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], and DESI Builders

[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 large-scale clustering of luminous red galaxies selected from the Dark Energy Spectroscopic Instrument Legacy Imaging Surveys Data Release 9 to constrain the local primordial non-Gaussianity (PNG) parameter $f_{\rm NL}$. We thoroughly investigate the impact of various photometric systematic effects, such as those caused by Galactic extinction, local stellar density, varying survey depth, and astronomical seeing using spherical harmonics cross power spectrum and mean galaxy density statistics. Lognormal density fields are simulated with and without PNG to construct covariance matrices, evaluate the robustness of power spectrum modeling code, assess whether spurious fluctuations are properly mitigated, and calibrate imaging systematics cleaning methods. With harmonic modes from $\ell = 2$ to 300, we find $36.07(25.03) < f_{\rm NL} < 61.44(75.64)$ with a conservative cleaning approach and $13.09(-15.95) < f_{\rm NL} < 69.14(91.84)$ with an extreme treatment of imaging systematics, both at 68%(95%) confidence. We find significant remaining systematic error raised by calibration issues in the South Galactic Cap and local stellar density in the North Galactic Cap, which induce noticeable biases in $f_{\rm NL}$ constraints. While our constraints are consistent with zero PNG at 95% confidence for the extreme approach, we show that the characterization of stellar contamination and calibration issues are crucial to derive unbiased constraints on $f_{\rm NL}$ in the era of DESI and LSST cosmology.

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

1 INTRODUCTION

Characteristics of the cosmic microwave background (CMB), large-scale structure (LSS), and supernovae (SN) Hubble diagrams are explained to remarkable extents by a cosmological model for the Universe that consists of dark energy, dark matter, and ordinary luminous matter, and has experienced a period of rapid expansion, known as *inflation*, in its early stages (see, e.g., Weinberg et al. 2013). The paradigm of inflation elegantly addresses fundamental issues, such as the isotropy, flatness, and homogeneity of the Universe as well as the absence of magnetic monopole (see, e.g., Weinberg 2008). At the end of inflation, the Universe went through a reheating process, and primordial fluctuations were generated to seed the subsequent growth of structure (Kofman et al. 1994; Bassett et al. 2006; Lyth & Liddle 2009). Even though current observations imply that inflation has certainly happened, characteristics of the field or the fields driving the inflationary expansion remain vastly unknown, and statistical properties of primordial fluctuations pose an intriguing problem in modern observational cosmology. Early studies of cosmological observables have suggested that initial con-

ditions of the Universe are consistent with Gaussian fluctuations (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 gravitational potential ϕ, and it is parameterized by a nonlinear coupling constant f_{NL} (Komatsu & Spergel 2001),

$$\Phi = \phi + f_{NL}[\phi^2 - <\phi^2>]. \tag{1}$$

Standard slow-roll inflation predicts f_{NL} to be of order 10^{-2} (see, e.g., Alvarez et al. 2014, for a review), while multi-field inflationary scenarios anticipate considerably higher than unity f_{NL} values (see, e.g., de Putter et al. 2017). Therefore, getting robust constraints on f_{NL} is the first stepping stone toward better understanding the dynamics of the early Universe. PNG 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 scale-dependent shift in galaxy bias (see, e.g., Dalal et al. 2008; Slosar et al. 2008),

$$\Delta b \sim f_{NL} \frac{(b-p)}{k^2}, \tag{2}$$

where p determines the response of galaxies to the halo gravitational field. If only mass determines how galaxies populate a halo, $p = 1$, which is often referred to as the universality of the halo occupation distribution. 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 might take different values, e.g., $p = 1.6$ (Slosar et al. 2008). Because of the dependence of Δb on k^{-2}, the signature of local primordial non-Gaussianity is more visible on small wavenumbers (or large scales) in the two-point clustering statistics.

The current tightest bound on f_{NL} comes from the three-point clustering measurement of the CMB temperature anisotropies by the Planck satellite, $f_{NL} = 0.9 \pm 5.0$ (Akrami et al. 2019). Upcoming generations of CMB experiments will improve this constraint, but since CMB is limited by cosmic variance, its data alone cannot enhance statistical precision of f_{NL} measurements enough to break the degeneracy amongst various inflationary paradigms (see, e.g., Ade et al. 2019). Combining CMB with LSS data could cancel cosmic variance, partially if not completely, and improve these results to a precision level needed to differentiate between alternative inflationary scenarios (see, e.g., Schmittfull & Seljak 2018). Constraining f_{NL} with the three-point clustering of LSS is likewise hindered by the late-time nonlinear effects raised from structure growth, which is non-trivial to account for (Baldauf et al. 2011b,a). UV Luminosity Function is a novel approach for constraining f_{NL} by probing galaxy abundances and structure formationon small scales (e.g., $k \sim 2$ 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 $-166 < f_{NL} < 497$. 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_{NL} constraints from UV Luminosity Function. Given these limitations, the scale-dependent bias technique is the smoking gun for constraining local PNG with LSS.

Measuring f_{NL} with the scale-dependent bias effect is nonetheless incredibly challenging due to various systematic effects that modulate clustering power on scales where there is a high sensitivity

to f_{NL}. These systematics are broadly classified into theoretical and observational. For instance, survey geometry entangles clustering power on different angular modes (Beutler et al. 2014; Wilson et al. 2017). Relativistic effects also generate scale-dependent signatures on large scales, identical to local PNG, which hinder measuring f_{NL} with the scale-dependent bias effect using higher order multipoles of power spectrum (Wang et al. 2020). Similarly, matter density fluctuations with wavelengths larger than survey volume, 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 the survey size to zero (Peacock & Nicholson 1991; De Mattia & Ruhlmann-Kleider 2019). Accounting for these effects in modeling power spectrum is crucial to derive unbiased f_{NL} constraints (see, e.g., Riquelme et al. 2022).

On the other hand, observational systematics are driven primarily by varying imaging properties 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 model and mitigate, compared to integral constraint and survey geometry, and it has hampered previous studies of local PNG with the scale-dependent bias effect in the large-scale clustering of galaxies and quasars (see, e.g., Ho et al. 2015). For instance, Pullen & Hirata (2013) found that the contribution of stellar density and astronomical seeing is too high for a robust f_{NL} measurement using the quasar sample from the Sloan Digital Sky Survey DR6. 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).

The Dark Energy Spectroscopic Instrument (DESI) uses robotically-driven fibers to collect 5000 spectra simultaneously and is designed to deliver an unparalleled volume of spectroscopic data that will enable future analyses to deepen our understanding of the energy contents of the Universe, specifically, the equation of state for Dark Energy (Aghamousa et al. 2016). Even though the primary focus of DESI is dark energy, assuming imaging systematics are under control, DESI along with other upcoming surveys, such as Rubin Observatory and SphereX, are expected to yield unprecedented constraints on f_{NL} as well (see, e.g., Heinrich & Doré 2022). DESI survey targets galaxies and quasars to construct a 3D map of LSS up to redshifts around 4. DESI preselects its targets from its dedicated imaging surveys, known as the DESI Legacy Imaging Surveys, which are collected from ground-based observations conducted between 2014 and 2019 from three telescopes in Chile and the US. The characterization of potential sources for systematic error in DESI imaging data is of paramount importance to DESI success since spectroscopic catalogs could inherit these issues from imaging catalogs, and thus negatively impact DESI science goals.

The effects of observational systematics in DESI targets have been studied in great detail (see, e.g., Kitanidis et al. 2020; Zhou et al. 2021; Chaussidon et al. 2022). Improving techniques to characterize systematic error in these tracers is crucial for the science beyond dark energy, such as constraining f_{NL} and other features in the primordial power spectrum (Beutler et al. 2019). Some of the current methods seek to mitigate systematic effects by either cross-correlating target density and imaging maps (mode deprojection) or solving a least-square optimization to estimate the contribution from each imaging property to target density (template-based re-

gression), ultimately to regress out the modes affected by imaging properties from target density. Another class of methods aims to cross-correlate different tracers of dark matter to enhance inferences by canceling cosmic variance and by reducing the effect of systematic error, as each tracer might respond differently to a source of systematic error. Giannantonio et al. (2014) presents improved f_{NL} constraints using the integrated Sachs-Wolfe effect. These methods have their limitations and strengths (see, e.g., Weaverdyck & Huterer 2021, for a review). For instance, mode deprojection yields an unbiased clustering but can be employed for angular clustering only, and its involved matrix algebra could prove time-consuming for large survey sizes. Template-based regression is on the other hand computationally economic, but it returns biased clustering by removing some of clustering power, depending on the number of input templates and the flexibility of the regression model. Specifically related to the template-based regression method, there is little effort to calibrate and characterize the amount of clustering power removed during the cleaning process. For studies like BAO and RSD, these effects are demonstrated to be negligible (Merz et al. 2021); however, these effects introduce significant biases in f_{NL} constraints (Mueller et al. 2022) as they highly impact galaxy clustering on large scales (Rezaie et al. 2021).

This paper presents robust constraints on f_{NL} from DESI imaging data with exquisite treatment of imaging systematic error and mitigation biases. We measure the significance of residual systematic error in our data using angular cross-power spectrum (between galaxy density and imaging properties) and mean density contrast of galaxies. Specifically, the robustness of our results is validated against various sources of systematic error, including but not limited to photometric calibration and Milky Way extinction. We cross-correlate the density map of galaxies with the template maps of imaging properties to evaluate the effectiveness of different treatment methods and to characterize the significance of remaining systematic error. Various linear and nonlinear data cleaning approaches are applied with different combinations of imaging templates to quantify the sensitivity of f_{NL} constraints to alternative cleaning methods. The redshift distribution of tracers is determined from early spectroscopic data of the DESI Survey Validation for modeling angular power spectrum. This paper is structured as follows. Section 2 describes the DESI imaging galaxy sample and simulations with and without PNG and imaging systematic effects, and Section 3 outlines the theoretical framework for modeling angular power spectrum and analysis techniques to account for various observational and theoretical systematic error. Finally, we present f_{NL} constraints in Section 4, and conclude with a comparison to previous f_{NL} studies in Section 5.

2 DATA

Luminous red galaxies (LRGs) are massive galaxies that occupy massive halos, lack active star formation, and are one of the highly biased tracers of dark matter gravitational field. A distinct break around 4000 Å in the spectrum of LRGs 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 of $0.4 < z < 1.0$ throughout its five-year mission (Aghamousa et al. 2016). DESI spectroscopy selects its targets from photometry of three ground-
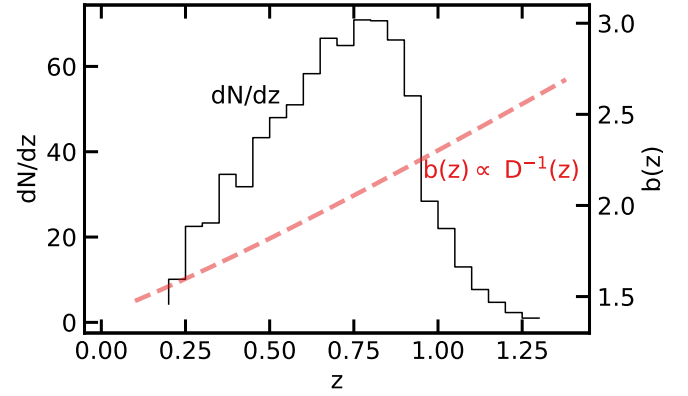

Figure 1. Redshift distribution (solid) and bias evolution (dashed) of the DR9 LRG sample (Zhou et al. 2021, 2022). The redshift distribution is inferred from spectroscopy in the DESI Survey Validation phase, and the model for bias assumes a constant clustering amplitude.

based surveys that observed 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 same instrument under the Dark Energy Survey (Abbott et al. 2016), which constitute about 1130 deg² 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 LRG sample

We work with the photometric LRGs selected from the DESI Legacy Imaging Surveys Data Release 9 (DR9; Dey et al. 2018) using the selection criteria designed for the DESI 1% survey (CITE), 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 Tab. 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). Fig. 1 shows the redshift distribution of the DR9 LRGs (solid black), inferred from the spectroscopic DESI Survey Validation data (CITE), and the evolution of halo bias (red dashed), adopted from Zhou et al. (2021).

DESI-like LRGs are selected brighter than the imaging survey depth limits; therefore, the DR9 LRG density field is nearly homogeneous, unlike the other DESI tracers. To further reduce stellar contamination, the LRG sample is masked rigorously for foreground

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
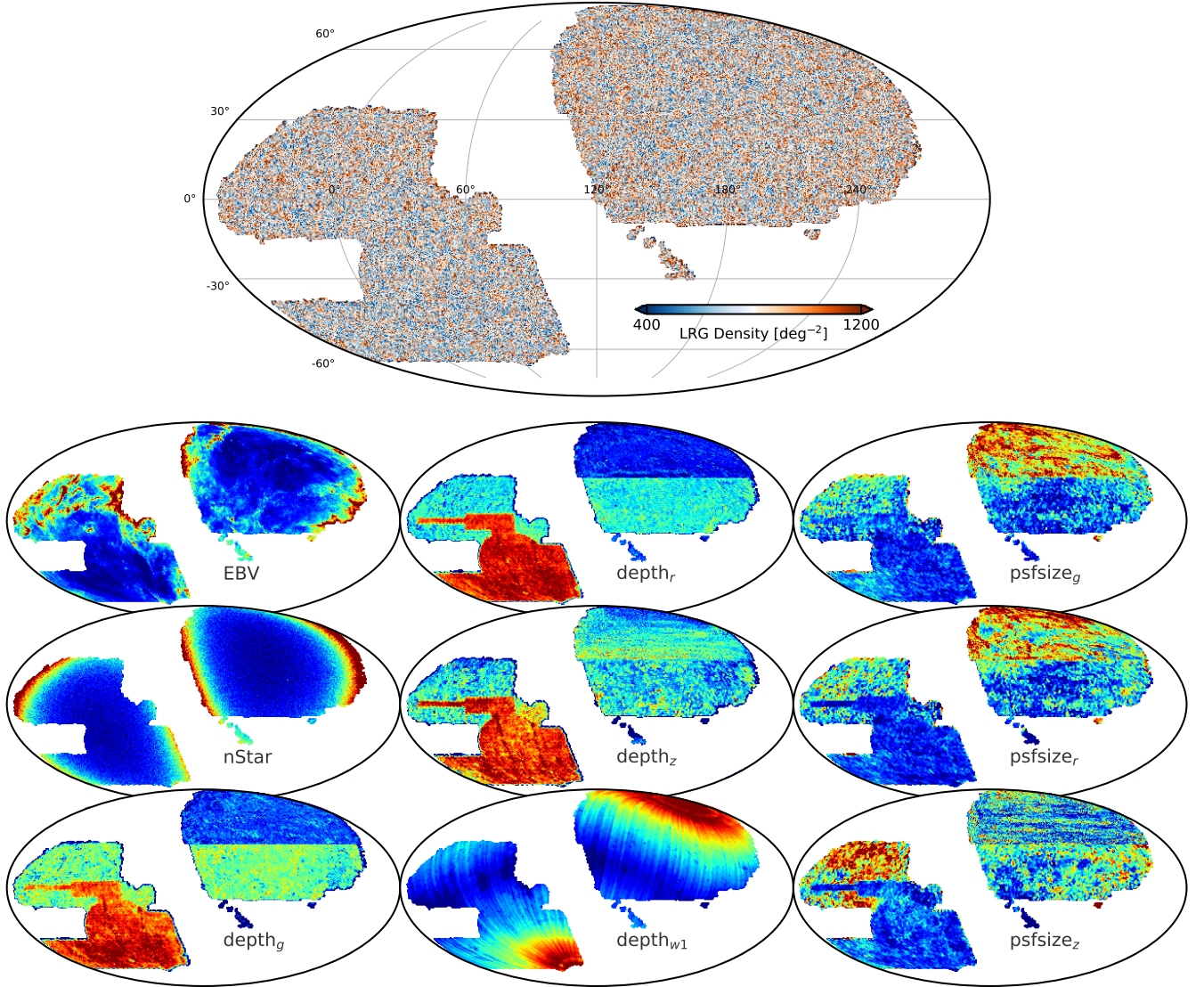
Figure 2. DESI Legacy Imaging Survey Data Release 9 Luminous Red Galaxies and imaging properties (Dey et al. 2018). Top: Observed target density field in deg^{-2}. Spurious disconnected islands from the DECaLS North footprint at Declination below -11 and parts of the DECaLS South with Declination below -30 are rejected from the DR9 sample due to potential calibration issues. Bottom: Mollweide projections of the DR9 catalog imaging properties (survey depth and astronomical seeing/psfsize) and MW foregrounds (extinction and local stellar density) in celestial coordinates.

bright stars, galaxies, and clusters of galaxies[1]. Then, the sample is binned into HEALPIX (Gorski et al. 2005) at NSIDE = 256 to construct the 2D density map with an average surface density of 800 deg^{-2} with sky coverage around 14000 square degrees. We correct for the pixel incompleteness and lost areas in the density field of LRGs using a catalog of random points, hereafter referred to as randoms, uniformly scattered over the footprint with the same cuts and masks applied. Fig. 2 (top) shows the observed density field of the DR9 LRGs in deg^{-2} before applying any correction weights to account for imaging systematic effects. The DR9 LRG density exhibits large-scale spurious fluctuations, which are unlikely to be of cosmological origin. 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 Correlation coefficient analysis

We study the correlation between the DR9 LRG sample and imaging properties as potential sources of systematic error, mapped into HEALPIX at the same NSIDE. Following Zhou et al. (2022), imaging properties investigated in this work are 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); and survey-related imaging properties include survey depth (galaxy depth in the g, r, and z bands and PSF depth in W1) and astronomical seeing (psfsize) in the g, r, and z bands. Templates for these survey-related imaging properties are produced by making the histogram of randoms, which are painted with imaging properties, in HEALPIX and averaging over the attributes of randoms in each pixel.

Fig. 2 (bottom) illustrates the imaging templates investigated as potential sources of systematic error. Each map shows its own characteristic large-scale spurious fluctuations. For instance, the under-dense part of the DR9 LRG sample in the SGC can be associated with survey depth, while the over-density in the NGC can be linked to the extinction map. We reject some parts of the DR9 sample to minimize the potential for photometric calibration systematics. There are some disconnected islands, hereafter referred to as *spurious islands*, in the DECaLS North region at Dec < −11. Additionally, some parts of the DECaLS South footprint with Dec < −30 are removed from the sample, because a different catalog of standard stars is employed to calibrate images below that region. We discuss how these quality cuts influence f_{NL} constraints from the DR9 LRG sample in Section 4.

Fig. 3 shows the Pearson correlation coefficient between the DR9 LRG density and DESI imaging properties for the three imaging surveys (DECaLS North, DECaLS South, and BASS+MzLS) in the top panel. The horizontal curves are constructed from lognormal simulations (see, subsection 2.2) to quantify the significance of correlations. Fig. 3 (bottom) shows the correlation matrix among imaging properties for the DESI footprint. There is a strong correlation between the LRG density and depth maps, and next correlated properties seem to be Galactic foregrounds. There is a small correlation between the LRG density and the W1 depth and psfsize properties. We observe a significant inner correlation among the imaging properties themselves, especially between the local stellar
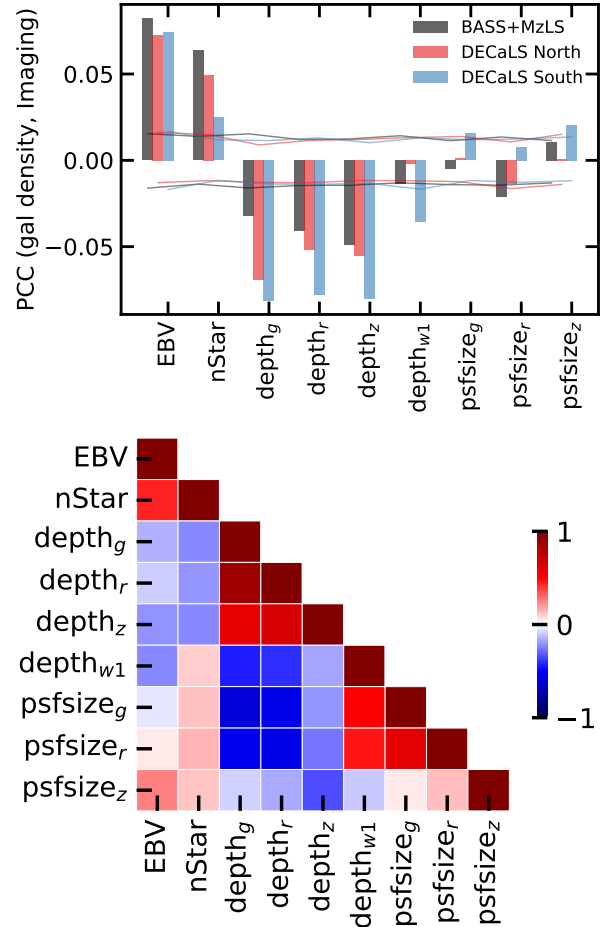
Figure 3. Top: Pearson correlation coefficients between galaxy density and imaging properties in the three imaging regions. Solid curves represent the 95% spread of correlation coefficients observed in 100 randomly selected lognormal mock density realizations. Bottom: Pearson correlation matrix from imaging properties for the full DESI footprint.

density and Milky Way extinction; also, the r-band and g-band survey properties are more correlated with each other than with the z-band.

2.1.2 Imaging weights

We follow a template-based approach to derive a set of weights for removing the effects of imaging systematics. The weights are organized in HEALPIX and the LRG density map is multiplied by these weight maps in an attempt to reduce the level of spurious fluctuations in the LRG data. We regress the DR9 density map against a set of imaging maps, referred to as templates or predictors. Because of the inner correlation amongst the maps, a few subsets of imaging maps are considered as well for a conservative cleaning approach. These subsets are selected to minimize the correlations among the predictors while having a maximum correlation with the observed LRG density map:

- Conservative I: Extinction, depth$_z$
- Conservative II: Extinction, depth$_z$, psfsize$_r$
- All Maps: Extinction, depth in grz and $W1$, psfsize in grz

We also investigate whether including external maps for neutral

hydrogen column density (HI4PI Collaboration et al. 2016) and photometric calibration (e.g., in the z band; *CALIBZ*) could provide more insights on remaining systematic effects. The results are discussed in Section 4.

Linear and nonlinear models (approximated using a neural network) are applied to assess the potential of nonlinear systematic error. Parameters of the models are estimated by optimizing the negative Poisson log-likelihood, $= \sum \lambda - \rho \log(\lambda)$, where the summation runs over pixels, ρ is the observed galaxy density, and λ is the predicted galaxy density given imaging properties $\mathbf{x}$ as input, $\lambda(\mathbf{x}) = \log(1 + e^{f(\mathbf{x})})$. We investigate the use of linear multivariate and nonlinear regression to approximate f. For finding the parameters of the linear model, we perform a Monte Carlo Markov Chain (MCMC) search using the EMCEE package (Foreman-Mackey et al. 2013). For the nonlinear approximation, we employ the implementation of artificial neural networks from Rezaie et al. (2021); specifically, the nonlinear model is an ensemble of 20 neural network models. Each neural network consists of three hidden layers and 20 rectifier units on each layer. The rectifier is the identity function for positive input and zero for negative, and it introduces nonlinearities in the neural network. For the linear model, we use all data for computing the log of posterior during MCMC while for the nonlinear approach, we use 60% of data for training, 20% for validation, and 20% for testing; the training-validation-testing split is applied to minimize the chance of over-fitting by the nonlinear model. We test the nonlinear model on the entire DR9 data by changing the permutation of the training, validation, and testing sets. The neural networks are trained for up to 70 training epochs with ADAM optimizer, which is a variant of stochastic gradient descent. We tune the initial learning rate by minimizing the loss on the validation set and adjust it to dynamically vary between two boundary values of 0.001 and 0.1 to avoid local minima during gradient descent. The best neural network model is then selected from the lowest prediction error when applied to the validation set. Finally, we apply the ensemble of 20 best-fit models to the test set and average over the predictions to construct the predicted galaxy density map in HEALPIX. The predicted galaxy density is employed as weights to down-weight observed density map for removing imaging systematics. We train the linear and nonlinear models for different sets of input imaging maps.

Upon inspecting the predicted density maps, we find that while most of the large-scale spurious fluctuations are explained by just the extinction map and depth in the z band, adding the r-band psfsize results in a finer structure in the predicted density map. We observe that using all imaging maps as input features for regression does not add much information. Comparing the linear to the nonlinear prediction for the same set of input maps, we find that the nonlinear approach yields finer structures due to higher flexibility. Overall, both models predict higher galaxy density near the boundaries where the DESI imaging surveys observed the high extinction regions of the Milky Way. These over-dense regions are likely contaminated artifacts entering the selection, e.g., stellar contaminats or other artifacts because of obscured photometry by MW extinction.

2.2 Synthetic lognormal density fields

Density fluctuations of galaxies can be approximated with lognormal distributions on large scales (Coles & Jones 1991). Unlike N-body simulations, simulating lognormal density fields is not computationally intensive. We use FLASK (Full-sky Lognormal Astro-fields Simulation Kit; Xavier et al. 2016) to generate ensembles of synthetic galaxy density fields that mimic the redshift and angular distributions of the DR9 LRG sample. We create 1000 realizations with $f_{NL} = 0$ and 76.92 using a redshift dependent bias $b(z) = 1.43/D(z)$. We adapt the fiducial cosmology from a flat ΛCDM universe, including one massive neutrino with $m_\nu = 0.06$ eV, and 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 employed throughout the rest of the manuscript. We demonstrate later in Section 4 that our f_{NL} constraints are quite robust against the choice of our fiducial cosmology.

3 ANALYSIS TECHNIQUES

This section describes the estimator for measuring the angular power spectrum and the methodology for modeling it in the presence of PNG. We also demonstrate how we correct for the effects of survey geometry and integral constraint in the modeling. The statistical tools for measuring the remaining systematic error are also presented.

3.1 Measuring Power Spectrum

The quantity that carries cosmological information is galaxy density contrast, δ, which in pixel i is constructed as,

$$\hat{\delta}_i = \frac{\rho_i}{\hat{\bar{\rho}}} - 1, \tag{3}$$

where ρ is the density of galaxies accounted for pixel area $f_{\text{pix},i}$, which is determined by uniformly distributed random galaxies over footprint, and $\hat{\bar{\rho}}$ is the mean galaxy density directly estimated from the data,

$$\hat{\bar{\rho}} = \frac{\sum_i \rho_i f_{\text{pix},i}}{\sum_i f_{\text{pix},i}}. \tag{4}$$

By definition, Eqs. 3 and 4 ensure that the integral of the observed quantity over the footprint vanishes:

$$\sum_i \hat{\delta}_i f_{\text{pix},i} = 0, \tag{5}$$

and this constraint causes an effect commonly referred to as *integral constraint*, which needs to be accounted for in the model. To estimate power spectrum, the galaxy density contrast is expanded in terms of Legendre polynomials,

$$\hat{\delta}_i = \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\theta_i, \phi_i), \tag{6}$$

where θ, ϕ represent the polar and azimuthal angular coordinates of pixel i center, respectively. The cutoff at $\ell = \ell_{\max}$ assumes that modes with $\ell > \ell_{\max}$ do not contribute significantly to signal power. The coefficients $a_{\ell m}$ are then obtained by integrating the density contrast field over the total number of non-empty pixels N_{pix} and using the orthogonality of Legendre polynomials:

$$\hat{a}_{\ell m} = \frac{4\pi}{N_{\text{pix}}} \sum_{i=1}^{N_{\text{pix}}} \hat{\delta}_i f_{\text{pix},i} Y_{\ell m}^*(\theta_i, \phi_i), \tag{7}$$

where * represents the complex conjugate. Then, the angular power spectrum estimator is defined as the variance of $\hat{a}_{\ell m}$ coefficients:

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \hat{a}_{\ell m} \hat{a}_{\ell m}^*. \tag{8}$$

In order to extract $\hat{a}_{\ell m}$ and compute the angular power spectrum, C_ℓ, the function ANAFAST is called from HEALPix (Gorski et al. 2005) with the third order iteration of the quadrature to increase the accuracy[2]. Due to the survey geometry implicit in the summation over the non-empty pixels and explicit in $f_{\text{pix,i}}$, our estimator does not return an unbiased estimate of power spectrum, as different harmonic modes are no longer independent. Therefore, we account for the same effect in the modeling of angular power spectrum. The survey geometry treatment is validated against synthetic datasets.

3.2 Modeling Power Spectrum

Angular power spectrum is a projection of 3D clustering along the line of sight over all possible wavenumbers k. With redshift space distortions included, the projected angular power spectrum of galaxies is related to the 3D linear power spectrum $P(k)$ and shotnoise N_{shot} by (see, e.g., Padmanabhan et al. 2007),

$$C_\ell = \frac{2}{\pi} \int_0^\infty \frac{dk}{k} k^3 P(k) |\Delta_\ell(k)|^2 + N_{\text{shot}}, \quad (9)$$

where shotnoise is assumed to scale-independent, and kernel $\Delta_\ell(k) = \Delta_\ell^{\text{g}}(k) + \Delta_\ell^{\text{RSD}}(k)$ determines how much each wavenumber k contributes to mode ℓ by integrating over all comoving scales r,

$$\Delta_\ell^{\text{g}}(k) = \int \frac{dr}{r} r b(r,k) D(r) \frac{dN}{dr} j_\ell(kr), \quad (10)$$

$$\Delta_\ell^{\text{RSD}}(k) = -\int \frac{dr}{r} r f(r) D(r) \frac{dN}{dr} j_\ell''(kr), \quad (11)$$

where $D(r)$ is the normalized growth factor such that $D(0) = 1$, $f(r)$ is the growth rate, and dN/dr is the normalized redshift distribution of galaxies[3] (see, Fig. 1). In the presence of local primordial non-Gaussianity, $b(r)$ is the linear bias b (see, Fig. 1) plus the scale-dependent shift due to PNG (see, Eq. 2),

$$b(k,r) = b + \frac{2(b-p) f_{\text{NL}} \delta_c}{\alpha(k,r)}, \quad (12)$$

where $\alpha(k,r) \propto k^2 D(r)$, $\delta_c = 1.686$ is the critical density above which gravitational collapse occurs (Fillmore & Goldreich 1984), and the parameter p determines the response of the tracer to halo's gravitational field, e.g., 1 for luminous red galaxies following universality and 1.6 for tracers that are results of recent mergers like quasars. In order to overcome rapid oscillations in spherical Bessel functions, FFTLog[4] algorithm and its extension as implemented in Fang et al. (2020) are employed to evaluate the inner integrations over $d \ln r$.

For a galaxy survey that observes the sky partially, the measured power spectrum is convolved with the survey geometry. This means that the pseudo-power spectrum $\hat{C}_\ell$ obtained by the direct Spherical Harmonic Transforms of a partial sky map, differs from the full-sky angular spectrum C_ℓ. However, their ensemble average is related by a mixing matrix (Hivon et al. 2002),

$$< \hat{C}_\ell > = \sum_{\ell'} M_{\ell\ell'} < C_{\ell'} >, \quad (13)$$

where $M_{\ell\ell'}$ represents the mode-mode coupling from the partial sky coverage. This is known as the Window Function effect and a proper assessment of this effect is crucial for a robust measurement

of the large-scale clustering of galaxies. This window effect is a source of observational systematic error and impacts the measured galaxy clustering, especially on scales comparable to survey size.

We follow a similar approach to that of Chon et al. (2004) to model the window function effect on the theoretical power spectrum C_ℓ rather than correcting the measured pseudo-power spectrum from data. First, we use HEALPix to compute the pseudo-power spectrum of the window $\hat{C}_\ell^{\text{window}}$, which is defined by a mask file in ring ordering format with NSIDE= 256. Then, we transform it to correlation function by,

$$\omega^{\text{window}}(\theta) = \frac{1}{4\pi} \sum_\ell (2\ell + 1) \hat{C}_\ell^{\text{window}} P_\ell(\cos\theta). \quad (14)$$

Next, we normalize ω^{window} such that it is normalized to one at $\theta = 0$. Finally, we multiply the theory correlation function by ω^{window} and transform the result back to ℓ-space,

$$\hat{\omega}^{\text{model}} = \omega^{\text{model}} \omega^{\text{window}} \quad (15)$$

$$\hat{C}_\ell^{\text{model}} = 2\pi \int d\theta \hat{\omega}^{\text{model}}(\theta) P_\ell(\cos\theta). \quad (16)$$

The integral of the galaxy density contrast δ on the footprint is bound to zero, which is often referred to as the *Integral Constraint*. We account for this effect in the modeling by,

$$\hat{C}_\ell^{\text{model,IC}} = \hat{C}_\ell^{\text{model}} - \hat{C}_{\ell=0}^{\text{model}} \left(\frac{\hat{C}_\ell^{\text{window}}}{\hat{C}_{\ell=0}^{\text{window}}} \right) \quad (17)$$

Fig. 4 shows the mean measured power spectrum of 1000 lognormal density fields (dashed) and best fit theory prediction. light and dark shades represent the 68% error on the mean and one single realization, respectively. DESI footprint mask is applied to the mocks, and even though DESI covers around 40% of the sky, but the window effect is affecting modes down to $\ell = 200$. On the other hand, integral constraint only alters the power in the first two bins.

3.3 Parameter estimation

Signature of local PNG is unique and cannot be reproduced with other cosmological parameters. We allow three parameters to vary; f_{NL}, shotnoise, and bias at $z = 0$. Throughout this manuscript, we bin each mode with $\Delta\ell = 2$ between $\ell = 2$ and 20 and $\Delta\ell = 10$ from $\ell = 20$ to 300, while weighting each mode by $2\ell + 1$. We also find that the distribution of power spectrum at the lowest bin, $2 \le \ell < 12$, is not Gaussian and its standard deviation varies significantly from mocks with $f_{\text{NL}} = 0$ to 76.9 (see, Fig. 5). Therefore, we attempt to fit $\log C_\ell$ to make our constraints insensitive to the choice of covariance matrix. The parameter f_{NL} is constrained by maximizing a posterior defined as,

$$-2 \ln \mathcal{L} = (\log C(\Theta) - \log \hat{C})^\dagger \mathbb{C}^{-1} (\log C(\Theta) - \log \hat{C}) + \chi_{\text{priors}}^2, \quad (18)$$

where Θ represents the parameters, f_{NL}, bias at $z = 0$, and shotnoise, all of which are associated with a flat prior, χ_{priors}^2; $C(\Theta)$ is the (binned) theoretical power spectrum including the effects for survey geometry and integral constraint; $\hat{C}$ is the (binned) measured power spectrum; and $\mathbb{C}$ is the covariance matrix constructed from simulations.

3.4 Characterization of residual systematic error

We use the diagnostic tests first applied to SDSS data in Rezaie et al. (2021) based on cross power spectrum between galaxy density field

---

[2] We refer the reader to https://healpix.sourceforge.io/pdf/subroutines.pdf, p. 104-105.
[3] $dN/dr = (dN/dz) * (dz/dr) \propto (dN/dz) * H(z)$
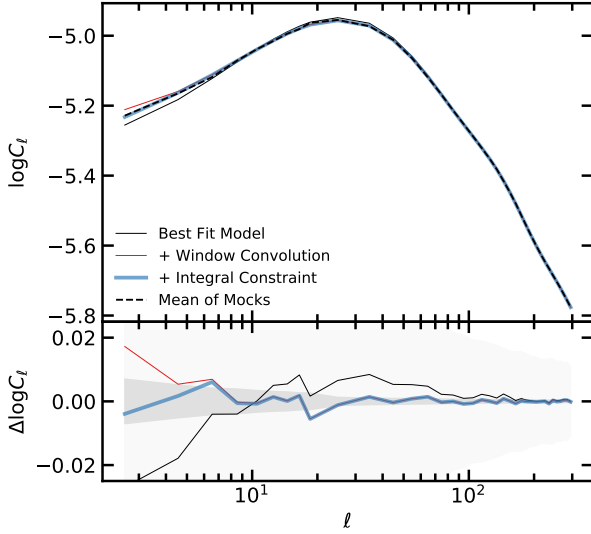[4] github.com/xfangcosmo/FFTLog-and-beyond

Figure 4. Mean power spectrum of 1000 mocks with $f_{NL} = 0$ and best fit theoretical prediction after accounting for various theoretical systematic effects, namely survey geometry effect and integral constraint. Dark and light shades represent 1σ error on the mean and one realization, respectively. Bottom panel shows the residual relative to the mean of mocks. No imaging systematic effects are added to these mocks.
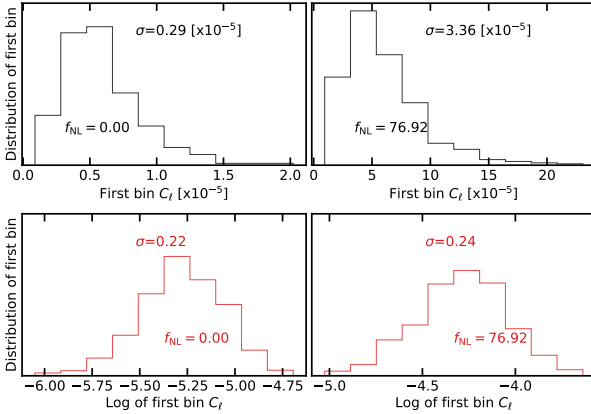

Figure 5. Distribution of the first bin power spectrum constructed from $f_{NL} = 0$ (left) and 76.92 (right) mocks. Differences in the standard deviations become less significant, and power spectrum measurements follow a more symmetric distribution after the log transformation.

and imaging maps and mean density contrast as a function of imaging properties to quantify the significance of imaging systematic effects.

3.4.1 Cross Spectrum

Taking $C_\ell^{g,x}$ as the cross power spectrum between galaxy density contrast field and imaging map, one can normalize this quantity by auto power spectrum of imaging map itself:

$$\hat{C}_{x,\ell} = \frac{(\hat{C}_\ell^{g,x})^2}{\hat{C}_\ell^{x,x}}, \tag{19}$$

and then construct a vector from cross spectra against all other imaging maps:

$$\hat{C}_{X,\ell} = [\hat{C}_{x_1,\ell}, \hat{C}_{x_2,\ell}, \hat{C}_{x_3,\ell}, ..., \hat{C}_{x_9,\ell}]. \tag{20}$$

Finally, cross power spectrum χ^2 can be defined as,

$$\chi^2 = C_{X,\ell}^T \mathbb{C}_X^{-1} C_{X,\ell}, \tag{21}$$

where covariance matrix $\mathbb{C}_X = < C_{X,\ell} C_{X,\ell'} >$ is constructed from mocks without systematic effects. This statistics is measured for every mock realization with the leave-one-out technique to construct a histogram, which is then compared to the χ^2 value observed from the DR9. Fig. 6 (top) shows the measured C_X the DR9 sample before and after applying various imaging weights, relative to that of the mocks with $f_{NL} = 0$. The dispersions of mocks with and without PNG are shown with the shade regions for comparison. We bin the C_X measurements from $\ell = 2$ to 20 with $\Delta\ell = 2$. The mean and standard deviation of $\hat{C}_{X,\ell}$ for 1000 mocks with and without f_{NL} are shown in Fig. 6. Extinction and stellar density have the highest cross power spectrum, and then depth in the z band. After applying the first version of weights with linear conservative I which includes extinction and depth-z, the cross power increases against psftsize in the r band. This indicates that only two maps are not sufficient to null out all of the cross correlations. With linear model there is residual power against extinction, depth-z, and psfsize-z; therefore, we apply weights based on a nonlinear model to account for more complex systematic effects.

Fig. 8 (top) shows the histogram of cross spectrum χ^2 from mocks with and without f_{NL}. Comparing with the data, the residual is 20014.8 before correction, and after applying the first set of weights, it reduces to 375.1 with p-value XXX. Adding psfsize-z, the linear model reduces the error dow to 195.9 (p-value = XXX). Although using all maps gives the lowest error i.e., 129.2, but it could potentially lead to over-fitting true clustering given how correlated the imaging maps are (see, Fig. 3). On the other hand, the nonlinear method with three maps yields a χ^2 value of 79.3 and p-value of XXX, and adding the stellar density map reduces the error to 70.9 (p-value=XXX). This test clearly shows that a nonlinear approach is desired to get a null test. 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. 9). 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.

3.4.2 Mean Density

In the absence of systematic error, the integral of mean density contrast over the footprint should be zero As an alternative test, we calculate the histogram of the density contrast field relative to each imaging map.

$$\delta_x = (\hat{\bar{\rho}})^{-1} \frac{\sum_i \rho_i f_{pix,i}}{\sum_i f_{pix,i}}, \tag{22}$$

where the summations are over pixels in each bin of imaging map x. Similarly, we construct the mean density contract vector against all imaging maps,

$$\delta_X = [\delta_{x_1}, \delta_{x_2}, \delta_{x_3}, ..., \delta_{x_9}], \tag{23}$$

and the total residual error as,

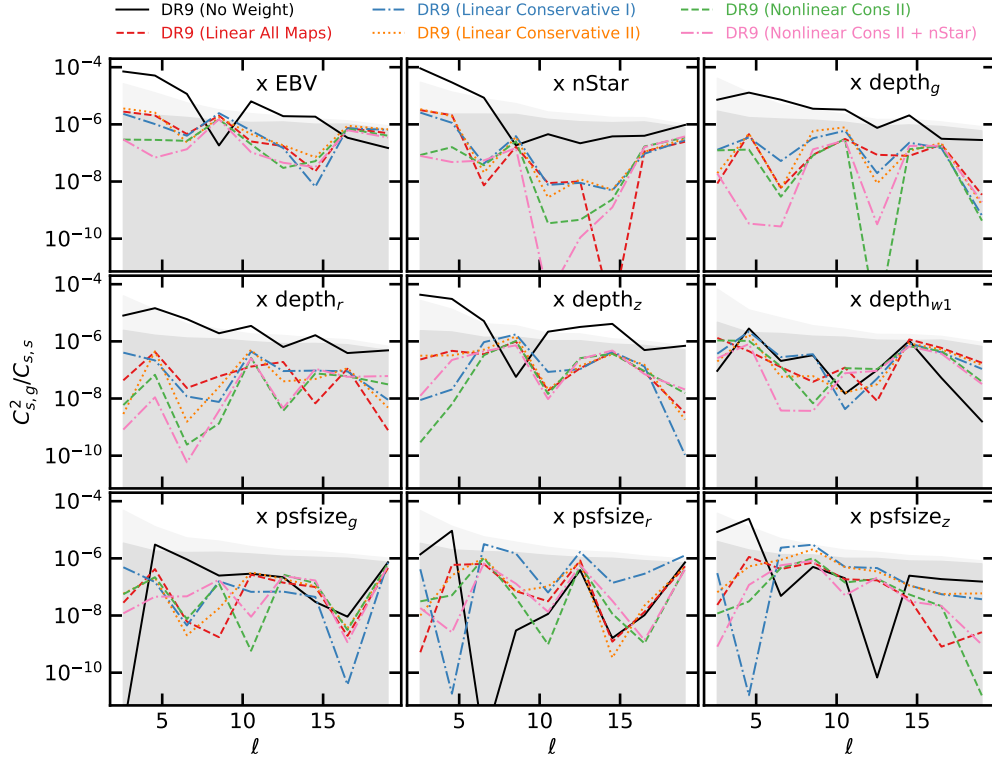$$\chi^2 = \delta_X^T \mathbb{C}_\delta^{-1} \delta_X, \tag{24}$$

Figure 6. Cross power spectra between the DR9 LRG sample and imaging maps. Dark and light shades represent the 97.5 percentile of 1000 lognormal mocks without and with PNG, respectively.
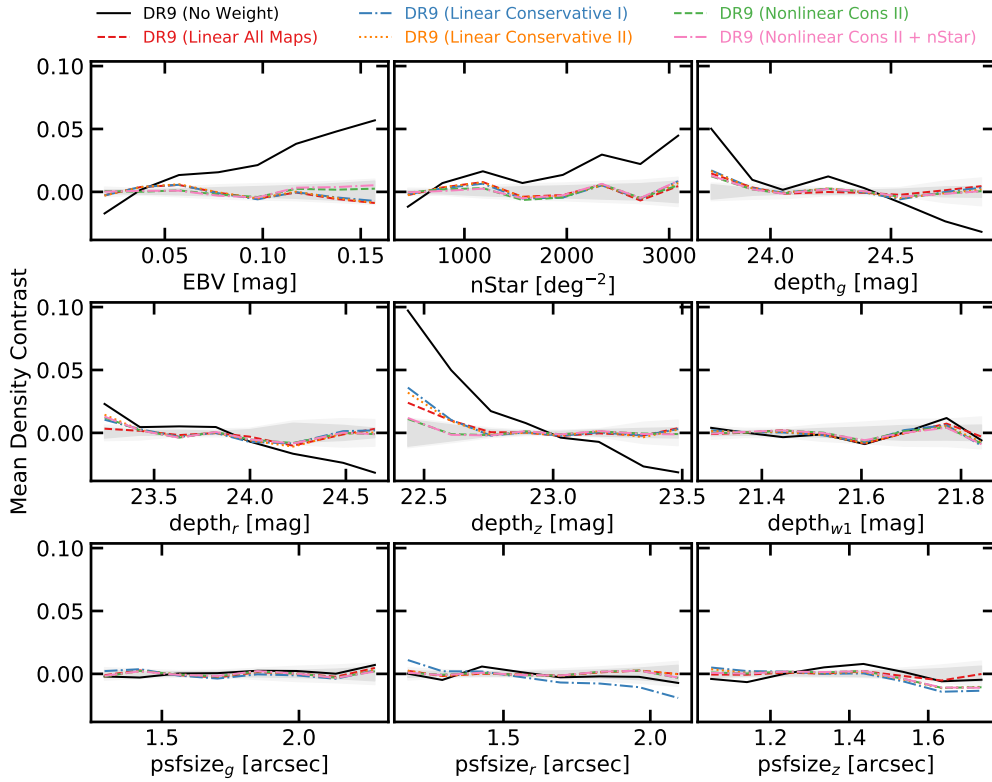

Figure 7. Mean density contrast of the DR9 LRGs as a function of imaging maps. Dark and light shades represent the 1σ dispersion of 1000 lognormal mocks without and with PNG, respectively.
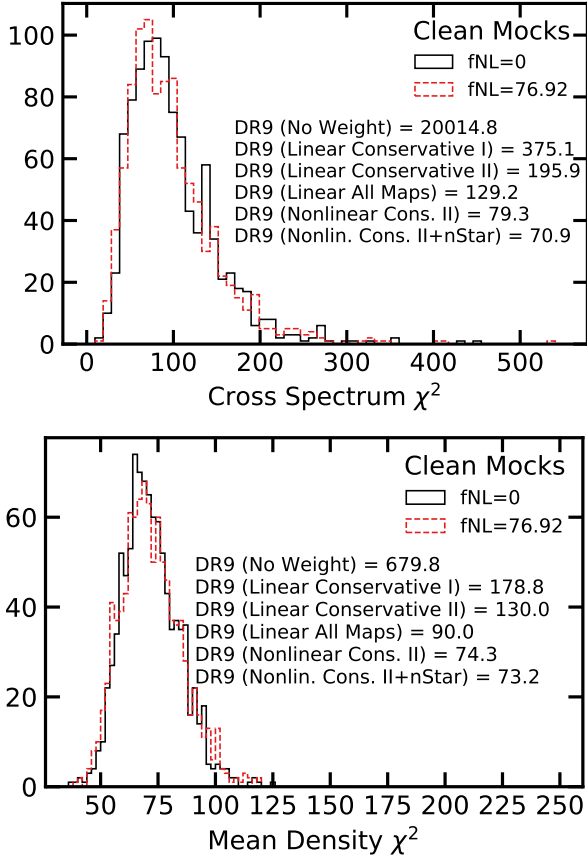
Figure 8. Remaining systematic error χ^2. Top: Cross power spectrum. Bottom: Mean density contrast. The values observed in the DR9 sample before and after linear and nonlinear treatments are quoted, and the histograms are constructed from 1000 realizations of clean lognormal mocks with $f_{\rm NL} = 0$ and 76.92.
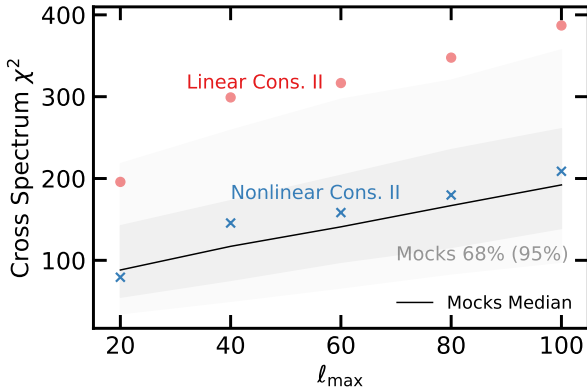

Figure 9. Cross Spectrum χ^2 for DR9 LRG density and imaging maps with linear and nonlinear cleaning (both with conservative II maps) as a function of the highest mode $\ell_{\rm max}$ while the lowest mode is fixed at $\ell_{\rm min} = 2$. Solid curve and dark (light) shade represent the median estimate and 68% (95%) confidence constructed from the $f_{\rm NL} = 0$ mocks.

where the covariance matrix $\mathbb{C}_\delta = <\delta_X \delta_X>$ is constructed from mocks without systematic effects. Fig. 6 (bottom) shows the mean density contrast for the DR9 LRG sample. The shades represent the 1σ level fluctuations observed in 1000 clean mocks with $f_{\rm NL} = 0$ and 76.92. Before treatment (solid) shows strong correlation around 10% against depth in the z band which is consistent with the cross power spectrum. Beside that, there are strong positive trends against extinction and stellar density at about $5 - 6\%$. The linear model is able to mitigate most of the systematic effects with only the extinction and depth-z maps as input, however a new trend appears against psfsize-r which is indicative of psfsize dependence in the sample. This finding is in agreement with the cross power spectrum. Even after applying the linear weights there is some residual against depth-z at around 2%, which indicates the systematic effects might be more complex than what can be removed using a linear model. Nonlinear model with three maps (or four maps including the stellar density) is capable of reducing the fluctuations below 2%.

Fig. 8 (bottom) shows the mean density χ^2 observed in the mocks vs DR9 sample before and after applying imaging weights. Linear weights with two maps reduce the chi-2 value from 679.8 (before) to 178.8. The p-value is indicative of remaining systematic effects. Adding psfsize-r does not help much with the p-value even though it reduces the chi-2 to 130. Using all maps with the linear model gives a more reasonable value however it leaves the analysis susceptible to over-fitting true clustering signal. With nonlinear approach two maps as input, the chi-2 is reasonable 74.3 with p-value XXX, and adding the stellar map does not change the p-value much, indicating the trend against stellar density can be explained with the extinction to a great extend.

4   RESULTS

This sections presents the constraint on $f_{\rm NL}$ from the DR9 LRG sample. The robustness of constraints are tested against various assumptions and details in survey mask, imaging weights, and calibration of data. The default analysis uses the covariance from $f_{\rm NL} = 0$ mocks. We also validate the modeling pipeline and characterize the amount of mitigation bias introduced in $f_{\rm NL}$ after cleaning for systematics.

4.1   DR9 LRGs

Fig. 10 shows the measured power spectrum of the DESI imaging DR9 LRG sample with different imaging weights applied, the best fit theory predictions, and the mean and 1σ error on angular power spectrum measured from the $f_{\rm NL} = 0$ lognormal simulations. Power spectra are similar with differences less than ??% for small scales ($\ell > 2$), and as we go to larger scales, the differences become more significant. There is a little difference between linear cons II and linear all maps. This proves that our feature selection procedure has worked to identify the important maps. Comparing linear cons I to linear cons II, modes with $6 \leq \ell < 10$ are different, indicating scales where psfsize-r is affecting the signal. Comparing nonlinear cons II to linear cons II, modes at the second bin ($4 \leq \ell < 6$) are very different, indicating the nonlinear approach is more flexible to reduce fluctuations on small scales as well large scales.

4.1.1   Calibrated constraints

Tab. 2 summarizes the best fit and marginalized mean estimates of $f_{\rm NL}$ from fitting power spectrum of the DR9 LRG sample us-

Table 2. 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).

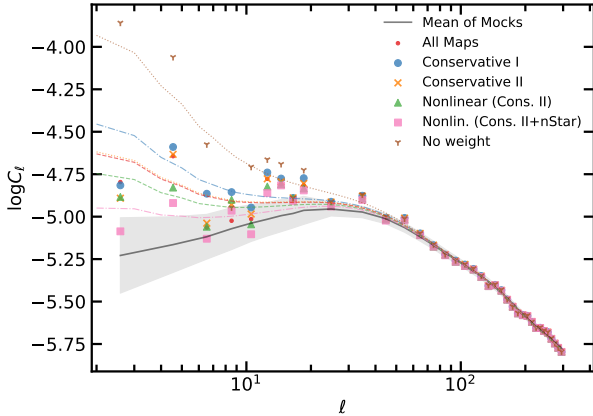| Footprint | Method | f_{NL} | | | | χ^2 |
		Best fit	Mean	68% CL	95% CL	
DESI	No Weight	113.18	115.49	$98.14 < f_{NL} < 132.89$	$83.51 < f_{NL} < 151.59$	44.4
DESI	Nonlinear (Cons. II)	47.38	48.81	$36.08 < f_{NL} < 61.44$	$25.03 < f_{NL} < 75.64$	34.6
DESI	Nonlin. (Cons. II+nStar)	48.92	50.10	$36.88 < f_{NL} < 63.31$	$24.87 < f_{NL} < 77.78$	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 10. Measured power spectrum of the DR9 LRG sample before and after correcting for systematics with their corresponding best fit theory predictions. The shade represents 1σ error constructed from the $f_{NL} = 0$ mocks.

ing the nonlinear approach with various combinations of imaging templates. All constraints are accounted for the mitigation bias using the lognormal mocks, which is further discussed in §4.2.2. With the corrections applied, we obtain $36.07(25.03) < f_{NL} < 61.44(75.64)$ for nonlinear (conservative II), $36.88(24.87) < f_{NL} < 63.31(77.78)$ for nonlinear (conservative II + nStar), and $13.09(-15.95) < f_{NL} < 69.14(91.84)$ for nonlinear (all maps + nStar) at $68\%(95\%)$ confidence. Fig 11 shows the 2D constraints with 68% and 95% confidence on f_{NL} and b for the DESI footprint from the DR9 sample before (no weight) and after applying different cleaning schemes. No weight constraint at 68% is $98.14 < f_{NL} < 132.89$ with a best fit of 113.18 and marginalized mean of 115.49, and is more than 2σ off from zero. Applying imaging weights shifts constraints to lower f_{NL} values, and b is slightly pulled upward since excess clustering due to systematics is removed. Using all maps with the linear model does not change the results, showing that three maps are sufficient at the linear level to mitigate systematics. As an alternative, using a nonlinear model with three maps shows around 1σ shift, with 68% confidence at $18.91 < f_{NL} < 40.59$, inconsistent with zero for more than 2σ. Adding a template for the local stellar density shift constraints by 1σ, making it consistent with zero. As the most rigorous approach, using all maps and stellar density included results in more than 2σ shift. We emphasize that these shifts to lower f_{NL} are somewhat expected as more input maps results in regressing more modes from cosmological clustering signal. Therefore, we use lognormal mocks to calibrate the amount of signal that is removed in each case and attempt to undo the effect.

4.1.2 Robustness tests

These results are subject to mitigation bias. f_{NL} constraints from DR9 LRG sample is summarized in Table 3. First, we focus on the DESI footprint and then compare constraints obtained from each sub-survey. We also evaluate the robustness of constraints against various cuts and configurations. First, we compare how constraints from whole DESI footprint compares to those from each survey individually, namely BASS+MzLS, DECaLS Nouth, and DECaLS South. Fig. 12 shows 68% and 95% confidence on f_{NL} and b from each individual survey or all combined as DESI. Constraints from all surveys are consistent and agree with each other within 68%. Both BASS+MzLS and DECaLS South are consistent with zero PNG, but DECaLS North deviates from zero at more than 2σ. Adding the stellar density template does not change constraints from BASS+MzLS much, but it shifts DECaLS North and DECaLS South by 0.5σ and σ, respectively. This might indicate that there are some unresolved issues with stellar contamination in DECaLS North and DECaLS South. We note that differences are more significant when all maps and stellar density are used as input. This is expected as more maps mean the model has more freedom to take out clustering modes.

- **Pixel completeness** We remove pixels with low completeness from the DESI footprint by applying $f_{pix} > 0.5$, and find that the impact is negligible. Specifically, the cut removes .6% survey area and causes best fit f_{NL} shifts only around 2%, from 28.58 to 28.07, see *comp cut* Table 3. When investigated this impact on each region separately, BASS+MzLS increases around 10%, DECaLS North decreases 1%, and DECaLS South decreases around 5%.
- **Imaging quality** We remove pixels with poor imaging from the DESI footprint 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 $psfsize_z < 2$. Overall the constraints are consistent despite best fit and marginalized mean estimates shift. Quantitatively, we lose about 8.2% survey area, and the best fit f_{NL} estimate changes about 2% from 28.58 to 29.16. See *imag cut* in Table 3. For BASS+MzLS only, the imaging cut increases the best fit by 62% from 15.43 to 25.03. For DECaLS North and DECaLS South, the best fit increases by 5% and 15% respectively.
- **Covariance** We now use the mocks with $f_{NL} = 76.92$ to construct a covariance matrix, and with the new covariance we observe a 12% increase in the f_{NL} constraint uncertainties and 11% increase in the best fit estimate of f_{NL}.
- **Lowest ℓ** We decrease the largest mode (or increase the lowest ℓ) used in estimating the best fit and 68% confidence intervals. Fig. 13 illustrates the results for the DESI footprint and how they compared to BASS+MzLS, DECaLS North, or DECaLS South only results. Points represent marginalized mean estimates of f_{NL} and errorbars represent 68% confidence from MCMC results. Overall we find that the constraints are robust against the largest mode.
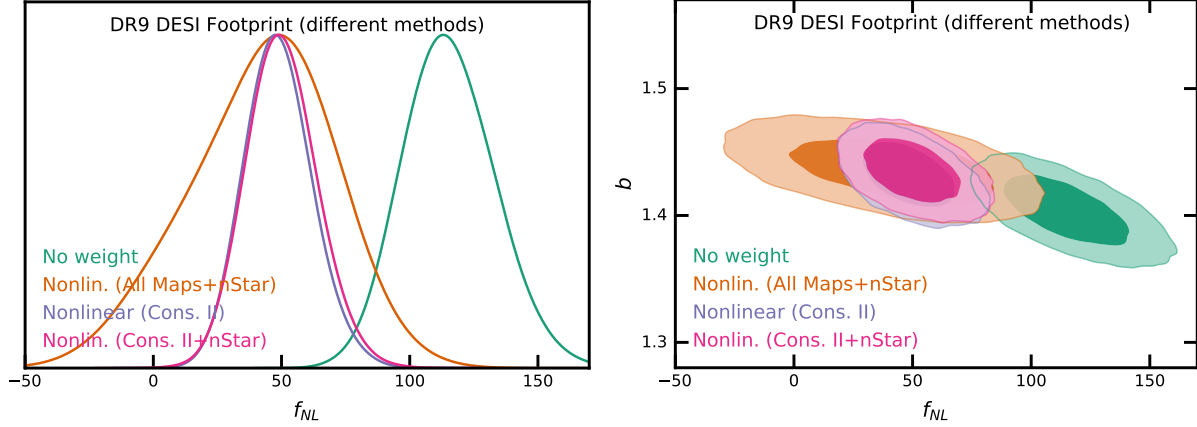
Figure 11. Calibrated 1D and 2D f_{NL} constraints from the DR9 LRG sample before and after applying nonlinear cleaning methods. The dark and light 2D contours represent 68% and 95% confidence intervals, respectively.

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

		f_{NL}				χ^2
Footprint	Method	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)+$f_{NL} = 76.92$ 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 + islands	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 + DEC < -30	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

- **External maps** We also derive imaging weights using additional external maps for the neutral hydrogen column density (HI) and magnitude calibration errors in the z band (CALIBZ). With the new weights, we find the best fit estimates increase from 41.02 to 55.46 for DECaLS North and from 31.24 to 33.79 for DECaLS South.

- **Declination cut** Our default analysis do not use the spurious islands in DECaLS North and DECaLS South below DEC = -30 to avoid potential calibration issues. PANASTARS are used for calibration below DEC of -30. Without these cuts, best fit f_{NL} estimates

increase from 31.24 to 43.79 for DECaLS South and decrease from 41.02 to 41.05 for DECaLS North. This indicates that indeed there is an issue with DECaLS South below DEC of -30.

Overall we find that the declination cut is necessary for DECaLS South, while adding external templates for HI and CALIBZ, using a different covariance, or applying imaging and completeness cuts do not alter the constraints significantly.
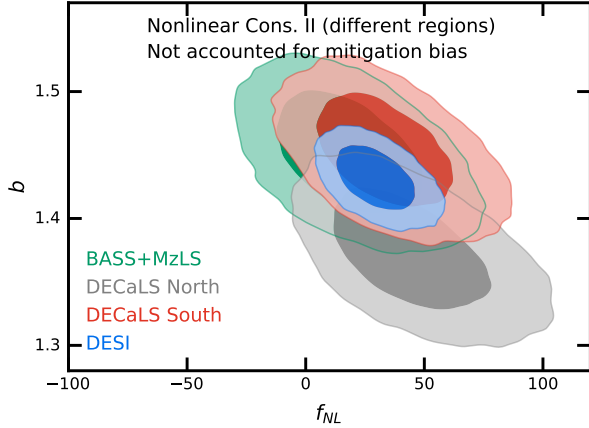
Figure 12. Uncalibrated 2D constraints from the DR9 LRG sample for each imaging survey compared with that for the whole DESI footprint. The dark and light shades represent the 68% and 95% confidence intervals, respectively.
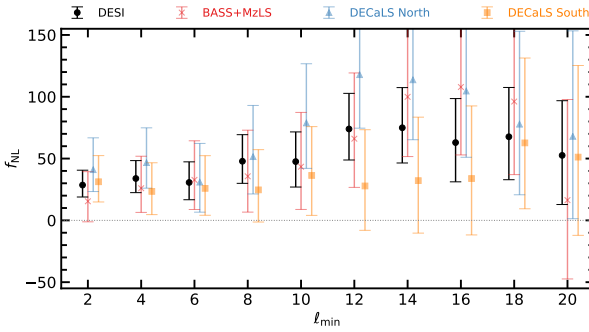

Figure 13. Robustness of the uncalibrated DR9 constraints w.r.t. the largest scale (lowest ℓ mode) used in MCMC regression. Points represent marginalized mean estimates of $f_{\rm NL}$ and errorbars represent 68% confidence.

4.2 Lognormal Mocks

4.2.1 Clean mocks

Corner plots of the PNG parameter $f_{\rm NL}$ and bias coefficient are shown in Fig. **??** for fitting the mean spectrum of mocks, with and without $f_{\rm NL}$. Best fit estimates, marginalized mean, 1σ and 2σ confidence intervals are summarized in Tab. 4. 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 $f_{\rm NL} = 76.9$ mocks or its log transformation using covariance matrices constructed from the same set of mocks or from the $f_{\rm NL} = 0$ 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 $f_{\rm NL} = 76.9$ mocks is fit using the covariance matrix estimated from the $f_{\rm NL} = 0$ 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 $f_{\rm NL}$-dependent covariance matrices and make the constraints less sensitive to covariance construction. Fig. **??** (left) shows the confidence contours for $f_{\rm NL} = 0$ mocks when fit is done to the log of mean spectra of $f_{\rm NL} = 0$ mocks for the different
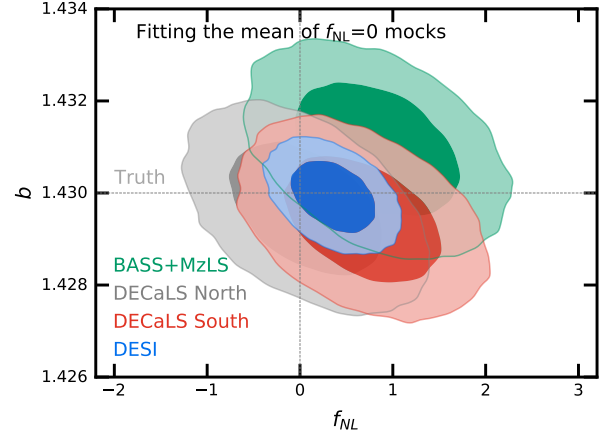
Figure 14. 68% and 95% confidence contours from the mean power spectrum of the $f_{\rm NL} = 0$ mocks for the DESI footprint and sub-imaging surveys. The truth values are represented by vertical and horizontal lines.
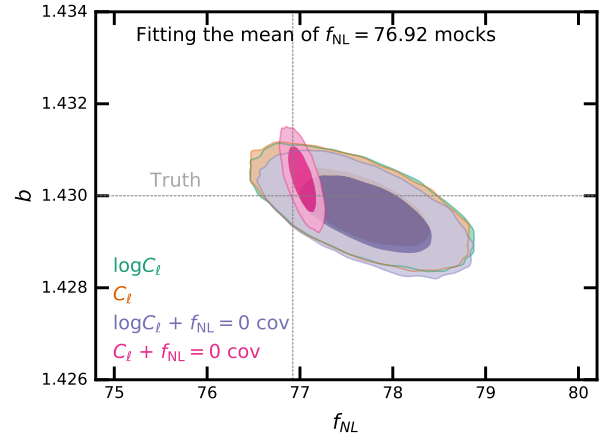

Figure 15. 68% and 95% confidence contours of fitting the mean power spectrum or its log transformation from the $f_{\rm NL} = 76.92$ mocks for the DESI footprint. Using the log C_ℓ fitting yield constraints that are insensitive to the covariance used. The truth values are represented by vertical and horizontal lines.

regions. We find that the underlying true $f_{\rm NL}$ value is recovered within 2σ confidence. Add a paragraph for the constraining power vs fsky.

Fig 16 shows the best fit estimates for b vs $f_{\rm NL}$ for $f_{\rm NL} = 0$ 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 $f_{\rm NL}$ estimates are plotted in the background. We obtain $f_{\rm NL} = MU \pm STD$ and $= MU \pm STD$ for the left and right panels, respectively.

4.2.2 Contaminated mocks

With template based mitigation, the measured power spectrum is biased and the amount of bias depends on the number of input templates. We use a linear model with the set of conservative II maps to simulate imaging systematics in our lognormal density fields. Then, we apply the cleaning method based on nonlinear model with Conservative II, Conservative II + nStar, or All Maps

Table 4. 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} Best fit	Mean	68% CL	95% CL	χ^2
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 5. 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} Best fit	Mean	68% CL	95% CL	χ^2
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

+ nStar sets of imaging maps to both set of mocks; with or without systematic effects (dashed or soild), and with and without f_{NL}. The marginalized mean, confidence intervals, and best fit estimates are presented in Tab 5. This test indicates that as more maps are input to the nonlinear model, more power is removed, and thus the constraints are systematically shifted to lower f_{NL} values. Fig 17 shows the true f_{NL} value and the measured f_{NL} value from fitting the mean of mocks. The results for the contaminated mocks before cleaning (No weight) is not shown for clarity. From this graph, then a pair of linear coefficients are to be found for mapping measured f_{NL} to true f_{NL} values. At the first iteration, we think these coefficients should be applied to the values presented in Tab 3.

5 CONCLUSIONS

We have presented constraints on f_{NL} using the scale-dependent bias effect in the large-scale clustering of DESI imaging DR9 LRGs. Methods from linear and nonlinear regressions are applied for data cleaning from foreground and imaging systematic effects. Same tools are tested on lognormal density fields to evaluate the sensitivity of signal to systematic error. As summarized in Table ??, we find that fitting $\log C_\ell$ rather than C_ℓ minimizes the dependence of constraints to the choice of covariance, and we are able to recover the truth f_{NL} at 95% confidence in both simulations with and without PNG. Table ??? summarizes the constraints from mocks undergone cleaning for systematics. We find that template-based regression removes clustering and thus biased constraints are obtained. We use the mock constraints to calibrate DR9 constraints. We obtain $36.07(25.03) < f_{NL} < 61.44(75.64)$ when cleaning is

performed with the nonlinear model using only three imaging maps. With more extreme cleaning using all maps and stellar density, we obtain $13.09(-15.95) < f_{NL} < 69.14(91.84)$ at 68%(95%).

Various tests are performed to assess the robustness of constraints against analysis assumptions, and the results are summarized in Table ??. We find

- constraints from individual surveys are consistent with each other
- no significant shift observed in constraints after applying imaging cut, completeness cut
- no significant shift after including additional imaging templates for hydrogen column density or calibration in the z-band.
- region below dec of -30 indicates some issues probably due to unaccounted for calibration issues
- constraints are robust against the largest scales (lowest ℓ mode) used in fitting for f_{NL}. Some signs of systematics on $10 < \ell < 18$.
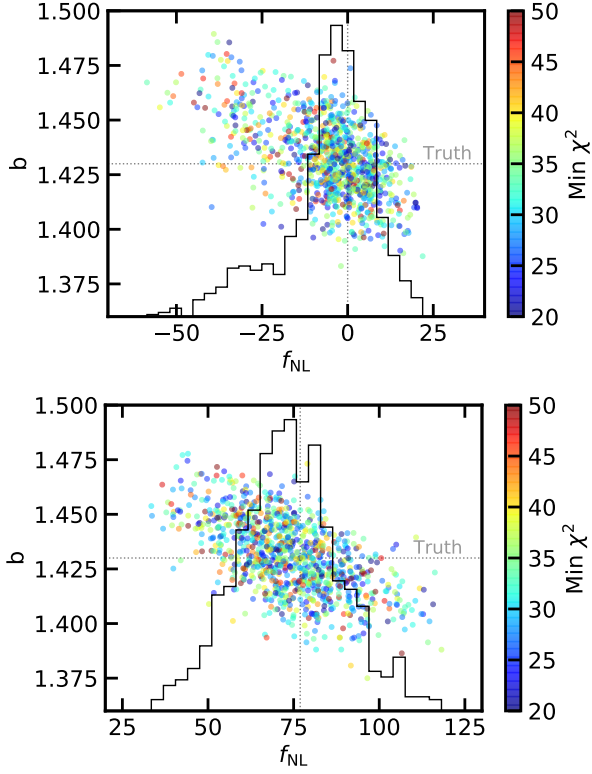
Figure 16. Top: 68% and 95% confidence contours for $f_{NL} = 0$ (left) and 76.92 (right) mocks. Using the $\log C_\ell$ fitting yield constraints that are insensitive to the covariance used. Bottom: best fit estimates from fitting 1000 lognormal mocks with $f_{NL} = 0$ (left) and 76.92 (right) in the DESI footprint. The truth values are represented by vertical and horizontal lines.

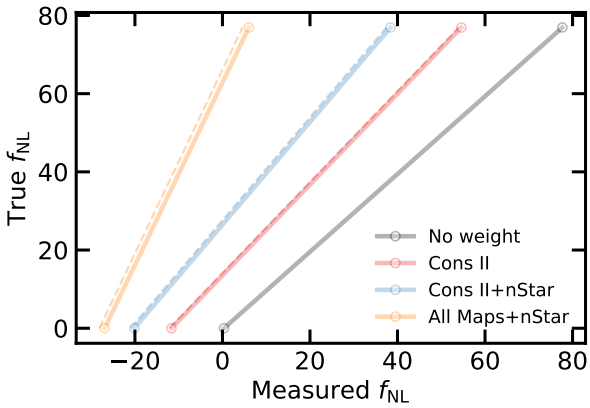**Figure 17.** True f_{NL} vs measured f_{NL} from the mean of the mocks with (dashed) and without systematics (solid).

ACKNOWLEDGEMENTS

DATA AVAILABILITY

DESI imaging DR9 catalogs are publicly available at https://www.legacysurvey.org/dr9/. Software for cleaning imaging data is available at https://github.com/mehdirezaie/sysnetdev. 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
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
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
Bouwens R. J., et al., 2015, ApJ, 803, 34
Castorina E., Moradinezhad Dizgah A., 2020, J. Cosmology Astropart. Phys., 2020, 007
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
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
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
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
Wang M. S., Beutler F., Bacon D., 2020, MNRAS, 499, 2598
Weaverdyck N., Huterer D., 2021, MNRAS, 503, 5061
Weinberg S., 2008, Cosmology. OUP Oxford
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

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