Local primordial non-Gaussianity from the large-scale clustering of photometric DESI luminous red galaxies

Mehdi Rezaie[1], Ashley J. Ross[2], Hee-Jong Seo[3], Hui Kong[3], Edmond Chaussidon[4],
Anna Porredon[3], Lado Samushia[1], Rongpu Zhou[5], Alex Krolewski[6,7,8], Arnaud de Mattia[4],
Jose Bermejo[?], Florian Beutler[9], Christophe Yeche[4], Nathalie Palanque-Delabrouille[4,5],
Klaus Honscheid[2,10], Santi Avila, Eva Mueller, Will Percival, Shadab Alam,
Benedict Bahr-Kalus, Chia-Hsun Chuang, Noah Weaverdyck, Simone Ferraro,
Tanveer Karim, and Adam Myers

[1]*Department of Physics, Kansas State University, 116 Cardwell Hall, Manhattan, KS 66506, USA*

[2]*Department of Physics and Astronomy, Ohio University, Athens, OH 45701, USA*

[3]*Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*

[4]*IRFU, CEA, Universite Paris-Saclay, F-91191 Gif-sur-Yvette, France*

[5]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA*

[6]*Department of Physics and Astronomy, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*

[7]*Perimeter Institute for Theoretical Physics, 31 Caroline St. North, Waterloo, ON N2L 2Y5, Canada*

[8]*Waterloo Centre for Astrophysics, University of Waterloo, 200 University Ave W, Waterloo, ON N2L 3G1, Canada*

[9]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK*

[10]*Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA*

ABSTRACT

This paper uses the angular power spectrum of luminous red galaxies selected from the Dark Energy Spectroscopic Instrument (DESI) imaging surveys to probe local primordial non-Gaussianity (PNG). The FFTLog algorithm is applied to enable modeling the angular power spectrum on large scales. Linear multivariate regression and artificial neural networks are applied to alleviate excess clustering signals caused by varying survey conditions such as depth, extinction, and astronomical seeing. Feature selection and cross-correlation techniques are implemented to find various levels of systematic mitigation ranging from conservative to extreme. Assuming that only halo mass determines halo occupation, we find $36.08(25.03) < f_{\rm NL} < 61.44(75.64)$ and $13.10(-15.96) < f_{\rm NL} < 69.14(91.84)$ at $68\%(95\%)$ confidence, respectively, with our conservative and rigorous methods. The methods are tested and calibrated rigorously against lognormal density simulations with $f_{\rm NL} = 0$ and 76.9, and realistic systematic effects included. When tested against the simulations without systematics, significant issues are observed regarding photometric calibration in the South Galactic Cap below DEC= −30 and stellar contamination in the North Galactic Cap. Assuming Planck's measurement is accurate a priori, our results provide evidence for some unknown calibration or stellar contamination issues. While more theoretical work and simulations are required to better understand halo assembly bias, our results motivate followup studies of local PNG and observational systematics with DESI spectroscopic samples, which are less prone to calibration and contamination issues.

Key words: cosmology: inflation - large-scale structure of the Universe

1 INTRODUCTION

Cosmological observations support the paradigm of inflation in which the universe experienced a phase of rapid expansion in its early stage. Inflation elegantly addresses fundamental issues, such

as the isotropy, flatness, and homogeneity of the universe as well as the absence of magnetic monopoles (see, e.g., Weinberg et al. 2013, for a review). At the end of inflation, the universe was reheated, and primordial perturbations seeded the subsequent growth of structure (Kofman et al. 1994; Bassett et al. 2006; Lyth & Liddle 2009). De-

spite inflation being considered as a compelling scenario, properties of the field or fields driving the inflationary expansion still remain vastly unexplored in modern observational cosmology (see, e.g., Biagetti 2019, for a review). Early studies of the cosmic microwave background (CMB) and large-scale structure (LSS) suggested that initial conditions of the universe are consistent with Gaussian and scale-invariant fluctuations (Tegmark et al. 2004; Guth & Kaiser 2005); however, alternative classes of inflationary models predict some levels of non-Gaussianities in the primordial gravitational field. In its simplest form, primordial non-Gaussianity (PNG) depends on the local value of the primordial potential ϕ (Komatsu & Spergel 2001),

$$\Phi = \phi + f_{\rm NL} [\phi^2 - <\phi^2>], \qquad (1)$$

where $f_{\rm NL}$ is a nonlinear coupling constant and ϕ is assumed to be a Gaussian random field. Local-type PNG generates a primordial bispectrum, which peaks in the squeezed limit. It alters the local number density of galaxies by coupling the long and small wavelength modes of the dark matter gravitational field, and consequently, it induces a very distinct scale-dependent bias (Dalal et al. 2008).

Getting robust constraints on parameter $f_{\rm NL}$ is the first stepping stone toward better understanding the dynamics of the early universe. Standard slow-roll inflation predicts $f_{\rm NL} \sim 10^{-2}$ (see, e.g., Alvarez et al. 2014, for a review), while multi-field inflationary scenarios anticipate considerably larger values (see, e.g., de Putter et al. 2017). The current tightest bound on $f_{\rm NL}$ comes from Planck's bispectrum measurement of CMB anisotropies, $f_{\rm NL} = 0.9 \pm 5.0$ (Akrami et al. 2019). Limited by cosmic variance, CMB data alone cannot enhance statistical precision of $f_{\rm NL}$ measurements enough to break the degeneracy amongst various inflationary paradigms (see, e.g., Ade et al. 2019). However, combining CMB with LSS data could cancel cosmic variance, partially if not completely, and improve these numbers significantly to $\sigma(f_{\rm NL}) \sim 1$ (see, e.g., Schmittfull & Seljak 2018). Also, constraining $f_{\rm NL}$ with the three-point clustering of LSS is hindered by nonlinearities raised from structure formation, which is non-trivial to model and disentangle from the primordial signal (Baldauf et al. 2011b,a). As a novel approach, UV Luminosity Function probes galaxy abundances and structure formation on small scales (e.g., $k \sim 2~{\rm Mpc}^{-1}$), which are otherwise impossible to explore with the scale-dependent bias. Sabti et al. (2021) used UV Luminosity Function from the Hubble Space Telescope catalogs (Bouwens et al. 2015) to find a 2σ bound of $-166 < f_{\rm NL} < 497$. Even though this is still not competitive with the current bounds from CMB and LSS, upcoming surveys such as the James Webb Space Telescope and the Nancy Grace Roman Space Telescope are forecast to yield up to four times improvements on $f_{\rm NL}$ constraints from UV Luminosity Function. Given these observational limitations and theoretical challenges, the scale-dependent bias technique is the smoking gun for constraining local PNG with LSS. From theoretical point of view, further simulation-based studies of halo-assembly bias are required to fully utilize the scale-dependent bias (Barreira et al. 2020; Barreira 2020, 2022; Lazeyras et al. 2023). Nevertheless, the current best $f_{\rm NL}$ constraints from LSS achieve $\sigma(f_{\rm NL}) \sim 20 - 30$, and most of the constraining power comes from the two point statistics utilizing the scale-dependent bias rather than higher order statistics (Castorina et al. 2019; Mueller et al. 2022; Cabass et al. 2022; D'Amico et al. 2022).

In spite of theoretical challenges, measuring $f_{\rm NL}$ with the scale-dependent bias effect is nonetheless incredibly challenging due to various systematic effects that modulate clustering power on large scales, where the primordial signal is sensitive. For instance, survey geometry mixes clustering power on different angular scales (Beutler et al. 2014; Wilson et al. 2017). Relativistic effects also generate PNG-like scale-dependent signatures on large scales, which interfere with measuring $f_{\rm NL}$ with the scale-dependent bias effect (Wang et al. 2020). Similarly, matter density fluctuations with wavelengths larger than survey size, known as super-sample modes, modulate galaxy power spectrum (Castorina & Moradinezhad Dizgah 2020). Another source for systematic error is raised because the mean galaxy density for constructing the density contrast field is estimated from data directly rather than being known a priori. This integral constraint effect pushes clustering power on modes near survey size to zero (Peacock & Nicholson 1991; De Mattia & Ruhlmann-Kleider 2019). Accounting for these theoretical systematic effects in the galaxy power spectrum is crucial to obtain unbiased inference (see, e.g., Riquelme et al. 2022). Other observational systematics are driven by varying photometric conditions across the sky (Ross et al. 2011) and photometric calibration issues that manifest as spurious fluctuations in the observed density field of galaxies (Huterer et al. 2013). This type of systematic error is much more complicated to mitigate, compared to integral constraint and survey geometry, and it has hampered previous studies of local PNG with the scale-dependent bias effect (see, e.g., Pullen & Hirata 2013; Ho et al. 2015). These imaging systematic issues are expected to be severe for wide-area galaxy surveys that observe the night sky closer to the Galactic plane and attempt to implement more relaxed selection criteria to include fainter galaxies (see, e.g, Kitanidis et al. 2020).

As an ongoing wide-area galaxy survey, the Dark Energy Spectroscopic Instrument (DESI) uses 5000 robotically-driven fibers to simultaneously collect spectra. DESI is designed to deliver an unparalleled volume of spectroscopic data over its five year mission that will deepen our understanding of the equation of state for dark energy (Aghamousa et al. 2016). Assuming systematic errors are under control, DESI has the potential to provide competitive constraints on local PNG with $\sigma(f_{\rm NL}) = 5$ (Aghamousa et al. 2016). Unprecedented precision with $\sigma(f_{\rm NL}) \sim 1$ can be achieved by combining datasets from DESI and other upcoming surveys, such as Rubin Observatory and SphereX (see, e.g., Heinrich & Doré 2022).

The purpose of this paper is to probe local PNG using the scale-dependent bias signature in the angular power spectrum of luminous red galaxies selected from DESI imaging data. Linear multivariate regression and neural networks are applied to clean and prepare our sample for such a delicate signal. Our results are validated against suites of lognormal simulations with $f_{\rm NL} = 0$ and 76.9, and potential sources of systematic error such as survey depth, astronomical seeing, photometric calibration, Milky Way extinction, and local stellar density. A particular objective is to quantify the sensitivity of $f_{\rm NL}$ signal to various treatment methods. The cross power spectra between the galaxy density and imaging property maps are calculated to evaluate the effectiveness of different strategies for cleaning contaminations and to characterize the significance level of remaining systematic error.

This paper is structured as follows. Section 2 describes the galaxy sample from DESI imaging and lognormal simulations with, or without, PNG and synthetic systematic effects. Section 3 outlines the theoretical framework for modeling the angular power spectrum, strategies for handling various observational and theoretical systematic effects, and statistical techniques for measuring the significance of remaining systematics in our sample after mitigation. Our results are presented in Section 4, and Section 5 summarizes our conclusions and directions for future work.

2 DATA

Luminous red galaxies (LRGs) are massive galaxies that populate massive halos, lack active star formation, and are highly biased tracers of dark matter gravitational field. A distinct break around 4000 Å in the LRG spectrum is often utilized to determine their redshifts accurately. LRGs are widely targeted in previous galaxy redshift surveys (see, e.g., Eisenstein et al. 2001; Prakash et al. 2016), and their clustering and redshift properties are well studied (see, e.g., Ross et al. 2020; Gil-Marín et al. 2020; Bautista et al. 2021; Chapman et al. 2022).

DESI is designed to collect spectra of millions of LRGs covering the redshift range $0.4 < z < 1.0$. DESI spectroscopy selects its targets from the DESI Legacy Imaging Surveys. The Legacy Surveys consists of three ground-based surveys that provide photometry of the sky in the optical g, r, and z bands between 2014 and 2019: the Mayall z-band Legacy Survey using the Mayall telescope at Kitt Peak (MzLS; Dey et al. 2018), the Beijing–Arizona Sky Survey using the Bok telescope at Kitt Peak (BASS; Zou et al. 2017), and the Dark Energy Camera Legacy Survey on the Blanco 4m telescope (DECaLS Flaugher et al. 2015). The BASS and MzLS surveys observed the same footprint in the North Galactic Cap (NGC) while the DECaLS observed both caps around the galactic plane; the BASS+MzLS footprint is separated from the DECaLS NGC at DEC > 32.375 degrees, although there is an overlap between the two regions for calibration purposes (Dey et al. 2018). Additionally, the DECaLS program integrates observations executed from the Blanco instrument under the Dark Energy Survey (Abbott et al. 2016), which cover about $1130 \deg^2$ of the South Galactic Cap (SGC) footprint. The DESI imaging catalogs also integrate the 3.4 (W1) and 4.6 μm (W2) infrared photometry from the Wide-Field Infrared Explorer (WISE; Wright et al. 2010; Meisner et al. 2018).

2.1 DESI imaging LRGs

Our sample of LRGs is drawn from the DESI Legacy Imaging Surveys Data Release 9 (DR9; Dey et al. 2018) using the color-magnitude selection criteria designed for the DESI 1% survey, described as the SV3 selection in more detail in Zhou et al. (2022). The color-magnitude selection cuts are defined in the g, r, z bands in the optical and $W1$ band in the infrared, as summarized in Table 1. The selection cuts are developed differently for each imaging survey to reach an almost uniform target surface density despite different survey efficiency and photometric calibration between DECaLS and BASS+MzLS. The implementation of these selection cuts in the DESI data processing pipeline is explained in Myers et al. (2022). The redshift distribution of our galaxy sample is inferred from DESI spectroscopy during the Survey Validation phase. A constant clustering amplitude is assumed for linear bias which is supported by data (Zhou et al. 2021). Figure 1 shows the redshift distribution (solid black line) and the evolution of linear bias (dashed red line) for our sample of LRGs.

The LRG sample is masked rigorously for foreground bright stars, bright galaxies, and clusters of galaxies[1] to further reduce stellar contamination (Zhou et al. 2022). Then, the sample is binned into HEALPix (Gorski et al. 2005) pixels at NSIDE = 256 to construct the 2D density map with an average surface density of 800 galaxies per square degree covering around 14000 square degrees (as shown in
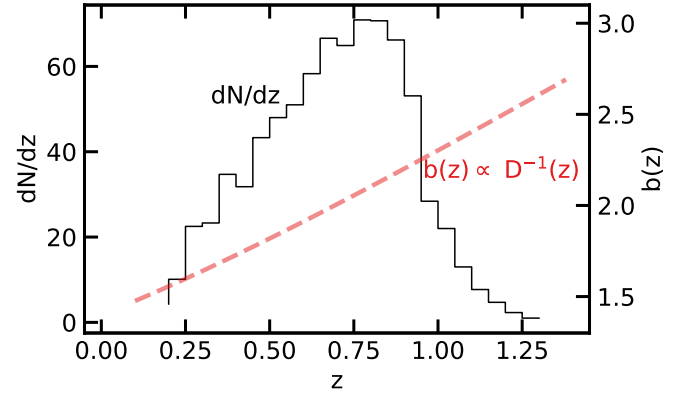
---

[1] See the maskbits at https://www.legacysurvey.org/dr9/bitmasks/

Figure 1. The redshift distribution (solid line) and bias evolution (dashed line) of DESI LRG targets. The redshift distribution is determined by DESI spectroscopy, and a constant clustering amplitude is assumed for linear bias.

Figure 2). The LRG density is corrected for the pixel incompleteness and lost areas using a catalog of random points, hereafter referred to as randoms, uniformly scattered over the footprint with the same cuts and masks applied. The LRG density map exhibits some large-scale spurious fluctuations, even though that DESI-like LRGs are selected brighter than the imaging survey depth limits. Specifically, the SGC footprint exhibits some systematic under-density while there is some systematic over-density near the survey boundaries in the NGC.

2.1.1 Imaging properties

In this paper, various imaging properties, which can be considered as potential sources of systematic error, are mapped into HEALPix of NSIDE= 256 (Figure 2). The maps include local stellar density constructed from point-like sources with a g-band magnitude in the range $12 \leq g < 17$ from the Gaia DR2 (see, Gaia Collaboration et al. 2018; Myers et al. 2022); Galactic extinction E[B-V] from Schlegel et al. (1998); survey depth (galaxy depth in g, r, and z and PSF depth in W1) and astronomical seeing (psfsize) in g, r, and z. These maps have been previously identified as potential sources of imaging systematics in DESI-like LRGs Zhou et al. (2022).

Each imaging map carries its characteristic fluctuations, which correlate with the LRG density map. For instance, large-scale fluctuations can be associated with stellar density, extinction, or survey depth; while small scale-fluctuations can be related to psfsize variations. Upon visual inspection, the under-dense regions of LRGs in the SGC can be associated with varying survey depth, while the over-dense regions in the NGC can be connected to galactic dust or extinction. Some regions of the DR9 footprint are removed from our analysis to avoid potential photometric calibration issues. Some of these regions are disconnected from the main footprint (e.g., spurious islands in the NGC) or different types of standard starts were used for their calibration (DEC< −30 in the SGC). The potential impact of these declination cuts on our f_{NL} constraints is explored in Section 4.

Figure 3 shows the Pearson correlation coefficient between the DESI LRG target density and the imaging systematics maps for the three imaging regions (DECaLS North, DECaLS South, and BASS+MzLS) in the top panel. The horizontal curves represent the 95% confidence regions, and are constructed by cross correlating 100 synthetic lognormal density fields and the imaging maps. Fig-

Table 1. Selection criteria for the DESI-like LRG targets (Zhou et al. 2022). Magnitudes are corrected for MW extinction. z_{fiber} represents the z-band fiber magnitude which corresponds to the expected flux within a DESI fiber.

| Footprint	Criterion	Description
DECaLS	$z_{fiber} < 21.7$	Faint limit
	$z - W1 > 0.8 \times (r - z) - 0.6$	Stellar rejection
	$[(g - r > 1.3) \text{ AND } ((g - r) > -1.55 * (r - W1) + 3.13)] \text{ OR } (r - W1 > 1.8)$	Remove low-z galaxies
	$[(r - W1 > (W1 - 17.26) * 1.8) \text{ AND } (r - W1 > W1 - 16.36)] \text{ OR } (r - W1 > 3.29)$	Luminosity cut
BASS+MzLS	$z_{fiber} < 21.71$	Faint limit
	$z - W1 > 0.8 \times (r - z) - 0.6$	Stellar rejection
	$[(g - r > 1.34) \text{ AND } ((g - r) > -1.55 * (r - W1) + 3.23)] \text{ OR } (r - W1 > 1.8)$	Remove low-z galaxies
	$[(r - W1 > (W1 - 17.24) * 1.83) \text{ AND } (r - W1 > W1 - 16.33)] \text{ OR } (r - W1 > 3.39)$	Luminosity cut
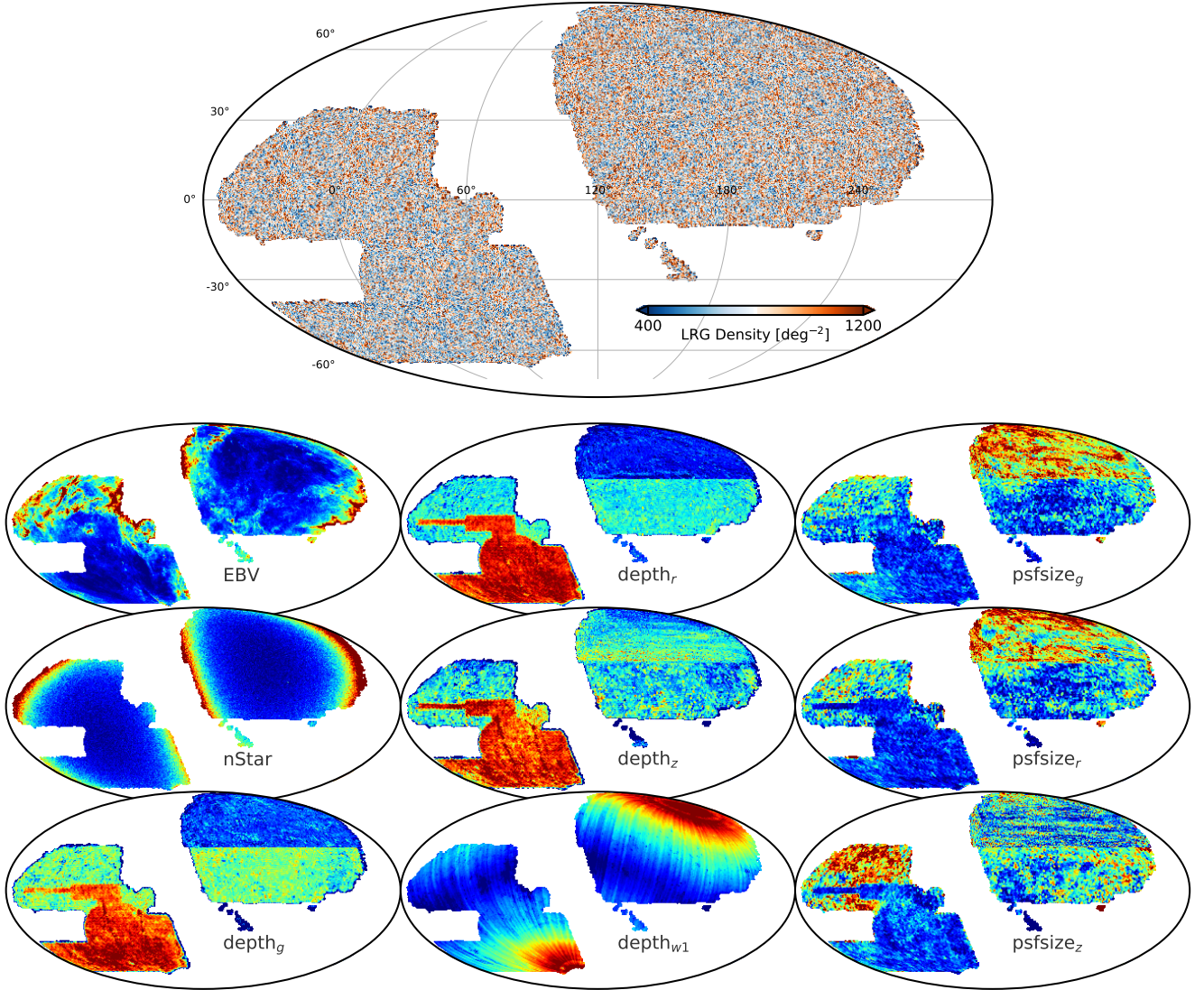
Figure 2. Top: The DESI LRG target density map before imaging systematics correction in Mollweide projection. Spurious disconnected islands from the North footprint and declination below -30 from the South footprint are removed for the analysis due to potential calibration issues (see text). Bottom: Mollweide projections of the DESI DR9 imaging properties (survey depth and astronomical seeing/psfsize) and MW foregrounds (extinction and local stellar density) in celestial coordinates. Two external imaging maps are neural hydrogen column density and photometric calibration maps (not shown), which only used for robustness tests.
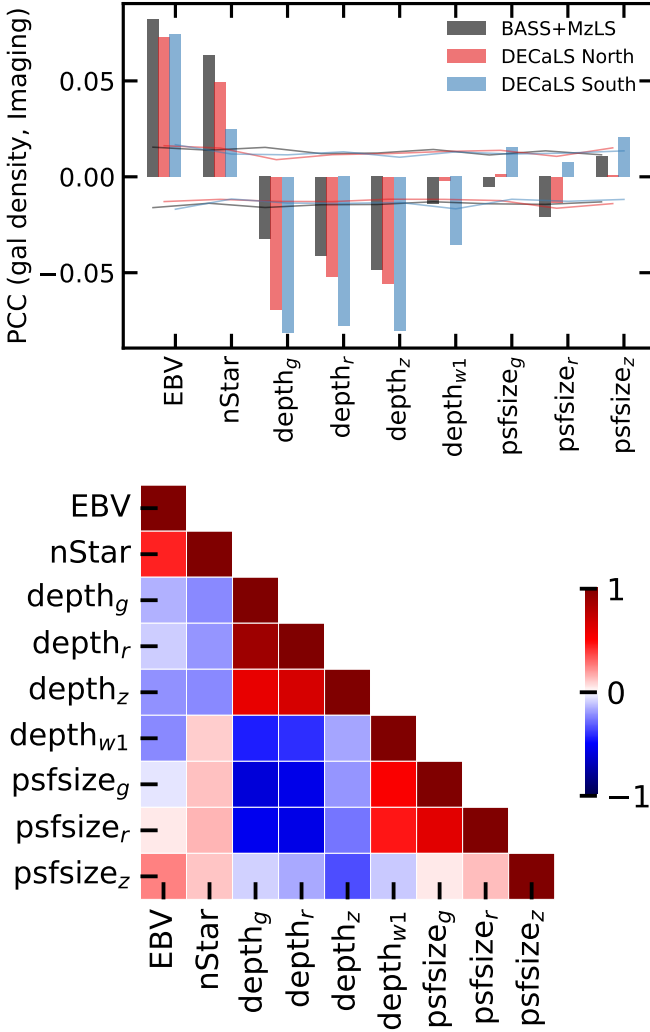
Figure 3. Top: The Pearson correlation coefficient between the DESI LRG target density and imaging properties in BASS+MzLS, DECaLS North, and DECaLS South. Solid horizontal curves represent the 95% confidence regions estimated from 100 simulated lognormal density maps. Bottom: The Pearson correlation matrix of imaging properties for the DESI footprint.

ure 3 (bottom) shows the correlation matrix among the imaging systematics maps for the DESI footprint. There is a strong correlation between the LRG density and depth, extinction, and stellar density. There is a small correlation between the LRG density and the $W1$-band depth and psfsize. Significant inner correlations exist among the imaging maps themselves, especially between the local stellar density and Galactic extinction; also, the r-band and g-band survey properties are more correlated with each other than with the z-band. The Spearman-r correlation is applied for robustness test but no significant differences are found.

2.1.2 Imaging weights

The effects of observational systematics in the DESI targets have been studied in great detail (see, e.g., Kitanidis et al. 2020; Zhou et al. 2021; Chaussidon et al. 2022). There are several approaches for handling imaging systematic errors, broadly classified into data-driven backward modeling and simulation-based forward modeling.

The general idea behind these approaches is to use the available data or simulations to learn or forward model the relationship between the observed target density and the imaging maps, and to use this relationship, which is often described by a set of *imaging weights*, to correct for systematic errors. Another techniques for mitigating the effect of imaging systematics relies on cross-correlating different tracers of dark matter to reduce excess clustering signals, as each tracer might respond differently to a source of systematic error (see, e.g., Giannantonio et al. 2014). These methods have their limitations and strengths (see, e.g., Weaverdyck & Huterer 2021, for a review).

In this paper, a backward modeling approach is applied to the data. Both linear multivariate and nonlinear neural network regressions are used to correct for imaging systematics. One potential problem that can arise in the mitigation process is *over-correction*, which occurs when the corrections applied to the data are too strong, leading to the removal of the clustering signal and the introduction of additional biases in the inferred parameter. These effects are demonstrated to be negligible for observables like BAO and RSD (Merz et al. 2021); however, they can significantly impact the clustering power on large scale, and thus induce biases in f_{NL} constraints (Mueller et al. 2022; Rezaie et al. 2021). The other problem is that the imaging maps are strongly correlated (Figure 3). It is crucial to develop techniques to control over-correction and to reduce dimensionality of the problem, in the hope of ensuring that the constraints are as accurate and reliable as possible, in particular for problems such as local primordial non-Gaussianity and other features in the primordial power spectrum (Beutler et al. 2019). Using more maps increases the input noise and the likelihood of overcorrection. With the objective of getting the smallest cross-correlations between the target density (after correction) and the imaging maps, our plan is to obtain various levels of corrections ranging from conservative to extreme by using the following combinations:

- **Conservative I**: Extinction and z-band depth.
- **Conservative II**: Extinction, z-band depth, and r-band psfsize.
- **All Maps**: Extinction, depth in $grzW1$, and psfsize in grz.

This will enable our analysis to indentify and calibrate mitigation biases using simulations (Section 3).

Both linear and neural network models are trained on each imaging region separately since the Pearson correlation coefficient analysis indicated that each imaging region responds differently to each imaging map (Figure 3). The optimal parameters associated with the linear and neural networks models are found by optimizing the negative Poisson log-likelihood, $\lambda - \rho \log(\lambda)$, measured between the observed target density ρ and the model output for the predicted density λ given only imaging properties **x** as input. No spatial coordinates are included in input **x** to avoid over-correction. The expected mean number of galaxies is defined as $\lambda(\mathbf{x}) \equiv \log(1 + e^{f(\mathbf{x})})$, where either linear multivariate regression or nonlinear artificial neural networks is used to estimate f.

The linear multivariate model only uses the imaging maps up to the linear power. The parameters are optimized with a random search method called Monte Carlo Markov Chain (MCMC) using the implementation of EMCEE (Foreman-Mackey et al. 2013). As our data size dominates the number of parameters for the linear model, no training-validation split is performed on the linear model. Our neural network-based mitigation approach uses the implementation from Rezaie et al. (2021); which trains an ensemble of 20 neural network models on the data. Each neural network is a fully-connected feed forward architecture that starts with the imaging values as input in the first layer, three hidden layers with 20 rectifier activation function units, and a single identity function unit

in the output layer. The rectifier activation function is defined as $\max(0, x)$. This simple form of nonlinearity is very effective in enabling deep neural networks to learn complex functions (Nair & Hinton 2010). Unlike linear regression, neural networks are prone to fitting noise, i.e., excellent performance on training data and poor performance on unseen data. Therefore, our analysis uses a training-validation-testing split to ensure that the network is well-optimized and generalizes well to unseen data. Specifically, 60% of the LRG data is used for training, 20% is used for validation, and 20% is used for testing. The neural network models are tested on the entirety of the LRG sample with the technique of permuting the choice of the training, validation, or testing sets (Arlot & Celisse 2010). The neural networks are trained for up to 70 training epochs with the gradient descent ADAM optimizer (Loshchilov & Hutter 2017). With this approach, the neural network parameters are updated iteratively following the gradient of the negative Poisson log-likelihood. The step size of the parameter updates is controlled via the learning rate hyper-parameter, which is initialized with a grid search and is set to dynamically vary between two boundary values of 0.001 and 0.1 to avoid local minima (Loshchilov & Hutter 2016). At each training epoch, the neural network models are applied to the validation set, and ultimately the models with the best performance are saved as the best models and applied to the test set. With cross-validation technique, the model predictions from the different test sets are aggregated together to form the predicted target density map into HEALPix of NSIDE = 256. The imaging weights are then defined as the inverse of the normalized predicted target density, and applied to the target density to correct for imaging systematics.

A visual inspection of the predicted target density maps reveals that most of the large-scale spurious fluctuations are explained by just extinction and depth in the z-band. Adding psfsize in the r-band results in more small-scale spurious fluctuations in the predicted target density maps. Our inspection shows that using all of the imaging maps as input for regression does not provide more information, which is expected from highly correlated predictors (Figure 3). Given the same set of inputs, the neural network-based weights show more small-scale spurious fluctuations which can be associated with the higher flexibility of the underlying model. Both approaches yield large-scale spurious fluctuations consistent with the LRG target density; for instance, both models predict a higher density of LRGs near the boundaries where the DESI imaging surveys observed regions of high extinction or high stellar density near the galactic plane. These over-dense regions are likely contaminated artifacts entering the LRG selection, e.g., stellar contaminants or other artifacts because of obscured photometry from extinction. This is interesting since the target selection cuts use extinction corrected magnitudes. Further robustness tests on these weights are discussed in Section 3, and their impacts on f_{NL} are presented in Section 4.

2.2 Synthetic lognormal density fields

Density fluctuations of galaxies on large scales can be approximated with lognormal distributions (Coles & Jones 1991). Unlike N-body simulations, simulating lognormal density fields is not computationally intensive, and allows quick and robust validation of data analysis pipelines. These mocks are therefore considered useful for our study since the signature of local PNG appears on large-scales and small scale clustering is not used. The package FLASK (Full-sky Lognormal Astro-fields Simulation Kit; Xavier et al. 2016) is employed to generate ensembles of synthetic lognormal density maps that mimic the redshift and angular distributions of the DR9 LRG targets. Two universes with $f_{NL} = 0$ and 76.92 are considered.

For each f_{NL}, 1000 realizations are generated assuming a redshift dependent bias $b(z) = 1.43/D(z)$ (see, e.g., Zhou et al. 2021). The analysis adapts a fiducial cosmology from a flat ΛCDM universe, including one massive neutrino with $m_\nu = 0.06$ eV, while the rest of cosmological parameters is deducted from Planck 2018 (Aghanim et al. 2020),

$$h = 0.67, \Omega_M = 0.31, \sigma_8 = 0.8, \text{ and } n_s = 0.97.$$

The same fiducial cosmology is used throughout this paper unless specified otherwise. The robustness of our constraints against the fiducial cosmology is explored in Section 4.

2.2.1 Contaminated mocks

The linear multivariate model with the *conservative II* set, which containts extinction, depth$_z$, psfsize$_r$, is applied to introduce synthetic spurious fluctuations in the lognormal density fields. The motivation for choosing a linear model for the contamination is to assess how much of the clustering signal can be removed by applying more flexible and rigorous models (based on neural networks) for correcting the effect of imaging systematics. The parameters of the linear model are fit on the DESI LRG target sample, separately on each imaging survey. The contamination model for each simulation is uniquely and randomly drawn from the parameter space constraint by MCMC. The same contamination model is used for both the $f_{NL} = 0$ and 76.92 simulations.

Similar to the analysis for the DESI LRG targets, the linear and neural network methods are applied to each simulation, with and without PNG or with and without systematics, to derive the imaging weights. No prior knowledge is used regarding the underlying model or the input imaging maps for contamination, to simulate a realistic scenario. Section 3 presents how the simulation results are incorporated to calibrate f_{NL} biases due to over-correction.

3 ANALYSIS TECHNIQUES

3.1 Power spectrum estimator

The pseudo angular power spectrum (Hivon et al. 2002) is utilized to extract information from the galaxy density contrast field, δ_g,

$$\delta_g = \frac{\rho - \overline{\rho}}{\overline{\rho}}, \qquad (2)$$

by decomposing it into spherical harmonics, $Y_{\ell m}$,

$$a_{\ell m} = \int d\Omega \, \delta_g W Y_{\ell m}^*, \qquad (3)$$

where $\overline{\rho}$ is the mean galaxy density (estimated from the data) and W is the survey window (estimated from the randoms). The angular power spectrum is estimated by ,

$$\tilde{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2. \qquad (4)$$

We use the implementation of anafast from the HEALPix package (Gorski et al. 2005) to do fast harmonic transforms and estimate the pseudo angular power spectrum and cross power spectrum. This power spectrum estimator is biased for a partial sky coverage. Because of survey mask, different harmonic modes are correlated and the measured clustering power on scales near the survey size is pushed to zero. As these scales are sensitive to PNG, correcting for these systematic errors is important to obtain unbiased constraints.

3.2 Modeling

3.2.1 Angular power spectrum

The projected angular power spectrum of galaxies is related to the 3D power spectrum $P(k)$ (see, e.g., Padmanabhan et al. 2007) and shotnoise N_{shot} by,

$$C_\ell = \frac{2}{\pi} \int_0^\infty \frac{dk}{k} k^3 P(k) |\Delta_\ell(k)|^2 + N_{\text{shot}}, \qquad (5)$$

where N_{shot} is assumed to scale-independent, and $\Delta_\ell(k) = \Delta_\ell^{\text{g}}(k) + \Delta_\ell^{\text{RSD}}(k)$ is the projection kernel which includes redshift space distortions and determines how much each wavenumber k contributes to the ℓ-mode power. $\Delta_\ell(k)$ is computed by integrating over the lth order spherical Bessel functions, $j_\ell(kr)$, and its second derivatives,

$$\Delta_\ell^{\text{g}}(k) = \int \frac{dr}{r} r b(k,z) D(r) \frac{dN}{dr} j_\ell(kr), \qquad (6)$$

$$\Delta_\ell^{\text{RSD}}(k) = -\int \frac{dr}{r} r f(r) D(r) \frac{dN}{dr} j_\ell''(kr), \qquad (7)$$

where $D(z)$ represents the linear growth factor scaled as $D(z = 0) = 1$, $f(r)$ is the growth rate, and dN/dr is the normalized redshift distribution of galaxies[2] (Figure 1). The galaxy bias, $b(k,z)$, is given by the redshift dependent linear bias (Figure 1) and the PNG-induced shift (Slosar et al. 2008),

$$b(k,z) = b(z) + 3(b(z) - p) f_{\text{NL}} \frac{\delta_c \Omega_m H_0^2}{k^2 T(k) D(z) c^2} \frac{g(\infty)}{g(0)}, \qquad (8)$$

where Ω_m is the matter density, H_0 is the Hubble constant[3], $T(k)$ is the transfer function, $\delta_c = 1.686$ represents the critical density for spherical collapse (Fillmore & Goldreich 1984), and $g(\infty)/g(0) \sim 1.3$ with $g(z) \equiv (1+z) D(z)$ (see, e.g., Mueller et al 2018). The parameter p is a correction due to galaxy selection beyond a Poisson sampling of the haloes of a given mass; if only mass determines how galaxies populate a halo, $p = 1$, which is often referred to as the universality of the halo occupation distribution (see, e.g., Slosar et al. 2008). However, numerical simulations indicate that the halo occupation distribution for other tracers, e.g., quasars, which are from recent mergers, could depend on other properties besides mass, and thus p can take different values. The theoretical uncertainty on p is not very well constrained, and Barreira (2022) showed that marginalizing over this parameter even with wide priors leads to biased constraints because of parameter space projection effects. Lazeyras et al. (2023) used N-body simulations to investigate secondary halo properties such as concentration, spin and sphericity of haloes, and found that halo spin and sphericity preserve the universality of the halo occupation function while halo concentration significantly alters the halo function. Without better-informed priors on p, it is argued that the scale-dependent bias effect can only be used to constrain the $p f_{\text{NL}}$ term (see, e.g., Barreira et al. 2020; Barreira 2020). Nevertheless, the significance of detection of nonzero PNG is not affected by different assumptions on p, i.e., a nonzero detection of $p f_{\text{NL}}$ will still imply a nonzero detection of f_{NL}. This paper is focused on how a careful treatment of imaging systematic effects, or lack thereof, can bias our PNG constraints. Therefore, we choose $p = 1$ for our sample of DESI-like LRGs. We use the FFTLog[4] algorithm and its extension as implemented in Fang et al. (2020) to handle the Bessel function integrals. Next, we explain the corrections for the effects of survey geometry.

[2] $dN/dr = (dN/dz) * (dz/dr) \propto (dN/dz) * H(z)$
[3] Because k is in unit of h/Mpc, $H_0 = 100 \,(\text{km s}^{-1})/(h^{-1}\text{Mpc})$
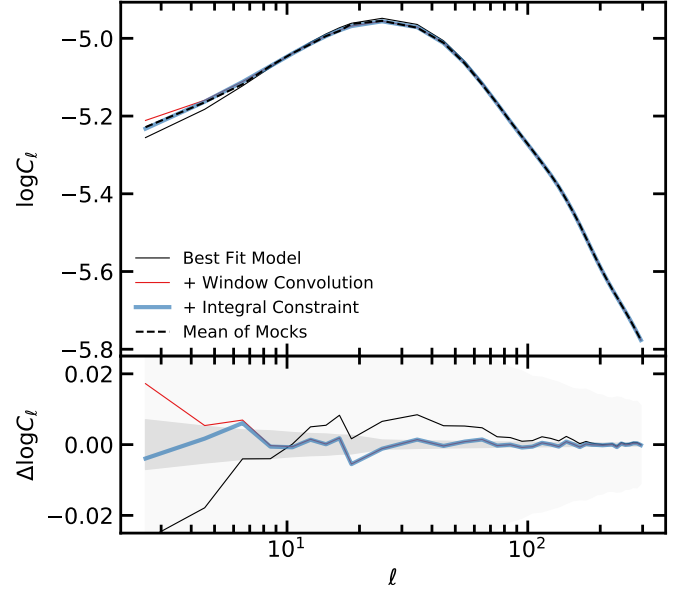[4] github.com/xfangcosmo/FFTLog-and-beyond

Figure 4. The mean power spectrum from the $f_{\text{NL}} = 0$ mocks and best fit theoretical prediction after accounting for the survey geometry and integral constraint effects. The dark and light shades represent 1σ error on the mean and one realization, respectively. Bottom panel shows the residual power spectrum relative to the mean power spectrum. No imaging systematic effects are added to these mocks.

3.2.2 Survey geometry and integral constraint

The ensemble average for the partial sky power spectrum is related to that of the full sky power spectrum via a mode-mode coupling matrix, $M_{\ell\ell'}$,

$$< \tilde{C}_\ell > = \sum_{\ell'} M_{\ell\ell'} < C_{\ell'} > . \qquad (9)$$

In general, $M_{\ell\ell'}$ is singular for a large footprint, but a common technique to remedy this issue is to use a discrete set of ℓ bins and assume that the angular power spectrum is constant in each bin. We following a similar approach to Chon et al. (2004) to do the convolution in Equation 9; The theoretical galaxy power spectrum C_ℓ is convolved with the survey mask power spectrum to obtain the expected pseudo-power spectrum. The convolution in the spherical harmonics space becomes a multiplication in the correlation function space. First, we obtain the survey mask paircount, and multiply it by the model correlation function. Then, the model pseudo-power spectrum is given by,

$$\tilde{C}_\ell^{\text{model}} = 2\pi \int \hat{\omega}^{\text{model}} \hat{\omega}^{\text{mask}} P_\ell(\cos\theta) d\theta. \qquad (10)$$

The integral constraint is another systematic effect which is induced because the mean galaxy density is estimated from the observed galaxy density. The estimate of the mean density is biased by the limited sky coverage. This issue was first raised in Peacock & Nicholson (1991). To account for the integral constraint, the survey mask power spectrum is used to introduce a scale-dependent correction factor that needs to be subtracted from the power spectrum. Finally, the pseudo power spectrum with integral constraint is given by,

$$\tilde{C}_\ell^{\text{model,IC}} = \tilde{C}_\ell^{\text{model}} - \tilde{C}_{\ell=0}^{\text{model}} \left(\frac{\tilde{C}_\ell^{\text{window}}}{\tilde{C}_{\ell=0}^{\text{window}}} \right). \qquad (11)$$
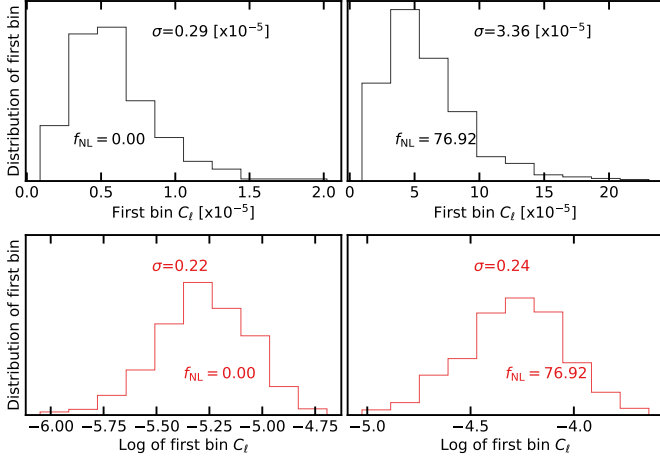
Figure 5. The distribution of the first bin power spectra and its log transformation from the simulations with $f_{NL} = 0$ (left) and 76.92 (right). The log transformation alleviates the asymmetry in the distributions.

The lognormal simulations are used to validated our modeling pipeline. Figure 4 shows the mean power spectrum of 1000 simulations (dashed) and the best fit theory prediction before and after accounting for the survey mask and integral constraint. The light and dark shades represent the 68% error on the mean and one single realization, respectively. The DESI mask is applied to the simulations, and even though the mask covers around 40% of the sky, but the survey window effect is affecting the clustering power on $\ell < 200$. On the other hand, the integral constraint only modulates the clustering power on $\ell < 6$.

3.3 Parameter estimation

Our parameter inference uses standard Monte-Carlo Markov Chain (MCMC) sampling, as implemented in EMCEE (Foreman-Mackey et al. 2013). Our MCMC allows f_{NL}, N_{shot}, and b to vary, but all other cosmological parameters are fixed at the fiducial values (see §2.2). For the linear bias, a constant clustering amplitude is assumed for our sample, $b(z) = b/D(z)$, and fit for b. Throughout this manuscript, we use a discrete set of bandpower bins with $\Delta\ell = 2$ between $\ell = 2$ and 20 and $\Delta\ell = 10$ from $\ell = 20$ to 300. Each mode is weighted by $2\ell + 1$.

We find that the distribution of power spectra at the lowest bin, $2 \leq \ell < 4$, is asymmetric and its standard deviation varies significantly from the simulations with $f_{NL} = 0$ to those with 76.9 (Figure 5). Therefore, we attempt to fit $\log C_\ell$ to make our f_{NL} constraints less sensitive to the choice of covariance matrix. The parameter f_{NL} is constrained by minimizing a negative likelihood defined as,

$$-2\ln\mathcal{L} = (\log\tilde{C}(\Theta) - \log\tilde{C})^\dagger \mathbb{C}^{-1}(\log\tilde{C}(\Theta) - \log\tilde{C}), \quad (12)$$

where Θ represents a container for the parameters f_{NL}, b, and N_{shot}; $\tilde{C}(\Theta)$ is the (binned) expected pseudo-power spectrum; $\tilde{C}$ is the (binned) measured pseudo-power spectrum; and $\mathbb{C}$ is the covariance matrix constructed from the lognormal simulations. Flat priors are implemented for all parameters.

3.4 Remaining systematic errors

In the absence of systematic effects, a) the mean galaxy density should be uniform across the footprint within the statistical fluctuations regardless of imaging conditions and b) the cross power spectrum between the galaxy density and imaging properties should be consistent with zero within the statistical fluctuations. Two statistical tests are implemented and applied to quantify remaining systematic effects in our sample.

3.4.1 Cross power spectrum

We calculate the cross power spectra between the galaxy density contrast and imaging maps,

$$\tilde{C}_{X,\ell} = [\tilde{C}_{x_1,\ell}, \tilde{C}_{x_2,\ell}, \tilde{C}_{x_3,\ell}, ..., \tilde{C}_{x_9,\ell}], \quad (13)$$

where $\tilde{C}_{x_i,\ell}$ represents the cross power spectrum between the galaxy density contrast and i^{th} imaging map, normalized by the auto power spectrum of x_i:

$$\tilde{C}_{x_i,\ell} = \frac{(\tilde{C}_{gx_i,\ell})^2}{\tilde{C}_{x_i x_i,\ell}}. \quad (14)$$

Then, the χ^2 value for the cross power spectra is calculated via,

$$\chi^2 = \tilde{C}_{X,\ell}^T \mathbb{C}_X^{-1} \tilde{C}_{X,\ell}, \quad (15)$$

where the covariance matrix $\mathbb{C}_X =< \tilde{C}_{X,\ell}\tilde{C}_{X,\ell'} >$ is constructed from the lognormal mocks. This χ^2 error is measured for every mock realization with the *leave-one-out* technique[5], and then is compared to the χ^2 value observed from the DR9 LRGs. We only include the bandpower bins from $\ell = 2$ to 20 with $\Delta\ell = 2$, although later higher ℓ modes are considered as well for a robustness test.

Figure 6 shows $\tilde{C}_X$ from the DR9 LRG sample before and after applying various corrections for imaging systematics. The dark and light shades show the 97.5th percentile from the $f_{NL} = 0$ and 76.9 mocks, respectively. The LRG sample has the highest cross correlation against extinction, stellar density, and the z-band depth (solid black curve). As the most conservative cleaning approach, a linear model is trained with the extinction and z-band depth maps (*linear conservative I*). With the correction applied, we find that the cross power spectrum against the r-band psfsize increases (dot-dashed blue curve), which indicates that only extinction and depth in the z-band are not sufficient to regress out all of the residual cross correlations between the LRG density and imaging. With the linear model re-trained using the three maps (*linear conservative II*), we are able to reduce the cross power spectra (dotted orange curve). For comparison, we also show the cross spectra for a case in which the DR9 is cleaned using the linear model trained with all imaging maps (dashed red curve). There are minor differences between the *linear conservative II* and *linear all maps* cross spectra, further supporting the idea that only three imaging maps are sufficient to regress out spurious correlations.

Figure 8 (top) shows the histogram of the cross spectrum χ^2 from 1000 mocks with and without f_{NL}. The χ^2 values observed in the DR9 LRG sample are quoted for comparison. Before cleaning, our LRG sample has a cross power spectrum χ^2 error of 20014.8. After correction with the linear conservative I approach, the cross spectrum χ^2 is reduced to 375.1 with p-value = 0.002. Adding the r-band psfsize, the linear model reduces the χ^2 down to 195.9 with

---

[5] 999 realizations are used to estimate the covariance matrix and applied to measure the χ^2 for the left-out realization.
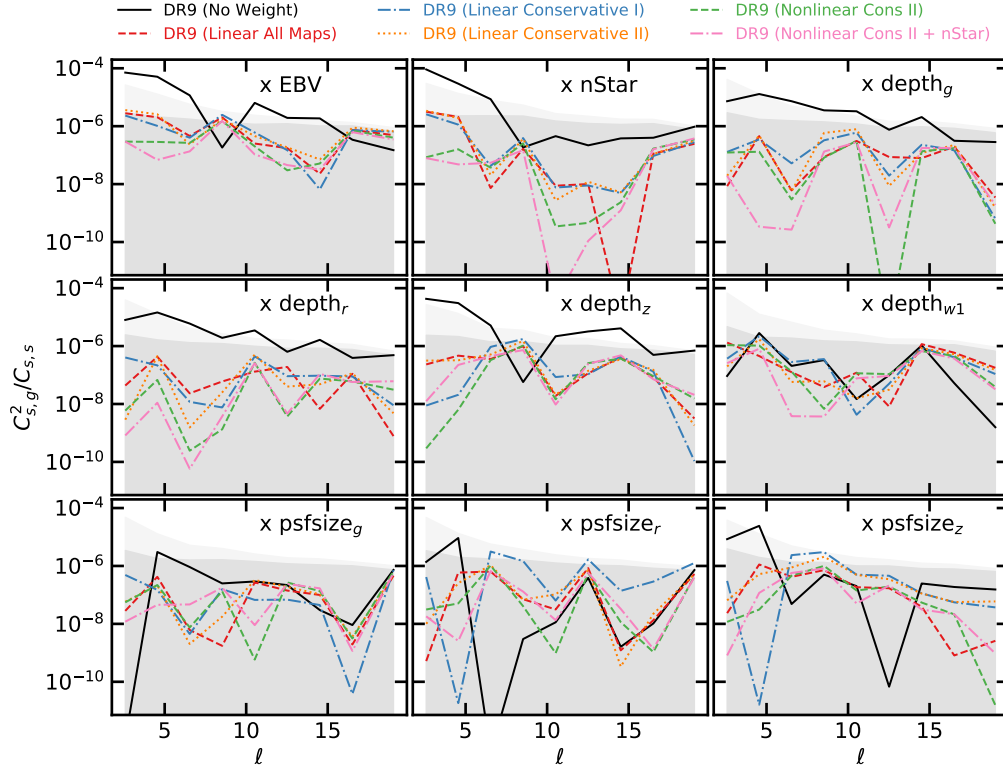
Figure 6. The cross power spectra between the DR9 LRG sample and imaging maps. The dark and light shades represent the 97.5 percentile of 1000 lognormal mocks without and with PNG, respectively.
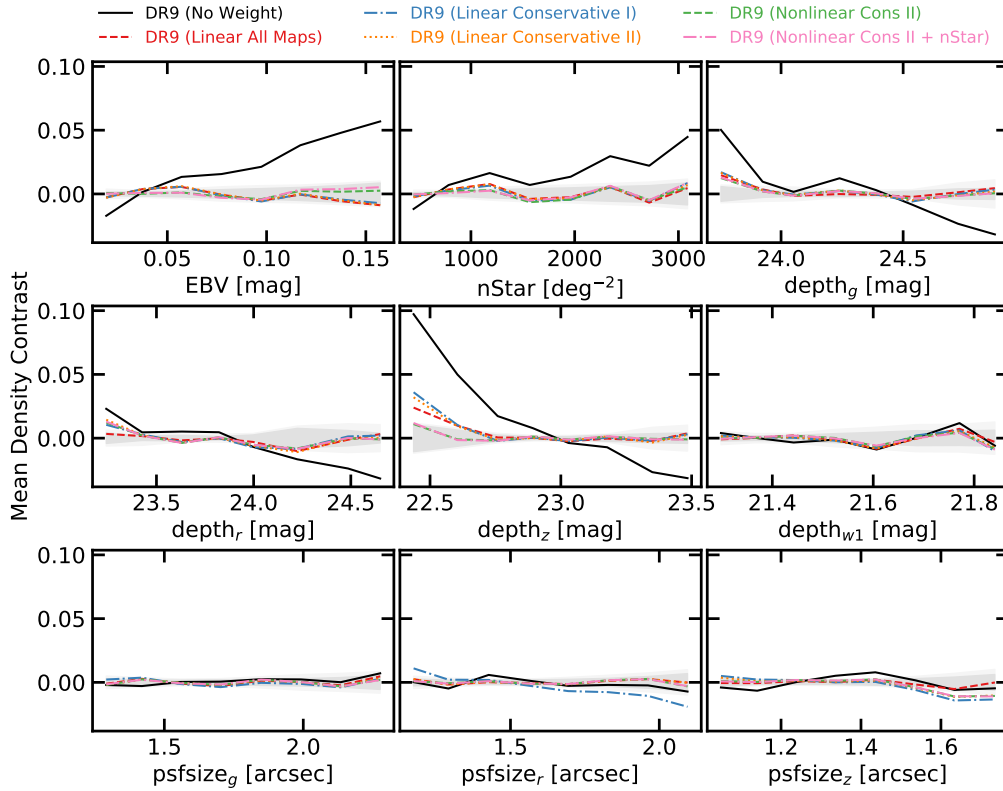

Figure 7. The mean galaxy density contrast of the DR9 LRG sample as a function of imaging maps. The dark and light shades represent the 1σ dispersion of 1000 lognormal mocks without and with PNG, respectively.
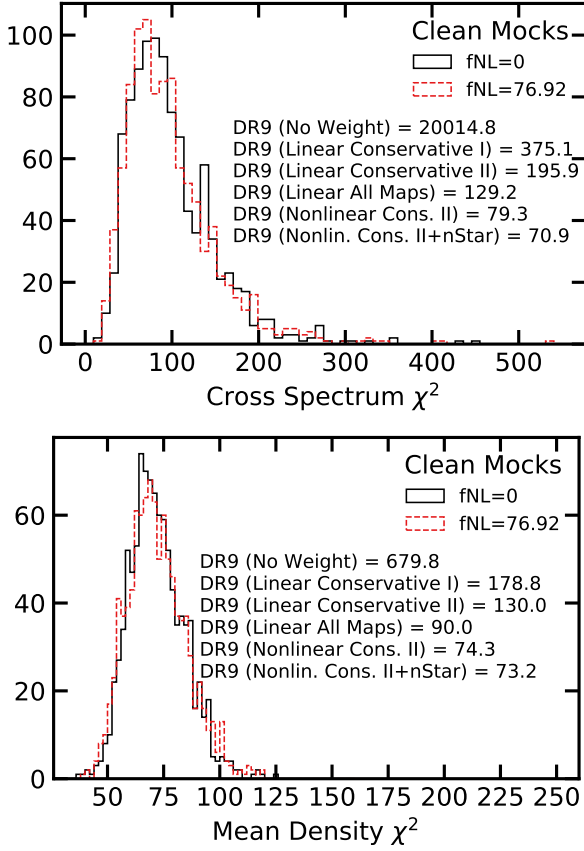
Figure 8. The remaining systematic error χ^2 from the galaxy-imaging cross power spectrum (top) and the mean galaxy density contrast (bottom). The values observed in the DR9 LRG sample before and after linear and nonlinear treatments are quoted, and the histograms are constructed from 1000 realizations of clean lognormal mocks with $f_{NL} = 0$ and 76.92.

p-value = 0.044; we can reject the null hypothesis that the DR9 sample with the linear conservative II is properly cleaned at 95% confidence. Although training the linear model with all imaging maps as input gives the lowest cross spectrum χ^2 of 129.2 (and p-value = 0.239), but it potentially makes the analysis more prone to over-fitting and regressing out the true clustering signal, given the inner correlations among the imaging properties (Figure 3). As an alternative, we train the nonlinear method with the extinction, z-band depth, and r-band psfsize maps (*nonlinear conservative II*) and clean our sample. The cross power spectrum χ^2 is reduced to 79.3 with p-value = 0.594. Adding the stellar density map reduces the cross power spectrum χ^2 error to 70.9 (p-value = 0.687). This cross power spectrum diagnostic supports the idea that a nonlinear cleaning approach is needed to properly regress out the remaining spurious fluctuations. We investigate the test with the cross power spectrum up to higher multipoles but find no evidence of remaining systematic errors (see Appendix A).

3.4.2   *Mean density contrast*

We calculate the histogram of the mean density contrast relative to the jth imaging property:

$$\delta_{x_j} = (\hat{\bar{\rho}})^{-1} \frac{\sum_i \rho_i f_{\text{pix},i}}{\sum_i f_{\text{pix},i}}, \tag{16}$$

where the summations are over HEALPIX pixels in the bin with similar imaging values. For instance, In the absence of systematic error, the mean density contrast from one part of the sky with high extinction should be consistent with that from another patch with low extinction. We compute the histograms against all other imaging properties (see Figure 2), and construct the total mean density contract as,

$$\delta_X = [\delta_{x_1}, \delta_{x_2}, \delta_{x_3}, ..., \delta_{x_9}], \tag{17}$$

and the total residual error as,

$$\chi^2 = \delta_X^T \mathbb{C}_\delta^{-1} \delta_X, \tag{18}$$

where the covariance matrix $\mathbb{C}_\delta = <\delta_X \delta_X>$ is constructed from the lognormal mocks. Figure 7 shows the mean density contrast against the imaging properties for the DR9 LRG sample. The dark and light shades represent the 1σ level fluctuations observed in 1000 lognormal density fields respectively with $f_{NL} = 0$ and 76.92. The DR9 LRG sample before treatment (solid curve) exhibits a strong trend around 10% against the z-band depth which is consistent with the cross power spectrum. Additionally, there are significant spurious trends against extinction and stellar density at about 5 − 6%. The linear approach is able to mitigate most of the systematic fluctuations with only extinction and depth in the z-band as input; however, a new trend appears against the r-band psfsize map with the *linear conservative I* approach (dot-dashed blue curve), which is indicative of the psfsize-related systematics in our sample. This finding is in agreement with the cross power spectrum. We re-train the linear model with three maps identified as *conservative II*, but we still observe around 2% residual spurious fluctuations in the low end of the z-band depth, which implies nonlinear systematic effects exist. We find that the nonlinear model trained with the three identified maps (or four maps including the stellar density) is capable of reducing the fluctuations below 2%. We experiment with different binning schemes but find consistent results.

Figure 8 (bottom) shows the mean density χ^2 observed in the mocks with or without f_{NL}. We find consistent results regardless of the underlying f_{NL}, which supports that our diagnostic is not sensitive to the fiducial cosmology. The values measured in the DR9 LRG sample before and after applying imaging weights are quoted for comparison. The *linear conservative I* weights reduce the χ^2 value from 679.8 (before correction) to 178.8. The p-value = 0 indicates severe remaining systematic effects. Adding the r-band psfsize does not reduce the p-value enough (e.g., greater than 0.05) even though the cleaning method yields a lower $\chi^2 = 130$. Training the linear model with all imaging maps returns a more reasonable $\chi^2 = 90$ and p-value of 0.084; however, regression with all imaging maps as input can lead to the removal of the true clustering signal. We try the *nonlinear conservative II* approach. We obtain a χ^2 value of 74.3 with p-value = 0.392. Re-training the nonlinear approach while adding the stellar density map yields minor improvement: $\chi^2 = 73.2$ and p-value = 0.422. This indicates that the stellar density trend in the mean LRG density can be explained via the extinction map.

3.5   Calibration of mitigation bias

The template-based mitigation of imaging systematics removes some of the true clustering signal, and the amount of the removed signal increases as more maps are fed to the regression. Below we describe an approach to calibrate our cleaning methods and de-bias our f_{NL} constraints.
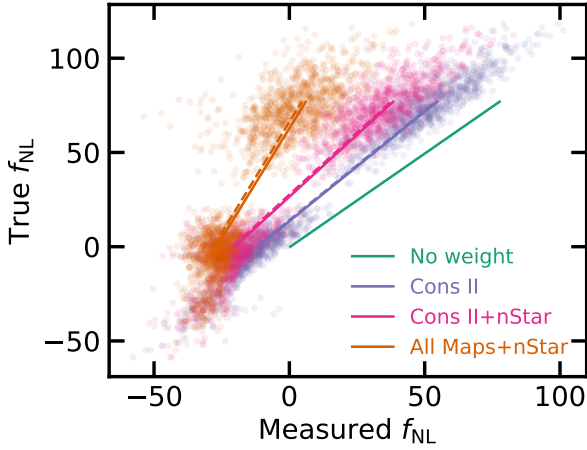
Figure 9. The true vs measured f_{NL} values from the $f_{NL} = 0$ and 76.9 mocks. The solid (dashed) lines represent the best fit estimates from fitting the mean power spectrum of the clean (contaminated) mocks. The scatter points show the best fit estimates from fitting the individual spectra for the clean mocks.

Table 2. Linear parameters employed to de-bias the f_{NL} constraints to account for the over-correction issue.

Cleaning Method	m_1	m_2
Nonlinear Conservative II	1.17	13.95
Nonlinear Conservative II+nStar	1.32	26.97
Nonlinear All Maps+nStar	2.35	63.5

To calibrate for over-correction, we utilize our series of lognormal density fields with and without PNG, with and without systematic effects. The contamination model is based on the linear multivariate approach with the extinction, z-band depth, and r-band psfsize maps as input and parameters drawn from the likelihood constrained by the DR9 LRG sample. The idea is to simulate systematic effects that reflect spurious fluctuations as realistic as the DR9 LRG sample. For correction, the neural network model is trained and applied to the simulations with various sets of imaging maps as input. Particularly, we consider *nonlinear conservative II*, *nonlinear conservative II + nStar*, and *nonlinear All Maps + nStar*. We fit both the mean power spectrum and each individual power spectrum of 1000 realizations. The best fit estimates from the mocks without systematics (and no mitigation applied) are considered as the true f_{NL} values and the estimates from the mocks (with the mitigation procedure applied) are considered as the measured f_{NL} values. Figure 9 shows the true f_{NL} values and the measured f_{NL} values from fitting the mean power spectrum (solid lines) and individual spectra (points). The dashed lines show the best fit estimates from the contaminated realizations. The best fit estimates for the contaminated mocks before cleaning (*no weight*) or per realizations are not shown for visual clarity.

Then, a pair of linear parameter are found to map the measured values to the true values, $f_{NL,true} = m_1 f_{NL,measured} + m_2$. These m_1 and m_2 coefficients for the mean power spectrum are summarized in Table 2. The coefficients for the *All Maps+nStar* approach are the highest, as more mitigation bias is expected if more imaging maps are used. We find that $m_1 - 1$ determines the added uncertainty in the f_{NL} constraints, once the correction coefficients are applied. We expect this effect to be maximum for *All Maps+nStar* and minimum for *Conservative II*.
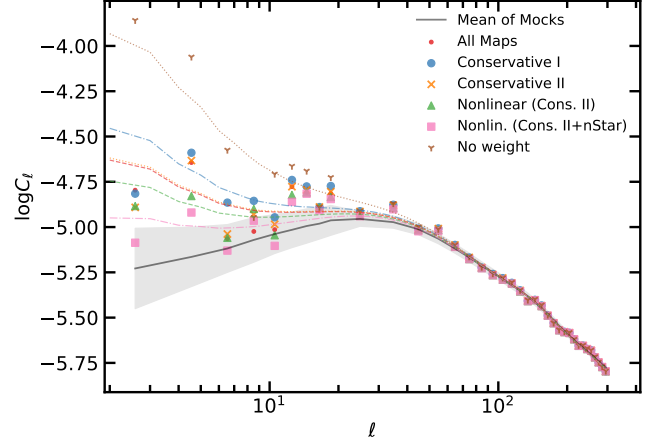
Figure 10. The angular power spectrum of the DR9 LRG sample before (*No weight*) and after correcting for imaging systematics using various methods with their corresponding best fit theory curves. The shade represents 1σ error constructed from the $f_{NL} = 0$ mocks.

4 RESULTS

4.1 DESI imaging LRG sample

Fig. 10 shows the measured power spectrum of the DESI imaging DR9 LRG sample before and after applying imaging weights, the best fit theory curves, and the mean power spectrum and 1σ error estimated from the $f_{NL} = 0$ lognormal simulations. The power spectra are similar on small scales ($\ell > 20$), but the differences between various cleaning methods are significant on large scales. By comparing *linear conservative I* to *linear conservative II*, we find that the measure power spectra on modes with $6 \leq \ell < 10$ are noticeably different between the two methods. We associate the differences to the r-band psfsize template in *linear conservative II*. On other scales, the differences between the spectra after the linear-based cleaning are negligible, supporting the idea that our feature selection procedure has worked to identify the primary maps causing excess clustering signal. Comparing *nonlinear conservative II* to *linear conservative II*, we find that the measured spectra on $4 \leq \ell < 6$ are very different, probably pointing at nonlinear spurious fluctuations with large-scale characteristics due to the extinction. Adding stellar density to the nonlinear approach (*nonlinear conservative II + nStar*) results in less excess power relative to the mock power spectrym, with the modes on $2 \leq \ell < 4$ reflecting the biggest change. The flexibility of the nonlinear approach to correct for these effects ameliorates the clustering measurements on these scales.

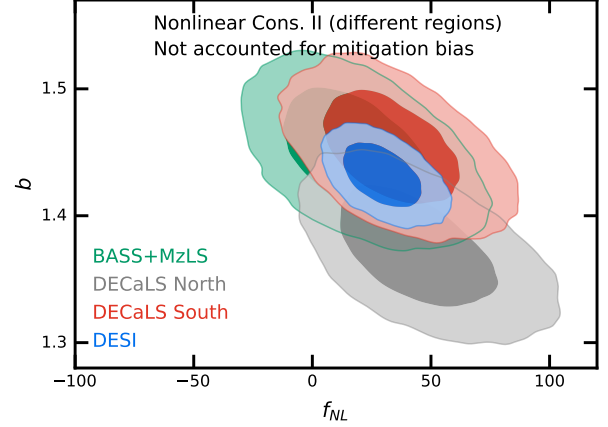#### 4.1.1 Calibrated constraints

All f_{NL} constraints presented here are calibrated for the effect of mitigation bias using the lognormal simulations. Tab. 3 describes the best fit and marginalized mean estimates of f_{NL} from fitting the power spectrum of the DR9 LRG sample before and after cleaning with the nonlinear approach given various combinations of imaging templates. Fig 11 shows the marginalized probability distribution for f_{NL} (top) and the 68% and 95% probability contours for the bias parameter and f_{NL} from the DR9 sample before (*no weight*) and after applying various imaging weights. With the imaging weights applied, we obtain $36.08(25.03) < f_{NL} < 61.44(75.64)$ with $\chi^2 = 34.6$ for *nonlinear conservative II*, $36.88(24.87) < f_{NL} <$

Table 3. Calibrated best fit and marginalized mean estimates for f_{NL} from fitting power spectrum of the DESI DR9 LRG sample before and after correcting for systematics. Degree of freedom is 34 (37 data points - 3 parameters).

| Footprint | Method | f_{NL} | | | | χ^2 |
		Best fit	Mean	68% CL	95% CL	
DESI	No Weight	113.18	115.49	$98.14 < f_{NL} < 132.89$	$83.51 < f_{NL} < 151.59$	44.4
DESI	Nonlinear (Cons. II)	47.38	48.81	$36.08 < f_{NL} < 61.44$	$25.03 < f_{NL} < 75.64$	34.6
DESI	Nonlin. (Cons. II+nStar)	48.92	50.10	$36.88 < f_{NL} < 63.31$	$24.87 < f_{NL} < 77.78$	35.2
DESI	Nonlin. (All Maps+nStar)	49.69	41.91	$13.10 < f_{NL} < 69.14$	$-15.96 < f_{NL} < 91.84$	39.5
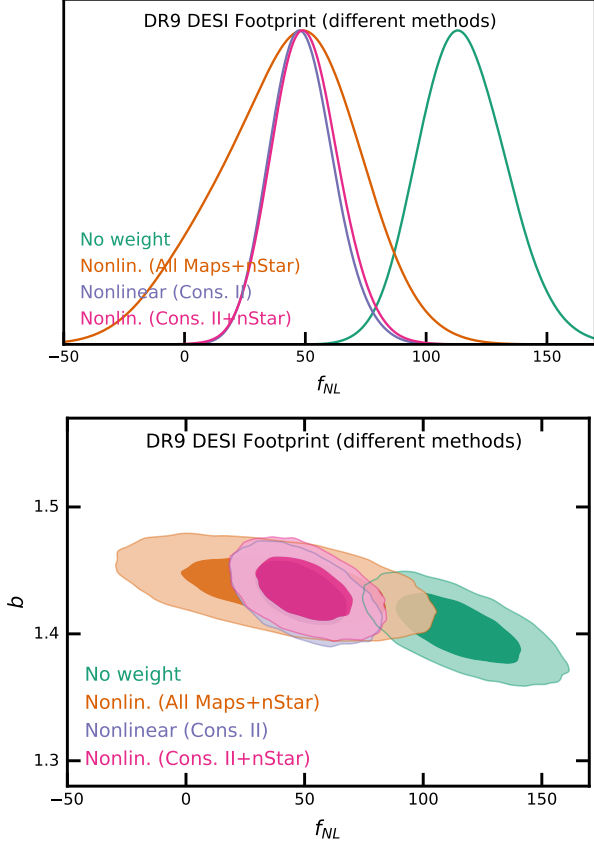
Figure 11. Calibrated constrains from the DR9 LRG sample. *Top*: probability distribution for f_{NL} marginalized over the shotnoise and bias. *Bottom*: 68% and 95% probability distribution contours for the bias and f_{NL} from the DR9 LRG sample before and after applying nonlinear cleaning methods. The lognormal mocks are used to correct these distributions for mitigation bias.

Figure 12. Uncalibrated 2D constraints from the DR9 LRG sample for each imaging survey compared with that for the whole DESI footprint. The dark and light shades represent the 68% and 95% confidence intervals, respectively.

4.1.2 Robustness tests

Now we proceed to perform some robustness tests and assess how sensitive the f_{NL} constraints are to the assumptions made in the analysis or the quality cuts applied to the data. For each case, we retrain the cleaning methods and derive new sets of imaging weights. Similarly, for the cases where a cut is applied on the data, we recalculate the covariance matrices using the new masked footprint to account for the changes in the window and integral constraint effect. calibrating the mitigation bias for all of these experiments is beyond the scope of this work. Therefore, the f_{NL} constraints presented here are subject to the mitigation bias effect. Table 4 describes the uncalibrated f_{NL} constraints from the DR9 LRG sample. Our tests are as follows:

- **Linear methods**: We find consistent f_{NL} constraints from *linear all maps* and *linear conservative II* which support the idea that all imaging maps are not needed to completely mitigate systematic effects. We find $\sigma(f_{NL}) \sim 25$ for the linear methods.
- **Imaging regions**: We compare how the f_{NL} constraints from fitting the power spectrum of the whole DESI footprint compares to that from the power spectrum for each sub-survey individually, namely BASS+MzLS, DECaLS North, and DECaLS South. Fig. 12 shows 68% and 95% probability contours on f_{NL} and b from each individual region, compared with that from DESI. The cleaning method here is *nonlinear conservative II*. Overall, we find that the constraints from all imaging surveys are consistent with each other within 68% confidence. Both BASS+MzLS and DECaLS South yield constraints consistent with $f_{NL} = 0$ within 95%, but DECaLS North deviates from zero PNG at more than 2σ.
- **Stellar density template (*nStar*)**: Adding the stellar density tem-

$63.31(77.78)$ with $\chi^2 = 35.2$ for *nonlinear conservative II + nStar*, and $13.10(-15.96) < f_{NL} < 69.14(91.84)$ with $\chi^2 = 39.5$ for *nonlinear all maps + nStar* at 68%(95%) confidence over 34 degrees of freedom. Overall, we find the maximum likelihood estimates to be consistent between among the various cleaning methods. For the conservative methods, the confidence intervals with or without *nStar* are consistent and more than 3σ off from zero. The method *nonlinear all maps+nStar* has asymmetric probability distribution, a larger uncertainty, and consistent with $f_{NL} = 0$ within the 95% confidence interval. For comparison, we obtain $98.14(83.51) < f_{NL} < 132.89(151.59)$ at 68%(95%) confidence with $\chi^2 = 44.4$ for the *no weight* approach. The uncalibrated probability contours are presented in Appendix B.

Table 4. Uncalibrated best fit and marginalized mean estimates for f_{NL} from fitting power spectrum of the DR9 LRG sample before and after correcting for systematics. Degree of freedom is 34 (37 data points - 3 parameters).

Footprint	Method	f_{NL}				χ^2
		Best fit	Mean	68% CL	95% CL	
DESI	No Weight	113.18	115.49	$98.14 < f_{NL} < 132.89$	$83.51 < f_{NL} < 151.59$	44.4
DESI	Linear (All Maps)	36.05	37.72	$26.13 < f_{NL} < 49.21$	$16.31 < f_{NL} < 62.31$	41.1
DESI	Linear (Conservative I)	49.58	51.30	$38.21 < f_{NL} < 64.33$	$27.41 < f_{NL} < 78.91$	38.8
DESI	Linear (Conservative II)	36.63	38.11	$26.32 < f_{NL} < 49.86$	$16.36 < f_{NL} < 63.12$	39.6
DESI	Nonlinear (Cons. II)	28.58	29.79	$18.91 < f_{NL} < 40.59$	$9.47 < f_{NL} < 52.73$	34.6
DESI	Nonlin. (Cons. II+nStar)	16.63	17.52	$7.51 < f_{NL} < 27.53$	$-1.59 < f_{NL} < 38.49$	35.2
DESI	Nonlin. (All Maps+nStar)	-5.87	-9.19	$-21.45 < f_{NL} < 2.40$	$-33.81 < f_{NL} < 12.06$	39.5
DESI (imag. cut)	Nonlin. (Cons. II)	29.16	30.57	$19.05 < f_{NL} < 42.18$	$9.01 < f_{NL} < 54.81$	35.8
DESI (comp. cut)	Nonlin. (Cons. II)	28.07	29.48	$18.38 < f_{NL} < 40.50$	$8.81 < f_{NL} < 53.10$	34.5
DESI	Nonlin. (Cons. II) + cov	31.62	33.11	$20.94 < f_{NL} < 45.24$	$10.56 < f_{NL} < 59.16$	33.5
BASS+MzLS	Nonlin. (Cons. II)	15.43	19.01	$-1.17 < f_{NL} < 39.43$	$-19.19 < f_{NL} < 63.56$	35.6
BASS+MzLS	Nonlin. (Cons. II+nStar)	13.12	15.39	$-4.59 < f_{NL} < 35.56$	$-24.88 < f_{NL} < 59.31$	34.7
BASS+MzLS	Nonlin. (All Maps+nStar)	-3.73	-6.34	$-27.11 < f_{NL} < 13.75$	$-47.44 < f_{NL} < 33.94$	36.8
BASS+MzLS (imag. cut)	Nonlin. (Cons. II)	25.03	29.12	$6.16 < f_{NL} < 52.44$	$-14.22 < f_{NL} < 80.54$	36.2
BASS+MzLS (comp. cut)	Nonlin. (Cons. II)	16.99	20.90	$0.26 < f_{NL} < 41.76$	$-18.30 < f_{NL} < 67.12$	35.8
DECaLS North	Nonlin. (Cons. II)	41.02	44.89	$23.33 < f_{NL} < 66.78$	$4.96 < f_{NL} < 93.02$	41.1
DECaLS North	Nonlin. (Cons. II+CALIBZ+HI)	55.46	60.44	$36.78 < f_{NL} < 84.05$	$17.86 < f_{NL} < 112.81$	38.4
DECaLS North	Nonlin. (Cons. II+nStar)	31.45	34.78	$14.14 < f_{NL} < 55.79$	$-5.81 < f_{NL} < 80.80$	41.2
DECaLS North	Nonlin. (All Maps+nStar)	0.81	-5.68	$-29.73 < f_{NL} < 16.71$	$-53.15 < f_{NL} < 36.19$	45.1
DECaLS North (no DEC cut)	Nonlin. (Cons. II)	41.05	44.82	$23.58 < f_{NL} < 66.08$	$6.40 < f_{NL} < 91.42$	40.7
DECaLS North (imag. cut)	Nonlin. (Cons. II)	43.27	48.39	$24.60 < f_{NL} < 72.50$	$4.71 < f_{NL} < 101.42$	35.1
DECaLS North (comp. cut)	Nonlin. (Cons. II)	40.55	44.63	$22.41 < f_{NL} < 67.11$	$3.95 < f_{NL} < 94.06$	41.4
DECaLS South	Nonlin. (Cons. II)	31.24	33.21	$14.89 < f_{NL} < 52.40$	$-5.11 < f_{NL} < 74.35$	30.2
DECaLS South	Nonlin. (Cons. II+CALIBZ+HI)	33.79	37.50	$17.71 < f_{NL} < 57.42$	$-0.31 < f_{NL} < 80.94$	30.8
DECaLS South	Nonlin. (Cons. II+nStar)	14.34	6.28	$-21.19 < f_{NL} < 30.01$	$-53.63 < f_{NL} < 49.51$	31.9
DECaLS South	Nonlin. (All Maps+nStar)	-36.76	-32.01	$-49.38 < f_{NL} < -13.61$	$-65.26 < f_{NL} < 7.52$	31.5
DECaLS South (no DEC cut)	Nonlin. (Cons. II)	43.79	46.79	$30.16 < f_{NL} < 63.41$	$16.38 < f_{NL} < 82.72$	23.8
DECaLS South (imag. cut)	Nonlin. (Cons. II)	26.47	23.36	$3.18 < f_{NL} < 47.84$	$-57.69 < f_{NL} < 71.39$	30.0
DECaLS South (comp. cut)	Nonlin. (Cons. II)	29.62	31.76	$13.00 < f_{NL} < 51.58$	$-9.78 < f_{NL} < 74.28$	29.7

plate (*nonlinear conservative II+ nStar*) does not change the constraints from BASS+MzLS much, but it shifts the f_{NL} distributions to lower values in DECaLS North and DECaLS South by 0.5σ and σ, respectively, reconciling all constraints with $f_{NL} = 0$. This might indicate that there are either some unresolved issues with the stellar contamination in DECaLS North and DECaLS South. Although our mock test indicates that we could expect a mitigation bias around $\Delta f_{NL} \sim 10-16$. So we can argue that some of the shift in constraints can be associated with the fact that the stellar density template is correlating with large-scale structure. We note that differences are more significant when all maps and stellar density are used as input. This is expected as cleaning the data with more maps is more prone to the over-correction issue.

- **Pixel completeness (*comp. cut*)**: We discard pixels with $f_{pix} < 0.5$ from the analysis to assess the effect of partially complete pixels on f_{NL}. The f_{pix} cut removes only .6% of the survey area, but no changes in the f_{NL} constraints are observed.

- **Imaging quality (*imag. cut*)**: Pixels with poor photometry are removed the DR9 sample by applying the following cuts on imaging properties; $E[B-V] < 0.1$, $nStar < 3000$, $depth_g > 23.2$, $depth_r > 22.6$, $depth_z > 22.5$, $psfsize_g < 2.5$, $psfsize_r < 2.5$, and $psfsize_z < 2$. Despite losing 8% of the survey area, we find a negligible impact on the best fit f_{NL} estimate from fitting the DESI power spectrum. When we fit each region separately, we find that the BASS+MzLS constraint increases by $\Delta f_{NL} \sim 10$ while the constraints from DECaLS North and DECaLS South do not show significant changes after this cut.

- **Covariance matrix (*cov*)**: We fit the DR9 power spectrum but use a covariance matrix constructed from the $f_{NL} = 76.92$ mocks. With

the new covariance, a 12% increase in the σf_{NL} is observed. We also find that the best fit and marginalized mean estimates of f_{NL} increase by $10 - 11\%$. Overall, we find that the differences are less than the statistical error.

- **External maps (*CALIBZ+HI*)**: The nonlinear *conservative II* method is re-trained using two additional external maps for the neutral hydrogen column density (HI) and the z-band photometric calibration error (CALIBZ). This test is only performed for the DECaLS region due to the limitation and issues with the maps. With this correction, we find that the best fit f_{NL} increases from 41.02 to 55.46 for DECaLS North and from 31.24 to 33.79 for DECaLS South.

- **Declination mask (*no DEC cut*)**: The fiducial mask removes the spurious islands in DECaLS North and DECaLS South below DEC = -30, where different standard stars are used for photometric calibrations, to avoid potential systematic effects. We analyze the DR9 without these cuts, and find that the best fit and marginalized f_{NL} mean estimates from DECaLS South shift significantly by $\Delta f_{NL} \sim 10$ which supports the likelihood of photometric issues in the DECaLS South region below DEC = -30. While the constraints from DECaLS North do not change significantly.

- **Scale dependence (*varying* ℓ_{min})**: We decrease the largest scale (or increase the lowest harmonic mode ℓ_{min}) while fitting the power spectrum. Fig. 13 illustrates the f_{NL} results for the DESI, BASS+MzLS, DECaLS North, and DECaLS South regions. The points represent the marginalized mean estimates and the errorbars represent 68% confidence derived from MCMC chains. We find noticeable systematic trends in DECaLS North and BASS+MzLS
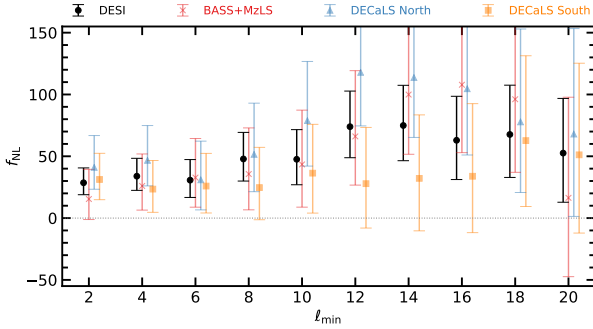
Figure 13. Robustness of the uncalibrated DR9 constraints w.r.t. the largest scale (lowest ℓ mode) used in MCMC regression. Points represent marginalized mean estimates of f_{NL} and errorbars represent 68% confidence.

on scales $12 < \ell < 18$, but the overall constraints are robust against ℓ_{min}.

4.2 Summary

In summary, we find that the nonlinear methods outperform the linear methods in removing the excess clustering signal on large scales. Adding the stellar density map results in significant changes, however when accounted for the mitigation bias, all methods recover the same maximum likelihood estimate. With calibration on the lognormal mocks, the conservative approaches that use a small subset of imaging maps show f_{NL} detection at more than 2σ confidence. The most flexible nonlinear method with all maps and stellar density returns a bigger associated error is larger and consistent with $f_{NL} = 0$. We also run various tests with cuts on the DR9 sample or changing the configuration or details of the analysis. Overall, we find consistent results across sub imaging surveys within DESI. However, our results show that a declination cut at DEC = −30 is necessary for DECaLS South to avoid potential calibration issues. Our analysis does not show a statistical demand for including external templates for HI and CALIBZ, using a different covariance matrix, or imposing additional cuts on the DR9 based on imaging and pixel completeness. We also obtain robust results regardless of the largest scale used for constraining f_{NL}.

5 CONCLUSIONS

We have presented constraints on the local primordial non-Gaussianity parameter f_{NL} from the angular power spectrum of LRGs from the DESI imaging DR9. We infer the redshift distribution of LRGs from early spectroscopy during DESI Survey Validation. The data set covers around 14,000 square degrees in the redshift range of $0.2 < z < 1.1$. Our analysis utilizes the scale-dependent bias effect that primarily comes from large scales; thus, it is sensitive to systematic errors caused by photometric calibration issues, survey depth variations, and Milky Way foregrounds.

We use the FFTLog algorithm to model the angular clustering on large scales. We simulate lognormal density fields with similar angular and redshift distributions to validate the pipeline, estimate covariance matrices, and characterize remaining systematic errors. Our mock test reveals that the distribution of power spectra on large scales is asymmetric. We demonstrate our analysis can benefit from fitting the log transformation of the power spectrum.

Multivariate linear and neural network-based regression models are applied to regress out spurious fluctuations in the LRG density field against various maps for the extinction, survey depth, astronomical seeing, neural hydrogen column density, and stellar density. Feature selection uses the Pearson correlation and the Spearman correlation coefficients to reduce the likelihood of over-correction, i.e., removing the clustering signal. The LRG density map is cross-correlated against the imaging maps. We quantify the remaining systematic fluctuations using the mean density and cross-power spectrum and run null tests against lognormal density fields. Our simulation-based tests reveal that the DR9 LRG sample with linear correction suffers from remaining systematic error primarily due to depth variations. We observe that the nonlinear mitigation approach reduces the excess clustering signal more effectively. We identify the extinction, z-band depth, and r-band seeing as the primary sources of systematic error.

We apply our cleaning methods to the lognormal mocks with and without PNG, with and without systematic effects, to calibrate the level of mitigation biases introduced in our constraints. With this conservative set of maps, we obtain values inconsistent with zero at more than 95% confidence. Adding a template for local stellar density does not change the limits. However, using all imaging maps and stellar density yields an asymmetric likelihood distribution. Nevertheless, our nonlinear cleaning methods return consistent maximum likelihood estimates. Assuming Planck's measurement is accurate a priori, our results indicate some unknown systematic error. Our results suggest a follow-up investigation of stellar contamination and depth-related variations in the spectroscopic sample of DESI LRGs. A second source of theoretical uncertainty is that our analysis considers the halo bias depends on halo mass only. So, a better characterization of observational uncertainties, coupled with more simulation-based studies of halo-assembly bias, is crucial to constrain local primordial non-Gaussianity with large-scale structure.

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 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."

DATA AVAILABILITY

The DR9 catalogs from the DESI Legacy Imaging Surveys are publicly available at https://www.legacysurvey.org/dr9/. The software and codes used for cleaning the imaging data are available at https://github.com/mehdirezaie/sysnetdev. The lognormal mock catalogs can be made available upon reasonable request.

REFERENCES

Abbott T., et al., 2016, Monthly Notices of the Royal Astronomical Society, 460, 1270
Ade P., et al., 2019, Journal of Cosmology and Astroparticle Physics, 2019, 056
Aghamousa A., et al., 2016, arXiv preprint arXiv:1611.00036
Aghanim N., et al., 2020, Astronomy & Astrophysics, 641, A6
Akrami Y., et al., 2019, arXiv preprint arXiv:1905.05697
Alvarez M., et al., 2014, arXiv e-prints, p. arXiv:1412.4671
Arlot S., Celisse A., 2010, Statistics Surveys, 4, 40
Baldauf T., Seljak U., Senatore L., 2011a, Journal of Cosmology and Astroparticle Physics, 2011, 006
Baldauf T., Seljak U., Senatore L., Zaldarriaga M., 2011b, Journal of Cosmology and Astroparticle Physics, 2011, 031
Barreira A., 2020, J. Cosmology Astropart. Phys., 2020, 031
Barreira A., 2022, J. Cosmology Astropart. Phys., 2022, 013
Barreira A., Cabass G., Schmidt F., Pillepich A., Nelson D., 2020, J. Cosmology Astropart. Phys., 2020, 013
Bassett B. A., Tsujikawa S., Wands D., 2006, Reviews of Modern Physics, 78, 537
Bautista J. E., et al., 2021, MNRAS, 500, 736
Beutler F., et al., 2014, Monthly Notices of the Royal Astronomical Society, 443, 1065
Beutler F., Biagetti M., Green D., Slosar A., Wallisch B., 2019, Physical Review Research, 1, 033209

Biagetti M., 2019, Galaxies, 7, 71
Bouwens R. J., et al., 2015, ApJ, 803, 34
Cabass G., Ivanov M. M., Philcox O. H. E., Simonović M., Zaldarriaga M., 2022, Phys. Rev. D, 106, 043506
Castorina E., Moradinezhad Dizgah A., 2020, J. Cosmology Astropart. Phys., 2020, 007
Castorina E., et al., 2019, J. Cosmology Astropart. Phys., 2019, 010
Center O. S., 1987, Ohio Supercomputer Center, http://osc.edu/ark:/19495/f5s1ph73
Chapman M. J., et al., 2022, MNRAS, 516, 617
Chaussidon E., et al., 2022, Monthly Notices of the Royal Astronomical Society, 509, 3904
Chon G., Challinor A., Prunet S., Hivon E., Szapudi I., 2004, Monthly Notices of the Royal Astronomical Society, 350, 914
Coles P., Jones B., 1991, Monthly Notices of the Royal Astronomical Society, 248, 1
D'Amico G., Lewandowski M., Senatore L., Zhang P., 2022, arXiv e-prints, p. arXiv:2201.11518
Dalal N., Dore O., Huterer D., Shirokov A., 2008, Physical Review D, 77, 123514
De Mattia A., Ruhlmann-Kleider V., 2019, Journal of Cosmology and Astroparticle Physics, 2019, 036
Dey A., et al., 2018, arXiv preprint arXiv:1804.08657
Eisenstein D. J., et al., 2001, The Astronomical Journal, 122, 2267
Fang X., Krause E., Eifler T., MacCrann N., 2020, Journal of Cosmology and Astroparticle Physics, 2020, 010
Fillmore J. A., Goldreich P., 1984, Astrophysical Journal, 281, 1
Flaugher B., et al., 2015, The Astronomical Journal, 150, 150
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Gaia Collaboration et al., 2018, A&A, 616, A1
Giannantonio T., Ross A. J., Percival W. J., Crittenden R., Bacher D., Kilbinger M., Nichol R., Weller J., 2014, Physical Review D, 89, 023511
Gil-Marín H., et al., 2020, MNRAS, 498, 2492
Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, The Astrophysical Journal, 622, 759
Guth A. H., Kaiser D. I., 2005, Science, 307, 884
HI4PI Collaboration et al., 2016, A&A, 594, A116
Heinrich C., Doré O., 2022, in American Astronomical Society Meeting Abstracts. p. 202.03
Hivon E., Górski K. M., Netterfield C. B., Crill B. P., Prunet S., Hansen F., 2002, The Astrophysical Journal, 567, 2
Ho S., et al., 2015, J. Cosmology Astropart. Phys., 2015, 040
Huterer D., Cunha C. E., Fang W., 2013, Monthly Notices of the Royal Astronomical Society, 432, 2945
Kitanidis E., et al., 2020, Monthly Notices of the Royal Astronomical Society, 496, 2262
Kofman L., Linde A., Starobinsky A. A., 1994, Physical Review Letters, 73, 3195
Komatsu E., Spergel D. N., 2001, Physical Review D, 63, 063002
Lazeyras T., Barreira A., Schmidt F., Desjacques V., 2023, J. Cosmology Astropart. Phys., 2023, 023
Loshchilov I., Hutter F., 2016, arXiv e-prints, p. arXiv:1608.03983
Loshchilov I., Hutter F., 2017, arXiv e-prints, p. arXiv:1711.05101
Lyth D. H., Liddle A. R., 2009, The primordial density perturbation: Cosmology, inflation and the origin of structure. Cambridge University Press
Meisner A. M., Lang D., Schlegel D. J., 2018, Research Notes of the American Astronomical Society, 2, 1
Merz G., et al., 2021, Monthly Notices of the Royal Astronomical Society, 506, 2503
Mueller E.-M., et al., 2022, Monthly Notices of the Royal Astronomical Society
Myers A. D., et al., 2022, arXiv e-prints, p. arXiv:2208.08518
Nair V., Hinton G. E., 2010, in Proceedings of the 27th international conference on machine learning (ICML-10). pp 807–814
Padmanabhan N., et al., 2007, MNRAS, 378, 852

Peacock J., Nicholson D., 1991, Monthly Notices of the Royal Astronomical Society, 253, 307

Prakash A., et al., 2016, The Astrophysical Journal Supplement Series, 224, 34

Pullen A. R., Hirata C. M., 2013, Publications of the Astronomical Society of the Pacific, 125, 705

Rezaie M., et al., 2021, Monthly Notices of the Royal Astronomical Society, 506, 3439

Riquelme W., et al., 2022, arXiv preprint arXiv:2209.07187

Ross A. J., et al., 2011, Monthly Notices of the Royal Astronomical Society, 417, 1350

Ross A. J., et al., 2020, MNRAS, 498, 2354

Sabti N., Muñoz J. B., Blas D., 2021, J. Cosmology Astropart. Phys., 2021, 010

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, The Astrophysical Journal, 500, 525

Schmittfull M., Seljak U., 2018, Phys. Rev. D, 97, 123540

Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, Journal of Cosmology and Astroparticle Physics, 2008, 031

Tegmark M., et al., 2004, Phys. Rev. D, 69, 103501

Wang M. S., Beutler F., Bacon D., 2020, MNRAS, 499, 2598

Weaverdyck N., Huterer D., 2021, MNRAS, 503, 5061

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Physics reports, 530, 87

Wilson M. J., Peacock J. A., Taylor A. N., de la Torre S., 2017, Monthly Notices of the Royal Astronomical Society, 464, 3121

Wright E. L., et al., 2010, AJ, 140, 1868

Xavier H. S., Abdalla F. B., Joachimi B., 2016, Monthly Notices of the Royal Astronomical Society, 459, 3693

Zhou R., et al., 2021, Monthly Notices of the Royal Astronomical Society, 501, 3309

Zhou R., et al., 2022, arXiv preprint arXiv:2208.08515

Zou H., et al., 2017, Publications of the Astronomical Society of the Pacific, 129, 064101

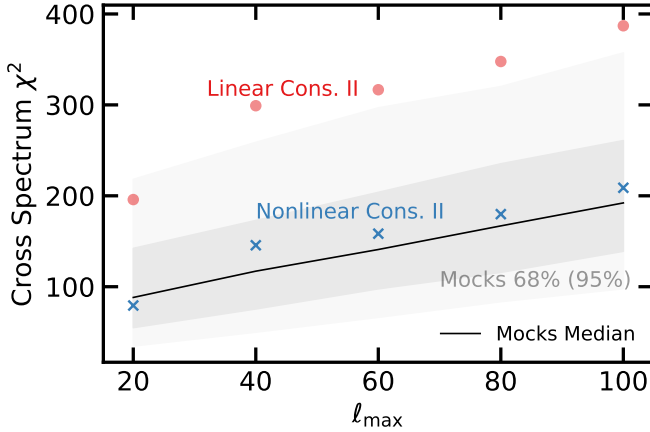de Putter R., Gleyzes J., Doré O., 2017, Physical Review D, 95, 123507

Figure A1. Cross power spectrum χ^2 as a function of the highest mode ℓ_{max} for the DR9 LRG sample using the linear and nonlinear imaging weights with the conservative II maps. The lowest mode is fixed at $\ell_{min} = 2$. Solid curve and dark (light) shade represent the median estimate and 68% (95%) confidence constructed from the $f_{NL} = 0$ mocks.

APPENDIX A: SCALE DEPENDENT SYSTEMATICS

HERE We further test the stability of our results by extending the highest mode from $\ell = 20$ to 100, or fluctuations over scales as small as 1.8 degrees (see, Fig. A1). The solid line shows how the median of 1000 mocks changes as the highest ℓ increases from 20 to 100. The red circles show the chi2 for the linear approach with three maps and the blue crosses show the chi2 for the nonlinear approach with three maps. However, as we show later, our second diagnostic based on the mean density contrast reveals that there is a residual systematic error against the z-band depth with the linear cleaning approach even though the z-band depth was an input for training.

APPENDIX B: UNCALIBRATED f_{NL} CONSTRAINTS FROM DR9

APPENDIX C: LOGNORMAL MOCKS

C1 Clean density fields

Corner plots of the PNG parameter f_{NL} and bias coefficient are shown in Fig. **??** for fitting the mean power spectrum of mocks, with and without f_{NL}. Best fit estimates, marginalized mean, 1σ and 2σ confidence intervals are summarized in Tab. C1. Fig **??** (right) shows confidence contours for different combinations of target variable (e.g., either power spectrum or its log transform) and covariance matrix. First we attempt to understand the impact of covariance on confidence intervals. We fit the mean power spectrum of $f_{NL} = 76.9$ mocks or its log transformation using covariance matrices constructed from the same set of mocks or from the $f_{NL} = 0$ mocks. When covariance is consistent with mean, the difference between fitting power spectrum and log of it is only 2%. If a wrong covariance is used for the log power, the effect is only 7%. However, when mean power spectrum of the $f_{NL} = 76.9$ mocks is fit using the covariance matrix estimated from the $f_{NL} = 0$ mocks, the constraints improve by a factor of 5, simply due to a false higher signal to noise ratio. Therefore, we argue that fitting logarithm of power spectrum would remove the need for having f_{NL}-dependent covariance matrices and make the constraints less sensitive to covariance construction. Fig.
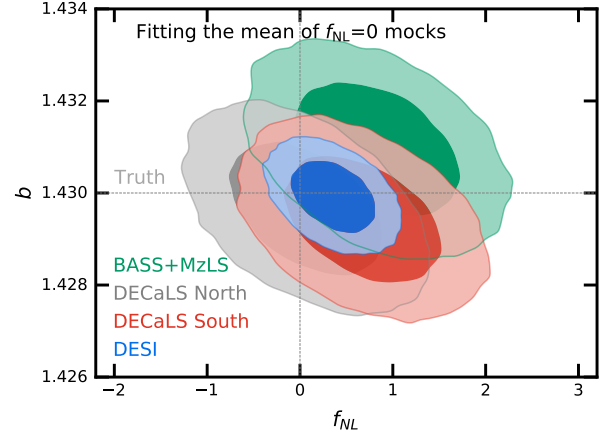
Figure C1. 68% and 95% confidence contours from the mean power spectrum of the $f_{NL} = 0$ mocks for the DESI footprint and sub-imaging surveys. The truth values are represented by vertical and horizontal lines.
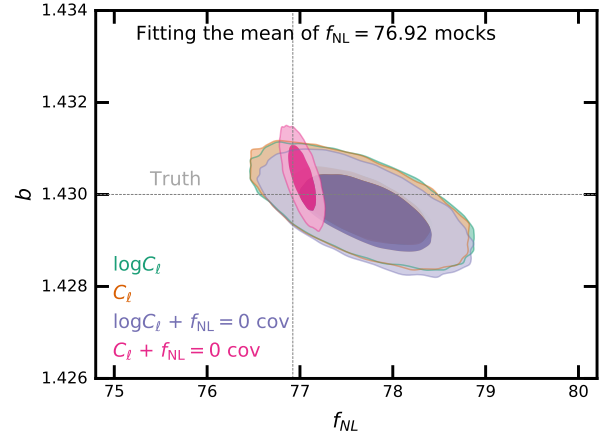

Figure C2. 68% and 95% confidence contours of fitting the mean power spectrum or its log transformation from the $f_{NL} = 76.92$ mocks for the DESI footprint. Using the log C_ℓ fitting yield constraints that are insensitive to the covariance used. The truth values are represented by vertical and horizontal lines.

?? (left) shows the confidence contours for $f_{NL} = 0$ mocks when fit is done to the log of mean spectra of $f_{NL} = 0$ mocks for the different regions. We find that the underlying true f_{NL} value is recovered within 2σ confidence. Add a paragraph for the constraining power vs fsky.

Fig C3 shows the best fit estimates for b vs f_{NL} for $f_{NL} = 0$ and $= 76.92$ mocks in the left and right, respectively. Truth values are represented via the dotted lines. The points are color-coded with the minimum χ^2 from fit for each realization. The histograms of best fit f_{NL} estimates are plotted in the background. We obtain $f_{NL} = MU \pm STD$ and $= MU \pm STD$ for the left and right panels, respectively.

C1.1 Contaminated density fields

Table C1. Best fit and marginalized mean estimates for f_{NL} from fitting the mean power spectrum of the mocks. Degree of freedom is 34 (37 data points - 3 parameters).

Mock / f_{NL}	Footprint	Observable	f_{NL}				χ^2
			Best fit	Mean	68% CL	95% CL	
Clean 76.92	DESI	$\log C_\ell$	77.67	77.67	$77.17 < f_{NL} < 78.16$	$76.71 < f_{NL} < 78.64$	38.8
Clean 76.92	DESI	C_ℓ	77.67	77.65	$77.17 < f_{NL} < 78.14$	$76.70 < f_{NL} < 78.60$	39.0
Clean 76.92	DESI	$\log C_\ell + f_{NL} = 0$ cov	77.70	77.71	$77.25 < f_{NL} < 78.17$	$76.81 < f_{NL} < 78.63$	39.9
Clean 76.92	DESI	$C_\ell + f_{NL} = 0$ cov	77.03	77.02	$76.93 < f_{NL} < 77.12$	$76.83 < f_{NL} < 77.22$	207.6
Clean 0	DESI	$\log C_\ell$	0.36	0.36	$0.06 < f_{NL} < 0.65$	$-0.23 < f_{NL} < 0.94$	35.7
Clean 0	DECaLS North	$\log C_\ell$	0.07	0.06	$-0.47 < f_{NL} < 0.60$	$-1.00 < f_{NL} < 1.12$	26.7
Clean 0	DECaLS South	$\log C_\ell$	0.67	0.67	$0.13 < f_{NL} < 1.22$	$-0.40 < f_{NL} < 1.75$	34.3
Clean 0	BASS+MzLS	$\log C_\ell$	0.83	0.82	$0.25 < f_{NL} < 1.40$	$-0.31 < f_{NL} < 1.96$	39.4

Table C2. Best fit and marginalized estimates for f_{NL} from fitting the mean power spectrum of the mocks before and after applying imaging weights.

| Mock / f_{NL}	Method	f_{NL}				χ^2
		Best fit	Mean	68% CL	95% CL	
Clean 0	No Weight	0.36	0.36	$0.06 < f_{NL} < 0.65$	$-0.23 < f_{NL} < 0.94$	35.7
Clean 0	ConsII	-11.64	-11.65	$-12.00 < f_{NL} < -11.30$	$-12.34 < f_{NL} < -10.97$	86.8
Clean 0	ConsII+nStar	-20.14	-20.13	$-20.44 < f_{NL} < -19.82$	$-20.74 < f_{NL} < -19.52$	472.8
Clean 0	All Maps+nStar	-26.91	-26.92	$-27.16 < f_{NL} < -26.68$	$-27.39 < f_{NL} < -26.46$	5481.0
Contaminated 0	ConsII	-12.12	-12.13	$-12.48 < f_{NL} < -11.78$	$-12.83 < f_{NL} < -11.44$	94.0
Contaminated 0	ConsII+nStar	-20.97	-20.98	$-21.28 < f_{NL} < -20.67$	$-21.58 < f_{NL} < -20.37$	556.3
Contaminated 0	All Maps+nStar	-28.13	-28.13	$-28.36 < f_{NL} < -27.90$	$-28.59 < f_{NL} < -27.67$	6760.5
Clean 76.92	No Weight	77.67	77.67	$77.17 < f_{NL} < 78.16$	$76.71 < f_{NL} < 78.64$	38.8
Clean 76.92	ConsII	54.57	54.57	$54.14 < f_{NL} < 55.01$	$53.72 < f_{NL} < 55.45$	603.5
Clean 76.92	ConsII+nStar	38.38	38.38	$37.99 < f_{NL} < 38.78$	$37.60 < f_{NL} < 39.16$	537.0
Clean 76.92	All Maps+nStar	6.04	6.04	$5.72 < f_{NL} < 6.36$	$5.41 < f_{NL} < 6.67$	694.0
Contaminated 76.92	ConsII	54.01	54.00	$53.57 < f_{NL} < 54.44$	$53.15 < f_{NL} < 54.86$	588.0
Contaminated 76.92	ConsII+nStar	37.48	37.49	$37.09 < f_{NL} < 37.88$	$36.70 < f_{NL} < 38.27$	510.7
Contaminated 76.92	All Maps+nStar	4.59	4.58	$4.26 < f_{NL} < 4.90$	$3.95 < f_{NL} < 5.22$	649.7
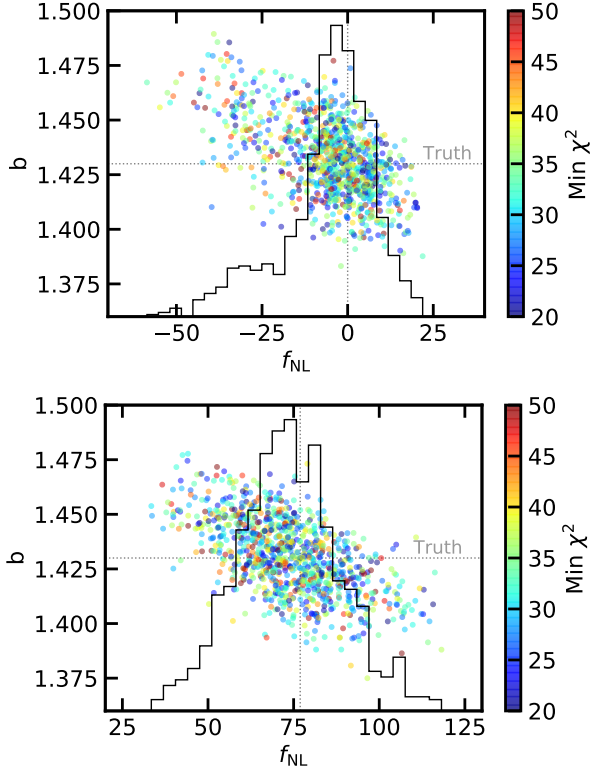
Figure C3. Top: 68% and 95% confidence contours for $f_{\rm NL} = 0$ (left) and 76.92 (right) mocks. Using the $\log C_\ell$ fitting yield constraints that are insensitive to the covariance used. Bottom: best fit estimates from fitting 1000 lognormal mocks with $f_{\rm NL} = 0$ (left) and 76.92 (right) in the DESI footprint. The truth values are represented by vertical and horizontal lines.
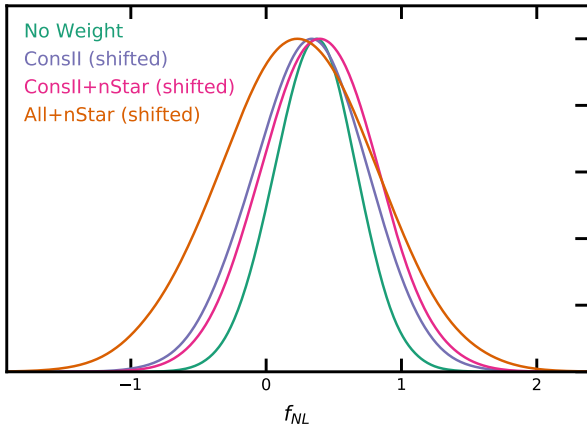

Figure C4. Shifted probability distributions of $f_{\rm NL}$ from the $f_{\rm NL} = 0$ mocks.

This paper has been typeset from a TEX/LATEX file prepared by the author.