The Probability Distribution for Non-Gaussianity Estimators

Tristan L. Smith, 1 Marc Kamionkowski, 2 and Benjamin D. Wandelt 3

1 Berkeley Center for Cosmological Physics, Physics Department, University of California, Berkeley, CA 94720
2 California Institute of Technology, Mail Code 350-17, Pasadena, CA 91125
3 UPMC Univ Paris 06, Institut d’Astrophysique de Paris, 98 bis, bld Arago, 75014 Paris, France
(Dated: January 19, 2013)

One of the principle efforts in cosmic microwave background (CMB) research is measurement of the parameter \( f_{\text{nl}} \) that quantifies the departure from Gaussianity in a large class of non-minimal inflationary (and other) models. Estimators for \( f_{\text{nl}} \) are composed of a sum of products of the temperatures in three different pixels in the CMB map. Since the number \( \sim N_{\text{pix}}^2 \) of terms in this sum exceeds the number \( N_{\text{pix}} \) of measurements, these \( \sim N_{\text{pix}}^2 \) terms cannot be statistically independent. Therefore, the central-limit theorem does not necessarily apply, and the probability distribution function (PDF) for the \( f_{\text{nl}} \) estimator does not necessarily approach a Gaussian distribution for \( N_{\text{pix}} \gg 1 \). Although the variance of the estimators is known, the significance of a measurement of \( f_{\text{nl}} \) depends on knowledge of the full shape of its PDF. Here we use Monte Carlo realizations of CMB maps to determine the PDF for two minimum-variance estimators: the standard estimator, constructed under the null hypothesis \( (f_{\text{nl}} = 0) \), and an improved estimator with a smaller variance for \( f_{\text{nl}} \neq 0 \). While the PDF for the null-hypothesis estimator is very nearly Gaussian when the true value of \( f_{\text{nl}} \) is zero, the PDF becomes significantly non-Gaussian when \( f_{\text{nl}} \neq 0 \). In this case we find that the PDF for the null-hypothesis estimator \( \hat{f}_{\text{nl}} \) is skewed, with a long non-Gaussian tail at \( f_{\text{nl}} > |f_{\text{nl}}| \) and less probability at \( f_{\text{nl}} < |f_{\text{nl}}| \) than in the Gaussian case. We provide an analytic fit to these PDFs. On the other hand, we find that the PDF for the improved estimator is nearly Gaussian for observationally allowed values of \( f_{\text{nl}} \). We discuss briefly the implications for trispectrum (and other higher-order correlation) estimators.

PACS numbers:

I. INTRODUCTION

The simplest single-field slow-roll inflation models predict that primordial perturbations should be nearly Gaussian, but with predictably small departures from Gaussianity. This is often quantified through the non-Gaussian parameter \( f_{\text{nl}} \) defined by

\[
\Phi = \phi + f_{\text{nl}} \left( \phi^2 - \langle \phi^2 \rangle \right),
\]

where \( \Phi \) is the gravitational potential and \( \phi \) a Gaussian random field. Standard single-field slow-roll inflation predicts \( f_{\text{nl}} \ll 1 \) for the primordial field (although nonlinear evolution of the density field may produce \( f_{\text{nl}} \sim 1 \) at the time of recombination; see, e.g., Ref. [4]). However, multi-field [5] or curvaton [6] models, or models with sharp features [7] or wiggles [8] may produce larger values of \( f_{\text{nl}} \). Measurement of \( f_{\text{nl}} \) has thus become one of the primary goals of cosmic microwave background (CMB) and large-scale-structure (LSS) research. Current limits from the CMB/LSS are in the ballpark of \( |f_{\text{nl}}| \lesssim 100 \) [9] [10]. The plot has thickened with a suggestion [11] (not universally accepted) that WMAP data prefers (at the 2.8\( \sigma \) level) \( f_{\text{nl}} \neq 0 \), with a best-fit value \( f_{\text{nl}} \simeq 35 \). The Planck satellite [12] is expected to achieve a sensitivity of \( f_{\text{nl}} \sim 5 \).

In this paper, we address the following question: What is the probability distribution function (PDF) \( P(\hat{f}_{\text{nl}}) \) for an estimator \( \hat{f}_{\text{nl}} \) that is constructed from a CMB map? If the PDF departs from the Gaussian distribution that is often assumed, then the 99.7% confidence level (C.L.) interval for \( f_{\text{nl}} \) may be different than three times the standard deviation for \( f_{\text{nl}} \). The interpretation of measurements thus requires knowledge of this PDF.

The question arises as the theory predicts not only the mean value of the estimator \( \hat{f}_{\text{nl}} \) but it also makes a prediction for the detailed functional form of the PDF \( P(\hat{f}_{\text{nl}}) \). The consistency of a given measurement of \( \hat{f}_{\text{nl}} \) with a theoretical prediction for \( f_{\text{nl}} \) depends on knowledge of the shape of \( P(\hat{f}_{\text{nl}}) \). Thus, for example, we often evaluate or forecast the standard error \( \sigma_{\text{nl}} \) with which a given measurement will recover the true value of \( f_{\text{nl}} \) and then simply assume that the error is Gaussian. If so, then with \( \sigma_{\text{nl}} = 10 \), for example, a measurement of \( \hat{f}_{\text{nl}} = 30 \) would represent a 3\( \sigma \) departure from \( f_{\text{nl}} = 0 \) and a measurement \( \hat{f}_{\text{nl}} = 0 \) would represent a 3\( \sigma \) departure from \( f_{\text{nl}} = 30 \). However, if the PDF depends on the true value \( f_{\text{nl}} \), and if that distribution is non-Gaussian, then it may be that a measurement \( \hat{f}_{\text{nl}} = 30 \) could be easily consistent with a true value \( f_{\text{nl}} = 0 \), while a measurement \( \hat{f}_{\text{nl}} = 0 \) could be inconsistent with \( f_{\text{nl}} = 30 \) with a confidence greater than “3\( \sigma \.” We will see below that something like this actually occurs with measurements of \( f_{\text{nl}} \).

This question is particularly important for measurements of non-Gaussianity (as opposed, for example, for the CMB power spectrum), because \( \hat{f}_{\text{nl}} \) is a sum over products of three temperature measurements (unlike the power spectrum, which sums over squares of temperature measurements). Suppose the temperature is measured in...
that the variance of this NHMV estimator increases as the "null-hypothesis minimum-variance under the null hypothesis estimate of the estimator designed to have the minimum variance is usually limited to the number, required to determine a 99.7% C.L. detection or sometimes even fewer if it is just the variance that is being estimated. Although with only 1000 realizations Fig. 8 in Ref. [13] shows hints of a non-Gaussian shape when \( f_{nl} \neq 0 \) with a long non-Gaussian tail for \( f_{nl} \gg |f_{nl}| \) and less probability at \( f_{nl} < |f_{nl}| \) than in the Gaussian case. As an example, taking \( f_{nl} = 100 \) for an experiment which measures multipoles out to \( l_{\text{max}} = 3000 \) (such as Planck) and assuming a Gaussian PDF for the NHMV this experiment measures 74 \( \leq f_{nl} \leq 148 \) at the 99.7% C.L.; the actual PDF shows that this experiment measures 68 \( \leq f_{nl} \leq 143 \) at the 99.7% C.L. Applying the CSZ estimator to the data we find it has a PDF which is well approximated by a Gaussian with \( f_{nl} = 100 \pm 12.5 \) at 99.7% C.L.

This paper is organized as follows. In Sec. [III we construct the standard minimum-variance estimator \( f_{nl} \) under the null hypothesis \( f_{nl} = 0 \) and discuss why the PDF for this estimator is not necessarily Gaussian, even in the limit of a large number of pixels. In Sec. [III.A we use Monte Carlo calculations to evaluate the PDF \( P(f_{nl}) \) for this estimator if the null hypothesis is indeed valid, i.e., if \( f_{nl} \) is indeed zero. We find that the PDF in this \( f_{nl} = 0 \) case is well approximated by a Gaussian, for \( N_{\text{pix}} \gg 1 \), even though the central-limit theorem does not apply. In Sec. [III.B we calculate the PDF assuming that the null hypothesis is not valid, i.e., if \( f_{nl} \neq 0 \). We find the PDFs in this case can be highly non-Gaussian, skewed to large \( |f_{nl}| \), with long large-\( f_{nl} \) non-Gaussian tails and less likelihood at \( f_{nl} \leq |f_{nl}| \) relative to the Gaussian distribution of the same variance. We provide fitting formulas for the PDF as a function of the estimator \( f_{nl} \), the true value of \( f_{nl} \), and the maximum multipole moment \( l_{\text{max}} \) of the map. In Sec. [IV we discuss the PDF of the CSZ estimator. We show that this estimator is well approximated by a Gaussian for values of \( f_{nl} \) still allowed by observations. In Sec. [V we summarize and discuss some possible implications of the work for other bispectra and also for the trispectrum and other higher-order statistics. An Appendix discusses the computational techniques we used in order to perform our Monte Carlo simulations.

---

1 We note that the CSZ estimator, which is defined under the Sachs-Wolfe limit, has yet to be generalized so that it can be applied to actual data. On the other hand a Bayesian approach, discussed in Ref. [13], allows for an \( f_{nl} \) inference that saturates the Cramer-Rao bound even in the presence of non-Gaussianity.
II. NON-GAUSSIANITY ESTIMATORS

A. Formalism

We assume a flat sky to avoid the complications (e.g., spherical harmonics, Clebsch-Gordan coefficients, Wigner 3j and 6j symbols, etc.) associated with a spherical sky, and we further assume the Sachs-Wolfe limit. We denote the fractional temperature perturbation at position \( \hat{\theta} \) on a flat sky by \( T(\hat{\theta}) \) and refer to it hereafter simply as the temperature.

The field \( T(\hat{\theta}) \) has a power spectrum \( C_l \) given by

\[
\langle T_{i_1} T_{i_2} \rangle = \Omega \delta_{i_1+i_2,0} C_l, \tag{2}
\]

where \( \Omega = 4\pi f_{\text{sky}} \) is the survey area (in steradian),

\[
T_i = \int d^2 \hat{\theta} e^{-i \hat{\theta} \cdot \vec{l}} T(\hat{\theta}) \approx \frac{\Omega}{N_{\text{pix}}} \sum_\delta e^{-i \vec{l} \cdot \delta} T(\hat{\delta}), \tag{3}
\]

is the Fourier transform of \( T(\hat{\theta}) \), and \( \delta_{i_1+i_2,0} \) is a Kronecker delta that sets \( i_1 = -i_2 \). The power spectrum for \( T(\hat{\theta}) \) is given by

\[
C_l = \frac{2\pi A}{l^2}, \tag{4}
\]

where the amplitude, \( A \approx 10^{-10} \). The bispectrum \( B(l_1, l_2, l_3) \) is defined by

\[
\langle T_{i_1} T_{i_2} T_{i_3} \rangle = \Omega \delta_{i_1+i_2+i_3,0} B(l_1, l_2, l_3). \tag{5}
\]

The Kronecker delta insures that the bispectrum is defined only for \( \vec{l}_1 + \vec{l}_2 + \vec{l}_3 = 0 \); i.e., only for triangles in Fourier space. Statistical isotropy then dictates that the bispectrum depends only on the magnitudes \( l_1, l_2, l_3 \) of the three sides of this Fourier triangle.

B. The null-hypothesis minimum-variance estimator

We now review how to construct the minimum-variance estimator for \( f_{\text{nl}} \) under the null hypothesis. This is the quantity that one would first determine from the data to check for consistency of the measurement with the null hypothesis \( f_{\text{nl}} = 0 \).

From Eq. (5), each triangle \( \vec{l}_1 + \vec{l}_2 + \vec{l}_3 = 0 \) gives an estimator,

\[
(\hat{f}_{\text{nl}})_{123} = \frac{T_{i_1} T_{i_2} T_{i_3}}{\Omega B(l_1, l_2, l_3) / f_{\text{nl}}}, \tag{6}
\]

and under the null hypothesis this has a variance proportional to

\[
\frac{\Omega^3 C_{l_1} C_{l_2} C_{l_3}}{[\Omega B(l_1, l_2, l_3) / f_{\text{nl}}]^2}. \tag{7}
\]

The null-hypothesis minimum-variance estimator is constructed by adding all of these estimators with inverse-variance weighting. It is

\[
\hat{f}_{\text{nl}} \equiv \sigma_{\text{nl}}^2 \sum_{i_1+i_2+i_3=0} \frac{T_{i_1} T_{i_2} T_{i_3} B(l_1, l_2, l_3) / f_{\text{nl}}}{6\Omega^2 C_{l_1} C_{l_2} C_{l_3}}, \tag{8}
\]

and it has inverse variance,

\[
\sigma_{\text{nl}}^{-2} = \sum_{i_1+i_2+i_3=0} \frac{[B(l_1, l_2, l_3) / f_{\text{nl}}]^2}{6\Omega C_{l_1} C_{l_2} C_{l_3}}. \tag{9}
\]

C. Non-gaussianity of the PDF

If the number of pixels in the CMB map is \( N_{\text{pix}} \), then there are also \( N_{\text{pix}} \) statistically independent \( T_i \). But there are a much larger number, \( \propto N_{\text{pix}}^2 \), of triplets \( T_{i_1} T_{i_2} T_{i_3} \), included in the estimator (cf., Eq. (8)), and so the number of individual “data points” (i.e., triplets) used in the minimum-variance estimator scales like \( N_{\text{pix}}^2 \gg N_{\text{pix}} \)

Since the number of terms included in the estimator is greater than the number of independently measured data points the standard central-limit theorem does not apply. Thus, we cannot assume that the PDF of the estimator will approach a Gaussian in the \( N_{\text{pix}} \to \infty \) limit.

This contrasts with the estimator \( \hat{C}_l \propto \sum |T_i|^2 \) of the power spectrum \( C_l \). While the PDF for \( \hat{C}_l \) is not necessarily Gaussian (it has a \( \chi^2 \) distribution), it is the sum of the squares of statistically independent quantities. The central-limit theorem therefore applies, and the distribution for \( \hat{C}_l \) does indeed approach a Gaussian for large \( l \).

The problems we address here for \( f_{\text{nl}} \) estimators parallel those discussed in the literature for the quadrupole moment \( C_2 \), as the distribution for quadrupole-moment estimators will be highly non-Gaussian and will also depend on the underlying theory (see, e.g., Ref. [18]).

III. THE PDF OF \( \hat{f}_{\text{nl}} \) FOR THE LOCAL MODEL

We now restrict our attention to a family of non-Gaussian models in which the temperature \( T(\hat{\theta}) \) has a non-Gaussian component; i.e.,

\[
T(\hat{\theta}) = t(\hat{\theta}) + 3 f_{\text{nl}} \left\{ |t(\hat{\theta})|^2 - \langle |t(\hat{\theta})|^2 \rangle \right\}, \tag{10}
\]

where \( t(\hat{\theta}) \) is a Gaussian random field with a power spectrum \( C_l \) given in Eq. (4). To zero-th order in \( f_{\text{nl}} \), the power spectrum and correlation function for \( T(\hat{\theta}) \) are the same as those for \( t(\hat{\theta}) \). Note that \( T(\hat{\theta}) \) is, strictly speaking, the temperature fluctuation, so \( \langle T(\hat{\theta}) \rangle = 0 = T_{\vec{r}=0} \).

The bispectrum for this model is

\[
B(l_1, l_2, l_3) = 6 f_{\text{nl}} (C_{l_1} C_{l_2} + C_{l_1} C_{l_3} + C_{l_2} C_{l_3}). \tag{11}
\]
The temperature Fourier coefficients can be written $T_{\ell} = t_{\ell} + f_{nl} \delta t_{\ell}$ with
\[ \delta t_{\ell}^2 \equiv \frac{3}{\Omega} \sum_{\ell'} t_{\ell'}^2 t_{\ell'}^2. \] (12)

Formally, the sum goes from $0 < |\vec{l}| < \infty$, but for a finite-resolution map, the sum is truncated at some $l_{\text{max}}$ such that the number of Fourier modes equals the number of data points.

We now proceed to evaluate $P(\hat{f}_{nl}; f_{nl}=0, l_{\text{max}})$, the PDF that arises if the true value is $f_{nl}$ for the NHMV estimator $\hat{f}_{nl}$ and for a map with $l_{\text{max}}$. To do so, we generated large numbers of Monte Carlo realizations of maps according to Eq. (12), for some assumed value of $f_{nl}$, and then applied the estimator in Eq. (3) to these maps. Each map is simulated in harmonic space from $l_{\text{min}} = 2$ up to a maximum multipole $l_{\text{max}}$. In order to produce a large number of realizations we re-expressed the generation of maps and implementation of the estimator in terms of fast Fourier transforms as discussed in Appendix A.

**A. The PDF of the null hypothesis minimum-variance estimator with $f_{nl} = 0$**

First we consider the shape of $P(\hat{f}_{nl}; f_{nl}=0, l_{\text{max}})$, the PDF for the NHMV estimator in Eq. (3) applied to a purely Gaussian ($f_{nl} = 0$) map. To do this we generated $10^6$ Gaussian realizations and applied the estimator in Eq. (3) to generate a histogram of values of $\hat{f}_{nl}$. From this histogram we determined $P(\hat{f}_{nl}; f_{nl}=0, l_{\text{max}})$ out to four times the root-variance, as shown in Fig. 1.

First we note that our simulations verify that the variance of the distribution for the null case is well approximated by the analytic expression [16] [17].
\[ \sigma^2_{\hat{f}_{nl}} \approx \frac{1}{8 A l_{\text{max}}^2 \ln(l_{\text{max}})}. \] (13)

Additionally our simulations show that out to at least four times the root-variance, the PDF $P(\hat{f}_{nl}; f_{nl}=0, l_{\text{max}})$ is well approximated by a Gaussian for $l_{\text{max}} \gtrsim 25$, even though the conditions for the central-limit theorem to apply are not satisfied. Therefore, a measurement of $f_{nl}$ that differed from 0 at more than three times the root-variance would indeed constitute a ‘99.7% confidence level’ inconsistency with the $f_{nl} = 0$ hypothesis.

**B. The PDF of the null hypothesis minimum-variance estimator with $f_{nl} \neq 0$**

We now consider the form of $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ when $f_{nl} \neq 0$, the PDF for the null-hypothesis minimum-variance estimator if the null hypothesis is in fact not valid. In this case, the non-Gaussian statistics of the $T_l$s impart some non-Gaussianity to the $\hat{f}_{nl}$ PDF.

In Fig. 2 we show $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ calculated using $10^6$ realizations with $f_{nl} = 1500$ and $l_{\text{max}} = 25$. Clearly the PDF in this case is highly non-Gaussian.

Non-Gaussianity of $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ for a central value $f_{nl} \neq 0$ may be significant for the interpretation of data. Suppose, for example, that a CMB measurement returns $\hat{f}_{nl} = 0$ with a root-variance $\sigma_{\hat{f}_{nl}} = 40$. If the PDF was assumed to be Gaussian the measurement $\hat{f}_{nl} = 0$ would rule out $f_{nl} = 100$ at the 2.5σ level, but given the asymmetric PDF of Fig. 2 it may rule out $f_{nl} = 100$ at a much higher significance.

In order to better understand the origin of the non-Gaussian PDF, it is useful to expand the minimum-variance estimator in Eq. (3) to linear order in $f_{nl}$ [13]:
\[ \hat{f}_{nl} \approx E_0 + f_{nl} E_1 + \cdots, \] (14)
where
\[ E_0 = \sigma_{f_{nl}}^2 \sum_{l_1+l_2+l_3=0} \frac{t_{l_1} t_{l_2} t_{l_3}}{6 \Omega^2 f_{nl} C_{l_1} C_{l_2} C_{l_3}} B(l_1, l_2, l_3) \] (15)
\[ E_1 = \sigma_{f_{nl}}^2 \sum_{l_1+l_2+l_3=0} \frac{\delta t_{l_1} t_{l_2} t_{l_3}}{2 \Omega^2 C_{l_1} C_{l_2} C_{l_3}} B(l_1, l_2, l_3). \] (16)

Since $E_0 \sim t^3$ and $E_1 \sim t^4$, it is clear that $\langle E_0 \rangle = 0$ and $\langle E_0 E_1 \rangle = 0$, and the normalization guarantees that
FIG. 2: The PDF $P(\hat{f}_{nl})$ when $\hat{f}_{nl} = 1500$ using the estimator in Eq. (8) with $l_{\text{max}} = 25$. The upper (lower) panel shows the PDF on a linear (log) scale. We can see that the PDF is significantly non-Gaussian with an exponential drop-off to the left of mean and a power-law to the right. We provide a fitting formula for $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ in the text.

\[ \langle E_1 \rangle = 1. \]  Furthermore, since we have already established that $P(\hat{f}_{nl})$ approaches a Gaussian in the large $l_{\text{max}}$ limit if $f_{nl} = 0$, we know that, to leading order, the non-Gaussian shape of $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ for $f_{nl} \neq 0$ is being generated by $E_1$.

Some of the statistics associated with $E_1$ have already been explored in Ref. [14]. There it is noted that the variance of $\hat{f}_{nl}$ is dominated by $E_1$ in the high $S/N$ limit leading to a slower scaling of the $S/N$ than the $l_{\text{max}}^{-2} \ln^{-1}(l_{\text{max}})$ scaling expected if the estimator saturated the Cramer-Rao bound [14]. We explored the same limit using our Monte Carlo realizations, as shown in Fig. 3 and find the same qualitative trend but with a different dependence on $l_{\text{max}}$. Ref. [14] found $\langle (\Delta E_1)^2 \rangle \propto \ln^{-2}(l_{\text{max}})$ whereas our simulations show $\langle (\Delta E_1)^2 \rangle \propto \ln^{-3}(l_{\text{max}})$. We have checked the scaling found with our simulations by computing the variance analytically, as we further discuss in Appendix B. Fig. 3 shows the agreement between our analytic calculation (solid curve) and simulations (data points).

Our simulations allow us to generate the full PDF for $E_1$, not just the variance. Fig. 4 shows this PDF for

\[ \langle E_1 \rangle = 1. \]  Furthermore, since we have already established that $P(\hat{f}_{nl})$ approaches a Gaussian in the large $l_{\text{max}}$ limit if $f_{nl} = 0$, we know that, to leading order, the non-Gaussian shape of $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ for $f_{nl} \neq 0$ is being generated by $E_1$.

Some of the statistics associated with $E_1$ have already been explored in Ref. [14]. There it is noted that the variance of $\hat{f}_{nl}$ is dominated by $E_1$ in the high $S/N$ limit leading to a slower scaling of the $S/N$ than the $l_{\text{max}}^{-2} \ln^{-1}(l_{\text{max}})$ scaling expected if the estimator saturated the Cramer-Rao bound [14]. We explored the same limit using our Monte Carlo realizations, as shown in Fig. 3 and find the same qualitative trend but with a different dependence on $l_{\text{max}}$. Ref. [14] found $\langle (\Delta E_1)^2 \rangle \propto \ln^{-2}(l_{\text{max}})$ whereas our simulations show $\langle (\Delta E_1)^2 \rangle \propto \ln^{-3}(l_{\text{max}})$. We have checked the scaling found with our simulations by computing the variance analytically, as we further discuss in Appendix B. Fig. 3 shows the agreement between our analytic calculation (solid curve) and simulations (data points).

Our simulations allow us to generate the full PDF for $E_1$, not just the variance. Fig. 4 shows this PDF for

\[ \langle E_1 \rangle = 1. \]  Furthermore, since we have already established that $P(\hat{f}_{nl})$ approaches a Gaussian in the large $l_{\text{max}}$ limit if $f_{nl} = 0$, we know that, to leading order, the non-Gaussian shape of $P(\hat{f}_{nl}; f_{nl}, l_{\text{max}})$ for $f_{nl} \neq 0$ is being generated by $E_1$.

Some of the statistics associated with $E_1$ have already been explored in Ref. [14]. There it is noted that the variance of $\hat{f}_{nl}$ is dominated by $E_1$ in the high $S/N$ limit leading to a slower scaling of the $S/N$ than the $l_{\text{max}}^{-2} \ln^{-1}(l_{\text{max}})$ scaling expected if the estimator saturated the Cramer-Rao bound [14]. We explored the same limit using our Monte Carlo realizations, as shown in Fig. 3 and find the same qualitative trend but with a different dependence on $l_{\text{max}}$. Ref. [14] found $\langle (\Delta E_1)^2 \rangle \propto \ln^{-2}(l_{\text{max}})$ whereas our simulations show $\langle (\Delta E_1)^2 \rangle \propto \ln^{-3}(l_{\text{max}})$. We have checked the scaling found with our simulations by computing the variance analytically, as we further discuss in Appendix B. Fig. 3 shows the agreement between our analytic calculation (solid curve) and simulations (data points).

Our simulations allow us to generate the full PDF for $E_1$, not just the variance. Fig. 4 shows this PDF for
various choices of $l_{\text{max}}$ (thin solid lines). An important conclusion from Fig. 4 is that the shape of the PDF approaches a universal form in the $l_{\text{max}} \gg 1$ limit. We provide a fit to the PDF (thick red dashed line), accurate to $\sim 10\%$ ($40\%$) out to three (four) times the root variance, using the fitting formula

$$\log[F(x)] = N - \begin{cases} -\frac{(x-x_p)^2}{2\sigma^2}, & x \leq x_p \\ -\frac{c}{\sigma^2} \left( \frac{(x-x_p)^2}{\sigma^2} + c^2 - c \right), & x > x_p, \end{cases}$$

where $N = \sqrt{2/\pi \sigma + c \exp[c^2/\sigma^2]} K_1(c^2/\sigma^2)$ and $K_1(x)$ is a modified Bessel function of the first kind, $c$ quantifies the non-Gaussianity of the distribution (and approaches a Gaussian in the $c \to \infty$ limit) and $x_p$ is the value of $(E_1 - \langle E_1 \rangle)/\sigma E_1$ at the peak of the distribution. The red curve in Fig. 4 shows Eq. (17) with parameter values $x_p = -0.22$, $\sigma = 0.80$, and $c = 0.91$.

We are now in a position to write down a semi-analytic expression for $P(\widehat{f}_{nl}; f_{nl}, l_{\text{max}})$, accurate to $\sim 10\%$ ($40\%$) out to three (four) times the root variance, as a function of $f_{nl}$ and $l_{\text{max}}$. Letting $\sigma_0$ and $\sigma_1$ denote the standard deviations of the distributions for $E_0$ and $E_1$ respectively we have

$$\sigma_0^2 \approx \frac{1}{8A f_{nl}^2 \ln(l_{\text{max}})}$$

$$\sigma_1^2 \approx \frac{9f_{nl}^2}{2 \ln^3(l_{\text{max}})}$$

A good approximation to the PDF of $\widehat{f}_{nl}$ is provided by the convolution of the PDF of $E_0$ and $f_{nl}E_1$:

$$P(\widehat{f}_{nl}; f_{nl}, l_{\text{max}}) \approx \frac{4}{9\sqrt{2}\pi \sigma_0 \sigma_1} \times \int_{-\infty}^{\infty} G_0(\widehat{f}_{nl} - x) F([x - f_{nl}]/\sigma_1) dx,$$

where $G_0(x)$ is a Gaussian with zero mean and standard deviation $\sigma_0$ and $F([x - f_{nl}]/\sigma_1)$ is given by Eq. (17) with $x_p = -0.22$, $\sigma = 0.80$, and $c = 0.91$.

To obtain an analytic expression for the PDF we can approximate the convolution in Eq. (21) to write

$$P(\widehat{f}_{nl}; f_{nl}, l_{\text{max}}) \approx \frac{2}{9} \exp\left[-\frac{X^2}{2(\sigma_0^2 + \sigma_1^2 \sigma_1^2)}\right] \sqrt{\frac{1}{\sigma_1^2}} \left\{ \sigma_1 \left(1 + \text{erf}\left[\frac{\sigma_0^2 + \sigma_1^2 \sigma_1^2(X + \sigma_1)}{\sqrt{2\sigma_0 \sigma_1 \sigma_1^2}}\right]\right) \right\} \exp\left[\frac{1}{2} \left(\frac{c^2 \sigma_0^2}{\sigma_1 \sigma_1^2} + \frac{2c(X + \sigma_1^2)}{\sigma_1 \sigma_1^2} + \frac{X^2}{\sigma_0^2 + \sigma_1^2 \sigma_1^2}\right)\right],$$

where $X = f_{nl} + x_p \sigma_1 - \widehat{f}_{nl}$.

Another useful way of quantifying the non-Gaussian shape of $P(\widehat{f}_{nl}; f_{nl}, l_{\text{max}})$ is to measure its skewness, $\langle (\Delta \widehat{f}_{nl})^3 \rangle / \sigma_{\widehat{f}_{nl}}^3$, as a function of $f_{nl}$ and $l_{\text{max}}$. We show this in Fig. 5 for $f_{nl} = 100$. An analytic fit to the skewness is given by

$$\langle (\Delta \widehat{f}_{nl})^3 \rangle / \sigma_{\widehat{f}_{nl}}^3 \equiv \left(\frac{f_{nl}}{100}\right)^3 \times \left(1 + 3.7 \exp\left[-(l_{\text{max}} - 5.1)/740\right] - 0.26\right),$$

with the variance of the distribution, $\sigma_{\widehat{f}_{nl}}^2$, given by

$$\sigma_{\widehat{f}_{nl}}^2 \approx \frac{1}{8A f_{nl}^2 \ln(l_{\text{max}})} \left[1 + \frac{36Af_{nl}^2 \sigma_{\widehat{f}_{nl}}^2}{\ln^3(l_{\text{max}})}\right].$$

Finally, we note that the shape of $P(\widehat{f}_{nl}; f_{nl}, l_{\text{max}})$ departs significantly from a Gaussian when $\sigma_0 \approx \sigma_1$. This occurs when

$$f_{nl} A^{1/2} \gtrsim \frac{\ln(l_{\text{max}})}{6l_{\text{max}}}$$

Therefore, for the Planck satellite (i.e., $l_{\text{max}} = 3000$) the non-Gaussian features of $P(\widehat{f}_{nl}; f_{nl}, l_{\text{max}})$ for the NHMV estimator are significant if $f_{nl} \gtrsim O(10)$. Thus, given that Planck is expected to measure $f_{nl}$ with a variance $\sigma \approx 5$, these PDFs may need to be taken into account to assign a precise confidence region with Planck data.

IV. THE PDF OF AN IMPROVED ESTIMATOR WHEN $f_{nl} \neq 0$

As we saw in the previous Section the standard (null-hypothesis) minimum-variance estimator $\widehat{f}_{nl}$ is constructed under the null hypothesis, so its variance is strictly minimized only when applied to maps with $f_{nl} = 0$. In particular, the variance of $\widehat{f}_{nl}$ is given in Eq. (23) so that when $36Af_{nl}^2 \sigma_{\widehat{f}_{nl}}^2 \ln^3(l_{\text{max}})$ is $\gtrsim 1$, the variance scales as the $\ln^3(l_{\text{max}})$, as opposed to the $\ln^2(l_{\text{max}})$ when $f_{nl} = 0$. This indicates that when $f_{nl} \neq 0$ there may be other estimators with smaller variances.

For a flat-sky and under the Sachs-Wolfe approximation Ref. [14] introduced an improved estimator for
f_{nl} \neq 0$ which has a variance that continues to decrease as $1/\ln(\ln(l_{max}))$ in the high signal-to-noise limit. To achieve this scaling they introduced a realization-dependent normalization,

$$N \equiv \sigma_{f_{nl}}^2 \sum_{l_1+l_2+l_3=0} \chi_{l_1}^* T_{l_1} T_{l_2} T_{l_3} B(l_1,l_2,l_3),$$

where

$$\chi_{l_1}^* \equiv \sum_k T_{l_1-k} T_k.$$

By construction $\langle N \rangle = 1$. They then define a new estimator constructed under the non-null hypothesis:

$$\hat{f}_{nl}^n \equiv \frac{\hat{f}_{nl}}{N}.$$ (27)

To explore the properties of the PDF of $\hat{f}_{nl}^n$, we expand the normalization as $N \approx N_0 + f_{nl} N_1 + \cdots$ and write

$$\hat{f}_{nl}^n \approx \frac{E_0}{N_0} + f_{nl} \frac{E_1 N_0 - E_0 N_1}{N_0^2} + \cdots,$$ (28)

$$\equiv \bar{E}_0 + f_{nl} \bar{E}_1 + \cdots.$$ (29)

In order to determine the shape of $P(\hat{f}_{nl})$, we computed $P(\bar{E}_0)$ and $P(\bar{E}_1)$ for various values of $l_{max}$. We found, as in the $\hat{f}_{nl}$ case, that these PDFs approach asymptotic shapes in the $l_{max} \gg 1$ limit. We show these PDFs in Fig. 6 determined by $10^6$ realizations for $l_{max} = 25$. It is clear that $P(\bar{E}_0)$ is very well approximated by a Gaussian, whereas $P(\bar{E}_1)$ has significant non-Gaussian wings. As in the $P(\hat{f}_{nl})$ case, this implies that the level of non-Gaussianity in $P(\hat{f}_{nl}^n)$ is significant only when the ratio $f_{nl}^2 \langle (\Delta E_1)^2 \rangle / \langle (\Delta E_0)^2 \rangle > 1$. Our simulations show

$$\langle (\Delta E_0)^2 \rangle \approx \frac{1}{8A r_{max}^2 \ln(l_{max})},$$ (30)

$$\langle (\Delta E_1)^2 \rangle \approx \frac{9 \ln^2(l_{max})}{l_{max}^3},$$ (31)

so that the PDF will be significantly non-Gaussian when

$$f_{nl} A^{1/2} \gtrsim \frac{1}{3} \left[ \frac{l_{max}}{8 \ln(l_{max})} \right]^{1/2}.$$ (32)

Therefore, for Planck (with $l_{max} = 3000$) $P(\hat{f}_{nl}^n; f_{nl}, l_{max})$ will be significantly non-Gaussian only if $f_{nl} \gtrsim O(1000)$. Since this has already been ruled out by observations [10], we conclude that $P(\hat{f}_{nl}^n; f_{nl}, l_{max})$ will be effectively Gaussian.

V. DISCUSSION

Here we have argued that the PDF for non-Gaussianity estimators cannot be assumed to be Gaussian, since the number of triplets used to construct these estimators may greatly exceed the number $N_{pix}$ of measurements. The 99.7% confidence-level interval cannot safely be assumed to be $3 \times$ the 66.5% confidence-level interval. We
found, however, that the standard minimum-variance estimator $f_{nl}$ constructed under the null hypothesis is well-approximated by a Gaussian distribution in the $l_{\text{max}} \gg 1$ limit if the null hypothesis is correct (i.e., when applied to purely Gaussian maps).

We then calculated the same PDF $P(f_{nl})$ under the hypothesis that the true value of $f_{nl}$ is non-zero. We find that the PDF is non-Gaussian in this case, skewed to large $f_{nl}$ if $f_{nl} > 0$ and vice versa for $f_{nl} < 0$. The PDF for small positive or negative $f_{nl}$ is significantly smaller for $f_{nl} > 0$ than the Gaussian PDF with the same variance. Thus, for example, if the NHMV estimator gives $f_{nl} > 0$, it may actually rule out $f_{nl} = 0$ with a smaller statistical significance than would be inferred assuming a Gaussian distribution of the same variance. For Planck (with $l_{\text{max}} \approx 3000$) we find that the non-Gaussian shape of $P(f_{nl})$ is significant if $f_{nl} \gtrsim O(10)$. Thus, the non-Gaussian shape of the PDF may need to be taken into account, even in case of a null result, to assign a precise 99.7% confidence-level upper (or lower, for $f_{nl} < 0$) limit to $f_{nl}$. We also provide, in Eq. (21), an analytic fit to these PDFs.

The non-Gaussian shape of $P(f_{nl})$ when $f_{nl} \neq 0$ is accompanied by a variance that decreases only logarithmically with increasing $l_{\text{max}}$. Because of this, Ref. [14] constructed an improved estimator under the $f_{nl} \neq 0$ hypothesis with a variance that saturates the Cramer-Rao bound and continues to decrease as $1/l_{\text{max}}^2 \log(l_{\text{max}})$. We found that for observationally allowed values of $f_{nl}$ this improved estimator has a PDF that is well approximated by a Gaussian shape. However, this estimator has only been defined under the Sachs-Wolfe limit and it is not immediately clear how it should be generalized to be applied to actual data. An alternative, Bayesian, approach to measuring $f_{nl}$ which also saturates the Cramer-Rao bound in the presence of $f_{nl} \neq 0$ is presented in Ref. [15].

We have restricted our attention to the bispectrum in the local model, but the PDF must be similarly determined for the non-Gaussianity parameter for bispectra with other shape dependences; e.g., the equilateral model [19, 20] or that which arises with self-ordering scalar fields [21]. It should also be interesting to explore the PDF for this estimator would take a prohibitively long time to evaluate for a significant number of realizations, especially since we are interested in probing the shape of the PDF far into the tail of the distribution ($\sim 3-4\sigma$).

As discussed at length in Ref. [25] this is even more of a problem when measuring non-Gaussianity on the full sky where the number of operations scales as $N_{\text{pix}}^{5/2}$. In order to make the problem tractable Ref. [25] rewrites $\hat{f}_{nl}$ in terms of real-space quantities reducing the number of operations to $N_{\text{pix}}^{3/2}$.

We can do the same for $\hat{f}_{nl}$ in the flat-sky approximation. Noting that

$$\delta_{l_1+l_2+l_3,0} = \int d^2\theta \Theta(\vec{l}_1+\vec{l}_2+\vec{l}_3),$$

Appendix A: Computing non-Gaussianity estimators using FFTs

We are interested in using Monte Carlo simulations to determine the shape of the PDF of $\hat{f}_{nl}$ as a function of the fiducial choice of $f_{nl}$ and the number $N_{\text{pix}}$ of pixels measured in a given observation. Applying the estimator in Eq. (6) to the local-model bispectrum [Eq. (11)] it can be rewritten

$$\hat{f}_{nl} = \sigma_{f_{nl}}^2 \sum_{|i_1+i_2+i_3|=0} \frac{T_{i_1}T_{i_2}T_{i_3}}{12\Omega^2} C_{l_1}.$$  \hspace{1cm} (A1)$$

The estimator in Eq. (A1) takes $N_{\text{pix}}^2$ operations to evaluate. Since current CMB observations have $N_{\text{pix}} \sim 10^6$ this estimator would take a prohibitive amount of time to evaluate for a significant number of realizations, especially since we are interested in probing the shape of the PDF far into the tail of the distribution ($\sim 3-4\sigma$).

Acknowledgments

We thank D. Babich, C. Hirata, and I. Wehus for useful discussions. TLS is supported by the Berkeley Center of Cosmological Physics. MK thanks the support of the Miller Institute and the hospitality of the Department of Physics at the University of California, where part of this work was completed. MK was supported at Caltech by DoE DE-FG03-92-ER40701, NASA NNX10AD04G, and the Gordon and Betty Moore Foundation. BDW was supported by NASA/JPL subcontract 1413479, and NSF grants AST 07-08849, AST 09-08693 ARRA, and AST 09-08902 during this work.
and writing

\[ A(\hat{\theta}) = \frac{1}{2\pi} \sum_f e^{i\hat{\theta} \hat{f}_l} \]  \hspace{1cm} (A3)\]

\[ B(\hat{\theta}) = \frac{1}{2\pi} \sum_f e^{i\hat{\theta} \hat{f}_l} \hat{C}_l \]  \hspace{1cm} (A4)\]

\[ \hat{f}_{nl} \] can be written

\[ \hat{f}_{nl} = \Omega \sigma_{f_{nl}}^2 \int \frac{d^3 \hat{\theta}}{\Omega} A^2(\hat{\theta}) B(\hat{\theta}). \]  \hspace{1cm} (A5)\]

Next, in order to compute the integral in Eq. (A5) we use the Nyquist sampling theorem and the fact that both \( A(\hat{\theta}) \) and \( B(\hat{\theta}) \) have finite Fourier spectra (truncated at a maximum frequency \( l_{\text{max}} \)). This allows us to rewrite the integral as a discrete sum

\[ \hat{f}_{nl} = \frac{\Omega \sigma_{f_{nl}}^2}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} A^2 \left( 2\pi \frac{i-1}{N}, 2\pi \frac{j-1}{N} \right) \times B \left( 2\pi \frac{i-1}{N}, 2\pi \frac{j-1}{N} \right), \]  \hspace{1cm} (A6)\]

where \( N \equiv 2(2l_{\text{max}} + 1) \).

Since Eqs. (A3) and (A4) are discrete inverse Fourier transforms we can use a fast Fourier transform (FFT) algorithm so that the number of operations scale as \( N_{\text{pix}} \ln(N_{\text{pix}}) \).

We can use the same computational trick when evaluating the non-Gaussian contribution for each realization by also employing a forward FFT in order to compute the convolution in Eq. (12).

Appendix B: Analytic calculation of \( \langle (\Delta E_1)^2 \rangle \)

In order to verify that our simulations are correct we performed an analytic calculation of the variance of \( E_1 \) [Eq. (16)] defined by

\[ E_1 = \sigma_{f_{nl}}^2 \sum_{i_1+i_2+i_3=0} \frac{\delta_{l_1}^2 \delta_{l_2}^2 \delta_{l_3}^2}{2 C_{l_1} C_{l_2} C_{l_3}} B(l_1, l_2, l_3). \]  \hspace{1cm} (B1)\]

A straightforward but tedious calculation shows that the variance is given by

\[ \langle (\Delta E_1)^2 \rangle = 9 \sigma_{f_{nl}}^4 (A_1 + 8 A_2 + A_3 + 4 A_4), \]  \hspace{1cm} (B2)\]

where

\[ A_1 = \sum_{(i),(k)} \frac{B(l) B(k)}{C_l C_k (l_1 + k_1, 0^+)} \]  \hspace{1cm} (B3)\]

\[ A_2 = \sum_{(i),(k)} \frac{B(l) B(k)}{C_l C_k (l_3 + k_3, 0^+)} \]  \hspace{1cm} (B4)\]

\[ A_3 = \sum_{(i)} \frac{B(l)^2}{C_l C_{l_1} C_{l_3} \sum_{|m|<1} C_{|l_1-m|} C_m} \]  \hspace{1cm} (B5)\]

\[ A_4 = \sum_{(i),(k)} \frac{B(l) B(k)}{C_l C_k C_{|l_1+k_1|} C_{|l_3+k_3|}} \]  \hspace{1cm} (B6)\]

where \( \{i\} \) indicates the sum is over \( l_1 + l_2 + l_3 = 0 \) and \( B(l) \equiv B(l_1, l_2, l_3) \). Computing these terms as a function of \( l_{\text{max}} \) we find that the variance is well-fit by the function

\[ \langle (\Delta E_1)^2 \rangle = \frac{14.0 l_{\text{max}}^{4.33}}{\ln^5 l_{\text{max}}} \]  \hspace{1cm} (B7)\]

In Fig. 3 we show how that this analytic calculation of the \( \langle (\Delta E_1)^2 \rangle \) is reproduced by the results of the Monte Carlo simulations.

[1] A. H. Guth and S. Y. Pi, Phys. Rev. Lett. 49, 1110 (1982); A. A. Starobinsky, Phys. Lett. B 117, 175 (1982); J. M. Bardeen, P. J. Steinhardt and M. S. Turner, Phys. Rev. D 28, 679 (1983).

[2] T. Falk, R. Rangarajan and M. Srednicki, Astrophys. J. 403, L1 (1993) [arXiv:astro-ph/9208001]; A. Gangui et al., Astrophys. J. 430, 447 (1994) [arXiv:astro-ph/9312033]; A. Gangui, Phys. Rev. D 50, 3684 (1994) [arXiv:astro-ph/9406014]; J. M. Maldacena, JHEP 0305, 013 (2003) [arXiv:astro-ph/0210603]; V. Acquaviva, N. Bartolo, S. Matarrese and A. Riotto, Nucl. Phys. B 667, 119 (2003) [arXiv:astro-ph/0209156]; D. Babich, P. Creminelli and M. Zaldarriaga, JCAP 0408, 009 (2004) [arXiv:astro-ph/0405356]; P. Creminelli et al., JCAP 0605, 004 (2006) [arXiv:astro-ph/0509029]; P. Creminelli et al., JCAP 0703, 005 (2007) [arXiv:astro-ph/0610600].

[3] X. c. Luo, Astrophys. J. 427, L71 (1994) [arXiv:astro-ph/9312004]; L. Verde et al., Mon. Not. Roy. Astron. Soc. 313, L141 (2000) [arXiv:astro-ph/9906301]; E. Komatsu and D. N. Spergel, Phys. Rev. D 63, 063002 (2001) [arXiv:astro-ph/0005036].

[4] N. Bartolo, E. Komatsu, S. Matarrese and A. Riotto, Phys. Rept. 402, 103 (2004) [arXiv:astro-ph/0406309].

[5] T. J. Allen, B. Grinstein and M. B. Wise, Phys. Lett. B 197, 66 (1987); L. A. Kofman and D. Y. Pogosian, Phys. Lett. B 214, 508 (1988); D. S. Salopek, J. R. Bond and J. M. Bardeen, Phys. Rev. D 40, 1753 (1989); A. D. Linde and V. F. Mukhanov, Phys. Rev. D 56, 535 (1997) [arXiv:astro-ph/9610219]; P. J. E. Peebles,
N. Dalal, O. Dore, D. Huterer and A. Shirokov, Astrophys. J. 510, 523 (1999) [arXiv:astro-ph/9805194]; P. J. E. Peebles, Astrophys. J. 510, 531 (1999) [arXiv:astro-ph/9805212].

S. Mollerach, Phys. Rev. D 42, 313 (1990); A. D. Linde and V. F. Mukhanov, Phys. Rev. D 56, 535 (1997) [arXiv:astro-ph/9610219]; D. H. Lyth and D. Wands, Phys. Lett. B 524, 5 (2002) [arXiv:hep-ph/0110002]; T. Moroi and T. Takahashi, Phys. Lett. B 522, 215 (2001) [Erratum-ibid. B 539, 303 (2002)] [arXiv:hep-ph/0110096]; D. H. Lyth, C. Ungarelli and D. Wands, Phys. Rev. D 67, 023503 (2003) [arXiv:astro-ph/0208055]; K. Ichikawa et al., arXiv:0802.4138 [astro-ph]; K. Enqvist, S. Nurmi, O. Tuulia and T. Takahashi, arXiv:0912.4557 [astro-ph.CO]; K. Enqvist and T. Takahashi, JCAP 0912, 001 (2009) [arXiv:0909.5362 [astro-ph.CO]]; K. Enqvist and T. Takahashi, JCAP 0809, 012 (2008) [arXiv:0807.3069 [astro-ph]]; K. Enqvist and S. Nurmi, JCAP 0510, 013 (2005) [arXiv:astro-ph/0508573]; A. L. Erickcek, M. Kamionkowski and S. M. Carroll, arXiv:0806.0377 [astro-ph]; A. L. Erickcek, C. M. Hirata and M. Kamionkowski, Phys. Rev. D 80, 083507 (2009) [arXiv:0907.0705 [astro-ph.CO]].

L. M. Wang and M. Kamionkowski, Phys. Rev. D 61, 063504 (2000) [arXiv:astro-ph/9907431].

S. Hannestad, T. Haugboelle, P. R. Jarnhus and M. S. Sloth, JCAP 1008, 001 (2010) [arXiv:0912.3527 [hep-ph]].

E. Komatsu et al. [WMAP Collaboration], Astrophys. J. Suppl. 148, 119 (2003) [arXiv:astro-ph/0302223]; E. Komatsu et al. [WMAP Collaboration], Astrophys. J. Suppl. 192, 18 (2011) [arXiv:1001.4538 [astro-ph.CO]].

N. Dalal, O. Dore, D. Huterer and A. Shirokov, Phys. Rev. D 77, 123514 (2008) [arXiv:0710.4560 [astro-ph]]; A. Slosar, C. Hirata, U. Seljak, S. Ho and N. Padmanabhan, JCAP 0808, 031 (2008) [arXiv:0805.3580 [astro-ph]]; S. Matarrese and L. Verde, Astrophys. J. 677, L77 (2008) [arXiv:0801.4826 [astro-ph]]; C. Carbone, L. Verde and S. Matarrese, Astrophys. J. 684, L1 (2008) [arXiv:0806.1950 [astro-ph]]; J. Q. Xia, M. Viel, C. Baccigalupi, G. De Zotti, S. Matarrese and L. Verde, Astrophys. J. 717, L17 (2010) [arXiv:1003.3451 [astro-ph.CO]]; J. Q. Xia, A. Bonaldi, C. Baccigalupi, G. De Zotti, S. Matarrese, L. Verde and M. Viel, JCAP 1008, 013 (2010) [arXiv:1007.1969 [astro-ph.CO]]; L. Verde and S. Matarrese, Astrophys. J. 706, L91 (2009) [arXiv:0909.3224 [astro-ph.CO]]; F. Schmidt and M. Kamionkowski, Phys. Rev. D 82, 103002 (2010) [arXiv:1008.6638 [astro-ph.CO]].

A. P. S. Yadav and B. D. Wandelt, Phys. Rev. Lett. 100, 181301 (2008) [arXiv:0712.1148 [astro-ph]].

P. A. R. Ade et al. [Planck Collaboration], arXiv:1101.2022 [astro-ph.IM].

F. Elsner and B. D. Wandelt, Astrophys. J. Suppl. 184, 264 (2009) [arXiv:0909.0009 [astro-ph.CO]].

P. Creminelli, L. Senatore and M. Zaldarriaga, JCAP 0703, 019 (2007) [arXiv:astro-ph/0606001].

F. Elsner and B. D. Wandelt, Astrophys. J. 724, 1262 (2010) [arXiv:1010.1254 [astro-ph.CO]].

D. Babich and M. Zaldarriaga, Phys. Rev. D 70, 083005 (2004) [arXiv:astro-ph/0408455].

M. Kamionkowski, T. L. Smith and A. Heavens, Phys. Rev. D 83, 023007 (2011) [arXiv:1010.0251 [astro-ph.CO]].

G. Efstathiou, Mon. Not. Roy. Astron. Soc. 348, 885 (2004) [arXiv:astro-ph/0310207]; C. J. Copi, D. Huterer, D. J. Schwarz and G. D. Starkman, [arXiv:1103.3505 [astro-ph.CO]].

G. R. Dvali and S. H. H. Tye, Phys. Lett. B 450, 72 (1999) [arXiv:hep-ph/9812483]; P. Creminelli, JCAP 0310, 003 (2003) [arXiv:astro-ph/0306122]; M. Alishahiha, E. Silverstein and D. Tong, Phys. Rev. D 70, 123504 (2004) [arXiv:hep-th/0404084].

D. Babich, P. Creminelli and M. Zaldarriaga, JCAP 0408, 009 (2004) [arXiv:astro-ph/0405356]; P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark and M. Zaldarriaga, JCAP 0605, 004 (2006) [arXiv:astro-ph/0509029]; P. Creminelli, L. Senatore, M. Zaldarriaga, and M. Tegmark, JCAP 0703, 005 (2007) [arXiv:astro-ph/0610600].

D. G. Figueroa, R. R. Caldwell and M. Kamionkowski, Phys. Rev. D 81, 123504 (2010) [arXiv:1003.0672 [astro-ph.CO]].

N. Kogo and E. Komatsu, Phys. Rev. D 73, 083007 (2006) [arXiv:astro-ph/0602099].

J. Smidt, A. Amblard, A. Cooray, A. Heavens, D. Munshi and P. Serra, Phys. Rev. D 81, 123007 (2010) [arXiv:1011.5026 [astro-ph.CO]].

T. L. Smith and M. Kamionkowski, in prep.

E. Komatsu, D. N. Spergel and B. D. Wandelt, Astrophys. J. 634, 14 (2005) [arXiv:0305189].