Mon. Not. R. Astron. Soc. **000**, 1–5 (0000) Printed 28 April 2014 (MN LaTeX style file v2.2)

Stochasticity in halo formation and the excursion set approach

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*

28 April 2014

ABSTRACT
The simplest stochastic halo formation models assume that the traceless part of the shear field acts to increase the initial overdensity (or decrease the underdensity) that a protohalo (or protovoid) must have if it is to form by the present time. Equivalently, it is the difference between the overdensity and (the square root of the) shear that must be larger than a threshold value. To estimate the effect this has on halo abundances using the excursion set approach, we must solve for the first crossing distribution of a barrier of constant height by the random walks associated with the difference, which is now a non-Gaussian variate. The correlation properties of such non-Gaussian walks are inherited from those of the density and the shear, and, since they are independent processes, the solution is in fact remarkably simple. We show that this provides an easy way to understand why earlier heuristic arguments about the nature of the solution worked so well. In addition to modeling halos and voids, this potentially simplifies models of the abundance and spatial distribution of filaments and sheets in the cosmic web.

Key words: large-scale structure of Universe

1 INTRODUCTION

The abundance and spatial distribution of gravitationally bound objects is a sensitive probe of the nature of the initial conditions, the expansion history of the universe, and the nature of gravity. The simplest models of such objects, which we will call halos, assume that they form from the spherically symmetric collapse of sufficiently overdense spherical patches in the primordial fluctuation field (Gunn & Gott 1972). Building on insights from Press & Schechter (1974), the excursion set approach (Bond et al. 1991) provides a framework for linking halos to such overdense regions in the primordial field. In this approach, concentric spheres are assumed to remain concentric as the protohalo collapses around its centre of mass, so one is interested in the largest sphere whose mean overdensity $\delta_{\rm L}$ (assumed to be a Gaussian variate) exceeds a critical value δ_c.

However, halos are not spherical, and in the simplest models of non-spherical collapse, the shear field is assumed to play an important role (Bond & Myers 1996). Measurements of halo formation in simulations show that the shear field does indeed matter (Sheth et al. 2001): it acts to increase the overdensity required for collapse, approximately as

$$\delta_{\rm L} > \delta_c\,(1+\sqrt{q^2/q_c^2}) \qquad (1)$$

where q^2 is the traceless shear associated with the protohalo patch, and q_c is a parameter that determines how important the effects of the shear are relative to the spherical collapse model. Large q_c means that the shear must be large if it is to affect halo formation, and spherical collapse is recovered in the $q_c \to \infty$ limit. Measurements of protohalo patches in simulations suggest that $q_c^2 \sim 6\delta_c^2$ (Despali et al. 2013; Sheth et al. 2013).

The effect of the shear can be incorporated into the excursion set approach by searching for the largest scale on which equation (1) is satisfied. The analysis is simplified by the fact that, in a Gaussian random field, q^2 is not correlated with $\delta_{\rm L}$ (Sheth & Tormen 2002). However, analytic progress has been hampered by the fact that on each scale q^2 is not a Gaussian variate – if it were, the analysis would be simple (see Castorina & Sheth 2013) – but is drawn from a χ^2_5 distribution.

The first crossing problem can be solved numerically of course, by noting that $q^2 = \sum_{i=1}^5 g_i^2/5$ where the g_i's are independent Gaussian variates with $\langle g_i^2 \rangle = \langle \delta_{\rm L}^2 \rangle$, so the task reduces to generating the 6 Gaussian walks (one for $\delta_{\rm L}$ and the other five to obtain q), and checking at each

⋆ E-mail: marcello.musso@uclouvain.be
† E-mail: sheth@ictp.it

step if equation (1) is satisfied. This makes the first crossing problem appear to be six-dimensional, since it depends on six Gaussian walks. Accounting for the fact that each of these walks has correlated steps is an additional complication.

The main goal of the present work is to show that significant analytic progress can be made by noting that, if one defines $\delta \equiv \delta_{\rm L} - q(\delta_c/q_c)$, the multi-dimensional Gaussian problem reduces to that of the single non-Gaussian variate δ first exceeding δ_c. One can therefore make use of recent progress in our understanding of the correlated steps problem for non-Gaussian walks (Musso & Sheth 2012, 2014a). In Section 2 we show that, in fact, this particular problem is even simpler than that for generic non-Gaussian walks, because the walks which make up δ are themselves Gaussian. A final section compares our analysis with previous more heuristic approximations, and summarizes.

2 FIRST CROSSING DISTRIBUTION WITH CORRELATED STEPS

In the excursion set approach, one is interested in the probability that the average $\delta_{\rm L}(r)$ of the overdensity field over a sphere of radius r exceeds the threshold b, while for all $R > r$ it remains below b. As r changes, $\delta_{\rm L}(r)$ describes a random trajectory, whose value at given r has a Gaussian distribution with variance

$$s(r) \equiv \langle\, \delta_{\rm L}^2(r)\,\rangle = \int \frac{{\rm d}k}{k}\, \frac{k^3 P(k)}{2\pi^2}\, W^2(kr)\,, \tag{2}$$

where $P(k)$ is the power spectrum of δ, and $W(kr)$ is the Fourier transform of the filter that one uses to compute the mean value. The variance s grows monotonically as r gets smaller, starting from $s = 0$ at very large r.

In practice, it is 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. One then wants the probability $f(s)$ that $\delta(s) > b(s)$ at s but $\delta(S) < b(S)$ for all $S < s$. In general, imposing the first constraint is straightforward, whereas the second one is difficult to treat analytically. This difficulty is due to the fact that, for any choice of $W(kr)$ other than a step function in Fourier space, the steps of the walks are correlated with each other.

2.1 Up-crossing rather than first-crossing

Since a walk that is first crossing is necessarily reaching the barrier from below, one may begin to approach the problem imposing the less restrictive constraint that $\delta = b$ and the increment $v \equiv {\rm d}\delta/{\rm d}s$ of the walk with scale (the "velocity" of the walk) is larger than the increment $b' = {\rm d}b/{\rm d}s$ of the barrier. This formulation correctly discards walks that are crossing downwards at s, although clearly fails to discard those walks that are crossing upwards at s but had already done so at a some larger scale S (i.e. walks with more than one upcrossing). However, Musso & Sheth (2012) showed that at small s the fraction of such walks is tiny, since the correlations between steps make sharp turns very unlikely, and walks with more than one upcrossing necessarily take at least two turns. Therefore, the upwards approximation already provides a good approximation to $f(s)$ on the range of scales of interest for Cosmology. Corrections at small s, if needed can be computed as a a perturbative expansion in the number of times a walk crosses the barrier going upwards (Musso & Sheth 2014a) or, non perturbatively, imposing that $f(s)$ is normalized to unity (Musso & Sheth 2014b).

If earlier upcrossings can be neglected, then $f(s)$ can be computed from the joint probability $p(\delta, v; s)$ that a walk reaches δ at scale s with velocity v. In particular, since one only wants walks that are crossing the barrier upwards, that is $\delta = b(s)$ and $v \geq b'$ (for a barrier of constant height, this is just $v \geq 0$), the first crossing probability is well approximated by

$$f(s) \simeq f_{\rm up}(s) \equiv \int_{b'}^{\infty} {\rm d}v\, (v - b')\, p(b, v; s)\,, \tag{3}$$

where the factor of $v - b'$ can be understood as the density current of the upcrossing walks (Musso & Sheth 2012). This expression correctly reduces to that of Press & Schechter (1974) in the small-s limit.

Although for a Gaussian distribution evaluating this integral is straightforward, it is in general convenient to use the rescaled stochastic quantities

$$\Delta \equiv \frac{\delta}{\sqrt{s}}, \quad \Delta' \equiv \frac{{\rm d}\Delta}{{\rm d}s} \quad \text{and} \quad \xi \equiv -\frac{\Delta'}{\sqrt{\langle\, \Delta'^2\,\rangle}} \equiv -2\Gamma s\, \Delta' \tag{4}$$

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

$$\langle\, \Delta^2\,\rangle = \langle \xi^2\,\rangle = 1 \qquad \text{and} \qquad \langle\, \Delta\, \xi\,\rangle = 0\,; \tag{5}$$

i.e., Δ and ξ are uncorrelated (although in a generic non-Gaussian case not independent) random variables. Similarly, we will work with

$$B(s) \equiv \frac{b(s)}{\sqrt{s}} \qquad \text{and} \qquad X \equiv -\frac{{\rm d}B/{\rm d}s}{\sqrt{\langle\, \Delta'^2\,\rangle}} = -2\Gamma s\, B', \tag{6}$$

where $B' \equiv {\rm d}B/{\rm d}s$. The sign of X is chosen so that a barrier that does not vary much, as it is typically the case at small s, has $X > 0$.

2.2 Gaussian walks

If δ is a Gaussian process, then Δ and ξ are independent random variables and their joint distribution factorizes:

$$p(B, \xi) = p(B)\, p(\xi) = \frac{{\rm e}^{-B^2/2}}{\sqrt{2\pi}}\, \frac{{\rm e}^{-\xi^2/2}}{\sqrt{2\pi}}\,. \tag{7}$$

Inserting this in equation (3) shows that $f(s)$ will be

$$f_{\rm up}(s) = -B'\, p(B) \left[\frac{1 + {\rm erf}(X/\sqrt{2})}{2} + \frac{{\rm e}^{-X^2/2}}{\sqrt{2\pi} X}\right]. \tag{8}$$

where $p(B) = {\rm e}^{-B^2/2}/\sqrt{2\pi}$, which reduces to $-B' p(B)$ (the result of Press & Schechter 1974) when $X \gg 1$.

For a wide variety of smoothing filters, power-spectra and barrier shapes, $f_{\rm up}(s)$ remains a good approximation to $f(s)$ also down to scales on which a substantial fraction of the walks cross with negative slopes (Musso & Sheth 2012). However, it cannot be accurate to arbitrarily small scales since, for a constant barrier, the integral of $f_{\rm up}(s)$ over all s diverges. This is, of course, a consequence of the fact that multiple upcrossings of the barrier may become important as s increases; they need to be accounted for with the techniques formally described by Musso & Sheth (2014a), which are however difficult to evaluate exactly, or with the excellent and efficient numerical approximation of Musso & Sheth (2014b). However, roughly speaking one may expect these corrections – from walks with two or more turns – to be of the order of the square of those introduced by the square bracket term in equation (8), which mostly accounts for walks with just one turn. Since these are no larger than $10-15\%$ over most of the range of interest in Cosmology, then $f_{\rm up}$ is accurate up to $1-2\%$ on the small mass side of this range (and exact for large masses), so we will continue with this simpler case.

2.3 Non-Gaussian walks

Motivated by equation (1) we now consider the problem of finding the first crossing distribution of a barrier of constant height δ_c by the non-Gaussian variate

$$\delta \equiv \delta_{\rm L} - \beta q_n\,, \qquad \text{with } q_n^2 \equiv \sum_{i=1}^{n} \frac{g_i^2}{n}\,, \tag{9}$$

where $\beta \equiv \delta_c/q_c$, and $\delta_{\rm L}$ and the g_i's are zero-mean Gaussian variates with

$$\langle g_i^2\rangle = \langle \delta_{\rm L}^2\rangle \equiv s_{\rm L} \qquad \text{and} \qquad \langle g_i g_j\rangle = 0\,; \tag{10}$$

the mean and second moment of δ are thus

$$\langle\delta\rangle = -\beta\,\langle q_n\rangle \qquad \text{and} \qquad s \equiv \langle\delta^2\rangle = s_{\rm L}\,(1+\beta^2)\,. \tag{11}$$

Note that the variance of δ is $\sigma^2 \equiv s - \langle\delta\rangle^2$, so this differs from the usual definition of s as the variance of a zero-mean variate. We found it convenient to work in terms of s rather than σ, as its expression is simpler. Of course, covariance guarantees that the two choices are equivalent, with $f(\sigma^2) = {\rm d}s/{\rm d}\sigma^2 f(s)$. We will comment further on this point later.

One can then compute

$$\Delta = \frac{\delta}{\sqrt{s}} = \frac{\Delta_{\rm L} - \beta\, Q_n}{\sqrt{1+\beta^2}}\,, \tag{12}$$

where $\Delta_{\rm L} \equiv \delta_{\rm L}/\sqrt{s_{\rm L}}$ and $Q_n \equiv q_n/\sqrt{s_{\rm L}}$. While $\Delta_{\rm L}$ is a unit variance Gaussian process, it follows from its definition that Q_n is a Chi-variate with n degrees of freedom, whose distribution is

$$p_{\chi_n}(Q_n) = \frac{2}{Q_n}\left(\frac{nQ_n^2}{2}\right)^{n/2} \frac{e^{-nQ_n^2/2}}{\Gamma(n/2)} \tag{13}$$

where here Γ (not to be confused with the parameter of the walks!) denotes the Gamma function, and its mean value is

$$\langle Q_n\rangle = \sqrt{\frac{2}{n}}\,\frac{\Gamma(n/2+1/2)}{\Gamma(n/2)}\,. \tag{14}$$

Being the convolution of a Gaussian with a χ_n variate, the distribution of Δ is manifestly non-Gaussian:

$$p(\Delta) = \int_0^\infty {\rm d}Q_n\, \frac{e^{-\left(\sqrt{1+\beta^2}\Delta + \beta Q_n\right)^2/2}}{\sqrt{2\pi/(1+\beta^2)}}\, p_{\chi_n}(Q_n)\,. \tag{15}$$

Although the integral can be evaluated exactly, for $\beta < 1$ and $n > 1$ it is very well-approximated by a Gaussian with the same mean $\langle\Delta\rangle$ and variance $1-\langle\Delta\rangle^2$; this will be useful in the next section. In addition, it is worth noting that $\langle q_n\rangle \propto \sqrt{s_{\rm L}}$, so that the variance of δ is linearly proportional to $s_{\rm L}$.

The correlation structure of the Δ walks is inherited from those of $\Delta_{\rm L}$ and Q_n. Since $\Delta_{\rm L}$ is Gaussian, its correlation structure is simple (the joint distribution of $\Delta_{\rm L}$ and $\Delta'_{\rm L}$ factorizes as in equation 7), so the issue is the correlation structure of Q_n, which is determined by that of n independent Gaussian walks. Differentiating it, one gets

$$Q'_n \equiv \frac{{\rm d}Q_n}{{\rm d}s_{\rm L}} = \sum_{i=1}^{n} \frac{G_i G'_i}{n Q_n} \tag{16}$$

in terms of the unit variance Gaussian variates $G_i \equiv g_i/\sqrt{s_{\rm L}}$. Since $\langle G'_i G'_j\rangle = \delta_{ij}\langle \Delta_{\rm L}'^2\rangle$ and $\langle G'_i G_j\rangle = \langle G'_i Q_n\rangle = 0$, it follows that

$$\langle Q'_n\rangle = 0 \qquad \text{and} \qquad \langle Q_n'^2\rangle = \langle\Delta_{\rm L}'^2\rangle/n \tag{17}$$

or equivalently that

$$\Gamma_q^2 = n\,\Gamma_{\rm L}^2\,. \tag{18}$$

In order to compute the first crossing distribution, one needs the conditional of Q'_n given Q_n. As noted by Musso & Sheth (2014a) this will always be Gaussian, since Q'_n depends linearly on G'_i. However, the relations above imply that $\langle\, Q_n'^{2m} | Q_n\,\rangle = (2m-1)!!\,\langle Q_n'^2\rangle^m$ at fixed Q_n is actually independent of Q_n, and $\langle\, Q_n'^{2m+1} | Q_n\,\rangle = 0$. Therefore, not only are Q'_n and Q_n uncorrelated variables (as always), but they are also independent: their joint probability distribution factorizes, just like in the Gaussian case (equation 7). Furthermore, Q'_n is Gaussian, even though Q_n is not. Although this is a new and interesting result in its own right, in the present context it is just a step towards the quantity of real interest.

Now that we know the correlation structure of the non-Gaussian variate Q_n, we can address the problem of the Δ walks. Recalling that $s \equiv \langle\,\delta^2\,\rangle = s_{\rm L}\,(1+\beta^2)$, which makes ${\rm d}s/{\rm d}s_{\rm L} = s/s_{\rm L}$, one has

$$\Delta' \equiv \frac{{\rm d}\Delta}{{\rm d}s} = \left(\frac{s_{\rm L}}{s}\right)^{3/2} (\Delta'_{\rm L} - \beta Q'_n) \tag{19}$$

which means that, being the sum of two zero mean Gaussian variable, also Δ' is Gaussian with zero mean and variance

$$\langle\Delta'^2\rangle = (s_{\rm L}/s)^3 (1+\beta^2/n)\langle\Delta_{\rm L}'^2\rangle\,. \tag{20}$$

The relation between the parameter Γ for the non-Gaussian walks and the corresponding $\Gamma_{\rm L}$ is therefore

$$\Gamma^2 = \frac{1+\beta^2}{1+\beta^2/n}\,\Gamma_{\rm L}^2\,. \tag{21}$$

Furthermore, since the stochastic variables $\Delta'_{\rm L}$ and Q'_n are independent of $\Delta_{\rm L}$ and Q_n, so is Δ' of Δ. This means that $p(\Delta, \Delta')$ factorizes, like in the Gaussian case, with $p(\Delta')$ being Gaussian, even though $p(\Delta)$ is not. Therefore, $f_{\rm up}(s)$ is given by equation (8) with $p_{\rm G}(B)$ replaced by $p(B)$ from equation (15), and $\Gamma_{\rm L}$ replaced by Γ of equation (21). It is remarkable that the structure of the first crossing solution for these non-Gaussian walks is so similar to the Gaussian case. In particular, for these walks, neglecting the Hermite polynomial terms in equation (23) of Musso & Sheth (2014a) leads to the exact result.

Note that had we chosen to work with $B_\sigma \equiv b/\sigma$ rather than B (i.e. normalizing by the square root of the variance of δ instead of $\sqrt{s}$), then we would have defined $\Delta'_\sigma \equiv {\rm d}(\delta/\sigma)/{\rm d}\sigma^2$, and hence $\Gamma_\sigma^{-1} \equiv 2\sigma^2\sqrt{\langle{\Delta'_\sigma}^2\rangle}$. Although $\Gamma_\sigma \neq$

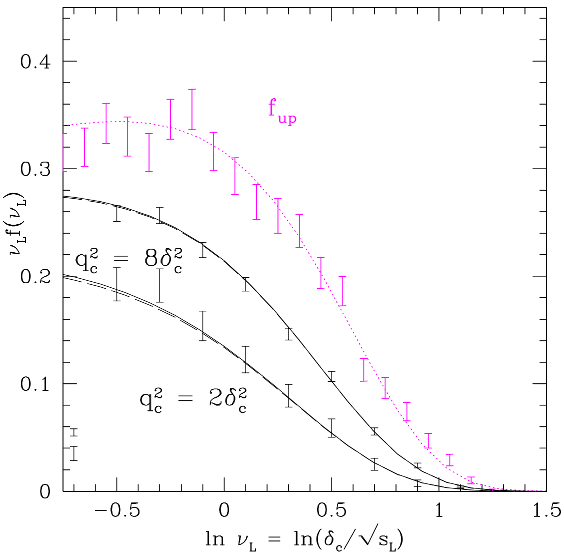

Figure 1. First crossing distribution of a barrier of constant height δ_c by Gaussian (upper) and non-Gaussian walks (lower) having correlated steps due to tophat smoothing of a CDM $P(k)$. Results for two values of the parameter q_c/δ_c, which determines the strength of the non-Gaussian component, are shown. The dotted curve shows equation (8) for a Gaussian distribution and $\Gamma_{\rm L}$, and the solid curves show equation (8) with the appropriate non-Gaussian distribution (equation 15) and Γ (equation 21). Dashed curves, almost indistinguishable from solid ones, show the result of setting $\Gamma = \Gamma_{\rm L}$ when defining X in equation (8).

Γ, $f_{\rm up}(s)$ depends not on Γ_σ but on $X_\sigma \equiv -2\Gamma_\sigma\sigma^2\,{\rm d}B_\sigma/{\rm d}\sigma^2$. A little algebra shows that $X_\sigma = ({\rm d}B_\sigma/{\rm d}\sigma^2)/\sqrt{\langle\Delta_\sigma'^2\rangle} = ({\rm d}B/{\rm d}s)/\sqrt{\langle\Delta'^2\rangle} = X$, so the final answer for $f_{\rm up}(s)$ does not depend on the normalization convention, up to the overall factor ${\rm d}s/{\rm d}\sigma^2$ needed to preserve the covariance of the distribution.

Finally, we stress that in order to derive this result we have assumed that β is constant. Had β depended on s, then Δ' would have been no longer independent of Δ, because of additional terms appearing in equation (19). Yet, expressing all occurrencies of Q_n in terms of Δ using equation (12), we could still have said that $p(\Delta'|\Delta)$ is Gaussian, although with a mean value that depends on Δ and a modified variance. However, a scale dependent $\beta(s)$ would likely signal that there are hidden stochastic variables in the problem that need to be made explicit (much like integrating over q_n in the present case would induce a non-vanishing scale dependent mean value for $\delta_{\rm L}$, as done by Sheth & Tormen 2002). Hence, we did not investigate this possibility further.

2.4 Comparison with previous work

The structure of our solution makes it easy to see why previous approximations to $f(s)$, based on heuristics, worked rather well. For instance, equation (A1) of Sheth et al. (2013) is motivated by equation (13) of Musso & Sheth (2012). The integral on the right hand side of their equation (A1) is the same as our equation (15). Therefore, their expression for $f(s)$ is the same as ours for $f_{\rm up}(s)$, except that they ignore the difference between $\Gamma_{\rm L}$ and Γ. This difference is small in the $\beta \ll 1$ limit where they were working, so they found good agreement with their Monte-Carlo solutions of this problem. This approximation is intuitively simple: the full $f_{\rm up}$ is a sum over the upcrossing distributions for walks crossing a constant barrier of height $\delta_c + \beta\,q_n$ with the contribution associated with q_n being weighted by $p_{\chi_n}(q_n)$.

A slightly different approximation, which yields additional insight, is that of Sheth & Tormen (2002). They argued that requiring $\delta \geq \delta_c$ is the same as requiring that $\delta - \beta\,\epsilon_q \geq \delta_c + \beta\,\langle q_n\rangle$, where $\epsilon_q \equiv q_n - \langle q_n\rangle$. At large n, $p(\epsilon_q)$ becomes approximately Gaussian with mean zero and variance $\langle\epsilon_q^2\rangle = s_{\rm L} - \langle q_n\rangle^2 \to s_{\rm L}/2n$ as $n\to\infty$. As a result $\delta - \beta\,\epsilon_q$ is approximately a Gaussian variate with mean zero and variance $\sigma^2 = s_{\rm L}\,(1+\beta^2) - \beta^2\langle q_n\rangle^2 \to s_{\rm L}\,(1+\beta^2/2n)$. Since $\langle q_n\rangle \propto \sqrt{s_{\rm L}}$, the quantity on the right hand side is like a deterministic 'moving' barrier $b_{\rm eff}(s_{\rm L})$, which the (approximately) Gaussian walks must cross. Therefore, $f_{\rm up}(s)$ should be well-approximated by equation (8) with $B = b_{\rm eff}(s_{\rm L})/\sqrt{\sigma^2(s_{\rm L})}$, $X = -2\Gamma_{\rm L}\sigma^2\,{\rm d}B/{\rm d}\sigma^2$ and $\Gamma_{\rm L}$ equal to the value for the purely Gaussian $\delta_{\rm L}$ walks. E.g., when $n = 5$ then $\langle q_5\rangle = (8/3)\sqrt{2s_{\rm L}/5\pi}$, so $\sigma^2 = s_{\rm L}\,[1+\beta^2 - \beta^2(128/45\pi)]$.

This line of reasoning led to equation (31) of Sheth et al. (2013), who showed that it was indeed in reasonable agreement with their Monte-Carlos. Like the previous approximation, this one ignores the fact that $\Gamma_\sigma \neq \Gamma_{\rm L}$, but in the limit where $\beta^2 \ll 1$ this difference matters little. Strictly speaking, equation (31) of Sheth et al. 2013 also ignores the fact that $\langle q_5\rangle/\sqrt{s_{\rm L}}$ is slightly different from unity, and that $s \neq s_{\rm L}$, since $\beta \ll 1$ anyway. Using the correct mean value removes most of the difference between the solid and dashed curves in their Fig.A1. And correctly using σ^2 rather than $s_{\rm L}$ makes this differ from their A1 only because it uses the Gaussian approximation instead of the full $p(B)$ – but we know this is a very good approximation.

The symbols in Figure 1 show the first crossing distribution, obtained by Monte-Carlo methods of walks whose steps are correlated because of TopHat smoothing of a CDM power spectrum, for which $\Gamma_{\rm L} \approx 1/3$. We show results for $q_c/\delta_c = 2, 8$ and ∞ (the limit in which the shear does not matter, so the non-Gaussian component does not contribute). The curves in Figure 1 show that equation (8), with the appropriate values of Γ, does indeed provide a very good description of the first crossing distribution. The dashed curves which lie slightly below the solid ones show equation (A1) of Sheth et al. (2013). We argued above that they differ from the solid ones only because they ignore the difference between Γ and $\Gamma_{\rm L}$: evidently, this only makes a small difference.

Stochasticity in the halo formation due to additional variables has also been considered, under the name of stochastic (or diffusive) barriers, by Maggiore & Riotto (2010) and Corasaniti & Achitouv (2011). In their work, the barrier is assumed to diffuse with uncorrelated steps around a mean value given by the barrier for ellipsoidal collapse studied by Sheth & Tormen (2002), and its distribution is Lognormal. Since there is no compelling reason for the Lognormal assumption (which is then approximated with a Gaussian anyway), and it is hard to justify the lack of correlations between the steps, we find the approach presented here more complete, and more closely tied to the physics of halo collapse.

3 DISCUSSION

We showed that the correlation structure of χ_n walks is remarkably similar to that of Gaussian walks: the probability of walk height and slope factorize similarly to how they do for the Gaussian case (equation 7). Therefore, when expressed in terms of the mean-square walk height s, the upcrossing distribution $f_{\rm up}(s)$ for χ_n walks is the same as that for Gaussian walks (equation 8), but with $p_G \to p_{\chi_n}$ and $\Gamma_{\rm L} \to n\,\Gamma_{\rm L}$ (equation 18).

We then used this result to derive an exact expression for the upcrossing distribution for walks which are obtained by convolving a Gaussian and a χ_n variate (equation 9); we argued that this is the problem which is relevant to stochastic barrier models such as equation (1), in which the (traceless) shear plays an important role in halo formation. As for the χ_n walks, our expression boils down to replacing the Gaussian pdf with the non-Gaussian one (equation 15), and rescaling the parameter Γ which describes the strength of the correlations between steps (equation 21), in the expression for Gaussian walks (equation 8).

The structure of our solution made it easy to see why previous approximations worked as well as they did, and their limitations (Figure 1 and related discussion). In contrast to what previous work assumed, our analysis shows that Γ is not the same as the corresponding $\Gamma_{\rm L}$ for Gaussian walks. However, the difference happens to be small for the stochastic barrier models of most interest, since these tend to have the shear (which contributes the non-Gaussian component) being less important than the density, so the effect on $f_{\rm up}$ is small (Figure 1).

However, even when the non-Gaussian component is large, our expression for $f_{\rm up}$ is much simpler than the corresponding expression for generic non-Gaussian walks (equations 23 or 25 in Musso & Sheth 2014a). Since it yields an excellent approximation to the first crossing distribution (Figure 1) we expect it will find use in analytic models of the cosmic web. In particular, Shen et al. (2006) have argued that whereas the 'moving' barrier associated with halo formation scales approximately as $\delta_c + 0.5\sqrt{s_{\rm L}}$, that for sheets scales as $\delta_c - 0.5\sqrt{s_{\rm L}}$. They argued that, as a result, the most massive halos are expected to be a substantial fraction of the mass of the filaments they populate, and these filaments a substantial fraction of the sheets in which they are embedded. Thus, massive halos will appear to be accreting most of their mass from filaments, whereas the growth of lower mass halos may be less obviously anisotropic. Recently, upon noting that the $\sqrt{s_{\rm L}}$ dependence is primarily due to the shear q (recall that $\langle q \rangle \propto \sqrt{s_{\rm L}}$), Sheth et al. (2013) have shown that this anisotropic growth manifests as nonlocal stochastic bias in the spatial distribution of massive halos. Since the $\delta_c - 0.5\sqrt{s_{\rm L}}$ scaling for sheets in the Shen et al. (2006) model is also primarily due to a dependence on q, our analysis here allows one to extend theirs to allow for walks with correlated steps, thus enabling more realistic models of the abundance and clustering of sheets as well as halos.

ACKNOWLEDGMENTS

MM is supported by the ESA Belgian Federal PRODEX Grant No. 4000103071 and the Wallonia-Brussels Federation grant ARC No. 11/15-040. He is grateful to the University of Pennsylvania for its hospitality in December 2013. RKS was supported in part by NSF-AST 0908241. He is grateful to the Perimeter Institute for its hospitality in September 2013, G. Dollinar and D. Bujatti for their hospitality in Trieste during October 2013, and the Institut Henri Poincare for its hospitality in November 2013.

REFERENCES

Bond J., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440, ADS
Bond J. R., Myers S. T., 1996, Astrophys.J.Supp., 103, 1, ADS
Castorina E., Sheth R. K., 2013, MNRAS, 433, L102, 1301.5128
Corasaniti P., Achitouv I., 2011, Phys.Rev.Lett., 106, 241302, 1012.3468
Despali G., Tormen G., Sheth R. K., 2013, MNRAS, 431, 1, 1212.4157
Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1, ADS
Maggiore M., Riotto A., 2010, ApJ, 717, 515, 0903.1250
Musso M., Sheth R. K., 2012, MNRAS, 423, L102, 1201.3876
Musso M., Sheth R. K., 2014a, MNRAS, 439, 3051, 1305.0724
Musso M., Sheth R. K., 2014b, MNRAS, 438, 2683, 1306.0551
Press W. H., Schechter P., 1974, ApJ, 187, 425, ADS
Shen J., Abel T., Mo H., Sheth R. K., 2006, ApJ, 645, 783, 0511365
Sheth R. K., Chan K. C., Scoccimarro R., 2013, Phys. Rev. D, 87, 083002, 1207.7117
Sheth R. K., Mo H., Tormen G., 2001, MNRAS, 323, 1, 9907024
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61, astro-ph/0105113