

RECEIVED: April 26, 2024

REVISED: August 14, 2024

ACCEPTED: August 16, 2024

PUBLISHED: January 30, 2025

DARK ENERGY SPECTROSCOPIC INSTRUMENT (DESI) SURVEY YEAR 1 RESULTS

Production of alternate realizations of DESI fiber assignment for unbiased clustering measurement in data and simulations



DARK ENERGY
SPECTROSCOPIC
INSTRUMENT

U.S. Department of Energy Office of Science

J. Lasker et al.

Full author list at the end of the paper

E-mail: james.lasker3@gmail.com

ABSTRACT: A critical requirement of spectroscopic large scale structure analyses is correcting for selection of which galaxies to observe from an isotropic target list. This selection is often limited by the hardware used to perform the survey which will impose angular constraints of simultaneously observable targets, requiring multiple passes to observe all of them. In SDSS this manifested solely as the collision of physical fibers and plugs placed in plates. In DESI, there is the additional constraint of the robotic positioner which controls each fiber being limited to a finite patrol radius. A number of approximate methods have previously been proposed to correct the galaxy clustering statistics for these effects, but these generally fail on small scales. To accurately correct the clustering we need to upweight pairs of galaxies based on the inverse probability that those pairs would be observed (Bianchi & Percival 2017). This paper details an implementation of that method to correct the Dark Energy Spectroscopic Instrument (DESI) survey for incompleteness. To calculate the required probabilities, we need a set of alternate realizations of DESI where we vary the relative priority of otherwise identical targets. These realizations take the form of alternate Merged Target Ledgers (AMTL), the files that link DESI observations and targets. We present the method used to generate these alternate realizations and how they are tracked forward in time using the real observational record and hardware status, propagating the survey as though the alternate orderings had been adopted. We detail the first applications of this method to the DESI One-Percent Survey (SV3) and the DESI year 1 data. We include evaluations of the pipeline outputs, estimation of survey completeness from this and other methods, and validation of the method using mock galaxy catalogs.

KEYWORDS: galaxy surveys, galaxy clustering, cosmological simulations, dark matter simulations

ARXIV EPRINT: [2404.03006](https://arxiv.org/abs/2404.03006)

Contents

1	Introduction	1
2	Survey	4
3	Mocks	6
4	Alternate Merged Target Ledgers	7
4.1	Pairwise Inverse Probability weighting	7
4.2	The Merged Target Ledgers	8
4.3	Generating the alternate Merged Target Ledgers	9
4.4	The observation loop	11
4.5	Outputs of the pipeline	13
4.6	Validation of the pipeline	13
5	Clustering measurements	14
5.1	Weights	14
6	Results	16
6.1	Code resource usage	17
6.2	Alternate MTL statistics data	18
6.3	Alternate MTL statistics SV3 mocks	21
6.4	SV3 2PCF results	21
7	Discussion	23
8	Data availability	27
9	Code availability	27
Author List		32

1 Introduction

Large scale spectroscopic sky surveys have been a primary way of studying the universe over the past twenty five years [1–3]. They obtain redshifts of many millions of galaxies in order to measure their position in three dimensional space and extract information regarding the clustering of galaxies and the dark matter fields which they trace throughout cosmic time. These measurements not only provide insight into the composition of the universe and the physics which drive the evolution of its expansion, but also the formation of the galaxies themselves.

The Dark Energy Spectroscopic Instrument (DESI) survey, like the extended Baryon Oscillation Spectroscopic Survey (eBOSS) and the Sloan Digital Sky Survey (SDSS) before it, utilizes a multi object spectrograph (MOS) which enables multiplexing of spectroscopic observations. However, DESI is able to multiplex at a scale ~ 5 x larger than that of eBOSS, capturing $\gtrsim 4000$ spectra of astronomical objects per pointing. This increase in efficiency is largely due to the fact that SDSS and eBOSS hand-placed the fibers which fed their spectrographs into plates, while the fibers which feed the DESI spectrographs are controlled robotically. These positioners place all 5000 fibers on targets within the focal plane [4]. DESI also utilizes a dynamic exposure time calculator [5, 6] which takes information on observing conditions including seeing, transparency, and sky brightness to estimate the true signal-to-noise accumulated during an exposure. The capability to optimize exposure times and to construct new “plates” as needed, combined with a fivefold increase in available fibers and a ~ 2.5 x increase in the mirror area, allowed DESI to obtain ~ 10 x as many extragalactic redshifts in its first year as eBOSS did in its entire 5 year survey.

While robotically controlled fibers have provided many gains to the DESI survey, they have not eliminated the problem of fiber collisions. A fiber collision occurs when two targets are close enough on the sky that two fibers and their plugs (in SDSS and eBOSS) or their positioners (in DESI) would have collided were a survey to attempt to simultaneously observe both [7]. DESI also has a limitation similar to that of fiber collisions where, in most cases, only one fiber may be placed within a single fiber patrol radius. The lone exception is in a small region of overlap towards the edges of the patrol radii of neighboring fibers. This enforces uniform areal observation density within a single field of view which can only begin to reflect the non-uniform density of targets after multiple passes. Since they have similar effects on the ability to observe close pairs, we will refer to both effects collectively as ‘fiber collisions’.

Additionally, robotic fiber positioners can suffer from time-variable issues which were not present in the hand-plugged fiber systems used by SDSS, such as losses to positioner power and software-based failures which interrupt the positioning process. Fiber-collision issues combine with more traditional problems such as sky subtraction failures, atmospheric variability, and redshift determination errors, to create incompleteness effects which vary spatially and temporally across the DESI survey.

Fiber collisions affect the smallest scale clustering since they are limited to angular scales on the sky of approximately < 1.48 arcminutes which projects to a distance of $0.99 \text{ Mpc } h^{-1}$ at an intermediate redshift of $z = 1.0$. Real time hardware events like power failures to the devices which control groups of fiber positioners (CAN busses) can also cause missing observations at larger sets of separations. These smaller scales are critical to analyses of the galaxy-halo connection. These analyses, whether they use Halo Occupation Distribution (HOD) methods [8–13] or SubHalo Abundance Matching (SHAM) methods [14–16], rely on clustering measurements at small scales to determine the placement and quantity of satellite galaxies.

Prior to the usage of pairwise inverse probability weighting, (e)BOSS and SDSS used several different methods to correct for incompleteness. The BOSS DR9 BAO analysis [17] used the method developed by Guo, Zehavi, and Zheng in 2012 [18] to correct for fiber collisions. These weights are determined by dividing the galaxy sample into those galaxies which are not subject to any fiber collision effects and those which are subject to them. The

relative fraction of galaxies within the collided sample which are observed is calculated. Then the pair counts which include galaxies from the collided sample are upweighted based on that fraction. This method does not account for the fact that unobserved pairs are not statistically equivalent to those which are observed. The Reid et al. 2014 [19] analysis of the growth of structure using small scale clustering utilizes a combination of angular upweighting on small scales, and nearest neighbor upweighting, which accounts for unobserved targets by adding 1 to a “close pair” weight for an observed target for each unobserved target which is closer to it than any other observed member of its target class, at large scales, and a truncation of the measured multipoles at small separation perpendicular to the line of sight. Other standard SDSS and eBOSS analyses [20–24] have used nearest neighbor upweighting. This method assumes that all fiber collisions occur between galaxies within clusters or are otherwise at similar redshifts.

To address this issue of measuring small scale clustering, Bianchi and Percival 2017 (hereafter BP17; [25]) developed a new method of correction for incompleteness by upweighting each pair of galaxies rather than individual galaxies or the total number of pair counts. They devise a method of Pairwise Inverse Probability (PIP) weights which maintains consistent numbers of both weighted pair counts and weighted individual galaxy counts on all scales as long as there is some area of the survey footprint which is observed multiple times.

Bianchi and Percival note that, despite the nominal possibility of computing such pairwise weights analytically, for modern surveys like DESI which will observe $\mathcal{O}(10^7)$ galaxies, such calculations will be practically impossible. They do note that these weights can be effectively estimated by randomly repeating the galaxy selection process a sufficient number of times and estimating the probability of each pair based on the frequency at which each pair is selected. The remainder of this paper will show a new method which allows us to estimate those probabilities, show its effect on the clustering measurements, and validate the weights using mocks.

Mohammad et al. 2018 [26] showed the first application of the BP17 PIP weighting to data using the VIPERS survey. Additionally, they compute angular upweighting as described in Percival and Bianchi 2017 (hereafter PB17; [27]). This application to a slit-based survey is not strictly the same as the DESI application, but showed the validity of the BP17 PIP + PB17 angular upweighting method when applied to mocks and data. Mohammad et al. 2020 [28] showed the first application of the BP17 PIP weighting + PB17 angular upweighting method to MOS data using the eBOSS DR16 catalogs. They used a version of the eBOSS tiling software which was modified to enable easier parallelization of fiber assignment across 1860 realizations of the eBOSS survey in which they varied only the initial random number generation used by the software. They store the results of the realizations as bitweights as recommended in BP17 and then use them to compute pairwise inverse probability weights. The authors then use EZ Mocks to validate that their application of PIP weighting successfully recovers one-halo clustering down to a scale of $0.1 \text{ Mpc } h^{-1}$.

In this paper we present a new method of calculating PIP weights to correct the DESI SV3 and Y1 clustering measurements for incompleteness. This method utilizing Alternate MTLs (AMTLs) is a nearly-maximally realistic Monte Carlo estimation of the likelihood that any target or collection of targets would be observed over an infinite repetition of the DESI experiment. This method incorporates more realism than prior treatments in eBOSS

by utilizing real observation results, hardware states, and sequential changes in priority as overlapping tiles are observed. This method will enable unbiased measurement of clustering of galaxies down to scales $\leq 0.1 \text{ Mpc } h^{-1}$.

In section 2 we describe the DESI survey strategy and how this affects the observation of different classes of targets, in section 3 we briefly discuss the mock galaxy catalogs used to validate this method, in section 4 we describe both the regular and alternate MTLs and the procedure for how we create and propagate the alternate MTLs, in section 5 we describe the DESI clustering code, PYCORR and how it implements these weights, in section 6 we show results from the data and mocks to validate the method and show the differences with other methods of incompleteness determination, and finally in section 7 we discuss the results and their applicability to other analyses within DESI.

2 Survey

The Dark Energy Spectroscopic Instrument (DESI: [29–31]) is the Stage IV successor to large scale sky surveys such as the Sloan Digital Sky Survey (SDSS: [1]) and the (extended) Baryon Oscillation Spectroscopic Survey ((e)BOSS: [3]). The DESI experiment is currently in its third year of a projected five year program and has already recorded redshifts for over 30 million galaxies out to redshifts over $z \sim 3.5$. The primary goal of this survey is to further constrain cosmological parameters such as the dark energy equation of state parameter (w) and the Hubble Constant (H_0). The survey plans to do this by performing the most precise measurement of the Baryon Acoustic Oscillation (BAO) feature and of Redshift Space Distortions (RSD) in its clustering measurements. However, the science goals of DESI are not limited to dark energy. Already-published analyses from DESI include explorations of the Galaxy Halo Connection, measurements of the local stellar population in the Milky Way, studies of Quasistellar Object (QSO) astrophysics, and many tests of the survey and pipeline performances. This paper is part of a set of analyses surrounding the first Data Release of DESI (DR1, [32]) which includes two point clustering measurements and validation [33], measurements of the BAO feature in Galaxies, QSOs, and the Lyman- α forest [34, 35], full shape measurements of Galaxy and QSO clustering [36], as well as cosmological constraints measured from the above features including tests of primordial non-Gaussianity [37–39].

DESI selects the galaxies and stars that it observes from a parent population of Milky Way targets (MWS; [40]), Bright, low redshift, Galaxies (BGS; [41]), Luminous Red Galaxies (LRGs; [42]), Emission Line Galaxies (ELGs; [43]), and Quasistellar Objects (Quasars/QSOs; [44]). Targets are obtained via prior imaging of the DESI footprint [45]. These targets are divided into two concurrent programs. In one program, BGS and MWS targets were observed in bright or otherwise suboptimal conditions. In the other program LRG, ELG, and QSO targets were observed in sufficiently good dark time conditions [5].

During the One-Percent Survey (also referred to as SV3), these targets were observed within 20 rosettes [46]. SV3 was designed to optimize the observation of complete samples of tracers over an area approximately equal to one percent of the final survey area. This survey also provided a final test for the survey operations procedures and software prior to the main survey. These rosettes were selected to be evenly distributed between the northern and southern galactic caps as well as to overlap with various fields of interest from

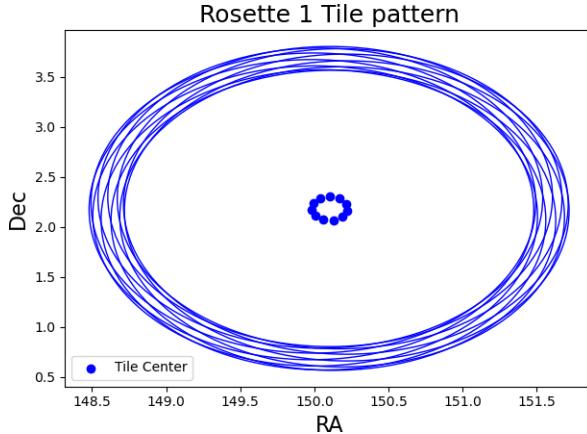


Figure 1. This plot shows the pattern in which an SV3 rosette is tiled, using Rosette 1 as an example. The points represent field centers and the circles represent the approximate (uncorrected for projection distortion) DESI FOV around each center.

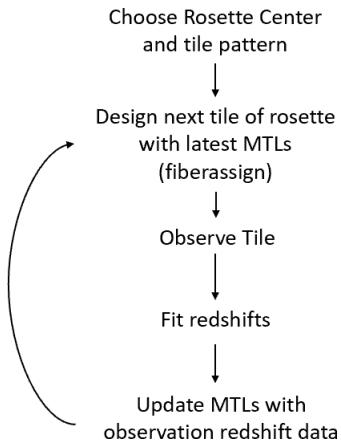


Figure 2. This flowchart sketches out the order in which SV3 operations happen. This highlights that MTLs are always updated in-between fiber assignment of one tile and fiber assignment of the next overlapping rosette tile.

CFHTLS, COSMOS, DES, GAMA, Euclid, ELAIS, GOODS, HSC, KiDS, VVDS, XBootes, and XDEEP as well as overlapping with the Coma cluster and the ecliptic pole. For each rosette, 12/10 (dark/bright) tiles were designed with tile centers spaced evenly around the rosette center at a distance of 0.12° as shown in figure 1. Observation of these rosettes followed a sequence where, for each rosette, one tile was designed and fiberassigned in the afternoon before observing or earlier, observed at night, and then the Merged Target Ledgers (MTLs¹) containing the targets which were observed in that tile² were updated. Then the next overlapping tile on the same rosette was designed, incorporating the information from the previous night’s observations, and observed in the same manner until all 12/10 tiles were observed. A flowchart of this process is shown in figure 2.

¹MTLs will be further explained in section 4.3 and more details can be found in section 6 of [5].

²MTLs are divided into HEALPixels (<https://healpix.jpl.nasa.gov/>) with $nside = 32$ (pixel size $\sim 1.8^\circ$) for ease of reading and storage.

This strategy enabled high completeness of 96.2% for MWS targets, 98.9% for BGS targets, 98.6% for LRG targets, 95.2% for ELG targets, and 99.4% for QSO targets within a range of rosette radii of 0.2 and 1.45 degrees [46]. Despite the nearly 100% completeness in that range, there is still an inner and outer region in which we can study the incompleteness measurements and weights from the AMTL method. Additionally, some regions within the central complete area are still subject to fiber collision effects due to high target number density and nightly focal plane positioning issues.

The observed SV3 spectra were processed by the DESI spectroscopic pipeline [47] in a homogeneous processing run denoted as ‘Fuji’, with the resulting redshift and large-scale structure (LSS) catalogs used in this paper released in the early data release (EDR [48]). The results of the AMTL pipeline we describe in this paper were included in the EDR LSS catalogs.

The main survey maintained the strategy of not observing overlapping areas of sky before updating the MTLs, but is no longer confined to rosettes. The tiles for the main survey were chosen based on the Hardin, Sloane, and Smith 2001 icosahedral tiling³[49] and are often fiber assigned “on-the-fly” as survey observations are occurring. This allows for greater flexibility in where one can point the telescope during a night of observing due to changing conditions including wind and clouds. The main survey is ongoing, but the DESI data release 1 (DR1 [32]) will include data from just over the first year of its operations, May 14, 2021 through June 14, 2022. Analyses of this dataset are ongoing and the resulting LSS catalogs have completeness 62% in BGS, 69% in LRG, 35% in ELG, and 88% in QSOs [33]. While we do not focus this analysis on the DR1 data, we do discuss some of the Y1 pipeline outputs and validation that the pipeline is working as expected. The relative incompleteness of the DR1 sample makes the AMTL pipeline a vital piece in the analysis, especially for the creation of realistic mocks [50, 51] and analyses of small-scale clustering.

3 Mocks

In order to validate the AMTL method, we use a set of 25 mock realizations processed in a similar manner to the data, reproducing the SV3 processing. They are built from the dedicated ABACUSSUMMIT⁴ N-body simulations [52, 53] and include different ensembles for LRG, ELG and QSO. These are then combined into a single file that mimics the structure of the observation data and is suitable to be processed through fiber assignment and LSS pipelines.

Mocks are made from Abacus realizations with fiducial cosmology (Planck 2015 Λ CDM cosmology [54]). For each tracer type (ELG, LRG and QSO) we build cubic mocks of 2 Gpc h^{-1} size each, at different primary redshifts following a HOD modelling [10, 12] tuned to the clustering amplitude of SV3 data. In SV3, we chose the following Abacus snapshots: LRGs at $z=0.8$, ELGs at $z=1.1$ and QSOs at $z=1.4$.

Once the cubic box mocks are made, we transform them to CutSky mocks, translating from Cartesian coordinates to sky coordinates and including the effect of redshift-space distortions, resulting in galaxy catalogs already in redshift space. These catalogs are resampled to match the same $n(z)$ distribution and footprint as SV3 data and prepared such that we add the

³<http://neilsloane.com/icosahedral.codes/>.

⁴<https://abacussummit.readthedocs.io/en/latest/>.

columns needed to go through fiber assignment. In addition, we also make random catalogs following the same footprint and redshift distribution as the mock catalogs.

These mock catalogs are then processed through the AMTL pipeline (described in section 4) and LSS pipeline (described in section 4 in [48]), creating clustering catalogs analogous to those of the data. Not all the observational effects are modelled, and therefore we adapt the AMTL and LSS pipelines to accommodate these differences.

The only significant difference in the AMTL pipeline is that we do not pull any information from the redshift-fitting process [47] and the effect of redshift failures is ignored through the entire AMTL loop (fixing $ZWARN = 0$ in all steps). In summary, we assume that all targets assigned have a good redshift. In future releases we expect to model this effect more robustly, although we expect negligible effect on the clustering signal, since redshift failures will depend on the individual physical properties of the targets and are assumed to be independent of the underlying clustering.

Finally, the output products of the mock AMTL method are run through the LSS pipeline to create clustering ready catalogs with the same completeness properties as the data. Together with the mock data, we process the associated randoms. Here we have modified the pipeline to ignore any effect coming from low data quality (assuming a 100% success rate) and to ignore any angular mask applied to the data. These implied subtle changes in the pipeline at the full catalog stage (see section 4.2.1 in [48]).

4 Alternate Merged Target Ledgers

4.1 Pairwise Inverse Probability weighting

Before describing the Alternate MTLs, we will first describe how they are used in the context of the Pairwise Inverse Probability (PIP) weighting of PB17.

PIP weighting, as the name suggests, involves weighting pairs of galaxies based on the inverse of the probability that each individual pair would be observed over the course of the entire survey. This probability can theoretically be calculated analytically, but the large number of galaxies observed in DESI ($\mathcal{O}(10^7)$) prohibits this from being practical. Instead of this, we can run Monte Carlo realizations of DESI, varying only those selection properties which are truly random.

When deciding how many realizations to generate for the calculation of PIP weights, we consider that in order for the PIP weights to be formally unbiased, there must not be any zero probability pairs. While we cannot guarantee the absence of zero probability pairs in a finite number of realizations, we show in table 4 that in SV3 there are no zero probability QSOs and BGS_BRIGHT targets with $\mathcal{O}(10^{-3})\%$ zero probability LRGs and ELGs, while in Y1 the fraction of zero probability objects is less than 1%.

To mitigate the effect of zero probability pairs, we also employ the angular upweighting of PB17. The combination of PIP weighting with angular upweighting was shown in Bianchi et al. 2018 [55] as well as in subsequent analyses of data and simulations [26, 28, 56] to provide practically good results in the presence of zero probability pairs despite a lack of perfect theoretical rigor.

Parameter Name	Parameter Description
TARGETID	Unique identifier for each DESI target (see [57]).
RA	Right Ascension in decimal degrees.
DEC	Declination in decimal degrees.
REF_EPOCH	Reference epoch for stellar targets used to adjust coordinates for proper motion.
PARALLAX	Component of stellar coordinate change due to parallax.
PMRA/PMDEC	Component of stellar coordinate change due to true (proper) motion of stars.
SV3_DESI_TARGET	Bitmasks described completely in <code>desitarget.targetmask</code> (see [57]) which record which target class(es) a target falls into.
/SV3_BGS_TARGET	
/SV3_MWS_TARGET	
/SV3_SCND_TARGET	
PRIORITY	The priority which fiberassign uses to determine which targets to assign fibers. This is determined by target class and observation status.
PRIORITY_INIT	The priority for a target that has yet to be observed.
SUBPRIORITY	A small additional priority which fiberassign uses to reproducibly resolve collisions. This is a random number in the range 0 to 1.
OBSCONDITIONS	Bitmask which represents the observing conditions (bright, dark or backup) in which an object can be targeted (see [5]).
NUMOBS_INIT	The number of observations required for a target that has yet to be observed.
NUMOBS_MORE	The number of successful observations until the target is considered “done”.
NUMOBS	The current number of observations of the target.
Z	The redshift of the target.
ZWARN	Bitmask which records Redrock warning flags for the redshift determination.
ZTILEID	The ID(s) of the tiles on which a target obtained a good observation.
TARGET_STATE	A human-readable representation of the current state of the target factoring in its various target classes and current status.
TIMESTAMP	The timestamp at which the MTL entry was made.
VERSION	The version of the <code>desitarget</code> code which was used to make the MTL update.

Table 1. List of parameters in a merged target ledger (see also, e.g., https://desidatamodel.readthedocs.io/en/stable/DESI_SURVEYOPS/mtl/sv3/dark/sv3mtl-dark-hp-HPX.html).

4.2 The Merged Target Ledgers

DESI stores potential targets in a set of files referred to as the merged target lists or ledgers (MTLs; see section 6 of [5]). These MTLs contain information generated and updated by the DESI targeting software [57] and used by the DESI assignment code [58], to place fibers on targets. This information includes a unique TARGETID field to identify objects, sky coordinates expressed as Right Ascension (RA) and Declination (DEC), a PRIORITY (determined by a target’s class), a randomly determined SUBPRIORITY (in the range 0 to 1), and bitmasks indicating the different target classes and programs to which an object belongs. A summary of all the parameters in the MTLs is given in table 1 and a more detailed description is given in section 6 of [5].

TARGETID	PRIORITY	NUMOBS	Z	ZWARN	TARGET_STATE
1	103200	0	-1.0	-1	LRG UNOBS
2	103100	0	-1.0	-1	ELG_HIP UNOBS
3	103400	0	-1.0	-1	QSO UNOBS

Table 2. Simplified initial merged target ledger.

TARGETID	PRIORITY	NUMOBS	Z	ZWARN	TARGET_STATE
1	103200	0	-1.0	-1	LRG UNOBS
2	103100	0	-1.0	-1	ELG_HIP UNOBS
3	103400	0	-1.0	-1	QSO UNOBS
1	2	1	0.684	0	LRG DONE
2	103100	1	1.45	4	ELG_HIP MORE_ZWARN
3	103350	1	1.9	0	QSO MORE_MIDZQSO

Table 3. Simplified updated merged target ledger.

When an observation is performed, each target for which good data was obtained has an entry added to the end of ledger that records the results of the observation. A highly simplified example of the initial ledger is presented in table 2 and the result of updating the states in the simplified initial ledger is presented in table 3. As shown in tables 2 and 3, the updates made to a target after a successful observation include an increment in NUMOBS, a drop in PRIORITY, an associated change in TARGET_STATE, and updates to the Z and ZWARN fields based on the output from the DESI spectroscopic pipeline [47] and redshift-fitting software.

The updating of MTLs in-between repeat observations of the same patches of sky combined with the resulting drop in priority for successfully observed targets highlights the utility of the AMTL method. Crucially, a single change in which a target was observed on the first pass through a region can cause cascading effects due to fiber placement on subsequent tiles.

4.3 Generating the alternate Merged Target Ledgers

Having described the regular MTLs, we will now describe how we generate the alternate MTLs.

The primary difference between the regular and alternate MTLs is the randomization of the subpriority field. The other major difference is, for both SV3 and main survey, a reproduction of the BGS promotion of 20% of BGS_FAINT targets to the higher priority class of BGS_BRIGHT, and for the main survey only, a reproduction of the ELG promotion of 10% of all ELG targets to the higher priority ELG_HIP class. DESI promotes a fraction of BGS_FAINT targets to the higher BGS_BRIGHT priority class and, in Y1 only, a fraction of ELG_LOP and ELG_VLO targets to the higher ELG_HIP priority class (same class as LRGs) in order to better measure the small scale cross-correlation between target classes in redshift regions where they overlap. This is especially important for the BGS samples, as the FAINT selection primarily extends the flux limit and a natural goal for the sample is to understand the distribution of satellite galaxies all the way to that flux limit. Without the

Tracer type	% of Zero Probability	% of Zero Probability
	Objects in SV3	Objects in Y1
BGS_BRIGHT	0	0.21
LRG	3E-3	0.22
ELG	3E-3	0.17
QSO	0	0.28

Table 4. Percentage of zero probability objects over 128 realizations of SV3 and Y1.

priority promotion, only the BRIGHT sample would be assigned fibers in the most dense regions. The LRG and ELG samples both have high densities in the range $0.6 < z < 1.0$ and thus boosting the priority of the ELGs improves the ability to determine the likelihood of ELGs and LRGs occupying the same halos.

To begin the alternate MTL generation process, the survey MTLs are loaded and then truncated so that only the initial entries are maintained. For the initial entries, the SUBPRIORITY field is replaced with new random numbers from the basic numpy uniform random number generator. The random number generation is seeded based on the healpixel number of the initial MTL, the realization number, and a user-provided random seed. The SV3 MTLs were randomized with a version of the code which did not add in the healpixel number to the random seed. This randomization is still valid, but we include this information, as well as a tagged version of the code as it was for the SV3 MTLs⁵ to allow for reproducibility. See section 9 for more information. The promotion of the BGS targets is a multiple step process where first, the initially promoted 20% are demoted back to the standard BGS_FAINT priority and then a desitarget convenience function (`random_fraction_of_trues`) is used to select 20% of the total BGS_FAINT targets to promote to BGS_FAINT_HIP. The promotion of ELG targets is similar with two additional considerations. First, targets are being drawn from two starting classes: ELG_LOP and ELG_VLO and must be returned to their respective classes. Second, not all demoted objects will see a drop in priority since it is not forbidden to have an object be targeted as both an ELG and an LRG (same priority as ELG_HIP) or as both an ELG and QSO (higher priority than ELG_HIP).

This procedure is done in parallel for up to 128 alternate MTLs per node on the Perlmutter supercomputer at NERSC.⁶ For the SV3 alternate MTLs, we used the now-retired Cori supercomputer also at NERSC⁷ which only had 32 physical cores per node, so we were only able to parallelize over 32 realizations per node. For SV3 and Y1, we have chosen to only have the 128 realizations of alternate MTLs for the data and 256 for the SV3 mocks shown in this work. This choice was made from a consideration of the number of zero probability objects that result from a given number of realizations as shown in figure 3 for SV3 and 4 for Y1. Each panel of both figures shows the number of zero probability objects from each target type (from left to right: BGS_BRIGHT, LRG, ELGs with QSO targets removed

⁵<https://github.com/desihub/LSS/releases/tag/v2.0.0-EDR>.

⁶<https://www.nersc.gov/perlmutter>.

⁷<https://www.nersc.gov/assets/Uploads/Intro-to-Cori.pdf>.

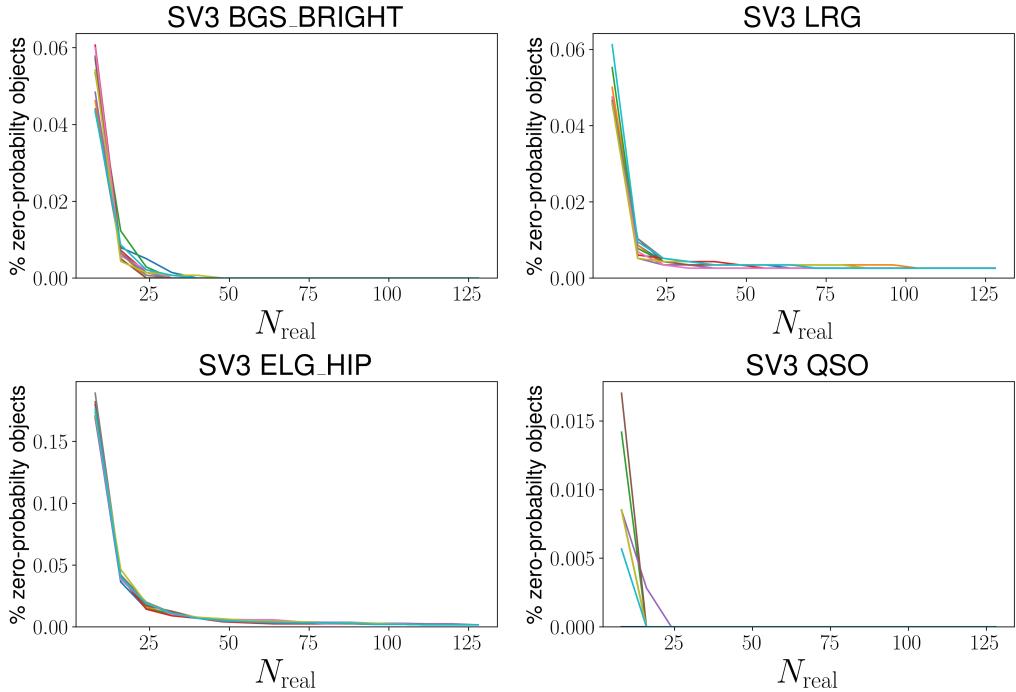


Figure 3. The percentage of objects not observed in any alternate realization of SV3 as the number of realizations increases from 8 to 128. Each line within a plot represents a random reordering of the 128 realizations to confirm that there is no dependence of the convergence to zero on the (random) order of the realizations.

(ELG_notqso), and QSO) as the number of realizations increases from 8 to 128. The multiple lines in each figure are random reorderings of the 128 realizations to confirm that there is no dependence of the convergence on the (random) ordering of realizations. The exact percentages of zero probability objects for SV3 data and Y1 data for the chosen number of 128 realizations is given in table 4. A zero probability object is a target that is never observed over a given number of realizations. BP17 states that their method of weighting is not completely valid where there are zero probability pairs. While we have essentially eliminated zero probability objects in SV3, the trend in figure 4 shows that it would be prohibitive to run enough realizations to eliminate all zero probability objects in Y1. However, we will show that the percentages of zero probability objects shown in table 4 from 128 realizations will sufficiently mitigate the problem.

4.4 The observation loop

The core of the utility of the AMTL method as compared to other implementations of PIP weighting is the ability to update the alternate MTLs using real observation information. In order to do this, we create an observation loop that follows the same ordering as the survey data processing and observation.

Following the creation of the initial randomized ledgers, the SV3 AMTL loop would use the `desitarget.mtl` function `tiles_to_be_processed` to check the list of observed SV3 tiles from the data against the list of already processed tiles for each realization. Then the set of tiles corresponding to the earliest remaining observation date were selected. All

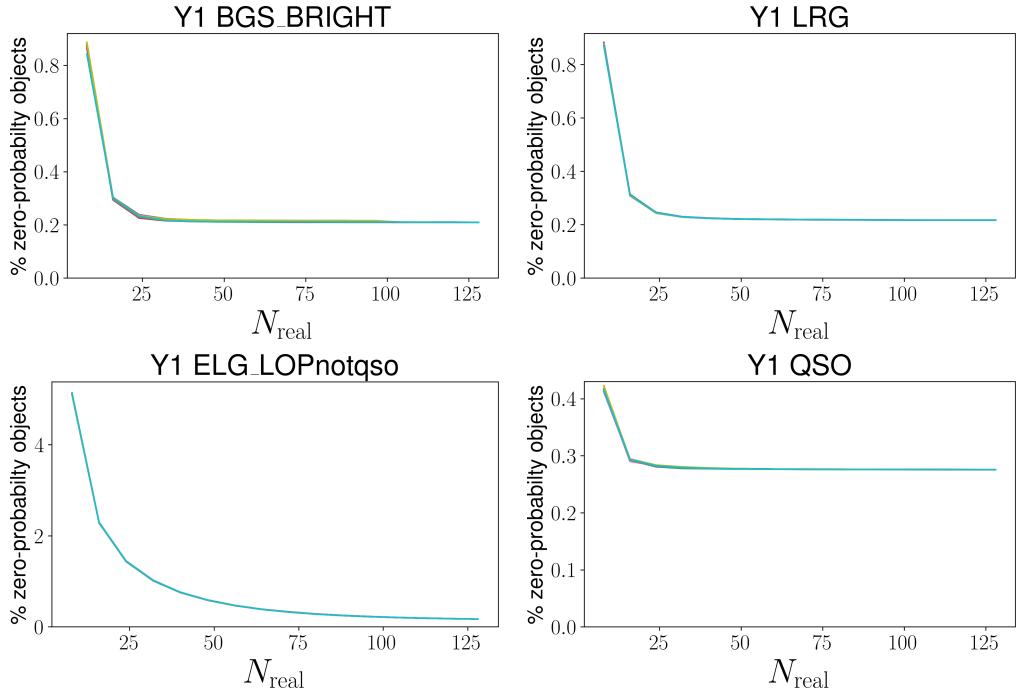


Figure 4. The same as figure 3 but for Y1 targets.

tiles which were observed on that date were put through an observation loop designed to replicate the SV3 observing process.

The observation loop begins by running fiberassign [58] using the alternate ledgers as the primary target ledgers. Other than the primary target ledgers, this procedure uses the same configuration as the corresponding fiber assignment run in the real survey including the list of sky fibers, calibration standards, secondary targets, and the hardware configuration at the exact time at which fiberassign was run for the main survey. This produces a set of files which map the alternate targets to DESI fibers for the alternate observations.

In order to accurately track the completeness effects due to improperly positioned fibers, fiber positioning system failures, and other focal plane-level effects, we map the observation result from each fiber in the original survey to the same fiber in the alternate realization. For example, if fiber 1 were pointed at target 1 in tile 1 of the real survey and fiber 1 were pointed at target 2 in tile 1 of an alternate realization, target 2 would receive the observation result of target 1 in that alternate realization. Inclusion of the observational results from the survey means any real-time positioning failures are automatically incorporated. The survey pipeline marks observations as bad if they come from fibers with flags indicating bad hardware status.

This can, and does, lead to the assignment of observations of one target class to observations of another target class. However, since the primary consideration is whether or not a target has been placed on a fiber which received a good observation rather than whether it received a successful redshift, this approach is still mostly valid. For the case of QSOs where the value of the redshift determines the number of future observations this is not exactly correct. For this purpose we are developing a method to assign results from a “twin fiber” which is nearby the true fiber and matches in as many properties as possible, primarily the

target class. For the full five year sample, we also plan to substitute QSO redshifts from original observations with the final survey redshift results. This will be possible since QSOs are the highest priority target class and will be nearly fully complete by the end of the survey. In this analysis, we will note that the fraction of observations whose primary target information changes is 6.7 % (13.0%) of all SV3 (Y1) dark time observations and 13.0 % (31.8%) of SV3 (Y1) dark time observations which result in a change in target.

The main survey data added an additional complication to the observation loop, the addition of reprocessed tiles. A reprocessing occurs when the results of the daily pipeline are modified due to (e.g.) a run of the offline pipeline discovering that a “good” redshift was not actually a good fit. Since reprocessings can change the state of already observed targets and they occur independently of the observations themselves, they necessitated a separation of the alternate fiber assignment step and the alternate MTL update step where we check for reprocessings of tiles. This also resulted in the storage of the maps between the real and alternate fibers from the regular observation loops so that the reprocessed redshift catalogs/ledgers can be modified in the same way as the original ones.

4.5 Outputs of the pipeline

The AMTL pipeline produces two primary outputs. First is the collection of N (for SV3 and Y1 data, $N=128$) realizations of MTLs which includes information on every (alternate) observation of every target. The columns for these MTLs are the same as in the real survey MTLs, which you can find in table 1.

The second output is a bitweight file with three columns, it can be attached to the LSS catalogs (and is indeed already available for the EDR catalogs.⁸ [48]) First is the unique TARGETID which is necessary to match the information to other catalogs. Second is two columns with information on how frequently each target was observed over the 128 realizations. One of those two columns is the probability that the individual target would be observed over those 128 realizations, PROB_OBS. PROB_OBS is calculated as the number of realizations in which a target was observed divided by the total number of realizations. The other remaining column is the bitweights, a shorthand encoding of whether each target was observed in each realization as first outlined in BP17. Bitweights are stored as an ordered list of 64 bit integers where each bit is set to 1 if a target was observed ($\text{NUMOBS} \geq 1$) in a given realization and set to 0 if it was not observed in that realization.⁹

4.6 Validation of the pipeline

Validation of the pipeline has been a multiple step process. Prior to running the full set of 128 realizations, we have run a single realization with the part of the pipeline which randomizes the subpriorities turned off and confirmed that this version of the pipeline is able to reproduce both SV3 and Y1 exactly. We have also validated that the pipeline produces a statistically equivalent set of promotions to ELG_HIP and BGS_BRIGHT out of the initial randomized alternate ledgers.

⁸<https://data.desi.lbl.gov/doc/>.

⁹In order to manipulate these bitweights, we provide the functions within desihub’s `LSS.bitweights` module.

In section 6 we show clustering measurements validating the PIP weights generated from the pipeline on the mocks described in section 3.

5 Clustering measurements

We will measure clustering based on pair-counts as a function of the redshift-space transverse, s_{\perp} , and radial, s_{\parallel} separation, measured in co-moving h^{-1} Mpc using the fiducial cosmology. The 2-point correlation function probes the excess probability to find in a galaxy sample two galaxies at a given separation. We use the so-called Landy-Szalay estimator [59]

$$\hat{\xi}(s_{\perp}, s_{\parallel}) = \frac{DD(s_{\perp}, s_{\parallel}) - DR(s_{\perp}, s_{\parallel}) - RD(s_{\perp}, s_{\parallel}) + RR(s_{\perp}, s_{\parallel})}{RR(s_{\perp}, s_{\parallel})} \quad (5.1)$$

where $XY(s_{\perp}, s_{\parallel})$ correspond to the weighted number of pairs of objects X and Y in a s_{\perp}, s_{\parallel} bin divided by the total weighted number of pairs. Specifically, $DD(s_{\perp}, s_{\parallel})$ is the weighted number of data (galaxy and quasar) pairs and $RR(s_{\perp}, s_{\parallel})$ is the weighted number of pairs of objects in the random catalog which samples the selection function. For the autocorrelation (single-tracer) measurements performed in our analysis, by symmetry $DR(s_{\perp}, s_{\parallel}) = RD(s_{\perp}, s_{\parallel})$. For standard estimates pair weights are the product of the total individual weights of the two objects in the pair w_1 and w_2 , themselves obtained as the product of systematic correction weights and FKP weights. We use 48 logarithmically spaced s_{\perp} bins between 0.01 and 100 Mpc/ h and 40 linearly spaced s_{\parallel} -bins between 0 and 40 Mpc/ h .

To optimize computing time we use the technique of Keihänen et al. 2019 [60], which consists in summing DR , RD and RR pair counts obtained using several random catalogs of a size approximately matching that of the data catalog. In practice, we use the maximum number available for the SV3 LSS catalogs, which is 18 and corresponds to 28, 67, 220, and $54 \times$ the data number for ELGs, LRGs, QSOs, and BGS_BRIGHT respectively.

In this work, we will show the results of the projected correlation function $w_p(r_p)$, with $r_p \equiv s_{\perp}$ and

$$w_p(r_p) = 2 \int_0^{40} ds_{\parallel} \xi(r_p, s_{\parallel}) \quad (5.2)$$

The projected correlation function w_p integrates the full two point correlation function ξ over a range in the line-of-sight distance spanning 40 Mpc h^{-1} . This distance is chosen to sufficiently remove all rsd effects while not introducing significant noise.

Correlation function measurements are performed with PYCORR¹⁰ which wraps a version of the CORRFUNC package [61, 62] modified to also run on GPU and support alternative lines-of-sight (first-point, end-point) and weight definitions (PIP and angular upweights) and the θ -cut, along with jackknife error estimation.

5.1 Weights

The set of weights applied in the DESI SV3 TPCFs include the “default” (IIP) weights, FKP weights, angular upweights, and bitwise (PIP) weights. The PIP and IIP weights

¹⁰<https://github.com/cosmodesi/pycorr>.

are never applied simultaneously since they are derived from the same set of AMTLs and correct for the same incompleteness.

The ‘‘default’’ IIP weights are calculated from the bitweights output from the AMTL pipeline with the ‘‘efficient estimator’’ of Bianchi and Verde 2020 [63].

$$w_{i,\text{IIP}} = \frac{N_{\text{real}} + 1}{(\text{popcnt } w_i^{(b)}) + 1} \quad (5.3)$$

FKP weights were derived in Feldman, Kaiser, and Peacock 1994 [64] as a variance-minimizing weight for a given scale of a power spectrum. These FKP weights are given by

$$w_{\text{FKP}}(z) = \frac{1}{1 + n(z)P_0} \quad (5.4)$$

where $n(z)$ is the number density of galaxies at redshift z , and P_0 is the clustering power at the scale in the power spectrum at which we desire to minimize the variance. In DESI SV3, we chose the following P_0 values for each tracer:

LRG:	10000
ELG:	4000
QSO:	6000
BGS:	7000

Angular upweighting, as first outlined in PB17 [27] and applied in Mohammad et al. 2018 [26] and Mohammad et al. 2020 [28], is derived from the angular clustering of the parent catalog of all potential targets of a tracer type and the angular clustering of the targets which have been observed. These pairwise weights are generated separately for the Data-Data pairs and the Data-Random pairs with the bitweights factoring into both sets of counts as shown below.

$$w_{\text{ANG}}^{\text{DD}} = \frac{DD^{\text{par}}(\theta)}{DD_{\text{PIP}}^{\text{fib}}(\theta)} \quad (5.5)$$

$$w_{\text{ANG}}^{\text{DR}} = \frac{DR^{\text{par}}(\theta)}{DR_{\text{PIP}}^{\text{fib}}(\theta)} \quad (5.6)$$

In equations (5.5) and (5.6), ‘‘par’’ and ‘‘fib’’ refer to the parent sample and fiber-assigned samples respectively. The fiber-assigned Data-Data pairs are weighted by the inverse of the probability that the pair would be observed (PIP). Similarly, the fiber-assigned Data-Random pairs are weighted by the inverse of the probability that the data galaxy would be observed (IIP).

Finally, the bitwise PIP weights are calculated using the efficient estimator from Bianchi and Verde 2020 [63]:

$$w_{ij} = \frac{N_{\text{real}} + 1}{\text{popcnt}(w_i^{(b)} \text{ and } w_j^{(b)}) + 1} \quad (5.7)$$

where w_{ij} is the weighting applied to the pair of galaxies i and j , N_{real} is the number of AMTL realizations, and $\text{popcnt}(w_i^{(b)} \text{ and } w_j^{(b)})$ is the count of the bitwise ‘‘and’’ operation

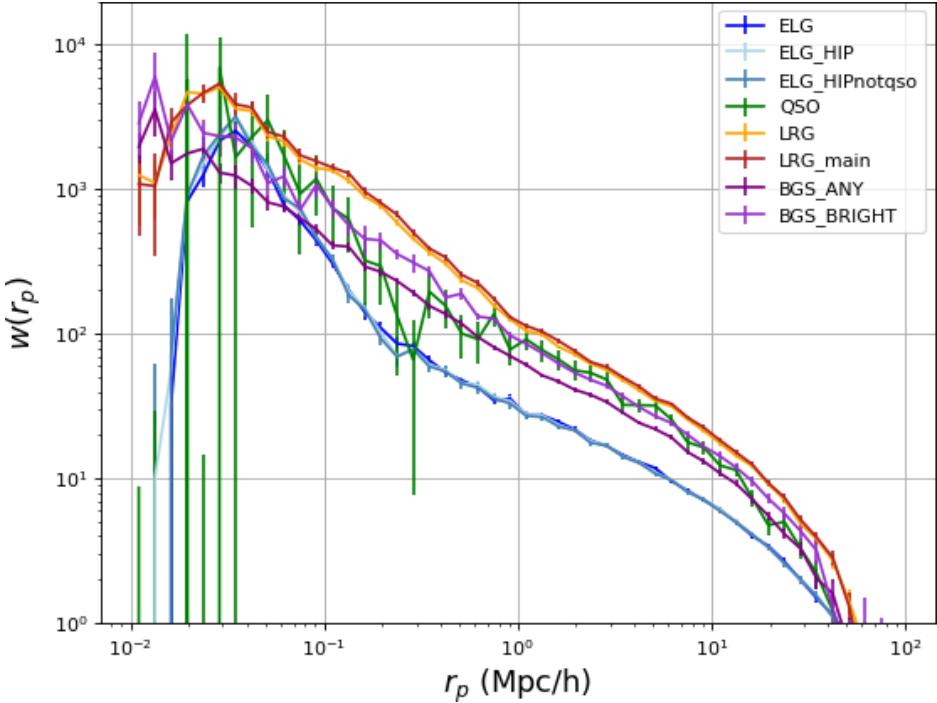


Figure 5. The projected clustering of all tracer types that have a ‘clustering’ catalog produced from DESI SV3 data.

between the bitweights of galaxies i and j . In the case of PIP weighting, the count of DD pairs is also normalized as in equation 22 of Bianchi and Verde 2020 [63].

These weights are provided to pycorr by multiplicatively combining all non-bitwise weights and then adding them to a (list of) bitweight(s).

Figure 5 shows the projected clustering results using the above methods for all tracer types which have a “clustering” catalog produced from DESI SV3 Data.

Figure 6 shows the clustering amplitude of four tracers (BGS_ANY, LRG, ELG, and QSO) as a function of their separation in bins of distance parallel and perpendicular to the line of sight (s_{\parallel} and s_{\perp} respectively).

6 Results

In this section we display results from the first two runs of the AMTL pipeline, one run covering SV3, and the other run covering Y1. Statistics which we will show include memory, disk space, and time usage of the code, a comparison of the number of targets in each class observed in the data and averaged over the AMTL runs, the estimation of completion from the AMTL pipeline, the difference in the completeness weights between the AMTL pipeline and methods which bootstrap completeness estimation from only the data, and finally the effects of weighting using PIP weights from this pipeline and other weighting methods on clustering statistics.

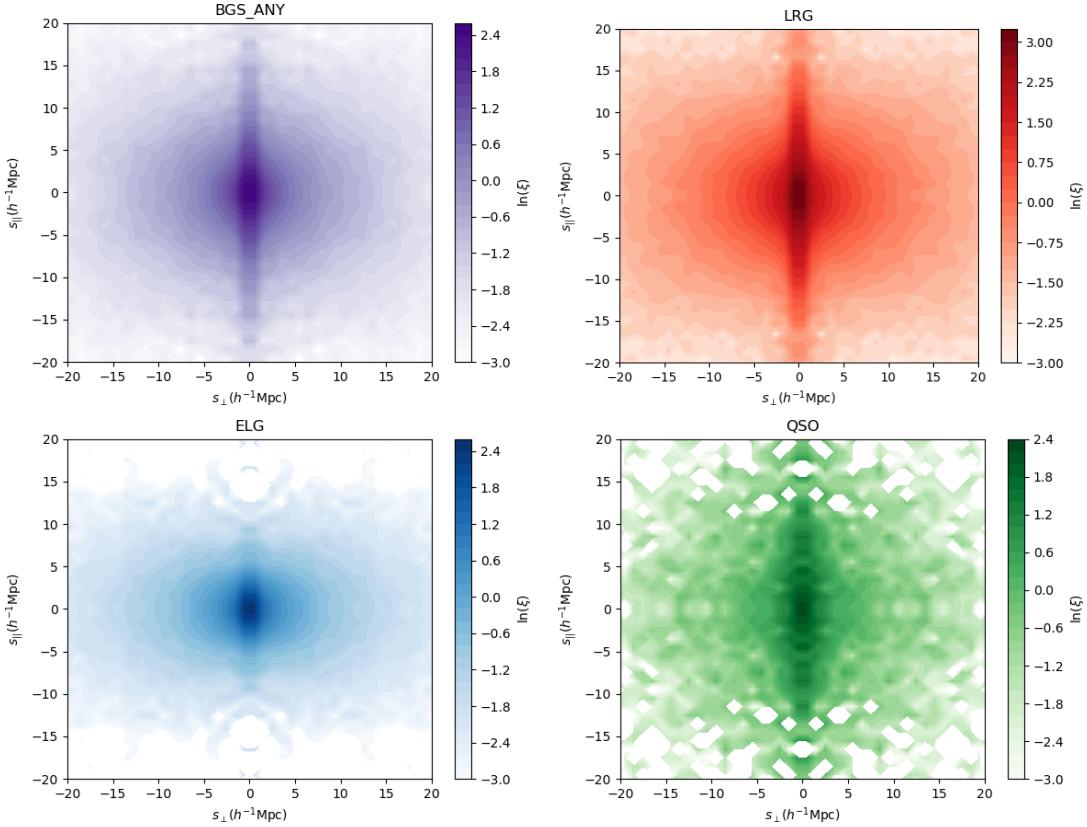


Figure 6. Contour plots showing the clustering amplitude as a function of s_{\perp} and s_{\parallel} .

6.1 Code resource usage

In order to estimate the resources required to generate the EDR AMTLs and corresponding bitweights, we reran the code which was originally run on the now-retired Cori cluster on the Perlmutter cluster at NERSC. In order to accurately estimate the resource usage of the EDR analysis in this paper, we divided the 128 realizations over 4 nodes as was required on Cori, despite Perlmutter having enough physical cores to run all 128 on a single node. The memory and computation time statistics below are for the dark time run only. Since there are comparable numbers of tiles between the bright and dark time programs in SV3, the time utilization would approximately double when including both and the memory usage should be similar.

The initialization step of the AMTL process took 11 wall-clock minutes, which converted to 0.73 node-hours. The peak memory usage of any one process generating one realization was 616 Mb and the peak total memory usage over all 128 processes was 78.5 Gb. The observation loop step took 16.04 wall-clock hours which converted to 64.1 node-hours. The peak memory usage of any one process handling one realization was 760 Mb and the peak total memory usage over all 128 processes was 77.6 Gb. Creating the bitweights from the 128 data realizations took 32 wall-clock minutes which converted to 0.54 node-hours as only one node is required to create the bitweights. The peak memory usage for the process was 2.73GB.

The full set of EDR MTL outputs including the bright and dark time AMTLs, fiberassign outputs, and bitweight files totals 917GB, with the vast majority coming from the MTLs

Tracer type	# Observed	Sum of PROB_OBS	%	# Observed	Sum of PROB_OBS	%
	in SV3	in SV3	difference	in Y1	in Y1	difference
BGS_BRIGHT	147721	148956	0.8	4234056	4274506	1.0
LRG	127020	128506	1.2	2477439	2503030	1.0
ELG	354398	358432	1.1	3985829	4021208	0.9
QSO	57060	57341	0.5	2036499	2051294	0.72

Table 5. Number of targets of each class observed in data and averaged over 128 AMTL realizations of SV3 and Y1.

themselves. Due to the nature of the ledgers, each MTL increases in size with each additional observation. More significantly, each additional run of fiberassign produces more outputs, some of which must be maintained for analysis. Between these increases and the much larger sky area covered by the main survey, disk space usage is rapidly becoming the limiting constraint in running the AMTL pipeline, with the Y1 outputs totaling 12.7TB. To mitigate this, we are currently developing several space saving measures, including storage of only the final ledger entry for each unique target and reducing the size of the alternate ledgers to only contain the columns which vary between realizations.

6.2 Alternate MTL statistics data

6.2.1 Comparison of number of observed targets by class

Since the AMTL pipeline functions by assigning real observation results from the DESI survey to alternate targets, we should end up with similar numbers of total observed targets between the average AMTL realization and in the data for the major target classes.

In table 5 we show the number of observed targets in the classes BGS_BRIGHT, LRG, ELG, and QSO in the SV3 Data and the Y1 Data as well as the sum of the PROB_OBS values for each of those target classes in the 128 realizations of the AMTLs.

Table 5 reveals a small, but consistent surplus in the number of observations in the AMTLs vs the data. This difference is also present and of similar magnitude in the mocks. Since the mock “observation” sample is produced in the same way as a random AMTL realization (see table 6), we believe this is reasonable behavior.

6.2.2 SV3 completeness estimation from AMTL

Figure 7 shows the completeness of one of the SV3 rosettes (rosette 16; rosettes are described in section 2 and [46]) as measured by the average PROB_OBS for targets of (left to right) LRGs, ELGs with QSO targets removed (ELG_notqso), and QSOs in RA/Dec bins within the rosette. The binning is varied in width based on the number of targets of each type with ELGs having the most targets and QSOs having the fewest. The tiling pattern of the rosette leads to the high, nearly 100% completeness in the intermediate parts of the rosette between 0.2 and 1.4 degrees from the center. Regions near the center of the rosette are subject to hardware failures more frequently since they are always covered by the same part of the focal plane and regions towards the exterior of the rosette are only tiled a few times.

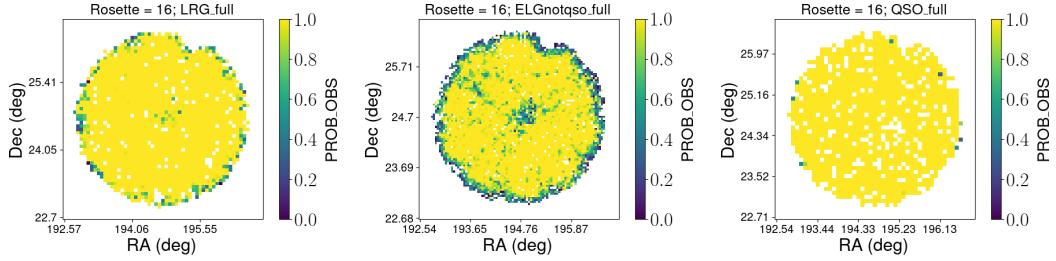


Figure 7. 2D histogram showing the mean value of PROB_OBS in bins of RA and Dec within Rosette 16 of SV3 for (from left to right) LRGs, ELGs with QSO targets removed, and QSOs.

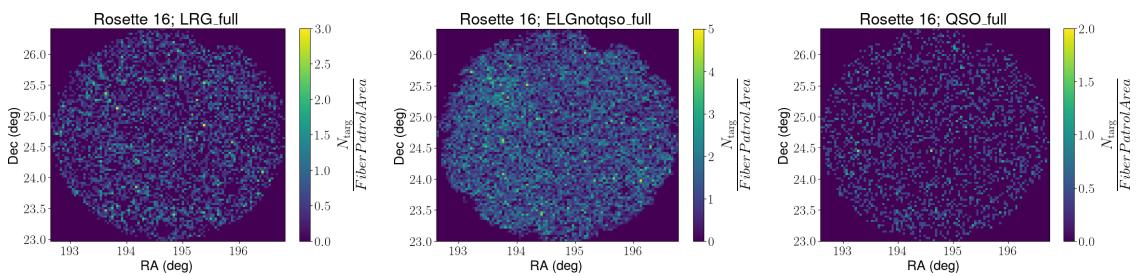


Figure 8. 2D histogram showing the areal number density (in units of targets per fiber patrol area) of targets in bins of RA and Dec within Rosette 16 of SV3 for (from left to right) LRGs, ELGs with QSO targets removed, and QSOs.

In order to study the potential for the AMTL-derived bitweights to impact clustering relative to weighting derived from other measures of incompleteness, we will compare the AMTL-derived observation probabilities with those other measures of incompleteness, focusing our comparisons on the most central and furthest exterior regions where completeness is not 100%.

Figure 8 shows the number density of targets within rosette 16 for each of the same groups of targets in the same order as figure 7. This is shown in units of number of targets per fiber patrol radius to better express the likelihood that a given bin in RA/Dec will experience fiber collisions. Focusing specifically on the upper left portion of the middle “ELGnotqso_full” subplot, we can see a correlation between regions of high number density and additional areas of low completeness where we can display the impact of the AMTL incompleteness weights.

Figures 9 and 10 are the equivalents of figures 7 and 8 for rosette 1. We show this rosette as a counterpoint to the rosette 16 since this rosette has much less pronounced differences in areal target number density and represents a more typical rosette.

Figures 11 and 12 show the mean value of the difference between PROB_OBS, the average probability that a target is observed as determined via the AMTL pipeline presented in this paper and FRACZ_TILELOCID,¹¹ the fraction of similar objects within a DESI tile which were successfully observed and received a redshift. These differences show that a utility like the AMTL pipeline is necessary. A simple Monte Carlo simulation of each target

¹¹The way SV3 catalogs were constructed [48], this is the most appropriate comparison. In the main survey LSS catalog processing, we have defined an additional column `FRAC_TLOBJS_TILES`, which accounts for additional incompleteness and should be multiplied by `FRACZ_TILELOCID`; see [65].

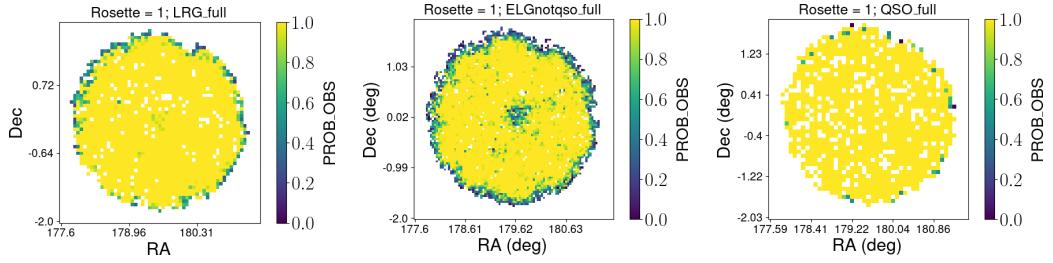


Figure 9. The same as figure 7 but for Rosette 1.

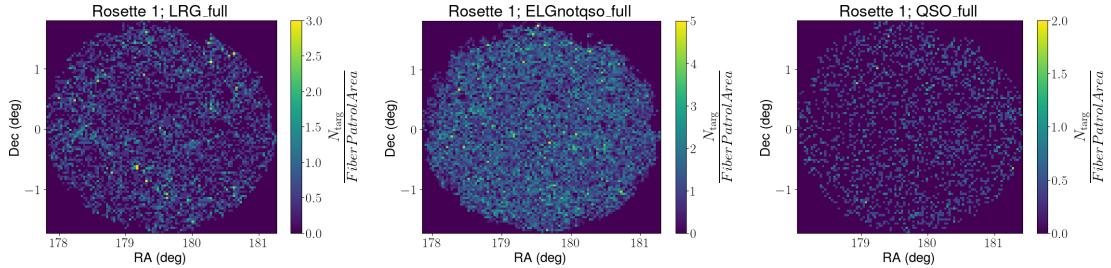


Figure 10. The same as figure 8 but for Rosette 1.

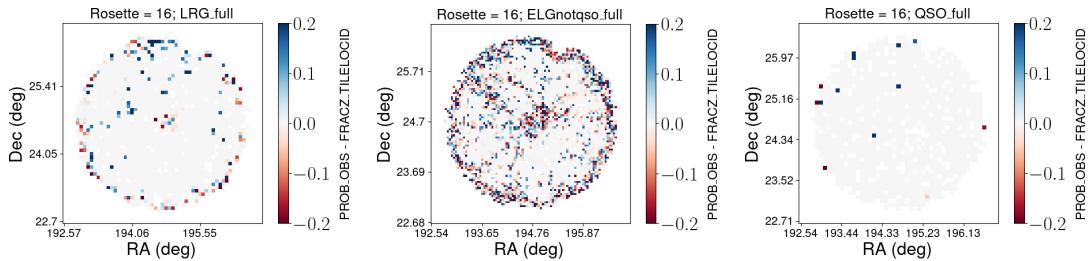


Figure 11. 2D histogram showing the mean value of the difference between PROB_OBS and FRACZ_TILELOCID in bins of RA and Dec within Rosette 16 of SV3 for (left to right) LRGs, ELGs with QSO targets removed, and QSOs.

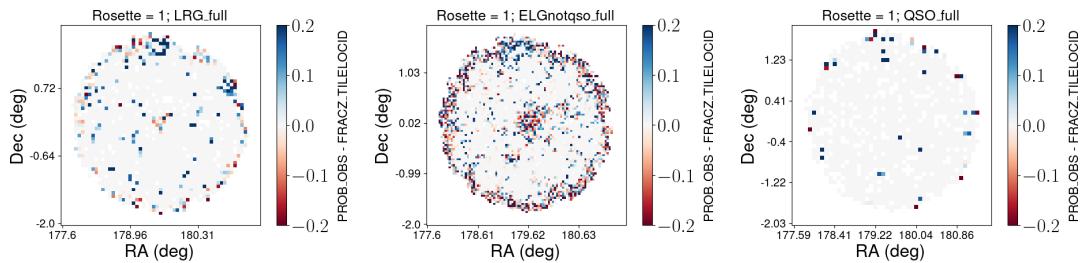


Figure 12. The same as figure 11 but for Rosette 1.

independently using what is, at first glance, a reasonable proxy for the probability that it is observed, can result in differences in probability for groups of objects on low completeness pixels of up to, occasionally exceeding, 20%.

Tracer type	# “Observed” in SV3 mocks	Sum of PROB_OBS in SV3 mocks	% difference
LRG	140952	141995	0.73
ELG	386709	388964	0.58
QSO	60129	60385	0.42

Table 6. Number of targets of each class marked as “observed” in mocks and average sum of PROB_OBS over 256 mock AMTL realizations of SV3.

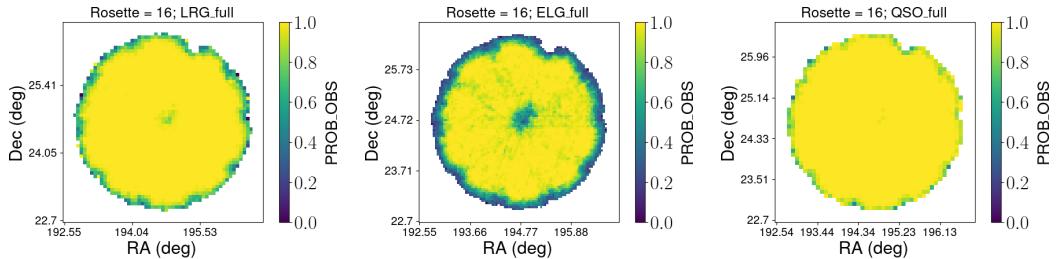


Figure 13. 2D histogram showing the mean value of PROB_OBS in bins of RA and Dec from 25 stacked mocks within Rosette 16 of SV3 for (from left to right) LRGs, ELGs, and QSOs.

6.3 Alternate MTL statistics SV3 mocks

We run the AMTL pipeline over 25 mocks generated from the AbacusSummit simulations as described in section 3, reproducing the same algorithm we apply on the observed data using 256 realizations instead of 128. The statistics averaged over the 25 mock realizations, as shown in table 6, show that the number of targets marked as “observed” in the mocks is $\sim 0.5\%$ smaller than the sum of the PROB_OBS in the mocks. Due to a larger number of available targets in the mocks, the mocks have significantly more observed targets than the data (shown in table 5) at a $\sim 10\%$ level.

We show the same plots as for the data AMTLs but with a stacked set of all 25 mocks to show that the mock AMTL process successfully replicates the data PROB_OBS distribution as well as the difference between PROB_OBS and FRACZ_TILELOCID within the limits of the similarity between the mock and data target density. These plots are the 2D histograms of PROB_OBS (figures 13 and 15), target number density (figures 14 and 16), and the difference between PROB_OBS and FRACZ_TILELOCID (figures 17 and 18).

6.4 SV3 2PCF results

Figures 19, 20, and 21 show the mean projected TPCF over 25 SV3 Abacus mocks with error bars coming from the standard deviation over the 25 mocks in four cases. The first case, which we treat as truth is the projected TPCF parent catalog of all possible mock targets and is shown in black. The remaining cases are different weights on what was chosen to be the “observed” set of targets. The “default” weight (orange) includes the IIP completeness weight, a combination of the default weights and the angular upweighting of PB17 (green; [27] and section 5.1 for more details), and the combination of the angular weights with the bitwise PIP weights discussed here (blue).

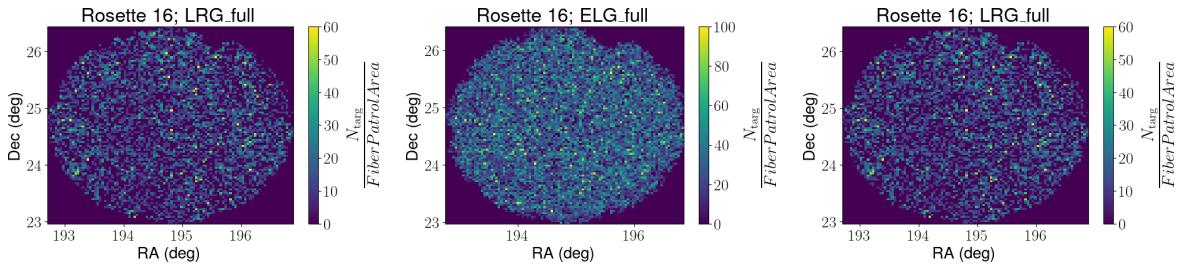


Figure 14. 2D histogram showing the areal number density (in units of targets per fiber patrol area) of targets in bins of RA and Dec of 25 stacked mocks within Rosette 16 of SV3 for (from left to right) LRGs, ELGs, and QSOs.

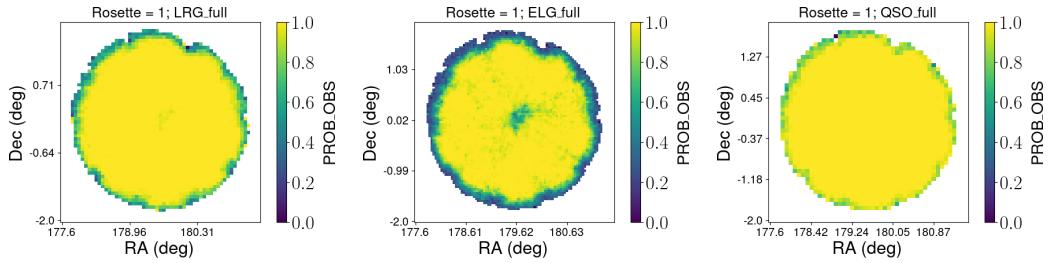


Figure 15. The same as figure 13, but for mock rosette 1.

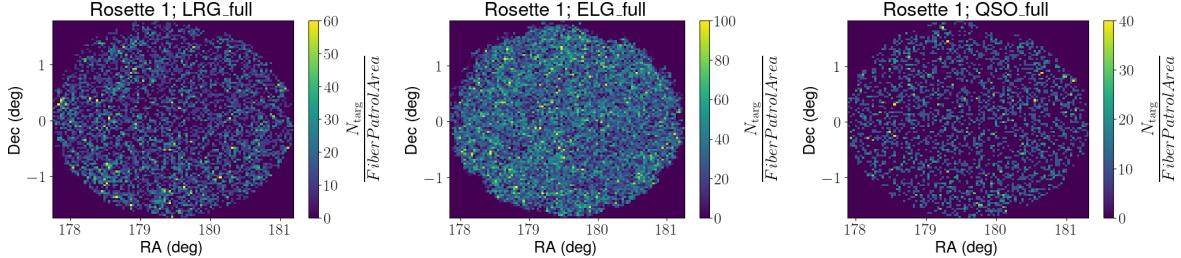


Figure 16. The same as figure 14 but for mock Rosette 1.

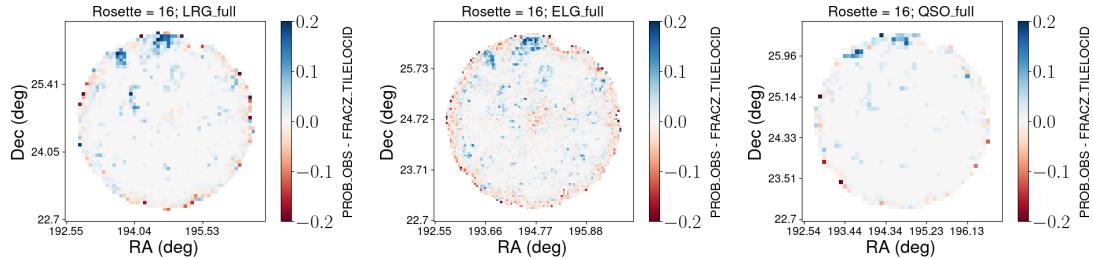


Figure 17. 2D histogram showing the mean value of the difference between PROB_OBS and FRACZ_TILELOCID in bins of RA and Dec from 25 stacked mocks within Rosette 16 of SV3 for (left to right) LRGs, ELGs, and QSOs.

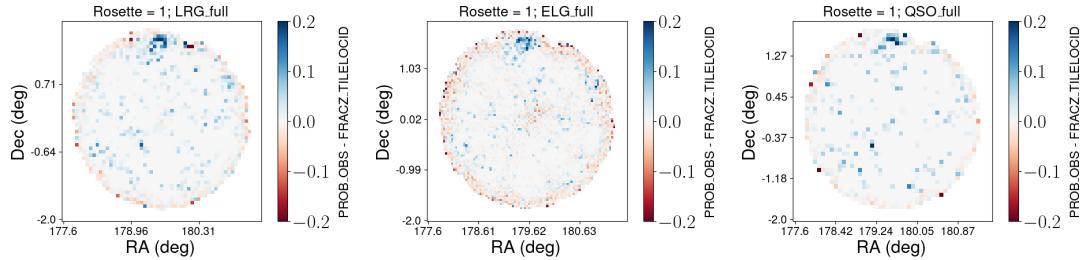


Figure 18. The same as figure 17 but for Rosette 1.

The inset on each plot shows the fractional difference between each weighting scheme with the “truth” with the blue (PIP weighted) line having error bars equal to the ratio of the “truth” errorbar and its value in each bin.

These plots show that the bitweights produced by the AMTL method are able to recover the parent clustering well within $1 - \sigma$ for all tracers down to $0.05 \text{ Mpc } h^{-1}$. Furthermore, the PIP weighting improves the recovery of the parent catalog clustering in all but a few separation bins.

The case of ELGs is most informative, since ELGs are the least complete target class. Clustering measured using only the default weighting is off by 20% uniformly down to a scale of nearly $1 \text{ Mpc } h^{-1}$. Additionally, the addition of angular upweights causes a spurious shift which pushes the 20% discrepancy down to a scale over $10 \text{ Mpc } h^{-1}$. The PIP weights when combined with the above schemes shows agreement within a few percent between scales of $0.05 \text{ Mpc } h^{-1}$ and $25 \text{ Mpc } h^{-1}$ with agreement within $1 - \sigma$ at all scales shown.

7 Discussion

Comparing the statistics of the data AMTL with the data shows a small, but persistent $\sim 1\%$ excess in the number of observed targets. The mocks also show a similar excess. This is likely due to the AMTL pipeline not fully accurately sampling observation results from the data for the alternate targets. As mentioned in section 4.4, we take observation results in the AMTL from the exact same fiber in the observations without regard to the target type which is assigned in the alternate realizations. In a future release of the AMTL pipeline we will have a fiber twin method implemented which will select observation results from a nearby fiber with nearly identical observational properties including target type. While the number of total observations shows some disagreements, the shape of the PROB_OBS distribution is recovered very well. For Rosette 16, the data PROB_OBS distributions in figure 7 and mock PROB_OBS distributions in figure 13 show nearly identical structure with the mocks showing much less fluctuation due to higher statistics and the data showing a small decrease in PROB_OBS towards the lower right of the rosette not present in the mocks due to a high target number density in that region. Rosette 1 also shows good agreement between figures 9 and 15 with the significantly higher PROB_OBS in the top left of the data rosette as compared to the mock coming from the very low data target number density.

The difference between PROB_OBS and the more generic FRACZ_TILELOCID in figures 11, 12, 17, and 18 shows similar patterns. There is very little difference between the two completeness metrics in the centers of both rosettes since completeness is at or near 100%

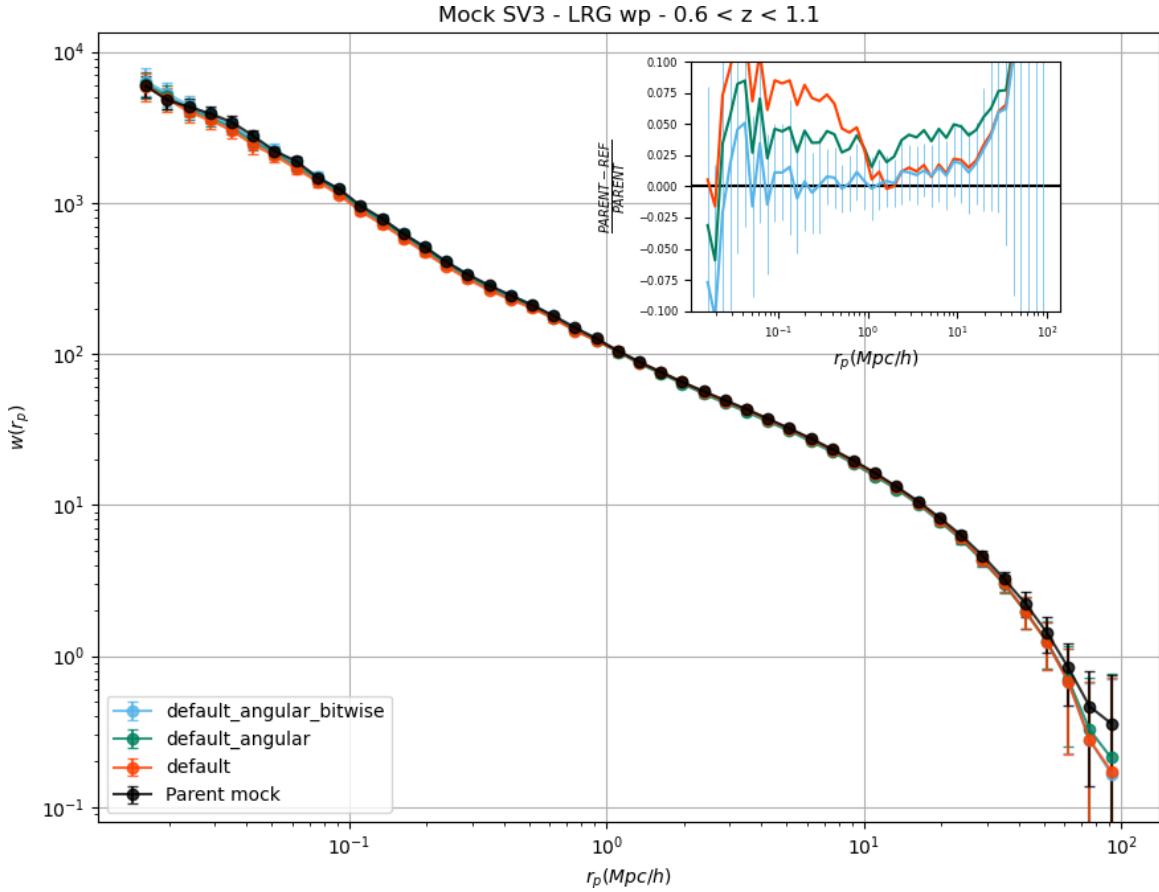


Figure 19. w_p from mean of 25 abacus mocks which were put through the AMTL pipeline with errorbars from the std. deviation of the 25 w_p measurements. The black line shows the true w_p from the mock parent catalog, the orange line is the observed subset with only default IIP completeness weights, the green line is the observed subset with default IIP completeness weights and the angular upweighting of PB17, and the blue line is the observed subset with angular upweighting and the PIP weights developed in BP17 and computed in this work. The inset shows the fractional difference of the clustering of each of the three “observed” samples with that of the parent catalog. The PIP weighted measurement has error bars equal to the std. deviation over the 25 mock parent w_p measurements divided by the mean parent w_p measurement.

at all locations. Towards the edges, we see a significant amount of fluctuation between the PROB_OBS and FRACZ_TILELOCID suggesting that completeness is somewhat different when measured through more approximate methods than an AMTL method.

Figures 19, 20, and 21 show the validation of the AMTL method by using the bitweights obtained by using the method on mocks to recover the clustering of the parent mocks. Since the true redshift of all targets in the mock are known, this enables us to validate against a known “truth” projected two point correlation function.

The validation results show that the bitwise weights calculated from the AMTL method are a strict improvement over the default completeness weights at small scales. For all three dark time tracers, the (blue) PIP+ANG upweighted mock projected two point correlation functions are an improvement of the default completeness weights at nearly all small (≤ 1 Mpc

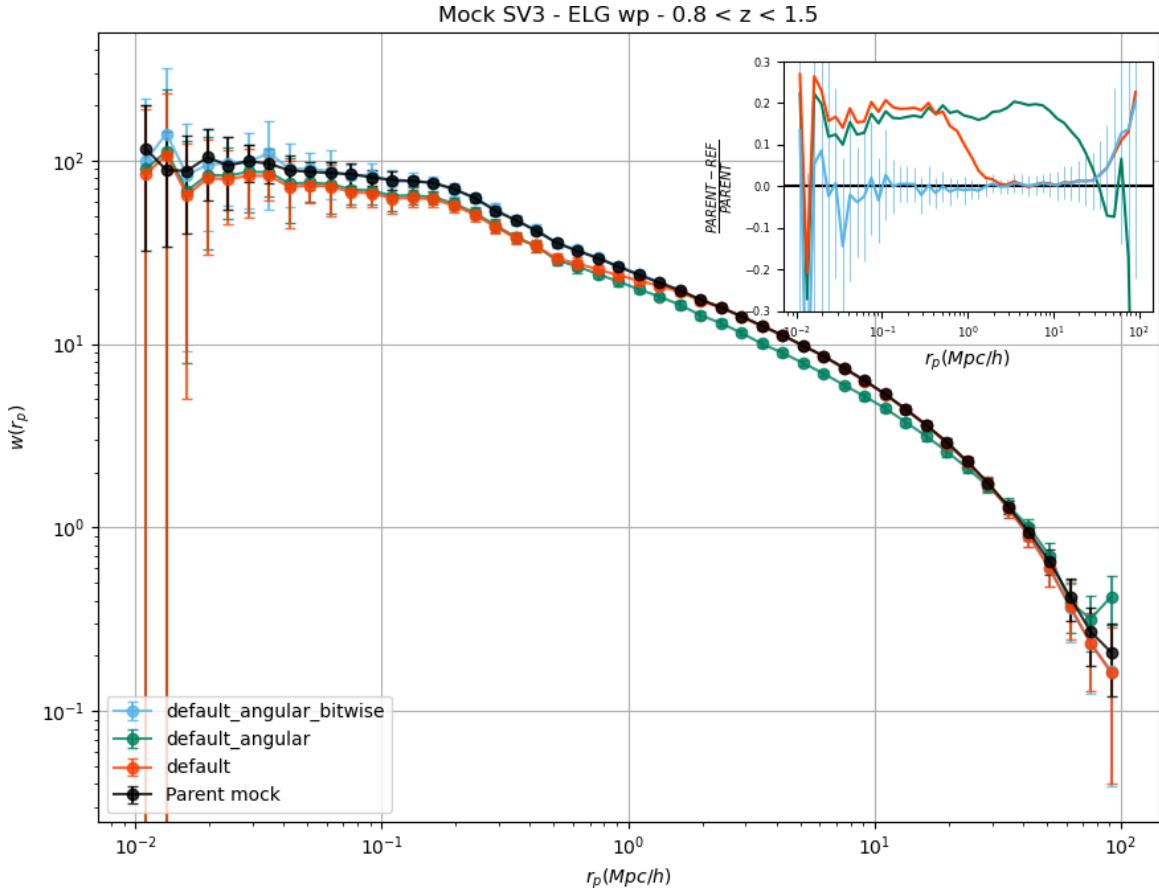


Figure 20. Same as figure 19 for ELGs.

h^{-1}) scales. Furthermore, down to scales of 0.02, 0.01, and 0.1 $\text{Mpc } h^{-1}$ (for LRGs, ELGs, and QSOs respectively), the PIP+ANG upweighted clustering from the mocks is consistent with the parent mock clustering within the scatter among the 25 mocks. These results, especially the ELG clustering, show that the PIP weighting and angular upweighting are both necessary to correct the clustering measurements for incompleteness. For a sample like SV3 which is highly complete, IIP weights function similarly to nearest neighbor weights, indicating that those weights would be insufficient for incompleteness correction at small scales.

The larger scale clustering ($s \geq 1 \text{ Mpc } h^{-1}$ and especially $s \geq 10 \text{ Mpc } h^{-1}$), while statistically consistent with the parent clustering within the mock variance, is not as well recovered. However, there are several reasons why this should be the case. First, the SV3 rosette design strongly limits the number of pairs which we can observe at separations nearing $100 \text{ Mpc } h^{-1}$. The size of an individual rosette at a redshift $z=1.0$ is only $\sim 70 \text{ Mpc } h^{-1}$. While DESI did place some rosettes adjacent to each other enabling the observation of some pairs at those separations, the vast majority of rosettes were isolated.

In conclusion, we have demonstrated the generation of 128 realizations of DESI SV3 and Y1 of the main survey using a nearly-maximally-realistic method utilizing alternative merged target ledgers. We have validated the AMTL method on 256 realizations of DESI SV3-like mocks showing the recovery of clustering down to scales of $0.01 \text{ Mpc } h^{-1}$ for the least complete tracer, Emission Line Galaxies.

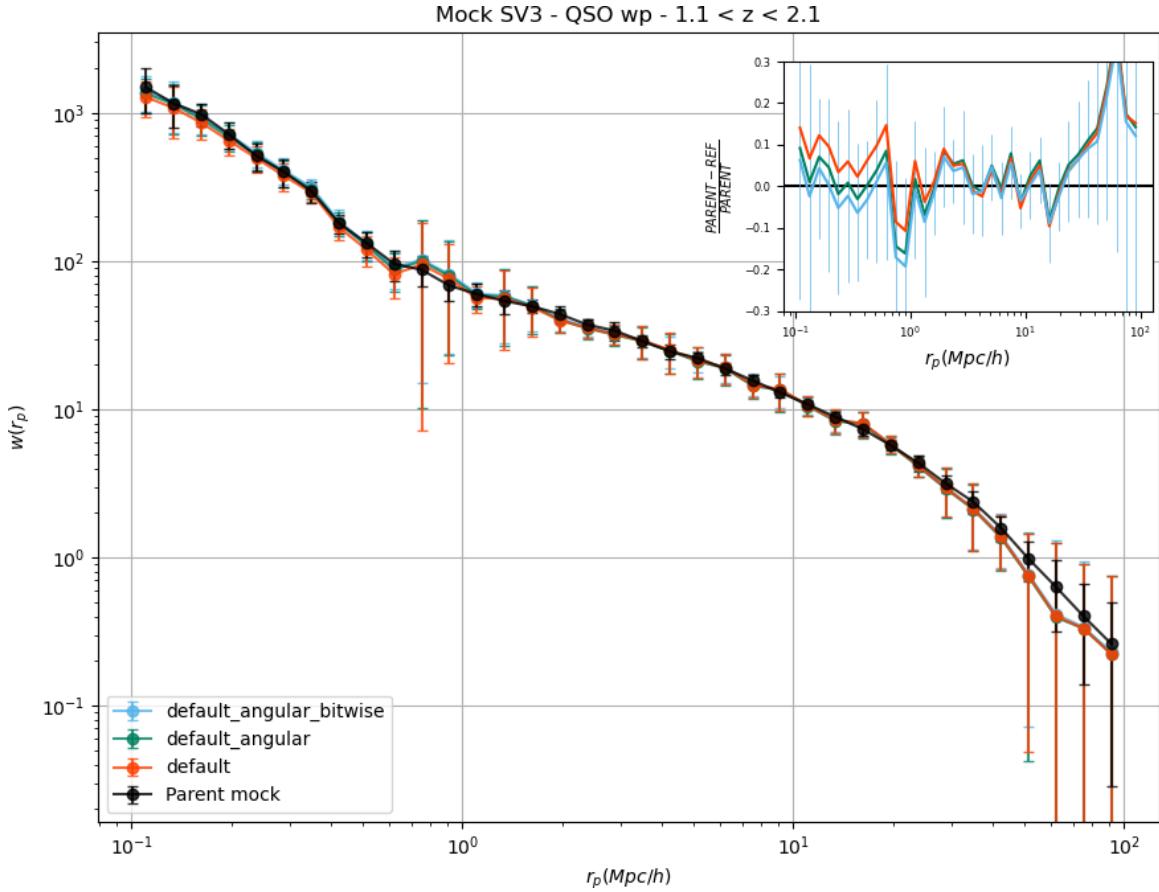


Figure 21. Same as figure 19 for QSOs.

Acknowledgments

The authors would like to thank the reviewer, Dr. Martin Reinecke, for his careful review of our paper and his useful and well-formatted comments.

JL and RK acknowledge support of the U.S. Department of Energy (DOE) in funding grant DE-SC0010129 for the work in this paper. Computational resources for JL were provided by SMU's Center for Research Computing. This work made use of Astropy:¹² a community-developed core Python package and an ecosystem of tools and resources for astronomy [66–68] ADM was supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0019022.

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

¹²<http://www.astropy.org>.

Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National 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>. 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.

8 Data availability

The SV3 data used in this analysis is currently available to the public at <https://data.desi.lbl.gov/doc/releases/edr/>. As per the DESI Data Management policy, all data points shown in the figures in this paper will be available in a machine-readable form on Zenodo with DOI 10.5281/zenodo.13308269 The Y1 data used in this analysis will be made public along the Data Release 1 (details in <https://data.desi.lbl.gov/doc/releases/>)

9 Code availability

The code used to generate the 128 realizations of the SV3 data and 256 realizations of the SV3 AbacusSummit mocks is tagged in the desihub/LSS github repository: <https://github.com/desihub/LSS/releases/tag/v2.0.0-EDR>. A script is provided which uses the same parameters as were used in the 128 data realizations presented in this paper to run the pipeline (SurveyAltMTLScript.sh). Two scripts are provided which use the same parameters as were used to generate the 256 mock realizations presented in this paper to run the pipeline (MockAltMTLScript_firstpart.sh and MockAltMTLScript_secondpart.sh).

This code was designed to work on the NERSC Cori and Perlmutter computational systems and was never tested on other computers. In order to run this on another computer/cluster, you will need to change directory paths in the provided scripts and also install the requisite DESI modules. The versions of these modules must be the same as used in the original EDR analyses. On Perlmutter, these are collected under the module version desimodules/fuji. The list of libraries and versions which comprise that module is provided here: <https://data.desi.lbl.gov/doc/releases/edr/software-version/>. The relevant modules for this analysis are LSS, desitarget, desihub, and fiberassign.

Due to the usage of fiberassign at varying times throughout the survey, versions 2.2.0, 2.3.0, 2.4.0, 2.5.0, 2.5.1, 3.0.0, and 4.0.0 are needed in addition to the final EDR version 5.4.0. The default pipeline uses the swap feature of the Environment Module¹³ system on Perlmutter to change between versions of fiberassign as necessary. In order to run on another computer, users would need to include the relevant fiberassign versions above into their own module system.

¹³<https://modules.sourceforge.net/>.

References

- [1] SDSS collaboration, *The Sloan Digital Sky Survey: Technical Summary*, *Astron. J.* **120** (2000) 1579 [[astro-ph/0006396](#)] [[INSPIRE](#)].
- [2] BOSS collaboration, *The Baryon Oscillation Spectroscopic Survey of SDSS-III*, *Astron. J.* **145** (2013) 10 [[arXiv:1208.0022](#)] [[INSPIRE](#)].
- [3] EBOSS collaboration, *The SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Overview and Early Data*, *Astron. J.* **151** (2016) 44 [[arXiv:1508.04473](#)] [[INSPIRE](#)].
- [4] DESI collaboration, *Overview of the Instrumentation for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **164** (2022) 207 [[arXiv:2205.10939](#)] [[INSPIRE](#)].
- [5] DESI collaboration, *Survey Operations for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **166** (2023) 259 [[arXiv:2306.06309](#)] [[INSPIRE](#)].
- [6] D. Kirkby et al., in preparation (2024).
- [7] M.R. Blanton et al., *An Efficient Algorithm for Positioning Tiles in the Sloan Digital Sky Survey*, *Astron. J.* **125** (2003) 2276 [[astro-ph/0105535](#)] [[INSPIRE](#)].
- [8] H. Gao et al., *The DESI One-Percent Survey: Constructing Galaxy-Halo Connections for ELGs and LRGs Using Auto and Cross Correlations*, *Astrophys. J.* **954** (2023) 207 [[arXiv:2306.06317](#)] [[INSPIRE](#)].
- [9] A.N. Pearl et al., *The DESI One-percent Survey: Evidence for Assembly Bias from Low-redshift Counts-in-cylinders Measurements*, *Astrophys. J.* **963** (2024) 116 [[arXiv:2309.08675](#)] [[INSPIRE](#)].
- [10] A. Rocher et al., *The DESI One-Percent survey: exploring the Halo Occupation Distribution of Emission Line Galaxies with AbacusSummit simulations*, *JCAP* **10** (2023) 016 [[arXiv:2306.06319](#)] [[INSPIRE](#)].
- [11] A. Smith et al., *Generating mock galaxy catalogues for flux-limited samples like the DESI Bright Galaxy Survey*, *Mon. Not. Roy. Astron. Soc.* **532** (2024) 903 [[arXiv:2312.08792](#)] [[INSPIRE](#)].
- [12] S. Yuan et al., *The DESI one-per cent survey: exploring the halo occupation distribution of luminous red galaxies and quasi-stellar objects with AbacusSummit*, *Mon. Not. Roy. Astron. Soc.* **530** (2024) 947 [[arXiv:2306.06314](#)] [[INSPIRE](#)].
- [13] DESI collaboration, *Unraveling emission line galaxy conformity at $z \sim 1$ with DESI early data*, [arXiv:2310.09329](#) [[INSPIRE](#)].
- [14] J. Yu et al., *The DESI One-Percent Survey: Exploring A Generalized SHAM for Multiple Tracers with the UNIT Simulation*, *Mon. Not. Roy. Astron. Soc.* **527** (2024) 6950 [[arXiv:2306.06313](#)] [[INSPIRE](#)].
- [15] H. Gao et al., *The DESI One-Percent Survey: A Concise Model for the Galactic Conformity of Emission-line Galaxies*, *Astrophys. J.* **961** (2024) 74 [[arXiv:2309.03802](#)] [[INSPIRE](#)].
- [16] F. Prada et al., *The DESI One-Percent Survey: Modelling the clustering and halo occupation of all four DESI tracers with Uchuu*, [arXiv:2306.06315](#) [[INSPIRE](#)].
- [17] BOSS collaboration, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: Baryon Acoustic Oscillations in the Data Release 9 Spectroscopic Galaxy Sample*, *Mon. Not. Roy. Astron. Soc.* **427** (2013) 3435 [[arXiv:1203.6594](#)] [[INSPIRE](#)].
- [18] H. Guo, I. Zehavi and Z. Zheng, *A New Method to Correct for Fiber Collisions in Galaxy Two-Point Statistics*, *Astrophys. J.* **756** (2012) 127 [[arXiv:1111.6598](#)] [[INSPIRE](#)].

- [19] B.A. Reid et al., *A 2.5 per cent measurement of the growth rate from small-scale redshift space clustering of SDSS-III CMASS galaxies*, *Mon. Not. Roy. Astron. Soc.* **444** (2014) 476 [[arXiv:1404.3742](https://arxiv.org/abs/1404.3742)] [[INSPIRE](#)].
- [20] BOSS collaboration, *The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: baryon acoustic oscillations in the Data Releases 10 and 11 Galaxy samples*, *Mon. Not. Roy. Astron. Soc.* **441** (2014) 24 [[arXiv:1312.4877](https://arxiv.org/abs/1312.4877)] [[INSPIRE](#)].
- [21] BOSS collaboration, *SDSS-III Baryon Oscillation Spectroscopic Survey Data Release 12: galaxy target selection and large scale structure catalogues*, *Mon. Not. Roy. Astron. Soc.* **455** (2016) 1553 [[arXiv:1509.06529](https://arxiv.org/abs/1509.06529)] [[INSPIRE](#)].
- [22] EBOSS collaboration, *The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Large-scale structure catalogues for cosmological analysis*, *Mon. Not. Roy. Astron. Soc.* **498** (2020) 2354 [[arXiv:2007.09000](https://arxiv.org/abs/2007.09000)] [[INSPIRE](#)].
- [23] EBOSS collaboration, *The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release*, *Astrophys. J. Suppl.* **250** (2020) 8 [[arXiv:2007.09001](https://arxiv.org/abs/2007.09001)] [[INSPIRE](#)].
- [24] EBOSS collaboration, *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*, *Mon. Not. Roy. Astron. Soc.* **500** (2020) 3254 [[arXiv:2007.09007](https://arxiv.org/abs/2007.09007)] [[INSPIRE](#)].
- [25] D. Bianchi and W.J. Percival, *Unbiased clustering estimation in the presence of missing observations*, *Mon. Not. Roy. Astron. Soc.* **472** (2017) 1106 [[arXiv:1703.02070](https://arxiv.org/abs/1703.02070)] [[INSPIRE](#)].
- [26] F.G. Mohammad et al., *The VIMOS Public Extragalactic Redshift Survey (VIPERS): Unbiased clustering estimate with VIPERS slit assignment*, *Astron. Astrophys.* **619** (2018) A17 [[arXiv:1807.05999](https://arxiv.org/abs/1807.05999)] [[INSPIRE](#)].
- [27] W.J. Percival and D. Bianchi, *Using angular pair upweighting to improve 3D clustering measurements*, *Mon. Not. Roy. Astron. Soc.* **472** (2017) L40 [[arXiv:1703.02071](https://arxiv.org/abs/1703.02071)] [[INSPIRE](#)].
- [28] EBOSS collaboration, *The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Pairwise-Inverse-Probability and Angular Correction for Fibre Collisions in Clustering Measurements*, *Mon. Not. Roy. Astron. Soc.* **498** (2020) 128 [[arXiv:2007.09005](https://arxiv.org/abs/2007.09005)] [[INSPIRE](#)].
- [29] DESI collaboration, *The DESI Experiment, a whitepaper for Snowmass 2013*, [arXiv:1308.0847](https://arxiv.org/abs/1308.0847) [[INSPIRE](#)].
- [30] DESI collaboration, *The DESI Experiment Part I: Science, Targeting, and Survey Design*, [arXiv:1611.00036](https://arxiv.org/abs/1611.00036) [[INSPIRE](#)].
- [31] DESI collaboration, *The DESI Experiment Part II: Instrument Design*, [arXiv:1611.00037](https://arxiv.org/abs/1611.00037) [[INSPIRE](#)].
- [32] DESI collaboration, *DESI 2024 I: Data Release 1 of the Dark Energy Spectroscopic Instrument*, in preparation (2025).
- [33] DESI collaboration, *DESI 2024 II: Sample definitions, characteristics and two-point clustering statistics*, in preparation (2024).
- [34] DESI collaboration, *DESI 2024 III: Baryon Acoustic Oscillations from Galaxies and Quasars*, [arXiv:2404.03000](https://arxiv.org/abs/2404.03000) [[INSPIRE](#)].
- [35] DESI collaboration, *DESI 2024 IV: Baryon Acoustic Oscillations from the Lyman Alpha Forest*, [arXiv:2404.03001](https://arxiv.org/abs/2404.03001) [[INSPIRE](#)].

- [36] DESI collaboration, *DESI 2024 V: Analysis of the full shape of two-point clustering statistics from galaxies and quasars*, in preparation (2024).
- [37] DESI collaboration, *DESI 2024 VI: Cosmological Constraints from the Measurements of Baryon Acoustic Oscillations*, [arXiv:2404.03002](https://arxiv.org/abs/2404.03002) [INSPIRE].
- [38] DESI collaboration, *DESI 2024 VII: Cosmological constraints from full-shape analyses of the two-point clustering statistics measurements*, in preparation (2024).
- [39] DESI collaboration, *DESI 2024 VIII: Constraints on Primordial Non-Gaussianities*, in preparation (2024).
- [40] C. Allende Prieto et al., *Preliminary Target Selection for the DESI Milky Way Survey (MWS)*, *Res. Notes AAS* **4** (2020) 188 [[arXiv:2010.11284](https://arxiv.org/abs/2010.11284)] [INSPIRE].
- [41] O. Ruiz-Macias et al., *Preliminary Target Selection for the DESI Bright Galaxy Survey (BGS)*, *Res. Notes AAS* **4** (2020) 187 [[arXiv:2010.11283](https://arxiv.org/abs/2010.11283)] [INSPIRE].
- [42] R. Zhou et al., *Preliminary Target Selection for the DESI Luminous Red Galaxy (LRG) Sample*, *Res. Notes AAS* **4** (2020) 181 [[arXiv:2010.11282](https://arxiv.org/abs/2010.11282)] [INSPIRE].
- [43] A. Raichoor et al., *Preliminary Target Selection for the DESI Emission Line Galaxy (ELG) Sample*, *Res. Notes AAS* **4** (2020) 180 [[arXiv:2010.11281](https://arxiv.org/abs/2010.11281)] [INSPIRE].
- [44] C. Yèche et al., *Preliminary Target Selection for the DESI Quasar (QSO) Sample*, *Res. Notes AAS* **4** (2020) 179 [[arXiv:2010.11280](https://arxiv.org/abs/2010.11280)] [INSPIRE].
- [45] DESI collaboration, *Overview of the DESI Legacy Imaging Surveys*, *Astron. J.* **157** (2019) 168 [[arXiv:1804.08657](https://arxiv.org/abs/1804.08657)] [INSPIRE].
- [46] DESI collaboration, *Validation of the Scientific Program for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **167** (2024) 62 [[arXiv:2306.06307](https://arxiv.org/abs/2306.06307)] [INSPIRE].
- [47] DESI collaboration, *The Spectroscopic Data Processing Pipeline for the Dark Energy Spectroscopic Instrument*, *Astron. J.* **165** (2023) 144 [[arXiv:2209.14482](https://arxiv.org/abs/2209.14482)] [INSPIRE].
- [48] DESI collaboration, *The Early Data Release of the Dark Energy Spectroscopic Instrument*, *Astron. J.* **168** (2024) 58 [[arXiv:2306.06308](https://arxiv.org/abs/2306.06308)] [INSPIRE].
- [49] R.H. Hardin, N.J.A. Sloane and W.D. Smith, *Tables of Spherical Codes with Icosahedral Symmetry*, AT&T Shannon Lab., Florham Park (2001), accessed 02-12-2024.
- [50] C. Zhao et al., *Mock catalogues with survey realism for the DESI DR1*, in preparation (2024).
- [51] D. Bianchi et al., *Characterization of DESI fiber assignment incompleteness effect on 2-point clustering and mitigation methods for 2024 analysis*, in preparation (2024).
- [52] N.A. Maksimova et al., *AbacusSummit: a massive set of high-accuracy, high-resolution N-body simulations*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 4017 [[arXiv:2110.11398](https://arxiv.org/abs/2110.11398)] [INSPIRE].
- [53] L.H. Garrison et al., *The abacus cosmological N-body code*, *Mon. Not. Roy. Astron. Soc.* **508** (2021) 575 [[arXiv:2110.11392](https://arxiv.org/abs/2110.11392)] [INSPIRE].
- [54] PLANCK collaboration, *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13 [[arXiv:1502.01589](https://arxiv.org/abs/1502.01589)] [INSPIRE].
- [55] D. Bianchi et al., *Unbiased clustering estimates with the DESI fibre assignment*, *Mon. Not. Roy. Astron. Soc.* **481** (2018) 2338 [[arXiv:1805.00951](https://arxiv.org/abs/1805.00951)] [INSPIRE].
- [56] A. Smith et al., *Correcting for Fibre Assignment Incompleteness in the DESI Bright Galaxy Survey*, *Mon. Not. Roy. Astron. Soc.* **484** (2019) 1285 [[arXiv:1809.07355](https://arxiv.org/abs/1809.07355)] [INSPIRE].

Author List

J. Lasker ^{id}^{a,*}, A. Carnero Rosell ^{id}^{b,c}, A.D. Myers^d, A.J. Ross ^{id}^{e,f,g}, D. Bianchi ^{id}^h,
 M.M. S Hanif ^{id}ⁱ, R. Kehoe^a, A. de Mattia^j, L. Napolitano ^{id}^d, W.J. Percival ^{id}^{k,l,m}, R. Staten^a,
 J. Aguilarⁿ, S. Ahlen ^{id}^o, L. Bigwood^p, D. Brooks^q, T. Claybaughⁿ, S. Cole ^{id}^p, A. de la Macorra ^{id}^r,
 Z. Ding ^{id}^s, P. Doel^q, K. Fanning ^{id}^{t,u}, J.E. Forero-Romero ^{id}^{v,w}, E. Gaztañaga^{x,y,z},
 S. Gontcho A Gontcho ^{id}ⁿ, G. Gutierrez^{aa}, K. Honscheid^{e,ab,g}, C. Howlett ^{id}^{ac}, S. Juneau^{ad},
 A. Kremin ^{id}ⁿ, M. Landriau ^{id}ⁿ, L. Le Guillou ^{id}^{ae}, M.E. Levi ^{id}ⁿ, M. Manera ^{id}^{af,ag}, A. Meisner ^{id}^{ad},
 R. Miquel^{ah,ag}, J. Moustakas ^{id}^{ai}, E. Mueller^{aj}, J. Nie ^{id}^{ak}, G. Niz ^{id}^{al,am}, M. Oh^{an},
 N. Palanque-Delabrouille ^{id}^{j,n}, C. Poppett^{n,ao,ap}, F. Prada ^{id}^{aq}, M. Rezaie ^{id}^{ar}, G. Rossi^{as},
 E. Sanchez ^{id}^{at}, D. Schlegelⁿ, M. Schubnell^{au,i}, H. Seo ^{id}^{av}, D. Sprayberry^{ad}, G. Tarlé ^{id}ⁱ,
 M. Vargas-Magaña ^{id}^r, B.A. Weaver^{ad}, Michael J. Wilson^p, Y. Zheng^s

^a Department of Physics, Southern Methodist University,
 3215 Daniel Avenue, Dallas, TX 75275, U.S.A.

^b Departamento de Astrofísica, Universidad de La Laguna (ULL),
 E-38206, La Laguna, Tenerife, Spain

^c Instituto de Astrofísica de Canarias, C/ Vía Láctea, s/n, E-38205 La Laguna, Tenerife, Spain

^d Department of Physics & Astronomy, University of Wyoming,
 1000 E. University, Dept. 3905, Laramie, WY 82071, U.S.A.

^e Center for Cosmology and AstroParticle Physics, The Ohio State University,
 191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.

^f Department of Astronomy, The Ohio State University,
 4055 McPherson Laboratory, 140 W 18th Avenue, Columbus, OH 43210, U.S.A.

^g The Ohio State University, Columbus, 43210 OH, U.S.A.

^h Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano,
 Via Celoria 16, I-20133 Milano, Italy

ⁱ University of Michigan, Ann Arbor, MI 48109, U.S.A.

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

^k Department of Physics and Astronomy, University of Waterloo,
 200 University Ave W, Waterloo, ON N2L 3G1, Canada

^l Perimeter Institute for Theoretical Physics,
 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada

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

ⁿ Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, U.S.A.

^o Physics Department, Boston University, 590 Commonwealth Avenue, Boston, MA 02215, U.S.A.

^p Institute for Computational Cosmology, Department of Physics, Durham University,
 South Road, Durham DH1 3LE, U.K.

^q Department of Physics & Astronomy, University College London,
 Gower Street, London, WC1E 6BT, U.K.

^r Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, México

^s Department of Astronomy, School of Physics and Astronomy, Shanghai Jiao Tong University,
 Shanghai 200240, China

^t Kavli Institute for Particle Astrophysics and Cosmology, Stanford University,
 Menlo Park, CA 94305, U.S.A.

^u SLAC National Accelerator Laboratory, Menlo Park, CA 94305, U.S.A.

^v Departamento de Física, Universidad de los Andes,
 Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia

^w Observatorio Astronómico, Universidad de los Andes,
 Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia

^x Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain

- ^y Institute of Cosmology and Gravitation, University of Portsmouth,
Dennis Sciama Building, Portsmouth, PO1 3FX, U.K.
- ^z Institute of Space Sciences, ICE-CSIC, Campus UAB,
Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain
- ^{aa} Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, IL 60510, U.S.A.
- ^{ab} Department of Physics, The Ohio State University,
191 West Woodruff Avenue, Columbus, OH 43210, U.S.A.
- ^{ac} School of Mathematics and Physics, University of Queensland, 4072, Australia
- ^{ad} NSF NOIRLab, 950 N. Cherry Ave., Tucson, AZ 85719, U.S.A.
- ^{ae} Sorbonne Université, CNRS/IN2P3, Laboratoire de Physique Nucléaire et de Hautes Energies (LPNHE), FR-75005 Paris, France
- ^{af} Departament de Física, Serra Húnter, Universitat Autònoma de Barcelona,
08193 Bellaterra (Barcelona), Spain
- ^{ag} Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology,
Campus UAB, 08193 Bellaterra Barcelona, Spain
- ^{ah} Institució Catalana de Recerca i Estudis Avançats,
Passeig de Lluís Companys, 23, 08010 Barcelona, Spain
- ^{ai} Department of Physics and Astronomy, Siena College,
515 Loudon Road, Loudonville, NY 12211, U.S.A.
- ^{aj} Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, U.K.
- ^{ak} National Astronomical Observatories, Chinese Academy of Sciences,
A20 Datun Rd., Chaoyang District, Beijing, 100012, P.R. China
- ^{al} Departamento de Física, Universidad de Guanajuato – DCI, C.P. 37150, Leon, Guanajuato, México
- ^{am} Instituto Avanzado de Cosmología A. C.,
San Marcos 11 – Atenas 202. Magdalena Contreras, 10720, Ciudad de México, México
- ^{an} Korea Astronomy and Space Science Institute,
776, Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Republic of Korea
- ^{ao} Space Sciences Laboratory, University of California, Berkeley,
7 Gauss Way, Berkeley, CA 94720, U.S.A.
- ^{ap} University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, U.S.A.
- ^{aq} Instituto de Astrofísica de Andalucía (CSIC),
Glorieta de la Astronomía, s/n, E-18008 Granada, Spain
- ^{ar} Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, U.S.A.
- ^{as} Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Korea
- ^{at} CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain
- ^{au} Department of Physics, University of Michigan, Ann Arbor, MI 48109, U.S.A.
- ^{av} Department of Physics & Astronomy, Ohio University, Athens, OH 45701, U.S.A.

* Corresponding author