generate new data, we would like to revert the forward diffusion process in Equation B4 by gradually denoising the corrupted data at  $t = 1$  over a series of discrete timesteps  $T$ . This involves training a generative model that can sample a sequence of latent variables  $\mathbf{z}_t$  with time moving backward from  $t = 1$  to  $t = 0$ . Define the time steps to be  $s(i) = (i-1)/T$  and  $t(i) = i/T$  where  $i \in [0, T]$ , our generative model for data  $\mathbf{y}$  can be written as:

$$p(\mathbf{y}) = p(\mathbf{z}_1)p(\mathbf{y}|\mathbf{z}_0)\prod_{i=1}^T p(\mathbf{z}_{s(i)}|\mathbf{z}_{t(i)}). \quad (\text{B7})$$

Familiar readers may recognize that this formalism is similar to a Markovian hierarchical variational autoencoder (VAE) (Kingma et al. 2016; Kaae Sønderby et al. 2016), a generalization of VAEs that incorporates multiple levels of latent hierarchies and a Markov chain generative process. Indeed, VDMs can be thought of as Markovian hierarchical VAEs where the latent dimension is the same as the data dimension and the latent encoding process is pre-defined as a linear Gaussian model (see Luo (2022)). This interpretation of VDMs is especially useful for understanding its optimization objective.

As stated in Appendix B.1, with sufficiently large  $T$  and small SNR(1), we expect  $q(\mathbf{z}_1)$  to be a standard Gaussian distribution. We thus also model the first term  $p(\mathbf{z}_1)$  as a standard Gaussian, i.e.

$$p(\mathbf{z}_1) \approx \mathcal{N}(\mathbf{z}_1|0, \mathbf{I}). \quad (\text{B8})$$

Similarly, as in the VAE analog, the second term represents the reconstructed data likelihood, which we will simply model as:

$$p(\mathbf{y}|\mathbf{z}_0) \approx \mathcal{N}(\mathbf{y}|\mathbf{z}_1, \sigma^2 \mathbf{I}). \quad (\text{B9})$$

Note that  $\sigma$  is a hyperparameter of the data likelihood and *not* related to the noise schedule. It determines how accurately the data is to be reconstructed during the training process, as well as the relative scaling between each term of the optimization objective (see Appendix B.3). In our framework, we set  $\sigma = 0.001$ . Finally, we choose the last term to be:

$$p(\mathbf{z}_s|\mathbf{y}) = q(\mathbf{z}_s|\mathbf{z}_t, \mathbf{y}), \quad (\text{B10})$$

where we have rewritten the ground truth  $q(\mathbf{z}_s|\mathbf{z}_t)$  to be conditioned on the target data  $q(\mathbf{z}_s|\mathbf{z}_t, \mathbf{y})$ .

## B.3. Variational optimization objective

To train the VDM, we minimize the negative Evidence Lower Bound (ELBO, Kingma & Welling (2013)),

$$-\log p(\mathbf{y}) \leq -\text{ELBO}(\mathbf{y}) = -\mathbb{E}_{q(\mathbf{z}_1|\mathbf{y})} [D_{\text{KL}}(q(\mathbf{z}_1|\mathbf{y})||p(\mathbf{z}_1))] \quad (\text{B11})$$

$$+ \mathbb{E}_{q(\mathbf{z}_0|\mathbf{y})} [\log p(\mathbf{y}|\mathbf{z}_0)] \quad (\text{B12})$$

$$+ \mathcal{L}_{\text{diff}}(\mathbf{y}), \quad (\text{B13})$$

where  $D_{\text{KL}}$  is the Kullback–Leibler (KL) divergence. The first two terms are referred to as the *prior matching loss* and the *reconstruction loss*, paralleling their counterparts in the VAE analog. The prior matching term minimizes the discrepancy between the final latent distribution and the Gaussian prior, while the reconstruction term ensures the accuracy between the data  $\mathbf{y}$  and its reconstruction from the latent variable  $\mathbf{z}_0$ . It is worth noting that in a common setup where the noise schedule is linear, these two terms are ignored since they do not depend on the parameters of the noise prediction model (see Ho et al. (2020) for example). However, as our noise schedule (Equation B6) consists of trainable parameters  $\eta$ , we do not ignore these terms.

Unlike in the VAE analog, the VDM loss includes an additional term, known as the *forward-reverse consistency loss*, which can be written as:

$$\mathcal{L}_{\text{diff}}(\mathbf{y}) = \sum_{i=1}^T \mathbb{E}_{q(\mathbf{z}_i|\mathbf{y})} D_{\text{KL}} [q(\mathbf{z}_{s(i)}|\mathbf{z}_{t(i)}, \mathbf{y})||p(\mathbf{z}_{s(i)}|\mathbf{z}_{t(i)})] \quad (\text{B14})$$

in case of finite  $T$ . This term is minimized when the forward diffusion process  $q(\mathbf{z}_s|\mathbf{z}_t, \mathbf{y})$  matches the reverse generative process  $p_\varphi(\mathbf{z}_s|\mathbf{z}_t)$  across *all time steps*.

In practice, we parameterize the denoising model as a noise prediction model  $\hat{\epsilon}_\varphi(\mathbf{z}_t, t)$ , which is a neural network with parameters  $\varphi$ . Additionally, the sum over the timesteps in Equation B14 can be written as an expectation over the timesteps  $t(i)$ , which is sampled uniformly (i.e., over  $i \sim U(1, T)$ ). We can thus simplify Equation B14 to be:

$$\mathcal{L}_{\text{diff}}(\mathbf{y}) = \frac{T}{2} \mathbb{E}_{\epsilon \sim \mathcal{N}(0, \mathbf{I}), i \sim U(1, T)} [w_\eta(t) \|\epsilon - \hat{\epsilon}_\varphi(\mathbf{z}_t, t)\|_2^2], \quad (\text{B15})$$

where the pre-factor  $w(t)$  can be written as:

$$w_\eta(t) = \exp(\gamma_\eta(t) - \gamma_\eta(s)) - 1. \quad (\text{B16})$$

A step-by-step derivation of the diffusion loss can be found in Luo (2022). Here, we note that rewriting  $\mathcal{L}_{\text{diff}}(\mathbf{y})$  as an expectation over the timesteps  $t(i)$  allows for the negative ELBO through a Monte Carlo estimator. This greatly increases the tractability and stability of the optimization process, as we do not need to simulate the entire trajectory from  $\mathbf{y}$  to  $\mathbf{z}_0$ . Equation B15 optimizes the parameter  $\varphi$  of the noise prediction model and  $\eta = \{\gamma_{\max}, \gamma_{\min}\}$  of the noise schedule simultaneously.

Lastly, in case of a conditional generative process, we simply include the conditioning features  $\theta$  into the noise prediction model, i.e.  $\hat{\epsilon}_\varphi(\mathbf{z}_t, t, \theta)$ .

## B.4. Continuous-time diffusion

The number of diffusion steps  $T$  is an important hyperparameter of the VDM, with a higher  $T$  generally leading to better performance (see Appendix F of Kingma et al. (2021)).

In our framework, we employ a continuous-time VDM, corresponding to setting  $T \rightarrow \infty$  and also reducing the number of hyperparameters by one. In this limit, the summation over the timesteps in Equation B14 is simply replaced by an integration over  $t$  from  $t = 0$  to  $t = 1$ . The final diffusion loss becomes:

$$\mathcal{L}_{\text{diff}}(\mathbf{y}) = \frac{1}{2} \mathbb{E}_{\epsilon \sim \mathcal{N}(0, \mathbf{I}), t \sim U(0, 1)} [\gamma'_\eta(t) \|\boldsymbol{\epsilon} - \hat{\boldsymbol{\epsilon}}_\varphi(\mathbf{z}_t, t)\|_2^2], \quad (\text{B17})$$

where the time derivative  $\gamma'_\eta(t) = d\gamma_\eta(t)/dt$  can be evaluated via automatic differentiation. It is worth noting that in this limit, the diffusion process is governed by a stochastic differential equation; see Song et al. (2020) for a more detailed discussion.

### C. ADDITIONAL RESULTS

#### C.1. Halo and central galaxies

Figure 12 shows the rest of the output features of the normalizing flows, including distributions of the halo virial mass  $M_{\text{halo}}$ , halo stellar mass  $M_\star$ , central galaxy total mass  $M_{\text{cent}}$ , and the velocity offset between the halo and central galaxies  $|\mathbf{v}_{\text{cent}}|$ , as a function of the simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ .

We do not observe strong variations of  $M_{\text{halo}}$  and  $M_{\text{cent}}$ , as expected due to the mass selection criterion  $(1.58 - 1.61) \times 10^{12} M_\odot$ , described in Section 2.2. However, since the target halo is selected from the initial DM-only, low-resolution uniform box, its mass could change when increasing the resolution and adding baryons in the final zoom-in simulation. Indeed, we observe minor but notable variations of  $M_{\text{halo}}$  and  $M_{\text{cent}}$  to  $A_{\text{SN1}}$  and  $A_{\text{SN2}}$ . These variations are captured by the flows and agree with the simulations.

The distribution of the halo stellar mass  $M_\star$  has similar trends to the that of the central galaxy, i.e.  $M_{\text{cent},\star}$ . This is expected given the strong correlations between  $M_\star$  and  $M_{\text{cent},\star}$  (see Figure 4). Similar to  $M_{\text{cent},\star}$ , we see that the distribution predicted by NeHOD generally align well with the simulations, with some discrepancies in the 16–84th percentiles. We expect the agreement will improve with more training samples.

Lastly, the distribution of the velocity offset  $|\mathbf{v}_{\text{cent}}|$  also shows good agreement between NeHOD and the simulations. The velocity offset  $|\mathbf{v}_{\text{cent}}|$  shows little variations with respect to the simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . Notably, the distribution has a long tail extending toward higher values of  $|\mathbf{v}_{\text{cent}}|$ . We see from the slight disagreement in the 84th percentiles that the flows have some difficulties capturing this tail. This is likely due to the limited number of training samples in this high-velocity range, and we anticipate improvement with additional training data.

#### C.2. Satellite halo mass functions

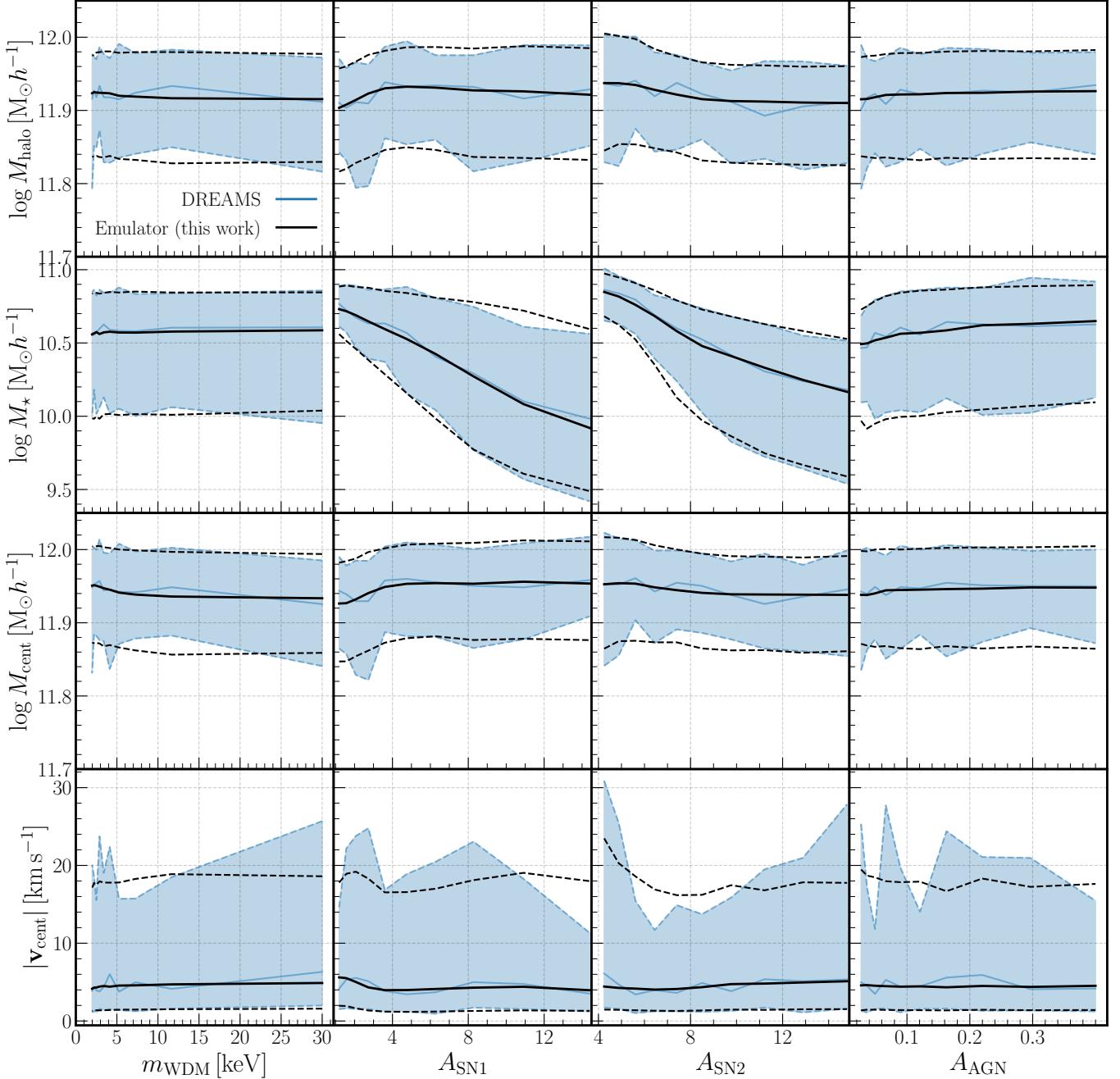
Figure 13 and Figure 14 show the satellite halo mass functions (SHMFs) and their differential forms, respectively.

#### C.3. Velocity distribution of satellites

We compute the distributions of radial velocity  $v_{\text{rad}}$ , the pairwise velocities  $v_{\text{pair}}$ , and the nearest-neighbor velocities  $v_{\text{NN}}$ . For simplicity, the radial velocity of each satellite is calculated from the center of the *halo* (using the reference frame described in Section 3). In other words, we do not take into account the velocity offsets between the halos and their central galaxies. These offsets are typically much smaller compared to the velocity of the satellites (order of  $10 \text{ km s}^{-1}$  compared to  $100 \text{ km s}^{-1}$ ), so we do not expect them to significantly affect our results. The nearest-neighbor velocity of each satellite is determined by identifying its nearest neighbor (in position space) and calculating the velocity between them. For each satellite, we assign its velocity a positive (negative) sign if it moves away from (towards) the central galaxy for  $v_{\text{rad}}$ , or another satellite satellite for  $v_{\text{pair}}$  and  $v_{\text{NN}}$ .

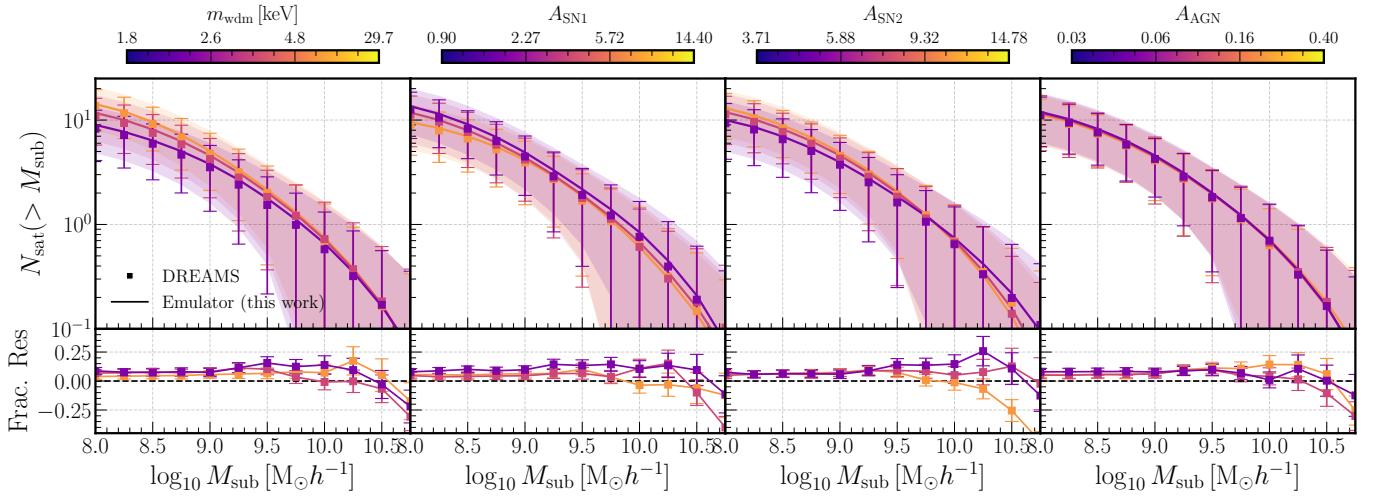
Figure 15 shows the distributions of the radial velocities  $v_{\text{rad}}$  (left), the pairwise velocities  $v_{\text{pair}}$  (middle), and the nearest-neighbor velocities  $v_{\text{NN}}$  (right) of satellites of the NeHOD samples (blue) and the simulations (orange). Similar to Figure 10, the medians, 16th, and 84th percentiles are shown. Note that, unlike the nearest-neighbor separations, we do not observe strong variations of the nearest-neighbor velocities with the simulation parameters, and thus do not show results for individual bins. For  $v_{\text{pair}}$  and  $v_{\text{NN}}$ , we show their probability distribution functions (PDFs) instead CDFs as the comparison is easier to interpret with PDFs due to the bimodality in the distributions (described below). Additionally, we note that although both peaks are at the same absolute velocity ( $\pm 200 \text{ km s}^{-1}$ , the negative peak is slightly more prominent, suggesting that the overall dynamics of the system have more infall.

Overall, NeHOD can predict the general trend in both the median and the  $1\sigma$  percentiles of the velocity distributions. However, we note similar discrepancies to those observed in the spatial distributions. In particular, NeHOD underpredicts the distributions of satellites moving at  $-100 \text{ km s}^{-1} < v_{\text{pair}}, v_{\text{NN}} < 100 \text{ km s}^{-1}$ . This implies that generated satellites are typically moving relatively to each other faster than those in the simulations. Additionally, NeHOD also underpredicts the distributions of  $v_{\text{pair}}$  at high negative velocity, approximately  $-400 \text{ km s}^{-1}$ . Since this occurs in the tail of the distribution, the discrepancy could be attributed to the limited number of training samples in this velocity range. Similarly, we observe a better agreement between the radial velocity  $v_{\text{rad}}$  distributions of the NeHOD samples and simulations. We discuss the limitations of the

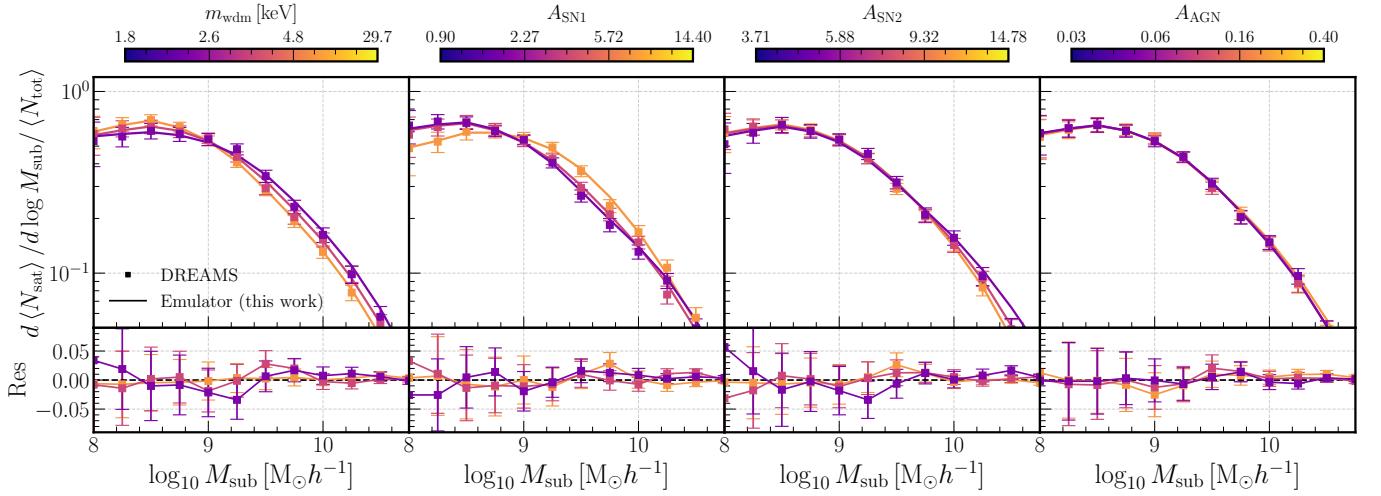


**Figure 12.** The properties of halos and central galaxies as a function of the simulation parameters  $\{m_{\text{WDM}}, A_{\text{SN1}}, A_{\text{SN2}}, A_{\text{AGN}}\}$  (left to right). From top to bottom, the panels show the halo virial mass, halo stellar mass, central galaxy total mass, and the velocity offset between the halo and central galaxies. Panels are the same as in Figure 3.

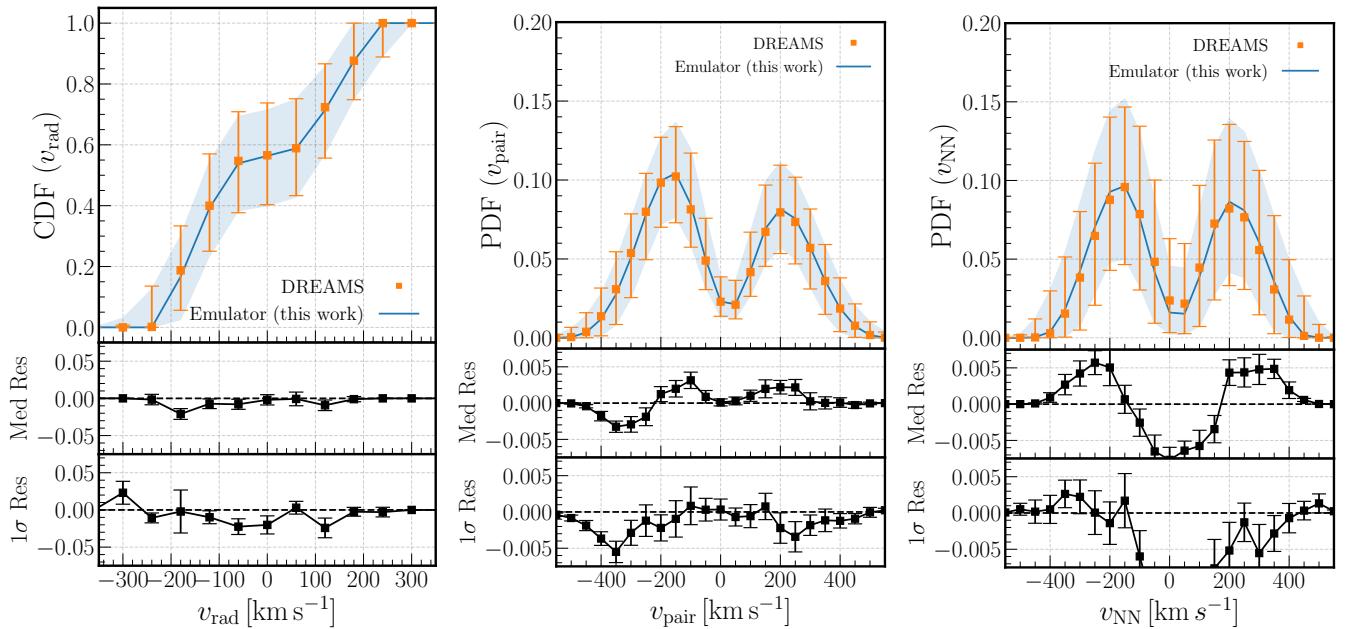
current model and potential ways to overcome them in Section 5.



**Figure 13.** *Top:* The halo stellar mass functions (SHMFs) generated by NeHOD and extracted from the simulations. The columns show the variations of the SHMFs over the WDM mass  $m_{\text{WDM}}$  and astrophysical parameters  $\{A_{\text{SNI}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . In each column, the color denotes the SHMFs of a bin of the corresponding parameter. The average SHMFs, along with the  $1-\sigma$  scatter, are shown as solid lines and shaded regions for NeHOD and error bars for the simulations. *Bottom:* The fractional residuals of the SHMFs, estimated as the differences between NeHOD and the simulation values.



**Figure 14.** *Top:* The differential SHMFs  $d \langle N_{\text{sat}} \rangle / d \log M_{\text{sub}}$ , normalized by the average total number of satellites  $\langle N_{\text{tot}} \rangle$ . The columns show the variations of the SHMFs over the WDM mass  $m_{\text{WDM}}$  and astrophysical parameters  $\{A_{\text{SNI}}, A_{\text{SN2}}, A_{\text{AGN}}\}$ . The NeHOD and simulation differential SHMFs are shown as solid lines and squares, respectively. The error bars denote the standard errors of the simulations. *Bottom:* The residuals between NeHOD and simulation differential SHMFs.



**Figure 15.** The distributions of radial velocities (left), pairwise velocities (middle), and the nearest-neighbor velocities (right). The top panels show the medians and  $1\sigma$  intervals of the CDFs for the NeHOD samples (blue band) and the simulations (error bars). The middle and bottom panels show the residuals between the NeHOD samples and the simulations of the medians and the  $1\sigma$  intervals, respectively.