LIKELIHOOD OF THE POWER SPECTRUM IN COSMOLOGICAL PARAMETER ESTIMATION

lei sun, qiao wang, and hu zhan

key laboratory of optical astronomy, national astronomical observatories, chinese academy of sciences,
20a datun road, chaoyang district, beijing 100012, china; sunl@nao.cas.cn

received 2013 june 21; accepted 2013 september 16; published 2013 october 17

abstract

The likelihood function is a crucial element of parameter estimation. In analyses of galaxy overdensities and weak lensing shear, one often approximates the likelihood of the power spectrum with a Gaussian distribution. The posterior probability derived from such a likelihood deviates considerably from the exact posterior on the largest scales probed by any survey, where the central limit theorem does not apply. We show that various forms of Gaussian likelihoods can have a significant impact on the estimation of the primordial non-Gaussianity parameter $f_{NL}$ from the galaxy angular power spectrum. The Gaussian plus log-normal likelihood, which has been applied successfully in analyses of the cosmic microwave background, outperforms the Gaussian likelihoods. Nevertheless, even if the exact likelihood of the power spectrum is used, the estimated parameters may be still biased. As such, the likelihoods and estimators need to be thoroughly examined for potential systematic errors.

key words: cosmology: observations – cosmology: theory – methods: statistical

online-only material: color figures

1. introduction

Bayesian inference is widely practiced in cosmological parameter estimation. The posterior distribution of the parameters (given the observed data) is mapped from the product of the likelihood of the data (given the parameters) and the prior of the parameters according to Bayes’ theorem. The likelihood extracts the information from the observed data, whereas the prior is from external sources and often assumed to be flat for the lack of external knowledge. Therefore, the likelihood function plays a crucial role in parameter estimation and needs to be determined accurately.

Cosmic fluctuations are expected to be well described by a Gaussian random field on scales where nonlinearity is negligible, so that the Fourier modes of the fluctuations follow independent complex Gaussian distributions characterized by the power spectrum (Bardeen et al. 1986; Bond & Efstathiou 1987, referenced herein). The observed power of fluctuations at a given scale then follow a $\Gamma$ distribution that is also determined by the power spectrum. For a Gaussian random field, the power spectrum encapsulates all the information in a small set of numbers and, hence, can be analyzed in place of the fluctuations far more efficiently without loss of information (Tegmark 1997).

It is noted in cosmic microwave background (CMB) analyses that the Gaussian approximation of the power spectrum likelihood leads to parameter biases, and better approximations have been developed (Bond et al. 2000; Bartlett et al. 2000; Verde et al. 2003). In analyses of galaxy density fluctuations and weak lensing shear fluctuations, however, the Gaussian approximation remains the backbone of the standard practice, and the covariance of the observables is often taken to be independent of cosmology (e.g., Tegmark et al. 2006; Percival et al. 2010; Ho et al. 2012; Hoekstra et al. 2006; Semboloni et al. 2006; Massey et al. 2007; Benjamin et al. 2007). Although the central limit theorem guarantees Gaussianity of the power spectrum likelihood on scales much smaller than the dimensions of a survey, there is always considerable deviation at the largest scales probed by the survey.

Recently, the model dependence of the covariance of the Gaussian likelihood has drawn some attention. It is found to have a significant impact on weak lensing shear analyses (Eifler et al. 2009; Jee et al. 2013; but cf. Kilbinger et al. 2013) and a mild effect on baryon acoustic oscillation (BAO) analyses (Labatie et al. 2012). Separately, Wilking & Schneider (2013) propose a quasi-Gaussian method by applying the Gaussian approximation on an unconstrained variable that is transformed from the constrained correlation functions (Keitel & Schneider 2011; Schneider & Hartlap 2009). They find it a better approximation than the usual Gaussian approximation. Although these studies are all based on correlation functions, they motivate a closer examination of the approximations in the likelihood analyses of power spectra. In fact, Carron (2013) has shown, based on Fisher information, that including the model dependent covariance would underestimate the parameter uncertainties.

To demonstrate the effect of approximate likelihood functions, we generate mock galaxy angular power spectra and estimate cosmological parameters using several approximate likelihoods. Particular attention is given to the primordial non-Gaussianity parameter $f_{NL}$, whose effect is most prominent on the largest scales (Dalal et al. 2008; Matarrese & Verde 2008).

This paper is organized as follows. In Section 2, we introduce the likelihood of the power spectrum and its four approximations. We study the impact of the approximate likelihoods on parameter estimation in Section 3. Section 4 is a summary.

2. likelihood of the power spectrum

Bayes’ theorem relates the posterior probability $P(\theta|D)$ of the parameters $\theta$ given the data $D$ to the likelihood of the data $L(D|\theta)$ given the parameters

$$P(\theta|D) \propto P(\theta)L(D|\theta),$$

(1)

where $P(\theta)$ is the prior of the parameters, and a normalization factor depending only on the data has been dropped. With a flat prior, the task of parameter estimation is essentially mapping $P(\theta|D)$ from $L(D|\theta)$. In this paper, our dataset is taken to be the galaxy angular power spectrum.

On scales where the cosmic density field can be treated as a Gaussian random field, the real and imaginary parts of
the spherical harmonic coefficients \( a_{\ell m} \) of the density fluctuations both follow a Gaussian distribution with zero mean and variance \((1/2)P_\ell\), where \( P_\ell \equiv \langle |a_{\ell m}|^2 \rangle \) is the angular power spectrum at multipole \( \ell \). For an ideal full-sky survey without measurement noise, the mean power of the modes \( \hat{P}_\ell = (1/(2\ell + 1)) \sum_m |a_{\ell m}|^2 \) (hereafter we refer to it less rigorously as the “observed” power spectrum) is an unbiased estimator of the angular power spectrum, and it follows a \( \Gamma \) distribution\(^1\) at each \( \ell \):

\[
\mathcal{L}_\Gamma(\hat{P}_\ell | P_\ell) \propto \frac{1}{P_\ell} \left( \frac{\hat{P}_\ell}{P_\ell} \right)^{\frac{3\ell+1}{2}} \exp \left[ -\frac{(2\ell + 1)\hat{P}_\ell}{2P_\ell} \right].
\]  

whose mean and variance are \( P_\ell \) and \( (2/(2\ell + 1))P_\ell^2 \), respectively. One may roughly account for the effect of partial sky coverage by replacing \( 2\ell + 1 \) with \( (2\ell + 1) f_{\text{sky}} \) (Scott et al. 1994; Hobson & Magueijo 1996). For galaxy surveys, the power spectra include a contribution from shot noise \( n_g^{-1} \), where \( n_g \) is the surface number density of galaxies.

With a single Gaussian random field one can express the likelihood of the whole power spectrum as a product of likelihoods of each multipole, i.e., \( \mathcal{L}_\Gamma(\hat{P} | P) = \prod_\ell \mathcal{L}_\Gamma(\hat{P}_\ell | P_\ell) \). In reality, significant correlations between different multipoles can arise from various sources such as a sky cut. Tomographic analyses of galaxies and shear in multiple photometric redshift bins also need to account for the correlations between different bins. Computing the exact likelihood in such cases is not practical, so approximations are necessary (e.g., Verde et al. 2003; Carron 2013).

We consider the following approximate likelihoods: Gaussian without the determinant

\[
\mathcal{L}_G(\hat{P}_\ell | P_\ell) \propto \sqrt{\frac{2\ell + 1}{2P_\ell}} \exp \left[ -\frac{(2\ell + 1)(\hat{P}_\ell - P_\ell)^2}{4P_\ell^2} \right];
\]  

and Gaussian plus log-normal (Verde et al. 2003)

\[
\mathcal{L}_{G+LN}(\hat{P}_\ell | P_\ell) \propto \exp \left[ -\frac{(2\ell + 1)(\hat{P}_\ell - P_\ell)^2}{4P_\ell^2} \right].
\]  

It is customary to replace \((2/(2\ell + 1))P_\ell^2 \) in the exponent of Equation (4) with a constant variance (or constant covariance in multivariate case, e.g., Tegmark 1997). We refer to such an approximation as \( \mathcal{L}_{G,cc} \).

Figure 1 shows the true likelihood function of the angular power spectrum \( \mathcal{L}_\Gamma(\hat{P}_\ell | P_\ell) \) and its four approximations (left panels) along with corresponding posterior probabilities \( P( P_\ell | \hat{P}_\ell) \) (right panels). None of the approximate likelihoods matches \( \mathcal{L}_\Gamma \) for the quadrupole. However, from the point of view of parameter estimation, it is most important to reproduce the posterior probability accurately. Indeed, the upper right panel demonstrates that the posterior mapped from the Gaussian plus log-normal likelihood matches that from the true likelihood fairly well.

---

\(^1\) The combined term \((2\ell + 1)\hat{P}_\ell/P_\ell \) follows a \( \chi^2 \) distribution, a special case of the \( \Gamma \) distribution.
The Gaussian approximation $\mathcal{L}_G$ results in a biased maximum-likelihood estimate of the quadrupole. Although its posterior $\mathcal{P}_G$ appears to underestimate the uncertainty of the quadrupole, all moments of $\mathcal{P}_G$ diverge with increasing range of $P_\ell$ (see Section 3.1). The posterior probabilities of the quadrupole mapped from $\mathcal{L}_{G,\text{nd}}$ and $\mathcal{L}_{G,\text{cc}}$ recover the correct maximum likelihood value $P_\ell = \hat{P}_\ell$, but their shapes deviate significantly from $P_\ell$. In addition, $\mathcal{P}_{G,\text{nd}}$ also suffers from divergent moments.

For higher multipoles, the true likelihood function approaches Gaussian because of the central limit theorem. The approximate quadrupole mapped from Figure 1 with $\hat{P}_\ell$ by the true likelihood $\mathcal{L}_G$ is given by the ensemble (i.e., the likelihood in the integrand above) is given as $\mathcal{L}_G$. Since $\hat{\mathcal{L}}_G$ appears to underestimate the uncertainty of the quadrupole in Figure 1, all multipoles suffer from an infinite mean value and infinite variance with $\mathcal{P}_G$ and $\mathcal{P}_{G,\text{nd}}$. As such, a prior should be applied to limit the parameter range when using $\mathcal{L}_G$ and $\mathcal{L}_{G,\text{nd}}$ to estimate parameters that can drive $P_\ell$ to infinity (e.g., the normalization of the power spectrum).

What about parameters that are generic functions of $P_\ell$? In this case, one usually obtains the posterior of the parameter via $P(\theta | \hat{P}_\ell) \propto P(\theta) \mathcal{L}(\hat{P}_\ell | \theta) \equiv P(\theta) \mathcal{L}(\hat{P}_\ell | P_{\ell})$, where $P(\theta)$ is the prior on $\theta$. Taking $\theta = \ln P_\ell$ as an example and applying a flat prior on $\theta$, one gets a biased estimator using the mode of $\mathcal{P}(\theta | \hat{P}_\ell)$.

$$\langle \hat{\theta} \rangle = \int \hat{\theta} \mathcal{L}(\hat{P}_\ell | P_{\ell}) d\hat{P}_\ell = \int \ln \hat{P}_\ell \mathcal{L}(\hat{P}_\ell | P_{\ell}) d\hat{P}_\ell = \psi \left( 2\ell + 1 \right) + \ln \left( \frac{2}{2\ell + 1} \right) + \theta^a,$$  

(11)

where $\psi$ is the digamma function, and $\theta^a \equiv \ln P_{\ell}^a$. The difference $\langle \hat{\theta} \rangle - \theta^a$ scales as $-2(\ell + 1)$ for large $\ell$. One could also map the posterior of $\theta$ from that of $P_{\ell}$ via $P(\theta | \hat{P}_\ell) = \mathcal{P}(\hat{P}_\ell | P_{\ell}) d\hat{P}_\ell / d\theta$, which is equivalent to applying $P(\theta) = P(\hat{P}_\ell) d\hat{P}_\ell / d\theta$ with the conventional approach. The result would still be biased.

Maximum likelihood estimators are asymptotically unbiased, meaning that the biases, if exist, decrease with the sample size. This is indeed seen in Equation (11): the bias of $\hat{\theta}$ estimated from a particular $P_{\ell}$ is inversely proportional to the number of modes available at $\ell$. The asymptotic behavior seems to guarantee unbiased parameter estimation with abundant small-scale data. However, some parameters might not be sensitive to small scales at all, and the rate at which the bias drops with the sample size may also vary with the parameters. Given that much work has been done to improve approximations of the power spectrum likelihood function, it is necessary as well to quantify potential biases on the parameters estimated with even the most accurate power spectrum likelihood.

Determining the uncertainties is an integral part of parameter estimation. For an observed quadrupole $\hat{P}_\ell$ in a full-sky noise-free survey, the 1$\sigma$ confidence interval of the estimated $P_{\ell}$ from the exact posterior $\hat{P}_\ell$, i.e., the range enclosing 68% posterior probability and having the same probability density at its two end points, is $[0.42\hat{P}_\ell; 3.41\hat{P}_\ell]$ (see the upper-right panel of Figure 1). Given the WMAP nine year maximum likelihood

Since the mean of the posterior distribution, $\hat{P}_\ell$, does not necessarily coincide with $\hat{P}_\ell$, it can be less ideal as an estimator. For instance, $P_T$ leads to

$$\hat{P}_\ell = \int P_T \mathcal{L}_T (\hat{P}_\ell | \hat{P}_\ell) d\hat{P}_\ell = \frac{2\ell + 1}{2\ell - 3} \hat{P}_\ell$$  

(8)

$$\langle \hat{P}_\ell \rangle = \int \hat{P}_\ell \mathcal{L}_T (\hat{P}_\ell | \hat{P}_\ell^a) d\hat{P}_\ell = \frac{2\ell + 1}{2\ell - 3} \hat{P}_\ell^a$$  

(9)

$$\sigma^2_{\hat{P}_\ell} = \int \hat{P}_\ell^2 \mathcal{L}_T (\hat{P}_\ell | \hat{P}_\ell^a) d\hat{P}_\ell - \langle \hat{P}_\ell \rangle^2 = \frac{2}{2\ell + 1} (\hat{P}_\ell^a)^2$$  

(10)

for $\ell \geq 2$, which perform poorly at low $\ell$. Moreover, Equations (3) and (4) show that as $P_T$ approaches infinity, the posterior $\mathcal{P}_G$ behaves as $P_{\ell}^{-1}$, and $\mathcal{P}_{G,\text{nd}}$ becomes a constant of $P_\ell$. Consequently, all the moments of $P_T$ diverge with $\mathcal{P}_G$ and $\mathcal{P}_{G,\text{nd}}$. Even though $\mathcal{P}_G$ appears to underestimate the uncertainty of the quadrupole in Figure 1, all multipoles suffer from an infinite mean value and infinite variance with $\mathcal{P}_G$ and $\mathcal{P}_{G,\text{nd}}$.
value of $151 \mu K^2$ for the quadrupole (Bennett et al. 2013), the 1σ interval is then $[63.4 \mu K^2, 515 \mu K^2]$ in the best case, which is slightly narrower than that in Bennett et al. (2013).

Even with the ideal survey considered above, the uncertainties of the power spectrum inferred from its posterior can differ considerably from the cosmic variance on scales where the number of modes available is small. For instance, the ensemble averaged 1σ confidence interval of the quadrupole, i.e., $\left[0.42 P_m^u, 3.41 P_m^u \right]$, is more than twice the range of cosmic variance, i.e., $\left[0.37 P_m^u, 1.63 P_m^u \right]$. When the scales of interest are much smaller than the survey dimensions, the uncertainties become essentially the same as the cosmic variance (in practice, noise dominates the small-scale uncertainties).

### 3.2. Estimation with Fiducial Data

To study the effect of the power spectrum likelihood in a more realistic way, we employ a Markov Chain Monte Carlo (MCMC) code, CosmoMC$^2$ (Lewis & Bridle 2002), to estimate parameters from mock galaxy angular power spectra and compare the results. Because $L_{G,\text{fid}}$ is not widely used, hereafter we only consider $L_T$, $L_G$, $L_{G,G}$ and $L_{G,LN}$.

We assume an imaging survey covering half of the sky and a Gaussian galaxy redshift distribution centered at $z_m = 1$ with a dispersion of $\sigma_z = 0.15$. The surface number density is taken to be $n_g = 10\text{arcmin}^{-2}$. The calculation of the angular power spectrum is described in Zhan (2006) with modification to include the damping of the BAO signal (Eisenstein et al. 2007). Since the difference between various approximate likelihoods and the exact one is most pronounced on the largest scales, we limit the multiple range to $2 \leq \ell \leq 1000$.

The parameters include the matter fraction $\Omega_m$, tilt of the matter power spectrum $n_s$, reduced Hubble constant $h$, rms density fluctuation within 8 $h^{-1}\text{Mpc}$ $\sigma_8$, primordial non-Gaussianity parameter of the local type $f_{NL}$, and linear clustering bias $b_g$. The fiducial values of these parameters are $(\Omega_m, n_s, h, \sigma_8, f_{NL}, b_g) = (0.27, 0.96, 0.72, 0.78, 0, 2)$, consistent with recent measurements (Larson et al. 2011). The baryon fraction is fixed at $\Omega_b = 0.0446$ and the cosmological constant parameter $\Omega_m = 1 - \Omega_b$.

The primordial non-Gaussianity is one of the few probes that can help shed light on the physics of inflation. Therefore, the parameter $f_{NL}$ is of great interest to future galaxy surveys. It gives rise to a scale-dependent effective bias (Dalal et al. 2008; Matarrese & Verde 2008)

$$b(k, f_{NL}) = b_g + f_{NL}(b_g - 1)u(k, z), \quad (12)$$

where $u(k, z) \propto k^{-2}T^{-1}(k)$ with $T(k)$ being the transfer function of fluctuations (see, e.g., Hamaus et al. 2011). Because the effect of $f_{NL}$ is most pronounced on the largest scales, its estimation may be prone to the errors of the approximate likelihoods at low multipoles.

In this subsection, we set the “observed” power spectrum $P_l$ to the power spectrum of the fiducial cosmological model, i.e., $P_{l,\text{fid}}$, which in some sense represents the best case scenario. The galaxy angular power spectrum with the effective bias can be written as

$$P_\ell = P_{\ell,\text{fid}} + f_{NL}Q_\ell + \tilde{f}_{NL}R_\ell, \quad (13)$$

where $Q_\ell$ and $R_\ell$, like $P_{\ell,\text{fid}}$, are determined by parameters other than $f_{NL}$. Equation (13) is convenient when $f_{NL}$ is the only parameter to be estimated.

---

2 http://cosmologist.info/cosmomc/

---

We first estimate all six parameters simultaneously. A Gaussian prior is applied to the linear bias $\sigma(b_g)/b_g = 0.2$, and flat priors are applied in the following parameter ranges: $0 < \Omega_m < 1$, $0.5 < n_s < 1.5$, $0 < h < 2$, $0 < \sigma_8 < 2$, and $-200 < f_{NL} < 200$. For each of the four likelihoods, we run four chains with about $10^5$ samples per chain after burn-in and merge them into one sample. The marginalized constraints on $n_s$ and $f_{NL}$ are shown in Figure 2. The black solid contours indicate the 1σ and 2σ regions estimated from the MCMC samples (red dots, thinned to leave ~3000 points for display), with the mode and mean values marked by triangles and squares, respectively. The green dashed contours denote predictions by a Fisher matrix analysis centered at the fiducial values (green crosses), which is based on the likelihood form of $L_{G,G}$ (Carron 2013).

When all the other parameters are marginalized, all the four likelihoods are able to recover the fiducial values of $n_s$ and $f_{NL}$ without much bias. However, the shape of the posterior contours based on $L_G$ and $L_{G,G}$ differ significantly from that based on the correct likelihood $L_T$, leading to mis-estimation of the uncertainties of $f_{NL}$. Contours given by $L_{G,LN}$, meanwhile, agree with those given by $L_T$.

Equation (12) means that a positive $f_{NL}$ would increase the power spectrum as long as $b_g > 1$. Since the posterior probabilities of low multipoles based on $L_G$ and $L_{G,G}$ are significantly lower than those based on $L_{G,LN}$ and $L_T$ at $P_l > \tilde{P}_l$ (e.g., Figure 1), the $f_{NL}$ contours in the $L_G$ and $L_{G,G}$ panels of Figure 2 are less extended toward the positive $f_{NL}$ direction than those in the $L_{G,LN}$ and $L_T$ panels. Similarly, the contours in the $L_{G,G}$ panel is more extended in the negative $f_{NL}$ direction than those in the other panels.

The effective bias could be driven below zero numerically by $f_{NL}$ so that one might expect a second peak in the posterior of $f_{NL}$. This is actually true for a single mode in $k$ ($\sim k/D_A$ with $D_A$ being the angular diameter distance). When all the modes are included, the probability of a large negative $f_{NL}$ is strongly suppressed by low multipoles. As a related test, we impose a cut $b > 0$ in MCMC. The result of $L_{G,G}$ becomes much more consistent with that of the Fisher Matrix analysis, and there is essentially no change in the results of the other likelihoods.

Next, we fix all the other parameters except $f_{NL}$. In this case, the posterior can be scanned efficiently using Equation (13). The scan range is $[-200, 200]$ with a step size of 0.01. The results are shown in Figure 3. The Gaussian likelihood $L_G$ causes a significant bias on $f_{NL}$ and underestimates the uncertainty. The Gaussian with a constant covariance likelihood $L_{G,G}$ does not bias $f_{NL}$ much but causes a spurious tail on the negative side. The Gaussian plus log-normal likelihood $L_{G,LN}$ again reproduces the result of the correct likelihood $L_T$ with less than 10% errors in the mean value and the two types of confidence level intervals in Figure 3.

### 3.3. Ensemble Behavior of the Estimators

We generate $10^4$ sets of galaxy angular power spectra to examine the ensemble behavior of the estimators of the primordial non-Gaussianity parameter. The power of each multipole ($2 \leq \ell \leq 1000$) is randomly drawn from its underlying $\Gamma$ distribution. We then scan the posterior probability of $f_{NL}$ from these power spectra using the same likelihood functions as in the last subsection. The scan range and step size are also kept the same.

The mean value $\bar{f}_{NL}$ and the maximum likelihood value $\tilde{f}_{NL}$ are determined for each set of the galaxy angular power spectrum. Their ensemble distributions are shown in Figure 4.
One can see minor peaks of $\tilde{f}_{NL}$ near $-10$. This is caused by the same negative effective bias issue noted in the last subsection where the posterior of a very special realization is examined. These minor peaks disappear when a cut of $b > 0$ is imposed.

The distribution of $\tilde{f}_{NL}$ from $\mathcal{L}_{G,cc}$ is mildly skewed toward $f_{NL} < 0$, and all the other results display no significant bias. The maximum likelihood estimator of $\mathcal{L}_{G,cc}$, $\mathcal{L}_{G+LN}$ and $\mathcal{L}_{\Gamma}$ are not biased, if the minor peaks are ignored. However, it is worth mentioning that the ensemble behavior is different from the behavior of a particular realization. As the upper left panel of Figure 3 shows, the Gaussian likelihood $\mathcal{L}_{G}$ underestimates $f_{NL}$ by more than 1σ even with $\hat{P}_t$ set to $P^u_\ell$. 

(A color version of this figure is available in the online journal.)
ized. A significant bias and underestimated uncertainty on survey size. Parameters that are constrained mainly by scales close to the three Gaussian approximations may lead to biases on parameters regardless of the physical dimensions of the survey because the distribution of the power spectrum is intrinsically not Gaussian. The distribution approaches Gaussian only on scales much smaller than the survey size where there are enough modes for the central limit theorem to take effect. Applying much smaller than the survey size where there are enough Gaussian likelihoods in the study (Gaussian \( L_{G} \), Gaussian without the determinant \( L_{G,nd} \), and Gaussian with a constant covariance \( L_{G,cc} \)) can accurately reproduce the posterior probability of the power spectrum on scales where the number of modes contained in the survey is small. This issue exists regardless of the physical dimensions of the survey because the distribution of the power spectrum is intrinsically not Gaussian. The distribution approaches Gaussian only on scales much smaller than the survey size where there are enough modes for the central limit theorem to take effect. Applying the three Gaussian approximations may lead to biases on parameters that are constrained mainly by scales close to the survey size.

In our tests with the primordial non-Gaussianity parameter \( f_{NL} \), the Gaussian likelihoods \( L_{G} \) and \( L_{G,cc} \) lead to distorted \( n_{s} \cdot f_{NL} \) error contours when four other parameters are marginalized. A significant bias and underestimated uncertainty on \( f_{NL} \) are obtained with \( L_{G} \) when all other parameters are fixed. The Gaussian plus log-normal likelihood \( L_{G+LN} \) can reproduce the true posterior probability of the power spectrum and that of \( f_{NL} \) accurately.

Analyses with real data often have to work with far more complex likelihood functions than we have considered. It is also worth emphasizing that using the exact likelihood of the power spectrum does not guarantee unbiased estimates of the parameters. Although perfection of the likelihood approximations has been pursued (mostly in CMB analyses), relatively less effort has been made on quantifying potential biases of the maximum likelihood estimator itself, which would be specific to each unique set of parameters and data. Given the unprecedented statistical power of future surveys, it is especially important to thoroughly examine the likelihood functions and estimators for potential systematics.

4. SUMMARY

We have examined several approximate likelihood functions for parameter estimation with the galaxy angular power spectrum in an idealized case. None of the three forms of Gaussian likelihoods in the study (Gaussian \( L_{G} \), Gaussian without the determinant \( L_{G,nd} \), and Gaussian with a constant covariance \( L_{G,cc} \)) can accurately reproduce the posterior probability of the galaxy angular power spectrum as labeled. Given the unprecedented statistical power of future surveys, it is especially important to thoroughly examine the likelihood functions and estimators for potential systematics.

Figure 4. Ensemble distributions of the mean \( f_{NL} \) (filled bars) and the maximum likelihood \( f_{NL} \) (open bars) estimated from 10⁴ realizations of the “observed” power spectrum. Each panel presents the results from the likelihood of the galaxy angular power spectrum as labeled.

(A color version of this figure is available in the online journal.)

We thank the referee for the very constructive comments and suggestions. This work was supported by the National Natural Science Foundation of China grant No. 11033005, the National Key Basic Research Science Foundation of China grant No. 2010CB833000, the Bairen program from the Chinese Academy of Sciences.

REFERENCES

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15
Barlett, J. G., Douspis, M., Blanchard, A., & Le Dour, M. 2000, A&A, 146, 507
Benjamin, J., Heymans, C., Semboloni, E., et al. 2007, MNRAS, 381, 702
Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, ApJS, 208, 20
Bond, J. R., & Efstathiou, G. 1987, MNRAS, 226, 655
Bond, J. R., Jaffe, A. H., & Knox, L. 2000, ApJ, 533, 19
Carron, J. 2013, A&A, 551, A88
Dalal, N., Döré, O., Huterer, D., & Shirokov, A. 2008, PhRvD, 77, 123514
Eifler, T., Schneider, P., & Hartlap, J. 2009, A&A, 502, 721
Eisenstein, D. J., Seo, H.-J., & White, M. 2007, ApJ, 664, 660
Hamaus, N., Seljak, U., & Desjacques, V. 2011, PhRvD, 84, 083509
Hamimeche, S., & Lewis, A. 2008, PhRvD, 77, 103013
Hinshaw, G., Nolta, M. R., Bennett, C. L., et al. 2007, ApJS, 170, 288
Ho, S., Cuesta, A., Seo, H.-J., et al. 2012, ApJ, 761, 14
Hobson, M. P., & Magueijo, J. 1996, MNRAS, 283, 1133
Hoekstra, H., Mellier, Y., van Waerbeke, L., et al. 2006, ApJ, 647, 116
Jee, M. J., Tyson, J. A., Schneider, M. D., et al. 2013, ApJ, 765, 74
Keitel, D., & Schneider, P. 2011, A&A, 534, A76
 Kilbinger, M., Fu, L., Heymans, C., et al. 2013, MNRAS, 430, 2200
Labatie, A., Starck, J. L., & Lachieze-Rey, M. 2012, ApJ, 760, 97
Larson, D., Dunkley, J., Hinshaw, G., et al. 2011, ApJS, 192, 16
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Massey, R., Rhodes, J., Leauthaud, A., et al. 2007, ApJS, 172, 239
Matarrese, S., & Verde, L. 2008, ApJL, 677, L77
Percival, W. J., Réid, B. A., Eisenstein, D. J., et al. 2010, MNRAS, 401, 2148
Planck collaboration, Ade, P. A. R., Aghanim, N., et al. 2013, arXiv:1303.5075
Schneider, P., & Hartlap, J. 2009, A&A, 504, 705
Scott, D., Srednicki, M., & White, M. 1994, ApJL, 421, L5
Semboloni, E., Mellier, Y., van Waerbeke, L., et al. 2006, A&A, 452, 51
Tegmark, M. 1997, PhRvD, 55, 5895
Tegmark, M., Eisenstein, D. J., Strauss, M. A., et al. 2006, PhRvD, 74, 123507
Verde, L., Peiris, H. V., Spergel, D. N., et al. 2003, ApJS, 148, 195
Wilking, P., & Schneider, P. 2013, A&A, 556, A70