Primordial non-Gaussianity from the covariance of galaxy cluster counts

Carlos Cunha,[1] Dragan Huterer,[1] and Olivier Doré[2,3]

[1]*Department of Physics, University of Michigan, 450 Church Street, Ann Arbor, Michigan 48109-1040, USA*
[2]*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California 91109, USA*
[3]*California Institute of Technology, Pasadena, California 91125, USA*

It has recently been proposed that the large-scale bias of dark matter halos depends sensitively on primordial non-Gaussianity of the local form. In this paper we point out that the strong scale dependence of the non-Gaussian halo bias imprints a distinct signature on the covariance of cluster counts. We find that using the full covariance of cluster counts results in improvements on constraints on the non-Gaussian parameter f_{NL} of 3 (1) orders of magnitude relative to cluster counts (counts + clustering variance) constraints alone. We forecast f_{NL} constraints for the upcoming Dark Energy Survey in the presence of uncertainties in the mass-observable relation, halo bias, and photometric redshifts. We find that the Dark Energy Survey can yield constraints on non-Gaussianity of $\sigma(f_{NL}) \sim 1$–5 even for relatively conservative assumptions regarding systematics. Excess of correlations of cluster counts on scales of hundreds of megaparsecs would represent a smoking-gun signature of primordial non-Gaussianity of the local type.

I. INTRODUCTION

Primordial non-Gaussianity provides cosmology one of the precious few connections between primordial physics and the present-day universe. Standard inflationary theory with a single-field, slowly rolling scalar field, predicts that the spatial distribution of structures in the Universe today is very nearly Gaussian random (e.g. [1–5]; for an excellent recent review, see [6]). Departures from Gaussianity, barring contamination from systematic errors or late-time non-Gaussianity due to secondary processes, can therefore be interpreted as violation of this "vanilla" inflationary assumption. Constraining or detecting primordial non-Gaussianity is therefore an important and basic test of the cosmological model.

Constraints on primordial non-Gaussianity have been traditionally obtained from observations of the cosmic microwave background, as nonzero non-Gaussianity generates a nonzero three-point correlation function (or its Fourier transform, the bispectrum) of density fluctuations [7–13]. Increasingly sophisticated algorithms have been developed to constrain non-Gaussianity [14–18] and, to the extent that it can be measured, Gaussianity has so far been confirmed [19–21]. For example, the most recent constraints from the Wilkinson Microwave Anisotropy Probe (WMAP) indicate $f_{NL} \approx 32 \pm 21$ (1σ; [22]), where the exact constraints depend somewhat on the choice of the statistical estimator applied to the data, the CMB map used, and details of the foreground subtraction. Here, f_{NL} is the parameter describing non-Gaussianity in the widely studied "local" model, where the non-Gaussian (NG) potential Φ_{NG} is defined by

$$\Phi_{NG}(x) = \Phi_G(x) + f_{NL}(\Phi_G^2(x) - \langle \Phi_G^2 \rangle), \qquad (1)$$

and where Φ_G is the Gaussian potential. Corresponding

constraints can be obtained on other classes of non-Gaussian models. For example, for "equilateral" models where most power comes from equilateral triangle configurations, $f_{NL}^{eq} = 26 \pm 140$ (1σ; [22]).

The CMB is not the only cosmological probe to be sensitive to the presence of primordial non-Gaussianity. It has been known for a relatively long time that the abundance of dark matter halos [23–29] (or voids [30,31]) is sensitive to the presence of primordial non-Gaussianity. This dependence is easy to understand: halos populate the high tail of the probability density distribution of structures in the Universe, and the shape of this distribution is sensitive to departures from Gaussianity. However, while the halo abundance is rather powerful in constraining models that are non-Gaussian in the density (rather than the potential) [32], for the popular models of the local type [cf. Eq. (1)] the abundance is much less constraining than the CMB anisotropy and not competitive with the CMB constraints (e.g. [33,34]).

Some of us [35] have recently shown that the clustering of dark matter halos is very sensitive to primordial non-Gaussianity of the local type. This exciting development paves way to using the large-scale structure to probe primordial non-Gaussianity nearly 3 orders of magnitude more accurately than using the abundance of halos. Dalal *et al.* [35] found, analytically and numerically, that the bias of dark matter halos acquires strong scale dependence

$$b(k) = b_0 + f_{NL}(b_0 - 1)\delta_c \frac{3\Omega_m H_0^2}{a g(a) T(k) c^2 k^2}, \qquad (2)$$

where b_0 is the usual Gaussian bias (on large scales, where it is constant), $\delta_c \approx 1.686$ is the collapse threshold, a is the scale factor, Ω_M is the matter fraction relative to critical, H_0 is the Hubble constant, k is the wavenumber, $T(k)$ is the

transfer function, and $g(a)$ is the growth suppression factor.[1] This result has been confirmed by other researchers using a variety of methods, including the peak-background split [36–38], perturbation theory [39–41], and numerical (N-body) simulations [42–44]. Astrophysical measurements of the scale dependence of the large-scale bias, using galaxy and quasar clustering as well as the cross-correlation between the galaxy density and CMB anisotropy, have recently been used to impose constraints on f_{NL} already comparable to those from the cosmic microwave background (CMB) anisotropy [36,38], giving $f_{NL} = 28 \pm 23$ (1σ), with some dependence on the assumptions made in the analysis [38]. In the future, constraints on f_{NL} are expected to be of order a few [35,45,46]. The sensitivity of the large-scale bias to other models of primordial non-Gaussianity has not been investigated yet (though see preliminary analyses in [47,48]).

Clustering of galaxy clusters, in particular, can very strongly constrain primordial non-Gaussianity. Clusters have an advantage of being large, relatively simple objects that are easy to find using either optical or x-ray light, or else from their Sunyaev-Zeldovich signature. Clusters already provide interesting constraints on dark energy [49,50] and they hold promise for precision measurements of cosmological and dark energy parameters (e.g. [51]). Since clusters are massive and hence significantly biased objects, their counts (via the mass function) and clustering (via the mass function and bias) are both sensitive to primordial non-Gaussianity. Recently, Oguri [52] has argued that the variance of cluster counts (i.e. scatter measured in each cell individually), in combination with the cluster counts, leads to interesting improvements on f_{NL} constraints relative to the counts-only case.

In this paper we point out that including the *covariance* of cluster counts in angle and redshift leads to very significant further improvements in the cluster constraints on local primordial non-Gaussianity. The principal reason for the improvement is simple: covariance is determined by the cluster power spectrum, which is proportional to the halo bias squared. At large scales, the non-Gaussian contribution to the halo bias dominates [cf. Eq. (2)], and this results in a strong f_{NL} signal in the covariance. Furthermore, we explore the sensitivity of the constraints to various assumptions about statistical and systematic errors in modeling the cluster mass-observable relation, as well as the presence of other cosmological parameters. We find that the bulk of the information about local non-Gaussianity comes from the far-separation covariances of cluster counts-in-cells.

This paper is organized as follows. In Sec. II, we describe the methodology that we use to obtain constraints from both counts and clustering of galaxy clusters. In Sec. III, we describe our fiducial assumptions about the cosmological model and data as well as solutions to various challenges in calculating the constraints. In Sec. IV, we describe the forecasted constraints on f_{NL} from the Dark Energy Survey. We discuss our results in Sec. V and conclude in Sec. VI.

II. METHODOLOGY

We address the following problem: how well can the cosmological parameters be recovered using counts of galaxy clusters in pixels distributed in angle and radius on the sky? We largely follow the formalism of Hu and Cohn [53] and Lima and Hu [54].

Assume that clusters are counted in square pixels of fixed angular size θ_{pix}, corresponding to comoving size $L_{pix}(z) = \theta_{pix} r(z)$, where r is the comoving distance. The clusters are also binned in the mass-observable (i.e. the observable proxy for cluster mass), with intervals $[M_{obs}^{\alpha}, M_{obs}^{\alpha+1}]$ where α refers to a specific mass-observable bin. The number density of clusters at a given redshift z with observable in the range $M_{obs}^{\alpha} \leq M_{obs} \leq M_{obs}^{\alpha+1}$ is given by

$$\bar{n}_{\alpha}(z) \equiv \int_{M_{obs}^{\alpha}}^{M_{obs}^{\alpha+1}} \frac{dM_{obs}}{M_{obs}} \int \frac{dM}{M} \frac{d\bar{n}}{d\ln M} p(M_{obs}|M), \quad (3)$$

where $p(M_{obs}|M)$ is the observable-mass relation (explained in Appendix A) and $d\bar{n}/d\ln M$ is the mass function. Uncertainties in the redshifts distort the volume element; we fully take into account the photometric redshift uncertainties following [55]; details are shown in Appendix B.

We adopt the mass function from Dalal *et al.* [35] who used N-body simulations to parametrize the shift in mass of a typical halo in the presence of non-Gaussianity. The mass shift, $M_G \rightarrow M$, is adequately described by a Gaussian with mean and variance, respectively, given by

$$\left\langle \frac{M}{M_G} \right\rangle - 1 = 1.3 \times 10^{-4} f_{NL} \sigma_8 \sigma(M_G, z)^{-2}, \quad (4)$$

$$\mathrm{var}\left(\frac{M}{M_G} \right) = 1.4 \times 10^{-4} (f_{NL} \sigma_8)^{0.8} \sigma(M_G, z)^{-1}, \quad (5)$$

where $\sigma(M, z)$ is the amplitude of mass fluctuations on mass scale M and at redshift z. The final non-Gaussian mass function is given by [35]

$$\frac{dn}{dM} = \int dM_G \frac{dn}{dM_G} \frac{dP}{dM}(M_G), \quad (6)$$

where $dP/dM(M_G)$ is the probability distribution that a Gaussian halo of mass M_G maps to a non-Gaussian halo of mass M, and is given by the Gaussian with the mean and variance given in Eqs. (4) and (5). For dn/dM_G, we adopt the Jenkins mass function [56].

---

[1]The usual linear growth $D(a)$, normalized to be equal to a in the matter-dominated epoch, is related to the suppression factor $g(a)$ via $D(a) = ag(a)$.

On large scales, the number counts of clusters $m(\mathbf{x})$ trace the linear density perturbation $\delta(\mathbf{x})$

$$m_i(M_\alpha, \mathbf{x}) \equiv m_{i\alpha} = \bar{m}_i(1 + b(M_\alpha, z)\delta(\mathbf{x})), \quad (7)$$

where i refers to the pixel (i.e. its angular and radial coordinates), and α indicates the mass bin. The spatial covariance of cluster counts is [57]

$$S_{ij}^\alpha = \langle (m_{i\alpha} - \bar{m}_{i\alpha})(m_{j\alpha} - \bar{m}_{j\alpha}) \rangle \equiv \bar{m}_{i\alpha}\bar{m}_{j\alpha}\xi_{ij}^\alpha, \quad (8)$$

where ξ_{ij}^α is the pixel real-space correlation function

$$\xi_{ij}^\alpha \equiv \int \frac{d^3k}{(2\pi)^3} |W_i(\mathbf{k})W_j(\mathbf{k})| \cos(k_x \Delta x_{ij}) \times \cos(k_y \Delta y_{ij})$$
$$\times \cos(k_z \Delta z_{ij}) b_{i\alpha} b_{j\alpha} P(k; z). \quad (9)$$

If i and j come from different redshift bins, the geometric mean of the two redshifts[2] is adopted for z. In the limit $r_{ij} \gg L_{\text{pix}}$, $\xi_{ij}^\alpha \rightarrow \xi(r_{ij})$, where the latter quantity is the standard two-point correlation function in real space. $\Delta x_{ij} = L_{\text{pix}} n_{xij}$ is the physical separation between i and j in the x direction (transverse to the line of sight), and n_{xij} is the number of pixels separating them; Δy_{ij} is defined equivalently. Finally, the window function W is the Fourier transform of the square pixel in the presence of redshift errors

$$W(\mathbf{k})_i = \exp\left(\frac{-\sigma_{z,i}^2 k_z^2}{2H_i^2}\right)$$
$$\times j_0(k_x L_{\text{pix}}/2) j_0(k_y L_{\text{pix}}/2) j_0(k_z \Delta z/2H_i), \quad (10)$$

where the index i refers to the redshift bin, $\sigma_{z,i}$ is the redshift scatter at the radial distance corresponding to the ith pixel, and H_i is the Hubble parameter. The photo-z bias is implicit in the Δz_{ij} term in Eq. (9).

The expression for the full Fisher matrix for galaxy cluster counts and their covariance is quite complicated (see [53]), but a reasonable approximation is given by [58]

$$F_{\mu\nu} = \bar{\mathbf{m}}_{,\mu}^t \mathbf{C}^{-1} \bar{\mathbf{m}}_{,\nu} + \frac{1}{2}\text{Tr}[\mathbf{C}^{-1}\mathbf{S}_{,\mu}\mathbf{C}^{-1}\mathbf{S}_{,\nu}], \quad (11)$$

where the first term encodes information from cluster counts, and the second from the covariance. Here, μ and ν are indices that refer to both cosmological and nuisance parameters (including f_{NL}). The cluster counts have been arranged as the vector $\bar{\mathbf{m}}$. $\mathbf{S} = \{S_{ij}^\alpha\}$ is the sample covari-

---

[2]In the linear regime, the correlation between pixels i and j contains the product of the growth factors corresponding to z_i and z_j. Therefore, the corresponding power spectrum, $P(k, z)$ in Eq. (9), should use the growth function equal to the geometric mean of the two growth functions. Instead, we effectively use the growth function which is evaluated at the redshift equal to the geometric mean of the two *redshifts* z_i and z_j. Results are insensitive to this approximation, especially because most of the information comes from relatively close redshift pairs.

ance matrix from Eq. (8), and $\mathbf{C} \equiv \mathbf{N} + \mathbf{S}$ is the total covariance. $N_{ij} = \bar{m}_i \delta_{ij}$ is the (shot) noise matrix. The derivative with respect to f_{NL} can be computed analytically, using the fact that $P(k, z) \propto b^2(k, z)$ and Eq. (2).

III. FIDUCIAL ASSUMPTIONS AND CALCULATIONAL CHALLENGES

We implement the procedure outlined above for optically selected clusters. In our fiducial setup we divide the sky into the 11×11 field of pixels of 41.32 sq. deg each, for a total of 5000 sq. deg which matches expectations for the Dark Energy Survey (DES). The facing surface of each pixel is a square with a side $L_{\text{pix}}(z) = \theta_{\text{pix}} r(z)$ (see Sec. II). Each pixel has redshift depth $\Delta z = 0.2$, and we assume a maximum redshift of 1.0 so that there are five redshift bins. We adopt the mass threshold of $M^{\text{th}} = 10^{13.7} h^{-1} M_\odot$ and also bin in mass, using 5 mass bins of width $\Delta \ln M^{\text{th}} = 0.2$, with the exception of the highest-mass bin, which we extend to infinity. Using smaller bins in angle or redshift yields better results, up to the point where the covariance matrix becomes dominated by shot-noise (which occurs for bins with area around 0.1–1 sq. deg). For very large number of pixels, the Gaussian approximation used to define the covariance used in our Fisher matrix would break down. In our fiducial case we have about 1.7×10^5 clusters subdivided into 3025 pixels, so that we are well within the Gaussian regime. In addition, results for large angular pixels are less sensitive to systematics due to nonlinear physics or angular mask uncertainties. In Sec. IV we consider departures from the fiducial assumption, namely, variations in the mass threshold, maximum redshift range, and pixel area.

We assume fiducial cosmological parameters based on the fifth year data release of the Wilkinson Microwave Anisotropy Probe [59]. Thus, we set the baryon density, $\Omega_b h^2 = 0.0227$, the dark matter density, $\Omega_m h^2 = 0.1326$, the normalization of the power spectrum at $k = 0.05 \text{ Mpc}^{-1}$, $\delta_\zeta = 4.625 \times 10^{-5}$, the tilt, $n = 0.963$, the optical depth to reionization, $\tau = 0.087$, the dark energy density, $\Omega_{\text{DE}} = 0.742$, and the dark energy equation of state, $w = -1$. In this cosmology, $\sigma_8 = 0.796$. We use CMBFAST [60], version 4.5.1, to calculate the transfer functions, and add Planck priors [81] when calculating the marginalized constraints on parameters.

To study systematic errors in cluster cosmology, we add a generous set of nuisance parameters described in Appendix A (see also Cunha [61] and Cunha *et al.* [51]), with 10 nuisance parameters describing the bias and variance in the mass-observable relation and 3 parameters describing uncertainty in the halo bias [a_c, p_c, and δ_c, cf. Eq. (B5)]. The assumption of 3 nuisance parameters describing the Gaussian halo bias is somewhat *ad hoc* but conservative since for a given mass function the halo bias can be predicted to roughly 10% accuracy in the range of scales we are interested [62]. We fix the photo-z scatter to

0.02 everywhere except in Sec. IV D where we consider the effects of including 10 additional nuisance parameters describing photometric redshift errors. In this exploratory paper we do not consider models for non-Gaussianity other than the one from Eq. (1), or observational systematic errors (e.g. atmospheric blurring or completeness variations across the sky). The study of these effects is left for future work.

Evaluating the expression for the Fisher matrix with the signal matrix of this size is clearly challenging: the total size of the matrix S [see Eq. (9)] is $N \times N$, where $N = N_{\text{pixels}} \times N_{\text{mass}} \times N_{\text{redshift}} = 121 \times 5 \times 5 = 3,025$ in our fiducial case. The bottleneck is in calculating the $\sim 10^7$ elements of the matrix, each of which involves the numerical computation of a rapidly oscillating triple integral; see Eqs. (8) and (9). Unlike previous works which studied constraints on dark energy [53,55,58], we cannot ignore the off-diagonal elements (i.e. the pixel *covariance*) of the matrix S since those elements, while being very small for the Gaussian case, become significant for $f_{\text{NL}} \neq 0$ (see the right panel of Fig. 1) due to the $f_{\text{NL}}^2 k^{-4}$ dependence scaling of the power spectrum as $k \to 0$. To reduce the size of the covariance matrix, we assume that the information from the different mass bins is independent, so that we can estimate the Fisher matrix for each mass bin separately and then add them in the end. The scatter in the mass-observable relation can generate correlations between mass bins at a given pixel. In addition, as Seljak [63] and McDonald and Seljak [64] noticed, correlating the halos of different masses at large separations would lead to improved constraints in our analysis, making our assumption conservative.

A. Regularization of the covariance

As Wands and Slosar [65] pointed out, the two-point correlation function for biased tracers of structure has an infrared divergence if f_{NL} is not zero. However, the *measured* correlation function from any survey is of course finite, because one cannot measure variance of the density field on scales larger than the survey. To that effect, Wands and Slosar [65] suggest regularizing the correlation function $\xi(r)$ by subtracting from it the variance of the density field evaluated at the scale of the survey. However, Cunha and Slosar (private communication) found out that the regularization of Wands and Slosar [65] contains a typo; the correct regularization term is given by

$$\Sigma^2(R) \equiv \int \frac{d^3k}{(2\pi)^3} |W_R(\mathbf{k})| b_{i\alpha} b_{j\alpha} P(k; z), \quad (12)$$

where we use the mass bin α and redshift bin i corresponding to those of the correlation function ξ_{ij}^α from which this is being subtracted. If i and j come from different redshift bins, the geometric mean of the two redshifts is taken. The difference from what is presented in Wands and Slosar [65] is that our expression has $|W_R(\mathbf{k})|$ instead of $|W_R(\mathbf{k})|^2$ [cf. Eqs. (47, 49, 50) in [65]]. Using the above expression, the observed 2-point correlation at a given survey volume has the desirable property that it integrates to zero over the survey volume.

We approximate the window function $|W_R(\mathbf{k})|$ as the Fourier transform of a spherical top hat, and adopt $R = 2h^{-1}$ Gpc as the linear dimension of our survey. For the main analysis in this paper, the effects of the divergence are not significant, since all of our results (except in Sec. V) assume zero fiducial f_{NL}, and the analytic expression for
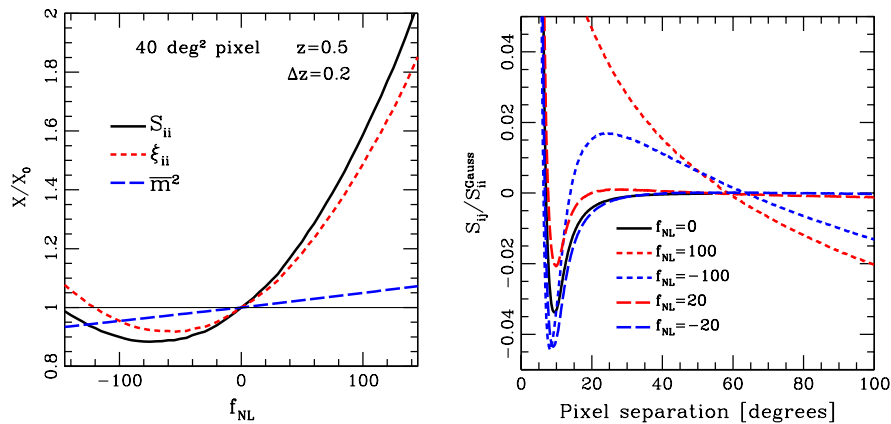

FIG. 1 (color online).   *Left panel*: Sensitivity of the variance of cluster counts to non-Gaussianity. The black lines shows the variance S_{ii}, the short dashed red line shows the (auto)correlation function ξ_{ii}^α, and the long-dashed blue line shows the squared mean counts. Note that $S_{ii} = (\bar{m})^2 \xi_{ii}^\alpha$. We assumed a pixel with area 40 sq. deg and radial redshift extent $\Delta z = 0.2$, centered at $z = 0.5$. *Right panel*: Sensitivity of the covariance of cluster counts to non-Gaussianity. We show the off-diagonal elements of the clustering matrix, normalized by the variance of $f_{\text{NL}} = 0$ case (S_{ii}^{Gauss}) as a function of angle between the ith and jth pixel. We use the same pixelization as for the left panel. We show the Gaussian case ($f_{\text{NL}} = 0$), and four non-Gaussian models ($f_{\text{NL}} = \pm 20$ and $f_{\text{NL}} = \pm 100$). Note that, because of the regularization, the results depend on the size of the survey. The larger the survey, the larger the effect of non-Gaussianity.

the derivative dS_{ij}/df_{NL} is weakly sensitive to the integration boundary. The divergence of the two-point correlation does affect the covariance for nonzero f_{NL} and for pixel separation greater than a few hundred Mpc. We use the lower boundary of integration $k_{min} = 10^{-4}$, and check that results are stable *vis-a-vis* variations in this value, or whether the regularization mentioned above has been applied or not. For Fig. 1 and the results in Sec. V, we do apply the corrected Wands-Slosar regularization prescription [cf. Eq. (12)].

Besides its impact on the regularization, the choice of survey geometry is important since the *distribution* of pixel-pixel separations depends on the geometry. We assume that the survey itself has square shape (and implicitly work in a flat-sky approximation), and assume a 11×11 field of square-shaped pixels for each redshift bin. To populate the covariance matrix, we precompute S_{ij} as a function of pixel separation for integer values of the separation along a row of pixels in Eq. (9)—that is, we set $\Delta y_{ij} = 0$ and vary Δx_{ij} at each redshift. We use linear interpolation to estimate the covariance for pixels whose physical separation, in units of $\Delta x_{i(i+1)}$, is noninteger. We find that the effects of disregarding the pixel orientation are negligible (by changing the orientation of bins and finding little change in the results). Precomputation of the covariance matrix elements as a function of pixel separation greatly reduces the number of covariance terms we need to calculate.

As the right panel of Fig. 1 shows, in the Gaussian case the off-diagonal terms of S_{ij} fall off very fast. We find that covariance terms for pixels in different redshifts to be negligible, because we use broad redshift bins. Hence, we only calculate covariance between different redshift bins when estimating the derivative of the covariance with respect to f_{NL}. To save time, for the results shown in Sec. IV, we only calculate terms in adjacent redshift bins. We checked that including larger redshift separations improves unmarginalized constraints by about 30%. But including the regularization removes most of the improvement (for fiducial $f_{NL} = 0$). To calculate the derivatives of the covariance with respect to f_{NL}, we use the fact that the derivative of the bias with respect to f_{NL} is analytic so that

$$\frac{d\xi_{ij}^{\alpha}}{df_{NL}} \equiv \int \frac{d^3k}{(2\pi)^3} |W(\mathbf{k})|^2 \cos(k_x \Delta x_{ij}) \cos(k_y \Delta y_{ij})$$
$$\times \cos(k_z \Delta z_{ij}) \frac{d(b_{i\alpha} b_{j\alpha})}{df_{NL}} P(k; z). \quad (13)$$

In calculating dS_{ij}/df_{NL}, we only keep the dominant term, which is the one with derivative with respect to ξ_{ij}^{α}. That is, we assume that

$$\frac{dS_{ij}^{\alpha}}{df_{NL}} \simeq \bar{m}_{i\alpha} \bar{m}_{j\alpha} \frac{d\xi_{ij}^{\alpha}}{df_{NL}}. \quad (14)$$

The terms we ignore correspond to the sensitivity of cluster counts to non-Gaussianity, and they would only *enhance*

the impact of f_{NL}, though slightly, as will be shown in the following sections. In a real survey one actually has to calculate the covariance at nonzero values of f_{NL} for which our approach of evaluating the derivative analytically at $f_{NL} = 0$ would be insufficiently general. For this sensitivity study, however, the analytic derivative is perfectly acceptable. We examine the sensitivity to the constraints around different fiducial values of f_{NL} in Sec. V.

IV. RESULTS

Our results are presented as follows. First, we discuss the sensitivity of cluster counts and clustering of counts to f_{NL}, and examine unmarginalized constraints on f_{NL}. Second, we examine the degeneracies with cosmological parameters and nuisance parameters due to modeling uncertainties in the observable-mass relation and in the halo bias. Third and last, we look at the impact of photometric redshift uncertainties.

A. Sensitivity of cluster covariance

The effect of non-Gaussianity on clustering is a combination of several effects, which can be identified from Eq. (8). The dominant effect is due to the explicit modification of the halo bias [Eq. (2)] which affects ξ_{ij}^{α} [cf. Eq. (9)] In addition, non-Gaussianity affects the mass function, which affects the mean cluster counts [cf. Eqs. (3) and (B1)], and the average cluster linear bias [cf. Eq. (B7)]. The left panel of Fig. 1 shows the dependence of the different terms that make up the clustering covariance S_{ij}, as a function of f_{NL}. For this sensitivity plot, we assume a 40 sq. deg pixel with redshift thickness $\Delta z = 0.2$ centered around $z = 0.5$ and a mass threshold $M^{th} = 10^{13.7} h^{-1} M_{\odot}$, and show only the diagonal elements $i = j$ for clarity. The relation between the functions plotted in this figure is $S_{ij} = \bar{m}^2 \xi_{ij}^{\alpha}$. It is apparent from the figure that ξ_{ij}^{α} encodes most of the dependence of the clustering signal on f_{NL}, and that the clustering covariance (S_{ij}, or ξ_{ij}^{α}) is much more sensitive to f_{NL} than the mean counts $\bar{m}$. As mentioned previously, we neglect the implicit mass function dependence of f_{NL} when calculating the covariance. Including it would only enhance the impact of f_{NL}, albeit slightly.

In the right panel of Fig. 1 we plot the absolute value of the clustering covariance as a function of angular separation between the centroids of two pixels. For reference, at $z = 0.5$, a 1 deg separation corresponds to about $23.4 h^{-1}$ Mpc. For $f_{NL} = 0$, the clustering covariance is large and positive at small separations, but becomes negative at intermediate pixel separations (~ 6 deg or $\sim 150 h^{-1}$ Mpc at $z = 0.5$); this behavior corresponds to a similar behavior of the two-point correlation function $\xi(r)$ (see e.g. Ref. [66]). The effect of nonzero f_{NL} depends on its sign as well as on the scale. For positive f_{NL}, the covariance increases monotonically with f_{NL} roughly up to

the scale of the survey. Beyond that scale ($\sim 60°$ in our example), the covariance reverses its trend with f_{NL} and becomes negative due to the integral constraint imposed by the regularization. For negative f_{NL}, the dependence of the covariance S_{ij} on f_{NL} is more complicated because the total bias becomes negative at large enough scales; thus, for $f_{NL} < 0$ the covariance depends monotonically on $|f_{NL}|$ only on scales ($\lesssim 7°$ in the right panel of Fig. 1) for which the bias correction—second term in Eq. (2)—is subdominant. Note that Fig. 1 hides the fact that the number of pixels at a given separation increases with separation: the number of off-diagonal elements in the covariance is much bigger than the number of diagonal elements, and this gives a "geometric boost" to the covariance.

B. Unmarginalized constraints from clustering and counts

Both panels of Fig. 2 show f_{NL} constraints as a function of the maximum pixel separation allowed in the covariance [cf. Eq. (8)] used to generate the Fisher matrix constraints [cf. Eq. (11)].

In the left panel of Fig. 2 we see that the cluster counts yield better unmarginalized constraints than the *variance* of cluster counts alone; however, once the covariances (i.e. off-diagonal terms of the signal matrix S_{ij}) are included, the clustering information rapidly beats that from the counts. In Table I we show the *unmarginalized* f_{NL} constraints for a variety of survey expectations. Changes in the constraints improve in the direction expected: the lower the mass threshold and the higher the maximum redshift, the better. This Table also shows that decreasing the angular area of the pixels to 12.5 deg^2 results in substantial

TABLE I. Unmarginalized constraints on f_{NL}. The fiducial case assumes no nuisance parameters, 5 bins in mass and redshift each, and other assumptions as in the text. Variations in the assumptions are shown in the first column, followed by the total number of clusters in the 5000 deg^2 survey we assumed, while cluster counts, covariance, and combined projected 1-σ constraints on f_{NL} are given in the following three columns.

| Assumption | Unmarginalized error $\sigma(f_{NL})$ | | | |
	Number	Counts	Covariance	Both
Fiducial	1.7×10^5	9.1	1.8	**1.7**
12.5 deg^2 pix	1.7×10^5	9.2	1.1	**1.1**
$z_{max} = 0.8$	1.3×10^5	13	2.3	**2.2**
$z_{max} = 1.4$	2.4×10^5	6.0	1.4	**1.4**
$M^{th} = 10^{13.5}$	3.6×10^5	8.3	1.4	**1.4**
$M^{th} = 10^{13.9}$	7.7×10^4	10	2.3	**2.3**

[$O(50\%)$] improvements. The improvement with decreasing pixel size, for f_{NL} constraints, does not happen if we consider only the variance in counts. For other parameters, that are sensitive to small scale information, such as Ω_{DE} and w, the smaller pixels do translate into better constraints even if only the sample variance is used. Further refinements of the pixelization leads to improvement up to the regime of shot-noise domination (which occurs for pixels of ~ 0.1–1 deg^2). Unmarginalized constraints are of order 10^{-1} in this regime, though observational systematics are likely to dominate statistical errors of this size.

C. Degeneracies with cosmological and nuisance parameters

In the right panel of Fig. 2 we show the *marginalized* constraints on f_{NL} assuming Planck priors and fixed nui-
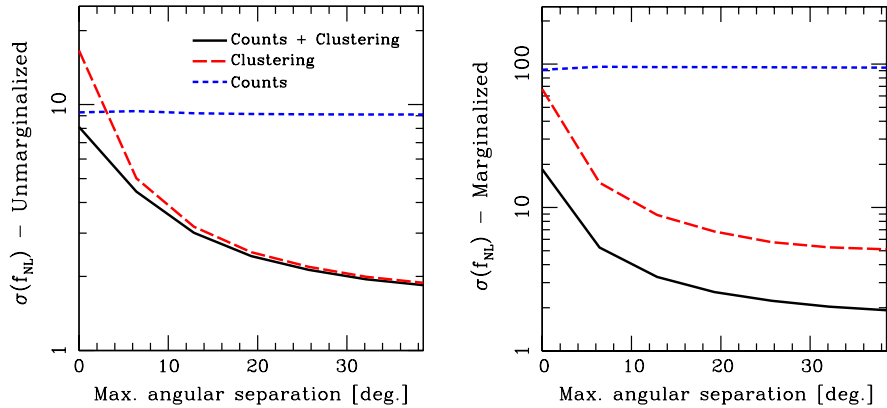

FIG. 2 (color online). 1-σ uncertainties in the parameter f_{NL} as a function of the maximum angular separation between pixel centroids in the covariance matrix. The *left* panel shows the unmarginalized constraints, while the *right* panel shows marginalized constraints assuming Planck priors and fixed halo bias and observable-mass nuisance parameters. Zero separation indicates the case of pure variances (as considered by Oguri [52]). The maximum angular separation between pixels for a 5000 sq. deg survey divided into 41.3 sq. deg pixels is about 90 deg (or $10\sqrt{2}$ pixel widths). This case would correspond to taking the full covariance into account for the calculation of f_{NL}, but disregarding the covariance between different redshift bins. The *blue short dashed* line corresponds to constraints derived using only cluster counts. The *red dashed* line shows the constraints when only the clustering of clusters is used, and the *solid black* line shows the combined constraints from counts and clustering.

TABLE II. Marginalized constraints on f_{NL} and dark energy with cluster counts, variance of the counts, and the two combined. The fiducial case assumes 5 bins in mass and redshift each with a mass threshold $M^{th} = 10^{13.7}$, maximum redshift $z_{max} = 1.0$, and other assumptions as in the text. Assumptions about the nuisance parameters are varied, and are shown in the first two columns. Entries with ∞ indicate that the method was unable to constrain the parameters.

| | | Marginalized errors—Variance only | | | | | | | | |
| Nuisance parameters | | Counts | | | Variance | | | Counts + Variance | | |
Halo bias	M_{obs}	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$
Marginalized	Marginalized	∞	∞	∞	∞	∞	∞	0.075	0.25	**55**
Known	Marginalized	0.095	0.32	$\mathbf{3.4 \times 10^3}$	∞	∞	∞	0.061	0.21	**27**
Marginalized	Known	∞	∞	∞	0.077	0.26	**98**	0.0037	0.016	**44**
Known	Known	0.0046	0.021	**91**	0.053	0.18	**67**	0.0035	0.014	**19**

TABLE III. Marginalized constraints on f_{NL} and dark energy with cluster counts, covariance of the counts, and the two combined. The fiducial case assumes 5 bins in mass and redshift each with a mass threshold $M^{th} = 10^{13.7}$, maximum redshift $z_{max} = 1.0$, and other assumptions as in the text. Assumptions about the nuisance parameters are varied, and are shown in the first two columns. Entries with ∞ indicate that the method was unable to constrain the parameters.

| | | Marginalized errors—Full Covariance | | | | | | | | |
| Nuisance parameters | | Counts | | | Covariance | | | Counts + Covariance | | |
Halo bias	M_{obs}	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$
Marginalized	Marginalized	∞	∞	∞	∞	∞	∞	0.069	0.23	**6.0**
Known	Marginalized	0.097	0.33	$\mathbf{2.1 \times 10^3}$	0.13	0.43	**12**	0.065	0.22	**5.4**
Marginalized	Known	∞	∞	∞	0.099	0.34	**7.0**	0.0036	0.014	**3.8**
Known	Known	0.0051	0.023	**94**	0.042	0.13	**5.1**	0.0036	0.014	**1.8**

sance parameters (both halo bias and mass-observable). We see that the change in the constraints from combined counts[3] and clustering is even more remarkable than the unmarginalized constraints shown in the right panel. The full clustering covariance yields about 1 order of magnitude better constraints than if only the variance is used. As we shall see, this fractional improvement remains even when we include nuisance parameters.

Tables II and III show f_{NL} constraints using the variance of cluster counts, and the full covariance, respectively. The results assumed Planck priors on the cosmological parameters, 10 nuisance parameters describing the mass-observable relation, and 3 nuisance parameters describing uncertainties in the Gaussian halo bias.

Comparing the last columns of Tables II and III, we see that the counts + covariance combination yields about an order of magnitude improvement over simply using counts + variance. For the counts + variance, the uncertainties in the halo bias parameters are the main source of degradation to f_{NL} constraints. Without the information from large separations provided by the full covariance, the Fisher matrix cannot disentangle the effects due to the Gaussian bias from the f_{NL} contribution. When the full

covariance is used (cf. Table III), the errors in the mass-observable relation are the dominant source of degradation. Marginalizing over all nuisance parameters, assuming flat priors, yields a degradation of ~ 3 in $\sigma(f_{NL})$. This is not large, considering we added 13 nuisance parameters, but not negligible either. Even modest prior information can improve the marginalized constraints significantly.

There are two principal reasons for the strong improvement of errors when the covariance is added:
(1) The strong scale dependence of the bias as a function of non-Gaussianity implies that most signal comes from the covariances, since the covariances have longer lever arms in k than the variance alone (and are much more sensitive than counts which only depend on non-Gaussianity via the mass function);
(2) The signature of f_{NL} in the covariance is unique, as no other cosmological parameter leads to a similar effect—therefore, the degeneracy with other cosmological parameters is very small, as first noted by [35].

Comparing the f_{NL} constraints for the full covariance for fixed nuisance parameters (Table III) to the unmarginalized constraints (Table I), we see that degeneracies with cosmological parameters only result in a small degradation of f_{NL} constraints (from 1.7 to 1.8).

Tables II and III also show the constraints obtained using counts alone, or (co)variance by itself. The information about f_{NL} from the *counts* is very degenerate with the cosmological and nuisance parameters. The "∞" symbols

---

[3]The slight degradation in f_{NL} constraints from counts seen in the right panel is real, and is due to adding the (positive) covariance matrix elements to the counts noise; see the first term on the right-hand side of Eq. (11). Using the full covariance therefore yields very slightly worse constraints.

indicate that the Fisher matrix could not be inverted, i.e., that particular technique was unable to simultaneously constrain all of the parameters. From the last row of both tables, we see that cluster counts are effective at constraining the cosmological parameters and mass-observable relation (from the mass binning); whereas, the (co)variance constrains mainly the nuisance parameters and f_{NL}.

Marginalization degrades the counts + covariance f_{NL} constraints roughly independently of the different survey assumptions, so one can use Table I to infer marginalized constraints. For example, from Table I, we see that using 12.5 deg^2 pixels yields about 60% better constraints. The full marginalized constraints are also improved by a similar factors so that, for example, $\sigma(f_{NL}) \sim 3.9$ for 12.5 deg^2 marginalized over the 13 nuisance parameters [compared to $\sigma(f_{NL}) = 6.0$ for 40 deg^2 pixels].

D. Photometric redshift errors

To study the effects of photometric redshift errors, we add 10 nuisance parameters to the analysis, namely, two parameters—one each describing the photo-z scatter and bias—in each of the five redshift bins. The results are summarized in Table IV.

If either the halo bias or the mass-observable nuisance parameters are fixed, then the degradation from the inclusion of photo-z's is not very damaging. In other words, the additional correlations between either photo-z and halo bias parameters, or between photo-z and mass-observable parameters, do not cause substantial additional degradation to f_{NL} constraints (relative to the case where only the photo-z parameters are unknown).

However, when all 23 nuisance parameters (10 for the photo-z's, 10 for the mass-observable relation, and 3 for halo bias) are left free, one cannot simultaneously constrain dark energy and f_{NL}, and the constraints on both drastically degrade. We traced the biggest source of degradation to the redshift evolution parameters in the mass-observable relation and to the photo-z bias nuisance parameters. Simply adding a 33% prior to the one parameter describing the evolution of the bias in $P(M_{obs}|M)$ [parameter a_1 in Eq. (A3)] was enough to reclaim respectable accuracy, with $\sigma(f_{NL}) = 18.8$ (see the bottom row of Table IV). Alternatively, if the bias in each photo-z bin is known to the absolute accuracy of 0.01 with all other parameters free, then $\sigma(f_{NL}) = 7.0$, which is just $\sim$15% worse than when photo-z parameters are fixed.[4] For a survey such as the DES, these requirements should be relatively easy to satisfy, given that spectroscopic samples of 10^4–10^5 galaxies will be available to calibrate the photo-

TABLE IV. Effect of photometric redshift uncertainties on the marginalized constraints on f_{NL}. The fiducial case assumes 5 bins in mass and redshift each with a mass threshold $M^{th} = 10^{13.7}$ and maximum redshift $z_{max} = 1.0$, and other assumptions as in the text. Variations are in the first two columns, while cluster, covariance, and combined projected 1-σ constraints on f_{NL} are given in the following three columns.

The effects of photo-z uncertainties				
Nuisance parameters				
Halo bias	M_{obs}	$\sigma(\Omega_{DE})$	$\sigma(w)$	$\sigma(f_{NL})$
Known	Known	0.016	0.041	**6.5**
Marginalized	Known	0.021	0.053	**6.7**
Known	Marginalized	0.11	0.36	**9.4**
Marginalized[a]	Marginalized[a]	0.23[a]	0.77[a]	**19**[a]

[a]In the bottom row, superscript ^{a} signals that a Fisher matrix prior of $F_{a_1,a_1} = 10$ is added to the nuisance parameter a_1 defined in Eq. (A3), which describes the redshift evolution of the bias in the mass-observable relation.

metric redshift errors (see e.g. Eqs. (19, 20) in Hearin *et al.* [67]).

V. DISCUSSION

A. Choice of the fiducial model

In our fiducial approach we estimated errors in f_{NL} around $f_{NL} = 0$. However, it is a slightly different matter to estimate the *detectability* of non-Gaussianity, which requires estimating the signal-to-noise at which a *nonzero* fiducial value of f_{NL} can be differentiated from zero.[5] The detectability is independent of the fiducial value if the observable quantity is linear in the parameter(s); this is clearly not the case here since the clustering signal is a quadratic function of the bias, which itself depends linearly on f_{NL}.

Figure 3 shows the fiducial unmarginalized constraints on f_{NL} as a function of its fiducial value. Unlike in the results shown previously, here we calculate all elements of the covariance matrix and its derivative with respect to f_{NL} (which is why the constraints for $f_{NL} = 0$ shown in the plot are slightly better than what is shown in Table I). The figure shows tightest constraints for $|f_{NL}| \simeq 10$—more than 4 times stronger than those for our fiducial assumption of $f_{NL} = 0$. The "witch's hat" shape shown in Fig. 3 can be understood by examining the second term on the right-hand side of Eq. (11) that contains the Fisher information from the covariance of cluster counts. The f_{NL} constraints are set by the competition between the signal, represented by the derivative of the covariance with respect to f_{NL}, $\mathbf{S}_{,\mu}$, and the noise, given by the total covariance, $\mathbf{C}$. These two

---

[4]Unlike f_{NL}, the dark energy constraints are sensitive to both bias and scatter of the photo-z's. For a prior uncertainty in the photo-z bias of 0.01 per bin, the photo-z scatter needs to be known to 0.025 *per bin* to achieve small ($\lesssim$ 15%) degradation in $\sigma(\Omega_{DE})$ and $\sigma(w)$ relative to the case of perfectly known photo-z errors.

[5]Arguably the best approach might be to use the Bayesian model selection techniques and, for a range of f_{NL} values, test if the hypothesis $f_{NL} = 0$ can be rejected. We do not pursue such an approach in this paper.
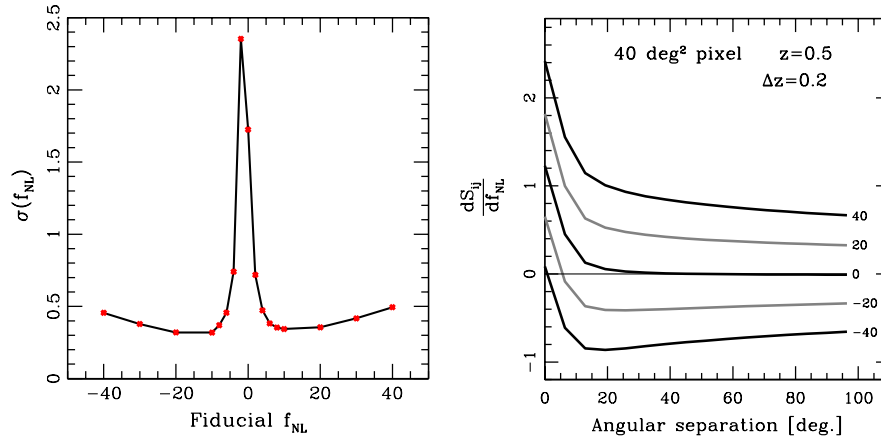
FIG. 3 (color online).　*Left panel*: Unmarginalized 1-σ constraints on f_{NL} as a function of the fiducial value of this parameter, assuming five redshift and five mass bins. The witch's hat shape can be explained from the competition between the derivative of the covariance with respect to f_{NL}, and the total covariance at the fiducial f_{NL}; see text. *Right panel*: Derivative of the signal matrix elements S_{ij} with respect to f_{NL} as a function of angular separation between pixels i and j, for $f_{NL} = -40, -20, 0, 20$, and 40. Recall that, at $z = 0.5$, a separation of 1 deg corresponds to about $23h^{-1}$ Mpc.

quantities vary with f_{NL} at different rates; the total covariance depends (roughly) quadratically on f_{NL}; whereas, $\mathbf{S}_{,\mu}$ only has a linear dependence. In addition, the matrix elements of $\mathbf{S}_{,\mu}$ and $\mathbf{C}$ have different sensitivity to f_{NL} at each angular separation, and it is the relative importance of the off-diagonal matrix elements relative to the diagonal elements that sets the shape of the curve in Fig. 3.

For very small values of $|f_{NL}|$ ($\ll 10$), the off-diagonal elements of the covariance are very small, and hence do not contribute much to the signal, or to $\mathbf{C}$. This can be seen in the $f_{NL} = 0$ curves in the right panel of Fig. 1 and in the right panel of Fig. 3. Note that the plots hide the fact that the number of pixels at a given separation increases with separation: the number of off-diagonal elements in the covariance is much bigger than the number of diagonal elements, and this gives a geometric boost to the covariance.

For large values of $|f_{NL}|$ ($\gg 10$), the off-diagonal elements of the covariance matrix can be significant relative to the diagonal elements (see the $f_{NL} = \pm 100$ curves in the right panel of Fig. 1). Therefore, the constraints on f_{NL} now worsen with the increasing value of $|f_{NL}|$, albeit slowly.

Finally, in the intermediate range of $|f_{NL}| \sim 10$, the off-diagonal elements of $\mathbf{C}$ are small relative to the diagonal and near-diagonal elements. For example, the right panel of Fig. 1 shows that, for $f_{NL} = 20$, the far-separation covariances are much smaller than the variances. However, the derivatives of the sample covariance, $d\mathbf{S}/df_{NL}$, are only moderately smaller for the off-diagonal pixels than for the diagonal ones (e.g. a factor of ~ 4 for $f_{NL} = 20$; see the right panel of Fig. 3). Therefore, it is at these intermediate values of $|f_{NL}| \sim 10$ that we find the best signal-to-noise, and best constraints on f_{NL}.

In summary, the dependence on the fiducial value of f_{NL} can be understood rather simply. For small f_{NL}, the large-scale covariances do not add much signal. For large f_{NL}, the covariances add too much noise. At intermediate f_{NL}, the signal-to-noise relation is "just right." We caution that the shape of the curve in Fig. 3 depends on the volume (and geometry) of the survey as well as in the number density of sources. The width of the pixels affect the width of the central part of the "hat" slightly. Smaller bins tend to shift the minima to smaller values of $|f_{NL}|$. We conclude that the power of a DES-like cluster surveys to rule out the Gaussian hypothesis may be even greater than indicated in Tables in this paper, since the error at $f_{NL} \neq 0$ nearly always smaller than that for $f_{NL} = 0$. This is another exciting development, but warrants further investigation, and, in particular, a more detailed study of the dependencies on the overall survey volume and selection. In this initial study we simply adopt the conservative errors, and show the $f_{NL} = 0$ results everywhere except in Fig. 3.

B. Clusters vs galaxies

It is useful to compare cluster constraints obtained here with the expected constraints from a similar, DES-type, galaxy survey. Forecasts of constraints on primordial non-Gaussianity from galaxy clustering were studied recently [35,45,68] using the Fisher matrix and a simple, Feldman-Kaiser-Peacock (FKP [69]) estimator that counts modes of $P(k)$ and combines them with the survey volume and its galaxy density. Perhaps counterintuitively, our constraints are a factor of a few better than those from galaxies estimated previously. We now explain the origin of this apparent discrepancy.

Both clusters and galaxies probe the power spectrum of dark matter halos (and thus the halo bias). However, there are some important differences

(i) Clusters additionally probe the mass function, which determines the counts, and also weakly affects the bias $b_0(M, z)$; see Eqs. (B6) and (B7);

(ii) The number density of galaxies may be significantly higher, depending on how they and the clusters are selected. However, as mentioned in Sec. III, the larger size of galaxy samples may not bring much additional information, since the constraints on f_{NL} benefit from very large-scale halo separations, and not from intrahalo correlations;

(iii) Clusters reside in more massive halos than galaxies, and thus have a higher bias. The higher the bias, the stronger is the correlation [cf. Eq. (9)];

(iv) With regards to systematics, clusters can naturally be binned by the mass-observable, which helps break degeneracies with nuisance parameters. This allows utilization of the cross-correlation between different mass bins to reduce the impact of sample variance (e.g. [63,64]), which we do not exploit in this paper.

(v) Large spectroscopic samples of galaxies are expected in the near future; whereas, clusters will rely on photometric redshifts; therefore, galaxy redshifts are likely to be more accurate than cluster redshifts.

Given all these differences, it is difficult to predict whether clusters or galaxies will give a stronger constraint on primordial non-Gaussianity without a direct calculation. We have verified that the FKP estimator of galaxy constraints on f_{NL} indeed gives a weaker result, and is in rough agreement with previous estimates in [35,45,46].

However, as discussed in Tegmark *et al.* [70], the FKP estimator is only optimal and lossless on scales much smaller than the linear size of the survey. Since good constraints on f_{NL} benefit from precisely the large-wavelength modes, it is not surprising that the FKP estimator for galaxies indicates worse constraints than our pixel-based estimator for clusters. We have additionally verified that constraints on the constant part of the bias, b_0 [see Eq. (2)], or the dark energy equation of state w, which do not benefit as much from large-wavelength modes, are comparable when estimated from the pixel-based formalism (from this paper) and the FKP approach assuming the same survey volume and number density of objects.

C. Comparison to previous work

Numerous papers have studied the power of cluster counts alone to probe primordial non-Gaussianity (e.g. [26,32–35]). To the extent that such constraints are generally weak due to degeneracies, and strongly depend on the priors and nuisance parameters varied, our results (see the ''counts'' columns in Table III) are in broad agreement with these studies.

A more interesting comparison can be made with the recent work of Oguri [52] who studied the counts + variance case of clusters, corresponding to results in our Table II. The main difference between the two studies is that we additionally considered the covariance of cluster counts, and found that it leads to a huge further improvement in the constraints. However, even for the counts + variance case only, our results differ substantially, and we forecast a much *weaker* constraint on non-Gaussianity than Ref. [52]. For example, we get $\sigma(f_{NL}) \sim 20$–30 compared to $\sigma(f_{NL}) \sim 8$ in Ref. [52] in the most fair comparison with their DES survey case and our assumptions with either no nuisance parameters or full mass-observable nuisance parameters.[6] These discrepancies could probably be explained by a number of other differences in the analyses: mass functions [Ref. [52] uses the LoVerde *et al.* [71] mass function with analytic fit for skewness, while we use Dalal *et al.* mass function from Eqs. (4)–(6)]; cosmological parameter priors (Ref. [52] uses the diagonal priors on some parameters while we use the full, off-diagonal Planck prior Fisher matrix), etc. We have not attempted to reproduce results from Ref. [52] using the assumptions made in that paper.

D. Issues for future study

There are a number of effects that remain to be studied in detail, but are beyond the scope of this preliminary analysis. We now list them here:

(i) *Fisher matrix approximation:* in this paper we have assumed the fiducial value of $f_{NL} = 0$ and calculated the errors on f_{NL} by taking the derivatives of observables with respect to this parameter. This ''Fisher error'' will be a good approximation to the true error if the error itself is small. Therefore, at least in the cases where the f_{NL} error is tight, we expect the Fisher approximation is a good one, though this should eventually be checked with Markov chain Monte Carlo methods.

(ii) *Calculational issues:* The computation of the cluster covariance is time consuming, particularly for small but nonzero values of f_{NL}. In this work we have largely avoided this issue by using the Fisher matrix approximation and taking analytic derivatives around $f_{NL} = 0$ (and a few other values), which enabled us to only evaluate the covariance at the fiducial Gaussian model. With real data, however, a full exploration of parameter space will be necessary, which might be sufficiently time consuming to warrant analysis using a smaller set of observable parameters. For example, one could resort to using larger pixels and a coarser binning in redshift, or perhaps using no pixels at all. One could also explore

---

[6]Reference [52] assumes only two mass-observable nuisance parameters.

speeding up the covariance calculations with various mathematical tricks.

(iii) *Mass function:* we have assumed the Dalal *et al.* [35] mass function which has been calibrated from numerical simulations and simply shifts the mass of halos with non-Gaussianity. A number of alternative mass functions have recently been proposed in the literature and studied numerically [42,71]. While the agreement in the relevant quantity $n_{\rm NG}(M, z)/n_{\rm G}(M, z)$ is becoming good, there is still no uniform agreement in the community about the convergence. The overall constraints are expected to be robust given that most of the effect of non-Gaussianity comes from the bias scaling as $f_{\rm NL}k^{-2}$ and not the mass function. Nevertheless, we expect constraints in this paper to be on the conservative side: given that the Dalal *et al.* mass function predicts a smaller effect due to non-Gaussianity than some of the other popular functions, use of these other mass functions would only increase the effects due to non-Gaussianity and thus improve the error bars on $f_{\rm NL}$.

(iv) *Corrections to the bias formula:* While the dependence of bias on $f_{\rm NL}$ is established to follow Eq. (2) both analytically and numerically, it could be that there are second-order corrections to the bias formula. These have been discussed in the literature; for example, it appears that a small constant offset in bias is warranted by the simulations and some analytical results [41,43,44]. Study of these higher-order corrections is very important but, given that there is no convergence in the community on this issue as of yet, we leave their inclusion for future work.

(v) *Relativistic corrections and gauge dependence:* Wands and Slosar [65] have shown that, to first-order, the scale-dependent bias does not receive relativistic corrections at large scales, using a spherical collapse model. However, other authors have shown that higher-order corrections in the matter perturbations can produce non-Gaussianity (see e.g. [72,73]). How the higher-order correction propagate to the halo bias is yet to be understood in detail.

(vi) *Observational systematics:* In this paper we have modeled the systematic uncertainties in understanding of the Gaussian bias $b_0(M, z)$ and the relation between cluster mass and its observational proxy by introducing nuisance parameters that describe uncertainty in these relations. However, we have not attempted to model observational uncertainties, such as variations in atmospheric seeing or photometric calibration. Clearly, knowledge of such uncertainties over large angular scales will be important if measurements of non-Gaussianity are not to be substantially degraded. We leave the study of observational systematics for future work.

VI. CONCLUSIONS

In this paper we studied how well primordial non-Gaussianity of the local type can be probed with galaxy clusters. We took into account cluster number counts, as well as the full covariance of cluster counts-in-cells. We allowed generous uncertainties in the knowledge of the cluster mass-observable relation, the photometric redshifts, and the Gaussian halo bias (we did not consider systematics due to uncertainties in angular selection, which may be important.) As we discuss at length in Sec. III, the Fisher matrix calculation is computationally challenging, and we resorted to a number of conservative approximations, the most important of which is using very large pixels. Since angular selection issues are expected to be most significant at small angular scales, our pixel choices partly justify neglecting angular uncertainties.

We found that most information on primordial non-Gaussianity comes from the previously neglected covariance of counts. The covariance links cluster overdensities across large distances, and thus benefits the constraints on primordial non-Gaussianity of the local type. The reason is easy to understand: the non-Gaussian parameter $f_{\rm NL}$ enters through the term proportional to k^{-2} in the bias, and correlates cluster counts in bins separated by hundreds of megaparsecs. Other cosmological parameters do not lead to these far-separation correlations in cluster counts (see the right panel of Fig. 1). Correlations of cluster counts across vast spatial distances of hundreds of megaparsecs therefore represent a smoking-gun signature of primordial non-Gaussianity of the local type.

The combination of counts and clustering is particularly effective at breaking degeneracies of $f_{\rm NL}$ with cosmological and nuisance parameters, since the two statistical probes complement each other very well. While our full set of 23 freely varying nuisance parameters can degrade $f_{\rm NL}$ constraints by factors of a few, even modest prior uncertainties on some of them break degeneracies and restore the accuracy in $f_{\rm NL}$. For example, the bias in each photo-z bin needs to be known to 0.01 to keep $f_{\rm NL}$ constraints within 15% of their values for the case of perfectly known photo-z's.

We investigated the sensitivity of our results to the choice of fiducial value of $f_{\rm NL}$ and found that the uncertainty in $f_{\rm NL}$ at $f_{\rm NL} \neq 0$ is smaller than that for $f_{\rm NL} = 0$. In other words, a nonzero small value of $f_{\rm NL}$ may even be more sensitively differentiated from the $f_{\rm NL} = 0$ case than indicated in our Tables; the reason for this is explained in Sec. V A.

Our forecasts indicate very strong constraints on primordial non-Gaussianity, which is perhaps surprising. However, closer inspection reveals a number of effects that help clusters achieve these numbers; we discuss these in Sec. V B. In particular, we use the pixel-based estimator, which is well suited for extracting signal from very large scales. Previous error forecasts of non-Gaussianity from

galaxy clustering used the suboptimal FKP estimator; dark energy studies that did use the pixel-based estimator only considered variance of cluster counts.

To achieve the full potential of forecasted constraints discussed here, a few more issues need to be carefully studied. Particularly important are theoretical uncertainties in linking dark matter halos to observed clusters of galaxies, and observational systematics across large angular scales. While constraints on primordial non-Gaussianity have improved 2 orders of magnitude between COBE [74] and WMAP [22], another 1 or even 2 orders of magnitude improvement may be possible with upcoming surveys of large-scale structure, especially if they include both dark matter halo counts and their clustering covariance.

ACKNOWLEDGMENTS

APPENDIX A: PARAMETRIZATION OF MASS-OBSERVABLE RELATION

We assume a log-normal form for the probability of measuring an observable signal, denoted $M_{\rm obs}$, given true mass M,

$$p(M_{\rm obs}|M) = \frac{1}{\sqrt{2\pi}\sigma_{\ln M}} \exp[-x^2(M_{\rm obs})], \quad (A1)$$

where

$$x(M_{\rm obs}) \equiv \frac{\ln M_{\rm obs} - \ln M - \ln M^{\rm bias}(M_{\rm obs}, z)}{\sqrt{2}\sigma_{\ln M}(M_{\rm obs}, z)}. \quad (A2)$$

For the optical survey, the mass threshold of the observable is set to $M^{\rm th} = 10^{13.7} h^{-1} M_\odot$ and the redshift limit is $z = 1$, corresponding to the projected sensitivity of the Dark Energy Survey. Different studies suggest a wide range of scatter for optical observables, ranging from a constant $\sigma_{\ln M} = 0.5$ [75] to a mass-dependent scatter in the range $0.75 < \sigma_{\ln M} < 1.2$ [76]. Using weak lensing and x-ray analysis of MAXBCG selected optical clusters, Ref. [77] estimated a log-normal scatter of ~0.45 for $P(M|M_{\rm obs})$, where M was determined using weak lensing and $M_{\rm obs}$ was an optical richness estimate. We choose a fiducial mass scatter of $\sigma_{\ln M} = 0.5$ and allow for a cubic

evolution in redshift and mass:

$$\ln M^{\rm bias}(M_{\rm obs}, z) = \ln M_0^{\rm bias} + a_1 \ln(1 + z)$$
$$+ a_2(\ln M_{\rm obs} - \ln M_{\rm pivot}), \quad (A3)$$

$$\sigma_{\ln M}^2(M_{\rm obs}, z) = \sigma_0^2 + \sum_{i=1}^{3} b_i z^i$$
$$+ \sum_{i=1}^{3} c_i(\ln M_{\rm obs} - \ln M_{\rm pivot})^i. \quad (A4)$$

We set $M_{\rm pivot} = 10^{15} h^{-1} M_\odot$. In all, we have 10 nuisance parameters for the optical mass errors ($\ln M_0^{\rm bias}$, a_1, a_2, σ_0^2, b_i, c_i).

There are few, if any, constraints on the number of parameters necessary to realistically describe the evolution of the variance and bias with mass. Reference [54] shows that a cubic evolution of the mass scatter with redshift captures most of the residual uncertainty when the redshift evolution is completely free [as assumed in the Dark Energy Task Force report [78]]. While generous, this parametrization assumes a log-normal distribution of the mass-observable relation that may fail for low-masses (see e.g. [79]). However, [51] shows that more complex distributions do not degrade results substantially ($\sim 20\%$–30% for the test case assumed by the authors). We have also implicitly assumed that selection effects can be described by the bias and scatter of the mass-observable relation. By the year 2016, we expect significant progress in simulations of cluster surveys that will allow us to better parametrize the cluster selection errors.

APPENDIX B: PHOTOMETRIC REDSHIFT ERRORS AND GAUSSIAN HALO BIAS

Uncertainties in the redshifts distort the volume element. Assuming photometric techniques are used to determine the redshifts of the clusters (hereafter photo-z's), and a perfect angular selection, the mean number of clusters in a photo-z bin $z_i^{\rm p} \le z^{\rm p} \le z_{i+1}^{\rm p}$ is

$$\bar{m}_{\alpha,i} = \int_{z_i^{\rm p}}^{z_{i+1}^{\rm p}} dz^{\rm p} \int dV \bar{n}_\alpha W_i^{\rm th}(\Omega) p(z^{\rm p}|z), \quad (B1)$$

where $W_i^{\rm th}(\Omega)$ is an angular top hat window function. We parametrize the probability of measuring a photometric redshift, $z^{\rm p}$, given the true cluster redshift z as [55]

$$p(z^{\rm p}|z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp[-y^2(z^{\rm p})], \quad (B2)$$

where

$$y(z^{\rm p}) \equiv \frac{z^{\rm p} - z - z^{\rm bias}}{\sqrt{2\sigma_z^2}}, \quad (B3)$$

z^{bias} is the photometric redshift bias, and σ_z is the scatter in the photo-z's.

On large scales, the number counts of clusters $m(\mathbf{x})$ trace the linear density perturbation $\delta(\mathbf{x})$

$$m_i(M_\alpha, \mathbf{x}) \equiv m_{i\alpha} = \bar{m}_i(1 + b(M_\alpha, z)\delta(\mathbf{x})), \qquad \text{(B4)}$$

where M_α denotes a bin in mass and i refers to the pixel on the sky defined by its angular location and redshift. The (Gaussian) halo bias may be very roughly approximated by [80]

$$b_0(M; z) = 1 + \frac{a_c \delta_c^2/\sigma^2 - 1}{\delta_c} + \frac{2p_c}{\delta_c[1 + (a\delta_c^2/\sigma^2)^{p_c}]} \tag{B5}$$

with $a_c = 0.75$, $p_c = 0.3$, and $\delta_c = 1.69$. Here, $\sigma(M, z)$ is the amplitude of mass fluctuations on scale M, defined as

usual by

$$\sigma^2 = \int \frac{k^3}{2\pi^2} P(k) W^2(kR) \frac{dk}{k}, \tag{B6}$$

where $W(x) = 3j_1(x)/x$ (the top hat window), $R = (3M/4\pi\bar{\rho}_m)^{1/3}$, and $P(k)$ and $\bar{\rho}_m$ are the matter power spectrum and energy density, respectively.

Integrating the expression above yields the average cluster linear bias:

$$b_{\alpha,i}(z) = \frac{1}{\bar{n}_{\alpha,i}(z)} \int_{M_{\text{obs}}^\alpha}^{M_{\text{obs}}^{\alpha+1}} \frac{dM_{\text{obs}}}{M_{\text{obs}}} \int \frac{dM}{M} \\ \times \frac{d\bar{n}_{\alpha,i}(z)}{d\ln M} b(M; z) p(M_{\text{obs}}|M). \tag{B7}$$

[1] J. Maldacena, J. High Energy Phys. 5 (2003) 013.
[2] V. Acquaviva, N. Bartolo, S. Matarrese, and A. Riotto, Nucl. Phys. **B667**, 119 (2003).
[3] P. Creminelli, J. Cosmol. Astropart. Phys. 10 (2003) 003.
[4] D. H. Lyth and Y. Rodriguez, Phys. Rev. Lett. **95**, 121302 (2005).
[5] D. Seery and J. E. Lidsey, J. Cosmol. Astropart. Phys. 06 (2005) 003.
[6] X. Chen, arXiv:1002.1416.
[7] T. Falk, R. Rangarajan, and M. Srednicki, Astrophys. J. **403**, L1 (1993).
[8] X.-c. Luo and D. N. Schramm, Phys. Rev. Lett. **71**, 1124 (1993).
[9] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach, Astrophys. J. **430**, 447 (1994).
[10] L.-M. Wang and M. Kamionkowski, Phys. Rev. D **61**, 063504 (2000).
[11] E. Komatsu and D. N. Spergel, Phys. Rev. D **63**, 063002 (2001).
[12] N. Bartolo, S. Matarrese, and A. Riotto, arXiv:1001.3957.
[13] M. Liguori, E. Sefusatti, J. R. Fergusson, and E. P. S. Shellard, arXiv:1001.4707.
[14] D. Babich, Phys. Rev. D **72**, 043003 (2005).
[15] D. Babich, P. Creminelli, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 08 (2004) 009.
[16] P. Creminelli, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 03 (2007) 019.
[17] K. M. Smith and M. Zaldarriaga, arXiv:astro-ph/0612571.
[18] J. R. Fergusson and E. P. S. Shellard, Phys. Rev. D **76**, 083523 (2007).
[19] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page et al., Astrophys. J. Suppl. Ser. **170**, 377 (2007).
[20] P. Creminelli, L. Senatore, M. Zaldarriaga, and M. Tegmark, J. Cosmol. Astropart. Phys. 03 (2007) 005.

[21] L. Senatore, K. M. Smith, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 01 (2010) 028.
[22] E. Komatsu, arXiv:1001.4538.
[23] F. Lucchin and S. Matarrese, Astrophys. J. **330**, 535 (1988).
[24] J. Robinson and J. E. Baker, arXiv:astro-ph/9905098.
[25] A. J. Benson, C. Reichardt, and M. Kamionkowski, Mon. Not. R. Astron. Soc. **331**, 71 (2002).
[26] S. Matarrese, L. Verde, and R. Jimenez, Astrophys. J. **541**, 10 (2000).
[27] L. Verde, M. Kamionkowski, J. J. Mohr, and A. J. Benson, Mon. Not. R. Astron. Soc. **321**, L7 (2001).
[28] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga, Phys. Rev. D **69**, 103513 (2004).
[29] E. Komatsu et al. (WMAP Collaboration), Astrophys. J. Suppl. Ser. **148**, 119 (2003).
[30] M. Kamionkowski, L. Verde, and R. Jimenez, J. Cosmol. Astropart. Phys. 01 (2009) 010.
[31] H. Song and J. Lee, Astrophys. J. Lett. **701**, L25 (2009).
[32] L. Verde, R. Jimenez, M. Kamionkowski, and S. Matarrese, Mon. Not. R. Astron. Soc. **325**, 412 (2001).
[33] L. Verde, L.-M. Wang, A. Heavens, and M. Kamionkowski, Mon. Not. R. Astron. Soc. **313**, L141 (2000).
[34] E. Sefusatti, C. Vale, K. Kadota, and J. Frieman, Astrophys. J. **658**, 669 (2007).
[35] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, Phys. Rev. D **77**, 123514 (2008).
[36] N. Afshordi and A. J. Tolley, Phys. Rev. D **78**, 123507 (2008).
[37] S. Matarrese and L. Verde, Astrophys. J. **677**, L77 (2008).
[38] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, J. Cosmol. Astropart. Phys. 08 (2008) 031.
[39] P. McDonald, Phys. Rev. D **78**, 123519 (2008).
[40] A. Taruya, K. Koyama, and T. Matsubara, Phys. Rev. D **78**, 123534 (2008).

[41] T. Giannantonio and C. Porciani, Phys. Rev. D **81**, 063530 (2010).

[42] M. Grossi, K. Dolag, E. Branchini, S. Matarrese, and L. Moscardini, Mon. Not. R. Astron. Soc. **382**, 1261 (2007).

[43] V. Desjacques, U. Seljak, and I. Iliev, arXiv:0811.2748.

[44] A. Pillepich, C. Porciani, and O. Hahn, arXiv:0811.4176.

[45] C. Carbone, L. Verde, and S. Matarrese, Astrophys. J. **684**, L1 (2008).

[46] C. Carbone, O. Mena, and L. Verde, arXiv:1003.0456.

[47] V. Desjacques and U. Seljak, Phys. Rev. D **81**, 023006 (2010).

[48] L. Verde and S. Matarrese, Astrophys. J. **706**, L91 (2009).

[49] A. Mantz, S. W. Allen, H. Ebeling, and D. Rapetti, Mon. Not. R. Astron. Soc. **387**, 1179 (2008).

[50] A. Vikhlinin et al., Astrophys. J. **692**, 1060 (2009).

[51] C. Cunha, D. Huterer, and J. A. Frieman, Phys. Rev. D **80**, 063532 (2009).

[52] M. Oguri, Phys. Rev. Lett. **102**, 211301 (2009).

[53] W. Hu and J. D. Cohn, Phys. Rev. D **73**, 067301 (2006).

[54] M. Lima and W. Hu, Phys. Rev. D **72**, 043006 (2005).

[55] M. Lima and W. Hu, Phys. Rev. D **76**, 123013 (2007).

[56] A. Jenkins, C. S. Frenk, S. D. M. White, J. M. Colberg, S. Cole, A. E. Evrard, H. M. P. Couchman, and N. Yoshida, Mon. Not. R. Astron. Soc. **321**, 372 (2001).

[57] W. Hu and A. V. Kravtsov, Astrophys. J. **584**, 702 (2003).

[58] M. Lima and W. Hu, Phys. Rev. D **70**, 043504 (2004).

[59] E. Komatsu et al. (WMAP Collaboration), Astrophys. J. Suppl. Ser. **180**, 330 (2009).

[60] U. Seljak and M. Zaldarriaga, Astrophys. J. **469**, 437 (1996).

[61] C. E. Cunha, Phys. Rev. D **79**, 063009 (2009).

[62] M. Manera, R. K. Sheth, and R. Scoccimarro, arXiv:0906.1314 [Mon. Not. Roy. Astron. Soc. (to be published)].

[63] U. Seljak, Phys. Rev. Lett. **102**, 021302 (2009).

[64] P. McDonald and U. Seljak, J. Cosmol. Astropart. Phys. 10 (2009) 007.

[65] D. Wands and A. Slosar, Phys. Rev. D **79**, 123507 (2009).

[66] J. Estrada, E. Sefusatti, and J. A. Frieman, Astrophys. J. **692**, 265 (2009).

[67] A. P. Hearin, A. R. Zentner, Z. Ma, and D. Huterer, arXiv:1002.3383.

[68] B. Sartoris et al., arXiv:1003.0841.

[69] H. A. Feldman, N. Kaiser, and J. A. Peacock, Astrophys. J. **426**, 23 (1994).

[70] M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay, Astrophys. J. **499**, 555 (1998).

[71] M. LoVerde, A. Miller, S. Shandera, and L. Verde, J. Cosmol. Astropart. Phys. 04 (2008) 014.

[72] A. L. Fitzpatrick, L. Senatore, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 05 (2010) 004.

[73] N. Bartolo, S. Matarrese, O. Pantano, and A. Riotto, Classical Quantum Gravity **27**, 124009 (2010).

[74] E. Komatsu, arXiv:astro-ph/0206039.

[75] H.-Y. Wu, E. Rozo, and R. H. Wechsler, Astrophys. J. **688**, 729 (2008).

[76] M. R. Becker, T. A. McKay, B. Koester, R. H. Wechsler, E. Rozo, A. Evrard, D. Johnston, E. Sheldon, J. Annis, E. Lau et al., Astrophys. J. **669**, 905 (2007).

[77] E. Rozo, E. S. Rykoff, A. Evrard, M. Becker, T. McKay, R. H. Wechsler, B. P. Koester, J. Hao, S. Hansen, E. Sheldon et al., Astrophys. J. **699**, 768 (2009).

[78] A. J. Albrecht et al., arXiv:astro-ph/0609591.

[79] J. D. Cohn and M. White, Mon. Not. R. Astron. Soc. **393**, 393 (2009).

[80] R. K. Sheth and G. Tormen, Mon. Not. R. Astron. Soc. **308**, 119 (1999).

[81] W. Hu, private communication.