Halo clustering and $g_{NL}$-type primordial non-Gaussianity

Kendrick M. Smith$^1$, Simone Ferraro$^1$, and Marilena LoVerde$^2$

$^1$ Princeton University Observatory, Peyton Hall, Ivy Lane, Princeton, NJ 08544 USA
$^2$ Institute for Advanced Study, Einstein Drive, Princeton, NJ 08540, USA

Abstract

A wide range of multifield inflationary models generate non-Gaussian initial conditions in which the initial adiabatic fluctuation is of the form $(\zeta_G + g_{NL} \zeta_G^3)$. We study halo clustering in these models using two different analytic methods: the peak-background split framework, and brute force calculation in a barrier crossing model, obtaining agreement between the two. We find a simple, theoretically motivated expression for halo bias which agrees with $N$-body simulations and can be used to constrain $g_{NL}$ from observations. We discuss practical caveats to constraining $g_{NL}$ using only observable properties of a tracer population, and argue that constraints obtained from populations whose observed bias is $\lesssim 2.5$ are generally not robust to uncertainties in modeling the halo occupation distribution of the population.
1 Introduction

In the last few decades, increasingly precise observations (e.g. [1–6]) have led to a standard cosmological model in which small initial fluctuations evolve in a ΛCDM background to give rise to the observed universe. Current data are consistent with initial fluctuations which are adiabatic, scalar, Gaussian, with weak deviations from scale invariance ($n_s < 1$ at 3σ).

The statistics of the initial fluctuations, i.e. deviations from Gaussian initial conditions, provide a powerful probe of the physics of the early universe. In the context of inflation [7–13], the simplest models (single-field, minimally coupled slow-roll) predict initial curvature perturbations with negligible deviations from Gaussianity. However, there is a rich phenomenology of non-Gaussian initial conditions in models with multiple fields, self-interactions near horizon crossing, or speed of sound $c_s \ll 1$ during inflation. In this paper, we will focus on non-Gaussianity of the so-called local type [14–17], in which the primordial potential\(^1\) is of the form

$$\Phi(x) = \Phi_G(x) + f_{NL}(\Phi_G(x)^2 - \langle \Phi_G^2 \rangle) + g_{NL}(\Phi_G(x)^3 - 3\langle \Phi_G^2 \rangle \Phi_G(x))$$ \hspace{1cm} (1)

where $\Phi_G$ is a Gaussian field and $f_{NL}$, $g_{NL}$ are free parameters.\(^2\)

Local non-Gaussianity can be generated by physical mechanisms involving multiple fields, such as light spectator fields during inflation which evolve to generate the initial adiabatic fluctuations (the curvaton scenario) [18–20], or models where the inflaton decay rate is modulated by a second field [21, 22]. Non-Gaussianity of local type is also naturally generated in non-inflationary models of the early universe such as the new ekpyrotic/cyclic scenario [23–25]. There is a theorem [26, 27] which states that any single-field model of inflation cannot generate detectable levels of local non-Gaussianity without violating observed limits on deviation from a scale-invariant power spectrum. Thus, detection of either $f_{NL}$ or $g_{NL}$ would rule out all single field models of inflation and place powerful constraints on the physics of the early universe. Current observational constraints on these parameters are consistent with zero [1,28–30], but are expected to improve by an order of magnitude or more in the near future.

In models of inflation in which $|g_{NL}| = \mathcal{O}(f_{NL}^2)$, it is unlikely that observational constraints on $g_{NL}$ will be competitive with constraints on $f_{NL}$. However, there are a number of examples where $f_{NL}^2 \ll |g_{NL}|$, making the $g_{NL}$ term in Eq. (1) the dominant source of primordial non-Gaussianity. This situation arises in curvaton models where non-quadratic terms in the potential are important [31–35] or in multifield models in which $(\Delta N)$ varies rapidly at the end of inflation [36,37]. The existence of these scenarios makes searching for $g_{NL}$ just as important as $f_{NL}$ and measurements provide important constraints on the microphysical parameter space.

In a pioneering paper [38], Dalal et al showed that large-scale clustering of halos depends sensitively on $f_{NL}$. More precisely, a sample of halos (or tracers such as galaxies or quasars) with

\(^1\)In studies of primordial non-Gaussianity, it is conventional to define a primordial potential $\Phi = \frac{i}{3} \zeta$, where $\zeta$ is the initial adiabatic curvature perturbation. Note that $\Phi$ is not the conformal Newtonian potential, which is given by $\frac{2}{3} \Phi$ deep in the radiation-dominated epoch where Eq. (1) applies.

\(^2\)We define $g_{NL}$-type non-Gaussianity including the term $-3\langle \Phi_G^2 \rangle \Phi_G$; this term simply renormalizes $\Phi_G$ so that its power spectrum $P_{\Phi_G}$ is equal to the observed power spectrum $P_k$ (to first order in $g_{NL}$).
constant bias \( b_1 \) in a Gaussian cosmology will have scale-dependent bias given by

\[
b(k) \approx b_1 + 2\delta_c(b_1 - 1)\frac{f_{NL}}{\alpha(k, z)} \tag{2}
\]

in an \( f_{NL} \) cosmology. Here, \( \delta_c \) is the spherical collapse threshold and \( \alpha(k, z) \) is defined by

\[
\alpha(k, z) = \frac{2k^2 T(k) D(z)}{3\Omega_m H_0^2} \tag{3}
\]

so that the linear density field and the primordial potential are related by \( \delta_{\text{lin}}(k, z) = \alpha(k, z)\Phi(k) \). Large-scale structure constraints on \( f_{NL} \) from scale-dependent bias are currently competitive with the CMB (e.g. [28,39]) and may ultimately provide constraints which are stronger (e.g. [40,41]). The key identity (2) has been derived using several different analytic frameworks [28, 42, 43] and agrees with \( N \)-body simulations (e.g. [30,38,44,45]).

In this paper we study the related issue of large-scale halo-clustering in a \( g_{NL} \) cosmology. We consider the large-scale halo bias in two analytic frameworks: the peak-background split (§3) and a barrier crossing model (§4). We find consistency between the two formalisms (in disagreement with [46]) and obtain an expression analogous to Eq. (2) for the scale-dependent halo bias in a \( g_{NL} \) cosmology. Our main results are a universal relation between the scale-dependent halo bias in a \( g_{NL} \) cosmology and the mass function in an \( f_{NL} \) cosmology,

\[
b(k) \approx b_1 + \beta_g g_{NL} \frac{\alpha(k, z)}{\alpha(k, z)} \tag{4}
\]

and expressions for \( \beta_g \) (Eqs. (50), (51)) which can be used in practice to constrain \( g_{NL} \) from data. We also discuss caveats when estimating the \( g_{NL} \) bias from observable quantities (§5.4) and argue that constraints obtained from tracer populations which are not highly biased (\( b_1 \gtrsim 2.5 \)) are generally not robust to uncertainties in HOD modeling.

Throughout this paper we use the WMAP5+BAO+SN fiducial cosmology [47], with baryon density \( \Omega_b h^2 = 0.0226 \), CDM density \( \Omega_c h^2 = 0.114 \), Hubble parameter \( h = 0.70 \), spectral index \( n_s = 0.961 \), optical depth \( \tau = 0.080 \), and power-law initial curvature power spectrum \( k^3P_\zeta(k)/2\pi^2 = \Delta^2_\zeta(k/k_{\text{piv}})^{n_s-1} \) where \( \Delta^2_\zeta = 2.42 \times 10^{-9} \) and \( k_{\text{piv}} = 0.002 \) Mpc\(^{-1} \). All power spectra and transfer functions have been computed using CAMB [48].

2 Definitions and notation

We will sometimes model halos of mass \( \geq M \) with peaks in a smoothed density field \( \delta_M \) defined as follows. Let \( \delta_M(\mathbf{x}) \) be the linear density field smoothed by a tophat filter with radius \( R(M) = (3M/4\pi \rho_m)^{1/3} \), i.e.

\[
\delta_M(x) = \int \frac{d^3k}{(2\pi)^3} e^{-ik\cdot x} \delta_{\text{lin}}(k)W_M(k) \tag{5}
\]

where

\[
W_M(k) = 3 \frac{\sin(kR(M)) - kR(M) \cos(kR(M))}{(kR(M))^3} \tag{6}
\]
Let $\sigma_M = (\delta_M^2)^{1/2}$ be the RMS amplitude of the smoothed density field, and let $\kappa_n(M)$ be its $n$-th non-Gaussian cumulant, defined by:

$$\kappa_n(M) = \frac{\langle \delta_M^n \rangle_{\text{conn}}}{\sigma_M^n}.$$  

(7)

Since $\delta_M$ and $\sigma_M$ are defined via linear theory, $\kappa_n(M)$ is independent of redshift as implied by the notation. To first order in $f_{NL}$ and $g_{NL}$, we have

$$\kappa_3(M) = \kappa_3^{(1)}(M)f_{NL}$$

$$\kappa_4(M) = \kappa_4^{(1)}(M)g_{NL}$$

(8)  (9)

with higher cumulants equal to zero, where $\kappa_3^{(1)}(M), \kappa_4^{(1)}(M)$ are the values of the cumulants at $f_{NL} = 1$ and $g_{NL} = 1$ respectively. These values are given explicitly by:

$$\kappa_3^{(1)}(M) = \frac{6}{\sigma_M^3} \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} W_M(k)W_M(k')W_M(|k + k'|) \frac{P_{mm}(k)P_{mm}(k')\alpha(|k + k'|)}{\alpha(k)\alpha(k')}$$

$$\kappa_4^{(1)}(M) = \frac{24}{\sigma_M^4} \int \frac{d^3k}{(2\pi)^3} \int \frac{d^3k'}{(2\pi)^3} \int \frac{d^3k''}{(2\pi)^3} W_M(k)W_M(k')W_M(k'')W_M(|k + k' + k''|)$$

$$\times \frac{P_{mm}(k)P_{mm}(k')P_{mm}(k'')\alpha(|k + k' + k''|)}{\alpha(k)\alpha(k')\alpha(k'')}$$

(10)  (11)

where $\alpha(k)$ was defined previously in Eq. (3) and $P_{mm}(k)$ is the power spectrum of the linear density field, $\langle \delta_{\text{lin}}(k)\delta_{\text{lin}}(k') \rangle = (2\pi)^3 P_{mm}(k)\delta^{(3)}(k + k')$. For numerical calculation, the following fitting functions (from [49]) are convenient:

$$\kappa_3^{(1)}(M) = (6.6 \times 10^{-4}) \left(1 - 0.016 \log \left(\frac{M}{h^{-1}M_\odot}\right)\right)$$

$$\kappa_4^{(1)}(M) = (1.6 \times 10^{-7}) \left(1 - 0.021 \log \left(\frac{M}{h^{-1}M_\odot}\right)\right).$$

(12)  (13)

This paper is mainly concerned with calculating halo bias $b(k) = P_{nh}(k)/P_{nm}(k)$ to first order in $f_{NL}$ and $g_{NL}$, so let us establish notation from the outset, by writing the large-scale bias in the general form:

$$b(k) = b_1 + b_1f_{NL} + b_1g_{NL} + \frac{\beta_ff_{NL} + \beta_gg_{NL}}{\alpha(k)}$$

(14)

where unlike Eq. (2) and Eq. (4) we have allowed for scale-independent corrections $b_1f$ and $b_1g$ from $f_{NL}$ and $g_{NL}$ primordial non-Gaussianity. Equation (14) defines the coefficients $b_1, b_1f, b_1g, \beta_f, \beta_g$. This equation assumes that the $k$-dependence is of the functional form $(\text{constant}) + (\text{constant})/\alpha(k)$, but we will derive this analytically (Eq. (35)) and show that it agrees with simulations ($\S$5.1). In this notation, the Dalal et al formula (2) can be written as $\beta_f = 2\delta_c(b_1 - 1)$.

### 3 Peak-background split

The peak-background split formalism is a procedure for predicting halo clustering statistics on large scales. The basic idea is that a long-wavelength fluctuation in the initial curvature alters the
local abundance of halos in a way which is equivalent to perturbing parameters of the background cosmology, e.g. the matter density $\rho_m$ or the amplitude $\Delta_\Phi$ of the initial fluctuations. The use of this formalism to study halo bias in non-Gaussian cosmologies was pioneered in [28]; we will review this calculation of the bias in an $f_{NL}$ cosmology ($\S$3.1) and then perform an analogous calculation in the $g_{NL}$ case ($\S$3.2).

3.1 $f_{NL}$ cosmology

In an $f_{NL}$ cosmology, the initial conditions are given by:

$$\Phi(x) = \Phi_G(x) + f_{NL}(\Phi_G(x)^2 - \langle \Phi_G^2 \rangle)$$

(15)

To analyze the effect of a long-wavelength mode, let us decompose the Gaussian potential as a sum $\Phi_G = \Phi_l + \Phi_s$ of long-wavelength and short-wavelength contributions. The long/short-wavelength decomposition of the non-Gaussian potential $\Phi$ is then

$$\Phi(x) = \Phi_l(x) + f_{NL}(\Phi_l(x)^2 - \langle \Phi_l^2 \rangle) + (1 + 2f_{NL}\Phi_l(x))\Phi_s(x) + f_{NL}(\Phi_s(x)^2 - \langle \Phi_s^2 \rangle)$$

(16)

and contains explicit coupling between long and short wavelength modes of the Gaussian potential.

Let us consider how the term $(1 + 2f_{NL}\Phi_l(x))\Phi_s(x)$ in Eq. (16) affects $n_l(x)$, the long-wavelength part of the halo number density field. In a local region where the long-wavelength potential takes some value $\Phi_l$, the amplitude $\Delta_\Phi$ of the small-scale modes is perturbed: $\Delta_\Phi \rightarrow (1 + 2f_{NL}\Phi_l)\Delta_\Phi$. This modifies the local halo abundance, in the same way that the global abundance would be modified if the cosmological parameter $\Delta_\Phi$ were perturbed, i.e. we get a term in the long-wavelength halo density of the form $\Delta n(x) = 2f_{NL}\Phi_l(x)(\partial n/\partial \log \Delta_\Phi)$. In addition, even in a Gaussian cosmology, there is a perturbation to the local halo abundance which is proportional to the long-wavelength part $\delta_l(x)$ of the density fluctuation, i.e. a term of the form $\Delta n(x) = \delta_l(x)(\partial n/\partial \delta_l)$. Putting this together, the long-wavelength part of the halo density is given by:

$$\bar{n}_l(x) = \bar{n} + \frac{\partial n}{\partial \delta_l} \delta_l(x) + 2f_{NL} \frac{\partial n}{\partial \log \Delta_\Phi} \Phi_l(x)$$

(17)

$$= \bar{n}(1 + b_1 \delta_l(x) + \beta_f f_{NL} \Phi_l(x))$$

where

$$b_1 = \frac{\partial \log n}{\partial \delta_l}$$

(18)

$$\beta_f = 2 \frac{\partial \log n}{\partial \log \Delta_\Phi}.$$ 

(19)

---

3In this derivation, we have swept two terms in Eq. (15) under the rug; let us now argue that these are indeed negligible. The term $f_{NL}(\Phi_s(x)^2 - \langle \Phi_s^2 \rangle)$ alters the statistics of the small scale modes; this does perturb the halo abundance (by generating skewness in the density field) but the perturbation is independent of the long-wavelength fluctuation $\Phi_l$. Therefore, this term does not contribute to the large-scale halo bias. The term $f_{NL}(\Phi_l(x)^2 - \langle \Phi_l^2 \rangle)$ perturbs the long-wavelength modes and decorrelates them (to order $\mathcal{O}(f_{NL})$) from both the linear density fluctuation $\delta(x)$ and the field $(2f_{NL}\Phi_l)$ which modulates the local power spectrum amplitude $\Delta_\Phi$. In principle, this should generate stochastic bias at order $\mathcal{O}(f_{NL}^2)$, but we will neglect this, since we are only calculating to order $\mathcal{O}(f_{NL})$. 

4
Intuitively, in an $f_{NL}$ cosmology, the local power spectrum amplitude $\Delta_\phi$ is not spatially constant, but varies throughout the universe in a way which is proportional to the long-wavelength potential $\Phi_l$.

Computing the halo bias $b(k) = P_{mh}(k)/P_{mm}(k)$ from Eq. (17) for $n_l(x)$, we get:

$$b(k) = \frac{b_1 P_{mm}(k) + \beta_f P_{m\Phi}(k)}{P_{mm}(k)} = b_1 + \frac{\beta_f f_{NL}}{\alpha(k,z)}.$$  

From the preceding argument, we predict that the scale-dependent $f_{NL}$ bias is given by $\beta_f = 2(\partial \log n/\partial \log \Delta_\phi)$. We will refer to this as a “weak” prediction for the bias: it cannot be used to constrain $f_{NL}$ from real data, since $\beta_f$ has not been expressed in terms of observable quantities.

To make further progress, we need to evaluate the derivative $(\partial \log n/\partial \log \Delta_\phi)$, by making additional assumptions. If we assume that the halo mass function is universal, then one can calculate the derivative, obtaining $(\partial \log n/\partial \log \Delta_\phi) = \delta_c(b_1 - 1)$, where $b_1$ is the Gaussian bias [28], so that:

$$\beta_f = 2\delta_c(b_1 - 1).$$

We will refer to this as a “strong” prediction for the scale-dependent bias in an $f_{NL}$ cosmology, since $\beta_f$ has been expressed in terms of the observable quantity $b_1$. The strong form is essential for constraining $f_{NL}$ from observations.

3.2 $g_{NL}$ cosmology

Let us now generalize the analysis of large-scale clustering in the previous subsection to the case of a $g_{NL}$ cosmology, with initial conditions given by:

$$\Phi(x) = \Phi_G(x) + g_{NL}(\Phi_G(x))^3 - 3\langle \Phi_G^2 \rangle \Phi_G(x).$$

Separating the Gaussian field into long and short wavelength pieces $\Phi_G = \Phi_l + \Phi_s$, we decompose $\Phi$ as follows:

$$\Phi(x) = \underbrace{\Phi_l(x) + g_{NL}(\Phi_l(x))^3 - 3\langle \Phi_l^2 \rangle \Phi_l(x)}_{\text{long}} + \underbrace{\Phi_s(x) + 3g_{NL}(\Phi_l(x))^2 - \langle \Phi_l^2 \rangle)\Phi_s(x) + 3g_{NL}\Phi_l(x)(\Phi_s(x)^2 - \langle \Phi_s^2 \rangle) + g_{NL}(\Phi_s(x)^3 - 3\langle \Phi_s^2 \rangle \Phi_s(x))}_{\text{short}}$$

As in the $f_{NL}$ case, we’ll consider the perturbation to the long-wavelength halo density $n_l(x)$ generated by each of these terms.

The term $3g_{NL}(\Phi_l(x)^2 - \langle \Phi_l^2 \rangle)\Phi_s(x)$ can be interpreted as a local modulation in the small-scale power spectrum amplitude, given by $\Delta_\phi \rightarrow (1 + 3g_{NL}(\Phi_l(x)^2 - \langle \Phi_l^2 \rangle))\Delta_\phi$. This generates a term $\Delta_{nl}(x) = 3g_{NL}(\Phi_l(x)^2 - \langle \Phi_l^2 \rangle)(\partial n/\partial \log \Delta_\phi)$ in the long-wavelength halo density, in close analogy with the $f_{NL}$ case (the modulation is proportional to $g_{NL}(\Phi_l^2 - \langle \Phi_l^2 \rangle)$ in this case, rather than $f_{NL}\Phi_l$).

The term $3g_{NL}\Phi_l(x)(\Phi_s(x)^2 - \langle \Phi_s^2 \rangle)$ can be interpreted as follows. In a local region where the long-wavelength potential takes the value $\Phi_l$, the small-scale modes are perturbed in the same way.
as in an $f_{NL}$ cosmology where the global value of $f_{NL}$ is given by $(3g_{NL} \Phi_l)$. This generates a term
\[ \Delta n_l(x) = 3g_{NL} \Phi_l(x) \frac{\partial n}{\partial f_{NL}} \] in the long-wavelength halo density.

Finally, there is the usual term $\Delta n_l(x) = \delta_l(x)(\partial n/\partial \delta_l)$ due to changes in mean background density (as in the Gaussian case).

Putting this all together, we find that the long-wavelength halo density field in a $g_{NL}$ cosmology is given by:
\[ n_l(x) = \bar{n} + \frac{\partial n}{\partial \delta_l} \delta_l(x) + 3g_{NL} \frac{\partial n}{\partial \log \Delta_\Phi} (\Phi_l(x)^2 - \langle \Phi_l^2 \rangle) + 3g_{NL} \frac{\partial n}{\partial f_{NL}} \Phi_l(x) \]
\[ = \bar{n} \left( 1 + b_1 \delta_l(x) + \frac{3}{2} \beta_f g_{NL} (\Phi_l(x)^2 - \langle \Phi_l^2 \rangle) + \beta_g g_{NL} \Phi_l(x) \right) \]  
(24)

where $b_1$ and $\beta_f$ were defined previously (Eqs. (18), (19)), and:
\[ \beta_g = 3 \frac{\partial \log n}{\partial f_{NL}} \]  
(25)

The large-scale halo bias $b(k) = P_{nh}(k)/P_{nm}(k)$ is given by:
\[ b(k) = b_1 + \frac{\beta_g g_{NL}}{\alpha(k, z)} . \]  
(26)

Note that the $(\beta_f g_{NL})$ term in Eq. (24) does not contribute to the bias, since the field $(\Phi_l(x)^2 - \langle \Phi_l^2 \rangle)$ and the long-wavelength density field $\delta_l$ are uncorrelated (their cross correlation is a three-point function of Gaussian fields, which vanishes). This term should generate stochastic bias, but we defer a systematic study of halo stochasticity in non-Gaussian cosmologies to a future paper [50].

We have now arrived at the peak-background split prediction (26) for halo bias in a $g_{NL}$ cosmology, which relates the scale-dependent $g_{NL}$ bias to the derivative $(\partial \log n/\partial f_{NL})$ of the halo mass function in an $f_{NL}$ cosmology. In the terminology of the previous subsection, this is a "weak" prediction: we have shown that the problem of computing the $g_{NL}$ bias is naturally related to the problem of understanding the mass function in an $f_{NL}$ cosmology, but the coefficient $\beta_g$ has not been expressed in terms of observable quantities.

To obtain a "strong" prediction, we need to evaluate the derivative $(\partial \log n/\partial f_{NL})$, which requires making additional assumptions. This has been done in [49], assuming a barrier crossing model for the mass function and using the Edgeworth expansion to calculate the derivative (see also [51–57]). The result is:
\[ \frac{\partial \log n(M)}{\partial f_{NL}} = \frac{\kappa_3(M)}{6} H_3(\nu(M)) - \frac{1}{6} \frac{d \kappa_3}{d \nu} H_2(\nu(M)) \]  
(27)

where $\nu = \delta_c/\sigma_M$, and $H_3(x) = x^2 - 1$ and $H_3(x) = x^3 - 3x$ are Hermite polynomials. We will compare this prediction with N-body simulations in §5.

---

4 Analogously to the $f_{NL}$ case, we have neglected two terms in Eq. (23). The term $g_{NL}(\Phi_l(x)^3 - 3(\Phi_l^2)\Phi_l(x))$ only alters power spectra at order $O(g_{NL}^2)$, and we will neglect terms of this order. The term $g_{NL}(\Phi_l(x)^3 - 3(\Phi_l^2)\Phi_l(x))$ generates kurtosis in the density field and modifies the halo mass function [49], but in a way which is independent of $\Phi_l$ and therefore does not contribute to large-scale clustering.
4 Barrier crossing model

In this section, we will study large-scale bias using a barrier crossing model, obtaining results which are consistent with the peak-background split formalism from the previous section. The two approaches are complementary: the barrier model has the advantage that it generates complete predictions for halo statistics (such as the mass function or bias) via an algorithmic calculational procedure, but obscures the physical intuition of the peak-background split. For completeness, the calculations in this section will be sufficiently general to include the cases of Gaussian, $f_{NL}$-type, and $g_{NL}$-type initial conditions.

4.1 Setting up the calculation

The barrier crossing model is an old, widely influential idea in cosmology, in which halos of mass $\geq M$ are identified with peaks in the smoothed linear density field [58]. Although more complex versions have been proposed, we will use the simplest version: a spherical collapse model with constant barrier height, defined formally as follows.

We model halos of mass $\geq M$ as regions where the smoothed linear density field $\delta_M(x)$ (defined in Eq. (5)) exceeds the threshold value $\delta_c$, i.e. the halo number density $n_h(x)$ is given by:

$$n_h(x) = \frac{\rho_m}{M} \theta(\delta_M(x) - \delta_c)$$

(28)

where $\theta$ is the step function

$$\theta(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 & \text{if } x \geq 0 
\end{cases}$$

(29)

Throughout this paper, we take $\delta_c = 1.42$; this value produces somewhat improved agreement between the barrier model and simulations, compared to the Press-Schechter value $\delta_c = 1.69$.

To study halo bias in this model, we define the following notation. Let $x, x'$ be two points separated by distance $r$, let $\delta_{\text{lin}}$ denote the (unsmoothed) linear density field at $x$, and let $\delta'_M$ denote the smoothed linear density field at $x'$. We denote the joint PDF of these random variables by $p(\delta_{\text{lin}}, \delta'_M)$, and denote the 1-variable PDF of $\delta'_M$ by $p(\delta'_M)$. We define

$$p_0 = \int_{\delta_c}^{\infty} d\delta'_M p(\delta'_M)$$

(30)

$$\xi_0(r) = \int d\delta_{\text{lin}} d\delta'_M p(\delta_{\text{lin}}, \delta'_M) \delta_{\text{lin}} \theta(\delta'_M - \delta_c)$$

(31)

These quantities are related to the halo mass function $n(M)$ and matter-halo correlation function $\xi_{mh}(r)$, but there is one wrinkle. In the barrier crossing model, the field $n_h$ defined in Eq. (28) represents the number density of halos with mass $\geq M$, whereas we want to consider a sample of halos with mass in a narrow mass range $(M, M + dM)$. Thus $n(M)$ and $\xi_{mh}(r)$ are obtained by

---

5We experimented with using a mass-dependent barrier $\delta_c(\nu)$ chosen for consistency with a universal mass function such as Sheth-Tormen [59] or Warren [60], but found that this did not result in further improvement.
taking derivatives as follows:

\[ n(M) = -\frac{2\rho_m}{M} \left( \frac{dp_0}{dM} \right) \]  
\[ \xi_{mh}(r) = \frac{d\xi_0(r)/dM}{dp_0/dM} \]  

(32) \quad (33)

4.2 Mass function, halo bias, and interpretation

In principle, calculation of the halo mass function and large-scale bias in the barrier crossing model
has now been reduced to evaluation of Eqs. (30)–(33). We defer details of the calculation to Ap-
pendix A and quote the final results. The halo mass function is given by:

\[ n(M) = 2\rho_m \left( \frac{d\log \sigma^{-1}}{dM} \right) e^{-\nu^2/2} \left[ \nu + \int f_{NL} \left( \frac{\kappa^3_3(M)\nu H_3(\nu)}{6} - \frac{d\kappa^3_3/\nu}{d(\log \sigma^{-1})/dM} H_2(\nu) \right) \right. 
\left. + g_{NL} \left( \frac{\kappa^4_4(M)\nu H_4(\nu)}{24} - \frac{d\kappa^4_4/\nu}{d(\log \sigma^{-1})/dM} H_3(\nu) \right) \right] \]  

(34)

The halo bias \( b(k) = P_{mh}(k)/P_{mm}(k) \) is given by (in the large-scale limit \( k \to 0 \)):

\[ b(k) = b_1 + b_1 f_{NL} + b_1 g_{NL} + \frac{\beta f f_{NL} + \beta g g_{NL}}{\alpha(k)} \]  

(35)

where:

\[ b_1 = 1 + \frac{\nu^2 - 1}{\delta_c} \]  
\[ b_1 f = -\kappa^3_3(M) \left( \frac{\nu^3 - \nu}{2\delta_c} \right) + \frac{d\kappa^3_3/\nu}{d(\log \sigma^{-1})/dM} \left( \frac{\nu + \nu^{-1}}{6\delta_c} \right) \]  
\[ b_1 g = -\kappa^4_4(M) \left( \frac{\nu^4 - 3\nu^2}{6\delta_c} \right) + \frac{d\kappa^4_4/\nu}{d(\log \sigma^{-1})/dM} \left( \frac{\nu^2}{12\delta_c} \right) \]  
\[ \beta f = 2\nu^2 - 2 \]  
\[ \beta g = \kappa^3_3(M) \left( \frac{\nu^3 - 3\nu}{2} \right) - \frac{d\kappa^3_3/\nu}{d(\log \sigma^{-1})/dM} \left( \frac{\nu - \nu^{-1}}{2} \right) \]  

(36) \quad (37) \quad (38) \quad (39) \quad (40)

Although the above expressions are the result of a purely formal calculation, we will now show that
each term has a natural interpretation.

Considering first the halo mass function (34), we have found a Press-Schechter mass function
(with \( \delta_c = 1.42 \)) in the Gaussian case, plus first-order corrections in \( f_{NL} \) and \( g_{NL} \) which agree with
those found in [49, 52] using the Edgeworth expansion. This agreement is expected since the two
calculations are based on the same barrier crossing model.

Moving on to halo bias, in the Gaussian case, we predict that \( b(k) \) is constant on large scales,
with value \( b_1 \) given by Eq. (36). The peak-background split argument suggests a general relation
between the large-scale halo bias and the halo mass function which applies generally to a universal
mass function of the form:

\[ n(M) = \frac{\rho_m}{M} f(\nu) \frac{d\log \sigma^{-1}}{dM} \]  

(41)
On large scales, the bias is predicted to be scale-independent and given by [61]:

\[ b_1 = 1 - \frac{\nu}{\delta_c} \frac{d \log f}{d \nu} \]  

(42)

Comparing our predictions (34), (36) for \( n(M) \) and \( b_1 \), we find agreement, i.e. Eq. (36) for \( b_1 \) can be interpreted as the general peak-background split expression for halo bias, specialized to the Press-Schechter mass function.

More generally, the \( b_{1f} \) and \( b_{1g} \) contributions to the bias (Eqs. (37), (38)) represent shifts in the scale-independent part of the bias due to primordial non-Gaussianity. It is straightforward to check that these terms can be obtained by plugging the non-Gaussian mass function in Eq. (34) into the peak-background split prediction (42) for scale-independent bias, i.e. the \( b_{1f} \) and \( b_{1g} \) terms can be interpreted as changes to the bias which are entirely due to the mass function being perturbed in a non-Gaussian cosmology. This type of term (scale-independent bias proportional to \( f_{NL} \)) was first found for \( f_{NL} \) cosmologies in [30]. Note that a scale-independent shift is unobservable in practice, and cannot be used to constrain non-Gaussianity, since the bias of a real tracer population, such as galaxies or quasars, is a free parameter.

The \( \beta_f \) contribution to the bias is the well-known scale-dependent bias in an \( f_{NL} \) cosmology. Comparing Eq. (39) for \( \beta_f \) with Eqs. (34), (36), this term can be written either as \( \beta_f = 2\partial(\log n)/\partial(\log \Delta \Phi) \) or \( \beta_f = 2\delta_c(b_1 - 1) \). (In §3, we referred to these as “weak” and “strong” predictions.)

The \( \beta_g \) contribution to the bias is the focus of this paper: scale-dependent bias in a \( g_{NL} \) cosmology. Eq. (40) gives this term in the “strong” form that was found previously (Eq. (27)) using the peak-background split argument. Alternately, we can write this term in the “weak” form \( \beta_g = 3\partial(\log n)/\partial f_{NL} \) using Eq. (34).

In summary, we have found that the complete expression for large-scale halo bias in the barrier crossing model (Eq. (35)) agrees perfectly with the peak-background split calculation from §3. The bias contains a scale-independent part \( (b_1 + b_{1f}f_{NL} + b_{1g}g_{NL}) \) which can be obtained from the halo mass function, via the general relation (42). The scale-independent bias depends on \( f_{NL} \) and \( g_{NL} \), because the halo mass function depends on these parameters. The bias also contains a scale-dependent part \( (\beta_{fNL} + \beta_{gNL})/\alpha(k) \) whose coefficients can be calculated explicitly and agree with the peak-background split predictions.

4.3 Comparison with previous work

It is interesting to compare the above calculations with the results of [62] (see also [43]), where \( \beta_g \) was calculated using the MLB formula [63], which gives \( N \)-point functions of halos as an asymptotic series in \( \nu \). The scale-dependent \( g_{NL} \) bias was found to be (in our notation):

\[ \beta_g^{\text{MLB}} = \kappa_3^{(1)}(M) \frac{\delta_c \nu (b_1 - 1)}{2} \]  

(43)

When this prediction was compared to \( N \)-body simulations, it was found to be a poor fit.

Comparing \( \beta_g^{\text{MLB}} \) with our calculation (40) for \( \beta_g \), it is seen that the two agree in the high-mass limit \( \nu \rightarrow \infty \), but disagree in subleading terms. This is expected since the MLB formula is based on
the same barrier crossing model that we have used, but it is an asymptotic result, whereas we have done an exact calculation (to first order in $f_{NL}$, $g_{NL}$). For realistic halo masses, the “subleading” terms neglected in the MLB formula are of order one (to quantify this better, $\beta_g$ and $\beta_g^{\text{MLB}}$ agree to 10% only when the halo bias $b_1 \geq 15$), so in practice the two predictions are quite different.

Recently, ref. [46] argued that the barrier crossing model cannot generate correct predictions for general non-Gaussian initial conditions such as the $g_{NL}$ model, but we found the opposite conclusion: brute-force calculation in the barrier crossing model, collecting all terms of order $O(g_{NL})$, agrees precisely (i.e. to all orders in $1/\nu$) with the peak-background split. It seems intuitively plausible that two must be consistent, since the peak-background split argument depends only on the assumption that halo formation is determined by the local density field, and the barrier crossing model is a concrete example of a model in which this assumption is satisfied.

5 Results from $N$-body simulations

In the last two sections, we have obtained complete analytic predictions for large-scale bias in a $g_{NL}$ cosmology, finding agreement between the peak-background split formalism (§3) and a barrier crossing model based on spherical collapse (§4).

To compare these predictions with simulation, we performed collisionless $N$-body simulations using the GADGET-2 TreePM code [64]. Simulations were done using periodic box size $R_{\text{box}} = 1600 \ h^{-1} \ \text{Mpc}$, particle count $N_p = 1024^3$, and force softening length $R_s = 0.05(R_{\text{box}}/N_p^{1/3})$. With these parameters and the fiducial cosmology from §1, the particle mass is $m_p = 2.92 \times 10^{11} \ h^{-1} \ M_\odot$.

We generate initial conditions by simulating a Gaussian primordial potential $\Phi$, and applying $f_{NL}$ or $g_{NL}$ corrections by straightforward use of Eq. (1). We linearly evolve to redshift $z_{\text{ini}} = 100$ using the transfer function$^6$ from CAMB [48], and obtain initial particle positions at this redshift using the Zeldovich approximation [65]. (At $z_{\text{ini}} = 100$, transient effects due to use of this approximation should be negligible [66].)

After running the $N$-body simulation, we group particles into halos using an MPI parallelized implementation of the friends-of-friends algorithm [67] with link length $L_{\text{FOF}} = 0.2R_{\text{box}}N_p^{-1/3}$. For a halo containing $N_{\text{FOF}}$ particles, we assign a halo position given by the mean of the individual particle positions. We estimate halo bias $b(k) = P_{mh}(k)/P_{mm}(k)$ using the procedure described in Appendix A of [68]. The statistical error $\Delta b(k)$ obtained using this procedure is smaller than the error that would be obtained assuming uncorrelated estimates of the power spectra $P_{mm}$ and $P_{mh}$, since shared sample variance is taken into account.

Results in this paper are based on 4 simulations with Gaussian initial conditions, 5 simulations with $g_{NL} = \pm 2 \times 10^6$, and 3 simulations with $f_{NL} = \pm 250$ (for a total of 20 simulations).

---

$^6$One subtlety here: straightforward use of CAMB’s transfer function at redshift 100 leads to inconsistencies since CAMB includes radiation (which is not negligible at $z = 100$) in its expansion history, while GADGET does not. For this reason we use CAMB’s linear transfer function at low redshift and extrapolate back to $z = 100$ using the growth function in an $\Omega_{\text{rad}} = 0$ universe.
Figure 1: An example to illustrate that halo bias in a $g_{NL}$ cosmology takes the functional form
$b(k) = b_1 + \beta_g g_{NL} / \alpha(k)$. This figure corresponds to redshift $z = 0.5$ and halo mass range $1.15 \leq M \leq 1.83 \times 10^{14} \ h^{-1} M_\odot$, but we find the same functional form for all redshifts and halo masses.

5.1 Fitting the functional form $b(k) = b_1 + \beta_g g_{NL} / \alpha(k)$

We now compare our analytic prediction for $b(k)$ to simulation in several steps, corresponding to increasingly strong versions of the prediction.

First, consider the weakest possible question: our analytic prediction for the bias is of the functional form
$$b(k) = b_1 + \beta_g g_{NL} / \alpha(k)$$

(44)

Is this a good fit to simulation, if we treat the coefficients $b_1$ and $\beta_g$ as free parameters? (We will compare our analytic prediction for $\beta_g$ to simulation in the next subsection; for now we are just testing whether the functional form (44) is correct.)

In Fig. 1, we show some example fits of this form, for redshift $z = 0.5$ and halo mass range $1.15 \leq M \leq 1.83 \times 10^{14} \ h^{-1} M_\odot$. Each fit was performed using bias estimates from 4 independent simulations with $L_{box} = 1600 \ h^{-1}$ Mpc and wavenumbers $k \leq 0.04 \ h \ Mpc^{-1}$. We find good $\chi^2$ values for the fits, with recovered parameters:

$$b_1 = 3.653 \pm 0.026$$

$b_1, 10^3 \beta_g = (3.575 \pm 0.038, 0.581 \pm 0.056)$ for $g_{NL} = 0$

$10^3 \beta_g = (3.824 \pm 0.039, 0.935 \pm 0.060)$ for $g_{NL} = -2 \times 10^6$

(45)

We note that the recovered bias parameters (45) in this example show that both $b_1$ and $\beta_g$ are $g_{NL}$-dependent. In the barrier crossing model, we made a prediction for the $g_{NL}$ dependence of
We find good agreement between this prediction and our simulations. Note that in practice, the $g_{NL}$ dependence of $b_1$ is unobservable since for a real tracer population, the halo occupation distribution is not known precisely and $b_1$ must be treated as a free parameter to be determined from data.

The observed $g_{NL}$ dependence of $\beta_g$ corresponds to scale-dependent bias of order $O(g_{NL}^2)$ or higher (note that $\beta_g$ is defined in such a way that constant $\beta_g$ corresponds to scale-dependent bias which is linear in $g_{NL}$). This complicates comparison with our analytic predictions, since we have only calculated the bias to order $O(g_{NL})$. We address this by estimating $\beta_g$ by averaging the estimates obtained from simulations with $g_{NL} = \pm 2 \times 10^6$, thus removing contributions to $b(k)$ which are proportional to $g_{NL}^2$. Note that this does not remove $O(g_{NL}^3)$ contributions to the bias, but we have checked that such contributions are negligible for $g_{NL} = \pm 2 \times 10^6$, by comparing with simulations with halved step size.

Repeating this fitting procedure for redshifts $z \in \{2, 1, 0.5, 0\}$ and a range of halo masses (the precise set of halo mass bins used is shown in Fig. 2 below), we find $\chi^2$ values which are consistent with their expected distribution, i.e. we find that the functional form (44) is a good fit to the simulations for a wide range of redshifts and halo masses. For this reason, in subsequent sections, we will “compress” the estimates of $b(k)$ in each simulation (as shown in Fig. 1) to two numbers ($b_1$ and $\beta_g$), with statistical errors given by the fitting procedure.

### 5.2 Comparison with analytic predictions

Now that we have established the functional form $b(k) = b_1 + \beta_g g_{NL}/\alpha(k)$ of the bias, and a procedure for estimating $\beta_g$ from simulation as a function of redshift and halo mass, we would like to compare with our analytic predictions for $\beta_g$.

First, consider the “weak” form of the prediction ($\beta_g = 3(\partial \log n/\partial f_{NL})$) obtained from the peak-background split argument. We can test this prediction by estimating the derivative ($\partial \log n/\partial f_{NL}$) directly from simulations, by taking finite differences of $\log(n)$ in simulations with $f_{NL} = \pm 250$. (We checked convergence in the step size.) We find that the prediction holds perfectly (within the statistical errors of the simulations) for all redshifts and halo masses (Fig. 2).

Second, consider the “strong” Edgeworth prediction (Eq. (40)), in which an explicit formula for $\beta_g$ is given. In this case, we find reasonable agreement at high mass ($M \gtrsim 10^{14} \, h^{-1} M_\odot$), but the prediction breaks down at low halo mass (Fig. 2).

Our interpretation is as follows. The peak-background split prediction $\beta_g = 3(\partial \log n/\partial f_{NL})$ is a universal relation between bias in a $g_{NL}$ cosmology and the mass function in an $f_{NL}$ cosmology. Although “weak” in the sense that it does not supply a closed-form expression for $\beta_g$, the derivation makes few assumptions, and one expects it to be exact. In order to constrain $g_{NL}$ from real data, we need a “strong” prediction which expresses $\beta_g$ in closed form, using only observable quantities (i.e. the analog of the Dalal et al formula $\beta_f = 2\delta_c(b_1 - 1)$ for an $f_{NL}$ cosmology). Using the Edgeworth expansion, one can make such a prediction in the context of the barrier crossing model (Eq. (40)), and obtain rough agreement with simulations, but the level of agreement is not really good enough for doing precision cosmology. Therefore, we next propose a slightly modified version of the Edgeworth prediction.
Figure 2: Comparison of the “weak” and “strong” predictions for the scale-dependent bias in a $g_{NL}$ cosmology. **Blue squares:** Direct estimates of the bias, extracted from simulations with $g_{NL} = \pm 2 \times 10^6$ as described in §5.1. **Green circles:** “Weak” analytic prediction for the bias ($\beta_g = 3(\partial \log n / \partial f_{NL})$) from the peak-background split formalism, showing perfect agreement. The estimates of $(\partial \log n / \partial f_{NL})$ shown in the figure were obtained directly from simulations with $f_{NL} = \pm 250$. **Red dotted curve:** Edgeworth prediction for the bias (Eq. (40)). Good agreement is seen at high mass, but at low masses Edgeworth underpredicts $3(\partial \log n / \partial f_{NL})$. We will find an improvement in §5.3.
5.3 A simple universal formula for the bias in a $g_{NL}$ cosmology

We would like to slightly modify the Edgeworth prediction (40) for $\beta_g$ so that it agrees better with $N$-body simulations. It is also convenient to have a prediction in which $\beta_g$ is given as a function of observable quantities: Gaussian bias $b_1$ (rather than halo mass, which is unobservable) and redshift $z$.

We start by rewriting the Edgeworth prediction (40) for $\beta_g$ in terms of variables $(b_1, z)$. The following fitting functions for $\kappa_3$ and $d\kappa_3/d\log(\sigma^{-1})$ are convenient:

\[
\kappa_3 = 0.000329(1 + 0.09z)b_1^{-0.09} \tag{46}
\]
\[
\frac{d\kappa_3}{d\log(\sigma^{-1})} = -0.000061(1 + 0.22z)b_1^{-0.25} \tag{47}
\]

For purposes of this subsection, we define the quantity $\nu$ to be given in terms of $b_1$ and $z$ by:

\[
\nu = [\delta_c(b_1 - 1) + 1]^{1/2} \quad \text{(where } \delta_c = 1.42) \tag{48}
\]

The Edgeworth prediction for $\beta_g$ can be written in the following form:

\[
\beta_{g_{\text{Edge}}} = \kappa_3 \left[ -1 + \frac{3}{2}(\nu - 1)^2 + \frac{1}{2}(\nu - 1)^3 \right] - \frac{d\kappa_3}{d\log(\sigma^{-1})} \left( \frac{\nu - \nu^{-1}}{2} \right) \tag{49}
\]

Empirically, we find that if we tweak the Edgeworth prediction by changing the coefficients of the polynomial in brackets as follows:

\[
\beta_g = \kappa_3 \left[ -0.7 + 1.4(\nu - 1)^2 + 0.6(\nu - 1)^3 \right] - \frac{d\kappa_3}{d\log(\sigma^{-1})} \left( \frac{\nu - \nu^{-1}}{2} \right) \tag{50}
\]

then we obtain good agreement with simulations (Fig. 3). The expression (50) for $\beta_g$ (with quantities $\kappa_3$, $d\kappa_3/d\log(\sigma^{-1})$, $\nu$ defined by Eqs. (46)–(48)) is one of the main results of this paper and is our observational “bottom line” when constraining $g_{NL}$ from real data.

We have motivated our “tweak” to the Edgeworth prediction as essentially a fitting function for the $\nu$ dependence (although it is worth noting that the $z$ dependence is correctly predicted by the barrier crossing model). A speculative interpretation of this tweak, which we will defer for future work, is as follows. In the barrier crossing model, the second-order halo bias is given by $b_2 = (\nu^3 - 3\nu)/\langle \delta_c \sigma_M \rangle$. It is tempting to conjecture that the expression in brackets in Eq. (50) is generally equal to $\langle \delta_c \sigma_M b_2 \rangle$, and interpret our “tweak” to the Edgeworth prediction (49) as perturbing the relation between $b_1$ and $b_2$, relative to the barrier crossing model. This opens up the possibility of directly measuring the second-order bias and determining $\beta_g$ directly. To study the viability of this idea, one would need to compare $\beta_g$ in simulation to some other estimate of second-order halo bias, such as the halo bispectrum in the squeezed limit.

5.4 An important caveat

There is an important caveat when using Eq. (50), or indeed any fitting function for the $g_{NL}$ bias, to constrain $g_{NL}$ from real data. It is tempting to compute $\beta_g$ by simply plugging the observed bias $b_1$
and redshift $z$ into Eq. (50). (Since the $z$-dependence is very mild, a rough estimate for the redshift suffices.) However, we have only shown that this procedure is correct in the limit of a narrow bin in halo mass and redshift, and a real tracer population will be a weighted average over $M$ and $z$.

For example, consider the case in which the “tracers” are the dark matter particles themselves, i.e. each halo is weighted in proportion to its mass (assuming all mass is in halos). This tracer population has bias $b_1 = 1$ (for the trivial reason that we are back to the dark matter field), so straightforward use of Eq. (50) would suggest that $\beta_g \approx -0.00025$. (This value would make the low-$k$ power spectrum a reasonably sensitive probe of $g_{NL}$.) In fact, the true $\beta_g$ of this tracer is zero, since the matter power spectrum $P_{mm}(k)$ does not contain a term proportional to $g_{NL}/\alpha(k)$. This example shows that the true $g_{NL}$ bias of a tracer population can differ significantly from the value obtained by straightforward use of Eq. (50). In general, the $g_{NL}$ bias will depend on the full HOD (halo occupation distribution) of the tracer population, not only on the Gaussian bias $b_1$.\footnote{Note that there is no analogous caveat in the $f_{NL}$ case. Because the relation $\beta_f = 2\delta_c(b_1 - 1)$ is linear, it applies to both a tracer population which is narrowly selected in $(M,z)$ and to a population which is an arbitrary weighted average over $(M,z)$.}

One popular approach to modeling the HOD is to assume that halos below some minimum mass $M_{\text{min}}$ do not host tracers, whereas the mean number of tracers in a halo of mass $M \geq M_{\text{min}}$ is proportional to the total mass $M$. For reference, we give a fitting function for the $g_{NL}$ bias for this HOD:

$$
\beta_g = \kappa_3 \left[ -0.4(\nu - 1) + 1.5(\nu - 1)^2 + 0.6(\nu - 1)^3 \right] \tag{51}
$$

where for purposes of this equation, $\kappa_3$ and $\nu$ are defined as functions of the observables $b_1$ and $z$. 

---

**Figure 3**: Scale-dependent $g_{NL}$ bias coefficient $\beta_g$ as a function of redshift $z$ and halo bias $b_1$, showing excellent agreement between our final analytic result (Eq. (50), dashed curves) and $N$-body simulations (error bars).
by Eqs. (46), (48) above.

Eq. (51) applies to a mass-weighted population of halos above $M_{\text{min}}$, whereas Eq. (50) applies to a population which is narrowly selected in mass. The two agree for $b_1 \gtrsim 2.5$, suggesting that HOD dependence is small in practice for highly biased samples, but disagree qualitatively for $b_1 \lessgtr 2.5$. For example, the $g_{NL}$ bias $\beta_g$ changes sign at $b_1 \approx 2.1$ for the narrowly selected sample (Eq. (50)), whereas $\beta_g$ is always positive for the mass-weighted sample (Eq. (51)).

Our perspective is that, in order to obtain $g_{NL}$ constraints which are robust to HOD modeling uncertainty, one should use highly biased samples ($b_1 \gtrsim 2.5$), where this uncertainty will be minimized. Samples which are not highly biased do not give robust constraints; for example, a tracer population with $b_1 \approx 1.8$ can have a $g_{NL}$ bias $\beta_g$ which is negative, zero, or positive, depending on the HOD.

For highly biased samples, it is useful to make the following observation: the $g_{NL}$ bias $\beta_{g}^{\text{fit}}$ which is obtained from straightforward use of Eq. (50) is always less than the true $g_{NL}$ bias $\beta_{g}^{\text{true}}$. This follows from positivity of the second derivative $d^2 \beta_g/db_1^2$. It follows that a $g_{NL}$ constraint obtained using $\beta_{g}^{\text{fit}}$ is always valid, but slightly overestimates the statistical error that could be obtained if $\beta_{g}^{\text{true}}$ were known. This effectively treats HOD uncertainty as an extra source of systematic error.

6 Discussion

We have computed large-scale halo bias for non-Gaussian initial conditions, using two analytic frameworks: the peak-background split formalism (§3) and a barrier crossing model (§4), finding agreement between the two. Although our emphasis has been on the constant-$f_{NL}$ and constant-$g_{NL}$ models, our calculational machinery should apply to more general non-Gaussian initial conditions.

The peak-background split formalism is simpler and also suggests a simple physical picture of non-Gaussian cosmologies on large scales. In an $f_{NL}$ cosmology, the amplitude $\Delta_{\Phi}$ of the initial fluctuations is not spatially constant, but is proportional to $(1 + 2f_{NL}\Phi_l)$. Thus, $\Delta_{\Phi}$ has fluctuations on large scales which are 100% correlated with the long-wavelength potential, generating halo bias of the form $(\beta_{f}f_{NL}/\alpha(k))$. In a $g_{NL}$ cosmology, the small-scale skewness is nonzero and proportional to $(g_{NL}\Phi_l)$, leading to halo bias of the form $(\beta_{g}g_{NL}/\alpha(k))$. The peak-background split argument is very useful for generating universal relations such as $\beta_g = 3\partial(\log n)/\partial f_{NL}$, which are “weak” in the sense that the RHS has not been expressed in terms of observable quantities, but have the advantage of being exact (as can be seen by comparing the two sets of errorbars in Fig. 2).

The barrier crossing model generates all terms in the large-scale bias, including terms such as $b_1 f$ and $b_1 g$ which are easy to miss, by a purely algorithmic calculational procedure. In addition, the barrier crossing model generates “strong” forms of the bias coefficients (e.g. the Edgeworth expression (40) for $\beta_g$), which are closed-form expressions in $M$ and $z$. However, these expressions are not exact because the barrier crossing model is approximation to the true process of halo formation.

---

8This statement assumes that the probability that a halo hosts a tracer is a function only of the mass and redshift. If the probability depends strongly on additional variables such as merger history, triaxiality, etc. then this will generate additional contributions to $\beta_g$, in analogy to the $f_{NL}$ case [28,69]. In principle, selection biases to $\beta_g$ can be addressed by folding the selection into the mass function when computing $\partial(\log n)/\partial f_{NL}$, but detailed study is beyond the scope of this paper.
To obtain a “bottom line” expression for the scale-dependent $g_{NL}$ bias $\beta_g$ in terms of redshift $z$ and Gaussian bias $b_1$, we found it necessary to tweak slightly the $b_1$ dependence of the Edgeworth prediction, arriving at the expression (50) which agrees very well with simulations. The caveat is that Eq. (50) applies only to a halo population which has been selected in a narrow halo mass and redshift range. In principle, one can calculate $\beta_g$ for a tracer population by multiplying by the halo occupation distribution and integrating over mass and redshift. In practice, the HOD is not known precisely and we have argued in §5.4 that the best approach is to only use highly biased populations ($b \gtrsim 2.5$) for constraining $g_{NL}$. Since $\beta_g$ is a rapidly increasing function of $b_1$, this strategy makes sense both from the perspective of minimizing statistical errors, and systematic errors due to HOD uncertainty. In data analysis, it may be useful to impose cuts which increase the mean halo bias at the expense of reducing the number of tracers. Another advantage of subdividing tracer populations is that this may permit $f_{NL}$ and $g_{NL}$ to be constrained simultaneously (with a single tracer population, the two are degenerate).

Acknowledgements

We thank Neal Dalal, Vincent Desjacques, Chris Hirata, Donghui Jeong, Brant Robertson, Fabian Schmidt, Neelima Sehgal, David Spergel and Matias Zaldarriaga for helpful discussions. K. M. S. is supported by a Lyman Spitzer fellowship in the Department of Astrophysical Sciences at Princeton University. M. L. is supported as a Friends of the Institute for Advanced Study Member and by the NSF through AST-0807444. S. F. is supported by the Martin Schwarzchild Fund in Astronomy at Princeton University. Simulations in this paper were performed at the TIGRESS high performance computer center at Princeton University which is jointly supported by the Princeton Institute for Computational Science and Engineering and the Princeton University Office of Information Technology.

References

[1] E. Komatsu et al., “Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation,” arXiv:1001.4538 [astro-ph.CO].

[2] SDSS Collaboration, W. J. Percival et al., “Baryon Acoustic Oscillations in the Sloan Digital Sky Survey Data Release 7 Galaxy Sample,” Mon. Not. Roy. Astron. Soc. 401 (2010) 2148–2168, arXiv:0907.1660 [astro-ph.CO].

[3] B. A. Reid et al., “Cosmological Constraints from the Clustering of the Sloan Digital Sky Survey DR7 Luminous Red Galaxies,” Mon. Not. Roy. Astron. Soc. 404 (2010) 60–85, arXiv:0907.1659 [astro-ph.CO].

[4] A. G. Riess et al., “A 3% Solution: Determination of the Hubble Constant with the Hubble Space Telescope and Wide Field Camera 3,” Astrophys. J. 730 (2011) 119, arXiv:1103.2976 [astro-ph.CO].
[5] R. Kessler et al., “First-year Sloan Digital Sky Survey-II (SDSS-II) Supernova Results: Hubble Diagram and Cosmological Parameters,” *Astrophys. J. Suppl.* **185** (2009) 32–84, arXiv:0908.4274 [astro-ph.CO].

[6] A. Vikhlinin et al., “Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints,” *Astrophys. J.* **692** (2009) 1060–1074, arXiv:0812.2720 [astro-ph].

[7] A. H. Guth, “The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems,” *Phys. Rev. D23* (1981) 347–356.

[8] A. D. Linde, “Chaotic Inflation,” *Phys. Lett. B129* (1983) 177–181.

[9] A. Albrecht and P. J. Steinhardt, “Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking,” *Phys. Rev. Lett.* **48** (1982) 1220–1223.

[10] A. H. Guth and S. Y. Pi, “Fluctuations in the New Inflationary Universe,” *Phys. Rev. Lett.* **49** (1982) 1110–1113.

[11] S. W. Hawking, “The Development of Irregularities in a Single Bubble Inflationary Universe,” *Phys. Lett. B115* (1982) 295.

[12] A. A. Starobinsky, “Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations,” *Phys. Lett. B117* (1982) 175–178.

[13] J. M. Bardeen, P. J. Steinhardt, and M. S. Turner, “Spontaneous Creation of Almost Scale-Free Density Perturbations in an Inflationary Universe,” *Phys. Rev. D28* (1983) 679.

[14] D. S. Salopek and J. R. Bond, “Nonlinear evolution of long wavelength metric fluctuations in inflationary models,” *Phys. Rev. D42* (1990) 3936–3962.

[15] A. Gangui, F. Lucchin, S. Matarrese, and S. Mollerach, “The Three point correlation function of the cosmic microwave background in inflationary models,” *Astrophys. J.* **430** (1994) 447–457, arXiv:astro-ph/9312033.

[16] E. Komatsu and D. N. Spergel, “Acoustic signatures in the primary microwave background bispectrum,” *Phys. Rev. D63* (2001) 063002, arXiv:astro-ph/0005036.

[17] T. Okamoto and W. Hu, “The Angular Trispectra of CMB Temperature and Polarization,” *Phys. Rev. D66* (2002) 063008, arXiv:astro-ph/0206155.

[18] A. D. Linde and V. F. Mukhanov, “Nongaussian isocurvature perturbations from inflation,” *Phys. Rev. D56* (1997) 535–539, arXiv:astro-ph/9610219.

[19] D. H. Lyth and D. Wands, “Generating the curvature perturbation without an inflaton,” *Phys. Lett. B524* (2002) 5–14, arXiv:hep-ph/0110002.

[20] D. H. Lyth, C. Ungarelli, and D. Wands, “The primordial density perturbation in the curvaton scenario,” *Phys. Rev. D67* (2003) 023503, arXiv:astro-ph/0208055.
[21] G. Dvali, A. Gruzinov, and M. Zaldarriaga, “A new mechanism for generating density perturbations from inflation,” *Phys. Rev.* **D69** (2004) 023505, arXiv:astro-ph/0303591.

[22] L. Kofman, “Probing string theory with modulated cosmological fluctuations,” arXiv:astro-ph/0303614.

[23] E. I. Buchbinder, J. Khoury, and B. A. Ovrut, “Non-Gaussianities in New Ekpyrotic Cosmology,” *Phys. Rev. Lett.* **100** (2008) 171302, arXiv:0710.5172 [hep-th].

[24] P. Creminelli and L. Senatore, “A smooth bouncing cosmology with scale invariant spectrum,” *JCAP* **0711** (2007) 010, arXiv:hep-th/0702165.

[25] J.-L. Lehners and P. J. Steinhardt, “Non-Gaussian Density Fluctuations from Entropically Generated Curvature Perturbations in Ekpyrotic Models,” *Phys. Rev.* **D77** (2008) 063533, arXiv:0712.3779 [hep-th].

[26] J. M. Maldacena, “Non-Gaussian features of primordial fluctuations in single field inflationary models,” *JHEP* **05** (2003) 013, arXiv:astro-ph/0210603.

[27] P. Creminelli and M. Zaldarriaga, “Single field consistency relation for the 3-point function,” *JCAP* **0410** (2004) 006, arXiv:astro-ph/0407059.

[28] A. Slosar, C. Hirata, U. Seljak, S. Ho, and N. Padmanabhan, “Constraints on local primordial non-Gaussianity from large scale structure,” *JCAP* **0808** (2008) 031, arXiv:0805.3580 [astro-ph].

[29] J. R. Fergusson, D. M. Regan, and E. P. S. Shellard, “Optimal Trispectrum Estimators and WMAP Constraints,” arXiv:1012.6039 [astro-ph.CO].

[30] V. Desjacques, U. Seljak, and I. Iliev, “Scale-dependent bias induced by local non-Gaussianity: A comparison to N-body simulations,” arXiv:0811.2748 [astro-ph].

[31] M. Sasaki, J. Valiviita, and D. Wands, “Non-gaussianity of the primordial perturbation in the curvaton model,” *Phys. Rev.* **D74** (2006) 103003, arXiv:astro-ph/0607627.

[32] K. Ichikawa, T. Suyama, T. Takahashi, and M. Yamaguchi, “Non-Gaussianity, Spectral Index and Tensor Modes in Mixed Inflaton and Curvaton Models,” *Phys. Rev.* **D78** (2008) 023513, arXiv:0802.4138 [astro-ph].

[33] K. Enqvist and T. Takahashi, “Signatures of Non-Gaussianity in the Curvaton Model,” *JCAP* **0809** (2008) 012, arXiv:0807.3069 [astro-ph].

[34] Q.-G. Huang, “Curvaton with Polynomial Potential,” *JCAP* **0811** (2008) 005, arXiv:0808.1793 [hep-th].

[35] P. Chingangbam and Q.-G. Huang, “The Curvature Perturbation in the Axion-type Curvaton Model,” *JCAP* **0904** (2009) 031, arXiv:0902.2619 [astro-ph.CO].
[36] Q.-G. Huang, “A geometric description of the non-Gaussianity generated at the end of multi-field inflation,” *JCAP* **0906** (2009) 035, arXiv:0904.2649 [hep-th].

[37] C. T. Byrnes and G. Tasinato, “Non-Gaussianity beyond slow roll in multi-field inflation,” *JCAP* **0908** (2009) 016, arXiv:0906.0767 [astro-ph.CO].

[38] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, “The imprints of primordial non-gaussianities on large-scale structure: scale dependent bias and abundance of virialized objects,” *Phys. Rev.* **D77** (2008) 123514, arXiv:0710.4560 [astro-ph].

[39] J.-Q. Xia, C. Baccigalupi, S. Matarrese, L. Verde, and M. Viel, “Constraints on Primordial Non-Gaussianity from Large Scale Structure Probes,” arXiv:1104.5015 [astro-ph.CO].

[40] C. Cunha, D. Huterer, and O. Dore, “Primordial non-Gaussianity from the covariance of galaxy cluster counts,” *Phys. Rev.* **D82** (2010) 023004, arXiv:1003.2416 [astro-ph.CO].

[41] N. Hamaus, U. Seljak, and V. Desjacques, “Optimal Constraints on Local Primordial Non-Gaussianity from the Two-Point Statistics of Large-Scale Structure,” arXiv:1104.2321 [astro-ph.CO].

[42] S. Matarrese and L. Verde, “The effect of primordial non-Gaussianity on halo bias,” *Astrophys. J.* **677** (2008) L77, arXiv:0801.4826 [astro-ph].

[43] T. Giannantonio and C. Porciani, “Structure formation from non-Gaussian initial conditions: multivariate biasing, statistics, and comparison with N-body simulations,” *Phys. Rev.* **D81** (2010) 063530, arXiv:0911.0017 [astro-ph.CO].

[44] M. Grossi, K. Dolag, E. Branchini, S. Matarrese, and L. Moscardini, “Evolution of Massive Haloes in non-Gaussian Scenarios,” *Mon. Not. Roy. Astron. Soc.* **382** (2007) 1261, arXiv:0707.2516 [astro-ph].

[45] A. Pillepich, C. Porciani, and O. Hahn, “Universal halo mass function and scale-dependent bias from N-body simulations with non-Gaussian initial conditions,” arXiv:0811.4176 [astro-ph].

[46] V. Desjacques, D. Jeong, and F. Schmidt, “Non-Gaussian Halo Bias Re-examined: Mass-dependent Amplitude from the Peak-Background Split and Thresholding,” arXiv:1105.3628 [astro-ph.CO].

[47] WMAP Collaboration, J. Dunkley *et al.*, “Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Likelihoods and Parameters from the WMAP data,” *Astrophys. J. Suppl.* **180** (2009) 306–329, arXiv:0803.0586 [astro-ph].

[48] A. Lewis, A. Challinor, and A. Lasenby, “Efficient Computation of CMB anisotropies in closed FRW models,” *Astrophys. J.* **538** (2000) 473–476, arXiv:astro-ph/9911177.

[49] M. LoVerde and K. M. Smith, “The Non-Gaussian Halo Mass Function with $f_{NL}$, $g_{NL}$ and $\tau_{NL}$,” arXiv:1102.1439 [astro-ph.CO].
[50] S. Ferraro, M. LoVerde, and K. M. Smith, “Large-scale halo stochasticity and primordial non-Gaussianity,” to appear.

[51] S. Matarrese, L. Verde, and R. Jimenez, “The abundance of high-redshift objects as a probe of non-Gaussian initial conditions,” *Astrophys. J.* 541 (2000) 10, arXiv:astro-ph/0001366.

[52] M. LoVerde, A. Miller, S. Shandera, and L. Verde, “Effects of Scale-Dependent Non-Gaussianity on Cosmological Structures,” *JCAP* 0804 (2008) 014, arXiv:0711.4126 [astro-ph].

[53] S. Chongchitnan and J. Silk, “A Study of High-Order Non-Gaussianity with Applications to Massive Clusters and Large Voids,” *Astrophys. J.* 724 (2010) 285–295, arXiv:1007.1230 [astro-ph.CO].

[54] T. Y. Lam and R. K. Sheth, “Halo abundances in the $f_{nl}$ model,” arXiv:0905.1702 [astro-ph.CO].

[55] M. Maggiore and A. Riotto, “The halo mass function from the excursion set method. III. First principle derivation for non-Gaussian theories,” arXiv:0903.1251 [astro-ph.CO].

[56] A. De Simone, M. Maggiore, and A. Riotto, “Excursion Set Theory for generic moving barriers and non-Gaussian initial conditions,” arXiv:1007.1903 [astro-ph.CO].

[57] G. D’Amico, M. Musso, J. Norena, and A. Paranjape, “An Improved Calculation of the Non-Gaussian Halo Mass Function,” arXiv:1005.1203 [astro-ph.CO].

[58] W. H. Press and P. Schechter, “Formation of galaxies and clusters of galaxies by selfsimilar gravitational condensation,” *Astrophys. J.* 187 (1974) 425–438.

[59] R. K. Sheth and G. Tormen, “Large scale bias and the peak background split,” *Mon. Not. Roy. Astron. Soc.* 308 (1999) 119, arXiv:astro-ph/9901122.

[60] M. S. Warren, K. Abazajian, D. E. Holz, and L. Teodoro, “Precision Determination of the Mass Function of Dark Matter Halos,” *Astrophys. J.* 646 (2006) 881–885, arXiv:astro-ph/0506395.

[61] S. Cole and N. Kaiser, “Biased clustering in the cold dark matter cosmogony,” *Mon. Not. Roy. Astron. Soc.* 237 (1989) 1127–1146.

[62] V. Desjacques and U. Seljak, “Signature of primordial non-Gaussianity of $\phi^3$-type in the mass function and bias of dark matter haloes,” *Phys. Rev.* D81 (2010) 023006, arXiv:0907.2257 [astro-ph.CO].

[63] S. Matarrese, F. Lucchin, and S. A. Bonometto, “A Path Integral Approach to Large Scale Matter Distribution Originated by Non-Gaussian Fluctuations,” *Astrophys. J.* 310 (1986) L21–L26.
A Barrier model calculations

In this appendix, we give details of the calculation of the halo mass function and large-scale bias (Eqs. (34)–(40)) in the barrier crossing model, to first order in $f_{NL}, g_{NL}$.

First, consider evaluation of the integrals in Eqs. (30), (31). Primordial non-Gaussianity enters the calculation by perturbing the PDFs which appear from Gaussian distributions. This perturbation can be written down explicitly using the Edgeworth expansion, which represents the PDF as a power series in cumulants. The Edgeworth expansion for the 1-variable PDF $p(\delta_M')$ is:

$$p(\delta_M') = \exp \left( \sum_{n=3}^{\infty} \frac{(-1)^n}{n!} \kappa_n(M) \sigma_M^n \frac{\partial^n}{\partial \delta_M^n} \right) \frac{1}{(2\pi)^{1/2} \sigma_M} e^{-\delta_M'^2/(2\sigma_M^2)}$$

$$= \frac{1}{(2\pi)^{1/2} \sigma_M} e^{-\delta_M'^2/(2\sigma_M^2)} \left( 1 + \frac{\kappa_3(M)}{6} H_3(\nu) + \frac{\kappa_4(M)}{24} H_4(\nu) + \cdots \right)$$

$$= \frac{1}{(2\pi)^{1/2} \sigma_M} e^{-\delta_M'^2/(2\sigma_M^2)} \left( 1 + f_{NL} \frac{\kappa_3^{(1)}(M)}{6} H_3(\nu) + g_{NL} \frac{\kappa_4^{(1)}(M)}{24} H_4(\nu) + \cdots \right)$$

where we have kept terms of first order in $f_{NL}, g_{NL}$. We can now compute $p_0$ by plugging into the definition (30):

$$p_0 = \frac{1}{2} \text{erfc} \left( \frac{\nu}{\sqrt{2}} \right) + f_{NL} \frac{\kappa_3^{(1)}(M)}{6} \frac{e^{-\nu^2/2}}{(2\pi)^{1/2}} H_2(\nu) + g_{NL} \frac{\kappa_4^{(1)}(M)}{24} \frac{e^{-\nu^2/2}}{(2\pi)^{1/2}} H_3(\nu)$$

Armed with this expression, it is easy to compute $n(M) = -2\rho_m/M(dp_0/dM)$, obtaining the form of the mass function in Eq. (34).
Moving on to the 2-variable PDF \( p(\delta_{\text{lin}}, \delta'_{M}) \), the Edgeworth expansion is:

\[
p(\delta_{\text{lin}}, \delta'_{M}) = \exp \left( \sigma_{\text{lin}} \sigma_{M} \kappa_{1,1} \frac{\partial^2}{\partial \delta_{\text{lin}} \partial \delta'_{M}} + \sum_{m,n \geq 3} (-1)^{m+n} m! n! \frac{\partial^{m+n}}{\partial \delta_{\text{lin}}^m \partial \delta'_{M}^n} \right)
\times \frac{1}{2 \pi \sigma_{\text{lin}} \sigma_{M}} \exp \left( -\frac{\delta_{\text{lin}}^2}{2 \sigma_{\text{lin}}^2} - \frac{\delta'_{M}^2}{2 \sigma_{M}^2} \right)
\]

(54)

where \( \sigma_{\text{lin}} = \langle \delta_{\text{lin}}^2 \rangle^{1/2} \) and the cumulant \( \kappa_{m,n} \) is defined by:

\[
\kappa_{m,n}(M, r) = \frac{\langle (\delta_{\text{lin}}^m (\delta'_{M})^n)_{\text{conn}} \rangle}{\sigma_{\text{lin}}^m \sigma_{M}^n}
\]

(55)

Note that the cumulant \( \kappa_n(M) \) defined previously in Eq. (7) is equal to \( \kappa_{0,n}(M, r) \).

Keeping the first few terms in the Edgeworth expansion:

\[
p(\delta_{\text{lin}}, \delta'_{M}) = \frac{1}{2 \pi \sigma_{\text{lin}} \sigma_{M}} \exp \left( -\frac{\delta_{\text{lin}}^2}{2 \sigma_{\text{lin}}^2} - \frac{\delta'_{M}^2}{2 \sigma_{M}^2} \right)
\times \left( 1 + \frac{\kappa_{1,1}(M, r)}{\sigma_{\text{lin}}} \delta_{\text{lin}} \left( \frac{\delta'_{M}}{\sigma_{M}} \right) + \frac{\kappa_{1,2}(M, r)}{2 \sigma_{\text{lin}}} \delta_{\text{lin}} H_2 \left( \frac{\delta'_{M}}{\sigma_{M}} \right)
+ \frac{\kappa_{1,3}(M, r)}{6 \sigma_{\text{lin}}} \delta_{\text{lin}} H_3 \left( \frac{\delta'_{M}}{\sigma_{M}} \right)
+ \frac{\kappa_{1,1}(M, r) \kappa_{0,4}(M)}{24 \sigma_{\text{lin}}} \delta_{\text{lin}} H_5 \left( \frac{\delta'_{M}}{\sigma_{M}} \right) + \ldots \right)
\]

(56)

we compute \( \xi_0(r) \) by integrating Eq. (31) term by term, obtaining:

\[
\xi_0(r) = \frac{\sigma_{\text{lin}} e^{-\nu^2/2}}{(2 \pi)^{1/2}} \left( \kappa_{1,1}(M, r) + \frac{\kappa_{1,2}(M, r)}{2} \nu + \frac{\kappa_{1,3}(M, r)}{6} H_2(\nu)
+ \frac{\kappa_{1,1}(M, r) \kappa_{3}(M)}{6} H_3(\nu) + \frac{\kappa_{1,1}(M, r) \kappa_{4}(M)}{24} H_4(\nu) \right)
\]

(57)

To make further progress, we convert the correlation function to a power spectrum \( P_0(k) = \int d^3r \, e^{i \mathbf{k} \cdot \mathbf{r}} \xi_0(r) \), and keep only the leading behavior of each term in the long-wavelength limit \( k \to 0 \).

\[
\int d^3r \, e^{i \mathbf{k} \cdot \mathbf{r}} \kappa_{1,2}(M, r) = \frac{1}{\sigma_{\text{lin}} \sigma_{M}^3} \int d^3\mathbf{q} d^3\mathbf{q}' W_M(q) W_M(q') \langle \delta(\mathbf{k}) \delta(\mathbf{q}) \delta(-\mathbf{q}') \rangle
\rightarrow 4 f_{NL} P(k) \frac{1}{\sigma_{\text{lin}} \alpha(k)}
\]

(58)

\[ ^9 \text{A technical point: } \sigma_{\text{lin}} \text{ is formally infinite, but it will cancel from the final results in Eqs. (36)–(40). One could make } \sigma_{\text{lin}} \text{ finite by introducing a smoothing scale } R \text{ for the matter field, and take the limit } R \to 0 \text{ at the end of the calculation.} \]

\[ ^{10} \text{The choice of terms to keep was dictated by the following considerations. Only terms with precisely one } \delta_{\text{lin}} \text{ derivative will give nonzero contributions to the integral } \int_{-\infty}^{\infty} d\delta_{\text{lin}} p(\delta_{\text{lin}}, \delta'_{M}) \text{ appearing in } \xi_{nh}(r), \text{ so we have only kept these terms. (Terms with two or more derivatives would contribute to the halo-halo correlation function } \xi_{hh}(r), \text{ so they may be relevant for halo stochasticity.) We have also omitted terms whose leading contribution is second-order or higher in } f_{NL} \text{ and } g_{NL}. \]

23
\[ \int d^3r e^{i \mathbf{k} \cdot \mathbf{r}} \kappa_{1,3}(M, r) = \frac{1}{\sigma_{\text{lin}} \sigma_M^3} \int \frac{d^3q d^3q' d^3q''}{(2\pi)^9} W_M(q) W_M(q') W_M(q'') \times \langle \delta(\mathbf{k}) \delta(\mathbf{q}) \delta(\mathbf{q}') \delta(-\mathbf{q}'') \rangle_{\text{conn}} \\
\rightarrow \frac{18 g_{\text{NL}} P(k)}{\sigma_{\text{lin}} \sigma_M^3} \int \frac{d^3q d^3q'}{(2\pi)^6} W_M(q) W_M(q') W_M(|\mathbf{q} + \mathbf{q}'|) \times \frac{P(q) P(q') \alpha(|\mathbf{q} + \mathbf{q}'|)}{\alpha(q) \alpha(q')} \\
= \frac{3 g_{\text{NL}} \kappa_3^{(1)}(M)}{\sigma_{\text{lin}}} \left( \frac{P(k)}{\alpha(k)} \right) \] (59)

where \( \rightarrow \) denotes the \( k \to 0 \) limit, and we have used Eq. (10) to simplify the last line. Putting this together, we find the following expression for \( P_0(k) \) in the \( k \to 0 \) limit:

\[ P_0(k) = \frac{e^{-\nu^2/2}}{(2\pi)^{1/2}} \left[ \frac{P(k)}{\sigma_M} \left( 1 + f_{\text{NL}} \frac{\kappa_3^{(1)}(M)}{6} H_3(\nu) + g_{\text{NL}} \frac{\kappa_4^{(1)}(M)}{24} H_4(\nu) \right) \right. \\
\left. + 2\nu f_{\text{NL}} \frac{P(k)}{\alpha(k)} + \kappa_3^{(1)}(M) \frac{H_2(\nu)}{2} g_{\text{NL}} \frac{P(k)}{\alpha(k)} \right] \] (60)

The halo bias in a narrow mass range is given by the derivative:

\[ b(k) = \frac{dP_0(k)/dM}{(dp_0/dM) P(k)} + 1 \] (61)

where the \( +1 \) converts Lagrangian to Eulerian bias. Plugging in the forms of \( p_0, P_0 \) in Eqs. (53), (60), a long but straightforward calculation now gives the halo bias in the form given in the text (Eqs. (35)-(40)).