Halo assembly bias in the early Universe: a clustering probe of the origin of the Little Red Dots

Zihao Wang[1], Fangzhou Jiang[2]★, Haonan Zheng[2]†, Xuejian Shen[3], Zixiang Jia[1], Luis C. Ho[2], Kohei Inayoshi[2], Linhua Jiang[2]

[1] *Department of Astronomy, Peking University, Beijing 100871, China*
[2] *Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China*
[3] *Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*

arXiv:2603.15736v1 [astro-ph.CO] 16 Mar 2026

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

The clustering of galaxies encodes key information about the structure and assembly history of their host dark matter (DM) haloes, providing a powerful probe of the origin of extreme high-redshift systems. While halo assembly bias has been extensively studied at low redshift, its behavior in the early Universe remains poorly explored. Using the large-volume, high-resolution Shin-Uchuu cosmological N-body simulation, we characterize halo assembly bias associated with formation time, concentration, and angular momentum across a wide range of halo masses and redshifts. We find that the sign and amplitude of assembly bias depend on halo mass for both concentration and spin. High-concentration and low-spin haloes are more strongly clustered below characteristic peak heights of $\nu \sim 1.5$ and ~ 0.75, respectively, while the trends weaken or reverse at higher masses. Halo age bias persists at all redshifts but decreases toward higher masses and earlier cosmic times. We apply these results to assess whether clustering can distinguish competing formation scenarios for the Little Red Dots (LRDs). We find that the direct-collapse-black-hole (DCBH) scenario predicts the strongest large-scale bias and enhanced pair fractions, the self-interacting-dark-matter (SIDM) core-collapse scenario and low-spin compact-galaxy scenarios yield weaker clustering due to lower characteristic halo masses and spin-related secondary bias, and a primordial-black-hole (PBH) scenario predicts unbiased clustering. Our results demonstrate that halo assembly bias and characteristic host masses provide powerful diagnostics for constraining the physical origin of LRDs, offering testable predictions for upcoming clustering measurements with *JWST* and future deep surveys.

Key words: galaxies:high-redshift – galaxies:haloes – galaxies: active – dark matter

1 INTRODUCTION

In the standard ΛCDM paradigm, dark matter (DM) governs the formation of the large-scale structure (LSS) of the Universe and forms gravitationally bound haloes that provide the sites for galaxy formation and evolution (Press & Schechter 1974; White & Frenk 1991). Galaxies and their host halo trace the otherwise invisible DM distribution; However, they do not constitute random sampling of the underlying smooth matter density field. Instead haloes are clustered and preferentially reside in overdense regions. This effect is quantified by the *halo bias* parameter b, defined through

$$\xi_{\rm hh}(r|M_{\rm vir}, z) = b^2 \xi_{\rm mm}(r, z), \quad (1)$$

where $\xi_{\rm hh}(r|M_{\rm vir}, z)$ denotes the two-point correlation function of haloes of virial mass $M_{\rm vir}$ at redshift z separated by distance r, and $\xi_{\rm mm}(r, z)$ is that of the underlying DM field. Halo bias depends primarily on halo mass, with more massive haloes forming preferentially in denser environments (e.g., Kaiser 1984; Sheth & Tormen 1999). At fixed mass, however, the clustering strength exhibits systematic dependence on secondary halo properties, such as formation time and internal structures (e.g., Gao et al. 2005; Gao & White 2007; Wechsler et al. 2006; Li et al. 2008; Salcedo et al. 2018). This dependence of halo clustering on properties beyond mass is referred to as *assembly bias*.

A number of secondary halo properties are known to influence the halo bias parameter, as revealed by cosmological N-body simulations. Notably, at fixed mass, haloes that form earlier are more strongly clustered than those that form later (Gao et al. 2005). The amplitude of this effect is most pronounced for low-mass haloes and progressively weakens towards higher masses (e.g., Gao & White 2007; Salcedo et al. 2018; Sato-Polito et al. 2019). Related, assembly bias effects are found for halo internal structure: at fixed mass, higher-concentration haloes exhibit stronger clustering (e.g., Wechsler et al. 2006; Gao & White 2007; Salcedo et al. 2018) below a characteristic mass scale of $M_{\rm vir} \simeq 10^{13} M_\odot$ at $z = 0$. Interestingly, this trend reverses at higher masses (Sato-Polito et al. 2019). A similar phenomenon has been identified for halo angular momentum,

★ Corresponding author: fangzhou.jiang@pku.edu.cn
† hnzheng@pku.edu.cn

© 2026 The Authors

with higher-spin haloes being more strongly clustered, particularly above $M_{\rm vir} \simeq 10^{12} M_\odot$ (e.g., Bett et al. 2007; Faltenbacher & White 2009; Lacerna & Padilla 2012). These results demonstrate that the strength of secondary bias effects depends sensitively on halo mass. It is therefore natural to expect that their magnitude, and potentially even their qualitative behavior, may evolve with cosmic time. However, this redshift dependence remains relatively unexplored.

Assembly bias provides valuable insight into the formation pathways of galaxies with extreme morphologies. For example, Zhang et al. (2025c) showed that isolated ultra-diffuse galaxies (UDGs) exhibit surprisingly stronger clustering than that expected for dwarf-mass haloes, which are commonly assumed to host such systems. This finding has motivated scenarios in which UDGs preferentially reside in earliest-forming haloes composed of self-interacting dark matter, allowing sufficient time for a shallow DM core to develop.

The *James Webb Space Telescope* (*JWST*) has revealed a diverse population of high-redshift galaxies with morphologies markedly distinct from those in the nearby Universe, ranging from compact and clumpy systems (e.g., Tanaka et al. 2024a; Mowla et al. 2024; Fujimoto et al. 2025; Vanzella et al. 2026) to extremely extended disks (e.g., Wang et al. 2025a; Umehata et al. 2025). Their formation, taking the early giant disks as an example, has been hypothesized to be linked to particular halo structures and environments (Jiang et al. 2025a). At the most compact end of this morphological diversity lie the so-called Little Red Dots (LRDs), which have attracted intense attention (e.g., Matthee et al. 2024; Greene et al. 2024; Furtak et al. 2024; Labbe et al. 2023; Kocevski et al. 2025). LRDs have been proposed as a potentially new class of active galactic nuclei (AGN) host galaxies. They are characterized by a distinctive V-shaped spectral energy distribution and broad Balmer emission lines (e.g., Greene et al. 2024; Kocevski et al. 2025; Hainline et al. 2024), yet lack several canonical AGN signatures, such as strong X-ray emission, prominent near-to-mid-infrared excess, or powerful radio jets (e.g., Maiolino et al. 2024a; Yue et al. 2024; Akins et al. 2024; Pérez-González et al. 2024; Gloudemans et al. 2025). Based on these observed signatures, LRDs are inferred to host supermassive black holes (SMBHs) with masses of $\sim 10^6 - 10^8\, M_\odot$ (Matthee et al. 2024; Greene et al. 2024; Taylor et al. 2025), while exhibiting a striking deficiency of stellar components (Chen et al. 2025a). The ubiquity of such systems at $z \gtrsim 5$ presents significant challenges to conventional baryonic seeding scenarios and to the standard framework of galaxy-black hole coevolution established at lower redshift Universe (e.g., Kormendy & Ho 2013; Reines & Volonteri 2016).

Although the physical nature and the internal structure of LRDs remains under debate (e.g., Naidu et al. 2025; Lin et al. 2025a; Inayoshi et al. 2025; Kido et al. 2025), a number of scenarios has been proposed to explain their cosmological statistical origin, particularly in terms of the properties required of their host DM haloes. First, within the standard ΛCDM paradigm, the direct-collapse-black-hole (DCBH) scenario posits that LRDs originate from massive seeds formed in chemically pristine haloes with sufficiently deep potential wells, exposed to strong Lyman-Werner (LW) radiation that suppresses molecular hydrogen cooling and fragmentation (e.g., Bromm & Loeb 2003; Begelman et al. 2006; Lodato & Natarajan 2006). While this mechanism provides a viable pathway for forming massive black hole seeds, its ability to reproduce the observed abundance and mass distribution of LRDs remains to be fully quantified in a cosmological statistical framework.

Second, in alternative cosmologies where DM possesses significant self-interactions, LRDs may arise from gravothermal core collapse in self-interacting dark-matter (SIDM) haloes (Jiang et al. 2025b; Shen et al. 2025; Feng et al. 2025). In this picture, early formation and high halo concentration are key conditions required to trigger sufficiently rapid core collapse and SMBH seeding by the observed epoch (Jiang et al. 2025b).

Beyond scenarios centered on massive BH seed formation, additional interpretations do not necessarily explicitly require LRD to host SMBHs. For example, Pacucci & Loeb (2025) argue that, based on the observed abundance, compactness, and redshift distribution of LRDs, these systems may preferentially inhabit haloes in the extreme low-spin tail of the angular momentum distribution. In this model, LRDs emerge as compact galaxies lacking rotational support due to the intrinsically low spin of their host haloes (Mo et al. 1998).

Although many of the aforementioned scenarios can reproduce halo number densities broadly consistent with the observed abundance of LRDs, they imply markedly different host halo properties, such as formation time, concentration, or spin. The clustering strength of their host haloes, therefore, may provide a powerful and complementary diagnostic of LRD origins, particularly in light of forthcoming large-scale clustering measurements from *JWST* surveys. However, while assembly bias has been extensively studied in cosmological simulations at low redshift ($z \lesssim 3$; e.g., Contreras et al. 2019), its behavior in the early Universe ($z \gtrsim 5$) remains poorly understood. With the advent of N-body simulations that simultaneously achieve high mass resolution and large cosmological volume, it is timely to revisit halo assembly bias at high redshift and explore how halo clustering can be used to constrain the formation pathways of LRDs.

In this work, we extend measurements of halo assembly bias associated with halo age, concentration, and spin – the properties that are closely related to proposed formation pathways of LRDs – to redshifts as high as $z = 10$. Using the high-resolution dark-matter-only (DMO) simulation Shin-Uchuu (Ishiyama et al. 2021), we quantify how these bias signals evolve with mass and redshift, providing an important reference for future clustering analyses. We then apply these results, for the first time, to LRD formation scenarios by populating haloes with LRDs according to the aforementioned models.

This paper is organized as follows. In Section 2, we introduce the simulations used in this work and list the definitions of relevant halo properties. In Section 3, we describe the measurements of halo assembly bias signals and examine their redshift evolution. In Section 4, we outline how LRD host haloes are selected in the simulation under different physical scenarios, compare their cross-correlations with those of normal galaxies, and discuss the implications for LRD formation pathways. We summarize the main conclusions in Section 5. Throughout the paper, we adopt a cosmology used in the Shin-Uchuu simulation, and define haloes as spherical overdensities that are 200 times the critical density of the Universe.

2 SIMULATION AND METHODS

2.1 Simulation

To characterize the clustering signal of the haloes of interest, we utilize the Shin-Uchuu N-body simulation (Ishiyama et al. 2021; Oogi et al. 2023; Dong-Páez et al. 2024; Aung et al. 2023; Prada et al. 2023). The simulation has a comoving box of side-length $140\, h^{-1}$ Mpc, particle mass of $8.97 \times 10^5\, h^{-1} M_\odot$, and a gravitational softening length of $0.4\, h^{-1}$ kpc. DM haloes are identified with the ROCKSTAR[1] halo finder (Behroozi et al. 2012a), which employs six-dimensional phase-space and temporal information to separate

[1] https://bitbucket.org/gfcstanford/rockstar/

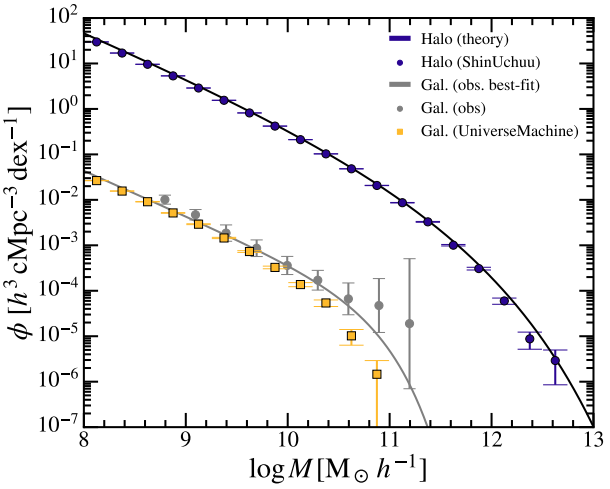

Figure 1. Mass functions at $z \simeq 5$ for distinct haloes in Shin-Uchuu (blue circles) and galaxies from the UniverseMachine catalogue (yellow squares). The halo mass function is in good agreement with theoretical expectations from Sheth et al. 2001 (blue solid), while the galaxy mass function exhibits a good agreement at the low-mass end and a deficit at the massive end relative to recent *JWST* results from Wang et al. 2025b (gray).

overlapping structures and to track subhaloes across snapshots. The halo center corresponds to the position of the particle with the minimum potential. In this work, we use the group catalogue provided by the Shin-Uchuu team[2], which includes pre-computed halo properties. Merger trees are constructed with a modified, parallelized implementation of the CONSISTENT TREE[3] algorithm (Behroozi et al. 2012b). Throughout the analysis, we consider haloes with more than 100 particles as resolved. A halo is regarded as relaxed if it satisfies $X_{\rm off} < 0.07$ and $2T/|U| < 1.5$, where $X_{\rm off} \equiv |\vec{r}_{\rm center} - \vec{r}_{\rm com}|/R_{\rm vir}$ denotes the halo's center-of-mass offset, and $2T/|U|$ is the virial ratio, where U and T are the potential and kinetic energies, respectively. In Fig. 1, we show that the halo mass function measured from Shin-Uchuu at $z \simeq 5$, compare it with the theoretical prediction from the Press-Schechter theory with the ellipsoidal-collapse critical overdensity Sheth et al. (2001).

2.2 Measurement of halo properties

We investigate halo assembly bias across several halo properties, focusing in particular on concentration, spin, and formation time. Since multiple definitions of these parameters are provided in the halo catalogue, here we specify the definition that we adopt.

We use the concentration obtained by fitting the Navarro-Frenk-White (NFW; Navarro et al. 1997) density profile

$$\rho(r) = \frac{\rho_{\rm s}}{(r/r_{\rm s})\,(1 + r/r_{\rm s})^2}, \tag{2}$$

where $\rho_{\rm s}$ is the characteristic (scale) density and $r_{\rm s}$ is the scale radius. Halo particles are divided into up to 50 radial equal-mass bins, with a minimum of 15 particles per bin, and an NFW profile is directly fitted to the density distribution to obtain the maximum-likelihood estimate of ρ_s and $r_{\rm s}$. We then construct concentration as $c = R_{\rm vir}/r_{\rm s}$. The scale density $\rho_{\rm s}$ is not saved in the catalogue –when needed, we derive it from the relation,

$$\rho_{\rm s} = \frac{M_{\rm vir}}{4\pi r_{\rm s}^3\,[\ln(1 + c) - c/(1 + c)]}. \tag{3}$$

We use the spin parameter in the Bullock et al. (2001) definition, given by

$$\lambda = \frac{J_{\rm vir}}{\sqrt{2} M_{\rm vir} R_{\rm vir} V_{\rm vir}} \tag{4}$$

where $J_{\rm vir}$ is the total angular momentum of all bound DM particles within the virial radius, $V_{\rm vir}$ is its circular velocity at the virial radius.

We define the halo formation epoch as the (earliest) redshift when the halo mass exceeds half its present virial mass, $z_{1/2}$. The corresponding look-back time of formation is then

$$t_{\rm form} = t(z) - t(z_{1/2}), \tag{5}$$

where $t(z)$ is the cosmic time at redshift z.

The prominence of a DM halo relative to the cosmic density fluctuations is commonly quantified by the peak height of the associated overdensity, defined as

$$\nu \equiv \delta_{\rm c}/\sqrt{S(M_{\rm vir})}, \tag{6}$$

where $\delta_{\rm c}(z)$ denotes the critical linearly extrapolated overdensity for halo collapse, and $S(M_{\rm vir})$ is the variance of the density field smoothed on the mass scale $M_{\rm vir}$.

Fig. 2 presents the distributions of concentration, spin, and formation time as functions of halo mass for all resolved distinct haloes at $z = 5$. In contrast to the concentration-mass relation at low redshift (e.g., Prada et al. 2012; Ludlow et al. 2012; Dutton & Macciò 2014; Ishiyama et al. 2021), which exhibits a negative slope, the halo concentration at cosmic dawn is nearly mass-independent, with a median of $\simeq 3$, consistent with recent findings in other cosmological simulations (e.g., Ishiyama et al. 2021; Yung et al. 2024). A pronounced low-c tail extending to $c \sim 1$ is evident and is primarily associated with perturbed haloes, for which concentration is ill-defined. Excluding these systems shifts the distribution towards higher c values. The median spin parameter shows little correlation with mass, with a median value of $\lambda \simeq 0.04$ and no significant redshift evolution, consistent with previous findings (e.g., Bullock et al. 2001). The formation time of low-mass haloes are generally earlier than more massive ones, reflecting the hierarchical nature of structure formation.

2.3 UniverseMachine galaxy catalogue

To model the galaxy population, we use the publicly available UniverseMachine galaxy catalogue (Behroozi et al. 2019). UniverseMachine is a forward modeling framework that populates galaxies within cosmological N-body simulations by assigning star formation histories, stellar masses, and UV luminosities through an empirically calibrated model of galaxy growth. The framework describes star formation as a function of halo mass, assembly history, and redshift using a flexible parameterization, with parameters constrained by jointly fitting a wide range of observational data, including stellar mass functions, UV luminosity functions, the cosmic star formation rate density, quenched fractions, and, in particular, galaxy clustering over the redshift range $z = 0$ to $z \simeq 10$. The explicit calibration to clustering statistics ensures that the spatial distribution of galaxies is consistently modeled alongside their internal properties, thereby providing a robust foundation for subsequent clustering analyses.

In this work, we adopt the $z \simeq 5$ mock galaxy catalogue constructed on the Shin-Uchuu halo merger trees. As shown in Fig. 1,

[2] https://www.skiesanduniverses.org/Simulations/Uchuu/
[3] https://bitbucket.org/pbehroozi/consistent-trees/

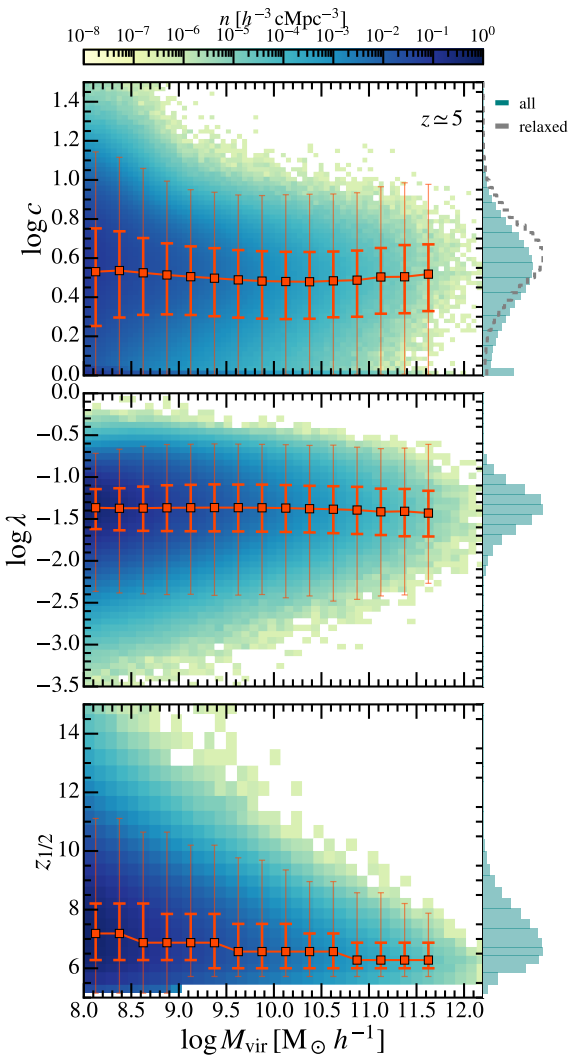

Figure 2. Distributions of halo properties examined in this study, at $z \simeq 5$ in the Shin-Uchuu simulation. From top to bottom, the panels display halo concentration, spin, and formation redshift as functions of halo mass. The colour indicates number density, while the solid curves indicate the median relation. The thicker and thinner error bars denote the 1σ and 4σ intervals, respectively. Side panels show the corresponding one-point distributions. For concentration, the histogram of the subsample of relaxed haloes is included for comparison, as indicated by the dashed grey line. At high redshift, halo concentration exhibits no clear dependence on mass. Spin is nearly mass-independent. Lower-mass haloes tend to form earlier than more massive ones.

the stellar mass function from UniverseMachine at $z \simeq 5$ is broadly consistent with recent *JWST* constraints (Wang et al. 2025b), with only a sight deficit at the massive end. We use the predicted stellar masses, star formation rates (SFRs), UV magnitudes, and galaxy-halo connections to identify candidate LRD hosts under certain formation scenarios discussed in Sect. 4.

3 REDSHIFT EVOLUTION OF HALO ASSEMBLY BIAS

In this section, we investigate how the clustering of haloes depends on concentration, spin, and formation time, and characterize the redshift evolution of the halo assembly bias.

3.1 Auto-correlation function

We base our analysis on all resolved distinct haloes, and compute the auto-correlation function (ACF) using

$$\xi(r) = \frac{DD(r)}{RR(r)} - 1 \tag{7}$$

where $DD(r)$, and $RR(r)$ denote the normalized numbers of halo-halo and random-random pairs, respectively, counted in bins of separation r. These quantities are defined as

$$\begin{aligned} DD(r) &= \frac{N_{DD}}{N_D(N_D-1)/2}, \\ RR(r) &= \frac{N_{RR}}{N_R(N_R-1)/2}, \end{aligned} \tag{8}$$

where N_D (N_R) is the number of haloes (random points); N_{DD} and N_{RR} are the numbers of halo-halo and random-random pairs, respectively. The calculation is performed for all haloes in the simulation box that satisfy the selection criteria, using the publicly available CORRFUNC code (Sinha & Garrison 2019). Haloes are further grouped in bins of width 0.25 in $\log(M_{\rm vir}/M_\odot)$ spanning the range 8 to 13. We require at least 1000 pairs per mass bin to obtain statistically robust measurements.

Fig. 3 presents the ACFs at $z \simeq 5$ in three halo mass bins, subdivided by percentile ranges of each halo parameter under consideration. The results reveal both similarities and differences relative to the trends observed in the local Universe.

At the low-mass end, $M_{\rm vir} \simeq 10^8 M_\odot$, corresponding to an overdensity peak height of $\nu \simeq 1.3$, haloes with higher concentrations are marginally more strongly clustered than their low-concentration counterparts.[4] This trend exhibits a mass dependence, weakening toward higher masses and eventually reversing: by $M_{\rm vir} \simeq 10^{11} M_\odot$, haloes in the lowest 10% of the concentration distribution are nearly twice as strongly clustered as those in the highest 10%. In the local Universe, this nonlinear dependence is often connected to the statistics of density peaks in the initial field (Dalal et al. 2008). At fixed ν, peaks of lower curvature are more strongly clustered and tend to assemble their mass earlier than high-curvature peaks, suggesting a connection to the concentration dependence of halo bias. At the low-mass end, however, assembly bias is generally attributed to nonlinear environmental effects that may depend on the location of haloes within the cosmic web (e.g., Hahn et al. 2009; Borzyszkowski et al. 2017), for which a comprehensive theoretical framework has yet to be established and is beyond the scope of this work.

High-spin haloes exhibit consistently stronger clustering across the mass range probed here, and the strength of this spin-dependent bias shows little variation with halo mass, in contrast to the trend observed at $z = 0$ (e.g., Sato-Polito et al. 2019). In the local Universe, dwarf low-spin haloes are more strongly clustered at low redshift, a trend commonly attributed to splashback haloes (e.g., Tucci et al. 2021). These systems reside near massive hosts and inherit the large-scale bias of their neighbors. At $z = 5$, however, splashback haloes are expected to be extremely rare. The trend we observe therefore likely reflects the intrinsic spin-dependent bias, which may arise from tidal torques exerted by nearby large-scale structures: because

[4] At such high redshift, more than half of the haloes are classified as perturbed under our criteria, with a non-negligible fraction exhibiting $c = 1$. In our default analysis, these systems belong to the low-concentration subsamples and could bias the clustering of the low-c population. However, excluding these perturbed haloes yields statistically consistent results, indicating that this effect does not alter our conclusions.

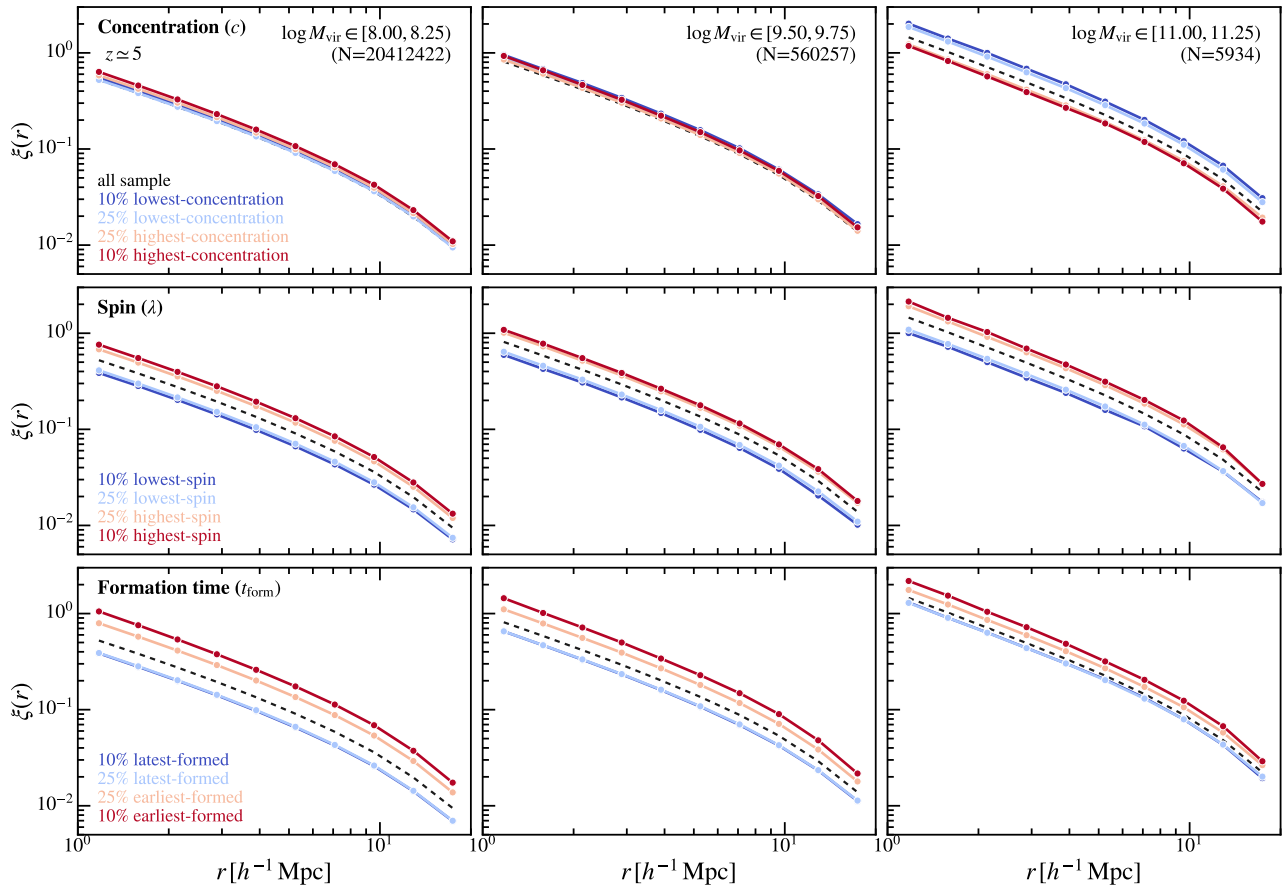

Figure 3. Auto-correlation functions (ACFs) of haloes from the Shin-Uchuu simulation at $z \simeq 5$. haloes are grouped into mass bins of width 0.25 dex, and we show three representative cases: low mass ($\log(M_{\rm vir}\,[h^{-1}{\rm M}_\odot]) \simeq 8$), intermediate mass ($\log(M_{\rm vir}\,[h^{-1}{\rm M}_\odot]) \simeq 9.5$), and high mass ($\log(M_{\rm vir}\,[h^{-1}{\rm M}_\odot]) \simeq 11$). The dashed black curves denote the ACFs of the full populations in each mass bin, while the colored curves correspond to subsamples selected by percentiles of concentration, spin, or formation time, as indicated. The three parameters exhibit distinct assembly-bias behaviors and mass trends. Concentration shows weak bias at the low-mass end and a reversal at high masses, where low-c haloes cluster more strongly. Spin exhibits a mass-independent trend in which low-spin haloes are more strongly clustered. Formation time shows enhanced clustering for the oldest haloes, with the strength of this trend weakening towards higher masses.

the tidal field that generates halo angular momentum correlates with the surrounding density field, high-spin haloes preferentially reside in more strongly clustered environments (Lacerna & Padilla 2012; Paranjape et al. 2018).

In the bottom row, the oldest haloes are more strongly clustered than the youngest ones, and this age-dependent signal also displays a notable mass dependence consistent with the local relation. A possible interpretation is that older haloes preferentially form in denser environments (Borzyszkowski et al. 2017; Contreras et al. 2021), while at the low-mass end the age-dependent bias is further driven by environmental quenching of halo growth, such as tidal suppression of mass accretion near massive neighbours (Mansfield & Kravtsov 2020).

3.2 Redshift evolution of the assembly bias

Fig. 3 has demonstrated that the assembly bias signal exhibits halo mass dependence. We now systematically quantify the mass dependence using the *relative bias* factor, defined as the ratio of the ACF of haloes selected by a given secondary property to that of the unconditional halo population at the same mass,

$$\tilde{b}^2_X(r|M_{\rm vir}) = \frac{\xi_X(r|M_{\rm vir})}{\xi(r|M_{\rm vir})}. \tag{9}$$

Here, $\xi_X(r|M_{\rm vir})$ denotes the ACF of haloes selected the secondary property, X (e.g., concentration, spin, or formation time), while $\xi(r|M_{\rm vir})$ represents the ACF of the full halo population within the same virial mass bin. Unless otherwise specified, the relative bias is evaluated over scales of 5-15 h^{-1} Mpc to facilitate direct comparison with previous studies (e.g., Tucci et al. 2021; Sato-Polito et al. 2019). We estimate uncertainties using 1000 bootstrap re-samplings with replacement, and find that in most cases the resulting uncertainties are small. The bias factors are then computed as the error-weighted mean of $\tilde{b}^2_X(r|M_{\rm vir})$ over the adopted scale range, where the uncertainties are estimated via bootstrap re-sampling.

Fig. 4 presents the relative bias of the upper and lower quartiles of each parameter as a function of halo mass, evaluated over redshifts $0 \le z \le 8$. The relative bias associated with halo concentration exhibits a strong dependence on both mass and redshift. At $z \gtrsim 6$, low-concentration haloes are marginally more strongly clustered than the full population, while high-concentration haloes are less clustered, across the entire mass range probed. As cosmic time progresses, the signal weakens and reverses sign at a redshift-dependent transition

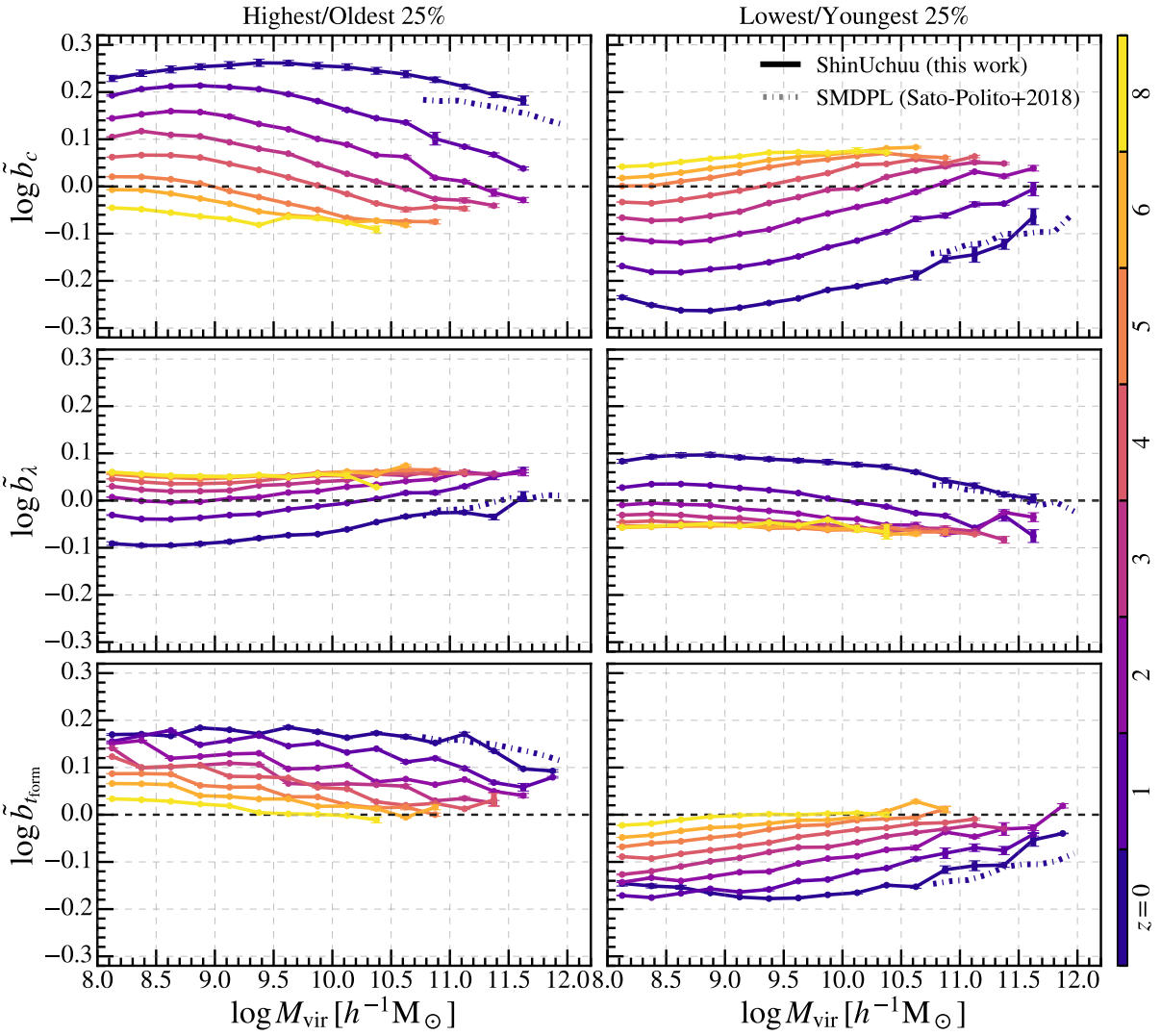

Figure 4. Redshift evolution of secondary bias as a function of halo mass for concentration (top), spin (middle), and formation time (bottom), from $z = 0$ to $z \simeq 8$. In each panel, the left-hand column shows the bias factors for the upper 25% of the parameter distribution, and the right-hand column shows the corresponding lower 25%. Curves are evaluated over scales of $5 - 15\,h^{-1}$ Mpc. At $z = 0$, we assess consistency by comparing our measurements with results from Sato-Polito et al. (2019). The amplitudes of all the secondary biases here exhibit dependence on halo mass and clear evolution with redshift.

mass. For example, at $z \simeq 5$ (4), the sign flip for the high-c subsample occurs at $M_{\rm vir} \sim 10^{9.0} h^{-1} M_\odot$ ($10^{9.8} h^{-1} M_\odot$). At later times, this crossover mass steadily shifts toward higher values. Meanwhile, at fixed halo mass below the transition scale, the amplitude of the concentration bias grows monotonically. By $z = 0$, concentration exhibits the strongest secondary-bias signal among the three parameters considered here.

The relative bias associated with halo spin exhibits marginal dependence on mass or redshift at $z \gtrsim 4$, where low-spin haloes are consistently less strongly clustered than high-spin haloes across all masses. Beginning around $z = 3$, the clustering of low-spin haloes progressively increases, eventually reversing sign at the low-mass end. The sign flip of the assembly bias of the low-λ subsample occurs at $M_{\rm vir} \simeq 10^{10.0} h^{-1} M_\odot$ ($10^{11.5} h^{-1} M_\odot$) at $z \simeq 1$ (0). As discussed above, this inversion could be attributed to splashback haloes. Ma et al. (2025) further argued that the reversal may also be driven by tidal effects in dense environments, including cosmic filaments, where low-mass haloes can lose high-angular-momentum material from their outer regions, thereby enhancing the clustering of the remaining low-spin haloes. Our results provide further support for these interpretations: the reversal first appears among low-mass haloes and emerges around the epoch when proto-clusters begin to form.

The oldest haloes are more strongly clustered than the youngest ones across nearly all redshifts. The strength of this age bias decreases with increasing halo mass and with increasing redshift. Overall, all bias signals are rather weak at cosmic dawn.

As a sanity check, our measurements at $z = 0$ based on the high-resolution Shin-Uchuu simulation are broadly consistent with previous studies that used larger-volume simulations with coarser resolution (e.g., Sato-Polito et al. 2019).

Since the assembly-bias signals associated with both concentration and spin exhibit clear crossover behaviors, it is useful to characterize the redshift evolution of the crossover mass scales, $M_{\rm c}$. As shown in Fig. 4, the secondary bias signals display a strong symmetry between the two ends of each parameter distribution. We therefore quantify the crossover mass using the *high-concentration* and *low-spin* pop-

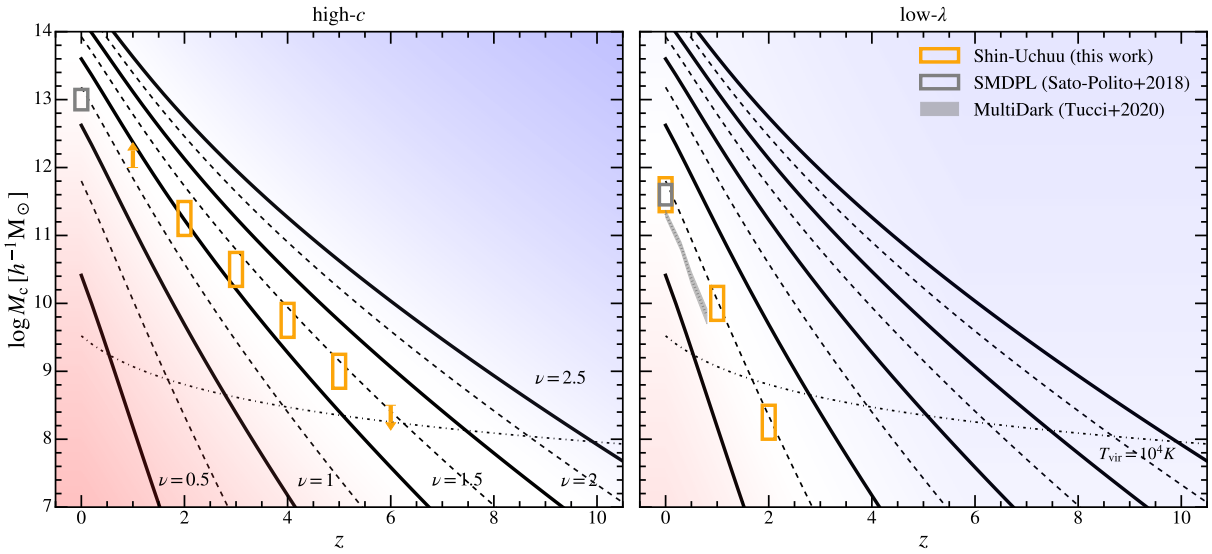

Figure 5. Redshift evolution of the critical mass, M_c, at which the assembly-bias signals reverses sign for high-concentration (left) and low-spin (right) haloes. The critical masses are extracted from Fig. 4 as the halo-mass intervals within which the relative bias (Eq. (9)) of the top 25% highest-c or bottom 25% lowest-λ haloes crosses unity. Overlaid for comparison are the contours of peak height (ν), measurements at low redshift from previous studies (grey symbols or band), and the atomic-cooling threshold, which serves as a proxy for the minimum halo mass for galaxy formation. A red (blue) background colour indicates the regimes of more (less) strong clustering, with the colour saturation schematically illustrates the strength of deviating from a relative bias of unity. The critical masses for high-c bias and low-λ bias broadly correspond to $1.5-1.75\sigma$ and 0.75σ density peaks, respectively.

ulations that are associated with the formation of LRDs, noting that the opposite selections exhibit similar characteristic mass scales. We identify the crossover mass as the halo-mass range over which the relative bias crosses unity, using mass bins of a width of 0.5 dex. For redshifts at which no clear crossing is detected within the probed mass ranges, we indicate the crossover scale as an upper or lower limit. The results are shown in Fig. 5. These measurements are compared with contours of peak height.

As shown in the left panel of Fig. 5, the crossover mass for the assembly bias associated with concentration broadly corresponds to 1.5σ density peaks, ranging between $\sim 1.25\sigma$ at $z = 1$ to $\sim 1.75\sigma$ at $z = 6$. Extrapolation toward lower redshift yields values consistent with the results of Sato-Polito et al. (2019) at $z = 0$.

We further compare the crossover mass with the halo mass that corresponds to the atomic-cooling threshold, given by the virial temperature criteria:

$$T_{\rm vir} \equiv \frac{\mu m_{\rm p}}{2k_{\rm B}} V_{\rm vir}^2 > 10^4\,{\rm K}, \tag{10}$$

where $\mu \simeq 0.6$ is the mean molecular weight, $m_{\rm p}$ is the proton mass, and $k_{\rm B}$ is the Boltzmann constant. This approximately corresponds to a halo mass above which haloes are able to form galaxies. This comparison reveals that galaxies residing in high-concentration haloes at $z \gtrsim 6$ tend to be less strongly clustered than the general galaxy population at the same mass scale.

The right panel of Fig. 5 presents an analogous analysis for the spin bias, and identifies the crossover masses below which low-spin haloes become more strongly clustered than their high-spin counterparts. The crossover masses broadly corresponds to 0.75σ density peaks. The low-redshift results are in good agreement with previous measurements (e.g., Sato-Polito et al. 2019; Tucci et al. 2021). Comparison with the atomic cooling threshold indicates that at $z \gtrsim 2$, the low-spin haloes that can host galaxies are less clustered than the general population.

4 IMPLICATIONS ON THE ORIGIN OF LRDS

The halo assembly bias at high redshifts, as characterized above, may provide a useful handle for identifying the host haloes of galaxies with unusual properties. For example, the various proposed scenarios for the LRDs attribute their origin to distinctive host-halo properties and therefore may predict distinct clustering signatures. In this section, we quantify these signatures. We emphasize that the discussion here focuses on where LRDs preferentially emerge in terms of their host halo properties, while the physical nature of LRDs themselves lies beyond the scope of this work.

We consider four representative scenarios, each imposing distinct constraints on the properties of their host DM haloes or environments. We briefly recap these scenarios and present the corresponding prescriptions for assigning LRDs to host haloes in Section 4.1. We then analyze the resulting clustering signals on large and small scales in Section 4.2.2. The uncertainties associated with the tagging methods are discussed in Section 4.3.

Before diving into specific models, we summarize some general statistics for reference: the LRD number density remains nearly constant from $z \simeq 4$ to $z \simeq 8$, at $\phi_{\rm LRD} \simeq 6 \times 10^{-5}\,{\rm Mpc}^{-3}$ (e.g., Inayoshi 2025; Kocevski et al. 2025; Kokorev et al. 2024). This corresponds to approximately 480 LRDs within the Shin-Uchuu simulation volume of $140^3\,h^{-3}\,{\rm Mpc}^3$. These values guide our LRD tagging strategy.

4.1 LRD formation mechanisms and tagging strategies

DCBH scenario: A possible pathway to produce the massive BHs and the weak or absent stellar components inferred for LRDs within ΛCDM framework is the formation of heavy seeds via DCBHs (e.g., Jeon et al. 2025; Cenci & Habouzit 2025; Pacucci et al. 2026), followed by sustained accretion at super-Eddington rates in the quasi-star phase (e.g., Begelman et al. 2008). In this scenario, gas collapses in rare atomic-cooling haloes and forms a massive BH seed of mass $M_{\rm BH} \sim 10^{4-6} M_\odot$. The formation conditions of a DCBH can be char-

acterized as follows: (i) the host halo must have a virial temperature $T_{\rm vir} > 10^4$ K, such that cooling proceeds primarily through atomic hydrogen; (ii) the gas must remain metal-free in order to suppress gas fragmentation led by metal-line and dust cooling; (iii) the halo must be exposed to a sufficiently strong Lyman-Werner (LW) radiation field to dissociate molecular hydrogen, since otherwise H_2 cooling would reduce the gas temperature to ~ 200 K, lower the Jeans mass, and induce fragmentation and star formation rather than monolithic collapse (Bromm & Loeb 2003; Shang et al. 2010; Omukai et al. 2008; Inayoshi et al. 2014; Regan et al. 2014; Wise et al. 2019).

Guided by these criteria, we adopt a simplified toy model to identify potential DCBH host haloes in the Shin-Uchuu simulation. We consider all haloes with a virial temperature of $T_{\rm vir} > 10^4$ K, classify a halo as pristine if it hosts no galaxy in the associated UniverseMachine catalogue, and require the total LW intensity, $J_{\rm LW} = J_{\rm loc} + J_{\rm bkg}$, to be higher than a critical value[5], $J_{\rm crit}$, where $J_{\rm loc}$ represents the contribution from nearby galaxies and $J_{\rm bkg}$ denotes the cosmological background. Following Greif & Bromm (2006) and Dijkstra et al. (2008), we approximate

$$J_{\rm bkg} = 10^{2-z/5}, \tag{11}$$

$$J_{\rm loc} = \frac{1}{4\pi}\sum_i f_{\rm es}\frac{L_{{\rm LW},i}({\rm SFR}_i)}{4\pi r_i^2}, \tag{12}$$

where the index i runs over nearby galaxies within 100 pkpc, r_i denotes their distances to the target halo, and $f_{\rm es} = 0.5$ is the escape fraction of LW photons (Agarwal et al. 2012; Dijkstra et al. 2014). Here J is expressed in units of $10^{-21}\,{\rm erg\,s^{-1}\ cm^{-2}\,Hz^{-1}\,sr^{-1}}$. For the LW luminosity of neighboring galaxies, we adopt an empirical UV-to-SFR conversion (Kennicutt 1998) and approximate the LW specific luminosity as

$$L_{\rm LW} \sim 8\times10^{27}\left(\frac{\rm SFR}{M_\odot/{\rm yr}}\right){\rm erg\,s^{-1}\,Hz^{-1}}, \tag{13}$$

where SFR denotes the star formation rate provided by the UniverseMachine catalogue. We set $J_{\rm crit} = 1000$ (Sugimura et al. 2014). To account for metal pollution driven by galactic winds, we cease DCBH seeding at redshifts below $z = 10$ (Dijkstra et al. 2014). For any halo that satisfies the above criteria and maintains $J_{\rm LW} > J_{\rm crit}$ for about at least one dynamical time (approximated by one tenth of instantaneous Hubble time, $0.1t_{\rm H}$), we classify it as a potential DCBH host. This procedure yields 4,452 candidates at $z = 5$. We caution against over-interpreting the predicted abundance, as our toy model is subject to several uncertainties, including the duty cycle of LRDs, the LW intensity estimates, external metal contamination, and the contribution from Pop III stars. To match the observed abundance of LRDs, we randomly down-sample the candidates to 10%.

SIDM scenario: LRDs are believed to have SMBHs that are overmassive relative to their host galaxies compared to local scaling relations (Matthee et al. 2024; Maiolino et al. 2024b), and some studies even argue that some LRDs may lack substantial host galaxies altogether (e.g., Chen et al. 2025a,b). This motivates scenarios in which BH seeding has a non-baryonic origin. Jiang et al. (2025b) explored the feasibility of seeding the massive BHs inferred for LRDs through gravothermal core collapse of SIDM haloes. In this framework, LRDs are expected to reside in highly concentrated haloes that form sufficiently early, which enables core collapse to proceed rapidly and complete by the epoch when LRDs are observed. More specifically, the requirement for a halo to seed a massive BH reduces to the condition that the gravothermal core-collapse timescale, $t_{\rm cc}$, be shorter than the time elapsed since halo formation, $t_{\rm form}$.

The core collapse timescale $t_{\rm cc}$ is approximated as (e.g., Essig et al. 2019)

$$t_{\rm cc} \approx \frac{150}{C}\frac{1}{\sigma_{\rm eff}\rho_{\rm s}r_{\rm s}}\frac{1}{\sqrt{4\pi G\rho_{\rm s}}}, \tag{14}$$

where $C = 0.75$ is an empirical constant calibrated using N-body simulations (e.g., Nishikawa et al. 2020), and $\rho_{\rm s}$ and $r_{\rm s}$ denote the scale density and scale radius of the NFW profile followed by the halo at its formation time. The factor $\sigma_{\rm eff}$ denotes the effective cross section per particle mass of the self-interactions. We follow Yang & Yu (2022); Yang et al. (2023) to define it as the integration of the velocity-dependent cross section over the velocity distribution of a halo, via

$$\sigma_{\rm eff} = \frac{1}{2}\int \tilde{v}^2{\rm d}\tilde{v}\sin^2\theta{\rm d}\cos\theta\frac{\sigma_0\omega^4}{2[\omega^2 + v^2\sin^2(\theta/2)]^2}\tilde{v}^5 e^{-\tilde{v}^2}, \tag{15}$$

where $\tilde{v} \equiv v/(2v_{\rm eff})$, and $v_{\rm eff}$ is a characteristic one-dimensional velocity dispersion. For an isotropic velocity field, it can be approximated by $v_{\rm eff} \approx v_{\rm max}/\sqrt{3}$, with $v_{\rm max}$ being the maximum circular velocity of the halo. For the differential cross section, we consider elastic Rutherford scattering through Yukawa potential where σ_0 is the normalization at low-velocity end, and ω is the velocity scale above which the cross section decreases sharply.

To evaluate the core-collapse time, we traverse the halo merger trees to obtain the halo mass $M_{\rm vir}$, scale radius $r_{\rm s}$, and concentration c at each redshift, and compute $\rho_{\rm s}$ from Eq. 3. We adopt a cross section of of $\sigma_0 = 10\,{\rm cm^2/g}$, $\omega = 200\,{\rm km/s}$ that are broadly consistent with constraints from the rotation curves of nearby galaxies (e.g., Jia et al. 2026), and use these parameters to derive the effective cross section $\sigma_{\rm eff}$.

However, the numerical resolution of the Shin-Uchuu simulation is just marginally sufficient for probing this scenario. Since both the core-collapse time, $t_{\rm cc}$, and the formation time, $t_{\rm form}$, depend on halo mass, there exists a characteristic halo mass where BH seeding is most efficient, $M_{\rm vir} \sim 10^8 h^{-1}M_\odot$, for the cross section we consider. This corresponds to roughly 100 DM particles in the Shin-Uchuu simulation. The marginal numerical resolution impacts the accuracy of concentration measurements (Wang et al. 2024). To ensure robustness against such uncertainties, we require the condition, $t_{\rm cc} < t_{\rm form}$, to be satisfied continuously for at least one dynamical time, $t_{\rm dyn} \sim 0.1t_{\rm H}$, before identifying a halo as having undergone core collapse and seeded a BH. This procedure yields 437 LRD candidates, primarily seeded over the redshift range $z \simeq 8-18$.

PBH scenario: Primordial black holes (PBHs), if they exist, may provide another non-baryonic origin for LRDs (e.g., Liu & Bromm 2024; Dayal 2024; Zhang et al. 2025a,b). PBHs can form from sufficiently large primordial inhomogeneities, either seeded in the initial conditions of the Universe or generated dynamically in the early Universe, and they may span a wide range of masses.

Two classes of scenarios have been discussed in the context of LRDs. In the first, PBHs are born massive, with $M_{\rm PBH} \gtrsim 10^6 M_\odot$ (e.g., Zhang et al. 2025a). In this case, existing bounds require a small PBH contribution to the total dark matter budget, $f_{\rm PBH} \lesssim 10^{-4}$ (e.g., Carr & Kühnel 2022). The low number density implies that the PBH two-point correlation function approaches a white noise plateau on comoving scales below $L_{\rm w} \sim (f_{\rm PBH}\,\rho_c/M_{\rm PBH})^{-1/3} \sim 0.1-1$ cMpc. On larger scales, once cosmic

[5] We note that additional dynamical heating from halo mergers is not considered in this work. Such heating may further enhance the efficiency of DCBH formation in overdense regions (Li et al. 2021; Lupi et al. 2021).

structure formation becomes nonlinear, PBHs are expected to trace the underlying matter clustering, ξ_{mm} (Desjacques & Riotto 2018; Luca et al. 2020). We note that, if one assumes PBHs are created from the high-σ peaks of nearly Gaussian fluctuations, constraints from second-order tensor perturbations and μ-distortions in CMB rule out such massive PBH scenario (e.g., Carr & Lidsey 1993; Chluba et al. 2012, 2021).

In the second, more complex class of scenarios, PBHs form initially at $M_{\rm PBH} \lesssim 10^2 M_\odot$ with strong small-scale clustering in the early Universe and subsequently undergo runaway mergers that produce a population of larger PBHs before Hawking evaporation becomes important (e.g., Holst et al. 2025; Zhang et al. 2025b). This scenario evades the CMB μ-distortion constraints discussed above. The characteristic mass and number density of the final merged PBHs are comparable to those in the first scenario. Therefore, for simplicity, we adopt the same large-scale clustering approximation, while neglecting additional small-scale structure imprinted by the early time clustering and merger history. Finally, PBH dynamics within DM haloes can further modify the one-halo term, so we truncate our clustering prediction on scales around and below the white noise scale $L_{\rm w}$.

Low-spin scenario: Pacucci & Loeb (2025) interpret LRDs as compact galaxies with exceptionally small effective radii ($R_{\rm eff} \sim 50 - 300$ pc), which they attribute to low halo spin following a simplified version of the model by Mo et al. (1998), in which galaxy size scales with the spin and virial radius of the host halo, $R_{\rm eff} \propto \lambda R_{\rm vir}$.

In this scenario, LRDs correspond to the lowest 1% tail of the halo spin distribution. The redshift evolution of the LRD abundance is governed by cosmological surface-brightness dimming and by the evolution of the critical density, which affects halo virial radii: at higher redshift, compact systems are more common but increasingly more difficult to detect, whereas at lower redshift they are brighter but rarer.

Guided by this picture, we use the UniverseMachine mock galaxy catalogue to select candidate LRD hosts as the lowest 1% in halo spin among haloes hosting UV-bright galaxies with $M_{\rm UV} < -17$. The UV brightness threshold follows Kocevski et al. 2025 and the practice of Pacucci & Loeb 2025. For subhaloes, we adopt the halo spin evaluated at the epoch of first infall, assuming that the size of a satellite galaxy is set by the spin and virial radius of its host halo when it was still a distinct halo. This selection yields a total of 850 objects within our simulation volume.

4.2 Clustering of LRDs

4.2.1 Measurements

We now evaluate the auto-correlation functions of LRD host haloes and normal galaxies, as well as their cross-correlation functions. Here, normal galaxies as defined as UniverseMachine galaxies with $M_{\rm UV} < -17$. The method used to measure the auto-correlation function is described in Section 3.1. The cross-correlation function is defined as

$$\xi_{12}(r) = \frac{D_1D_2(r) - D_1R_2(r) - D_2R_1(r) + R_1R_2(r)}{R_1R_2(r)}, \tag{16}$$

where $D_1D_2(r)$ is the number of cross pairs between the two data catalogues (i.e., LRDs and normal galaxies), while $D_1R_2(r)$, $D_2R_1(r)$, and $R_1R_2(r)$ denote the corresponding pair counts involving random catalogues.

To facilitate comparison with observations, we compute the projected correlation function, defined as

$$w_p(r_p) = 2 \times \int_0^{r_{\pi,\rm max}} \xi(r_p, r_\pi)\, {\rm d}\, r_\pi, \tag{17}$$

where r_p and r_π are the transverse and line-of-sight separations, respectively. We adopt $r_{\pi,\rm max} = 8\,h^{-1}$Mpc, corresponding to a velocity difference of approximately 1000 km s^{-1} at $z \simeq 4-6$, which effectively captures the clustering signal from large-scale structure.

Fig. 6 presents the projected correlation functions, where the black curves show the auto-correlation function of the normal galaxy sample, while the colored curves denote the galaxy-LRD cross-correlation functions for the three LRD-selection scenarios, as labeled. Uncertainties for the SIDM and low-spin scenarios are estimated via bootstrap resampling with replacement. For the DCBH scenario, we resample the candidate catalogue according to the down-sampling factor in each realization, and repeat this procedure 1000 times to estimate the uncertainties. For the PBH scenario, we assume that PBHs trace the matter distribution on large scales and compute the galaxy-PBH cross-correlation in the two-halo regime using $\xi_{\rm mm}(r, z)$. We evaluate $\xi_{\rm mm}$ with HALOMOD[6], adopting the halo-bias calibration of Tinker et al. (2010), the transfer function from CAMB[7], and the linear growth factor from Challinor & Lewis (2011). The upper axis indicates the corresponding angular separations at $z = 5$, converted from comoving distances. The bottom panel of Fig. 6 shows the ratio between the LRD-galaxy cross-correlation function and the galaxy auto-correlation function.

The clustering strength of the DCBH model is the strongest across most scales, followed by the SIDM and low-spin models. At large scales, the low-spin model shows slightly stronger clustering than the SIDM model, while in the one-halo regime the SIDM signal becomes stronger. The PBH model exhibits the weakest clustering signal in the two-halo regime, while truncated by hand at smaller scales.

Some observational studies perform clustering analyses of *JWST*-selected AGN samples by cross-correlating them with specific galaxy populations (Arita et al. 2024; Lin et al. 2025b). However, these galaxy samples differ from the galaxy catalogue adopted in this work and are difficult to reproduce with the UniverseMachine catalogue, which prevents a direct comparison. Nevertheless, the dependence on the chosen galaxy sample can be removed by subtracting the galaxy auto-correlation, yielding a clean estimate of the bias factor, which we compare in Fig. 7.

4.2.2 Large-scale clustering of LRDs

To interpret the results, we first quantify the clustering strength in the the two-halo term regime. Following Arita et al. (2024), we estimate the linear bias parameter b as the ratio between the clustering amplitude of a given tracer and that of the underlying DM, evaluated at a fixed comoving scale of $8\,h^{-1}$Mpc,

$$b = \sqrt{\frac{\xi(8, z)}{\xi_{\rm mm}(8, z)}}, \tag{18}$$

where $\xi(8, z)$ is obtained by fitting the real-space auto- or cross-correlation function with a power-law form over the range $r = 1-10\,h^{-1}$Mpc and evaluating the best-fit model at $8\,h^{-1}$Mpc. We estimate the linear bias parameter for the galaxy-galaxy auto-correlation function, $b_{\rm gal}$, and for the galaxy-LRD cross-correlation function,

[6] https://halomod.readthedocs.io/en/latest/index.html
[7] https://camb.readthedocs.io/en/latest/

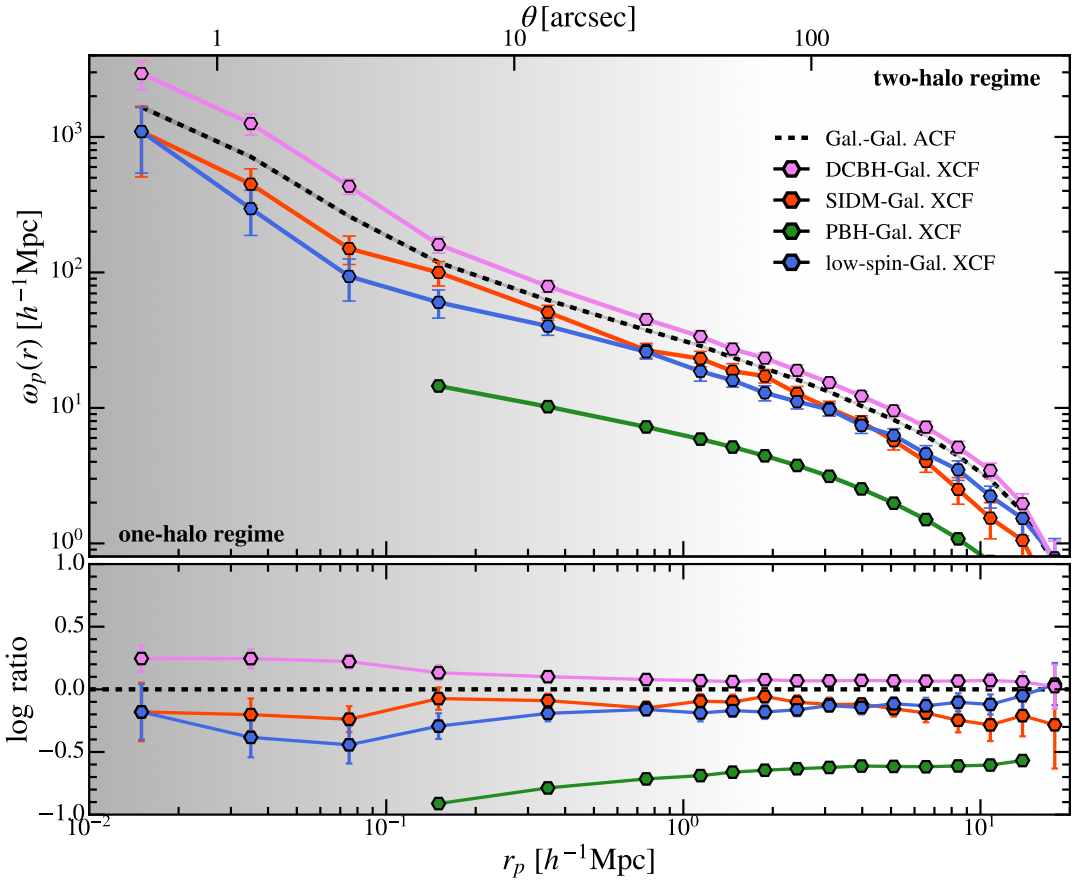

Figure 6. Projected correlation functions between galaxies with $M_{\rm UV} < -17$ and LRDs at $z = 5$. The black curve shows the auto-correlation function (ACF) of the normal galaxy sample, while colored curves denote the galaxy-LRD cross-correlation functions (XCFs) for the SIDM, DCBH, low-spin, and PBH scenarios. Shaded regions and error bars indicate uncertainties estimated from 1000 bootstrap realizations, with the DCBH case additionally resampled according to the down-sampling factor. For the PBH scenario, the two-halo term is computed assuming PBHs trace the matter distribution. The top axis shows the corresponding angular separation at $z = 5$. The lower panel displays the ratio of the LRD-galaxy cross-correlation to the galaxy auto-correlation. Among the four scenarios, the DCBH case exhibits stronger clustering than the galaxy-galaxy ACF on all scales. The low-spin case yields systematically weaker clustering across all scales. The SIDM case also shows weaker clustering on all scales compared to the galaxy-galaxy ACF, but exhibits a mild enhancement on small scales relative to the low-spin case.

$b_{\rm gL}$. Under the assumption of linear bias on these scales, the cross-correlation amplitude satisfies

$$b_{\rm LRD} \simeq b_{\rm gL}^2 / b_{\rm gal}, \tag{19}$$

which allows us to infer the bias of the LRD population, $b_{\rm LRD}$ (Mountrichas & Georgakakis 2011).

Fig. 7 presents the inferred large-scale linear bias for different LRD-host scenarios, together with their corresponding host-halo mass distributions at $z \simeq 5$. These measurements are compared with the standard halo bias-mass relation from Tinker et al. (2010), as well as with the bias-mass relations considering haloes of the highest 10% concentration, lowest 10% spin, and earliest 10% formation time.

Since the different scenarios naturally select LRD hosts with different halo-mass distributions, as illustrated in the upper panel of Fig. 7, part of the variation in clustering strength can be attributed trivially to mass differences, rather than to assembly bias specific to each scenario. To quantify this mass-driven contribution, we also compute, for each scenario, a mass-only "effective" bias, $b_{\rm eff}$, by convolving the scenario-specific mass distribution with the standard Tinker et al. 2010 bias-mass relation. This mass-only baseline provides a reference basis against which deviations can be attributed to secondary biases.

For comparison with observations, we overlay host-halo mass estimates from clustering analyses of *JWST*-selected AGN samples, including both LRD-like and unobscured populations (Arita et al. 2024; Lin et al. 2025b).

The DCBH scenario predicts the strongest clustering, with a bias of $b_{\rm LRD} \simeq 4.50$, reaching the regime of the observed of AGN samples (Lin et al. 2025b). A slight enhanced clustering relative to the standard bias-mass relation likely arises because the LW radiation criterion preferentially selects haloes in overdense environments. The resulting population spans a relatively broad halo-mass range of $M_{\rm vir} \gtrsim 10^{9.5}\, h^{-1} M_\odot$, with a median of $M_{\rm vir} \sim 10^{10.8} h^{-1} M_\odot$.

The low-spin scenario gives a bias of $b \simeq 3.16$, which is lower than the standard bias for haloes of comparable masses, driven primarily by the secondary bias associated with halo spin. The halo mass distribution is relatively narrow and massive, peaked around $M_{\rm vir} \sim 10^{10.6} h^{-1} M_\odot$.

The SIDM scenario likewise yields a weak clustering signal, with $b_{\rm LRD} \simeq 2.92$. However, this is primarily driven by its host haloes spanning a broad but systematically lower mass distribution. The halo mass at the time of core-collapse is typically $M_{\rm vir} \sim 10^{8-9} h^{-1} M_\odot$, after which the systems can continue to grow to higher masses.

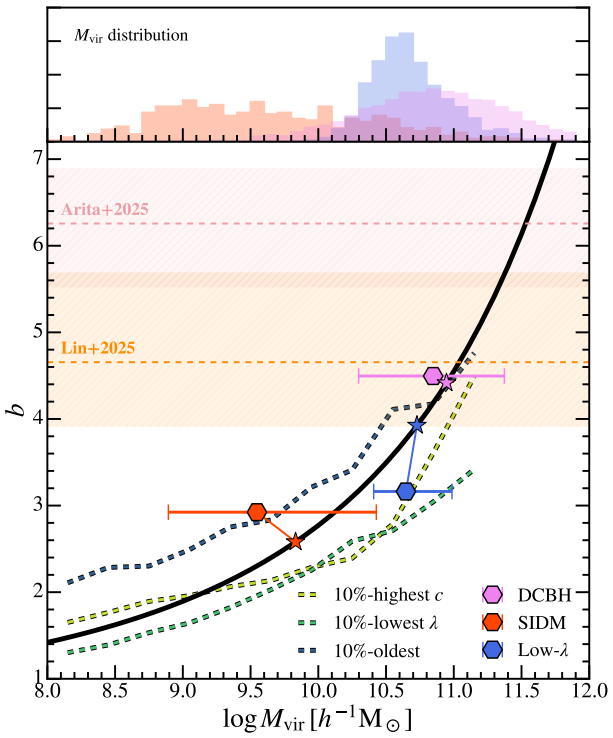

Figure 7. Large-scale linear bias inferred for LRD hosts under different formation scenarios observed at $z \simeq 5$. The main panel shows the measured bias b as a function of host halo virial mass, for the DCBH, SIDM, and low-spin scenarios, with symbols indicating the median and horizontal bars denoting the 1σ scatter of the corresponding host-halo mass distributions. The solid black curve shows the standard halo bias-mass relation from Tinker et al. (2010), while the dotted curves represent the bias-mass relations for haloes of the highest 10% concentration, lowest 10% spin, and earliest 10% formation time, as labeled. For each scenario, the star marker denotes the effective bias $b_{\rm eff}$ expected from halo-mass selection alone, computed by convolving the host-halo mass distribution with the standard halo bias relation. Shaded vertical bands indicate host-halo mass ranges inferred from clustering analyses of *JWST*-selected AGN samples (e.g., Lin et al. 2025b; Arita et al. 2024). The upper panel shows the normalized host-halo mass distributions for each scenario. Differences in the inferred bias reflect the combined effects of halo mass and secondary bias. The DCBH scenario yields relatively high bias of $b \gtrsim 4$. The low-spin and SIDM scenarios produce similar biases of $b \sim 3$, but correspond to rather different halo mass ranges.

For the PBH scenario, assuming that PBHs trace the underlying matter distribution on large scales and that they become LRDs stochastically, the expected large-scale bias is $b_{\rm LRD} = 1$ by definition. The halo-mass distribution depends on theoretical details such as the initial PBH mass spectrum and lies beyond the scope of this work.

4.2.3 Small-scale clustering and satellite fraction

An important observational feature of LRDs emerging from preliminary clustering analyses is their enhanced small-scale clustering (e.g., Zhuang et al. 2025), a high incidence of offset companion[8] (e.g., Rinaldi et al. 2025; Baggen et al. 2026; Chen et al. 2025b), and a significant pair fraction (e.g., Tanaka et al. 2024b; Mérida et al. 2025). Notably, four dual-LRD systems identified from photometric data have been reported (Tanaka et al. 2024b; Mérida et al. 2025) among a total number of about ~ 400, with additional candidates likely unresolved. This suggests that a non-negligible fraction of LRDs may currently reside in subhaloes or have been brought into close proximity to another LRD through subhalo dynamics. These findings motivates a closer examination of the small-scale clustering of LRDs and an assessment of whether the observed pair statistics can be reproduced under different formation scenarios.

[8] While companions of LRDs have been commonly reported, their nature remains uncertain. Some resolved systems exhibit clear galaxy-like features (e.g., Baggen et al. 2026), whereas in other cases the emission appears to arise from nebular gas associated with the LRD itself (e.g., Chen et al. 2025b).

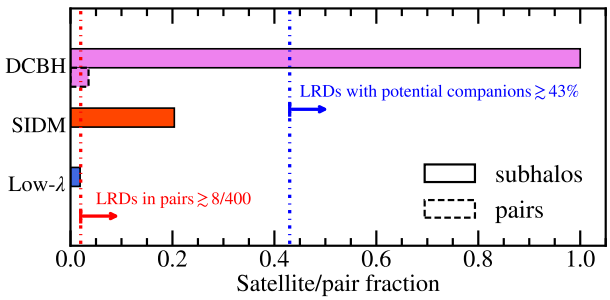

Figure 8. Fraction of LRDs associated with subhaloes and close pairs, predicted by the different formation scenarios. For comparison, we adopt the fraction of LRDs exhibiting spatially offset UV emission from Baggen et al. (2026), which could be suggestive of companion galaxies. We also compile close-pair statistics from Tanaka et al. (2024b); Mérida et al. (2025). The DCBH scenario, by construction, places LRDs around massive centrals and therefore produces a higher fraction of close pairs, while the SIDM model, owing to its lower halo mass range, predicts a modest satellite fraction.

As already shown in Fig. 6, the DCBH scenario predicts the strongest small-scale clustering relative to normal galaxies on scales of 10-100h^{-1} kpc, followed by the SIDM and low-spin scenarios.

At even smaller spatial scales, where pairs and close companions become relevant, subhalo statistics provide useful complement to the clustering analysis. One notable trend from the subhalo mass function is that the abundance of satellites decreases dramatically with increasing subhalo mass. In this sense, the abundance of satellite LRD hosts is ~ 3 dex higher in haloes with masses of order $10^{8-9}\,h^{-1}M_\odot$ than in haloes with masses near $10^{11}\,h^{-1}M_\odot$. Indeed, low-mass haloes of order $10^{8-9}\,h^{-1}M_\odot$ are also more naturally consistent with the lack of prominent stellar components in LRDs, since haloes with masses of $\sim 10^{11}\,h^{-1}M_\odot$ are expected to host substantial and spatially extended stellar components at these epochs. In most cosmological hydrodynamical simulations capable of resolving systems with $M_{\rm vir} \sim 10^{11}\,h^{-1}M_\odot$ systems, galaxy sizes are generally of order kpc (e.g., Shen et al. 2024; McClymont et al. 2025).

To this end, we further investigate the predicted abundances of satellite LRDs and close LRD pairs under the different scenarios. In Fig. 8, we show the fraction of satellite LRDs and that of LRDs in pairs. To estimate the pair fraction, we randomly draw 480 LRDs from each scenario, or use the full sample when fewer are available, and repeat this procedure 1000 times, counting the number of systems that satisfy the pair criteria in each realization. Here, a satellite LRD is defined as an LRD associated with a subhalo, either at the observed epoch or at any earlier time prior to subhalo disruption. An LRD pair is defined as two (or more) LRDs that satisfy at least one of the following criteria: (i) their host haloes have undergone a merger; (ii) their hosting subhaloes reside within the same host halo; or (iii) their projected separation at the observed epoch is smaller than 5 pkpc.

The results are further compared with compiled observations (e.g., Tanaka et al. 2024b; Mérida et al. 2025; Baggen et al. 2026)

The DCBH scenario most efficiently produces LRD satellites and pairs. This is, to some extent, by construction, since DCBH hosts preferentially form as satellites of UV-bright central galaxies, allowing multiple LRDs to arise within the same halo.

The SIDM scenario yields the second-highest satellite abundance, primarily because it preferentially selects lower-mass host haloes, which have a higher probability of being accreted as subhaloes. However, no resolved LRD pairs are found in this case, likely because high-concentration haloes at the relevant masses exhibit relatively weak clustering at the seeding epoch.

The low-spin scenario produces the lowest number of satellite LRDs and likewise yields no LRD pairs.

4.3 Evaluation and limitation of the results

It is important to clarify some uncertainties in our results, which mainly arise from the tagging criteria adopted for LRD hosts.

For the DCBH scenario, our implementation only captures DCBH formation as a channel linked to close halo pairs (e.g., Dijkstra et al. 2008; Visbal et al. 2014), while alternative pathways, such as DCBH formation following prior first star formation within the same halo (e.g., Susa 2007), are not considered. The inferred host halo mass distribution does not directly reflect the subhalo masses at which DCBH seeds initially form. The majority of these seed-hosting subhaloes are fully disrupted by the observed epoch, such that the merger trees point to their surviving distinct host haloes. This behavior is a natural outcome of hierarchical structure formation rather than a numerical artifact. Instead, it highlights that DCBH formation is preferentially associated with massive central haloes that can both provide sufficiently strong LW radiation fields and host intermediate-mass atomic-cooling subhaloes with $T_{\rm vir} > 10^4$ K. As a result, the large-scale clustering of DCBHs primarily traces the clustering of their central host haloes.

Into the one-halo term, the enhanced clustering signal also depends sensitively on the treatment of the LW radiation field and the assumed critical intensity $J_{\rm crit}$. Previous studies have shown that adopting a more stringent LW threshold can substantially amplify the small-scale clustering signal, with secondary effects that also propagate to larger scales (e.g., Agarwal et al. 2012). In addition, external metal enrichment from neighboring galaxies and Pop-III stars is not modeled here, which would further suppress the number of viable DCBH hosts and likely weaken the small-scale clustering (e.g. Dijkstra et al. 2014).

In the SIDM scenario, first, we note that the choice of the self-interaction cross section does not qualitatively affect our conclusions, although it does influence the characteristic halo mass range over which core collapse and BH seeding occur (Jiang et al. 2025b). Specifically, lower σ_0 and higher ω shift the characteristic halo mass for BH seeding toward higher values. The parameters adopted here are chosen such that core collapse can occur in haloes resolved in the simulation, with $M_{\rm vir} \gtrsim 10^8 h^{-1} M_\odot$. Even at this cross section, we can explicitly identify seed hosts only above this mass, while a non-negligible fraction of BHs may form in lower-mass, unresolved haloes. The inferred host-halo mass distribution should therefore be regarded as incomplete and missing a lower-mass tail. If lower-mass hosts were included, the true large-scale clustering of SIDM-seeded LRDs would likely be even weaker, although this could be partially offset by a higher satellite fraction and an increased incidence of close pairs.

Second, our analysis assumes that SIDM does not significantly modify the formation of large-scale structure or the overall concentration distribution, and instead adopts concentration measurements from the CDM simulations. This assumption is supported by both physical considerations and existing numerical results, as SIDM effects become pronounced primarily within radii $r \lesssim r_{\rm s}$ (Yang et al. 2024; Rocha et al. 2013). On small scales, however, recent simulations (e.g., Romanello et al. 2025) have shown that SIDM can suppress the one-halo term of the clustering signal, which may reduce the small-scale clustering amplitude relative to our predictions.

Overall, we can safely conclude that the SIDM scenario features the lowest halo-mass distributions among the different scenarios considered here. Interestingly, the secondary bias associated with early formation and high concentration largely cancels out.

For the low-spin scenario, the reduced clustering relative to the standard mass-based bias is a direct consequence of the secondary bias associated with halo spin. This qualitative behavior is robust and largely insensitive to the detailed host halo mass distribution.

In addition, halo spin may be physically connected to DCBH formation, although its role remains debated. On the one hand, some studies suggest that low angular momentum more easily avoids centrifugal support and facilitates large-scale inflows, allowing a larger fraction of the gas to collapse into a DCBH and subsequently grow (e.g., Lodato & Natarajan 2006; Agarwal et al. 2012; Suazo et al. 2019; Bhowmick et al. 2021). On the other hand, high angular momentum may also accelerate core collapse through the so-called "gravo-gyro" instability (e.g., Hachisu 1979; Dekel et al. 2025). Other studies, however, argue that angular momentum is not a critical factor in the collapse process of DCBH-forming haloes (e.g., Chon et al. 2016; Mone et al. 2025), partly because the angular momentum of the gas does not necessarily trace that of the dark matter halo, especially on small scales where the dynamics of the nuclear disc dominate (Dubois et al. 2012; Bonoli et al. 2014; Danovich et al. 2015). Regardless of the exact mechanism, if DCBH formation preferentially occurs in haloes with certain spin properties, the resulting population would inherit the clustering signal associated with halo spin bias.

Finally, we note that the discussion above is based on the assumption that LRDs arise from a single formation channel. In reality, LRDs may originate from multiple channels and therefore exhibit a superposition of different clustering signatures. In addition, uncertainties related to the subsequent dynamics of host haloes and BH seeds, including tidal disruption of subhaloes and wandering BHs, are not considered here. Nevertheless, our results still provide useful constraints on the relative contributions of these channels and offer a framework for interpreting future clustering measurements.

5 CONCLUSIONS

In this work, we provide the first systematic characterization of halo assembly bias in the early Universe, examining its dependence on formation time, concentration, and spin across redshifts, using the Shin-Uchuu simulation. We then explore how secondary bias can be leveraged in the *JWST* era, using LRDs as a case study to predict their clustering signatures and host-halo mass ranges under different formation scenarios. Our main results are summarized below.

- At all redshifts, earlier-forming haloes are more strongly clustered than later-forming haloes at fixed mass. High-concentration haloes show enhanced clustering at the low-mass end, whereas this trend reverses toward higher masses. Low-spin haloes show reduced clustering across the full mass range at $z \gtrsim 4$. Toward intermediate and lower redshifts, the clustering signal of low-spin haloes increases progressively at the low-mass end.

- There are clear sign reversals in both concentration-related and spin-related bias signals, and the critical mass at which the assembly-bias signal flips sign depends on redshift. For concentration, the characteristic crossover mass above which high-concentration haloes become less clustered broadly follows a peak height of $\nu \simeq 1.5$. Notably, in contrast to the local Universe, this implies that the majority of galaxies associated with high-concentration haloes at high redshift are expected to exhibit weaker clustering. For spin, the crossover mass below which low-spin haloes become more strongly clustered occurs near $\nu \simeq 0.75$, but this only holds for $z \lesssim 3$. At higher redshifts, low-spin haloes are always less clustered, regardless of mass.
- Different proposed LRD formation channels populate systematically different halo-mass regimes, which strongly influences their clustering signatures. The DCBH scenario selects relatively massive haloes ($M_{\rm vir} \sim 10^{10.8} h^{-1} M_\odot$) and predicts the strongest large-scale bias ($b \gtrsim 4$). The low-spin scenario occupies similarly massive haloes but exhibits reduced bias ($b \sim 3$) due to spin-related secondary bias. The SIDM scenario favors substantially lower characteristic halo masses, peaking at $M_{\rm vir} \sim 10^{9-9.5} h^{-1} M_\odot$ and exhibits the broadest halo mass distribution, extending down to the lowest masses resolved in the simulation ($M_{\rm vir} \gtrsim 10^8 h^{-1} M_\odot$). Its large-scale clustering is correspondingly weak and overall comparable to that of the low-spin scenario. However, this arises from a combination of effects, namely, the lower halo masses and the partial cancellation of secondary bias associated with early formation and high concentration. In the PBH scenario, where seeds trace the matter distribution by construction, no large-scale bias is expected. These differences imply that host-halo mass distributions constitute a primary discriminant among LRD formation models. Despite the halo-mass differences, assembly bias associated with secondary halo properties still plays a non-negligible role in interpreting clustering-based halo masses.
- The different scenarios make distinct predictions for satellite fractions and close LRD pairs. The DCBH scenario most naturally produces multiple LRDs within the same massive environment and hence predicts enhanced pair counts. The SIDM scenario, despite exhibiting weaker large-scale clustering, can yield a relatively high satellite fraction because it preferentially selects lower-mass haloes, which are more likely to be accreted as satellites than more massive systems. The low-spin scenario predicts the lowest satellite abundance. These small-scale signatures offer an additional observational avenue for distinguishing formation models.

Overall, we conclude that halo assembly bias persists into the early Universe, albeit with a reduced amplitude relative to later times. This underscores the importance of accounting for secondary bias when inferring halo masses of high-z galaxies from clustering measurements. Neglecting assembly bias may lead to systematic offsets in halo mass estimates of up to 1 dex.

For LRDs, our study predicts that different formation scenarios lead to distinct clustering signatures and host-halo mass ranges. These predictions can be tested with upcoming clustering measurements. If the cross-correlation between LRDs and normal galaxies with $M_{\rm UV} < -17$ on large spatial scales ($\gtrsim 1\,h^{-1}$ Mpc) is stronger than the auto-correlation of the normal galaxy population, the DCBH scenario would be favored. If instead the cross-correlation is weaker, the SIDM and low-spin scenarios would be preferred.

These two scenarios can be further distinguished using small-scale clustering measurements and satellite/pair statistics, as the SIDM model predicts stronger small-scale clustering and a higher satellite fraction than the low-spin model. We emphasize that the broader and generally lower halo mass range, which extends to values comparable to or only slightly above the atomic-cooling threshold, is a distinctive feature of the SIDM model, making it particularly appealing given that a substantial fraction of LRDs seem to lack prominent galaxy components.

Finally, if the cross-correlation amplitude is an order of magnitude lower than the auto-correlation of normal galaxies, the PBH scenario may be favored. We leave it to forthcoming observational studies to discriminate among these possibilities.

ACKNOWLEDGEMENTS

We thank Liang Gao, Qiao Wang, Huangyu Xiao, and Wei-Xiang Feng for useful discussions. FJ acknowledges support by the National Natural Science Foundation of China (NSFC, Grant No. 12473007), China Manned Space Program with Grant No. CMS-CSST-2025-A03, National Key R&D Program of China (Grant No. 2025YFA1614103). ZJ acknowledges support by the Beijing Natural Science Foundation (Grant No. QY23018). LCH was supported by the China Manned Space Program (CMS-CSST-2025-A09) and the National Science Foundation of China (12233001). KI was supported by the National Natural Science Foundation of China (12573015, W2532003, 1251101148, 12233001, 12473037), the Beijing Natural Science Foundation (IS25003), and the China Manned Space Program (CMS-CSST- 2025-A09). We acknowledge the High-performance Computing Platform of Peking University for providing computational resources and support.

DATA AVAILABILITY

The analysis pipeline and data products are available upon requests to the corresponding authors.

REFERENCES

Agarwal B., Khochfar S., Johnson J. L., Neistein E., Dalla Vecchia C., Livio M., 2012, MNRAS, 425, 2854
Akins H. B., et al., 2024, COSMOS-Web: The over-abundance and physical nature of "little red dots"–Implications for early galaxy and SMBH assembly (arXiv:2406.10341), https://arxiv.org/abs/2406.10341
Arita J., Kashikawa N., Onoue M., Yoshioka T., Takeda Y., Hoshi H., Shimizu S., 2024, The nature of low-luminosity AGNs discovered by JWST based on clustering analysis: progenitors of low-z quasars? (arXiv:2410.08707), https://arxiv.org/abs/2410.08707
Aung H., et al., 2023, MNRAS, 519, 1648
Baggen J. F. W., Scoggins M. T., van Dokkum P., Haiman Z., Torralba A., Matthee J., 2026, arXiv e-prints, p. arXiv:2602.02702
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Begelman M. C., Rossi E. M., Armitage P. J., 2008, MNRAS, 387, 1649
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2012a, The Astrophysical Journal, 762, 109
Behroozi P. S., Wechsler R. H., Wu H.-Y., Busha M. T., Klypin A. A., Primack J. R., 2012b, The Astrophysical Journal, 763, 18
Behroozi P., Wechsler R. H., Hearin A. P., Conroy C., 2019, Monthly Notices of the Royal Astronomical Society, 488, 3143–3194
Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, MNRAS, 376, 215
Bhowmick A. K., Blecha L., Torrey P., Kelley L. Z., Vogelsberger M., Nelson D., Weinberger R., Hernquist L., 2021, Monthly Notices of the Royal Astronomical Society, 510, 177–196
Bonoli S., Mayer L., Callegari S., 2014, MNRAS, 437, 1576
Borzyszkowski M., Porciani C., Romano-Díaz E., Garaldi E., 2017, Monthly Notices of the Royal Astronomical Society, 469
Bromm V., Loeb A., 2003, ApJ, 596, 34

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559

Carr B., Kühnel F., 2022, SciPost Physics Lecture Notes

Carr B. J., Lidsey J. E., 1993, Phys. Rev. D, 48, 543

Cenci E., Habouzit M., 2025, Monthly Notices of the Royal Astronomical Society, 542, 2597–2609

Challinor A., Lewis A., 2011, Physical Review D, 84

Chen C.-H., Ho L. C., Li R., Zhuang M.-Y., 2025a, The Host Galaxy (If Any) of the Little Red Dots (arXiv:2411.04446), https://arxiv.org/abs/2411.04446

Chen C.-H., Ho L. C., Li R., Inayoshi K., 2025b, The Physical Nature of the Off-centered Extended Emission Associated with the Little Red Dots (arXiv:2505.03183), https://arxiv.org/abs/2505.03183

Chluba J., Erickcek A. L., Ben-Dayan I., 2012, ApJ, 758, 76

Chluba J., et al., 2021, Experimental Astronomy, 51, 1515

Chon S., Hirano S., Hosokawa T., Yoshida N., 2016, The Astrophysical Journal, 832, 134

Contreras S., Zehavi I., Padilla N., Baugh C. M., Jiménez E., Lacerna I., 2019, Monthly Notices of the Royal Astronomical Society, 484, 1133–1148

Contreras S., Chaves-Montero J., Zennaro M., Angulo R. E., 2021, Monthly Notices of the Royal Astronomical Society, 507, 3412–3422

Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12

Danovich M., Dekel A., Hahn O., Ceverino D., Primack J., 2015, Monthly Notices of the Royal Astronomical Society, 449, 2087–2111

Dayal P., 2024, Astronomy & Astrophysics, 690, A182

Dekel A., Stone N. C., Chowdhury D. D., Gilbaum S., Li Z., Mandelker N., van den Bosch F. C., 2025, Astronomy & Astrophysics, 695, A97

Desjacques V., Riotto A., 2018, Phys. Rev. D, 98, 123533

Dijkstra M., Haiman Z., Mesinger A., Wyithe J. S. B., 2008, Monthly Notices of the Royal Astronomical Society, 391, 1961–1972

Dijkstra M., Ferrara A., Mesinger A., 2014, Monthly Notices of the Royal Astronomical Society, 442, 2036–2047

Dong-Páez C. A., et al., 2024, MNRAS, 528, 7236

Dubois Y., Pichon C., Haehnelt M., Kimm T., Slyz A., Devriendt J., Pogosyan D., 2012, Monthly Notices of the Royal Astronomical Society, 423, 3616–3630

Dutton A. A., Macciò A. V., 2014, MNRAS, 441, 3359

Essig R., McDermott S. D., Yu H.-B., Zhong Y.-M., 2019, Physical Review Letters, 123

Faltenbacher A., White S. D. M., 2009, The Astrophysical Journal, 708, 469–473

Feng W.-X., Yu H.-B., Zhong Y.-M., 2025, Dark Bondi Accretion Aided by Baryons and the Origin of JWST Little Red Dots (arXiv:2506.17641), https://arxiv.org/abs/2506.17641

Fujimoto S., et al., 2025, Primordial Rotating Disk Composed of ≥15 Dense Star-Forming Clumps at Cosmic Dawn (arXiv:2402.18543), https://arxiv.org/abs/2402.18543

Furtak L. J., et al., 2024, A high black hole to host mass ratio in a lensed AGN in the early Universe (arXiv:2308.05735), https://arxiv.org/abs/2308.05735

Gao L., White S. D. M., 2007, MNRAS, 377, L5

Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66

Gloudemans A. J., Duncan K. J., Eilers A.-C., Farina E. P., Harikane Y., Inayoshi K., Lambrides E., Vardoulaki E., 2025, The Astrophysical Journal, 986, 130

Greene J. E., et al., 2024, ApJ, 964, 39

Greif T. H., Bromm V., 2006, MNRAS, 373, 128

Hachisu I., 1979, PASJ, 31, 523

Hahn O., Porciani C., Dekel A., Carollo C. M., 2009, MNRAS, 398, 1742

Hainline K. N., et al., 2024, An Investigation Into The Selection and Colors of Little Red Dots and Active Galactic Nuclei (arXiv:2410.00100), https://arxiv.org/abs/2410.00100

Holst I., Krnjaic G., Xiao H., 2025, Phys. Rev. D, 112, 083527

Inayoshi K., 2025, Little Red Dots as the Very First Activity of Black Hole Growth (arXiv:2503.05537), https://arxiv.org/abs/2503.05537

Inayoshi K., Omukai K., Tasker E., 2014, MNRAS, 445, L109

Inayoshi K., Murase K., Kashiyama K., 2025, Spectral Uniformity of Little Red Dots: A Natural Outcome of Coevolving Seed Black Holes and Nascent Starbursts (arXiv:2509.19422), https://arxiv.org/abs/2509.19422

Ishiyama T., et al., 2021, Monthly Notices of the Royal Astronomical Society, 506, 4210–4231

Jeon J., et al., 2025, arXiv e-prints, p. arXiv:2508.14155

Jia Z., Jiang F., Li S., Li R., Wang J., Zhu L., 2026, An Enhanced Isothermal Jeans Approach to Constraining Dark Matter Self-Interactions from Galactic Kinematics (arXiv:2601.17118), https://arxiv.org/abs/2601.17118

Jiang F., et al., 2025a, Formation and Environmental Context of Giant Bulgeless Disk Galaxies in the Early Universe: Insights from Cosmological Simulations (arXiv:2504.01070), https://arxiv.org/abs/2504.01070

Jiang F., Jia Z., Zheng H., Ho L. C., Inayoshi K., Shen X., Vogelsberger M., Feng W.-X., 2025b, Formation of the Little Red Dots from the Core-collapse of Self-interacting Dark Matter Halos (arXiv:2503.23710), https://arxiv.org/abs/2503.23710

Kaiser N., 1984, ApJ, 284, L9

Kennicutt Jr. R. C., 1998, ARA&A, 36, 189

Kido D., Ioka K., Hotokezaka K., Inayoshi K., Irwin C. M., 2025, Black Hole Envelopes in Little Red Dots (arXiv:2505.06965), https://arxiv.org/abs/2505.06965

Kocevski D. D., et al., 2025, The Rise of Faint, Red AGN at $z > 4$: A Sample of Little Red Dots in the JWST Extragalactic Legacy Fields (arXiv:2404.03576), https://arxiv.org/abs/2404.03576

Kokorev V., et al., 2024, A Census of Photometrically Selected Little Red Dots at 4 < z < 9 in JWST Blank Fields (arXiv:2401.09981), https://arxiv.org/abs/2401.09981

Kormendy J., Ho L. C., 2013, Annual Review of Astronomy and Astrophysics, 51, 511–653

Labbe I., et al., 2023, UNCOVER: Candidate Red Active Galactic Nuclei at 3<z<7 with JWST and ALMA (arXiv:2306.07320), https://arxiv.org/abs/2306.07320

Lacerna I., Padilla N., 2012, MNRAS, 426, L26

Li Y., Mo H. J., Gao L., 2008, MNRAS, 389, 1419

Li W., Inayoshi K., Qiu Y., 2021, ApJ, 917, 60

Lin X., et al., 2025a, The Discovery of Little Red Dots in the Local Universe: Signatures of Cool Gas Envelopes (arXiv:2507.10659), https://arxiv.org/abs/2507.10659

Lin X., et al., 2025b, The Large-scale Environments of Low-luminosity AGNs at $3.9 < z < 6$ and Implications for Their Host Dark Matter Halos from a Complete NIRCam Grism Redshift Survey (arXiv:2505.02896), https://arxiv.org/abs/2505.02896

Liu B., Bromm V., 2024, Impact of primordial black holes on the formation of the first stars and galaxies (arXiv:2312.04085), https://arxiv.org/abs/2312.04085

Lodato G., Natarajan P., 2006, MNRAS, 371, 1813

Luca V. D., Desjacques V., Franciolini G., Riotto A., 2020, Journal of Cosmology and Astroparticle Physics, 2020, 028–028

Ludlow A. D., Navarro J. F., Li M., Angulo R. E., Boylan-Kolchin M., Bett P. E., 2012, MNRAS, 427, 1322

Lupi A., Haiman Z., Volonteri M., 2021, MNRAS, 503, 5046

Ma Q., Li C., Chen Y., Mo H., 2025, Interpreting the strong clustering of ultra-diffuse galaxies by halo spin bias (arXiv:2512.17742), https://arxiv.org/abs/2512.17742

Maiolino R., et al., 2024a, JWST meets Chandra: a large population of Compton thick, feedback-free, and intrinsically X-ray weak AGN, with a sprinkle of SNe (arXiv:2405.00504), https://arxiv.org/abs/2405.00504

Maiolino R., et al., 2024b, Astronomy & Astrophysics, 691, A145

Mansfield P., Kravtsov A. V., 2020, Monthly Notices of the Royal Astronomical Society, 493, 4763–4782

Matthee J., et al., 2024, ApJ, 963, 129

McClymont W., et al., 2025, Monthly Notices of the Royal Astronomical Society, 544, 1732–1747

Mo H. J., Mao S., White S. D. M., 1998, Monthly Notices of the Royal Astronomical Society, 295, 319–336

Mone E., Pries B., Wise J., Ferrans S., 2025, Beyond the Goldilocks Zone: Identifying Critical Features in Massive Black Hole Formation (arXiv:2412.08829), https://arxiv.org/abs/2412.08829
Mountrichas G., Georgakakis A., 2011, Monthly Notices of the Royal Astronomical Society, 420, 514–525
Mowla L., et al., 2024, The Firefly Sparkle: The Earliest Stages of the Assembly of A Milky Way-type Galaxy in a 600 Myr Old Universe (arXiv:2402.08696), https://arxiv.org/abs/2402.08696
Mérida R. M., et al., 2025, Astronomy & Astrophysics, 698, A317
Naidu R. P., et al., 2025, A "Black Hole Star" Reveals the Remarkable Gas-Enshrouded Hearts of the Little Red Dots (arXiv:2503.16596), https://arxiv.org/abs/2503.16596
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Nishikawa H., Boddy K. K., Kaplinghat M., 2020, Physical Review D, 101
Omukai K., Schneider R., Haiman Z., 2008, The Astrophysical Journal, 686, 801–814
Oogi T., et al., 2023, MNRAS, 525, 3879
Pacucci F., Loeb A., 2025, The Astrophysical Journal Letters, 989, L19
Pacucci F., Ferrara A., Kocevski D. D., 2026, The Little Red Dots Are Direct Collapse Black Holes (arXiv:2601.14368), https://arxiv.org/abs/2601.14368
Paranjape A., Hahn O., Sheth R. K., 2018, Monthly Notices of the Royal Astronomical Society, 476, 3631–3647
Prada F., Klypin A. A., Cuesta A. J., Betancort-Rijo J. E., Primack J., 2012, MNRAS, 423, 3018
Prada F., Behroozi P., Ishiyama T., Klypin A., Pérez E., 2023, arXiv e-prints, p. arXiv:2304.11911
Press W. H., Schechter P., 1974, ApJ, 187, 425
Pérez-González P. G., et al., 2024, What is the nature of Little Red Dots and what is not, MIRI SMILES edition (arXiv:2401.08782), https://arxiv.org/abs/2401.08782
Regan J. A., Johansson P. H., Wise J. H., 2014, The Astrophysical Journal, 795, 137
Reines A. E., Volonteri M., 2016, in American Astronomical Society Meeting Abstracts #227. p. 119.01
Rinaldi P., et al., 2025, Not Just a Dot: the complex UV morphology and underlying properties of Little Red Dots (arXiv:2411.14383)
Rocha M., Peter A. H. G., Bullock J. S., Kaplinghat M., Garrison-Kimmel S., Oñorbe J., Moustakas L. A., 2013, Monthly Notices of the Royal Astronomical Society, 430, 81–104
Romanello M., Despali G., Marulli F., Giocoli C., Moscardini L., 2025, The AIDA-TNG project: abundance, radial distribution, and clustering properties of halos in alternative dark matter models (arXiv:2512.15883), https://arxiv.org/abs/2512.15883
Salcedo A. N., Maller A. H., Berlind A. A., Sinha M., McBride C. K., Behroozi P. S., Wechsler R. H., Weinberg D. H., 2018, Monthly Notices of the Royal Astronomical Society, 475, 4411–4423
Sato-Polito G., Montero-Dorta A. D., Abramo L. R., Prada F., Klypin A., 2019, MNRAS, 487, 1570
Shang C., Bryan G. L., Haiman Z., 2010, Monthly Notices of the Royal Astronomical Society, 402, 1249–1262
Shen X., et al., 2024, The THESAN project: galaxy sizes during the epoch of reionization (arXiv:2402.08717), https://arxiv.org/abs/2402.08717
Shen T., Shen X., Xiao H., Vogelsberger M., Jiang F., 2025, Massive Black Holes Seeded by Dark Matter – Implications for Little Red Dots and Gravitational Wave Signatures (arXiv:2504.00075), https://arxiv.org/abs/2504.00075
Sheth R. K., Tormen G., 1999, Monthly Notices of the Royal Astronomical Society, 308, 119–126
Sheth R. K., Mo H. J., Tormen G., 2001, Monthly Notices of the Royal Astronomical Society, 323, 1–12
Sinha M., Garrison L. H., 2019, Monthly Notices of the Royal Astronomical Society, 491, 3022–3041
Suazo M., Prieto J., Escala A., Schleicher D. R. G., 2019, The Astrophysical Journal, 885, 127
Sugimura K., Omukai K., Inoue A. K., 2014, MNRAS, 445, 544
Susa H., 2007, The Astrophysical Journal, 659, 908–917
Tanaka T. S., et al., 2024a, Crimson Behemoth: a Massive Clumpy Structure Hosting a Dusty AGN at z = 4.91 (arXiv:2410.00104), https://arxiv.org/abs/2410.00104
Tanaka T. S., et al., 2024b, Discovery of dual "little red dots" indicates excess clustering on kilo-parsec scales (arXiv:2412.14246), https://arxiv.org/abs/2412.14246
Taylor A. J., et al., 2025, Broad-Line AGN at 3.5<z<6: The Black Hole Mass Function and a Connection with Little Red Dots (arXiv:2409.06772), https://arxiv.org/abs/2409.06772
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, The Astrophysical Journal, 724, 878–886
Tucci B., Montero-Dorta A. D., Abramo L. R., Sato-Polito G., Artale M. C., 2021, MNRAS, 500, 2777
Umehata H., et al., 2025, ADF22-WEB: A giant barred spiral starburst galaxy in the z = 3.1 SSA22 protocluster core (arXiv:2410.22155), https://arxiv.org/abs/2410.22155
Vanzella E., et al., 2026, Astronomy & Astrophysics, 705, A171
Visbal E., Haiman Z., Bryan G. L., 2014, Monthly Notices of the Royal Astronomical Society, 445, 1056–1063
Wang K., Mo H. J., Chen Y., Schaye J., 2024, MNRAS, 527, 10760
Wang W., et al., 2025a, Nature Astronomy, 9, 710–719
Wang T., et al., 2025b, The Astrophysical Journal Letters, 988, L35
Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
White S. D. M., Frenk C. S., 1991, ApJ, 379, 52
Wise J. H., Regan J. A., O'Shea B. W., Norman M. L., Downes T. P., Xu H., 2019, Nature, 566, 85–88
Yang D., Yu H.-B., 2022, Journal of Cosmology and Astroparticle Physics, 2022, 077
Yang S., Du X., Zeng Z. C., Benson A., Jiang F., Nadler E. O., Peter A. H. G., 2023, The Astrophysical Journal, 946, 47
Yang S., et al., 2024, A quantitative comparison between velocity dependent SIDM cross sections constrained by the gravothermal and isothermal models (arXiv:2305.05067), https://arxiv.org/abs/2305.05067
Yue M., Eilers A.-C., Ananna T. T., Panagiotou C., Kara E., Miyaji T., 2024, Stacking X-ray Observations of "Little Red Dots": Implications for their AGN Properties (arXiv:2404.13290), https://arxiv.org/abs/2404.13290
Yung L. Y. A., Somerville R. S., Nguyen T., Behroozi P., Modi C., Gardner J. P., 2024, Characterising ultra-high-redshift dark matter halo demographics and assembly histories with the GUREFT simulations (arXiv:2309.14408), https://arxiv.org/abs/2309.14408
Zhang B., Feng W.-X., An H., 2025b, Little Red Dots from Small-Scale Primordial Black Hole Clustering (arXiv:2507.07171), https://arxiv.org/abs/2507.07171
Zhang S., Liu B., Bromm V., Kühnel F., 2025a, Primordial Black Holes as Seeds for Extremely Overmassive AGN Observed by JWST (arXiv:2512.14066), https://arxiv.org/abs/2512.14066
Zhang Z., Chen Y., Rong Y., Wang H., Mo H., Luo X., Li H., 2025c, Nature, 642, 47–52
Zhuang M.-Y., Li J., Shen Y., Lin X., Shapley A. E., Wang F., Wu Q., Yang Q., 2025, NEXUS: A Spectroscopic Census of Broad-line AGNs and Little Red Dots at $3 \lesssim z \lesssim 6$ (arXiv:2505.20393), https://arxiv.org/abs/2505.20393

This paper has been typeset from a TEX/LATEX file prepared by the author.