



The Galaxy–Halo Connection of DESI Luminous Red Galaxies with Subhalo Abundance Matching

Angela M. Berti¹ , Kyle S. Dawson¹ , and Wilber Dominguez²¹ Department of Physics & Astronomy, University of Utah, Salt Lake City, UT 84112, USA² Department of Physics & Astronomy, Swarthmore College, 500 College Ave., Swarthmore, PA 19081, USA; aberti@ucsd.edu

Received 2023 March 28; revised 2023 June 23; accepted 2023 July 12; published 2023 August 29

Abstract

We use subhalo abundance and age distribution matching to create magnitude-limited mock galaxy catalogs at $z \sim 0.43, 0.52$, and 0.63 with z -band and $3.4 \mu\text{m}$ W1-band absolute magnitudes and $r - z$ and $r - \text{W1}$ colors. From these magnitude-limited mocks, we select mock luminous red galaxy (LRG) samples according to the $(r - z)$ -based (optical) and $(r - \text{W1})$ -based (infrared) selection criteria for the LRG sample of the Dark Energy Spectroscopic Instrument (DESI) survey. Our models reproduce the number densities, luminosity functions, color distributions, and projected clustering of the DESI Legacy Surveys that are the basis for DESI LRG target selection. We predict the halo occupation statistics of both optical and IR DESI LRGs at fixed cosmology and assess the differences between the two LRG samples. We find that IR-based SHAM modeling represents the differences between the optical and IR LRG populations better than using the z band and that age distribution matching overpredicts the clustering of LRGs, implying that galaxy color is uncorrelated with halo age in the LRG regime. Both the optical and IR DESI LRG target selections exclude some of the most luminous galaxies that would appear to be LRGs based on their position on the red sequence in optical color–magnitude space. Both selections also yield populations with a nontrivial LRG–halo connection that does not reach unity for the most massive halos. We find that the IR selection achieves greater completeness ($\gtrsim 90\%$) than the optical selection across all redshift bins studied.

Unified Astronomy Thesaurus concepts: [Large-scale structure of the universe \(902\)](#); [Galaxies \(573\)](#)

1. Introduction

The Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016a, 2016b) survey is a spectroscopic galaxy redshift survey of unprecedented scale that will classify tens of millions of galaxies in four target classes over $\sim 14,000 \text{ deg}^2$ at the lowest redshifts of this program. DESI’s Bright Galaxy Survey (BGS; Hahn et al. 2022) extends to $z \sim 0.4$, while the luminous red galaxy (LRG; Zhou et al. 2023) sample will reach to $z \sim 1$ and cover 20 times the volume of the BGS.³

The DESI target samples are optimized for precision measurements of cosmological parameters. However, DESI also offers novel opportunities to study galaxy evolution and the high-mass end of the stellar-to-halo mass relation (SHMR), provided that sample selection effects are well understood.

Unlike the magnitude-limited BGS sample of relatively nearby galaxies, the DESI LRG sample is selected with a comparatively complex set of magnitude and color cuts, creating incompleteness that may depend on any combination of galaxy color, stellar mass, and redshift. However, the LRG sample covers a volume 20 times larger than that of the BGS sample and will contain a much higher number density of spectroscopic redshifts at $0.4 < z < 1$ than any previous spectroscopic galaxy redshift survey, making it a good sample for statistical studies with negligible sample and cosmic variance uncertainties. The rarity and associated low number density of massive galaxies ($\sim 2 \times 10^{-5} \text{ Mpc}^{-3}$ for

$\log(M_*/M_\odot) > 11.5$) means that such large volumes are essential for obtaining sample sizes large enough for statistically significant measurements of the high-mass end of the galaxy stellar mass function (SMF) and the SHMR of LRGs.

Existing studies of the galaxy–halo relationship for “DESI-like” LRG samples include halo occupation distribution (HOD) modeling (e.g., Seljak 2000; Berlind & Weinberg 2002; Bullock et al. 2002; Berlind et al. 2003; Zheng & Weinberg 2007; Zheng et al. 2007) in Zhou et al. (2021) and semianalytic modeling in Hernández-Aguayo et al. (2021) with GALFORM (Cole et al. 2000), which predicts absolute magnitudes with dust attenuation.

In its simplest form, the HOD model provides a statistical description for how galaxies occupy dark matter halos solely as a function of halo mass. Zhou et al. (2021) fit the projected clustering of DESI-like LRGs measured in five redshift bins in the range $0.4 < z < 0.9$ with a five-parameter HOD model plus a sixth nuisance parameter to account for photometric redshift uncertainties. They found similar HOD parameters at $0.4 < z < 0.8$ and statistically significant differences in the model parameters for only the highest redshift bin ($0.8 < z < 0.9$). While Zhou et al. (2021) demonstrated that clustering measurements using photometric redshifts are sufficient to constrain the HOD parameters of DESI-like LRGs, the standard HOD framework they used does not accommodate galaxy populations with halo occupation fractions that do not reach unity for the most massive halos.

Hernández-Aguayo et al. (2021) applied the GALFORM semianalytic model (SAM) to the Planck-Millennium cosmological N -body simulation (Baugh et al. 2018). They then converted predicted absolute magnitudes to apparent magnitudes at various redshift snapshots, enabling the DESI LRG target selection to be applied directly to GALFORM mock

³ DESI will observe emission line galaxies to $z \sim 1.6$ and quasars to $z \sim 3.5$.



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

galaxy catalogs to select mock LRG samples. Hernández-Aguayo et al. (2021) found that the DESI LRG selection criteria exclude a small but important fraction of the most massive galaxies ($\log(M_*/M_\odot) > 11.15$). Consequently, their model predicts that the halo occupation fraction of LRGs does not reach unity for the most massive halos and actually drops with increasing mass, indicative of a nontrivial LRG–halo connection that is not modeled well with a standard HOD. By comparing the HOD and subhalo mass functions of stellar mass–selected mock galaxies against those of mock LRG samples, Hernández-Aguayo et al. (2021) showed that the DESI LRG selection cuts likely affect the selection of subhalos populated by LRGs; i.e., (sub)halo mass is, by itself, insufficient to determine whether a subhalo hosts an LRG.

Subhalo abundance matching (SHAM; e.g., Vale & Ostriker 2004; Conroy et al. 2006; Behroozi et al. 2010; Reddick et al. 2013) is an empirical technique for assigning galaxies to dark matter halos in numerical N -body simulations by assuming a correlation between a galaxy property—usually luminosity or stellar mass—and a halo property such as mass or circular velocity. In the simplest application of SHAM, mock galaxy catalogs are constructed to reproduce the number density and luminosity or SMF of a target data set with a single free parameter to allow scatter in the galaxy–halo property correlation.

Extensions to the SHAM framework to incorporate dependencies between additional galaxy and halo properties are broadly referred to as conditional abundance matching (CAM; e.g., Hearin et al. 2014; Zentner et al. 2014). In addition to a primary galaxy–halo property correlation, CAM models assume a correlation between secondary galaxy and halo properties at a fixed value of the primary property. Age distribution matching (Hearin & Watson 2013) is a form of CAM that equates galaxy color or a similar property with a proxy for the age of dark matter halos. Unlike the standard HOD framework, in which the statistical relationship between galaxies and dark matter halos is solely a function of stellar and halo mass, CAM can naturally accommodate galaxy assembly bias, the dependence of galaxy properties (besides stellar mass) on the mass accretion history of their host halos (e.g., Zentner et al. 2014; Wechsler & Tinker 2018, and references therein).

The SHAM framework has been used to study the dependence of sample completeness on redshift and stellar mass for LRGs from the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), which obtained spectroscopic redshifts of 1.5 million galaxies with $\log(M_*/M_\odot) > 11$ to $z \sim 0.7$. BOSS contains two color- and magnitude-selected samples of massive galaxies ($\log(M_*/M_\odot) > 11$; Reid et al. 2015): the LOWZ sample of LRGs at $0.15 < z < 0.43$ and the approximately stellar mass-limited constant mass (CMASS) sample, which includes galaxies of all colors at $0.43 < z < 0.8$.

Saito et al. (2016) used SHAM to construct $z \sim 0.5$ mock galaxy catalogs for the BOSS CMASS sample with and without added assembly bias effects. They used the Stripe 82 Massive Galaxy Catalog (Bundy et al. 2015) to replicate the total galaxy SMF above $\log(M_*/M_\odot) > 10.5$ over $0.43 < z < 0.7$ and assign galaxies to halos in the MultiDark simulation (Riebe et al. 2013). Saito et al. (2016) found that assembly bias does factor into the galaxy–halo connection for high-mass galaxies; i.e., the SHMR for these galaxies has some dependence on galaxy color and should not be inferred from the clustering signal without any consideration of color.

Yu et al. (2022) modeled LRGs from BOSS and eBOSS (Dawson et al. 2016) at $0.2 < z < 1.0$ with a SHAM framework that includes two additional free parameters to account for redshift uncertainty and sample incompleteness.

SHAM has also been used to model DESI-like samples at low redshift. Safonova et al. (2021) used SHAM and age distribution matching to create $z \sim 0.1$ mock galaxy catalogs representative of the DESI BGS sample with r -band luminosities and $g - r$ colors. The low-redshift range of the BGS sample allowed them to utilize spectroscopic redshifts from the Sloan Digital Sky Survey (SDSS; York et al. 2000) and Galaxy and Mass Assembly (Loveday et al. 2012) project.

In this work, we use SHAM and age distribution matching to create magnitude-limited mock galaxy catalogs at multiple redshifts within the redshift range of the DESI LRG sample. Two distinct target selection algorithms were considered for the DESI LRG sample: an optical selection that uses $r - z$ color and an infrared selection that uses $r - W1$ ⁴ color. We select mock LRG samples from our magnitude-limited mocks based on both the optical and IR DESI target selections that match the number density and two-dimensional color–magnitude space distribution of each DESI LRG target sample. We then predict the clustering signal and halo occupation statistics of these samples as a function of redshift.

Our method is a novel approach to modeling DESI LRGs that complements existing SAM and HOD models by utilizing the full photometric samples that are the basis of DESI LRG target selection. We offset the precision of the training data lost to photometric redshift errors by driving down cosmic variance uncertainties with complete photometric samples from an unprecedented survey volume. Finally, this work offers a comparative study of the samples selected by the optical and IR selection algorithms considered for DESI LRGs.

The structure of this paper is as follows. Section 2 describes the cosmological simulation and photometric galaxy samples used in this work. Section 3 describes our modeling procedure and the two-point statistics used to constrain our models. In Section 4, we present the predicted properties of the LRG samples, and in Section 5, we summarize the conclusions of this work. Where applicable, we assume the cosmological parameter values of Planck Collaboration et al. (2016), $h = 0.6777$ and $\Omega_m = 0.307115$, under the assumption of a flat, Λ CDM cosmological model.

2. Simulations and Data

In this section, we describe the cosmological simulation, halo finder, and associated halo properties used for our models. We also present the data from which we select parent galaxy samples for training our models, as well as the DESI LRG target selection functions.

2.1. Simulations

We use halo catalogs and merger histories obtained with the publicly available ROCKSTAR phase-space temporal halo finder (Behroozi et al. 2012b) for the MultiDark Planck 2 (MDPL2) simulation⁵ (Klypin et al. 2016). The MDPL2 assumes Planck cosmology ($h = 0.6777$, $\Omega_m = 0.307115$; Planck Collaboration et al. 2016) and evolves 3840³ dark matter particles in a $1 h^{-1}$ Gpc cubic volume beginning at $z = 120$. The particle

⁴ Here W1 is the 3.4 μm band of WISE (Wright et al. 2010).

⁵ www.cosmosim.org

resolution is $1.51 \times 10^9 h^{-1} M_\odot$. In total, 126 snapshots are available between $z \sim 15$ and $z = 0$. For this work, we use three snapshots at $z = 0.425$, 0.523 , and 0.628 .

The ROCKSTAR halo finder is designed to preserve particle-halo membership and identify accurate halo merger trees across multiple time steps of a simulation. For MDPL2, the ROCKSTAR halo catalogs are mass-complete for halos (including subhalos) above $M_{\text{vir}} \gtrsim 11.4 \times 10^{11} h^{-1} M_\odot$.

2.1.1. Subhalo Abundance Matching

In SHAM modeling, mock galaxies are assigned to dark matter halos by exploiting the correlation between some galaxy property—usually stellar mass or luminosity—and a halo property, such as virial mass or circular velocity. Circular velocity is defined as $v_{\text{circ}}(r, z) \equiv \sqrt{GM(<r, z)/r}$, where $M(<r, z)$ is the enclosed mass within a radius r at redshift z . SHAM has been tested with several versions of halo circular velocity. In this work, we use v_{peak} , the peak value of the maximum circular velocity ($v_{\text{max}}(z) = \max\{v_{\text{circ}}(r, z)\}$) achieved throughout a halo’s entire assembly history.

2.1.2. Age Distribution Matching

Age distribution matching assumes a correlation at fixed luminosity between galaxy color and some proxy (at fixed model luminosity) for the age of the halo in which each mock galaxy resides. Redder colors (i.e., older, quenched galaxies) are generally assigned to older halos. Model colors are assigned at fixed luminosity (in practice in narrow luminosity bins) because the galaxy color distribution is highly dependent on luminosity.

We equate the cumulative distribution $\mathcal{D}_{\text{gal}}$ of galaxies of k -corrected color $\mathcal{C}$ at fixed absolute magnitude M_X to the cumulative distribution $\mathcal{D}_{\text{halo}}$ of halo age proxy A at fixed model absolute magnitude:

$$\mathcal{D}_{\text{gal}}(<\mathcal{C} | M_X) = \mathcal{D}_{\text{halo}}(<A | \text{model } M_X). \quad (1)$$

In Equation (1), $(\mathcal{C}, M_X) = (r - z, M_z)$ or $(r - W1, M_{W1})$ for this work.

Implementations of age distribution matching at $z \sim 0$ using spectroscopic galaxy redshifts from SDSS have used the halo starvation redshift, z_{starve} , for the halo age proxy A in Equation (1) (Hearin & Watson 2013; Hearin et al. 2014; Safonova et al. 2021). In general, z_{starve} represents the redshift at which a galaxy loses its supply of cold gas, which leads to the quenching of star formation and the reddening of the galaxy. Multiple physical processes relevant to a halo’s assembly history can affect the value of z_{starve} for a given halo, which Hearin & Watson (2013) incorporated into the following definition:

$$z_{\text{starve}} \equiv \max\{z_{\text{char}}, z_{\text{acc}}, z_{\text{form}}\}. \quad (2)$$

Equation (2) is described as follows.

1. The z_{char} is either the redshift at which a halo’s mass first exceeds some characteristic value, M_{char} , or the redshift of the relevant simulation snapshot (z_{sim}) for halos that never achieve M_{char} .
2. The z_{acc} is the redshift at which a subhalo accretes onto a parent halo (for host halos $z_{\text{acc}} = z_{\text{sim}}$). Hearin & Watson (2013) followed Behroozi et al. (2012c), who defined z_{acc} as the snapshot after which a subhalo always remains a subhalo. They noted that alternative definitions, such as

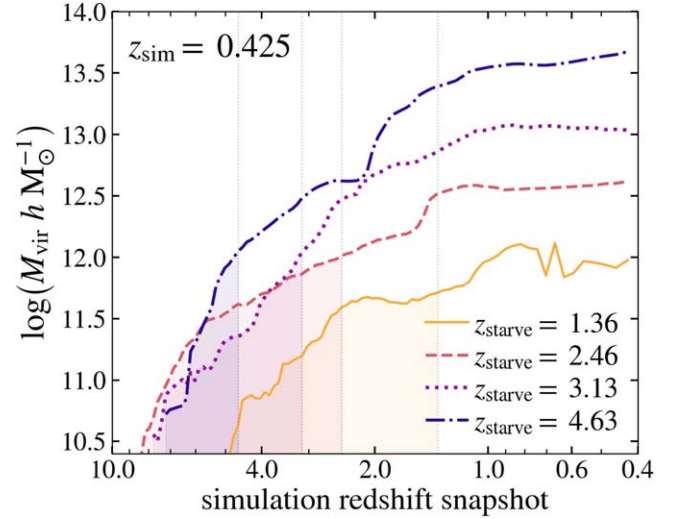


Figure 1. Mass accretion histories and z_{starve} values of four randomly selected halos from the $z_{\text{sim}} = 0.425$ snapshot of the MDPL2 simulation. The most massive halo (purple dashed-dotted line) has the earliest starvation redshift ($z_{\text{starve}} = 4.63$), corresponding to when its mass first reaches the characteristic value $M_{\text{char}} = 10^{12} h^{-1} M_\odot$. The other three halos also reach M_{char} (at later redshifts), but in accordance with Equation (2), the redshift at which this occurs for each halo is not necessarily the same as its z_{starve} value; e.g., the least massive halo (orange solid line) has $z_{\text{starve}} = 1.36$, although its mass does not exceed M_{char} until $z \lesssim 1$. The shaded region below each curve corresponds to a halo’s mass accretion history before it reaches z_{starve} .

that of Wetzel et al. (2014), where z_{acc} is the snapshot at which a subhalo has been identified as such for two consecutive snapshots, have little impact on their results.

3. The z_{form} is the “formation” redshift at which a halo transitions from the fast to slow accretion regime.

We use same definition of z_{form} as Hearin & Watson (2013), motivated by Wechsler et al. (2002),

$$z_{\text{form}} \equiv \frac{c_{\text{vir}}}{4.1a_0} - 1, \quad (3)$$

where $c_{\text{vir}} = R_{\text{vir}}/R_s$ is a halo’s concentration at the time of observation, indicated by a_0 . For host halos, a_0 is the scale factor of the relevant simulation snapshot, while for subhalos, a_0 is the scale factor at the time of accretion: $z_{\text{acc}} = 1/a_0 - 1$. Here R_{vir} is the virial radius of a halo, and R_s is the NFW scale radius (Navarro et al. 1997).

We adopt the value of $M_{\text{char}} = 10^{12} h^{-1} M_\odot$ used in Hearin & Watson (2013), who noted that their results are insensitive to the precise value of M_{char} used. The empirical and physical motivations for M_{char} are described in detail in Section 6.3 of Hearin & Watson (2013). Briefly, there is empirical support for a characteristic halo mass above which star formation is highly inefficient: $\sim 10^{12} h^{-1} M_\odot$ is the halo mass at which the SHMR peaks, falling off rapidly at higher halo masses (Yang et al. 2012, 2013; Behroozi et al. 2013; Moster et al. 2013; Watson & Conroy 2013), and Behroozi et al. (2012a) showed that this mass remains essentially constant throughout much of cosmic history.

We compute z_{starve} for all halos in our model from the publicly available ROCKSTAR halo merger trees for MDPL2. Figure 1 shows sample halo mass accretion histories and corresponding z_{starve} values for four randomly selected halos from the $z_{\text{sim}} = 0.425$ snapshot of MDPL2.

2.2. Photometry and Redshift Estimates

We use publicly available catalogs from the ninth data release (DR9) of the DESI Legacy Imaging Surveys⁶ (Dey et al. 2019). The Legacy Surveys provide optical imaging in the g , r , and z bands from a combination of three public surveys: the DECam Legacy Survey (DECaLS; Flaugher et al. 2015; Blum et al. 2016), the Beijing–Arizona Sky Survey (BASS; Zou et al. 2017), and the Mayall z -band Legacy Survey (MzLS; Silva et al. 2016). The Legacy Surveys also include four mid-infrared bands from the Wide-field Infrared Survey Explorer (WISE; Wright et al. 2010), although only the 3.4 μm W1 band is relevant for DESI LRG target selection.

In total, the Legacy Surveys cover 14,000 deg^2 visible from the northern hemisphere, comprised of two contiguous regions within the northern and southern galactic caps. To avoid effects from systematic differences among the data from the three component optical surveys, we limit our study to the approximately 9000 deg^2 covered by DECaLS.

Zhou et al. (2021) computed photometric redshifts for the full catalog of DECaLS DR7 objects using the random forest regression machine-learning algorithm in `Scikit-Learn` (Pedregosa et al. 2011) and a “truth” data set of spectroscopic and many-band photometric redshifts for objects within DR7. They quantified the accuracy of their photometric redshifts with the normalized median absolute deviation (NMAD, $\sigma_{\text{NMAD}} = 1.48 \times \text{median}(|\Delta z|/(1 + z_{\text{spec}})$; Dahlen et al. 2013), where $\Delta z = z_{\text{phot}} - z_{\text{spec}}$ and z_{spec} are the redshift truth values used to train the random forest algorithm, and reported $\sigma_{\text{NMAD}} = 0.021$ for LRGs. Their outlier rate for LRGs is 1%, where outliers are objects with $|\Delta z| > 0.1 \times (1 + z_{\text{spec}})$. Additionally, Zhou et al. (2021) estimated that their redshifts are accurate for objects with an apparent z -band magnitude of $z < 21$, well beyond the $z < 20.7$ cut we use to select our target galaxy samples, described in Section 2.4 below.

2.3. DESI LRG Target Selection

The DESI LRG target sample is intended to serve as a cosmological tracer spanning a redshift range of $\sim 0.4 < z \lesssim 1.0$. The sample lies between the low-redshift BGS (Hahn et al. 2022) tracer sample at $z \lesssim 0.4$ and the emission line galaxy (Raichoor et al. 2022) sample, optimized to trace the density field over the approximate range $1.0 < z < 1.6$.

Two different selection algorithms were considered for the DESI LRG sample: an optical selection function based on z -band magnitude and $r - z$ color and an infrared selection function based on W1-band magnitude and $r - W1$ color. Both selections are tuned to yield a constant LRG target density of ~ 600 objects deg^{-2} and a comoving number density around $5 \times 10^{-4} h^3 \text{Mpc}^{-3}$ at $0.4 < z < 0.8$. Both the optical and IR selections were tested in DESI’s Survey Validation (DESI et al. 2023, in preparation) observations. Based largely on the calibration of W1-band imaging, the DESI Main Survey exclusively uses the IR selection. A complete description of DESI LRG target selection is given in Zhou et al. (2023). Here we cover the details most relevant for this work.

Due to slight differences in photometry among BASS, MzLS, and DECaLS, the optical and IR DESI LRG target selections use slightly different cuts for the north (BASS and

MzLS) and south (DECaLS) galactic caps. As this work uses g , r , and z magnitudes from DECaLS, we use the corresponding optical LRG target selection cuts (Zhou et al. 2020):

$$z - W1 > 0.8 \times (r - z) - 0.6, \quad (4a)$$

$$((g - W1 > 2.6) \text{ AND } (g - r > 1.4)) \text{ OR } (r - W1 > 1.8), \quad (4b)$$

$$(r - z > 0.45 \times (z - 16.83)) \text{ AND } (r - z > 0.7) \quad (4c)$$

$$\text{AND } (r - z > 0.19 \times (z - 13.80)), \quad z_{\text{fiber}} < 21.5. \quad (4d)$$

The relevant IR LRG target selection cuts for this work are (Zhou et al. 2023)

$$z - W1 > 0.8 \times (r - z) - 0.6, \quad (5a)$$

$$(g - W1 > 2.9) \text{ OR } (r - W1 > 1.8), \quad (5b)$$

$$((r - W1 > 1.8 \times (W1 - 17.14)) \text{ AND } (r - W1 > W1 - 16.33)) \text{ OR } (r - W1 > 3.3), \quad (5c)$$

$$z_{\text{fiber}} < 21.6. \quad (5d)$$

Equations 4(a) and 5(a) are designed to reject stars, Equations 4(b) and 5(b) remove blue and low-redshift objects, Equations 4(c) and 5(c) are color-dependent magnitude limits that select only the most luminous objects at a given redshift, and z_{fiber} in Equations 4(d) and 5(d) is the expected z -band flux within a DESI fiber. All magnitudes in Equations (4a) and (5a) use the AB system and are corrected for galactic extinction using the relevant `MW_TRANSMISSION` values from the Legacy Surveys DR9.

2.4. Parent Galaxy Samples

A primary goal of this work is to create mock galaxy catalogs that are both statistically complete and represent a superset of the color–magnitude space occupied by DESI LRG targets. Selection of DESI LRG targets is based entirely on g , r , z , and W1 apparent magnitudes (see Section 2.3), so we would ideally create mock catalogs where every mock galaxy has an apparent magnitude in each of these bands. SHAM, however, exploits the correlation between some physical halo property (e.g., circular velocity) and a physical galaxy property independent of redshift (e.g., luminosity). We therefore train our mock catalogs on galaxy samples that are complete to an absolute magnitude threshold that includes all DESI LRG targets (in the relevant redshift bin; see below).

To select suitable parent galaxy samples, we first apply an apparent z -band magnitude cut of $z < 20.7$ and take an additional step to remove stars by excluding catalog sources with `TYPE = PSF`. We also apply the masks described in Table 1, which are provided with DECaLS DR9, to remove sources affected by bad pixels or contamination from bright stars. Finally, a geometric mask is applied to ensure complete angular coverage by the catalogs of random points provided with DR9 (see Section 3.1). The resulting sample contains $\mathcal{O}(10^8)$ galaxies, sufficient to divide it into redshift bins and maintain a low statistical error.

We initially tested six redshift bins of width $\Delta z_{\text{phot}} = 0.1$ between $z_{\text{phot}} = 0.4$ and 1.0 but found that DECaLS photometry is only deep enough to apply our model up to $z_{\text{phot}} \sim 0.7$. At $z_{\text{phot}} \gtrsim 0.7$, the data are incomplete above the absolute magnitude threshold that encompasses DESI LRG targets. We therefore limit our study to three redshift bins of $\Delta z_{\text{phot}} = 0.1$ within $0.4 < z_{\text{phot}} < 0.7$.

⁶ www.legacysurvey.org/dr9

For each redshift bin, we compute z - and $W1$ -band absolute magnitudes using photometric redshifts to obtain distance moduli. We k -correct the absolute magnitudes to the redshift of the relevant simulation snapshot, z_{sim} (see Table 2), with the IDL package `kcorrect` (Blanton & Roweis 2007) using DECam g , r , and z and WISE $W1$ and $W2$ filter responses. For each redshift bin, we use the simulation snapshot closest to the median z_{phot} of the data; e.g., galaxies in the $0.4 < z_{\text{phot}} < 0.5$ bin are k -corrected to $z_{\text{sim}} = 0.425$.

The final step in selecting parent galaxy samples is to identify z - and $W1$ -band absolute magnitude cuts in each redshift bin that yield complete samples that also include the full absolute magnitude range of DESI LRG targets in that bin. Figure 2 shows k -corrected absolute magnitude distributions for all $z < 20.7$ DECaLS galaxies in each redshift bin. For each bandpass and redshift bin combination, we identify an absolute magnitude cut (dashed black lines in Figure 2) that eliminates fainter galaxies where DECaLS becomes incomplete while preserving $\geq 99\%$ of the DESI LRG targets (solid orange and hatched purple histograms).

Table 2 lists the details of each parent galaxy sample, including the relevant absolute magnitude cut, sample size, effective number density, and included fractions of IR and optical DESI LRG targets. Besides enforcing statistically complete parent samples, these magnitude cuts also eliminate galaxies with larger photometric redshift errors, increasing the accuracy of clustering measurements (see Section 3.1).

Figure 3 shows example optical ($r - z$ vs. M_z) and IR ($r - W1$ vs. M_{W1}) color-magnitude diagrams of the magnitude-limited parent galaxy samples for the $0.4 < z_{\text{phot}} < 0.5$ redshift bin. Also shown in Figure 3 are the color-magnitude distributions of the optical and IR DESI LRG targets.

3. Modeling

In this section, we describe our modeling procedure and the two-point statistics we use to constrain the model parameters.

3.1. Projected Correlation Functions

One goal of this work is to exploit the completeness and enormous volume of DECaLS data, which comes at the expense of the precision clustering measurements achievable with spectroscopic redshifts. Zhou et al. (2021) demonstrated the constraining power of the projected correlation function, $\omega_p(r_p)$, of DECaLS galaxies computed with line-of-sight distances derived from photometric redshifts (they used this statistic to compute HOD parameters of DESI-like LRGs selected from DECaLS DR7). The projected correlation function conveys a three-dimensional correlation function, $\xi(r)$, integrated along the line of sight, effectively eliminating the effects of radial distance uncertainty due to photometric redshift errors:

$$\omega_p(r_p) \equiv \int_{-\pi_{\text{max}}}^{\pi_{\text{max}}} \xi(r_p, \pi) d\pi = 2 \int_0^{\pi_{\text{max}}} \xi(r_p, \pi) d\pi, \quad (6)$$

where r_p is the projection of r into the plane perpendicular to the line-of-sight separation, π .

We use the `Corrfunc` package (Sinha & Garrison 2017, 2019) to calculate $\omega_p(r_p)$ for both our target galaxy samples and mock catalogs in 19 logarithmic bins between $r_p > 0.1$ and $\lesssim 44 h^{-1}$ Mpc. We also compute $\omega_p(r_p)$ in

Table 1
Bitmasks Applied to Photometry for Parent Galaxy Sample Selection

MASKBIT	Description
5, 6, 7	Bad pixel in all of a set of overlapping g -, r -, or z -band images
8, 9	Bad pixel in a WISE $W1$ or $W2$ bright star mask
11	Pixel within locus of a radius-magnitude relation for Gaia ^a DR2 stars to $G < 16$
12	Pixel in a Siena Galaxy Atlas ^b large galaxy
13	Pixel in a globular cluster
FITBIT	Description
6	Source is a medium-bright star
7	Gaia source ^c
8	Tycho-2 star ^d

Notes. Additional details at legacysurvey.org/dr9/bitmasks.

^a Gaia Collaboration (2018).

^b Moustakas et al. (2021).

^c Gaia Collaboration et al. (2016).

^d Høg et al. (2000).

additional bins at $r_p < 0.1 h^{-1}$ Mpc but do not use these measurements for model fitting.

As our data samples are confined to narrow redshift bins of width $\Delta z_{\text{phot}} = 0.1$, photometric redshift errors will cause some galaxies that belong to a given redshift bin to scatter into an adjacent bin and be excluded from the calculation of $\omega_p(r_p)$ for their true bin. To account for this, we adopt the method used by Zhou et al. (2021; see their Figure 8); we use the Landy-Szalay estimator (Landy & Szalay 1993) for the cross-correlation of two samples, D_1 and D_2 :

$$\omega_p(r_p) = \sum_{-\pi_{\text{max}}}^{\pi_{\text{max}}} \left(\frac{D_1 D_2 - D_1 R_2 - D_2 R_1}{R_1 R_2} + 1 \right). \quad (7)$$

Each term of Equation (7) denotes pair counts between two samples, where D and R respectively indicate samples of data (i.e., galaxies) and random points with the same angular and redshift distributions as the corresponding data sample. Here D_1 is all galaxies within a given redshift bin, $z_{\text{min}} < z_{\text{phot}} < z_{\text{max}}$, where z_{min} and z_{max} are the limits of the bin, while D_2 is all galaxies within a wider redshift range defined by $(z_{\text{min}} - \pi_{\text{max}}) < z_{\text{phot}} < (z_{\text{max}} + \pi_{\text{max}})$, where $\pi_{\text{max}} = 150 h^{-1}$ Mpc. We verify our implementation of this method with `Corrfunc` by reproducing the projected correlation functions of DECaLS LRGs from Zhou et al. (2021; see their Figure 9) using a different clustering code.

The DECaLS data include catalogs of random points with the same angular sky coverage and mask information as the survey footprint, which we use to construct our random samples. We use 20 times as many random as data points for each galaxy sample and draw redshifts for random points from the redshift distribution of the corresponding data sample.

To measure the $\omega_p(r_p)$ of our mock catalogs, we take advantage of the `Corrfunc` theory module, which can quickly calculate the autocorrelation function of a sample within a periodic volume using analytic random pair counts. We confirmed that this method produces the expected result by calculating the $\omega_p(r_p)$ of several mock catalogs directly from pair counts between mock galaxies and catalogs of random points constructed for the simulation volume.

Table 2
Properties of Parent Galaxy Samples and Corresponding Simulation Snapshot Redshifts (z_{sim})

Redshift Bin	z_{sim}	Luminosity Cut ^a	N_{gal}	n_{eff} [$\times 10^{-3} h^3 \text{Mpc}^{-3}$]	Included Fraction of DESI LRG Targets	
					Optical Selection	IR Selection
$0.4 < z_{\text{phot}} < 0.5$	0.425	$^{0.43}M_z < -21.60$	8,314,309	5.60	0.992	0.982
		$^{0.43}M_{W1} < -22.25$	7,565,153	4.33	0.992	0.992
$0.5 < z_{\text{phot}} < 0.6$	0.523	$^{0.52}M_z < -21.60$	7,804,346	2.86	0.992	0.988
		$^{0.52}M_{W1} < -22.85$	4,909,857	2.01	0.992	0.992
$0.6 < z_{\text{phot}} < 0.7$	0.628	$^{0.63}M_z < -21.85$	6,548,126	1.46	0.994	0.990
		$^{0.63}M_{W1} < -23.15$	4,758,470	1.12	0.993	0.994

Note.

^a The k -correction redshifts are rounded to two decimal places for clarity; e.g., $^{0.43}M_z$ indicates that the absolute z -band magnitudes are k -corrected to $z_{\text{sim}} = 0.425$.

3.2. Jackknife Error Estimation and Goodness of Fit

To estimate the uncertainty of the $\omega_p(r_p)$ measurements of our target galaxy samples, we use `healpy`⁷ (Górski et al. 2005; Zonca et al. 2019) with $N_{\text{side}} = 6$ to divide the angular sky coverage of each galaxy sample into N_{jk} regions of roughly equal area, suitable for jackknife resampling. We then measure $\omega_p(r_p)$ in each jackknife sample, where each jackknife sample consists of the entire galaxy sample with one jackknife region removed, and compute the covariance matrix as follows:

$$\text{Cov}_{ij} = \frac{N_{\text{jk}} - 1}{N_{\text{jk}}} \sum_{\ell=1}^{N_{\text{jk}}} (\omega_i^\ell - \bar{\omega}_i)(\omega_j^\ell - \bar{\omega}_j), \quad (8)$$

where ω_i^ℓ and ω_j^ℓ are the $\omega_p(r_p)$ of the ℓ th jackknife region for the i th and j th r_p bins, respectively, and $\bar{\omega}_i$ and $\bar{\omega}_j$ are the mean values of $\omega_p(r_p)$ across all jackknife regions for the i th and j th r_p bins, respectively.

With the covariance matrix in hand, we quantify how successful any instance of our model is at fitting the projected correlation function of the data by computing χ^2 per degree of freedom (χ_ν^2),

$$\chi_\nu^2 = \frac{1}{\nu} \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} (\omega_i - \omega_i^{\text{mod}})(\text{Cov}^{-1})_{ij}(\omega_j - \omega_j^{\text{mod}}), \quad (9)$$

where N_p is the number of r_p bins used for fitting, ν is equal to N_p minus the number of free model parameters, ω_i and ω_j are the data $\omega_p(r_p)$ values in the i th and j th r_p bins, respectively, and ω_i^{mod} and ω_j^{mod} are the $\omega_p(r_p)$ values of the relevant mock catalog in the i th and j th r_p bins, respectively.

3.3. Luminosity Assignment

We use SHAM to create mock galaxy populations with the same number density and luminosity distribution as the parent galaxy sample by assuming the relation

$$n_{\text{eff}}(<M) = n_h(>v_{\text{peak}}); \quad (10)$$

i.e., the (effective) number density of galaxies, n_{eff} , of magnitude M or brighter equals the number density of halos, n_h , with circular velocity v_{peak} or greater.

To assign absolute magnitudes to halos in each redshift bin, we implement the following procedure.

1. Compute for each parent galaxy sample the effective galaxy number density as a function of absolute magnitude in band $X \in \{z, W1\}$, $n_{\text{eff}}(<M_X)$. We compute $n_{\text{eff}}(<M_X)$ for each parent galaxy sample as follows. For each galaxy in the sample, we calculate an effective volume, V_{eff} ,

$$V_{\text{eff}}(M_X) = f_\Omega (V_{\text{max}}(M_X) - V_{\text{min}}), \quad (11)$$

where f_Ω is the fractional solid angle covered by the parent galaxy sample, V_{min} is the comoving volume of the lower limit of the redshift bin, and $V_{\text{max}}(M_X)$ is the comoving volume of either the upper limit of the redshift bin or the maximum possible redshift a galaxy of magnitude M_X could have and still be observed at the magnitude limit of the sample, whichever is smaller. The effective galaxy number density is the sum of the inverse of $V_{\text{eff}}(M_X)$ over all galaxies in the sample:

$$n_{\text{eff}}(<M_X) = \sum_i [V_{\text{eff}}^i(M_X)]^{-1}. \quad (12)$$

2. Assign absolute magnitudes to mock galaxies with no scatter in the luminosity- v_{peak} relation according to Equation (10). For each halo, we find the cumulative number density $n_h(>v_{\text{peak}})$ corresponding to its value of v_{peak} . We then assign to each halo a mock galaxy with the absolute magnitude M_X at which the effective number density $n_{\text{eff}}(<M_X)$ of the parent galaxy sample equals $n_h(>v_{\text{peak}})$.
3. To incorporate magnitude-dependent scatter into the luminosity- v_{peak} relation, we assign to each mock galaxy a new absolute magnitude, M'_X , where M'_X is drawn from a Gaussian distribution centered at M_X with width σ_{mag} . The σ_{mag} is proportional to the absolute value of M_X , and the constant of proportionality is a free parameter in the model. We then rank order all mock galaxies (including their v_{peak} values) by M'_X , rank order the original distribution of mock magnitudes M_X , and assign the ordered original mock magnitude distribution to the mock galaxy catalog ordered by M'_X . This method incorporates scatter into the luminosity- v_{peak} relation while exactly replicating the target luminosity function in the resulting mock galaxy catalog.
4. Scatter the positions of the mock galaxies along one of the three axes of the simulation volume to mimic the uncertainty in the radial (line-of-sight) position of our target galaxy samples due to photometric redshift errors. For each mock galaxy, we draw a “scattered” coordinate

⁷ healpix.sourceforge.net

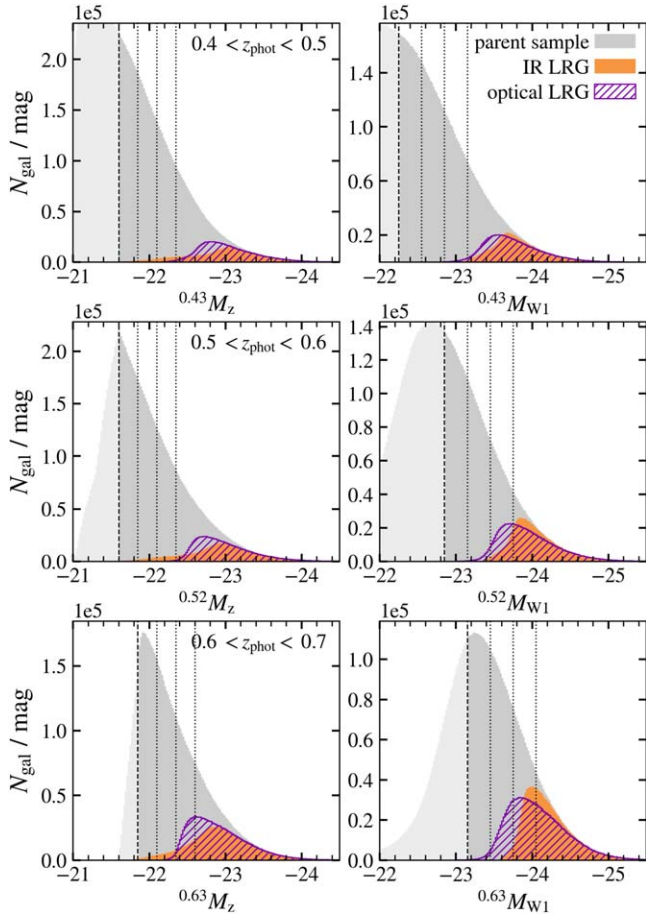


Figure 2. Distribution of k -corrected z -band (left column) and $W1$ -band (right column) absolute magnitudes for the three redshift bins in our study: $0.4 < z_{\text{phot}} < 0.5$ (top), $0.5 < z_{\text{phot}} < 0.6$ (middle), and $0.6 < z_{\text{phot}} < 0.7$ (bottom). Each panel shows the distribution of all galaxies with $z < 20.7$ (solid gray), as well as the distributions of optical (hatched purple; see Equation (4a)) and IR (solid orange; see Equation (5a)) DESI LRG targets. The leftmost dashed black line in each panel is the absolute magnitude cut that defines each absolute magnitude–limited parent galaxy sample for our models. At fainter magnitudes (to the left of the cut in each panel), the DECaLS sample is incomplete. The three dotted black lines in each panel denote the luminosity bins used to constrain the model parameters (see Section 3.2 and Tables 2 and 3).

x' from a Gaussian distribution of width σ_{LOS} centered at the galaxy’s original x position. Mock galaxies that scatter out of the simulation volume of $1 h^{-3} \text{Gpc}^3$ are wrapped back in to preserve the periodic boundary conditions; e.g., a mock galaxy at $x = 25 h^{-1} \text{Mpc}$ that scatters to $x' = -50 h^{-1} \text{Mpc}$ is placed at $x = 950 h^{-1} \text{Mpc}$. We repeat this process for the other two simulation axes (i.e., y and z).

5. For each mock catalog, compute the projected correlation function, $\omega_p(r_p)$ (see Section 3.1), averaged over the three axes of the simulation volume as described in the previous step. We then compute the goodness of fit per degree of freedom, χ_ν^2 (see Section 3.2), of the model fit to the mean $\omega_p(r_p)$ of the relevant parent galaxy sample. We measure χ_ν^2 for each mean $\omega_p(r_p)$ in bins of absolute magnitude (M_z or M_{W1}).
6. Repeat steps 3–5 for additional values of σ_{mag} and σ_{LOS} as needed to minimize χ_ν^2 in each luminosity bin. In practice, we first coarsely sample a wide range of values

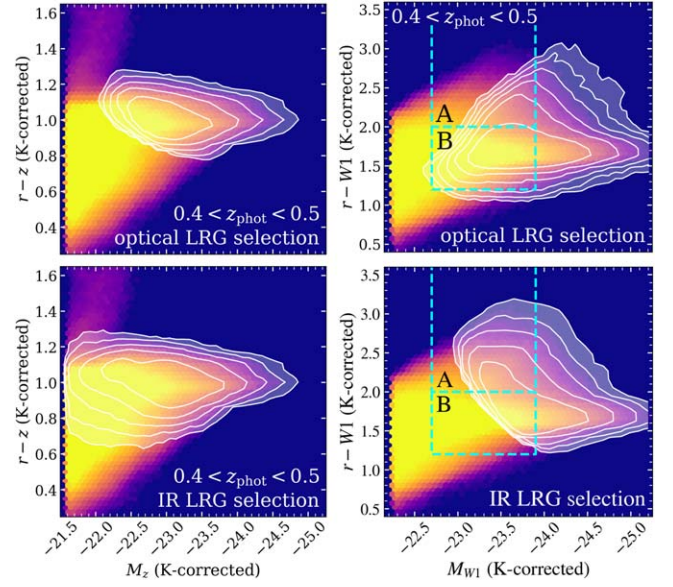


Figure 3. Color–magnitude diagrams of the parent galaxy samples and DESI LRG target samples (overlaid white contours) in the $0.4 < z_{\text{phot}} < 0.5$ redshift bin. The left column shows optical (k -corrected $r - z$ color vs. M_z magnitude), while the right column shows IR (k -corrected $r - W1$ color vs. M_{W1} magnitude). The top (bottom) row shows the distribution of optical (IR) LRGs in each color–magnitude space. The boxed regions labeled “A” and “B” are referenced in Section 4.1 below.

of σ_{mag} and σ_{LOS} :

$$0 \leq \frac{\sigma_{\text{mag}}}{|M_X|} \leq 1.0, \Delta \left[\frac{\sigma_{\text{mag}}}{|M_X|} \right] = 0.1, \quad (13a)$$

$$0 \leq \sigma_{\text{LOS}} \leq 150 h^{-1} \text{Mpc}, \Delta \sigma_{\text{LOS}} = 10 h^{-1} \text{Mpc}. \quad (13b)$$

We then more densely sample (using $\Delta(\sigma_{\text{mag}}/|M_X|) = 0.01$ and $\Delta\sigma_{\text{LOS}} = 5 h^{-1} \text{Mpc}$) narrower ranges of both parameters around the initial coarse-grained values that minimize χ_ν^2 .

7. Finally, identify the value of σ_{mag} that coincides with the region of $\sigma_{\text{mag}} - \sigma_{\text{LOS}}$ parameter space containing the minimum χ_ν^2 across all luminosity bins (see Table 3) and parameterize the dependence of σ_{LOS} on absolute magnitude, M_X , as follows:

$$\sigma_{\text{LOS}}(M_X) = s_{\text{LOS}} M_X + \sigma_{\text{LOS},0}. \quad (14)$$

The best-fit values of σ_{mag} , s_{LOS} , and $\sigma_{\text{LOS},0}$ are given in Table 4. The values of σ_{LOS} and $\sigma_{\text{LOS},0}$ are determined by linear fits to σ_{LOS}^i versus $\langle M_X^i \rangle$, where σ_{LOS}^i are the values that minimize the χ_ν^2 of $\omega_p(r_p)$ at fixed σ_{mag}^8 in the i th luminosity bin, M_X^i .

The luminosity assignment stage of our modeling procedure involves three free parameters, σ_{mag} , s_{LOS} , and $\sigma_{\text{LOS},0}$, which account for scatter in the luminosity– v_{circ} relation and the photometric redshift errors of our target galaxy samples. We constrain these parameters by fitting the projected correlation functions of mock galaxy catalogs created from our model to those of the corresponding parent galaxy samples.

Figures 4 and 5 show $\omega_p(r_p)$ for the $W1$ - and z -band absolute magnitude–limited parent galaxy samples, respectively, and

⁸ We tested using σ_{mag} with linear dependence on absolute magnitude instead of a constant value across all luminosity bins and found that fixing σ_{mag} while allowing σ_{LOS} to scale linearly with luminosity reproduces $\omega_p(r_p)$ (Section 3.1) better than if both parameters have linear luminosity dependence.

Table 3
Luminosity Bins Used to Constrain the Magnitude Dependence of Model Parameters

0.4 < $z_{\text{phot}} < 0.5$		0.5 < $z_{\text{phot}} < 0.6$		0.6 < $z_{\text{phot}} < 0.7$	
Luminosity Bin	N_{gal}	Luminosity Bin	N_{gal}	Luminosity Bin	N_{gal}
z -band					
$-21.60 > {}^{0.43}M_z > -21.85$	2,458,920	$-21.60 > {}^{0.52}M_z > -21.85$	2,304,853	$-21.85 > {}^{0.63}M_z > -22.10$	2,040,336
$-21.85 > {}^{0.43}M_z > -22.10$	1,880,295	$-21.85 > {}^{0.52}M_z > -22.10$	1,732,820	$-22.10 > {}^{0.63}M_z > -22.35$	1,512,940
$-22.10 > {}^{0.43}M_z > -22.35$	1,317,269	$-22.10 > {}^{0.52}M_z > -22.35$	1,205,545	$-22.35 > {}^{0.63}M_z > -22.60$	1,039,573
${}^{0.43}M_z < -22.35$	1,919,594	${}^{0.52}M_z < -22.35$	1,855,014	${}^{0.63}M_z < -22.60$	1,401,662
$W1$ -band					
$-22.25 > {}^{0.43}M_{W1} > -22.55$	2,324,816	$-22.85 > {}^{0.52}M_{W1} > -23.15$	1,771,554	$-23.15 > {}^{0.63}M_{W1} > -23.45$	1,634,860
$-22.55 > {}^{0.43}M_{W1} > -22.85$	1,837,859	$-23.15 > {}^{0.52}M_{W1} > -23.45$	1,245,107	$-23.45 > {}^{0.63}M_{W1} > -23.75$	1,273,814
$-22.85 > {}^{0.43}M_{W1} > -23.15$	1,277,192	$-23.45 > {}^{0.52}M_{W1} > -23.75$	754,154	$-23.75 > {}^{0.63}M_{W1} > -24.05$	798,012
${}^{0.43}M_{W1} < -23.15$	1,565,914	${}^{0.52}M_{W1} < -23.75$	693,491	${}^{0.63}M_{W1} < -24.05$	685,799

Table 4
Best-fit Values for SHAM Model Parameters

Redshift Bin	Model	σ_{mag}	s_{LOS}	$\sigma_{\text{LOS},0}$
0.4 < $z_{\text{phot}} < 0.5$	W1-band	0.66	−22.3	−439.2
	z -band	0.57	−11.4	−181.5
0.5 < $z_{\text{phot}} < 0.6$	W1-band	0.76	21.3	584.0
	z -band	0.76	60.9	1435.6
0.6 < $z_{\text{phot}} < 0.7$	W1-band	0.80	31.2	839.4
	z -band	0.77	50.8	1248.4

corresponding mock galaxy catalogs. The agreement between the data and model increases with increasing luminosity within each redshift bin and increasing redshift overall. Note that the shaded regions in Figures 4 and 5 at $r_p < 0.1 h^{-1} \text{Mpc}$ denote measurements not used for model fitting.

The clustering of the full magnitude-limited parent galaxy samples and corresponding mock catalogs is shown for each redshift bin in Figure 6, with $\omega_p(r_p)$ for each redshift bin offset by 0.15 dex for clarity. Agreement between the model and data increases with increasing redshift. We emphasize that while the model $\omega_p(r_p)$ deviates from that of the data for the full magnitude-limited parent samples in the $0.4 < z_{\text{phot}} < 0.5$ redshift bin, there is still good data–model agreement within the highest luminosity bins, where the vast majority of LRGs reside, for both the W1- and z -band models (Figures 4 and 5, respectively).

3.4. Color Assignment

Figure 7 shows an illustration of our color assignment algorithm (Equation (1)) with age distribution matching for the $0.4 < z_{\text{phot}} < 0.5$ redshift bin of the z -band model. The dotted purple curve is the cumulative distribution of ${}^{0.43}(r - z)$ color for galaxies in the $-23.0 > {}^{0.43}M_z > -23.05$ luminosity bin, and the solid gray curve is the halo z_{starve} distribution of $z_{\text{sim}} = 0.425$ mock galaxies in the same model luminosity bin. The magenta arrows in Figure 7 indicate that a mock galaxy in this luminosity bin in a halo with $z_{\text{starve}} \approx 2.6$ is assigned a ${}^{0.43}(r - z)$ color of 0.96.

Implementation of the color assignment algorithm with no scatter in the color– z_{starve} relationship yields mock LRG samples (see Section 3.5 below) that generally overpredict the clustering amplitude compared to the data, with the exception of optical LRGs selected from the W1-band model

(see Figure 10). We tested the effect of introducing scatter into the color– z_{starve} relation on the clustering of mock LRGs using a method analogous to how scatter is incorporated into the luminosity– v_{peak} relation (see Section 3.3). After assigning model colors $\mathcal{C}$ without scatter according to Equation (1) as described above, we assign to each mock galaxy a new color $\mathcal{C}'$, where $\mathcal{C}'$ is drawn from a Gaussian distribution centered at $\mathcal{C}$ with width σ_{color} . We then rank order all mock galaxies by $\mathcal{C}'$, rank order the original distribution of mock colors $\mathcal{C}$, and assign the ordered original color distribution to the mock galaxy catalog ordered by $\mathcal{C}'$.

For mock LRGs (both optical and IR) from the z -band model, as well as mock IR LRGs from the W1-band model, the clustering amplitude decreases with increasing σ_{color} until σ_{color} is sufficiently large that model colors are effectively assigned at random, at which point the decrease in clustering amplitude levels off at a constant value that still overpredicts $\omega_p(r_p)$ relative to the data. The exception to this is the predicted clustering of mock optical LRGs from the W1-band model, which largely agrees with the data with no scatter in the color– z_{starve} relation. Additionally, $\sigma_{\text{color}} > 0$ in this instance does not affect the predicted clustering amplitude of mock optical LRGs. This implies that color is uncorrelated with halo age for LRGs, which we discuss further in Section 4.1, and motivates our decision to proceed with two versions of our models (rather than introduce and constrain σ_{color} as an additional parameter): (1) a “default” age distribution model, in which galaxy color increases monotonically with increasing halo z_{starve} at fixed luminosity (i.e., $\sigma_{\text{color}} = 0$), and (2) a “random color” model, in which there is no correlation between galaxy color and halo z_{starve} . Both models yield magnitude-limited mock galaxy catalogs with identical color–magnitude distributions that reproduce the target distribution of the relevant data; the only difference is whether the galaxy color is maximally correlated (default model) or entirely uncorrelated (random color model) with halo z_{starve} .

3.5. Selecting Mock LRGs

DESI LRG target selection (see Section 2.3) is entirely a function of apparent g , r , z , and W1 (and z_{fiber}) magnitudes. As such, it would be ideal to have mock galaxy catalogs with each of these model magnitudes for every galaxy; mock LRG samples could then be obtained simply by applying the DESI LRG target selection functions (Equations (4a) and (5a)) to the full mock for each redshift bin. However, the SHAM and age distribution matching techniques we employ create mock

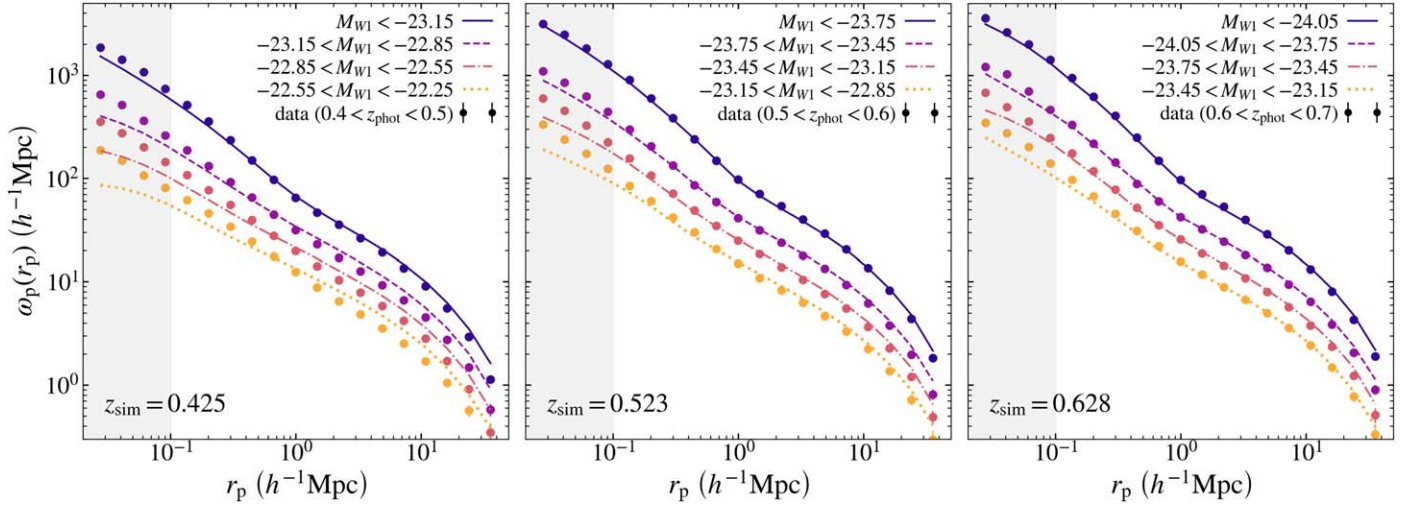


Figure 4. Luminosity-binned projected clustering of the W1-band absolute magnitude-limited parent galaxy samples and corresponding mock galaxy catalogs, with different colors and line styles representing different luminosity bins as denoted by the legend in each panel. Luminosity bins are offset by 0.15 dex for clarity. Each panel shows a different redshift bin. The shaded region at $r_p < 0.1 h^{-1} \text{Mpc}$ denotes measurements not used for modeling.

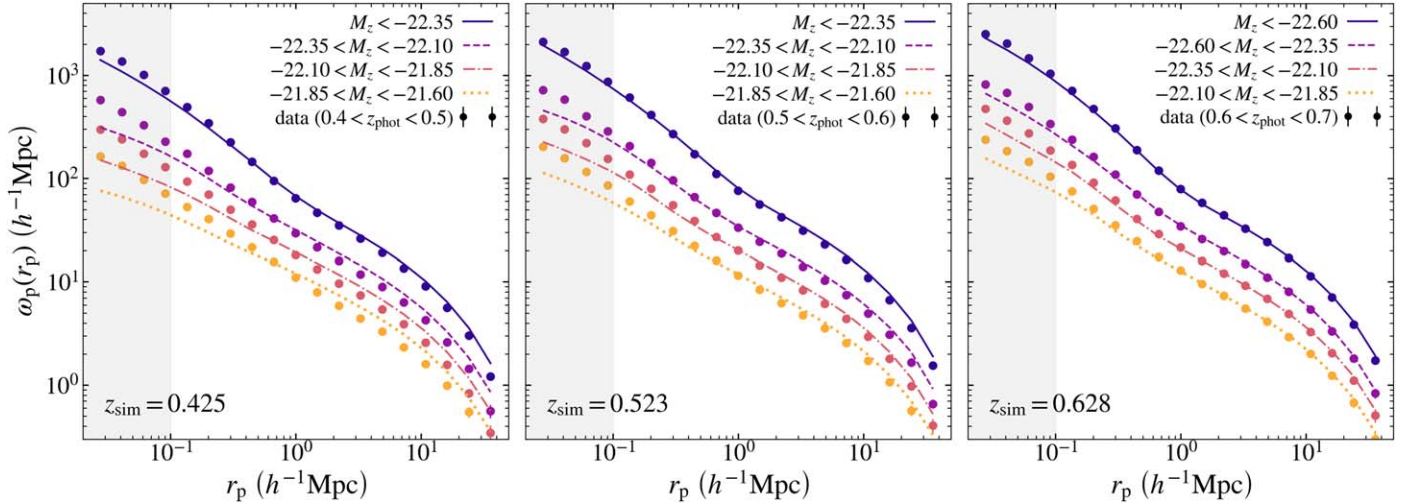


Figure 5. Same as Figure 4 but for the z -band parent galaxy samples and corresponding mock galaxy catalogs.

galaxy catalogs with absolute z - or W1-band magnitudes and corresponding $r - z$ or $r - W1$ colors.

To select mock DESI LRG targets using only these model quantities, we first identify where the actual DESI LRG targets reside in k -corrected color-magnitude space in each redshift bin for both optical and IR DESI LRG targets. In each redshift bin, we compute the fractions of galaxies that are optical and IR DESI LRG targets in narrow two-dimensional bins in color-magnitude space. This is shown in Figure 8.

We then use these optical and IR LRG fractions in color-magnitude space to statistically select mock optical and IR LRGs from our mock galaxy catalogs in the corresponding model color-magnitude space; i.e., in each color-magnitude bin, we draw mock galaxies and flag them as LRGs until the LRG fraction in that bin matches the value from the data.

It is worth noting that the LRG fraction in most of color-magnitude space is either $\sim$ zero or $\sim$ 1 by design, especially for optical LRGs in $r - z$ versus M_z (i.e., optical) space and IR LRGs in $r - W1$ versus M_{W1} (i.e., IR) space. This is clearly shown in Figure 8.

The first and second columns of Figure 8 show the fractions of optical and IR LRG targets, respectively, in optical space, i.e., $r - z$ color versus M_z magnitude, and the red sequence is clearly visible in each panel. For example, in the $0.4 < z_{\text{phot}} < 0.5$ redshift bin, the red sequence corresponds almost entirely to the region where the LRG fraction is ~ 1 , i.e., where $M_z \lesssim -22.5$ and $r - z$ is between ~ 0.9 and ~ 1.3 (with a shift toward redder colors in higher redshift bins).

Interestingly, the LRG fraction begins to deviate from ~ 1 in the most luminous region of optical space ($M_z \lesssim -23.7$), and this is true for both optical and IR LRG targets. Both selections (Equations (4a) and (5a)) exclude some of the most luminous red-sequence objects in optical space but not in IR space (third and fourth columns of Figure 8).

To investigate why these luminous red-sequence objects in the first two columns of Figure 8 are not selected as DESI LRG targets, we looked at their positions in color-color and color-magnitude space relative to the full set of DESI LRG target selection cuts (Equations (4a) and (5a)). This is shown in Figure 9, where light blue and gray contours in each panel

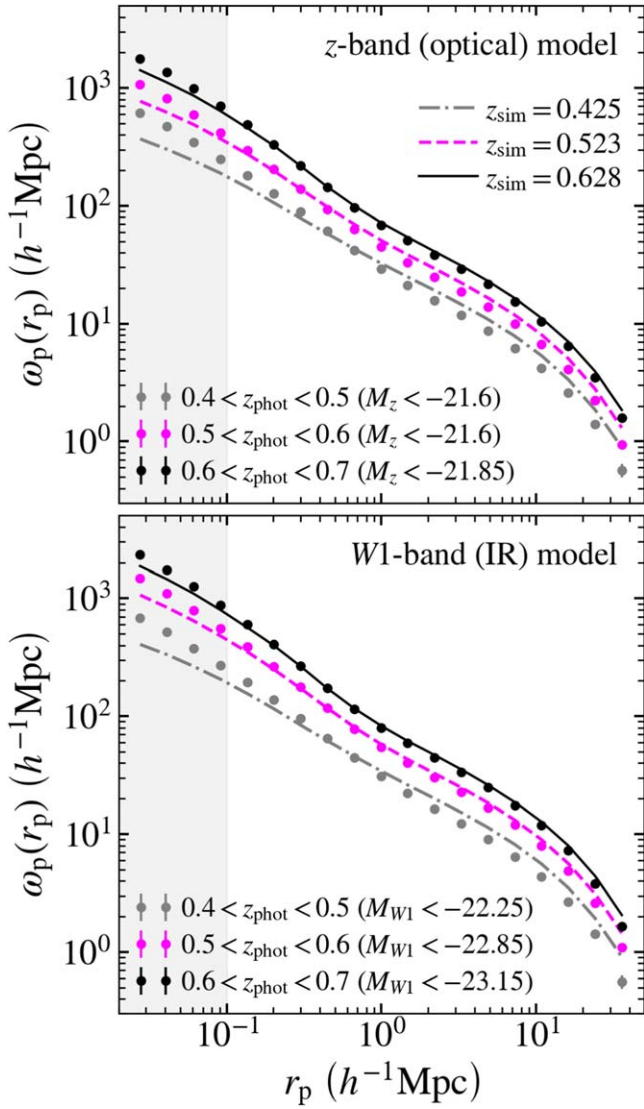


Figure 6. Projected clustering of the z -band (top panel) and $W1$ -band (bottom panel) absolute magnitude-limited parent galaxy samples and corresponding mock galaxy catalogs, with different colors and line styles representing different redshift bins as denoted by the legend in each panel. Redshift bins are offset by 0.15 dex for clarity.

show the distributions of all optical and IR DESI LRG targets with photometric redshifts between 0.4 and 0.7.

The color-coded points in each panel are luminous ($M_z < -23.7$) red-sequence objects in the same redshift range that are not DESI LRG targets despite passing the stellar cut (Equations 4(a) and 5(a); solid black line in panel (a)) and the optical and IR color-magnitude cuts (Equation 5(c), dashed black line in panel (e); and Equation 5(c), solid black line in panel (b)), as well as the z_{fiber} cut (Equations 4(d) and 5(d); not shown).

Panel (d) of Figure 9 shows that the low-redshift or blue color cut (Equations 4(d) and 5(b)) excludes these objects from selection as DESI LRG targets.

4. Predicted LRG Properties

In this section, we present the results of our z - and $W1$ -band models. We report the HOD parameters of mock LRGs and compare the predicted clustering to the data and relevant studies in the literature.

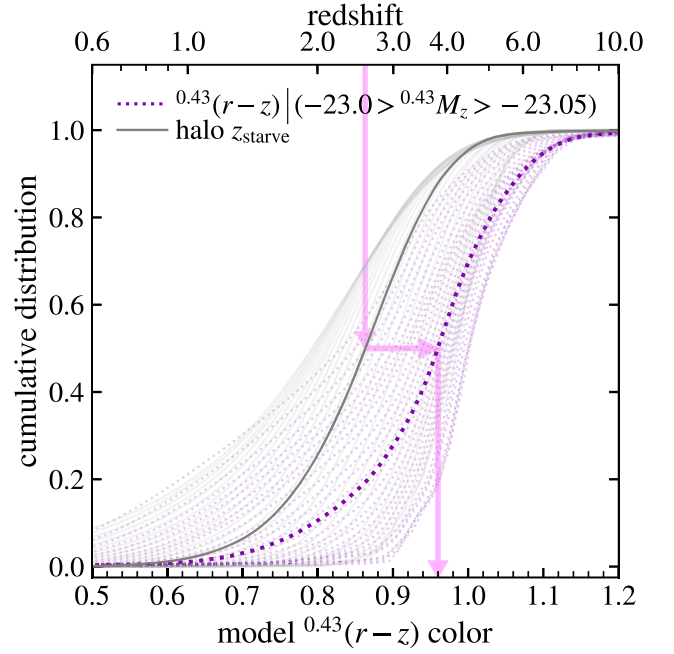


Figure 7. Illustration of the model color assignment procedure (Section 3.4), which equates halo z_{starve} with galaxy color at fixed luminosity. The dotted purple curve is the cumulative distribution of $^{0.43}(r-z)$ color for galaxies in the $-23.0 > ^{0.43}M_z > -23.05$ luminosity bin, while the solid gray curve is the cumulative distribution of halo z_{starve} for mock galaxies in the same luminosity bin. The magenta arrows indicate that a mock galaxy in this luminosity bin in a halo with $z_{\text{starve}} \approx 2.6$ is assigned a $^{0.43}(r-z)$ color of 0.96. The faint purple and gray curves are the galaxy color and halo z_{starve} distributions of additional luminosity bins.

4.1. LRG Clustering

Our z - and $W1$ -band models yield magnitude-limited mock galaxy catalogs from which we select mock LRGs. From these results, we predict the clustering of DESI LRG target samples.

Figure 10 shows the predicted clustering of IR and optical mock LRGs compared to the relevant DESI LRG target sample. In addition to the data (black points with error bars), each panel of Figure 10 shows the clustering of mock LRGs according to the z - and $W1$ -band models, with and without the addition of scatter to the model color- z_{starve} relation (Section 3.4).

As shown in the top row of Figure 10, the z - and $W1$ -band models predict very similar clustering signals for IR LRGs. The predicted amplitude of both the one- and two-halo terms exceeds that of the data in the two lowest redshift bins and is in better agreement in the highest redshift bin ($z_{\text{sim}} = 0.628$). Across all redshift bins, the addition of scatter to the color- z_{starve} relation brings the predicted clustering amplitude of IR LRGs closer to the data, reducing the clustering amplitude of the one-halo term by $\sim 15\%$ – 20% and the two-halo term by $\sim 10\%$, depending on the redshift bin. However, this reduction in the discrepancy between the model and data is a ceiling achieved only by adding such a high degree of scatter that color is uncorrelated with our proxy for halo assembly history⁹ (z_{starve}).

⁹ We also tested using proxies for IR and optical LRGs by selecting the N brightest mock galaxies from the relevant magnitude-limited mock galaxy catalog to recreate the number densities of the IR and optical DESI LRG target samples, instead of applying the DESI LRG target selections to the magnitude-limited mocks as described in Section 3.5. The result was a slightly larger predicted clustering amplitude compared to the “random colors” models that do use the DESI LRG target selections, shown in Figure 10.

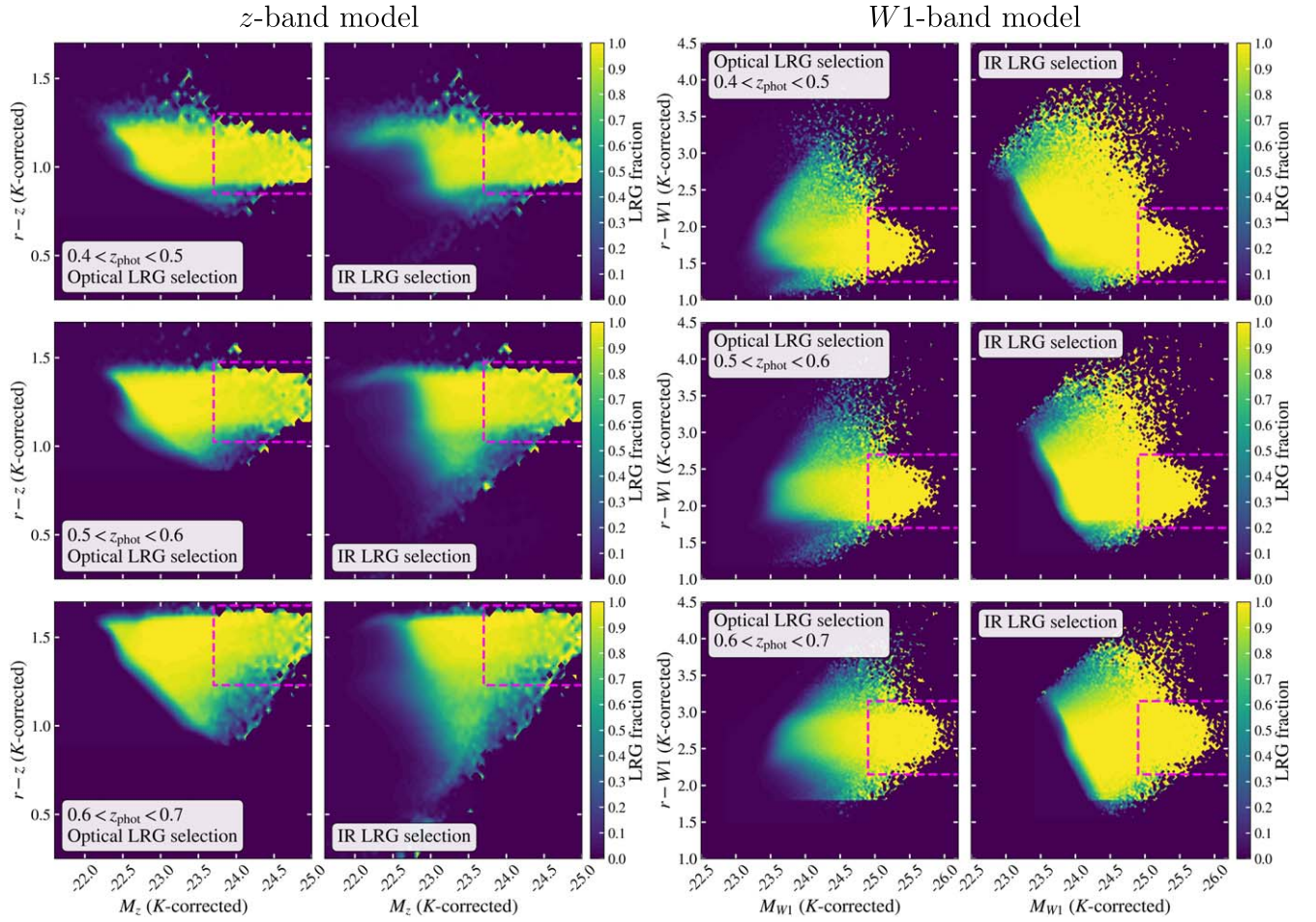


Figure 8. The LRG fractions in color–magnitude space used to select optical and IR LRG samples from our magnitude-limited mock galaxy catalogs. The first and second columns show optical and IR LRG targets, respectively, in $r - z$ color vs. M_z magnitude space. The third and fourth columns show optical and IR LRG targets, respectively, in $r - W1$ color vs. M_{W1} magnitude space. Each row shows a different redshift bin. Note that the distributions of optical LRGs in the z -band model (first column) and IR LRGs in the $W1$ -band model (fourth column) are compact by design; i.e., the fraction of galaxies that are LRGs is either ~ 0 or ~ 1 nearly everywhere in color–magnitude space. In contrast, the distributions of IR LRGs in the z -band model (second column) and optical LRGs in the $W1$ -band model (third column) have a broader gradient between zero and 1, especially toward fainter magnitudes. The boxed region in each panel marks the luminous end of the red sequence, where essentially all galaxies are expected to be LRGs. The LRG fraction in this boxed region is at least 98% for the $W1$ -band model (third and fourth columns) but only 92%–94% for the z -band model (first and second columns).

The bottom row of Figure 10 shows that the z - and $W1$ -band models predict different clustering amplitudes for optical LRGs. The z -band model overpredicts the clustering amplitude of optical LRGs to a similar degree as the overprediction for IR LRGs. Assigning color at random such that galaxy color is uncorrelated with halo assembly history again reduces this discrepancy of the one-halo (two-halo) term by $\sim 20\%$ – 25% ($\sim 10\%$ – 15%).

On the other hand, the $W1$ -band model prediction for the clustering of optical LRGs is in very good agreement with the data, especially the two-halo term (the one-halo term is actually slightly underpredicted). Additionally, adding any amount of scatter to the color– z_{starve} relation has no effect on the predicted clustering amplitude in this case.

The distributions of optical and IR LRGs in color–magnitude space are helpful for interpreting Figure 10. As Figure 3 shows, in optical space ($M_r - M_z$ versus M_z), both optical and IR LRGs are largely confined to the red sequence, with the distribution extending below it toward bluer color, especially at higher redshift.

In contrast, in IR space ($M_r - M_{W1}$ versus M_{W1}), the galaxy distribution does extend significantly above the red sequence.

For example, in the $0.4 < z_{\text{phot}} < 0.5$ redshift bin, the red sequence is at $M_r - M_{W1} \sim 1.6$, but the distribution extends to $M_r - M_{W1} \gtrsim 3.0$. The IR LRGs occupy the entire region of excess $M_r - M_{W1}$ color across the full range of M_{W1} , while optical LRGs occupy only the part of this region that corresponds to more luminous M_{W1} ; i.e., the optical LRG selection excludes a population of galaxies with very red $M_r - M_{W1}$ colors and moderately luminous $W1$ -band luminosities that are included by the IR selection.

Figure 11 shows optical ($r - z$) versus IR ($r - W1$) color for both IR and optical LRGs in the $0.4 < z_{\text{phot}} < 0.5$ redshift bin, while the top three panels of Figure 12 explore the differences in the color distributions of optical and IR LRGs located above the red sequence (box A in Figure 3) in $r - W1$ versus M_{W1} color–magnitude space. The bottom three panels (box B in Figure 3) show the color distributions of optical and IR LRGs from the same M_{W1} range along the red sequence.

Along the infrared red sequence, the spectral energy distributions of the optical and IR LRGs are quite similar. However, the two selections diverge above the red sequence. Table 5 lists the mean, median, and skewness of each color distribution, as well as the number of LRGs in each population.

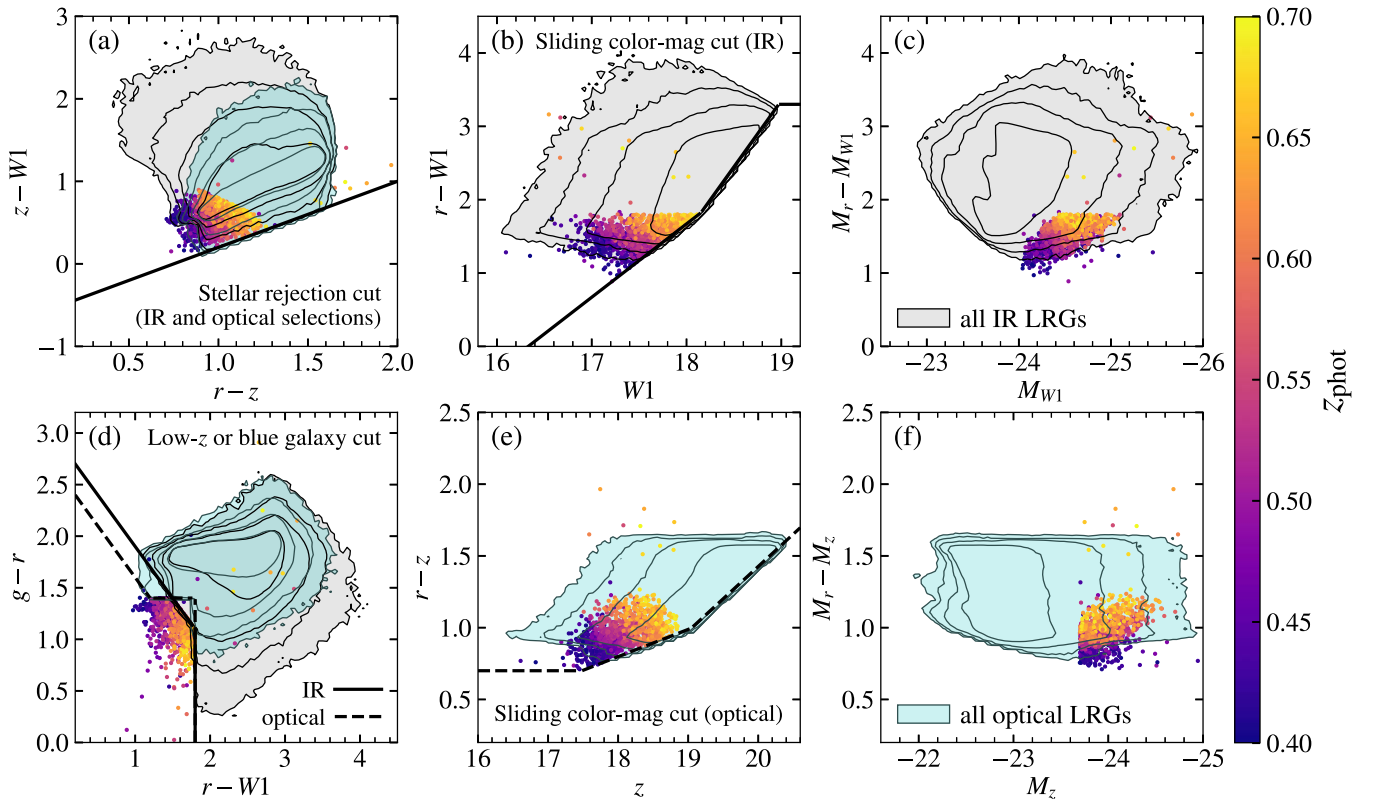


Figure 9. The color-coded points in each panel are bright ($M_z < -23.7$) red-sequence objects with photometric redshifts between 0.4 and 0.7 that are not DESI LRG targets, i.e., objects with $M_z < -23.7$ in the first two columns of Figure 8 that cause the LRG fraction at the bright end of the red sequence to be < 1 . The dashed and solid black lines in panels (a), (b), (d), and (e) are the various color-color and color-magnitude cuts that define the optical and IR DESI LRG target samples (Equations (4a) and (5a), respectively). The light blue and gray contours show the distributions of all optical and IR DESI LRG targets with photometric redshifts between 0.4 and 0.7.

Above the red sequence, IR LRGs have bluer $r-z$ and $g-r$ colors than optical LRGs in the same region of color-magnitude space, while IR LRGs have redder $r-W1$ colors than their optical counterparts from the same color-magnitude region. The same general trend is seen along the red sequence, but the color differences between the optical and IR LRG selections are smaller; i.e., along the red sequence, IR LRGs also have bluer $r-z$ and $g-r$ colors than optical LRGs but by only half as much as above the red sequence. Similarly, IR LRGs along the red sequence have redder $r-W1$ colors than their optical counterparts, but the color difference is two to three times smaller than for IR and optical LRGs above the red sequence.

Activity from active galactic nuclei (AGNs) may artificially inflate the observed IR (W1-band) luminosities and redden the $r-W1$ colors of some galaxies that pass the IR LRG selection (e.g., Webster et al. 1995; Georgakakis et al. 2009; Banerji et al. 2012; Glikman et al. 2012; Kim & Im 2018; Klindt et al. 2019; Rivera et al. 2021) relative to galaxies with comparable IR luminosities and $r-W1$ colors but no AGN activity. The bluer $r-z$ colors of the IR LRG targets cause these objects to be excluded by the optical LRG selection (e.g., panel (e) of Figure 9). Our IR (W1-band) model would assign mock IR LRGs with artificially red $r-W1$ colors and high IR luminosities due to AGN activity to halos with higher bias than if their $r-W1$ colors and W1-band magnitudes were not influenced by AGN activity. This could explain why our W1-band model overpredicts the clustering amplitude of IR LRGs relative to optical LRGs, as shown in Figure 10.

The influence of AGN activity may also explain the lack of correlation between IR and optical color when assigning colors to mock galaxies based solely on a proxy for halo age (here z_{starve}). This modeling assumption attributes galaxy color entirely to halo mass accretion history and does not account for how baryonic effects such as AGN activity may contribute to galaxy color. In other words, optical color is not necessarily a reliable proxy for IR color. This is also illustrated by Figure 11, which shows optical ($r-z$) versus IR ($r-W1$) color for both IR and optical LRGs in the $0.4 < z_{\text{phot}} < 0.5$ redshift bin. The optical and IR colors of optical LRGs are more closely correlated than for IR LRGs, but in both cases, there is considerable scatter in $r-W1$ at fixed $r-z$. Xu et al. (2022) also found a lack of correlation between halo assembly properties and galaxy color using an SAM.

4.2. The LRG-Halo Connection

We compute the mean central and satellite halo occupation statistics as a function of halo mass directly from the mock galaxy catalogs and halo abundances from the MDPL2 simulation. The results are shown in Figure 13. Each panel of Figure 13 shows $\langle N|M_{\text{vir}} \rangle$ for optical and IR LRGs in purple and orange, respectively.

The key conclusions of Figure 13 are summarized in Table 6, which gives the peak value of $\langle N_{\text{cen}} \rangle$ for optical and IR LRGs predicted by both the z - and W1-band models, as well as the mean value of $\langle N_{\text{cen}} \rangle$ for the most massive halos, which is universally less than the corresponding peak value.

Table 5Statistics of Distributions of $r - z$, $g - r$, and $r - W1$ Colors of Optical and IR LRGs Selected from above (Box A) and along (Box B) the Red Sequence in IR Color–Magnitude Space ($r - W1$ vs. M_{W1})

Box A (LRGs above Red Sequence)							Box B (LRGs along Red Sequence)						
0.4 < z _{phot} < 0.5													
	(r − z)		(g − r)		(r − W1)			(r − z)		(g − r)		(r − W1)	
Mean	1.04	1.10	1.64	1.74	2.30	2.18	Mean	0.98	1.01	1.72	1.78	1.74	1.68
Median	1.06	1.10	1.65	1.74	2.25	2.14	Median	0.98	1.00	1.78	1.80	1.73	1.67
Skewness	−1.86	0.44	−0.36	1.15	1.44	1.30	Skewness	−0.63	0.33	−1.20	−0.73	0.12	−0.02
N _{IR}	170,492		N _{IR} /N _{opt} = 3.44				N _{IR}	313,311		N _{IR} /N _{opt} = 0.60			
N _{opt}	49,517						N _{opt}	521,970					
0.5 < z _{phot} < 0.6													
	(r − z)		(g − r)		(r − W1)			(r − z)		(g − r)		(r − W1)	
Mean	1.18	1.29	1.71	1.87	2.64	2.53	Mean	1.13	1.18	1.71	1.80	2.14	2.11
Median	1.20	1.30	1.73	1.88	2.61	2.49	Median	1.16	1.20	1.79	1.81	2.16	2.12
Skewness	−0.99	−1.17	−0.60	0.23	1.11	1.68	Skewness	−0.83	−0.82	−1.30	−1.14	−0.46	−0.41
N _{IR}	179,348		N _{IR} /N _{opt} = 3.51				N _{IR}	441,691		N _{IR} /N _{opt} = 0.67			
N _{opt}	51,093						N _{opt}	660,486					
0.6 < z _{phot} < 0.7													
	(r − z)		(g − r)		(r − W1)			(r − z)		(g − r)		(r − W1)	
Mean	1.39	1.47	1.77	1.90	2.96	2.90	Mean	1.35	1.41	1.69	1.79	2.56	2.54
Median	1.42	1.48	1.79	1.90	2.92	2.87	Median	1.40	1.43	1.76	1.81	2.59	2.56
Skewness	−1.31	−0.84	−0.39	0.04	1.36	1.90	Skewness	−1.43	−1.23	−0.88	−0.61	−0.85	−0.74
N _{IR}	249,983		N _{IR} /N _{opt} = 1.88				N _{IR}	748,591		N _{IR} /N _{opt} = 0.72			
N _{opt}	132,656						N _{opt}	1,044,175					

Note. The boxed regions A and B are shown for the $0.4 < z_{\text{phot}} < 0.5$ redshift bin in Figure 3, and the color distributions are shown in Figure 12. The mean, median, and skewness of each distribution are shown for IR (optical) LRGs in bold (plain) text.

The z -band model does not predict significant differences between the populations selected by the IR and optical LRG selection functions. According to this model, the central LRG halo occupation fraction peaks at $\sim 90\%$ at $z \sim 0.43$ and drops to $\sim 70\%$ at the largest halo masses, while by $z \sim 0.63$, $\langle N_{\text{cen}} \rangle$ peaks at 82% – 84% and falls to $\sim 35\%$ for the most massive halos regardless of LRG selection function.

In contrast, the $W1$ -band model does predict a different LRG–halo relationship for IR versus optical LRGs. In all redshift bins, this model has $\langle N_{\text{cen}} \rangle$ for IR LRGs peaking at around 92% and remaining at $\gtrsim 75\%$ for the most massive halos, while $\langle N_{\text{cen}} \rangle$ for optical LRGs peaks at 85% – 92% and falls as low as 62% ($z \sim 0.43$) to 69% ($z \sim 0.52$) at the largest halo masses, although it remains at 92% for the highest redshift bin ($z \sim 0.63$).

We fit a standard five-parameter HOD model to each optical and IR mock LRG sample (Tables 7 and 8). For ease of comparison with their results, we use the same functional form as Zhou et al. (2021), who derived HOD parameters from clustering measurements of DESI LRG targets using photometric redshifts in bins that approximately match the redshift bins we use. This parameterization is described in detail in Zheng et al. (2005, 2007). Briefly, the probability $\langle N_{\text{cen}} \rangle$ that a halo of virial mass M_{vir} hosts a central galaxy is given by

$$\langle N_{\text{cen}} | M_{\text{vir}} \rangle = \frac{1}{2} \left(1 + \operatorname{erf} \left[\frac{\log(M_{\text{vir}}) - \log(M_{\text{min}})}{\sigma_{\log M}} \right] \right), \quad (15)$$

where the parameter $\log(M_{\text{min}})$ is the minimum halo mass for hosting a central galaxy, and the parameter $\sigma_{\log M}$ defines the steepness of the transition of $\langle N_{\text{cen}} \rangle$ from $\sim$ zero at low M_{vir} to ~ 1 at high M_{vir} .

The mean number of satellite galaxies, $\langle N_{\text{sat}} \rangle$, hosted by a halo of virial mass M_{vir} is approximated by a power law given by

$$\langle N_{\text{sat}} | M_{\text{vir}} \rangle = \left(\frac{M_{\text{vir}} - M_0}{M_1} \right)^\alpha, \quad (16)$$

where M_0 , M_1 , and α are the remaining free parameters of the HOD fit. The HOD fits of Zhou et al. (2021) incorporate a sixth nuisance parameter to account for photometric redshift uncertainty, which in this work is addressed by the magnitude-dependent line-of-sight scatter parameter, σ_{LOS} (Equation (14)). Table 7 gives the prior interval for each HOD parameter. The best-fit HOD parameters and the satellite fraction of each mock LRG sample are given in Table 8.

4.3. Comparison with Other DESI-like LRG Studies

There are two main studies of the LRG–halo connection of DESI-like LRGs suitable for comparison with our results. Zhou et al. (2021; who provided the photometric redshifts used here) measured the clustering of LRGs at $0.41 < z_{\text{phot}} < 0.93$ selected from DECaLS DR7 photometry using the optical target selection (Equation (4a)) and fit a standard five-parameter analytic HOD model in five redshift bins.

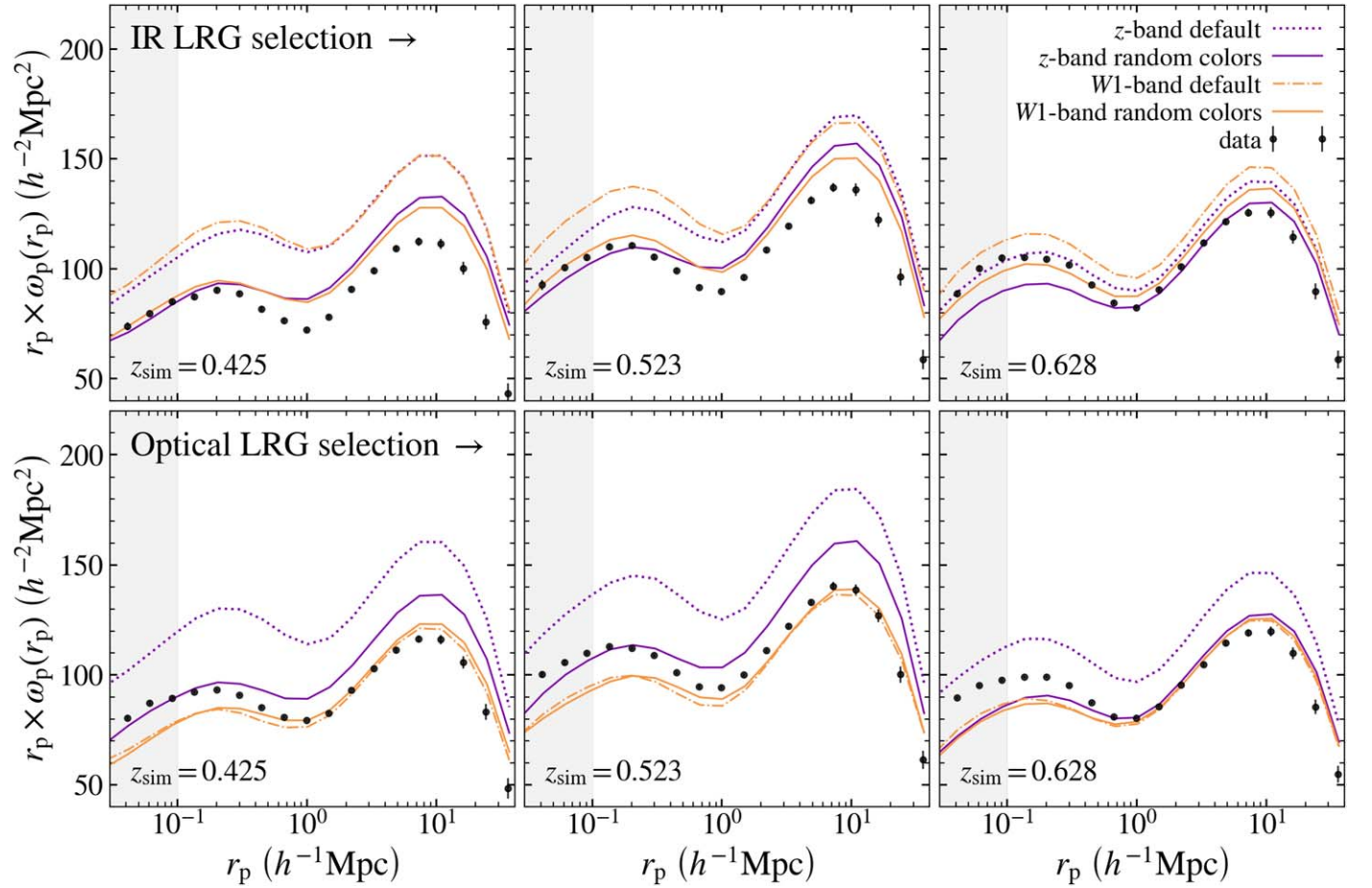


Figure 10. Predicted clustering of IR (top row) and optical (bottom row) mock LRGs compared to the clustering of the relevant LRG target samples from the data (black points in each panel). Dotted purple lines in each panel show mock LRGs from the default z -band model (monotonic correspondence between $r - z$ color and z_{starve} at fixed luminosity), while solid purple lines show LRGs from the z -band model with randomized colors. Dashed orange lines in each panel show mock LRGs from the default W1-band model, and solid orange lines show LRGs from the W1-band model with randomized colors. Each column shows a different redshift bin.

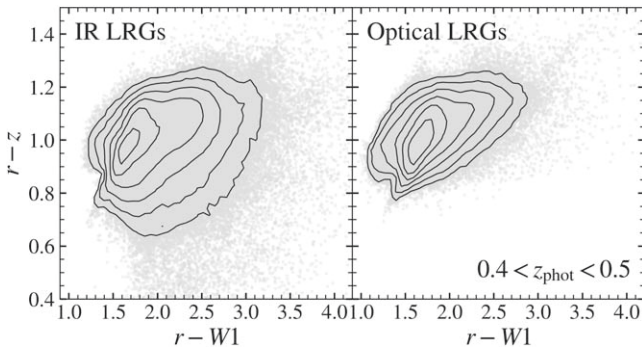


Figure 11. The IR ($r - W1$) vs. optical ($r - z$) color for IR (left panel) and optical (right panel) LRGs at $0.4 < z_{\text{phot}} < 0.5$.

Zhou et al. (2021) found little to no evolution of HOD parameters for LRGs across $0.4 \lesssim z \lesssim 0.8$. Our results are consistent with a lack of HOD parameter evolution across the subset of this redshift range that we model, although some of our predicted parameter values themselves differ from those of Zhou et al. (2021), particularly $\sigma_{\log M}$. Zhou et al. (2021) found a very steep transition from $\langle N_{\text{cen}} \rangle \sim 0$ to 1 ($\sigma_{\log M} \sim 0-0.28$; see their Figure 13), while our models predict a more gradual transition, with $\sigma_{\log M} \sim 1$.

Figure 14 compares our predicted LRG halo occupation statistics at $z \sim 0.63$ with the best-fit HOD parameterization of

Zhou et al. (2021) in the relevant redshift bin from their study ($0.61 < z_{\text{phot}} < 0.72$). Of the redshift bins Zhou et al. (2021) used that overlap with this work, this is the bin in which they found the greatest deviation of $\langle N_{\text{cen}} \rangle$ from a step function, although we also select this bin for Figure 14 to enable comparison with an additional study, discussed later in this section.

Our HOD predictions for satellite LRGs agree with those of Zhou et al. (2021) at halo masses below $\log(M_{\text{vir}}/h^{-1}M_{\odot}) \sim 14$, and our overall satellite fractions of $f_{\text{sat}} \sim 0.14-0.15$ across both model bands and LRG selections is consistent with Zhou et al. (2021), who found $f_{\text{sat}} \sim 0.13-0.16$, depending on redshift.

Zhou et al. (2021) obtained best-fit parameters for the analytic forms of central and satellite LRG HODs (Equations (15) and (16)), which will necessarily correspond to 100% of massive halos containing a central LRG due to the functional form of Equation (15) (this is also the case for analytic HOD fits to our mock LRG samples, shown in Table 8). Enforcing 100% halo occupation at the high-mass end by LRGs also suppresses accounting for potential contamination by lower-mass halos.

As Figure 13 shows, our models do not predict that central LRG halo occupation reaches unity, instead peaking at 82%–96% and falling in most cases to $\sim 60\%-77\%$ for the most massive halos ($\log(M_{\text{vir}}/h^{-1}M_{\odot}) \gtrsim 14, 75$), although this depends on both model band (z or W1) and LRG selection (IR or optical; see Table 6).

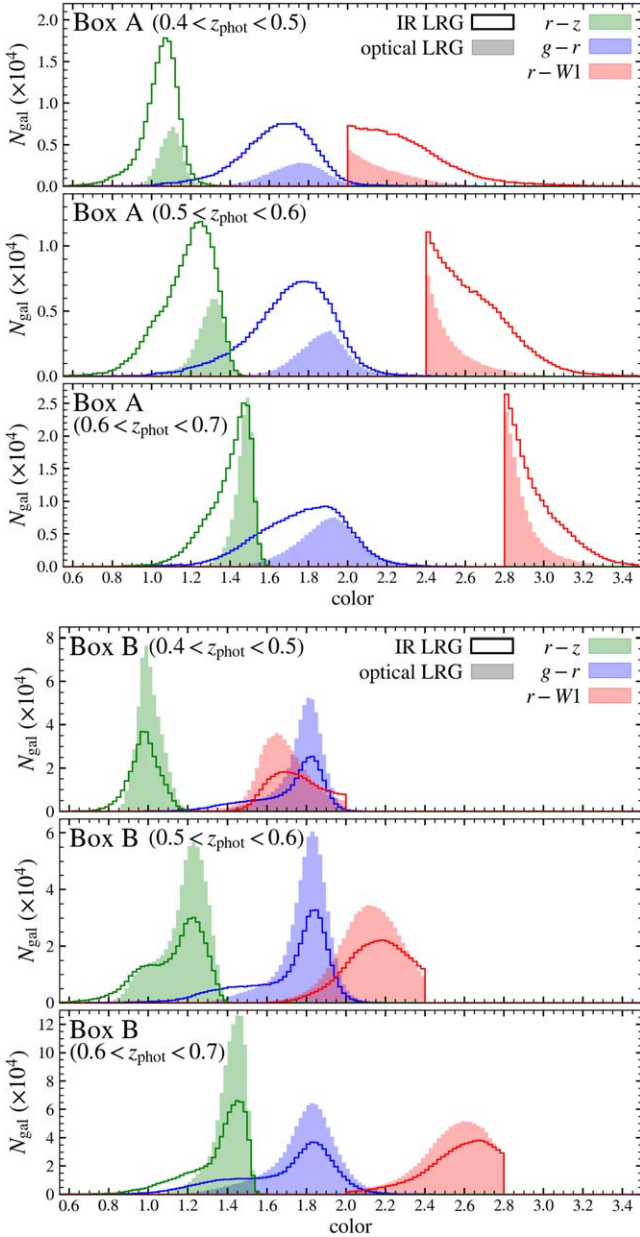


Figure 12. Distributions of $r-z$, $g-r$, and $r-W1$ colors of optical (shaded histograms) and IR (open histograms) LRGs from above (box A; top three panels) and along (box B; bottom three panels) the red sequence in $r-W1$ vs. M_{W1} color-magnitude space. The regions of color-magnitude space corresponding to boxes A and B are shown in Figure 3. The mean, median, and skewness of each distribution, as well as the number of LRGs in each population, are given in Table 5. Optical and IR LRGs from along the red sequence have similar distributions of each color, while optical LRGs above the red sequence have redder $r-z$ and $g-r$ and bluer $r-W1$ colors compared to IR LRGs in the same $W1$ -band luminosity range.

The other study to which we can compare our results is Hernández-Aguayo et al. (2021), who used the GALFORM semianalytic galaxy formation model (Cole et al. 2000) and nine snapshots from the Planck-Millennium N -body simulation (Baugh et al. 2018) to study the galaxy-halo connection of DESI-like LRGs at redshifts between 0.6 and 1.0. Their LRG HOD results for $z = 0.64$ are overlaid with our $z_{\text{sim}} \sim 0.63$ results in Figure 14.

Hernández-Aguayo et al. (2021) predicted central and satellite LRG abundances below our model predictions, although they noted that GALFORM underpredicts the

abundance of LRGs across the entire redshift range they studied, with the greatest discrepancy at $z \sim 0.6-0.7$. The shape of their central LRG HOD is notably similar to ours, consistent with a gradual transition to a peak central LRG halo occupation of less than unity that decreases at the largest halo masses. Specifically, Hernández-Aguayo et al. (2021) found that the $\langle N_{\text{cen}} | M_{\text{vir}} \rangle$ for DESI-like LRGs reaches a maximum of $\sim 70\%$ at $\log(M_{\text{vir}}/h^{-1}M_{\odot}) \sim 13.75$ and is as low as $\sim 40\%$ at $\log(M_{\text{vir}}/h^{-1}M_{\odot}) \sim 14.5$.

5. Summary and Conclusion

We have used a SHAM and age distribution matching framework to construct magnitude-limited mock galaxy catalogs at $z \sim 0.43, 0.52$, and 0.63 with $r-z$ and $r-W1$ colors. From these catalogs, we select mock LRG samples according to both the optical and IR DESI LRG target selection functions. This work is the first application SHAM modeling in the infrared, complimenting the few existing studies of the galaxy-halo connection of DESI-like LRGs.

Our models reproduce the number densities, luminosity functions, color distributions, and magnitude-dependent projected clustering (in the luminosity range of DESI LRG targets) of the parent galaxy samples from the Legacy Surveys DR9 photometry that serve as the basis for DESI LRG target selection. With the mock LRG samples selected from our magnitude-limited mock galaxy catalogs, we predict the halo occupation statistics of both optical and IR DESI LRGs at a fixed cosmology. We assess the differences between these two LRG populations, as well as the effect of using the z band versus the $3.4 \mu\text{m}$ $W1$ band for SHAM and age distribution matching.

The main results of this work are as follows.

1. Both the optical and IR DESI LRG target selections exclude some of the most luminous galaxies that would appear to be LRGs based on their position on the red sequence in optical color-magnitude space (Figure 8). This is a result of the specific DESI LRG target selection cut intended to exclude blue, low-redshift ($z \lesssim 0.4$) galaxies (Figure 9, Section 3.5).
2. Optical and IR LRGs occupy similar regions of optical color-magnitude space ($r-z$ versus M_z), where the red sequence corresponds to the very reddest $r-z$ colors (Figure 3). In IR color-magnitude space ($r-W1$ versus M_{W1}), there is a sizable galaxy population at significantly redder colors than the red sequence. The IR LRGs occupy most of this region of excess $r-W1$ color, but optical LRGs are largely excluded. This could be a result of AGN activity artificially inflating the $W1$ -band luminosities for some of these objects (Section 4.1).
3. There are clear distinctions between the LRG samples obtained from the optical versus IR selections that are apparent from the data alone, namely, among their optical ($r-z$ versus M_z) and IR ($r-W1$ versus M_{W1}) color-magnitude distributions, clustering, and color-color distributions. Our IR-based ($W1$ -band) model predicts greater differences between the halo occupation statistics of optical and IR LRGs than the z -band model (Figure 13, Table 6) and is therefore the preferred regime for this comparative study of the two selections.
4. Age distribution matching, which assumes a monotonic correlation between halo age and galaxy color at fixed luminosity, tends to overpredict the clustering amplitude of

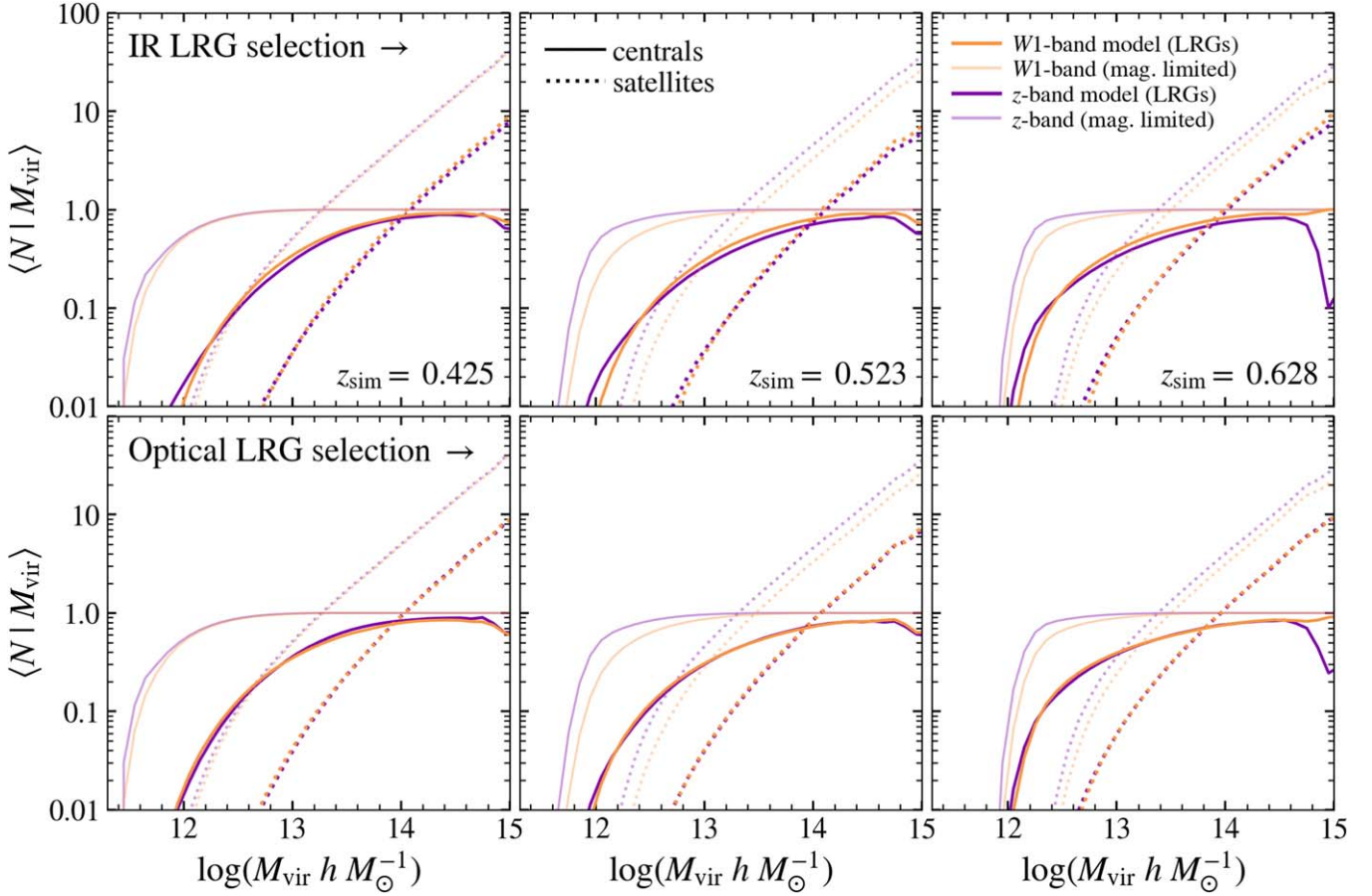


Figure 13. Predicted HODs of IR (top row) and optical (bottom row) mock LRGs from both the z -band (thicker purple lines) and W1-band (thicker orange lines) models. Also shown in fainter purple and orange lines are the HODs of the z - and W1-band magnitude-limited mock galaxy catalogs from which mock LRGs are selected. Solid (dotted) lines show results for central (satellite) galaxies and LRGs. Each column shows a different redshift bin. The best-fit values of a standard five-parameter HOD model, as well as the satellite fraction, for each mock LRG sample are given in Table 8.

Table 6
Predicted Halo Occupation Completeness for Mock Central LRGs by Selection (IR or Optical)

z_{sim}	LRG Selection	Peak $\langle N_{\text{cen}} M_{\text{vir}} \rangle^{\text{a}}$	Min. M_{vir} of Peak $\langle N_{\text{cen}} M_{\text{vir}} \rangle^{\text{b}}$	$\langle N_{\text{cen}} M_{\text{vir}} > 10^{14.75} h^{-1} M_{\odot} \rangle^{\text{c}}$
0.425	IR	0.92 0.91	14.45 14.85	0.74 0.70
	Optical	0.85 0.90	14.35 14.55	0.62 0.71
0.523	IR	0.91 0.87	14.45 14.65	0.77 0.61
	Optical	0.85 0.84	14.75 14.65	0.69 0.64
0.628	IR	0.96 0.82	14.85 14.55	0.96 0.31
	Optical	0.92 0.84	14.95 14.55	0.92 0.38

Notes. Results for the W1-band (z -band) model are in bold (plain) text.

^a Highest central occupation fraction reached across all halo M_{vir} .

^b Minimum halo M_{vir} at which the highest central occupation fraction is achieved.

^c Mean central occupation fraction at the largest halo masses ($M_{\text{vir}} \gtrsim 10^{14.75} h^{-1} M_{\odot}$), which is universally less than or equal to the peak central occupation fraction.

DESI LRGs (Figure 10). Introducing scatter into the assumed age–color relation improves the agreement between the predictions and the data somewhat, although the greatest improvement comes from increasing this scatter to a level that is equivalent to assigning colors at random. Our models therefore suggest that either galaxy color is uncorrelated with halo age in the LRG regime or there is some additional model parameter that has been neglected and is possibly related to our use of photometric redshifts.

- Both DESI LRG target selections yield populations with a nontrivial LRG–halo connection that does not reach

Table 7
Prior Ranges for HOD Fit Parameters (Equations (15) and (16))

Parameter	Prior Interval
$\log(M_{\text{min}})$	(11.0, 14.0)
$\sigma_{\log M}$	(0.001, 1.5)
α	(0.0, 2.0)
$\log(M_0)$	(11.0, 14.0)
$\log(M_1)$	(11.5, 15.5)

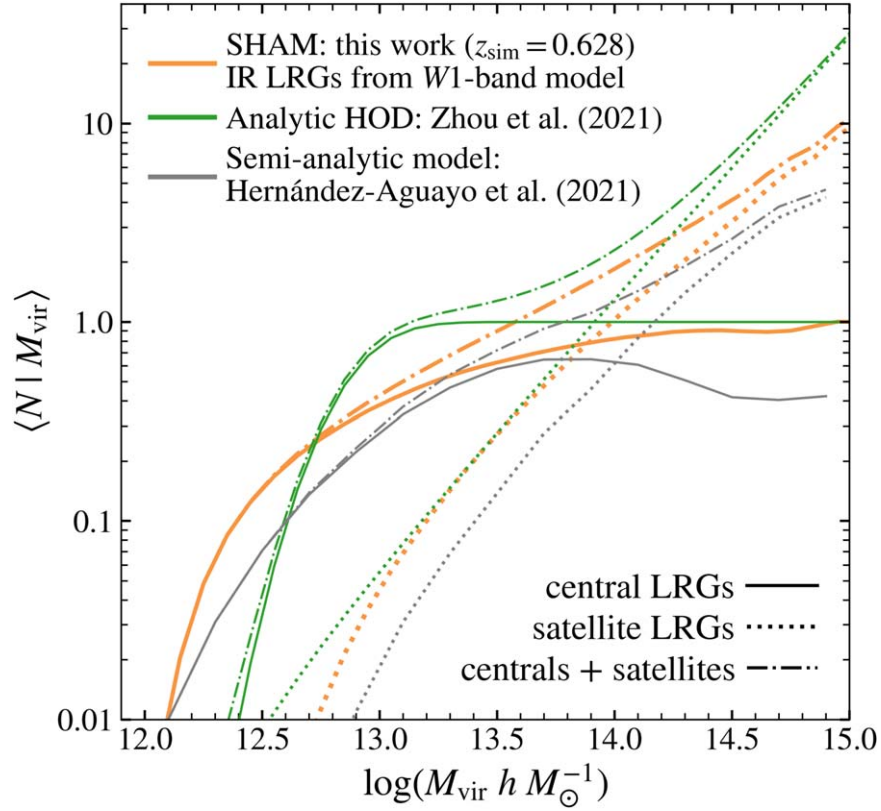


Figure 14. Comparison of LRG HOD parameters from this work at $z \sim 0.63$ (orange lines) with the analytic HOD fits of Zhou et al. (2021; green lines) and the SAM of Hernández-Aguayo et al. (2021; gray lines) in comparable redshift bins. Solid, dotted, and dashed–dotted lines show results for, respectively, central LRGs, satellite LRGs, and the combination of both.

Table 8

Best-fit Values of the Standard Five-parameter HOD Model (Equations (15) and (16)) and the Satellite Fraction, f_{sat} , of IR and Optical Mock LRG Samples

z_{sim}	LRG Selection	$\log(M_{\text{min}}/h^{-1}M_{\odot})$	$\sigma_{\log M}$	α	$\log(M_0/h^{-1}M_{\odot})$	$\log(M_1/h^{-1}M_{\odot})$	f_{sat}
0.425	IR	13.28 13.35	0.93 0.95	0.98 1.00	12.93 12.87	14.02 14.07	14.8 14.5
	Optical	13.32 13.26	1.05 0.93	0.95 0.98	12.91 12.89	14.00 14.00	14.9 14.8
0.523	IR	13.37 13.52	0.99 1.14	0.94 0.91	12.94 12.98	14.05 14.09	14.4 14.5
	Optical	13.45 13.44	1.15 1.11	0.95 0.96	12.84 12.85	14.05 14.05	14.7 14.6
0.628	IR	13.26 13.40	1.02 1.18	0.95 0.95	12.88 12.82	13.95 13.99	14.1 14.4
	Optical	13.30 13.31	1.19 1.16	0.96 0.94	12.78 12.82	13.93 13.93	14.4 14.7

Note. Results for the W1-band (z-band) model are in bold (plain) text.

unity ($\langle N_{\text{cen}} \rangle \sim 1$) for the most massive halos (Figure 13, Section 4.2). However, the IR selection achieves greater completeness ($\langle N_{\text{cen}} \rangle \geq 91\%$) than the optical selection ($\langle N_{\text{cen}} \rangle \sim 85\%–92\%$) across all redshift bins studied. Our results of $\langle N_{\text{cen}} \rangle < 1$ are qualitatively consistent with the SAM predictions of Hernández-Aguayo et al. (2021; Figure 14, Section 4.3), although they found a lower maximum completeness of $\sim 70\%$ at $z \sim 0.64$.

A natural extension of this work would be to utilize spectroscopic redshifts from the DESI BGS sample (Hahn et al. 2022) to conduct a similar study at lower redshift. The BGS¹⁰ is nearing completion and obtaining spectroscopic redshifts for a magnitude-limited ($r < 19.5$) sample of over 800 galaxies deg^{-2} at $z < 0.4$. While the LRG target selection algorithms

used in this work are designed to identify LRGs at $0.4 \lesssim z \lesssim 1.0$, these cuts could easily be modified to select LRGs in the same redshift range as the BGS. Mock LRGs could then be selected from SHAM-based magnitude-limited mock galaxy catalogs tuned to recreate the magnitude and color distributions of the BGS sample.

Acknowledgments

The authors thank Zheng Zheng for comments on an earlier draft of this work.

The Legacy Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS; Proposal ID 2014B-0404; PIs: David Schlegel and Arjun Dey), the Beijing–Arizona Sky Survey (BASS; NOAO Proposal ID 2015A-0801; PIs: Zhou Xu and Xiaohui Fan), and the Mayall z-band Legacy Survey (MzLS; Proposal ID 2016A-0453; PI: Arjun Dey). DECaLS, BASS,

¹⁰ Specifically, the BGS Bright sample; BGS also contains BGS Faint and AGN samples with different selection criteria.

and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. The Legacy Surveys project is honored to be permitted to conduct astronomical research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

This project used data obtained with the Dark Energy Camera (DECam), which was constructed by the Dark Energy Survey (DES) collaboration. Funding for the DES Projects has been provided by the U.S. Department of Energy, the U.S. National Science Foundation, the Ministry of Science and Education of Spain, the Science and Technology Facilities Council of the United Kingdom, the Higher Education Funding Council for England, the National Center for Supercomputing Applications at the University of Illinois at Urbana-Champaign, the Kavli Institute of Cosmological Physics at the University of Chicago, the Center for Cosmology and Astro-Particle Physics at the Ohio State University, the Mitchell Institute for Fundamental Physics and Astronomy at Texas A&M University, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo, Financiadora de Estudos e Projetos, Fundacao Carlos Chagas Filho de Amparo a Pesquisa do Estado do Rio de Janeiro, Conselho Nacional de Desenvolvimento Científico e Tecnológico and the Ministerio da Ciencia, Tecnologia e Inovacao, the Deutsche Forschungsgemeinschaft, and the Collaborating Institutions in the Dark Energy Survey. The Collaborating Institutions are Argonne National Laboratory, the University of California at Santa Cruz, the University of Cambridge, Centro de Investigaciones Energeticas, Medioambientales y Tecnologicas-Madrid, the University of Chicago, University College London, the DES-Brazil Consortium, the University of Edinburgh, the Eidgenössische Technische Hochschule (ETH) Zurich, Fermi National Accelerator Laboratory, the University of Illinois at Urbana-Champaign, the Institut de Ciències de l'Espai (IEEC/CSIC), the Institut de Física d'Altes Energies, Lawrence Berkeley National Laboratory, Ludwig Maximilians Universität München and the associated Excellence Cluster Universe, the University of Michigan, NSF's NOIRLab, the University of Nottingham, the Ohio State University, the University of Pennsylvania, the University of Portsmouth, SLAC National Accelerator Laboratory, Stanford University, the University of Sussex, and Texas A&M University.

BASS is a key project of the Telescope Access Program (TAP), which has been funded by the National Astronomical Observatories of China, the Chinese Academy of Sciences (the Strategic Priority Research Program "The Emergence of Cosmological Structures" grant No. XDB09000000), and the Special Fund for Astronomy from the Ministry of Finance. BASS is also supported by the External Cooperation Program of the Chinese Academy of Sciences (grant No. 114A11KYSB20160057) and the Chinese National Natural Science Foundation (grant No. 11433005).

The Legacy Surveys team makes use of data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), which is a project of the Jet Propulsion

Laboratory/California Institute of Technology. NEOWISE is funded by the National Aeronautics and Space Administration.

The Legacy Surveys imaging of the DESI footprint is supported by the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy under contract No. DE-AC02-05CH1123; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract; and the U.S. National Science Foundation, Division of Astronomical Sciences under contract No. AST-0950945 to NOAO.

The Photometric Redshifts for the Legacy Surveys (PRLS) catalog used in this paper was produced thanks to funding from the U.S. Department of Energy Office of Science, Office of High Energy Physics via grant DE-SC0007914.

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) and the Partnership for Advanced Supercomputing in Europe (PRACE; www.prace-ri.eu) for funding the MultiDark simulation project by providing computing time on the GCS Supercomputer SuperMUC at the Leibniz Supercomputing Centre (LRZ; www.lrz.de).

Some of the results in this paper have been derived using the `healpy` and `HEALPix` package.

ORCID iDs

Angela M. Berti  <https://orcid.org/0000-0003-3582-6649>

Kyle S. Dawson  <https://orcid.org/0000-0002-0553-3805>

References

- Banerji, M., McMahon, R. G., Hewett, P. C., et al. 2012, *MNRAS*, **427**, 2275
- Baugh, C. M., Gonzalez-Perez, V., Lagos, C. D. P., et al. 2018, *MNRAS*, **483**, 4922
- Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, *ApJ*, **717**, 379
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2012a, *ApJL*, **762**, L31
- Behroozi, P. S., Wechsler, R. H., & Conroy, C. 2013, *ApJ*, **770**, 57
- Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. 2012b, *ApJ*, **762**, 109
- Behroozi, P. S., Wechsler, R. H., Wu, H.-Y., et al. 2012c, *ApJ*, **763**, 18
- Berlind, A. A., & Weinberg, D. H. 2002, *ApJ*, **575**, 587
- Berlind, A. A., Weinberg, D. H., Benson, A. J., et al. 2003, *ApJ*, **593**, 1
- Blanton, M. R., & Roweis, S. 2007, *AJ*, **133**, 734
- Blum, R. D., Burleigh, K., Dey, A., et al. 2016, AAS Meeting, **228**, 317.01
- Bullock, J. S., Wechsler, R. H., & Somerville, R. S. 2002, *MNRAS*, **329**, 246
- Bundy, K., Leauthaud, A., Saito, S., et al. 2015, *ApJS*, **221**, 15
- Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, *MNRAS*, **319**, 168
- Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, *ApJ*, **647**, 201
- Dahlen, T., Mobasher, B., Faber, S. M., et al. 2013, *ApJ*, **775**, 93
- Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, *AJ*, **151**, 44
- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, **145**, 10
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv:1611.00036
- DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016b, arXiv:1611.00037
- Dey, A., Schlegel, D. J., Lang, D., et al. 2019, *AJ*, **157**, 168
- Flaugher, B., Diehl, H. T., Honscheid, K., et al. 2015, *AJ*, **150**, 150
- Gaia Collaboration, Brown, A. G. A., Vallenari, A., et al. 2018, *A&A*, **616**, A1
- Gaia Collaboration, Prusti, T., de Bruijne, J. H. J., et al. 2016, *A&A*, **595**, A1
- Georgakakis, A., Clements, D. L., Bendo, G., et al. 2009, *MNRAS*, **394**, 533
- Glikman, E., Urrutia, T., Lacy, M., et al. 2012, *ApJ*, **757**, 51
- Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, *ApJ*, **622**, 759
- Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023, *AJ*, **165**, 253
- Hearin, A. P., & Watson, D. F. 2013, *MNRAS*, **435**, 1313
- Hearin, A. P., Watson, D. F., Becker, M. R., et al. 2014, *MNRAS*, **444**, 729
- Hernández-Aguayo, C., Prada, F., Baugh, C. M., & Klypin, A. 2021, *MNRAS*, **503**, 2318
- Høg, E., Fabricius, C., Makarov, V. V., et al. 2000, *A&A*, **355**, L27
- Kim, D., & Im, M. 2018, *A&A*, **610**, A31
- Klindt, L., Alexander, D. M., Rosario, D. J., Lusso, E., & Fotopoulou, S. 2019, *MNRAS*, **488**, 3109
- Klypin, A., Yepes, G., Gottlöber, S., Prada, F., & Heß, S. 2016, *MNRAS*, **457**, 4340

- Landy, S. D., & Szalay, A. S. 1993, [ApJ](#), **412**, 64
- Loveday, J., Norberg, P., Baldry, I. K., et al. 2012, [MNRAS](#), **420**, 1239
- Moster, B. P., Naab, T., & White, S. D. M. 2013, [MNRAS](#), **428**, 3121
- Moustakas, J., Lang, D., Schlegel, D. J., et al. 2021, AAS Meeting, **237**, 527.04
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, [ApJ](#), **490**, 493
- Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, [JMLR](#), **12**, 2825
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, [A&A](#), **594**, A13
- Raichoor, A., Moustakas, J., Newman, J. A., et al. 2023, [AJ](#), **165**, 126
- Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2013, [ApJ](#), **771**, 30
- Reid, B., Ho, S., Padmanabhan, N., et al. 2015, [MNRAS](#), **455**, 1553
- Riebe, K., Partl, A. M., Enke, H., et al. 2013, [AN](#), **334**, 691
- Rivera, C. G., Alexander, D. M., Rosario, D. J., et al. 2021, [A&A](#), **649**, A102
- Safonova, S., Norberg, P., & Cole, S. 2021, [MNRAS](#), **505**, 325
- Saito, S., Leauthaud, A., Hearin, A. P., et al. 2016, [MNRAS](#), **460**, 1457
- Seljak, U. 2000, [MNRAS](#), **318**, 203
- Silva, D. R., Blum, R. D., Allen, L., et al. 2016, AAS Meeting, **228**, 317.02
- Sinha, M., & Garrison, L. 2017, Corrfunc: Blazing fast correlation functions on the CPU, Astrophysics Source Code Library, ascl:1703.003
- Sinha, M., & Garrison, L. 2019, in Software Challenges to Exascale Computing, ed. A. Majumdar & R. Arora (Singapore: Springer), 3
- Vale, A., & Ostriker, J. P. 2004, [MNRAS](#), **353**, 189
- Watson, D. F., & Conroy, C. 2013, [ApJ](#), **772**, 139
- Webster, R. L., Francis, P. J., Petersont, B. A., Drinkwater, M. J., & Masci, F. J. 1995, [Natur](#), **375**, 469
- Wechsler, R. H., Bullock, J. S., Primack, J. R., Kravtsov, A. V., & Dekel, A. 2002, [ApJ](#), **568**, 52
- Wechsler, R. H., & Tinker, J. L. 2018, [ARA&A](#), **56**, 435
- Wetzel, A. R., Tinker, J. L., Conroy, C., & Bosch, F. C. v. d. 2014, [MNRAS](#), **439**, 2687
- Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, [AJ](#), **140**, 1868
- Xu, X., Zheng, Z., & Guo, Q. 2022, [MNRAS](#), **516**, 4276
- Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2013, [ApJ](#), **770**, 115
- Yang, X., Mo, H. J., van den Bosch, F. C., Zhang, Y., & Han, J. 2012, [ApJ](#), **752**, 41
- York, D. G., Adelman, J., John E. Anderson, J., et al. 2000, [AJ](#), **120**, 1579
- Yu, J., Zhao, C., Chuang, C.-H., et al. 2022, [MNRAS](#), **516**, 57
- Zentner, A. R., Hearin, A. P., & van den Bosch, F. C. 2014, [MNRAS](#), **443**, 3044
- Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, [ApJ](#), **633**, 791
- Zheng, Z., Coil, A. L., & Zehavi, I. 2007, [ApJ](#), **667**, 760
- Zheng, Z., & Weinberg, D. H. 2007, [ApJ](#), **659**, 1
- Zhou, R., Dey, B., Newman, J. A., et al. 2023, [AJ](#), **165**, 58
- Zhou, R., Newman, J. A., Dawson, K. S., et al. 2020, [RAAS](#), **4**, 181
- Zhou, R., Newman, J. A., Mao, Y.-Y., et al. 2021, [MNRAS](#), **501**, 3309
- Zonca, A., Singer, L., Lenz, D., et al. 2019, [JOSS](#), **4**, 1298
- Zou, H., Zhou, X., Fan, X., et al. 2017, [PASP](#), **129**, 064101