Testing Lyα Emitters and Lyman-break Galaxies as Tracers of Large-scale Structures at High Redshifts

Sang Hyeok Im, Ho Seong Hwang, Jaehong Park, Jaehyun Lee, Hyunmi Song, Stephen Appleby, Yohan Dubois, C. Gareth Few, Brad K. Gibson, Juhan Kim, et al.

The Astrophysical Journal, 972:196 (17pp), 2024 September 10 https://doi.org/10.3847/1538-4357/ad67d2

Testing Lyα Emitters and Lyman-break Galaxies as Tracers of Large-scale Structures at High Redshifts

Sang Hyeok Im[1], Ho Seong Hwang[1,2], Jaehong Park[3,4], Jaehyun Lee[5], Hyunmi Song[6], Stephen Appleby[7,8], Yohan Dubois[9], C. Gareth Few[10], Brad K. Gibson[11], Juhan Kim[12], Yonghwi Kim[13], Changbom Park[3], Christophe Pichon[3,9,14], Jihye Shin[5], Owain N. Snaith[15], M. Celeste Artale[16], Eric Gawiser[17], Lucia Guaita[16], Woong-Seob Jeong[5], Kyoung-Soo Lee[18], Nelson Padilla[19], Vandana Ramakrishnan[18], Paulina Troncoso[20], and Yujin Yang[5]

[1] Department of Physics and Astronomy, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea; hhwang@astro.snu.ac.kr
[2] SNU Astronomy Research Center, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul 08826, Republic of Korea
[3] Korea Institute for Advanced Study, 85 Hoegi-ro, Dongdaemun-gu, Seoul 02455, Republic of Korea
[4] Space Science Exploration Directorate, Korea AeroSpace Administration (KASA), 537, Haeansaneop-ro, Sanam-myeon, Sacheon-si, Gyeongsangnam-do, Republic of Korea
[5] Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea
[6] Department of Astronomy and Space Science, Chungnam National University, 99 Daehak-ro, Yuseong-gu, Daejeon 34134, Republic of Korea
[7] Asia Pacific Center for Theoretical Physics, Pohang, 37673, Republic of Korea
[8] Department of Physics, POSTECH, Pohang 37673, Republic of Korea
[9] CNRS and Sorbonne Université, UMR 7095, Institut d'Astrophysique de Paris, 98 bis, Boulevard Arago, F-75014 Paris, France
[10] E.A. Milne Centre for Astrophysics, University of Hull, Hull HU6 7RX, UK
[11] Woodmansey Primary School, Hull Road, Woodmansey, HU17 0TH, UK
[12] Center for Advanced Computation, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea
[13] Korea Institute of Science and Technology Information, 245 Daehak-ro, Yuseong-gu, Daejeon, 34141, Republic of Korea
[14] IPhT, DRF-INP, UMR 3680, CEA, L'Orme des Merisiers, Bât 774, F-91191 Gif-sur-Yvette, France
[15] Department of Physics and Astronomy, University of Exeter, Exeter EX4 4QL, UK
[16] Universidad Andres Bello, Facultad de Ciencias Exactas, Departamento de Fisica, Instituto de Astrofisica, Fernandez Concha 700, Las Condes, Santiago RM, Chile
[17] Department of Physics and Astronomy, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA
[18] Department of Physics and Astronomy, Purdue University, 525 Northwestern Avenue, West Lafayette, IN 47906, USA
[19] Instituto de Astronomía Teórica y Experimental (IATE), CONICET-UNC, Laprida 854, X500BGR, Córdoba, Argentina
[20] Escuela de Ingeniería, Universidad Central de Chile, Avenida Francisco de Aguirre 0405, 171-0614 La Serena, Coquimbo, Chile
Received 2024 March 12; revised 2024 July 17; accepted 2024 July 24; published 2024 September 10

Abstract

We test whether Lyα emitters (LAEs) and Lyman-break galaxies (LBGs) can be good tracers of high-z large-scale structures, using the Horizon Run 5 cosmological hydrodynamical simulation. We identify LAEs using the Lyα emission line luminosity and its equivalent width, and LBGs using the broadband magnitudes at $z \sim 2.4$, 3.1, and 4.5. We first compare the spatial distributions of LAEs, LBGs, all galaxies, and dark matter around the filamentary structures defined by dark matter. The comparison shows that both LAEs and LBGs are more concentrated toward the dark matter filaments than dark matter. We also find an empirical fitting formula for the vertical density profile of filaments as a binomial power-law relation of the distance to the filaments. We then compare the spatial distributions of the samples around the filaments defined by themselves. LAEs and LBGs are again more concentrated toward their filaments than dark matter. We also find the overall consistency between filamentary structures defined by LAEs, LBGs, and dark matter, with the median spatial offsets that are smaller than the mean separation of the sample. These results support the idea that the LAEs and LBGs could be good tracers of large-scale structures of dark matter at high redshifts.

Unified Astronomy Thesaurus concepts: Large-scale structure of the universe (902); Hydrodynamical simulations (767); Lyα galaxies (978); Lyman-break galaxies (979)

1. Introduction

Studying the physical properties of large-scale structures in the Universe is important for constraining cosmological models (e.g., Park 1990; Lin et al. 1996; Hawkins et al. 2003; Eisenstein et al. 2005; Percival et al. 2007; Park et al. 2012; Hong et al. 2020; Dong et al. 2023). Since the first large redshift surveys of galaxies in the 1970s and 1980s (e.g., Gregory & Thompson 1978; Kirshner et al. 1978; Davis et al. 1982; Geller & Huchra 1989), the spatial distribution of galaxies has been a useful tool for studying of cosmology and the structure formation. Statistical analyses of the galaxy distribution (e.g., two-point and angular correlation functions) (Peebles 1975; Landy & Szalay 1993; Hamilton 1993) allow us to make a direct comparison between observations and theoretical predictions from cosmological simulations. However, most of the early observational studies of galaxy distribution were conducted with single-slit spectroscopy and were therefore inefficient in terms of the number of galaxies observed simultaneously.

Thanks to the development of the multiobject spectrographs in the 1990s, which allow simultaneous observations for hundreds of galaxies, the number of galaxies with measured redshifts has increased rapidly (e.g., Shectman et al. 1996; York 2000; Colless et al. 2001; Hwang et al. 2016). In addition, the appearance of advanced detectors and larger telescopes

enables deeper redshift surveys over larger areas. Therefore, we are now in a good position to easily measure the spatial distribution of high-redshift galaxies, which is important for understanding cosmology and structure formation (e.g., Lilly et al. 2009; Silverman et al. 2015; Sohn et al. 2021).

Despite these improvements in the observing techniques, spectroscopic observations for a large number of galaxies still require a large amount of telescope time and effort. The situation becomes worse if one tries to observe galaxies only in a certain redshift range because one cannot know the exact redshift of a galaxy before obtaining the spectrum of the galaxy. In this regard, narrowband imaging observations for Lyα emitters (LAEs; Hu & McMahon 1996; Ouchi et al. 2003; Gawiser et al. 2007; Lee et al. 2014) and the dropout technique for Lyman-break galaxies (LBGs; Steidel et al. 1998; Giavalisco et al. 2004; Bouwens et al. 2007; Toshikawa et al. 2016) are common ways to select galaxies within a certain redshift range at high redshifts without spectroscopic observations. One can reduce the number of target galaxies for follow-up spectroscopic observations using these kinds of approaches, making the redshift survey of high-redshift galaxies more efficient. Especially the typical redshift range of LAEs from narrowband imaging observations ($\Delta z \lesssim 0.04-0.08$), which is much smaller than that of LBGs from using the dropout technique ($\Delta z \sim 0.5$), enables us to select galaxies within a narrow redshift range very efficiently.

One of the ongoing observational projects using such narrowband imaging to study galaxy distribution at high redshifts is the One-hundred-deg^2 DECam Imaging in Narrow-bands (ODIN) survey (Lee et al. 2024; Ramakrishnan et al. 2023). We will use the data from this survey to identify LAEs at three redshifts $z \sim 2.4$, 3.1, and 4.5 with three custom narrowband filters that have central wavelengths of 419, 501, and 673 nm. We will also use these LAEs to identify and study the large-scale structure.[21] Because the target redshifts of the ODIN survey are around the epoch of peak mass accretion in galaxy clusters, the data will play an important role in improving our understanding of cosmic structure formation and evolution. The observation is currently ongoing with the Dark Energy Camera (DECam) mounted at Víctor M. Blanco 4 m telescope of Cerro Tololo Inter-American Observatory. The very first paper about the relation between Lyα blobs and the distribution of LAEs from the ODIN observed data of the COSMOS field at $z \sim 3.1$ has recently been published (Ramakrishnan et al. 2023).

Although using LAEs and LBGs is a convenient method to identify high-z galaxies, they are a subset of the entire galaxy populations. Therefore, it may not be obvious whether they can represent well the spatial distribution of the entire galaxy population or dark matter at high redshifts. In this study, we would like to examine the capability of LAEs and LBGs as tracers of the large-scale structures at high redshifts, by considering the large-scale structures of dark matter as a reference. We focus on the filamentary structures of LAEs, LBGs, all galaxies, and dark matter using the data from a new cosmological hydrodynamical simulation, the Horizon Run 5 (Lee et al. 2021; Park et al. 2022). We will examine the spatial distribution of these tracers around filaments and how those tracers give different shapes of filamentary structures. Please note that our analyses are in three-dimensional space, which may not be sufficient for direct comparison with observations. We plan to present a direct comparison with observations using two-dimensional mock data as a future study.

In Section 2, we describe the Horizon Run 5 data along with the selection methods for LAE and LBG samples. Section 3 includes our results on the filamentary structures of our samples and their spatial distribution around the filaments. In Section 4, we compare our results with previous studies and discuss the capability of LAEs and LBGs as tracers of large-scale structures of dark matter at high redshifts. Finally, we present the summary in Section 5.

2. Data and Sample Selection

2.1. Data

We use the data of the Horizon Run 5 (HR5: Lee et al. 2021) cosmological hydrodynamical simulation for our analysis. HR5 has a simulation box size[22] of $\sim$1 cGpc^3 with a high-resolution cuboid zoom-in region of $1049 \times 119 \times 127$ (cMpc)^3, where the spatial resolution reaches down to $\sim$1 proper kpc. This unique geometry of the zoom-in region is designed to be optimized for generating mock light cone data of deep field surveys such as ODIN (Lee et al. 2024), LSST (LSST Dark Energy Science Collaboration 2012), and DESI (DESI Collaboration et al. 2016). The linear power spectrum and the initial conditions are generated at $z = 200$ and evolved using the adaptive mesh refinement code RAMSES (Teyssier 2002) until $z = 0.625$. Various subgrid physics are also implemented during the simulation to model galaxy formation and evolution. After the simulation, galaxies (i.e., subhalos) are identified using PGalF, which is an extended galaxy finder based on the friends-of-friends algorithm (see Appendix A of Kim et al. 2023 for details). The physical properties of each galaxy are calculated with the particles belonging to the galaxy. For more details about the HR5 simulation, please see Lee et al. (2021). We use three snapshot data of HR5 at $z \sim 2.4$, 3.1, and 4.5, which are similar to the three target redshifts of the ODIN survey. Throughout HR5 simulation and this study, we adopt the ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_b = 0.047$, $\sigma_8 = 0.816$, and $h_0 = 0.684$ as compatible with Planck data (Planck Collaboration et al. 2016).

HR5 has particles and grids with different resolutions to simultaneously achieve the high spatial resolution at the zoom-in region and obtain the structures with very large scales ($\sim$1 cGpc) throughout the simulation box. Those particles and grids with different resolutions may be mixed near the boundary between the regions with different resolutions. This can affect the estimated physical properties of the halos and subhalos (i.e., galaxies) near the boundary. To deal with this issue, we use only the galaxies and dark matter particles in the $90 \times 90 \times 1000$ (cMpc)^3 region at the center of the highest resolution region. Hereafter, we will refer to this region as the central zoom-in region. Inside this region, all galaxies and dark matter particles are at least 3 cMpc away from the lower-resolution particles.

We generate the samples of dark matter particles and all galaxies at the central zoom-in region for three redshifts: $z \sim 2.4$, 3.1, and 4.5. We randomly select 0.1% of the total dark matter particles to reduce the computational load of further analysis (e.g., generating a density map or extracting filaments

[21] The distribution of galaxies or dark matter larger than the galaxy cluster scale ($\gtrsim$a few Mpc).

[22] Throughout this paper, we present all spatial quantities in comoving scale unless otherwise mentioned.

from those particles). We also apply a minimum stellar mass cut of $2.137 \times 10^8\,M_\odot$, corresponding to ∼100 star particles, to the galaxy samples to select galaxies with statistically reliable physical properties. We use the same stellar mass cut when generating the LAE and LBG samples in Section 2.2. The sample of all galaxies above the mass threshold is divided further into several subsamples with different stellar mass ranges. The stellar mass range of each subsample is available in Table 1. These ranges are selected to make the number of galaxies in each mass bin at least larger than 10,000 and to be similar to each mass bin in each redshift. From these subsamples, we examine how the stellar masses of galaxies affect their spatial distribution and resulting filamentary structures. In Table 1, we present the number of dark matter particles and galaxies in each sample and subsample.

2.2. LAE and LBG Selection

2.2.1. LAEs

We assign the Lyα emission line luminosity ($L_{\mathrm{Ly}\alpha}$) and the rest-frame equivalent width (REW) of each galaxy to select LAEs (Weinberger et al. 2019; J. Park et al. 2024, in preparation). To achieve this, we use the conditional probability function of REW given the UV luminosity, motivated by the empirical models (e.g., Dijkstra & Wyithe 2012; Gronke et al. 2015; Oyarzún et al. 2017; Mason et al. 2018). In this study, we adopt Equation (4) of Mason et al. (2018), and adjust the free parameters to match the observed Lyα luminosity functions at $z \sim 2.4$, 3.1, and 4.5 (J. Park et al. 2024, in preparation). We then assign a REW value to each galaxy according to the UV luminosity of the galaxy and the conditional probability that we have calculated. The REW value of each galaxy is then converted to $L_{\mathrm{Ly}\alpha}$ using the following equation:

$$L_{\mathrm{Ly}\alpha} = \mathrm{REW} \times L_{\mathrm{UV}} \times \frac{\nu_\alpha}{\lambda_\alpha}\left(\frac{\lambda_\alpha}{\lambda_{\mathrm{UV}}}\right)^{-\beta-2}, \quad (1)$$

where ν_α and λ_α are the frequency and wavelength of the Lyα emission line, respectively. The L_{UV} represents the UV luminosity, and λ_{UV} is the wavelength of the UV emission. Here, we assume that the UV continuum of each galaxy follows a power law with a slope of β. This semiempirical modeling does not capture the exact physical conditions of each galaxy as is done with radiative transfer calculations (e.g., Zheng et al. 2014; Behrens et al. 2018). However, it allows us to populate LAEs in a large simulation box and to study their statistical properties such as luminosity function and spatial distribution.

We summarize the selection criteria for LAEs at $z \sim 2.4$, 3.1, and 4.5 in Equations (2), (3), and (4), respectively. Here, we consider the 5σ detection limit (Lee et al. 2024) of the three narrowband filters of the ODIN survey as the minimum $L_{\mathrm{Ly}\alpha}$ values. We also give a minimum REW value of 20 Å to select those galaxies with significant emission line luminosity compared to the continuum level. We adopt these minimum REW values from the LAE selection criteria of the ODIN survey (see Ramakrishnan et al. 2023 and Firestone et al. 2023).

$$z \sim 2.4:\ \log\left(\frac{L_{\mathrm{Ly}\alpha}}{\mathrm{erg\ s^{-1}}}\right) > 42.18 \text{ and } \mathrm{REW} > 20\ \text{Å}, \quad (2)$$

$$z \sim 3.1:\ \log\left(\frac{L_{\mathrm{Ly}\alpha}}{\mathrm{erg\ s^{-1}}}\right) > 42.20 \text{ and } \mathrm{REW} > 20\ \text{Å}, \quad (3)$$

$$z \sim 4.5:\ \log\left(\frac{L_{\mathrm{Ly}\alpha}}{\mathrm{erg\ s^{-1}}}\right) > 42.39 \text{ and } \mathrm{REW} > 20\ \text{Å}. \quad (4)$$

Finally, ∼3.8, 7.4, and 8.1% of galaxies with stellar masses larger than $2.137 \times 10^8\,M_\odot$ are selected as LAEs for $z \sim 2.4$, 3.1, and 4.5, respectively. We mainly use the sample of all LAEs at each redshift, but divide it into three stellar mass bins if necessary. The relevant sample sizes and the mass ranges of the LAE samples and their subsamples are summarized in Table 1.

2.2.2. LBGs

We select LBGs using the selection criteria of apparent magnitudes of galaxies as done for observations (e.g., Steidel et al. 2003; Toshikawa et al. 2016; Kawamata et al. 2018). In this study, we are interested in comparing LBGs, LAEs, all galaxies, and dark matter particles at each redshift. Thus, we use the snapshot data of HR5 for each redshift instead of light cone data with continuous redshift ranges. Therefore, it is important to note that LBG samples in our study cannot be directly compared to photometrically selected LBGs from observations, which have redshift ranges of $\Delta z \sim 0.5$.

To calculate the apparent magnitudes of each galaxy, we first model the spectral energy distribution (SED) of the galaxy. To do that, we consider each star particle belonging to a galaxy as a simple stellar population. Then, we calculate the intrinsic SED of each galaxy as the summation of the SEDs of all the star particles and apply the Milky Way dust attenuation model to the intrinsic SED. The simple stellar population model and the dust attenuation law are selected to give the well-matched prediction of observational properties such as UV luminosity function and dust mass–stellar mass relation. The details about the choice of models for simple stellar population and dust attenuation law will be provided by H. Song et al. (2024, in preparation). We apply the Madau (1995) intergalactic medium absorption model to the intrinsic SED of each galaxy to obtain the observed SED. The apparent magnitudes of each galaxy are calculated by integrating the SED over the filter set of Subaru Strategic Program (SSP; Aihara et al. 2018b, 2018a, 2022) that we are going to use for the analysis of the ODIN data (Ramakrishnan et al. 2023; Firestone et al. 2023). Here, we calculate all the magnitudes in the AB magnitude system (Oke & Gunn 1983).

We present the selection criteria for LBGs at $z \sim 2.4$, 3.1, and 4.5 in Equations (5), (6), and (7), respectively. The basic idea of the $z \sim 3.1$ and 4.5 criteria is to identify u- and g-dropout galaxies using the magnitude difference across the Lyman limit. However, the Lyman limit for $z \sim 2.4$ is outside the u band, so our criteria for $z \sim 2.4$ is set to capture the suppressed u-band magnitude due to the Lyα forest absorption lines. Therefore, we set the upper limit of the $u - g$ color for $z \sim 2.4$. These selection criteria are optimized to make the selection efficiency peaks at each ODIN target redshift. For three redshifts, we also consider the 5σ limiting magnitudes of SSP as the maximum broadband magnitudes. These values are measured with randomly placed 2″ diameter apertures on the SSP image of the COSMOS field (see Section 2.1 of Ramakrishnan et al. 2023 for details). With these

Table 1
The Number of Galaxies or Dark Matter Particles in Each Sample for Three Redshifts

		$z \sim 4.5$	$z \sim 3.1$	$z \sim 2.4$
DM particles (0.1% sampled)		456,716	453,766	450,738
All galaxies	$2.137 \times 10^8 < M_\star/M_\odot < 3.5 \times 10^8$	29,289	70,113	72,651
	$3.5 \times 10^8 < M_\star/M_\odot < 6.0 \times 10^8$	23,233	63,961	76,826
	$6.0 \times 10^8 < M_\star/M_\odot < 1.5 \times 10^9$	22,700	72,912	97,243
	$1.5 \times 10^9 < M_\star/M_\odot$	12,335	56,784	87,905
	Total	87,557	263,770	334,625

		$z \sim 4.5$	$z \sim 3.1$	$z \sim 2.4$
LAEs	$2.137 \times 10^8 < M_\star/M_\odot < 1.5 \times 10^9$	3904	6245	1590
	$1.5 \times 10^9 < M_\star/M_\odot < 2.5 \times 10^9$	1379	3850	1620
	$2.5 \times 10^9 < M_\star/M_\odot$	1833	9457	9505
	Total	7116	19,552	12,715
LBGs	$2.137 \times 10^8 < M_\star/M_\odot < 1.5 \times 10^9$	2486	5693	1292
	$1.5 \times 10^9 < M_\star/M_\odot < 2.5 \times 10^9$	3374	6961	3402
	$1.5 \times 10^9 < M_\star/M_\odot$	5664	17,996	12,411
	Total	11,524	30,650	17,105
	LAEs ∩ LBGs	3179	8679	2740

Note. The stellar mass ranges for the subsamples of galaxy samples are also represented. Note that LAEs and LBGs are also included in all galaxy samples. We also present the number of galaxies that are selected as both LAE and LBG.

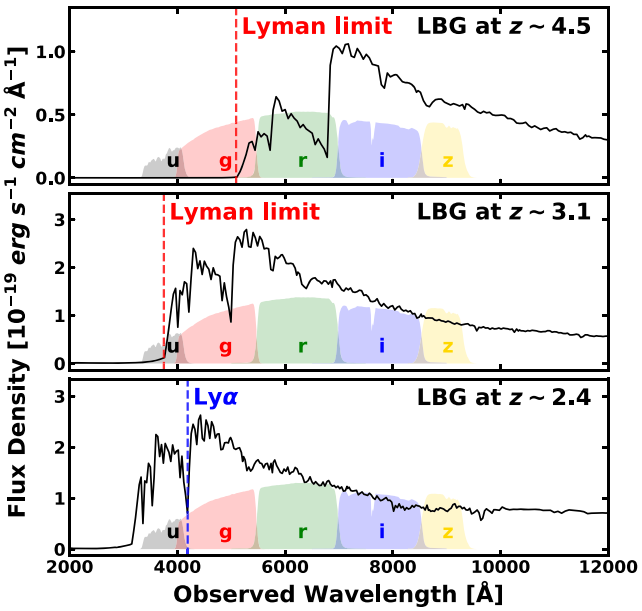

Figure 1. Example SEDs of LBG at z ∼ 2.4, 3.1, and 4.5, which are selected using the criteria of Equations (5), (6), and (7), respectively. They show clear Lyman-break features for $z \sim 3.1$ and 4.5, and the suppressed u-band fluxes due to the Lyα forest (see the text for the details). For comparison, we also present the filter response functions for the broadband filters of the SSP in arbitrary units.

criteria, ∼5.11, 11.6, and 13.2% of galaxies with stellar masses larger than $2.137 \times 10^8\,M_\odot$ are selected as LBGs for $z \sim 2.4$, 3.1, and 4.5, respectively. The LBG sample is further divided into three subsamples based on their stellar masses. Similar to the LAEs, we mainly use the sample of all LBGs at each redshift unless it is mentioned. The total number of selected LBGs and the mass ranges of the subsamples for three redshifts are summarized in Table 1. We also present the example SEDs of our LBGs for each redshift in Figure 1. They show clear Lyman-break features (for $z \sim 3.1$ and 4.5) or suppressed fluxes in the u band (for $z \sim 2.4$). It should be noted that some galaxies in our simulation could be selected as both LAEs and LBGs; we also list the number of such galaxies in Table 1.

$z \sim 2.4$:
$0.6 < (u - g) < 1.2$ and $-0.2 < (r - i) < 0.1$ and
$(u - r) > 2.1 \times (r - i) + 0.5$ and $(g - r) < 0.7$ and
$(u - r) < 2.1 \times (r - i) + 1.1$ and
$g < 26.3$ and $r < 26.0$ and $i < 25.9$, (5)

$z \sim 3.1$:
$(u - g) > 1.2$ and $(r - i) < 0.3$ and
$0.1 < (g - r) < 0.7$ and $(u - g) > 2.5 \times (g - r) + 0.5$ and
$g < 26.3$ and $r < 26.0$ and $i < 25.9$, (6)

$z \sim 4.5$:
$(g - i) > 2.3$ and $(g - i) > 2.1 \times (i - z) + 2.2$ and
$(r - i) > 0.8$ and $(g - i) > 0.4 \times (i - z) + 1.2$ and
$-0.5 < (i - z) < 0.3$ and $i < 25.9$. (7)

2.3. Physical Properties of the Sample Galaxies

2.3.1. Spatial Distribution

We examine the spatial distribution of our samples of LAEs, LBGs, and dark matter particles. For convenience, we plot the two-dimensional distribution for our samples in a sliced region with the line-of-sight thickness of 15 cMpc of the central zoom-in region at $z \sim 2.4$, 3.1, and 4.5 in Figure 2. We present the distribution of dark matter particles at the background of each panel to compare it with the distribution of LAEs and LBGs. The plot clearly shows that LAEs and LBGs trace the dense regions of dark matter particles at the three redshifts, as expected. We can also see the growth of structures as the Universe evolves.

2.3.2. Relation between Star Formation Rate and Stellar Mass

To have an idea about the difference in physical properties between our samples, we examine the relation between star formation rate (SFR) and stellar mass of LAE, LBG, and all galaxy samples. In Figure 3, we present the relation for the three redshifts along with the stellar mass distribution of our samples. Here, we use the average SFR over 100 Myr ($\mathrm{SFR}_{100\,\mathrm{Myr}}$) to refer to the SFR of each galaxy. The relations between SFR and stellar mass of all galaxies are well fitted with a single power law at three redshifts as well known in previous studies (e.g., Elbaz et al. 2007; Speagle et al. 2014; Salmon et al. 2015; Iyer & Gawiser 2018). Here, we perform least-squares fitting and obtain the best-fit power-law slopes of 0.85, 0.82, and 0.91 for $z \sim 2.4$, 3.1, and 4.5, respectively.

To quantitatively compare the SFRs among the samples without the dependence of stellar mass on the SFR, we calculate the starburstiness (R_{SB}; Elbaz et al. 2011) of each galaxy using the following equation:

$$R_{\mathrm{SB}} = \frac{\mathrm{sSFR}}{\mathrm{sSFR}_{\mathrm{MS}}}. \quad (8)$$

Here, sSFR is the specific SFR of each galaxy, and $\mathrm{sSFR}_{\mathrm{MS}}$ is the sSFR of the star-forming main-sequence galaxy with the same stellar mass. We use the best-fit power-law relation between SFR and stellar mass from the least-squares fitting to calculate the $\mathrm{sSFR}_{\mathrm{MS}}$ of each galaxy. We then compare the distribution of R_{SB} of the LAE, LBG, and all galaxy samples in Figure 4. LAE and LBG samples have more skewed distributions of R_{SB} toward larger values than all galaxy samples. We also perform the Kolmogorov–Smirnov (KS) test and the Anderson–Darling (AD) k-sample tests on the R_{SB} distributions. The results are shown in each panel of Figure 4 with the p-values indicating the probability that the two samples are drawn from the same parent distribution. All the p-values are small (i.e., <0.001), indicating that the LAEs and LBGs are distinctive from the entire galaxy samples. This result, along with the skewed R_{SB} distributions of LAEs and LBGs in Figure 4, implies that LAEs and LBGs tend to be more star forming than the entire galaxy populations at the three redshifts.

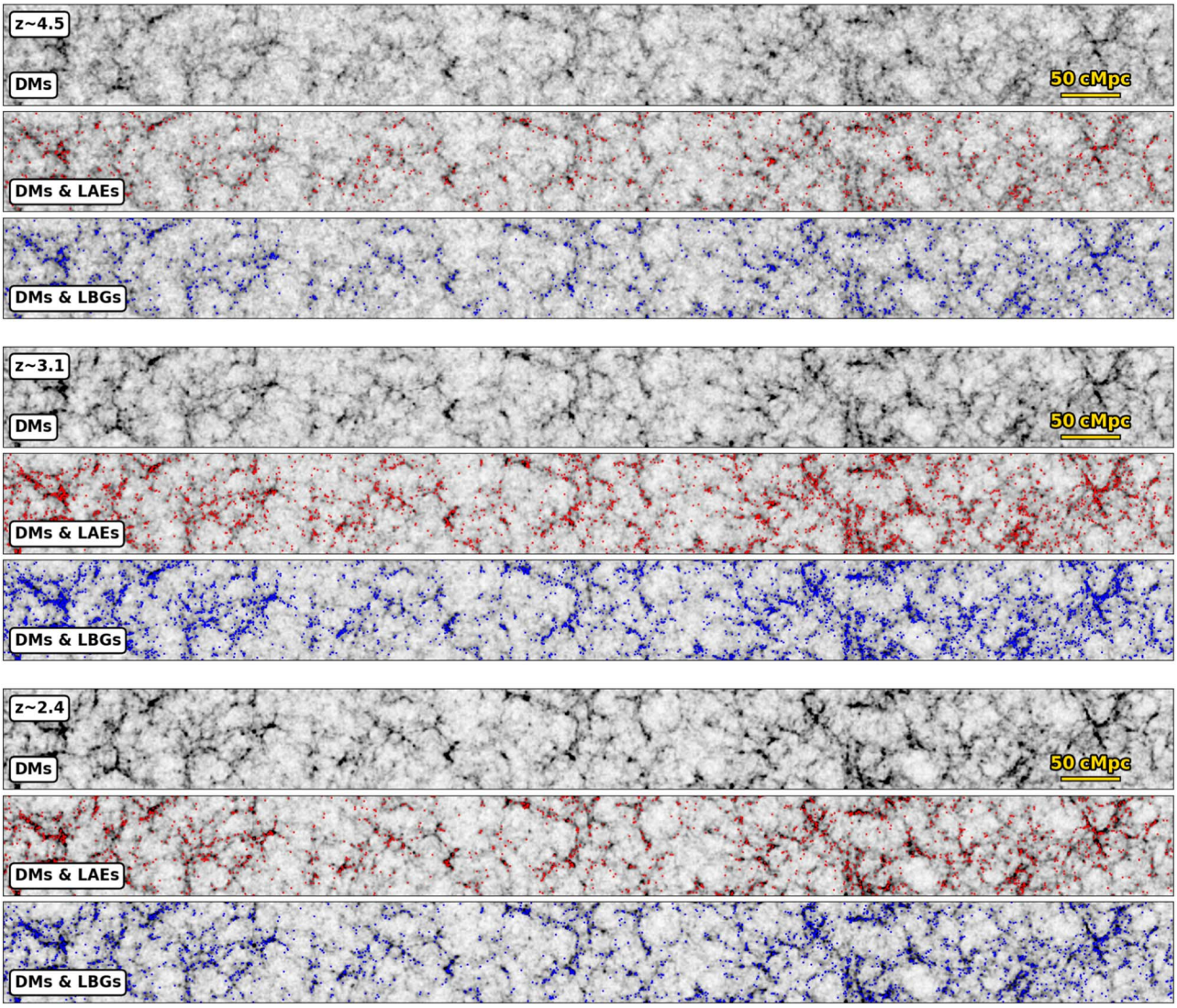

Figure 2. Spatial distribution of dark matter particles, LAEs, and LBGs at $z \sim 4.5$ (top three panels), 3.1 (middle three panels), and 2.4 (bottom three panels). For visualization, only a slice with the line-of-sight thickness of 15 cMpc is shown. The yellow horizontal bar at the upper panel of each redshift represents the comoving length of 50 Mpc. LAEs and LBGs seem to follow well the dense regions of dark matter.

3. Results

3.1. Filamentary Structures

To identify the filamentary structures of our samples, we apply the Discrete Persistent Structure Extractor code[23] (DisPerSE; Sousbie 2011; Sousbie et al. 2011) to those samples. A brief description of how DisPerSE extracts filamentary structure is as follows. First, DisPerSE generates a density map of the sample using the Delaunay tessellation. Then, DisPerSE finds sets of small segments, connecting saddle points and local maxima of the density field. These sets of segments are extracted as the filamentary structure of the sample. More details on DisPerSE are available in Sousbie (2011) and Sousbie et al. (2011).

When running DisPerSE, the lack of data points outside the boundary would affect the density field estimation and the resulting filamentary structures near the boundary. DisPerSE provides several options to solve this problem by creating artificial data points outside the boundary. We test these options of DisPerSE and find that each option gives slightly different filamentary structures near the boundary. To be conservative, we decide not to use the filamentary structures within 5 cMpc from the boundary for our analysis.

Another important parameter to consider is the choice of the minimum persistence level of the extracted filamentary structures. The persistence level of a filament is defined as the density difference between its end points at local maxima and saddle point. The filaments with larger persistence levels are considered more robust features than those with smaller persistence levels. We can extract more robust but less detailed filamentary structures by selecting those with higher persistence levels. There are many studies based on DisPerSE, which adopt 3σ of the noise level of random data as the minimum persistence level (e.g., Malavasi et al. 2016; Kraljic et al. 2017). This means that the selected filamentary structures would not be generated by random fluctuations within the 3σ confidence level. However, we use a higher level (5σ) for dark matter particles because 3σ gives much more detailed filamentary structures than the distribution of the galaxy samples (i.e., LAEs, LBGs, and all galaxies within stellar mass ranges).

[23] http://www2.iap.fr/users/sousbie/web/html/indexd41d.html

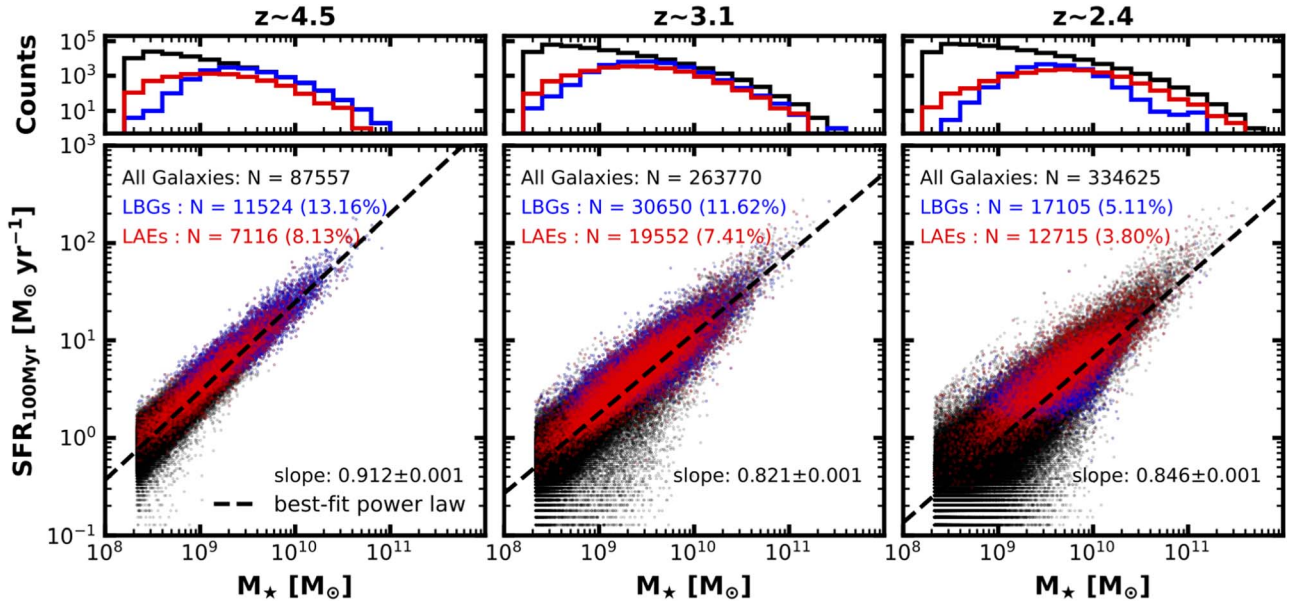

Figure 3. (Upper panels) Stellar mass distributions of LAE, LBG, and all galaxy samples. (lower panels) The relations between stellar mass and SFR of the samples. Here, we use the average SFR over 100 Myr to refer to the SFR of each galaxy. The best-fit power-law slopes for all galaxy samples are also presented.

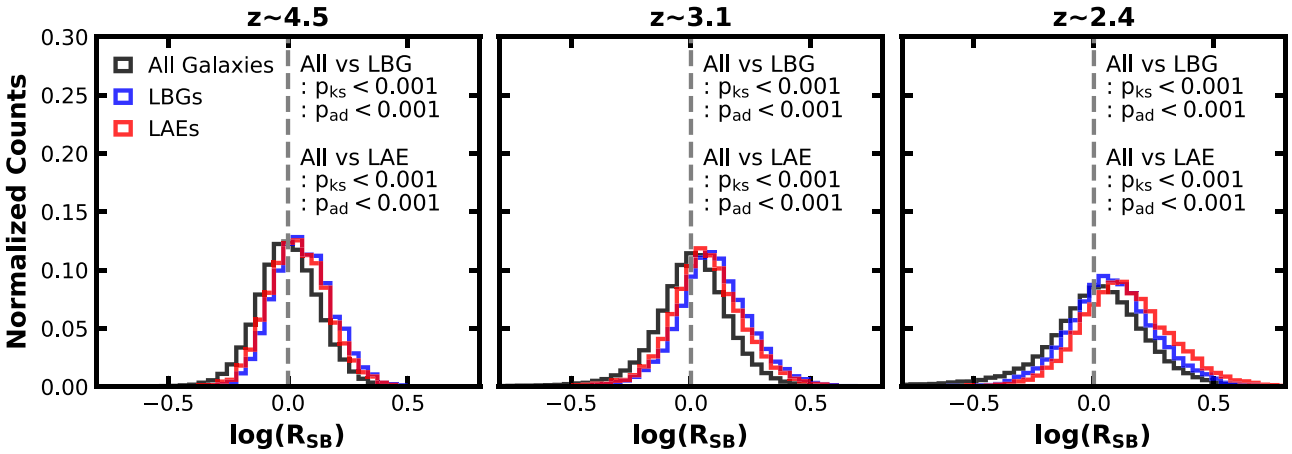

Figure 4. Distributions of starburstiness (R_{SB}) of the LAE, LBG, and all galaxy samples at three redshifts. LAE and LBG samples show more skewed distributions toward larger R_{SB} values than the entire galaxy population at three redshifts. The p-values of the KS test and AD test between LAE, LBG, and all galaxy samples at three redshifts are also presented in each panel.

Figure 5 shows the extracted filamentary structures of dark matter particles at $z \sim 2.4$. For comparison, we present the filamentary structures with different persistence levels as different colors (see labels in Figure 5).

3.2. Spatial Distributions of Samples around Filaments Defined by Dark Matter

We compare the spatial distributions of LAEs, LBGs, all galaxies, and dark matter particles around the filamentary structures defined by dark matter particles. We first calculate the distance from a galaxy or a dark matter particle to the closest filament for each sample. We will refer to this as $D_{\mathrm{fil,DM}}$; the distance to the filamentary structures defined by dark matter. We then calculate the vertical density profile around the filamentary structures using the following equation:

$$\Sigma(D_{\mathrm{fil,DM}}) = \frac{\#\text{ of objects at } r_i < D_{\mathrm{fil,DM}} < r_j}{\pi({r_j}^2 - {r_i}^2)} \times \frac{1}{N_{\mathrm{tot}}}. \quad (9)$$

Here, $\Sigma\,(D_{\mathrm{fil,DM}})$ is the two-dimensional probability density that a galaxy (or a particle) is located at the annulus between r_i and r_j from the filament, and N_{tot} is the total number of data in the sample.

In this calculation, the galaxies and dark matter particles close to the nodes (i.e., the intersections of the filaments) can affect the calculated density profiles around the filamentary structures. To remove such effects, we exclude the galaxies and dark matter particles close (<2 cMpc) to the nodes when calculating $\Sigma\,(D_{\mathrm{fil,DM}})$. We choose the minimum separation of 2 cMpc as the distance from filaments where the density profiles of our samples become smaller than those of randomly distributed particles (see Section 3.3).

Figure 6 shows the $\Sigma\,(D_{\mathrm{fil,DM}})$ of each sample at three redshifts. LAEs, LBGs, and all galaxies show steeper slopes and are more concentrated toward the filamentary structure defined by dark matter particles than dark matter particles themselves. This result is consistent with the visual impression of Figure 2, where LAEs, LBGs, and all galaxies reside in the

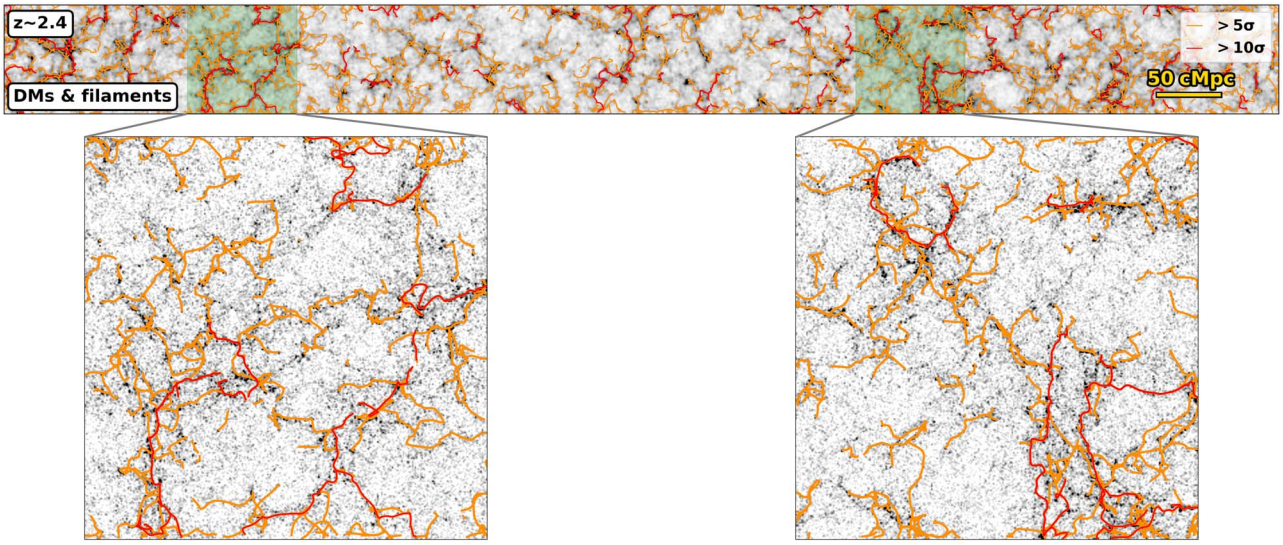

Figure 5. Filamentary structures defined by dark matter particles at $z \sim 2.4$ with the minimum persistence level of 5σ (orange) and 10σ (red). For convenience, only a slice with the line-of-sight thickness of 15 cMpc is presented. We note that we use three-dimensional filamentary structures for our analysis, while only the projected view of them is shown in here. Please also note that filaments with a minimum persistence level of 5σ include those with 10σ.

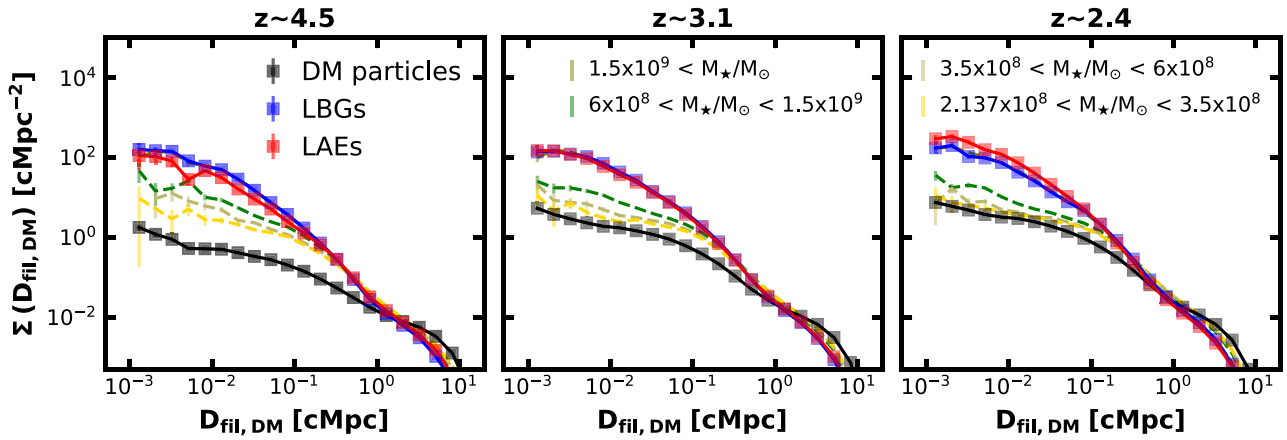

Figure 6. Vertical density profile around dark matter filaments as a function of distance to dark matter filaments ($D_{\rm fil,DM}$) for each sample at three redshifts. Dashed lines in each panel represent the subsamples of the entire galaxies with different stellar mass ranges (see labels in the middle and right panel). The error bars represent the standard deviations from bootstrapping with 1000 times. For three redshifts, LAEs, LBGs, and all galaxies are more concentrated toward the dark matter filaments than dark matter particles. Galaxies with larger stellar masses are also more concentrated than those with lower stellar masses.

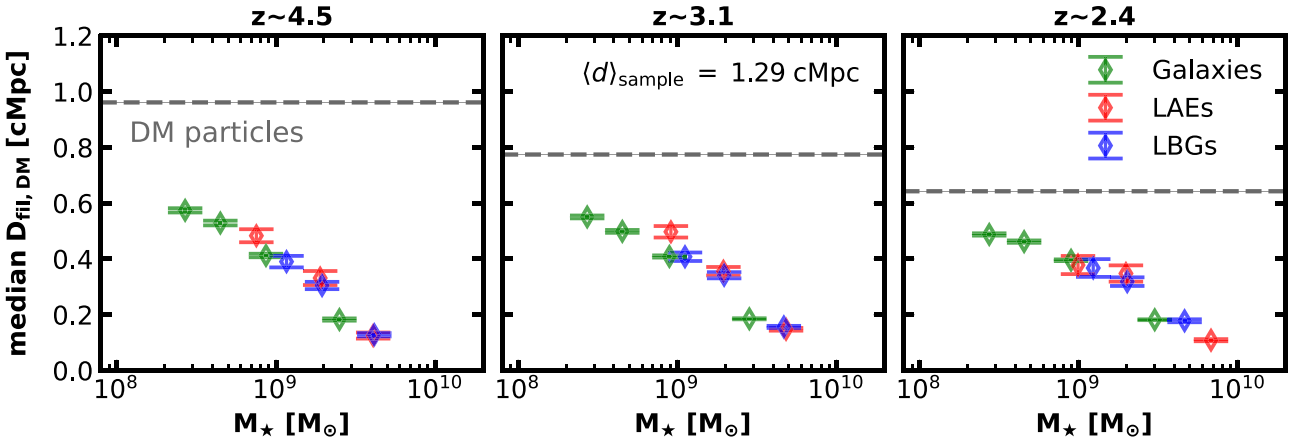

Figure 7. Median $D_{\rm fil,DM}$ (i.e., distance to dark matter filament) of each sample at three redshifts. Vertical error bars and shaded regions around dashed lines represent 1σ errors from bootstrapping with 1000 times. The LAE, LBG, and galaxy samples show smaller median values than dark matter particles. In the right panel, we also present the mean separation of the sample used for extracting filaments.

dense regions of dark matter particles. In addition, the galaxies with larger stellar masses are more concentrated toward the filaments than those with smaller stellar masses, indicating that more massive galaxies would form and evolve in the denser region of dark matter. These trends appear at all three redshifts.

Next, we compare the median values of $D_{\rm fil,DM}$ of our samples in Figure 7. Here, we use only the galaxies and dark matter particles with $D_{\rm fil,DM} < 2$ cMpc to consider those directly relevant to filamentary structures as described in Section 3.3. Here, we use the mass-binned subsamples of LAEs and LBGs at each redshift. LAEs, LBGs, and all galaxies have smaller median values than dark matter particles, regardless of redshift. This result again shows that LAEs, LBGs, and all galaxies are more concentrated toward the filamentary structures defined by dark matter than dark matter particles themselves. In addition, more massive galaxies have smaller median values than less massive galaxies, indicating that galaxies with larger stellar masses are more concentrated toward dark matter filaments than those with smaller stellar masses. It is also important to note that LAEs and LBGs follow well the trend of the entire galaxy population (i.e., decreasing median value as stellar mass increases). However, the absolute values of median $D_{\rm fil,DM}$ of the samples would differ for different choices of the minimum persistence level or the size of the sample size used to extract the filaments. Therefore, we also present the mean separation, $\langle d \rangle_{\rm sample}$, of the tracer sample (here, dark matter particles) for comparison with other studies.

To examine how the choice of the minimum persistence level affects our results, we also use a higher value (10σ) and perform similar analyses in Appendix A. We also use another filament finder T-ReX (Bonnaire et al. 2020) to examine whether there is any bias introduced by using only one filament finder and show the results in Appendix B. In both cases, the median values become higher than those in Figure 7. However, both tests show trends similar to Figure 7, indicating the robustness of our results to the choice of the persistence level and the filament finder.

3.3. Empirical Fitting Formula for Vertical Density Profiles

To have a quantitative description of the vertical density profile around dark matter filaments, $\Sigma\,(D_{\rm fil,DM})$, we introduce the following fitting formula:

$$\Sigma\,(D_{\rm fil,DM}) = \frac{\Sigma_0}{(1 + D_{\rm fil,DM}/D_c)^{\alpha}}. \quad (10)$$

Here, Σ_0 is the density at the center of the filamentary structure (i.e., where $D_{\rm fil,DM} = 0$) and D_c is a characteristic distance from the filament where the slope of the profile changes from zero to $-\alpha$.

The density profiles for our samples of LAEs, LBGs, all galaxies, and dark matter particles are well fitted with Equation (10) within $D_{\rm fil,DM} \lesssim 2$ cMpc regardless of redshift. We present the example cases with the best-fit curves in Figure 8. Here, we show only the results for dark matter particles at three redshifts (upper panels) and LAEs, LBGs, and galaxies with stellar masses larger than $1.5 \times 10^9\,M_{\odot}$ at $z \sim 2.4$ (lower panels). We also generate comparison samples of randomly distributed particles with the same size as each sample and calculate their $\Sigma(D_{\rm fil,DM})$. The density profile of these randomly distributed samples starts to decrease at some point because of the mean separation between the filaments. When $D_{\rm fil,DM} \gtrsim 2$ cMpc, the density profiles for all of our samples become smaller than those of randomly distributed particles (see Figure 8). This can be interpreted as the size of the overdense region around dark matter filaments in our study is $\sim$2 cMpc. This size of the overdense region could vary with the choice of the minimum persistence level. The fitting results for all of our samples are shown in Appendix C.

3.4. Spatial Distributions of Samples around Filaments Defined by Themselves

We also examine the spatial distributions of DMs, LAEs, LBGs, and all galaxies around the filamentary structures defined by themselves. The shape of the extracted filamentary structures depends not only on the choice of the minimum persistence level, but also on the size of the sample. Therefore, we generate 1000 subsamples with the same size of 7000 for each sample to make a fair comparison between them. The size of the subsamples is chosen to be slightly smaller than the size of the LAE samples at $z \sim 4.5$. Each subsample is constructed from a random selection of the members from each galaxy sample. We extract the filamentary structures of these random subsamples with a minimum persistence level of 3σ. We again calculate the distance from a galaxy or a dark matter particle in each sample to the closest filament defined by themselves; $D_{\rm fil,self}$. Here, the galaxies and dark matter particles close to the nodes ($\leqslant$2 cMpc) are ignored to remove the effect of the nodes. We calculate median values of $D_{\rm fil,self}$ for 1000 random subsamples of each sample. Then, we calculate the mean and the standard deviation value of those 1000 median $D_{\rm fil,self}$ values for each sample.

Figure 9 shows the results for the LAEs, LBGs, all galaxies, and dark matter particles at three redshifts. Here, we present the standard deviation for each sample as a shaded region or a vertical error bar. At all redshifts, LAEs, LBGs, and all galaxies have smaller median $D_{\rm fil,self}$ values than dark matter particles. This result implies that LAEs, LBGs, and all galaxies form thinner filamentary structures than dark matter. We also find that more massive galaxies tend to be more concentrated toward their filaments than less massive galaxies, which is consistent with the result of Section 3.2. LAEs and LBGs again follow this stellar mass trend well. Similar to Figure 7, the absolute values of the median $D_{\rm fil,self}$ of the samples would change if one uses a different minimum persistence level or different sample size. Therefore, we present the mean separation of the samples used to extract the filaments (here, each subsample with a size of 7000) in the right panel of Figure 9.

3.5. Difference between Filamentary Structures from Dark Matter Particles and Galaxies

We examine whether there is any systematic spatial offset between the filamentary structures defined by dark matter particles and those by LAEs, LBGs, and all galaxies. We again use the filamentary structures from the 1000 randomly selected subsamples of Section 3.4 for LAEs, LBGs, and all galaxies. We then compare them to the dark matter filaments in Section 3.2. Figure 10 shows a two-dimensional projected view of these filamentary structures at the same region of Figure 2 at $z \sim 2.4$. For convenience, we present only five cases out of 1000 subsamples constructed from each galaxy sample with random selection. The filamentary structures of the galaxy

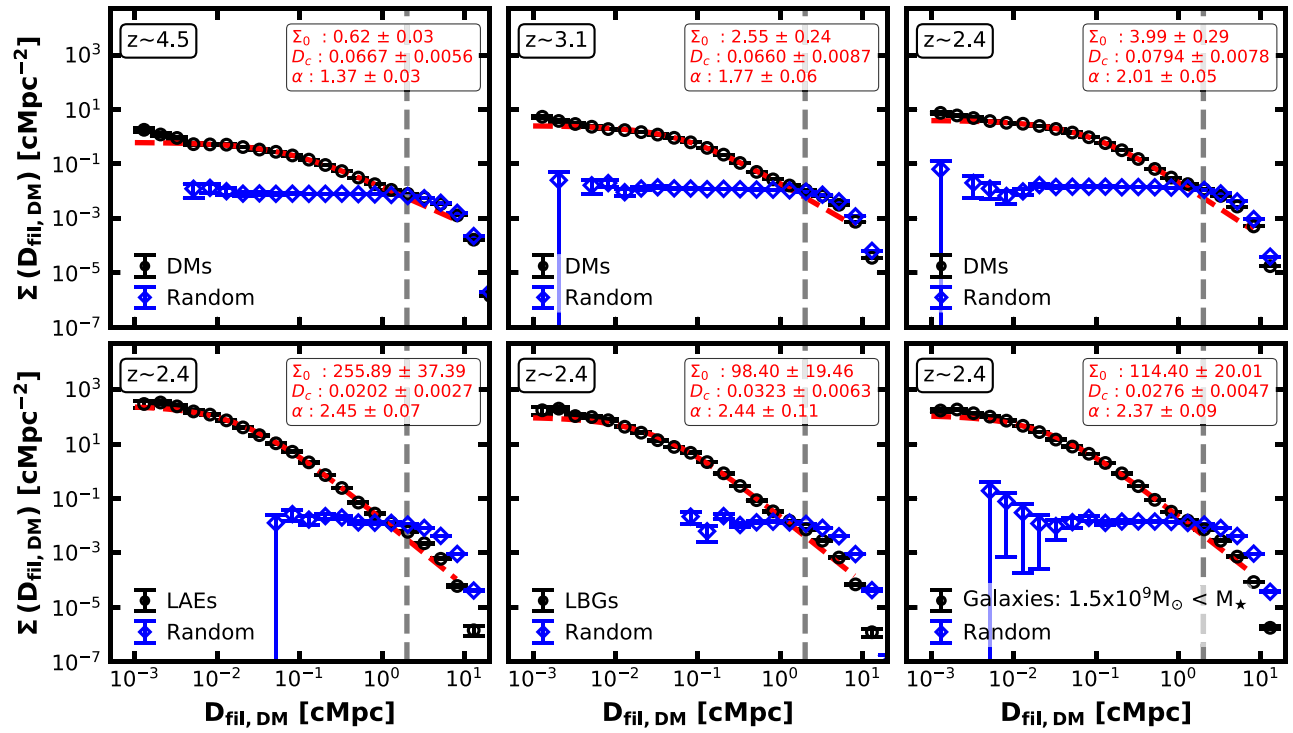

Figure 8. Fitting results for the vertical density profile of dark matter particles at three redshifts (upper panels) and LAEs, LBGs, and galaxies with stellar masses larger than $1.5 \times 10^9\,M_\odot$ at $z \sim 2.4$ (lower panels). We also present the best-fit values of Σ_0, D_c, and α for each sample at the upper right corner of each panel. The gray dashed line in each panel represents the maximum fitting range (2 cMpc). The blue markers show the vertical density profile of randomly distributed particles with the same size as the sample at each panel. The empirical fitting formula of Equation (10) shows good agreement at $D_{\rm fil,DM} \lesssim 2$ cMpc.

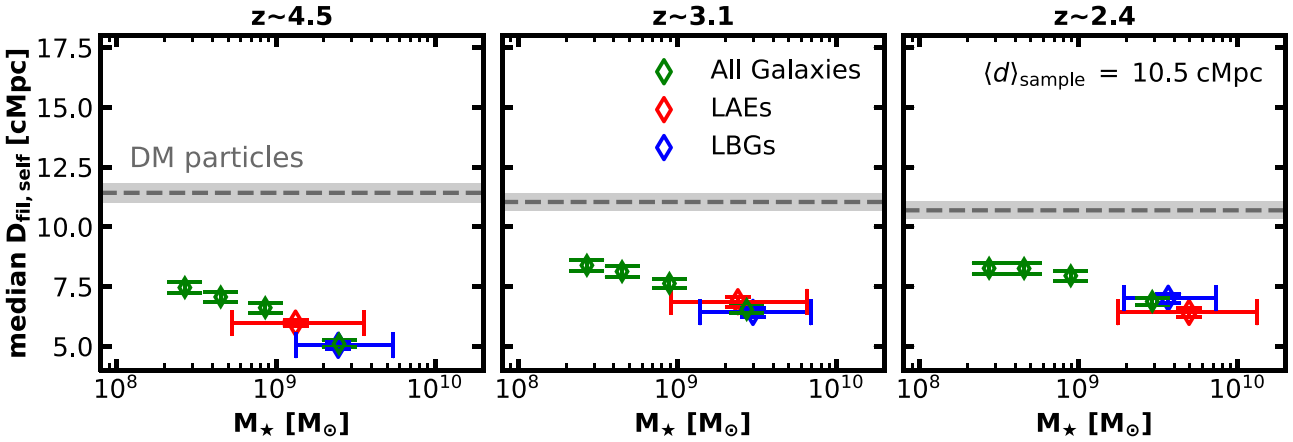

Figure 9. Median $D_{\rm fil,self}$ value of each sample at three redshifts. The shaded regions and vertical error bars show the standard deviations for the samples (see text for detail), while the horizontal error bars for LAEs and LBGs represent the 15.9 and 84.1 percentiles of stellar masses. The galaxy samples including LAEs and LBGs show smaller median values than dark matter particles. We also present the mean separation of the sample used for extracting filaments at the right panel.

samples show good agreement with those of dark matter particles in overall shapes. However, there are discrepancies in details between filamentary structures traced by dark matter particles and those by LAEs, LBGs, and all galaxies. Especially, there are many missing dark matter filaments when we identify filaments using galaxy samples. It would be because of the nature of the galaxies, which are biased tracers of dark matter distributions, and the effect of the random realization of the galaxy samples.

To quantify the difference between the filamentary structures of the galaxy samples and the dark matter particles, we calculate the median spatial offset between them. As mentioned in Section 3.1, the filamentary structures extracted by DisPerSE consist of small segments. We calculate the distance from each segment of the filamentary structures defined by galaxy samples (i.e., LAEs, LBGs, and all galaxies) to the filamentary structures defined by dark matter particles. We calculate the median value of these distances to refer to the median spatial offset between the filamentary structures of galaxy samples and dark matter.

Figure 11 shows the median offsets for three redshifts. Here, the vertical error bars represent the standard deviation of the median offset values from 1000 random subsamples of our samples. As in Figures 7 and 9, we present the mean separation of the sample, which is used for extracting filamentary structures. Galaxy samples, including LAEs and LBGs, have median values smaller than the mean separation of them (i.e., 10.5 cMpc). As in Sections 3.2 and 3.4, there is a trend that

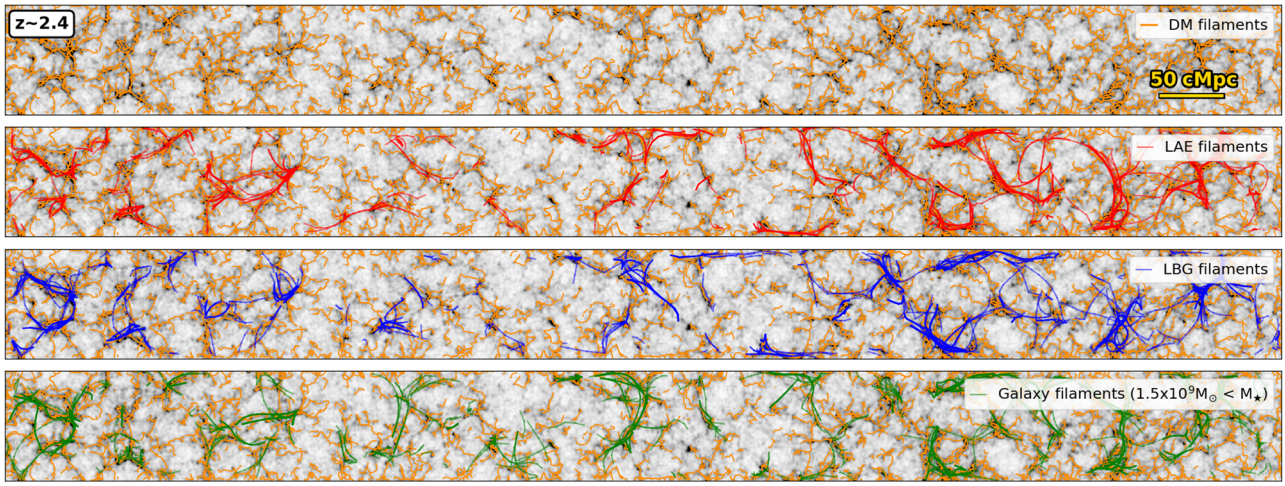

Figure 10. Two-dimensional view of filamentary structures from five random subsamples of LAEs, LBGs, and all galaxies with stellar masses larger than $1.5 \times 10^9\ M_\odot$ at $z \sim 2.4$. We present the distribution of dark matter and their filamentary structures as background in each panel. For comparison, we plot the filaments within the same region in Figure 2.

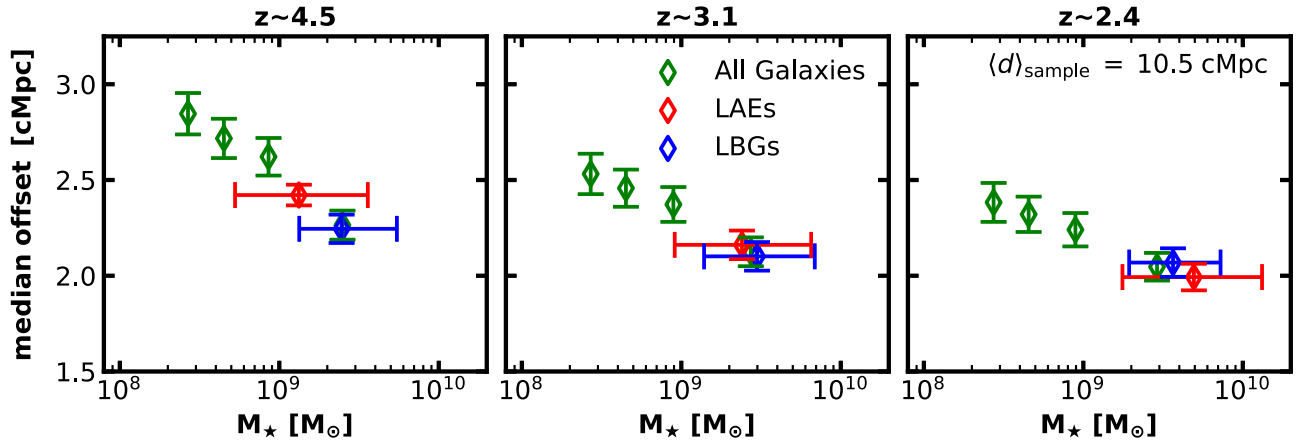

Figure 11. Median spatial offset between filamentary structures traced by galaxy samples (i.e., LAEs, LBGs, and all galaxies) and dark matter particles. Here, we also present the mean separation of the sample, which is used for extracting filamentary structures. The vertical error bars represent the standard deviation of mean offsets from 1000 different random subsamples, while the horizontal error bars for LAEs and LBGs represent their 15.9 and 84.1 percentiles of stellar mass.

massive galaxies have smaller median offset values than less massive galaxies. This is because more massive galaxies reside in denser regions than less massive galaxies. LAEs and LBGs again follow the stellar mass trend well, implying that they have no systematic bias compared to the entire galaxy population. We also perform a similar analysis for random subsamples with a larger sample size and find a similar result (see Appendix D).

4. Discussion

We perform three sets of analyses of the filamentary structures traced by LAEs, LBGs, all galaxies, and dark matter particles in Section 3. Here, we discuss the results of those analyses. First, in Section 4.1, we compare our results with previous studies based on observations and simulations. In Section 4.2, we discuss the capability of LAEs and LBGs as tracers of large-scale structures of dark matter at high redshifts.

4.1. Comparison with Other Studies

We first focus on the relation between the stellar mass of galaxies and their distance from filaments. Many observational studies examine how filamentary structures affect the growth of stellar mass in galaxies. For instance, Kraljic et al. (2017) analyze the local ($0.03 \leqslant z \leqslant 0.25$) galaxies from the Galaxy and Mass Assembly (Driver et al. 2009) redshift survey to examine how galaxy properties change within filamentary structures. They find that the galaxies with larger stellar masses are more concentrated toward the filaments than those with smaller stellar masses. Laigle et al. 2018 also find similar results with galaxies from the COSMOS2015 (Laigle et al. 2018) photometric redshift catalog at a redshift range of $0.5 < z < 0.9$. Some studies based on simulations also find a similar relation between the mass of halos and their distance to filaments (e.g., Jhee et al. 2022). Our simulation-based results in Sections 3.2 and 3.4 are consistent with these results. These results also support the idea of the hierarchical structure in the sense that more massive galaxies would be formed and evolved in denser regions of dark matter.

On the other hand, there are some observational studies focusing on the relationship between the star formation activity of galaxies and their distance from filaments. Again, Kraljic et al. (2017) and Laigle et al. (2018) find that passive galaxies reside closer to the filamentary structures than star-forming galaxies in the local universe. This environment dependence is expected to be reversed at high redshifts because of the larger

amount of gas in high-redshift galaxies (i.e., the reversal of the SFR–density relation, Elbaz et al. 2007; Hwang et al. 2019; Song et al. 2021; Martín-Navarro et al. 2024). Indeed, we could see a hint of such a trend (not shown here), which needs more detailed analysis by simultaneously considering the mass effect as well; this will be the topic of future studies.

4.2. LAEs and LBGs as Tracers of Large-scale Structures

We discuss the capability of LAEs and LBGs as tracers of large-scale structures of dark matter at high redshifts. In Section 3.2, we examine the spatial distribution of our samples around the filamentary structures of dark matter particles and find that LAEs and LBGs are more concentrated toward the dark matter filaments than dark matter particles themselves. We also examine the spatial distribution of our samples around the filamentary structures defined by themselves in Section 3.4. We find that LAEs and LBGs are again more concentrated toward filaments than dark matter. In both Sections 3.2 and 3.4, we find a stellar mass trend that massive galaxies are more concentrated toward filaments. This trend could be understood by the fact that massive galaxies have larger galaxy bias (i.e., higher probability to be formed in dense regions) than less massive ones (e.g., Kaiser 1984; Mo & White 1996). It is important to note that LAEs and LBGs follow this stellar mass trend of the entire galaxy population well without systematic bias. These results, combined with the fact that they are preferred tracers in observations, suggest that LAEs and LBGs can be good tracers of filamentary structures of dark matter at high redshifts.

In Section 3.5, we also analyze the difference between filamentary structures traced by galaxy samples (i.e., LAEs, LBGs, and all galaxies with various stellar mass ranges) and dark matter particles. The filamentary structures traced by galaxy samples look similar to those by dark matter particles in terms of overall shape. We also calculate the spatial offsets between filamentary structures and find that the median offsets of the galaxy samples are smaller than the mean separation of the sample. We also find a trend that massive galaxies have smaller median offsets than less massive galaxies. LAEs and LBGs again seem to follow this mass trend, suggesting that there is no systematic bias of LAEs and LBGs as tracers of filamentary structures of dark matter at high redshifts compared to the entire galaxy population. However, it is important to note that galaxy samples, including LAEs and LBGs cannot trace all the features of large-scale structures of dark matter. We confirm that galaxies, including LAEs and LBGs can trace the overall shape of dark matter filaments with spatial offsets smaller than the mean separation of them.

Furthermore, the LBG samples from each snapshot data in this study cannot be directly compared to the photometrically selected LBGs in observations, which have wide redshift ranges ($\Delta z \sim 0.5$). These wide redshift ranges would make it difficult to trace three-dimensional large-scale structures using the projected distribution of LBGs from observations. On the other hand, LAEs from narrowband observations have narrower redshift ranges ($\Delta z \lesssim 0.04-0.08$). Therefore, it would be more efficient to use LAEs to study high-z large-scale structures in observations, such as the ODIN survey, than to use LBGs.

5. Conclusions

We have tested the capability of LAEs and LBGs as tracers of large-scale structures of dark matter at high redshifts using the data from the Horizon Run 5 cosmological hydrodynamical simulation at three redshifts: $z \sim 2.4$, 3.1, and 4.5. To do that, we have generated samples of LAEs, LBGs, all galaxies in various stellar mass ranges, and dark matter particles. We have focused on the filamentary structures of these samples, which were extracted by the DisPerSE code. We have performed three different tests on the spatial distributions and filamentary structures of these samples and found the following results.

(I) We have examined the spatial distributions of our samples of LAEs, LBGs, all galaxies, and dark matter particles around the filamentary structure defined by dark matter particles. The result shows that LAEs and LBGs are more concentrated toward the dark matter filaments than dark matter particles. We also find that more massive galaxies are more concentrated than less massive galaxies, which is consistent with the previous results from observations and simulations.

(II) We find an empirical fitting formula as Equation (10) for the vertical density profile around the filamentary structures. The formula works well within 2 cMpc from the filamentary structures. Beyond 2 cMpc, the density profiles for all of our samples become smaller than those of randomly distributed particles.

(III) We have compared the spatial distribution of our samples around the filamentary structure defined by themselves. Again, LAEs and LBGs are more concentrated toward their filaments than dark matter, suggesting that they form thinner filamentary structures than dark matter. Furthermore, more massive galaxies are again more concentrated toward their filaments.

(IV) The filamentary structures traced by galaxy samples (i.e., LAEs, LBGs, and all galaxies) and dark matter particles show good agreement in terms of their overall shapes. The median spatial offsets between the filamentary structures of galaxy samples and dark matter particles are smaller than the mean separation of the galaxy sample used for extracting filaments. We also find a stellar mass trend that more massive galaxies have smaller median spatial offsets from dark matter filaments.

(V) We also find that our LAE and LBG samples follow well the stellar mass trend in the results (I), (II), and (IV). Although the stellar mass seems to be a dominant factor for the capability as the tracers of dark matter filaments, our results suggest that there is no systematic bias of LAEs and LBGs as such tracers compared to the entire galaxy population.

These results support the idea that LAEs and LBGs could be good tracers of large-scale structures of dark matter at high redshifts. It should be noted that we have analyzed three-dimensional distributions of LAEs and LBGs in Horizon Run 5. Because photometric surveys like the ODIN survey will provide the data for the two-dimensional distributions of LAEs and LBGs, we need to analyze the two-dimensional mock data of HR5 and compare it with the observational results. Furthermore, to generate more realistic LBG samples, we need the light cone data of cosmological simulations with a continuous redshift range. We plan to construct the light cone

data of HR5 and will perform analyses similar to this study and compare the results with the ODIN observed data.

Acknowledgments

H.S.H. acknowledges the support by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT), NRF-2021R1A2C1094577 and by Samsung Electronic Co., Ltd. (project No. IO220811-01945-01). We thank the support from the KISTI National Supercomputing Center and its Nurion Supercomputer through the Grand Challenge Program (KSC-2018-CHA-0003, KSC-2019-CHA-0002). Large data transfer was supported by KREONET, which is managed and operated by KISTI. This work is also supported by the Center for Advanced Computation at Korea Institute for Advanced Study. J.P. was supported by a KIAS Individual grant (PG078702) at the Korea Institute for Advanced Study. J.L. is supported by the National Research Foundation of Korea (NRF-2021R1C1C2011626). This work was partially supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2022M3K3A1093827). Y.K. is supported by the Korea Institute of Science and Technology Information (KISTI) under the institutional R&D project (K24L2M1C4). E.G. would also like to acknowledge support from NSF grant AST-2206222. O.N.S. acknowledges support from ERC Consolidator Grant "GAIA-BIFROST" (Grant Agreement ID 101003096) and the UK Science and Technology Facilities Council (Consolidated Grant ST/V000721/1; Small Award ST/Y002695/1 M.C.A. acknowledges support from ANID BASAL project FB210003. This work is partially supported by grant Segal ANR-19-CE31-0017 of the French Agence Nationale de la Recherche.

Appendix A
Results from a Higher Minimum Persistence Level

We adopt 5σ as the minimum persistence level when we extract filamentary structures of dark matter particles in Section 3.1. Here, we present how the choice of the minimum persistence level affects the results of our analysis in Section 3. We calculate again the median values of $D_{\rm fil,DM}$ of our samples with the filamentary structures of dark matter extracted with 10σ minimum persistence level. As mentioned in Section 3.1, filaments with higher persistence levels represent more robust and less detailed structures. We present the results in Figure 12 along with those of 5σ level. The absolute values of median $D_{\rm fil,DM}$ become larger ($\Delta D_{\rm fil,DM} \lesssim 0.1$ cMpc) than those with 5σ level because the mean separation between filaments becomes larger. Again, LAEs, LBGs, and all galaxies have smaller median values than dark matter particles. We also find LAEs and LBGs follow well the stellar mass trend of all galaxies as in Section 3.2. These indicate that the choice of the minimum persistent level of filamentary structures would not affect our main results.

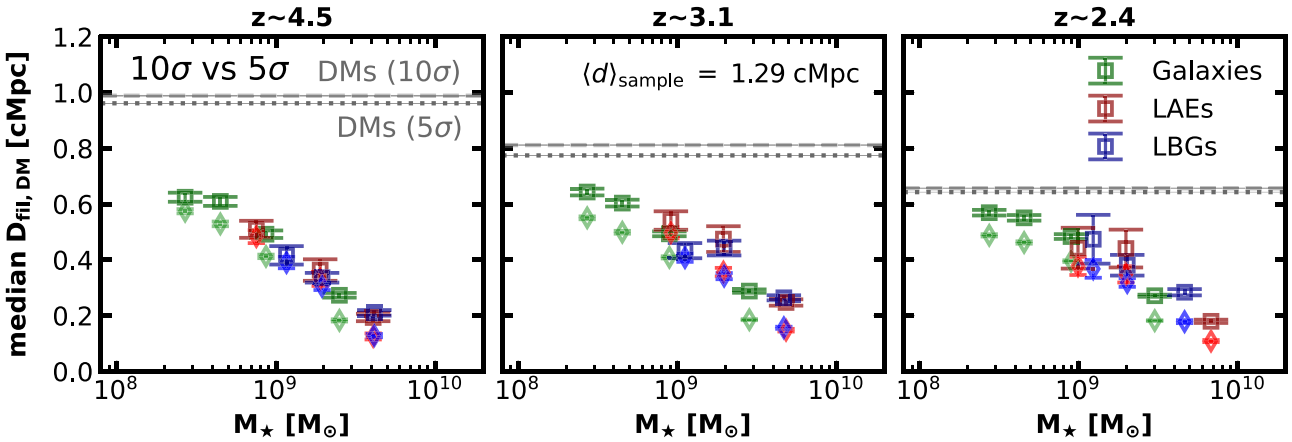

Figure 12. Median $D_{\rm fil,DM}$ value of each sample at three redshifts with 10σ (darker square) and 5σ (fainter diamond markers) minimum persistence levels. The meaning of the error bars and shaded regions are the same as in Figure 7.

Appendix B
Results with the T-ReX Filament Finder

We also perform a similar analysis to Section 3.2 with the filamentary structures extracted with the T-ReX filament finder (Bonnaire et al. 2020). T-ReX finds filamentary structures as the ridges of the given data points (here, galaxies or dark matter particles) using graph theory and the Gaussian mixture model. We would like to emphasize that the mathematical definition of filaments in T-ReX is different from that in DisPerSE. Therefore, comparing the results of the two algorithms could show the robustness of our results with respect to how the filaments are defined and identified. A comprehensive comparison between various filament finders is available in Libeskind et al. (2018).

Figure 13 compares the median value of $D_{\rm fil,DM}$ of the samples using the dark matter filaments from T-ReX and DisPerSE. Although the absolute values of the median $D_{\rm fil,DM}$ are larger than those with DisPerSE filaments ($\Delta D_{\rm fil,DM} \lesssim 0.5$ cMpc), the overall trend is similar. This result shows that our main conclusions would not change much when we use different filament finders for our analyses.

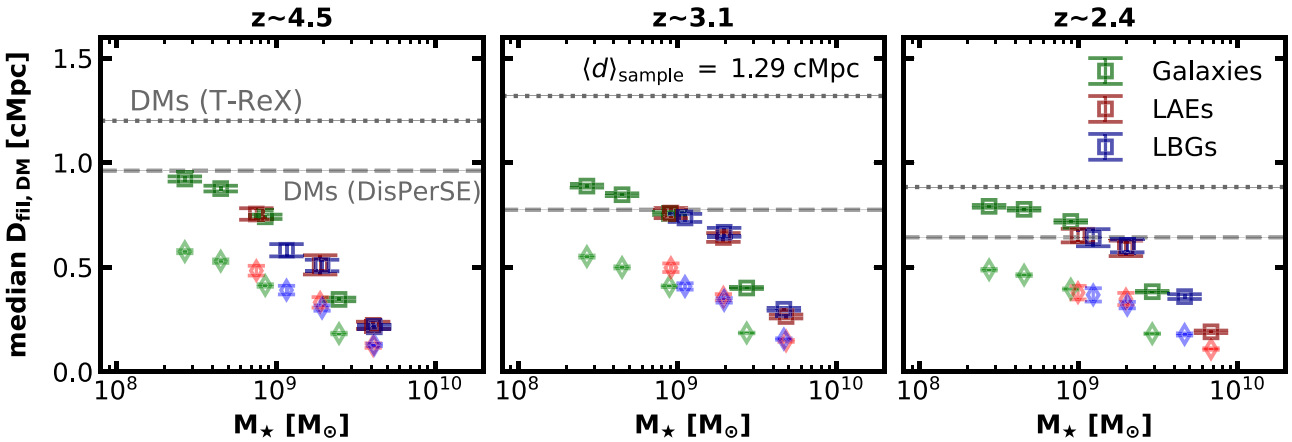

Figure 13. Median $D_{\rm fil,DM}$ value of each sample at three redshifts using T-Rex (darker square) and DisPerSE (fainter diamond markers) as the filament finder. The meaning of the error bars and shaded regions are the same as in Figure 7.

Appendix C
Fitting Results for All Samples

In Section 3.3, we find an empirical fitting formula for the vertical density profile around filamentary structures as Equation (10). Here, we present the fitting results of all of our samples at three redshifts in Figure 14. The results show that our empirical formula works well for all of our samples at three redshifts within ∼2 cMpc regardless of redshift. We also present the best-fit values for the fitting parameters (i.e., Σ_0, D_c, and α) in Table 2.

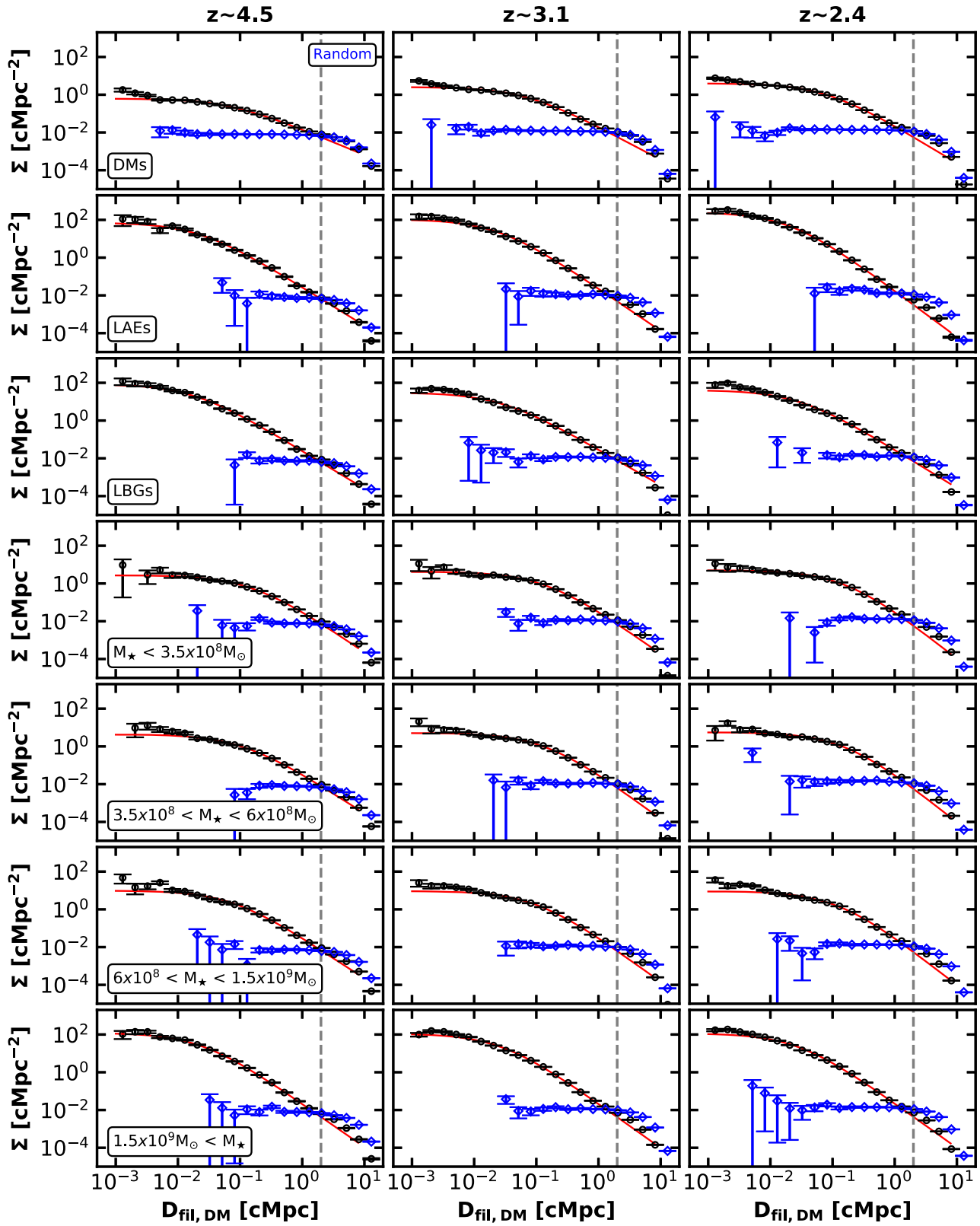

Figure 14. The vertical density profiles for all of our samples (black markers) and random particles (blue markers) for the case of the 5σ minimum persistence level. Above ∼2 cMpc (gray dashed line in each panel), the density profiles of our samples become smaller than those of random particles. The red solid line in each panel represents the best-fit curve for each sample using Equation (10).

Table 2
Best-fit Values of the Parameters in the Fitting Formula for the Vertical Density Distributions in Section 3

Redshift	Parameters	DM Particles	LAEs	LBGs	All Galaxies			
					Bin 1	Bin 2	Bin 3	Bin 4
$z \sim 2.4$	Σ_0 [cMpc^{-2}]	3.99 ± 0.29	256 ± 37	98.4 ± 19.5	5.02 ± 0.48	5.60 ± 0.57	8.98 ± 1.18	114 ± 20
	D_c [cMpc]	0.079 ± 0.008	0.020 ± 0.003	0.032 ± 0.006	0.158 ± 0.024	0.155 ± 0.025	0.120 ± 0.022	0.028 ± 0.005
	α	2.01 ± 0.05	2.45 ± 0.07	2.44 ± 0.12	2.59 ± 0.15	2.62 ± 0.16	2.58 ± 0.16	2.37 ± 0.09
$z \sim 3.1$	Σ_0 [cMpc^{-2}]	2.55 ± 0.24	106 ± 16	114 ± 18	4.17 ± 0.42	5.13 ± 0.46	9.31 ± 1.03	111 ± 18
	D_c [cMpc]	0.066 ± 0.009	0.024 ± 0.003	0.026 ± 0.004	0.133 ± 0.021	0.134 ± 0.018	0.098 ± 0.015	0.027 ± 0.004
	α	1.77 ± 0.06	2.26 ± 0.06	2.32 ± 0.08	2.31 ± 0.12	2.42 ± 0.11	2.39 ± 0.11	2.36 ± 0.08
$z \sim 4.5$	Σ_0 [cMpc^{-2}]	0.62 ± 0.03	71.1 ± 18.4	125 ± 20	2.71 ± 0.18	4.27 ± 0.61	9.82 ± 1.79	122 ± 20
	D_c [cMpc]	0.067 ± 0.006	0.022 ± 0.005	0.023 ± 0.004	0.149 ± 0.017	0.106 ± 0.022	0.075 ± 0.017	0.024 ± 0.004
	α	1.37 ± 0.03	2.04 ± 0.10	2.30 ± 0.07	2.22 ± 0.09	2.11 ± 0.14	2.19 ± 0.14	2.30 ± 0.07

Appendix D
Results from Random Subsamples with a Larger Sample Size

In Section 3.5, we analyze the median spatial offset values between filamentary structures traced by galaxy samples and dark matter particles using randomly selected subsamples with a size of 7000. Here, we perform a similar analysis but for the subsamples constructed with a larger sample size of 10,000. We exclude the LAE sample at $z \sim 4.5$ which has a smaller sample size than 10,000. Figure 15 shows the median spatial offsets between filamentary structures of galaxy samples and dark matter. Again, we find that the median offsets are smaller than the mean separation of the sample and the stellar mass trend that massive galaxies show smaller offsets than less massive ones. LAEs and LBGs follow this trend well, indicating that they have no systematic bias as tracers of filamentary structures of dark matter compared to the entire galaxy population.

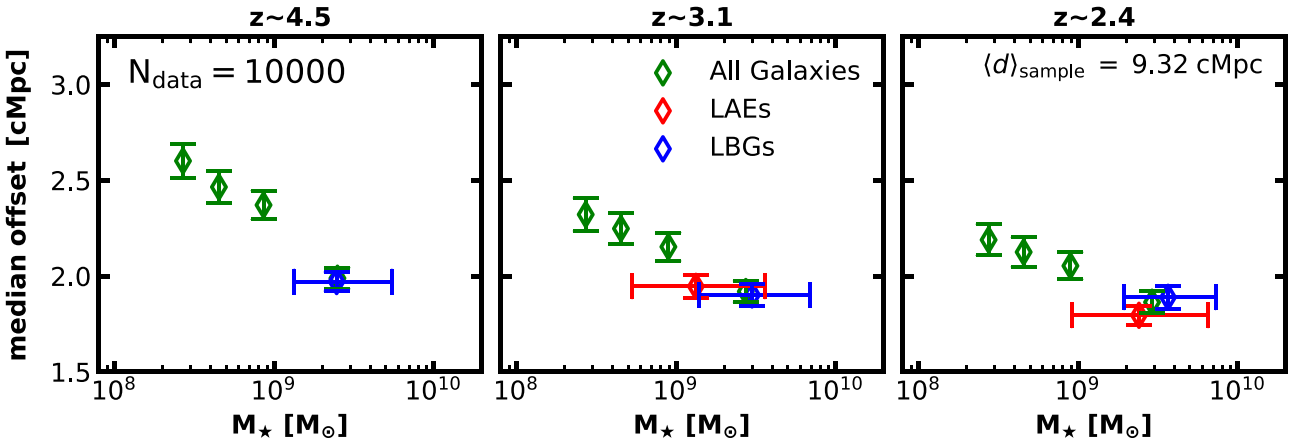

Figure 15. Median spatial offset between filamentary structures traced by galaxy samples (i.e., LBGs and all galaxies of several stellar mass ranges) and dark matter particles. The error bars and shaded regions represent the standard deviation of mean offsets from 1000 different random subsamples. The mean offset values of the samples are similar to each other and smaller than the mean separation of them.

ORCID iDs

Sang Hyeok Im https://orcid.org/0009-0003-9748-4194
Ho Seong Hwang https://orcid.org/0000-0003-3428-7612
Jaehong Park https://orcid.org/0000-0003-3095-6137
Jaehyun Lee https://orcid.org/0000-0002-6810-1778
Hyunmi Song https://orcid.org/0000-0002-4362-4070
Stephen Appleby https://orcid.org/0000-0001-8227-9516
Yohan Dubois https://orcid.org/0000-0003-0225-6387
C. Gareth Few https://orcid.org/0000-0003-3940-7687
Brad K. Gibson https://orcid.org/0000-0003-4446-3130
Juhan Kim https://orcid.org/0000-0002-4391-2275
Yonghwi Kim https://orcid.org/0000-0003-4164-5414
Changbom Park https://orcid.org/0000-0001-9521-6397
Christophe Pichon https://orcid.org/0000-0003-0695-6735
Jihye Shin https://orcid.org/0000-0001-5135-1693
Owain N. Snaith https://orcid.org/0000-0003-1414-1296
M. Celeste Artale https://orcid.org/0000-0003-0570-785X
Eric Gawiser https://orcid.org/0000-0003-1530-8713
Lucia Guaita https://orcid.org/0000-0002-4902-0075
Woong-Seob Jeong https://orcid.org/0000-0002-2770-808X
Kyoung-Soo Lee https://orcid.org/0000-0003-3004-9596
Nelson Padilla https://orcid.org/0000-0001-9850-9419
Vandana Ramakrishnan https://orcid.org/0000-0002-9176-7252
Paulina Troncoso https://orcid.org/0000-0001-6162-3023
Yujin Yang https://orcid.org/0000-0003-3078-2763

References

Aihara, H., AlSayyad, Y., Ando, M., et al. 2022, PASJ, 74, 247
Aihara, H., Arimoto, N., Armstrong, R., et al. 2018a, PASJ, 70, S4
Aihara, H., Armstrong, R., Bickerton, S., et al. 2018b, PASJ, 70, S8
Behrens, C., Byrohl, C., Saito, S., & Niemeyer, J. C. 2018, A&A, 614, A31
Bonnaire, T., Aghanim, N., Decelle, A., & Douspis, M. 2020, A&A, 637, A18
Bouwens, R. J., Illingworth, G. D., Franx, M., & Ford, H. 2007, ApJ, 670, 928
Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039
Davis, M., Huchra, J., Latham, D. W., & Tonry, J. 1982, ApJ, 253, 423
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016, arXiv:1611.00036
Dijkstra, M., & Wyithe, J. S. B. 2012, MNRAS, 419, 3181
Dong, F., Park, C., Hong, S. E., et al. 2023, ApJ, 953, 98
Driver, S. P., Norberg, P., Baldry, I. K., et al. 2009, A&G, 50, 5.12
Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560
Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33
Elbaz, D., Dickinson, M., Hwang, H. S., et al. 2011, A&A, 533, A119
Firestone, N. M., Gawiser, E., Ramakrishnan, V., et al. 2023, arXiv:2312.16075
Gawiser, E., Francke, H., Lai, K., et al. 2007, ApJ, 671, 278
Geller, M. J., & Huchra, J. P. 1989, Sci, 246, 897
Giavalisco, M., Dickinson, M., Ferguson, H. C., et al. 2004, ApJL, 600, L103
Gregory, S. A., & Thompson, L. A. 1978, ApJ, 222, 784
Gronke, M., Dijkstra, M., Trenti, M., & Wyithe, S. 2015, MNRAS, 449, 1284
Hamilton, A. J. S. 1993, ApJ, 417, 19
Hawkins, E., Maddox, S., Cole, S., et al. 2003, MNRAS, 346, 78
Hong, S., Jeong, D., Hwang, H. S., et al. 2020, MNRAS, 493, 5972
Hu, E. M., & McMahon, R. G. 1996, Natur, 382, 231
Hwang, H. S., Geller, M. J., Park, C., et al. 2016, ApJ, 818, 173
Hwang, H. S., Shin, J., & Song, H. 2019, MNRAS, 489, 339
Iyer, K., Gawiser, E., Davé, R., et al. 2018, ApJ, 866, 120
Jhee, H., Song, H., Smith, R., et al. 2022, ApJ, 940, 2
Kaiser, N. 1984, ApJL, 284, L9
Kawamata, R., Ishigaki, M., Shimasaku, K., et al. 2018, ApJ, 855, 4
Kim, J., Lee, J., Laigle, C., et al. 2023, ApJ, 951, 137
Kirshner, R. P., Oemler, A. J., & Schechter, P. L. 1978, AJ, 83, 1549
Kraljic, K., Arnouts, S., Pichon, C., et al. 2017, MNRAS, 474, 547
Laigle, C., Pichon, C., Arnouts, S., et al. 2018, MNRAS, 474, 5437
Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64
Lee, J., Shin, J., Snaith, O. N., et al. 2021, ApJ, 908, 11
Lee, K.-S., Dey, A., Hong, S., et al. 2014, ApJ, 796, 126
Lee, K.-S., Gawiser, E., Park, C., et al. 2024, ApJ, 962, 36
Libeskind, N. I., van de Weygaert, R., Cautun, M., et al. 2018, MNRAS, 473, 1195
Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, ApJS, 184, 218
Lin, H., Kirshner, R. P., Shectman, S. A., et al. 1996, ApJ, 471, 617
LSST Dark Energy Science Collaboration 2012, arXiv:1211.0310
Madau, P. 1995, ApJ, 441, 18
Malavasi, N., Arnouts, S., Vibert, D., et al. 2016, MNRAS, 465, 3817
Martín-Navarro, I., de Lorenzo-Cáceres, A., Gadotti, D. A., et al. 2024, A&A, 684, A110
Mason, C. A., Treu, T., Dijkstra, M., et al. 2018, ApJ, 856, 2
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Oke, J. B., & Gunn, J. E. 1983, ApJ, 266, 713
Ouchi, M., Shimasaku, K., Furusawa, H., et al. 2003, ApJ, 582, 60
Oyarzún, G. A., Blanc, G. A., González, V., Mateo, M., & Bailey, J. I. I. 2017, ApJ, 843, 133
Park, C. 1990, MNRAS, 242, 59P
Park, C., Choi, Y.-Y., Kim, J., et al. 2012, ApJL, 759, L7
Park, C., Lee, J., Kim, J., et al. 2022, ApJ, 937, 15
Peebles, P. J. E. 1975, ApJ, 196, 647
Percival, W. J., Nichol, R. C., Eisenstein, D. J., et al. 2007, ApJ, 657, 645
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13
Ramakrishnan, V., Moon, B., Im, S. H., et al. 2023, ApJ, 951, 119
Salmon, B., Papovich, C., Finkelstein, S. L., et al. 2015, ApJ, 799, 183
Shectman, S. A., Landy, S. D., Oemler, A., et al. 1996, ApJ, 470, 172
Silverman, J. D., Kashino, D., Sanders, D., et al. 2015, ApJS, 220, 12
Sohn, J., Geller, M. J., Hwang, H. S., et al. 2021, ApJ, 909, 129
Song, H., Laigle, C., Hwang, H. S., et al. 2021, MNRAS, 501, 4635
Sousbie, T. 2011, MNRAS, 414, 350
Sousbie, T., Pichon, C., & Kawahara, H. 2011, MNRAS, 414, 384
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Steidel, C. C., Adelberger, K. L., Dickinson, M., et al. 1998, ApJ, 492, 428
Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728
Teyssier, R. 2002, A&A, 385, 337
Toshikawa, J., Kashikawa, N., Overzier, R., et al. 2016, ApJ, 826, 114
Weinberger, L. H., Haehnelt, M. G., & Kulkarni, G. 2019, MNRAS, 485, 1350
York, D. G., Adelman, J., Anderson, J. E. J., et al. 2000, AJ, 120, 1579
Zheng, Z.-Y., Wang, J.-X., Malhotra, S., et al. 2014, MNRAS, 439, 1101