Optimal constraints on Primordial non-Gaussianity with the eBOSS DR16 quasars in Fourier space

Marina S. Cagliari,[a,1] **Emanuele Castorina,**[a,b] **Marco Bonici,**[c,d,e] **Davide Bianchi**[a]

[a]Dipartimento di Fisica 'Aldo Pontremoli', Università degli Studi di Milano,
  Via Celoria 16, 20133 Milan, Italy
[b]INFN, Sezione di Milano, Via Celoria 16, 20133 Milan, Italy
[c]Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, ON N2L 3G1, Canada
[d]Department of Physics and Astronomy, University of Waterloo, 200 University Ave W,
  Waterloo, ON N2L 3G1, Canada
[e]INAF-IASF Milano, Via Alfonso Corti 12, I-20133 Milano, Italy

  E-mail: marina.cagliari@unimi.it, emanuele.castorina@unimi.it,
  mbonici@uwaterloo.ca, davide.bianchi1@unimi.it

Abstract. We present constraints on the amplitude of local Primordial Non-Gaussianities (PNG), $f_{\rm NL}$, using the quasar sample in the Sloan Digital Sky Survey IV extended Baryon Oscillation Spectroscopic Survey Data Release 16. We analyze the power spectrum monopole, testing for the presence of scale dependent galaxy bias induced by local PNG. Our analysis makes use of optimal redshift weights that maximize the response of the quasar sample to the possible presence of non zero PNG. We find $-4 < f_{\rm NL} < 27$ at 68% confidence level, which is among the strongest bounds with Large Scale Structure data. The optimal analysis reduces the error bar by $\sim 10\%$ compared to the standard one, an amount which is smaller than what was expected from a Fisher matrix analysis. This, and the reduced improvement over previous releases of the same catalog, suggest the presence of still unknown systematic effects in the data. If the quasars have a lower response to local PNG, our optimal constraint becomes $-23 < f_{\rm NL} < 21$ at 68%, with an improvement of 30% over standard analyses. We also show how to use the optimal weights to put data-driven priors on the sample's response to local PNG.

---

[1]Corresponding author

Contents

1 Introduction and main results

The late time distribution of the Large Scale Structure (LSS) of the Universe is the result of the evolution, under gravitational interaction, of the set of primordial curvature perturbations. By measuring the n-point functions of a galaxy sample we have therefore the unique opportunity to test the statistical properties of the initial conditions of the Universe. Of particular relevance for LSS probes is the presence of possible Primordial Non-Gaussianities (PNG). The leading hypothesis for the dynamical generation of the primordial density fluctuations, Inflation (see [1] for a review), offers theoretical guidance to the most generic ways PNG could arise in cosmological correlators, also indicating that PNG are generically smaller than the dominant Gaussian term.

In this work we focus on the so-called local PNG, for which the primordial gravitational potential $\Phi_P(\mathbf{x})$ is a non-linear function of a Gaussian field φ, $\Phi_P = \varphi + f_{\rm NL}(\varphi^2 - \langle \varphi^2 \rangle)$. The amplitude of local PNG is parameterized by a single number $f_{\rm NL}$, and we immediately see that, if the primordial fluctuations are of $\mathcal{O}(10^{-5})$, local PNG are $\mathcal{O}(10^5)$ smaller than the Gaussian term for $f_{\rm NL} = 1$. Local PNG are among the most studied in the literature because they are exactly zero if the inflationary dynamics is driven by a single degree of freedom, the so-called single-field models [2–4]. A robust detection of $f_{\rm NL}$ will therefore exclude all such models and point to a more complicated inflationary sector. Conversely, multi-field models of inflation generically predict $f_{\rm NL} \sim \mathcal{O}(1)$ [5, 6], and could be severely constrained by a strong experimental bound. Measurements of the anisotropies of the Cosmic Microwave Background (CMB) from the Planck satellite put the stringent limit $f_{\rm NL} = 0.8 \pm 5$ [7], and upcoming instruments are expected to reduce this error bar by another 50% [8]. Differently than the CMB, which is sensitive to $f_{\rm NL}$ starting with the three-point function, LSS can probe PNG at

the two-point, or power spectrum in Fourier space, level. As first pointed out in ref. [9], the quadratic term in the definition of the primordial potential Φ_P induces a correlation between the long-wavelength gravitational field and the small scale fluctuations. The latter could very well be in the range corresponding to the formation of halos and galaxies, whose number density is therefore modulated by the large scale value of Φ_P. Mathematically, we say that the large scale bias of galaxies is modified in the presence of local PNG, and it reads

$$\delta_g(\mathbf{x}, z) = b(z)\,\delta_m(\mathbf{x}, z) + f_{\mathrm{NL}}\,b_\phi(z)\,\Phi_P(\mathbf{x}) \tag{1.1}$$

where δ_g is the galaxy density perturbation and b is the Gaussian linear bias. The new bias coefficient b_ϕ parameterizes the actual response of small scale fluctuations to the presence of local PNG, and it is subject to large theoretical uncertainties due to our incomplete knowledge of galaxy formation physics [10, 11].[1] Via Einstein's Equations, the presence of Φ_P in the above expression implies that, on large scales, the power spectrum acquires a distinct k^{-2} feature, which is then interpreted as the smoking gun of local PNG. In this respect, knowing the value of b_ϕ is not a fundamental limitation, since what ultimately matters to exclude single field models is a detection rather than the actual value of f_{NL}. The possibility to measure local PNG with the galaxy power spectrum has spurred a tremendous amount of research activities, and all major spectroscopic and photometric instruments like the Dark Energy Spectroscopic Instrument [DESI; 15], Euclid [16], SPHEREx [17] and the Vera C. Rubin Observatory [18], have the search for PNG as one of their primary science goals. It also serves as an important science case for future facilities [19]. Current LSS bounds are still far from the CMB one, $|f_{\mathrm{NL}}| \sim \mathcal{O}(20-30)$ [20–23], but are expected to improve down to $\sigma_{f_{\mathrm{NL}}} \sim 1$ with current and future observations [24–28].[2]

The main goal of this work is to provide the most stringent and robust constraints on local PNG with current data. We will use the extended Baryon Oscillation Spectroscopic Survey (eBOSS) Data Release 16 quasar (QSO) sample [DR16Q; 30, 31]. Our analysis takes advantage of optimal signal weighting that maximizes the response of a given galaxy sample to the presence of local PNG. These weights were first derived in ref. [20], and are based on optimal quadratic estimators [32–34]. The main reason to use optimal weights lies in eq. (1.1): the non-Gaussian contribution is proportional to the primordial potential, and therefore it does not evolve over time, while the linear bias term is proportional to the matter density, which grows over time. This suggests that, in a given sample, high-redshift objects should be given more weight than low-redshift ones, since the Gaussian piece is smaller at earlier times. As we will see in section 2 in more detail, the optimal analysis downweights the Gaussian signal by $w_0 \sim b(z)D(z)$, where $D(z)$ is the linear growth factor that decreases with increasing redshift, and upweights the PNG term by $\tilde{w} \sim b_\phi$.

The optimal redshift weighting therefore requires some prior knowledge of the response b_ϕ of a given sample to the presence of local PNG. For mass selected halos, analytical models [35, 36] and simulations [37, 38] suggest that $b_\phi \propto (b - p)$, where p is a number of $\mathcal{O}(1)$, is a very good approximation to the true response. As stated above, the picture is however much more complicated for observed galaxies. In this work we will present constraints on f_{NL} for the two values of p mostly used in the literature, $p = 1.0$ and $p = 1.6$. The fact that

---

[1]This point could however be turned the other way around and suggests that by carefully selecting the galaxy sample one could maximize the response, i.e. the value of b_ϕ, to improve the constraint [12–14].

[2]Recent work, ref. [29], presents evidence of non zero f_{NL} at more than 99% confidence level (c.l.) with DESI imaging data. However, CMB and LSS measure local PNG on the same range of scales, which suggests a non-cosmological origin for the signal reported in ref. [29].

the bound on local PNG depends on p could raise some concern about the robustness of our results. However, as we already pointed out, what matters is only a possible detection of f_{NL}, not what the actual value is. Moreover, galaxy selection based on luminosity, color, or magnitude, correlates reasonably well with host halo mass [39, 40], therefore we do not expect a large deviation from the values used in this work. Nevertheless, we will also show, in a novel application of our framework, how optimal signal weights can be used to put a data-driven prior on the value of b_ϕ or p. As mentioned above, the optimal weights $\tilde{w}$ are proportional to the response of the galaxy number density to the presence of local PNG. This implies that if the input value of b_ϕ we use for the weighting is very different from the true response, the optimal analysis will not improve the bound on f_{NL} over the un-weighted case, or will even worsen the constraints. As a first application of this idea, we will show in section 4 that a large value of $p \gtrsim 3$ for the response of the QSOs in DR16 is not favored by the data, without relying on any numerical simulations. We expect that as the uncertainty on f_{NL} reduces in the near future, our method will provide invaluable information on the most likely value of p to use in the data analysis. It should, however, be kept in mind that assuming a value of p implies that different data sets cannot be combined together or with the CMB. For this reason we will also show, in appendix A, constraints for $b_\phi f_{\mathrm{NL}}$ for all the data sets used in this work.

Our strongest bounds read

$$\begin{cases} -4 < f_{\mathrm{NL}} < 27\,, & 68\%\,\mathrm{c.l.}\,, \\ -18 < f_{\mathrm{NL}} < 42\,, & 95\%\,\mathrm{c.l.}\,, \end{cases} \quad \text{for } p = 1.0\,, \tag{1.2}$$

and

$$\begin{cases} -23 < f_{\mathrm{NL}} < 21\,, & 68\%\,\mathrm{c.l.}\,, \\ -43 < f_{\mathrm{NL}} < 44\,, & 95\%\,\mathrm{c.l.} \end{cases} \quad \text{for } p = 1.6\,, \tag{1.3}$$

which should be compared with a standard Feldman-Kaiser-Peacock [FKP; 41] analysis, see figure 8 and table 2 for the full results. The optimal analysis improves by 10% and 30% over the FKP one for the $p = 1.0$ and $p = 1.6$ cases respectively. It is worth stressing that the power spectrum of DR16Q catalog is, on all scales, dominated by the shot-noise, and we therefore did not expect much larger gains [20]. Our optimal constraints are robust to the treatment of systematic effects. The bounds using a linear method to remove known foregrounds are statistically indistinguishable from the ones obtained with a non-linear algorithm based on Neural-Network [NN; 42]. However, compared to the previous eBOSS data release [20], the improvement in the constraint is smaller than what was expected from the increase in volume, and it is most likely due to the presence of residual foregrounds in the maps. This could also be the reason for the more limited improvement of the optimal analysis with respect to the standard one in comparison to the improvement found in DR14 [20].

Our results are roughly comparable with the ones in ref. [21], which also used the DR16Q data set. However, there are a number of important differences with our analysis. First, the weights employed in ref. [21] are defined for pairs of galaxies, and cannot be automatically applied to individual galaxies, as relevant for a power spectrum analysis. To avoid imaginary weights for a single object, the Authors of [21] imposed, by hand, the positivity of the weights, which is not by itself an optimal procedure. In our case, the weights can very well be negative, precisely in the region where the signal is: in this way the product of the weights times the signal contributes positively to the total signal-to-noise. More generally, cosmological information is contained in the galaxy fields and not in its non-linear transformations, like for
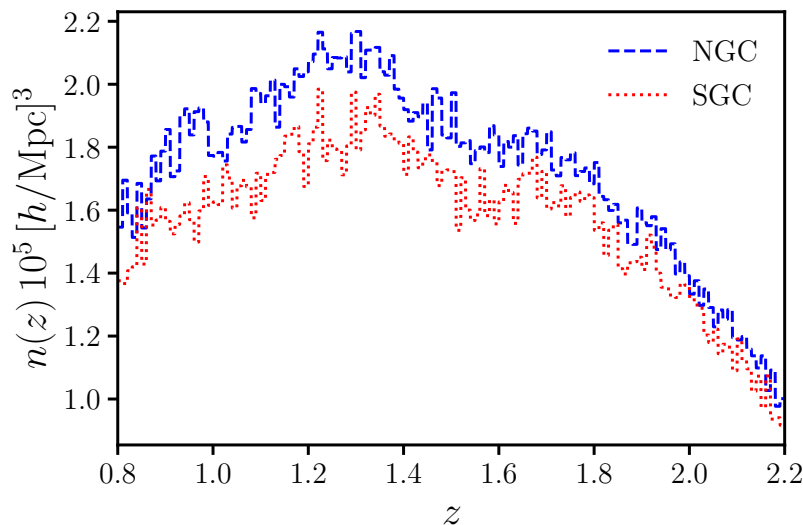
Figure 1. The quasar number density as a function of redshift for the DR16Q sample. The dashed blue line corresponds to the NGC and the dotted red line to the SGC. The SGC has a lower density in comparison to the NGC due to the difference in mean depth in the two regions [30].

example pairs of galaxies. The other main difference with the work of [21] is in the modeling of the signal. As discussed in section 2.4, we think the bound of ref. [21] is artificially tighter due to an incorrect choice for the effective redshift, $z_{\rm eff}$, at which the theoretical model is evaluated. For most applications the precise definition of $z_{\rm eff}$ does not matter, but it becomes important in searches for local PNG, where the signal is proportional to $b_\phi(z) \sim b(z) - p$. For the DR16Q analysis in ref. [21], a too high value of $z_{\rm eff}$ results in a higher linear bias $b(z)$, which artificially reduces the uncertainty on $f_{\rm NL}$ to keep the product $b_\phi f_{\rm NL} \sim$ constant. In section 2.4 we will also clarify on this issue and on what the more accurate definition of $z_{\rm eff}$ is.

The rest of this paper is organized as follows: in section 2 we present the DR16Q data set and the measurements of the power spectrum; in section 3 we discuss the modeling of the QSO power spectrum, its convolution with the window function and the definition of the effective redshift; in section 4 we present and discuss the constraints on $f_{\rm NL}$; section 5 concludes and summarizes our results.

All the codes, scripts, measurements, and MonteCarlo Markov chains used in this work are freely accessible at `https://github.com/mcagliari/eBOSS-DR16-QSO-OQE`.

2 Data

2.1 The eBOSS QSO Sample

In this work we use the eBOSS DR16Q sample [30, 31]. As part of the Sloan Digital Sky Survey IV (SDSS-IV) experiment [43], the eBOSS data were acquired at the Apache Point Observatory in New Mexico.

The DR16Q sample contains 343 708 quasars in the redshift range $0.8 < z < 2.2$. The sample is divided into two fields of view, the North Galactic Cap (NGC), which covers an area of $2\,924\,{\rm deg}^2$, and the South Galactic Cap (SGC), with an area of $1\,884\,{\rm deg}^2$; in comparison to Data Release 14 (DR14) the area is approximately doubled. The whole sample has a volume

of $\sim 20\,(\mathrm{Gpc}/h)^3$. The NGC has a mean density of $n \approx 1.8 \times 10^{-5}\,(\mathrm{Mpc}/h)^{-3}$, while the SGC has a slightly lower density of $n \approx 1.6 \times 10^{-5}\,(\mathrm{Mpc}/h)^{-3}$. The number densities as a function of the redshift of the NGC and the SGC quasars are shown in figure 1. The number density of SGC is about 10% lower than the NGC number density because of the lower mean depth of the survey in the SGC region. The data of the North and South Galactic Cap were released in two separate catalogs, each one with the corresponding random catalog. The random catalogs are 50 times more dense than the data catalogs, and their redshift distributions are produced by sampling from the observed data redshifts [30], a procedure known as shuffling. The use of the shuffling scheme to produce the random catalogs introduces a systematic effect called radial integral constraint [RIC; 44]. We discuss how to estimate and correct for the RIC effect in section 3.2.

Both the data and random catalogs contain three weights for each data point. First, there are the close pairs weights, w_{cp}, which take into account fiber collisions. The second weights are related to the spectroscopic completeness, w_{noz}, and correct for the expected redshift failure rate. Third, there are the imaging systematic weights, w_{sys}. These weights correct for systematic effects at large angular scales, and they are therefore especially important for f_{NL} measurements. In the official data release of eBOSS DR16 [30], these weights are computed with linear regression of the imaging properties and Galactic foregrounds. Additional catalogs were released for the quasars [42]. In these catalogs, the imaging systematic weights were computed using neural networks. Neural networks are able to approximate non-linear functions, hence they could in principle produce a better correction than the weights computed with the linear regression. NN methods will however remove part of the signal as well, and could bias negative the constraint on local PNG. Hereafter we will refer to the official DR16Q catalogs as the linear weight catalog, and to the catalogs with NN systematic weights as NN weight catalog. The completeness weight contribution to any data point is [30]

$$w_{\mathrm{c}} = w_{\mathrm{cp}}\,w_{\mathrm{noz}}\,w_{\mathrm{sys}}\,, \tag{2.1}$$

where w_{sys} can either be from the linear weight catalog or the NN weight catalog. The same weighting procedure of eq. (2.1) applies to the objects in the random catalogs.

2.2 Mocks

A set of 1000 synthetic clustering catalogs for each Galactic cap [45] was simulated using the effective Zel'dovich approximation mock method [EZmock; 46]. The EZmock catalogs were produced assuming a flat ΛCDM cosmology with $\Omega_m = 0.307115$, $\Omega_\Lambda = 0.62885$, $\Omega_b = 0.048206$, $h = 0.6777$, $\sigma_8 = 0.8225$, $n_s = 0.9611$, and $f_{\mathrm{NL}} = 0$. The mocks reproduce the two and three-point clustering statistics of DR16Q.

In the official release of the eBOSS DR16 EZmock catalogs, three sets of mock catalogs are provided. Each set of EZmocks consists of 1000 pairs of data and random catalogs. We refer to the first set of EZmocks as EZmock *realistic*, since the data and random catalogs of this set contain all the known observational systematic effects. Each data catalog has a corresponding random catalog, whose redshift distribution is produced by shuffling the redshift of the data catalog. The 1000 data-random pairs of the EZmock realistic set are used to estimate the covariance matrix, $\boldsymbol{\Sigma}$. The imaging systematic weights of the realistic EZmock are computed only with the linear regression method and not with the neural network. The covariance matrix will thus only contain information about the linear imaging weights. The other two sets are the EZmocks *complete* and the EZmocks *shuffled*. These two sets share the same data catalogs, which do not have any observational systematic effect, while the random
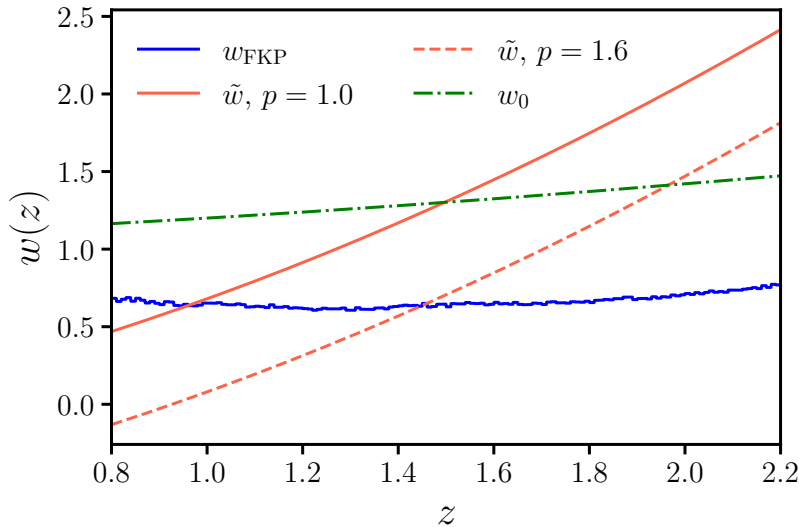
Figure 2. Weights as a function of redshift for the eBOSS DR16Q. The solid blue line corresponds to the FKP weights, eq. (2.3); its almost flat behavior as a function of redshift is due to $n(z)\,P_{\rm fid} \ll 1$. The optimal weights to estimate $f_{\rm NL}$ are shown in red and dot-dashed green. The solid and dashed red lines respectively correspond to $\tilde{w}(z)$ for $p = 1.0$ and $p = 1.6$. These weights have a clear dependence on the quasar response to the $f_{\rm NL}$ signal.

catalog redshift distributions were produced in different ways. In the EZmock shuffled set each data catalog has a corresponding random catalog the redshift distribution of which is produced with the shuffling scheme. The EZmock complete set only has two random catalogs, one for each Galactic cap. The redshift distributions of these random catalogs were sampled from the same $n(z)$ interpolation used for the EZmock data catalogs. We use the EZmock complete and shuffled sets to estimate the RIC effect.

2.3 Power spectrum estimation

As shown in [20], the optimal power spectrum is the cross-correlation of two different fields produced by weighting the underlying catalog. To arrive at the power spectrum estimator we start from the two quasar density fields [41],

$$\tilde{F}(\mathbf{r}) = \tilde{w}_{\rm tot}\,[w_{\rm c}^{\rm qso}\,n_{\rm qso}(\mathbf{r}) - \alpha_{\rm s}w_{\rm c}^{\rm s}\,n_{\rm s}(\mathbf{r})]\,, \quad F_0(\mathbf{r}) = w_{\rm tot,\,0}\,[w_{\rm c}^{\rm qso}\,n_{\rm qso}(\mathbf{r}) - \alpha_{\rm s}w_{\rm c}^{\rm s}\,n_{\rm s}(\mathbf{r})]\,, \quad (2.2)$$

where $n_{\rm qso}$ and $n_{\rm s}$ respectively are the number density of the quasar sample and the corresponding random catalog, $w_{\rm c}^{\rm qso}$ and $w_{\rm c}^{\rm s}$ are the completeness weights from eq. (2.1) for the quasars and the randoms. The total weights, $\tilde{w}_{\rm tot}$ and $w_{\rm tot,\,0}$, are the product of the FKP weights [41],

$$w_{\rm FKP}(z) = \frac{1}{1 + \bar{n}(z)\,P_{\rm fid}}\,, \quad (2.3)$$

and the optimal weights for a $f_{\rm NL}$ measurements with power spectrum data [20],

$$\tilde{w}(z) = b(z) - p\,, \quad w_0(z) = D(z)\left(b(z) + \frac{f(z)}{3}\right)\,. \quad (2.4)$$

Therefore the total weights read [20]

$$\tilde{w}_{\rm tot}(z) = w_{\rm FKP}(z)\,\tilde{w}(z)\,, \quad w_{\rm tot,\,0}(z) = w_{\rm FKP}(z)\,w_0(z)\,. \quad (2.5)$$

In eq. (2.3), $\bar{n}(z)$ is the mean density as a function of redshift, and $P_{\mathrm{fid}} = 3 \times 10^4 \, (\mathrm{Mpc}/h)^3$, which corresponds to the expected power on the scales affected by PNG in the sample. In eq. (2.4), $b(z)$ is the fiducial value of the QSO bias model [47],

$$b(z) = 0.278 \left((1+z)^2 - 6.565 \right) + 2.393 \,, \tag{2.6}$$

and $D(z)$ and $f(z)$ are respectively the growth factor and growth rate as functions of redshift. As mentioned in section 1, in this work we use $p = 1.0$ and $p = 1.6$. The dependence on the redshift of the weights defined in eq. (2.3) and eq. (2.4) is plotted in figure 2. It shows the difference between the FKP weighting scheme, which is almost constant in redshift, as $n(z)P_{\mathrm{fid}} \ll 1$, and the optimal weights, which have a strong dependence on redshift. Finally, the factor α_{s} in eq. (2.3) is defined as

$$\alpha_{\mathrm{s}} = \frac{\sum^{\mathrm{qso}} w_{\mathrm{c}}}{\sum^{\mathrm{s}} w_{\mathrm{c}}} \,, \tag{2.7}$$

and it properly normalizes the number density of the random catalog.

Following ref. [48], we write the monopole of the cross-correlation between the two weighted fields in eq. (2.2) as

$$\widetilde{P}_0(k) = A_0^{-1} \int \frac{\mathrm{d}\Omega_k}{4\pi} \left[\int \mathrm{d}\mathbf{r}_1 \, \tilde{F}(\mathbf{r}_1) \, e^{i\mathbf{k}\cdot\mathbf{r}_1} \int \mathrm{d}\mathbf{r}_2 \, F_0(\mathbf{r}_2) \, e^{-i\mathbf{k}\cdot\mathbf{r}_2} \mathcal{L}_0(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}_2) \right] - S_0 \,, \tag{2.8}$$

where $\mathcal{L}_0$ is the first Legendre polynomial. The normalization factor A_0 and the shot noise contribution, S_0, are respectively defined as

$$A_0 = \int \mathrm{d}\mathbf{r} \, w_{\mathrm{tot},0}(\mathbf{r}) \, \tilde{w}(\mathbf{r}) \, [w_{\mathrm{c}} \, n_{\mathrm{qso}}(\mathbf{r})]^2 \,, \tag{2.9}$$

$$S_0 = A_0^{-1} \int \mathrm{d}\mathbf{r} \, w_{\mathrm{c}} \, n_{\mathrm{qso}}(\mathbf{r}) \, (w_{\mathrm{c}}(\mathbf{r}) + \alpha_{\mathrm{s}}) \, w_{\mathrm{tot},0}(\mathbf{r}) \, \tilde{w}(\mathbf{r}) \, \mathcal{L}_0(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}}) \,. \tag{2.10}$$

We calculate the monopole of the power spectrum using `nbodykit` [49], which implements eq. (2.8) as follows [50, 51],

$$\widetilde{P}_0(k) = A_0^{-1} \int \frac{\mathrm{d}\Omega_k}{4\pi} \tilde{F}(\mathbf{k}) \, F_0(-\mathbf{k}) \,, \tag{2.11}$$

with

$$F_0(\mathbf{k}) = \int \mathrm{d}\mathbf{r} \, F_0(\mathbf{r}) \, e^{i\mathbf{k}\cdot\mathbf{r}} \mathcal{L}_0(\hat{\mathbf{k}}\cdot\hat{\mathbf{r}})$$

$$= 4\pi \, Y_{00}(\hat{\mathbf{k}}) \int \mathrm{d}\mathbf{r} \, F_0(\mathbf{r}) \, Y_{00}^*(\hat{\mathbf{r}}) \, e^{i\mathbf{k}\cdot\mathbf{r}} \,, \tag{2.12}$$

where Y_{00} is the first spherical harmonic. The normalization and the shot noise are computed as discrete sums over the quasars and the randoms. The normalization is

$$A_0 = \alpha_{\mathrm{s}} \sum_i^{N_{\mathrm{s}}} n_{\mathrm{s}}(\mathbf{r}_i) \, w_{\mathrm{c}}(\mathbf{r}_i) \, w_{\mathrm{tot},0}(\mathbf{r}_i) \, \tilde{w}_{\mathrm{tot}}(\mathbf{r}_i) \,. \tag{2.13}$$

The shot noise contribution becomes

$$S_0 = A_0^{-1} \left[\sum_i^{N_{\rm qso}} w_c^2(\mathbf{r}_i) \, w_{\rm tot,0}(\mathbf{r}_i) \, \tilde{w}_{\rm tot}(\mathbf{r}_i) + \alpha_s^2 \sum_i^{N_{\rm s}} w_c(\mathbf{r}_i) \, w_{\rm tot,0}(\mathbf{r}_i) \, \tilde{w}_{\rm tot}(\mathbf{r}_i) \right] . \qquad (2.14)$$

To compute the power spectrum estimator in eq. (2.11) we use a mesh of 512^3 cells. The quasars and the random objects are projected onto the mesh using a triangular shaped cloud interpolation [52]. In this interpolation each quasar and random is weighted by both its completeness and total weight and we assume Planck [53] as fiducial cosmology. The power spectrum is estimated on a logarithmic grid from $k_{\rm min} = 3.75 \times 10^{-3} \, ({\rm Mpc}/h)^{-1}$ to $k_{\rm max} = 2.23 \times 10^{-1} \, ({\rm Mpc}/h)^{-1}$ for NGC, and $k_{\rm max} = 2.78 \times 10^{-1} \, ({\rm Mpc}/h)^{-1}$ for SGC.

A final remark about the normalization of the power spectrum estimator in eq. (2.11) and the shot noise contribution in eq. (2.14). Since eq. (2.13) is just an approximation of the exact definition in eq. (2.9), it has been pointed out that this could lead to biased constraints on cosmological parameters [54]. We, therefore, decided to re-normalize the measured power spectra by the limit, at small separation, of the monopole of the window function, $Q_0(0)$, see section 2.4. The final estimator of the power spectrum is

$$\hat{P}_0(k) = \frac{A_0}{Q_0(0)} \left(\widetilde{P}_0(k) - S_0 \right) . \qquad (2.15)$$

We note that we used, both for the linear and NN catalog, the value of $Q_0(0)$ computed from the randoms of the linear catalog, since the two catalogs are expected to have the same response at small scales.

In figure 3 we show the NGC and SGC (left and right columns) observed power spectra of the linear and NN catalog (dots and squares). The three rows correspond to the different weights used to estimate the $\hat{P}_0(k)$, from top to bottom the rows correspond to the FKP weights, the optimal weight with $p = 1.0$, and with $p = 1.6$. We also plotted the best fit model for the linear (dashed line) and NN (dash-dotted line) catalogs (see section 4). The best fit of the FKP weight case is for the model with $p = 1.6$. For the NN catalog power spectra, the band powers have been shifted by 5% along the k-axis for better visualization of the corresponding error bars. The observed power spectra estimated from the two catalogs differ only in the first two bins. The NN catalog power spectra have less power in the first bin than their linear counterparts. The excess of power in the linear catalog power spectra is expected to be related to large scale systematic effects that the linear regression weights are not able to correct. The solid horizontal line in each panel shows the amplitude of the power spectrum at its peak, and it serves to guide the eye to the fact that the optimal weighted measurements are larger than the corresponding FKP ones. This was expected because the optimal weights upweight high redshift galaxies (see figure 2), which have a higher bias and therefore a larger clustering amplitude.

2.4 Window functions

The window function, $W(\mathbf{s})$, represents the footprint on the sky and the redshift selection function of the survey. It is an essential ingredient to compare a power spectrum model with the observed power spectrum.

In order to evaluate the model of the observed power spectrum we need the multipoles of the window function (see section 3.2), defined as follows,

$$Q_\ell(s) \equiv (2\ell + 1) \int {\rm d}\Omega_s \int {\rm d}^3\mathbf{s}_1 \, W(\mathbf{s}_1) \, W(\mathbf{s} + \mathbf{s}_1) \, \mathcal{L}_\ell(\hat{\mathbf{s}}_1 \cdot \hat{\mathbf{s}}) \equiv \int {\rm d}s_1 \, s_1^2 \, Q_\ell(s; s_1) . \qquad (2.16)$$
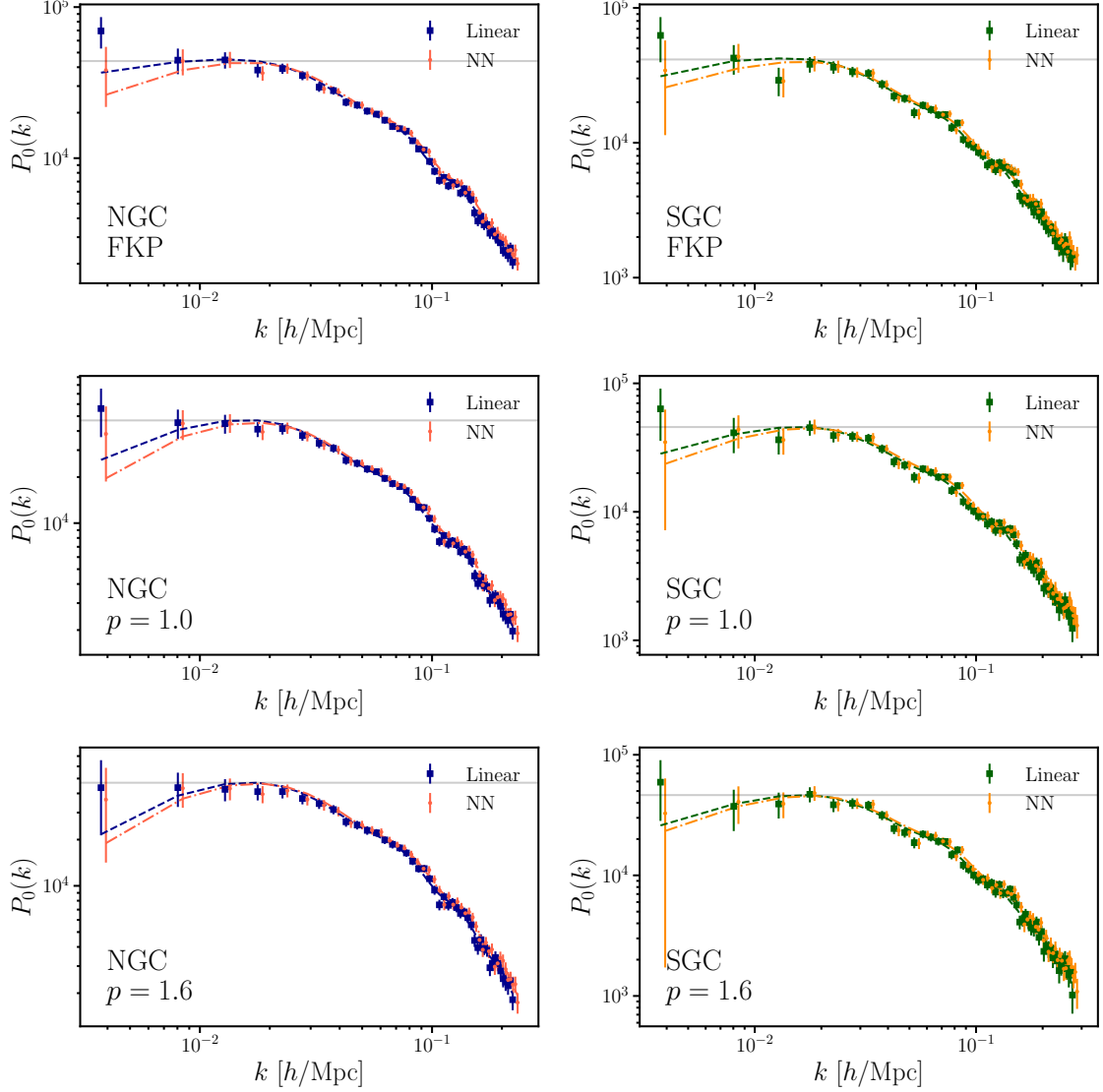
Figure 3. Observed power spectra and best fit model for the NGC (right column), and SGC (left column), and for the three weighting scheme: FKP (top row), optimal weights with $p = 1.0$ (middle row), and with $p = 1.6$ (bottom row). For the NGC (SGC) blue (green) dots and the dashed line correspond to the linear catalog results, red (yellow) squares, and dash-dotted line to the NN catalog. The horizontal gray line marks the turn-around point in the power spectrum. To better visualize the power spectra error bars, the NN $P(k)$ was shifted by 5%.

To compute the window function multipoles we use the pair counting approach introduced in ref. [55]. First, with `nbodykit` we calculate the weighted pair counts of the random catalogs as a function of the three-dimensional separation and the cosine of the line-of-sight angle, $RR^{\mathrm{w}}(s, \mu)$. That is done by cross-correlating the random catalog weighted by $w_{\mathrm{c}} \, \tilde{w}_{\mathrm{tot}}$, and the random catalog weighted by $w_{\mathrm{c}} \, w_{\mathrm{tot}, 0}$. Second, we compute its multipoles,

$$RR_{\ell}^{\mathrm{w}}(s) = (2\ell + 1) \int \mathrm{d}\mu \, RR^{\mathrm{w}}(s, \mu) \, \mathcal{L}_{\ell}(\mu) \,. \tag{2.17}$$

Finally, in order to obtain the window function multipoles the quantity above needs to be normalized to take into account the width of the shell over which the pair counting is performed and the density of the random catalog in comparison to the data catalog. The window function multipole ℓ is finally defined as

$$Q_{\ell}(s) = \frac{RR_{\ell}^{\mathrm{w}}(s)}{4\pi \, s^3 \, \mathrm{d} \ln s} \frac{\left(\sum^{\mathrm{qso}} w_{\mathrm{c}}\right)^2 - \sum^{\mathrm{qso}} w_{\mathrm{c}}^2}{\left(\sum^{\mathrm{s}} w_{\mathrm{c}}\right)^2 - \sum^{\mathrm{s}} w_{\mathrm{c}}^2} \,, \tag{2.18}$$

with $\mathrm{d} \ln s = \frac{s_{n+1} - s_n}{s}$, where s is the center of the n-th separation bin. We stress that each set of optimal weights requires its own multipoles of the window function $Q_{\ell}(s)$.

To reduce the computational time of the pair counting algorithm we divided the random catalogs into five subsets and computed $RR^{\mathrm{w}}(s, \mu)$ for each of them. These subsets are 10 times denser than the data catalog. We calculated the window function multipoles of each random subset using eq. (2.18), with the caveat that the sum over the random $\sum^{\mathrm{s}}$ is now over the subset. The final $Q_{\ell}(s)$ is the mean of the five subsets.

3 Analysis Methods

3.1 The Power spectrum Model

To model the quasar power spectrum in redshift-space we use linear theory. Linear theory is enough to make the prediction for two reasons: first, the local f_{NL} signal is at low k, where structures are still growing with a linear regime. Second, the smaller scales of this sample are dominated by redshift errors, which dominate over the non-linearities. We write the power spectrum model as follows [56],

$$P_{\mathrm{qso}}(k, \mu; z) = G(k, \mu; \sigma_{\mathrm{FoG}})^2 \left[b_{\mathrm{tot}}(k; z) + f(z) \, \mu^2\right]^2 \, P_m(k; z) + N \,, \tag{3.1}$$

where P_m is the matter power spectrum in real-space, N is the residual shot noise free parameter, and $f(z)$ is the growth rate. The total quasar bias includes the PNG [9, 35],

$$b_{\mathrm{tot}}(k; z) = b_1 + \Delta b = b_1 + f_{\mathrm{NL}} \, (b_1 - p) \, \tilde{\alpha}(k; z) \,, \tag{3.2}$$

where b_1 is the quasar linear bias, and $\tilde{\alpha}(k; z)$ is

$$\tilde{\alpha}(k; z) = \frac{3 \, \Omega_m \, H_0^2 \, \delta_{\mathrm{c}}}{c^2 \, k^2 \, T(k) \, D(z)} \,. \tag{3.3}$$

In eq. (3.3), $\delta_{\mathrm{c}} = 1.686$ is the critical density in the spherical collapse in an Einstein-De Sitter Universe, Ω_m is the matter density parameter, and H_0 the Hubble parameter, both at $z = 0$, and c is the speed of light. Then, $T(k)$ is the matter transfer function normalized to 1 at

low-k, and $D(z)$ is the growth factor normalized to $(1+z)^{-1}$ in the matter dominated era. Finally, the damping of the power spectrum due to nonlinear redshift-space distortions is included with a Lorentzian function,

$$G(k, \mu; \sigma_{\text{FoG}}) = \left[1 + \frac{(k\,\mu\,\sigma_{\text{FoG}})^2}{2}\right]^{-1}, \tag{3.4}$$

where σ_{FoG} accounts for both the typical velocity dispersion of QSOs, as well as their redshift error, which is estimated to be $\sigma_z = 300\,\text{km s}^{-1}$ for DR16Q, with no significant dependence on the redshift [31].

The quasar power spectrum multipoles are then easily computed

$$P_{\ell,\text{qso}}(k; z) = \frac{2\ell+1}{2} \int_{-1}^{1} \mathrm{d}\mu\, P_{\text{qso}}(k, \mu; z)\, \mathcal{L}_\ell(\mu). \tag{3.5}$$

To evaluate the cosmological quantities in eq. (3.1) and eq. (3.3) we assume a Planck fiducial cosmology [53], and fix the redshift to an effective value. In the following section, we discuss how the effective redshift is defined and computed. We calculate the cosmological functions with `classy`, the Python wrapper of the `CLASS` CMB Boltzmann solver [57].

3.2 Convolution with the window function and the effective redshift

The ensemble average of the power spectrum estimator in eq. (2.8) is [55, 58]

$$\langle \hat{P}_0(k) \rangle = \sum_{\ell, L} \begin{pmatrix} \ell & L & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 \int \mathrm{d}s\, s^2 j_0(ks) \int \mathrm{d}s_1\, s_1^2\, \xi_\ell(s; s_1(z))\, Q_L(s; s_1(z)), \tag{3.6}$$

where $Q_L(s; s_1)$ is defined in eq. (2.16), $\xi_\ell(s; s_1)$ is the multipole ℓ of the QSO correlation function, $j_A(ks)$ is the spherical Bessel function of order A, and $\begin{pmatrix} \ell & L & 0 \\ 0 & 0 & 0 \end{pmatrix}$ is a Wigner 3-j symbol.

In eq. (3.6) the redshift evolution of the signal is taken into account by $s_1(z)$ in $\xi_\ell(s; s_1)$, which should be integrated against the $Q_L(s; s_1)$ for a proper model comparison. This could be a time consuming step if repeated for every point in the parameter space exploration, and it is therefore often approximated. Noticing that in most applications the correlation function is factorizable in time and space, $\xi(s, z) \sim g(s)\, h(z)$, a possible way to speed up the computation of the theoretical model is to separate the integrals over s and s_1. An even more useful approximation is to assume that the model can be evaluated only at some effective redshift, z_{eff}, defined by the radial selection function. The expression for the power spectrum then simplifies to

$$P_0(k; z_{\text{eff}}) = \sum_{\ell, L} \begin{pmatrix} \ell & L & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 \int \mathrm{d}s\, s^2\, j_0(ks)\, \xi_\ell(s; z_{\text{eff}})\, Q_L(s), \tag{3.7}$$

with the multipole of the correlation function of the random catalog $Q_L(s)$ defined in eq. (2.16). The final integral is a simple Hankel Transform that can be computed quite efficiently [55].

The question becomes, then, what is the most accurate definition of z_{eff}. The estimator of the power spectrum in eq. (2.8), and therefore the multiples of the window functions as well, contains two powers of the radial selection function, which suggests the following definition of z_{eff} for a sample with given $n(z)$ and weights $w(z)$,

$$z_{\text{eff}} = \frac{\int \mathrm{d}z\, n(z)^2\, [\chi(z)^2/H(z)]\, w(z)^2\, z}{\int \mathrm{d}z\, n(z)^2\, [\chi(z)^2/H(z)]\, w(z)^2}, \tag{3.8}$$
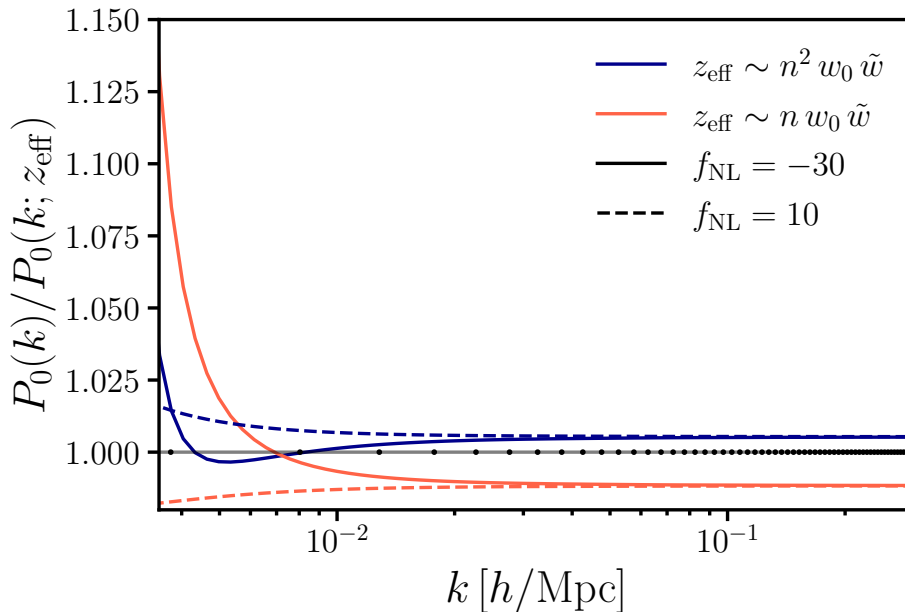
Figure 4. The accuracy of the effective redshift approximation. In red, and for two different values of f_{NL}, the plot shows the ratio between the theoretical model integrated over the full radial selection function of the DR16Q sample and the same model evaluated at a z_{eff} defined in eq. (3.9), see table 1. The accuracy of the other most common definition of z_{eff} in the literature is shown with blue lines. The black points correspond to the effective wavenumbers of the measurements of the QSO power spectrum.

	FKP	$p = 1.0$	$p = 1.6$
NGC	1.49	1.65	1.76
SGC	1.50	1.66	1.76

Table 1. The effective redshift for the different weights and the two sky regions. The optimal weights increase the z_{eff} of the sample for $p = 1.0$ and $p = 1.6$.

where $\chi(z)$ is the comoving distance and $H(z)$ is the Hubble parameter. In practice, the integral above can be estimated via Monte-Carlo methods as

$$z_{\mathrm{eff}} = \frac{\sum^{\mathrm{qso}} z\, n(z)\, w_{\mathrm{c}}^2\, w_{\mathrm{FKP}}(z)^2\, \tilde{w}(z)\, w_0(z)}{\sum^{\mathrm{qso}} n(z)\, w_{\mathrm{c}}^2\, w_{\mathrm{FKP}}(z)^2\, \tilde{w}(z)\, w_0(z)}, \tag{3.9}$$

or with the analogous expression written in terms of the random catalog. The values of z_{eff} for the samples used in this work are shown in table 1. We see that the optimal weighting scheme increases z_{eff}, since it up weights high redshift objects, which have a higher response to the presence of PNG.

The accuracy of our definition of z_{eff} is presented in figure 4, with the blue lines showing the ratio between the power spectrum model fully integrated over redshift and the monopole evaluated at z_{eff}. The dashed line corresponds to $f_{\mathrm{NL}} = 10$, while the continuous one to $f_{\mathrm{NL}} = -30$, with $p = 1.6$ in both cases. The black points on the horizontal axis show the effective values of the wavenumbers of the measurements. We find that our approximation

is sub-percent accurate at high-k, and better than 2.5% accurate on very large scales, thus much smaller than the sample variance of the measurements.

On the other hand, the DR16Q analysis of ref. [21] adopts a definition of $z_{\rm eff}$ with one less power of $n(z)$ than eq. (3.9).[3] This choice produces the red set of curves in figure 4, which we find are more than a percent off at high-k, a number that could become significant over many data points, and more than 10% inaccurate at large scales. In particular, ref. [21] reports $z_{\rm eff} = 1.83$ for the weights optimized with $p = 1.6$, a number significantly higher than our $z_{\rm eff} = 1.76$. At fixed value of $f_{\rm NL}$, the product $b_\phi f_{\rm NL}$ is 10% larger at $z = 1.83$ than at $z = 1.76$. This suggests that the authors of [21] would have at least gotten a 10% weaker constraint on $f_{\rm NL}$, had they used the more accurate definition of $z_{\rm eff}$ in eq. (3.8).

Finally, we write the convolution of the window function in the more convenient form

$$P_0(k; z_{\rm eff}) = \sum_{\ell,\, L} i^\ell \begin{pmatrix} \ell & L & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 \int \frac{\mathrm{d}q}{2\pi^2}\, q^2\, P_{\ell,\,{\rm qso}}(q; z_{\rm eff}) \int \mathrm{d}s\, s^2 j_0(ks)\, j_\ell(qs)\, Q_L(s) \qquad (3.10)$$

$$= \sum_{\ell,\, L} i^\ell \begin{pmatrix} \ell & L & 0 \\ 0 & 0 & 0 \end{pmatrix}^2 \int \frac{\mathrm{d}q}{2\pi^2}\, q^2\, P_{\ell,\,{\rm qso}}(q; z_{\rm eff})\, Q_{\ell,\, L}(k, q)\,, \qquad (3.11)$$

where we defined

$$Q_{\ell,\, L}(k, q) = \int \mathrm{d}s\, s^2 j_0(ks)\, j_\ell(qs)\, Q_L(s)\,. \qquad (3.12)$$

The integral in eq. (3.11) is then evaluated as a simple matrix multiplication. This choice allows us to never compute the correlation function multipoles, which are formally divergent in the presence of local PNG.

The multipoles of the window function corresponding to the optimal weights with $p = 1.0$ are shown in figure 5, left panel. In our model we use only the even multipoles up to $\ell = 4$, and neglect possible odd ones [58].[4]

3.3 The Integral Constraint

The final step to model the observed power spectrum is to correct the convolved power spectrum with the integral constraint effects. Integral constraint effects arise when the survey selection function is estimated from the data themselves [44]. Fluctuations over the whole survey average to zero when the observed galaxy mean density is used as the true cosmological mean. This causes a suppression of power at large scales, which we call Global Integral Constraint [GIC; 55, 62]. On the other hand, an additional radial integral constraint is produced when the radial $n(z)$ is inferred from the data [44]. That is the case for eBOSS data, where the random catalog redshift distribution is obtained by shuffling the data redshift distribution. The RIC causes a suppression of the large scale fluctuations along the line-of-sight.

The global integral constraint depends on the Hankel transform, $|\tilde{W}_\ell(k)|^2$, of the window function multipole $Q_\ell(s)$. The Hankel transforms are normalized so that $|\tilde{W}_0(0)|^2 = 1$ [55]. To correct for the RIC we need to estimate the effect that the shuffling of the random produces on the measured power spectrum. To do so we used both the complete and shuffled EZmocks.

---

[3]The published version of ref. [21] contains the following definition of the effective redshift, right below their eq. (11), $z_{\rm eff} = \sum_i z_i\, w_{\rm tot} / \sum_i w_{\rm tot}$, where $w_{\rm tot}^2 = w_{\rm FKP}^2\, w_{\rm c}^2\, |\tilde{w}\, w_0|$. This is a typo, as the analysis of ref. [21] actually used $z_{\rm eff} = \sum_i z_i\, w_{\rm tot}^2 / \sum_i w_{\rm tot}^2$, corresponding to the red set of curves in fig. 4. We thank Eva-Maria Mueller for the correspondence about this point.

[4]Wide angle effects and other projection effects are negligible for the DR16Q volume [58–61].
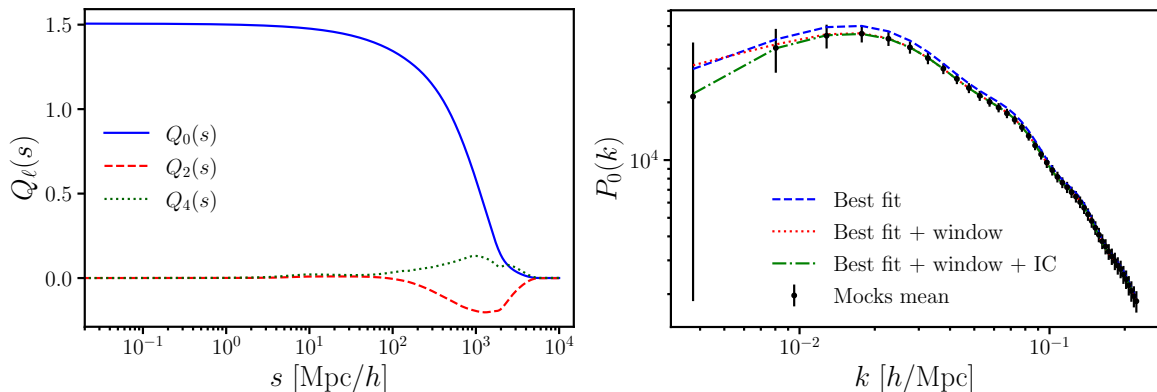
Figure 5. *Left:* multipoles of the window function in configuration space. The blue solid line, red dashed, and dotted green respectively are the monopole, quadrupole, and hexadecapole of the window function. *Right*: the effect of the different components of the observed power spectrum model compared to the mean power spectrum of the EZmock realistic catalogs (black circles with error bars). The blue dashed line is the monopole of the best fit model, the dotted red line is the best fit model convolved with the window function, and the dash-dotted green line is the observed power spectrum corrected with the integral constraint as in eq. (3.14).

First, we compute the mean of the power spectra of the complete EZmocks, $\bar{P}_{\rm c}(k)$, and the mean of the power spectra of the shuffled EZmock, $\bar{P}_{\rm r}(k)$. Then, the radial integral constraint correction is defined via

$$W_{\rm RIC}(k) = \frac{\bar{P}_{\rm c}(k) - \bar{P}_{\rm r}(k)}{\bar{P}_{\rm c}(k)} \, . \tag{3.13}$$

Given $|W(k)|^2$ and $W_{\rm RIC}(k)$ the final expression for the power spectrum is [21, 55]

$$P_0^{\rm IC}(k) = P_0(k) - P_0(0) \, |W(k)|^2 - P_0(k) \, W_{\rm RIC}(k) \, , \tag{3.14}$$

where we dropped the dependence on the effective redshift for brevity. In figure 5 we present the effect of the different components of the observed power spectrum model and compare it with the mean power spectrum of the EZmock realistic catalogs. First, the plot shows the importance of the window convolution (dotted red line), which removes power at $k \lesssim 5 \times 10^{-2} \, ({\rm Mpc}/h)^{-1}$. Second, the integral constraint correction (dash-dotted green line) only affects the first bin, but it is well within the 68% error bars.

3.4 Parameter estimation

To estimate the posterior distribution of the parameters $\boldsymbol{\theta}$ of our model, $\mathbf{T}(\boldsymbol{\theta})$, given our data vector $\mathbf{D}$, we assume a multi-variate Gaussian likelihood,

$$\mathcal{L}(\mathbf{D}|\boldsymbol{\theta}, \boldsymbol{\Sigma}) \propto \exp\left(-\frac{1}{2} \sum_{ij} (D_i - T_i(\boldsymbol{\theta})) \, \Sigma_{ij}^{-1} (D_j - T_j(\boldsymbol{\theta}))\right) \, . \tag{3.15}$$

The model parameters $\boldsymbol{\theta}$ are three for each field of view: $\sigma_{\rm FoG}$, b_1 and N (see section 3.1), and $f_{\rm NL}$. The rest of the cosmology is fixed to the Planck best fit values [53]. Therefore, when fitting the data of one sky patch (single field analysis) the total number of free parameters is four, and when fitting the two fields of view data (joint analysis) the free parameters are

seven. In the case of the joint analysis f_{NL} is common for the two fields, while the other three parameters of the model are unique for each field of view, for a total of six parameters. We assume a uniform prior distribution for all the parameters, with the following bounds

$$
\begin{aligned}
f_{\mathrm{NL}} &\in [-500, 500] \,, \\
b_1 &\in [0.1, 6] \,, \\
\sigma_{\mathrm{FoG}} &\in [0, 20] \,, \\
N &\in [-5000, 5000] \,.
\end{aligned}
\tag{3.16}
$$

In eq. (3.15), $\boldsymbol{\Sigma}^{-1}$ is the inverse of the covariance matrix estimated with the EZmock realistic catalogs (see section 2.2). As a finite number of mocks, $N_{\mathrm{m}} = 1000$, is used to estimate the covariance matrix, its inverse, the precision matrix, is biased [63]. This bias is corrected by re-scaling the covariance matrix with the inverse of the Hartlap factor,

$$
\boldsymbol{\Sigma}' = \frac{N_{\mathrm{m}} - 1}{N_{\mathrm{m}} - N_{\mathrm{b}} - 2} \, \boldsymbol{\Sigma} \,,
\tag{3.17}
$$

where N_{b} is the number of k-bins in the observed power spectrum. This correction is $\sim 5\%$ for the NGC and $\sim 6\%$ for the SGC.

The MonteCarlo Markov Chain (MCMC) algorithm employed in this analysis is the Hamiltonian MonteCarlo [HMC; 64–66]. HMC is a sampling technique that combines principles from Hamiltonian mechanics and MCMC methods to efficiently explore high-dimensional parameter spaces. Unlike traditional methods, HMC employs a dynamic integration of the target probability distribution. By introducing auxiliary momentum variables, HMC maps the trajectory of the walkers, which explore the posterior distribution, into a Hamiltonian system, whose potential energy is defined by the product of the likelihood and the prior distribution. Then, HMC exploits the gradients of the distribution to follow the Hamiltonian trajectories and efficiently sample the posterior even in a high-dimensional space. In this work we use the No-U-Turn Sampler [NUTS; 67, 68] implementation of HMC, which is able to automatically adapt critical parameters like the step size and the trajectory length. By setting the acceptance rate to 0.9, the chains quickly converge, within a few thousand steps, to $R - 1 \lesssim 10^{-3}$, where R is the Gelman-Rubin statistics [69].

4 Constraints and Discussion

In this section we present and discuss the constraints on f_{NL} we obtained with the analyses of DR16Q. Figure 3 shows the measured data points and error bars of the monopole of the power spectrum with the best fit model of the joint analyses. The plots are presented for the two sky regions, NGC (left column) and SGC (right column), and the three weighting schemes, the standard FKP with a model assuming $p = 1.6$ (top row), the optimal weights for $p = 1.0$ (middle row) and $p = 1.6$ (bottom row).

Figures 6 and 7 show the two-dimensional posterior of the joint analyses for the linear and NN catalog, respectively. Both figures are organized as follows: the left column corresponds to the analysis of the power spectrum monopoles measured with the standard FKP weights and the right column to the analysis of the optimally weighted power spectrum; the top row presents the results for $p = 1.0$, and the bottom row to $p = 1.6$. In the plots, we show three of the seven fit parameters: f_{NL}, and the linear bias b_1 of the two sky caps. In all the analyses there is almost no correlation between the biases of the two sky regions, which
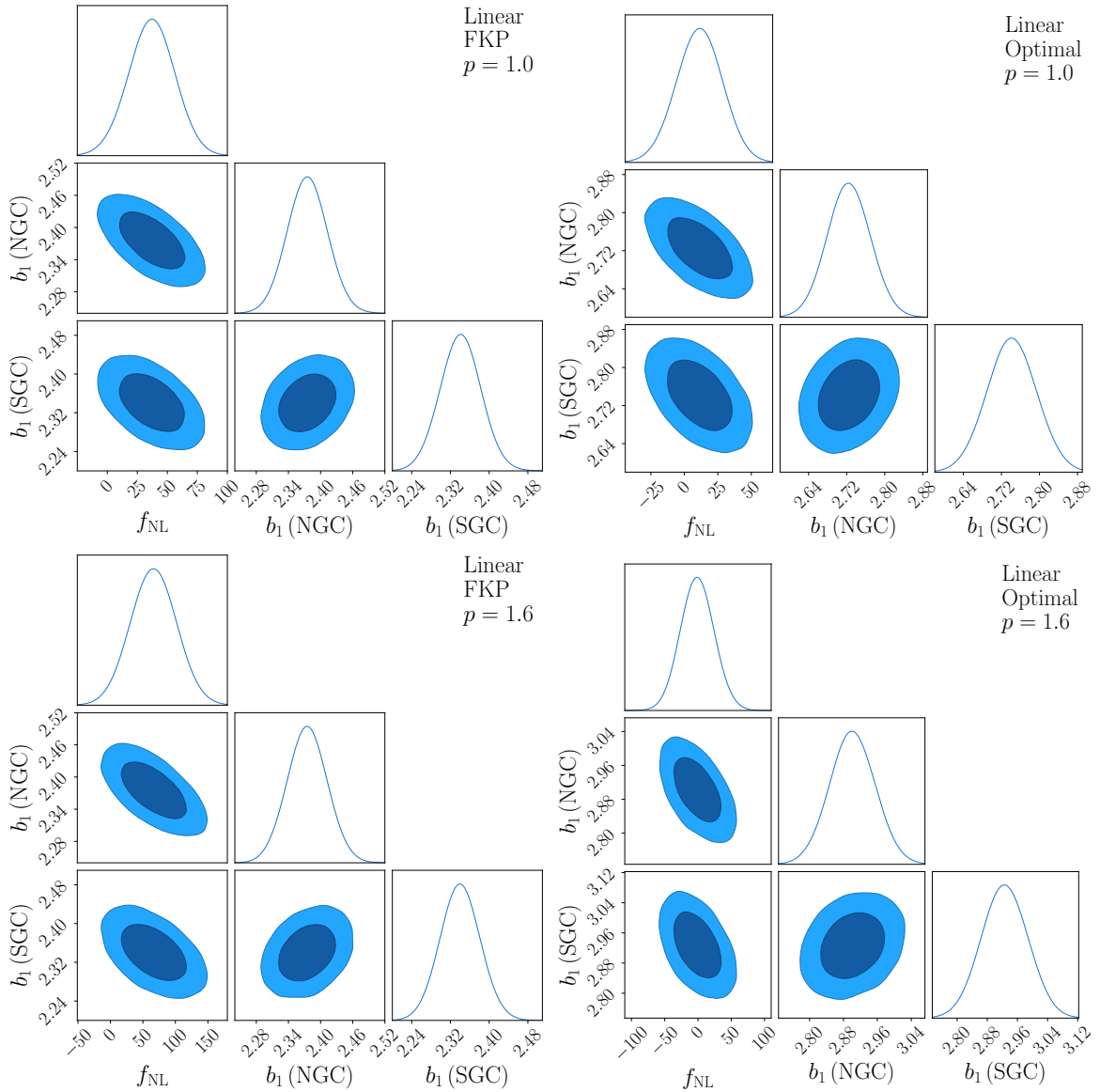
Figure 6. Two-dimensional posterior distributions for $f_{\rm NL}$ and the quasar bias, b_1, of NGC and SGC from the joint analysis of the linear catalog. The plots on the top correspond to $p = 1.0$, and the bottom plots to $p = 1.6$. On the left are the results for the FKP weighting scheme and on the right for the optimal weights.

was expected as they correspond to independent fields of view. Moreover, the linear biases of the two Galactic caps are always consistent with each other, and the bias estimated with the optimal weights is larger than the bias estimated using the standard FKP weights. The reason behind the latter behavior is the higher effective redshift of the sample when using the optimal weighting scheme, as discussed in section 2.3. Another effect visible in figures 6 and 7 is how the correlation between the linear bias and $f_{\rm NL}$ changes between the FKP weights and the optimal weights. Even though this effect is present for both p values it is more evident in the case with $p = 1.6$. The optimal weights reduce the correlation between $f_{\rm NL}$ and the linear bias. We observe this same behavior for the other fit parameters.
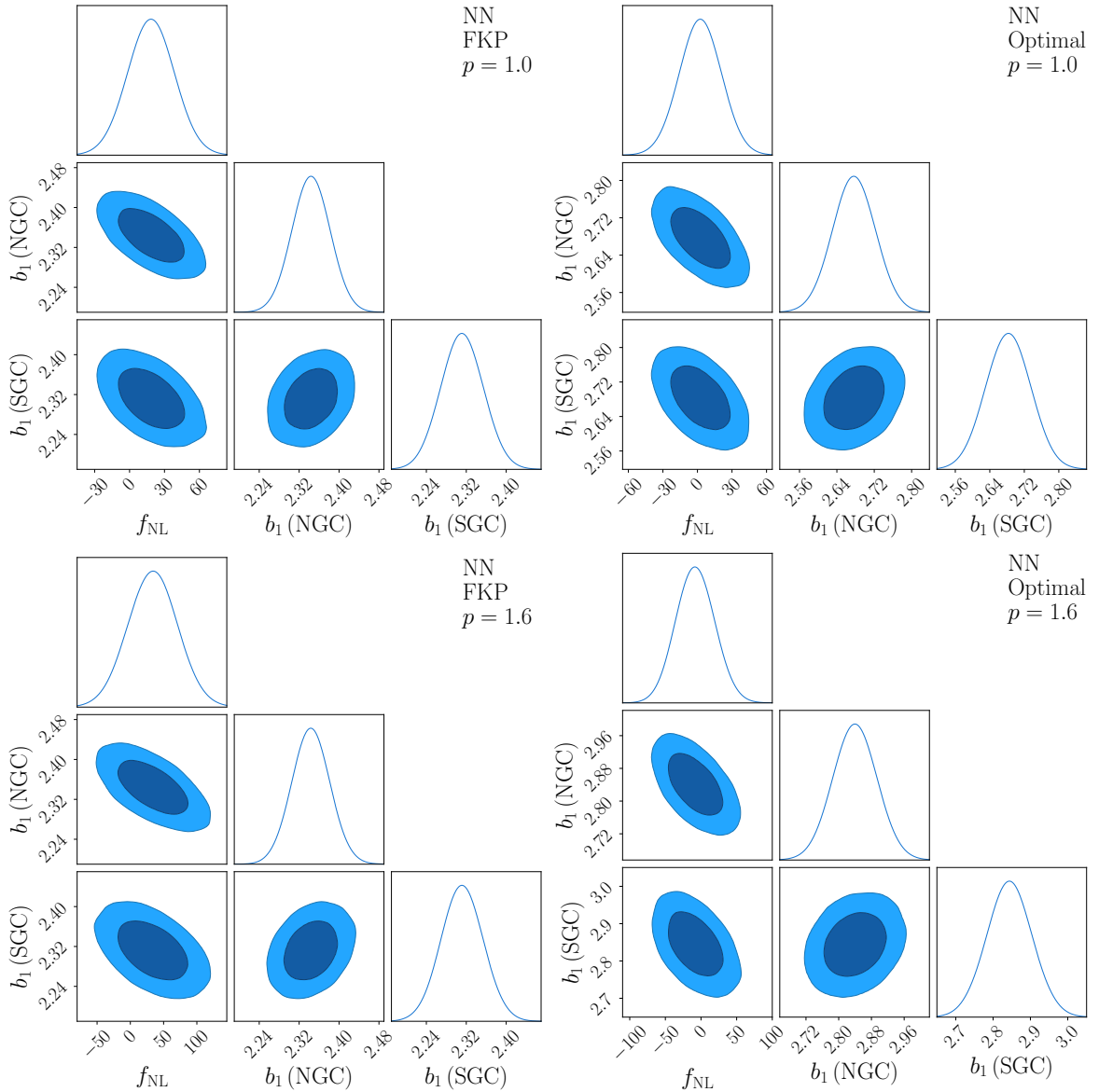
Figure 7. Same as figure 6, but for the NN catalog.

Figure 8 presents the comparison of the one-dimensional $f_{\rm NL}$ posterior distributions obtained with the joint analyses of the linear catalogs. The left panel corresponds to the case with $p = 1.0$ and the right panel to $p = 1.6$. In both panels, the red dashed line is the $f_{\rm NL}$ posterior of the FKP analysis and the dotted blue line represents the posterior of the optimal weight analysis. The corresponding vertical lines mark the 95% constraints. For both values of p the 95% constraints estimated with the standard FKP weights do not contain $f_{\rm NL} = 0$. Given that CMB, which measures $f_{\rm NL} = 0.8 \pm 5$, and LSS probe the primordial power spectrum over the same range of scales, the results suggest the presence of residual contamination in the FKP catalog produced with the linear systematic weights. The optimal weights shift the posterior to values more consistent with $f_{\rm NL} = 0$. Serendipitously, this suggests that higher redshifts QSOs in the samples might be less affected by systematic
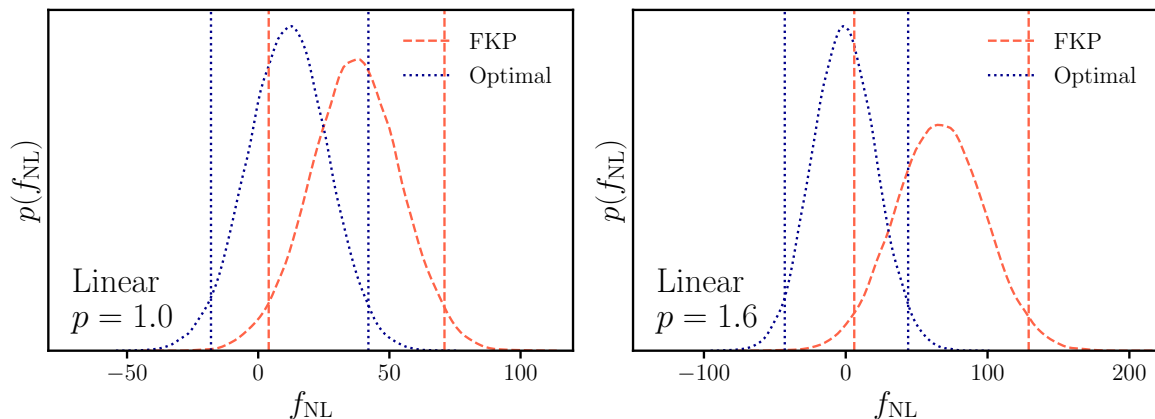
Figure 8. One dimensional posterior distribution for f_{NL} for the joint analysis of the linear catalog. The dashed red curve is the posterior distribution obtained with the FKP weight analysis and the dotted blue curve with the optimal weights analysis. The vertical lines mark the corresponding 95% constraints. On the left, the results for $p = 1.0$, and on the right for $p = 1.6$.

effects. Nevertheless, the most important difference between the standard FKP and optimal weights is that the optimal weights give tighter constraints. For $p = 1.0$ the optimal weights improve the 95% constraint of about 10%, for $p = 1.6$ this improvement is a little less than 30%. The comparison between the optimally weighted and the unweighted constraints is difficult, due to the large systematic effects still present in the FKP catalogs, especially the NGC ones. At 95% c.l. the joint analysis for $p = 1.0$ does not contain $f_{NL} = 0$ for the linear catalogs, and it barely contains it at the 68% c.l. for the NN ones, see table 2. Nevertheless, the larger improvements with $p = 1.6$ than with $p = 1.0$ point in the direction of $b_\phi \sim b_1 - 1.6$.

The 68% and 95% constraints on f_{NL} are written down in table 2, where we also make a comparison with the results of ref. [20] on the eBOSS DR14Q data. We always get tighter constraints with the linear catalogs rather than with the NN catalogs. The improvement in the constraints using the optimal weights in comparison to the standard FKP ones is the same for the two Galactic caps. On the other hand, the improvements with respect to the eBOSS DR14Q analysis are between 10% and 20% for the two catalogs and are smaller than the one expected by the doubling of the survey volume. This can be attributed to large scale systematic effects that are still present in the sample. In particular, the constraints from the single field analyses show smaller improvements with respect to the eBOSS DR14Q sample than the joint analysis. In the case of SGC, the constraints are even worse than in the older ones. Nevertheless, the single field analyses give posterior distributions consistent with each other and their combination in the joint analysis produces the tightest bounds. The best constraints of this work are from the joint analysis of the linear catalog with the optimal weights. The 95% constraints for $p = 1.0$ are $-18 < f_{NL} < 42$ corresponding to $\sigma_{f_{NL}} \sim 15$ and for $p = 1.6$ they are $-43 < f_{NL} < 44$ with $\sigma_{f_{NL}} \sim 22$.

We also repeated the parameters estimation assuming a value of $p = 3.0$. In this case, we expect a worse constraint on f_{NL} from both the FKP and the optimal analysis compared to the cases discussed above. However, as discussed in section 1, we can use these constraints to show how the optimal analysis can provide a data-driven estimate of the value of p. If the value of p used in the optimal weights is not close to the true one, then the optimal analysis will not improve over the standard case or will improve less than an analysis with a value of

	p		C.L.	Linear	NN	eBOSS DR14Q
joint	1.0	FKP	68%	$19 < f_{\mathrm{NL}} < 53$	$0 < f_{\mathrm{NL}} < 36$	
			95%	$4 < f_{\mathrm{NL}} < 71$	$-16 < f_{\mathrm{NL}} < 54$	$-39 < f_{\mathrm{NL}} < 41$
		Optimal	68%	$-4 < f_{\mathrm{NL}} < 27$	$-14 < f_{\mathrm{NL}} < 19$	
			95%	$-18 < f_{\mathrm{NL}} < 42$	$-30 < f_{\mathrm{NL}} < 34$	$-51 < f_{\mathrm{NL}} < 21$
	1.6	FKP	68%	$34 < f_{\mathrm{NL}} < 97$	$0 < f_{\mathrm{NL}} < 66$	
			95%	$6 < f_{\mathrm{NL}} < 129$	$-32 < f_{\mathrm{NL}} < 98$	$-74 < f_{\mathrm{NL}} < 81$
		Optimal	68%	$-23 < f_{\mathrm{NL}} < 21$	$-33 < f_{\mathrm{NL}} < 15$	
			95%	$-43 < f_{\mathrm{NL}} < 44$	$-54 < f_{\mathrm{NL}} < 40$	$-81 < f_{\mathrm{NL}} < 26$
NGC	1.0	FKP	68%	$26 < f_{\mathrm{NL}} < 67$	$1 < f_{\mathrm{NL}} < 44$	
			95%	$6 < f_{\mathrm{NL}} < 87$	$-23 < f_{\mathrm{NL}} < 63$	$-34 < f_{\mathrm{NL}} < 61$
		Optimal	68%	$-6 < f_{\mathrm{NL}} < 34$	$-20 < f_{\mathrm{NL}} < 21$	
			95%	$-27 < f_{\mathrm{NL}} < 52$	$-42 < f_{\mathrm{NL}} < 41$	$-56 < f_{\mathrm{NL}} < 38$
	1.6	FKP	68%	$49 < f_{\mathrm{NL}} < 125$	$-2 < f_{\mathrm{NL}} < 78$	
			95%	$10 < f_{\mathrm{NL}} < 159$	$-38 < f_{\mathrm{NL}} < 121$	$-67 < f_{\mathrm{NL}} < 112$
		Optimal	68%	$-39 < f_{\mathrm{NL}} < 28$	$-53 < f_{\mathrm{NL}} < 17$	
			95%	$-80 < f_{\mathrm{NL}} < 54$	$-95 < f_{\mathrm{NL}} < 47$	$-87 < f_{\mathrm{NL}} < 42$
SGC	1.0	FKP	68%	$-15 < f_{\mathrm{NL}} < 41$	$-17 < f_{\mathrm{NL}} < 45$	
			95%	$-36 < f_{\mathrm{NL}} < 72$	$-43 < f_{\mathrm{NL}} < 78$	$-64 < f_{\mathrm{NL}} < 31$
		Optimal	68%	$-17 < f_{\mathrm{NL}} < 29$	$-21 < f_{\mathrm{NL}} < 32$	
			95%	$-35 < f_{\mathrm{NL}} < 55$	$-40 < f_{\mathrm{NL}} < 63$	$-61 < f_{\mathrm{NL}} < 26$
	1.6	FKP	68%	$-22 < f_{\mathrm{NL}} < 78$	$-34 < f_{\mathrm{NL}} < 81$	
			95%	$-61 < f_{\mathrm{NL}} < 135$	$-82 < f_{\mathrm{NL}} < 146$	$-122 < f_{\mathrm{NL}} < 63$
		Optimal	68%	$-28 < f_{\mathrm{NL}} < 36$	$-37 < f_{\mathrm{NL}} < 40$	
			95%	$-51 < f_{\mathrm{NL}} < 74$	$-66 < f_{\mathrm{NL}} < 84$	$-92 < f_{\mathrm{NL}} < 42$

Table 2. Summary of the f_{NL} 68% and 95% constraints of this work. The results for NGC, SGC and the joint analysis are presented, and compared with the eBOSS DR14Q constraints [20].

$\tilde{w} \sim b_\phi$ closer to the actual response. Note that this approach assumes that the local PNG signal we are looking for is non-zero. In the case of $p = 3.0$, the NGC analysis always shows evidence for non-zero f_{NL} at the 95% c.l., thus we focus on SGC. For the linear catalog, the constraints are $-140 < f_{\mathrm{NL}} < 81$ for the FKP weights and $-129 < f_{\mathrm{NL}} < 76$ for the optimal analysis. This improvement by 8% should be compared to the 20% and 56% reduction of the error bar for $p = 1.0$ and $p = 1.6$ respectively in the SGC analysis. This implies that larger values of $p \gtrsim 3$ are disfavored for this sample.

5 Conclusions

In this work, we presented the most stringent constraint on the amplitude of local Primordial Non-Gaussianities with Large Scale Structure data, in particular with the eBOSS DR16Q data set. Assuming the QSOs response to $f_{\rm NL}$ is proportional to $b_1 - p$, where b_1 is the linear bias, our strongest bounds read

$$\begin{aligned}
-4 < f_{\rm NL} < 27\,, & \quad \text{for } p = 1.0\,, \\
-23 < f_{\rm NL} < 21\,, & \quad \text{for } p = 1.6\,,
\end{aligned} \tag{5.1}$$

at 68% c.l..

Our goal was to show that the optimal signal weighting reduces the error bars on $f_{\rm NL}$ compared to standard analysis, and we robustly find improvement between 10-30% depending on the analysis setup. While our optimal constraints are always consistent with no local PNG, the comparison with previous eBOSS data releases does not allow us to exclude the presence of residual systematic effects in the data. Nevertheless, the DR16Q catalog and the analysis presented here represent an important step forward in the direction of robust and optimal analysis of PNG with LSS data. We have also shown how optimal weights could provide a data-driven prior on the largely unknown value of p, and we were able to exclude $p \gtrsim 3$.

This work can be extended in several directions. First, we have not attempted an optimal noise weighting of the power spectrum data. This could be done using optimal quadratic estimators [34], for which new algorithms have been recently presented [70]. A fully optimal analysis will allow us to get closer to the full Fisher information contained in the power spectrum. Secondly, it is well known that the Bispectrum is the most sensitive probe to local PNG. A careful study of the optimal weights for higher-point statistics is still missing and could revolutionize the way we constrain $f_{\rm NL}$. We intend to return to these interesting problems in future work.

Acknowledgments

References

[1] D. Baumann, *Inflation*, in *Theoretical Advanced Study Institute in Elementary Particle Physics: Physics of the Large and the Small*, pp. 523–686, 2011, DOI [0907.5424].

[2] J.M. Maldacena, *Non-Gaussian features of primordial fluctuations in single field inflationary models*, *JHEP* **05** (2003) 013 [astro-ph/0210603].

[3] P. Creminelli and M. Zaldarriaga, *Single field consistency relation for the 3-point function*, *JCAP* **10** (2004) 006 [astro-ph/0407059].

[4] G. Cabass, E. Pajer and F. Schmidt, *How Gaussian can our Universe be?*, *JCAP* **01** (2017) 003 [1612.00033].

[5] L. Senatore and M. Zaldarriaga, *The Effective Field Theory of Multifield Inflation*, *JHEP* **04** (2012) 024 [1009.2093].

[6] M. Alvarez et al., *Testing Inflation with Large Scale Structure: Connecting Hopes with Reality*, 1412.4671.

[7] Planck Collaboration, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. IX. Constraints on primordial non-Gaussianity*, *Astron. Astrophys.* **641** (2020) A9 [1905.05697].

[8] CMB-S4 collaboration, *CMB-S4 Science Book, First Edition*, 1610.02743.

[9] N. Dalal, O. Dore, D. Huterer and A. Shirokov, *The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects*, *Phys.Rev.* **D77** (2008) 123514 [0710.4560].

[10] A. Barreira, *Predictions for local PNG bias in the galaxy power spectrum and bispectrum and the consequences for f_{NL} constraints*, *JCAP* **01** (2022) 033 [2107.06887].

[11] A. Barreira, *Can we actually constrain f_{NL} using the scale-dependent bias effect? An illustration of the impact of galaxy bias uncertainties using the BOSS DR12 galaxy power spectrum*, *JCAP* **11** (2022) 013 [2205.05673].

[12] E. Castorina, Y. Feng, U. Seljak and F. Villaescusa-Navarro, *Primordial non-Gaussianities and zero bias tracers of the Large Scale Structure*, *Phys. Rev. Lett.* **121** (2018) 101301 [1803.11539].

[13] J.M. Sullivan, T. Prijon and U. Seljak, *Learning to concentrate: multi-tracer forecasts on local primordial non-Gaussianity with machine-learned bias*, *JCAP* **08** (2023) 004 [2303.08901].

[14] A. Barreira and E. Krause, *Towards optimal and robust f_{NL} constraints with multi-tracer analyses*, 2302.09066.

[15] DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L.E. Allen et al., *The DESI Experiment Part I: Science,Targeting, and Survey Design*, *arXiv e-prints* (2016) arXiv:1611.00036 [1611.00036].

[16] R. Laureijs, J. Amiaux, S. Arduini, J.L. Auguères, J. Brinchmann, R. Cole et al., *Euclid Definition Study Report*, *arXiv e-prints* (2011) arXiv:1110.3193 [1110.3193].

[17] B.P. Crill, M. Werner, R. Akeson, M. Ashby, L. Bleem, J.J. Bock et al., *SPHEREx: NASA's near-infrared spectrophotometric all-sky survey*, in *Space Telescopes and Instrumentation 2020: Optical, Infrared, and Millimeter Wave*, M. Lystrup, M.D. Perrin, N. Batalha, N. Siegler and E.C. Tong, eds., vol. 11443, p. 114430I, International Society for Optics and Photonics, SPIE, 2020, DOI.

[18] LSST Science Collaboration, P.A. Abell, J. Allison, S.F. Anderson, J.R. Andrew, J.R.P. Angel et al., *LSST Science Book, Version 2.0*, *arXiv e-prints* (2009) arXiv:0912.0201 [0912.0201].

[19] A. Achúcarro et al., *Inflation: Theory and Observations*, 2203.08128.

[20] E. Castorina et al., *Redshift-weighted constraints on primordial non-Gaussianity from the clustering of the eBOSS DR14 quasars in Fourier space*, *JCAP* **09** (2019) 010 [1904.08859].

[21] E.-M. Mueller et al., *Primordial non-Gaussianity from the completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey II: measurements in Fourier space with optimal weights*, *Mon. Not. Roy. Astron. Soc.* **514** (2022) 3396.

[22] G. D'Amico, M. Lewandowski, L. Senatore and P. Zhang, *Limits on primordial non-Gaussianities from BOSS galaxy-clustering data*, 2201.11518.

[23] G. Cabass, M.M. Ivanov, O.H.E. Philcox, M. Simonović and M. Zaldarriaga, *Constraints on multifield inflation from the BOSS galaxy survey*, *Phys. Rev. D* **106** (2022) 043506 [2204.01781].

[24] N. Sailer, E. Castorina, S. Ferraro and M. White, *Cosmology at high redshift — a probe of fundamental physics*, *JCAP* **12** (2021) 049 [2106.09713].

[25] G. Cabass, M.M. Ivanov, O.H.E. Philcox, M. Simonovic and M. Zaldarriaga, *Constraining Single-Field Inflation with MegaMapper*, 2211.14899.

[26] Cosmic Visions 21 cm collaboration, *Inflation and Early Dark Energy with a Stage II Hydrogen Intensity Mapping experiment*, 1810.09572.

[27] D. Bragança, Y. Donath, L. Senatore and H. Zheng, *Peeking into the next decade in Large-Scale Structure Cosmology with its Effective Field Theory*, 2307.04992.

[28] D. Karagiannis, A. Lazanu, M. Liguori, A. Raccanelli, N. Bartolo and L. Verde, *Constraining primordial non-Gaussianity with bispectrum and power spectrum from upcoming optical and radio surveys*, *Mon. Not. Roy. Astron. Soc.* **478** (2018) 1341 [1801.09280].

[29] M. Rezaie et al., *Local primordial non-Gaussianity from the large-scale clustering of photometric DESI luminous red galaxies*, 2307.01753.

[30] A.J. Ross, J. Bautista, R. Tojeiro, S. Alam, S. Bailey, E. Burtin et al., *The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: Large-scale structure catalogues for cosmological analysis*, *Mon. Not. R. Astron. Soc.* **498** (2020) 2354 [2007.09000].

[31] B.W. Lyke, A.N. Higley, J.N. McLane, D.P. Schurhammer, A.D. Myers, A.J. Ross et al., *The Sloan Digital Sky Survey Quasar Catalog: Sixteenth Data Release*, *Astrophys. J. Suppl.* **250** (2020) 8 [2007.09001].

[32] M. Tegmark, *How to measure CMB power spectra without losing information*, *Phys. Rev. D* **55** (1997) 5895 [astro-ph/9611174].

[33] J.R. Bond, A.H. Jaffe and L. Knox, *Radical Compression of Cosmic Microwave Background Data*, *Astrophys. J.* **533** (2000) 19 [astro-ph/9808264].

[34] M. Tegmark, A.J.S. Hamilton, M.A. Strauss, M.S. Vogeley and A.S. Szalay, *Measuring the Galaxy Power Spectrum with Future Redshift Surveys*, *Astrophys. J.* **499** (1998) 555 [astro-ph/9708020].

[35] A. Slosar, C. Hirata, U. Seljak, S. Ho and N. Padmanabhan, *Constraints on local primordial non-Gaussianity from large scale structure*, *JCAP* **08** (2008) 031 [0805.3580].

[36] M. Biagetti, *The Hunt for Primordial Interactions in the Large Scale Structures of the Universe*, *Galaxies* **7** (2019) 71 [1906.12244].

[37] M. Biagetti, T. Lazeyras, T. Baldauf, V. Desjacques and F. Schmidt, *Verifying the consistency relation for the scale-dependent bias from local primordial non-Gaussianity*, *Mon. Not. Roy. Astron. Soc.* **468** (2017) 3277 [1611.04901].

[38] A. Barreira, G. Cabass, F. Schmidt, A. Pillepich and D. Nelson, *Galaxy bias and primordial non-Gaussianity: insights from galaxy formation simulations with IllustrisTNG*, *JCAP* **12** (2020) 013 [2006.09368].

[39] R. Scoccimarro, R.K. Sheth, L. Hui and B. Jain, *How many galaxies fit in a halo? Constraints on galaxy formation efficiency from spatial clustering*, *Astrophys. J.* **546** (2001) 20 [astro-ph/0006319].

[40] S. Yuan, B. Hadzhiyska, S. Bose and D.J. Eisenstein, *Illustrating galaxy–halo connection in the DESI era with illustrisTNG*, *Mon. Not. Roy. Astron. Soc.* **512** (2022) 5793 [2202.12911].

[41] H.A. Feldman, N. Kaiser and J.A. Peacock, *Power-Spectrum Analysis of Three-dimensional Redshift Surveys*, *Astrophys. J.* **426** (1994) 23 [astro-ph/9304022].

[42] M. Rezaie, A.J. Ross, H.-J. Seo, E.-M. Mueller, W.J. Percival, G. Merz et al., *Primordial non-Gaussianity from the completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey - I: Catalogue preparation and systematic mitigation*, *Mon. Not. R. Astron. Soc.* **506** (2021) 3439 [2106.13724].

[43] M.R. Blanton, M.A. Bershady, B. Abolfathi, F.D. Albareti, C. Allende Prieto, A. Almeida et al., *Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe*, *Astron. J.* **154** (2017) 28 [1703.00052].

[44] A. de Mattia and V. Ruhlmann-Kleider, *Integral constraints in spectroscopic surveys*, *Journal of Cosmology and Astro-Particle Physics* **2019** (2019) 036 [1904.08851].

[45] C. Zhao, C.-H. Chuang, J. Bautista, A. de Mattia, A. Raichoor, A.J. Ross et al., *The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: 1000 multi-tracer mock catalogues with redshift evolution and systematics for galaxies and quasars of the final data release*, *Mon. Not. R. Astron. Soc.* **503** (2021) 1149 [2007.08997].

[46] C.-H. Chuang, F.-S. Kitaura, F. Prada, C. Zhao and G. Yepes, *EZmocks: extending the Zel'dovich approximation to generate mock galaxy catalogues with accurate clustering statistics*, *Mon. Not. R. Astron. Soc.* **446** (2015) 2621 [1409.1124].

[47] P. Laurent, S. Eftekharzadeh, J.-M. Le Goff, A. Myers, E. Burtin, M. White et al., *Clustering of quasars in SDSS-IV eBOSS: study of potential systematics and bias determination*, *Journal of Cosmology and Astro-Particle Physics* **2017** (2017) 017 [1705.04718].

[48] K. Yamamoto, M. Nakamichi, A. Kamino, B.A. Bassett and H. Nishioka, *A Measurement of the Quadrupole Power Spectrum in the Clustering of the 2dF QSO Survey*, *Publ. Astron. Soc. Japan* **58** (2006) 93 [astro-ph/0505115].

[49] N. Hand, Y. Feng, F. Beutler, Y. Li, C. Modi, U. Seljak et al., *nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure*, *Astron. J.* **156** (2018) 160 [1712.05834].

[50] D. Bianchi, H. Gil-Marín, R. Ruggeri and W.J. Percival, *Measuring line-of-sight dependent Fourier-space clustering using FFTs*, *Mon. Not. Roy. Astron. Soc.* **453** (2015) L11 [1505.05341].

[51] N. Hand, Y. Li, Z. Slepian and U. Seljak, *An optimal FFT-based anisotropic power spectrum estimator*, *JCAP* **07** (2017) 002 [1704.02357].

[52] R.W. Hockney and J.W. Eastwood, *Computer Simulation Using Particles* (1981).

[53] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi et al., *Planck 2018 results. VI. Cosmological parameters*, *Astron. Astrophys.* **641** (2020) A6 [1807.06209].

[54] A. de Mattia et al., *The Completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: measurement of the BAO and growth rate of structure of the emission line galaxy sample from the anisotropic power spectrum between redshift 0.6 and 1.1*, *Mon. Not. Roy. Astron. Soc.* **501** (2021) 5616 [2007.09008].

[55] M.J. Wilson, J.A. Peacock, A.N. Taylor and S. de la Torre, *Rapid modelling of the redshift-space power spectrum multipoles for a masked density field*, *Mon. Not. R. Astron. Soc.* **464** (2017) 3121 [1511.07799].

[56] N. Kaiser, *Clustering in real space and in redshift space*, *Mon. Not. R. Astron. Soc.* **227** (1987) 1.

[57] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes*, *Journal of Cosmology and Astro-Particle Physics* **2011** (2011) 034 [1104.2933].

[58] F. Beutler, E. Castorina and P. Zhang, *Interpreting measurements of the anisotropic galaxy power spectrum*, *JCAP* **03** (2019) 040 [1810.05051].

[59] E. Castorina and M. White, *Beyond the plane-parallel approximation for redshift surveys*, *Mon. Not. Roy. Astron. Soc.* **476** (2018) 4403 [1709.09730].

[60] E. Castorina and M. White, *The Zeldovich approximation and wide-angle redshift-space distortions*, *Mon. Not. Roy. Astron. Soc.* **479** (2018) 741 [1803.08185].

[61] E. Castorina and E. Di Dio, *The observed galaxy power spectrum in General Relativity*, *JCAP* **01** (2022) 061 [2106.08857].

[62] J.A. Peacock and D. Nicholson, *The large-scale clustering of radio galaxies.*, *Mon. Not. R. Astron. Soc.* **253** (1991) 307.

[63] J. Hartlap, P. Simon and P. Schneider, *Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix*, *Astron. Astrophys.* **464** (2007) 399 [`astro-ph/0608064`].

[64] S. Duane, A.D. Kennedy, B.J. Pendleton and D. Roweth, *Hybrid Monte Carlo*, *Physics Letters B* **195** (1987) 216.

[65] R. Neal, *MCMC Using Hamiltonian Dynamics*, in *Handbook of Markov Chain Monte Carlo*, pp. 113–162 (2011), DOI.

[66] M. Betancourt, *A Conceptual Introduction to Hamiltonian Monte Carlo*, *arXiv e-prints* (2017) arXiv:1701.02434 [`1701.02434`].

[67] M.D. Hoffman and A. Gelman, *The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo*, *arXiv e-prints* (2011) arXiv:1111.4246 [`1111.4246`].

[68] H. Ge, K. Xu and Z. Ghahramani, *Turing: a language for flexible probabilistic inference*, in *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain*, pp. 1682–1690, 2018, http://proceedings.mlr.press/v84/ge18b.html.

[69] A. Gelman and D.B. Rubin, *Inference from Iterative Simulation Using Multiple Sequences*, *Statistical Science* **7** (1992) 457.

[70] O.H.E. Philcox, *Cosmology without window functions: Quadratic estimators for the galaxy power spectrum*, *Phys. Rev. D* **103** (2021) 103504 [`2012.09389`].

A   Fitting $b_\phi f_{\mathrm{NL}}$

Table 3 summarizes the constraint on $b_\phi f_{\mathrm{NL}}$ for the different measurements of the power spectrum. As in the main text, all bounds are compatible with zero PNG with the exception of the FKP analysis of NGC. Figure 9 shows the 2D posterior of the parameters in SGC.
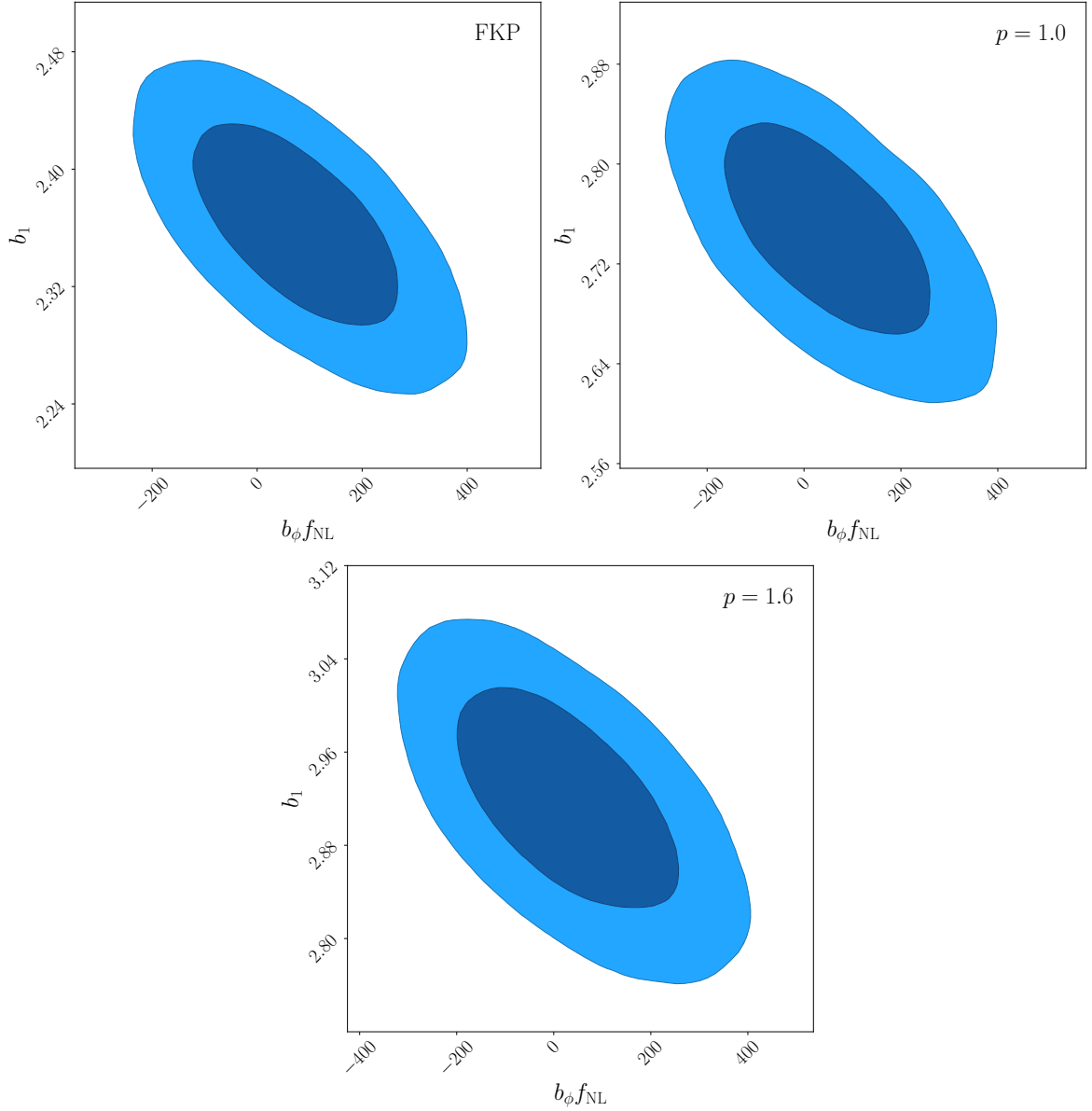
Figure 9. Two dimensional posterior distributions for $b_\phi f_{\rm NL}$ and and the quasar linear bias, b_1, from the SGC linear catalog analysis. The posterior distribution of the analysis with the FKP weighting scheme, the optimal weights with $p = 1.0$, and $p = 1.6$ are shown respectively on the top left, top right and bottom panel.

		C.L.	Linear	NN
NGC	FKP	68%	$131 < b_\phi\,f_{\mathrm{NL}} < 310$	$2 < b_\phi\,f_{\mathrm{NL}} < 192$
		95%	$35 < b_\phi\,f_{\mathrm{NL}} < 394$	$-102 < b_\phi\,f_{\mathrm{NL}} < 280$
	Optimal $p = 1.0$	68%	$-27 < b_\phi\,f_{\mathrm{NL}} < 200$	$-115 < b_\phi\,f_{\mathrm{NL}} < 123$
		95%	$-174 < b_\phi\,f_{\mathrm{NL}} < 288$	$-255 < b_\phi\,f_{\mathrm{NL}} < 227$
	Optimal $p = 1.6$	68%	$-176 < b_\phi\,f_{\mathrm{NL}} < 121$	$-222 < b_\phi\,f_{\mathrm{NL}} < 76$
		95%	$-371 < b_\phi\,f_{\mathrm{NL}} < 234$	$-409 < b_\phi\,f_{\mathrm{NL}} < 192$
SGC	FKP	68%	$-48 < b_\phi\,f_{\mathrm{NL}} < 198$	$-83 < b_\phi\,f_{\mathrm{NL}} < 188$
		95%	$-162 < b_\phi\,f_{\mathrm{NL}} < 322$	$-207 < b_\phi\,f_{\mathrm{NL}} < 323$
	Optimal $p = 1.0$	68%	$-93 < b_\phi\,f_{\mathrm{NL}} < 175$	$-129 < b_\phi\,f_{\mathrm{NL}} < 174$
		95%	$-202 < b_\phi\,f_{\mathrm{NL}} < 317$	$-230 < b_\phi\,f_{\mathrm{NL}} < 349$
	Optimal $p = 1.6$	68%	$-121 < b_\phi\,f_{\mathrm{NL}} < 165$	$-144 < b_\phi\,f_{\mathrm{NL}} < 182$
		95%	$-237 < b_\phi\,f_{\mathrm{NL}} < 311$	$-265 < b_\phi\,f_{\mathrm{NL}} < 354$

Table 3. Summary on the 68% and 95% constraints on $b_\phi\,f_{\mathrm{NL}}$ for the NGC and SGC.