

Fermilab

Forward modeling fluctuations in the DESI LRGs target sample using image simulations

FERMILAB-PUB-24-0755-PPD

arXiv:2405.16299

This manuscript has been authored by Fermi Research Alliance, LLC
under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy,
Office of Science, Office of High Energy Physics.

Forward modeling fluctuations in the DESI LRGs target sample using image simulations

H. Kong¹, A. J. Ross^{2,3,4}, K. Honscheid^{2,3,4}, D. Lang^{5,6},
 A. Porredon⁷, A. de Mattia⁸, M. Rezaie⁹, R. Zhou¹⁰,
 E. F. Schlafly¹⁰, J. Moustakas¹¹, A. Rosado-Marin¹²,
 J. Aguilar¹⁰, S. Ahlen¹³, D. Brooks¹⁴, E. Chaussidon¹⁰,
 T. Claybaugh¹⁰, S. Cole¹⁵, A. de la Macorra¹⁶, Arjun Dey¹⁷,
 Biprateep Dey¹⁸, P. Doel¹⁴, K. Fanning^{19,20},
 J. E. Forero-Romero^{21,22}, E. Gaztañaga^{23,24,25}, S. Gontcho A
 Gontcho¹⁰, G. Gutierrez²⁶, C. Howlett²⁷, S. Juneau¹⁷,
 A. Kremin¹⁰, M. Landriau¹⁰, M. E. Levi¹⁰, M. Manera^{1,28},
 P. Martini^{2,3,4}, A. Meisner¹⁷, R. Miquel^{1,28}, E. Mueller²⁹,
 A. D. Myers³⁰, J. A. Newman¹⁸, J. Nie³¹, G. Niz^{32,33},
 W. J. Percival^{5,6,34}, C. Poppett^{10,35,36}, F. Prada³⁷, G. Rossi³⁸,
 E. Sanchez³⁹, D. Schlegel¹⁰, M. Schubnell^{40,41}, H. Seo¹²,
 D. Sprayberry¹⁷, G. Tarlé^{40,41}, M. Vargas-Magaña¹⁶,
 B. A. Weaver¹⁷, H. Zou³¹

¹Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra Barcelona, Spain

²Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA

³Department of Astronomy, The Ohio State University, 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH

⁴The Ohio State University, Columbus, 43210 OH, USA

⁵Perimeter Institute for Theoretical Physics, 31 Caroline St N, Waterloo, ON N2L 2Y5, Canada

⁶Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada

⁷Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute, German Centre for Cosmological Lensing, 44780 Bochum, Germany

⁸IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

⁹Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, USA

¹⁰Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

- ¹¹Department of Physics & Astronomy, Siena College, 515 Loudon Road, Loudonville, NY, USA 12211
- ¹²Department of Physics & Astronomy, Ohio University, Athens, OH 45701, USA
- ¹³Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA
- ¹⁴Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
- ¹⁵Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
- ¹⁶Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México
- ¹⁷NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, USA
- ¹⁸Department of Physics & Astronomy and Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA
- ¹⁹Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, Menlo Park, CA 94305, USA
- ²⁰SLAC National Accelerator Laboratory, Menlo Park, CA 94305, USA
- ²¹Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia
- ²²Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia
- ²³Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
- ²⁴Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK
- ²⁵Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain
- ²⁶Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA
- ²⁷School of Mathematics and Physics, University of Queensland, 4072, Australia
- ²⁸Departament de Física, Serra Hünter, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain
- ²⁹Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, U.K
- ³⁰Department of Physics & Astronomy, University of Wyoming, 1000 E. University, Dept. 3905, Laramie, WY 82071, USA
- ³¹National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Rd., Chaoyang District, Beijing, 100012, P.R. China
- ³²Departamento de Física, Universidad de Guanajuato - DCI, C.P. 37150, Leon, Guanajuato, México
- ³³Instituto Avanzado de Cosmología A. C., San Marcos 11 - Atenas 202. Magdalena Contreras, 10720. Ciudad de México, México
- ³⁴Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada
- ³⁵Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA
- ³⁶University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA
- ³⁷Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain
- ³⁸Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea

³⁹CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain

⁴⁰Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA

⁴¹University of Michigan, Ann Arbor, MI 48109, USA

E-mail: hkong@ifae.es

Abstract. We use the forward modeling pipeline, *Obiwan*, to study the imaging systematics of the Luminous Red Galaxies (LRGs) targeted by the Dark Energy Spectroscopic Instrument (DESI). Imaging systematics refers to the false fluctuation of galaxy densities due to varying observing conditions and astrophysical foregrounds corresponding to the imaging surveys from which DESI LRG target galaxies are selected. We update the *Obiwan* pipeline, which we previously developed to simulate the optical images used to target DESI data, to further simulate WISE images in the infrared. This addition allows simulating the DESI LRGs sample, which utilizes WISE data in the target selection. Deep DESI imaging data combined with a method to account for biases in their shapes is used to define a truth sample of potential LRG targets. We inject these data evenly throughout the DESI Legacy Imaging Survey footprint at declinations between -30 and 32.375 degrees. We simulate a total of 15 million galaxies to obtain a simulated LRG sample (*Obiwan* LRGs) that predicts the variations in target density due to imaging properties. We find that the simulations predict the trends with depth observed in the data, including how they depend on the intrinsic brightness of the galaxies. We observe that faint LRGs are the main contributing source of the imaging systematics trend induced by depth. We also find significant trends in the data against Galactic extinction that are not predicted by *Obiwan*. These trends depend strongly on the particular map of Galactic extinction chosen to test against, implying systematic contamination in the Galactic extinction maps is a likely root cause (e.g., Cosmic-Infrared Background, dust temperature correction). We additionally observe a morphological change of the DESI LRGs population evidenced by a correlation between OII emission line average intensity and the size of the z -band PSF. This effect most likely results from uncertainties in background subtraction. The detailed findings we present should be used to guide any observational systematics mitigation treatment for the clustering of the DESI LRGs sample.

Contents

1	Introduction	1
2	Data	2
2.1	DESI Legacy Imaging Survey DR9	2
2.2	COSMOS Repeats dataset	3
2.3	COSMOS Deep Catalogs	5
3	The Obiwan Pipeline	5
3.1	Optical Bands	6
3.2	WISE bands	6
4	LRG-like target selection	7
5	Pre-processing	9
5.1	Truth Data Generation	9
5.2	Source injection strategy	9
6	Obiwan with DR9 images	9
6.1	Morphology test	10
6.2	Imaging Systematics	12
7	Testing the Obiwan procedure	13
7.1	Correlation With Galaxy Brightness	16
7.2	Galactic Extinction	19
7.3	Investigation of Trends With <code>psfsize</code>	24
7.3.1	Possible root causes	24
7.4	Other secondary effects	26
7.5	Discussion on imaging systematics mitigation	27
7.6	Implications for cosmological analysis	28
8	Conclusions	28
A	Computational Resources	37
A.1	Computational resources for ‘blobs’	37
A.2	Cosmos Deep	37
A.3	Obiwan runs	37
A.4	Methods for Increasing the Efficiency of Synthetic Galaxy Production	37
B	Extra Details on Cosmos Deep	38
B.1	cut on z -band depth	38
B.2	The WISE band in COSMOS Deep catalog	38
C	De-biasing procedure for Cosmos Deep catalog	39
D	Measurement bias	41

E	Testing impact of psfsize with COSMOS repeats	42
E.1	Origination of trend with psfsize	42
E.2	Morphological change of LRGs with psfsize	47

1 Introduction

Modern cosmology is built upon detecting signals from the distant universe. We collect light and gravity waves through various instruments at different frequencies [1–5]. For galaxy imaging surveys [6–8], in particular, we collect light profiles of galaxies by taking pictures of the night sky. These images help us infer the properties of galaxies, which can be used as fundamental input for model development in Cosmology. However, as light travels through the Milky Way and the Earth’s atmosphere to our camera, it is perturbed by different effects. These effects alter the light profile, making the galaxies in the images deviate from their true appearance. We call these effects imaging systematics [9–12]. It is an umbrella term containing any possible effects contributing to galaxy densities’ false fluctuation. In practice, we use survey property maps to quantify the strength of different effects on the galaxy densities. These maps include astronomical effects like Galactic dust extinction and stellar density (including stellar streams) and instrument-based effects like image background noise level and seeing.

The correlation between these property maps and the galaxy density indicates how each survey property contaminates our galaxy samples.

An accurate correction of imaging systematics lays a solid foundation for a stringent and unbiased measurement for all cosmological probes. It is essential for probes that are sensitive to the power spectrum. For example, primordial non-Gaussianity [13–15] is sensitive to the large-scale part of the power spectrum. Redshift Space distortion measurements [16] and full shape analysis [17, 18] are sensitive to the overall shape. If not treated properly, imaging systematics would leave an imprint on the galaxy density map that would be degenerate with the primordial non-Gaussianity signal. On the other hand, traditional probes like Baryon Acoustic Oscillations [19] have a characteristic bump when computing their correlation function, thus are less sensitive to imaging systematics; it is possible to obtain a robust measurement even without any treatment of imaging systematics.

As ongoing and future galaxy surveys yield increasingly precise data, the margin for error in imaging systematics needs to diminish to achieve greater constraining power. The Dark Energy Spectroscopic Survey (DESI) [20–28] is a Stage-4 [29] cosmology survey that aims at obtaining better cosmological constraints than the former surveys [30, 31]. DESI is conducted on the Mayall 4-meter telescope at Kitt Peak National Observatory. It is designed to build an unprecedented 3D map of the universe. It will obtain 47 million galaxy and star spectra, and is expected to achieve a precision of less than 0.4% precision measurement of the BAO distance scale, and 1.05% precision on the Hubble parameter [32].

Imaging systematics arises from varying observing conditions in imaging surveys. In the case of DESI, the galaxies are selected from the 9th data release (DR9) of the DESI Legacy Imaging Surveys [33–35]. The survey is comprised of 3 imaging projects on different telescopes: The Beijing-Arizona Sky Survey (BASS) [36], the DECam Legacy Survey (DECaLS) [37], and the Mayall z-band Legacy Survey (MzLS) [38]. These images are processed with the DESI image reduction pipeline LEGACYPipe¹. LEGACYPipe transforms these images into a catalog of astronomical objects.

We developed a forward modeling pipeline, Obiwan², to study the effects of imaging systematics. The idea is to add simulated galaxies into real images and extract these simulated galaxies with the

¹<https://github.com/legacysurvey/legacypipe>

²https://github.com/DriftingPig/obiwan_code/tree/ObiwanLrgPaper

image reduction pipeline `LEGACYPIPE` [39–41]. The pixel-to-object transformation is performed in the same way as real DESI galaxies. `Obiwan` traces the un-cleaned residuals in the image and provides a catalog with the same imaging systematics variation as the tracers we are interested in. `Obiwan` was first utilized to study the imaging systematics of eBOSS ELGs [42, 43], and it is further developed to fit our purpose here.

In this work, we use `Obiwan` to study the imaging systematics trend of DESI LRGs[44]: The bright and luminous galaxies selected from DESI Legacy Imaging Surveys. They will be targeted to get spectrum throughout the 5-year DESI main survey. We recover similar systematics trends on simulated LRGs that we call `Obiwan` LRGs. We also identify spurious trends that are possibly due to imperfect survey property maps or uncertainties in the image calibration stage. The results we present allow a deeper understanding of the sources of artificial variation in the density of DESI LRGs, which we expect will influence how this variation is corrected for in the ultimate clustering analyses of DESI LRGs.

This paper is organized as follows. In Section 2, we describe the data we used in our analysis. In Section 3 we describe the pipeline, as well as new updates. Section 4 describes our method to determine the target selection cut. In Section 5, we discuss the process needed before the simulation runs. In Section 6, we discuss the analysis of output products. In Section 7, we discuss the indication of our results.

2 Data

We use three datasets for our work. These datasets originate from the DR9 DESI Legacy Imaging Survey images, and only the image selection criteria are different. The images are compiled from 4 cameras: The Dark Energy Camera (DECam, [45]) on the Victor M. Blanco 4-meter Telescope at the Cerro Tololo Inter-American Observatory (CTIO); the MOSAIC-3 camera [46], formerly mounted on the Nicholas U. Mayall 4-meter Telescope at the Kitt Peak National Observatory; the 90Prime camera [47] on the Bok 2.3-meter telescope located on Kitt Peak; and the Wide-field Infrared Survey Explorer (WISE, [48]) satellite.

The three datasets described below are composed of these images and are processed with the pipeline `LEGACYPIPE`. The DESI Legacy Imaging Survey DR9 described in Section 2.1 uses most of the images, excluding a few defective images. Images were also excluded because the selected region had reached the depth requirement. This set is used to mass produce `Obiwan` LRGs to study imaging systematics. The COSMOS Deep catalog in Section 2.3 utilizes the regions of the sky with many more exposures than a typical region and has much deeper photometry. This set is used to select the truth sample that `Obiwan` injects into the images. The COSMOS Repeats in Section 2.2 uses a large number of observations of the COSMOS field. The images are separated into ten distinct sets, each individually processed. These repeats are used to understand photometric scattering on DESI LRGs to define extended color cut for the injected `Obiwan` LRG-like sample.

2.1 DESI Legacy Imaging Survey DR9

We use and simulate data from the DR9 DESI Legacy Imaging Survey³ [33], which was used to obtain the sample that DESI uses for target selection and follow-up spectroscopy [49–56]. We briefly describe these images here. This imaging survey consists of three optical bands: g , r , and z , and 4 infrared bands from WISE, $W1$, $W2$, $W3$, and $W4$. The optical band is separated into two parts, the South and the North, by a declination (Dec) of 32.375 in the North Galactic Cap. The South consists

³<https://www.legacysurvey.org/dr9/description/>

of exposures taken from DECam [57], with a typical seeing of less than 1.3". The North part consists of images from the Beijing-Arizona Sky Survey [58] using the 90Prime camera for the g and r bands, and the Mosaic z -band Legacy Survey (MzLS) [59] for the z -band. The BASS survey has a typical seeing of less than 1.7", and MzLS has a less than 1.3" seeing. The pipeline transforms these images into a common grid with a pixel scale of 0.262 arcsec. These pixels are grouped into 3600×3600 ‘bricks’, and image processing is performed at the individual brick level.

The infrared bands are imaged by the WISE survey [60]. WISE is a space telescope whose orbits intersect with its Ecliptic poles ($|\text{Dec}| \approx 66$) while taking snapshots of the sky during its journey. Because of this orbiting plan, the WISE images have stable orientation in regions far from the Ecliptic poles. Image orientation can be very different during different visits for regions close to the Ecliptic poles. Because of this difference, it is much easier to measure regions far from the Ecliptic poles: The point spread function (PSF) of WISE galaxies is a stable image, and its profile is determined by the orientation of the WISE telescope [61]. Our work focuses only on the Southern part of the imaging survey because we do not have a truth input sample in the northern imaging area, and more details are explained in Section 2.3. Though WISE is deeper near the Ecliptic poles, these regions are more challenging to model due to the different image orientations in different visits; however, we avoid these problematic regions since our footprint does not cover that region. The WISE survey has four bandpasses, but only two bands ($W1$ and $W2$) are used in DESI target selection, so we will focus on only $W1$ and $W2$ in what follows. These two bands have a typical point-spread function width of 6.5". The WISE images we process are the unWISE coadded data products [62, 63], which use a 2.75" pixel scale. Figure 1 is a representative source with typical PSF size in all bands. It is located in a region observed by all three cameras in the optical bands. The first column shows the real images in optical and infrared bands. The second column shows galaxy models of the real images. The third column shows the galaxy stamp for the simulated galaxy. The fourth column is the “blob map”: sources within this map are treated as blended sources and will be fit simultaneously.

2.2 COSMOS Repeats dataset

The COSMOS region has been imaged extensively by many instruments, including DECam. The COSMOS Repeats dataset we use is a set of catalogs covering 1 deg^2 around $(\text{RA}, \text{Dec}) = (150.1 \text{ deg}, 2.2 \text{ deg})$. This small region has a large number of exposures taken in all optical bands, which provides the ability to produce many independent subsets of data processed through LEGACYPipe and containing the same photometric sources. Ten such realizations were produced for the COSMOS repeats data. Gaussian noise was added to each image in each set so that the total `galdepth`⁴ in each set was equal to the DESI imaging requirements. The different image sets are grouped by seeing. The ten sets are labeled “set 0” through “set 9”, going from smallest to largest seeing (`psfsize`). The scatter in photometric measurements of sources common to each set of the COSMOS Repeats data thus provides a fully empirical determination of the expected variance in LEGACYPipe photometry in different seeing conditions. We simulate sources in the same sets of images with the same settings as COSMOS repeats catalog. We obtained ten sets of corresponding simulated Obiwan COSMOS repeat catalogs that mimic the scattering behavior of sources in real COSMOS repeat catalogs. Achieving consistency with the distributions recovered from the real data in the COSMOS Repeats is our main method of validating the LRG-Obiwan pipeline. The COSMOS Repeats data will also be used to determine our target selection for the extended color boundary that we draw our truth sample from and will be discussed in Section 4. It will also be used to study PSF error contributions in Section 7.3.

⁴It is defined as detection sensitivity of a round exponential galaxy with a radius of 0.45 arcsecond.

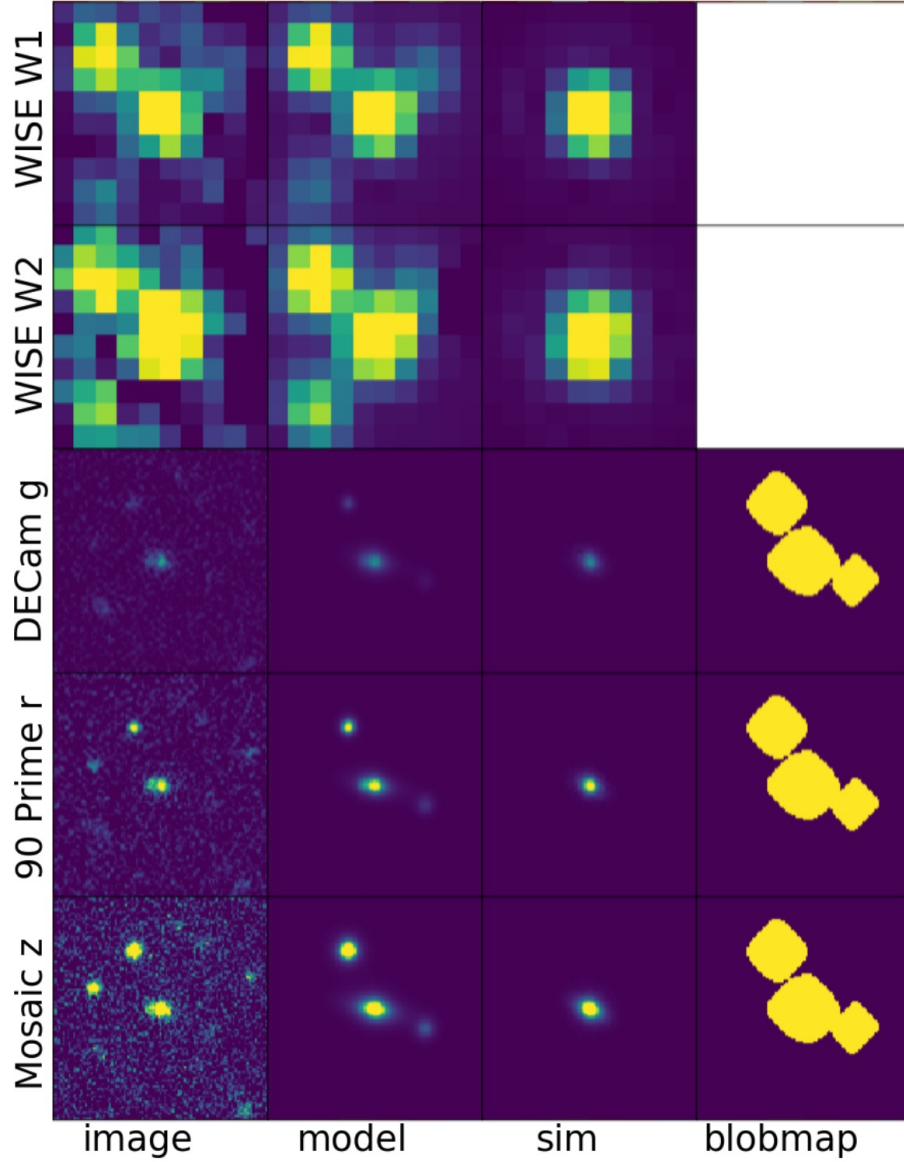


Figure 1: Simulated targets in different bands and cameras, processed through the LEGACYPIPE image reduction pipeline. The injected galaxy is located at the center of each image. The figures in the first column are the co-added images. The second column displays the galaxy/star profiles derived from fitting the image. The third column shows the simulated source injected into the images after convolution with the local PSF. The fourth column shows the contiguous area (known as a ‘blob’), where optical sources within the same blob are fit simultaneously (e.g., to resolve blending). The blobs are not used for WISE processing (see text for details) and, thus, are not shown for their images. The top rows show the WISE W1 and W2 band images, respectively. The third row shows the DECam image in the g - band, and the fourth row shows the r - band image from the BASS survey. The third row shows the z - band image from the MzLS survey.

2.3 COSMOS Deep Catalogs

The COSMOS Deep Catalogs are image products derived from DECam observations in the COSMOS region, processed with the same LEGACYPIPE procedures as the main DR9 data. The only difference is that the main survey catalog has a maximum depth limit. The depth is closely related to the total number of individual exposures in one processing unit chosen for measurement. Thus, standard Legacy Surveys DR9 release uses far fewer individual images in the COSMOS regions than have been observed by DECam. The Cosmos Deep catalog uses all individual images available and reaches a much higher depth in all optical bands. The higher depth enables the COSMOS deep catalog to reach a better precision in flux and shape measurement, and more detection of faint sources. Details for computational resources used are in Appendix A.2. In addition, we perform a cut of $\text{galdepth_z} > 24.5$ to ensure that the catalog is deep enough in z -band. Details about this cut are described in Appendix B.1.

The high depth in optical bands also affects the WISE band: the WISE band images remain the same as in the DR9 release, with a depth much lower than the optical band. The sources in the WISE images are measured in a process called “forced photometry”: We take the shape and location measured in the optical band, and fit these sources in the WISE images to obtain their flux values in the WISE bands. A higher depth in the optical band has pros and cons for understanding the WISE measurements. It helps improve the accuracy of WISE band fitting by providing more precise information about the galaxies’ location and shape. However, since more faint sources are detected in the optical band, the sources in the WISE images are more likely to be blended with other sources. In the presence of blending, the flux on one source could be falsely assigned to a nearby source due to a slight deviation in the shape, PSF, and astrometric position input values. We further test flux accuracy in the WISE band and describe the results in Appendix B.2. We find no significant discrepancy in the COSMOS deep catalog for the WISE flux, so we use it as our ‘truth catalog’.

As the COSMOS Deep data set is imaged with the DECam telescope, the filter properties differ from those used in the BASS/Mzls surveys. Therefore, the COSMOS Deep Catalog is not a representative truth sample for the BASS/MzLS region. Although it is potentially possible to map the truth catalog to the BASS/MzLS region with a filter-transmission function similar to what is described in equation 1-6 of [33], we were unable to perform a robust validation in limited tests. Thus, we focus on the DECaLS data in this study and leave simulating BASS/MzLS for future work. However, given the similarities in the sample selection between the North and South imaging regions, we expect that any conclusions on the properties that drive fluctuations in target density based on our analysis of DECaLS data would hold for BASS/MzLS data.

When sampling from the Deep catalogs, we de-bias the galaxy shapes (this will be discussed in Section 5.1). We also apply the LRG-like cut (described in Section 4), which reduces the sample to only galaxies with a reasonable likelihood of being selected as a DESI target. The final product yields 23,892 sources, with 4,986 that pass the DESI LRG target selection.

3 The Obiwan Pipeline

The synthetic source injection pipeline used in this work is Obiwan. It injects galaxies into the real images, and extracts them using the same procedure applied to real galaxies in the Legacy Surveys DR9 data. This pipeline can be divided into two stages: In the first stage, it injects sources and processes images on optical bands, obtaining the information on flux, shape, etc, for each source. In the second stage, it injects and processes images on the WISE bands. The location and shape of each source, obtained in the first stage, is used in this stage as input. The pipeline fits WISE images to get the WISE flux for each source. Figure 2 shows the workflow of this pipeline.

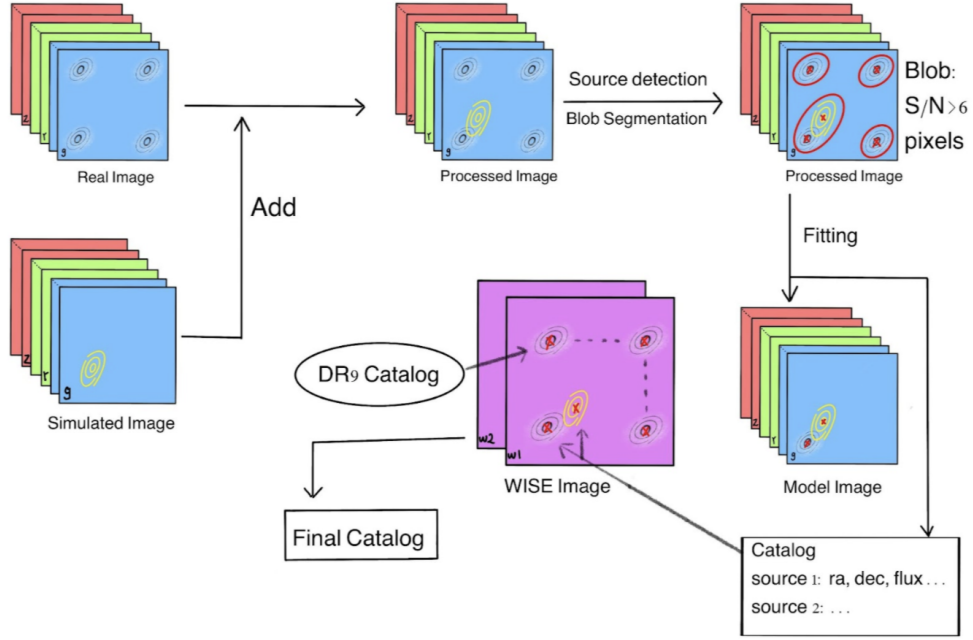


Figure 2: Workflow of Obiwan. The blue, green, red, and purple panel represents the g , r , z , $W1/W2$ band respectively. The black and white pattern represents the real galaxies, and the yellow pattern represents the simulated galaxies. The workflow shows how the simulated galaxies are added to real images on a per-exposure basis. Only patches with simulated galaxies are fitted. For the fitting of the WISE image, we also need extra information on galaxies from the DR9 data release.

3.1 Optical Bands

Obiwan works by adding simulated sources to real images. The sources are added on a per-exposure basis, at the level of each CCD detector. The RA, Dec location of the injected source is converted into pixel detector units, the added source is convolved with the PSF at that location, and Poisson noise is applied afterward. In the LEGACYPIPE processing, these individual exposures are transformed into common brick coordinates. Then, the pipeline proceeds to the source detection stage, where sources are detected and segmented into “blobs”. A “blob” is defined as a set of contiguous pixels with signal-to-noise greater than 6 (when including all images in the optical bands), plus an 8-pixel margin. The blob is defined to be the same region in brick pixel space across all of the optical bands. Each blob will contain at least one source and the photometry is fit independently in each blob. A key aspect that makes Obiwan more efficient is that in the source fitting step, we only fit blobs that overlap the true input RA,Dec of our simulated sources. This helps reduce processing time, and each iteration takes 75% less time than fitting over the full image.

3.2 WISE bands

Compared to processing a whole image, Obiwan uses much less processing time in the optical band, thanks to only processing a small fraction of sources within the “blobs” that have injected galaxies in them. To preserve this structure while computing WISE flux values for Obiwan sources, we need some extra operations on the part of the pipeline that computes WISE flux for galaxies.

When producing the DR9 catalog, all blobs were processed in optical bands. The WISE band fitting takes all of the source information from optical bands, and uses the information to perform “forced photometry”. This means that the source location and shape parameters are held fixed and the WISE flux is the only free parameter. “Forced photometry” simultaneously fits all sources detected in optical bands over the entire brick of WISE images.

In our Obiwan simulation run, we do not have the optical-band information for all sources. Due to the computationally efficient structure applied in optical bands, we only have the sources within the same blob as simulated sources; sources outside the selected blobs are not processed in previous steps; thus we do not know the details of the sources’ shapes and locations in these non-processed blobs. However, these sources are processed and recorded in the DR9 catalogs. These non-processed blobs have conditions identical to those of the sources processed in DR9, so their properties are identical to those recorded in DR9. As a result, we can take these sources’ entries in DR9 as input for the WISE flux forced photometry measurement.

In addition to these sources in blobs that are unaffected by Obiwan injected galaxies, we must also include bright stars from outside the brick, because bright enough stars will affect the WISE photometry. We look up and include nearby bright stars in the WISE fitting. We have validated that the WISE flux derived from this method is identical to a simulation run that processes all blobs in the optical bands.

4 LRG-like target selection

We define a new LRG-like cut used to perform target selection on the sample from the COSMOS Deep Catalog. The selected sample with this LRG-like cut served as the truth sample injected into DESI Legacy Imaging Survey DR9 images.

The DESI LRGs sample studied in this work is the LRG SV3 sample described in Table 4 of [44]. This selection was used during DESI’s Survey Validation phase and is similar to the main LRG target selection, with slightly extended cuts to include more high redshift LRGs. Our work can thus be trivially modified to study LRGs within the main selection cut.

The sources near selection boundaries can be scattered inside or outside of the LRG target selection boundary. To fully simulate the scattering feature of LRGs, we need to define a LRG-like color cut, which spans a larger color selection box than the LRG SV3 selection. At the same time, it also needs to be small enough not to spend too much computing time obtaining the photometry of sources that are unlikely to have measured photometry that matches the LRG SV3 selection. We describe our procedure to generate such a color selection box as follows.

We collect all targets that pass the LRG SV3 selection cuts in the 10 COSMOS Repeats sets. We record the position of all sources, and match these positions to all the other COSMOS Repeats sets. This new expanded sample is denoted the Potential-LRGs.

Due to photometric scattering, we need to develop a selection scheme that contains LRGs both inside and outside the LRG SV3 color/magnitude selection. We denote this selection scheme LRG-like color cut. We use Potential-LRGs to quantify the fraction of LRGs inside the LRG-like cut: For a given LRG-like cut, we define **scatter rate** as the amount of Potential-LRGs outside LRG-like cut, divided by the total number of Potential-LRGs. We also define **contamination rate** as the number of sources inside this color cut, divided by the number of Potential-LRGs inside the LRG-like cut. Our goal for the LRG-like cut is to make the selected sample have a low scatter and contamination rates.

The following process describes our method to develop the LRG-like cut: We start from the LRG SV3 selection cut, which is composed of 4 equations described in equation 4.1–4.4 excluding bold numbers. We iteratively extend each selection boundary from equations 4.1–4.4 by 0.01 magnitudes,

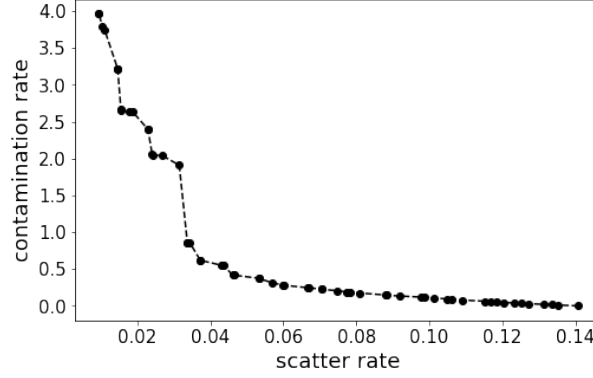


Figure 3: The change in scatter and contamination rates as we extend the color selection boundaries. Each point represents a set of color cuts that define the potential LRG-like sample. From the right side to the left side of the plot, we go to a more extended color cut. It would include more potential LRGs, as well as other non-LRGs. We eventually chose the left-most point as our final color cut. This color cut is wide enough to fully mimic the scattering of galaxies outside and inside the LRG SV3 selection boundary while keeping the ratio of input LRGs in Obiwan simulations high enough.

and compute the scatter rate for each equation. We choose the selection cut that corresponds to the lowest scatter rate, and update this selection cut to the new boundary (i.e., increased by 0.01 magnitudes). This would be a new data point in Figure 3. We perform this process iteratively, widening the selection cuts little by little. Each step is a new realization of a color cut, corresponding to different scatter and contamination rates. Figure 3 shows this variation: Each point represents one potential LRG-like color cut. As we extend the selection boundaries, we have a lower scatter rate and a greater contamination rate.

We select the color cut with the lowest scatter rate, as ideally, our simulation would include all sources potentially selected as LRGs. Our final selection corresponds to a scatter rate of 0.009, and a contamination rate of 4.03. These numbers imply that we will include more than 99% of potential LRG sources and that 20% of the injected sources will ultimately be selected as LRGs. Follow-up of less than 1%

Our final LRG-like target selection cut is:

$$z - w1 > 0.8 * (r - z) - 0.6 + \mathbf{0.00} \quad (4.1)$$

$$z_{fiber} < 21.7 + \mathbf{0.32} \quad (4.2)$$

$$\begin{aligned} & [(g - r > 1.3 - \mathbf{0.45}) \text{ AND} \\ & (g - r > -1.55 * (r - w1) + 3.13 - \mathbf{0.13})] \text{ OR} \\ & (r - w1 > 1.8 - \mathbf{0.37}) \end{aligned} \quad (4.3)$$

$$\begin{aligned} & [(r - w1 > (w1 - 17.26) * 1.8 - \mathbf{0.46}) \text{ AND} \\ & (r - w1 > (w1 - 16.36) * 1 - \mathbf{0.88})] \text{ OR} \\ & (r - w1 > 3.29 - \mathbf{1.65}) \end{aligned} \quad (4.4)$$

The bold number shows the number of boundaries moved in each color cut. The non-bold part is the original LRG-SV3 selection cut. Eventually, 23892 sources are selected as truth input.

5 Pre-processing

5.1 Truth Data Generation

We use the COSMOS Deep dataset to form our input truth set. We apply the LRG-like color cut described in Section 4, and form a LRG-like truth set with clean photometry measured from the Deep data. However, the shape information measured from the COSMOS Deep dataset is not reliable enough for our purposes. This is because the shape measurement in LEGACYPIPE is biased. LEGACYPIPE treats source fitting as a minimization problem, including terms for the chi-squared of the fit, a prior on the ellipticity, and a prior for the model complexity.

$$\begin{aligned} \text{Minimize } W = \sum & [(\text{image} - \text{model})^2 \times \text{inverse_variance}] \\ & + \text{ellipticity_prior}(e_1, e_2) \\ & + \text{parameters_prior}(N_{\text{parameters}}) \end{aligned} \quad (5.1)$$

The first term is summed over pixels of a galaxy stamp. We have pixelized data for the observed image (`image`), the modeled image (`model`), and a map that indicates the variance for each pixel (`inverse_variance`). The second term is a prior that is a function of galaxy ellipticity. The third term increases with the complexity of the galaxy model, and is used to determine the favorable model for a given galaxy. The prior term imposed by ellipticity and the number of parameters is independent of the depth of the image. This means that at higher depth, this prior term becomes less significant since the `inverse_variance` term increases and the first term gets bigger than the rest. The `ellipticity_prior` term also increases with the intrinsic ellipticity of a galaxy. For galaxies with low `inverse_variance`, this helps avoid fitting background noise into the galaxy. However, the measured shape will be biased. This bias also exists in the COSMOS Deep catalog. If we take the shape information as it is, and construct galaxy stamps directly from the COSMOS Deep catalog, the bias in shape would be applied twice: First, the input catalog is biased in shape. Second, our pipeline fitting is the same as LEGACYPIPE, so we would use the same ellipticity prior during the fitting. Our output result made with this method would have a different distribution in shape compared with the real LRGs. We would not be able to fully represent the LRG distribution morphologically. Thus, we develop a procedure to de-bias the shape distribution in the COSMOS Deep truth distribution. The procedure is described in Appendix C. This procedure modifies the input shape of truth catalogs. After this de-biasing procedure, the shape distribution of output Obiwan LRGs is consistent with DESI LRGs.

5.2 Source injection strategy

The injected galaxies on the images need to be far from each other while maintaining a large number density. The optimal pattern to achieve this is the hexagonal pattern. This method achieves the maximum number density for a given minimum distance between sources. Thus, we adopt this pattern for source injection. The distance between the closest sources is 360 pixels, or 94". We deliberately chose this large separation to avoid blending on WISE images. We also eliminate sources that touch pixels containing bright galaxies, bright stars, and clusters; the details are in Appendix A.1.

6 Obiwan with DR9 images

We inject galaxies in the DECaLS North Galactic Cap (NGC) and South Galactic Cap (SGC) regions, and the footprint is shown in Figure 4. If no pixel is masked, 100 galaxies are injected for each brick.

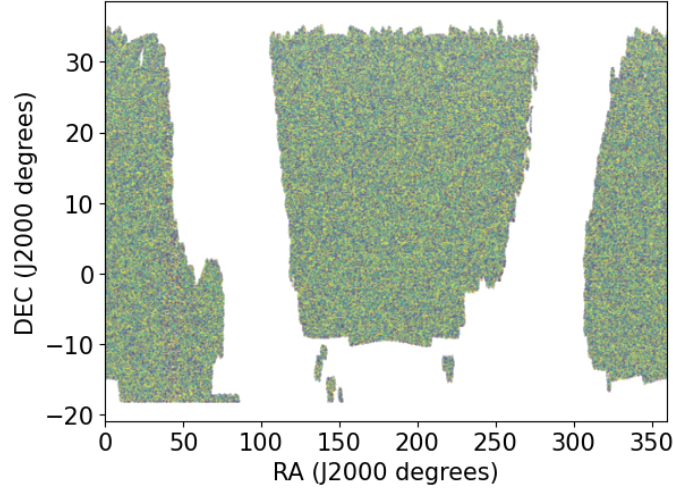


Figure 4: The footprint of Obiwan LRGs we simulated, displayed in Right Ascension and Declination. It is the same footprint as DESI DECaLS NGC and SGC above $\delta = -30$. DESI sources are only targeted above $\delta = -30$, so the region below is not scientifically important. The northern part imaged with BASS/MzLS is not processed, because we do not know the truth input in that region. The patch in the middle is the North Galactic Cap (NGC) region, and the patches on both sides are the South Galactic Cap (SGC) region.

In total, 6.6 million galaxies are simulated in the SGC, and 8.5 million galaxies are simulated in the NGC. The density of final Obiwan LRGs is about 30% that of DESI LRGs. The number density of Obiwan LRGs is limited by computational resources, and we put more details in Appendix A.3. We also found that the flux of output galaxies is systematically lower than the injected values. However, this flux bias has a negligible effect on LRG target selection. Appendix D provides an in-depth discussion over this issue.

6.1 Morphology test

Morphologically, six types of light profiles are included in LEGACYPipe, and 5 of them are used in our analysis. The unused one is the "DUP" type, a confirmed star by external surveys. Below is a list of the five types of profiles; each type needs a different number of parameters to describe its shape.

- **PSF**: point source: 0 parameters
- **REX**: round exponential: 1 parameter (radius). Ellipticity is held constant at zero (circular profile)
- **EXP**: elliptical: 3 parameters (radius and two ellipticity components)
- **DEV**: deVaucouleurs: 3 parameters (radius and two ellipticity components)
- **SER**: Sérsic: 4 parameters (Sérsic index, radius, and two ellipticity components). The light intensity profile follows $I(R) \propto e^{-bR^{1/n}}$, where R is the half-light radius, and n is the Sérsic index which controls the degree of curvature of the profile.

Figure 5 shows the proportion of different types of galaxies in DESI LRGs and Obiwan LRGs. We see good agreement between these two. This results from the debiasing method described in

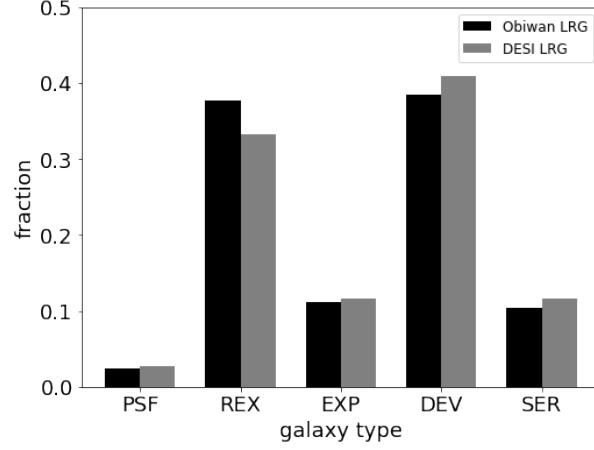


Figure 5: The proportion of different light profiles chosen as the best-fit model in DESI LRGs and Obiwan LRGs in Legacy Surveys DR9. We see that the Obiwan LRGs have similar a proportion in recovered profile types of galaxies as the DESI LRGs. See the text below for the definitions of the labeled profiles.

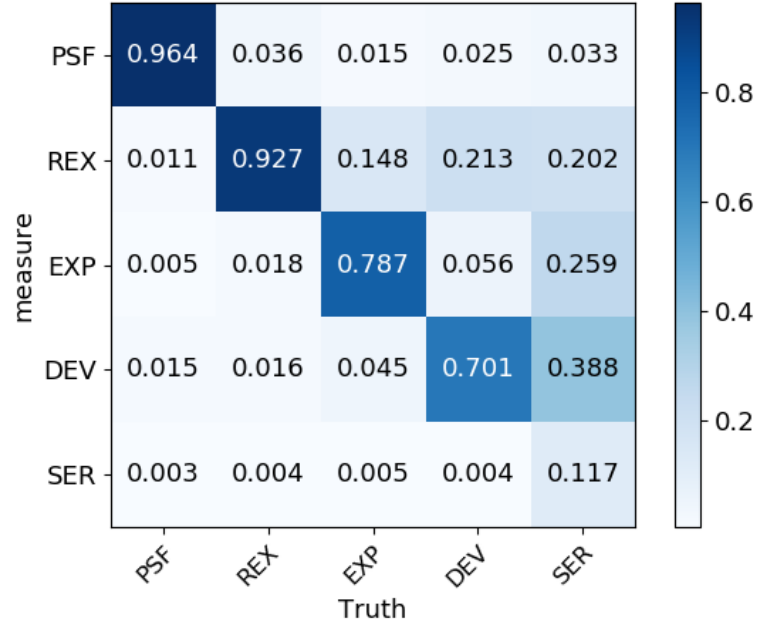


Figure 6: A matrix showing the proportion of sources measured given the light profile denoted on the y-axis given the input profile used to simulate the source, denoted on the x-axis. The fitting pipeline penalizes model complexity and thus will use a more complex model to measure flux only if the image signal to noise is high enough to detect a preference statistically. See the text for details on the specific profile and statistical thresholds.

Appendix C, as galaxy types are closely related to galaxy shapes. For example, PSF and REX sources have an ellipticity of 0, so we need the sample to have the right proportion of sources with an ellipticity of 0. With a precise estimate of morphological distribution from the input truth data, we obtain an accurate shape distribution in the output Obiwan LRGs. Figure 6 shows the scatter between the input and recovered output morphological type of our simulated galaxies. Most galaxies are recovered as their input type, or a simpler type (meaning a model of fewer parameters). As is shown in equation 5.1, the model used depends on several terms. The first term decides how much the image deviates from the model, and it is controlled by the noise level in that location. The third term does not depend on the variance in the image, but only on the number of parameters used in the model. When the image is more noisy, the first term becomes less important, and the pipeline favors a model with fewer parameters. Additionally, there is minimal misclassification of exponential galaxies (EXP) as de Vaucouleurs (DEV) galaxies or vice versa. This could be useful for identifying LRGs with star-forming disks. In Section 7.3, we found a correlation between the morphological type of LRGs and z -band `psfsize`. An accurate shape recovery helps us investigate this issue.

6.2 Imaging Systematics

We use survey property maps to demonstrate how measurements of galaxy properties are contaminated by numerous effects such as Galactic extinction, denoted as $E(B - V)$; the local density of stars at some flux threshold, denoted as `stardens`; the detection limit of a point source, denoted as `psfdepth`⁵; and the amount by which the Earth’s atmosphere disperses the light we gather with our telescopes, denoted as `psfsize`. For foreground contaminations like `stardens` and $E(B - V)$, we use maps⁶ produced as described in section 4.5.2 of [53]. The maps are produced using HEALPIX [64] at a resolution $N_{\text{side}} = 256$. For effects related to observing conditions, like `psfsize` and `psfdepth`, we use direct outputs from the catalog produced by the image reduction pipeline. We see a slight improvement in the matching between Obiwan LRGs and DESI LRGs, particularly in using the catalog level `psfdepth`, compared with using pixel values in HEALPIX. `psfdepth` depends on the list of images processed in a certain footprint. The combination of image coadds varies at a resolution smaller than HEALPIX resolution, so it is also conceptually favorable to test the trends of `psfdepth` on a catalog level.

We attach the values of `stardens` ($E(B - V)$ is already in the catalog) from the HEALPIX map to the catalogs of DESI LRGs and Obiwan LRGs. We use Obiwan input as randoms. Using the Obiwan input as randoms is equivalent to using a general random catalog. The general random catalog is uniformly distributed in the (RA,DEC) sphere. We also apply the same mask to it as DESI LRGs. The random catalog from Obiwan input has the advantage of reducing Poisson noise when tested with Obiwan LRGs. The data counts are binned into eight even percentiles of the given property. Then, the reference random sample is binned in the same way. We divide the data counts by the reference and normalize them by the total number of reference counts to data (so that the overall mean is 1). We obtain results separately for data in the NGC, shown in Figure 7, and SGC, shown in Figure 8.

To quantify the consistency between the data and the simulated trends via a χ^2 statistic, we first need covariance matrixes for DESI LRGs and Obiwan LRGs. We use:

$$Cov_{\text{LRG}} = Cov_{\text{mock}} \quad (6.1)$$

$$Cov_{\text{Obiwan}} = Cov_{\text{mock}} + Cov_{\text{randoms}} \quad (6.2)$$

⁵`psfdepth` is 5σ point source detection limit. `galdepth` is the 5σ detection limit for a fiducial galaxy. They are similar, and here we use `psfdepth`. The conclusions apply to both.

⁶We specifically used <https://data.desi.lbl.gov/public/ets/target/catalogs/dr9/1.1.1/pixweight/main/resolve/dark/pixweight-1-dark.fits>.

where the Cov_{mock} is obtained from 1000 Flask [65] mocks with the same redshift distribution and density of DESI LRGs. We computed imaging systematics trends for each of the 1000 mocks. For each survey property map m , we denote the density variation for mock i as sys_i^m , then the element in row j column k for Cov^m (mock or random) is

$$Cov_{ik}^m = \sum_{i=1}^{1000} (sys_i^m[j] - sys_{mean}^m[j]) * (sys_i^m[k] - sys_{mean}^m[k]) \quad (6.3)$$

Here j, k denotes the bin number of the survey property m . We have 8 bins for each survey property map, so j goes from 1 to 8. Similarly, Cov_{randoms} is obtained from uniformly distributed randoms with the same number density as the Obiwan LRGs, which thus accounts for the noise of their finite statistics. Cov_{mock} is much larger than Cov_{randoms} .

We calculated χ_{Obiwan}^2 values comparing the trends observed in the real LRGs compared to the Obiwan LRGs

$$\chi_{\text{Obiwan}}^2 = (sys_{\text{LRG}} - sys_{\text{Obiwan}} + c \cdot \mathbb{1}) \cdot Cov_{\text{Obiwan}} \cdot (sys_{\text{LRG}} - sys_{\text{Obiwan}} + c \cdot \mathbb{1})^T \quad (6.4)$$

and also compared to the null expectation of 1,

$$\chi_{\text{null}}^2 = (sys_{\text{LRG}} - \mathbb{1} + c \cdot \mathbb{1}) \cdot Cov_{\text{LRG}} \cdot (sys_{\text{LRG}} - \mathbb{1} + c \cdot \mathbb{1})^T \quad (6.5)$$

Here the constant c shifts systematic trend difference up or down to minimize the χ^2 value, sys_{LRG} and sys_{Obiwan} are the systematic trend in DESI LRGs and Obiwan LRGs. The constant c can be theoretically calculated as:

$$c = \frac{-(sys_{\text{LRG}} - sys_x) \cdot Cov_y \cdot \mathbb{1}}{\mathbb{1} \cdot Cov_y \cdot \mathbb{1}} \quad (6.6)$$

Here x is Obiwan or null, and y is Obiwan or LRG.

With a degree of freedom of 8, a trend that fully resembles DESI LRGs trend should have a $\chi^2 < 13.4$ at 90% confidence level, or $\chi^2 < 15.5$ at 95% confidence level. In the NGC, the plot shows that there is no clear systematic trend in `psfsize g`, `psfsize r`, `psfsize z`, and `psfdepth W1`, either predicted or observed. We see that the Obiwan LRGs successfully predict the trends observed in `psfdepth g`, `psfdepth r`, and `psfdepth z`. The images' depth indicates the galaxies' noise level, which suggests that the systematic trends are from photometric scattering. However, we do see trends in extinction and stellar density that are not predicted by Obiwan. We investigate these further in Section 7.2.

In the SGC, we see qualitative agreement in the slopes of the trends, but the χ^2 results suggest a poor match between the DESI LRGs trends and those that Obiwan predicts. The mismatch for `psfdepth` is pronounced in high-depth regions. We explain the reason for this mismatch in Section 7.1. In short, this mismatch does not result from failure in Obiwan simulation. It arises from not having a perfect truth input. The mismatch can be corrected if we match the color distribution of DESI LRGs with Obiwan LRGs.

7 Testing the Obiwan procedure

There are several trends between the observed LRG density and imaging properties that we could not simulate in the fiducial Obiwan settings. In this section, we discuss components not simulated by Obiwan, which causes a mismatch in imaging systematics trends that we see in Figure 7 and 8.

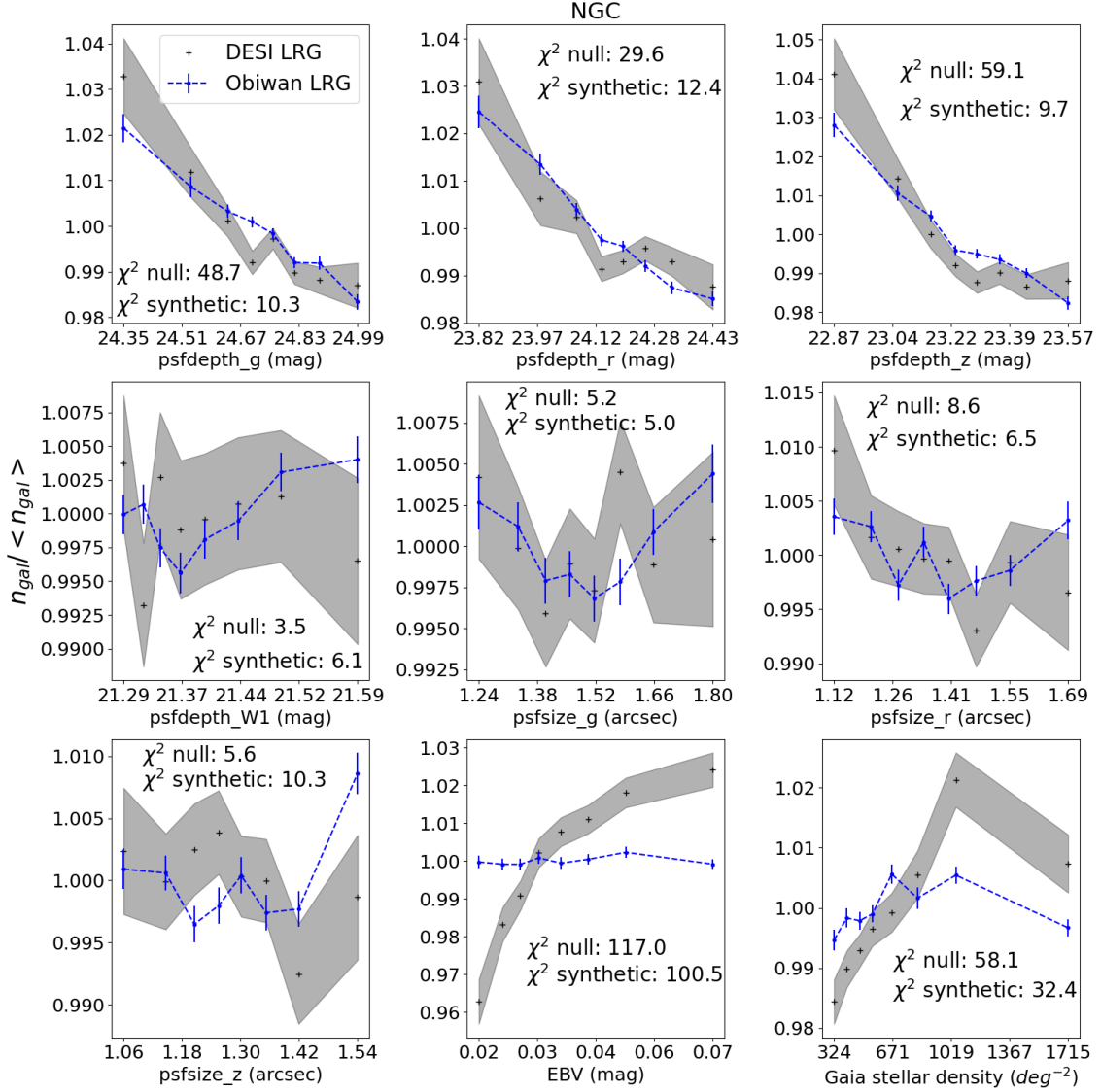


Figure 7: Predicted and observed trends between LRG density and imaging properties of the North Galactic Cap. The shaded curves are the results for DESI LRGs, while the blue lines are for synthetic Obiwan LRGs. From top-to-bottom and left-to-right, the first 4 plots are $psfdepth$ trend in g , r , z , $W1$ bands, respectively. The next three plots show the trends against $psfsize$ in the g , r , z bands. Next is the trend against Galactic extinction as measured by [66], and finally, the trend against stellar density. The observed trends with $psfdepth$ are predicted. Neither sample has an obvious trend with $psfsize$. The trends in $E(B - V)$ and stellar density are not predicted by the synthetic Obiwan-LRGs, and we believe this is due to contamination in the $E(B - V)$ map (see Section 7.2).

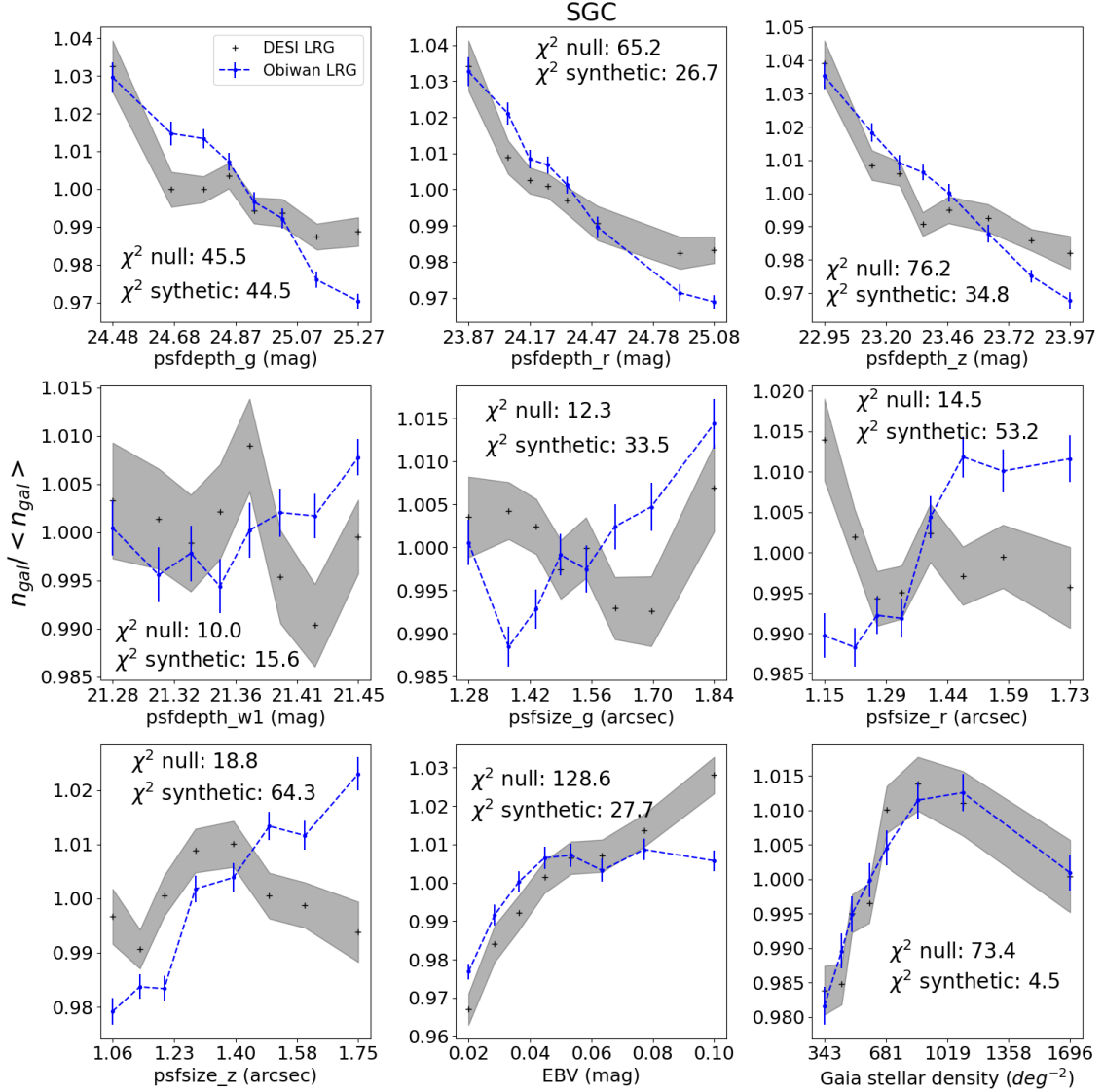


Figure 8: The plots are similar to that of Figure 7 except that it is processed on the SGC region. The conclusions are similar to those of NGC. One noticeable difference is that the trends with $psfdepth$ are not as well predicted for the NGC. We believe this is related to how the truth sample is constructed, which induces a mismatch between Obiwan LRGs and DESI LRGs. Details are discussed in Section 7.1.

Section 7.1 discusses the dependency of imaging systematics on intrinsic galaxy brightness, and how imperfections in the "truth catalog" cause mismatch by not having the perfect magnitude distribution. Section 7.2 discusses uncertainties in the $E(B - V)$ maps and its correlation with the number density of DESI LRGs. Section 7.3 discusses signs of un-simulated PSF error and that it could potentially change the characteristics of the LRG population near the color selection boundaries. Section 7.4 discusses other non-simulated secondary effects that could bias our results, but we do not see a clear sign of their impact. Section 7.5 discusses potential issues with using survey property maps for imaging systematics mitigation, the most widely used method for current galaxy surveys. We identify

two factors not commonly considered with this method and how forward modeling tools like Obiwan can help. Section 7.6 talks about how the systematics discussed in this work impact cosmological measurements by potentially introducing a varying galaxy bias that is not properly mitigated.

7.1 Correlation With Galaxy Brightness

The trends with `psfdepth` still have a small discrepancy between what is observed and predicted, especially at the greatest depths. This originates from the small mismatch in the magnitude distribution between DESI LRGs and Obiwan LRGs, shown in Figure 9. The panels show that Obiwan LRGs are generally fainter than DESI LRGs, especially in the z -band. We investigate possible explanations for this mismatch.

One possible reason is that the COSMOS Deep data is not strictly noiseless. Although COSMOS Deep is overall a clean sample in optical bands, the W1 flux in the WISE band does not have deep photometry, as is shown in Figure 19. The luminosity cut shown in equation 4.4 is an important source of photometric scattering, and it involves utilizing the WISE W1 band. We do not find evidence of any strong discrepancies induced by imperfect WISE band input, but we can not rule out the possibility that it could cause a mismatch in color distribution in other bands.

A second reason is the sample variance from the truth catalog. The truth catalog is from a small footprint, so the LRG-like sample selected in this region is slightly different from the LRG-like sample in the full footprint. This sample from the truth catalog serves as seeds that generate Obiwan inputs. We inject these Obiwan inputs to images and further produce Obiwan LRGs. Therefore, the features exclusively of the LRGs from the truth catalog can also be seen in Obiwan LRGs. We demonstrate this effect in Figure 9. In the left plot, we see LRGs selected from the COSMOS deep catalog have a dip at z -band fiber magnitude around 20.5. As we use COSMOS deep catalog as "truth input", this feature is also seen in Obiwan LRGs. Due to this sample variance, the COSMOS deep region LRGs also have a higher density LRGs at the fainter part of z -band fiber magnitude (>21.4) than the overall magnitude distribution of DESI LRGs.

Thus, our discrepancy comes from not having a large enough footprint of the COSMOS Deep catalog. The mismatch does not appear only on z -band, but compared with other bands, z -band is easier to see the difference by eye. Thus, we choose z -band data for illustration. The result here suggests that we should use a larger footprint to generate the "truth catalog" in future work. If everything worked correctly, we would see a better match between DESI LRGs in COSMOS deep (or any other deep data used), and the DESI LRGs in the NGC (or any other larger footprint). We should also see a much smaller difference between the z -band magnitude distribution of Obiwan LRGs and DESI LRGs.

So far, we have only looked at trends between the full LRG target sample and the survey property maps. As the strongest trends we predict are with the imaging depth, one could reasonably expect the significance of the trends to depend on the brightness of the LRG target. To investigate this, we split the DESI LRGs sample into two sub-samples: the bright sample (z band magnitude < 20.5), and the faint sample (z band magnitude > 20.5). With this split, we observe dramatically different trends, as shown in Figure 10.

For the brighter half of the data, we find no significant trend observed or predicted with the depth. Conversely, a quite strong trend is observed in the faint half of the data, with $\sim 15\%$ variation in number density as a function of depth. The trends are well-predicted by Obiwan, with similar performance in both the NGC and SGC. This suggests that a portion of the mismatch found for the SGC in Figure 8 was due to the excess of faint galaxies in the Obiwan sample predicting a somewhat stronger trend in the full sample than observed in the real data.

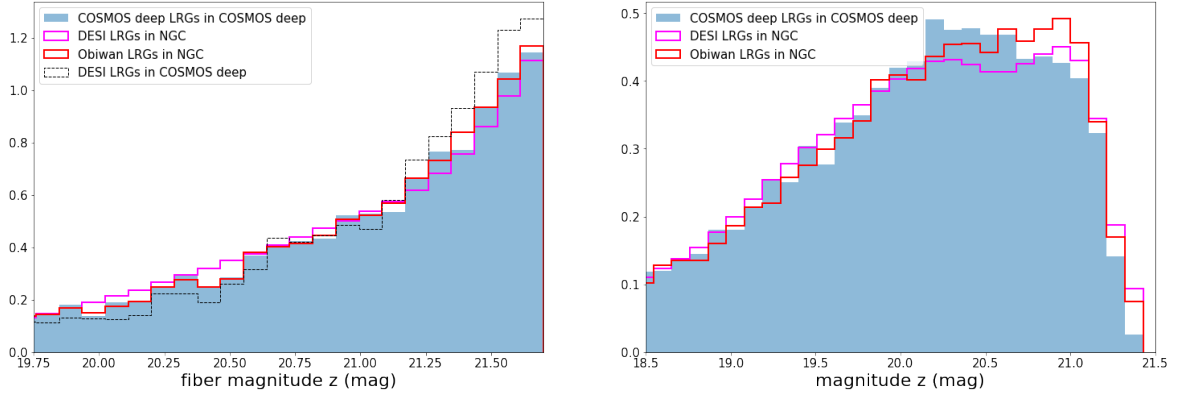


Figure 9: The z -band fiber magnitude and total magnitude distributions of LRGs in different catalogs and regions. The blue histograms show LRGs in the COSMOS Deep catalog. The black histogram shows DESI LRGs cut to the COSMOS Deep footprint. The magenta histograms show DESI LRGs in the whole North Galactic Cap (NGC) region. The red histograms show Obiwan LRGs in the NGC region. The left plot shows a histogram in z -band fiber magnitude. The right plot shows the total z -band magnitude. The left plot suggests that the COSMOS Deep region is a small region, and DESI LRGs in this same region (Black) have a very different distribution compared to DESI LRGs in a much larger footprint (Magenta). However, LRGs selected from the COSMOS Deep catalog (Blue) show a better resemblance to the histogram of LRGs in a larger footprint (Magenta). The COSMOS Deep LRGs (Blue) have an under-density feature at around z -band fiber magnitude 20.5. This feature is transferred to Obiwan LRGs (Red) with Obiwan simulation so we see the same under-density in Obiwan LRGs in the same z -band fiber magnitude range.

The results in this section suggest that the best modeling of the variation of the DESI LRGs density with imaging properties would include the brightness of the LRG target. The fiducial DR1 LRG corrections for imaging systematics only split the sample by redshift [67]. We recommend that the analysis of the sample eventually used for primordial non-Gaussianity, f_{NL} , (results of which are to be included in [68]) makes sure the results are robust to selections on the LRG brightness, as f_{NL} analyses are particularly sensitive to spurious large-scale clustering induced by imaging systematics [9, 14].

The Obiwan LRGs predict the brightness-dependent trends with imaging systematics. However, as is shown in Figure 9, it is difficult to achieve a perfect match in magnitude distribution between DESI LRGs and Obiwan LRGs. To mitigate this issue, we re-weight Obiwan LRGs so that their magnitude distribution matches with DESI LRGs. We use the basic bin-based reweighting method⁷ to match the magnitude distribution of g , r , z , $g - r$, and $r - z$ of Obiwan LRGs to DESI LRGs. The matching is successful in that the KS values obtained from the comparison of the Obiwan LRGs and DESI LRGs magnitude distributions are reduced from 0.018, 0.028, 0.038, 0.027, 0.016 to 0.008, 0.008, 0.007, 0.007, 0.006, respectively. We apply these weights to Obiwan LRGs and re-compute the systematics plot with these weights applied.

Figure 11 shows the trends between LRG density and the psfdepth in the SGC, after re-weighting the Obiwan LRGs. Compared to the result in Figure 8 (We repeated relevant plots in Figure 11), the χ^2 value in Obiwan LRGs is greatly reduced. This re-weighting process is a rudimentary study of how to improve the matching between real and synthetic sources. In regions of different

⁷<https://hsf-training.github.io/analysis-essentials/advanced-python/45DemoReweighting.html>

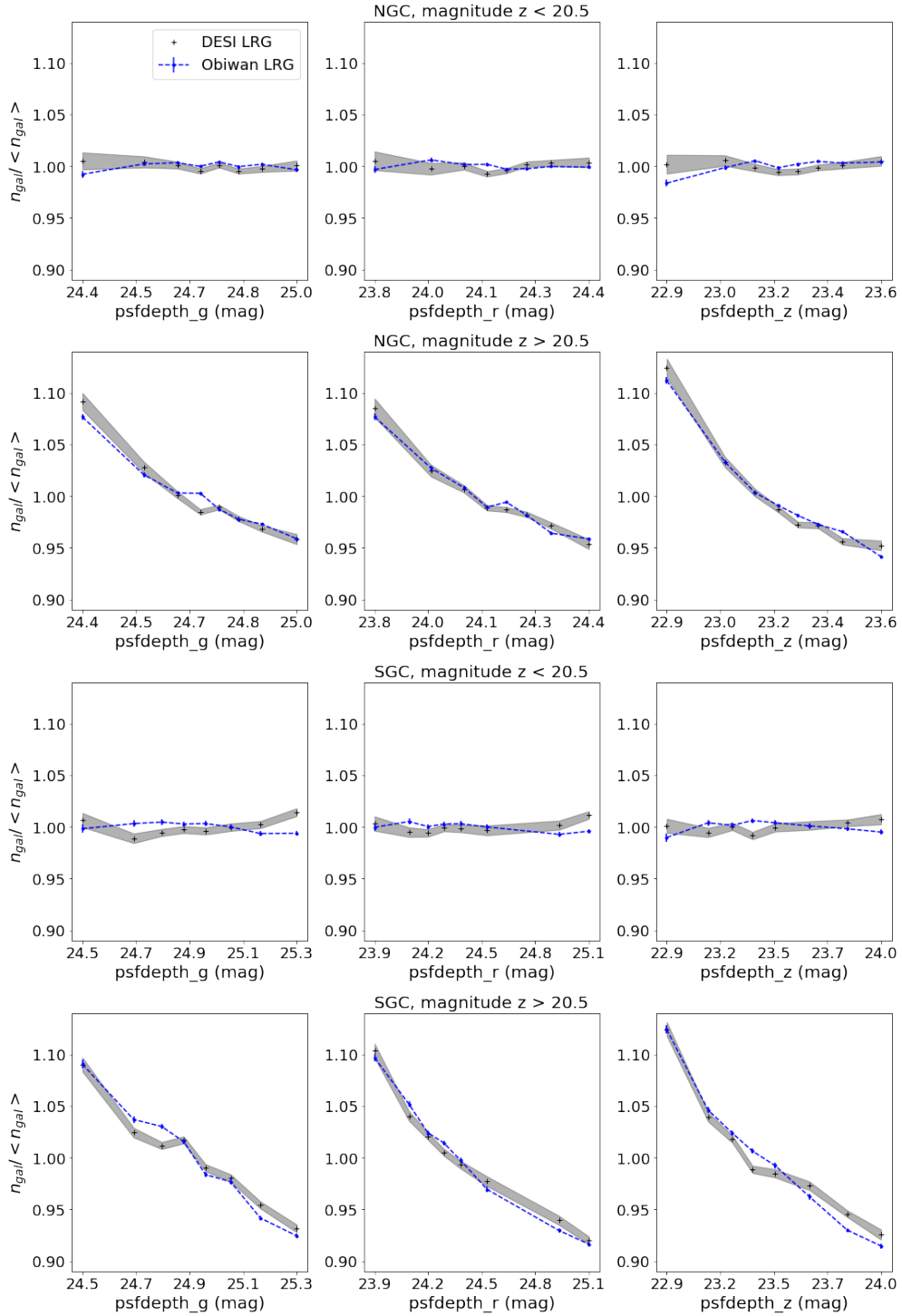


Figure 10: Systematics trends with psfdepth , when LRGs are split into a bright sample (rows 1&3) and a faint sample (rows 2&4), in the NGC (rows 1&2) and SGC (rows 3&4). The grey line represents DESI-LRGs, while the blue dashed lines represent Obiwan LRGs. The y-axis in all plots are set to the same range. The result here suggests that imaging systematics for DESI LRGs have a strong dependency with their intrinsic magnitude, and this variation is well predicated by Obiwan LRGs.

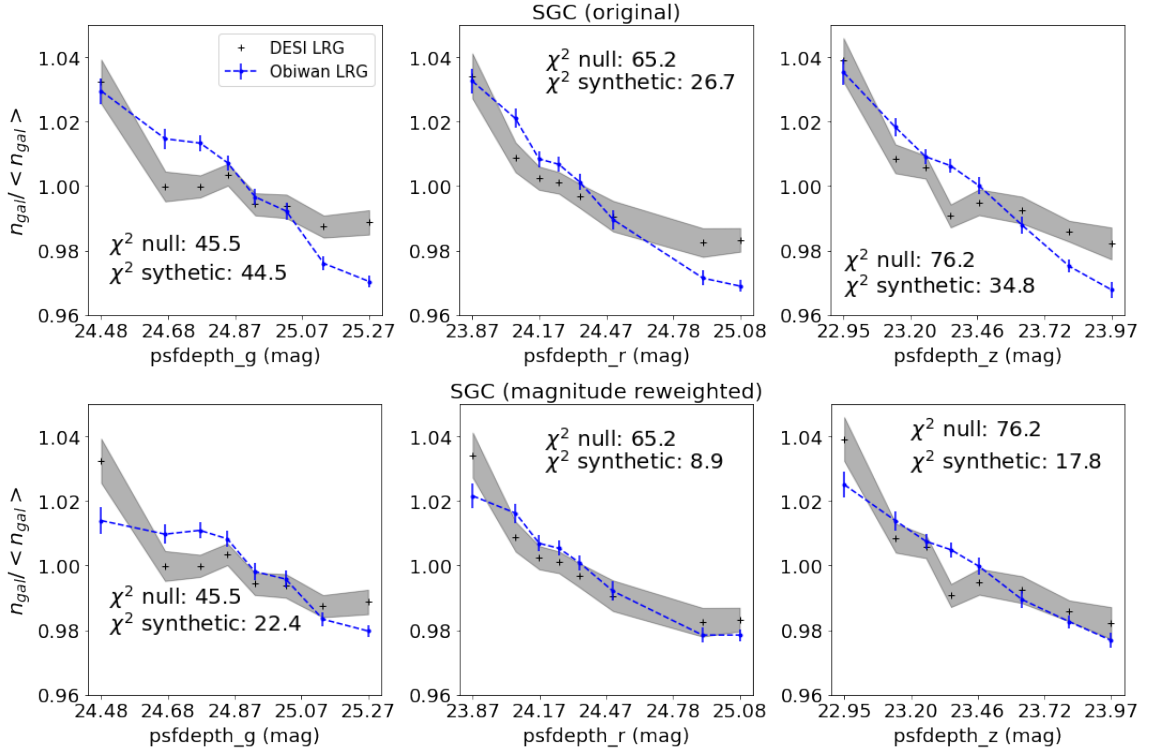


Figure 11: The trends in the SGC between DESI LRG and Obiwan LRG density and imaging depth before (upper row) and after (lower row) we apply magnitude re-weighted weights to Obiwan LRGs

survey properties, the magnitude distribution varies. This method ignores this variation. In addition, although the re-weighted distribution has a much smaller KS value, it is still not small enough to mathematically declare the distributions of Obiwan LRGs and DESI LRGs are drawn from the same underlying distribution. In our future work, we will address this issue thoroughly.

7.2 Galactic Extinction

Various maps have been developed to correct for galactic extinction. As detailed in [69], these maps can be divided into three categories: Thermal emission-based maps, e.g., [66], stellar reddening-based maps, e.g., [70, 71], and HI-based maps, e.g., [72]. The Galactic extinction is typically mapped as $E(B - V)$, in magnitudes, and a correction to apply to any measured magnitude is a constant coefficient [73] for the particular wavelength band, multiplied by the $E(B - V)$ at the given celestial coordinate. The SFD ([66]) map is the most commonly used map for the application of extinction corrections in current galaxy surveys, including the DESI LRGs targets. It is thermal-based, and the data are directly collected from dust emission. This gives it high accuracy. However, it suffers from extragalactic imprints. When stacking the SFD $E(B - V)$ value around reference objects, [69] see that the average $E(B - V)$ value correlates with the redshift of these objects, indicating leaks of Cosmic Infrared Background (CIB) signal. The stellar-reddening-based maps measure dust emission on a different wavelength range, in which CIB contamination is negligible. These maps measure $E(B - V)$ by comparing stellar spectra to templates while fitting for the $E(B - V)$ value at the same location. The accuracy of this method is limited by the number of usable stars at any given location. Moreover, the zero-point calibration uncertainties from star catalogs used for stellar-reddening-based $E(B - V)$ maps could bias the measurements for these maps [74]. However, they are free from CIB leaks.

HI-based maps trace the hydrogen column density, correlated with $E(B - V)$. HI-based maps are not strictly $E(B - V)$ maps. Due to the scope of this work, we did not use HI-based maps in our test.

Observations of distant galaxies are affected by $E(B - V)$. If we have a “truth map” for $E(B - V)$ (and per-band coefficients), it would not be a big problem, as the needed corrections to the photometry would be known. However, we do not have such a “truth map”. When we detect a correlation of our galaxies’ relative density with the $E(B - V)$ map, we are not certain how the error in our $E(B - V)$ map plays a role in such correlation.

We will consider three potential factors that could contribute to the trends we observe between $E(B - V)$ and DESI LRGs:

1. The inaccuracies in the map influence the target selection of DESI LRGs by shifting the dust-corrected flux of each galaxy.
2. The $E(B - V)$ values are correlated with the true background galaxy density due to some systematics in the $E(B - V)$ map, e.g., the CIB leaks. For LRGs, their own dust can absorb light and re-emit it in the infrared, contributing to the CIB. Thus, regions with a higher density of LRGs will have higher CIB and a positive trend should be expected between the LRG density and an $E(B - V)$ map with CIB contamination.
3. Dust extinction induces more photometric scattering. As galaxies are dimmer at a higher extinction, fewer photons from the galaxy reach the telescope given the same exposure time. This effectively reduces the depth of the image. Obiwan simulates this effect.

Factors 1 and 2 are not recovered by Obiwan because Obiwan assumes no error in the $E(B - V)$ map, while factor 3 should be automatically included during Obiwan’s image simulation procedure. In DESI LRGs, we observe a strong correlation of the relative number density of LRGs with the $E(B - V)$ SFD map, as shown in Figures 7 and 8. Meanwhile, Obiwan LRGs exhibit a weaker correlation. Under our assumption, the weaker correlation in Obiwan LRGs is due to not simulating factors 1 and 2 in the Obiwan simulation.

To test the impact of the three factors on the LRG density vs. $E(B - V)$ trend, we redo the DESI LRGs target selection based on two alternative maps. One of the alternative maps is the stellar-reddening-based map produced by [70], which we label ‘Mudur+23’. This map derives $E(B - V)$ map based on data from Pan-STARRS1 [75] and 2MASS [76] along with Gaia parallaxes [77]. The other alternative map is the Corrected-SFD (CSFD) map produced by [78]. It cross-correlated the SFD map with the galaxy clustering signal and separated the SFD map into two maps: the CIB residual map and the CSFD map. The CSFD dust map has a much lower residual CIB contamination than SFD. However, the correction is CIB only. Non-CIB systematics are not included in the CSFD map. Compared to the stellar reddening-based map, the $E(B - V)$ values on the CSFD map share more similarity with the SFD map. We define LRGs selected with these maps `real` LRGs, since they are all real galaxies we observe. When using an alternative $E(B - V)$ map, `real` LRGs is not always identical to DESI LRGs. When we use the alternative $E(B - V)$ maps, the (extinction-corrected) magnitudes of the galaxies change. When we apply LRG target selection to these galaxies, we obtain a sample slightly different from DESI LRGs. Meanwhile, if we use SFD map in LRG target selection, `real` LRGs are identical to DESI LRGs.

The trend for the density of the `real` LRGs against the three $E(B - V)$ maps (the original SFD map and the two alternative maps) are shown as blue, orange, and green curves in Figure 12. The result is plotted against the $E(B - V)$ assumed in the target selection (i.e., the x -axis is three different quantities). We also show an additional red curve. It uses DESI LRGs selected by the SFD map, but it

is plotted against the same stellar-reddening-based map as the green curve (the x -axis uses quantities in the stellar-reddening-based map). We define this procedure as “using a different binning”, as it effectively only changes which $E(B - V)$ bin each LRG goes to. Compared with the blue curve, one can observe that all trends are decreased when using an alternative map for target selection and/or binning. We can use the three factors to explain what we see in Figure 12.

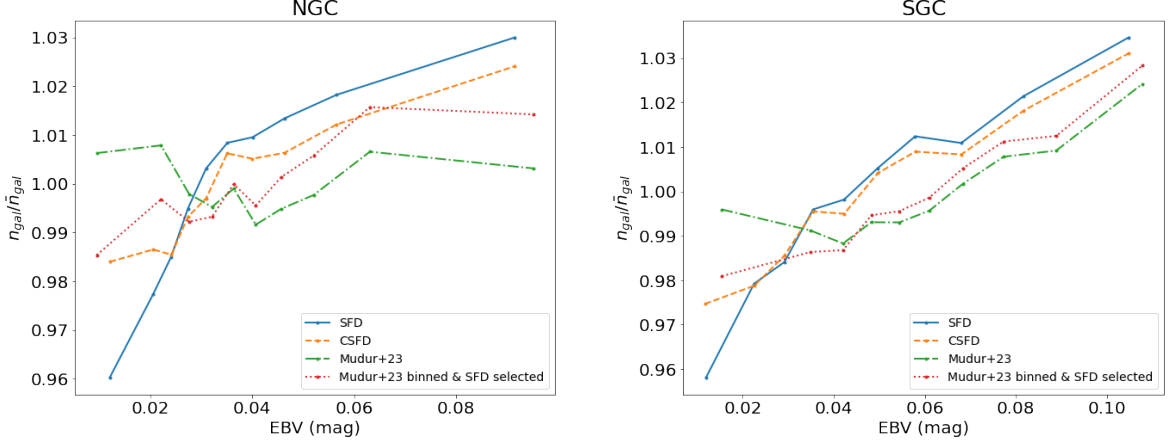


Figure 12: Density fluctuation of DESI LRGs based on different maps in the NGC (left) and SGC (right) region. The target selection process is also altered based on different $E(B - V)$ maps. The blue line is based on the SFD map [66], the green line (labeled ‘Mudur+23’) is based on the stellar reddening map from [70], the orange line is based on the Corrected-SFD (CSFD) map [78], the red line uses LRGs selected by the SFD map, but the binning is computed with the same stellar reddening map as the green line. For LRGs selected and plotted against SFD maps, we see a clear trend. In the stellar reddening map, NGC and SGC are slightly different. In NGC, we do not see any trend. In SGC, we see a trend that is weaker than the SFD map. The difference is also seen in Obiwan LRGs seen in Figure 13. In the CSFD map, this trend is similar to the SFD map. In the lower $E(B - V)$ region, the trend is weaker. The fact that the red lines exhibit a larger trend than the green line suggests that some of the $E(B - V)$ systematics trend is introduced by target selection, which uses the $E(B - V)$ to determine the colors of galaxies.

Factor 1 can be seen by comparing the red and green curves. The two samples are binned with the same stellar-reddening-based map. The red sample is DESI LRGs that uses the SFD map to determine their magnitudes and perform target selection. The green curve uses a stellar-reddening-based map for a similar procedure. Thus, the difference between these two curves is a result of the change in target selection. We find that 2.0% of LRGs are selected with the stellar-reddening based map, and not selected by the SFD map. 2.6% of LRGs are selected with the SFD map, but not selected by the stellar-reddening based map.

These two curves in Figure 12 show how LRG sample difference results in different systematics trends. The red curve has a larger trend than the green curve, which indicates that the CIB leaks affect the target selection of DESI LRGs.

Factor 2 can be seen by comparing the blue and red curves. The two trends are computed with identical DESI LRGs sample. The only difference lies in the binning. The blue curve is SFD-map-selected LRGs binned with the SFD map. It has a stronger trend. The red curve is the same SFD-map-selected LRGs but binned with the stellar-reddening-based map. It has a weaker trend. The reduction of trend in the red curve compared to the blue curve is partly a result of less

CIB contamination in the stellar-reddening-based map. However, it is worth noting that this trend reduction could also result from uncertainties in the Mudur+23 map: If we add some noise to the “ $E(B - V)$ truth map”, then some $E(B - V)$ values will be falsely assigned to other bins, resulting in a suppression of the actual trend. It could also result from SFD $E(B - V)$ map uncertainty: as the galaxies are selected with the SFD map, binning with the same map includes relevant unknown systematics.

This contamination within the SFD $E(B - V)$ map has the potential to impact measurements of galaxy clustering. Given that the CIB emanates from galaxies, it exhibits a correlation with galaxy clustering. Thus, the utilization of the SFD $E(B - V)$ map for imaging systematics correction would inadvertently null part of the true clustering signal, and lead to biased clustering measurements. Indeed, based on these considerations, the DESI large-scale structure (LSS) catalogs [67] in data release 1 (DR1 [79]) do not directly use the SFD or CSFD maps for the correction of systematic trends.

Factor 3 is predicted by Obiwan, since Obiwan simulates the dimming effect from dust extinction. As we discussed in Section 7.1, galaxies with different apparent magnitudes are affected by photometric scattering differently. Faint galaxies are more sensitive towards changes in the depth of images. Here $E(B - V)$ can generate an effect similar to depth by making galaxies fainter. Given its similarity with depth, we cannot see an isolated effect for Factor 3 on DESI LRGs. However, with Obiwan image simulations, we know that these effects are simulated as a whole. If we look at the $E(B - V)$ systematics trend in Figures 7 and 8, we see that the difference in systematics trend of Obiwan LRGs and DESI LRGs are larger in the NGC than the SGC, which is evidenced by χ^2 synthetic. This value is 100.5 in NGC and 27.7 in SGC. This indicates that a higher proportion of the $E(B - V)$ trend in the SGC is contributed by phenomena simulatable by Obiwan. Factor 3 is one of these phenomena.

We take a deeper look at these simulatable phenomena. We noticed that $E(B - V)$ in the SGC is negatively correlated with `psfdepth` with a correlation coefficient of about -0.24 . While in NGC it is positively correlated with a correlation coefficient of about 0.14 . The difference in the two regions is due to the survey design strategy. For the DECaLS data, the DECaLS team adjusted the exposure times to try to reach uniform depth after dust extinction [80]. The reported `psfdepth` are in observed magnitudes. Thus, the survey goes to deeper `psfdepth` where the $E(B - V)$ is larger. The positive correlation is expected in the DECaLS area. The SGC has more DES data, and DES did not adopt the depths in the same way as DECaLS. As lower `psfdepth` leads to higher relative LRG density, this partially explains why we see a stronger positive trend in Obiwan LRGs in the SGC than in the NGC. Due to its close relation with depth, this trend is also related to the magnitude of LRGs. In Figure 13, we split the sample into faint LRGs and bright LRGs in the same way as in Section 7.1. We see some trend in the SGC faint sample in Obiwan LRGs, while the trend is much weaker in the SGC bright sample. Since faint ones are more influenced by photometric scattering, this behavior is expected. In NGC, given the positive correlation between `psfdepth` and $E(B - V)$, which cancels out with each other, the trend is much weaker. We can compare the $E(B - V)$ trend of bright DESI LRGs and faint DESI LRGs by looking at the grey and black curve on the left side of Figure 13. We see that the difference is larger in the SGC than in the NGC. These all indicate that compared to the NGC, the $E(B - V)$ systematics trend in the SGC has a higher portion contributed by Obiwan simulatable phenomena that could be magnitude dependent. As the overall amplitude of the SFD $E(B - V)$ trend is similar in the NGC and the SGC, this suggests that the contributions of factors 1 and 2 (unknown $E(B - V)$ systematics, including the persistence of CIB contamination) are larger in the NGC than in the SGC. This result is consistent with the stellar-reddening-based map in Figure 12. The stellar-reddening-map-based $E(B - V)$ trend (green) which is devoid of CIB is more similar to the SFD-map-based $E(B - V)$ trend (blue) from the SFD map in the SGC than in the NGC. Our

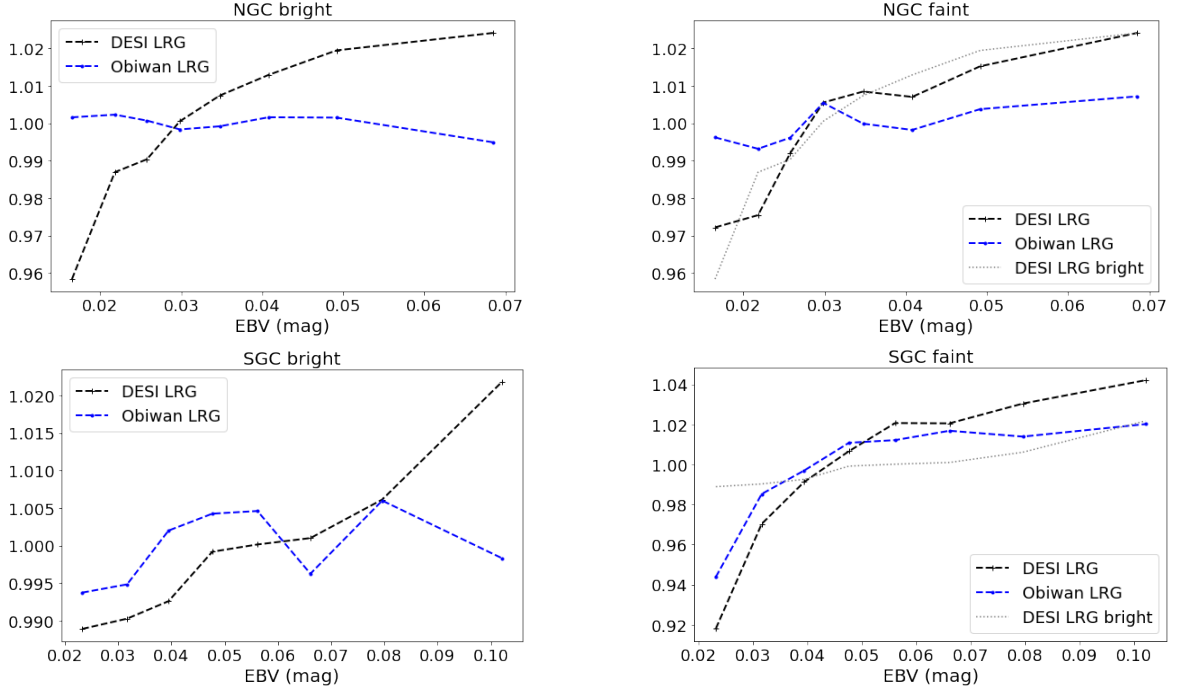


Figure 13: Systematics trend in $E(B - V)$ when we separate the sample into two subsamples: One bright sample with z band magnitude < 20.5 , One faint sample with z band magnitude > 20.5 . Blue curves are made with Obiwan-LRGs, and black curves are made with real LRGs. The grey curve on the left plots is the same black curve on the right plots, and they are plotted for a better visual comparison. We see some trend recovery in the SGC faint sample from Obiwan-LRGs, meaning part of the trend is due to simulatable phenomena by Obiwan (e.g., $E(B - V)$ in the SGC produces a shallower depth that is not corrected in the observing strategy design). The rest do not match well, indicating that the trend in real LRGs is likely due to the SFD $E(B - V)$ map systematics. The result here supports the observation in Figure 12, where we see more discrepancies between the SFD $E(B - V)$ and alternative maps in NGC than in SGC.

findings here are consistent with [81] (particularly Figure 15 in their paper), and we further discuss it in the next paragraph.

In Figure 12, the two alternative maps we use here have different trends, as shown in the green and orange curves. The CSFD map only corrects for the contributions of CIB in the SFD map, so any other systematics in the SFD map is not recovered by CSFD. As discussed in [70], there is a negative trend for star-reddening-based $E(B - V) - E(B - V)$ SFD as a function of $E(B - V)$ SFD. This trend supports the conclusion in [81]. In their work, they computed the reddening coefficient with the blue tip of the stellar locus and the measurement is based on the $E(B - V)$ SFD map. They found a dust map normalization difference of 15% between NGC and SGC that may be due to uncertainties in dust temperature correction. The dust temperature map has high uncertainty, particularly in the low $E(B - V)$ region. The angular resolution of the dust temperature map is much lower than the $E(B - V)$ map. As the $E(B - V)$ scales linearly with the dust temperature map, uncertainties in the dust temperature map bias the overall amplitude of $E(B - V)$ rather than the relative fluctuations. This effect generates a trend as a function of the absolute $E(B - V)$ value.

7.3 Investigation of Trends With `psfsize`

Obiwan LRGs predict that the LRG target density generally increases with `psfsize`, and if anything, the opposite seen in DESI LRGs (though the trends are consistent with the null expectation). To locate the problem, we first split the sample into faint LRGs and Bright LRGs in the same way as in Section 7.1. The discrepancy is mainly contributed by faint LRGs. We further split faint LRGs into a high-depth region with `psfdepth_z > 23.5` and a low-depth region with `psfdepth_z < 23.5`. We find that the trend is mainly contributed from the low-depth region.

Figure 14 shows the trend of DESI LRGs and Obiwan LRGs with these cuts (magnitude $z > 20.5$, `psfdepth_z < 23.5`). They are composed of 19.5% of all LRGs in the SGC. The trend of DESI LRGs goes down first. And then it goes up for `psfsize_g` and `psfsize_z`. However, Obiwan LRGs does not have a noticeable downward trend, while it does have an upward trend at large `psfsize`. Plots in Figure 15 indicate the reason behind the mismatch. Here, we split the sample into four regions by `psfsize_z`, then plot their histograms, binning in z magnitude. We normalize these histograms with the random galaxies in the same regions. Using this procedure, we can compare the relative densities of each region as a function of the z band magnitude. The plot for Obiwan LRGs is produced in the same way as DESI LRGs. The variation we see in DESI LRGs around z band magnitude 21 is not replicated in Obiwan.

The COSMOS repeat sets are ideal sets for pinning down this issue. We perform several tests documented in Appendix E. We eventually concluded that this effect is caused by a correlation of flux bias with `psfsize`, shown in Figure 25. The trend for this bias is larger for galaxies with Sérsic index ≥ 4 (we define them as disk LRGs). Because of this dependency with galaxy morphology, disk LRGs have a lower flux at high `psfsize_z`, so they are less likely to be selected as LRGs due to the color cut 4.2. Consequently, we see less disk LRGs in regions with large `psfsize`. This effect is further propagated to DESI LRGs after spectroscopic redshift cuts. As disk LRGs have a star-forming disk, they have stronger OII emission lines. We use the DESI LRGs with spectroscopic redshift from the iron catalog. We compute the median OII flux for each `psfsize_z` bin. This gives us an estimate of the median intensity of OII flux for each `psfsize_z` bin. In Figure 16, we found noticeable trends in all three redshift bins. The median OII intensity decreases with an increase in `psfsize_z`. We do not see similar trends in the north (BASS/MzLS region). As we have not thoroughly tested the north in this work, we do not know why similar trends are not seen in the north.

7.3.1 Possible root causes

Knowing where the discrepancy originates will help us develop treatments for this issue, so we did some further investigation. Our main suspects of root causes are background subtraction and PSF uncertainty.

We currently use the original background determination by LEGACYPipe as truth. Obiwan injects galaxies on images that are already background subtracted. Possible residuals or over-subtractions in the background subtraction are not replicated in Obiwan. The background subtraction procedure in the LEGACYPipe code masks out regions containing sources and fits a spline model [82] to the rest of the images.

In the case of PSF uncertainty, it is a widely studied subject. In Dark Energy Survey (DES) analysis [83] [84], people saw a bias between the PSF size and the size of stars (Brighter-Fatter effect). They also saw a color dependency in the fitting residuals (Differential Chromatic Refraction). The pipeline (PSFEx[85]) used for PSF determination in Legacysurvey is similar to DES (PIFF[86]), so similar systematics could also be present in Legacysurvey.

We design a test to separate the impacts of background subtraction uncertainty and PSF uncertainty. We apply a Gaussian filter to each individual image (CCD) to make them reach the same

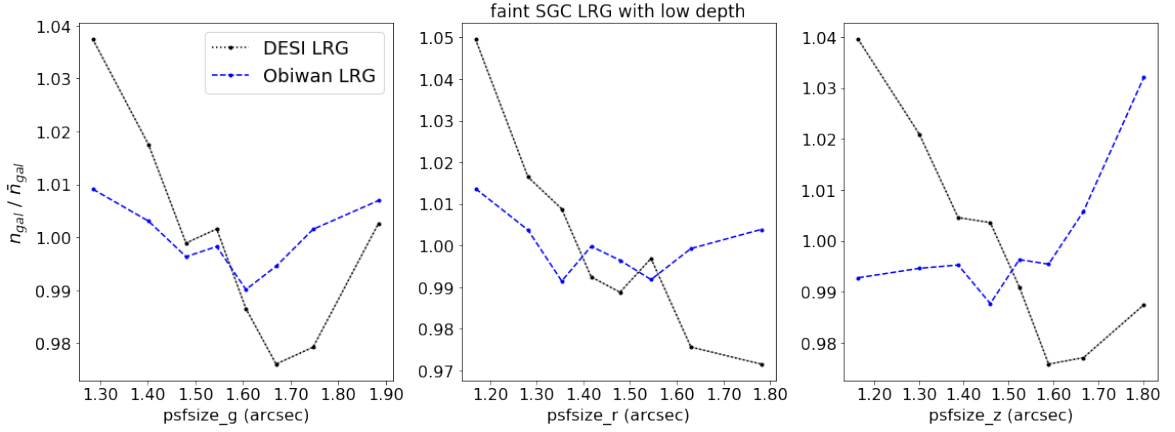


Figure 14: The DESI LRG psfsize systematics trend. The sample is in the SGC and has a magnitude cut of magnitude $z > 20.5$, and a depth cut of psfdepth $z < 23.5$. We perform these cuts because the trend is most evident in faint LRGs in low-depth regions. The black dashed line is the result from DESI LRGs, and the blue dashed line is the trend from Obiwan LRGs.

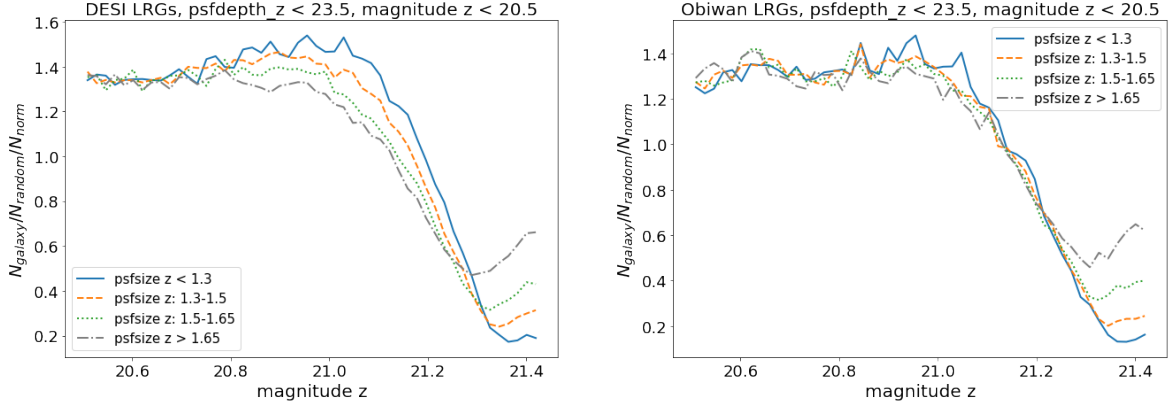


Figure 15: Histogram of z band magnitude of DESI LRGs (left) and Obiwan LRGs (right) in the SGC. The sample is split into 4 sets with different psfsize_z range. The histogram is normalized with the randoms in the same footprint so that the amplitude reflects their relative density.

psfsize. Then, we measure the 7" aperture flux of the same LRGs on different CCDs. A 7" aperture is large enough to contain most of the flux from LRGs. The 7" aperture flux is independent of any fitting methods or determination of PSF. On each CCD, we record the measured flux and the psfsize value before Gaussian filtering. These data are later used to test the correlation between the measured flux and the psfsize. We perform the same test with synthetic Obiwan LRGs on real images and also on fully synthetic images with Gaussian background noise for comparison.

We select all pairs of images with a difference in psfsize_z $> 0.12''$. With these data pairs, we found that the 7" aperture flux for the images with a larger PSF is 0.19 nanomaggies lower than the images with a smaller PSF. The equivalent value measured from the synthetic Obiwan LRGs with synthetic images is less than $\sim 10^{-3}$ nanomaggies. For Obiwan LRGs injected on real images, this number is 0.09 nanomaggies. Overall, we conclude that the background on a larger psfsize is lower than that on images with smaller psfsize. The flux loss in DESI LRGs is ~ 0.1 nanomaggies larger than the synthetic Obiwan LRGs. However, we do not find the flux loss for Disk LRGs different from

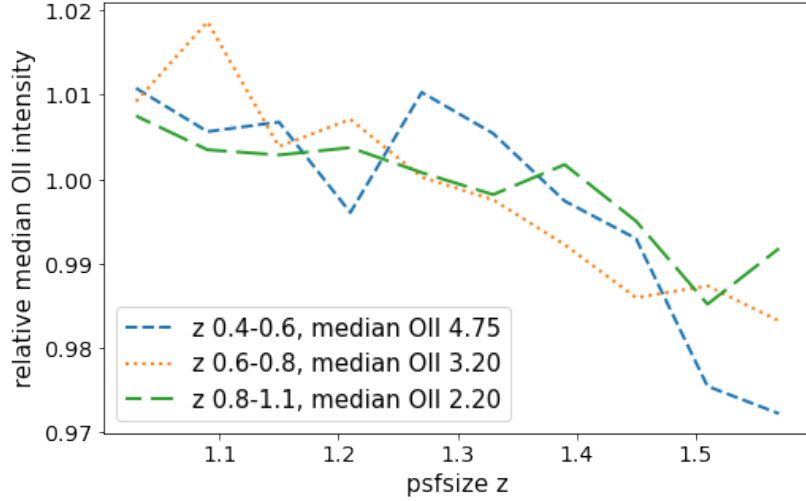


Figure 16: The median flux of OII emission line for LRGs at three redshift bins in different `psfsize_z` range, divided by the overall median OII flux. Blue is for the redshift range 0.4-0.6, orange is for the redshift range 0.6-0.8, and green is for the redshift range 0.8-1.1. The trend here originates from a higher loss in flux for disk LRGs at high `psfsize_z`.

Compact LRGs. The conclusion here applies to 3 optical (g, r, z) bands.

Our test on 7" aperture flux suggests that there could be some flux bias induced by background subtraction. In regions with larger `psfsize`, it becomes more difficult to distinguish between signals from real galaxies and signals from the background, as the galaxies on images are more spread out. The signal-to-noise is lower on each pixel. The procedure of source fitting uses a set of effective pixels. The fitting of Disk LRGs requires more pixels, so it is more sensitive to the cumulative effect of the flux bias from the background subtraction.

On the other hand, uncertainties from PSF determination may not be a dominant factor here, as our PSF-insensitive test is already sufficient to explain the trends with `psfsize`.

7.4 Other secondary effects

We have identified other effects that would potentially affect our ability to properly simulate the expected target density fluctuations in the real data. We do not see an evident mismatch when these effects are not being accounted for. However, we list these effects and potential treatments below.

Shape noise: Real galaxies are not perfect model galaxies. Since our injected galaxies are accurately described by the simple galaxy models, they would be slightly different from what we see in real galaxies. This effect is negligible for the precision we are worried about now. However, this would be a more significant issue when trying to study galaxies' shape response, e.g., as done for weak lensing measurements [87], and the studies of magnification [88]. Realistic galaxy stamps will be needed in this analysis.

Sagittarius stream [89]: No noticeable trend due to the Sagittarius stream being present in the LRGs. The contamination level from stars is extremely low for LRGs, so stars do not provide direct contamination. However, the Sagittarius stream, along with other stellar streams or Milky Way stars, can enter contamination by altering the background level of the image. As stellar streams originate from satellite galaxies, they contain many faint stars below the detection limit and merge into the

background. We may get a shifted flux measurement when falsely subtracting background noise that is faint Sagittarius stream stars.

Variation in extinction coefficients [73]: The extinction coefficient denotes the extent of extinction’s impact on each band. It is computed with stellar spectra. However, this coefficient is related to the wavelength of observation. This means that if the galaxy we observed differs from stellar spectra, the coefficient we assume would not be accurate.

7.5 Discussion on imaging systematics mitigation

Currently, studies on imaging systematics mitigation typically employ an approach based on survey property maps. The observed footprint for a survey is separated into small pixels. Each pixel is associated with a collection of quantities like depth, PSF, $E(B - V)$, airmass, etc. Thus, we have pixelized maps for each of these quantities. Studies on imaging systematics mitigation typically find a relationship between these survey property maps and the false fluctuations of galaxy densities. Then, these studies derive “imaging systematics weights”. The weights minimize the systematics trend on all these survey property maps. One weight is assigned to one pixel so that the weights form an imaging-systematics-weight map with the same pixel resolution. These kinds of methods rely on two assumptions that are not entirely true:

- **Assumption 1:** No errors exist in these survey property maps.
- **Assumption 2:** For sources at the same location, the imaging systematics associated with them are the same and can be represented as a number that we call “imaging systematics weight”.

Our study with galactic extinction shows flaws on **Assumption 1**. The SFD $E(B - V)$ map is commonly used as a survey property map. Therefore, such methods have the assumption that the SFD $E(B - V)$ map is perfectly accurate. After applying systematics weight, the sample tends to barely have any trend against the SFD $E(B - V)$ map. Our study challenges this approach of treating the SFD $E(B - V)$ map. Our findings imply that a portion of the observed trend within the SFD $E(B - V)$ map may be attributable to contamination stemming from CIB or other Large-Scale-Structure systematics. In Section 7.2, factor 1 could result in false fluctuations of galaxies that need to be corrected with weights. Since target selection defines color cut to select LRGs, this shift is most evident around the color selection boundaries. This effect needs to be carefully corrected since its impact on different magnitudes of LRGs may vary significantly. Factor 2 is purely due to false binning and should not be corrected at all. Ideally, systematics-weighted galaxies should still have a positive trend against the SFD $E(B - V)$ map, as we should not correct the trend that stems from factor 2.

Our study on galaxy brightness in Section 7.1 and `psfsize` in Section 7.3 challenges **Assumption 2**. We show that the imaging systematics trend for DESI LRGs changes significantly when we apply a cut on z -band magnitude. The mechanism behind this is well-understood: Faint galaxies have more photometric scattering than bright galaxies so they are affected by the variation of galaxy depth and seeing (`psfsize`) to a larger extent. If two galaxies of different magnitudes have the same configuration in a survey property map, they could be affected by imaging systematics differently. However, the current survey-property-map-based approach will assign them the same “imaging systematics weight”. Applying the same correction for galaxies with different brightness will result in an over-correction in brighter galaxies and an under-correction in fainter galaxies. In both cases, they introduce an artificial signal when computing the sample’s power spectrum.

One way to mitigate this issue is to separate galaxy samples into subsamples with different magnitude ranges. Then, we compute imaging systematics weights independently for each subsample.

However, with a lower density for each subsample, the computation of imaging systematics weight would be prone to fitting the noise of each subsample. Thus, such a step needs to be carefully validated.

Another way to deal with the issue is through forward modeling tools like *Obiwan*. Forward modeling can produce randoms with detailed scattering features for each source at any location. We can use such randoms to correct for magnitude-dependent imaging systematics. The randoms produced from forward modeling can predict density fluctuations at any color and brightness. Compared with making such predictions with real data, there are many advantages of predicting this density fluctuation with forward modeling randoms like *Obiwan* LRGs. These randoms do not contain clustering signals, ensuring recovery of imaging systematics trends solely from uniform random fluctuations. They also lack cosmic variance: Their covariance with survey property maps is smaller than real galaxies under the assumption of the same number density. Moreover, they can reach a number density larger than real galaxies with more simulations. These advantages would help make the predictions precise and free from over-fitting the true clustering signal.

Due to the scope of this work, we did not perform clustering measurements. Current *Obiwan* LRGs does not have a large enough density to be used as magnitude-dependent randoms. In Appendix A.4, we explore various methods to enhance the computational efficiency of generating *Obiwan* randoms.

7.6 Implications for cosmological analysis

For the robustness of cosmological analyses, accurately estimating and mitigating the effects of imaging systematics is crucial and will become increasingly challenging as statistical uncertainties decrease. Our work has discussed several phenomena that may go under the radar of many systematics mitigation methods: A sample could have a null systematics trend over a surveyed property in terms of number density, but its redshift distribution, galaxy bias, etc., could still change. [90] shows that galaxies’ morphology correlates with the Large-Scale-Structure environment they stay in, indicating that the morphology dependence possibly links to bias differences. The changes in the galaxy properties and thus (likely) galaxy bias parameters may complicate measurements of structure growth that combine galaxy auto correlations and galaxy-lensing cross-correlations (e.g., [91]), as these typically assume the same effective galaxy bias parameters in the modeling to recover the amplitude of matter clustering. The variation in galaxy bias is particularly problematic for primordial non-Gaussianity measurements. Many survey properties vary considerably on large scales (e.g., $E(B-V)$). When the galaxy bias shifts with these survey properties, it effectively introduces a scale-dependent bias that could degenerate with the primordial non-Gaussianity signal. [92] found that the need to include many maps in the modeling of imaging systematics inflated the statistical uncertainty by more than a factor of two. By correctly modeling how the intrinsic properties of galaxies change with survey property maps, we can reduce this statistical uncertainty induced by imaging systematics.

8 Conclusions

We conclude our findings as follows:

- To properly simulate the DESI LRGs sample, we further develop the pipeline, *Obiwan* to be compatible with both optical band images and the infrared band WISE images.
- We find shape bias in the COSMOS Deep catalog. We develop a procedure to eliminate this shape bias successfully, see Section 5.1 and Appendix C.

- Our inclusion of WISE images demonstrates how forward modeling can be used for combined survey analysis. We find significant differences in the WISE flux measured as a function of the imaging depth (see Figure 19), due to the increased number of total sources that the WISE flux can be attributed to in the deeper data. In the limited efforts of this work, we did not validate the optimal optical band depth to be used in the WISE flux measurement. However, it is possible to perform this cross-survey validation with forward modeling tools like *Obiwan*. For example, LSST can be combined with Euclid and Roman telescope to reach a broader color space. Such a combination may encounter similar issues, as the depth limit is different across surveys.
- We find a consistent negative bias in flux measurement in all bands. This bias is positive in synthetic images with artificial blending and Gaussian noise. Indicating that this bias is likely due to the background noise structures in the real images.
- We see imprints of systematics in extinction maps based on dust emission (e.g. CIB contamination, variation in dust properties that changes the optical-infrared relation). We find significant trends between the DESI LRGs density on the sky and the SFD $E(B - V)$ map [66], which are not predicted by *Obiwan*. These trends change significantly when we test against alternative $E(B - V)$ maps and are nearly removed when compared to the stellar-reddening-based $E(B - V)$ map [70]. Since CIB is correlated with galaxy clustering, we caution against any analyses of DESI LRGs eliminating this trend, as this would result in removing real large-scale structures. Meanwhile, some extent of the trends with $E(B - V)$ is real and related to the impact on the imaging depth (photometric scattering), and is predicted by our forward modeling.
- We find that *Obiwan* predicts the trends with `psfdepth` that we observe in the real data, which are similarly strong in each of the optical imaging bands. While the overall trends are predicted, statistically, they are not consistent in the SGC. The slight mismatch we see is due to cosmic variance in the truth input, which has a larger impact on the deepest data in the SGC. The mismatch in the color and magnitude distribution of *Obiwan* LRGs and DESI LRGs is the main source of the systematics trend deviation in this region. We further demonstrate that this issue can be greatly mitigated by applying a ‘color matching’ weight to *Obiwan* LRGs to match its color distribution with that of DESI LRGs. The details are discussed in Section 7.1.
- We find that the trends with imaging depth depend strongly on the z -band magnitude. When selecting the data to have $z < 20.5$, only 1% level variations are found in the observed data with the imaging depth, and *Obiwan* predicts no significant trend. However, with selection cut $z > 20.5$, up to 20% trends are observed in the real data and predicted by *Obiwan*. This suggests it is important to properly model the color/brightness dependence of systematic trends in the DESI LRGs target sample when producing any corrections for use in galaxy clustering measurements. Forward modeling randoms like DESI LRGs can be a powerful catalog to correct for imaging systematics that varies with intrinsic galaxy brightness.
- We find signs of non-simulated density variation against `psfsize`, which is especially prominent for extended galaxies. This is likely due to a bias in background subtraction that is correlated with `psfsize` for each CCD. This effect results in LRGs around color selection boundaries having different properties, and the impact is also seen in LRGs after spectroscopic redshift cut.
- We need to know how imaging systematics from each survey property are generated before developing methods to correct for these trends. Current imaging systematics methods commonly take the approach of developing one weight for each HEALPix pixel. We demonstrate that this

may not be enough if we want to reach a higher precision. Uncertainties in survey property maps, and the magnitude dependence of galaxies on some survey properties, could induce imperfect mitigation in imaging systematics that cannot be recovered with these methods. Signals like primordial non-Gaussianity are highly degenerate with imaging systematics signals. These factors can generate unknown bias in such cosmological measurements.

- By comparing systematic trends in real galaxies with those in synthetic galaxies generated by forward modeling tools like *Obiwan*, we can detect potential unknown imaging systematics and explore relevant resources to identify their root causes. Replicating imaging systematics in real galaxies through forward modeling tools like *Obiwan* is challenging work. A perfect match between real and synthetic galaxies is only possible when all potential contaminations are properly considered in the image simulation settings. However, a mismatch of density fluctuations in the synthetic galaxies also provides us valuable insights into what unknown systematics could possibly be missed. Identifying missing factors is vital for proper imaging systematics mitigation.

Indeed, the results we have described have already been used to help understand the LRG sample used for DESI DR1 [67, 79] and thus the cosmological analyses [68, 93–96].

Data Availability

We produce the *Obiwan* LRGs catalog that traces imprints of DESI LRGs in the real images and can be used for future studies. The product is currently available internally for DESI members⁸. The data will be uploaded to zenodo before publication⁹. In the future, the catalogs used in this analysis will be made public along with the Data Release 1¹⁰.

Acknowledgements

We thank Yi-Kuan Chiang for insightful discussions on dust extinction maps.

We thank Alex Krolewski and the anonymous reviewer for reading and providing feedback on this work.

HK thanks the people from the NERSC helpdesk for support in installing and compiling software in NERSC.

The results lead in this paper have received funds from MCIN/AEI/10.13039/501100011033 and UE NextGenerationEU/PRTR (JDC2022-049551-I).

AP acknowledges support from the European Union’s Horizon Europe program under the Marie Skłodowska-Curie grant agreement 101068581.

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under Contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the U.S. National Science Foundation (NSF), Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF’s National Optical-Infrared Astronomy Research Laboratory; the Science and Technology Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National

⁸https://data.desi.lbl.gov/desi/survey/catalogs/image_simulations/LRG/

⁹<https://zenodo.org/uploads/11260267>

¹⁰<https://data.desi.lbl.gov/doc/releases/>

Council of Humanities, Science and Technology of Mexico (CONAHCYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: <https://www.desi.lbl.gov/collaborating-institutions>.

The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall z-band Legacy Survey (MzLS). DECaLS, BASS 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. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory. Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by: the Director, Office of Science, Office of High Energy Physics of the U.S. Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the U.S. National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract to the U.S. Department of Energy. The complete acknowledgments can be found at <https://www.legacysurvey.org/>.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U. S. National Science Foundation, the U. S. Department of Energy, or any of the listed funding agencies.

The authors are honored to be permitted to conduct scientific research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

References

- [1] N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi et al., Planck 2018 results-i. overview and the cosmological legacy of planck, *Astronomy & Astrophysics* **641** (2020) A1.
- [2] D. Swetz, P.A. Ade, M. Amiri, J. Appel, E. Battistelli, B. Burger et al., Overview of the atacama cosmology telescope: receiver, instrumentation, and telescope systems, *The Astrophysical Journal Supplement Series* **194** (2011) 41.
- [3] A. Abramovici, W.E. Althouse, R.W. Drever, Y. Gürsel, S. Kawamura, F.J. Raab et al., Ligo: The laser interferometer gravitational-wave observatory, *science* **256** (1992) 325.
- [4] M. Lacy, S. Baum, C. Chandler, S. Chatterjee, T. Clarke, S. Deustua et al., The karl g. jansky very large array sky survey (vlass). science case and survey design, *Publications of the Astronomical Society of the Pacific* **132** (2020) 035001.
- [5] H. Aihara, C.A. Prieto, D. An, S.F. Anderson, É. Aubourg, E. Balbinot et al., The eighth data release of the sloan digital sky survey: first data from sdss-iii, *The Astrophysical Journal Supplement Series* **193** (2011) 29.
- [6] D.E.S. Collaboration:, T. Abbott, F. Abdalla, J. Aleksić, S. Allam, A. Amara et al., The dark energy survey: more than dark energy—an overview, *Monthly Notices of the Royal Astronomical Society* **460** (2016) 1270.
- [7] R. Laureijs, J. Amiaux, S. Arduini, J.-L. Augeres, J. Brinchmann, R. Cole et al., Euclid definition study report, *arXiv preprint arXiv:1110.3193* (2011) .

- [8] Ž. Ivezić, S.M. Kahn, J.A. Tyson, B. Abel, E. Acosta, R. Allsman et al., Lsst: from science drivers to reference design and anticipated data products, *The Astrophysical Journal* **873** (2019) 111.
- [9] M. Rezaie, A.J. Ross, H.-J. Seo, E.-M. Mueller, W.J. Percival, G. Merz et al., Primordial non-gaussianity from the completed sdss-iv extended baryon oscillation spectroscopic survey-i: Catalogue preparation and systematic mitigation, *Monthly Notices of the Royal Astronomical Society* **506** (2021) 3439.
- [10] E. Chaussidon, C. Yèche, N. Palanque-Delabrouille, A. de Mattia, A.D. Myers, M. Rezaie et al., Angular clustering properties of the desi qso target selection using dr9 legacy imaging surveys, *Monthly Notices of the Royal Astronomical Society* **509** (2022) 3904.
- [11] H. Johnston, A.H. Wright, B. Joachimi, M. Bilicki, N.E. Chisari, A. Dvornik et al., Organised randomness: Learning and correcting for systematic galaxy clustering patterns in kids using self-organising maps, *Astronomy & Astrophysics* **648** (2021) A98.
- [12] E.L. Wagoner, E. Rozo, X. Fang, M. Crocce, J. Elvin-Poole, N. Weaverdyck et al., Linear systematics mitigation in galaxy clustering in the dark energy survey year 1 data, *Monthly Notices of the Royal Astronomical Society* **503** (2021) 4349.
- [13] E.-M. Mueller, M. Rezaie, W.J. Percival, A.J. Ross, R. Ruggeri, H.-J. Seo et al., Primordial non-gaussianity from the completed sdss-iv extended baryon oscillation spectroscopic survey ii: measurements in fourier space with optimal weights, *Monthly Notices of the Royal Astronomical Society* **514** (2022) 3396.
- [14] M. Rezaie, A.J. Ross, H.-J. Seo, H. Kong, A. Porredon, L. Samushia et al., Local primordial non-gaussianity from the large-scale clustering of photometric desi luminous red galaxies, *arXiv preprint arXiv:2307.01753* (2023) .
- [15] A. Krolewski, W.J. Percival, S. Ferraro, E. Chaussidon, M. Rezaie, J.N. Aguilar et al., Constraining primordial non-gaussianity from desi quasar targets and planck cmb lensing, *arXiv preprint arXiv:2305.07650* (2023) .
- [16] A. De Mattia, V. Ruhlmann-Kleider, A. Raichoor, A.J. Ross, A. Tamone, C. Zhao et al., The completed sdss-iv extended baryon oscillation spectroscopic survey: measurement of the bao and growth rate of structure of the emission line galaxy sample from the anisotropic power spectrum between redshift 0.6 and 1.1, *Monthly Notices of the Royal Astronomical Society* **501** (2021) 5616.
- [17] O.H. Philcox, M.M. Ivanov, M. Simonović and M. Zaldarriaga, Combining full-shape and bao analyses of galaxy power spectra: a 1.6% cmb-independent constraint on h_0 , *Journal of Cosmology and Astroparticle Physics* **2020** (2020) 032.
- [18] Y. Kobayashi, T. Nishimichi, M. Takada and H. Miyatake, Full-shape cosmology analysis of the sdss-iii boss galaxy power spectrum using an emulator-based halo model: A 5% determination of σ_8 , *Physical Review D* **105** (2022) 083517.
- [19] A.J. Ross, F. Beutler, C.-H. Chuang, M. Pellejero-Ibanez, H.-J. Seo, M. Vargas-Magana et al., The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: Observational systematics and baryon acoustic oscillations in the correlation function, *Monthly Notices of the Royal Astronomical Society* **464** (2017) 1168.
- [20] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L.E. Allen et al., The DESI Experiment Part I: Science, Targeting, and Survey Design, *arXiv e-prints* (2016) arXiv:1611.00036 [[1611.00036](#)].
- [21] DESI Collaboration, B. Abareshi, J. Aguilar, S. Ahlen, S. Alam, D.M. Alexander et al., Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument, *Astronomical Journal* **164** (2022) 207 [[2205.10939](#)].
- [22] DESI Collaboration, A.G. Adame, J. Aguilar, S. Ahlen, S. Alam, G. Aldering et al., Validation of the Scientific Program for the Dark Energy Spectroscopic Instrument, *Astronomical Journal* **167** (2024) 62 [[2306.06307](#)].

- [23] DESI Collaboration, A.G. Adame, J. Aguilar, S. Ahlen, S. Alam, G. Aldering et al., The Early Data Release of the Dark Energy Spectroscopic Instrument, [arXiv e-prints \(2023\)](#) [arXiv:2306.06308](#) [2306.06308].
- [24] A. Adame, J. Aguilar, S. Ahlen, S. Alam, D. Alexander, M. Alvarez et al., Desi 2024 vi: Cosmological constraints from the measurements of baryon acoustic oscillations, [arXiv preprint arXiv:2404.03002](#) (2024) .
- [25] T.N. Miller, P. Doel, G. Gutierrez, R. Besuner, D. Brooks, G. Gallo et al., The Optical Corrector for the Dark Energy Spectroscopic Instrument, [arXiv e-prints \(2023\)](#) [arXiv:2306.06310](#) [2306.06310].
- [26] J.H. Silber, P. Fagrelus, K. Fanning, M. Schubnell, J.N. Aguilar, S. Ahlen et al., The Robotic Multiobject Focal Plane System of the Dark Energy Spectroscopic Instrument (DESI), [Astronomical Journal](#) **165** (2023) 9 [2205.09014].
- [27] M. Levi, C. Bebek, T. Beers, R. Blum, R. Cahn, D. Eisenstein et al., The DESI Experiment, a whitepaper for Snowmass 2013, [arXiv e-prints \(2013\)](#) arXiv:1308.0847 [1308.0847].
- [28] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L.E. Allen et al., The DESI Experiment Part II: Instrument Design, [arXiv e-prints \(2016\)](#) arXiv:1611.00037 [1611.00037].
- [29] A. Albrecht, G. Bernstein, R. Cahn, W.L. Freedman, J. Hewitt, W. Hu et al., Report of the dark energy task force, [arXiv preprint astro-ph/0609591](#) (2006) .
- [30] S. Alam, F.D. Albareti, C.A. Prieto, F. Anders, S.F. Anderson, T. Anderton et al., The eleventh and twelfth data releases of the sloan digital sky survey: final data from sdss-iii, [The Astrophysical Journal Supplement Series](#) **219** (2015) 12.
- [31] R. Ahumada, C.A. Prieto, A. Almeida, F. Anders, S.F. Anderson, B.H. Andrews et al., The 16th data release of the sloan digital sky surveys: first release from the apogee-2 southern survey and full release of eboss spectra, [The Astrophysical Journal Supplement Series](#) **249** (2020) 3.
- [32] M. Vargas-Magana, D.D. Brooks, M.M. Levi and G.G. Tarle, Unraveling the universe with desi, [arXiv preprint arXiv:1901.01581](#) (2019) .
- [33] A. Dey, D.J. Schlegel, D. Lang, R. Blum, K. Burleigh, X. Fan et al., Overview of the desi legacy imaging surveys, [The Astronomical Journal](#) **157** (2019) 168.
- [34] Schlegel et al.
- [35] J. Moustakas, D. Lang, A. Dey, S. Juneau, A. Meisner, A.D. Myers et al., Siena Galaxy Atlas 2020, [Astrophysical Journal, Supplement](#) **269** (2023) 3 [2307.04888].
- [36] H. Zou, X. Zhou, X. Fan, T. Zhang, Z. Zhou, J. Nie et al., Project overview of the beijing–arizona sky survey, [Publications of the Astronomical Society of the Pacific](#) **129** (2017) 064101.
- [37] R.D. Blum, K. Burleigh, A. Dey, D.J. Schlegel, A.M. Meisner, M. Levi et al., The decam legacy survey, in [American Astronomical Society Meeting Abstracts# 228](#), vol. 228, pp. 317–01, 2016.
- [38] D.R. Silva, R.D. Blum, L. Allen, A. Dey, D.J. Schlegel, D. Lang et al., The mayall z-band legacy survey, in [American Astronomical Society Meeting Abstracts# 228](#), vol. 228, pp. 317–02, 2016.
- [39] D. Lang, D.W. Hogg, K. Mierle, M. Blanton and S. Roweis, Astrometry. net: Blind astrometric calibration of arbitrary astronomical images, [The astronomical journal](#) **139** (2010) 1782.
- [40] D.W. Hogg and D. Lang, Replacing standard galaxy profiles with mixtures of gaussians, [Publications of the Astronomical Society of the Pacific](#) **125** (2013) 719.
- [41] D. Lang, A hybrid fourier–real gaussian mixture method for fast galaxy–psf convolution, [arXiv preprint arXiv:2012.15797](#) (2020) .

- [42] H. Kong, K.J. Burleigh, A. Ross, J. Moustakas, C.-H. Chuang, J. Comparat et al., Removing imaging systematics from galaxy clustering measurements with obiwan: application to the sdss-iv extended baryon oscillation spectroscopic survey emission-line galaxy sample, *Monthly Notices of the Royal Astronomical Society* **499** (2020) 3943.
- [43] A. Raichoor, A. De Mattia, A.J. Ross, C. Zhao, S. Alam, S. Avila et al., The completed sdss-iv extended baryon oscillation spectroscopic survey: large-scale structure catalogues and measurement of the isotropic bao between redshift 0.6 and 1.1 for the emission line galaxy sample, *Monthly Notices of the Royal Astronomical Society* **500** (2021) 3254.
- [44] R. Zhou, B. Dey, J.A. Newman, D.J. Eisenstein, K. Dawson, S. Bailey et al., Target selection and validation of desi luminous red galaxies, *The Astronomical Journal* **165** (2023) 58.
- [45] B. Flaugher, H.T. Diehl, K. Honscheid, T.M.C. Abbott, O. Alvarez, R. Angstadt et al., The Dark Energy Camera, *Astronomical Journal* **150** (2015) 150 [[1504.02900](#)].
- [46] A. Dey, D. Rabinowitz, A. Karcher, C. Bebek, C. Baltay, D. Sprayberry et al., Mosaic3: a red-sensitive upgrade for the prime focus camera at the Mayall 4m telescope, in Ground-based and Airborne Instrumentation for Astronomy VI, C.J. Evans, L. Simard and H. Takami, eds., vol. 9908 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, p. 99082C, Aug., 2016, [DOI](#).
- [47] G.G. Williams, E. Olszewski, M.P. Lesser and J.H. Burge, 90prime: a prime focus imager for the Steward Observatory 90-in. telescope, in Ground-based Instrumentation for Astronomy, A.F.M. Moorwood and M. Iye, eds., vol. 5492 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 787–798, Sept., 2004, [DOI](#).
- [48] E.L. Wright, P.R.M. Eisenhardt, A.K. Mainzer, M.E. Ressler, R.M. Cutri, T. Jarrett et al., The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance, *Astronomical Journal* **140** (2010) 1868 [[1008.0031](#)].
- [49] E. Chaussidon, C. Yèche, N. Palanque-Delabrouille, D.M. Alexander, J. Yang, S. Ahlen et al., Target selection and validation of desi quasars, *The Astrophysical Journal* **944** (2023) 107.
- [50] C. Hahn, M.J. Wilson, O. Ruiz-Macias, S. Cole, D.H. Weinberg, J. Moustakas et al., The desi bright galaxy survey: Final target selection, design, and validation, *The Astronomical Journal* **165** (2023) 253.
- [51] A. Raichoor, J. Moustakas, J.A. Newman, T. Karim, S. Ahlen, S. Alam et al., Target selection and validation of desi emission line galaxies, *The Astronomical Journal* **165** (2023) 126.
- [52] R. Zhou, J.A. Newman, K.S. Dawson, D.J. Eisenstein, D.D. Brooks, A. Dey et al., Preliminary Target Selection for the DESI Luminous Red Galaxy (LRG) Sample, *Research Notes of the American Astronomical Society* **4** (2020) 181 [[2010.11282](#)].
- [53] A.D. Myers, J. Moustakas, S. Bailey, B.A. Weaver, A.P. Cooper, J.E. Forero-Romero et al., The target-selection pipeline for the dark energy spectroscopic instrument, *The Astronomical Journal* **165** (2023) 50.
- [54] T.-W. Lan, R. Tojeiro, E. Armengaud, J.X. Prochaska, T.M. Davis, D.M. Alexander et al., The DESI Survey Validation: Results from Visual Inspection of Bright Galaxies, Luminous Red Galaxies, and Emission-L, *Astrophysical Journal* **943** (2023) 68 [[2208.08516](#)].
- [55] J. Guy, S. Bailey, A. Kremin, S. Alam, D.M. Alexander, C. Allende Prieto et al., The Spectroscopic Data Processing Pipeline for the Dark Energy Spectroscopic Instrument, *Astronomical Journal* **165** (2023) 144 [[2209.14482](#)].
- [56] Raichoor et al.
- [57] D. DePoy, T. Abbott, J. Annis, M. Antonik, M. Barceló, R. Bernstein et al., The dark energy camera (decam), in Ground-based and Airborne Instrumentation for Astronomy II, vol. 7014, pp. 190–198, SPIE, 2008.

- [58] H. Zou, X. Zhou, X. Fan, T. Zhang, Z. Zhou, J. Nie et al.,
Project Overview of the Beijing-Arizona Sky Survey, [Publications of the ASP](#) **129** (2017) 064101
[1702.03653].
- [59] D.R. Silva, R.D. Blum, L. Allen, A. Dey, D.J. Schlegel, D. Lang et al.,
The Mayall z-band Legacy Survey, in *American Astronomical Society Meeting Abstracts #228*, vol. 228
of *American Astronomical Society Meeting Abstracts*, p. 317.02, June, 2016.
- [60] E.L. Wright, P.R. Eisenhardt, A.K. Mainzer, M.E. Ressler, R.M. Cutri, T. Jarrett et al., The wide-field
infrared survey explorer (wise): mission description and initial on-orbit performance, *The Astronomical
Journal* **140** (2010) 1868.
- [61] D. Lang, unwise: unblurred coadds of the wise imaging, *The Astronomical Journal* **147** (2014) 108.
- [62] D. Lang, unWISE: Unblurred Coadds of the WISE Imaging, [Astronomical Journal](#) **147** (2014) 108
[1405.0308].
- [63] A.M. Meisner, D. Lang, E.F. Schlafly and D.J. Schlegel,
Full-sky unWISE Coadds at Seven Years’ Depth, [Research Notes of the American Astronomical Society](#)
5 (2021) 200.
- [64] K.M. Górski, E. Hivon, A.J. Banday, B.D. Wandelt, F.K. Hansen, M. Reinecke et al.,
HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere,
Astrophysical Journal **622** (2005) 759 [astro-ph/0409513].
- [65] H.S. Xavier, F.B. Abdalla and B. Joachimi, Improving lognormal models for cosmological fields,
Monthly Notices of the Royal Astronomical Society **459** (2016) 3693.
- [66] D.J. Schlegel, D.P. Finkbeiner and M. Davis, Maps of dust infrared emission for use in estimation of
reddening and cosmic microwave background radiation foregrounds, *The Astrophysical Journal* **500**
(1998) 525.
- [67] DESI Collaboration, DESI 2024 II: Two Point Clustering Measurements and Validation, in preparation
(2024) .
- [68] DESI Collaboration, DESI 2024 VIII: Constraints on Primordial Non-Gaussianities, in preparation
(2024) .
- [69] Y.-K. Chiang and B. Ménard, Extragalactic imprints in galactic dust maps, *The Astrophysical Journal*
870 (2019) 120.
- [70] N. Mudur, C.F. Park and D.P. Finkbeiner, Stellar-reddening-based extinction maps for cosmological
applications, *The Astrophysical Journal* **949** (2023) 47.
- [71] G.M. Green, E.F. Schlafly, D. Finkbeiner, H.-W. Rix, N. Martin, W. Burgett et al., Galactic reddening in
3d from stellar photometry—an improved map, *Monthly Notices of the Royal Astronomical Society* **478**
(2018) 651.
- [72] D. Lenz, B.S. Hensley and O. Doré, A new, large-scale map of interstellar reddening derived from h i
emission, *The Astrophysical Journal* **846** (2017) 38.
- [73] E.F. Schlafly and D.P. Finkbeiner, Measuring reddening with sloan digital sky survey stellar spectra and
recalibrating sfd, *The Astrophysical Journal* **737** (2011) 103.
- [74] E.S. Rykoff, D. Tucker, D. Burke, S. Allam, K. Bechtol, G. Bernstein et al., The dark energy survey
six-year calibration star catalog, *arXiv preprint arXiv:2305.01695* (2023) .
- [75] K.C. Chambers, E. Magnier, N. Metcalfe, H. Flewelling, M. Huber, C. Waters et al., The pan-starrs1
surveys, *arXiv preprint arXiv:1612.05560* (2016) .
- [76] M.F. Skrutskie, R.M. Cutri, R. Stiening, M.D. Weinberg, S. Schneider, J.M. Carpenter et al.,
The Two Micron All Sky Survey (2MASS), [Astronomical Journal](#) **131** (2006) 1163.
- [77] A. Vallenari, A.G. Brown, T. Prusti, J.H. De Bruijne, F. Arenou, C. Babusiaux et al., Gaia data release
3-summary of the content and survey properties, *Astronomy & Astrophysics* **674** (2023) A1.

- [78] Y.-K. Chiang, Corrected sfid: A more accurate galactic dust map with minimal extragalactic contamination, arXiv preprint arXiv:2306.03926 (2023) .
- [79] DESI Collaboration, DESI 2024 I: Data Release 1 of the Dark Energy Spectroscopic Instrument, in preparation (2024) .
- [80] E.F. Schlafly, D. Kirkby, D.J. Schlegel, A.D. Myers, A. Raichoor, K. Dawson et al., Survey Operations for the Dark Energy Spectroscopic Instrument, Astronomical Journal **166** (2023) 259 [2306.06309].
- [81] E.F. Schlafly, D.P. Finkbeiner, D.J. Schlegel, M. Jurić, Ž. Ivezić, R.R. Gibson et al., The blue tip of the stellar locus: Measuring reddening with the sloan digital sky survey, The Astrophysical Journal **725** (2010) 1175.
- [82] M.R. Blanton, E. Kazin, D. Muna, B.A. Weaver and A. Price-Whelan, Improved background subtraction for the sloan digital sky survey images, The Astronomical Journal **142** (2011) 31.
- [83] J. Zuntz, E. Sheldon, S. Samuroff, M.A. Troxel, M. Jarvis, N. MacCrann et al., Dark energy survey year 1 results: weak lensing shape catalogues, Monthly Notices of the Royal Astronomical Society **481** (2018) 1149.
- [84] M. Jarvis, E. Sheldon, J. Zuntz, T. Kacprzak, S. Bridle, A. Amara et al., The des science verification weak lensing shear catalogues, Monthly Notices of the Royal Astronomical Society **460** (2016) 2245.
- [85] E. Bertin, Automated morphometry with sextractor and psfex, in Astronomical data analysis software and systems xx, vol. 442, p. 435, 2011.
- [86] M. Jarvis, G. Bernstein, A. Amon, C. Davis, P. Léget, K. Bechtol et al., Dark energy survey year 3 results: point spread function modelling, Monthly Notices of the Royal Astronomical Society **501** (2021) 1282.
- [87] T. Abbott, M. Aguena, A. Alarcon, S. Allam, O. Alves, A. Amon et al., Dark energy survey year 3 results: Cosmological constraints from galaxy clustering and weak lensing, Physical Review D **105** (2022) 023520.
- [88] J. Elvin-Poole, N. MacCrann, S. Everett, J. Prat, E. Rykoff, J. De Vicente et al., Dark energy survey year 3 results: Magnification modeling and impact on cosmological constraints from galaxy clustering and galaxy-galaxy lensing, arXiv preprint arXiv:2209.09782 (2022) .
- [89] H.J. Newberg, B. Yanny, C. Rockosi, E.K. Grebel, H.-W. Rix, J. Brinkmann et al., The ghost of sagittarius and lumps in the halo of the milky way, The Astrophysical Journal **569** (2002) 245.
- [90] J. Domínguez-Gómez, I. Pérez, T. Ruiz-Lara, R.F. Peletier, P. Sánchez-Blázquez, U. Lisenfeld et al., Galaxies in voids assemble their stars slowly, Nature **619** (2023) 269.
- [91] M. White, R. Zhou, J. DeRose, S. Ferraro, S.-F. Chen, N. Kokron et al., Cosmological constraints from the tomographic cross-correlation of desi luminous red galaxies and planck cmb lensing, Journal of Cosmology and Astroparticle Physics **2022** (2022) 007.
- [92] M. Rezaie, A.J. Ross, H.-J. Seo, H. Kong, A. Porredon, L. Samushia et al., Local primordial non-gaussianity from the large-scale clustering of photometric desi luminous red galaxies, Monthly Notices of the Royal Astronomical Society (2024) stae886.
- [93] DESI Collaboration, DESI 2024 III: Baryon Acoustic Oscillations from Galaxies and Quasars, in preparation (2024) .
- [94] DESI Collaboration, DESI 2024 V: Fullshape from Galaxies and Quasars, in preparation (2024) .
- [95] DESI Collaboration, DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations, in preparation (2024) .
- [96] DESI Collaboration, DESI 2024 VII: Cosmological Constraints from the Fullshape Measurements, in preparation (2024) .

Appendix

A Computational Resources

A.1 Computational resources for ‘blobs’

Source measurements are segmented into units of ‘blobs’. The computational resources for each ‘blob’ increase sharply with the area and number of sources in each ‘blob’. ‘Blobs’ containing bright stars and galaxies often correspond to a large blob that contains many sources and is computationally expensive to process. Given the areas will not later be used for analysis, the Obiwan pipeline is made more efficient by ignoring them. The `LEGACYPIPE` code produces a bitmask map for each brick; we exclude areas that are near optically bright and “medium-bright” stars, WISE *W1*-bright stars, large galaxies, and globular clusters. These correspond to the “maskbits” data product bits 1, 11, 8, 12, and 13, respectively.

A.2 Cosmos Deep

Because of the large number of exposures used in the COSMOS Deep production run, the computational time used to process these data in the same footprint is much higher than that in the DR9 production run. For example, brick 1503p015 uses 113 hours of wall time in the COSMOS Deep production run on one CPU node on NERSC’s Cori machine, while in the DR9 production run it uses 8 hours of wall time on one node. The maximum numbers of overlapping exposures used in the DR9 production run in this region was 6, 7, 7 for g, r, z bands, respectively. Meanwhile, the same numbers for the COSMOS deep production run are 45, 52, 120. This results in a much deeper depth in the three optical bands: The maximum `psfdepth` for the DR9 production run in this region are 25.19, 24.65, and 23.80 for g, r and z , while the median depth are 22.99, 23.94, 24.47 for the same bands. The corresponding numbers for the COSMOS Deep production run are 26.53, 26.27, 25.73 for maximum depth, and 25.31, 25.53, 24.79 for median depth.

A.3 Obiwan runs

Figure 17 shows the time consumed for each brick. Obiwan runs on Permuter nodes at NERSC. With 128 threads on a single node, one node can process 8 bricks in parallel. On average it takes 11 minutes to finish one brick with 16 threads. Despite the fast speed of each brick, the pipeline is still inefficient. The problem is that the parallelization is limited by the memory (RAM) of the CPU available during the simulation. To read all the images needed for processing, we can only process a maximum of 8 bricks at a time. This means that with 16 threads available for computation, only 1 thread is active most of the time. The parallelization happens in the source fitting stage. As Obiwan fits many fewer sources than a normal reduction run, the efficiency on a per-source basis is therefore lower compared to the full image processing.

A.4 Methods for Increasing the Efficiency of Synthetic Galaxy Production

The current procedure of Obiwan is computationally expensive. However, this is still feasible with image simulations significantly faster than Obiwan. One idea that allows significant speed up in estimating and mitigating galaxies at one location repeatedly. This would avoid repeating time-consuming stages like reading images, detecting sources, etc, and focus on the stage of source fitting. Furthermore, the procedure of source fitting can be accelerated by operating the code on GPUs, as it involves intensive utilization of matrix operations.

Another practical way to address this problem is to develop a Monte Carlo procedure to estimate how galaxies are scattered under certain conditions analytically. This procedure numerically estimates

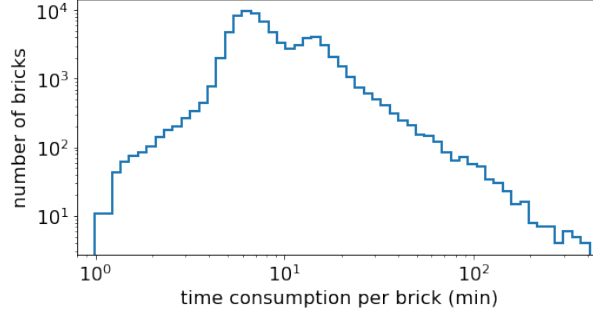


Figure 17: Obiwan simulation time consumption for bricks in the South Galactic Cap (SGC). One brick spans an area of 0.0625 square degrees. The median time to process one brick in the SGC is 7.56 minutes, and the average time consumed per brick is 10.89 minutes.

how galaxy magnitudes are altered by estimating errors generated by each survey property. It is much faster than image simulation. Such a method would need to be calibrated on image simulation to validate the accuracy of its numerical estimations. The Obiwan results we have presented could be used for such calibration.

B Extra Details on Cosmos Deep

B.1 cut on z -band depth

The depth of the COSMOS Deep Catalog allows us to use these data as a “truth” catalog, and our injected targets are sampled from it. However, even within the COSMOS Deep Catalog, there are significant variations in the depth and we first determine a minimum depth cut for the truth sample. The Legacypipe works in such a way that the complexity of the light profile chosen to obtain the photometry of a given source depends highly on the depth, based on simple criteria on the χ^2 improvement required to allow a more complex profile (shown in equation 5.1). We must make a trade-off in the truth catalog between its precision in shape, and its total galaxy counts. The “truth input” needs to be accurate enough in flux so that the output color distribution in simulated galaxies matches the real galaxies. Section 7.1 explains that the noise within the truth input leads to a mismatch in magnitude between real and synthetic LRGs. Figure 18 shows how these two variables compete given different depth cuts. We use the `galdepth` metric produced by Legacypipe, which gives the sensitivity to a round exponential-profile galaxy with a half-light radius of 0.45". We determine the depth cut by checking the percentage of Sérsic galaxies in the sample. A higher fraction of Sérsic galaxies in the sample means that the galaxies are better resolved in morphology. The determination of the type of a galaxy in Legacypipe has a prior that favors a simpler type (further explained in Section 5.1). Because of this prior, the images have to be clear enough for the pipeline to choose the type Sérsic for the galaxies in the images. By selecting regions with a higher fraction of Sérsic galaxies, we have a higher confidence in the truth sample’s LRG morphology, and the sample’s photometry will also be better. Based on the observed distribution of Sérsic fraction, we selected all data with `galdepth_z > 24.5` into our truth sample.

B.2 The WISE band in COSMOS Deep catalog

We test the difference in the WISE flux measurements in the presence of different image sets in the optical band. We test it on 3 sets: the official DR9 release, the COSMOS Deep set, and the COSMOS

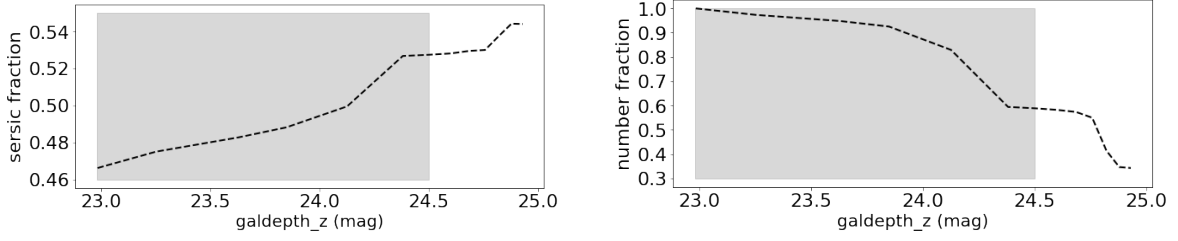


Figure 18: Depth cut on the COSMOS Deep catalog. The left plot shows the fraction of Sérsic galaxies when using different z -band galaxy depth (`galdepth_z`) cuts. The right plot shows the fraction of the remaining COSMOS deep catalog after the `galdepth_z` cut. The increase of the `galdepth_z` cutting limit will yield a higher fraction of Sérsic galaxies in the LRG SV3 sample, and at the same time, there will be a decrease in the number of the total sample selected. We select a `galdepth_z` cut of 24.5 as a trade-off between these two considerations.

Repeats set. We compare the WISE flux on the same sources measured with different optical images. They are shown in Figure 19. Figure 19 (left) suggests that there are not many false assignments of W1 flux to nearby sources since most sources in the two catalogs have similar W1 flux. However, Figure 19 (right) shows that the WISE flux dispersion between the COSMOS Deep and DR9 catalog is even bigger than between the COSMOS Repeats with the smallest and largest seeings (sets 0 and 9, respectively). One possible explanation is that the input catalogs from the optical bands have different number densities. The total number of sources detected mainly depends on `psfdepth`, as the extra faint sources detected are mostly point sources. For COSMOS repeat sets, despite having consistent `galdepth`, their `psfdepth` still increases with the set number. However, the difference between set 0 and set 9 is much smaller than the difference between DR9 and COSMOS deep. Thus, the difference we see for WISE flux in different sets is likely contributed by the difference in the number of sources given in the optical band. Potential errors in WISE PSF or morphological measurements in optical bands could induce uncertainty in flux W1 measurement, especially when the number of sources for "forced photometry" increases. However, no clear way exists to test which version has better precision for WISE flux measurement. We use the WISE flux from the COSMOS Deep measurement for consistency.

C De-biasing procedure for Cosmos Deep catalog

We compute the `coadd observed` image and the `coadd inverse_variance` image, as well as the average PSF (point spread function) image of the COSMOS Deep data in each brick. Then we replace the `coadd observed` image with Gaussian noise based on pixel values from the `inverse_variance` image. Next, we add one simulated COSMOS Deep source to its original pixel location, with the shape and flux information taken from the COSMOS Deep dataset. We define a blob that has this source centered in the middle of the image, with a size of 128×128 pixels. With these modifications, it is much faster for the pipeline to locate the designated blob, and fit the one galaxy that we injected. We then fit this blob using the same process in the `Legacypipe`, and `Legacypipe` returns the measured shape information. After we obtain the shape measured from `Legacypipe`, we compare it with the input shape. When they don't match, we modify the input shape and run the pipeline for more time. We repeat this process until the output shape reaches the desired precision. Although there is only a prior on ellipticity in equation 5.1, we find that the parameters of galaxy radius and Sérsic index are also biased when we check the shape distribution of Obiwan galaxies that directly samples from the

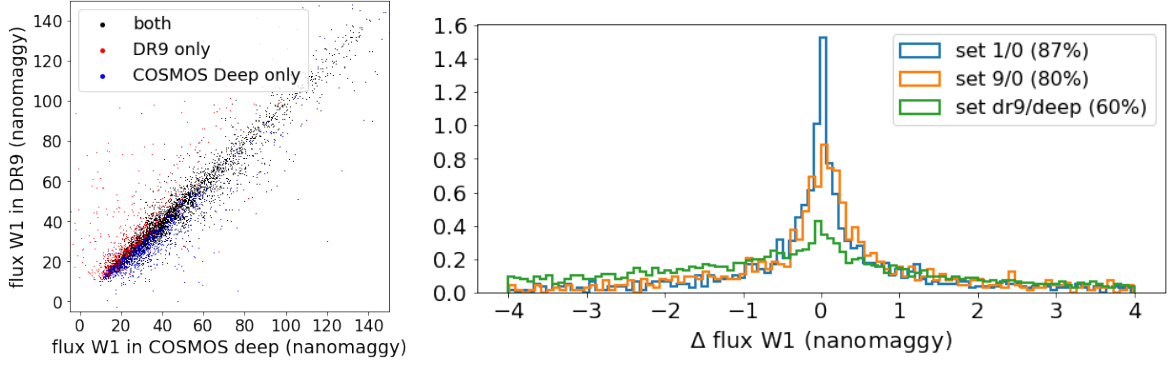


Figure 19: (Left) Flux W1 for sources that pass LRG sv3 color cut on either DR9 catalogs or COSMOS deep catalogs. The black dots are sources passing LRG sv3 color cut on both sets, the red dots are sources passing LRG sv3 color cut only on the DR9 catalog. The blue dots are sources passing LRG sv3 color cut only on the COSMOS deep catalog. We see that most sources have similar W1 flux, and their differences in flux W1 value are likely due to photometric scattering. (right) Histogram of flux W1 difference distribution between two sets. Sources shown in this plot pass LRG sv3 color cut on either set. Blue is the difference between COSMOS repeat set 1 and 0. 87% of the sources are within the range of in this plot. Orange is the difference between COSMOS set 9 and 0. 80% of the sources are within the range of in this plot. Green is the difference between the DR9 catalog and the COSMOS deep catalog. 60% are within the range of this plot. This plot suggests that flux W1 measured in COSMOS deep has a larger difference with DR9 catalog, compared with the difference that we see in between COSMOS repeat sets.

COSMOS Deep catalog. It is likely because parameters for galaxy shapes are correlated with each other. We define a quantity \mathbf{A} that includes 3 shape parameters: e (ellipticity), $shape_r$ (half-light radius), and $seraic_n$ (Sérsic index). It is defined by

$$\mathbf{A}_i = (\text{seraic_n}_i^O - \text{seraic_n}_s)^2/5.5 + (e_i^O - e_s)^2 + (\text{shape_r}_i^O - \text{shape_r}_s)^2 \quad (\text{C.1})$$

The upper index O here denotes the output value returned by Legacypipe. The lower index i denotes the number of iterations. The lower index s denotes values taken from the COSMOS Deep dataset, and it is also the starting value of the iteration. Our goal is to have a similar shape distribution between the output Obiwan LRG catalog and the DESI LRG catalog. To achieve this, we require that the shape parameter returned by Legacypipe is similar to the ones measured by the COSMOS Deep set. When the condition is achieved, we record the input shape parameters to use them as input for injected galaxies.

When the \mathbf{A} value does not reach our desired accuracy, we modify the parameters. Each time the parameters are modified by:

$$X_{i+1}^I = X_i^I - (X_i^O - X_s) * \text{step} \quad (\text{C.2})$$

The X here denotes one of the parameters from $seraic_n$, e or $shape_r$. The upper index "I" denotes the input value from i th iteration. The step is a constant number used to determine how much the new parameter is modified. The three parameters are all updated during each iteration. We added

some constraints on the ellipticity e . We force galaxies with $e = 0$ (round galaxies or point source) in the deep photometry to remain $e = 0$. Galaxies with non-zero ellipticity preserve their orientation.

These new set of parameters are accepted if

$$\mathbf{A}_{i+1} < \mathbf{A}_i \quad (\text{C.3})$$

Otherwise, we make the *step* size smaller and try again, with a maximum of 3 times per trial. The *step* value is defined as 0.2, 0.1, 0.05 for 0 fail, 1 fail, and 2 fails. The *step* value is chosen so that the code can reach the correct value quickly and precisely.

There are two criteria to stop the iteration. First, when the equation fails 3 times, it means that the current value is the smallest \mathbf{A} we can get. Second, If $\mathbf{A} < 10^{-4}$, it means that the error is small enough. On either occasion, we stop the iteration.

There are many advantages to a procedure like this. First, it makes debiasing computationally possible by working on the coadd images instead of individual exposures. A COSMOS Deep brick takes more than 40 hours to finish on NERSC’s one Cori node. Since we need to fit on the same spot for each source involved iteratively, it would take even more time than a COSMOS Deep run. Second, it automatically takes into account the bias from blending, without actually being contaminated by nearby sources: The blended region is more noisy and has a lower value on *inverse_variance* image. Also, when blending exists in a pixel, it would be brighter and has a higher Poisson noise.

We test our shape-modified set on the COSMOS Repeats set. We inject these shape-modified galaxies into COSMOS Repeats images and recover the output shape with *Obiwan*. We compare the shape distribution of our simulated LRGs with the real LRGs in the same set and perform the Kolmogorov–Smirnov (KS) test over the distribution; the results are shown in Figure 20. One can observe that based on the p -values of the KS test, the distribution of the shape-modified catalog (dot-dashed) is much more similar to the DESI LRGs when compared with the un-modified version (star-solid). The p -values of the un-modified data are near zero, thus clearly rejecting the possibility that the un-modified data are representative of DESI LRGs. Conversely, the p -values of the modified data increase in all cases and are acceptable in most, and thus the debiasing process yields considerable improvement in the output shape distribution of *Obiwan* LRGs.

D Measurement bias

Overall, the flux measurement is negatively biased. Figure 21 shows the measurement flux bias as a function of input magnitude in each band. The dashed line in the middle is the median value in each flux bin. It shows that the bias is consistent across all magnitude bins, and it exists in all optical and infrared bands. To understand whether these features are a result of pipeline structures, or some features in the images, we designed four parallel runs: On the first run, we first generate synthetic Gaussian background noise that is much smaller than a typical Legacy Survey image. We inject galaxies into these images. In this setting, there are very few galaxies in the image, and no blending would occur here. The image is also almost noiseless. The second run is similar to the first run. The difference is that we set the Gaussian noise to the level of a typical Legacy Survey Image. In the third setting, we take the images on the second run and add the real galaxies back with their model parameters recorded in the DR9 data release. With this setting, we have a fully simulated image without any contamination. Blending may occur in such images, but it is free of error from background subtraction, cosmic ray contamination, etc. In the final setting, we use the real images and add simulated galaxies to these images, which is the same way as our previous production runs. The 4 production runs use the same input galaxy setting on the same set of images. In total, 442 bricks are processed for this comparison. Table 1 shows the results of these production runs. The

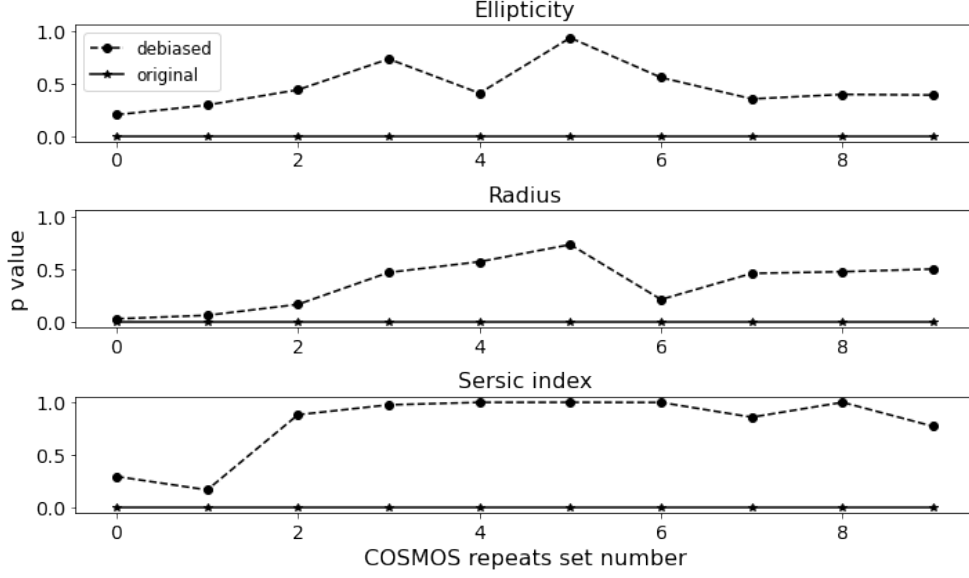


Figure 20: Histogram comparison between the simulated LRGs and the real LRGs in the same COSMOS Repeats footprint. This plot's dot and star curves represent different types of simulated galaxies. The dot-dashed curve uses input seeds from the de-biased sample, and the star solid curve uses input seeds from the un-modified sample. The x-axis denotes the ten sets of the COSMOS Repeats numbered 0 to 9, and the y-axis uses a p-value to indicate how close the histograms are between the simulated and the real galaxy shape distribution, using the KS test, 1 being the most similar and 0 being the least similar. These plots compare the shape parameters of ellipticity, Sérsic index, and the half-light radius.

results show that flux bias does not appear on these non-real images. This means that there is no algorithmic negative bias in the pipeline. The bias from real images is caused by other effects that can not be easily simulated. Fortunately, the bias is small enough that it has little effect on how we select galaxies as LRGs. If we correct the bias in our simulated outputs as a function of flux:

$$flux_{g_{corrected}} = 1.0123 * flux_{g_{measured}} + 0.0087 \quad (D.1)$$

$$flux_{r_{corrected}} = 1.0112 * flux_{r_{measured}} + 0.0100 \quad (D.2)$$

$$flux_{z_{corrected}} = 1.0166 * flux_{z_{measured}} + 0.0276 \quad (D.3)$$

$$flux_{W1_{corrected}} = 1.0070 * flux_{W1_{measured}} + 0.2483 \quad (D.4)$$

The population of galaxies that pass the target selection cuts with the new flux values only changes by 0.77%. Therefore, this bias has a negligible effect on determining the target selection of LRGs.

E Testing impact of psfsize with COSMOS repeats

E.1 Origination of trend with psfsize

The COSMOS Repeats data (described in Section 2.2) is suitable for this analysis because of its shallow and consistent galdepth and varying seeing (psfsize). We use COSMOS Deep as truth data for COSMOS Repeats as COSMOS Deep fully covers the region of COSMOS Repeats.

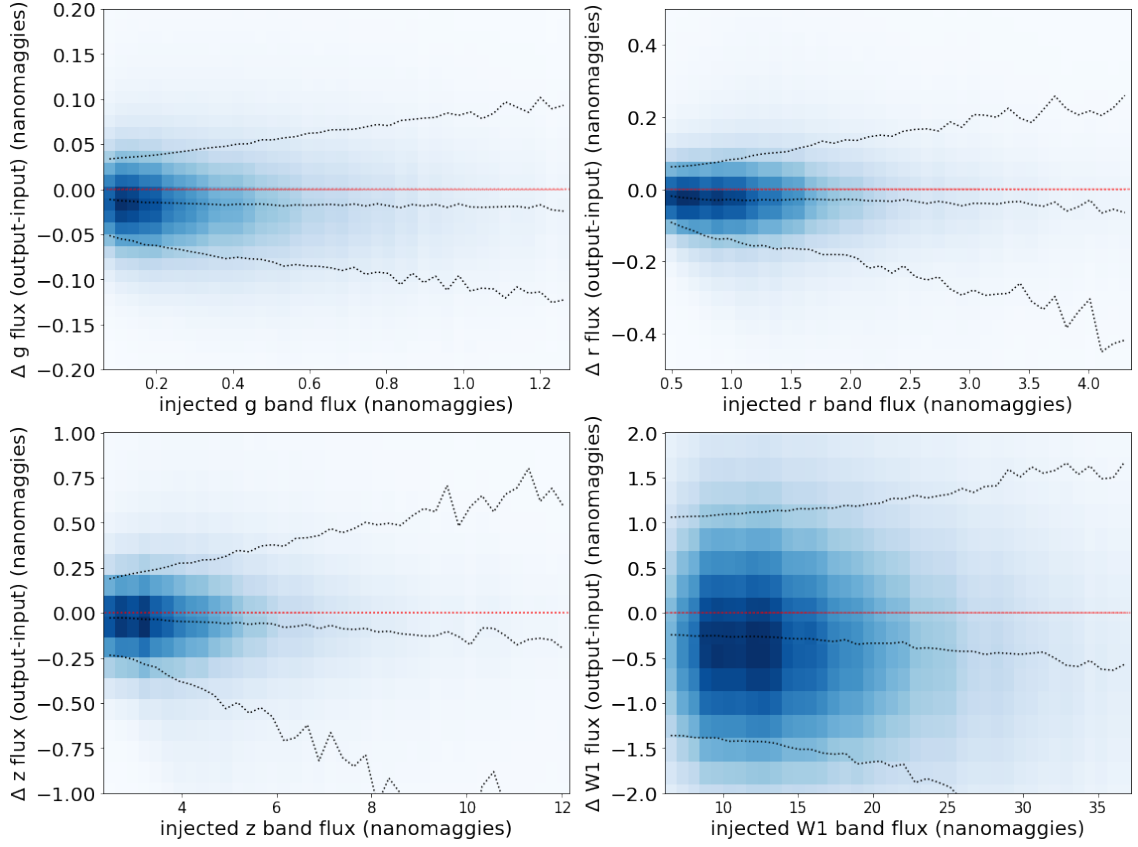


Figure 21: The density (blue) plot of output flux minus input flux. The black dashed line is the 84th, 50th, and 16th percentile for a given input magnitude. The red dashed line is centered at 0 on the y-axis. The difference between the red line and the 50th-percentile line shows that the flux measurement is consistently biased at all input flux levels in all bands. The four plots show results in the g , r , z , $W1$ band, respectively.

images median value	negligible noise	regular noise	Noise + model galaxies	Real images
$\Delta flux_g$ (nanomaggies)	0.005	0.013	0.017	-0.014
$\Delta flux_r$ (nanomaggies)	0.002	0.017	0.024	-0.03
$\Delta flux_z$ (nanomaggies)	-0.009	0.057	0.062	-0.063

Table 1: Median flux difference (output-input) for 4 production runs. In the first column, simulated galaxies are added to images containing tiny Gaussian noise (standard deviation of 10^{-4} nanomaggies per pixel). In the second column, simulated galaxies are added to images with Gaussian noise expected in the weight image. In the third column, galaxies are added to images with regular-level Gaussian noise, plus idealized model galaxy images for the DR9 catalog. In the fourth column, simulated galaxies are added to real DR9 images, as in our main Obiwan runs. In the real images, the fluxes are consistently biased low. This does not happen in other image sets. For the "negligible noise" set, there is approximately no bias present. In the "regular noise" and "Noise + model galaxies" sets, the flux values are biased high.

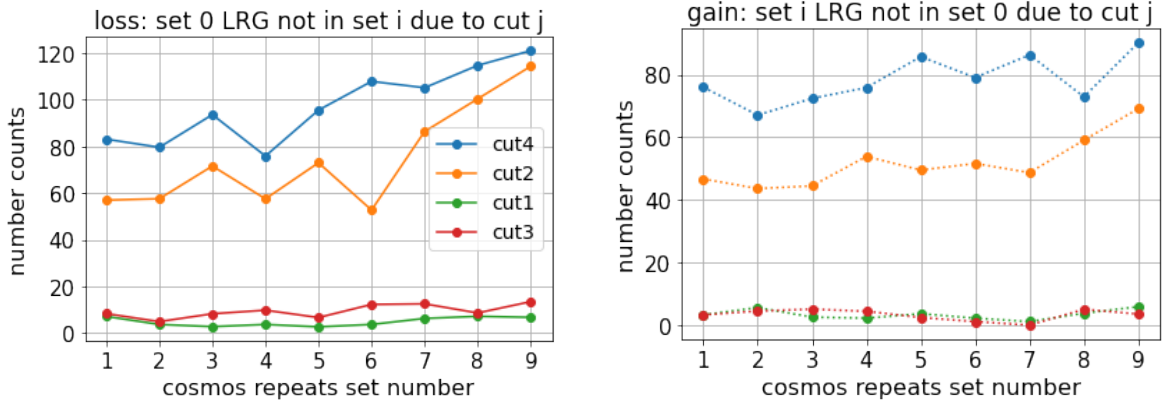


Figure 22: Number counts of LRG/LRG-like sources scattering outside or inside each color cut boundary. The left plot shows the number counts of LRGs in COSMOS Repeat set 0, but not in COSMOS Repeat set i (i is from 1 to 9), and does not pass the color cut j (j is from 1 to 4). We call these LRGs “lost” with respect to set i . The right plot shows the number counts of LRGs in COSMOS Repeat set i , but not in COSMOS Repeat set 0, and does not pass the color cut j . We call these LRGs “gained” with respect to set i . The cuts 1–4 are the same ordering as in equations 4.1–4.4. Two of the four cuts dominate the photometric scattering. Cut 2 is the “faint limit” cut on z band fiber magnitude. Cut 4 is the “luminosity cut” to select LRGs at higher redshift [44]. LRGs around these two cuts tend to have low magnitudes, so there is more photometric scattering around these two cuts.

We first test which LRG color cuts have large scattering among the 4 equations 4.1–4.4. In Figure 22, We use COSMOS Repeats set 0 as a reference set. For any set i in sets 1 through 9, we compute the number counts of “Loss” and “Gain”. “Loss” is defined as sources passing LRG color cut in reference set 0, but not in set i . We consider them “lost LRGs” in set i . For each color cut, we compute the number counts of “lost LRGs” that do not pass this cut. “Gain” is the opposite of “Loss”: They are the LRGs that pass the LRG color cut in set i , but not in set 0. We consider set i “gained” these LRGs. We check the number counts of these “gained LRGs” that do not pass a certain color cut in set 0. Cuts with high number counts in both the “loss” and “gain” plots are the dominant power to produce photometric scattering and, thus, potential density fluctuations. Here, we see that cut 2 (equation 4.2) and cut 4 (equation 4.4) have this dominant power in producing density fluctuation. Thus, we focus on these two cuts in the following analysis.

Density variation appears when number counts in “loss” and “gain” deviate from each other. In Figure 23, we see that the COSMOS Repeats show such deviation starting from COSMOS Repeat set 6. However, Obiwan LRGs simulated in the same setting do not show a clear deviation in all sets. This is a sign that some unknown factor not simulated by Obiwan is driving the deviation we see in the real COSMOS Repeats sets. As cut 2 (equation 4.2) is a simple cut on fiber magnitude z , it is easier to investigate. It is also possible to investigate cut 4 (equation 4.4), but due to its complicated functional form, we focus on testing cut 2 in the following paragraphs.

We first looked at the histogram of fiber mag z distribution of LRGs in a counterpart set that failed this color cut. The distribution looks like normal photometric scattering: Their values are close to the fiber mag z cut of 21.7. Thus, we suspect that the seeing (psfsize) is making the sample consistently fainter or brighter. Figure 24 validates our assumption. We match **all** sources in COSMOS Repeats set 9 to the COSMOS Deep data. We find that at fiber z magnitude < 22 in

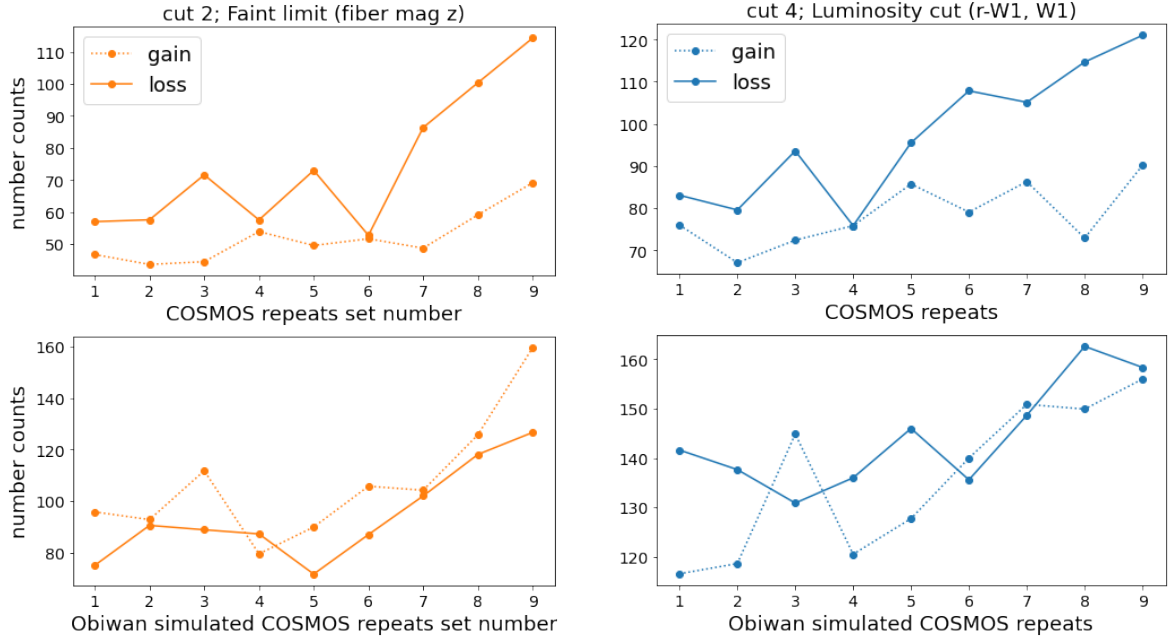


Figure 23: Comparison for LRGs “gained” and “lost” (described in Figure 22) for cut 2 (left, equation 4.2) and cut 4 (right, equation 4.4). The upper plot shows results from the original COSMOS Repeats sets. The lower plots show the result from Obiwan simulations with the same setting. The “gain” and “loss” curves in Obiwan simulations do not deviate from each other. Meanwhile, LRGs in the original COSMOS Repeats sets show deviation between “gain” and “loss” starting from set 6 and beyond. The result suggests signs of LRG density variation from seeing (psfsize) that is not simulated by Obiwan. We see a net loss of LRGs at high seeing that is not simulated by Obiwan, which is consistent with the trend seen in Figure 14, where we see a downward trend of LRGs as a function of psfsize.

COSMOS Deep, the flux in COSMOS Repeats set 9 is statistically lower. For sources around fiber z magnitude = 21.7, the scattering range is about fiber z magnitude around 21.4 to 21.9. We checked the median fiber z flux difference between each COSMOS Repeats set and COSMOS Deep. Figure 25 shows the median Δ flux value in all sets. There is a clear downward trend from set 0 to set 9, meaning that the higher seeing is causing the measured fiber z flux around 21.4 to 21.9 to appear fainter. Notably, this trend is greatly amplified when we exclusively choose more extended sources. When we look at sources with Sérsic index ≥ 4 , we find this trend to be much larger. According to COSMOS deep data, 43% of DESI LRGs have Sérsic index ≥ 4 , so this does have a strong impact. This process effectively shifts cut 2 (equation 4.2) to a higher threshold for cutting the true fiber flux value, as the observed flux is statistically lower than the true flux. The more “loss” of LRGs in high-seeing regions is a result of this shift. As it excludes a small fraction of LRGs close to the fiber z magnitude = 21.7 color cut.

A potential shift could also appear round cut 4 (equation 4.4). We can not test it due to the complexity of the color-cut equation. COSMOS repeat set is a small dataset and it is difficult to use it to test a color cut that involves two bands (r -band and $W1$ -band). However, we do see that LRGs close to the fiber z magnitude = 21.7 boundary are also close to the boundary in cut 4 (equation 4.4). Thus, potential density variations resulting from shifts around cut 4 (equation 4.4) could also introduce variation of LRGs around mag z = 21 at different seeing as shown in Figure 15.

We conclude that we find signs of PSF errors not simulated by Obiwan. This error shifts the

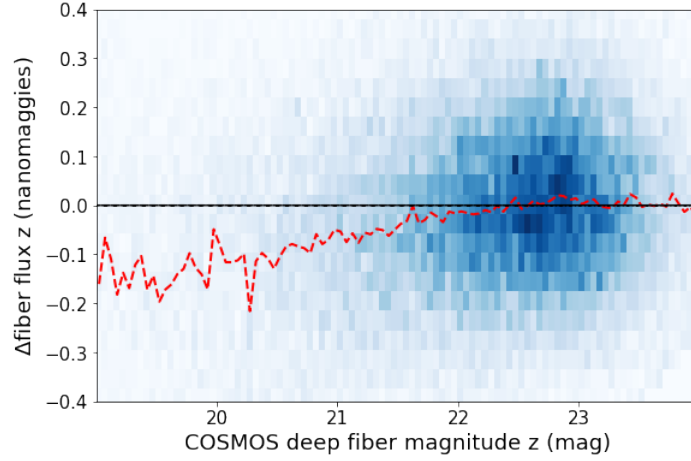


Figure 24: Flux difference between **all** sources measured in the COSMOS Repeats set 9 catalog, and sources measured in the COSMOS Deep catalog, as a function of fiber z -band magnitude in the COSMOS Deep catalog. The blue patches are the density of sources. The red line is the median value of Δ fiber flux as a function of z -band fiber magnitude in COSMOS Deep. Sources in the COSMOS Repeats set 9 catalog are statistically fainter than their counterparts in the COSMOS Deep catalog below fiber z magnitude 22.

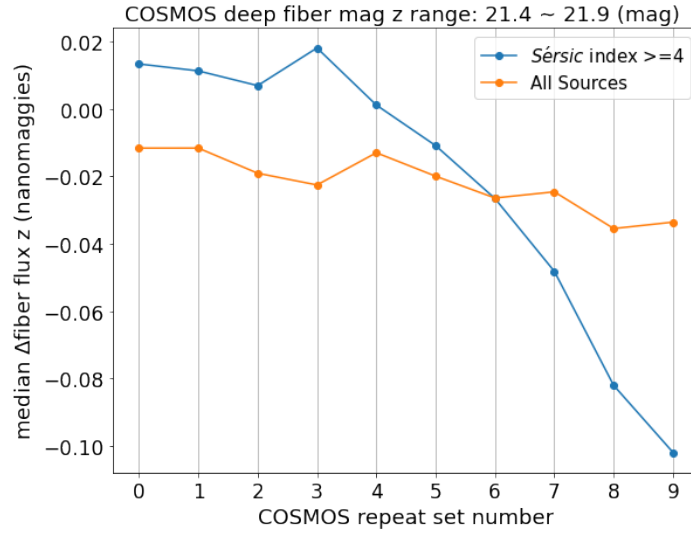


Figure 25: Median Δ fiber flux between the COSMOS Repeats set i (i from 1 to 9) catalogs and the COSMOS Deep catalog, with the COSMOS Deep catalog having a fiber z magnitude range of 21.4 to 21.9. Orange is for all sources, while blue is for sources with Sérsic index ≥ 4 . We chose this range because this region is sensitive to LRG color cut 4.2. With increased seeing, flux measured in the COSMOS Repeats set is statistically fainter. The trend is prominent for more extended sources. This could be an explanation for fewer LRGs observed in regions with worse seeing.

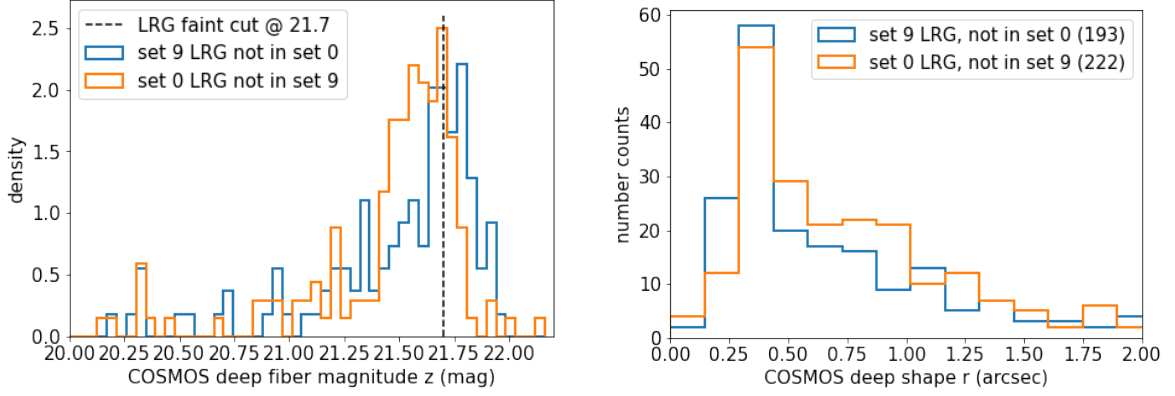


Figure 26: Histogram of fiber magnitude z (left) and half-light radius of the galaxy model for galaxies (right) in the COSMOS Deep catalog for LRGs in COSMOS Repeats set 9 but not in set 0 (blue), and LRGs in COSMOS Repeats set 0 but not in set 9 (orange). The numbers in the legend of the right plot denote the total count of sources plotted in each histogram. The black dashed line in the left plot is the LRG color cut of fiber magnitude < 21.7 . 31% of LRGs in set 9 but not in set 0 are outside the LRG faint cut, while 20% of LRGs in set 0 but not in set 9 are outside the LRG faint cut. They suggest that LRGs gained in a better seeing are statistically brighter. The right plot shows that LRGs gained in a better seeing have a larger half-light radius, which suggests that LRGs gained in a better seeing have more extended profiles.

statistical distribution of sources to a slightly fainter flux in regions of worse seeing. This shift induces a loss in the number density of LRGs. It could also be the underlying reason for the density dispersion of LRGs around mag $z = 21$ at different seeing that we see in Figure 15.

E.2 Morphological change of LRGs with `psfsize`

Two phenomena are going on in Figure 15. The first phenomenon is that with an increase of `psfsize` z , there is an increase in the LRG density at the faintest magnitudes, $z > 21.3$. This phenomenon is replicated in Obiwan LRGs. This phenomenon is likely because there is more photometric scattering at high `psfsize` z as these regions typically have lower depth. Sources in this mag $z > 21.3$ regions with the fiber mag $z < 21.7$ cuts are compact sources. According to results from Obiwan, their true color is more likely to be outside LRG color selection boundaries. The second phenomenon is described in Section E.1, which causes variation at around mag $z = 21$. The sources lost due to this phenomenon likely have a more extended shape because they have a more significant difference between mag z and fiber mag z .

The sources gained in the first phenomenon and the sources lost in the second phenomenon are not likely to have a large overlap because of their difference in morphology. Figure 26 shows the COSMOS Deep fiber magnitude and half-light radius of LRGs selected **only** in COSMOS Repeats set 0 or set 9. Set 0 LRGs are statistically brighter and have a more extended profile than set 9 LRGs. This difference validates our claim that LRGs have a slightly different population at different seeing. For DESI LRGs, we may not see a large density variation in seeing as the “gained” and “lost” in these two phenomena add up. However, it could pose potential risks in cosmological analysis as the two populations of LRGs have different sample statistics. Compact galaxies have less star-forming processing going on than extended galaxies. Even if they are similar in photometric color, they may differ in stellar mass, which results in a difference in galaxy bias.