Primordial Non-Gaussianity and Extreme-Value Statistics of Galaxy Clusters

Sirichai Chongchitnan* and Joseph Silk
Department of Physics, University of Oxford and
Beecroft Institute for Particle Astrophysics and Cosmology,
Denys Wilkinson Building, 1 Keble Road, Oxford, OX1 3RH, United Kingdom.

What is the size of the *most* massive object one expects to find in a survey of a given volume? In this paper, we present a solution to this problem using Extreme-Value Statistics. We calculate the probability density function (pdf) of extreme-mass clusters in a survey volume, and show how primordial non-Gaussianity shifts the peak of this pdf. We compare the extreme-value pdfs using various mass functions and find significant differences at high redshift. Applying our formalism to the recently reported massive high-redshift cluster XMMUJ0044.0-2-33, we find that $f_{\rm NL} \simeq 360$ is required to explain its existence as the most massive cluster observed in the redshift range $1.6 < z < 2.2$. Finally, we argue that the probability distribution of extreme-mass clusters belongs to the so-called Fréchet family of distribution, regardless of the presence of non-Gaussianity.

I. INTRODUCTION

The statistics of the primordial seeds that grew into the observed large-scale structures holds a wealth of information about the physics of the primordial Universe. In the simplest models of inflation, the primordial density fluctuations obey an almost Gaussian statistics (see [1] for a review). Tiny deviations from Gaussianity may be quantified, amongst other ways, by the 'local' non-Gaussianity parameter, $f_{\rm NL}$, defined via the expansion of the non-linear Newtonian potential

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

where ϕ is a Gaussian random field. This form of non-Gaussianity arises in simple models of single and multi-field inflation [2, 3, 4] as well as some curvaton models [5, 6]. Observational constraints on $f_{\rm NL}$ from the cosmic microwave background (CMB) anisotropies are currently consistent with $f_{\rm NL} = 32 \pm 42$ (2σ) [7]. However, if $f_{\rm NL}$ is in fact much smaller, its effects on the CMB would be difficult to extract and distinguished from non-Gaussianity arising from secondary sources such as gravitational lensing and instrumental noise [8, 9].

The statistics of large-scale structures offers a complementary probe of non-Gaussianity on much smaller scales than the CMB. Indeed, it is possible that $f_{\rm NL}$ measured on Gpc scales may be quite different from that measured on Mpc scales. Some hints at a possible scale-dependence of non-Gaussianity come from the numerous observations of massive high-redshift clusters [10, 11, 12, 13, 14] which may be in greater abundances than expected from a Gaussian statistics. Some authors have concluded that the level of non-Gaussianity on Mpc scale required to explain the existence of certain rare clusters is $f_{\rm NL}$ = a few $\times 10^2$ [15, 16]. In contrast, some have argued that these claims are based on misinterpretation of data, and that the occurrences of these rare objects are in fact consistent with a Gaussian statistics [17, 18, 19].

In this work, we offer our opinion on this debate by presenting an alternative approach to calculating the probability of observing rare objects based on extreme-value statistics. We begin by asking: what is the probability distribution of the *most* massive clusters found within a given volume at a given redshift range? Our technique relies on a basic application of the so-called void probability distribution introduced by White [20]. This approach was successfully used to study the abundances of massive clusters given a Gaussian initial condition in [21, 22]. In this work, we extend the groundwork laid by these authors to study the effect of $f_{\rm NL}$ on the distribution of extreme-mass objects. For other cosmological applications of extreme-value theory, see [23, 24, 25, 26, 27].

One of the main differences between previous calculations and ours is that we take into account the clustering of rare objects, as parametrized by the bias. It is well known that the presence of non-Gaussianity introduces a strong scale dependence in the bias [28]. For the calculation of probabilities, it will be necessary to calculate the bias in real space. Valageas [29, 30] showed how the real-space bias is related to $f_{\rm NL}$ and we shall make use of his formalism in this paper.

*Electronic address: siri@astro.ox.ac.uk

II. THE PRIMORDIAL DENSITY FLUCTUATIONS

In this section, we introduce the parameters needed to describe the primordial density fluctuations statistically. Some of our present conventions (such as the Fourier transform and the moments of the density fluctuations) slightly differ from our earlier work [31, 32]. In particular, smoothing by a window function will be kept explicit, in contrast with other work in which overdensities are defined to be implicitly smoothed.

Let ρ_c, ρ_b, ρ_r, ρ_Λ be the time-dependent energy densities of cold dark matter, baryons, radiation and dark energy. Let $\rho_m = \rho_c + \rho_b$. We define the density parameter for species i as

$$\Omega_i \equiv \frac{\rho_i(z=0)}{\rho_{\rm crit}}, \tag{2}$$

where $\rho_{\rm crit}$ is the critical density defined as $\rho_{\rm crit} \equiv 3H_0^2/8\pi G$. The Hubble constant, H_0, is parametrized by h via the usual formula $H_0 \equiv 100h$ km s^{-1}Mpc^{-1}. Results from a range of astrophysical observations are consistent with $h \simeq 0.7$, $\Omega_c \simeq 0.23$, $\Omega_b \simeq 0.046$, $\Omega_r \simeq 8.6 \times 10^{-5}$ and $\Omega_\Lambda = 1 - \Omega_m - \Omega_r$ (see *e.g.* [7, 33]).

The density fluctuation field, δ, is defined at redshift z as

$$\delta(\mathbf{x}, z) \equiv \frac{\rho_m(\mathbf{x}, z) - \langle \rho_m(z) \rangle}{\langle \rho_m(z) \rangle}, \tag{3}$$

where $\langle \rho_m \rangle$ is the mean matter energy density. To make the notation less cumbersome, we shall write $\delta(\mathbf{x})$ to mean $\delta(\mathbf{x}, z)$. The Fourier decomposition of $\delta(\mathbf{x})$ is given by

$$\delta(\mathbf{x}) = \int d\mathbf{k}\, \delta(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{4}$$

The gravitational Newtonian potential, Φ, is related to the density fluctuation by the cosmological Poisson equation

$$\delta(\mathbf{k}) = \mathcal{A}(k, z)\Phi(\mathbf{k}), \tag{5}$$

$$\mathcal{A}(k, z) \equiv \frac{2}{3\Omega_m} \left(\frac{k}{H_0}\right)^2 T(k_{\rm EH}) D(z), \tag{6}$$

where T is the transfer function and D is the linear growth factor calculated using the fitting formula of [34, 35] with $D(0) \approx 0.76$. We follow the approach outlined in [36] and use the transfer function

$$T(k) = \frac{\ln[1 + (0.124k)^2]}{(0.124k)^2} \left[\frac{1 + (1.257k)^2 + (0.4452k)^4 + (0.2197k)^6}{1 + (1.606k)^2 + (0.8568k)^4 + (0.3927k)^6}\right]^{1/2}. \tag{7}$$

In addition, we also incorporate the baryonic correction of Eisenstein and Hu [37], whereby the transfer function is evaluated at

$$k_{\rm EH} = \frac{k\Omega_r^{1/2}}{H_0\Omega_m} \left[\alpha + \frac{1-\alpha}{1 + (0.43ks)^4}\right]^{-1}, \tag{8}$$

with

$$\alpha = 1 - 0.328 \ln(431\Omega_m h^2) \frac{\Omega_b}{\Omega_m} + 0.38 \ln(22.3\Omega_m h^2) \left(\frac{\Omega_b}{\Omega_m}\right)^2,$$

and

$$s = \frac{44.5 \ln(9.83/\Omega_m h^2)}{\sqrt{1 + 10(\Omega_b h^2)^{3/4}}} \text{ Mpc}.$$

The power spectrum, $P(k)$, can be defined via the two-point correlation function in Fourier space as

$$\langle \delta(\mathbf{k_1}), \delta(\mathbf{k_2}) \rangle = \delta_D(\mathbf{k_1} + \mathbf{k_2}) P(k), \tag{9}$$

where δ_D is the 3-dimensional Dirac delta function. In linear perturbation theory, it is usually assumed that inflation laid down a primordial spectrum of the form $P(k) \propto k^{n_s}$, where n_s is the scalar spectral index (assumed to be 0.96 in this work).

We define the dimensionless power spectrum, $\mathcal{P}(k)$, as

$$\mathcal{P}(k) \equiv k^3 P(k) \propto (\mathcal{A}(k,z))^2 \left(\frac{k}{H_0}\right)^{n_s-1}. \tag{10}$$

Consequently, the variance of density fluctuations smoothed on scale R can be written as

$$\widetilde{\sigma}_R^2 = 4\pi \int_0^\infty \frac{dk}{k} W^2(kR)\mathcal{P}(k). \tag{11}$$

It will also be useful to define

$$\sigma_R^2 \equiv \frac{1}{4\pi}\widetilde{\sigma}_R^2, \tag{12}$$

which is a more common definition of the variance in the literature. We shall choose W to be the spherical top-hat function of radius R. In Fourier space, we have

$$W(kR) = 3\left[\frac{\sin(kR)}{(kR)^3} - \frac{\cos(kR)}{(kR)^2}\right]. \tag{13}$$

The mass, M, of matter enclosed by a top-hat window of radius R is given by

$$M \equiv \frac{4}{3}\pi R^3 \rho_m \approx 1.16 \times 10^{12} \left(\frac{R}{h^{-1}\mathrm{Mpc}}\right)^3 \; h^{-1} M_\odot. \tag{14}$$

With the above relation, the smoothed variance, σ_R, can be equivalently expressed as σ_M. Finally, the normalization of $\mathcal{P}(k)$ is such that

$$\sigma_8 \equiv \sigma(R = 8h^{-1}\mathrm{Mpc}, z = 0) = 0.8. \tag{15}$$

III. CLUSTER NUMBER COUNTS

The mean number density, n, of objects with mass greater than m, at redshift z can be calculated by

$$n(>m, z) = \int_m^\infty \frac{dn}{dM}\, dM, \tag{16}$$

where dn/dM is the differential number density. In the presence of local non-Gaussianity, Matarrese, Verde and Jimenez [38] used a saddle-point expansion to derive a correction factor for dn/dM of the form

$$\mathcal{R} \;=\; \exp\left(\frac{S_3\delta_c^3}{6\sigma_M^2}\right)\left[\frac{\delta_c^2}{6\Delta}\cdot\frac{dS_3}{d\ln\sigma_M} + \Delta\right], \tag{17}$$

$$\Delta \;\equiv\; \sqrt{1 - \frac{\delta_c S_3}{3}}, \tag{18}$$

where the third cumulant, S_3, is given by $S_3 = \langle\delta^3\rangle$ (assuming zero mean). S_3 can be calculated either from a 3-dimensional integral

$$S_3(M) = \frac{6 f_{R;RR}(0)}{\widetilde{\sigma}_M^4}, \tag{19}$$

with $f_{R;RR}$ defined in equation (33), or the fitting formula [31]

$$S_3(M) = \frac{3.15 \times 10^{-4} f_{\mathrm{NL}}}{\sigma_M^{0.838}}, \tag{20}$$

which gives the scaling $dS_3/d\ln\sigma_M \simeq -0.838 \times S_3$. There are alternative forms of the correction factor, $\mathcal{R}$, given by LoVerde *et al.* [39] based on a low-order Edgeworth expansion, and by Paranjape *et al.* [40] based on resumming terms in the saddle-point expansion of the mass function. We tested both of these alternative corrections and found

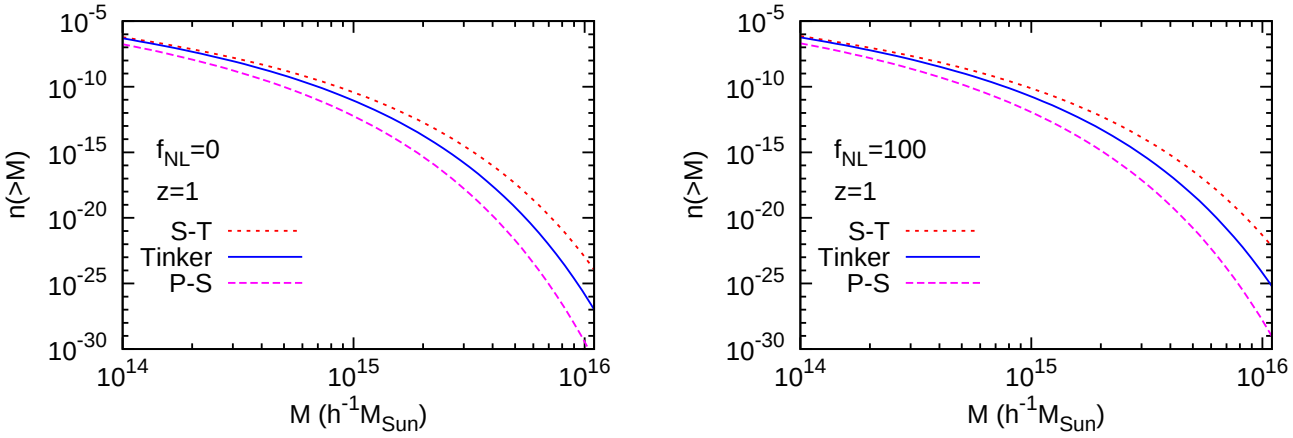

FIG. 1: The number density, n, of objects above mass M with $f_{\rm NL}=0$ (left) and 100 (right) calculated at $z=1$ using the Press-Schechter (P-S), Sheth-Tormen (S-T) and Tinker *et al.* mass functions. With $f_{\rm NL}=100$, the number density increases by roughly an order of magnitude at the high-mass end compared with $f_{\rm NL}=0$.

that, in the range of parameters used in this paper, there are only small differences between the various prescriptions and our main results are unaffected by the choice of the correction factor. In the rest of this work, we shall use only the MVJ correction factor (see [41] for a comparison between the correction factors).

In summary, we shall consider the non-Gaussian differential abundance of the form

$$\frac{dn}{dM}=\mathcal{R}\times F(\nu)\frac{\rho_m}{M}\frac{d\ln\sigma^{-1}}{dM}, \tag{21}$$

where $\nu\equiv\delta_c/\sigma_M$ and $F(\nu)$ is one of the following three standard mass-functions

Press-Schechter [42]
$$F_{\rm PS}=\sqrt{\frac{2}{\pi}}\nu e^{-\nu^2/2}, \tag{22}$$

Sheth-Tormen [43]
$$F_{\rm ST}=0.322\sqrt{\frac{2a}{\pi}}\nu\exp\left(-\frac{a\nu^2}{2}\right)\left[1+\left(a\nu^2\right)^{-0.3}\right],\quad a=0.707, \tag{23}$$

Tinker *et al.* [44, 45]
$$F_{\rm Tinker}=0.368\left[1+(\beta\nu)^{-2\phi}\right]\nu^{2\eta+1}e^{-\gamma\nu^2/2}, \tag{24}$$

$$\beta=0.589(1+z)^{0.2},\phi=-0.729(1+z)^{-0.08},$$
$$\eta=-0.243(1+z)^{0.27},\gamma=0.864(1+z)^{-0.01}. \tag{25}$$

Figure 1 shows the number density, n, of objects with mass above M calculated at $z=1$ using the mass functions given above. We see that the Sheth-Tormen gives the highest number density, followed by the Tinker and the Press-Schechter mass functions. Increasing $f_{\rm NL}$ to 100 (right panel) increases the number density at the high-mass end by roughly an order of magnitude. See *e.g.* [46, 47] for more comparisons between various mass functions.

IV. BIAS

In the seminal work of Dalal *et al.*[28], it was shown quantitatively how non-Gaussianity gives rise to characteristic changes in the clustering of density peaks corresponding to rare objects. At leading order, it is common to define the bias in Fourier space as the ratio of the power-spectra

$$b^2(k)=\frac{P_{\rm halo}(k)}{P_m(k)}, \tag{26}$$

which represents the amplitude at which density peaks ($P_{\rm halo}$) trace the underlying dark matter distribution (P_m). The Fourier space formalism was used by a majority of papers on non-Gaussian bias (*e.g.* [41, 48, 49, 50, 51]).

However, an arguably more intuitive measure of the bias is in real space, where the density fluctuation in peaks (*i.e.* luminous objects) is expressed as a non-linear function of the local dark-matter density fluctuation. On linear scales, the bias is given by the ratio of the correlation functions [52, 53]

$$b^2(r) = \frac{\xi_{\rm pk}(r)}{\xi(r)}, \tag{27}$$

where r is the comoving length in Eulerian space (throughout this work quantities with a subscript 'pk' are associated with density peaks). The correlation function, ξ, is defined as

$$\xi(\mathbf{x_1}, \mathbf{x_2}) = \langle \delta(\mathbf{x_1}), \delta(\mathbf{x_2}) \rangle, \quad r = |\mathbf{x_1} - \mathbf{x_2}|. \tag{28}$$

In the linear regime, we can write

$$\xi(r) = 4\pi \int_0^\infty \frac{dk}{k}\, \mathcal{P}(k) j_0(kr), \tag{29}$$

where $j_0(x) = \sin x/x$. The real-space bias tells us directly about the clustering amplitude of density peaks separated by distance r. We shall refer to r as the separation length.

Unfortunately, when comparing (26) and (27), we see that the real-space bias, $b(r)$, and the Fourier-space bias, $b(k)$, are not simply related via a Fourier transform but rather a complicated convolution. In [32], we avoided this problem by interpreting (27) as a ratio of joint probabilities of finding overdensities at two points distance r apart, and then applying a bivariate Edgeworth expansion. Due to the algebraic nature of the Edgeworth expansion, this technique was readily applied to non-Gaussianity parametrized by the cubic order parameter, $g_{\rm NL}$, but surprisingly the application is much less straightforward for $f_{\rm NL}$.

An alternative method for calculating the real-space bias in the presence of $f_{\rm NL}$ was presented by Valageas [29, 30] in which he showed that analytic calculations could be made as long as the separation length is sufficiently large. In this work, we shall follow this formalism, of which we give a simplified account here.

A crucial element in the real-space approach is the mapping between the separation length, s, in Lagrangian coordinates (associated with linear density fluctuations) and that in Eulerian coordinates (associated with non-linear fluctuations). This relation is given by

$$s \simeq r\left(1 + \frac{2\delta_R(r)}{3}\right), \tag{30}$$

accurate at large distances where $\delta_R(r) \ll 1$. Here $\delta_R(r)$ can be interpreted as the radial profile of the linear density contrast from the centre of the halo. The profile is given by

$$\delta_R(r) = \frac{\delta_c}{\widetilde{\sigma}_R^2}\, \widetilde{\sigma}_{R,0}^2(r) + \frac{\delta_c^2}{\widetilde{\sigma}_R^4}\, \left[f_{0;RR}(r) + 2g_{R;0R}(r) - 3\frac{\widetilde{\sigma}_{R,0}^2(r)}{\widetilde{\sigma}_R^2} f_{R;RR}(0) \right]. \tag{31}$$

In this expression, the functions $\widetilde{\sigma}_{R_1,R_2}(r)$, $f_{R;R_1R_2}(r)$ and $g_{R;R_1R_2}(r)$ are defined by the following integrals[1]

$$\widetilde{\sigma}_{R_1,R_2}^2(r) = 4\pi \int_0^\infty \frac{dk}{k} \mathcal{P}(k) W(kR_1) W(kR_2) j_0(kr), \tag{32}$$

$$f_{R;R_1R_2}(r) = 8\pi^2 D(0) f_{\rm NL} \int_0^\infty \frac{dk_1}{k_1} \mathcal{P}(k_1) W(k_1R_1) \int_0^\infty \frac{dk_2}{k_2} \mathcal{P}(k_2) W(k_2R_2) \int_{-1}^1 d\mu W(kR) \frac{\mathcal{A}(k)}{\mathcal{A}(k_1)\mathcal{A}(k_2)} j_0(kr), \tag{33}$$

$$g_{R;R_1R_2}(r) = 8\pi^2 D(0) f_{\rm NL} \int_0^\infty \frac{dk_1}{k_1} \mathcal{P}(k_1) W(k_1R_1) j_0(k_1r) \int_0^\infty \frac{dk_2}{k_2} \mathcal{P}(k_2) W(k_2R_2) \int_{-1}^1 d\mu W(kR) \frac{\mathcal{A}(k)}{\mathcal{A}(k_1)\mathcal{A}(k_2)}, \tag{34}$$

where μ is the cosine of the angle between $\mathbf{k_1}$ and $\mathbf{k_2}$, and $k = \sqrt{k_1^2 + k_2^2 + 2k_1k_2\mu}$.

With these definitions, Valageas showed via a saddle-point expansion that the bias for objects smoothed over scale R is given by

$$b_R^2(r) = \frac{1}{\widetilde{\sigma}_{0,0}^2(r)} \left[(1 + \delta_R(s)) e^{\Delta(s)} - 1 \right], \tag{35}$$

[1] In this paper $f_{\rm NL}$ is defined in the 'LSS' convention. The 'CMB' convention, as used in [29], satisfies $f_{\rm NL}^{\rm CMB} = D(0) f_{\rm NL}^{\rm LSS}$.

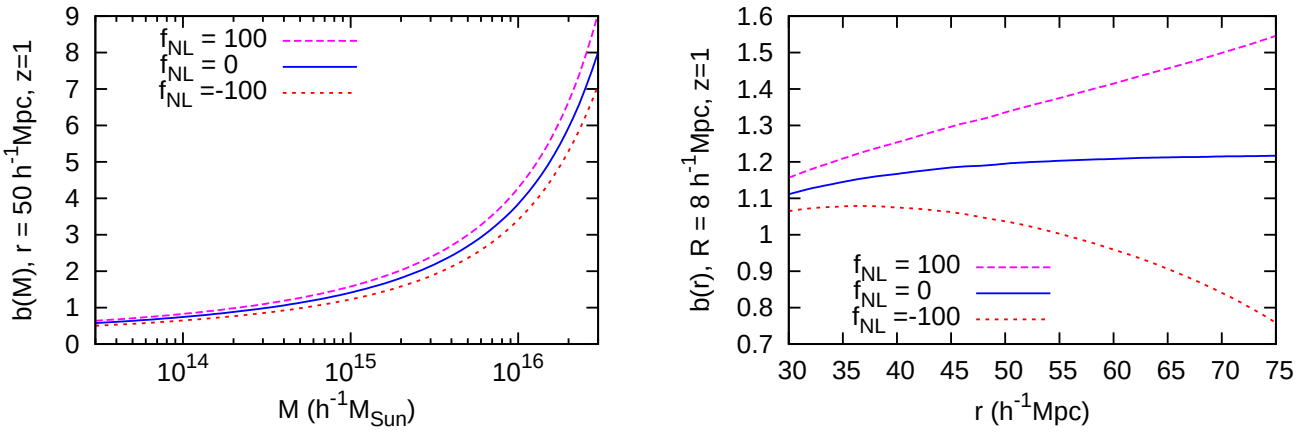

FIG. 2: The effect of $f_{\rm NL}$ on the real-space bias, b, at $z = 1$. In the panel on the left, the bias is shown as a function of smoothing mass-scale, M, with separation length $r = 50h^{-1}$Mpc. The other panel shows the effect of varying r with $R = 8h^{-1}$Mpc, illustrating the scale-dependence of the bias when $f_{\rm NL} = \pm 100$.

where

$$\Delta(s) = \frac{\widetilde{\sigma}^2_{R,R}(s)\delta_c^2}{u\widetilde{\sigma}^2_R} + \frac{2\delta_c^3}{u^3}\left[f_{R;RR}(s) + 2g_{R;RR}(s) + \left(1 - \frac{u^3}{\widetilde{\sigma}^6_R}\right) f_{R;RR}(0)\right], \tag{36}$$

$$u = \widetilde{\sigma}^2_R + \widetilde{\sigma}^2_{R,R}(s). \tag{37}$$

Figure 2 shows the real space bias for a range of smoothing mass-scale M and separation length r. Keeping r fixed and varying M (panel on the left), non-Gaussianity shifts $b(M)$ up or down (depending on the sign of $f_{\rm NL}$). On the other hand, keeping M fixed and varying r, we see that nonzero $f_{\rm NL}$ introduces a striking scale dependence on $b(r)$ ($b(r)$ is roughly constant on large scale if $f_{\rm NL} = 0$). This scale-dependence is similar to that seen in [32] for $g_{\rm NL}$.

V. EXTREME-VALUE DISTRIBUTIONS

In this section, we present the calculation of the distribution[2] of extreme-mass clusters. The necessary ingredients are the non-Gaussian number density and real-space bias calculated in the previous sections.

A. Distribution function

White [20] derived the following expression for the cumulative probability that a region of volume V contains no object of mass M and above

$$P(M) = \exp\left[\sum_{k=1}^{\infty} \frac{(-n)^k}{k!}\left(\prod_{i=1}^{k}\int_V \mathbf{dx_i}\right)\xi_k^{\rm pk}(\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_k})\right], \tag{38}$$

where $\xi_1^{\rm pk} \equiv 1$, n is given by (16) and $\xi_k^{\rm pk}$ is the k-point correlation function of density peaks in V. As in [21, 22], we shall at times refer to V as a 'patch'. If we take V to be a sphere of comoving radius L, the volume-averaged correlation then simplifies to the cumulant (connected moment) smoothed by a top-hat window of radius L [54, 55]

$$\left(\prod_{i=1}^{k}\int_V \frac{\mathbf{dx_i}}{V}\right)\xi_k^{\rm pk}(\mathbf{x_1}, \mathbf{x_2}, \ldots \mathbf{x_k}) = \langle\delta_{\rm pk}^k\rangle_c(L) = (\sigma^{\rm pk})_L^{2k-2} S_k^{\rm pk}(L). \tag{39}$$

[2] We use the word 'distribution' in the strict sense, referring to the cumulative distribution and not the pdf.

The cumulants for density peaks have been calculated in the context of hierarchical structure formation with Gaussian initial condition [54, 56, 57, 58]. In the presence of non-Gaussianity, however, the perturbation theory required to calculate the cumulants for galaxy clusters becomes much more difficult (see *e.g.* [54, 59, 60]).

To make analytic progress, we shall assume that the cumulant for density peaks and that for dark matter are related, at leading order, by a power-law scaling of the bias

$$S_k^{\rm pk} \approx b^k S_k, \tag{40}$$

where b is calculated using the real-space approach outlined in section IV (similarly, we take $\sigma^{\rm pk} \approx b\sigma$). The bias depends on the mass threshold, M, and the separation length, r, which should be chosen to reflect the clustering scale of density peaks within the spherical patch. For density peaks in the form of galaxy clusters, it is reasonable to choose $r \sim 50{-}100h^{-1}$Mpc, corresponding to typical separation of observed clusters. It is important to note the relative sizes of the length scales involved, namely, the smoothing scale, R, corresponding to the mass threshold, the separation length, r, and the patch size, L. They must obey the inequality

$$R < r < L, \tag{41}$$

for the calculations to be valid. For example, for a cluster of mass $10^{16}h^{-1}M_\odot$, the lower bound R in the above inequality is $18h^{-1}$Mpc, in which case it is sufficient to take, for instance, $r \sim 50h^{-1}$Mpc and $L \sim 100h^{-1}$Mpc.

When the series (38) is truncated at three terms, only $S_2 = 1$ and S_3 (calculated using (19) or (20)) are needed. Altogether, we find the following approximation for the extreme-value distribution

$$\ln P(M) = -nV + \frac{b^4}{2}(nV\sigma_L)^2 - \frac{b^7}{6}(nV)^3\sigma_L^4 S_3(L), \tag{42}$$

valid in a sufficiently large volume where only low-order correlations are important.

The distribution function (42) is not easy to analyse and it is much more intuitive to look instead at its derivative, *i.e.* the probability density function. Nevertheless, the shape of the distribution function holds valuable statistical information to which we shall return when we consider the extremal-type distributions in the final section.

B. PDF of extreme-mass objects

We can obtain the probability density function (pdf) for the most massive objects expected in a volume by differentiating the distribution function (42) with respect to M, noting that the only dependence on M is in the number density, n, and the bias, b. The result is

$$p(M) = \frac{dP}{dM} = VP(M)\left[\frac{dn}{dM}\left(-1 + nVb^4\sigma_L^2 - \frac{1}{2}n^2V^2b^7\sigma_L^4 S_3(L)\right) + n^2Vb^3\sigma_L^2\frac{db}{dM}\left(2 - \frac{7}{6}nVb^3\sigma_L^2 S_3(L)\right)\right]. \tag{43}$$

Thus, we see explicitly that the pdf of extreme-mass objects not only depends on the bias, but also on its scale dependence, db/dM, which changes significantly in the presence of non-Gaussianity.

C. $f_{\rm NL}$ and extreme-value pdf

The main results of this paper are shown in Figure 3. The panels show the probability density function (43) for the three mass functions at redshift $z = 1, 1.6$ and 3 (corresponding to the left, middle and right column) with $f_{\rm NL} = 200, 100$ and 0 (top, middle and bottom row respectively). The survey volume is taken to be a sphere of radius $100h^{-1}$Mpc and the bias is calculated at $r = 50h^{-1}$Mpc. To display the correct scaling on the horizontal log scale, we plot $dP/d\log M_{\rm max}$ on the vertical axis whilst the actual probability value is $dP/dM_{\rm max}$, which does not exceed unity. From these graphs, we make the following observations:

(a) Going from the bottom row to the top, we see that increasing $f_{\rm NL}$ shifts the peak of the pdf to the right and generally broadens the pdf. This has the effect of increasing the mass of the most probable extreme objects in a given volume.

(b) Going from the first column to the third, we see that at higher redshifts, the pdfs are more peaked and the peaks are located at lower $M_{\rm max}$.

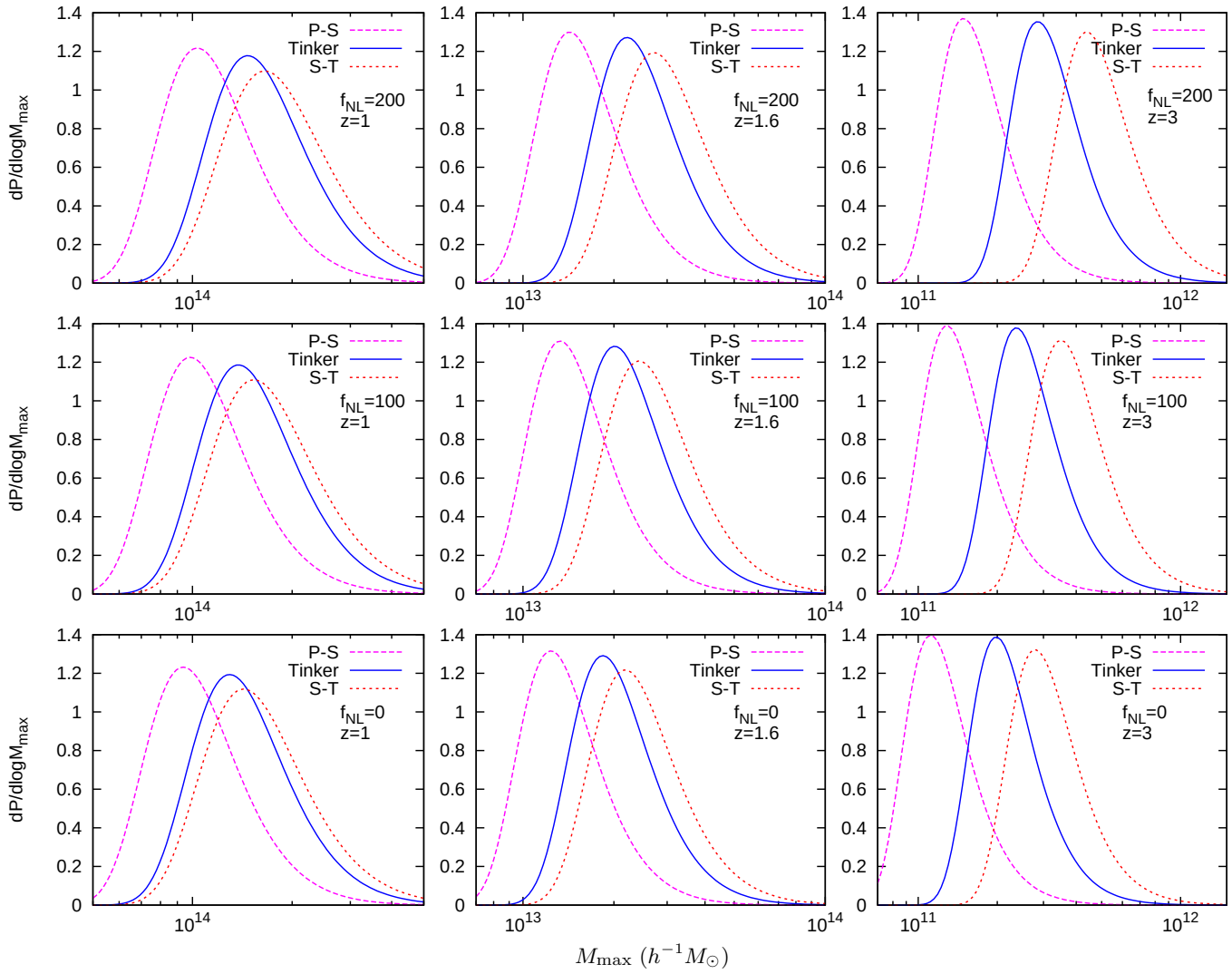

FIG. 3: The probability density function of the maximum mass, $M_{\rm max}$, of objects in a spherical volume of radius $L = 100h^{-1}$ Mpc. In each panel, the mass functions used are Press-Schechter (dashed/magenta), Tinker (solid/blue) and Sheth-Tormen (dotted/red). *Top row:* $f_{\rm NL} = 200$ with $z = 1, 1.6$ and 3 (from left to right). The pdf at the same redshifts are shown for $f_{\rm NL} = 100$ (middle row) and $f_{\rm NL} = 0$ (bottom row). The non-Gaussian effects are most visible in the third column in which the peaks can be seen to move to higher $M_{\rm max}$ with increasing $f_{\rm NL}$.

(c) The Sheth-Tormen mass function predicts the largest mass of extreme objects, followed by the Tinker and the Press-Schechter mass functions. This is a consequence of their predicted number densities as seen in Figure 1.

(d) The differences between the mass-functions become much more pronounced at high redshifts. In the third column, we see a clear separation of the peaks for different mass functions, with non-Gaussianity further enhancing the differences.

In addition, the effect of varying the patch radius, L, is shown in Figure 4, in which L varies from 100 to 1000 h^{-1}Mpc for a range of redshifts ($z = 0$, 1, 1.6 and 3). By increasing L, the peak of the pdf shifts significantly to higher $M_{\rm max}$ similar to increasing $f_{\rm NL}$. However, unlike the broadening effect of $f_{\rm NL}$, the pdf becomes more peaked with increasing L. This is expected as $\sigma \to 0$ for an infinitely large volume. By increasing the redshift, the pdfs shift to a lower range of $M_{\rm max}$ as before, whilst there does not appear to be a general trend on the change in the heights of the peaks.

Strictly speaking, for a patch size as large as a Gpc, the redshift variation within the patch must be taken into account (as emphasised in [21]). This requires replacing the number density, bias and cumulants by their average within a comoving volume. We shall demonstrate this in the next section.

We also repeated the calculation of the pdf for a range of separation length r. We find that as long as r is small compared to the patch size, the effect on the pdf is negligible. For example, with $L = 500h^{-1}$Mpc, varying r between

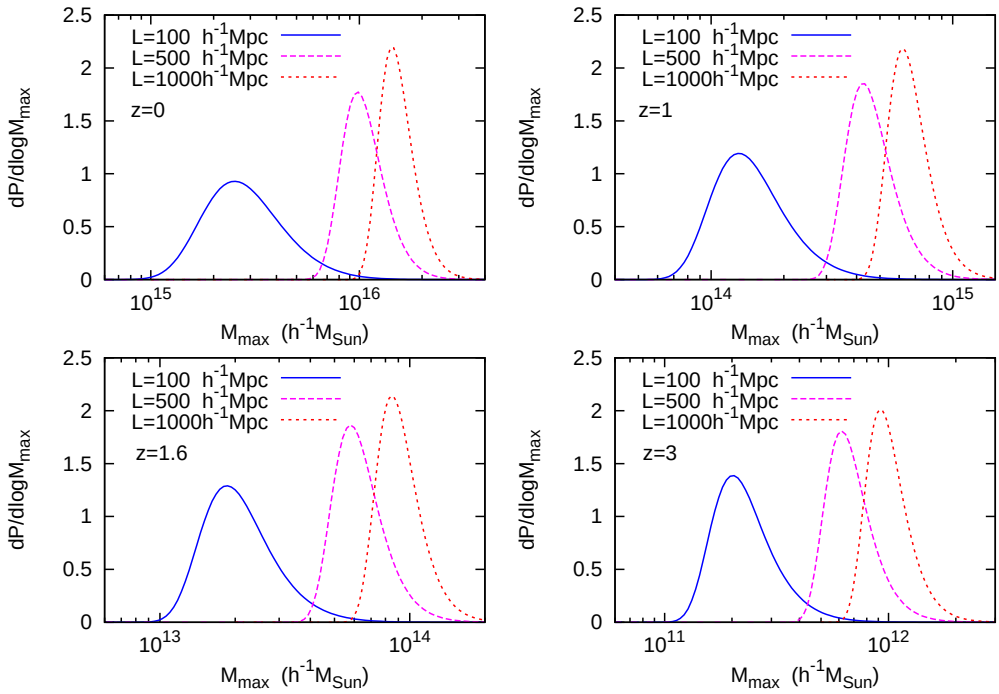

FIG. 4: The extreme-value pdf as described in Figure 3 for patch sizes $L = 100h^{-1}$Mpc (solid/blue), $500h^{-1}$Mpc (dashed/magenta) and $1000h^{-1}$Mpc (dotted/red). The patches are centred at redshifts (from top left) 0, 1 ,1.6 and 3 . The Tinker mass function is used, with $z = 1$, $r = 50h^{-1}$Mpc and $f_{\rm NL} = 10$. Noting the changes in the horizontal scale, we see that the location of the peak is clearly very sensitive to a change in L.

50 and 100 h^{-1}Mpc produces virtually the same pdfs.

VI. APPLICATION TO A MASSIVE CLUSTER AT $z = 1.6$

Santos *et al.*[13] recently reported the discovery of a cluster XMMUJ0044.0-2-33 (hereafter XMMUJ0044) at $z = 1.579$, detected in the X-ray data of the XMM-Newton telescope and later followed up spectroscopically. The cluster mass was estimated to be $\sim 3.5-5\times10^{14}M_\odot$, far greater than the previous X-ray cluster of mass 5.7×10^{13} at $z = 1.62$ reported by Tanaka, Finoguenov and Ueda [14]. We shall now use extreme-value statistics to study the probability of finding XMMUJ0044 as the maximum-mass cluster. In particular, we ask, a) is the existence of XMMUJ0044 consistent with $f_{\rm NL} = 0$? b) what is the value of $f_{\rm NL}$ needed to make this cluster the *most* probable extreme-mass cluster expected from such a survey?

The XDCP survey covers a sky area of 80 deg^2 [13] and redshift up to ~ 2.2. We adopt a conservative value of $f_{\rm sky} = 1$ since, as pointed out in [19], the value of $f_{\rm sky}$ should take into account all surveys that have covered the particular part of the sky, regardless of whether a positive detection is reported.

We are interested in the probability that XMMUJ0044 (taken conservatively to be of mass $3.5\times10^{14}M_\odot$) is the most massive cluster in the redshift range $1.6 \lesssim z \lesssim 2.2$. To account for the redshift variation within the patch, we replace the number density, n, in equation (43) by the redshift average

$$\langle n\rangle = \frac{1}{V}\int_{1.6}^{2.2} dz\; n(z)\frac{dV}{dz}, \tag{44}$$

where the survey volume, V, satisfies

$$\frac{dV}{dz} = f_{\rm sky}\frac{4\pi}{H(z)}\left(\int_0^z \frac{dz'}{H(z')}\right)^2, \tag{45}$$

$$H(z) \approx H_0\left[\Omega_m(1+z)^3+\Omega_\Lambda\right]^{1/2}. \tag{46}$$

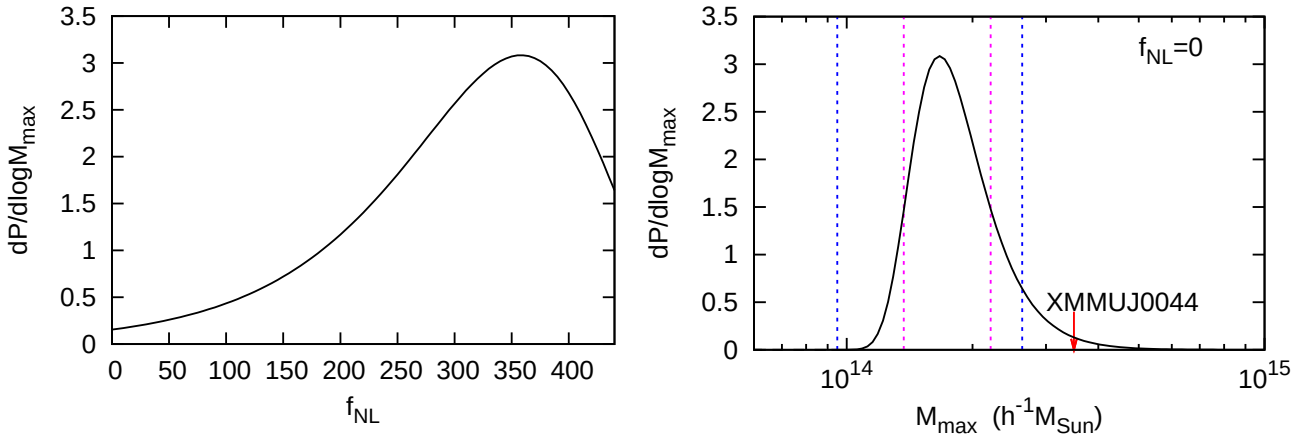

FIG. 5: *Left:* The probability that XMMUJ0044.0-2-33 is the most massive cluster in a survey (as specified in the text) plotted as a function of $f_{\rm NL}$. Using the Sheth-Tormen mass function, $f_{\rm NL} \approx 360$ is needed to make this cluster the most probable extreme-mass object in the survey. *Right:* The extreme-value pdf with $f_{\rm NL} = 0$, using the Sheth-Tormen mass function. The vertical dashed lines around peaks indicate 1σ and 2σ intervals. The cluster lies 4σ away from the mean.

Integrating (45) gives the patch volume, V, (and hence L) as required in (43). Similarly, we replace the bias and the cumulants by their respective redshift averages. In practice, we evaluate the integral (44) by an average over redshifts with bin size $\Delta z = 0.1$.

We shall use the Sheth-Tormen mass function since, as seen in the previous section, it predicts a greater abundance of rare objects compared with the other mass functions, thus yielding the most conservative estimate for $f_{\rm NL}$.

The Eddington bias refers to the apparent boost in the number of high-mass clusters due to the fact that it is more likely for lower-mass objects to scatter to high luminosity than it is for rarer massive objects to scatter to lower luminosity. We account for this effect by performing the mass shift [19]

$$\ln M \longrightarrow \ln M + \frac{1}{2}\gamma\sigma^2_{\rm lnM}, \tag{47}$$

where $\sigma_{\rm lnM} \sim 0.3$ is the error estimated from the observation and γ is the local slope of the mass function determined using the relation $dn/d\ln M \propto M^\gamma$.

Putting all this together, the result is shown in Figure (5). The panel on the left shows the probability $dP/d\log M$ that XMMUJ0044 is the maximum-mass cluster in the redshift range $1.6 \leq z \leq 2.2$. The probability is plotted as function of $f_{\rm NL}$, peaking at $f_{\rm NL} \simeq 360$. This is roughly consistent with previous estimates of $f_{\rm NL}$ on cluster scales [15, 16, 18]. Comparing the ratio of extreme-value probabilities for $f_{\rm NL} = 0$ and 360, we find

$$\frac{\text{Prob(cluster is of maximum mass, } f_{\rm NL} = 360)}{\text{Prob(cluster is of maximum mass, } f_{\rm NL} = 0)} \approx 25:1 \tag{48}$$

We also computed the extreme-value pdf for $f_{\rm NL} = 0$ (Figure 5, right panel). The vertical dashed lines around the peaks represent the 1σ and 2σ intervals for the pdf. We find that XMMUJ0044 lies at about 4σ away from the mean (which occurs at around $1.8 \times 10^{14} h^{-1} M_\odot$). The relatively high number of σ reflects the fact that the pdf is strongly peaked over a narrow range of $M_{\rm max}$ when the survey volume is large (see Figure 4). We therefore conclude that, according to extreme-value statistics, the existence of XMMUJ0044 poses some tension with Gaussian initial condition.

There are certainly additional factors which we have not considered, namely, the effects of systematics, selection bias [17], error in the cluster-mass determination [61] and the uncertainty in σ_8, all of which may reduce the magnitude of $f_{\rm NL}$ needed to explain the existence of this cluster.

VII. EXTREMAL TYPES

The shape of the extreme-value distribution function holds valuable information about the statistical nature of galaxy clusters. The following theorem, which roughly states that extreme-value distributions converge to one of only three possible types, lies at the heart of extreme-value theory.

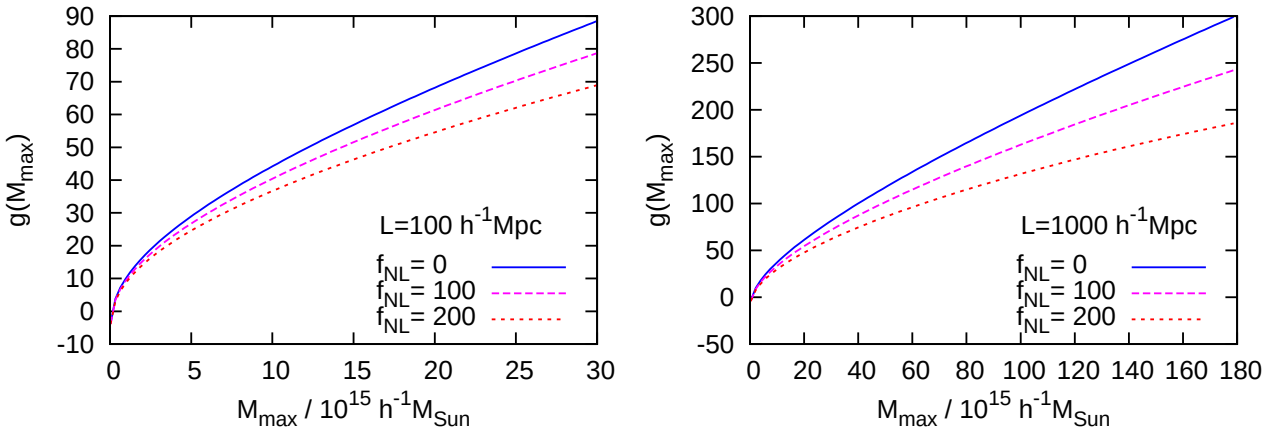

FIG. 6: The quantile plots for the distributions of extreme-mass clusters for a range of values of $f_{\rm NL}$. The patch size is $L = 100h^{-1}$Mpc (left) and $L = 1000h^{-1}$Mpc (right). The parameters used are $z = 1$ and $r = 50h^{-1}$Mpc and the Tinker mass function was used. The concavity of these curves suggests that they belong to the Fréchet class of distribution though they approach the Gumbel distribution at the high-mass end. This behaviour is insensitive to changes in all other parameters.

Theorem (*Extremal Types Theorem*)**.** *Let X_i be independent and identically distributed random variables. Define the block maximum as $M_n \equiv \max_{1\le i\le n}\{X_i\}$. If, for some constants $a_n > 0$, b_n, we have*

$$P(a_n M_n + b_n \le x) \longrightarrow G(x) \quad as \;\; n \longrightarrow \infty,$$

(in other words, if the distribution of rescaled maxima converges to a distribution G for large sample size), then G is one of the following distributions:
I. Gumbel type, $G(x) = \exp(-e^{-y})$
II. Fréchet type, $G(x) = \begin{cases} 0, & x \le b \\ \exp(-y^{-\alpha}), & x > b \end{cases}$
III. Weibull type, $G(x) = \begin{cases} \exp(-(-y)^{\alpha}), & x \le b \\ 1, & x > b \end{cases}$
where $y = ax + b$, a, b, α are constants, $a > 0$ and $\alpha > 0$.

See, for example, [62] for the proof. In this section, we investigate which of these extremal types do the distributions of extreme-mass clusters belong to.

The following function will be useful in distinguishing between the three cases:

$$g(x) = -\ln(-\ln P(x)). \tag{49}$$

In the case of the Gumbel distribution, $g(x) = P^{-1}(x) = \inf\{y : P(y) \ge x\}$, which means that $g(x)$ is the *x-quantile* of P. We shall refer to $g(x)$ as the quantile function [63, 64].

To see which extremal type a given extreme-value distribution, $P(x)$, belongs to, one simply plots the quantile function and analyse its curvature for increasing patch size L. If the quantile is a linear, the distribution is of Gumbel type. If it concaves up (*i.e.* $g''(x) > 0$), the distribution is of Weibull type. If the quantile concaves down, it is of Fréchet type. Note that the quantiles must be plotted on linear and not logarithmic scales.

Figure 6 shows the quantile plot, $g(M_{\rm max})$, of the distribution of extreme-mass clusters, $P(M_{\rm max})$, with $f_{\rm NL}$ in the range 0-200. The parameters used are $z = 1, r = 50h^{-1}$Mpc and the Tinker mass function was used. The patch sizes used are $L = 100h^{-1}$Mpc (left panel) and $1000h^{-1}$Mpc (right). The concavity of these graphs clearly shows that the distribution of extreme-mass clusters are of the Fréchet type, although the tails of the quantile graphs show an almost linear (*i.e.* Gumbel) behaviour.

The Fréchet distribution[3] arises in situations when there is a natural lower limit in the distribution function ($P(x) = 0$ for $x \le$ some constant). In our case, the definition of a galaxy cluster (*e.g.* via M_{200}) translates to a

[3] Some applications of the Fréchet distribution to environmental sciences are summarised in [64, 65]

loose lower bound on $M_{\rm max}$ and this may explain why the distribution of extreme-mass clusters appears to be of the Fréchet type. If only the high-mass tail of the distribution is taken into account, the Gumbel distribution is a reasonable approximation. As pointed out in [22], if the underlying distribution is exactly Gaussian, the distribution can be shown to converge to the Gumbel type, albeit very slowly. In any case, we find that $P(M_{\rm max})$, for all practical purposes, belongs to the Fréchet family.

This conclusion is remarkably robust against changes in $f_{\rm NL}$, mass function, redshift, separation length and patch size. It may be possible that this insensitivity stems from the truncation of the series (42). A more rigorous approach to studying the extremal-type convergence is to fit the distribution to some functional form (*e.g.* see [25] in which the extreme CMB-temperature distribution is fitted to a generalised extreme-value distribution) or prove the convergence using one of the criteria given, for example, in [62, 65].

We note that, contrary to the observation in [22], we found no combination of parameters which give rise to a Weibull distribution, which arises when there is a natural upper bound for the distribution function. Moreover, it is worth noting that if the pdfs such as those in Figures 3 and 4 are well-approximated by 'skew-symmetric' functions (*e.g.* an Edgeworth expansion) then the distribution cannot converge to the Weibull type as proven in [66].

Finally, we point out an interesting fact that if the coefficients in the expansion (38) conspire to make $P(x)$ an exactly Poissonian distribution

$$P(x,\lambda) = e^{-\lambda}\sum_{k=0}^{x}\frac{\lambda^k}{k!}, \tag{50}$$

then the limiting distribution $G(x)$ will completely degenerate to $G = 1$ or 0. This is one of the rare examples where the extreme-value distribution does not converge to any of the three standard distributions. Of course, we do not expect a realistic distribution of galaxy clusters to be exactly Poissonian.

VIII. CONCLUSION AND DISCUSSION

Our results can be summarized in three main points as follows:

1. We have presented a simple procedure to calculate the statistics of extreme-mass galaxy clusters in the presence of primordial non-Gaussianity parametrized by $f_{\rm NL}$. Our main results are the expressions for the cumulative probability distribution (42) and the probability density function (43) for the most massive object in a survey of a given volume and redshift range. These expressions enable us to deduce the most probable extreme-mass cluster in a survey of a given specification. The effects of changing the mass function and varying the value of $f_{\rm NL}$, survey volume and redshift are summarised in Figures 3 and 4.

2. Next, we applied our formalism to investigate whether the recently reported cluster XMMUJ0044.0-2-33 (mass $\sim 3.5\times10^{14}M_\odot$ at redshift 1.6) is consistent with $f_{\rm NL} = 0$. We found that if $f_{\rm NL} = 0$, this cluster lies at 4σ away from the most probable extreme-mass object predicted by extreme-value statistics. If this cluster *is* indeed the most probable extreme-mass object, then we require $f_{\rm NL} \sim 360$, which increases the extreme-value probability by a factor of 25 compared to $f_{\rm NL} = 0$. This conclusion is based on the Sheth-Tormen mass function, which gives the most conservative value of $f_{\rm NL}$ out of all the mass functions analysed in this paper. The value of $f_{\rm NL}$ obtained here is roughly in agreement with previous estimates of $f_{\rm NL}$ on cluster scales.

3. Finally, we found that the distribution of extreme-mass clusters belongs to the Fréchet family of distribution regardless of the presence of $f_{\rm NL}$. This conclusion is robust against changes in other parameters.

An important ingredient in our calculation is the mass function. In the presence of primordial non-Gaussianity, it remains to be seen what the correct mass function should be. Our investigation showed that the Press-Schechter, Sheth-Tormen and Tinker mass functions give similar extreme-value statistics at low redshift, but there are large differences at high redshift and large $f_{\rm NL}$. It remains for further numerical simulations along the lines of [41, 46] to establish the validity of the various mass functions and non-Gaussian correction factors in the presence of non-Gaussianity.

Another crucial ingredient is the bias which, in this work, was calculated using the real-space formalism given in [29, 30]. As pointed out in these papers, it is possible to extend the calculation to other types of non-Gaussianity (non-local or higher-order local type). It will be an interesting extension to study extreme-value statistics in the presence of different types of non-Gaussianity.

Our calculation relies on the approximation (40), which relates the reduced cumulants of density peaks to those of dark matter via a power-law scaling of the linear bias. Whilst we believe that this is a reasonable leading-order

approximation, a more rigorous approach is to calculate these cumulants from either nonlinear perturbation theory or use values obtained from numerical simulations. Both of these are challenging routes towards the understanding of the statistics of large-scale structures in the presence of primordial non-Gaussianity.

Finally, the value of $f_{\rm NL} \sim 360$ needed to explain the observed massive cluster at $z = 1.6$ is obviously in tension with the CMB constraint of $f_{\rm NL} = \mathcal{O}(10)$, hence suggesting that $f_{\rm NL}$ may be scale-dependent. Such a scenario generically arises in multi-field inflation [67, 68] and single-field inflation with pathological features [69, 70] (see [4, 71] and references therein for other possibilities). However, as previously mentioned, selection effects and error in the mass determination may simply be behind the apparent rareness of massive clusters. More extreme mechanisms such as modified gravity can also influence the number of high-redshift clusters [72, 73].

Acknowledgment

We are grateful to Olaf Davis for helpful discussions, and to Christopher Gordon and Aseem Paranjape for comments on an early version of the manuscript. SC supported by Lincoln College, Oxford.

[1] N. Bartolo, E. Komatsu, S. Matarrese, and A. Riotto, Phys. Rept. **402**, 103 (2004).
[2] J. M. Maldacena, JHEP **05**, 013 (2003).
[3] G. I. Rigopoulos, E. P. S. Shellard, and B. J. W. van Tent, Phys. Rev. **D73**, 083522 (2006).
[4] C. T. Byrnes and K. Choi, ArXiv e-prints (2010), 1002.3110.
[5] N. Bartolo, S. Matarrese, and A. Riotto, Phys. Rev. D **69**, 043503 (2004).
[6] M. Sasaki, J. Väliviita, and D. Wands, Phys. Rev. D **74**, 103003 (2006).
[7] E. Komatsu *et al.*, ArXiv e-prints (2010), 1001.4538.
[8] A. Cooray, D. Sarkar, and P. Serra, Phys. Rev. D **77**, 123006 (2008).
[9] B. Yu and T. Lu, Phys. Rev. D **78**, 063008 (2008).
[10] M. J. Jee *et al.*, Astrophys. J. **704**, 672 (2009).
[11] J. P. Stott *et al.*, Astrophys. J. **718**, 23 (2010).
[12] M. Brodwin *et al.*, Astrophys. J. **721**, 90 (2010).
[13] J. S. Santos *et al.*, ArXiv e-prints (2011), 1105.5877.
[14] M. Tanaka, A. Finoguenov, and Y. Ueda, ApJ**716**, L152 (2010).
[15] K. Enqvist, S. Hotchkiss, and O. Taanila, JCAP **4**, 17 (2011).
[16] B. Hoyle, R. Jimenez, and L. Verde, Phys. Rev. D **83**, 103502 (2011).
[17] S. Hotchkiss, ArXiv e-prints (2011), 1105.3630.
[18] L. Cayon, C. Gordon, and J. Silk, ArXiv e-prints (2010), 1006.1950.
[19] M. J. Mortonson, W. Hu, and D. Huterer, Phys. Rev. **D83**, 023015 (2011).
[20] S. D. M. White, MNRAS**186**, 145 (1979).
[21] O. Davis, J. Devriendt, S. Colombi, J. Silk, and C. Pichon, MNRAS**413**, 2087 (2011).
[22] S. Colombi, O. Davis, J. Devriendt, S. Prunet, and J. Silk, MNRAS, 552 (2011).
[23] J. P. Bernstein and S. P. Bhavsar, MNRAS**322**, 625 (2001).
[24] T. Antal, F. Sylos Labini, N. L. Vasilyev, and Y. V. Baryshev, Europhysics Letters **88**, 59001 (2009).
[25] G. Mikelsons, J. Silk, and J. Zuntz, MNRAS**400**, 898 (2009).
[26] R. K. Sheth and A. Diaferio, ArXiv e-prints (2011), 1105.3378.
[27] J. . Waizmann, S. Ettori, and L. Moscardini, ArXiv e-prints (2011), 1105.4099.
[28] N. Dalal, O. Dore, D. Huterer, and A. Shirokov, Phys. Rev. **D77**, 123514 (2008).
[29] P. Valageas, Astron. Astrophys. **514**, A46 (2010).
[30] P. Valageas, Astron. Astrophys. **508**, 93 (2009).
[31] S. Chongchitnan and J. Silk, Astrophys. J. **724**, 285 (2010).
[32] S. Chongchitnan and J. Silk, Phys. Rev. D **83**, 083504 (2011).
[33] O. Lahav and A. R. Liddle, ArXiv e-prints (2010), 1002.3488.
[34] S. M. Carroll, W. H. Press, and E. L. Turner, Ann. Reviews of Astron. Astrophys. **30**, 499 (1992).
[35] O. Lahav, P. B. Lilje, J. R. Primack, and M. J. Rees, MNRAS **251**, 128 (1991).
[36] S. Weinberg, *Cosmology* (Oxford University Press, 2008).
[37] D. J. Eisenstein and W. Hu, Astrophys. J. **496**, 605 (1998).
[38] S. Matarrese, L. Verde, and R. Jimenez, Astrophys. J. **541**, 10 (2000).
[39] M. LoVerde, A. Miller, S. Shandera, and L. Verde, JCAP **0804**, 014 (2008).
[40] A. Paranjape, C. Gordon, and S. Hotchkiss, ArXiv e-prints (2011), 1104.1145.
[41] C. Wagner and L. Verde, ArXiv e-prints (2011), 1102.3229.
[42] W. H. Press and P. Schechter, Astrophys. J. **187**, 425 (1974).
[43] R. K. Sheth and G. Tormen, MNRAS **308**, 119 (1999).

[44] J. Tinker *et al.*, Astrophys. J. **688**, 709 (2008).
[45] J. L. Tinker *et al.*, Astrophys. J. **724**, 878 (2010).
[46] A. Pillepich, C. Porciani, and O. Hahn, MNRAS **402**, 191 (2010).
[47] S. Bhattacharya *et al.*, Astrophys. J. **732**, 122 (2011).
[48] S. Matarrese and L. Verde, Astrophys. J. **677**, L77 (2008).
[49] V. Desjacques, D. Jeong, and F. Schmidt, ArXiv e-prints (2011), 1105.3628.
[50] F. Schmidt and M. Kamionkowski, ArXiv e-prints (2010), 1008.0638.
[51] S. Shandera, N. Dalal, and D. Huterer, ArXiv e-prints (2010), 1010.3722.
[52] J. N. Fry and E. Gaztanaga, Astrophys. J. **413**, 447 (1993).
[53] N. Kaiser, Astrophys. J. Lett. **284**, L9 (1984).
[54] F. Bernardeau, S. Colombi, E. Gaztañaga, and R. Scoccimarro, Phys. Rep.**367**, 1 (2002).
[55] J. Gaite, JCAP **11**, 4 (2009).
[56] M. Manera and E. Gaztanaga, ArXiv e-prints (2009), 0912.0446.
[57] R. Juszkiewicz, F. R. Bouchet, and S. Colombi, ApJ**412**, L9 (1993).
[58] A. Cooray and R. Sheth, Phys. Rep.**372**, 1 (2002).
[59] T. Giannantonio and C. Porciani, Phys. Rev. D **81**, 063530 (2010).
[60] T. Matsubara, Phys. Rev. D **83**, 083518 (2011).
[61] A. Mantz, S. W. Allen, H. Ebeling, D. Rapetti, and A. Drlica-Wagner, MNRAS**406**, 1773 (2010).
[62] M. Leadbetter, G. Lindgren, and H. Rootzén, *Extremes and related properties of random sequences and processes* (Springer-Verlag, New York, 1982).
[63] E. Jondeau, S.-H. Poon, and M. Rockinger, *Financial Modeling Under Non-Gaussian Distributions* (Springer-Verlag, London, 2007).
[64] R.-D. Reiss and M. Thomas, *Statistical Analysis of Extreme Values*, 3rd ed. (Birkhauser Verlag, Basel, 2007).
[65] S. Kotz and S. Nadarajah, *Extreme Value Distributions: Theory and Applications* (Imperial College Press, London, 2000).
[66] C. Sheng-Mao and M. G. Genton, Communications in Statistics: Theory & Methods **36**, 1705 (2007).
[67] C. T. Byrnes, S. Nurmi, G. Tasinato, and D. Wands, JCAP **2**, 34 (2010).
[68] J. Kumar, L. Leblond, and A. Rajaraman, JCAP **4**, 24 (2010).
[69] X. Chen, R. Easther, and E. A. Lim, JCAP **6**, 23 (2007).
[70] F. Arroja, A. Enea Romano, and M. Sasaki, ArXiv e-prints (2011), 1106.5384.
[71] A. Becker, D. Huterer, and K. Kadota, JCAP **1**, 6 (2011).
[72] M. C. Martino, H. F. Stabenau, and R. K. Sheth, Phys. Rev. D **79**, 084013 (2009).
[73] K. C. Chan and R. Scoccimarro, Phys. Rev. D **80**, 104005 (2009).