Constraints on primordial non-Gaussianity from 800,000 photometric quasars

Boris Leistedt,1 Hiranya V. Peiris,1 and Nina Roth1

1Department of Physics and Astronomy, University College London, London WC1E 6BT, U.K.

(Dated: August 29, 2018)

We derive robust constraints on primordial non-Gaussianity (PNG) using the clustering of 800,000 photometric quasars from the Sloan Digital Sky Survey in the redshift range 0.5 < z < 3.5. These measurements rely on the novel technique of extended mode projection to control the impact of spatially-varying systematics in a robust fashion, making use of blind analysis techniques. This allows the accurate measurement of quasar halo bias at the largest scales, while discarding as little as possible of the data. The standard-local type PNG parameters fNL and gNL both imprint a k−2 scale-dependent effect in the bias. Constraining these individually, we obtain −49 < fNL < 31 and −2.7 × 105 < gNL < 1.9 × 105, while their joint constraints lead to −105 < fNL < 72 and −4.0 × 105 < gNL < 4.9 × 105 (all at 95% CL). Introducing a running parameter nNL to constrain b(k) ∝ k−2+nNL and a generalised PNG amplitude fNL, we obtain −45.5 exp(3.7 nNL) < fNL < 34.4 exp(3.3 nNL) at 95% CL. These results incorporate uncertainties in the cosmological parameters, redshift distributions, shot noise, and the bias prescription used to relate the quasar clustering to the underlying dark matter. These are the strongest constraints obtained to date on PNG using a single population of large-scale structure tracers, and are already at the level of pre-Planck constraints from the cosmic microwave background. A conservative forecast for a Large Synoptic Survey Telescope-like survey incorporating mode projection yields σ(fNL) ∼ 5 – competitive with the Planck result – highlighting the power of upcoming large scale structure surveys to probe the initial conditions of the universe.

PACS numbers: 98.80.Cq, 98.80.Es, 98.65.Dx, 98.54.Aj

Canonical single-field slow-roll inflation predicts initial conditions for structure formation that are essentially Gaussian [1,2]. Any measurement of deviations from this prediction — summarised by the term primordial non-Gaussianity (PNG) — can thus provide evidence for non-standard inflationary physics. One of the most physically interesting forms of PNG is the so-called local model, where the primordial potential φ is modified by including higher order terms,

\[ \Phi = \phi + f_{NL}[\phi^2 - \langle \phi^2 \rangle] + g_{NL}[\phi^3 - 3\phi \langle \phi^2 \rangle], \]

where all fields are evaluated at the same spatial coordinate, and fNL and gNL are real-valued constants (often called the skewness and kurtosis parameters).

The most stringent constraints on PNG currently come from higher-order statistics of the cosmic microwave background (CMB). Most recently, the Planck collaboration reported −8.9 < fNL < 14.3 (95% CL) [13], while constraints on the kurtosis have been obtained from the WMAP satellite: −7.4 × 105 < gNL < 8.2 × 105 (WMAP5, 95% CL) [14], or gNL = (−4.3 ± 2.3) × 105 (WMAP9, 68% CL) [15], gNL = (−3.3 ± 2.2) × 105 (WMAP9, 68% CL) [16]. While these results are compatible with Gaussian initial conditions, their uncertainties still leave room for non-standard inflation models.

The unknown relation (bias) between the dark matter density field and a set of observed tracers (which inhabit dark matter halos) is generally considered to be a complication in constraining cosmological parameters from large-scale structure (LSS) data. However, in the case of PNG, the bias is actually an advantage that can be used to distinguish between non-Gaussianity in the initial conditions and that generated through late-time non-linear structure formation. PNG introduces a distinctive k-dependence into the halo bias; qualitatively, the bias for local-type PNG scales as b(k) ∝ k−2 [17,21].

This implies that the strongest signal can be expected on large scales (small k), accessible to wide-area galaxy surveys. At these scales, the bias can be well-approximated by a multiplicative factor between the dark matter- and galaxy power spectra. LSS clustering constraints on PNG provide an independent validation of the CMB results, and are predicted to improve significantly with on-going and future LSS surveys, eventually surpassing CMB constraints if systematic errors can be controlled [22,29]. Quasars – the bright nuclei at the centre of the most active galaxies – are highly-biased tracers of the LSS, spanning large volumes and covering extended redshift ranges: in principle, quasar surveys are ideal for constraining PNG [19,30]. However, previous analyses have been complicated by the presence of spurious excess power at large scales due to systematics, which mimic the signature of PNG [30,35].

In this Letter, we use a large sample of quasars [36] from the Sloan Digital Sky Survey (SDSS) [37] to constrain PNG. Using a novel technique for blind mitigation of systematics described in Leistedt & Peiris (2014) [34], we are able to significantly enhance the constraining power of the dataset, resulting in PNG constraints from a single LSS dataset which are competitive with those from the CMB. Our results represent a significant step toward achieving the exquisite control of systematics necessary.
to exploit future LSS surveys to measure PNG.

**PNG with photometric quasars.** The first LSS constraints on PNG were derived in Ref. [19] using a combination of tracers from early SDSS releases, leading to \(f_{NL} = 28^{+23}_{-24} (68\% \, \text{CL})\). Among these tracers, the photometric quasars — candidate quasars identified using imaging data only — have the highest bias and probe the largest volume. Therefore, they had the most constraining power \(f_{NL} = 8^{+28}_{-27} (68\% \, \text{CL})\), demonstrating their potential to constrain PNG and complement CMB experiments. However, subsequent analyses [30–32] of photometric quasars from the Sixth SDSS Data Release (using the catalogue from Ref. [33]) also revealed systematic effects, such as spatially-varying depth and stellar contamination, which could strongly bias the clustering measurements on the largest scales and jeopardise cosmological inferences if not properly mitigated [33–34]. Refs. [30–33] obtained \(f_{NL} = 5 \pm 21 (68\% \, \text{CL})\) using a range of LSS probes, discarding the auto-correlation of quasars to avoid the main systematic contamination. Ref. [40] used the latest catalogue of SDSS photometric quasars, XDQSOz [30–31], to obtain \(f_{NL} = 103^{+148}_{-146} (68\% \, \text{CL})\), and \(f_{NL} = 2^{-66}_{+86} (68\% \, \text{CL})\) when combined with constraints from the clustering of Luminous Red Galaxies, relying however on stringent quality cuts on the survey maps and power spectra to limit the impact of systematics. Therefore, photometric quasars have thus far remained underexploited for PNG, as systematics are handled by removing hard-won data.

**XDQSOz power spectrum measurements.** A different approach was adopted by two of us in Ref. [35], where the SDSS XDQSOz catalogue was analysed with minimal quality and sky cuts, leading to a sample of \(\sim 800,000\) quasars covering \(\sim 8300 \, \text{deg}^2\) (compared to \(\sim 400,000–500,000\) objects used by previous analyses). This base sample was further separated into four redshift bins by selecting objects with photometric redshift estimates \(\hat{z}_p\) in top-hat windows \([0.5,1.35], [1.35,1.7], [1.7,2.2], [2.2,3.5]\). A quadratic maximum likelihood [12,13] method was used to simultaneously estimate the auto- and cross-angular power spectra of the \(z\)-binned data. To mitigate the impact of systematics in these power spectra, Ref. [35] introduced a novel technique, extended mode projection, relying on the fact that most potential systematics (e.g., observing conditions, calibration) were also measured during SDSS observations, and could therefore be mapped onto the sky.

We constructed a non-linear, data-driven model of systematics, using \(\sim 3,700\) orthogonal templates obtained by decorrelating \(\sim 20,000\) maps of potential contaminants, including 200 base templates constructed from SDSS data and products of pairs. These orthogonal templates were cross-correlated with the XDQSOz data, yielding null tests which are used to select the most significant systematics following the principles of blind analysis. These were then marginalised over via mode projec-

\[
\text{FIG. 1: Redshift distributions of the four quasar samples used in this analysis, parametrised as superpositions of normal distributions. The shaded regions show the 1, 2 and 3\(\sigma\) regions explored when adding 5\% uncertainty to the parameters of these fits, which are included in the MCMC analysis.}
\]

The shot noise was measured in each sample from the average surface density of photometric quasars, but is also subject to uncertainties due to the unknown fraction of stars in the samples — between 0 and 20\% from the quality cuts applied to XDQSOz [35]. Since non-zero stellar contamination reduces the shot noise, we also marginalised over this effect when constraining PNG.

**Halo bias from PNG.** The impact of local-type PNG is to modify the halo bias by adding a \(k\)-dependent term to the Gaussian bias \(b^G(z)\) [17,21]

\[
b^{NG}(k,z) = b^G(z) + \frac{\beta_f(z)f_{NL} + \beta_g(z)g_{NL}}{\alpha(k,z)}.
\]

Here, we neglect an additional small contribution induced by the effect of PNG on the halo mass function, which is independent of \(k\) and can thus be absorbed in \(b^G(z)\). Note that we have also suppressed the implicit mass dependence of \(b^G, \beta_f, \text{and} \beta_g\) in the previous equation. The exact expression for \(\alpha(k,z)\) and fitting functions for \(\beta_f\) and \(\beta_g\) can be found in Ref. [21].
A simple extension of the local model is the introduction of a spectral index $n_{f_{NL}}$ in the $f_{NL}$-generated scale-dependent bias $b_{0,β}$, i.e., changing its scaling from $k^{-2}$ into $k^{-2+n_{f_{NL}}}$ by using

$$\alpha(k, z) \rightarrow \alpha(k, z) \left( \frac{k}{k_{\text{piv}}} \right)^{-n_{f_{NL}}},$$

where we choose $k_{\text{piv}} = 0.06$ Mpc$^{-1}$. Note that this parametrisation is not equivalent to an intrinsically scale-dependent $f_{NL}$ as described in Refs. 51–52. Instead, it allows us to extend our analysis to other types of PNG, like that generated by single-field inflation with a modified initial state 48, or models with several light fields 19.

The quasar bias is known to evolve strongly with redshift (e.g., Refs. 53–57), and thus one cannot use a constant linear bias per redshift bin due to the extended and complicated redshift distributions shown in Fig. 1. For the Gaussian bias $b(z)$ in Eq. 2, we used

$$b(z) = b_0 \left[ 1 + \left( \frac{1 + z}{2.5} \right)^{β} \right],$$

which is in good agreement with previous studies of SDSS quasars (e.g., Ref. 58–60).

**Monte Carlo Markov Chain (MCMC) analysis.** We built a Gaussian likelihood 34, jointly using the 10 auto- and cross-angular power spectra (between redshift bins) estimated in Ref. 53, at multipole resolution $Δℓ = 15$. The theoretical predictions were calculated using CAMB-sources 61, modified to support PNG and our quasar bias model. We used *emcee* 62 to run an MCMC analysis, and sample combinations of the following parameters: Cosmological parameters (‘cosmo’): parameters of the base ΛCDM model, with fiducial values and uncertainties corresponding to the constraints from Planck combined with Baryon Acoustic Oscillations (BAO), as in Ref. 63. Bias model: the model described above, with uniform priors $b_0 \in [0, 2]$ and $β \in [4, 6]$. Redshift distributions (‘n(z)’): the amplitude and width of the Gaussian functions used to fit the $n(z)$ estimates, with Gaussian priors of 5% 1σ uncertainties around the fiducial values. Additionally, we sampled the slope of number counts, which controls magnification bias, with Gaussian priors centred at the measured value with 5% 1σ uncertainty. Shot noise: we marginalised over the shot noise with a prior $[0, 1]$ times the value measured from the photometric quasar surface density, in order to account for the unknown (but bounded) amount of stellar contamination.

**Results.** We first test the robustness of the bias model by examining the bias measured in the four redshift samples individually and jointly. Therefore, in addition to the ‘coupled’ model presented above, used to connect all power spectra to the theory predictions, we consider an alternative, ‘decoupled’ case where the bias amplitude of each redshift sample is fit separately, using four parameters $b_1, b_2, b_3, b_4$. In this case, we used $β = 5$ and uniform priors $b_i \in [0, 2]$. The constraints on the bias parameters from the XDQSOz power spectra are shown in Fig. 2, and demonstrate that the separate bias amplitudes $b_i, i = 1 \ldots 4$ of the four samples are in good agreement with each other, with the fiducial model with $b_0 = 1$ and $β = 5$ (black line), and also with the results obtained with the coupled model (shaded band). Note that the slope parameter $β$ is used to capture the uncertainty in the evolution of the bias at $z > 2.5$. This
redshift range is not as strongly constrained by the data, but is nevertheless crucial since it has the greatest bias, and therefore is expected to produce the strongest PNG signature.

Having confirmed that Eq. 3 is an adequate model for the quasar bias, we can now advance towards constraining the PNG parameters. Unless stated otherwise, all values are quoted at 95% CL.

Fig. 3 shows the constraints on \( f_{\text{NL}} \) only (with \( n_{f_{\text{NL}}} = 0 \)) for various combinations of parameters and sources of uncertainties. As expected, adding parameters (e.g., using the ‘decoupled’ bias model) increases the error estimates. When only varying \( b_0, \beta \), we obtain \(-26 < f_{\text{NL}} < 34 \) (top row), and adding the uncertainties in the cosmological parameters and redshift distributions relaxes the constraints to \(-49 < f_{\text{NL}} < 31 \) (bottom row). Note that this case is more conservative but introduces degeneracies between parameters, in particular those affecting the overall amplitude of the power spectra, such as the bias \( b_0 \), the amplitude of matter fluctuations \( \sigma_8 \), and the relative amplitudes of the peaks in \( n(z) \). However, \( f_{\text{NL}} \) does not suffer from any direct degeneracies with these parameters given its specific signature on large scales, yielding constraints which are robust to the choice of model.

When only using \( g_{\text{NL}} \) as a source of scale-dependent bias, we obtain \(-2.7 \times 10^5 < g_{\text{NL}} < 1.9 \times 10^5 \). However, as apparent from Eq. 2 and as shown in Ref. [64], \( f_{\text{NL}} \) and \( g_{\text{NL}} \) leave similar signatures on the scale-dependent bias of LSS tracers, and are not easily separable. Despite the large volume and extended redshift range spanned by our data, this degeneracy is confirmed by their joint constraints, shown in Fig. 3 leading to \(-105 < f_{\text{NL}} < 72 \) and \(-4.0 \times 10^5 < g_{\text{NL}} < 4.9 \times 10^5 \).

For the extended model of Eq. 5 we can introduce a generalised parameter \( \tilde{f}_{\text{NL}} \), which only corresponds to \( f_{\text{NL}} \) as defined in Eq. 4 when \( n_{f_{\text{NL}}} = 0 \). In this model, the scale-dependent bias scales as \( b(k) \sim k^{-2 + n_{f_{\text{NL}}}} \), so the constraints on \( \tilde{f}_{\text{NL}} \) widen as \( n_{f_{\text{NL}}} \) increases, due to the less pronounced signature of PNG on large scales. Nevertheless, \( \tilde{f}_{\text{NL}} \) can be constrained at any \( n_{f_{\text{NL}}} < 2 \). Fig. 5 shows the joint constraints on \( \tilde{f}_{\text{NL}} \) and \( n_{f_{\text{NL}}} \), summarised as \(-45.5 \exp(3.7 n_{f_{\text{NL}}}) < \tilde{f}_{\text{NL}} < 34.4 \exp(3.3 n_{f_{\text{NL}}}) \), fully compatible with \( f_{\text{NL}} = 0 \).

**Future prospects.** We performed a Fisher forecast of a Large Synoptic Survey Telescope-like photometric survey [55], using angular power spectra of galaxies in 20 tomographic bins in the redshift range \( 0.5 < z < 3.5 \), with bias modelled by \( b_0 = 1, \beta = 3 \) in Eq. 4. This idealised forecast, without any contamination by systematics, yields a 1\( \sigma \) uncertainty on \( f_{\text{NL}} \) of \( \sim 1 \), consistent with the 3D power spectrum analysis of Ref. [22].

We then incorporated realistic estimates of the main 50 systematics identified in Ref. [35]. This translates into a \( \sim 10\% \) clustering contamination for \( \ell > 100 \), up to factors of a few on the largest scales where the PNG signal is strongest. Without any mitigation in the subsequent analysis, the resulting estimate for \( f_{\text{NL}} \) is biased by \( \Delta f_{\text{NL}} \sim 30 \), which is highly significant compared to the expected uncertainty. Including a harmonic-space mode projection in the forecast removes this bias, while consistently increasing the parameter uncertainty to \( \sigma(f_{\text{NL}}) = 5 \), comparable to current Planck constraints. However, this increase in uncertainty will be significantly reduced when performing the mode projection in pixel

---

**TABLE I: PNG constraints for different models.** The base model is \( (b_0, \beta) + \text{cosmo} + n(z) \). The first super/subscript corresponds to the 68% CL, and the 95% CL can be obtained by adding the second number.

| Model | \( f_{\text{NL}} \) | \( g_{\text{NL}}/10^5 \) |
|-------|------------------|-----------------|
| base+\( f_{\text{NL}} \) | \(-4^{+27+8}_{-35-10} \) | --- |
| base+\( g_{\text{NL}} \) | --- | \( 0.3^{+1.1+0.5}_{-2.2-0.8} \) |
| base+\( f_{\text{NL}}+g_{\text{NL}} \) | \(-10^{+54+28}_{-60-35} \) | \( 0.2^{+2.9+1.8}_{-2.7-1.5} \) |
space, using sky maps of the systematics which will be available for a real survey. 

**Conclusion.** Photometric quasar surveys, while potentially constituting ideal datasets for probing PNG, have thus far been systematics-limited. We derived constraints on PNG using the results of a novel power-spectrum estimation method applied to ~ 800,000 photometric quasars from SDSS. This approach self-consistently marginalises over a non-linear data-driven model of spatially-varying systematics. Our results, summarised in Table I, incorporate uncertainties in the cosmological parameters and the parameters of a realistic bias model, while marginalising over uncertainties in the redshift distributions of the quasars. The resulting constraints on $f_{\text{NL}}$ and $g_{\text{NL}}$ are the tightest obtained using a single population of LSS tracers, and are at the level of pre-Planck CMB constraints. Our results demonstrate the potential of future LSS surveys to reach the $f_{\text{NL}} \sim 1$ levels predicted by the simplest models of inflation.

**Acknowledgments.** We are grateful to Jo Bovy, David Hogg and Adam Myers for sharing the XDQSOz catalogue with us, and Jo Bovy, Masahiro Takada for useful conversations.

BL is supported by the Perren Fund and the IMPACT Fund. HVP and NR are supported by STFC and the European Research Council under the European Community’s Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no 306478-CosmicDawn.

We acknowledge use of the following public software packages: healpix [65] and camb_sources [61]. We acknowledge use of the Legacy Archive for Microwave Background Data Analysis (LAMBDA). Support for LAMBDA is provided by the NASA Office of Space Science. This work is based on observations obtained with SDSS. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is [http://www.sdss.org/](http://www.sdss.org/)

---

* Electronic address: boris.leistedt.11@ucl.ac.uk
† Electronic address: h.peiris@ucl.ac.uk
‡ Electronic address: n.roth@ucl.ac.uk

[1] T. J. Allen, B. Grinstein, and M. B. Wise, Physics Letters B 197, 66 (1987).
[2] D. S. Salopek and J. R. Bond, Phys. Rev. D 42, 3936 (1990).
[3] T. Falk, R. Rangarajan, and M. Srednicki, Astrophys. J. Lett. 403, L1 (1993), astro-ph/9208001.

---

[4] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach, Astrophys. J. 430, 447 (1994), astro-ph/9312033.
[5] L. Verde, L. Wang, A. F. Heavens, and M. Kamionkowski, Mon. Not. Roy. Astron. Soc. 313, 141 (2000), astro-ph/9906301.
[6] L. Wang and M. Kamionkowski, Phys. Rev. D 61, 063504 (2000), astro-ph/9907431.
[7] A. Gangui and J. Martin, Mon. Not. Roy. Astron. Soc. 313, 323 (2000), astro-ph/9908009.
[8] E. Komatsu and D. N. Spergel, Phys. Rev. D 63, 063002 (2001), astro-ph/0005036.
[9] J. Maldacena, Journal of High Energy Physics 5, 013 (2003), astro-ph/0210003.
[10] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nuclear Physics B 667, 119 (2003), astro-ph/0209156.
[11] D. Babich, P. Creminelli, and M. Zaldarriaga, Journal of Cosmology and Astroparticle Physics 8, 009 (2004), astro-ph/0405356.
[12] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, Physics Reports 402, 103 (2003), astro-ph/0406399.
[13] Planck Collaboration, ArXiv e-prints (2013), 1303.5084.
