Snowmass2021 Cosmic Frontier White Paper: High Density Galaxy Clustering in the Regime of Cosmic Acceleration

Kyle Dawson¹, Andrew Hearin², Katrin Heitmann², Mustapha Ishak³, Johannes Ulf Lange^{4,5}, Martin White^{6,7}, Rongpu Zhou⁷ for the Snowmass 2021 Cosmic Frontier 4 Topical Group

¹Department of Physics and Astronomy, University of Utah, Salt Lake City, UT 84112, USA ²High Energy Physics Division, Argonne National Laboratory, 9700 South Cass Avenue, Lemont, IL 60439, USA ³Department of Physics, The University of Texas at Dallas, Richardson, TX 75080, USA

⁴Kavli Institute for Particle Astrophysics and Cosmology and Department of Physics, Stanford University, CA 94305, USA ⁵Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95064, USA ⁶Berkeley Center for Cosmological Physics, Department of Physics, University of California, Berkeley, CA 94720, USA ⁷Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA

Abstract

Joint studies of imaging and spectroscopic samples, informed by theory and simulations, offer the potential for comprehensive tests of the cosmological model over redshifts z < 1.5. Spectroscopic galaxy samples at these redshifts can be increased beyond the planned DESI program by at least an order of magnitude, thus offering significantly more constraining power for these joint studies. Spectroscopic observations of these galaxies in the latter half of the 2020's and beyond would leverage the theory and simulation effort in this regime. In turn, these high density observations will allow enhanced tests of dark energy, physics beyond the standard model, and neutrino masses that will greatly exceed what is currently possible. Here, we present a coordinated program of simulations, theoretical modeling, and future spectroscopy that would enable precise cosmological studies in the accelerating epoch where the effects of dark energy are most apparent.

Introduction 1

Surveys to measure fluctuations in the cosmic microwave background (CMB) radiation and large-scale structure (LSS) through galaxies and quasars provide our best windows into the fundamental physics of the cosmos. The tightest constraints on dark energy, modifications to General Relativity, mass limits on light dark matter particles, and neutrino masses all come from one or both of these measurements. Within the next few years, we expect to see a dramatic leap in our ability to probe fundamental physics through cosmology with the Dark Energy Spectroscopic Instrument [DESI; 1, 2], Rubin Observatory's Legacy Survey of Space and Time (LSST), and CMB-S4. DESI has already begun, and is quickly mapping the three dimensional distribution of galaxies and the intergalactic medium. The LSST on Rubin will begin in a few years, providing deep imaging over unprecedented areas of the sky. With CMB-S4 closing out the decade, our view of the cosmic microwave background will advance every bit as much as in the optical.

With five years of spectroscopy from DESI, the expansion history of the universe will be well mapped. However, growth of structure, the full shape of the power spectrum, higher order clustering statistics, and cosmology from joint lensing and spectroscopic surveys will be less well constrained. The five-year DESI samples will allow initial investigations using these techniques, but advances in theory, simulations, and spectroscopic samples will allow significant advances.

In the latter half of this decade, DESI will remain the most competitive instrument in the world for large-scale spectroscopic surveys. New spectroscopic facilities have been proposed with even faster survey speeds, offering the potential to increase the number density of DESI galaxies by an order of magnitude or more. Long-term investments in spectroscopic surveys will allow enhanced tests of dark energy, physics beyond the standard model, and neutrino masses that will greatly exceed what is currently possible. A new sample at a much higher number number density that covers the full epoch of cosmic acceleration back to the matter dominated era will enable a unique spectroscopic probe of the following cosmological questions:

- 1. Under General Relativity, structure grows at a rate that can be predicted at any time given the Hubble parameter, H(z). New growth measurements over a range of redshifts will thus allow improved constraints on the expansion history and the equation of state for dark energy.
- 2. By comparing the expansion history derived from growth measurements to that from geometric measurements, we can test the self-consistency of General Relativity or whether additional physics is required.
- 3. The long redshift baseline afforded by growth measurements to z < 1.5 will further break degeneracies of standard parameters with common dark energy models, and hence improve constraints on the dark energy equation of state.
- 4. The amplitude of clustering with redshift depends on the sum of neutrino mass. A survey of structure growth over a wide redshift range would tighten constraints on the neutrino masses. If cosmology results favor the lowest allowed neutrino masses, then these measurements will also help resolve the question of mass hierarchy.
- 5. Precise measurements of the power spectrum afforded by a large number of tracers would provide tight limits on dark matter interactions, light relic particles, early dark energy, and potential features in the primordial power spectrum that appear on small scales.
- 6. Current lensing measurements of the amplitude of clustering indicate potential tension with predictions from the CMB under a Λ CDM model. As highlighted in the

report from the Astro2020 Decadal Survey, resolving this tension with higher precision growth measurements could reinforce the Λ CDM model or reveal extensions in the dark sector.

Order of magnitude increases over the initial DESI samples of z < 1.5 galaxies are within reach of several proposed spectroscopic programs. A survey of galaxies at a high number density would allow high precision measurements down to small scales, with the statistical power to constrain the clustering amplitude at better than 0.1% precision on scales comparable to a typical dark matter halo. With a large sample of spectroscopicallyconfirmed galaxies, multiple tracers can be identified for clustering and cross-correlation studies, thus alleviating shot noise and offering robust checks on systematic errors. Vastly increased sample sizes will also facilitate studies of higher order statistics that remain in the early stages of development. Finally, by probing the same volume with spectroscopy and imaging data, such a survey will reap the full benefits of lensing and three-dimensional clustering.

In this white paper, we describe a program to use full-scale clustering (where both non-linear and linear scales are jointly modeled) to test the cosmological model over the redshift interval 0 < z < 1.5. In Section 2, we provide a brief overview of current Stage-IV Dark Energy programs and expectations for constraints on fundamental physics. In Section 3, we describe the effort in theory and simulations that is required to reach the precision allowed by these new techniques. Investment in that theory and simulation work should begin now for future spectroscopic observations to be most effective. Finally, in Section 4, we describe the science drivers and new spectroscopic samples that are possible with existing and proposed facilities.

2 Stage-IV Dark Energy Surveys

Large-scale cosmology surveys spanning imaging, spectroscopy, and lensing will independently advance four of the five key science drivers identified by the 2014 Particle Physics Project Prioritization Panel (P5) [3]. Analyses that jointly utilize imaging and spectroscopy will substantially enhance the science returns that can be reaped from the data, as shown in Table 1.

Stage-IV Dark Energy surveys will pursue these science drivers in three approximate redshift intervals using complementary techniques. At the lowest redshifts (z < 1), Rubin will assess the distribution of dark matter on all scales through weak lensing, while DESI will assess the three dimensional clustering of galaxies with precise spectroscopy. Over moderate redshifts (1 < z < 2), Rubin will measure the distribution of galaxies through photometric redshifts, while DESI will provide three-dimensional positions with fiber-based spectroscopy. Finally, at the highest redshifts (z > 2), Rubin will measure the distribution of galaxies through photometric redshifts, and DESI will measure the three-dimensional distribution of matter through quasar and Lyman- α forest spectroscopy. In each of these redshift regimes, imaging and spectroscopy play critical, complementary roles. DESI in particular will provide essential fiber-based, optical spectroscopy over the full redshift range 0 < z < 4.

Question	Optimized approach	
What is the physical origin of cosmic acceleration at	z < 1.5 spectroscopy at	
late times (Dark energy)?	high number density + Ru-	
	bin lensing measurements	
What is the sum and hierarchy of the neutrino masses?	Full-shape power spectrum	
	measurements from z $<$	
	1.5 spectroscopy at high	
	density + characterization	
	of underlying dark matter	
	distribution through Rubin	
	lensing measurements	
What is the particle nature of dark matter?	three-dimensional cluster-	
	ing on small scales with	
	z < 1.5 spectroscopy at	
	high number density +	
	galaxy-galaxy lensing from	
	Rubin, Roman, & Euclid	
Does the rate of cosmic expansion and the growth of	All of the above	
structure indicate new particle or field content?		

Snowmass2021 Cosmic Frontier White Paper Template

Table 1: P5 advances enabled by efforts at z < 1.5 with theory, simulations, Rubin imaging, and DESI spectroscopy. Spectroscopic sample sizes significantly larger than those planned with DESI are possible in the near future, allowing tighter constraints on each of these science drivers.

2.1 DESI at z < 1.5

The currently planned DESI program will obtain a spectroscopic sample of roughly 13 million bright galaxies (BGS), seven million luminous red galaxies (LRG) at redshifts z < 1, and 16 million emission line galaxies at redshifts 0.6 < z < 1.6. Covering roughly 2/3 of the observable extragalactic sky, the baryon acoustic oscillation (BAO) distance measurements from DESI over redshifts z < 1.0 will be near the sample-variance limit. Increasing the footprint or adding new z < 1 targets will only have marginal returns for these BAO measurements. The DESI BAO distance measurements over 1 < z < 1.5 will be within a factor of two of the sample-variance limit for this same survey footprint, lending potential for future surveys of new areas and higher target densities.

These spectroscopic samples will also enable measurements of the growth of structure through redshift space distortions (RSD). RSD result from the imprint of gravitational infall on the measured redshifts of galaxies. These peculiar velocities record the history of structure growth which can be modelled with clustering measurements performed along and transverse to the line-of-sight. RSD introduce anisotropy to this three-dimensional clustering, typically parameterized as $f\sigma_8 = \frac{\partial \sigma_8}{\partial \ln a}$, where $a = (1+z)^{-1}$ is the dimensionless cosmic expansion factor. Evaluating the amplitude of clustering at redshift z = 0 under an assumption of a flat Λ CDM cosmology, σ_8 has been measured to a precision of 3.5% [4] using BAO and RSD in the Stage-III Baryon Oscillation Spectroscopic Survey [BOSS; 5]

and its successor, eBOSS [6]. Under the same Λ CDM model, with the same priors on the baryon density from Big Bang Nucleosynthesis (BBN), and the sample priors on the spectral index of the primordial power spectrum, the precision on σ_8 is expected to improve to 0.5% in DESI when using similar algorithms and scales to $0.20 \ h Mpc^{-1}$ to perform the RSD measurements. While a significant improvement over Stage-III, theoretical and modeling work will be needed to reach this goal. Additional data will be needed to reach the precision of 0.2% identified as a goal in the Astro2020 Decadal Survey to test consistency with the *Planck* [7] $\sigma_8 = 0.811 \pm 0.006$ constraint.

2.2 Rubin observatory at z < 2

The Vera C. Rubin Legacy Survey of Space and Time (LSST) uses the Simonyi Survey Telescope, which is designed to observe a 18,000 deg² region of the sky in six optical bands [8]. LSST will run for a decade allowing the detection of ~20 billion galaxies with photometric redshifts extending well-above unity. It will measure shapes for ~2 billion galaxies allowing lensing tomography to constrain the growth rate of structure and the equation of state of dark energy to reach the requirements for a Stage IV dark energy experiments as defined in the Dark Energy Task Force report [9]. It will also measure clustering of galaxies and their cross correlation with lensing measurement in the so-called $3 \times 2pt$ analysis. LSST will also discover and measure at least 500 Supernovae Type Ia (SNe Ia) per season which will provide tens of thousands of well-measured SNe Ia light curves up to $z \sim 1$ during the survey. LSST will discover ~100,000 clusters of galaxies. Finally, LSST will measure strong gravitational lensing and time delay for the multiple images providing a sample of ~2600 time-delayed lensing systems. A comprehensive analysis of the expected dark energy constraining power combining the probes can be found in Ref. [10].

2.3 New Approaches with Stage-IV Dark Energy Surveys

Whereas the expansion history will be well mapped by DESI at redshifts z < 1.5, new techniques will be developed to improve clustering measurements at smaller scales. As a step toward this goal, several projects within DESI have begun to characterize DESI galaxies through galaxy-galaxy lensing from public weak lensing surveys (KiDS, DES, and HSC) and CMB lensing in order to deduce cosmological constraints from full-scale clustering studies. Such cross-correlations often rely on a much higher number density than the BAO measurements, but may not need a spectrum of every object. Ongoing analyses will help establish the balance between photometric and spectroscopic sample sizes. Additional projects are also underway to compute higher order clustering statistics to better characterize scale-dependent bias that is degenerate with cosmological information at small scales. These studies will also help establish the relationship between number density and cosmological constraints.

3 Immediate Analysis Needs

Theoretical forecasts of the cosmological constraining power of the nonlinear regime of structure formation now date back many years [e.g., 11–14], and suggest that an improved ability to use smaller-scale information can result in factors of 2-4 improvement on dark energy constraints. The natural question arises as to whether these gains can be realized in practice, or whether the need to marginalize over nuisance parameters capturing systematic uncertainty results in an excessive loss of constraining power. Recent work analyzing BOSS galaxy samples has shown that these potential gains could indeed be a reality [15–17]; by including measurements from nonlinear scales, these recent analyses have achieved a full factor of 2 improvement beyond previous BOSS analyses [e.g. 18] that restricted attention to the quasi-linear regime.

In addition to pushing the boundaries of a single survey by improving analysis capabilities, cross-correlations of data from different wavebands will enhance the science returns from ongoing and upcoming surveys (see Ref. [19] for a detailed analysis). Therefore, a modest investment into an overarching simulation and modeling program that will enable the exploitation of small scales and cross-correlations has the potential to substantially extend the scientific reach of ongoing and upcoming surveys (see also Ref. [20]).

3.1 High-Fidelity Simulations

Simulations play an important role in fully exploiting the information available from cosmological surveys, in particular on smaller length scales. The simulation needs for ongoing and upcoming surveys span a wide range, from large numbers of realizations for covariance estimates, to detailed simulations that allow the creation of high-quality synthetic sky catalogs to enable tests of analysis pipelines and evaluation of systematic effects, to simulations that allow the exploration of physics beyond the confines of Λ CDM. Simulations beyond Λ CDM are important for many reasons. They provide predictions across cosmological parameters for a range of cosmological observables and enable the exploration of signatures of new physics beyond the current cosmological paradigm for both dark energy and dark matter.

The importance of simulation campaigns has been recognized widely in the community. Critical simulation campaigns have been carried out for contemporary surveys, including the MICE [21] and Buzzard simulations [22] for DES, the DC2 simulated catalogs for LSST DESC [23, 24], and the AbacusSummit suite for DESI [25]. A related White Paper [26] provides details regarding the roles of simulations in Cosmic Frontier science.

Conventional N-body simulations are now extremely mature tools; over the years, many code comparison and verification projects have been carried out, e.g., Refs. [27–32]. These studies have shown that contemporary N-body codes with appropriate settings for initial conditions, force and mass resolution and coverage of large enough volumes, agree at the 1% level out to scales of $k \sim 1h^{-1}$ Mpc, beyond which baryonic effects need to be taken into account. The statistical power of future spectroscopic surveys can be be 0.1% or better at these scales, offering substantial power to constrain the assumptions in N-body simulations. Simulation development over the course of the next decade is required to



Figure 1: The real-space power spectra of halos from the N-body simulations using GADGET-2 [33] and SWIFT [34] relative to the ABACUS [35] code. Figure reproduced from "The DESI *N*-body Simulation Project I: Testing the Robustness of Simulations for the DESI Dark Time Survey" [32].

keep pace with the statistical precision allowed by these potential surveys.

As a very recent example, in Ref. [32], several N-body simulation codes have been considered within DESI to test theoretical models and systematic errors in two-point clustering and other measurement techniques. These simulations must accurately predict the growth of the dark matter halos that host the galaxies that are used as tracers of the underlying dark matter distribution. The simulations codes under consideration make different approximations to gravitational interactions between particles and therefore slightly different predictions for the halo mass distribution and its clustering properties. At scales out to $k_{\rm max} = 0.20h \,{\rm Mpc}^{-1}$, where previous RSD measurements have been performed, there is almost perfect agreement between the high resolution SWIFT simulations and the ABACUS simulations (Figure 1). However, even for these two simulations, the agreement degrades to the 0.2% level around $k = 1h \,{\rm Mpc}^{-1}$. While these halos have not yet been populated with galaxies, the exercise demonstrates that a theoretical floor of at least 0.2% precision currently exists in the first step of modeling halo clustering at small scales.

3.2 Synthetic Skies from High-Fidelity Simulations

Carrying out sufficiently accurate simulations is only the first step, however, establishing robust connections to the observable universe is essential for making useful predictions for galaxy clustering on nonlinear scales – the process of galaxy formation is not directly modeled in N-body simulations. It is important, therefore, to develop improved techniques for relating the galaxy population to the underlying N-body simulations, a quantitative improvement in the so-called 'galaxy-halo' connection.

Many contemporary efforts to derive cosmological constraints from nonlinear scales are

built upon simplistic empirical models such as the Halo Occupation Distribution (HOD). Due to the very formulation of HOD-type models, incorporating new constraints from multiple redshifts and more than a single tracer galaxy population requires a significant expansion of the parameter space, and/or reliance upon plausibly-violated assumptions about the galaxy-halo connection. Thus conventional halo occupation models actually penalize attempts to incorporate new constraining data. Moreover, this older generation of models was devised at a time when consensus in the field had not yet been reached on the reliability of cosmological simulations to resolve halo substructure, and so the HOD and related models are founded merely upon host halos identified at a particular simulated snapshot. But the field of computational cosmology has seen dramatic progress in the quality of simulated data products over the last fifteen years: subhalo catalogs with merger trees are becoming widely available for high-resolution, survey-scale simulations [36–39]; thus the continued use of HOD-type models as the basis of our cosmological predictions fails to capitalize upon the now well-established ability of contemporary codes to track the evolution of halos and their substructure across cosmic time. These core limitations of traditional halo occupation models highlight how they bear the mark of the era in which they were developed. Hydrodynamical simulations or Semi-Analytic Models (SAMs) have also been used to generate predictions across redshift in a physically motivated way. These models remain irreplaceable in the effort to understand the detailed physics of galaxies and clusters, but the full scientific potential of data in the 2020s can only be delivered with expansive explorations of parameter space based on high-resolution, Gpc-scale simulations, and so new techniques beyond the traditional implementations of these models are needed.

Considerable recent progress has been made by a promising new generation of empirical models that bridge the gap between the level of complexity achieved by SAMs and the computational efficiency of empirical models, e.g., UniverseMachine [40] and EMERGE [41]. The ability of these models to make predictions for multiple tracers across redshift is promising, but further advances are needed on both the modeling and computation side for this new approach to conduct cosmological inference with survey-scale simulations. To meet the predictive needs and to maximize the scientific returns of the upcoming surveys, we consider it critical to invest in the development of a new generation of modeling approaches that address the limitations of contemporary techniques.

3.3 Lack of Inter-collaboration structures

The general requirements on simulations to analyze surveys are similar for the different surveys. While some surveys need larger volumes, other surveys possibly require higher mass resolution. However, developing an overarching simulation program that allows surveys to share simulations and derived data products such as synthetic catalogs would be very valuable. In Ref. [42] a strong case was made for a joint simulation program across the Vera Rubin Observatory (Rubin) Dark Energy Science Collaboration (DESC), the Nancy Grace Roman Space Telescope, and Euclid. This case clearly carries over to DESI, CMB-S4, and other future surveys as well.



Figure 2: Number of spectroscopic tracers required to reach $\bar{n}P(k) = 1$ as a function of redshift for scales $k_{\text{max}} = 0.2$, 0.5, and 1.0 $h \,\text{Mpc}^{-1}$. The DESI BGS, LRG, and ELG number densities are as expected based on current survey operations. Left: Number density as a function redshift. Right: Cumulative surface density as a function of redshift. The dependence of large-scale bias on number density is based on extrapolations of the HOD models [43] derived from clustering in the photometric catalogs of the DESI LRG sample [44], an assumed halo mass function [45], and an assumed halo mass-halo bias relation [46]. The large-scale bias decreases by a factor of roughly 1.7 between the number densities reported for $k_{\text{max}} = 0.2 h \,\text{Mpc}^{-1}$ and those reported for $k_{\text{max}} = 1.0 h \,\text{Mpc}^{-1}$.

4 Future Survey Prospects

Advances in simulations and the theory connecting galaxies to dark matter will allow more accurate modeling of the cosmological signal. Meanwhile, future spectroscopic programs have the potential to significantly improve the statistical precision allowed in these analyses. After DESI completes, there will still be a very large number of spectroscopic modes available to probe clustering at small scales at z < 1.5. New spectroscopic galaxy samples at these redshifts will allow precise measurements of structure growth and the full-shape of the power spectrum during the critical time when the Universe transitioned from a matter-dominated expansion to a dark energy-dominated expansion.

Number density, redshift range, spectroscopic completeness, and volume for a future spectroscopic program will need to be matched to analysis techniques and a systematic error budget that will steadily improve over the next decade. In the absence of a specific systematic error budget or well-accepted metrics to optimize area and number density for cross-correlation studies or higher order statistics, we refer to the product $\bar{n}P(k) = 1$ for characterizing the potential of future samples. Here, \bar{n} is the average number density of the sample and P(k) is the amplitude of the observed power spectrum at a scale k. This parameterization reflects a fairly optimal balance between area and number density for two-point statistics in a sample of a fixed size [for further discussion, see 47]. As shown

in the left panel of Figure 2, the DESI luminous red galaxy (LRG) and emission line galaxy (ELG) samples will achieve number densities that exceed the limit $\bar{n}P(k) = 1$ for scales $k_{\text{max}} = 0.20 h \text{ Mpc}^{-1}$.

The bright galaxy sample (BGS) will reach $\bar{n}P(k) = 1$ for scales $k_{\text{max}} = 1 h \text{ Mpc}^{-1}$ at redshifts z < 0.3. A scale of $k_{\text{max}} = 1 h \text{ Mpc}^{-1}$ corresponds roughly to the transition between sampling the halo population (two halo term) and sampling the galaxies within a halo (one halo term). It is this region where the nuisance terms resulting from connecting galaxies to halos are best constrained, thus breaking degeneracies with the cosmological information.

The DESI BGS sample has such a high number density in part to serve as a testbed for new cosmological analysis techniques. Future programs have the potential to expand these techniques to higher redshifts and larger volumes. As shown in the right panel of Figure 2, a program that achieves $\bar{n}P(k) = 1$ at $k_{\text{max}} = 1 h \text{ Mpc}^{-1}$ to a redshift z = 1 would require roughly 4000 galaxies deg⁻². A program that extends a clustering sample of this scope to z = 1.5 would require roughly 10,000 galaxies deg⁻². With samples of this size extending to higher redshift, the techniques developed on the BGS sample can eventually be applied over a redshift range that includes the matter dominated era and the transition to the accelerating epoch.

4.1 Science Drivers for Low-z clustering with future surveys

Current lensing measurements indicate tension with predictions from the CMB under a Λ CDM model (see, e.g., Ref. [48] for recent results). These tensions may be explained by an additional field or interaction that distorts the shape of the power spectrum, modifications to General Relativity, other extensions to the Standard Model, a chance statistical realization of the analyses, or systematic errors in the data analysis. Three-dimensional clustering data from DESI and lensing from Rubin observatory will allow us to improve the precision of these measurements. Advances in simulations will allow us to improve the theoretical modeling of these measurements. New spectroscopic data samples that sample clustering at the densities presented in Figure 2 over a wide redshift range can advance our understanding of cosmic acceleration as follows:

- 1. Growth of structure: probe origin of cosmic acceleration at late times
- 2. Growth of structure: constrain neutrino mass
- 3. Full-shape power spectrum measurements: provide limits on models for new physics that distort the shape of the power spectrum at scales $k < 1 h \,\mathrm{Mpc}^{-1}$
- 4. Calibrating samples of voids, clusters, peculiar velocities, and other techniques: perform comprehensive tests of the cosmologial model
- 5. Higher-order statistics: resolving degeneracies or identifying cosmological signatures not present in the power spectrum
- 6. Multiple probes: test for new particle or field contents in the cosmological model



Figure 3: Expected redshift distribution for the DESI LRG samples compared to various extended selection algorithms. Left: A broad extension to the DESI algorithm that excludes criteria intended to reduce the number of low redshift galaxies. Right: An extended selection that incorporates criteria intended to reduce the number of z < 0.3 galaxies. In both cases, the redshifts are from COSMOS photo-z's [49] for the $z_{\rm fiber}$ -limited samples and DESI spectroscopic redshifts for DESI LRGs. Pixel-level simulations indicate that the samples described by a selection to $z_{\rm fiber} < 21.6$ can be spectroscopically classified in the same exposure time as planned for the ongoing DESI program. Simulations indicate that galaxies with $21.6 < z_{\rm fiber} < 22.0$ require exposures twice as long while galaxies with $22.0 < z_{\rm fiber} < 22.4$ require exposures four times as long.

The techniques to perform these cosmological constraints will likely arise from a combination of three-dimensional clustering at small scales, higher order clustering statistics, and cross-correlations with lensing surveys.

4.2 Clustering Samples from DESI with Extended LRG Selections

The reddest, most luminous LRGs exhibit relatively high large-scale-structure bias, which enhances the amplitude of their clustering. Because of their strong 4000 Å break and their well-behaved, red spectral energy distributions, LRGs can be selected from imaging catalogs and spectrally classified with high efficiency. Pixel-level simulations indicate that more than 95% of these targets can be assigned reliable redshifts using a selection to $z_{\rm fiber} < 21.6$, where $z_{\rm fiber}$ is the z-band magnitude in an aperture matched to the DESI fiber diameter.

The currently planned DESI program will obtain an LRG spectroscopic sample of roughly six million galaxies at redshifts z < 1. Beyond this approximate redshift, the strong 4000 Å break and characteristic spectral features move into a region of the spectrum with strong sky line contamination, thus complicating redshift classification. As shown in Figure 3, the

DESI LRG sample will cover the redshift range 0.3 < z < 1 with a surface density of approximately 600 deg⁻². These observations will be obtained over five years, using a subset of the 50M fibers that will be available over all pointing centers.

In the latter half of this decade, DESI will remain the most competitive instrument for large-scale spectroscopic surveys. DESI will be able to maintain a high redshift efficiency for targets identified in extensions to the LRG selections that include bluer galaxies with lower bias. Potential selections have been explored using infrared and optical photometry, with simulations revealing consistently high redshift efficiency at the same exposure times planned for the current DESI survey. Two such selections are presented in orange in Figure 3. At 6600 deg⁻², the first selection (left panel) includes a high fraction of low redshift galaxies and very high completeness as a function of stellar mass. The second selection (right panel) is tuned for minimal redshift overlap with the BGS sample, leading to a density of 3000 deg⁻². This second selection exceeds a number density of $1 \times 10^{-3} h^3 Mpc^{-3}$ out to z < 0.9. As demonstrated in Figure 2, this number density would be sufficient to reach the threshold $\bar{n}P(k) = 1$ at $k_{\text{max}} = 1 h \text{ Mpc}^{-1}$. The galaxies in these two selections are sufficiently bright that spectra can be classified at high efficiency using the same exposure times as planned for the current LRG sample. At this exposure time, roughly 10M spectra can be obtained per year.

Longer exposure times could be used to include even fainter galaxies, as shown in the green and red curves of Figure 3. Extending the magnitude limit to $z_{\rm fiber} < 22.0$ (green) would require a factor of two increase in exposure time for the additional 1700 deg⁻² galaxies included in the selection. Doing so would extend the redshift coverage to z < 1 where the density allows sampling of clustering beyond the shotnoise limit. A factor of four increase in exposure time (red) would allow observations of a sample to $z_{\rm fiber} < 22.4$, thus vastly exceeding the required number densities over 0.3 < z < 1 and extending coverage to z < 1.1.

Continuing DESI observations in the area that overlaps the Vera Rubin Observatory survey footprint would allow joint studies of a high density spectroscopic sample and high signal-to-noise lensing sample. In Table 2, we demonstrate several potential survey strategies that can be completed with the DESI instrument in five years or less. With only an investment in DESI operations, such a survey would allow enhanced tests of dark energy, physics beyond the standard model, and neutrino masses using very high density probes of large-scale structure out to an approximate redshift of z = 1.

4.3 Future Clustering Samples at 1 < z < 1.5

As shown in Figure 2, roughly 6000 galaxies deg⁻² are required to reach a threshold of $\bar{n}P(k) = 1$ at $k_{\text{max}} = 1 h \text{ Mpc}^{-1}$ over the redshift range 1 < z < 1.5. At these redshifts, the absorption features that allow robust classification of passive, LRG spectra become harder to detect due to increased sky background and decreased continuum signal. For this reason, the DESI survey relies on the detection of [O II] line flux for galaxy redshift classification in this redshift range. At a wavelength of 3727 Å, the [O II] emission line that is often associated with star formation is visible with Silicon detectors to approximately z < 1.6. However, selecting these targets from imaging data to have both significant line

Snowmass2021 Cosmic Frontier White Paper Template

Survey Boundaries	Area	Survey Time (years)	Sample Size	
Selection at 6600 deg ⁻² without low redshift cut ($z_{\text{fiber}} < 21.6$)				
$0 < \delta < 12$	3372 deg^2	2 years	22M	
$-10 < \delta < 12$	6024 deg^2	4 years	40M	
Selection at 3000 deg ⁻² with low redshift cut ($z_{\text{fiber}} < 21.6$)				
$0 < \delta < 12$	3372 deg^2	1 year	10M	
$-10 < \delta < 12$	6024 deg^2	2 years	18M	
Selection at 4700 deg ⁻² with low redshift cut ($z_{\text{fiber}} < 22.0$)				
$0 < \delta < 12$	3372 deg^2	2 years	16M	
$-10 < \delta < 12$	6024 deg^2	4 years	28M	
Selection at 7100 deg ⁻² with low redshift cut ($z_{\text{fiber}} < 22.4$)				
$0 < \delta < 12$	3372 deg^2	5 years	24M	

Table 2: Potential survey footprints for four different selection algorithms using DESI as shown in Figure 3. The reported area corresponds to the overlap with Vera Rubin imaging in regions at galactic latitudes that avoid excessive Galactic extinction and stellar contamination.

strength and redshifts in the range 1.0 < z < 1.5 is far more challenging than the selection of LRG targets described in Section 4.2. As can be seen in the right panel of Figure 2, the DESI ELG selection produces only around 400 deg⁻² robust redshift classifications in this redshift range. These redshifts are obtained from a spectroscopic sample of roughly 1400 deg⁻², some of which are at lower redshift and some of which are not classified with high confidence. More than an order of magnitude increase in number density is required to reach the shot-noise limited densities over this redshift range.

If a selection algorithm with new imaging data is identified to select strong [O II] emitting galaxies over 1 < z < 1.5 with high efficiency, then it will be possible to improve clustering statistics with new DESI observations. However, it is most likely that new selections will experience similar contamination from low redshift galaxies and require longer exposure times to reach to fainter [O II] line fluxes. While such an investigation remains to be done, one can assume similar rates of redshift success. The [O II] luminosity distribution is also unknown, but for illustrative purposes, we assume that an order of magnitude increase in sample size will produce limiting [O II] line strengths that are between a factor of one and two smaller than in the DESI sample. Such a sample would require between 21,000 and 84,000 fiber-exposures deg⁻² at equivalent DESI exposure times. A pilot survey with DESI could complete between 120 and 500 deg² in one year under these assumptions.

Observations at this high density are more efficiently pursued with a future spectroscopic facility with a significantly higher survey speed. Three concepts for future spectroscopic programs were described in white papers submitted to the Astro2020 Decadal Survey: MegaMapper [50], Mauna Kea Spectroscopic Explorer (MSE) [51], and SpecTel [52]. Each of these offers the potential for even larger sample sizes than would be obtained with DESI. Both the proposed fiber positioner designs and the collecting areas are yet to be finalized for these three facilities. Accepting these uncertainties, we approximate the collecting area of the 6.5-meter MegaMapper telescope to be three times that of DESI and the collecting area of both the 12-meter class MSE and SpecTel telescopes to be ten times that of DESI. We assume a similar focal plane of 20,000 fibers for all three designs, and assume that sufficient targets are available to populate the fibers regardless of field of view. The survey areas available to each facility simply scales with the product of fiber number and collecting area.

Under these idealized assumptions, we find that in order to reach $\bar{n}P(k) = 1$ at $k_{\text{max}} = 1 h \text{ Mpc}^{-1}$ over the redshift range 1 < z < 1.5:

- MegaMapper could complete between 1,500 and 6,000 deg² per year;
- MSE could complete at least 5000 deg² per year; and
- SpecTel could complete at least 5000 deg² per year.

References

- [1] DESI Collaboration, Amir Aghamousa, et al. The DESI Experiment Part I: Science, Targeting, and Survey Design. *arXiv e-prints*, page arXiv:1611.00036, October 2016.
- [2] DESI Collaboration, Amir Aghamousa, et al. The DESI Experiment Part II: Instrument Design. *arXiv e-prints*, page arXiv:1611.00037, October 2016.
- [3] Steve Ritz et al. Building for Discovery: Strategic Plan for U.S. Particle Physics in the Global Context. 5 2014.
- [4] eBOSS Collaboration et al. Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Cosmological implications from two decades of spectroscopic surveys at the Apache Point Observatory. *Phys. Rev. D.*, 103(8):083533, April 2021.
- [5] Kyle S. Dawson et al. The Baryon Oscillation Spectroscopic Survey of SDSS-III. *AJ*, 145(1):10, January 2013.
- [6] Kyle S. Dawson et al. The SDSS-IV Extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data. *AJ*, 151(2):44, February 2016.
- [7] Planck Collaboration et al. Planck 2018 results. VI. Cosmological parameters. *A&A*, 641:A6, September 2020.
- [8] Željko Ivezić et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. *Astrophys. J.*, 873(2):111, March 2019.
- [9] Andreas Albrecht, Gary Bernstein, Robert Cahn, Wendy L. Freedman, Jacqueline Hewitt, Wayne Hu, John Huth, Marc Kamionkowski, Edward W. Kolb, Lloyd Knox, John C. Mather, Suzanne Staggs, and Nicholas B. Suntzeff. Report of the Dark Energy Task Force. *arXiv e-prints*, pages astro–ph/0609591, September 2006.

- [10] The LSST Dark Energy Science Collaboration, Rachel Mandelbaum, et al. The LSST Dark Energy Science Collaboration (DESC) Science Requirements Document. *arXiv e-prints*, page arXiv:1809.01669, September 2018.
- [11] A. R. Zentner, E. Semboloni, S. Dodelson, T. Eifler, E. Krause, and A. P. Hearin. Accounting for baryons in cosmological constraints from cosmic shear. *Phys. Rev. D*, 87(4):043509, February 2013.
- [12] Beth A. Reid, Hee-Jong Seo, Alexie Leauthaud, Jeremy L. Tinker, and Martin White. A 2.5 per cent measurement of the growth rate from small-scale redshift space clustering of SDSS-III CMASS galaxies. *MNRAS*, 444(1):476–502, October 2014.
- [13] E. Krause, T. F. Eifler, and the DES Collaboration. Dark Energy Survey Year 1 Results: Multi-Probe Methodology and Simulated Likelihood Analyses. arXiv:1706.09359, June 2017.
- [14] Andrés N. Salcedo, Benjamin D. Wibking, David H. Weinberg, Hao-Yi Wu, Douglas Ferrer, Daniel Eisenstein, and Philip Pinto. Cosmology with stacked cluster weak lensing and cluster-galaxy cross-correlations. *MNRAS*, 491(3):3061–3081, January 2020.
- [15] Benjamin D. Wibking, David H. Weinberg, Andrés N. Salcedo, Hao-Yi Wu, Sukhdeep Singh, Sergio Rodríguez-Torres, Lehman H. Garrison, and Daniel J. Eisenstein. Cosmology with galaxy-galaxy lensing on non-perturbative scales: emulation method and application to BOSS LOWZ. *MNRAS*, 492(2):2872–2896, February 2020.
- [16] Johannes U. Lange, Andrew P. Hearin, Alexie Leauthaud, Frank C. van den Bosch, Hong Guo, and Joseph DeRose. Five-percent measurements of the growth rate from simulation-based modelling of redshift-space clustering in BOSS LOWZ. *MNRAS*, *submitted*, page arXiv:2101.12261, January 2021.
- [17] Michael J. Chapman et al. The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: measurement of the growth rate of structure from the small-scale clustering of the luminous red galaxy sample. *arXiv:2106.14961*, page arXiv:2106.14961, June 2021.
- [18] BOSS Collaboration et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. MNRAS, 470(3):2617–2652, September 2017.
- [19] Authors et al. Snowmass2021 Cosmic Frontier White Paper: Cosmology from Combined Probes. Contribution to Snowmass 2021, 2022.
- [20] Kyle Dawson, Josh Frieman, Katrin Heitmann, Bhuvnesh Jain, Steve Kahn, Rachel Mandelbaum, Saul Perlmutter, and Anže Slosar. Cosmic Visions Dark Energy: Small Projects Portfolio. arXiv e-prints, page arXiv:1802.07216, February 2018.

- [21] P. Fosalba, M. Crocce, E. Gaztañaga, and F. J. Castander. The MICE grand challenge lightcone simulation - I. Dark matter clustering. *MNRAS*, 448(4):2987–3000, April 2015.
- [22] Joseph DeRose et al. The Buzzard Flock: Dark Energy Survey Synthetic Sky Catalogs. *arXiv e-prints*, page arXiv:1901.02401, January 2019.
- [23] The LSST Dark Energy Science Collaboration, Danila Korytov, et al. CosmoDC2: A Synthetic Sky Catalog for Dark Energy Science with LSST. *Astrophys. J.S.*, 245(2):26, December 2019.
- [24] LSST Dark Energy Science Collaboration (LSST DESC). The LSST DESC DC2 Simulated Sky Survey. *Astrophys. J.S.*, 253(1):31, March 2021.
- [25] Nina A. Maksimova, Lehman H. Garrison, Daniel J. Eisenstein, Boryana Hadzhiyska, Sownak Bose, and Thomas P. Satterthwaite. ABACUSSUMMIT: a massive set of highaccuracy, high-resolution N-body simulations. *MNRAS*, 508(3):4017–4037, December 2021.
- [26] Simon Birrer, Salman Habib, Katrin Heitmann, Zarija Lukić, , Yuuki Omori, and Yi-Min Zhong. Snowmass2021 Computational Frontier White Paper: Cosmological Simulations and Modeling. Contribution to Snowmass 2021, 2022.
- [27] Katrin Heitmann, Paul M. Ricker, Michael S. Warren, and Salman Habib. Robustness of Cosmological Simulations. I. Large-Scale Structure. *ApJS*, 160(1):28–58, September 2005.
- [28] Katrin Heitmann, Zarija Lukić, Patricia Fasel, Salman Habib, Michael S Warren, Martin White, James Ahrens, Lee Ankeny, Ryan Armstrong, Brian O'Shea, et al. The cosmic code comparison project. *Computational Science & Discovery*, 1(1):015003, 2008.
- [29] Katrin Heitmann, Martin White, Christian Wagner, Salman Habib, and David Higdon. The Coyote Universe. I. Precision Determination of the Nonlinear Matter Power Spectrum. *ApJ*, 715(1):104–121, May 2010.
- [30] Aurel Schneider, Romain Teyssier, Doug Potter, Joachim Stadel, Julian Onions, Darren S. Reed, Robert E. Smith, Volker Springel, Frazer R. Pearce, and Roman Scoccimarro. Matter power spectrum and the challenge of percent accuracy. *JCAP*, 2016(4):047, April 2016.
- [31] Lehman H. Garrison, Daniel J. Eisenstein, and Philip A. Pinto. A high-fidelity realization of the euclid code comparison n-body simulation with abacus. *MNRAS*, 485(3), 3 2019.
- [32] Cameron Grove et al. The DESI *N*-body Simulation Project I: Testing the Robustness of Simulations for the DESI Dark Time Survey. *arXiv e-prints*, page arXiv:2112.09138, December 2021.

- [33] Volker Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364(4):1105–1134, December 2005.
- [34] Matthieu Schaller, Pedro Gonnet, Aidan B. G. Chalk, and Peter W. Draper. SWIFT: Using task-based parallelism, fully asynchronous communication, and graph partitionbased domain decomposition for strong scaling on more than 100,000 cores. *arXiv e-prints*, page arXiv:1606.02738, June 2016.
- [35] Lehman H. Garrison, Daniel J. Eisenstein, Douglas Ferrer, Nina A. Maksimova, and Philip A. Pinto. The ABACUS cosmological N-body code. *MNRAS*, 508(1):575–596, November 2021.
- [36] Chia-Hsun Chuang et al. UNIT project: Universe N-body simulations for the Investigation of Theoretical models from galaxy surveys. *MNRAS*, 487(1):48–59, July 2019.
- [37] Katrin Heitmann, Nicholas Frontiere, Esteban Rangel, Patricia Larsen, Adrian Pope, Imran Sultan, Thomas Uram, Salman Habib, Hal Finkel, Danila Korytov, Eve Kovacs, Silvio Rizzi, Joe Insley, and Janet Y. K. Knowles. The Last Journey. I. An Extremescale Simulation on the Mira Supercomputer. *ApJS*, 252(2):19, February 2021.
- [38] Tomoaki Ishiyama, Francisco Prada, Anatoly A. Klypin, Manodeep Sinha, R. Benton Metcalf, Eric Jullo, Bruno Altieri, Sofía A. Cora, Darren Croton, Sylvain de la Torre, David E. Millán-Calero, Taira Oogi, José Ruedas, and Cristian A. Vega-Martínez. The Uchuu simulations: Data Release 1 and dark matter halo concentrations. *MNRAS*, 506(3):4210–4231, September 2021.
- [39] Sownak Bose, Daniel J. Eisenstein, Boryana Hadzhiyska, Lehman H. Garrison, and Sihan Yuan. Constructing high-fidelity halo merger trees in AbacusSummit. *arXiv:2110.11409*, page arXiv:2110.11409, October 2021.
- [40] P. Behroozi, R. Wechsler, A. Hearin, and C. Conroy. UniverseMachine: The Correlation between Galaxy Growth and Dark Matter Halo Assembly from z=0-10. arXiv:1806.07893, June 2018.
- [41] B. P. Moster, T. Naab, and S. D. M. White. EMERGE An empirical model for the formation of galaxies since $z \sim 10$. *ArXiv:1705.05373*, May 2017.
- [42] Nick Battaglia, Andrew Benson, Tim Eifler, Andrew Hearin, Katrin Heitmann, Shirley Ho, Alina Kiessling, Zarija Lukic, Michael Schneider, Elena Sellentin, and Joachim Stadel. Report from the Tri-Agency Cosmological Simulation Task Force. arXiv eprints, page arXiv:2005.07281, May 2020.
- [43] Zheng Zheng and David H. Weinberg. Breaking the Degeneracies between Cosmology and Galaxy Bias. *ApJ*, 659(1):1–28, April 2007.
- [44] Rongpu Zhou et al. The clustering of DESI-like luminous red galaxies using photometric redshifts. *MNRAS*, 501(3):3309–3331, March 2021.

- [45] Giulia Despali, Carlo Giocoli, Raul E. Angulo, Giuseppe Tormen, Ravi K. Sheth, Giacomo Baso, and Lauro Moscardini. The universality of the virial halo mass function and models for non-universality of other halo definitions. *MNRAS*, 456(3):2486– 2504, March 2016.
- [46] Jeremy L. Tinker, Brant E. Robertson, Andrey V. Kravtsov, Anatoly Klypin, Michael S. Warren, Gustavo Yepes, and Stefan Gottlöber. The Large-scale Bias of Dark Matter Halos: Numerical Calibration and Model Tests. *ApJ*, 724(2):878–886, December 2010.
- [47] Andreu Font-Ribera, Patrick McDonald, Nick Mostek, Beth A. Reid, Hee-Jong Seo, and Anže Slosar. DESI and other Dark Energy experiments in the era of neutrino mass measurements. *JCAP*, 2014(5):023, May 2014.
- [48] Catherine Heymans et al. KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints. *A&A*, 646:A140, February 2021.
- [49] C. Laigle et al. The COSMOS2015 Catalog: Exploring the 1. *The Astrophysical Journal Supplement Series*, 224(2):24, June 2016.
- [50] David Schlegel, Juna A. Kollmeier, and Simone Ferraro. The MegaMapper: a z¿2 spectroscopic instrument for the study of Inflation and Dark Energy. In *Bulletin of the American Astronomical Society*, volume 51, page 229, September 2019.
- [51] Jennifer Marshall et al. The Maunakea Spectroscopic Explorer. In *Bulletin of the American Astronomical Society*, volume 51, page 126, September 2019.
- [52] Richard Ellis and Kyle Dawson. SpecTel: A 10-12 meter class Spectroscopic Survey Telescope. In *Bulletin of the American Astronomical Society*, volume 51, page 45, September 2019.