Probability distribution for non-Gaussianity estimators

Tristan L. Smith, Marc Kamionkowski, Benjamin D. Wandelt

PHYSICAL REVIEW D **84,** 063013 (2011)

Probability distribution for non-Gaussianity estimators

Tristan L. Smith,[1] Marc Kamionkowski,[2] and Benjamin D. Wandelt[3,4]

[1]*Berkeley Center for Cosmological Physics, Physics Department, University of California, Berkeley, California 94720, USA*
[2]*California Institute of Technology, Mail Code 350-17, Pasadena, California 91125, USA*
[3]*UPMC Université Paris 06, Institut d'Astrophysique de Paris, 98 bis, boulevard Arago, 75014 Paris, France*
[4]*Departments of Physics and Astronomy, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801, USA*

(Received 25 April 2011; published 23 September 2011)

One of the principle efforts in cosmic microwave background (CMB) research is measurement of the parameter $f_{\rm nl}$ that quantifies the departure from Gaussianity in a large class of nonminimal inflationary (and other) models. Estimators for $f_{\rm nl}$ are composed of a sum of products of the temperatures in three different pixels in the CMB map. Since the number $\sim N_{\rm pix}^2$ of terms in this sum exceeds the number $N_{\rm pix}$ of measurements, these $\sim N_{\rm pix}^2$ terms cannot be statistically independent. Therefore, the central-limit theorem does not necessarily apply, and the probability distribution function (PDF) for the $f_{\rm nl}$ estimator does not necessarily approach a Gaussian distribution for $N_{\rm pix} \gg 1$. Although the variance of the estimators is known, the significance of a measurement of $f_{\rm nl}$ 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 ($f_{\rm nl} = 0$), and an improved estimator with a smaller variance for $f_{\rm nl} \neq 0$. While the PDF for the null-hypothesis estimator is very nearly Gaussian when the true value of $f_{\rm nl}$ is zero, the PDF becomes significantly non-Gaussian when $f_{\rm nl} \neq 0$. In this case we find that the PDF for the null-hypothesis estimator $\widehat{f_{\rm nl}}$ is skewed, with a long non-Gaussian tail at $\widehat{f_{\rm nl}} > |f_{\rm nl}|$ and less probability at $\widehat{f_{\rm nl}} < |f_{\rm nl}|$ 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 $f_{\rm nl}$. We discuss briefly the implications for trispectrum (and other higher-order correlation) estimators.

DOI: 10.1103/PhysRevD.84.063013 PACS numbers: 98.70.Vc, 98.80.Cq

I. INTRODUCTION

The simplest single-field slow-roll inflation models predict that primordial perturbations should be nearly Gaussian [1], but with predictably small departures from Gaussianity [2]. This is often quantified through the non-Gaussianity parameter $f_{\rm nl}$ defined by [3]

$$\Phi = \phi + f_{\rm nl}(\phi^2 - \langle\phi^2\rangle), \tag{1}$$

where Φ is the gravitational potential and ϕ a Gaussian random field. Standard single-field slow-roll inflation predicts $f_{\rm nl} \ll 1$ for the primordial field (although nonlinear evolution of the density field may produce $f_{\rm nl} \sim 1$ 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 $f_{\rm nl}$. Measurement of $f_{\rm nl}$ 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 $|f_{\rm nl}| \lesssim 100$ [9,10]. The plot has thickened with a suggestion [11] (not universally accepted) that the WMAP 3-year data prefer (at the 2.8σ level) $f_{\rm nl} \neq 0$, with a best-fit value $f_{\rm nl} \simeq 80$ (a less significant result of 1.5σ is found in an analysis of the most recent WMAP data release [9]). The Planck satellite [12] is expected to achieve a sensitivity of $f_{\rm nl} \sim 5$.

In this paper, we address the following question: What is the probability distribution function (PDF) $P(\widehat{f_{\rm nl}})$ for an estimator $\widehat{f_{\rm nl}}$ 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 $f_{\rm nl}$ may be different than 3 times the standard deviation for $f_{\rm nl}$. 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 $\widehat{f_{\rm nl}}$, but it also makes a prediction for the detailed functional form of the PDF $P(\widehat{f_{\rm nl}})$. The consistency of a given measurement of $\widehat{f_{\rm nl}}$ with a theoretical prediction for $f_{\rm nl}$ depends on knowledge of the shape of $P(\widehat{f_{\rm nl}})$. Thus, for example, we often evaluate or forecast the standard error $\sigma_{f_{\rm nl}}$ with which a given measurement will recover the true value of $f_{\rm nl}$ and then simply assume that the error is Gaussian. If so, then with $\sigma_{f_{\rm nl}} = 10$, for example, a measurement of $\widehat{f_{\rm nl}} = 30$ would represent a 3σ departure from $f_{\rm nl} = 0$ and a measurement $\widehat{f_{\rm nl}} = 0$ would represent a 3σ departure from $f_{\rm nl} = 30$. However, if the PDF depends on the true value $f_{\rm nl}$, and if that distribution is non-Gaussian, then it may be that a measurement $\widehat{f_{\rm nl}} = 30$ could easily be consistent with a true value $f_{\rm nl} = 0$, while a measurement $\widehat{f_{\rm nl}} = 0$ could be

inconsistent with $f_{\rm nl} = 30$ with a confidence greater than "3σ." We will see below that something similar to this actually occurs with measurements of $f_{\rm nl}$.

This question is particularly important for measurements of non-Gaussianity (as opposed, for example, for the CMB power spectrum), because $\widehat{f_{\rm nl}}$ 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 $N_{\rm pix}$ pixels. There are then $\sim N_{\rm pix}^2$ terms in the $f_{\rm nl}$ estimator (after restrictions imposed by statistical isotropy). While these terms may have zero covariance, they are not statistically independent; there is no way to construct $N_{\rm pix}^2$ statistically independent quantities from $N_{\rm pix}$ measurements. The conditions required for the validity of the central-limit theorem are therefore not met, and $P(\widehat{f_{\rm nl}})$ will not necessarily approach a Gaussian in the $N_{\rm pix} \gg 1$ limit.

The PDF can be obtained from Monte Carlo simulations, but the simulations are very computationally intensive (e.g., see Ref. [13]). The number of Monte Carlo realizations is thus usually limited to the number, $\lesssim 1000$, 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 the results shown in Fig. 8 of Ref. [13] show hints of a non-Gaussian PDF, simulations done up until now do not include enough realizations to precisely map the functional form of $P(\widehat{f_{\rm nl}})$. The number of realizations required to ultimately map the 4σ, 5σ, etc. ranges will be prohibitive, especially since the simulations will need to be rerun repeatedly to determine how the error ranges depend on cosmological parameters, instrument-noise properties, scanning strategies, etc., and they then must be run for multiple theoretical values $f_{\rm nl}$.

Work along these lines was begun in Ref. [14], wherein it was shown that the variance of the distribution $P(\widehat{f_{\rm nl}})$ may have a strong dependence on the true underlying value of $f_{\rm nl}$. More precisely, they evaluated the variance of the estimator designed to have the minimum variance under the null hypothesis $f_{\rm nl} = 0$ (which we refer to frequently below as the "null-hypothesis minimum-variance" estimator, or NHMV estimator), and showed that the variance of this NHMV estimator increases as $f_{\rm nl}^2$ increases. They then constructed an alternative estimator $\widehat{f_{nl}^n}$, which we call the CSZ estimator,[1] which has a PDF with a variance that saturates the Cramer-Rao bound up to corrections of order $f_{\rm nl}^2$. Still, as we have argued above, the consistency of a hypothesis with a measurement requires full knowledge of the PDF of whatever estimator is used in the analysis.

To address these questions, we calculate the PDF for an ideal (no-noise) map to understand the irreducible PDF introduced simply by cosmic variance under the Sachs-Wolfe approximation and on a flat sky. We hope that lessons learned about $P(\widehat{f_{\rm nl}})$ in this ideal case may help interpret and understand current/forthcoming results and assess the validity of full-experiment simulations.

We calculate these PDFs by using Monte Carlo realizations of numerous no-noise flat-sky CMB maps. The first order of business with a map will be to determine whether a given map is consistent or inconsistent with the null hypothesis $f_{\rm nl} = 0$. Therefore, we first calculate the PDF that arises if $f_{\rm nl}$ does indeed vanish, for the NHMV estimator $\widehat{f_{\rm nl}}$, and we also calculate the PDF that arises if the true value of $f_{\rm nl}$ is nonzero. We provide an analytic fit for these PDFs in Eq. (21). If the evidence from such a measurement were to show that $f_{\rm nl}$ is nonzero, then the next step would be to apply the CSZ estimator $\widehat{f_{\rm nl}^n}$ for $f_{\rm nl} \neq 0$ [14] to obtain a more precise value for $f_{\rm nl}$ or to test consistency of the data with a specific nonzero value of $f_{\rm nl}$. We therefore follow by calculating the PDF for these improved non-null-hypothesis estimators.

We find that, besides having a variance that increases with $f_{\rm nl}^2$, the PDF of the NHMV can have a significantly non-Gaussian shape when $f_{\rm nl} \neq 0$ with a long non-Gaussian tail for $\widehat{f_{\rm nl}} > |f_{\rm nl}|$ and less probability at $\widehat{f_{\rm nl}} < |f_{\rm nl}|$ than in the Gaussian case. As an example, taking $f_{\rm nl} = 100$ for an experiment which measures multipoles out to $l_{\rm max} = 3000$ (such as the Planck satellite [12]) and assuming a Gaussian PDF for the NHMV this experiment measures $74 \leq \widehat{f_{\rm nl}} \leq 148$ at the 99.7% C.L.; the actual PDF shows that this experiment measures $68 \leq \widehat{f_{\rm nl}} \leq 143$ 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 $\widehat{f_{\rm nl}} = 100 \pm 12.5$ at 99.7% C.L.

This paper is organized as follows. In Sec. II we construct the standard minimum-variance estimator $\widehat{f_{\rm nl}}$ under the null hypothesis $f_{\rm nl} = 0$ 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 $P(\widehat{f_{\rm nl}})$ for this estimator if the null hypothesis is indeed valid, i.e., if $f_{\rm nl}$ is indeed zero. We find that the PDF in this $f_{\rm nl} = 0$ case is well approximated by a Gaussian, for $N_{\rm pix} \gg 1$, 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 $f_{\rm nl} \neq 0$. We find the PDFs in this case can be highly non-Gaussian, skewed to large $|\widehat{f_{\rm nl}}|$, with long large-$\widehat{f_{\rm nl}}$ non-Gaussian tails and

[1]We note that the CSZ estimator, which is defined under the Sachs-Wolfe approximation, has yet to be generalized so that it can be applied to actual data. On the other hand a Bayesian approach, discussed in Ref. [15], allows for an $f_{\rm nl}$ inference that saturates the Cramer-Rao bound even in the presence of non-Gaussianity.

less likelihood at $\widehat{f_{\rm nl}} \leq |f_{\rm nl}|$ relative to the Gaussian distribution of the same variance. We provide fitting formulas for the PDF as a function of the estimator $\widehat{f_{\rm nl}}$, the true value of $f_{\rm nl}$, and the maximum multipole moment $l_{\rm max}$ 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 $f_{\rm nl}$ 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. Appendix A discusses the computational techniques we used in order to perform our Monte Carlo simulations.

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 approximation. We denote the fractional temperature perturbation at position $\vec{\theta}$ on a flat sky by $T(\vec{\theta})$ and refer to it hereafter simply as the temperature.

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

$$\langle T_{\vec{l}_1} T_{\vec{l}_2} \rangle = \Omega \delta_{\vec{l}_1+\vec{l}_2,0} C_l, \tag{2}$$

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

$$T_{\vec{l}} = \int d^2\vec{\theta} e^{-i\vec{l}\cdot\vec{\theta}} T(\vec{\theta}) \simeq \frac{\Omega}{N_{\rm pix}} \sum_{\vec{\theta}} e^{-i\vec{l}\cdot\vec{\theta}} T(\vec{\theta}) \tag{3}$$

is the Fourier transform of $T(\vec{\theta})$, and $\delta_{\vec{l}_1+\vec{l}_2,0}$ is a Kronecker delta that sets $\vec{l}_1 = -\vec{l}_2$. The power spectrum for $T(\vec{\theta})$ is given by

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

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

$$\langle T_{\vec{l}_1} T_{\vec{l}_2} T_{\vec{l}_3} \rangle = \Omega \delta_{\vec{l}_1+\vec{l}_2+\vec{l}_3,0} B(l_1, l_2, l_3). \tag{5}$$

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, and 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 $f_{\rm nl}$ 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 $f_{\rm nl} = 0$.

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

$$(\widehat{f_{\rm nl}})_{123} = \frac{T_{\vec{l}_1} T_{\vec{l}_2} T_{\vec{l}_3}}{\Omega B(l_1, l_2, l_3)/f_{\rm nl}}, \tag{6}$$

and under the null hypothesis this has a variance proportional to

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

The null-hypothesis minimum-variance estimator is constructed by adding all of these estimators with inverse-variance weighting. It is [16,17]

$$\widehat{f_{\rm nl}} \equiv \sigma^2_{f_{\rm nl}} \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{T_{\vec{l}_1} T_{\vec{l}_2} T_{\vec{l}_3} B(l_1, l_2, l_3)/f_{\rm nl}}{6\Omega^2 C_{l_1} C_{l_2} C_{l_3}}, \tag{8}$$

and it has inverse variance,

$$\sigma^{-2}_{f_{\rm nl}} = \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{[B(l_1, l_2, l_3)/f_{\rm nl}]^2}{6\Omega C_{l_1} C_{l_2} C_{l_3}}. \tag{9}$$

C. Non-Gaussianity of the PDF

If the number of pixels in the CMB map is $N_{\rm pix}$, then there are also $N_{\rm pix}$ statistically independent $T_{\vec{l}}$. But there are a much larger number, $\propto N^2_{\rm pix}$, of triplets $T_{\vec{l}_1} T_{\vec{l}_2} T_{\vec{l}_3}$, 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 similar to $N^2_{\rm pix} \gg N_{\rm pix}$. 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 $N_{\rm pix} \rightarrow \infty$ limit.

This contrasts with the estimator $\hat{C}_l \propto \sum |T_{\vec{l}}|^2$ of the power spectrum C_l. While the PDF for $\hat{C}_l$ is not necessarily Gaussian (it has a χ^2_{2l+1} 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 $f_{\rm nl}$ 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 $\widehat{f_{\rm nl}}$ FOR THE LOCAL MODEL

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

$$T(\vec{\theta}) = t(\vec{\theta}) + 3f_{\rm nl}\{[t(\vec{\theta})]^2 - \langle [t(\vec{\theta})]^2 \rangle\}, \tag{10}$$

where $t(\theta)$ is a Gaussian random field with a power spectrum C_l given in Eq. (4). To zeroth order in $f_{\rm nl}$, the power spectrum and correlation function for $T(\vec{\theta})$ are the same as those for $t(\vec{\theta})$. Note that $T(\vec{\theta})$ is, strictly speaking, the temperature fluctuation, so $\langle T(\vec{\theta})\rangle = 0 = T_{\vec{l}=0}$. The bispectrum for this model is

$$B(l_1, l_2, l_3) = 6 f_{\rm nl}(C_{l_1}C_{l_2} + C_{l_1}C_{l_3} + C_{l_2}C_{l_3}). \quad (11)$$

The temperature Fourier coefficients can be written $T_{\vec{l}} = t_{\vec{l}} + f_{\rm nl}\delta t^2_{\vec{l}}$ with

$$\delta t^2_{\vec{l}} \equiv \frac{3}{\Omega}\sum_{\vec{l}'} t_{\vec{l}-\vec{l}'}\, t_{\vec{l}'}. \quad (12)$$

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

We now proceed to evaluate $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$, the PDF that arises if the true value is $f_{\rm nl}$ for the NHMV estimator $\widehat{f_{\rm nl}}$ and for a map with $l_{\rm max}$. To do so, we generated large numbers of Monte Carlo realizations of maps according to Eq. (12), for some assumed value of $f_{\rm nl}$, and then applied the estimator in Eq. (8) to these maps. Each map is simulated in harmonic space from $l_{\rm min} = 2$ up to a maximum multipole $l_{\rm max}$. In order to produce a large number of realizations we reexpressed 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 $f_{\rm nl} = 0$

First we consider the shape of $P(\widehat{f_{\rm nl}}; f_{\rm nl} = 0, l_{\rm max})$, the PDF for the NHMV estimator in Eq. (8) applied to a purely Gaussian ($f_{\rm nl} = 0$) map. To do this we generated 10^6 Gaussian realizations and applied the estimator in Eq. (8) to generate a histogram of values of $\widehat{f_{\rm nl}}$. From this histogram we determined $P(\widehat{f_{\rm nl}}; f_{\rm nl} = 0, l_{\rm max})$ out to 4 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],

$$\sigma^2_{f_{\rm nl}} \approx \frac{1}{72 A l^2_{\rm max}\ln(l_{\rm max})}. \quad (13)$$

Additionally our simulations show that out to at least 4 times the root variance, the PDF $P(\widehat{f_{\rm nl}}; f_{\rm nl} = 0, l_{\rm max})$ is well approximated by a Gaussian for $l_{\rm max} \gtrsim 25$, even though the conditions for the central-limit theorem to apply are not satisfied. Therefore, a measurement of $\widehat{f_{\rm nl}}$ that differed from 0 at more than 3 times the root variance would indeed constitute a "99.7% confidence-level" inconsistency with the $f_{\rm nl} = 0$ hypothesis.

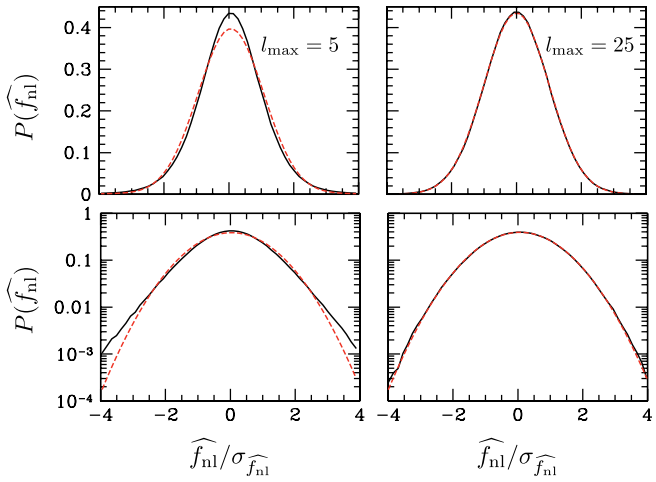

FIG. 1 (color online). Numerical evaluations of $P(\widehat{f_{\rm nl}}; f_{\rm nl} = 0, l_{\rm max})$. The left (right) two panels show the PDF for $l_{\rm max} = 5$ and $l_{\rm max} = 25$ for 10^6 realizations for a scale-invariant power spectrum. In all panels the PDF has been normalized to have a unit variance, and the corresponding Gaussian PDF (with the same variance) is shown as the red dashed curve. As $l_{\rm max}$ gets larger, the PDF tends toward a Gaussian. This is not guaranteed by the central-limit theorem since the majority of the terms that appear in the estimator are *not* statistically independent.

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

We now consider the form of $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ when $f_{\rm nl} \neq 0$, 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 $T_{\vec{l}}$'s impart some non-Gaussianity to the $\widehat{f_{\rm nl}}$ PDF.

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

Non-Gaussianity of $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ for a central value $f_{\rm nl} \neq 0$ may be significant for the interpretation of data. Suppose, for example, that a CMB measurement returns $\widehat{f_{\rm nl}} = 0$ with a root variance $\sigma_{f_{\rm nl}} = 40$. If the PDF was assumed to be Gaussian the measurement $\widehat{f_{\rm nl}} = 0$ would rule out $f_{\rm nl} = 100$ at the 2.5σ level, but given the asymmetric PDF of Fig. 2 it may rule out $f_{\rm nl} = 100$ 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. (8) to linear order in $f_{\rm nl}$ [14]:

$$\widehat{f_{\rm nl}} \approx E_0 + f_{\rm nl}E_1 + \cdots, \quad (14)$$

where

$$E_0 = \sigma^2_{f_{\rm nl}} \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{t_{\vec{l}_1} t_{\vec{l}_2} t_{\vec{l}_3}}{6\Omega^2 f_{\rm nl} C_{l_1} C_{l_2} C_{l_3}} B(l_1, l_2, l_3), \quad (15)$$

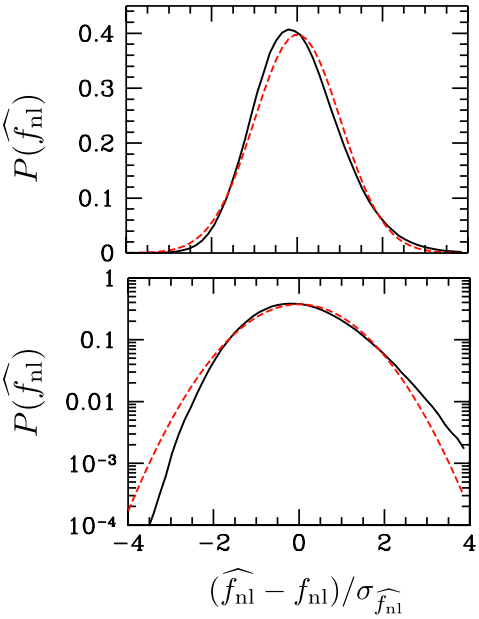

FIG. 2 (color online). The PDF $P(\widehat{f_{\rm nl}})$ when $\widehat{f_{\rm nl}} = 1500$ using the estimator in Eq. (8) with $l_{\rm max} = 25$. The upper (lower) panel shows the PDF on a linear (log) scale. We can see that the PDF is significantly non-Gaussian with an exponential drop-off to the left of mean and a power-law to the right. We provide a fitting formula for $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ in the text.

$$E_1 = \sigma^2_{f_{\rm nl}} \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{\delta t^2_{\vec{l}_1}\, t_{\vec{l}_2}\, t_{\vec{l}_3}}{2\Omega^3 C_{l_1} C_{l_2} C_{l_3}} B(l_1, l_2, l_3). \qquad (16)$$

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 $\langle E_1 \rangle = 1$. Furthermore, since we have already established that $P(\widehat{f_{\rm nl}})$ approaches a Gaussian in the large $l_{\rm max}$ limit if $f_{\rm nl} = 0$, we know that, to leading order, the non-Gaussian shape of $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ for $f_{\rm nl} \neq 0$ 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 $\widehat{f_{\rm nl}}$ is dominated by E_1 in the high S/N limit leading to a slower scaling of the S/N than the $l^{-2}_{\rm max}\ln^{-1}(l_{\rm max})$ 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 found the same qualitative trend but with a different dependence on $l_{\rm max}$. Reference [14] found $\langle(\Delta E_1)^2\rangle \propto \ln^{-2}(l_{\rm max})$, whereas our simulations show $\langle(\Delta E_1)^2\rangle \propto \ln^{-3}(l_{\rm max})$. We have checked the scaling found with our simulations by computing the variance analytically, as we further discuss in Appendix B. Figure 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. Figure 4 shows this PDF for various

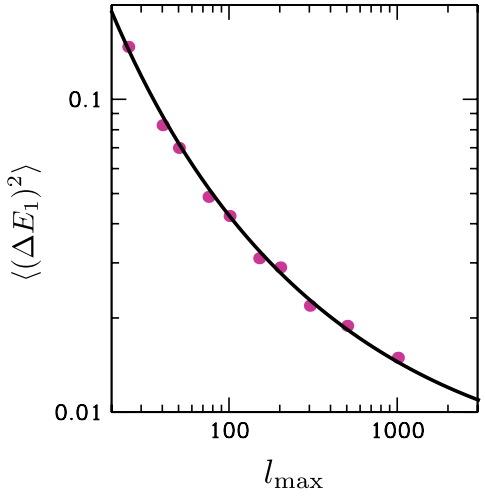

FIG. 3 (color online). The dependence of $\langle(\Delta E_1)^2\rangle$ on $l_{\rm max}$. The points correspond to the results of our Monte Carlo simulations for 1000 realizations at different values of $l_{\rm max}$. The solid curve shows the analytic calculation of the variance presented in Appendix B which is well fit by the function $\langle(\Delta E_1)^2\rangle = [14.0(l_{\rm max})^{0.433}]/[\ln^{5.1}(l_{\rm max})] \approx 4.5\ln^{-3}(l_{\rm max})$.

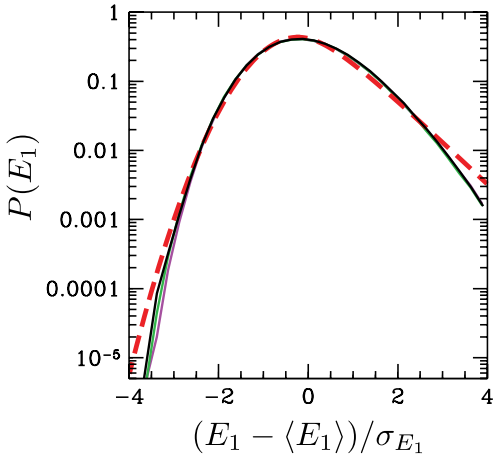

FIG. 4 (color online). The PDF of E_1 calculated using 10^6 realizations. The thin solid curves correspond to $P(E_1)$ with $l_{\rm max} = 25$ (green, middle curve), $l_{\rm max} = 50$ (purple, right curve), and $l_{\rm max} = 100$ (black, left curve). Since the functional form of the PDF for each choice of $l_{\rm max}$ is nearly identical, we conclude that $P(E_1)$ approaches an asymptotic functional form in the large-$l_{\rm max}$ limit. The thick red dashed curve corresponds to a fit to $P(E_1)$, accurate to $\sim$10% out to 3 times the root variance, using the fitting formula in Eq. (17) with parameter values $x_p = -0.22$, $\sigma = 0.80$, and $c = 0.91$.

choices of $l_{\rm max}$ (thin solid lines). An important conclusion from Fig. 4 is that the shape of the PDF approaches a universal form in the $l_{\rm max} \gg 1$ limit. We provide a fit to the PDF (thick red dashed line), accurate to $\sim$10% (40%) out to 3 (4) times the root variance, using the fitting formula

$$\log[F(x)] = N - \begin{cases} -(x-x_p)^2/(2\sigma^2), & x \le x_p, \\ -\frac{c}{\sigma^2}(\sqrt{(x-x_p)^2+c^2} - c), & x > x_p, \end{cases} \tag{17}$$

where $N \equiv \sqrt{2/\pi}\sigma + c\exp[c^2/\sigma^2]K_1(c^2/\sigma^2)$ 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 dashed 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 semianalytic expression for $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$, accurate to $\sim$10% (40%) out to 3 (4) times the root variance, as a function of $f_{\rm nl}$ and $l_{\rm max}$. Letting σ_0 and σ_1 denote the standard deviations of the distributions for E_0 and E_1, respectively, we have

$$\sigma_0^2 \approx \frac{1}{72Al_{\rm max}^2\ln(l_{\rm max})}, \tag{18}$$

$$\sigma_1^2 \approx \frac{f_{\rm nl}^2}{2\ln^3(l_{\rm max})}. \tag{19}$$

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

$$P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max}) \approx \frac{4}{9\sqrt{2\pi}\sigma_0\sigma_1}\int_{-\infty}^{\infty} G_0(\widehat{f_{\rm nl}} - x) \times F([x - f_{\rm nl}]/\sigma_1)dx, \tag{20}$$

where $G_0(x)$ is a Gaussian with zero mean and standard deviation σ_0 and $F([x - f_{\rm nl}]/\sigma_1)$ 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. (20) to write

$$P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max}) \approx \frac{2}{9}\exp\left[-\frac{X^2}{2(\sigma_0^2+\sigma_1^2\sigma^2)}\right]\sqrt{\frac{1}{\sigma_1^2(\sigma_0^2+\sigma_1^2\sigma^2)}}\left\{\sigma_1\sigma\left(1+\mathrm{erf}\left[\frac{\sigma_0^2+\sigma_1\sigma^2(X+\sigma_1)}{\sigma_0\sigma\sqrt{2(\sigma_0^2+\sigma_1^2\sigma^2)}}\right]\right)\right.$$
$$\left. + \sqrt{\sigma_0^2+\sigma_1^2\sigma^2}\left(1-\mathrm{erf}\left[\frac{c\sigma_0^2+\sigma_1\sigma^2(X+\sigma_1)}{\sqrt{2}\sigma_0\sigma_1\sigma^2}\right]\right)\exp\left[\frac{1}{2}\left(\frac{c^2\sigma_0^2}{\sigma_1^2\sigma^4}+\frac{2c[X+\sigma_1c^2]}{\sigma_1\sigma^2}+\frac{X^2}{\sigma_0^2+\sigma_1^2\sigma^2}\right)\right]\right\}, \tag{21}$$

where $X \equiv f_{\rm nl} + x_p\sigma_1 - \widehat{f_{\rm nl}}$.

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

$$\frac{\langle(\Delta\widehat{f_{\rm nl}})^3\rangle}{\sigma_{f_{\rm nl}}^3} = \left(\frac{f_{\rm nl}}{100}\right)^3\left(\frac{1}{1+3.7\exp[-(l_{\rm max}-5.1)/740]} - 0.26\right), \tag{22}$$

with the variance of the distribution, $\sigma_{f_{\rm nl}}^2$, given by

$$\sigma_{f_{\rm nl}}^2 \approx \frac{1}{72Al_{\rm max}^2\ln(l_{\rm max})}\left[1+\frac{36Af_{\rm nl}^2l_{\rm max}^2}{\ln^2(l_{\rm max})}\right]. \tag{23}$$

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

$$f_{\rm nl}A^{1/2} \gtrsim \frac{\ln(l_{\rm max})}{6l_{\rm max}}. \tag{24}$$

Therefore, for the Planck satellite (i.e., $l_{\rm max} = 3000$) the non-Gaussian features of $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ for the NHMV estimator are significant if $f_{\rm nl} \gtrsim \mathcal{O}(10)$. Thus, given that Planck is expected to measure $f_{\rm nl}$ with a variance $\sigma \approx 5$,

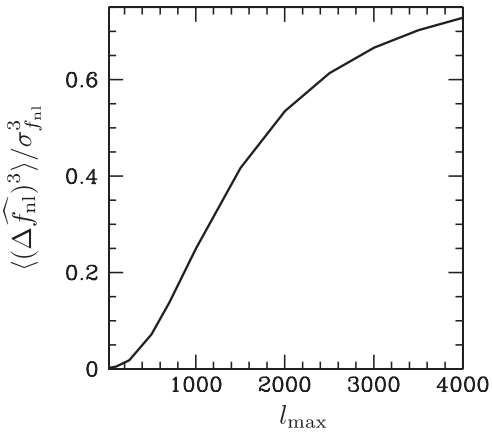

FIG. 5. The skewness, $\langle(\Delta\widehat{f_{\rm nl}})^3\rangle$, as a fraction of the variance of $P(\widehat{f_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ as a function of $l_{\rm max}$ for $f_{\rm nl} = 100$. We provide an analytic fitting formula in Eqs. (22) and (23) as a function of $f_{\rm nl}$ and $l_{\rm max}$.

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 $f_{\rm nl} \neq 0$

As we saw in the previous section the standard (null-hypothesis) minimum-variance estimator $\widehat{f_{\rm nl}}$ is constructed under the null hypothesis, so its variance is strictly minimized only when applied to maps with $f_{\rm nl} = 0$ [14]. In particular, the variance of $\widehat{f_{\rm nl}}$ is given in Eq. (23) so that when $36Af_{\rm nl}^2 l_{\rm max}^2/\ln^2(l_{\rm max}) \gtrsim 1$, the variance scales as the $\ln^{-3}(l_{\rm max})$, as opposed to $l_{\rm max}^{-2}\ln^{-1}(l_{\rm max})$. This indicates that when $f_{\rm nl} \neq 0$ 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 $f_{\rm nl} \neq 0$ which has a variance that continues to decrease as $1/[l_{\rm max}^2 \ln(l_{\rm max})]$ in the high signal-to-noise limit. To achieve this scaling they introduced a realization-dependent normalization,

$$\mathcal{N} \equiv \sigma^2_{f_{\rm nl}} \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{\chi_{\vec{l}_1} T_{\vec{l}_2} T_{\vec{l}_3}}{2C_{l_1}C_{l_2}C_{l_3}} B(l_1, l_2, l_3), \tag{25}$$

where

$$\chi_{\vec{l}} \equiv \sum_{\vec{k}} T_{\vec{l}-\vec{k}} T_{\vec{k}}. \tag{26}$$

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

$$\widehat{f^n_{\rm nl}} \equiv \frac{\widehat{f_{\rm nl}}}{\mathcal{N}}. \tag{27}$$

To explore the properties of the PDF of $\widehat{f^n_{\rm nl}}$, we expand the normalization as $\mathcal{N} \approx \mathcal{N}_0 + f_{\rm nl}\mathcal{N}_1 + \cdots$ and write

$$\widehat{f^n_{\rm nl}} \approx \frac{E_0}{\mathcal{N}_0} + f_{\rm nl}\frac{E_1\mathcal{N}_0 - E_0\mathcal{N}_1}{\mathcal{N}_0^2} + \cdots, \tag{28}$$

$$\equiv \mathcal{E}_0 + f_{\rm nl}\mathcal{E}_1 + \cdots. \tag{29}$$

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

$$\langle(\Delta\mathcal{E}_0)^2\rangle \approx \frac{1}{72Al_{\rm max}^2\ln(l_{\rm max})}, \tag{30}$$

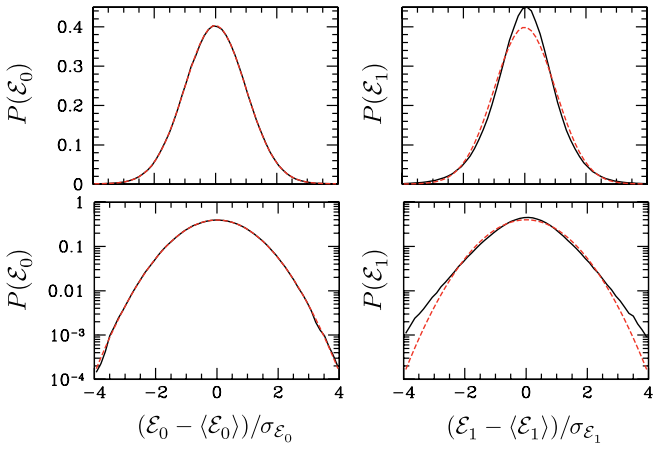

FIG. 6 (color online). The PDF $P(\mathcal{E}_0)$ (left panels) and $P(\mathcal{E}_1)$ (right panels) for $l_{\rm max} = 25$ determined with 10^6 non-Gaussian realizations. The top panels show the PDF on a linear scale; the bottom panels show the PDF on a log scale. We have confirmed that the shape of the PDF is unchanged for larger values of $l_{\rm max}$. The PDF of $\mathcal{E}_0$ (left panels) is well approximated by a Gaussian. However, the PDF of the first-order correction $\mathcal{E}_1$ (right panels) has significant non-Gaussian wings. This implies that the full PDF of $\widehat{f^n_{\rm nl}}$ is also non-Gaussian, even if the true value of $f_{\rm nl}$ matches that assumed in the construction of the CSZ estimator. Quantitatively, however, the level of non-Gaussianty will be small for Planck, as the variance of $\mathcal{E}_1$ is $\langle(\Delta\mathcal{E}_1)^2\rangle \approx 9\ln^2(l_{\rm max})/(l_{\rm max}^3)$.

$$\langle(\Delta\mathcal{E}_1)^2\rangle \approx \frac{\ln^2(l_{\rm max})}{l_{\rm max}^3}, \tag{31}$$

so that the PDF will be significantly non-Gaussian when

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

Therefore, for Planck (with $l_{\rm max} = 3000$) $P(\widehat{f^n_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ will be significantly non-Gaussian only if $f_{\rm nl} \gtrsim \mathcal{O}(1000)$. Since this has already been ruled out by observations [9,10], we conclude that $P(\widehat{f^n_{\rm nl}}; f_{\rm nl}, l_{\rm max})$ 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 $N_{\rm pix}$ 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 $\widehat{f_{\rm nl}}$ constructed under the null hypothesis is well approximated by a Gaussian distribution in the $l_{\rm max} \gg 1$ limit if the null hypothesis is correct (i.e., when applied to purely Gaussian maps).

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

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

The results presented here are made within the flat-sky, Sachs-Wolfe approximation. As such our conclusions should be taken as an order of magnitude estimate of $P(\widehat{f_{\rm nl}})$ calculated on the full sky and with the full transfer function (see Ref. [14] for a further discussion). However, we note that a comparison between the exact and approximate scaling of the signal-to-noise ratio with $l_{\rm max}$ shows the agreement to be better than an order of magnitude [16].

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 maximum likelihood, rather than quadratic, estimators (see, e.g., Ref. [14]). Ultimately, a variety of experimental effects and more precise power spectra and bispectra, rather than the Sachs-Wolfe-limit quantities used here, will need to be included in interpreting the results of realistic experiments.

There is also interest in using higher-order correlation functions to measure $f_{\rm nl}$ from CMB maps. Our arguments should apply also to these higher-order correlation functions, such as the trispectrum, etc. For example, the estimator for the amplitude of the n-point correlation function (e.g., $n = 3$ for the bispectrum, $n = 4$ for the trispectrum, etc.), will be constructed from $\sim N_{\rm pix}^{(n-1)}/n!$ combinations of n pixels, and this number of combinations scales even more rapidly with $N_{\rm pix}$ than that for the bispectrum. Thus, although the signal-to-noise ratio scales more rapidly with $N_{\rm pix}$ for these higher-order correlation functions than that for the bispectrum [17,22,23], concerns about the PDF for these estimators should be even more serious than for the bispectrum. It will thus be necessary to understand the PDF for these higher-order estimators to confidently forecast the statistical significance of measurements [24].

ACKNOWLEDGMENTS

We thank D. Babich, C. Hirata, and I. Wehus for useful discussions. T. L. S. is supported by the Berkeley Center of Cosmological Physics. M. K. thanks the Miller Institute for support and the Department of Physics at the University of California for hospitality, where part of this work was completed. M. K. was supported at Caltech by DoE DE-FG03-92-ER40701, NASA NNX10AD04G, and the Gordon and Betty Moore Foundation. B. D. W. was supported by NASA/JPL Subcontract No. 1413479, and NSF Grants No. AST 07-08849, No. AST 09-08693 ARRA, and No. AST 09-08902 during this work.

APPENDIX A: COMPUTING NON-GAUSSIANITY ESTIMATORS USING FAST FOURIER TRANSFORMS (FFT)

We are interested in using Monte Carlo simulations to determine the shape of the PDF of $\widehat{f_{\rm nl}}$ as a function of the fiducial choice of $f_{\rm nl}$ and the number $N_{\rm pix}$ of pixels measured in a given observation. Applying the estimator in Eq. (8) to the local-model bispectrum [Eq. (11)] it can be rewritten

$$\widehat{f_{\rm nl}} = \sigma^2_{f_{\rm nl}} \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{T_{\vec{l}_1} T_{\vec{l}_2} T_{\vec{l}_3}}{\Omega^2 C_{l_3}}. \tag{A1}$$

The estimator in Eq. (A1) takes $N_{\rm pix}^2$ operations to evaluate. Since current CMB observations have $N_{\rm pix} \sim 10^6$ 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 (~ 3–4σ).

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 $N_{\rm pix}^{5/2}$. In order to make the problem tractable Ref. [25] rewrites $\widehat{f_{\rm nl}}$ in terms of real-space quantities reducing the number of operations to $N_{\rm pix}^{3/2}$.

We can do the same for $\widehat{f_{\rm nl}}$ in the flat-sky approximation. Noting that

$$\delta_{\vec{l}_1+\vec{l}_2+\vec{l}_3,0} = \int \frac{d^2\theta}{\Omega} e^{i\vec{\theta}\cdot(\vec{l}_1+\vec{l}_2+\vec{l}_3)}, \tag{A2}$$

and writing

$$A(\vec{\theta}) \equiv \frac{1}{\Omega}\sum_{\vec{l}} e^{i\vec{l}\cdot\vec{\theta}} T_{\vec{l}}, \tag{A3}$$

$$B(\vec{\theta}) \equiv \frac{1}{\Omega}\sum_{\vec{l}} e^{i\vec{l}\cdot\vec{\theta}} \frac{T_{\vec{l}}}{C_l}, \tag{A4}$$

$\widehat{f_{\rm nl}}$ can be written

$$\widehat{f_{\rm nl}} = \Omega\sigma^2_{f_{\rm nl}} \int \frac{d^2\theta}{\Omega} A^2(\vec{\theta}) B(\vec{\theta}). \tag{A5}$$

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

$$\widehat{f_{\rm nl}} = \frac{\Omega\sigma^2_{f_{\rm nl}}}{N^2} \sum_{i=1}^{N}\sum_{j=1}^{N} A^2\left(2\pi\frac{i-1}{N}, 2\pi\frac{j-1}{N}\right) \times B\left(2\pi\frac{i-1}{N}, 2\pi\frac{j-1}{N}\right), \tag{A6}$$

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

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 $N_{\rm pix}\ln(N_{\rm pix})$.

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

$$E_1 = \sigma^2_{f_{\rm nl}} \sum_{\vec{l}_1+\vec{l}_2+\vec{l}_3=0} \frac{\delta t^2_{\vec{l}_1} t_{\vec{l}_2} t_{\vec{l}_3}}{2C_{l_1}C_{l_2}C_{l_3}} B(l_1, l_2, l_3). \tag{B1}$$

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

$$\langle(\Delta E_1)^2\rangle = 9\sigma^4_{f_{\rm nl}}(A_1 + 8A_2 + A_3 + 4A_4), \tag{B2}$$

where

$$A_1 \equiv \sum_{\{\vec{l}\},\{\vec{k}\}} \frac{B(l)}{C_{l_1}} \frac{B(k)}{C_{k_1}} \delta_{\vec{l}_1+\vec{k}_1,0}, \tag{B3}$$

$$A_2 \equiv \sum_{\{\vec{l}\},\{\vec{k}\}} \frac{B(l)}{C_{l_1}} \frac{B(k)}{C_{k_1}} \delta_{\vec{l}_3+\vec{k}_3,0}, \tag{B4}$$

$$A_3 \equiv \sum_{\{\vec{l}\}} \frac{B(l)^2}{C^2_{l_1}C_{l_2}C_{l_3}} \sum_{|\vec{m}|=1}^{l_{\rm max}} C_{|\vec{l}_1-\vec{m}|} C_m, \tag{B5}$$

$$A_4 \equiv \sum_{\{\vec{l}\},\{\vec{k}\}} \frac{B(l)B(k)}{C_{l_1}C_{k_1}C_{k_3}} C_{|\vec{l}_1+\vec{k}_2|} \delta_{l_3+k_3,0}, \tag{B6}$$

where $\{\vec{l}\}$ indicates the sum is over $\vec{l}_1+\vec{l}_2+\vec{l}_3=0$ and $B(l) \equiv B(l_1, l_2, l_3)$. Computing these terms as a function of $l_{\rm max}$ we find that the variance is well fit by the function

$$\langle(\Delta E_1)^2\rangle = \frac{14.0 l^{0.433}_{\rm max}}{\ln^{5.1}(l_{\rm max})}. \tag{B7}$$

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. **117B**, 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); A. Gangui *et al.*, Astrophys. J. **430**, 447 (1994); A. Gangui, Phys. Rev. D **50**, 3684 (1994); J. M. Maldacena, J. High Energy Phys. 05 (2003) 013; V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nucl. Phys. **B667**, 119 (2003); D. Babich, P. Creminelli, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 08 (2004) 009; P. Creminelli *et al.*, J. Cosmol. Astropart. Phys. 05 (2006) 004; 03 (2007) 005.

[3] X. c. Luo, Astrophys. J. **427**, L71 (1994); L. Verde *et al.*, Mon. Not. R. Astron. Soc. **313**, 141 (2000); E. Komatsu and D. N. Spergel, Phys. Rev. D **63**, 063002 (2001).

[4] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, Phys. Rep. **402**, 103 (2004).

[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**, R535 (1997); P. J. E. Peebles, Astrophys. J. **510**, 523 (1999); **510**, 531 (1999).

[6] S. Mollerach, Phys. Rev. D **42**, 313 (1990); D. H. Lyth and D. Wands, Phys. Lett. B **524**, 5 (2002); T. Moroi and T. Takahashi, Phys. Lett. B **522**, 215 (2001); **539**, 303(E) (2002); D. H. Lyth, C. Ungarelli, and D. Wands, Phys. Rev. D **67**, 023503 (2003); K. Ichikawa *et al.*, Phys. Rev. D **78**, 023513 (2008); K. Enqvist, S. Nurmi, O. Taanila, and T. Takahashi, J. Cosmol. Astropart. Phys. 04 (2010) 009; K. Enqvist and T. Takahashi, J. Cosmol. Astropart. Phys. 12 (2009) 001; 09 (2008) 012; K. Enqvist and S. Nurmi, J. Cosmol. Astropart. Phys. 10 (2005) 013; A. L. Erickcek, M. Kamionkowski, and S. M. Carroll, Phys. Rev. D **78**, 123520 (2008); A. L. Erickcek, C. M. Hirata, and M. Kamionkowski, Phys. Rev. D **80**, 083507 (2009).

[7] L. M. Wang and M. Kamionkowski, Phys. Rev. D **61**, 063504 (2000).

[8] S. Hannestad, T. Haugboelle, P. R. Jarnhus, and M. S. Sloth, J. Cosmol. Astropart. Phys. 06 (2010) 001.

[9] E. Komatsu *et al.* (WMAP Collaboration), Astrophys. J. Suppl. Ser. **148**, 119 (2003); E. Komatsu *et al.* (WMAP Collaboration), Astrophys. J. Suppl. Ser. **180**, 330 (2009); **192**, 18 (2011).

[10] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, Phys. Rev. D **77**, 123514 (2008); A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, J. Cosmol. Astropart. Phys. 08 (2008) 031; S. Matarrese and L. Verde, Astrophys. J. **677**, L77 (2008); C. Carbone, L. Verde, and S. Matarrese, Astrophys. J. **684**, L1 (2008); J. Q. Xia, M. Viel, C. Baccigalupi, G. De Zotti, S. Matarrese, and L. Verde, Astrophys. J. **717**, L17 (2010); J. Q. Xia, A. Bonaldi, C. Baccigalupi, G. De Zotti, S. Matarrese, L. Verde, and M. Viel, J. Cosmol. Astropart. Phys. 08 (2010) 013; L. Verde and S. Matarrese, Astrophys. J. **706**, L91 (2009); F. Schmidt and M. Kamionkowski, Phys. Rev. D **82**, 103002 (2010).

[11] A. P. S. Yadav and B. D. Wandelt, Phys. Rev. Lett. **100**, 181301 (2008).

[12] P. A. R. Ade *et al.* (Planck Collaboration), arXiv:1101.2022 [Astron. Astrophys. (to be published)].

[13] F. Elsner and B. D. Wandelt, Astrophys. J. Suppl. Ser. **184**, 264 (2009).

[14] P. Creminelli, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 03 (2007) 019.

[15] F. Elsner and B. D. Wandelt, Astrophys. J. **724**, 1262 (2010).

[16] D. Babich and M. Zaldarriaga, Phys. Rev. D **70**, 083005 (2004).

[17] M. Kamionkowski, T. L. Smith, and A. Heavens, Phys. Rev. D **83**, 023007 (2011).

[18] G. Efstathiou, Mon. Not. R. Astron. Soc. **348**, 885 (2004); C. J. Copi, D. Huterer, D. J. Schwarz, and G. D. Starkman, arXiv:1103.3505.

[19] G. R. Dvali and S. H. H. Tye, Phys. Lett. B **450**, 72 (1999); P. Creminelli, J. Cosmol. Astropart. Phys. 10 (2003) 003; M. Alishahiha, E. Silverstein, and D. Tong, Phys. Rev. D **70**, 123505 (2004).

[20] D. Babich, P. Creminelli, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 08 (2004) 009; P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 05 (2006) 004; P. Creminelli, L. Senatore, M. Zaldarriaga, and M. Tegmark, J. Cosmol. Astropart. Phys. 03 (2007) 005.

[21] D. G. Figueroa, R. R. Caldwell, and M. Kamionkowski, Phys. Rev. D **81**, 123504 (2010).

[22] N. Kogo and E. Komatsu, Phys. Rev. D **73**, 083007 (2006).

[23] J. Smidt, A. Amblard, A. Cooray, A. Heavens, D. Munshi, and P. Serra, Phys. Rev. D **81**, 123007 (2010).

[24] T. L. Smith and M. Kamionkowski (unpublished).

[25] E. Komatsu, D. N. Spergel, and B. D. Wandelt, Astrophys. J. **634**, 14 (2005).