Bispectrum of galaxies from high-redshift galaxy surveys:
Primordial non-Gaussianity and nonlinear galaxy bias

Emiliano Sefusatti[*]

Particle Astrophysics Center, Fermi National Accelerator Laboratory, Batavia, Illinois 60510-0500, USA

Eiichiro Komatsu

Department of Astronomy, University of Texas at Austin, 2511 Speedway, RLM 15.306, Texas 78712, USA
(Received 21 May 2007; published 17 October 2007)

The greatest challenge in the interpretation of galaxy clustering data from any surveys is galaxy bias. Using a simple Fisher matrix analysis, we show that the bispectrum provides an excellent determination of linear and nonlinear bias parameters of intermediate and high-z galaxies, when all measurable triangle configurations down to mildly nonlinear scales, where perturbation theory is still valid, are included. The bispectrum is also a powerful probe of primordial non-Gaussianity. The planned galaxy surveys at $z \gtrsim 2$ should yield constraints on non-Gaussian parameters, $f_{NL}^{\text{loc.}}$ and $f_{NL}^{\text{eq.}}$, that are comparable to, or even better than, those from cosmic microwave background experiments. We study how these constraints improve with volume and redshift range, as well as the number density of galaxies. Finally, we show that a halo occupation distribution may be used to improve these constraints further by lifting degeneracies between gravity, bias, and primordial non-Gaussianity.

I. INTRODUCTION

Why study high-z galaxy surveys? The recognition that baryon acoustic oscillations in the galaxy power spectrum [1–4] are an excellent probe of the nature of dark energy has led to several proposals for large-volume redshift surveys at $z \gtrsim 1$.

Galaxies are a biased tracer of the underlying matter distribution. The use of highly biased tracers, such as luminous red galaxies at lower z and Lyman break galaxies or Lyman-α emitters at higher z, requires a reliable modeling of nonlinearity and scale dependence of galaxy bias, even at relatively large spatial scales [5–7].

The distribution of galaxies is non-Gaussian. The galaxy bispectrum, the three-point correlation function in Fourier space, does not vanish. It has been known for more than a decade that the bispectrum is an excellent tool for measuring galaxy bias parameters, independent of the overall normalization of dark matter fluctuations [8–10]. This method has been applied successfully to existing galaxy surveys such as the 2dF Galaxy Redshift Survey (2dFGRS) and the Sloan Digital Sky Survey (SDSS), yielding constraints on nonlinearities in galaxy bias [11–14] as well as on the halo occupation distribution (HOD) [15]. Moreover, it has been shown that the galaxy bispectrum contains additional cosmological information that is not present in the power spectrum [16,17].

The galaxy bispectrum on large scales, or other statistical tools that are sensitive to the higher-order correlation of galaxies, are sensitive to statistical properties of primordial

fluctuations: primordial non-Gaussianity [18–24]. The common belief is that the cosmic microwave background (CMB) is most sensitive to primordial non-Gaussianity; however, the number of Fourier modes available in the galaxy survey data is much larger than that in the CMB data, as the former probes the full three-dimensional structure of density fields. The galaxy bispectrum can, therefore, outperform the CMB bispectrum, provided that the other sources of non-Gaussianity, such as nonlinear gravitational evolution and galaxy bias, are under control. Furthermore, it should always be emphasized that galaxy surveys provide information on the spatial scales that are much smaller than those probed by CMB; thus, these two measurements are complementary to each other.

Motivated by these considerations, in this paper we study how well the planned high-z galaxy surveys would constrain primordial non-Gaussianity and galaxy bias parameters. High-z surveys are more useful for this task than low-z ones, owing to much weaker nonlinearities in matter clustering and redshift-space distortion, which allows us to use the galaxy bispectrum to the smaller spatial scales that are inaccessible by low-z surveys.

This paper is organized as follows. In Sec. II we describe the leading contributions to the bispectrum of the observed galaxy distribution. In Sec. III we present our Fisher matrix analysis for the galaxy bispectrum, and in Sec. IV we show our predictions for constraints on galaxy bias and primordial non-Gaussianity from sample galaxy redshift survey designs, studying dependence on the survey volume, maximum wave number, number density of galaxies, and redshifts. We also present our predictions for the ongoing, upcoming, and planned galaxy surveys. In Sec. V we

*emiliano@fnal.gov

extend our Fisher matrix analysis to the halo occupation distribution as a tool to describe galaxy biasing at large scales. Finally, in Sec. VI we present our conclusions.

II. THE GALAXY BISPECTRUM

A. Primordial non-Gaussianity

We explore two parametrizations of primordial non-Gaussianity which are motivated by inflationary models. While representing a wide variety of non-Gaussian models, these parametrizations are by no means exhaustive. Our method can be applied to any other functional forms of the bispectrum (e.g., [25,26]) in a straightforward way.

1. Local model

The first one is described by the *local* expression for Bardeen's curvature perturbations during the matter era, $\Phi(x)$, in position space [24,27,28],

$$\Phi(\mathbf{x}) = \Phi_G(\mathbf{x}) + f_{NL}^{\text{loc.}}[\Phi_G^2(\mathbf{x}) - \langle\Phi_G^2(\mathbf{x})\rangle], \qquad (1)$$

where $\Phi_G(\mathbf{x})$ is a Gaussian field and $f_{NL}^{\text{loc.}}$ is a constant characterizing the amplitude of primordial non-Gaussianity. In this case, the leading contribution in the $f_{NL}^{\text{loc.}}$ expansion to the bispectrum, $B_\Phi^{\text{local}}(k_1, k_2, k_3)$, of the curvature field is given by

$$
\begin{aligned}
B_\Phi^{\text{local}} &\simeq 2f_{NL}^{\text{loc.}}[P_\Phi(k_1)P_\Phi(k_2) + \text{cyc.}] \\
&= 2f_{NL}^{\text{loc.}}C_\Phi^2\left[\frac{1}{k_1^{4-n_s}k_2^{4-n_s}} + \text{cyc.}\right],
\end{aligned} \qquad (2)
$$

where we approximate $P_\Phi(k) \simeq P_{\Phi_G}(k)$, and

$$C_\Phi \equiv \frac{P_\Phi(k)}{k^{n_s-4}}, \qquad (3)$$

which quantifies departure from a scale-invariant spectrum. We include n_s explicitly, as we are interested in determining how a departure from scale invariance, which has been hinted by WMAP [29], would affect detectability of primordial non-Gaussianity in the distribution of galaxies. For this local model, most of the signal is given by squeezed triangular configurations, $k_1 \ll k_2, k_3$. The *local* type of non-Gaussianity described by Eq. (2) is predicted in models such as the curvaton scenario [30], models with inhomogeneous reheating [31,32], multiple field inflationary models [33], or generically in models where the nonlinearities arise from the evolution of perturbations outside the horizon.

The best limits to date on possible values for the $f_{NL}^{\text{loc.}}$ parameter come from measurements of the microwave background bispectrum on the WMAP data [29,34,35] $-36 \leq f_{NL}^{\text{loc.}} \leq 100$ at 95% C.L., which corresponds to the 1-σ error of

$$\Delta f_{NL}^{\text{loc.}} = 34 \quad (\text{WMAP3}), \qquad (4)$$

which is a factor of 50 better than the limit from COBE

[36]. Upon completion, WMAP is expected to reach $\Delta f_{NL}^{\text{loc.}} \simeq 20$, while the Planck satellite would yield $\Delta f_{NL}^{\text{loc.}} \simeq 3$ [28,37,38].

Measurements of the galaxy bispectrum in the SDSS main sample are expected to yield $\Delta f_{NL}^{\text{loc.}} \simeq 150$ [23]. Recently [39] pointed out that a full-sky measurement of the bispectrum of fluctuations in the 21-cm background might reach $f_{NL}^{\text{loc.}} \simeq 1$, while a more aggressive analysis by Cooray [40] shows that the same observations could reach $\Delta f_{NL}^{\text{loc.}} \simeq 0.01$ in principle. Other large-scale structure probes such as cluster abundance, on the other hand, are unlikely to improve CMB limits on $f_{NL}^{\text{loc.}}$ [41]; however, it should provide an important cross-check of the results if a significant $f_{NL}^{\text{loc.}}$ was detected in the CMB, and it should not be forgotten that the spatial scales probed by the cluster abundance are smaller than those probed by the CMB.

2. Equilateral model

The second model for primordial non-Gaussianity is given by

$$
\begin{aligned}
B_\Phi^{\text{equil.}} = 6f_{NL}^{\text{eq.}}C_\Phi^2\bigg[&-\frac{1}{k_1^{4-n_s}k_2^{4-n_s}} - \frac{1}{k_1^{4-n_s}k_3^{4-n_s}} \\
&- \frac{1}{k_2^{4-n_s}k_2^{4-n_s}} - \frac{2}{(k_1k_2k_3)^{2(4-n_s)/3}} \\
&+ \left(\frac{1}{k_1^{(4-n_s)/3}k_2^{2(4-n_s)/3}k_3^{4-n_s}} + \text{cyc.}\right)\bigg].
\end{aligned} \qquad (5)
$$

Creminelli *et al.* and Babich *et al.* [34,42] have shown that this form provides a good approximation to the bispectra predicted by higher derivatives and Dirac-Born-Infeld (DBI) inflationary models [43,44].

The bispectrum in Eq. (5) is normalized in such a way that, for equilateral configurations ($k_1 = k_2 = k_3 = k$), it coincides with the local form given in Eq. (2). The important difference is that this form has the largest contribution from the equilateral configurations, as opposed to the local form in which the largest contribution comes from the squeezed configurations. The current limits from WMAP are $-256 \leq f_{NL}^{\text{eq.}} \leq 332$ at 95% C.L. [34], which corresponds to the 1-σ error of

$$\Delta f_{NL}^{\text{eq.}} = 147 \quad (\text{WMAP3}). \qquad (6)$$

3. The primordial density bispectrum

Density fluctuations in Fourier space, $\delta_{\mathbf{k}}$, are related to the curvature perturbations, $\Phi_{\mathbf{k}}$, via the Poisson equation, $\delta_{\mathbf{k}}(a) = M(k; a)\Phi_{\mathbf{k}}$, where

$$M(k, a) \equiv \frac{2}{3}\frac{D(a)}{H_o^2\Omega_m}k^2T(k). \qquad (7)$$

Here a is the scale factor, $T(k)$ is the matter transfer function, and $D(a)$ is the growth function [45].

This allows us to write the primordial contribution to a generic n-point function of the matter density fields in terms of the respective correlator of the curvature perturbations as

$$\langle \delta_{\mathbf{k}_1} \delta_{\mathbf{k}_2} \ldots \delta_{\mathbf{k}_N} \rangle_I = M(k_1; a) M(k_2; a) \ldots M(k_N; a)$$
$$\times \langle \Phi_{\mathbf{k}_1} \Phi_{\mathbf{k}_2} \ldots \Phi_{\mathbf{k}_N} \rangle. \qquad (8)$$

In particular, the initial (primordial) matter bispectrum, $B_I(k_1, k_2, k_3)$, is given by

$$B_I(k_1, k_2, k_3) = M(k_1) M(k_2) M(k_3) B_\Phi(k_1, k_2, k_3), \qquad (9)$$

where we have omitted for brevity the explicit dependence of $M(k; a)$ on a. We can also relate the linear density power spectrum $P_L(k)$ to the curvature power spectrum $P_\Phi(k)$ as

$$P_L(k) = M^2(k) P_\Phi(k). \qquad (10)$$

A hierarchical relation between the scale dependence of the initial bispectrum and the power spectrum such as $B_\Phi(k) \sim f_{NL} P_\Phi^2(k)$ with a constant f_{NL} is by no means generic or universal. A f_{NL} with a peculiar scale dependence unrelated to the power spectrum appears, for instance, in string-motivated models such as DBI inflation [43,46]. Our analysis can be applied to any models of primordial non-Gaussianity, provided that the bispectrum can be calculated from those models.

Note that a post-Newtonian effect can yield an additional contribution to nonlinearity of primordial perturbations and hence to non-Gaussianity [47]. Although we do not include this effect in our analysis, it would be interesting to study how important the post-Newtonian term would be for the future galaxy surveys.

B. Non-Gaussianity from nonlinear gravitational evolution

Even if the initial perturbations are Gaussian, the subsequent gravitational evolution makes the evolved density fields non-Gaussian. On large scales one can study the nonlinear evolution of matter density fluctuations by means of perturbation theory, and write the solution up to the second order in δ as

$$\delta_{\mathbf{k}} \simeq \delta_{\mathbf{k}}^{(1)} + \int d^3 q_1 d^3 q_2 \delta_D(\mathbf{k} - \mathbf{q}_{12}) F_2(\mathbf{q}_1, \mathbf{q}_2) \delta_{\mathbf{q}_1}^{(1)} \delta_{\mathbf{q}_2}^{(1)}, \qquad (11)$$

where $\delta^{(1)}$ is the linear solution, and $F_2(\mathbf{k}_1, \mathbf{k}_2)$ is a known mathematical function given by

$$F_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{x}{2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}x^2, \qquad (12)$$

with $x \equiv \hat{\mathbf{k}}_1 \cdot \hat{\mathbf{k}}_2$. Therefore, one obtains

$$B_G(k_1, k_2, k_3) = 2 F_2(\mathbf{k}_1, \mathbf{k}_2) P_L(k_1) P_L(k_2) + \text{cyc.} \qquad (13)$$

The bispectrum of matter density fluctuations (i.e., no galaxies yet) evolved from non-Gaussian primordial fluctuations on large scales is thus given by the sum of Eqs. (9) and (13):

$$B(k_1, k_2, k_3) = B_I(k_1, k_2, k_3) + B_G(k_1, k_2, k_3). \qquad (14)$$

As usual, we shall focus on the *reduced* bispectrum, defined as

$$Q(k_1, k_2, k_3) \equiv \frac{B(k_1, k_2, k_3)}{P(k_1)P(k_2) + \text{cyc.}} \qquad (15)$$

which has an advantage of being only mildly sensitive to cosmological parameters. That is to say, the dependence on cosmology has been "factored out" by a product of the power spectra in the denominator and the Q_G component is particularly insensitive to the amplitude of matter fluctuations (e.g., σ_8). The reduced bispectrum of matter density fluctuations is also given by the sum of two contributions:

$$Q(k_1, k_2, k_3) = Q_I(k_1, k_2, k_3) + Q_G(k_1, k_2, k_3)$$
$$= \frac{B_I(k_1, k_2, k_3)}{P(k_1)P(k_2) + \text{cyc.}} + \frac{B_G(k_1, k_2, k_3)}{P(k_1)P(k_2) + \text{cyc.}}. \qquad (16)$$

It is important to remember that, in the leading order, Q_G does not depend on the linear growth factor $D(a)$, and thus it is independent of redshifts. In other words, B_G is proportional to D^4, which cancels D^4 in $[P(k)]^2$ in the denominator exactly. On the other hand, Q_I is proportional to $1/D(a)$ because $B_I \propto D^3$, and thus it is larger at higher redshifts. Therefore, high-z galaxy surveys are expected to be more sensitive to primordial non-Gaussianity, relative to the gravitational bispectrum, than low-z ones.

In Fig. 1 we plot the equilateral configurations of the reduced bispectrum, $Q(k) \equiv Q(k, k, k)$, from nonlinear gravitational evolution and non-Gaussian initial conditions at $z = 0$, 1, and 4. As mentioned earlier, the local and equilateral model of primordial non-Gaussianity give the same results for these configurations. For Gaussian initial conditions, $f_{NL} = 0$, $Q(k) = 0.57$ at tree level in perturbation theory and is independent of scales [48]. On the other hand, $Q(k)$ exhibits a clear scale dependence for $f_{NL} \neq 0$. A positive f_{NL} enhances $Q(k)$ at large scales, i.e., $Q(k) > 0.57$, whereas a negative f_{NL} suppresses it. This is because a positive f_{NL} results in positively skewed density fluctuations. It is easy to show that the scale dependence of the primordial component $Q_I(k)$ is given by $1/M(k)$, where $M(k)$ is given by Eq. (7). As $1/M(k) \propto k^{-2}$ on large scales, we find that the primordial non-Gaussian signal is larger on large scales. This property makes it easier to find primordial non-Gaussianity in CMB observations; however, as there are many more modes available on smaller scales, the *cumulative* signal-to-noise ratio for higher-order correlation functions increases due to the large number of observable configurations on small scales [17].

How about other configurations? In Fig. 2 we plot the dark matter bispectrum with Gaussian or non-Gaussian
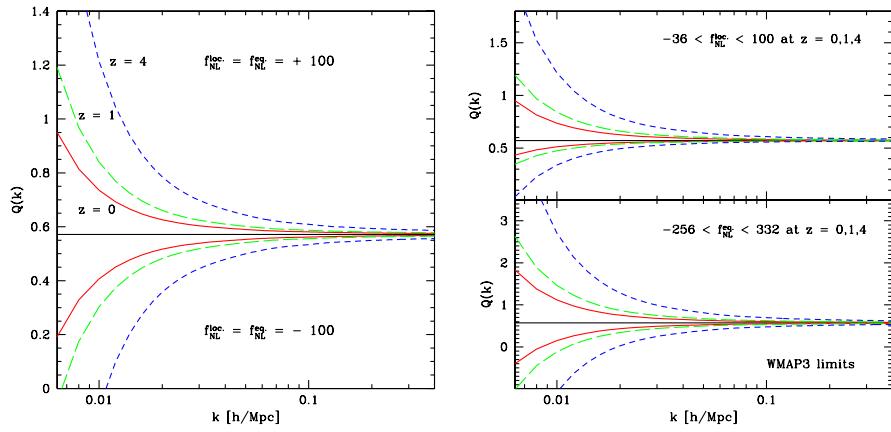
FIG. 1 (color online).   Equilateral configurations of the reduced bispectrum of dark matter distribution at the second order ("tree-level"). The horizontal lines at $Q(k) = 0.57$ show the gravitational contribution only, which corresponds to $f_{NL}^{loc.} = 0 = f_{NL}^{eq.}$. The solid, long-dashed, and short-dashed lines show the gravitational contribution with non-Gaussian initial perturbations at $z = 0$, 1, and 4, respectively. Note that $f_{NL}^{loc.} = f_{NL}^{eq.}$ for equilateral configurations. Left panel: The curves above $Q(k) = 0.57$ show $f_{NL} = +100$, while the curves below it show $f_{NL} = -100$. Right panel: The same as the left panel but for the current WMAP3 limits on $f_{NL}^{loc.}$ (top) and $f_{NL}^{eq.}$ (bottom).
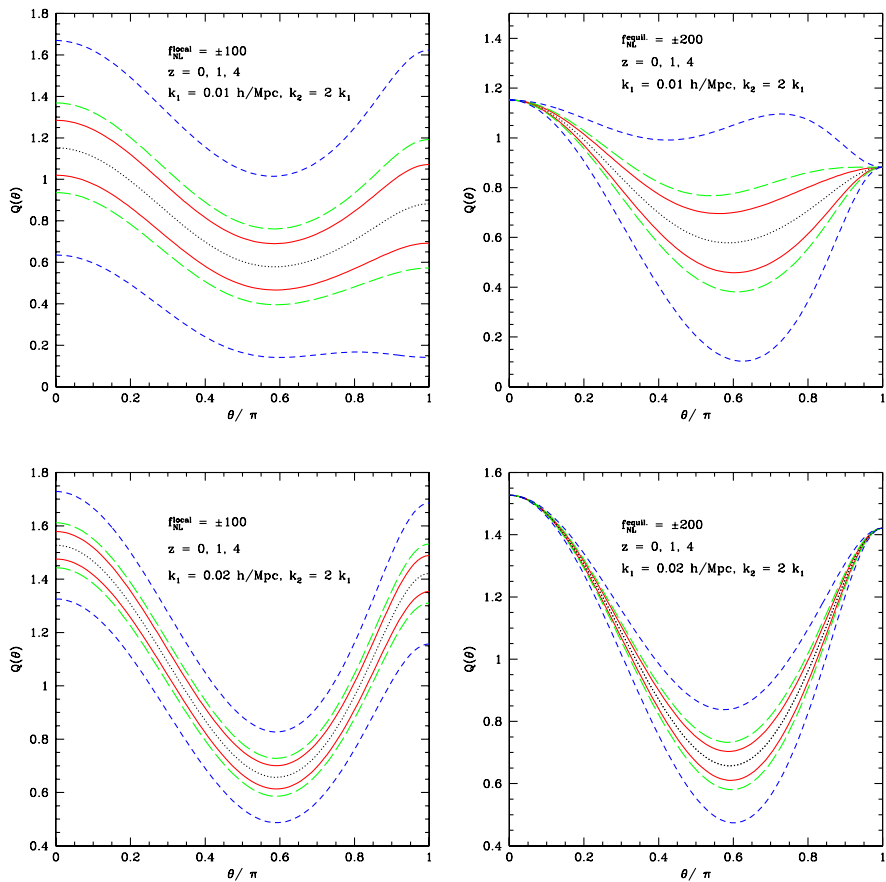

FIG. 2 (color online).   Configuration dependence of the reduced bispectrum of dark matter distribution from Gaussian and non-Gaussian initial conditions, as a function of an angle, θ, between two wave vectors, $\mathbf{k}_1$ and $\mathbf{k}_2$, where the magnitude satisfies $k_2 = 2k_1 = 0.02\, h\,\mathrm{Mpc}^{-1}$ (top panels) and $0.04\, h\,\mathrm{Mpc}^{-1}$ (bottom panels). Left panel: $f_{NL}^{loc.} = \pm 100$. Right panels: $f_{NL}^{eq.} = \pm 200$. The dotted black line shows a Gaussian case ($f_{NL} = 0$, redshift independent), while the solid, long-dashed, and short-dashed lines show non-Gaussian cases at $z = 0$, 1, and 4, respectively.

initial conditions at different redshifts as a function of the angle, θ, between $\mathbf{k}_1$ and $\mathbf{k}_2$, for $k_1 = 0.01\, h\,\text{Mpc}^{-1}$ (top panels) and $0.02\, h\,\text{Mpc}^{-1}$ (bottom panels) and $k_2 = 2k_1$. We find a marked difference in the configuration dependence of the two primordial bispectra, local (left panels) and equilateral (right panels), under consideration.

C. Non-Gaussianity from galaxy bias

The galaxy bispectrum is most useful for measuring galaxy bias. Assuming that galaxy formation is a local process and depends only on the local matter density field, one may expand the galaxy number overdensity, δ_g, in Taylor series of the underlying matter overdensity, δ, as [49]

$$\delta_g(\mathbf{x}) \simeq b_1 \delta(\mathbf{x}) + \tfrac{1}{2} b_2 \delta^2(\mathbf{x}), \qquad (17)$$

where b_1 and b_2 are the linear and nonlinear bias parameters, respectively. It has been shown that this model describes the bispectrum or three-point correlation functions from the SDSS and 2dFGRS (e.g., [11,12]), as well as from numerical simulations (e.g., [50]). The galaxy bispectrum is given by

$$\begin{aligned} B_g(k_1, k_2, k_3) &\simeq b_1^3 B(k_1, k_2, k_3) \\ &\quad + b_1^2 b_2 [P_L(k_1) P_L(k_2) + \text{cyc.}], \end{aligned} \qquad (18)$$

up to the second order in matter density fluctuations. Here, $B(k_1, k_2, k_3)$ is the intrinsic bispectrum of the underlying matter distribution. The reduced galaxy bispectrum is

$$Q_g(k_1, k_2, k_3) \simeq \frac{1}{b_1} Q(k_1, k_2, k_3) + \frac{b_2}{b_1^2}. \qquad (19)$$

With $Q(k_1, k_2, k_3) = Q_G(k_1, k_2, k_3) + f_{NL} \tilde{Q}_I(k_1, k_2, k_3)$, one obtains

$$Q_g(k_1, k_2, k_3) \simeq \frac{1}{b_1} Q_G(k_1, k_2, k_3) + \frac{f_{NL}}{b_1} \tilde{Q}_I(k_1, k_2, k_3) + \frac{b_2}{b_1^2}, \qquad (20)$$

where we have factorized f_{NL} out from Q_I introducing $\tilde{Q}_I \equiv Q_I(f_{NL} = 1)$.

Finally, the galaxy power spectrum is given by $P_g(k) \simeq b_1^2 P(k)$, up to the second order in density fluctuations; however, corrections due to nonlinear bias appear at the third-order level [5,6,51,52]. Therefore, the bias parameters, b_1 and b_2, affect both the galaxy power spectrum and bispectrum. The bispectrum helps us extract the cosmological information from the galaxy power spectrum by providing b_1 and b_2.

D. Redshift-space distortion

The bispectrum measured from redshift surveys is distorted along the line of sight direction by radial motion of galaxies. For our analysis in this paper we shall deal only with a spherically averaged power spectrum and bispec-

trum. The power spectrum in redshift space after averaging over angles in k space, $P_s(k)$, is related to the real space power spectrum by

$$P_s(k) = a_0^P(\beta) P_g(k), \qquad (21)$$

while the bispectrum is given by

$$B_s(k_1, k_2, k_3) = a_0^B(\beta) B_g(k_1, k_2, k_3), \qquad (22)$$

where [16,53]

$$a_0^P(\beta) = 1 + \tfrac{2}{3}\beta + \tfrac{1}{5}\beta^2, \qquad (23)$$

$$a_0^B(\beta) = 1 + \tfrac{2}{3}\beta + \tfrac{1}{9}\beta^2, \qquad (24)$$

with $\beta \equiv \Omega_m^{5/7}/b_1$. The reduced bispectrum in redshift space is thus given by

$$Q_s(k_1, k_2, k_3) = \frac{a_0^B(\beta)}{[a_0^P(\beta)]^2}\left[\frac{1}{b_1} Q(k_1, k_2, k_3) + \frac{b_2}{b_1^2}\right]. \qquad (25)$$

We remark that our treatment does not take into account a peculiar scale dependence of redshift distortions that reduces the amplitude of nonlinear corrections to the reduced matter bispectrum [54]. This fact only partially justifies the approximation of neglecting such corrections altogether.

III. FISHER MATRIX ANALYSIS

A. Method

In our analysis we shall consider a set of surveys characterized by their volume V, mean galaxy density n_g, and redshift range. We shall assume that these surveys have a simple survey geometry, i.e., a contiguous hexahedron.

Our bispectrum estimator is given by [55]

$$\hat{B} \equiv \frac{V_f}{V_B} \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \, \delta_D(\mathbf{q}_{123}) \delta_{\mathbf{q}_1} \delta_{\mathbf{q}_2} \delta_{\mathbf{q}_3}, \qquad (26)$$

where the integration is over the bin defined by $q_i \in (k_i - \delta k/2, k_i + \delta k/2)$, $V_f = (2\pi)^3/V$ is the volume of the fundamental cell in Fourier space, and

$$\begin{aligned} V_B &\equiv \int_{k_1} d^3 q_1 \int_{k_2} d^3 q_2 \int_{k_3} d^3 q_3 \, \delta_D(\mathbf{q}_{123}) \\ &\simeq 8\pi^2 k_1 k_2 k_3 \Delta k^3, \end{aligned} \qquad (27)$$

with Δk a multiple of the fundamental frequency, $k_f \equiv 2\pi/L$. We assume that two coincide, i.e., $\Delta k = k_f$, thereby taking into account all "fundamental" triangular configurations.

The variance for our estimator, to the leading order, is given by a triple product of the power spectra,

$$\Delta B^2 \simeq k_f^3 \frac{s_{123}}{V_B} P(k_1) P(k_2) P(k_3), \qquad (28)$$

where $s_{123} = 6, 2, 1$ for equilateral, isosceles, and general triangles, respectively. As for variance of the reduced bispectrum, we assume that variance from the bispectrum in the numerator dominates over that from the power spectra in the denominator:

$$\frac{\Delta Q^2}{Q^2} \simeq \frac{\Delta B^2}{B^2}. \tag{29}$$

We calculate variance of the redshift-space galaxy reduced bispectrum from Eqs. (28) and (29) with $P(k)$ given by

$$P_{\text{tot}}(k) \equiv P_s(k) + \frac{1}{(2\pi)^3} \frac{1}{\bar{n}}, \tag{30}$$

where the second term accounts for the shot noise. We finally obtain

$$\Delta Q_s^2(k_1, k_2, k_3) \simeq \frac{s_{123} k_f^3}{V_B} \frac{P_{\text{tot}}(k_1) P_{\text{tot}}(k_2) P_{\text{tot}}(k_3)}{[P_s(k_1) P_s(k_2) + \text{cyc.}]^2}. \tag{31}$$

Once the variance of the reduced bispectrum is given, the Fisher matrix for a given redshift bin can be expressed as

$$F_{\alpha\beta} \equiv \sum_{k_1, k_2, k_3 \leq k_{\max}} \frac{\partial Q_s(i)}{\partial p_\alpha} \frac{\partial Q_s(i)}{\partial p_\beta} \frac{1}{\Delta Q_s^2(i)}, \tag{32}$$

where the parameters p_α represent b_1, b_2, and f_{NL}. We use Eq. (25) for Q_s, which is valid only up to the second order in perturbations. Neglecting higher-order corrections would introduce systematic errors when dealing with the real data, particularly at low z and at small spatial scales, where nonlinearity is substantial and perturbation theory essentially breaks down. Since we consider high-z surveys on large scales, we expect that higher-order effects would not affect our results very much.

As a fiducial cosmological model we use a flat ΛCDM cosmology with matter density $\Omega_m = 0.3$, baryon density $\Omega_b = 0.04$, Hubble parameter $h = 0.7$, spectral index $n_s = 1$, and $\sigma_8 = 0.9$. Of these parameters, σ_8 and n_s affect our forecast for the projected errors on the bias parameters most. We thus also consider different values such as those suggested by the WMAP 3-yr results, $\sigma_8 = 0.75$ and $n_s = 0.95$ [29].

B. Comments on covariance matrix

We shall not include covariance between the cosmological parameters and b_1, b_2, and f_{NL}. (We do include covariance between b_1, b_2, and f_{NL}.) Sefusatti *et al.* [16] have shown that, for the SDSS main sample, an analysis with the full covariance among all parameters with a prior from the WMAP 3-yr results yields an error on b_1 that is twice as large as that from a simpler analysis without covariance. On the other hand, an error on b_2 is not affected significantly. Note that the effect on b_1 that they observed was due mainly to degeneracy between b_1 and the amplitude of matter fluctuations, as Sefusatti *et al.* [16] did not

use the reduced bispectrum. We expect that degeneracy would be lifted in our analysis, as we use the reduced bispectrum in which the overall amplitude of matter fluctuations cancels.

More importantly, we shall not include covariance between different triangular configurations in the bispectrum. The covariance arises from both observational selection functions (i.e., survey geometry and mask) and a connected six-point function generated by nonlinear gravitational evolution. This is a rather crude approximation. Scoccimarro *et al.* [23] have included the full reduced bispectrum covariance plus the peculiar survey geometry, when they calculated the constraints on galaxy bias and primordial non-Gaussianity from the SDSS main sample. Using the same realizations of the survey and the same estimator for the covariance matrix, Sefusatti and Scoccimarro [17] have compared the analysis with the full covariance matrix (including the observational selection function) and that with an approximate diagonal Gaussian variance. They have found that the latter simplified treatment overestimates the signal-to-noise ratio by a factor of 2 for $k_{\max} \sim 0.1\ h\,\text{Mpc}^{-1}$, and a factor of 8 for $k_{\max} \sim 0.3\ h\,\text{Mpc}^{-1}$ at redshift zero. It is not clear, however, how to separate the contribution from nonlinear evolution from the effect of the selection function. One would generically expect the radial contribution to be smaller at high z, as the six-point function from gravitational clustering becomes smaller than the nonconnected part of the six-point function (which consists of power spectra) at higher z. In any case, our results should be taken as a guide, and one needs to perform the full analysis including the selection functions peculiar to a given survey design.

C. Nonlinearity and maximum wave number

While one can measure the galaxy power spectrum or bispectrum down to very small spatial scales, say, 10 kpc, it is challenging to extract useful cosmological information from such small spatial scales owing to strong nonlinearity. Therefore, one has to decide on the maximum wave number, $k_{\max}$, below which theory may be trusted. Not surprisingly, since there are many more modes available on smaller spatial scales, the amount of cosmological information one can extract from data grows as $k_{\max}$ increases. It is therefore important to use a realistic $k_{\max}$ in order not to overestimate the statistical power of a given galaxy survey design.

How do we decide on $k_{\max}$? The first obvious thing to do would be to test our theory of the power spectrum and bispectrum against numerical simulations. A value of $k_{\max}$ can be found by comparing perturbation theory predictions with numerical simulations (see, e.g., [56], for an analysis for the matter power spectrum). It is likely that a simple model provided by the second-order (tree-level) bispectrum given by Eq. (25) breaks down at a relatively small k

due to nonlinearities of gravitational growth as well as nonlinear or even nonlocal bias. Therefore, a model of the bispectrum that takes into account higher-order perturbations would be necessary to push k_{max} further. New promising techniques such as a renormalized perturbation theory approach [57–61] may be used to obtain better predictions for the power spectrum and bispectrum. Further progress is required, particularly for understanding redshift distortions [62].

In this paper we use a very simple prescription for getting k_{max}. We choose k_{max} so that $\sigma(R_{min}, z) = 0.5$ and $k_{max} = \pi/(2R_{min})$. The main motivation for this choice is that for small perturbations in the matter distribution, say, $\sigma(R, z) < 1$, one may reasonably expect that an analytical model for nonlinearities is viable. Note that k_{max} derived in this way depends on z, as $\sigma(R, z) = \sigma(R, 0)D(z)/D(0)$.

An alternative, much more conservative estimate of k_{max} could be given by requiring that the error on b_1 derived from the tree-level bispectrum for some k_{max} always exceeds the higher-order ("1-loop," or fourth-order perturbation) corrections in perturbation theory to the reduced matter bispectrum at the same k_{max}. Specifically, one may use

$$\frac{\Delta b_1}{b_1} \geq \frac{\Delta Q_{eq}^{1\text{-loop}}(k_{max})}{Q_{eq}^{tree}(k_{max})}, \qquad (33)$$

to determine k_{max}. Here, Δb_1 is computed including all scales down to k_{max} and $\Delta Q^{1\text{-loop}}(k_{max})$ are the 1-loop corrections [55,63] to the tree-level reduced bispectrum Q^{tree} evaluated for equilateral configurations with $k_1 = k_2 = k_3 = k_{max}$. This approach, however, makes no use of a large amount of information on small scales and is far from being optimal. Also, the 1-loop correction to the bispectrum, as it is the case for the power spectrum, tends to overestimate the nonlinear behavior measured in simulations, thereby making this approach even more conservative than necessary. In Sec. IV A 2 we shall compare these two approaches as a function of volume and number density. For the expressions of the 1-loop corrections to the reduced bispectrum see, e.g., Bernardeau *et al.* [64].

D. Fiducial values for the galaxy bias parameters

The galaxy bias parameters b_1 and b_2 depend on a number of factors, including galaxy populations, luminosities, and redshifts. On the other hand, the bias of dark matter halos, which can be calculated from N-body simulations, is understood relatively well. Therefore, the galaxy bias can be calculated from the dark matter halo bias, if we assume that galaxies form in dark matter halos. To do this, one needs (at least) the following information: (i) the halo bias [65,66] and (ii) how each halo is populated with galaxies, that is, the HOD, $\langle N \rangle_M$.

We calculate the galaxy bias parameters from the large-scale expression

$$b_i \simeq \frac{1}{n_g} \int_{M_{min}} dM n_h(M, z) b_i^h(M, z) \langle N \rangle_M, \qquad (34)$$

for $i = 1$ and 2, where $n_h(M, z)$ is the mass function of dark matter halos of mass M at redshift z, $b_i^h(M, z)$ is the halo bias function, and the HOD $\langle N \rangle_M$ is the mean number of galaxies per halo of a given mass M. We shall use the Sheth and Tormen formula for $n_h(M, z)$ [66]:

$$n_h(M, z) = -\frac{\bar{\rho}}{M^2} \frac{d\ln\sigma}{d\ln M} f(\nu) \qquad (35)$$

where $\nu = \delta_c/\sigma(M, z)$ with $\delta_c = 1.686$, and

$$f(\nu) = A\sqrt{\frac{2q}{\pi}}[1 + (q\nu^2)^{-p}]\nu e^{-q\nu^2/2} \qquad (36)$$

with $A = 0.322$, $p = 0.3$, and $q = 0.707$. The halo bias parameters b_1^h and b_2^h are given by [65,67]

$$b_1^h(M, z|z_f) = 1 + \epsilon_1 + E_1, b_2^h(M, z|z_f)$$
$$= \frac{8}{21}(\epsilon_1 + E_1) + \epsilon_2 + E_2, \qquad (37)$$

where z refers to the redshift of observation, while z_f refers to the redshift of formation of halos of mass M, and

$$\epsilon_1 = \frac{q\nu^2 - 1}{\delta_f}, \qquad \epsilon_2 = \frac{q\nu^2}{\delta_f} \frac{q\nu^2 - 3}{\delta_f}, \qquad (38)$$

$$E_1 = \frac{2p/\delta_f}{1 + (q\nu^2)^p}, \qquad \frac{E_2}{E_1} = \frac{1 + 2p}{\delta_f} + 2\epsilon_1, \qquad (39)$$

with $\delta_f = \delta_c D(z)/D(z_f)$, where $D(z)$ is the linear growth function. In the left panel of Fig. 3 we show the halo bias functions $b_1(M, z)$ and $b_2(M, z)$, as a function of M and z, in the approximation that the formation redshift equals the observation redshift, $z = z_f$. We shall always assume this throughout the paper.

As for the HOD, we adopt the form proposed by Tinker *et al.* [68]:

$$\langle N \rangle_M = 1 + \frac{M}{M_1} \exp\left(-\frac{M_{cut}}{M}\right) \qquad (40)$$

for $M > M_{min}$ and zero otherwise. The parameter M_{min} represents the minimum mass above which we find a (central) galaxy in the halo, while M_1 represents the mass above which we can find a second (satellite) galaxy. Measuring the HOD parameters for subhalo populations from several N-body simulations at different redshifts and densities, Conroy *et al.* [69] found a correlation between M_{cut} and M_1 given by
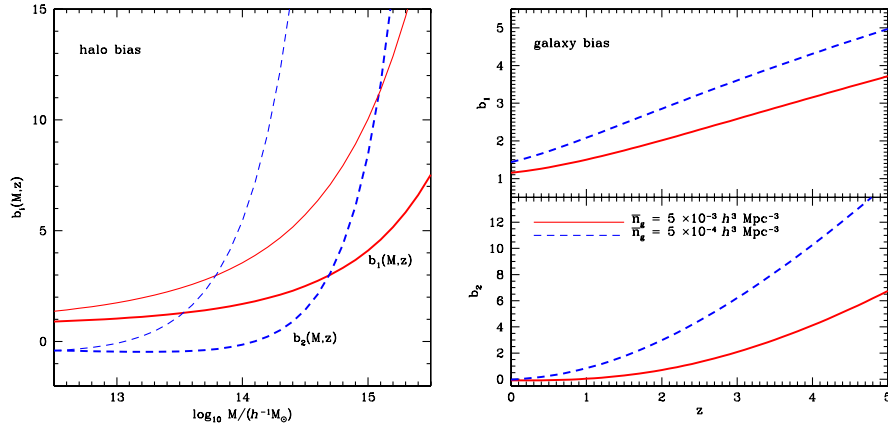
FIG. 3 (color online).   Left panel: The halo bias functions $b_1^h(M, z)$ (solid lines) and $b_2^h(M, z)$ (dashed lines), as a function of the mass M, for $z = 0$ (thick lines) and $z = 1$ (thin lines) in the approximation that the formation redshift equals the observation redshift, $z = z_f$. Right panel: The galaxy bias parameters b_1 and b_2, for the mean galaxy density of $n_g = 5 \times 10^{-3}$ (continuous lines) and $n_g = 5 \times 10^{-4} \ h^3 \, \text{Mpc}^{-3}$ (dashed lines).

$$\log_{10} M_{\text{cut}} = 0.76 \log_{10} M_1 + 2.3. \qquad (41)$$

One also finds from Table 2 in Conroy *et al.* [69] that M_1/M_{min} depends on redshift and density only weakly; thus, for simplicity we shall keep this ratio fixed at $\log_{10}(M_1/M_{\text{min}}) = 1.1$, and find M_{min} from

$$n_g = \int_{M_{\text{min}}} dM \, n_h(M, z) \langle N \rangle_M, \qquad (42)$$

for a given n_g.

In the right panel of Fig. 3 we show the galaxy bias parameters b_1 and b_2 from Eq. (34) as a function of redshift for two values of the mean galaxy density, $n_g = 5 \times 10^{-3}$ and $5 \times 10^{-4} \ h^3 \, \text{Mpc}^{-3}$. As expected, for a fixed galaxy number density the value of the linear bias, b_1, increases with redshift. We find that b_1 and b_2 are strongly correlated. We shall come back to this point in Sec. V.

We admit that these values are derived from very simplified models without much justification. We need observational data to determine the true bias parameters for high-z surveys eventually, although we do not have sufficient data for doing so yet. Nevertheless, we find our approach useful for our purpose of deriving the fiducial values of b_1 and b_2 with a realistic redshift evolution, particularly for redshift surveys spanning a wide range in redshift, for which one has to consider a set of redshift bins and assume different fiducial values for b_1 and b_2 at different z. We note that b_1 obtained from our method agrees with those obtained in the previous work by assuming $\sigma_{8,g} \simeq 1$ and $n_g \simeq 5 \times 10^{-4} \ h^3 \, \text{Mpc}^{-3}$ [70,71].

In Sec. V we shall show how one can extend this simple picture by introducing a redshift dependence in the HOD, and how one can make use of the information on galaxy bias derived from the bispectrum to constrain the HOD parameters directly.

IV. RESULTS

In this section we present the results from our Fisher matrix analysis of the galaxy bispectrum. We first study how the derived constraints on the galaxy bias parameters and primordial non-Gaussianity depend on the choice of k_{max}, taking into account the two approaches discussed above. We then study how the constraints depend on the survey volume and redshift. Finally we shall apply our method to make forecasts for several current and proposed redshift surveys.

A. Dependence on k_{max}, volume, redshift, and number density

1. k_{max}

As mentioned in the previous section, constraints on galaxy bias and primordial non-Gaussianity would depend strongly on k_{max}, the smallest scale included in the analysis.

As an example, we consider sample surveys at two redshifts: the median redshifts of (i) $\bar{z} = 1$ and (ii) $\bar{z} = 3$. Each has the volume of $V = 10 \ h^{-3} \, \text{Gpc}^3$ and the number density of $n_g = 5 \times 10^{-3} \ h^3 \, \text{Mpc}^{-3}$. (The total number of galaxies in the survey volume at each redshift is 50×10^6 galaxies.) The bias parameters are (i) $b_1 = 1.5$ and $b_2 = 0.035$ at $\bar{z} = 1$, and (ii) $b_1 = 2.6$ and $b_2 = 2.1$ at $\bar{z} = 3$.

In the upper panels of Fig. 4 we plot the marginalized, 1-σ, fractional errors on b_1 and b_2 at $\bar{z} = 1$ and 3, assuming Gaussian initial conditions, i.e., $f_{NL} = 0$. We observe an interesting effect: a fractional error on b_1 improves at lower z, while that on b_2 improves at higher z. (Note that this statement is true only when *the same k_{max} is used at both redshifts*. See discussion below.) This can be understood as follows. Let us recall the form of the galaxy reduced bispectrum [Eq. (19)]:
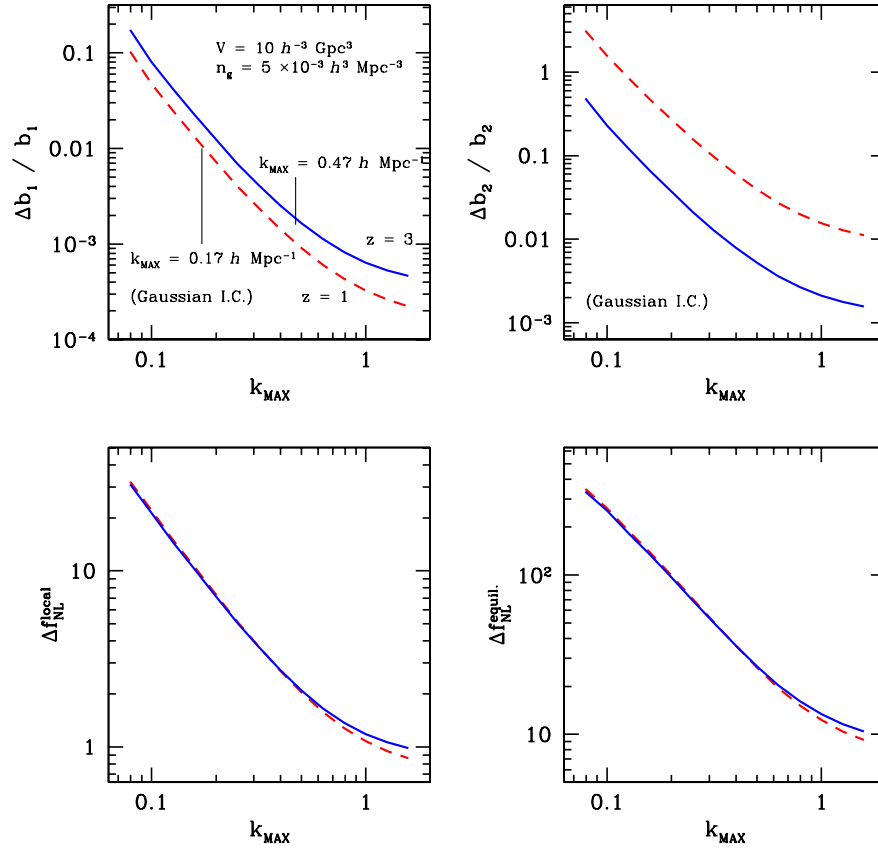
FIG. 4 (color online).   Upper panels: Predicted errors on galaxy bias parameters vs the maximum wave number k_{max}. The dashed and solid lines show the prediction for a galaxy survey at $z = 1$ and 3, respectively. Each survey is assumed to have the survey volume of $V = 10\ h^{-3}\,\mathrm{Gpc}^3$ and the number density of $n_g = 5 \times 10^{-3}\ h^3\,\mathrm{Mpc}^{-3}$. The left panel shows the marginalized 1-σ errors on the linear bias b_1, while the right panel shows the nonlinear bias b_2. Both assume Gaussian initial conditions, $f_{NL} = 0$. The vertical lines show k_{max} as determined from $\sigma(R; z) = 0.5$ for each redshift (see Sec. III B). Lower panels: Predicted errors on primordial non-Gaussian parameters vs k_{max}. The left panel shows the marginalized 1-σ errors on the local model $f_{NL}^{\mathrm{loc.}}$, while the right panel shows the equilateral model $f_{NL}^{\mathrm{eq.}}$. The bias parameters have been marginalized.

$$Q_g(k_1, k_2, k_3) \simeq \frac{1}{b_1} Q(k_1, k_2, k_3) + \frac{b_2}{b_1^2}.$$

Now, Q on the right-hand side is independent of z at the tree level when initial fluctuations are Gaussian. Therefore, the first term falls as $1/b_1$ at higher z where b_1 is larger (Fig. 3). On the other hand, the second term actually grows as z: for the current example $b_2/b_1^2 = 0.016$ at $z = 1$ and 0.31 at $z = 3$. Therefore, our sensitivity to b_2 grows with z, while our sensitivity to b_1 declines with z.

Let us study more quantitatively the sensitivity to b_1. In the limit of linear bias, $b_2 = 0$, a signal-to-noise ratio of the reduced bispectrum of equilateral configurations is given by

$$\left.\frac{Q_s^2(k)}{\Delta Q_s^2(k)}\right|_{b_2=0} \simeq \frac{[a_0^B(\beta)]^2}{[a_0^P(\beta)]^3}\frac{V_B}{k_f s_{123}}\frac{B_G^2(k, k, k; z)}{P_L^3(k; z)} \propto D^2(z).$$

(43)

We expect, therefore, that a signal-to-noise ratio of the

bispectrum from gravitational instability declines with z, resulting in an increasing error on b_1 at higher z.

In practice, however, we predict that galaxy surveys at higher z should result in better determinations of both b_1 and b_2. The reason is quite simple: k_{max} at higher z must be larger than that at lower z. In the upper left panel of Fig. 4 we show k_{max} as determined from $\sigma(R; z) = 0.5$: $k_{max} = 0.17\ h\,\mathrm{Mpc}^{-1}$ at $z = 1$ and $k_{max} = 0.47\ h\,\mathrm{Mpc}^{-1}$ at $z = 3$. The difference is clear: when the modes up to k_{max} are included, a survey at $z = 3$ yields an error on b_1 that is a factor of 5 better than that at $z = 1$. As for b_2, a survey at $z = 3$ does better by nearly 2 orders of magnitude.

How about primordial non-Gaussianity? In the lower panels of Fig. 4 we show the predicted errors on $f_{NL}^{\mathrm{loc.}}$ and $f_{NL}^{\mathrm{eq.}}$, marginalized over b_1 and b_2. We find that the difference between $z = 1$ and 3 is negligible at the same k_{max}. This is a consequence of the fact that a signal-to-noise ratio for the primordial bispectrum component is not, in the first approximation, redshift dependent. For equilateral configurations one finds

$$\frac{Q_s^2(k)}{\Delta Q_s^2(k)}\bigg|_I \simeq \frac{[a_0^B(\beta)]^2}{[a_0^P(\beta)]^3} \frac{V_B}{k_f s_{123}} \frac{B_I^2(k,k,k;z)}{P_L^3(k;z)} \propto \text{constant}$$

(44)

where we considered only the primordial term Q_I. Nevertheless, we still predict that galaxy surveys at higher z should result in better determinations of both $f_{NL}^{\text{loc.}}$ and $f_{NL}^{\text{eq.}}$, as k_{max} must be larger at higher z and therefore many more modes are available for the analysis at higher z.

Results from a much more conservative estimate of k_{max} [Eq. (33)] will be given near the end of the next section.

2. Volume, redshift, and number density of galaxies

Dependence of the predicted errors on volume is straightforward: it depends simply on $1/\sqrt{V}$. Dependence on z is a combination of two effects: (i) how a signal-to-noise ratio for a given k_{max} grows with z, and (ii) how k_{max} grows with z. Finally, the number density of galaxies determines a signal-to-noise ratio on small scales, where the shot noise plays an important role. In particular, very

high-z surveys at, e.g., $z \gtrsim 3$, do not add very much if the number density of galaxies is too low.

In Fig. 5 we show how the predicted constraints on the galaxy bias parameters b_1 and b_2 improve with the survey volume, as a function of the median redshift, $\bar{z}$. We used $n_g = 5 \times 10^{-3} \, h^3 \, \text{Mpc}^{-3}$ for the number density of galaxies. The fiducial values of b_1 and b_2 are calculated for each $\bar{z}$ from Fig. 3. We used k_{max} determined from $\sigma(R, \bar{z}) = 0.5$ for a given $\bar{z}$. Since k_{max} grows as $\bar{z}$ increases, the predicted constraints on b_1 and b_2 also improve as $\bar{z}$ increases. A spike at $z \sim 0.8$ in $\Delta b_2/b_2$ is a numerical artifact of b_2 being very close to zero. The dependence on volume is given simply by $1/\sqrt{V}$.

In the lower panels of Fig. 5 we show the predicted constraints on primordial non-Gaussianity, marginalized over b_1 and b_2. We find that a survey of the size $V \sim 1 \, h^{-3} \, \text{Gpc}^3$ at $z \sim 4$–6 or $V \sim 10 \, h^{-3} \, \text{Gpc}^3$ at $z \sim 1$–2 is as sensitive to $f_{NL}^{\text{loc.}}$ as the CMB data from Planck. A more ambitious design, e.g., $V \sim 10 \, h^{-3} \, \text{Gpc}^3$ at $z \gtrsim 2$, can achieve $\Delta f_{NL}^{\text{loc.}} \sim 1$, although it depends on the number density quite strongly. An even more ambitious design, $V \sim 100 \, h^{-3} \, \text{Gpc}^3$, would enable us to detect the primor-
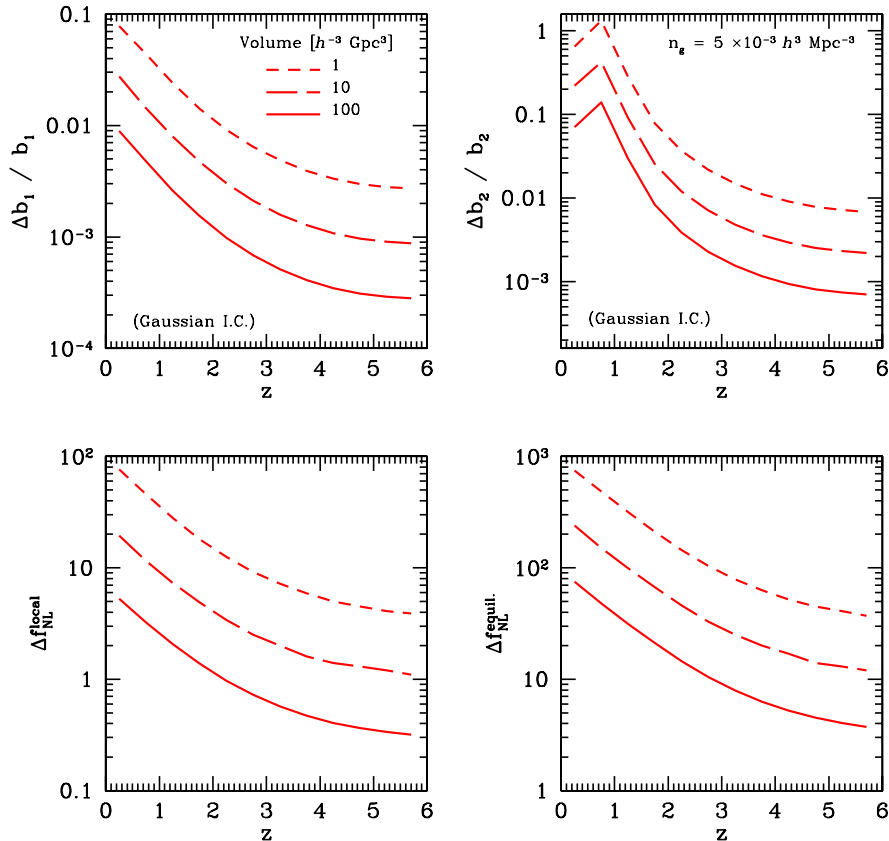

FIG. 5 (color online). Predicted 1-σ errors on galaxy bias and primordial non-Gaussianity vs the survey volume V and redshift z. The short-dashed, long-dashed, and solid lines show $V = 1$, 10, and 100 $h^{-3} \, \text{Gpc}^3$, respectively, with the galaxy number density of $n_g = 5 \times 10^{-3} \, h^3 \, \text{Mpc}^{-3}$. Upper panels: Fractional errors on the linear bias b_1 (left panel) and nonlinear bias b_2 (right panel), for Gaussian initial conditions, $f_{NL} = 0$. Lower panels: Errors on primordial non-Gaussian parameters $f_{NL}^{\text{loc.}}$ (left panel) and $f_{NL}^{\text{eq.}}$ (right panel), marginalized over b_1 and b_2.

dial bispectrum from the ubiquitous non-Gaussian "floor" from the second-order evolution of primordial fluctuations [47].

Constraints on the equilateral type of non-Gaussianity suffer from a stronger degeneracy between the primordial bispectrum and the nonlinear gravitational evolution as well as nonlinear bias, and thus the predicted errors on $f_{NL}^{eq.}$ are an order of magnitude larger than those on $f_{NL}^{loc.}$. (See Sec. IV B for more details.) Nevertheless, a survey of $V \sim 10~h^{-3}\,\text{Gpc}^3$ at $z \sim 1$ should provide a constraint that is comparable to that from the WMAP 3-yr data.

We also computed the Fisher matrix for an all-sky survey from $z = 0$ to 5. We divided the entire redshift range in bins of the size $\Delta z = 0.5$, and used $n_g = 5 \times 10^{-3}~h^3\,\text{Mpc}^{-3}$. We find that such a survey should provide $\Delta f_{NL}^{loc.} \sim 0.2$ and $\Delta f_{NL}^{eq.} \sim 2$. These values probably represent the best limits on f_{NL} one can ever hope to achieve from galaxy surveys.

How about the number density of galaxies? When the number density is low, the shot noise completely dominates at small scales, and thus one fails to improve a signal-to-noise ratio by increasing k_{max}. This suggests that very high-z galaxy surveys do not add much if the number density of galaxies is too low. In Fig. 6 we show the case for $n_g = 5 \times 10^{-4}~h^3\,\text{Mpc}^{-3}$. Clearly, our sensitivity to all of b_1, b_2, $f_{NL}^{loc.}$, and $f_{NL}^{eq.}$ does not improve at all beyond

$z \sim 3$. Therefore, it makes sense to conduct very high-z surveys, only if one can detect more than $n_g \sim 10^{-3}~h^3\,\text{Mpc}^{-3}$.

How robust are these results? The most uncertain parameter in our analysis is k_{max}. What if k_{max} is significantly lower than that from $\sigma(R, z) = 0.5$? To address this question, we have repeated our analysis using a much more conservative estimate of k_{max} given by Eq. (33). This estimate was derived by throwing away any information beyond k_{max} at which the tree-level bispectrum becomes inaccurate. This is a conservative estimate because we can certainly improve our theoretical prediction by going to the higher-order, "1-loop" (fourth-order) calculations [55,63]. We show the results in Fig. 7. We find significantly weaker constraints; for example, fractional errors on b_1 from a survey of $V = 10~h^{-3}\,\text{Gpc}^3$ now go from 9% to a few percent from $z = 0$ to $z = 5$, or a factor of ~ 8 and 20 weaker constraints at $z \sim 1$ and 5, respectively. For this choice of k_{max} the shot noise is not very important because we are restricted to a fairly small k already. Therefore we obtain similar results for a lower density, $n_g = 5 \times 10^{-4}~h^3\,\text{Mpc}^{-3}$.

B. Parameter degeneracy

Are galaxy bias and primordial non-Gaussianity independent? In Fig. 8 we show the 2-d joint constraints (95%
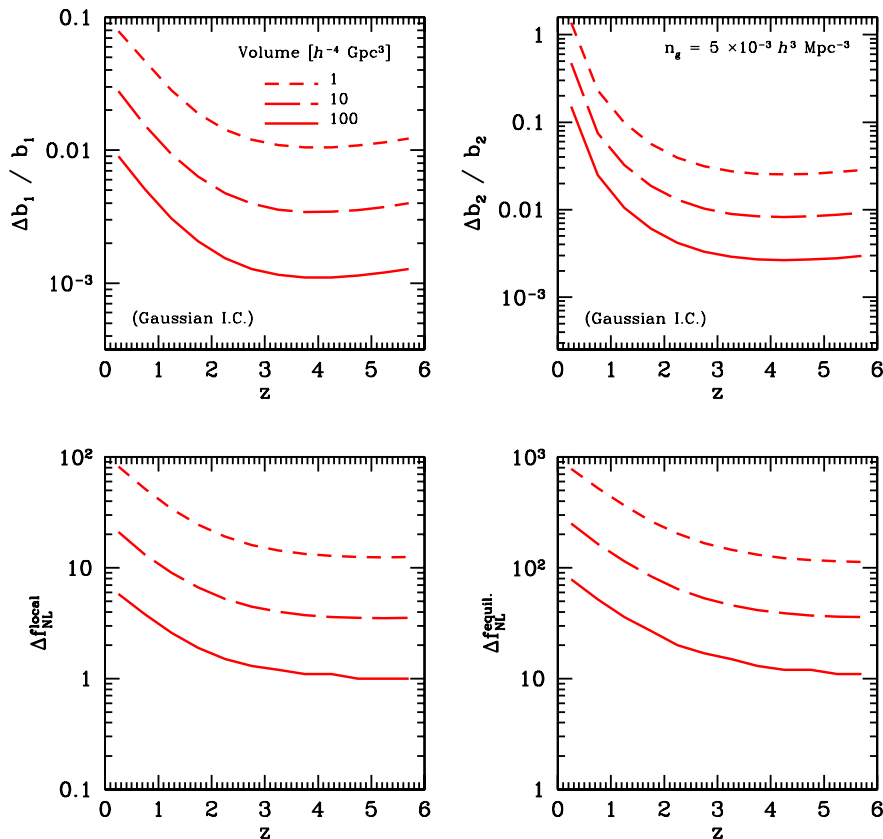
FIG. 6 (color online).   Same as Fig. 5 but for a smaller number density of galaxies, $n_g = 5 \times 10^{-4}~h^3\,\text{Mpc}^{-3}$.
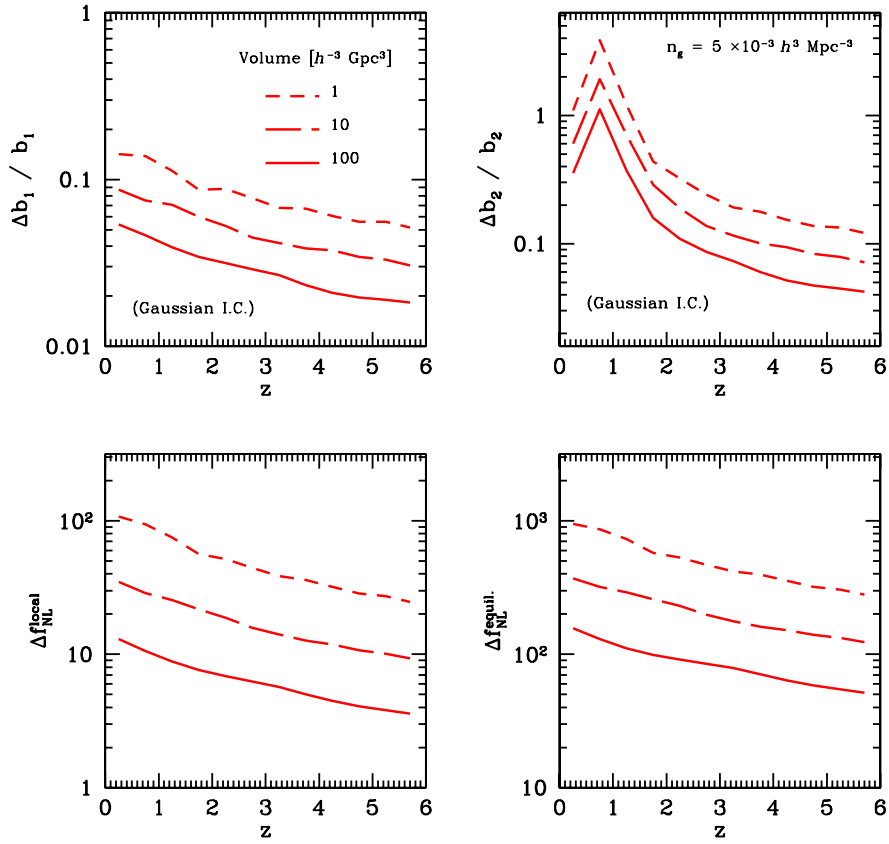
FIG. 7 (color online). Same as Fig. 5 but for a much more conservative estimate of k_{max} given by Eq. (33).
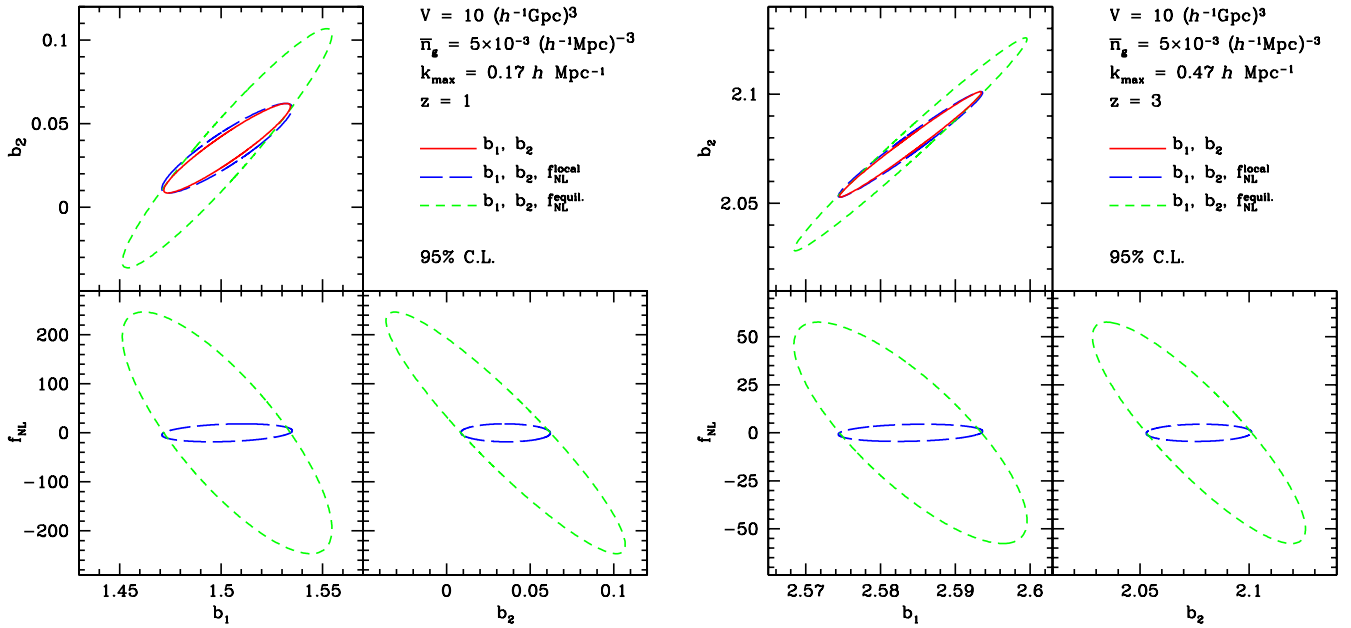
FIG. 8 (color online). Two-dimensional joint 95% C.L. constraints on galaxy bias and primordial non-Gaussianity from $z = 1$ (left figure) and $z = 3$ (right figure). The top left, bottom left, and bottom right panels show the joint constraints on (b_1, b_2), (b_1, f_{NL}), and (b_2, f_{NL}), marginalized over f_{NL}, b_2, and b_1, respectively. The assumed survey volume is $V = 10\ h^{-3}$ Gpc^3, while the number density is $n_g = 5 \times 10^{-3}\ h^3$ Mpc^{-3}.

C.L.) on (b_1, b_2), (b_1, f_{NL}), and (b_2, f_{NL}), marginalized over f_{NL}, b_2, and b_1, respectively. The survey parameters are $V = 10 \, h^{-3} \, \mathrm{Gpc}^3$, $n_g = 5 \times 10^{-3} \, h^3 \, \mathrm{Mpc}^{-3}$, $z = 1$ (left panels) and 3 (right panels). The fiducial values of bias parameters are $b_1 = 1.5$ and $b_2 = 0.035$ at $z = 1$, and $b_1 = 2.6$ and $b_2 = 2.1$ at $z = 3$.

We find that $f_{NL}^{\mathrm{loc.}}$ is not degenerate with b_1 or b_2, which is very good news; however, $f_{NL}^{\mathrm{eq.}}$ reveals a rather strong degeneracy with both b_1 and b_2. This can easily be understood by observing that, as shown in Fig. 2, the contribution due to equilateral non-Gaussianity can, for certain values of $f_{NL}^{\mathrm{eq.}}$, reduce the peculiar dependence on the shape of the triangular configuration proper to the bispectrum from gravitational instability. Since such dependence is crucial in breaking the degeneracy between linear and quadratic bias, this second corresponding to an overall constant in the galaxy reduced bispectrum, it is natural to expect weaker constraints on *all* the parameters considered. Therefore, the equilateral model turns out to be much harder to constrain by galaxy surveys than the local model. A way to partially reduce this degeneracy will be discussed in Sec. V.

C. Current and proposed redshift surveys

We are now in a position to apply our Fisher matrix analysis tools to several current and future redshift galaxy surveys, both at low and high redshifts. For the sake of simplicity, and to allow for easier comparison, we shall assume that *each survey is characterized uniquely by its survey volume V, redshift range, and galaxy number density n_g.* In other words, we shall ignore complications related to the specific geometry and selection functions.

A galaxy survey of a large volume and a relatively low galaxy density [but not too low—there should be at least $n_g \sim 1/P(k_{\mathrm{max}})$ galaxies for the optimal survey efficiency] is necessary to detect the baryon acoustic oscillations (BAO) in the power spectrum, which may be used to constrain the nature of dark energy. We point out that such a survey should also provide competitive constraints on primordial non-Gaussianity.

We shall consider the following survey designs. Table I tabulates V, n_g, z, k_{max}, b_1, and b_2 of these surveys. We calculate k_{max} from $\sigma(R, z) = 0.5$ and $k_{\mathrm{max}} = \pi/(2R)$ (Sec. III), and b_1 and b_2 from a halo approach given in Sec. III D.

(i) Low-z surveys include two ongoing surveys and one planned survey:
  (a) The SDSS main sample (for which a detailed analysis of the expected constraints on galaxy bias and primordial non-Gaussianity has been given in Scoccimarro *et al.* [23]) at $\bar{z} = 0$ and $V = 0.3 \, h^{-3} \, \mathrm{Gpc}^3$.
  (b) The SDSS Luminous Red Galaxy (LRG) sample at $\bar{z} = 0.35$ and $V = 0.72 \, h^{-3} \, \mathrm{Gpc}^3$.
  (c) A proposed extension of SDSS, Apache Point Observatory Large-Scale Structure (APO-LSS) survey at $\bar{z} = 0.35$ and $V = 3.8 \, h^{-3} \, \mathrm{Gpc}^3$.

TABLE I. Survey parameters (V in units of $h^{-3} \, \mathrm{Gpc}^3$, n_g in units of $h^3 \, \mathrm{Mpc}^{-3}$, z), maximum wave number assumed in the analysis (k_{max} in units of $h \, \mathrm{Mpc}^{-1}$), fiducial values of galaxy bias (b_1 and b_2), and marginalized 1-σ constraints from the Fisher matrix analysis of the reduced bispectrum in redshift space for $\sigma_8 = 0.9$ and $n_s = 1$.

	V	n_g	z	k_{max}	b_1	b_2	Δb_1	Δb_2	Δb_1	Δb_2	$\Delta f_{NL}^{\mathrm{loc.}}$	Δb_1	Δb_2	$\Delta f_{NL}^{\mathrm{eq.}}$
SDSS	0.3	30		0.09	1.19	−0.10	0.270	0.151	0.309	0.151	255.5	0.450	0.421	1775
LRG	0.72	1	0.35	0.11	2.14	0.96	0.209	0.348	0.223	0.353	113.4	0.338	0.726	998
APO-LSS	3.8	4	0.35	0.11	1.69	0.21	0.069	0.068	0.071	0.069	34.9	0.108	0.160	386
WFMOS1	1.6	5	0.7	0.14	1.87	0.45	0.076	0.096	0.080	0.096	41.0	0.123	0.216	435
	2.4	5	1.1	0.18	2.16	1.00	0.047	0.081	0.048	0.081	23.1	0.076	0.175	266
		combined					· · ·	· · ·	· · ·	· · ·	20.1	· · ·	· · ·	227
ADEPT	45	1	1.25	0.20	2.97	3.44	0.020	0.063	0.021	0.063	6.1	0.031	0.111	73
	55	1	1.75	0.26	3.44	5.43	0.017	0.066	0.017	0.067	4.5	0.025	0.112	53
		combined					· · ·	· · ·	· · ·	· · ·	3.6	· · ·	· · ·	43
WFMOS2	0.5	5	2.55	0.38	3.27	4.64	0.058	0.220	0.060	0.223	25.7	0.094	0.406	256
	0.5	5	3.05	0.48	3.64	6.39	0.056	0.253	0.058	0.255	22.1	0.087	0.439	215
		combined					· · ·	· · ·	· · ·	· · ·	16.8		· · ·	164
HETDEX	0.68	5	2.25	0.34	3.05	3.70	0.051	0.172	0.053	0.174	23.6	0.083	0.326	244
	0.69	5	2.75	0.42	3.42	5.32	0.049	0.199	0.050	0.201	20.0	0.077	0.357	202
	0.67	5	3.25	0.53	3.79	7.16	0.050	0.237	0.051	0.238	18.0	0.076	0.401	177
	0.64	5	3.75	0.65	4.14	9.20	0.053	0.291	0.054	0.292	17.1	0.079	0.469	163
		combined			· · ·	· · ·	· · ·	· · ·	· · ·	· · ·	9.6	· · ·	· · ·	95
CIP	1.26	50	4	0.71	3.16	4.12	0.010	0.036	0.010	0.037	4.7	0.016	0.066	51
	1.13	50	5	1.03	3.72	6.76	0.010	0.047	0.010	0.048	4.0	0.015	0.079	40
	1.02	50	6	1.46	4.26	9.90	0.011	0.066	0.012	0.066	3.8	0.016	0.102	36
		combined					· · ·	· · ·	· · ·	· · ·	2.4	· · ·	· · ·	24

(ii) Intermediate z surveys include three planned surveys:

(a) The Hobby-Eberly Dark Energy Experiment (HETDEX) [72], targeting Lyman-α emitters (LAE) between $z = 2$ and $z = 4$. We assume $V = 2.7 \ h^{-3} \, \text{Gpc}^3$ and $n_g = 5 \times 10^{-4} \ h^3 \, \text{Mpc}^{-3}$, following the "G2" design given in Takada *et al.* [71] for comparison.

(b) Two redshift surveys with the planned Wide-Field Multi-Object Spectrograph (WFMOS), one detecting 2×10^6 galaxies at $0.5 < z < 1.3$ in $2000 \, \text{deg}^2$ (WFMOS1; $V = 4 \ h^{-3} \, \text{Gpc}^3$), and the other detecting $600\,000$ galaxies at $2.3 < z < 3.3$ in $300 \, \text{deg}^2$ (WFMOS2; $V = 1 \ h^{-3} \, \text{Gpc}^3$) [73]. To facilitate comparison, we assume the same number density of galaxies for both, $n_g = 5 \times 10^{-4} \ h^3 \, \text{Mpc}^{-3}$.

(c) The Advanced Dark Energy Physics Telescope (ADEPT) mission, a space-based redshift survey of 100×10^6 galaxies at $1 < z < 2$ in $28\,600 \, \text{deg}^2$. We assume $V = 100 \ h^{-3} \, \text{Gpc}^3$ and $n_g = 10^{-4} \ h^3 \, \text{Mpc}^{-3}$.

(iii) High-z surveys are represented by the Cosmic Inflation Probe (CIP) mission, a space-based redshift survey targeting Hα emitters at $3.5 < z < 6.5$ [74]. We assume $V = 3.4 \ h^{-3} \, \text{Gpc}^3$ and $n_g = 5 \times 10^{-3} \ h^3 \, \text{Mpc}^{-3}$, following the "SG" design given in Takada *et al.* [71].

The 8th and 9th columns in Table I tabulate Δb_1 and Δb_2 for Gaussian initial conditions ($f_{NL} = 0$); the 10th, 11th, and 12th columns tabulate Δb_1, Δb_2, and $\Delta f_{NL}^{\text{loc.}}$; and the 13th, 14th, and 15th columns tabulate Δb_1, Δb_2, and $\Delta f_{NL}^{\text{eq.}}$. For HETDEX, WFMOS, ADEPT, and CIP we provide Δb_1, Δb_2, $\Delta f_{NL}^{\text{loc.}}$, and $\Delta f_{NL}^{\text{eq.}}$ from each redshift bin as well as $\Delta f_{NL}^{\text{loc.}}$ and $\Delta f_{NL}^{\text{eq.}}$ from a combined analysis of all bins.

The predicted constraints on galaxy bias range from a few percent for APO-LSS, WFMOS, and HETDEX to less than 1% for ADEPT, the latter being significantly better owing obviously to a larger survey volume. As we have mentioned already in Sec. IV B, inclusion of $f_{NL}^{\text{loc.}}$ does not degrade sensitivity to galaxy bias, whereas $f_{NL}^{\text{eq.}}$ does degrade it significantly.

Our prediction for SDSS should not be compared directly to those derived by Scoccimarro *et al.* [23], as we have ignored covariance between different bispectrum configurations due to nonlinear effects and survey geometry (Sec. III B), which was included in their work. In this sense our analysis is more optimistic; however, in the other sense, our analysis is in fact more realistic than theirs. We have used $k_{\text{max}} \simeq 0.09 \ h \, \text{Mpc}^{-1}$ for SDSS, which is significantly more conservative than their value, $k_{\text{max}} = 0.3 \ h \, \text{Mpc}^{-1}$. This has made our predicted errors much weaker than theirs: they obtained $\Delta b_1/b_1 = 0.04$ (ours is 0.26) and $\Delta f_{NL}^{\text{loc.}} = 145$ (ours is 256).

How would these results depend on σ_8 and n_s? A lower value for the rms density fluctuations, $\sigma_8 = 0.75$, as recently suggested by the WMAP 3yr data [29], increases k_{max} because nonlinearity is weaker for a lower σ_8, and thus improves the constraints. On the other hand, it follows from Eq. (43) that the signal-to-noise ratio for a given triangle configuration is proportional to σ_8^2, and thus a lower σ_8 results in worse constraints. When combined, these two effects result in 20% and 30% stronger constraints on galaxy bias and primordial non-Gaussianity, respectively, for low-redshift surveys, whereas these effects cancel for intermediate and high-z surveys. A departure from a scale-invariant spectrum has a smaller effect. For $n_s = 0.95$ we find only 10% improvement in bias and non-Gaussianity.

V. CONSTRAINING THE HOD

So far we have assumed that the linear and quadratic bias, b_1 and b_2, are completely free. For surveys spanning a large redshift range, therefore, we had to introduce multiple redshift bins, and, as a result, we had to have an excessive number of free parameters, 2 times the number of redshift bins.

In this section we attempt to reduce the number of free parameters by using a halo approach. The fiducial values of b_1 and b_2 were derived from a given form of the HOD (Sec. III D). If we can parametrize the HOD by fewer parameters than 2 times the number of redshift bins, the model has more constraining power.

In the limit that the evolution of bias is given by the mass function, we may approximate that the HOD is independent of z. We make a minimal extension of the HOD that we used in Sec. III D; namely, instead of fixing a ratio of M_1 and M_{min}, we assume

$$\log_{10} M_1 = a + \log_{10} M_{\text{min}}. \tag{45}$$

The fiducial value is $a = 1.1$, as before. We still assume that a relation between M_{cut} and M_1 is given by Eq. (41). We then replace b_1 and b_2 at different redshift bins with a single parameter a in the Fisher matrix analysis.

The expected 1-σ errors on a and primordial non-Gaussian parameters are given in Table II in the 3rd to 5th columns. We find that the error on $f_{NL}^{\text{loc.}}$ is unaffected: this is an expected result because $f_{NL}^{\text{loc.}}$ is not degenerate with galaxy bias. On the other hand, we find a significant, about a factor of 2, improvement in $f_{NL}^{\text{eq.}}$. This is due to the fact that the analysis in terms of the HOD is equivalent to introducing a theoretical prior on a relation between b_1 and b_2, lifting the degeneracy. In other words, while in the previous section we allowed b_1 and b_2 to vary independently, we are now making use of the fact that the halo model predicts them to be strongly correlated.

TABLE II. Marginalized 1-σ errors on the HOD parameters and primordial non-Gaussianity from the Fisher matrix analysis of the reduced bispectrum in redshift space for $\sigma_8 = 0.9$ and $n_s = 1$. The survey parameters and k_{max} are the same as in Table I.

Survey		1-parameter HOD							2-parameter HOD					
	z	Δa	Δa	$\Delta f_{NL}^{loc.}$	Δa	$\Delta f_{NL}^{eq.}$	Δa	Δb	Δa	Δb	$\Delta f_{NL}^{loc.}$	Δa	Δb	$\Delta f_{NL}^{eq.}$
SDSS	0	0.20	0.26	198	0.20	500	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
LRG	0	0.10	0.15	112	0.13	363	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
APO-LSS	0.25	0.064	0.077	35.1	0.066	120	⋯	⋯	⋯	⋯	⋯	⋯	⋯	⋯
WFMOS1	0.7	0.050	0.066	39.0	0.057	134	540	780	570	820	41	1800	2500	435
	1.1	0.024	0.031	22.7	0.029	91	1050	960	1070	970	23	3100	2800	266
	comb.	0.024	0.028	19.6	0.026	75	0.14	0.14	0.15	0.14	20	0.15	0.14	75
ADEPT	0.7	0.0050	0.0062	6.1	0.0078	35	56	45	56	45	6.1	116	93	73
	1.1	0.0033	0.0040	4.5	0.0050	28	1200	710	1200	710	4.5	2400	1300	53
	comb.	0.0033	0.0034	3.6	0.0042	22	0.019	0.012	0.020	0.012	3.6	0.020	0.012	22
WFMOS2	0.7	0.0147	0.020	26	0.021	120	83	33	83	33	26	177	70	256
	1.1	0.0120	0.016	22	0.017	110	789	259	789	259	22	1539	505	215
	comb.	0.0120	0.013	17	0.013	81	0.11	0.038	0.11	0.038	17	0.11	0.038	81
HETDEX	2.25	0.015	0.020	24.1	0.021	111	3500	1600	3500	1600	24.1	7900	3500	250
	2.75	0.011	0.015	19.9	0.016	99	1400	500	1400	500	19.9	2800	1000	201
	3.25	0.010	0.013	18.0	0.014	93	3200	990	3200	990	18.0	6100	1900	177
	3.75	0.010	0.013	17.2	0.013	92	730	200	730	200	17.2	1300	350	164
	comb.	0.005	0.007	9.7	0.008	49	0.034	0.11	0.034	0.011	9.7	0.035	0.011	49
CIP	4	0.0032	0.0041	4.7	0.0042	23	11.1	2.8	11.1	2.8	4.7	24	6.0	51
	5	0.0026	0.0034	4.0	0.0034	21	6.9	1.4	6.9	1.4	4.0	13	2.6	40
	6	0.0025	0.0033	3.7	0.0034	21	13.1	2.2	13.1	2.2	3.8	22	3.7	36
	comb.	0.0016	0.0021	2.3	0.0021	13	0.010	0.0020	0.011	0.0020	2.4	0.011	0.0020	13

When we have a survey that covers a wide range in z, we may be able to constrain more than one parameter in the HOD. To see how it works we extend the minimal model by introducing one more parameter:

$$\log_{10} M_1 = a + bz + \log_{10} M_{min}, \qquad (46)$$

and we assume $b = 0$ as the fiducial value. An analysis based on a single redshift bin would naturally lead to a large degeneracy between a and b. However, one can lift this degeneracy by including multiple redshift bins.

In Table II, columns 6 to 13, we present the expected 1-σ errors on a and b as well as on primordial non-Gaussianity. One can clearly see that a and b are degenerate within a single redshift bin: combined errors are orders of magnitude smaller than those from single bins. The determination of primordial non-Gaussianity is largely unaffected by the extra HOD parameter.

VI. CONCLUSIONS

The quest to understand the nature of dark energy has recently provided a further motivation for future large redshift surveys. It is certain that the study of higher-order correlation functions of galaxies will be required in order to extract maximum cosmological information from such large data sets. For instance, as it has been recognized for more than a decade, the bispectrum can be used to measure nonlinearities in the galaxy-mass relation. Nonlinearity in

galaxy bias must be understood better to meet the high accuracy required for precision measurements of the baryon acoustic oscillations and their interpretation as a standard ruler, particularly for highly biased tracers such as the luminous red galaxies and Lyman-α emitters.

We have shown that, with the most conservative assumption about the maximum wave number k_{max}, used in the analysis, the bispectrum measured from a galaxy redshift survey should yield a fractional error on the linear bias of order $0.1\sqrt{h^{-3}\,\mathrm{Gpc}^3/V}$ at $z = 0$ to $0.05\sqrt{h^{-3}\,\mathrm{Gpc}^3/V}$ at $z = 6$. This is an *extremely* conservative limit, however, as it assumes no understanding of even the mildest nonlinearities in the dark matter and galaxy biasing evolution.

An analysis that includes all configurations down to mildly nonlinear scales, within which $\sigma(R) \lesssim 0.5$ is satisfied, should yield more than an order of magnitude better determination of bias parameters, both linear and nonlinear, owing to the large number of configurations available at smaller scales. This is precisely where intermediate to high-z galaxy surveys play a leading role: nonlinearity is much weaker at higher z, and therefore we can access a large number of modes on small scales.

While we can study nonlinearities in the matter distribution by means of new techniques based on perturbation theory and of N-body simulations, a simple, local description of galaxy bias may have to be improved further. In this perspective, our results show that a large amount of information can be extracted from higher-order correlations

which, in turn, may be used to constrain more sophisticated models of galaxy biasing.

We have also shown that the bispectrum from large-volume, high-redshift surveys is highly sensitive to primordial non-Gaussianity. The CMB observations have been the best probe of the Gaussian nature of primordial perturbations so far, and the Planck satellite would be quite close to the ideal experimental limit. On the other hand, a redshift survey of the large-scale structure actually contains much more information than CMB, as the number of modes available from the three-dimensional fluctuations is vastly larger than that from the two-dimensional temperature and polarization anisotropies.

Not only can they provide independent constraints on scales smaller than those probed by CMB, but also their constraints can be comparable to, if not better than, those from CMB. The best limit one can achieve from an all-sky survey up to redshift ~ 5 should reach $f_{NL}^{\rm loc} \sim 0.2$ and $f_{NL}^{\rm eq.} \sim 2$, an order of magnitude better than the best limits achievable by CMB. The planned surveys such as HETDEX and ADEPT should reach the constraints that are comparable to those from the WMAP and Planck CMB experiments, respectively. It should also be understood that galaxy surveys provide the best limits on small scales that are not accessible by CMB. This is particularly important when probing scale-dependent non-Gaussian models. We remark that our simple analysis does not take into account the full covariance properties of the bispectrum. However, as mentioned above, given our choice for $k_{\rm max}$, for a survey like the SDSS the effect of both covariance *and* selection

function might account for a degradation of the constraints up to a factor of 2. This would not significantly change our conclusions. Moreover, an improved model of galaxy bias might allow us to consider even smaller scales than those assumed in our analysis, with the larger signal-to-noise ratio partially compensating for the inclusion of the bispectrum covariance.

Finally, we have shown that galaxy bias parameters modeled by the halo occupation distribution should be very useful for lifting parameter degeneracies between nonlinear gravity, galaxy bias, and primordial non-Gaussianity.

We believe that, while a lot of work still has to be done in order to achieve a satisfactory description of the evolution of nonlinearities and galaxy bias at small scales, higher-order correlation functions will play a crucial role in the analysis of future redshift surveys, providing indispensable information on galaxy bias as well as on the nature of primordial perturbations.

ACKNOWLEDGMENTS

[1] S. Cole *et al.*, Mon. Not. R. Astron. Soc. **362**, 505 (2005).
[2] D. J. Eisenstein *et al.*, Astrophys. J. **633**, 560 (2005).
[3] G. Hütsi, Astron. Astrophys. **449**, 891 (2006).
[4] W. J. Percival *et al.*, Astrophys. J. **657**, 51 (2007).
[5] P. McDonald, Phys. Rev. D **74**, 103512 (2006) **74**, 129901 (2006).
[6] R. E. Smith, R. Scoccimarro, and R. K. Sheth, Phys. Rev. D **75**, 063512 (2007).
[7] R. E. Smith, R. Scoccimarro, and R. K. Sheth, arXiv:astro-ph/0703620.
[8] J. N. Fry, Phys. Rev. Lett. **73**, 215 (1994).
[9] S. Matarrese, L. Verde, and A. F. Heavens, Mon. Not. R. Astron. Soc. **290**, 651 (1997).
[10] R. Scoccimarro, H. A. Feldman, J. N. Fry, and J. A. Frieman, Astrophys. J. **546**, 652 (2001).
[11] E. Gaztañaga, P. Norberg, C. M. Baugh, and D. J. Croton, Mon. Not. R. Astron. Soc. **364**, 620 (2005).
[12] T. Nishimichi, I. Kayo, C. Hikage, K. Yahata, A. Taruya, Y. P. Jing, R. K. Sheth, and Y. Suto, Publ. Astron. Soc. Jpn. **59**, 93 (2007).
[13] J. Pan and I. Szapudi, Mon. Not. R. Astron. Soc. **362**, 1363 (2005).

[14] L. Verde *et al.*, Mon. Not. R. Astron. Soc. **335**, 432 (2002).
[15] G. V. Kulkarni, R. C. Nichol, R. K. Sheth, H.-J. Seo, D. J. Eisenstein, and A. Gray, Mon. Not. R. Astron. Soc. **378**, 1196 (2007).
[16] E. Sefusatti, M. Crocce, S. Pueblas, and R. Scoccimarro, Phys. Rev. D **74**, 023522 (2006).
[17] E. Sefusatti and R. Scoccimarro, Phys. Rev. D **71**, 063001 (2005).
[18] M. J. Chodorowski and F. R. Bouchet, Mon. Not. R. Astron. Soc. **279**, 557 (1996).
[19] R. Durrer, R. Juszkiewicz, M. Kunz, and J.-P. Uzan, Phys. Rev. D **62**, 021301 (2000).
[20] J. N. Fry and R. J. Scherrer, Astrophys. J. **429**, 36 (1994).
[21] C. Hikage, E. Komatsu, and T. Matsubara, Astrophys. J. **653**, 11 (2006).
[22] R. Scoccimarro, Astrophys. J. **542**, 1 (2000).
[23] R. Scoccimarro, E. Sefusatti, and M. Zaldarriaga, Phys. Rev. D **69**, 103513 (2004).
[24] L. Verde, L. Wang, A. F. Heavens, and M. Kamionkowski, Mon. Not. R. Astron. Soc. **313**, 141 (2000).
[25] X. Chen, R. Easther, and E. A. Lim, J. Cosmol. Astropart. Phys. 06 (2007) 023.

[26] M. Liguori, F. K. Hansen, E. Komatsu, S. Matarrese, and A. Riotto, Phys. Rev. D **73**, 043505 (2006).

[27] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach, Astrophys. J. **430**, 447 (1994).

[28] E. Komatsu and D. N. Spergel, Phys. Rev. D **63**, 063002 (2001).

[29] D. N. Spergel *et al.*, Astrophys. J. Suppl. Ser. **170**, 377 (2007).

[30] D. H. Lyth, C. Ungarelli, and D. Wands, Phys. Rev. D **67**, 023503 (2003).

[31] G. Dvali, A. Gruzinov, and M. Zaldarriaga, Phys. Rev. D **69**, 083505 (2004).

[32] G. Dvali, A. Gruzinov, and M. Zaldarriaga, Phys. Rev. D **69**, 023505 (2004).

[33] F. Bernardeau and J.-P. Uzan, Phys. Rev. D **66**, 103506 (2002).

[34] P. Creminelli, L. Senatore, M. Zaldarriaga, and M. Tegmark, J. Cosmol. Astropart. Phys. 3 (2007) 5.

[35] E. Komatsu *et al.*, Astrophys. J. Suppl. Ser. **148**, 119 (2003).

[36] E. Komatsu, B. D. Wandelt, D. N. Spergel, A. J. Banday, and K. M. Górski, Astrophys. J. **566**, 19 (2002).

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

[38] A. P. S. Yadav, E. Komatsu, and B. D. Wandelt, arXiv:astro-ph/0701921.

[39] A. Pillepich, C. Porciani, and S. Matarrese, arXiv:astro-ph/0611126.

[40] A. Cooray, Phys. Rev. Lett. **97**, 261301 (2006).

[41] E. Sefusatti, C. Vale, K. Kadota, and J. Frieman, Astrophys. J. **658**, 669 (2007).

[42] D. Babich, P. Creminelli, and M. Zaldarriaga, J. Cosmol. Astropart. Phys. 8 (2004) 9.

[43] M. Alishahiha, E. Silverstein, and D. Tong, Phys. Rev. D **70**, 123505 (2004).

[44] P. Creminelli, J. Cosmol. Astropart. Phys. 10 (2003) 3.

[45] The function, $M(k; a)$, uses the same definition as in [24] for $\Omega_m = 1$, but it differs from $M(k)$ given in [21], where the dependence on the growth function has been taken out as $M(k; a) = M^{HKM}(k)D(a)$. Also, the same function is defined in [23] in terms of the gravitational potential (i.e., the 0-0 component of the metric perturbations) during radiation domination, so that $M(k; a) = -\frac{9}{10}M^{SSZ}(k; a)$, leading to a different definition of the non-Gaussian parameter, $f_{NL} = -\frac{10}{9}f_{NL}^{SSZ}$.

[46] X. Chen, Phys. Rev. D **72**, 123518 (2005).

[47] N. Bartolo, S. Matarrese, and A. Riotto, J. Cosmol. Astropart. Phys. 10 (2005) 10.

[48] The equilateral reduced bispectrum is independent of scales only in the second-order perturbations. A scale dependence arises when the higher-order terms are included [55].

[49] J. N. Fry and E. Gaztanaga, Astrophys. J. **413**, 447 (1993).

[50] F. Marín, R. Wechsler, J. Frieman, and R. Nichol, arXiv:0704.0255.

[51] A. F. Heavens, S. Matarrese, and L. Verde, Mon. Not. R. Astron. Soc. **301**, 797 (1998).

[52] A. Taruya, Astrophys. J. **537**, 37 (2000).

[53] N. Kaiser, Mon. Not. R. Astron. Soc. **227**, 1 (1987).

[54] R. Scoccimarro, H. M. P. Couchman, and J. A. Frieman, Astrophys. J. **517**, 531 (1999).

[55] R. Scoccimarro, S. Colombi, J. N. Fry, J. A. Frieman, E. Hivon, and A. Melott, Astrophys. J. **496**, 586 (1998).

[56] D. Jeong and E. Komatsu, Astrophys. J. **651**, 619 (2006).

[57] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063520 (2006).

[58] M. Crocce and R. Scoccimarro, Phys. Rev. D **73**, 063519 (2006).

[59] M. Crocce and R. Scoccimarro, arXiv:0704.2783.

[60] S. Matarrese and M. Pietroni, J. Cosmol. Astropart. Phys. 06 (2007) 026.

[61] P. McDonald, Phys. Rev. D **75**, 043514 (2007).

[62] R. Scoccimarro, Phys. Rev. D **70**, 083007 (2004).

[63] R. Scoccimarro, Astrophys. J. **487**, 1 (1997).

[64] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep. **367**, 1 (2002).

[65] H. J. Mo, Y. P. Jing, and S. D. M. White, Mon. Not. R. Astron. Soc. **284**, 189 (1997).

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

[67] R. Scoccimarro, R. K. Sheth, L. Hui, and B. Jain, Astrophys. J. **546**, 20 (2001).

[68] J. L. Tinker, D. H. Weinberg, Z. Zheng, and I. Zehavi, Astrophys. J. **631**, 41 (2005).

[69] C. Conroy, R. H. Wechsler, and A. V. Kravtsov, Astrophys. J. **647**, 201 (2006).

[70] H.-J. Seo and D. J. Eisenstein, Astrophys. J. **598**, 720 (2003).

[71] M. Takada, E. Komatsu, and T. Futamase, Phys. Rev. D **73**, 083520 (2006).

[72] G. J. Hill, K. Gebhardt, E. Komatsu, and P. J. Mac Queen, in *The New Cosmology: Conference on Strings and Cosmology*, edited by R. E. Allen, D. V. Nanopoulos, and C. N. Pope, AIP Conf. Proc. No. 743 (AIP, New York, 2004), pp. 224–233.

[73] K. Glazebrook, D. Eisenstein, A. Dey, B. Nichol, and WFMOS Feasibility Study Dark Energy Team, arXiv:astro-ph/0507457.

[74] G. J. Melnick, G. G. Fazio, V. Tolls, D. T. Jaffe, K. Gebhardt, V. Bromm, E. Komatsu, and R. A. Woodruff, Bull. Am. Astron. Soc. **36**, 1509 (2004).