MNRAS **455,** 2573–2581 (2016) doi:10.1093/mnras/stv2307

Cosmological parameter inference from galaxy clustering: the effect of the posterior distribution of the power spectrum

B. Kalus,[1]★ W. J. Percival[1] and L. Samushia[2,1,3]

[1]*Institute of Cosmology & Gravitation, Dennis Sciama Building, University of Portsmouth, Portsmouth PO1 3FX, UK*
[2]*Department of Physics, Kansas State University, 116, Cardwell Hall, Manhattan, KS 66506, USA*
[3]*National Abastumani Astrophysical Observatory, Ilia State University, 2A Kazbegi Ave, GE-1060 Tbilisi, Georgia*

Accepted 2015 October 2. Received 2015 September 1; in original form 2015 April 15

ABSTRACT

We consider the shape of the posterior distribution to be used when fitting cosmological models to power spectra measured from galaxy surveys. At very large scales, Gaussian posterior distributions in the power do not approximate the posterior distribution $\mathcal{P}_R$ we expect for a Gaussian density field δ_k, even if we vary the covariance matrix according to the model to be tested. We compare alternative posterior distributions with $\mathcal{P}_R$, both mode-by-mode and in terms of expected measurements of primordial non-Gaussianity parametrized by $f_{\rm NL}$. Marginalising over a Gaussian posterior distribution $\mathcal{P}_{\rm f}$ with fixed covariance matrix yields a posterior mean value of $f_{\rm NL}$ which, for a data set with the characteristics of Euclid, will be underestimated by $\triangle f_{\rm NL} = 0.4$, while for the data release 9 of the Sloan Digital Sky Survey-III Baryon Oscillation Spectroscopic Survey (BOSS DR9; Ahn et al.) it will be underestimated by $\triangle f_{\rm NL} = 19.1$. Adopting a different form of the posterior function means that we do not necessarily require a different covariance matrix for each model to be tested: this dependence is absorbed into the functional form of the posterior. Thus, the computational burden of analysis is significantly reduced.

Key words: methods: statistical – inflation – large-scale structure of Universe.

1 INTRODUCTION

Forthcoming galaxy surveys, such as the Dark Energy Spectroscopic Instrument (DESI; Schlegel et al. 2011), Euclid (Laureijs et al. 2011)[1] and the Square Kilometre Array (SKA),[2] will constrain the Universe's expansion history, geometry and the growth of structure with unprecedented accuracy. The basic statistics containing large-scale structure information are the two-point clustering measurements, the correlation function and the galaxy power-spectrum $P(\boldsymbol{k})$. As they form a Fourier pair, their information content is the same and we focus only on the latter in this work. The linear galaxy power spectrum encodes a wealth of information about the physics of the Universe, allowing us to constrain cosmological models with baryon acoustic oscillations (BAO), gravitational models with redshift space distortions (RSD) and inflationary models with primordial non-Gaussianity, parametrized to first order by $f_{\rm NL}$. In order to do so, one has to know the likelihood and/or posterior of power spectra, which for simple cases can be calculated analytically. For general cases, one usually assumes the likelihood or posterior to be multivariate Gaussian with a covariance matrix $C_{ij} \equiv \langle P(k_i)P(k_j)\rangle$. The estimation of the covariance matrix is a critical step in the analysis of data. Internal methods such as the subsample, jackknife and bootstrap methods have been widely used in the past, but Norberg et al. (2009) have shown that they are not able to faithfully reproduce variances. Robust estimates are often instead obtained from mock galaxy catalogues. In recent analyses of the Baryon Oscillation Spectroscopic Survey (BOSS; Ahn et al. 2012; Dawson et al. 2013), these were generated from second order Lagrangian Perturbation Theory matter fields using a friends-of-friends group finder (Davis et al. 1985) to find haloes (Scoccimarro & Sheth 2002; Manera et al. 2012). Their masses were calibrated by comparisons with N-body simulations. A Halo Occupation Distribution then prescribed how to populate these haloes with mock galaxies, and the geometry and the efficiency of the survey were sampled. Alternative methods for producing mocks include N-body simulations, comoving Lagrangian acceleration (COLA; Tassev, Zaldarriaga & Eisenstein 2013) simulations or alternative simpler methods such as pinpointing orbit-crossing collapsed hierarchical objects (PINOCCHIO; Monaco, Theuns & Taffoni 2002) or effective Zel'dovich approximation mocks (EZmocks; Chuang et al. 2015). The covariance matrix is then the sample variance of the power spectra from the different mocks (Manera et al. 2012; Taylor, Joachimi & Kitching 2013; Percival et al. 2014). The covariance matrix computed from the mocks will depend on the cosmological model that was used to generate them. It is computationally costly to produce mock catalogues for each possible cosmological

★E-mail: benedict.kalus+mnras@port.ac.uk
[1] www.euclid-ec.org
[2] www.skatelescope.org

Table 1. Notation used for probabilities.

Symbol	Name	Description
$\mathcal{Z}(\widehat{P})$	evidence	probability of the data $\widehat{P}$
$L(\widehat{P}\|P_{\rm H})$	likelihood	probability of the data $\widehat{P}$ given the hypothesis $P_{\rm H}$
$\mathcal{P}(P_{\rm H}\|\widehat{P})$	posterior	probability of the hypothesis P given the data $\widehat{P}$
$\Pi(P_{\rm H})$	prior	probability of the hypothesis

model and set of parameters to be tested, so one usually chooses a cosmological model which will produce a P(k) reasonably close to the measured one and uses the covariance matrix computed from the mocks created assuming that model. This approximation does not hold in general, especially at large scales. In this paper, we study other ways of approaching this problem, including using approximations to the true posterior distribution to obtain accurate inferences without requiring a covariance matrix for each cosmological model. We apply the most suitable of these approximation and the true distribution to provide a probability distribution function (PDF) for measurements of the non-Gaussianity parameter $f_{\rm NL}$. Our result will provide a complementary method to analysing $f_{\rm NL}$ directly from $\delta(x)$, as described in Verde et al. (2013).

We proceed as follows. In Section 2, we test the standard Gaussian posterior shapes mode-by-mode on a toy example where we measure the power spectrum itself. We do a similar test in Section 2.5 to study the impact of using different posterior distributions on a real survey, i.e. the Sloan Digital Sky Survey (SDSS)-III BOSS. In Section 3, we study alternative posterior shapes for power spectrum estimates inspired by cosmic microwave background (CMB) analyses. We test the most promising posterior distribution by postdicting a $f_{\rm NL}$-measurement for a data sample like the data release 9 (DR9) of BOSS and we make predictions of Euclid $f_{\rm NL}$-measurements in Section 4. We conclude in Section 5.

Throughout this work, we adopt a Bayesian framework and mark observed data with a hat, e.g. $\widehat{P}$ and quantities related to the hypothetical model with an H, e.g. $P_{\rm H}$. We denote probability distributions with different letters L, $\mathcal{P}$, Π and $\mathcal{Z}$, which we define in Table 1, to make clear whether they depend on data and/or the model. The ubiquitous Bayesian equation thus reads for this example

$$\mathcal{P}(P_{\rm H}|\widehat{P}) = \frac{L(\widehat{P}|P_{\rm H})\Pi(P_{\rm H})}{\mathcal{Z}(\widehat{P})}. \tag{1}$$

2 THE POWER SPECTRUM LIKELIHOOD

In this section, we elaborate the analytic likelihood and posterior functions of the galaxy clustering power spectrum assuming a Gaussian density field. We consider this posterior function as the 'truth' and compare it to commonly used approximations of the galaxy power spectrum posterior function for single modes, which we shall introduce in Section 2.3.

2.1 The True Distribution of $\widehat{|\delta_k|}$ Under the Assumption of a Gaussian Density Field

The positions of the galaxies in a survey can be transformed into a galaxy overdensity field

$$\delta(\boldsymbol{x}) \equiv \frac{n(\boldsymbol{x}) - \bar{n}(\boldsymbol{x})}{\bar{n}(\boldsymbol{x})}, \tag{2}$$

where $n(x)$ is the measured galaxy number density and $\bar{n}(\boldsymbol{x})$ the expected value. Fourier transforming $\delta(\boldsymbol{x})$ yields

$$\delta_k \equiv \frac{1}{V}\int \mathrm{d}^3\boldsymbol{x}\,\delta(\boldsymbol{x})\exp(i\boldsymbol{k}\boldsymbol{x}) \tag{3}$$

whose covariance matrix

$$\left\langle \delta_{k_1}\delta^*_{k_2}\right\rangle = \frac{(2\pi)^3}{V}\delta_D(\boldsymbol{k_1}-\boldsymbol{k_2})P(\boldsymbol{k_1}) \tag{4}$$

is given by the power spectrum $P(\boldsymbol{k})$. Following the standard assumption that δ_k forms a Gaussian random field, the probability of measuring a particular value of the real and imaginary parts $(\widehat{\delta_u}, \widehat{\delta_v})$ of a single $\widehat{\delta_k} = \widehat{\delta_u} + i\widehat{\delta_v}$ is a zero centred Gaussian distribution with standard deviation half the true power $\frac{1}{2}P_T(\boldsymbol{k})$:

$$\mathcal{Z}(\widehat{\delta_u}) = \frac{1}{\sqrt{\pi P_T(\boldsymbol{k})}}\exp\left(-\frac{\widehat{\delta_u}^2}{P_T(\boldsymbol{k})}\right),$$

$$\mathcal{Z}(\widehat{\delta_v}) = \frac{1}{\sqrt{\pi P_T(\boldsymbol{k})}}\exp\left(-\frac{\widehat{\delta_v}^2}{P_T(\boldsymbol{k})}\right). \tag{5}$$

We use the letter $\mathcal{Z}$ here, because we have assumed that the true power is known, i.e. the distribution only depends on the data (cf. Table 1). The distribution of the absolute value $\widehat{|\delta_k|} = \sqrt{\widehat{\delta_u}^2 + \widehat{\delta_v}^2}$ is given by a Rayleigh distribution:

$$\begin{aligned}\mathcal{Z}_R(\widehat{|\delta_k|}) &= \int \mathrm{d}\widehat{\delta_u}\int \mathrm{d}\widehat{\delta_v}\,\mathcal{Z}(\widehat{\delta_u})\mathcal{Z}(\widehat{\delta_v})\delta_D(\widehat{|\delta_k|} - \sqrt{\widehat{\delta_u}^2+\widehat{\delta_v}^2})\\ &= \frac{2\widehat{|\delta_k|}}{P_T(\boldsymbol{k})}\exp\left(-\frac{\widehat{|\delta_k|}^2}{P_T(\boldsymbol{k})}\right).\end{aligned} \tag{6}$$

Throughout this paper, we regard equation (6) as the 'true' distribution of $\widehat{|\delta_k|}$ to which we compare several approximations later.

Any model dependence enters the Rayleigh distribution only in the covariance of the density field, which is equal to the true power spectrum. The position of the distribution's peak equals the value of the true power. Measurements of $\widehat{\delta_k}$ have been used to make cosmological inferences when they have been further decomposed into spherical harmonics and spherical Bessel functions, because radial and angular modes can be distinguished, allowing an easy analysis of RSD. However, this method is rather complex and computationally expensive (Heavens & Taylor 1995; Percival et al. 2004). It is difficult to linearly compress $\widehat{\delta_k}$ efficiently maximally retaining information.

2.2 The posterior in terms of the power

We can rewrite the Rayleigh distribution in terms of the power. We replace $P_T(\boldsymbol{k})$ with $P_{\rm H}(\boldsymbol{k})$, and $\widehat{\delta_k}$ with $\sqrt{\widehat{P}(\boldsymbol{k})}$ in equation (6) which in this way depends on both data and model, and hence becomes a likelihood (cf. Table 1):

$$L_R(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k})) = \frac{2\sqrt{\widehat{P}(\boldsymbol{k})}}{P_{\rm H}(\boldsymbol{k})}\exp\left(-\frac{\widehat{P}(\boldsymbol{k})}{P_{\rm H}(\boldsymbol{k})}\right). \tag{7}$$

We can use Bayes' theorem (cf. equation 1) to find the posterior. It is standard to assume a uniform prior

$$\Pi\left(P_{\rm H}(\boldsymbol{k})\right) = \begin{cases}\frac{1}{P_{\max(k)}}, & \text{if } 0 \le P_{\rm H}(\boldsymbol{k}) \le P_{\max}(\boldsymbol{k}),\\ 0 & \text{otherwise,}\end{cases} \tag{8}$$

which requires an arbitrary choice of $P_{\rm max}(\boldsymbol{k})$. We assume that $P_{\rm max}(\boldsymbol{k})$ is far in the right tail of the likelihood such that $\frac{\Pi(P_{\rm H}(\boldsymbol{k}))}{\mathcal{Z}(\widehat{P}(\boldsymbol{k}))}$ is effectively constant and hence acts only as a normalization factor. Thus, for the 'true' posterior we have

$$\mathcal{P}_R(P_{\rm H}(\boldsymbol{k})|\widehat{P}(\boldsymbol{k})) = \frac{L_R(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k}))}{\int {\rm d}P_{\rm H}\, L_R(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k}))}$$
$$\propto \frac{2\sqrt{\widehat{P}(\boldsymbol{k})}}{P_{\rm H}(\boldsymbol{k})} \exp\left(-\frac{\widehat{P}(\boldsymbol{k})}{P_{\rm H}(\boldsymbol{k})}\right). \quad (9)$$

As $\widehat{P}$ is a constant in the posterior, one can rewrite equation (9) such that the log-posterior only depends on the ratio $\widehat{P}(\boldsymbol{k})/P_{\rm H}(\boldsymbol{k})$:

$$-2\ln(\mathcal{P}_R) = 2M \ln\left(\frac{P_{\rm H}(\boldsymbol{k})}{\widehat{P}(\boldsymbol{k})}\right) + 2M\frac{\widehat{P}(\boldsymbol{k})}{P_{\rm H}(\boldsymbol{k})} + \text{const.} \quad (10)$$

Then we follow the method of Hamimeche & Lewis (2008) and introduce

$$\gamma(x) \equiv \sqrt{-\ln(x) + x} \quad (11)$$

to make equation (10) look more quadratic:

$$-2\ln(\mathcal{P}_R) = 2M\left[\gamma\left(\frac{\widehat{P}(\boldsymbol{k})}{P_{\rm H}(\boldsymbol{k})}\right)\right]^2 + \text{const.} \quad (12)$$

We can also define

$$P_\gamma(\boldsymbol{k}) \equiv P_f(\boldsymbol{k})\gamma\left(\frac{\widehat{P}(\boldsymbol{k})}{P_{\rm H}(\boldsymbol{k})}\right) \quad (13)$$

for some fiducial model with power P_f. P_γ has then a symmetric Gaussian posterior with a fixed variance $\tilde{C}_k = \frac{2P_f^2(k)}{M}$ evaluated for our fiducial model

$$-2\ln(\mathcal{P}_R) = 4P_\gamma \tilde{C}_k^{-1} P_\gamma + \text{const.} \quad (14)$$

In general. things are more complicated than this simple picture. For example, the survey geometry leads to a convolution of δ_k, and non-linear effects distort the small scale mode distribution. Ideally, we would like to use a single distribution, and this should be matched to simulations (e.g. Blot et al. 2015). In order to broaden the choice, we also consider a number of forms for the likelihood inspired by CMB analyses.

2.3 Common Approximations of the Likelihood/Posterior of the Power Spectrum

Often, the power-spectrum is directly analysed, incorrectly assuming it follows a Gaussian distribution, thus the distribution of a finite empirical realization of the power spectrum $\widehat{P}(\boldsymbol{k})$ would read

$$\mathcal{Z}\left(\widehat{P}(\boldsymbol{k})\right) = \frac{\exp\left(-\frac{1}{2}\frac{\left[\widehat{P}(\boldsymbol{k})-P_T(\boldsymbol{k})\right]^2}{C_k}\right)}{\sqrt{2\pi C_k}}, \quad (15)$$

where $C_k \equiv \langle P_T^2(\boldsymbol{k})\rangle = \frac{2P_T^2(k)}{M}$ is the variance of the true power spectrum P_T at a bin centred around $\boldsymbol{k}$ comprising M independent modes. Note, that we assume that the widths and positions of the $\boldsymbol{k}$-bins are such that window effects are negligible (Feldman, Kaiser & Peacock 1994) and different modes are independent.

As in Section 2.2, we replace $P_T(\boldsymbol{k})$ with $P_{\rm H}(\boldsymbol{k})$ in equation (15) making it a likelihood (cf. Table 1)

$$L(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k})) = \frac{\exp\left(-\frac{1}{2}\frac{\left[\widehat{P}(\boldsymbol{k})-P_{\rm H}(\boldsymbol{k})\right]^2}{C_k^H}\right)}{\sqrt{2\pi C_k^H}}, \quad (16)$$

where $C_k^H \equiv \langle P_H^2(\boldsymbol{k})\rangle$ is the variance for the hypothetical power spectrum $P_{\rm H}(\boldsymbol{k})$.

However, in practice one chooses a fiducial model with power spectrum $\tilde{P}(\boldsymbol{k})$ and estimates the variance $\tilde{C}_k \equiv \langle \tilde{P}^2(\boldsymbol{k})\rangle$ for this particular choice:

$$L(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k}), \tilde{C}_k) = \frac{\exp\left(-\frac{1}{2}\frac{\left[\widehat{P}(\boldsymbol{k})-P_{\rm H}(\boldsymbol{k})\right]^2}{\tilde{C}_k}\right)}{\sqrt{2\pi\tilde{C}_k}}. \quad (17)$$

For mock based variance calculations, $\tilde{P}(\boldsymbol{k})$ is the cosmology of the mocks used in their analysis.

We can again use Bayes' theorem (cf. equation 1) and assume the same uniform prior as before to find the posterior. For the posterior assuming a Gaussian distribution in $\widehat{P}(\boldsymbol{k})$ with model-dependent covariance we have

$$\mathcal{P}_D(P_{\rm H}(\boldsymbol{k})|\widehat{P}(\boldsymbol{k})) = \frac{L(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k}))}{\int {\rm d}P_{\rm H} L(\widehat{P}(\boldsymbol{k})|P_{\rm H}(\boldsymbol{k}))}$$
$$\propto \frac{\exp\left(-\frac{1}{2}\frac{\left[\widehat{P}(\boldsymbol{k})-P_{\rm H}(\boldsymbol{k})\right]^2}{C_k^H}\right)}{\sqrt{2\pi C_k^H}}, \quad (18)$$

where we adopt the subscript notation $\mathcal{P}_D$ of Hamimeche & Lewis (2008). Note that both the exponential and the covariance matrix C_k^H depend on $P_{\rm H}(\boldsymbol{k})$.

If a fixed covariance is assumed, we have to apply the Bayesian equation (1) to equation (17) giving

$$\mathcal{P}_f(P_{\rm H}(\boldsymbol{k})|\widehat{P}(\boldsymbol{k}), \tilde{C}_k) \propto \frac{\exp\left(-\frac{1}{2}\frac{\left[\widehat{P}(\boldsymbol{k})-P_{\rm H}(\boldsymbol{k})\right]^2}{\tilde{C}_k}\right)}{\sqrt{2\pi\tilde{C}_k}}. \quad (19)$$

2.4 A simple test of the posterior shapes for the isotropically averaged power spectrum

In this subsection, we combine the single mode posterior functions to posterior functions of the band-power. We do not take any anisotropic effects, such as RSD, into account. This is conservative because the effective volume for higher multipole moments (cf. equation 25) is smaller, therefore containing fewer independent modes and thence amplifying the effect of choosing different posterior shapes.

In Gaussian cases, we suppose that our volume is large enough to accommodate M independent complex Gaussian distributed samples of δ_k such that we can use

$$C_{ab} = \frac{2}{M}\delta_D(\boldsymbol{k}_a - \boldsymbol{k}_b)P^2(\mathbf{k}_a). \quad (20)$$

to calculate the covariance matrices at higher numbers of modes M. We can obtain the band-power version of $\mathcal{P}_R(P_{\rm H}(\boldsymbol{k})||\widehat{\delta_k}|)$ by multiplying together the single mode expressions.

The three different posterior shapes of $P_{\rm H}$ are plotted in Fig. 1. In the top panel of Fig. 1, we plot single mode posterior distributions for which we adopt $|\widehat{\delta_k}| = 100$ and $P_T(\boldsymbol{k}) = \widehat{P}(\boldsymbol{k}) = |\widehat{\delta_k}|^2 = 10\,000$. Note, that a different choice would shift the peak positions and normalization factor, but preserve the shapes. We make two different choices for the fixed covariance to see the effect of making the wrong assumption. For the dotted red line, we choose the covariance matrix which corresponds to the true power spectrum $P_T(\boldsymbol{k})$, i.e. $\tilde{C}_k = 2P_T^2(\boldsymbol{k}) = 50\,000\,000$, and for the dashed-dotted line, we consider that our guess of the power spectrum is 5 per cent lower than the actual power spectrum, i.e. $\tilde{C}_k = 45\,125\,000$. The panels in the middle and at the bottom of Fig. 1 show the posterior distributions for 10 and 100 independent modes, respectively.

Figs 1 and 2 show that different choices of the covariance matrix provide very different posterior distributions for a small number of modes, but if we can increase the number of independent modes, we see the effect of the central limit theorem and the posterior distribution functions become more and more similar. We observe that the maximum of the fixed-covariance posterior always agrees with the true value, even if the wrong fiducial model has been chosen. However, if we choose the wrong covariance matrix, we over or underestimate the error of our measurements. If we do not fix the covariance, the best fit, i.e. the maximum of the posterior, has an offset with regard to the true value, which decreases as the number of modes increases. We also notice the long right tails of the varying-covariance Gaussian and the posterior measured from $|\widehat{\delta_k}|$. The logarithmic plot in Fig. 2 shows that the tails of all approximations disagree with the true posterior distribution. However, $\mathcal{P}_D$ is closest to the truth.

2.5 Application to a real survey

We have seen that a Gaussian distribution for $P_{\rm H}(\boldsymbol{k})$ is not a good approximation to the true Rayleigh distribution if the number of modes is small. In this section, we study whether this has an impact on a real survey. We will base our analysis on an analytic linear error for the power spectrum and errors, but use survey parameters for the data release 11 (DR11) of BOSS. For a real survey, we have to take into account that the discrete positions of the galaxies in a given survey are sampled from a continuous random field by a Poisson point process (Feldman et al. 1994). To take this sampling process into account,equation (4) becomes

$$\left\langle \delta_{k_1} \delta^*_{k_2} \right\rangle = \frac{(2\pi)^3}{V} \delta_D(\boldsymbol{k_1} - \boldsymbol{k_2}) \left[P(\boldsymbol{k_1}) + \bar{n}^{-1} \right], \tag{21}$$

and hence also

$$\mathcal{P}(P_{\rm H}(\boldsymbol{k}) | |\widehat{\delta_k}|) \propto \frac{|\widehat{\delta_k}| \exp\left(-\frac{|\widehat{\delta_k}|^2}{P_{\rm H}(\boldsymbol{k}) + \bar{n}^{-1}} \right)}{P_{\rm H}(\boldsymbol{k}) + \bar{n}^{-1}}. \tag{22}$$

The average number density $\bar{n} = 2 \times 10^{-4}\ \frac{h^3}{\rm Mpc^3}$ can be calculated from the number of galaxies contained in the BOSS DR11 CMASS sample (690,826) and its survey volume $V_{\rm S} = 10\,{\rm Gpc}^3$ (Anderson et al. 2013) assuming $h = 0.7$. For the covariance matrices of the Gaussians, we need to know the number of modes (Feldman et al. 1994; Tegmark 1997)

$$M = V_n V_{\rm eff}(\boldsymbol{k}), \tag{23}$$

where

$$V_n \equiv \frac{k_n^2 \triangle k_n}{2\pi^2} \tag{24}$$

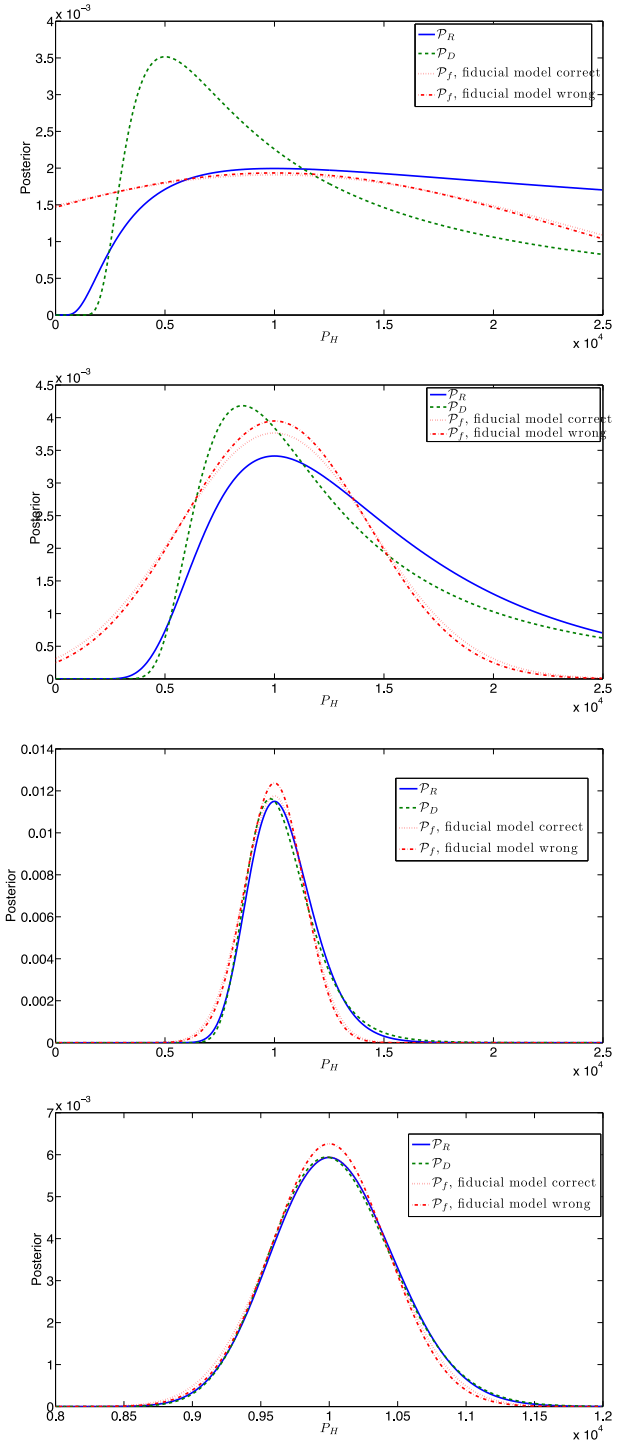

Figure 1. Comparison of different posterior distribution functions for 1, 10, 100 and 1000 independent modes (from top to bottom). The blue line represents the product of single Rayleigh distributed modes (true posterior distribution) and some of the approximations, such as the Gaussian posterior distribution with a model-dependent covariance (green), and the Gaussian posterior where the covariance is estimated for a fixed fiducial model (red). The posterior takes the form of the dotted red line if the fiducial and the true power spectra agree, the dashed-dotted line shows the effect of choosing a fiducial model of which the power spectrum is wrong by 5 per cent.

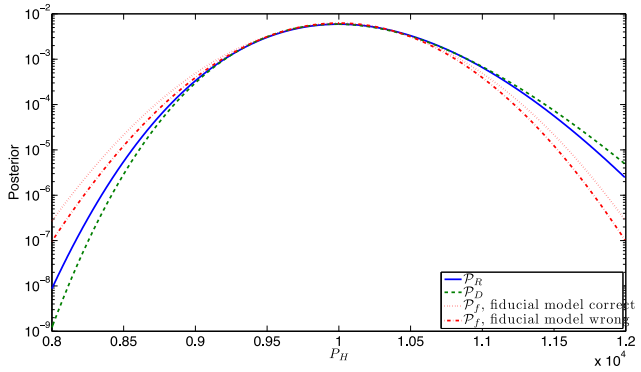

Figure 2. Same as the bottom panel of Fig. 1, but with a logarithmic ordinate.

is the k-space-'volume' of the nth k-bin centred at k_n with width $\triangle k_n$, and

$$V_{\rm eff}(\boldsymbol{k}) \equiv V_S \left[\frac{\bar{n}\,P(\boldsymbol{k})}{1+\bar{n}\,P(\boldsymbol{k})}\right]^2 \tag{25}$$

is the effective volume. Anderson et al. (2013) calculate the power spectrum in Fourier modes averaged over bin widths of $\triangle k = 0.008h\,{\rm Mpc}^{-1}$. The values of the k-bin centres and their corresponding number of modes M are M=18, 180 and 500 in the three lowest k-bins centred at $k = 0.004$, 0.012 and $0.02\,{\rm Mpc}\,h^{-1}$. We model the measured power spectrum as $\widehat{P}(\boldsymbol{k}) = b^2 P_{\rm lin}(\boldsymbol{k})$, where $b = 1.87$ is the large-scale bias and $P_{\rm lin}(\boldsymbol{k})$ is a linear power spectrum produced by CAMB (Lewis, Challinor & Lasenby 2000). For the other measurement we take $|\widehat{\delta_k}| = \sqrt{\widehat{P}(\boldsymbol{k}) + \bar{n}^{-1}}$. The resulting posterior distributions for the three lowest k-bins are plotted in Fig. 3. At the largest scales, i.e. $k = 0.004h\,{\rm Mpc}^{-1}$, neither $\mathcal{P}_D$ or $\mathcal{P}_f$ match $\mathcal{P}_R$. At $k = 0.012h\,{\rm Mpc}^{-1}$ and $k = 0.02h\,{\rm Mpc}^{-1}$ $\mathcal{P}_f$ and $\mathcal{P}_D$ become more similar, but neither of them features the asymmetric shape of $\mathcal{P}_R$. Additionally, $\mathcal{P}_f$ and $\mathcal{P}_D$ produce smaller error bars compared to $\mathcal{P}_R$. We can also numerically compare the distributions if we introduce the Kullback–Leibler (KL) divergence (Kullback & Leibler 1951). A distribution $\mathcal{P}_1$ is 'better' than $\mathcal{P}_2$, if the loss of information due to approximating the true distribution with $\mathcal{P}_1$ is less than the same loss caused by using $\mathcal{P}_2$ as an approximation. If we use a probability density function (PDF) g to approximate another PDF f, a measure of the loss of information is given by the KL divergence

$$D_{\rm KL}\,(g||f) \equiv \int_{-\infty}^{\infty} {\rm d}x f(x) \ln\left(\frac{f(x)}{g(x)}\right). \tag{26}$$

The KL divergences given in Table 2 tell us the same story as Fig. 3. The KL divergences of the Gaussian approximation with a varying covariance $\mathcal{P}_D$ is at all scales less than the KL divergence of $\mathcal{P}_f$, i.e. $\mathcal{P}_D$ is a better approximation to the true $\mathcal{P}_R$. On the downside, its best fit has an offset with respect to $\mathcal{P}_R$. We will therefore investigate alternative posterior shapes in the next section.

3 STUDYING ALTERNATIVE POSTERIOR SHAPES

We have seen in the previous sections that the true posterior distribution $\mathcal{P}_R$ is not well approximated by either $\mathcal{P}_f$ or $\mathcal{P}_D$ if the number of independent modes is low, which is the case at large scales, i.e. small values of k. A similar problem arises when cosmological models are fitted to CMB power spectra, which are Wishart distributed. Bond, Jaffe & Knox (2000), Smith, Challinor & Rocha (2006), Percival & Brown (2006) and Hamimeche & Lewis (2008) have studied alternative distribution shapes that approximate the Wishart distribution. We take a similar approach to Verde et al.

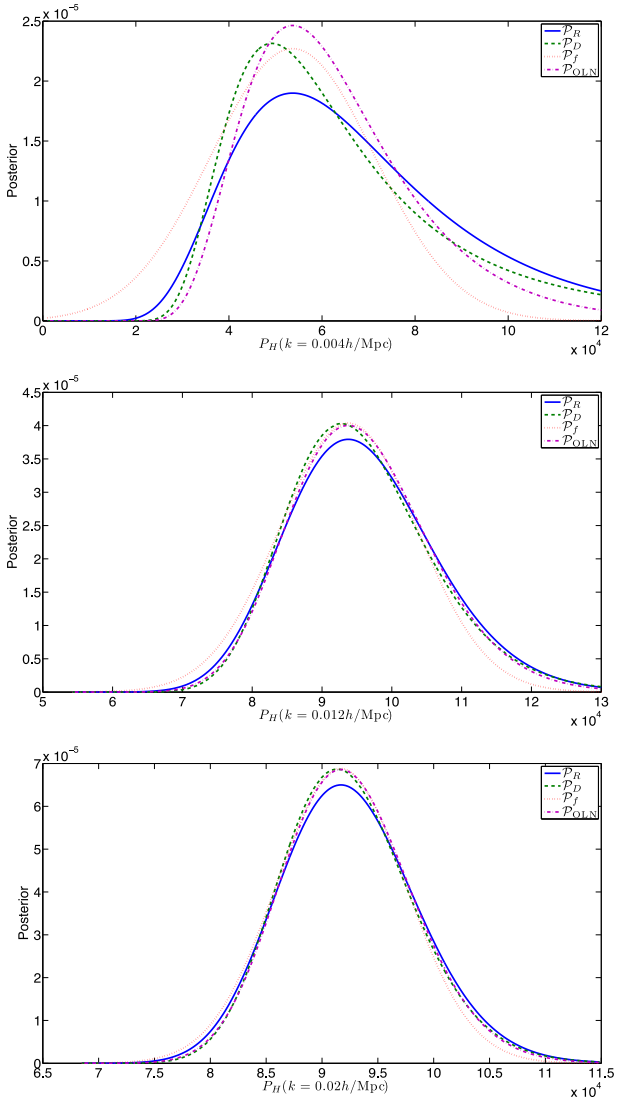

Figure 3. Posterior distribution functions of the hypothetical power spectrum $P_{\rm H}(\boldsymbol{k})$ for the three lowest $\boldsymbol{k}$-bins of BOSS DR11 CMASS. The colour coding is the same as in Fig. 1, with the addition of the OLN posterior distribution plotted in magenta.

Table 2. Kullback–Leibler divergences of the different approximations with respect to the true $\mathcal{P}_R$ at different scales k_n for BOSS DR11 CMASS.

k_n $\frac{\rm Mpc}{h}$	$D_{\rm KL}(\mathcal{P}_D\|\|\mathcal{P}_R)$	$D_{\rm KL}(\mathcal{P}_f\|\|\mathcal{P}_R)$	$D_{\rm KL}(\mathcal{P}_f^{\rm wrong}\|\|\mathcal{P}_R)$
0.004	0.0213 221	0.382 926	0.335 588
0.012	0.00 451 188	0.0341 685	0.0374 833
0.02	0.00 336 263	0.0138 535	0.0199 032

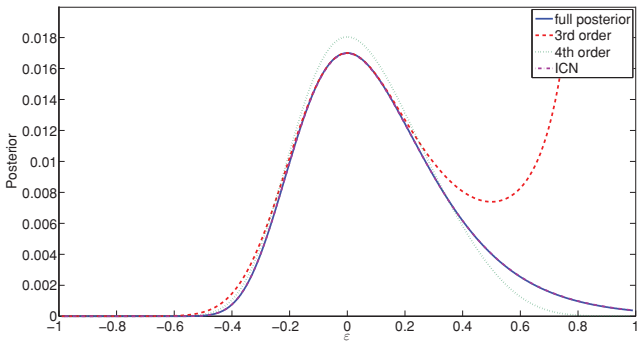

Figure 4. Third and fourth order Taylor expansion to the true posterior shape $\mathcal{P}_R$ with $M = 20$ modes. The x-axis is a perturbation $\varepsilon \equiv P_{\rm H}/\widehat{P} - 1$ of the model power spectrum $P_{\rm H}$ around the average recovered best-fitting value $\widehat{P}$. As the third order approximation is not normalizable, the normalization has been chosen such that it agrees with the 4th order at the maximum. The true posterior shape agrees very well with the ICN posterior shape.

(2003) and Percival & Brown (2006) and expand the natural logarithm of equation (9) around the maximum $P_{\rm H}(\boldsymbol{k}) \equiv (1+\varepsilon)|\widehat{\delta_{\boldsymbol{k}}}|^2$:[3]

$$-2\ln(\mathcal{P}_R) = 2M\left(\frac{\varepsilon^2}{2} - \frac{2\varepsilon^3}{3} + \frac{3\varepsilon^4}{4} + \mathcal{O}\left(\varepsilon^5\right)\right) + \text{const.} \quad (27)$$

This equation agrees to third order with the Taylor expansions of the logarithms of the following distributions:

(i) the inverse cubic normal (ICN) distribution (Smith et al. 2006)

$$-2\ln(\mathcal{P}_{\rm ICN}) = 18\tilde{C}_{\boldsymbol{k}}^{-1}\left[\widehat{P}(\boldsymbol{k}) - \widehat{P}(\boldsymbol{k})^{4/3}P_{\rm H}(\boldsymbol{k})^{-1/3}\right]^2, \quad (28)$$

(ii) the offset lognormal (OLN) distribution

$$-2\ln(\mathcal{P}_{\rm OLN}) = 2(1+a)\tilde{C}_{\boldsymbol{k}}^{-1}\left[\widehat{P}(\boldsymbol{k})\ln\left(\frac{P_{\rm H}(\boldsymbol{k}) + a\widehat{P}(\boldsymbol{k})}{\widehat{P}(\boldsymbol{k}) + a\widehat{P}(\boldsymbol{k})}\right)\right]^2 \quad (29)$$

if $a = -1/4$,

(iii) and combinations of any of the distributions given in chapter 5.1 of Percival & Brown (2006).

We can see from Fig. 4 that the 3rd order diverges for large values of the model power spectrum $P_{\rm H}$. Hence the optimal free parameter a might differ from $a = -1/4$. Therefore, we use the KL divergence to optimize a in the OLN distribution $\mathcal{P}_{\rm OLN}$. It can be found to be $a = -0.201$ at $k = 0.004\frac{\rm Mpc}{h}$, $a = -0.240$ at $k = 0.012\frac{\rm Mpc}{h}$ and $a = -0.242$ at higher values of k. $\mathcal{P}_{\rm OLN}$ peaks at the maximum of the true distribution $\mathcal{P}_R$ and it approximates the tails of the true distribution a bit better than the Gaussian approximations, but as Fig. 3 shows, it is still obviously different from $\mathcal{P}_R$.

The ICN distribution (Smith et al. 2006) fits the true distribution better. Fig. 4 shows a remarkable agreement between $\mathcal{P}_R$ and $\mathcal{P}_{\rm ICN}$. Writing both $-2\ln(\mathcal{P}_R)$ and $-2\ln(\mathcal{P}_{\rm ICN})$ as Taylor series, we see that their Taylor coefficients are equal for $k \leq 3$ and approximately equal for much higher orders (cf. Appendix A).

[3] For realistic, noisy measurements of $|\widehat{\delta_{\boldsymbol{k}}}|$ and $\widehat{P}(\boldsymbol{k})$, $P_{\rm H}(\boldsymbol{k})$ has to be replaced by $P_{\rm H}(\boldsymbol{k}) + \bar{n}^{-1}$ everywhere in this section. For simplicity, we do not write the noise explicitly.

4 THE EFFECT ON $f_{\rm NL}$ MEASUREMENTS

4.1 Physical model

In this section, we test the effect of using different posterior distribution shapes on the inference of a real observable. The largest deviations between the posteriors are at small k and we would therefore expect the largest effects for parameters dependent on these modes. At these scales, (local) primordial non-Gaussianity alters the biasing law between dark-matter haloes and the underlying mass–density field (Afshordi & Tolley 2008; Dalal et al. 2008; Matarrese & Verde 2008; Slosar et al. 2008; Giannantonio & Porciani 2010; Schmidt & Kamionkowski 2010; Valageas 2010; Desjacques, Jeong & Schmidt 2011), making $f_{\rm NL}$ a perfect test parameter of our analysis. The parameter arises in models where the potential has a local quadratic term

$$\Phi = \phi + f_{\rm NL}\left(\phi^2 - \left\langle\phi^2\right\rangle\right). \quad (30)$$

The resulting alteration of the bias can be written as

$$b(k, f_{\rm NL}) = b_0 + \delta b(f_{\rm NL}) + \Delta b(k, f_{\rm NL}), \quad (31)$$

where b_0 is the bias in a universe without primordial non-Gaussianity, $\delta b(f_{\rm NL})$ is the scale-independent modification to the bias from the non-Gaussian form of the mass functions and (Schmidt & Kamionkowski 2010; Desjacques et al. 2011)

$$\Delta b(k, f_{\rm NL}) \approx (b_0 - 1) f_{\rm NL} A(k) \quad (32)$$

is the local scale-dependent correction due to the easier halo formation with additional long-wavelength fluctuations, which depends on the critical density $\delta_c(z)$ in the peak- background split model, as well as the matter transfer function $T(k)$, the matter density $\Omega_{\rm m}$, the present-time Hubble parameter H_0 and the linear growth function $D(z)$ through the parameter

$$A(k, z) = \frac{3\Omega_{\rm m}\delta_c(z)}{k^2 T(k)}\left(\frac{H_0}{c}\right)^2. \quad (33)$$

As $\delta b(f_{\rm NL}) \ll \Delta b(k, f_{\rm NL})$ at our scales of interest (Slosar et al. 2008; Giannantonio & Porciani 2010), we neglect $\delta b(f_{\rm NL})$. Fig. 5 shows the effect of $f_{\rm NL}$ on the galaxy power spectrum at large scales. We plot the galaxy power spectrum $P_{\rm g}$ divided by the galaxy power spectrum at $f_{\rm NL} = 0$, hence what we plot is proportional to the square of equation (31). At lowest $\boldsymbol{k}$, negative $f_{\rm NL}$ enhances the power spectrum due to the fact that the term proportional to $f_{\rm NL}^2$ dominates the total bias. At slightly higher $\boldsymbol{k}$, but still at large

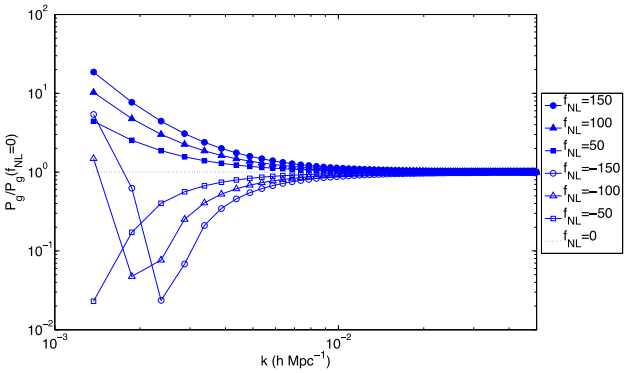

Figure 5. Galaxy power spectra $P_{\rm g}$ calculated for different values of $f_{\rm NL}$ divided by the galaxy power spectrum $P_{\rm g}(f_{NL} = 0)$ of a universe with a Gaussian primordial density field.

scales, the term linear in $f_{\rm NL}$ dominates, and the power is enhanced or decreased depending on the sign of $f_{\rm NL}$. At small, yet still linear, scales, $A(k, z)$ becomes small, thus initial local non-Gaussianities do not have an effect on the galaxy power spectrum at these scales.

Here, we work to first order in δ, so that we can continue to assume that δ_k is drawn from a Gaussian distribution, with an altered $P(\boldsymbol{k})$, i.e. the first order effect of non-Gaussianity is to $P_{\rm H}(\boldsymbol{k})$, keeping the distribution the same. Furthermore, we do not alter V_n (equation 24) to include any coupling between modes from the non-Gaussian signal. Where k is very small, higher order corrections to δ will become increasingly important (e.g. Tellarini et al. 2015), suggesting that the Gaussian limit for δ will break down here.

4.2 Boss results

We use BOSS DR9 parameters and the same CAMB linear matter power spectrum as Ross et al. (2013). We also assume $\delta_c = \frac{1.686}{D(z)}$ as expected from the spherical collapse model in an Einstein–de Sitter universe and a flat prior for $f_{\rm NL}$. We plot $f_{\rm NL}$ posterior functions in Fig. 6, assuming a measurement of a power spectrum with underlying $f_{\rm NL} = 0$. $\mathcal{P}_f$ is not symmetric, as both a linear and a quadratic term of $f_{\rm NL}$ enter the power spectrum. The ICN distribution agrees again very well with $\mathcal{P}_R$. $\mathcal{P}_R$, $\mathcal{P}_f$ and $\mathcal{P}_{\rm ICN}$ reproduce the true value as their best-fitting estimate. Using $\mathcal{P}_D$, the most likely value of $f_{\rm NL}$ is $f_{\rm NL} = -25.5$ considering the same k-bins as Ross et al. (2013) in their analysis of DR9 BOSS data, i.e. $0.004\frac{h}{\rm Mpc} \leq k \leq 0.05\frac{h}{\rm Mpc}$.

One has to keep in mind that there are different definitions of the measured value. The commonly published value is the posterior mean $\langle f_{\rm NL}\rangle$, due to the fact that if $f_{\rm NL}$ is fitted as part of a longer list of cosmological parameters, one has to rely on Markov chain Monte Carlo techniques (e.g. Lewis & Bridle 2002). In general, such techniques cannot provide accurate estimates of the best-fitting value. Hence, data analysis papers more often present $\langle f_{\rm NL}\rangle$ as their results. If the posterior is asymmetric, the best fit and posterior mean do not agree. Given a flat $f_{\rm NL}$-prior, we expect $f_{\rm NL} = 11.4$ using $\mathcal{P}_R$. Based on our arguments in Sections 2.1 and 4.1, we think of the mean of $\mathcal{P}_R$ as the correct estimate of $f_{\rm NL}$. This seems counter-intuitive because our input was that we measure a power spectrum which corresponds to $f_{\rm NL} = 0$, but we have to consider that $\hat{P}(\boldsymbol{k})$ is a finite empirical realization in our part of the Universe corresponding to the value of $f_{\rm NL} = 0$ we have assumed we would measure locally, but due to the non-Gaussian shape of the posterior distribution, the ensemble average of $f_{\rm NL}$ measured in other parts of the Universe is higher than the value we set as an input for our local environment.

Our results are summarized in Table 3. $\mathcal{P}_{\rm ICN}$ reproduces the correct estimate of $f_{\rm NL}$, whereas $\mathcal{P}_D$ and $\mathcal{P}_f$ estimate $f_{\rm NL} = -11.9$ and $f_{\rm NL} = -7.7$, respectively. The choice of the posterior distribution also affects the error estimation. If we use $\mathcal{P}_R$ or $\mathcal{P}_{\rm ICN}$, the length of our postdicted 95 per cent $f_{\rm NL}$-confidence interval (CI, cf. Table 3) is similar to the length of Ross et al. (2013)'s most naïve case ii 95 per cent CI, i.e. $32 < f_{\rm NL} < 198$.

Table 3. $f_{\rm NL}$-postdictions of the best-fitting $f_{\rm NL}^{\rm (BF)}$ and marginalized best fits $\langle f_{\rm NL}\rangle$, as well as its 95 per cent confidence interval, for BOSS DR9 using different shapes of the posterior distribution.

Posterior	$f_{\rm NL}^{\rm (BF)}$	$\langle f_{\rm NL}\rangle$	95 per cent Confidence interval
$\mathcal{P}_R$	0	11.4	$-71.5 < f_{\rm NL} < 100.7$
$\mathcal{P}_D$	−25.5	−11.9	$-68.2 < f_{\rm NL} < 53.4$
$\mathcal{P}_f$	0	−7.7	$-90.9 < f_{\rm NL} < 71.0$
$\mathcal{P}_{\rm ICN}$	0	11.4	$-71.9 < f_{\rm NL} < 101.2$

4.3 Euclid results

We make similar predictions for the Euclid survey (Laureijs et al. 2011). We assume bias values $b(z) = \sqrt{1+z}$, matched to simulations of Orsi et al. (2010) and also assumed in Amendola et al. (2013), and number densities $\bar{n}(z)$ predicted for Euclid by Pozzetti, Hirata & Geach (in preparation) and a survey covering 15 000 deg^2. We generate CAMB matter power spectra $P(\boldsymbol{k}, z)$ for the redshift range $0.9 < z < 1.74$. Note, that the aim of this paper is to test how the use of different posterior shapes influences cosmological measurements, but not primarily to make $f_{\rm NL}$-predictions. We refer to more rigorous predictions which can be found e.g. in Fedeli et al. (2011), Laureijs et al. (2011), Giannantonio et al. (2012), Yamauchi, Takahashi & Oguri (2014). These studies also include 3-point statistics, weak lensing tomography, measurements of the integrated Sachs-Wolfe effect and/or the use of the multitracer technique. Their constraints are therefore tighter than ours.

As Euclid will probe a much larger volume, it will accommodate many more k-modes and hence we see good agreement of $\mathcal{P}_f$ with $\mathcal{P}_R$ in Fig. 7. As against our results in Section 2.5, fixing the covariance provides better $f_{\rm NL}$ results than the inferences from a posterior with varying covariance. However, $\mathcal{P}_{\rm ICN}$ is still the best approximation and accurately reproduces the marginalized $f_{\rm NL}$-value of $\mathcal{P}_R$ and its 95 per cent CI, whereas using $\mathcal{P}_f$ yields the correct width of the 95 per cent CI, but its position and the marginalized value have an offset of 0.38 (cf. Table 4). We therefore still recommend either using $\mathcal{P}_{\rm ICN}$ or $\mathcal{P}_R$ when cosmological models are fitted to power spectra from galaxy surveys even as large as Euclid.

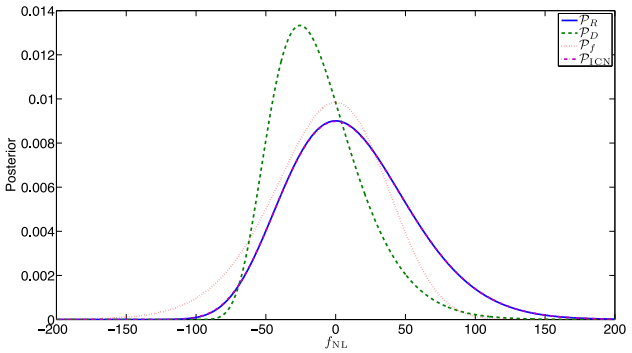

Figure 6. Analytic $f_{\rm NL}$-posterior functions for a BOSS like survey combining all k-bins.

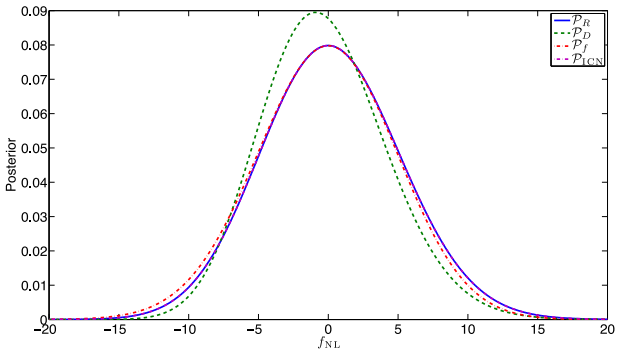

Figure 7. Analytic $f_{\rm NL}$-posterior functions for an Euclid like survey combining all k-bins.

Table 4. $f_{\rm NL}$-predictions similar to Table 3, but for Euclid.

Posterior	$f_{\rm NL}^{\rm (BF)}$	$\langle f_{\rm NL}\rangle$	95 per cent Confidence interval
$\mathcal{P}_R$	0	0.24	$-9.0<f_{\rm NL}<9.4$
$\mathcal{P}_D$	−1.0	−0.30	$-8.4<f_{\rm NL}<7.8$
$\mathcal{P}_f$	0	−0.14	$-9.4<f_{\rm NL}<9.0$
$\mathcal{P}_{\rm ICN}$	0	0.24	$-9.0<f_{\rm NL}<9.4$

5 CONCLUSIONS

We have studied different posterior shapes that can be used in the fitting process of cosmological models to power spectra from galaxy surveys. As the underlying matter density field is at least approximately Gaussian, we assume that the true posterior distribution $\mathcal{P}_R$ is based on a Rayleigh likelihood distribution in δ. Assuming Gaussian posteriors in $P(\boldsymbol{k})$, be it with a fixed or a varying covariance matrix, does not approximate $\mathcal{P}_R$ well and yields biased best-fitting values and wrong error estimates especially on large scales where statistics are not good enough to make use of the central limit theorem.

If one confines oneself to use Gaussian posterior shapes, it depends on the parameter one wants to constrain whether a fixed or varying covariance matrix provides more accurate results. We found that the posterior shape $\mathcal{P}_D$ with varying covariance follows $\mathcal{P}_R$ closer than $\mathcal{P}_f$ with a fixed covariance when the power spectrum $P_{\rm H}$ (or any parameter linear in the power spectrum) is fitted to the power spectrum $\widehat{P}$, but when $f_{\rm NL}$ is fitted to $\widehat{P}$ it is the other way round.

Due to these reasons, we advise against using Gaussian posterior distributions. Instead, we have found that posterior distributions, such as the ICN distribution $\mathcal{P}_{\rm ICN}$ (cf. equation 28) or applying Hamimeche & Lewis (2008) method to $\mathcal{P}_R$ (cf. equation 14), provide simple, more accurate alternatives. They confidently reproduce the correct width of the 95 per cent confidence intervals in our simplified predictions of $f_{\rm NL}$-measurements. However, the final decision about which posterior is the best to use should be done after testing these methods against simulations which account for the non-linear effects that we have ignored for simplicity in our analytic calculations. We leave this for future work.

A major advantage of the non-Gaussian posteriors presented in this paper, is the fact that their covariance matrices do not depend on the power spectrum of the model to be tested. The estimation of covariance matrices is a critical and computationally expensive step in the data analysis. Extensions to configuration-space analyses based on the correlation function $\xi(\boldsymbol{r})$ are left for future work.

ACKNOWLEDGEMENTS

The authors would like to thank Matteo Tellarini, Ashley Ross and David Wands for valuable discussions about primordial non-Gaussianity. We thank the referee Andrew Jaffe for his helpful comments.

Some of the results in this paper have been generated using the CAMB package (Lewis et al. 2000). For some of the other results, we made use of the facilities and staff of the UK Sciama High Performance Computing cluster supported by the ICG, SEPNet and the University of Portsmouth.

WJP acknowledges support from UK STFC through the consolidated grant ST/K0090X/1, and from the European Research Council through the Darksurvey grant. LS is grateful for support from SNSF grant SCOPES IZ73Z0-152581, GNSF grant FR/339/6-350/14, and DOE grant DEFG 03-99EP41093.

REFERENCES

Afshordi N., Tolley A. J., 2008, Phys. Rev. D, 78, 123507
Ahn C. P. et al., 2012, ApJS, 203, 21
Amendola L. et al. (Euclid Theory Working Group Collaboration), 2013, Living Rev. Relativ., 16, 6
Anderson L. et al. (BOSS Collaboration), 2013, MNRAS, 441, 24
Blot L., Corasaniti P. S., Alimi J. M., Reverdy V., Rasera Y., 2015, MNRAS, 446, 1756
Bond J. R., Jaffe A. H., Knox L. E., 2000, ApJ, 533, 19
Chuang C. H., Kitaura F. S., Prada F., Zhao C., Yepes G., 2015, MNRAS, 446, 2621
Dalal N., Dore O., Huterer D., Shirokov A., 2008, Phys. Rev. D, 77, 123514
Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, AJ, 292, 371
Dawson K. S. et al. [BOSS Collaboration], 2013, AJ, 145, 10
Desjacques V., Jeong D., Schmidt F. Phys. Rev. D, 84, 061301
Fedeli C., Carbone C., Moscardini L., Cimatti A., 2011, MNRAS, 414, 1545
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Giannantonio T., Porciani C., 2010, Phys. Rev. D, 81, 063530
Giannantonio T., Porciani C., Carron J., Amara A., Pillepich A., 2012, MNRAS, 422, 2854
Hamimeche S., Lewis A., 2008, Phys. Rev. D, 77, 103013
Heavens A. F., Taylor A. N., 1995, MNRAS, 275, 483
Kullback S., Leibler R. A., 1951, Ann. Math. Stat., 22, 79
Laureijs R. et al. [EUCLID Collaboration], 2011, preprint (arXiv:1110.3193)
Lewis A., Bridle S., 2002, Phys. Rev. D, 66, 103511
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Manera M. et al., 2012, MNRAS, 428, 1036
Matarrese S., Verde L., 2008, ApJ, 677, L77
Monaco P., Theuns T., Taffoni G., 2002, MNRAS, 331, 587
Norberg P., Baugh C. M., Gaztanaga E., Croton D. J., 2009, MNRAS, 396, 19
Orsi A., Baugh C. M., Lacey C. G., Cimatti A., Wang Y., Zamorani G., 2010, MNRAS, 405, 1006
Percival W. J. et al. [2dFGRS Collaboration], 2004, MNRAS, 353, 1201
Percival W. J., Brown M. L., 2006, MNRAS, 372, 1104
Percival W. J. et al., 2014, MNRAS, 439, 2531
Ross A. J. et al., 2013, MNRAS, 428, 1116
Schlegel D. et al. [BigBoss Collaboration], 2011, preprint (arXiv:1106.1706)
Schmidt F., Kamionkowski M., 2010, Phys. Rev. D, 82, 103002
Scoccimarro R., Sheth R. K., 2002, MNRAS, 329, 629
Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, JCAP, 0808, 031
Smith S., Challinor A., Rocha G., 2006, Phys. Rev. D, 73, 023517
Tassev S., Zaldarriaga M., Eisenstein D., 2013, JCAP, 1306, 036
Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928
Tegmark M., 1997, Phys. Rev. Lett., 79, 3806
Tellarini M., Ross A. J., Tasinato G., Wands D., 2015, J. Cosmol. Astropart. Astrophys., 07, 004
Valageas P., 2010, A&A, 514, A46
Verde L. et al. [WMAP Collaboration], 2003, ApJS, 148, 195
Verde L., Jimenez R., Alvarez-Gaume L., Heavens A. F., Matarrese S., 2013, J. Cosmol. Astropart. Astrophys., 1306, 023
Yamauchi D., Takahashi K., Oguri M., 2014, Phys. Rev. D, 90, 083520

APPENDIX A: COMPARISON OF $\mathcal{P}_R$ AND $\mathcal{P}_{\rm ICN}$ TAYLOR SERIES

In this appendix, we compare the Taylor Series of $\mathcal{P}_R$ and $\mathcal{P}_{\rm ICN}$ to explain why they are so similar. We write the hypothetical power

spectrum $P_{\rm H} \equiv (1+\varepsilon)|\widehat{\delta_k}|^2$ as a perturbation around the measured power. The Rayleigh posterior hence becomes

$$-2\ln(\mathcal{P}_R) = 2\ln(1+\varepsilon) + \frac{2}{1+\varepsilon}. \tag{A1}$$

Ignoring the irrelevant zero order contribution, the Taylor series reads

$$-2\ln(\mathcal{P}_R) = 2\sum_{\kappa=1}^{\infty}(-1)^{\kappa}\varepsilon^{\kappa}\frac{\kappa-1}{\kappa}. \tag{A2}$$

The ICN distribution in terms of ε is given by

$$\begin{aligned} -2\ln(\mathcal{P}_{\rm ICN}) &= 9\left[1-(1+\varepsilon)^{-1/3}\right]^2 \\ &= 9\left[1-2(1+\varepsilon)^{-1/3}+(1+\varepsilon)^{-2/3}\right]. \end{aligned} \tag{A3}$$

We make use of the generalized binomial series $(1+\varepsilon)^{\alpha} = \sum_{\kappa=0}^{\infty}\binom{\alpha}{\kappa}\varepsilon^{\kappa}$, where $\binom{\alpha}{\kappa} \equiv \frac{\Gamma(\alpha+1)}{\Gamma(\kappa+1)\Gamma(\alpha-\kappa+1}$ is the generalized binomial coefficient, and obtain the series

$$-2\ln(\mathcal{P}_{\rm ICN}) = 9\sum_{\kappa=1}^{\infty}\varepsilon^{\kappa}\left[\binom{-\frac{2}{3}}{\kappa} - 2\binom{-\frac{1}{3}}{\kappa}\right]. \tag{A4}$$

Again, we have ignored irrelevant constant terms. The negative entries in the binomial coefficients can be removed using $\binom{\alpha}{\kappa} = (-1)^{\kappa}\binom{\kappa-\alpha-1}{\kappa}$:

$$-2\ln(\mathcal{P}_{\rm ICN}) = 9\sum_{\kappa=1}^{\infty}(-1)^{\kappa}\varepsilon^{\kappa}\left[\binom{\kappa-\frac{1}{3}}{\kappa} - 2\binom{\kappa-\frac{2}{3}}{\kappa}\right]. \tag{A5}$$

If we insert values for $\kappa \leq 3$, we find the equality

$$2\frac{\kappa-1}{\kappa} = 9\left[\binom{\kappa-\frac{1}{3}}{\kappa} - 2\binom{\kappa-\frac{2}{3}}{\kappa}\right]. \tag{A6}$$

Thus, $\mathcal{P}_R$ and $\mathcal{P}_{\rm ICN}$ are the same to third order. What is even more striking is that for larger κ,the approximation

$$2\frac{\kappa-1}{\kappa} \approx 9\left[\binom{\kappa-\frac{1}{3}}{\kappa} - 2\binom{\kappa-\frac{2}{3}}{\kappa}\right] \tag{A7}$$

still holds. For $\kappa < 17$, the two sides differ by less than 20 per cent. Therefore, the agreement between $\mathcal{P}_R$ and $\mathcal{P}_{\rm ICN}$ is high.

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