The excursion set approach in non-Gaussian random fields

Marcello Musso$^1$ & Ravi K. Sheth$^{2,3}$

$^1$ CP3-IRMP, Université Catholique de Louvain, 2 Chemin du Cyclotron, 1348 Louvain-la-Neuve, Belgium
$^2$ The Abdus Salam International Center for Theoretical Physics, Strada Costiera, 11, Trieste 34151, Italy
$^3$ Center for Particle Cosmology, University of Pennsylvania, 209 S. 33rd St., Philadelphia, PA 19104, USA

ABSTRACT

Insight into a number of interesting questions in cosmology can be obtained by studying the first crossing distributions of physically motivated barriers by random walks with correlated steps: higher mass objects are associated with walks that take fewer steps before crossing the barrier. We show how to write the first crossing distribution as a formal series, ordered by the minimum number of times a walk upcrosses the barrier. Since walks with many upcrossings are negligible if the walk has not taken too many steps, the leading order term in this series is the most relevant for understanding the massive objects of most interest in cosmology. For walks associated with Gaussian random fields, this first term only requires knowledge of the bivariate distribution of the walk height and slope, and provides an excellent approximation to the first crossing distribution for all barriers and smoothing filters of current interest. We show that this simplicity survives when extending the approach to the case of non-Gaussian random fields. Although this second part of our analysis is motivated by the possibility that the primordial fluctuation field is non-Gaussian, our results are general. In particular, they do not assume the non-Gaussianity is small, so they may be viewed as the solution to an excursion set analysis of the late-time, nonlinear fluctuation field rather than the initial one.

Key words: large-scale structure of Universe

1 INTRODUCTION

The statistical distribution of gravitationally bound objects in the Universe is a powerful tool for constraining the amount of primordial non-Gaussianity, thus helping shed some light on the physics of the very early times. The dependence on mass of the abundance and spatial correlations of collapsed objects are useful and complementary tools for probing non-Gaussianity on different scales, in particular, scales that are smaller than those accessible with CMB observations.

The excursion set approach (Bond et al. 1991) provides an analytical framework for linking the statistics of haloes to fluctuations in the primordial density field. In this approach, one studies the overdensity field $\delta$ smoothed on the scale $R$, 

$$
\delta(x, R) = \int dy W_R(x - y) \delta(y)
$$

(1)

where $W_R$ is a filter that goes to zero for $|x - y| \gg R$. At a given (randomly chosen) position in space the evolution of $\delta_R$ as a function of (the inverse of) $R$ resembles a random trajectory. Repeating this for every position in space gives an ensemble of trajectories, each starting from zero (homogeneity demands $\delta = 0$ for infinitely large smoothing scales). For each trajectory, one looks for the largest $R$ (if any) for which the value of the smoothed density field lies above some threshold value (which may itself depend on $R$). An object of mass $M \sim R^3$ is then associated with that trajectory.

If $dn/dM$ denotes the comoving number density of haloes of mass $M$, then the mass fraction in such halos is $(M/\bar{\rho}) (dn/dM)$, where $\bar{\rho}$ is the comoving background density. The excursion set approach assumes that this mass fraction equals the fraction of walks which cross the threshold (the “barrier”) for the first time when the smoothing scale is $R$:

$$
f(R) dR = (M/\bar{\rho}) (dn/dM) dM.
$$

(2)

Although recent work has focussed on the shortcomings of this ansatz (Paranjape & Sheth 2012), the first crossing distribution is nevertheless expected to provide substantial insight into the dependence of $dn/dM$ on cosmological parameters. In any case, the question of how the first crossing distribution depends on the nature of the underlying fluctuation field is interesting in its own right.

A crucial part of the problem is to avoid double count-
ing trajectories, i.e., to discard at lower scales all trajectories that have already crossed at larger scales (since they are already associated with an object of larger mass — given by the largest scale on which the trajectory crossed the barrier). This is rather straightforward to implement numerically, but hard to deal with analytically. Indeed, exact solutions are known only for the unrealistic case of walks with uncorrelated steps (for Gaussian fields, this corresponds to a smoothing filter that is a sharp step function in Fourier space) and only for a few specific barriers. Considerable effort has been devoted to finding satisfactory analytical approximations, or fitting formulae, for the generic case in which steps are correlated.

The problem is potentially even harder for non-Gaussian initial conditions, since different Fourier modes of the field become coupled, and this introduces additional correlations between the steps, whatever the smoothing filter. Moreover, the most sizeable non-Gaussian deviations are likely to be in the (massive object) tail of the distribution. In this regime, perturbative expansions around the Gaussian result are likely to blow up, so they must be handled with care (D’Amico et al. 2011).

In this paper we provide a simple analytic approximation scheme that works for a broad variety of barriers and filters, and can be implemented up to an arbitrary precision level for any (Gaussian or non-Gaussian) distribution of the underlying matter density field. The general formalism is presented in Section 2 and explicit calculations are carried out in Section 3, where we summarize our previous work on Gaussian fields and show how to extend it to non-Gaussian fields. Section 4 shows how to use our results as the basis of an excursion set study based on the late-time, nonlinear (rather than initial) fluctuation field. A final section summarizes our results. Appendix A discusses how to go beyond the simplest approximation we present in the main text, and our use of the Edgeworth and related-expansions for approximating non-Gaussian distributions is summarized in Appendix B.

2 FIRST CROSSING DISTRIBUTION WITH CORRELATED STEPS

In hierarchical models, the variance \( \langle \delta^2(R) \rangle \) of the density field \( \delta \) when smoothed on scale \( R \) vanishes by definition for \( R = \infty \), and it grows monotonically for smaller \( R \) (note that \( \langle \delta \rangle \equiv 0 \) for any \( R \)), according to

\[
s(R) = \int_0^\infty \frac{dk}{k} \frac{k^3 P(k)}{2\pi^2} W^2(kR),
\]

where \( P(k) \) is the power spectrum of \( \delta \). Therefore, \( R \) and \( s \) can be used interchangeably, and it is in fact customary and convenient to study the walks as a function of \( s \) rather than \( R \), as this has the advantage of hiding the dependence on the power spectrum and the smoothing filter. These only appear when the actual relation between \( s \) and \( R \) is needed.

What we are after is the first crossing rate, i.e. the probability that a walk \( \delta \) crosses for the first time the barrier \( b(s) \) at some scale \( s \). In other words, we want to compute the probability that \( \delta(s) > b(s) \) at small \( s \) but \( \delta(s_1) < b(s_1) \) for all \( s_1 < s \), knowing the probability distribution \( p(\delta; s) \) of the walk values at any \( s \). In general, requiring \( \delta(s) > b(s) \) is straightforward, whereas the additional constraint on the walk heights for all \( s_1 < s \) is difficult to treat analytically.

2.1 Height alone

In one of the earliest works on this subject, Press & Schechter (1974) simply ignored this constraint, and estimated \( f(s) \) as

\[
f_{CC}(s) = -\frac{d}{ds} \int_{-\infty}^b d\delta \, p(\delta; s).
\]

(The reason for the subscript CC will become clear shortly.) Strictly speaking, Press & Schechter ended up multiplying the right hand side of this expression by a factor of 2, and they only studied the special case in which \( b = \delta_s \) is independent of \( s \). (The extension to barriers which decrease monotonically as \( s \) increases is trivial; if the barrier increases sufficiently rapidly with \( s \), then one must be a little more careful, as we discuss shortly.) That this does not impose any constraint on the walk values at larger scales (smaller \( s \)) is a point which was highlighted by Bond et al. (1991).

In fact, this formulation does not even distinguish between trajectories crossing the threshold upwards or downwards, a point to which we return shortly.

Recently Paranjape, Lam, & Sheth (2012) noted that there is an interesting and instructive limit in which equation (3) is exact. Consider the set of smooth deterministic curves having \( \delta \propto \sqrt{s} \). Each of these curves represents what they called a completely correlated walk: one which is a monotonic function of \( s \) whose amplitude is set by a single number, the constant of proportionality, which one may take to be the height on scale \( s = 1 \). If the distribution of this constant is specified on one scale (say \( s = 1 \)) then the distribution of \( \delta \) on another scale, \( p(\delta; s) \), is simply related to \( p(\delta; 1) \), and, for this family of curves, equation (3) is exact: hence the subscript CC. This limit is interesting because, regardless of the filter and the matter power spectrum, all correlated walks tend to deterministic trajectories with \( \delta \propto \sqrt{s} \) as \( s \to 0 \). Thus, in this (large mass) limit, equation (3) is exact, explaining the numerical results of Bond et al. (1991).

At larger values of \( s \), the completely correlated limit is no longer so accurate. However, although small fluctuations around these trajectories appear, most walks still retain monotonic functions of \( s \). Therefore the contribution to \( p(\delta; s) \) from walks criss-crossing the barrier multiple times is still negligible, so the constraint \( \delta(s_1) < b(s_1) \) for \( s_1 < s \) is automatically satisfied.

2.2 Upcrossing requires both height and slope

As \( s \) gets larger, one must account for larger and larger fluctuations away from the deterministic trajectories. Musso & Sheth (2012) argued that the most efficient way of doing so, at fixed height \( \delta \) on scale \( s \), is to consider fluctuations in the slope \( v \equiv d\delta/ds \) (\( v \) for velocity). (For the ensemble of deterministic walks, the distribution of the slope \( v \) at fixed height \( \delta \) is a delta-function centered on \( \delta/2s \).) Since one only wants to count walks that are crossing the barrier upwards, to the condition that \( \delta = b(s) \) one should add the requirement that \( v \geq db/ds \) (for a barrier of constant height, this is just \( v \geq 0 \)).
Thus, if earlier upcrossings can be neglected, Musso & Sheth showed that \( f(s) \) can be computed from the joint probability \( p(\delta, v; s) \) that a walk reaches \( \delta \) at scale \( s \) with velocity \( v \), as

\[
f(s) \simeq \int_{\nu(s)}^{+\infty} dv \left[ v - b(s) \right] p(b(s), v; s).
\]

(5)

Clearly, this formulation fails to discard those walks that were above threshold at some \( s_1 < s \), but with \( \delta(s_2) < b(s_2) \) for \( s_1 < s_2 < s \), i.e. walks with more than one upcrossing. However, at small \( s \), the fraction of such walks is tiny, since the correlations between the steps make sharp turns very unlikely.

Since \( p(b, v) = \langle \delta_D(\delta - b)\delta_D(\delta' - v) \rangle \), this approximation can also be written as

\[
f(s) \simeq \int_{\nu}^{+\infty} d\nu \left[ \frac{d}{ds} \theta(\delta - b) \right] \theta(\delta' - b'),
\]

(6)

which makes it clear that the condition to recover \( f_{CC} \) is that \( \delta' > b' \) for most realisations. This is exactly the case for correlated steps in the large mass regime, since \( \delta = b \) implies that typically \( \delta' \sim b/s \), and \( b' \ll b/s \) for small \( s \) (as long as the barrier is not receding too fast from its initial value).

In terms of the conditional probability \( p(v|b(s)) = p(b(s), v; s)/p(b(s); s) \), and omitting for ease of notation all explicit \( s \) dependences, the rate can also be written as

\[
f(s) \simeq p(b) \int_{\nu}^{+\infty} d\nu \left( v - b' \right) p(v|b).
\]

(7)

This allows a very intuitive explanation in terms of particles in a box: \( p(b) \) plays the rôle of a number density at \( b \) (the number of particles in the one-dimensional volume element \( d\delta \)), while the integral is the mean of \( \delta' - b' \) over all velocities larger than the barrier’s increment given that \( \delta = b \), that is the average escape velocity at \( b \). The product of the two evaluated at the boundary is by definition the escape rate from the box. This makes it also easy to see the connection to deterministic walks for which \( p(v|b) \to \delta_D(b/2s) \), and thus \( f(s) \to p(b)(b/2s - b') = -p(b/s) \sqrt{b/(b^2 - s^2)}/ds \), which is indeed \( f_{CC}(s) \). Of course, at larger \( s \), when \( p(v|\delta) \) is broader, equation (7) is a substantially more accurate approximation for \( f(s) \).

### 2.3 Accounting for multiple upcrossings

The approximation of equation (5) accounts for all walks that cross the barrier upwards at \( s \), including those that crossed the barrier previously, and thus it overestimates \( f(s) \). The error is expected to increase as \( s \) gets large, when such walks become increasingly common. Removing all the walks that crossed at \( s_1 < s \), i.e. walks with \( \delta(s_1) = b(s_1) \) and \( v(s_1) > b'(s_1) \), and then integrating over \( s_1 \), would account correctly for the trajectories with just one crossing before the last one. So, if we stopped here, and assuming for simplicity a constant barrier, then we would get

\[
f(s) = \int_{0}^{+\infty} dv \nu p(b, v)
- \int_{0}^{+\infty} ds_1 \int_{0}^{+\infty} dv_1 \nu_1 \int_{0}^{+\infty} dv \nu p(b_1, v_1, b, v) + \ldots,
\]

(8)

where \( p(b_1, v_1, b, v) \) is the quadrivariate distribution of \( \delta_1, \delta(s_1), \delta(s) \), and \( b_1 \equiv b(s_1) = b \). It is straightforward to include a moving barrier, simply inserting \( b' \) and \( b'_1 \) where needed (à la equation 3).

Trajectories crossing more than once would now be over-counted: for instance, a single walk crossing at \( s_1 \) and \( s_2 \) would be removed twice by this procedure, and needs to be reintroduced. This would call for an additional correction, for walks crossing twice or more, containing \( p(b_1, v_1, b_2, v_2) \) and integrals over \( s_1 \) and \( s_2 \), and so on. However, trajectories with more zigzags will be even more suppressed, making an expansion in the number of crossings meaningful in the sense of perturbation theory at small \( s \).

Similarly to equation 3 for the leading order term, the first subleading correction can also be written in a more evocative way as

\[
\langle \frac{d}{ds} \theta(\delta - b) \theta(\delta' - b') \rangle \equiv \frac{d}{ds_1} \theta(\delta_1 - b_1) \theta(\delta_1' - b_1'),
\]

(9)

and the same pattern holds for higher order corrections. A rigorous derivation of this expansion from a path integral expression is carried out in Appendix A. However, for most cosmological applications, the analysis of Musso & Sheth (2012) is sufficiently accurate, so one does not even need the second term of equation (5).

### 2.4 Gaussian or not?

Before moving on, it is worth noting that the logic above holds in full generality, regardless of the shape of the distribution: the completely correlated non-Gaussian walks have a modified \( s \) dependence, but the first crossing distribution in this limit is still given by equation (11), and this limit will still be a good approximation as \( s \to 0 \). As \( s \) becomes large, an expansion in the number of previous upcrossings is still sensible, where constraining the slope of the walk is the most natural and efficient way of ensuring it is upcrossing. So, equation (7) should remain a good approximation until \( s \) values where walks which can have previously upcrossed more than once dominate, at which point the next terms in the program (outlined in Section 2.3) will become important.

That said, there is one sense in which the non-Gaussian case is more complicated. For a Gaussian field, the probability distribution of \( \delta \) on scale \( s \) only depends on the ratio \( \delta/\sqrt{s} \). This makes

\[
f_{CC}(s) = -\left( \frac{d}{ds} \frac{b}{\sqrt{s}} \right) p(b/\sqrt{s}),
\]

(10)

where for ease of notation we have not written the scale dependence of \( b(s) \) explicitly. If the barrier is constant, \( b = \delta_c \), then \( \delta_c/\sqrt{s} \) is usually called \( \nu \) and one finds \( s f(s) = \nu p(\nu)/2 \): the final factor of 1/2 is the reason Press & Schechter multiplied by 2 so many years ago. But notice that, in this limit, the first crossing distribution is very simply related to the shape of the pdf. This would also apply to the non-Gaussian case, provided that the distribution of \( \delta \) is indeed a function of \( \delta/\sqrt{s} \) only. This is rarely the case, but as we will see it becomes a reasonable approximation at very large scales.
3 Explicit Calculation

In what follows, it will be convenient to use the rescaled stochastic quantities

\[ \Delta \equiv \frac{\delta}{\sqrt{s}} \quad \Delta' \equiv \frac{d\Delta}{ds} \quad \text{and} \quad \xi \equiv -\frac{\Delta'}{\sqrt{\langle \Delta^2 \rangle}} \equiv 2Gamma \Delta' \]

where \( \Gamma \), defined by \( (2Gamma s)^2 \equiv 1/\langle \Delta^2 \rangle \) is a weak function of \( s \) (e.g. Musso & Sheth 2012). Notice that

\[ \langle \Delta^2 \rangle = \langle \xi^2 \rangle = 1 \quad \text{and} \quad \langle \Delta \xi \rangle = 0; \] (12)

i.e., \( \Delta \) and \( \xi \) are independent. Similarly, we will work with

\[ B(s) \equiv \frac{b(s)}{\sqrt{s}} \quad \text{and} \quad X \equiv -\frac{dB/ds}{\sqrt{\langle \Delta^2 \rangle}} = -2Gamma B'. \] (13)

where \( B' \equiv dB/ds \). The sign of \( X \) is chosen so that a typical barrier has \( X > 0 \), since \( b(s) \) for most problems of current interest does not vary much with \( s \), and thus \( B' < 0 \). Since we are enforcing \( \delta = b \), in these rescaled variables \( f(s) \) reads

\[ f(s) \approx -B' \int_{-\infty}^{s} d\xi \left[ 1 - \frac{\xi}{X} \right] p(B, \xi; s). \] (14)

Equivalently, equation (10) becomes

\[ f(s) \approx \frac{d}{ds} \left[ \int_{-\infty}^{\infty} d\Delta \int_{-\infty}^{X(s)} d\xi p(\Delta, \xi; s, s_1) \right]_{s_1 = s}, \] (15)

where \( s_1 \) must be set equal to \( s \) after taking the derivative. Since \( X \sim B \), we see explicitly that we recover \( f_{CC}(s) \) in the large mass \( X \gg 1 \) limit.

3.1 Summary of the Gaussian result

If \( \delta \) is a Gaussian process, then the joint distribution of \( \Delta \) and \( \xi \) is particularly simple because \( \langle \Delta \xi \rangle = 0 \). When \( \Delta = B \), then

\[ pc(B, \xi) = pc(B) pc(\xi) = \frac{e^{-B^2/2}}{\sqrt{2\pi}} \frac{e^{-\xi^2/2}}{\sqrt{2\pi}}. \] (16)

Inserting this in equation (14) shows that \( f(s) \) will be proportional to \(-B'pc(B)\), which, for a Gaussian process is just \( f_{CC}(s) \) times a correction factor that is a function of \( X \) alone. Performing the integral yields

\[ f(s) = -B' e^{-B^2/2} \sqrt{2\pi} \left[ \frac{1}{2} \text{erf}(X/\sqrt{2}) + \frac{e^{-X^2/2}}{2\sqrt{\pi}} \right]. \] (17)

This reduces to equation (10) – and therefore to \( f_{CC}(s) \) – for \( X \gg 1 \) (the first term in the square brackets tends to unity while the second one is exponentially suppressed).

Musso & Sheth (2012) showed that, for a wide variety of smoothing filters, power-spectra and barrier shapes, this expression was substantially more accurate than \( f_{CC} \), and accurate down to scales on which a substantial fraction of the walks might have negative slopes. However, it cannot be accurate to arbitrarily small scales since the integral of \( f(s) \) over all \( s \) diverges. This is, of course, related to the fact that multiple upcrossings of the barrier become important as \( s \) increases. Appendix A describes how to account for these, but since we have not found a similarly simple analytic expression for the resulting \( f(s) \), and the range over which equation (14) is accurate covers most of the range which is of interest in cosmology, we will continue with this simpler case.

Before moving on, we think the special case of Gaussian walks with correlated steps crossing a constant barrier deserves further comment. This is because, once one accounts for differences in notation and presentation, our equation (17) turns out to be the same as equation (3.14a) of Bond et al. (1999) for the first crossing distribution of a barrier of constant height. The origin of this agreement is that the expression within angle brackets in our equation (6) is the same as their equation (3.12a). (The same is true for our equation (10) and their (13).) However, they appear to have made an error when comparing their equation (3.14a) with the Monte-Carlo solution of the constant barrier problem: their Figure 9 suggests that ignoring multiple crossings is a bad approximation, whereas Musso & Sheth (2012) showed that it is in fact rather good. (The Monte-Carlo solutions themselves are in good agreement.) This led them, and the rest of the field since, to dismiss the approximation in which one ignores multiple upcrossings, and to focus instead on what appeared to be a more tractable problem (in which one ignores correlations between steps). In this respect, one might view our analysis of the constant barrier problem as having corrected an error which went unnoticed for more than twenty years. Of course, our analysis is more general, since we have shown how to apply it (successfully) to arbitrary barriers. We now show how it can be generalized to arbitrary fluctuation fields.

3.2 Generic non-Gaussian case

The joint probability distribution of a generic stochastic process can always be written as an asymptotic expansion in Hermite polynomials around the Gaussian distribution obtained from the second moments. Since \( \Delta \) and \( \xi \) are independent

\[ p(B, \xi) = \sum_{n,k} \frac{\langle H_n(\hat{\Delta}) H_k(\hat{\xi}) \rangle}{n! k!} H_n(B) H_k(\xi) pc(B) pc(\xi). \] (18)

where we have used hats to distinguish the stochastic quantities from the continuous variables of the probability distribution, and \( H_n(x) \equiv \exp(x^2/2)(-d/dx)^n \exp(-x^2/2) \). This expression follows from the fact that the Hermite polynomials form an orthogonal basis with respect to the Gaussian weight. Rearranging the terms of the sum and factoring out \( p(B) \) shows that

\[ p(B, \xi) = p(B) \sum_k \frac{\langle H_k(\hat{\xi}) \rangle}{k!} \left( \frac{d}{d\xi} \right)^k pc(\xi), \] (19)

where

\[ \langle H_k(\hat{\xi}) \rangle = \sum_n \frac{\langle H_n(\hat{\Delta}) H_k(\hat{\xi}) \rangle H_n(B)/n!}{\sum_n \langle H_n(\hat{\Delta}) \rangle H_n(B)/n!} \] (20)

and \( \langle f(\hat{\xi}) \rangle \) is the expectation value of \( f(\hat{\xi}) \) given that \( \Delta = B \), i.e. the one computed from \( p(\xi|B) \).

This expression must be inserted into equation (14) and integrated over \( \xi \). The \( k = 0 \) term is just \( p(B) pc(\xi) \), and gives the same as equation (17), with \( pc(B) \) replaced by \( p(B) \). The following ones can be integrated by parts, and they pick up a factor of

© 0000 RAS, MNRAS 000, 1–9
that become non-perturbative on large scales) can be expressed as

\[(\partial^2 B / \partial X)^{k-1} [1 + erf(\sqrt{2\pi} X)] / (2\pi^2),\]

for \(k \geq 2\) becomes \(H_{k-2}(X)p_2(X) / (2\pi^2)\). For \(k = 1\), one has

\[\langle H_0(\Delta) H_0(\xi) \rangle = \langle X / B' \rangle \langle H_{1+1}(\Delta) \rangle / (n + 1); \quad (21)\]

furthermore, \(H_0(B)p(B) = \int_0^\infty d\Delta \Delta H_{1+1}(\Delta)p_2(\Delta),\) so that

\[p(B) \langle H_0(\xi) \rangle B = \langle X / B' \rangle \int B d\Delta \partial p(\Delta) / \partial s. \quad (22)\]

Putting all the contributions together yields

\[f(s) = \left[ \frac{d}{ds} \int_{\Delta(B)}^\infty d\Delta p(\Delta) \right] \left[ 1 + erf(\sqrt{2\pi} X) \right] / 2 \]

\[ - B' \langle B \rangle \frac{e^{-X^2/2}}{\sqrt{2\pi X}} \left[ 1 + \sum_{k=0}^{\infty} \frac{\langle H_{k+2}(\xi) \rangle}{(k + 2)!} \right] B s \).

Equation (23), the main result of this paper, is the non-Gaussian generalization of equation (17). The connection is most clearly seen by supposing that \(p(\delta; s)\) is a function of \(\delta / \sqrt{s}\), and in particular the derivative of the integral in the first line, which we could have written as \(f(\delta; s)\), becomes \(-B'p(B)\). Then equation (23) – apart from the polynomial corrections in the square brackets of the second line, which we suppose small – becomes equation (17), except that here \(p(B)\) is the full non-Gaussian distribution. In general, of course \(p(\delta; s)\) will not be a function of \(\delta / \sqrt{s}\) only; the associated departure from self-similarity will introduce an additional term, and this is why equation (23) involves an explicit derivative with respect to \(s\).

In the large mass regime, the entire second line of equation (23) is exponentially suppressed with respect to the first one, as it was in the Gaussian case, and the first crossing rate reduces to equation (4). The polynomial corrections in the second line, which would up for \(X \gg 1\), have a chance of becoming non-negligible only at large \(s\) when the exponential suppression is no longer effective. In this regime, however, the perturbative treatment of these non-Gaussian corrections is fully under control.

A remarkable feature of this result is that, although we started from the non-Gaussian bivariate distribution \(p(b, \nu; s)\), all the relevant non-Gaussian corrections (those that become non-perturbative on large scales) can be expressed in terms of the univariate non-Gaussian distribution \(p(b; s)\). We have thus managed to disentangle the problem of the first crossing of the barrier from that of the evaluation of the probability of the walks, which we deal with next.

3.3 Non-Gaussian case: large mass limit

We have already argued that in the large mass limit the formal expression for the first crossing rate coincides with Eq. (4). To see this explicitly, differentiate with respect to \(s\) to get

\[f(s) \approx -B' + \sum_{n=1}^{+\infty} \frac{\Delta_n^3}{n!} \left[ -\frac{\partial}{\partial B} \right]^{n-1} p(B; s), \quad (24)\]

where the infinite sum is the integral of the Kramers-Moyal expansion for \(\partial p / \partial s\). The crucial point is therefore to compute \(p(B)\) from the moments of the distribution.

The single point distribution \(p(B)\) can be written as

\[p(B; s) = e^{W(B; s)} / \sqrt{2\pi}, \quad (25)\]

where the full expression of \(W(B; s)\) in terms of the moments is given in Appendix A as a series of modified Hermite polynomials. This function, which corresponds to the logarithm of the Edgeworth expansion of \(p(B; s)\), has a straightforward interpretation in terms of connected Feynman diagrams constructed out of the connected moments of the distribution, and better convergence properties that the Edgeworth expansion itself. We also show in the Appendix that the large mass limit \((B \gg 1)\) of \(W(B; s)\) is obtained by keeping only the highest order term of each polynomial, which in diagrammatic language corresponds to discarding diagrams with loops. This approximation is also what one would get doing the analysis in Fourier space and transforming back to real space by mean of a stationary phase approximation. In this regime, the infinite series of polynomials turns into a simpler infinite power series, whose first terms are

\[W(B; s) \approx -B^2 / 2 + \left( \frac{\Delta_3}{3!} \right) B^3 + \left( \frac{\Delta_4}{4!} \right) - 3(\Delta_3 / 2) B^3 + \ldots \quad (26)\]

In the same spirit, we can approximate the \(n\)-th derivative as \(\partial^n / \partial B^n p(B) \approx \partial^n W / \partial B^n p(B)\), since higher derivatives of \(W(B; s)\) also correspond to loop diagrams and are subleading.

The consistency of the truncation of \(W(B; s)\) is a delicate subject. Clearly, as \(B\) becomes large one should keep adding more and more terms to Eq. (26), especially if the non-Gaussian moments are large, and a true \(B \to \infty\) limit would necessarily require resumming the whole series. Fortunately, the range of values of interest for Cosmology (where \(B\) increases both with mass and redshift) is not so extreme, since primordial non-Gaussianities are fairly small. As discussed by D’Amico et al. (2011), for values of \((\Delta_3)\) \(\sim .01\) and \(B \sim 10\) (corresponding to the most massive clusters of galaxies) the three terms listed in Eq. (26) are \(O(100), \ O(10)\) and \(O(1)\) respectively, while neglected terms start with \(O(10^{-1})\). These values are obtained for primordial non-Gaussianity with \(f_{NL} \sim 100\), which is now excluded by the Planck mission (Ade et al. 2013). However, even larger values of \(B\) can be attained at higher redshifts, or by the study of different objects like the reionisation pattern of cosmic structures (Joudaki et al. 2011, D’Alonso et al. 2013), so that the discussion about how to truncate \(W(B)\), besides having its own theoretical interest, is not unnecessary.

Truncating the Kramers-Moyal series in Eq. (24) is on the other hand less dangerous. The reduced moments typically tend to a constant on large scales, and the presence of their derivative in the coefficients of the series introduces an additional suppression. Furthermore, this series does not sit in an exponential, and errors in the truncation are potentially less harmful. Already for the \(n = 3\) term of the series, keeping just the leading term of \(\partial W / \partial B\) gives a \(O(1)\) result (or less, given the additional suppression due to the scale derivative). Within the range of parameters outlined above, a fair approximation for the first crossing rate is thus

\[f(s) \approx -B' + \left( \frac{\Delta_3}{3!} \right) B^3 \frac{W(B; s)}{\sqrt{2\pi}}, \quad (27)\]

with \(W(B; s)\) given by Eq. (26).

In many cases \((\Delta_3)\) (and more generally \((\Delta_n)\)) is only weakly scale dependent. If we can drop the \((\Delta^3)\) term, then the expression above simplifies even further, reducing to equation (10). This is just the Gaussian result,
−B′ exp(−B′²/2)/√π, times the non-Gaussian correction to \( p(B) \), \( \exp(\Delta^3/3! + \ldots) \). This factorisation justifies the common practice of obtaining the full non-Gaussian mass function as the product of the fit from Gaussian simulations times an analytically predicted non-Gaussian correction, known as the non-Gaussian to Gaussian ratio. As already pointed out by Musso & Paranjape (2012), this ratio is simply the ratio of the pdf's. Our results confirm this intu- tion, and at the same time highlight the conditions under which this result is true.

### 3.4 Relation to previous work

Most previous work on non-Gaussian excursion sets has considered a barrier of constant height, for which \( −B′ = B/2s \). This is for instance the case of Matarrese et al. (2001) and LoVerde et al. (2008), who also explicitly assume that \( f(s) = 2\sigma S/\langle \Delta S \rangle \), or our \( \langle \Delta S \rangle \equiv d\sigma S/\langle \Delta^c \rangle \), (see however the discussion on the fudge factor by Matarrese et al.). This assumption, together with the choice of a Top-Hat filter, is expected to be justified from the point of view of excursion sets. However, their main concern was reproducing the results of N-body simulations, rather than excursion sets, and multiplying by 2 was going in the correct direction. Also, this error disappears when considering the non-Gaussian to Gaussian ratio, as they did, with the aim of computing the correction that should multiply the result Gaussian simulations.

The correspondence between our expression in the large mass limit and theirs is helped by noting that it is conventional to define \( \sigma S \equiv \langle \Delta^3 \rangle \), so our \( \langle \Delta^3 \rangle \equiv d\sigma S/\langle \Delta^c \rangle \). Since equation (1) with an extra factor of 2 is the full story, they are in effect missing the \( X \) dependent corrections which matter at smaller masses. Our large mass limit differs slightly from the one of LoVerde et al. only because they keep the Edgeworth expansion, while we have been careful about how we write the large mass limit of equation (23). In this we followed D’Amico et al. (2011), who pointed out that perturba- tive non-Gaussian corrections blow up at small \( s \), and need to be resummed in an exponential, whose argument corre- sponds to equation (20) in this regime. The same approach is followed by LoVerde & Smith (2011).

If \( \sigma S \) is only weakly scale-dependent, so the \( \langle \Delta^3 \rangle \) term can be dropped, then the expression above simplifies even further: it is just the Gaussian result for \( f(s) \) times the non-Gaussian correction to \( p(B) \). Musso & Paranjape (2012) used this to argue that the large scale limit of the non-Gaussian mass function from correlated random walks is always one half of the one obtained without filter-induced correlations, finding very good agreement with Monte-Carlo simulations.

Moving barriers and weakly non-Gaussian fields were first considered by Lam & Sheth (2004), but only for a sharp-\( k \) filter. They found that, for moving barriers also, the large mass limit is just the Gaussian result times the non-Gaussian correction to the pdf, provided \( d\sigma S/\langle \Delta^c \rangle \) is small. Our more general analysis confirms this is true for other filters also, although the Gaussian result itself de- pends on the smoothing filter. Although writing \( f(s) \) this way is common practice, our analysis shows that it is not appropriate at lower masses, nor will it be accurate if \( \sigma S \) is scale-dependent, the latter being a point also made by Musso & Paranjape (2012).

A self-consistent treatment of excursion sets with cor- related steps was attempted by Maggiore & Riott (2011), who used a path integral formalism to compute the first crossing rate for barriers of constant height. Unfortunately, their choice to expand around the uncorrelated Gaussian solution makes the calculations very involved, and its reli- ability becomes problematic for large masses. Moreover, it only works for one specific choice of filter (Top-Hat) and matter power spectrum (\( \Lambda \)CDM), where their results are within 10% of the correct answer for Gaussian walks. The same is true for other works following the same approach like Corasaniti & Achitouv (2011) (who were only able to con- sider linear barriers with small slope) and D’Amico et al. (2011) (who did not consider moving barriers, but focussed on the safer non-Gaussian to Gaussian ratio, and assessed the range of validity of their results).

### 4 HALO ABUNDANCES DIRECTLY FROM THE NONLINEAR FIELD

Although the excursion set approach was formulated to pre- dict the abundance of nonlinear objects from the initial fluc- tuation field, we can use it to predict halo abundances from the late time field as well. This is because equation (23) is valid even if \( p(b) \) is highly non-Gaussian. The problem is particularly simple because halos are often identified in the nonlinear field by finding a spherical or triaxial patch which is a fixed multiple of the background density, independent of halo mass (Despali et al. 2013). In effect, this means our ex- cursion set approach, applied to the nonlinear non-Gaussian field with a constant barrier, is an analytic model of the numerical halo finding algorithm.

This has an important consequence for studies of the halo distribution which seek to approximate the smoothed halo field as a Taylor series in quantities derived from the underlying matter distribution. If the Taylor series is in the matter overdensity only, then the halos are said to be locally biased with respect to the mass. Our analysis shows that the bias must be nonlocal since the mass overdensity is not the only quantity which matters: at the very least, the first deriva- tive of the matter field with respect to smoothing scale plays an important role in determining halo abundances, and this is expected to make the halo-mass bias \( k \)-dependent (Musso & Sheth 2012).

That said, the critical nonlinear overdensity is of order 100× the background. This is substantially (at least 10×) larger than the rms value of the field when smoothed on the typical halo scale, so it may be that the additional terms which come from constraining the slope are irrelevant. Since in this limit, our equation (23) reduces to equation (1), the halo mass function is very simply related to the probabili- ty distribution function of the nonlinearly evolved field. We are in the process of exploring this nonlinear excursion set approach further.

Our analysis also allows one to address a related ques- tion, having to do with the self-consistency of the approach. Namely, suppose we estimate halo abundances not from the initial field (as is usually done), but from a weakly evolved one. How does the prediction compare with that based on

© 0000 RAS, MNRAS 000, II-3
the initial field (the usual estimate). If the approach is self-consistent, these two estimates should agree.

Let $p(M|\delta \nu)$ denote the probability that a cell of volume $V_e$ placed randomly in the evolved distribution contains mass $M$. This distribution has mean $\bar{M} \equiv \rho V_e$, where $\rho$ is the comoving background density, so the Eulerian density is $1 + \delta \nu = \bar{M}/M$. Local models of the evolution from the initial Lagrangian density $\delta_i$ to the Eulerian one assume that $1 + \delta \nu$ is a deterministic invertible function of $\delta_i$. In perturbation theory, this means that

$$\int_\nu^\infty d\nu p(m|\nu) (m/\bar{m}) = \int_{\delta_i(M,V_e)}^\delta d\delta p(\delta|s(M)) \tag{28}$$

[Bernardeau et al. 2002; Lam & Sheth 2008]. (If halos were made of discrete particles, then this mass weighting is similar to only counting cells which are centered on particles of the distribution.) Halos correspond to large $M/\bar{M}$, for which $\delta_i(M,V_e) \to \delta_i(M)$. Since the right hand side here is the same as the right hand side of equation (14), the extra factor of $M/\bar{M}$ on the left hand side here shows that it is the mass-weighted Eulerian distribution which is related to the halo mass function. This demonstrates self-consistency at least at the higher masses where equation (14) is the appropriate limit. It will be interesting to explore if this self-consistency survives in modified gravity models, where, because the linear theory growth factor becomes $k$-dependent, even just the linearly evolved field is rather different from the initial one.

## 5 DISCUSSION

We derived an intuitively simple formal expansion for the first crossing distribution of random walks with correlated steps, in which walks are ordered by the minimum number of times they cross the barrier from below (equation 5). The nature of the correlations between the steps is determined by the statistics of the field (i.e., Gaussian or non-Gaussian) when it is smoothed, which itself depend on the form of the smoothing filter. The leading order term of this expansion, equation (25), is particularly simple. It only requires that when walks cross the barrier, they do so crossing upwards. Therefore, it requires knowledge of only the joint distribution of the walk height and its first derivative: in appropriately scaled units, these turn out to be independent of one another, making the analysis particularly simple.

Previous work has shown that, for Gaussian initial conditions, this approximation (i.e. neglecting all the other terms associated with walks with multiple zig-zags) leads to equation (17), which works well for all filters of current interest, and for all barriers which are monotonic functions of smoothing scale. Our equation (23) is a straightforward generalization of equation (17) to non-Gaussian fields: again, only the bivariate distribution of height and slope is required. In the large mass regime, our formula reduces to the even simpler form of equation (14), which depends on the distribution of the walk heights alone. In spite of the fact that perturbative non-Gaussian corrections individually blow up in this regime, this result is completely non-perturbative and exact, and it simply reflects the fact that those walks that reach the barrier in very few steps are very unlikely to cross it multiple times, because of the correlations.

Equation (23) is useful for excursion set models which assume that the initial fluctuation field was non-Gaussian; indeed, this was the original motivation for this study. However, the analysis worked out so easily – in particular, the fact that equation (23) does not assume that the non-Gaussianity is weak – that it can be used to predict the abundance of nonlinear objects from the nonlinear rather than the initial fluctuation field. We argued that this means that halo bias must be nonlocal in principle, although local bias may be a good approximation in practice. We also argued that our formulation demonstrates self-consistency of the approach, in the sense that applying it to the initial or the late-time field (i.e., the Lagrangian or Eulerian fields) yields the same estimate of halo abundances, at least at the higher masses where equation (14) is the appropriate limit. It will be interesting to explore if this self-consistency survives in modified gravity models, where, because the linear theory growth factor becomes $k$-dependent, even just the linearly evolved field is rather different from the initial one.

## ACKNOWLEDGMENTS

MM is supported by the ESA Belgian Federal PRODEX Grant No. 4000103071 and the Wallonia-Brussels Federation grant ARC No. 11/15-040. RKS is supported in part by nsf-ast 0908241.
\[ p(B; s) = e^D e^{-B^2/2} \sqrt{2\pi s} \] (B1)

where \( \mathcal{D} = \sum_{i=3}^{\infty} (|\Delta_i|c_i/|t|)(-\partial/\partial B)^i \). Expanding the exponential gives

\[ e^D = 1 + \sum_{i=3}^{\infty} (\Delta_i)^i/i! (-\partial B)^i + \frac{1}{2!} \sum_{i,j=3}^{\infty} (\Delta_i)^i (\Delta_j)^j/j! (-\partial B)^{i+j} + \ldots \] (B2)

The probability distribution can then be written in terms of the Hermite polynomials \( H_n(B) \equiv e^{B^2/2}(-\partial B)^n e^{-B^2/2} \), as

\[ p(B; s) = e^{-B^2/2} \sqrt{2\pi s} \left[ 1 + \sum_{i=3}^{\infty} \frac{\langle \Delta_i \rangle^i}{i!} H_i(B) \right. \]

\[ + \sum_{i,j=3}^{\infty} \frac{\langle \Delta_i \rangle^i \langle \Delta_j \rangle^j}{2! j!} H_i(B) H_j(B) + \ldots \] (B3)

which is the Gram-Charlier expansion usually referred to in the literature.

Although this expression formally correct, this expression is not convenient to deal with very large masses. In this regime, \( B \) can be so large that one might also have \( \langle \Delta_i \rangle^i \equiv B^3 \gg 1 \) (see D’Amico et al. (2011) for a detailed discussion), and in order to make reliable predictions one cannot truncate the series but needs to sum an infinite number of terms. In order to avoid doing this, it is convenient to resum the series above into an exponential and write

\[ p(B; s) = \frac{e^{W(B, s)}}{\sqrt{2\pi s}}, \] (B4)

where the function \( W \) is

\[ W(B; s) = \frac{B^2}{2} + \sum_{i=3}^{\infty} \frac{\langle \Delta_i \rangle^i}{i!} H_i(B) \]

\[ + \frac{1}{2!} \sum_{i,j=3}^{\infty} \frac{\langle \Delta_i \rangle^i \langle \Delta_j \rangle^j}{i! j!} h_{ij}(B) \]

\[ + \frac{1}{3!} \sum_{i,j,k=3}^{\infty} \frac{\langle \Delta_i \rangle^i \langle \Delta_j \rangle^j \langle \Delta_k \rangle^k}{i! j! k!} h_{ijk}(B) + \ldots \] (B5)

and where we have defined the modified polynomials

\[ h_{ij} \equiv H_{i+j} - H_i H_j, \] (B6)

\[ h_{ijk} \equiv H_{i+j+k} - H_i H_j H_k - (H_{i+j+k} + \text{perms.}) \]

\[ = H_{i+j+k} + 2H_i H_j H_k - (H_{i+j+k} + \text{perms.}) \] (B7)

\[ h_{i+j+k} \equiv H_{i+j+k} - H_i H_j H_k - (H_{i+j+k} + \text{perms.}) \]

and so on. At a first sight this is hardly going to help, since we are still dealing with an infinite series of terms that diverge when \( B \gg 1 \). However, one can check that \( h_{ij} \) has
degree $i + j - 2$, $h_{ijk}$ has degree $i + j + k - 4$, and similarly for higher order ones, so that $W(B; s)$ is a better behaved expansion when $B$ is large. Moreover, thanks to the exponential representation, truncating the expansion at any order is guaranteed to return a positive definite probability distribution.

This result has a nice interpretation in terms of Feynman diagrams. If one assigns a power of $B$ to each external leg and uses $(-1)^n(\Delta^n)c/n!$ as vertices and -1 as propagator, each Hermite polynomial in Eq. (125) represents the sum of all possible ways to connect the vertices listed in its coefficient with all possible combinations of external and internal lines and the correct combinatorial factors. For instance, $\langle \Delta^3 \rangle c H_3(B)$ represents the one tree-level graph with three external legs (whence $B^3$) and the three one-loop graphs with one external leg (whence $-3B$) containing just one cubic vertex. In this language, $W$ becomes the generator of the connected graphs; these are obtained removing from each $H_{ijk}$... all the disconnected pieces, that is the products of two or more lower order connected terms.

In the large-$B$ limit, it is consistent to approximate this expansion keeping the leading term of each polynomial (which is equivalent to neglecting loop diagrams order by order). However, the smaller $s$ gets, the higher is the order at which one can safely truncate the expansion. Up to 4th order one recovers

$$W(B; s) \simeq -\frac{B^2}{2} + \frac{\langle \Delta^3 \rangle_c B^3}{3!} + \frac{\langle \Delta^4 \rangle_c - 3\langle \Delta^3 \rangle^2}{4!} B^4$$

which is enough to describe the mass function over the range of scales of interest, as discussed by D’Amico et al. (2011).

Here, if the combinations of connected moments which appear in the expansion above were functions of $B$ only, then the resulting pdf would be self-similar in the sense used in the previous sections. That fact that they are not, in general, functions of the scaling variable $B$, means that the first term in equation (23) will result in an additional contribution to $f(s)$, which must be added to equation (17).

REFERENCES

Ade P., et al., 2013
Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys.Rep., 367, 1
Bond J., Cole S., Efstathiou G., Kaiser N., 1991, Astrophys.J., 379, 440
Corasaniti P., Achitouv I., 2011, Phys.Rev.Lett., 106, 241302
D’Aloisio A., Zhang J., Shapiro P. R., Mao Y., 2013
D’Amico G., Musso M., Norena J., Paranjape A., 2011, JCAP, 1102, 001
Despali G., Tormen G., Sheth R. K., 2013, MNRAS, 431, 1
Joudaki S., Lore O., Ferramacho L., Kaplinghat M., Santos M. G., 2011, Phys.Rev.Lett., 107, 131304
Lam T. Y., Sheth R. K., 2008, MNRAS, 386, 407
Lam T. Y., Sheth R. K., 2009, MNRAS, 398, 2143
LoVerde M., Miller A., Shandera S., Verde L., 2008, JCAP, 0804, 014
LoVerde M., Smith K. M., 2011, JCAP, 1108, 003
Maggiore M., Riotto A., 2010, Astrophys.J., 717, 526
Matarrese S., Verde L., Jimenez R., 2000, Astrophys.J., 541, 10
Musso M., Paranjape A., 2012, MNRAS, 420, 369
Musso M., Sheth R. K., 2012, MNRAS, 423, L102
Paranjape A., Lam T. Y., Sheth R. K., 2012, MNRAS, 420, 1429
Paranjape A., Sheth R. K., 2012, MNRAS, 426, 2789
Press W. H., Schechter P., 1974, Astrophys. J., 187, 425

© 0000 RAS, MNRAS 000, 1–9