

On the Markovian assumption in the excursion set approach: the approximation of Markovian Velocities

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*

8 July 2014

ABSTRACT

The excursion set approach uses the statistics of the density field smoothed on a wide range of scales, to gain insight into a number of interesting processes in nonlinear structure formation, such as cluster assembly, merging and clustering. The approach treats the curve defined by the height of the overdensity fluctuation field when changing the smoothing scale as a random walk. The steps of the walks are often assumed to be uncorrelated, so that the walk heights are a Markov process, even though this assumption is known to be inaccurate for physically relevant filters. We develop a model in which the walk steps, rather than heights, are a Markov process, and correlations between steps arise because of nearest neighbour interactions. This model is a particular case of a general class, which we call Markov Velocity models. We show how these can approximate the walks generated by arbitrary power spectra and filters, and, unlike walks with Markov heights, provide a very good approximation to physically relevant models. We define a Markov Velocity Monte Carlo algorithm to generate walks whose first crossing distribution is very similar to that of TopHat smoothed Λ CDM walks. And we demonstrate that Markov Velocity walks generically exhibit a simple but realistic form of assembly bias, so we expect them to be useful in the construction of more realistic merger history trees.

Key words: large-scale structure of Universe

1 INTRODUCTION

Simulations of hierarchical gravitational clustering suggest that the abundance and clustering of gravitationally bound objects in the Universe can be powerful tools for constraining the nature of the initial fluctuation field. Since simulations are expensive, there is considerable interest in models that can provide a better understanding of how cluster abundances and clustering depend on cosmological parameters. The excursion set approach (Bond et al. 1991) is perhaps the most developed of these: motivated by the seminal work of Press & Schechter (1974) and Epstein (1983) it provides an analytical framework which relates the statistics of gravitationally bound dark matter haloes to fluctuations in the primordial density field, and the subsequent expansion history.

In this approach, at a given (randomly chosen) position in space one looks at the overdensity field smoothed on some scale R : plotting this smoothed δ as a function of (the in-

verse of) R resembles a random trajectory, the steps of which are, in general, correlated. The nature of the correlations depends on the smoothing filter (e.g. TopHat, Gaussian), and on the nature of the initial fluctuation field (Gaussian or non-Gaussian). Repeating this for every position in space gives an ensemble of trajectories, each one of which starts from $\delta(R = \infty) = 0$ (the Universe is homogeneous on large smoothing scales). For each trajectory, one searches 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), the value of which is determined by the expansion history of the background cosmology. 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 haloes is $(M/\bar{\rho}) dn/dM$, where $\bar{\rho}$ is the comoving background density. The excursion set approach assumes that this halo mass fraction equals the fraction of walks which cross the threshold (the “barrier”) for the first time when the smoothing scale is R :

$$f(R) dR = (M/\bar{\rho}) (dn/dM) dM. \quad (1)$$

* E-mail: marcello.musso@uclouvain.be

† E-mail: shethrk@physics.upenn.edu

Although recent work has focused on the shortcomings of this ansatz (Paranjape & Sheth 2012), not to mention the fact that variables other than the overdensity affect halo formation (Bond & Myers 1996; Sheth et al. 2001), and this is not evident in the simplest version of the approach outlined above (but see Sheth & Tormen 2002; Sheth et al. 2013), the first crossing distribution is nevertheless expected to provide substantial insight into the dependence of dn/dM on cosmological parameters.

In practice, one works not with $f(R)$ but with $f(s)$, where

$$s(R) \equiv \int \frac{dk}{k} \frac{k^3 P(k)}{2\pi^2} W^2(kR) \quad (2)$$

denotes the variance in the fluctuation field when smoothed on scale R with a filter of shape W . In hierarchical models, s is a monotonic function of R , so $f(R)dR = f(s)ds$. Working with s has the advantage of removing most of the dependence on the shape of the power spectrum: $P(k)$ mainly matters only through the dependence of s on R .

To solve the first crossing problem, we must be able to identify the fraction of trajectories for which $\delta = b(s)$ and $\delta < b(S)$ for all $S < s$. This is straightforward only when δ is a Markov process, allowing for exact analytic results for some barriers (Bond et al. 1991; Sheth 1998), and so much subsequent work has focused on this approximation. The more general case of walks with correlated steps has only recently seen real analytic progress, which is based on the realization that much of the constraint coming from the condition $\delta < b(S)$ for all $S < s$ is encapsulated by counting walks which are crossing the barrier from below (Bond et al. 1991; Musso & Sheth 2012). That is, if one defines the velocity $v \equiv d\delta/ds$, then a good approximation to $f(s)$, valid for all power-spectra, smoothing windows and barrier shapes of current interest, can be computed from the joint probability $p(\delta, v; s)$ that a walk reaches δ at scale s with velocity $v \geq db/ds$ (so that it overtakes the barrier). That is,

$$f(s) \simeq f_{\text{up}}(s) = p(b; s) \int_{b'}^{\infty} dv (v - b') p(v|b), \quad (3)$$

where $b' \equiv db/ds$. This simplicity also holds when the fluctuation field being smoothed is not Gaussian (Musso & Sheth 2014a).

In this approximation, the normalized quantity

$$\gamma^2 \equiv \frac{\langle v\delta \rangle^2}{\langle v^2 \rangle \langle \delta^2 \rangle} = \frac{1}{4s\langle v^2 \rangle} \quad (4)$$

plays an important role, because $p(b/\sqrt{s})$ is a normalized Gaussian with zero mean, and $p(v|b)$ is also Gaussian whose mean is $\langle v|b \rangle \equiv \langle v\delta \rangle b/s = b/2s$ and whose variance is $\langle v^2 \rangle (1 - \gamma^2) = 1/4\Gamma^2 s$, where $\Gamma^2 \equiv \gamma^2/(1 - \gamma^2)$. The additional constraint on v that comes from requiring walks to cross the barrier upwards makes equation (3) an essentially perfect description of the first crossing distribution down to scales of order $s \lesssim \Gamma^2 b^2(s)$.

Recently, Musso & Sheth (2014b) have shown that judicious use of the same constraint on v allows an even better approximation for $f(s)$. This approach sets

$$f(s) \simeq f_{\text{BS}}(s), \quad (5)$$

where f_{BS} is got by solving, via back-substitution, the ex-

pression

$$p(\delta \geq b(s); s) = \int_0^s dS f_{\text{BS}}(S) p(\delta \geq b(s)|\text{up } S), \quad (6)$$

where

$$p(\delta \geq b|\text{up } S) \equiv \frac{\int_{B'}^{\infty} dV (V - B') p(V|B) p(\delta \geq b|V, B)}{\int_{B'}^{\infty} dV (V - B') p(V|B)}, \quad (7)$$

and B and V denote the barrier height and the velocity of the walk on scale S . The new ingredient required by this approach is $p(b|B, V)$, the probability that the walk has height b on s when it is conditioned to have height B and velocity V on scale $S \leq s$. That is, in contrast to f_{up} , which only required knowledge of walk heights and velocities on the same scale, this one requires knowledge of correlations between two, but only two, different scales. To turn the approximation sign in our equation (5) into an equality, we would have to replace $p(b|\text{up } S)$ with $p(b|\text{first } S)$: the additional requirement that the walk was below B on all smoothing scales larger than V would introduce many more scales into the problem. Fortunately, Musso & Sheth (2014b) have shown that, in fact, $f_{\text{BS}}(s)$ is an essentially perfect description of $f(s)$ over all values of s .

The analysis that follows provides yet another example of why, in the context of random walks, thinking about v is so useful: this is because, generally speaking, the derivative of a non-Markovian diffusive process is often Markovian. Section 2 describes a toy model that is the next more complicated model to one in which there are no correlations between steps. We use this model to motivate a more general Markov Velocities model, which we formulate and develop in Section 3. Our Markovian Velocities model is particularly useful in cosmology since it provides a very good description of the first crossing distribution of walks having the full correlation structure resulting from TopHat smoothed Λ CDM power spectra: some of these applications are described in Section 4. A final section summarizes our results.

2 A SIMPLE ILLUSTRATIVE EXAMPLE

In this section we consider the first crossing distribution of discretized walks each having n infinitesimal steps, and heights $\delta_1, \dots, \delta_n$. If the underlying process is Gaussian, the joint distribution of the steps is entirely determined by the correlation matrix $C_{ij} \equiv \langle (\delta_i - \delta_{i-1})(\delta_j - \delta_{j-1}) \rangle$. Since $\delta_k = \sum_{i=1}^k (\delta_i - \delta_{i-1})$, the variance is $s_k = \sum_{i,j=1}^k C_{ij}$. Typically, one considers evenly spaced steps, for which $s_k = k \Delta s$. If the steps are equally spaced in variance, then the constraint $2 \sum_{i=1}^{k-1} C_{ik} + C_{kk} = s_k - s_{k-1} = \Delta s$ is always satisfied, and it simply means that in the continuum limit one always has $2\langle \delta v \rangle = \langle \delta^2 \rangle' = 1$. The analysis so far has been general.

2.1 Markov heights from uncorrelated steps

The usual approach to this problem assumes that C_{ij} is diagonal, with $C_{ij} = \Delta s \delta_{ij}$. The probability $p(\delta_1, \dots, \delta_n)$ of a walk can be equivalently expressed as the joint probability of the steps $\delta_i - \delta_{i-1}$ for $i \leq n$. The independence of the

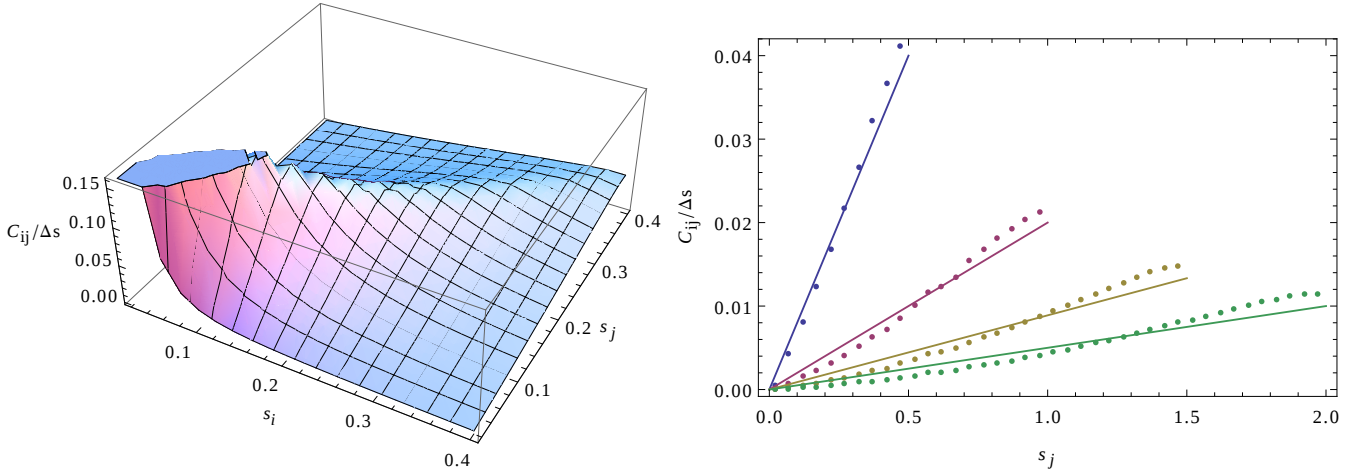


Figure 1. Left: correlation matrix $C_{ij} = \Delta s^2 \langle v_i v_j \rangle$ for the walk steps obtained by TopHat filtering a Λ CDM power spectrum. The correlation is maximal along the diagonal and decreases approximately linearly with distance from the diagonal. Right: the same C_{ij} , shown for a few fixed values of $s_i = 0.5, 1, 1.5, 2$ (symbols), and compared with the toy model $\langle v_i v_j \rangle = s_j / s_i^2$ described in Section 2.2 (lines).

steps makes the joint probability factorize as

$$p(\delta_1, \dots, \delta_n) = \prod_{i=1}^n p(\delta_i - \delta_{i-1}), \quad (8)$$

where each $\delta_i - \delta_{i-1}$ is an independent Gaussian variate with zero mean and variance $C_{ii} = \Delta s$.

Suppose we now consider a walk with k infinitesimal steps of width Δs and a finite one of width $(n - k)\Delta s$. Since these are $k + 1$ independent Gaussian variables, the conditional distribution of δ_n given all the other k steps is a Gaussian with mean δ_k and variance $s_n - s_k$. Since $p(\delta_n | \delta_k, \dots, \delta_1) = p(\delta_n | \delta_k)$, the heights at all steps previous to k do not matter, the walk heights are said to follow a Markov process. It is this Markovianity which vastly simplifies analysis of the first crossing distribution.

2.2 The next step: Markov steps

If the steps are not independent variables, as Musso & Sheth (2012) have emphasized, then in order to account for the correlations it is convenient to think in terms of the walk velocity $v_i = (\delta_i - \delta_{i-1}) / \Delta s$, for which the correlation matrix reads $C_{ij} = \Delta s^2 \langle v_i v_j \rangle$: each entry is now quadratic in Δs .

If C_{ij} is not diagonal, the next more complicated assumption is that the correlation between the velocities is maximal for $i = j$ and decreases to a minimum for $i = 0$ with constant slope: the correlation is directly proportional to the separation between the steps. Then the constraint implies $C_{ij} = \Delta s \min(i, j) / \max^2(i, j)$. This condition is equivalent to having $\langle v_i v_j \rangle = s_i / s_j^2$ for $s_i \leq s_j$. Figure 2 shows that, despite its simplicity, a matrix with this structure is actually rather similar to that associated with TopHat smoothing of a Λ CDM power spectrum. In the next section we use this similarity to motivate models which generalize this, our simplest toy model. But first, we will use this simple model to illustrate a number of rather interesting features of this Markov steps model.

First, we note that C_{ij} has determinant $|C| =$

$(\Delta s^n / n!) \prod_{i=1}^n [i^3 - (i-1)^3]$. Therefore, it is invertible, having inverse

$$C^{-1} = \frac{1}{\Delta s} \begin{bmatrix} x_1 & -y_2 & 0 & \dots \\ -y_2 & x_2 & -y_3 & \\ 0 & -y_3 & x_3 & \ddots \\ \vdots & & \ddots & \ddots \end{bmatrix}, \quad (9)$$

where $x_i = i^4 / [i^3 - (i-1)^3] + i^4 / [(i+1)^3 - i^3]$ for all $i < n$, $x_n = n^4 / [n^3 - (n-1)^3]$ and $y_i = i^2(i-1)^2 / [i^3 - (i-1)^3]$.

One can thus write the probability of the trajectory in a factorized form with only nearest-neighbour interactions:

$$p(\delta_1, \dots, \delta_n) = \prod_{i=1}^n p(\Lambda_i), \quad (10)$$

where we introduced the n independent Gaussian variates

$$\Lambda_i \equiv \delta_i - \delta_{i-1} - [(i-1)^2 / i^2] (\delta_{i-1} - \delta_{i-2}), \quad (11)$$

whose variance is

$$\langle \Lambda_i^2 \rangle = (\Delta s / i^4) [i^3 - (i-1)^3]. \quad (12)$$

Note that even though C_{ij} has no zero entries – so it is very different from the case of independent steps – the conditional distribution of one step only depends on the one immediately before, showing that this is a model where the steps rather than the heights are Markovian.

Suppose we again consider a walk with k infinitesimal steps of width Δs and a finite one of width $(n - k)\Delta s$. Since these are $k + 1$ Gaussian variables, the conditional distribution of $\delta_n - \delta_k$ is the Gaussian $p(\delta_n - \delta_k - w_i C_{ij}^{-1} (\delta_j - \delta_{j-1}))$, where $w_i \equiv \langle (\delta_n - \delta_k)(\delta_i - \delta_{i-1}) \rangle$ and $i, j \leq k$. In this particular case, $w_i \equiv \sum_{j=k+1}^n C_{ij} = C_{ik} \sum_{j=k+1}^n (k^2 / j^2)$, so that the mean still depends only on $\delta_k - \delta_{k-1} = \Delta s v_k$. This gives

$$p(\delta_n | \delta_k, \dots, \delta_1) = p(\delta_n - \delta_k - \psi_{nk} s_k v_k) \quad (13)$$

with no sum running over the k index, where

$$\psi_{nk} \equiv \frac{\langle (\delta_n - \delta_k) v_i \rangle}{s_k \langle v_k v_i \rangle} = \sum_{j=k+1}^n \Delta s \frac{s_k}{s_j^2} \rightarrow \left(1 - \frac{s_k}{s_n}\right); \quad (14)$$

notice that the formal definition of ψ_{nk} depends on i , but its explicit expression does not. Furthermore, the last statement holds in the continuum limit, when the sum can be replaced by an integral. The variance of the distribution is $\langle(\delta_n - \delta_k)^2\rangle - s_k \psi_{nk}^2$; in the continuum limit, since $\langle(\delta_n - \delta_k)^2\rangle = s_n - s_k - 2 \sum_{i=1}^k \sum_{j=k+1}^n C_{ij}$, it becomes $s_n(1 - s_k/s_n)^3$.

For future reference, we show that the same could have been proven by noting that, after some algebra, one has

$$\delta_n - \delta_k - \psi_{nk} s_k v_k = \sum_{i=k+1}^n \Lambda_i \sum_{j=i}^n \frac{i^2}{j^2}. \quad (15)$$

This expression is explicitly independent of any of the Λ_i 's with $i \leq k$, and therefore it factorizes from $p(\delta_1, \dots, \delta_k)$.

Also for future reference, we note that our requirement that $\langle v_i v_j \rangle = s_i/s_j^2$ for $s_i \leq s_j$ means that $\langle v_i^2 \rangle = 1/s_i$. This means that γ , the quantity which plays an important role in the analysis of f_{up} , equals $1/2$ for all scales, and $\Gamma^2 = \gamma^2/(1 - \gamma^2) = 1/3$.

Finally, note that the term on the right-hand side of equation (13) is the same object which plays a key role in equation (7) of the back-substitution approach. Since $p(b|B, V)$ here is independent of what happened on scales larger than S , it is in fact the same as $p(b|\text{first } S)$, which means that equation (5) should be an even better approximation for this toy model than it is in general.

2.3 Monte Carlo realizations

The factorized form of the joint probability (equation 10) makes the model introduced above particularly well suited for numerical simulation. This is because, as can be seen by setting $k = 0$ in the previous equation, the walk height after n steps is

$$\delta_n = \sum_{i=1}^n \Lambda_i \sum_{j=i}^n \frac{i^2}{j^2}, \quad (16)$$

where the Λ_i are the independent Gaussian variates of equation (10). This shows that δ_n depends only on the steps previous to it. In the large- n limit, the sum over j simplifies so that

$$\delta_n \simeq \sum_{i=1}^n i \Lambda_i \psi_{ni} \simeq \sum_{i=1}^n \sqrt{3\Delta s} g_i (1 - i/n) \quad (17)$$

where the final expression has used the large i limit for the variance $\langle \Lambda_i^2 \rangle \rightarrow 3\Delta s/i^2$, and the g_i are zero-mean unit-variance Gaussian random numbers. Thus, we may think of this as a process in which correlations between steps arise because of a smoothing window that has shape $\sqrt{3}(1 - S/s)$ when $S \leq s$ and is 0 otherwise. A smoothing window which is a step function leads to Markovian walk heights δ . For Gaussian fields, this suggests a deeper connection between the g_i and the Fourier modes of the field, which we explore in Section 3.5.

Before moving on, we note that the model above has $\langle \delta_i \delta_n \rangle = (3 - i/n)/2$ for $i \leq n$. One gets the same expression by setting $\kappa = 1/2$ in equation (90) of Maggiore & Riotto (2010). This is not just coincidence. Since $\kappa = 1/2$ is in any case required in their expressions to correctly reproduce the $S \rightarrow 0$ limit (in which f_{PS} is exact; their value $\kappa \simeq 0.45$ comes from a numerical fit), one might argue that our toy

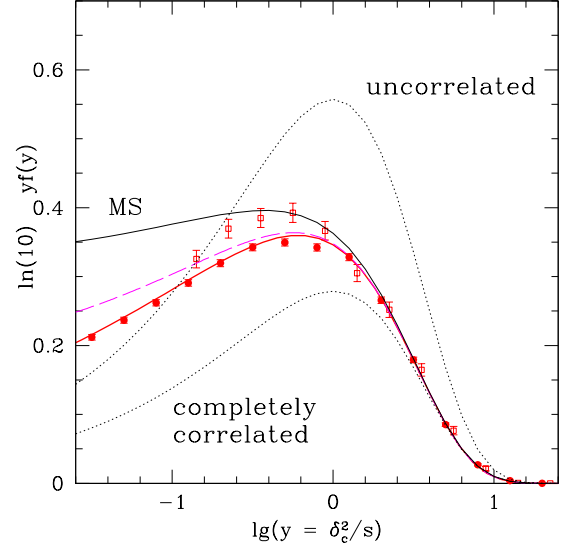


Figure 2. First crossing distribution $yf(y) = sf(s)$ of a constant barrier of height δ_c , by walks with $\langle v_i v_j \rangle = s_i/s_j^2$ for $s_i < s_j$. Filled circles show the Monte Carlo's distribution; thick solid curve shows the full back-substitution expression of Musso & Sheth (2014b); dashed curve shows equation (A5); and thinner solid curve shows the simpler approximation, $sf_{\text{up}}(s)$ of equation (3). The two dotted curves show the approximation of Press & Schechter (1974) and twice this value. Open squares show the result of Monte Carloing walks associated with TopHat smoothing of a Λ CDM power spectrum.

model is in fact the most natural leading order solution of the path integral approach; using walks with uncorrelated steps (Markovian heights) as the leading order approximation led them to significant complications. As we show below, our rather different approach to this problem is far more accurate and yields much more insight.

2.4 First crossing distribution

We noted that our simple model has $\gamma = 1/2$ for all scales. This value is especially significant for cosmology, because TopHat smoothing of a Λ CDM power spectrum has $\gamma \sim 1/2$. Therefore, one might ask if the first crossing distributions are similar. Figure 2 compares the first crossing distributions of a barrier of constant height obtained from Monte Carlo realizations of the walks for the process we have just described (filled circles) with several theoretical predictions for the same model and with Monte Carlo realizations of TopHat-smoothed Λ CDM walks (open squares). To guide the eye, the two dotted curves show $sf_{\text{PS}}(s) = (\nu/2) e^{-\nu^2/2}/\sqrt{2\pi}$ (the original result of Press & Schechter 1974) and $2sf_{\text{PS}}$ (the result of Bond et al. 1991 for uncorrelated steps), the thin solid curve shows f_{up} from equation (3) with $\gamma = 1/2$, and the thick solid curve shows the full back-substitution solution described by Musso & Sheth (2014b) (i.e., set $\gamma = 1/2$, $\xi = \sqrt{S/s}(3 - S/s)/2$ and $\Sigma/(\Gamma\xi) = (1 - S/s)/(1 - S/3s)$ in their equations 19 and 20). The agreement between the two sets of symbols at large δ_c^2/s is encouraging. Nevertheless, the fact that γ is mildly scale dependent for Λ CDM mo-

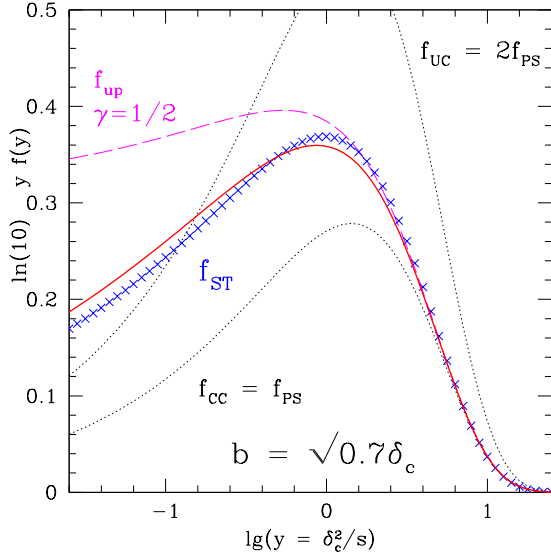


Figure 3. First crossing distribution $yf(y) = sf(s)$ of a constant barrier of height $\sqrt{0.7}\delta_c$, by walks with $\langle v_i v_j \rangle = s_i/s_j^2$ for $s_i < s_j$. Thick solid curve shows the full back-substitution expression of Musso & Sheth (2014b); the previous figure showed that it provides an excellent description of the Monte Carlo'd distribution. Dashed curve shows the simpler approximation, $sf_{\text{up}}(s)$ of equation (3). The two dotted curves show the approximation of Press & Schechter (1974) and twice this value. Crosses show the distribution associated with haloes in numerical simulations of structure formation in a Λ CDM universe. Our toy model provides a remarkably good description of the crosses.

tivates a generalization of the model to accommodate this scale dependence. We provide this in the next section.

Before doing so, we think it is interesting that our simple toy model actually provides a rather good description of the mass fraction in objects identified in numerical simulations of structure formation in a Λ CDM universe. The crosses in Figure 3 show the well-tested fitting formula for this quantity, from Sheth & Tormen (1999): $yf_{\text{ST}}(y) = 0.322 [1 + (s/b^2)^{0.3}] e^{-b^2/2s} / \sqrt{2\pi s/b^2}$. We argue later that this agreement motivates use of our toy model for generating fast Monte Carlo realizations of halo merger history trees because its walks have correlated steps whose first crossing distribution is a good match to measured halo abundances.

3 MORE GENERAL MODELS: MARKOVIAN VELOCITIES

The analysis of the previous section is specific to the particular form we assumed for the correlation matrix, for which $\langle v_i v_j \rangle = s_i/s_j^2$ for $s_i < s_j$. One could have instead assumed that, for instance, the velocity correlation does not decrease with s_i but remains constant. This happens if $C_{ij} = \Delta s/(2j+1)$ for $i \leq j$, or $\langle v_i v_j \rangle = 1/2s_j$ in the continuum limit. In this case too the same analysis would go through, the only difference being the value of ψ_{nk} . Therefore, the question arises as to whether the same can be done in more generality: one can see in fact that the same structure of equation (13) arises for any process for which ψ_{nk}

as defined by equation (14) does not depend on s_i , which happens for all models for which the dependence of $\langle v_i v_j \rangle$ on s_i and s_j factorizes.

We have already noticed that the factorization of the probability of the trajectory in equation (10) implies that the steps (and therefore in the continuum limit the velocities) are a Markov process, because the conditional probability of a step only depends on the one before; this is not true for the walk heights (the sums of the steps). In contrast, for sharp- k smoothing it is the walk heights that are a Markov process, while the steps are uncorrelated (the probability of each step is independent of the previous ones). We have also argued that the small differences between the toy model and TopHat-smoothed Λ CDM in Figure 2 are due to the fact that, for Λ CDM, $\langle v^2 \rangle$ differs from $1/s$. Therefore, it is interesting to see if one can capture a more general scaling of $\langle v^2 \rangle$, while preserving the nice features of a Markov process.

3.1 Langevin equation

A generic Markov process can be obtained as the solution of a Langevin equation sourced by a white noise η , such that $\langle \eta(s)\eta(S) \rangle = \delta_D(s-S)$. Let us consider a linear Langevin equation of the form

$$v' + \frac{\phi'}{\phi} v = \frac{\sqrt{\Phi'}}{\phi} \eta, \quad (18)$$

where the reason for the form of the functions in front of v and η will become clear shortly. This general equation is solved by

$$v(s) = \frac{1}{\phi(s)} \int_0^s dS \sqrt{\Phi'(S)} \eta(S), \quad (19)$$

for completely generic functions $\phi(s)$ and $\Phi(s)$. The white-noise properties of η mean that for $S \leq s$

$$\langle v(s)v(S) \rangle = \frac{1}{\phi(s)\phi(S)} \int_0^S dt \Phi'(t) = \frac{\Phi(S)}{\phi(s)\phi(S)}, \quad (20)$$

which indeed has the required factorized form.

To set Φ , recall that we are only interested in models that have

$$\langle v(s)\delta(s) \rangle = \int_0^s dS \langle v(s)v(S) \rangle = \frac{1}{2}. \quad (21)$$

This implies that

$$\Phi(S) = \frac{\phi(S)\phi'(S)}{2} = \frac{[\phi^2(S)]'}{4}, \quad (22)$$

making

$$\langle v(s)v(S) \rangle = \frac{\phi'(S)}{2\phi(s)}. \quad (23)$$

This expression for the correlation between velocities at two different scales generalizes the simple model discussed earlier for the off-diagonal elements of the correlation matrix C_{ij} , for which we had set $\langle v_i v_j \rangle = s_i/s_j^2$ (recall $s_i \leq s_j$). Since this expression is factorized, we can already see that equation (13) remains true in this very general case also. Of course, we have yet to specify ϕ .

To do so, we note that setting $S = s$ in equation (23) implies $2\langle v^2(s) \rangle = d \ln \phi(s)/ds$. Therefore

$$\phi(s) = \exp \left[2 \int_{s_0}^s dt \langle v^2(t) \rangle \right], \quad (24)$$

where s_0 is a constant whose exact value does not matter: it always cancels out because equation (18) really only depends on ϕ'/ϕ and $\Phi'/\phi^2 = [\phi^2]'/4\phi^2 = (\phi'/\phi)'/2 + (\phi'/\phi)^2$. This also means that equation (18) can be expressed entirely in terms of $\langle v^2(s) \rangle$:

$$v' + 2\langle v^2 \rangle v = \sqrt{4\langle v^2 \rangle^2 + \langle v^2 \rangle'} \eta, \quad (25)$$

showing how requiring the Markov process to reproduce the desired $\langle v^2 \rangle$ and to satisfy $\langle \delta v \rangle = 1/2$ determines both ϕ and Φ , and hence completely fixes $\langle v(s)v(S) \rangle$ and the evolution of v itself. The term under the square root sign on the right-hand is easy to remember: it is the variance of the quantity on the left-hand side. Hence, the most natural way to write equation (25) is to normalize the left-hand side by its variance, so that η is obviously unit variance as well.

What does all this imply for the walk heights $\delta(s) = \int_0^s ds v(S)$? Since v is itself an integral, integrating equation (19) and reversing the order of integrations yields

$$\delta(s) = \int_0^s ds \sqrt{\Pi(S)} \psi(s, S) \eta(S) \quad (26)$$

with

$$\psi(s, S) \equiv \frac{\phi(S)}{S} \int_s^S \frac{dt}{\phi(t)} = 2 \left[\frac{\langle \delta(s)\delta(S) \rangle}{S} - 1 \right], \quad (27)$$

where the last equality comes from integrating the two variables of equation (23) in the $[0, S]$ and $[S, s]$ range respectively, and

$$\Pi(S) \equiv \frac{\Phi'(S)}{\phi^2(S)} S^2 = [4\langle v^2(S) \rangle^2 + \langle v^2(S) \rangle'] S^2. \quad (28)$$

Equation (26) generalizes equation (16), where the correspondence is recovered upon setting $g_i/\sqrt{\Delta s} \rightarrow \eta(S)$; this suggests that it may be helpful to think of $\psi(s, S)$ as a window that smooths the white noise η to yield δ , and of $\Pi(S)$ as the analogue of the power spectrum. One can check that for the simple model described earlier, for which $\langle v^2 \rangle = 1/s$ and $\phi(s) \propto s^2$, one has $\Pi(S) = 3$ and $\psi(s, S) = 1 - S/s$.

Integrating v between S and s yields $\delta(s) - \delta(S)$. Using equation (19) for v one then gets

$$\delta(s) - \delta(S) - \psi(s, S) S v(S) = \int_S^s dt \sqrt{\Pi(t)} \psi(s, t) \eta(t), \quad (29)$$

which generalizes equation (15). Since the integral runs between S and s , and η is white noise, this quantity is independent of $\delta(t)$ and $v(t)$ for all $t \leq S$. Its probability is then the conditional probability of $\delta \equiv \delta(s)$ given $\Delta \equiv \delta(S)$ and $V \equiv v(S)$:

$$p(\delta|\Delta, V, \dots) = p(\delta - \Delta - \psi SV), \quad (30)$$

where the dots stand for any other combination of $\delta(t)$ and $v(t)$ at $t \leq S$. This is a Gaussian with mean

$$\langle \delta|\Delta, V, \dots \rangle = \Delta + \psi(s, S) SV. \quad (31)$$

Moreover, $\langle (\delta - \Delta - \psi SV)V \rangle = 0$ implies that

$$\psi(s, S) = \frac{\langle (\delta - \Delta)V \rangle}{S \langle V^2 \rangle} \quad (32)$$

just like in equation (14). The variance of $p(\delta|\Delta, V, \dots)$, which we call $C_{\delta\delta|\Delta V}$, is that of $\delta - \Delta - \psi SV$, that is

$$C_{\delta\delta|\Delta V} = \langle (\delta - \Delta)^2 \rangle - \frac{\langle (\delta - \Delta)V \rangle^2}{\langle V^2 \rangle}, \quad (33)$$

showing that for Markov Velocities the conditional distribution of δ at s given the whole trajectory prior to S is simply the conditional distribution of the step $\delta - \Delta$ given V . Therefore, equation (13) remains true, with ψ now given by equation (32).

Finally, we note that when $\psi \rightarrow 0$ then the walk height δ is Markovian, since $\delta - \Delta$ becomes an independent variable. (The most generic case, in which it is $\delta - [\langle \delta \Delta \rangle / S] \Delta$ that is independent, cannot be recovered from a Langevin equation for v alone.) If instead ψ is finite, equations (27) and (32) imply that

$$\langle \delta(s)\delta(S) \rangle = S [1 + 2\gamma_S^2 \langle (\delta - \Delta)V \rangle], \quad (34)$$

with $\gamma_S^2 = (4S\langle V^2 \rangle)^{-1}$ as defined in equation (4), which recovers the sharp- k correlation function $\langle \delta(s)\delta(S) \rangle = S$ when the second term vanishes. This happens not only when $\langle \delta V \rangle = \langle \Delta V \rangle = 1/2$, but also when $\gamma_S = 0$, that is when $\langle V^2 \rangle$ diverges. However, we stress that while if δ is Markovian then necessarily $\gamma = 0$, the converse is not always true: $\gamma = 0$ only implies that $p(\delta|\Delta, V) = p(\delta|\Delta)$, and not that $p(\delta|\Delta, V, \dots) = p(\delta|\Delta)$. Nevertheless, in order to compute $f(s)$ with the approximations used in this paper (quite good, as we have seen), processes with the same γ are barely distinguishable (especially for small γ , see Figure 4), and assuming that δ is Markovian is still satisfactory.

The expressions above show that $p(\delta|\Delta, V, \dots)$ differs from walks that are Markovian in δ because of the term ψSV . However, because $\langle V|\Delta \rangle = \Delta \langle V \Delta \rangle / \langle \Delta^2 \rangle = \Delta/2S$, we could also have written $\langle \delta|\Delta, V, \dots \rangle$ as $\langle \delta|\Delta \rangle + \psi(s, S) S (V - \langle V|\Delta \rangle)$, which instead emphasizes the connection to walks with completely correlated steps. In this case, the dependence of $p(\delta|\Delta, V, \dots)$ on $\delta - \langle \delta|\Delta \rangle$ is the piece identified by Bond et al. (1991) (their equation 5.8) and highlighted by Paranjape et al. (2012) (their equation 22); our expression shows that there is an additional piece which depends on how far V is from its mean value.

Similarly, one can write equation (19) as

$$v(s) - \frac{\phi(S)}{\phi(s)} v(S) = \int_S^s dt \frac{\sqrt{\Phi'(t)}}{\phi(s)} \eta(t). \quad (35)$$

The right hand side is independent of both Δ and V , indicating that also

$$p(v|\Delta, V, \dots) = p\left(v - \frac{\phi(S)}{\phi(s)} V\right) \quad (36)$$

is Gaussian with mean

$$\langle v|\Delta, V, \dots \rangle = \frac{\phi(S)}{\phi(s)} V = \frac{\langle vV \rangle}{\langle V^2 \rangle} V = \langle v|V \rangle, \quad (37)$$

as one could also have seen by differentiating equation (31) with respect to s , and variance

$$C_{vv|\Delta V} = \langle v^2 \rangle - \frac{\langle vV \rangle^2}{\langle V^2 \rangle}. \quad (38)$$

That $p(v|\Delta, V, \dots)$ does not depend on Δ indicates that the correlation between v and Δ arises entirely because of the individual correlations between v and V , and Δ and V . That is, in these models,

$$\langle v\Delta \rangle = \langle vV \rangle \langle V\Delta \rangle / \langle V^2 \rangle, \quad (39)$$

as it can be seen directly by taking the correlation of equation (35) with Δ .

Later on we will also need the conditional probability $p(\delta|v, \Delta, V, \dots)$, which is the distribution of a linear combination of $\delta - \Delta - \psi SV$ and $v - \langle v|V \rangle$ that is also uncorrelated with the latter. This combination is

$$\delta - \Delta - \psi SV - \frac{C_{\delta v|\Delta V}}{C_{vv|\Delta V}} \left(v - \frac{\langle vV \rangle}{\langle V^2 \rangle} V \right), \quad (40)$$

where $C_{\delta v|\Delta V}$ is the covariance of the two variables, and is a Gaussian variate with zero mean and variance

$$C_{\delta\delta|v\Delta V} = C_{\delta\delta|\Delta V} - \frac{C_{\delta v|\Delta V}^2}{C_{vv|\Delta V}}. \quad (41)$$

Similarly, $p(v|\delta, \Delta, V, \dots)$ is the distribution of

$$v - \frac{\langle vV \rangle}{\langle V^2 \rangle} V - \frac{C_{\delta v|\Delta V}}{C_{\delta\delta|\Delta V}} (\delta - \Delta - \psi SV), \quad (42)$$

having variance

$$C_{vv|\delta\Delta V} = C_{vv|\Delta V} - \frac{C_{\delta v|\Delta V}^2}{C_{\delta\delta|\Delta V}}. \quad (43)$$

Finally, since $v - \langle v|V \rangle$ is uncorrelated with both Δ and V , $C_{\delta v|\Delta V}$ is just its covariance with δ , that is

$$C_{\delta v|\Delta V} = \frac{1}{2} - \frac{\langle vV \rangle}{\langle V^2 \rangle} \langle \delta V \rangle = \frac{1}{2} - \langle vV \rangle S(2\gamma_s^2 + \psi). \quad (44)$$

3.2 Scale-invariant models

To see what all this implies, recall that the simplest Markovian velocities model had $\langle v^2 \rangle = 1/s$. We will call this a scale-invariant model because it makes $\gamma = (4s\langle v^2 \rangle)^{-1/2}$ independent of s . The analysis above shows that to produce models with other (constant) values of γ one simply sets $\langle v^2 \rangle = (4\gamma^2 s)^{-1}$ in equation (25). This means that $\phi = (s/s_0)^{1/2\gamma^2}$, so

$$\langle v(s)v(S) \rangle = \frac{(S/s)^{1/2\gamma^2-1}}{4\gamma^2 s} \quad (45)$$

and, for $\gamma^2 \neq 1/2$,

$$\psi(s, S) = \frac{1 - (S/s)^{1/2\gamma^2-1}}{1/2\gamma^2 - 1}, \quad (46)$$

while for $\gamma^2 = 1/2$ one finds $\psi(s, S) = \ln(s/S)$. In addition, $\Pi(S) = (1 - \gamma^2)/4\gamma^4 = (4\gamma^2\Gamma^2)^{-1}$ is also still independent of S , so that

$$\delta(s) = \int_0^s dS \eta(S) \frac{\psi(s, S)}{2\gamma\Gamma}, \quad (47)$$

which generalizes equation (16). Note that $\langle v(s)v(S) \rangle$ and $\psi(s, S)$ diverge when $S/s \rightarrow 0$ if $\gamma^2 > 1/2$. However, by construction $\gamma^2 \leq 1$, and $\langle \delta(s)\delta(S) \rangle \rightarrow 0$ in this limit.

This model provides a family of first crossing distributions, indexed by γ , which interpolate smoothly between the original formula of Press & Schechter (1974) (when $\gamma \rightarrow 1$) and twice this formula (when $\gamma \rightarrow 0$). Paranjape et al. (2012) identified these as being the limiting cases of walks with completely correlated and uncorrelated steps. Figure 4 shows this for a few representative choices of γ . Whereas f_{up} becomes an increasingly bad approximation to $f(s)$ as γ decreases (dashed curves), f_{BS} remains well behaved. Indeed, inserting equations (45) and (46) in the back-substitution

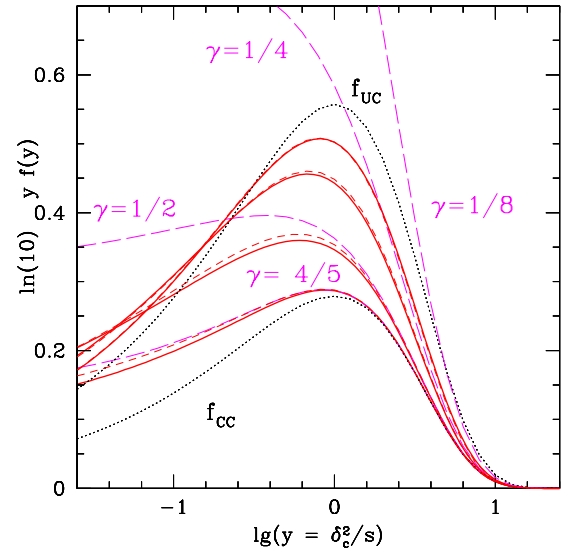


Figure 4. First crossing distribution $y f(y) = s f(s)$ of a constant barrier of height δ_c , by scale-invariant Markov velocity walks with the quoted values of γ . The two dotted curves show the approximation of Press & Schechter (1974) and twice this value (the distributions for walks with completely correlated and uncorrelated steps). Long-dashed curves show $s f_{\text{up}}(s)$ of equation (3) and thick solid curves show $s f_{\text{BS}}(s)$ of equation (6). The latter provide an excellent description of our Monte Carlos in all cases, correctly returning f_{PS} and $2f_{\text{PS}}$ as the limiting cases when $\gamma \rightarrow 1$ and 0, respectively, whereas f_{up} becomes an increasingly bad approximation as γ decreases. Short dashed curves show the back-substitution expression for Gaussian smoothed walks that have the same γ ; although they do not have Markov Velocities, their first crossing distributions are very well-approximated by those for Markov velocity walks.

algorithm of Musso & Sheth (2014b) yields excellent agreement with the Monte Carlo's $f(s)$ (so we have not shown the Monte Carlos).

We now turn to the question of how well such Markovian Velocity walks approximate walks with Gaussian or TopHat smoothing filters. One would in fact very much like to use Markovian velocity models to mimic generic walks, and, in particular, to provide a realistic approximation to TopHat smoothed Λ CDM walks.

A very important case is when $\gamma^2 = 1/6$, since this is also the value for TopHat smoothing of $P(k) \propto k^{-2}$. For this particular power spectrum and filter, the full correlation structure of the TopHat smoothed walk heights is given by

$$\langle \delta(s)\delta(S) \rangle = S[5 - (S/s)^2]/4 \quad \text{for } S \leq s. \quad (48)$$

This agrees *exactly* with the Markov Velocities model: set $\gamma^2 = 1/6$ in equation (46) and insert in (34). We therefore conclude that TopHat smoothing of $P(k) \propto k^{-2}$ is characterized by truly Markovian walk velocities. This is one of our key results. This case is however special: as we discuss in Section 3.5 below, TopHat smoothing of a different $P(k)$ does not lead to Markovian velocities.

Furthermore, Gaussian smoothing of $P(k) \propto k^n$ has $\langle \delta(s)\delta(S) \rangle = S[2/(1 + (S/s)^{1/\Gamma^2})]^\Gamma$, where $\Gamma^2 = (n+3)/2$. Since this is not of the form of equation (34) for any n , Gaussian smoothed walks never have Markovian velocities.

For instance, Gaussian smoothing of $P(k) \propto k^{-1}$ yields $\gamma^2 = 1/2$ and $\langle \delta \Delta \rangle = 2S/(1 + S/s)$, while the corresponding Markovian velocity model has $\psi = \ln(s/S)$ and thus $\langle \delta \Delta \rangle = S[1 - \ln(S/s)/2]$. The two differ by more than 10 per cent for $S/s < 0.2$.

Therefore, in order to reproduce a generic correlation structure with Markovian velocities, there are two possible choices: to match the velocity correlation parameter γ , or to match the amplitude of the spatial correlations – the $S/s \rightarrow 0$ limit of $\langle \delta(s)\delta(S) \rangle / S$ (the $S/s \rightarrow 1$ limit is, of course, always equal to unity). For Gaussian smoothing, the latter condition means $2\Gamma_G^2 = (1 - \Gamma_{MV}^2)^{-1}$: the two Γ s are not the same. We have found that matching γ produces slightly better agreement, most likely because the integral in equation (6) (which we solve numerically) receives little contribution at small S : so we have only shown these cases – using short dashed lines – in Figure 4. The agreement is rather impressive, suggesting that Markov Velocities are a rather good approximation to the full story.

We have thus shown that the first crossing distributions of scale-invariant Markov Velocity models are in good agreement with those of Gaussian or TopHat-smoothed walks having the same γ , over a wide range of scale-free $P(k)$. This raises the question of how well a Markovian velocity model can mimic TopHat-smoothed Λ CDM walks, for which $P(k)$ is not a simple power law. We address this in the next subsection.

3.3 Scale-dependent models

Although the scale-invariant models, being completely analytic, yield considerable insight, they are rather restrictive. Indeed, because equation (25) itself does not require $\langle v^2 \rangle \propto s^{-1}$, it allows us to model cases in which γ is a function of scale s . Thus, for example, we can choose the scale dependence of $\langle v^2(s) \rangle$ to be exactly the same as that of TopHat-smoothed Λ CDM walks. As before, this will not guarantee that the full correlation matrix $\langle v_i v_j \rangle$ will also be matched exactly, but, following the discussion above, we expect the resulting walks to provide a good approximation to the Λ CDM ones.

In particular, such walks have

$$\gamma(s) \approx a + b \ln(s/\delta_c^2) \quad (49)$$

with $a = 0.45$ and $b = -0.03$. Recalling that $\langle v^2(s) \rangle \equiv (4s\gamma^2)^{-1}$ means

$$\Pi(S) = \frac{1 - \gamma^2(S) - 2b\gamma(S)}{4\gamma^4(S)} \quad (50)$$

and

$$\psi(s, S) = \int_1^{s/S} d\tau \exp \left[-\frac{\ln \tau / \gamma(S)}{2(\gamma(S) + b \ln \tau)} \right]. \quad (51)$$

If we simply ignore the factor of $b \ln \tau$ in the denominator of the term in the exponential, then

$$\psi(s, S) \approx \frac{1 - (S/s)^{1/2\gamma^2(S)-1}}{1/2\gamma^2(S) - 1}, \quad (52)$$

which is equation (46) but with $\gamma \rightarrow \gamma(S)$. If we also drop the $2b\gamma(S)$ term from Π (recall $b \ll 1$), then even the ‘power spectrum’ is simply given by replacing $\gamma \rightarrow \gamma(S)$. Therefore,

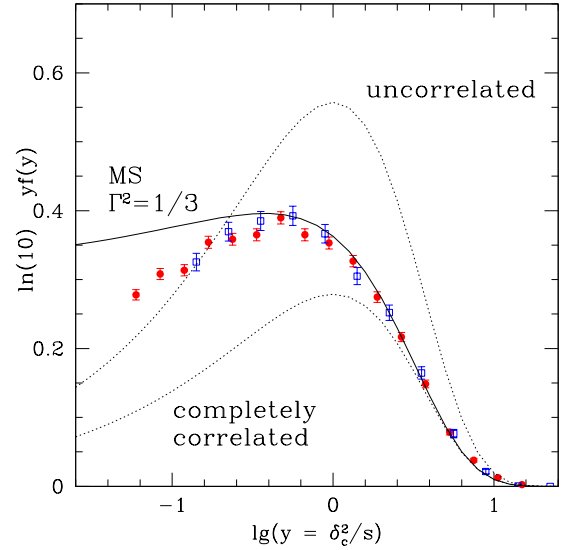


Figure 5. Comparison of the first crossing distribution of a constant barrier of height δ_c , by walks associated with the exact TopHat smoothing of a Λ CDM power spectrum (open squares, same as previous Figure), and by walks that have the same $\langle v^2(s) \rangle$ but Markovian velocities (filled dots). For comparison, the thin solid curve shows the simpler approximation of Musso & Sheth (2012) with $\Gamma^2 = 1/3$ (same as previous Figure) and the two dotted curves show the approximation of Press & Schechter (1974) and twice this value.

equation (47) with $\gamma \rightarrow \gamma(S)$ can be used to generate Monte Carlo realizations of walks.

Note that at $s = 1$, $\gamma = 0.48$ and $\Pi = 3.7$, which are not far from our original simplest model which had $\gamma = 1/2$ and $\Pi = 3$, and which Figure 2 showed worked reasonably well at $s \leq 1$. Since our γ decreases as s increases, we expect the associated first crossing distributions to become more like those for smaller γ (i.e., more negative n).

3.4 MVMC: Markov Velocity Monte Carlo

In practice, we do not generate walks using the approximate expression for ψ given in equation (52). Rather, we generate walks with the full structure (i.e. no dropping of $b \ll 1$ terms) by replacing equation (16) with

$$\delta_n = \sum_{j=1}^n \sum_{i=1}^j \sqrt{\Pi_i \Delta s} g_i \frac{i}{j^2} \frac{(S_i/\delta_c^2)^{1/2a\gamma_i-2}}{(S_j/\delta_c^2)^{1/2a\gamma_j-2}} \quad (53)$$

where $S_k = k \Delta s$.

Figure 5 shows a comparison of the first crossing distribution of a constant barrier of height δ_c , by walks associated with TopHat smoothing of a Λ CDM power spectrum (open squares, same as previous figure), and for walks which have the same $\langle v^2(s) \rangle$ but Markovian velocities. For comparison, the thin solid curve shows the simpler approximation of Musso & Sheth (2012) with $\Gamma^2 = 1/3$ (same as previous figure) and the two dotted curves show the approximation of Press & Schechter (1974) and twice this value. Clearly, our Markov velocities model provides an excellent approximation to the actual distribution.

Fundamentally, this is because the velocity correlations

are largest on the diagonal of C_{ij} , both for the Markov Velocities model and the actual TopHat smoothed walks. However, the toy model has a sharp edge while the real thing is smooth. By ensuring the Markov model has the same $\langle v^2(s) \rangle$, we are forcing the terms along the diagonal C_{ii} to be the same, but because $\langle \delta v \rangle = 1/2$, the Markov model C_{ij} is less than the real one for $S \sim s$, but is larger for $S \ll s$, so the two errors approximately compensate. This, – i.e., the need to set $\langle v^2(s) \rangle$ to that of the real walks – is why it was necessary to go beyond the simplest toy model we developed initially.

3.5 Markov velocities and Fourier space filters

We remarked earlier that one could think of the Markov Velocities model as one in which there is a smoothing kernel $\psi(s, S)$ applied to a power spectrum $\Pi(S)$. However, equation (2) shows that one can always think of $\delta(s)$ as a weighted sum over Fourier modes, with the variance on scale k given by the power spectrum. What does our Markov Velocities model imply for the window that smooths $P(k)$?

Let us first consider filters that are compact in Fourier space. For instance, if we set

$$W_\alpha(kR) \equiv [1 - (kR)^\alpha] \vartheta(1 - kR), \quad (54)$$

then for power-law power spectra, $P(k) \propto k^n$, the integrals which define $S(R)$ and $\langle \delta(s)\delta(S) \rangle$ can be done analytically, from which $\langle v(s)v(S) \rangle$ can also be derived. Comparison with the scale-invariant results derived previously shows that applying our truncated filters W_α of equation (54) to power-law $P(k)$ will result in a Markovian Velocity model having

$$1/\gamma^2 = 2 + 2\alpha/(3 + n). \quad (55)$$

Clearly, for more general $P(k)$, such as for the Λ CDM family, where the mapping between R and S is not a power law but is still monotonic, the shape of ψ depends on both s and S rather than S/s , as well as on α .

However, regardless of the shape of $P(k)$, differentiating equation (54) twice gives

$$\frac{d^2 W_\alpha}{dR^2} + \frac{1 - \alpha}{R} \frac{dW_\alpha}{dR} = \frac{\alpha}{R^2} \delta_D(1 - kR), \quad (56)$$

where now the right hand side only selects those Fourier modes with $k = 1/R$. It follows that the field filtered with $W_\alpha(kR)$ satisfies

$$\frac{d^2 \delta_\alpha}{dR^2} + \frac{1 - \alpha}{R} \frac{d\delta_\alpha}{dR} = \eta_\alpha, \quad (57)$$

where η_α is white noise. This is a Langevin equation for $d\delta_\alpha/dR$, which gives another Langevin equation for v when changing variables from R to s . Therefore, whatever the shape of $P(k)$, one can build a family of Markovian velocity models indexed by the value of α . The usual Markovian walk heights are recovered in the limit $\alpha \rightarrow \infty$, when the second derivative becomes irrelevant.

That Markovian velocities may arise from truncated Fourier filters makes intuitive sense, but the underlying reason is that the derivative is discontinuous. For example, filters of the form $(1 - |kR|)^\alpha \vartheta(1 - kR)$, while truncated, do not yield Markovian velocities because for them, the discontinuity appears in higher derivatives (e.g., $\alpha = 2$ yields Markovian accelerations, and so on). On the other hand,

filters of the form $|kR|^\alpha \vartheta(1 - kR)$ are already discontinuous and yield directly Markovian heights with $\langle \delta(r)\delta(R) \rangle = \langle \delta^2(R) \rangle (r/R)^\alpha$ (the usual sharp- k smoothing is the $\alpha \rightarrow 0$ limit of this more general Markov process). In general, if the n -th derivative of the filter is discontinuous, the n -th derivative of δ will be Markovian.

Even though the truncated filters in equation (54) can reproduce all possible (constant) values of γ , they are not the only ones which give rise to walks with Markov Velocities. For instance, smoothing $P(k) \propto k^{-2}$ with the TopHat filter

$$W_{\text{TH}}(kR) = (3/kR) j_1(kR) \quad (58)$$

yields the same correlation structure (equation 48 and $\gamma^2 = 1/6$) as the model obtained using equation (54) with $\alpha = 2$. It is then interesting to see when a TopHat filter, which has great physical interest but is clearly not truncated nor discontinuous in k , gives a Markov process.

Spherical Bessel functions satisfy the recurrence relation $x^m [-(1/x)(d/dx)]^m [j_1(x)/x] = j_{m+1}(x)/x$, from which it can be seen that a TopHat-filtered δ_{TH} obeys the differential equations

$$\frac{d\delta_{\text{TH}}}{dR} = \eta_1 \quad \text{and} \quad \frac{d^2 \delta_{\text{TH}}}{dR^2} - \frac{1}{R} \frac{d\delta_{\text{TH}}}{dR} = \eta_2. \quad (59)$$

Upon noting that $\delta_{\text{TH}} \equiv \eta_0$, the statistics of the field and its derivatives are described by the correlation functions

$$\langle \eta_a(r) \eta_b(R) \rangle = \int dk \frac{9(-k)^{a+b} P(k)}{2\pi^2 r R} j_{a+1}(kR) j_{b+1}(kr). \quad (60)$$

When $a = b = 0$, this returns $\langle \delta_{\text{TH}}(r) \delta_{\text{TH}}(R) \rangle$. This scheme can be extended to any higher derivative, yielding differential equations for $d\delta_{\text{TH}}/dR$ of arbitrary order.

For a power-law $P(k) \propto k^n$, one sees that η_1 behaves like white noise when $n = 0$ (as the integral above reduces to the closure equation for spherical Bessel functions), making δ_{TH} a Markov process. For $n = -2$, it is η_2 that becomes white noise, while the correlation function of η_1 has a factorized form: therefore, it is $d\delta_{\text{TH}}/dR$, the velocity, which is a Markov process. And indeed, in this case, the differential equation for η_2 is the same as $\alpha = 2$ in equation (57). (For $n = -4$ TopHat smoothing has Markovian accelerations, and so on.) For all other values of n , strictly speaking, neither process is Markovian. However, one can check that for $-1 \leq n < 0$ the variance $\langle \eta_1^2(R) \rangle$ diverges, making $\gamma = 0$. In this case, for single-scale approximations such as the up-cross and back-substitution approximations described in this paper, δ_{TH} effectively behaves like a Markov process. In contrast, for $-3 \leq n < 1$ it is $\langle \eta_2^2(R) \rangle$ that diverges: in these cases, the single-scale approximations will be well described by Markov Velocity models.

4 APPLICATIONS TO GALAXY FORMATION AND COSMOLOGY

As we have seen, Markov Velocities arise quite naturally when smoothing with a TopHat filter a field having power spectrum close to that of Λ CDM. This makes these models very useful for studying the evolution and clustering of structures. In this section, we explore several applications of Markov Velocity models, relying on their key property that the conditional probability of δ and v at one scale may

only depend on Δ and V at one larger scale. That is to say, in Markov Velocity models, at fixed halo mass and slope of the initial density profile (measured on the mass scale of the protohalo), there should be no correlation between formation history and the larger scale environment. We also show that the converse is also true: at fixed mass and slope, the density profile on large scales (i.e., the environment) is statistically the same regardless of the shape of the inner profile (i.e., the formation history).

4.1 Correlations with larger scales

Let $f(s|\Delta, S)$ denote the fraction of walks which first cross $b(s)$ on scale s , subject to the constraint that the walk had height Δ on scale $S < s$. Then $f(s|\Delta, S) \approx f_{\text{up}}(s|\Delta, S)$ where

$$f_{\text{up}}(s|\Delta, S) = p(b|\Delta) \int dv (v - b') p(v|b, \Delta) \quad (61)$$

(Musso et al. 2012). The similarity to equation (3) means that $f_{\text{up}}(s|\Delta, S)$ equals $p(b|\Delta)$ times a correction factor which depends on the mean and variance of $p(v|b, \Delta)$.

The conditional distribution $p(v|b, \Delta)$ is a Gaussian with mean

$$\langle v|b, \Delta \rangle = \langle v|\Delta \rangle + \frac{C_{v\delta|\Delta}}{C_{\delta\delta|\Delta}} (b - \langle \delta|\Delta \rangle) \quad (62)$$

and variance

$$C_{vv|\Delta} = C_{vv|\Delta} - \frac{C_{v\delta|\Delta}^2}{C_{\delta\delta|\Delta}}, \quad (63)$$

where, for our Markov Velocity models,

$$\begin{aligned} \langle v|\Delta \rangle &= \langle V|\Delta \rangle \langle vV \rangle / \langle V^2 \rangle, \\ C_{v\delta|\Delta} &= \frac{1}{2} \left(1 - \frac{\langle \delta\Delta \rangle \langle vV \rangle}{\langle \Delta^2 \rangle \langle V^2 \rangle} \right), \\ C_{\delta\delta|\Delta} &= s - \langle \delta\Delta \rangle^2 / S, \\ C_{vv|\Delta} &= \langle v^2 \rangle \left(1 - \frac{\langle vV \rangle^2}{\langle v^2 \rangle \langle V^2 \rangle 4S \langle V^2 \rangle} \right). \end{aligned} \quad (64)$$

The amount by which the ratio $f(s|\Delta, S)/f(s)$ differs from unity is often used to estimate the bias that comes from the large scale environment; the $S \ll s$ limit, in which $b \gg \langle \delta|\Delta \rangle$ yields what is known as the bias on large scales. As discussed in detail by Musso et al. (2012), this bias will be a constant in the $k \rightarrow 0$ limit, but the dependence on v will lead to k -dependent corrections (the leading order term being $\propto k^2$). Both the $k \rightarrow 0$ constant and the amplitude of the k -dependent corrections depend on s .

4.2 Conditional progenitor distributions

The expression above only requires that the walk have height Δ on scale S . Instead, let $f(s|S)$ denote the conditional distribution of first reaching $b(s)$ on scale s given that the walk first crossed $B(S)$ on scale S . Following Bond et al. (1991), this is the quantity which most excursion set approaches use to model the distribution of progenitors from an earlier time that become part of a more massive object later (e.g. Lacey & Cole 1993; Sheth & Tormen 2002).

For walks with uncorrelated steps, $f(s|S)$ has the same form as the unconditional distribution, upon rescaling $\delta \rightarrow$

$b - B$ and $s \rightarrow s - S$. Moreover, it also has the same form as $f(s|\Delta = B, S)$ of the previous section. For walks with correlated steps, however, the analysis is more complicated, and the two quantities are not the same – the additional ‘first crossing’ constraint on the larger scale matters.

The same logic that leads to equation (3) for the unconditional distribution suggests that

$$f(s|S) \simeq f_{\text{up}}(s|S), \quad (65)$$

where

$$f_{\text{up}}(s|S) \equiv \frac{p(B)}{f_{\text{up}}(S)} \int_{B'}^{\infty} dV (V - B') p(V|B) f_{\text{up}}(s|B, V) \quad (66)$$

and

$$f_{\text{up}}(s|B, V) = p(b|B, V) \int_{b'}^{\infty} dv (v - b') p(v|b, B, V) \quad (67)$$

is the upcrossing (rather than first crossing) distribution for walks conditioned to start from B with velocity V (see equation 10 and associated discussion in Musso & Sheth 2012). (Strictly speaking, one might replace $f_{\text{up}}(s|B, V)$ with the corresponding $f_{\text{BS}}(s|B, V)$, for the same reasons one might replace f_{up} with f_{BS} of equation (5), but we will not do so below.) This shows that $f_{\text{up}}(s|S)$ is a sum over upcrossing distributions $f_{\text{up}}(s|B, V)$ conditioned to have different slopes V when they upcrossed, and weighted by the probability of upcrossing with that slope.

To compute $f_{\text{up}}(s|B, V)$, we first note that equation (67) has the same structure as equation (3), with the unconditional distributions there being replaced by conditional ones: $p(b) \rightarrow p(b|B, V)$ and $p(v|b) \rightarrow p(v|b, B, V)$. These Gaussian distributions are given by equation (30) and equation (42) respectively. The similarity to equation (3) means that $f_{\text{up}}(s|B, V)$ equals $p(b|B, V)$ times a correction factor that depends on $[b' + \langle v|b, B, V \rangle] / C_{vv|\delta\Delta V}^{1/2}$.

Although equation (42) is written in a form which highlights the connection to walks with correlated steps, because $\langle v|\Delta, V \rangle = \langle v|V \rangle$ does not depend on Δ , $\langle v|\delta, \Delta, V \rangle$ is the sum of two terms, one of which is proportional to V and the other to $\delta - \Delta$. That δ and Δ do not appear separately, but only as $\delta - \Delta$ shows the close connection of our Markov Velocities process to one which is Markovian in δ ; the additional dependence on V indicates that the walks are non-Markovian in δ . We made a similar remark when discussing equation (31). As Musso & Sheth (2012) note, this dependence on V , this non-Markovian behaviour, shows explicitly that our Markovian Velocities model comes with assembly bias effects built-in.

Before we discuss these effects, to get some intuition about the shape of $f_{\text{up}}(s|S)$, it is instructive to set V equal to its mean value $B/2S$. This makes $\delta - \langle \delta|\Delta, V \rangle \rightarrow \delta - \langle \delta|\Delta \rangle$. If we now move $f_{\text{up}}(s|B, V = B/2S)$ out of the integral over V in equation (66), this makes the numerator there equal $f_{\text{up}}(s|B, V = B/2S) f_{\text{up}}(S)$, so that $f(s|S) = f_{\text{up}}(s|B, V = B/2S)$. Since $f_{\text{up}}(s|B, V = B/2S) \propto p(b|B, V = B/2S)$ at $S/s \ll 1$, the result depends on the scaling variable $(b - \langle \delta|B \rangle) / \sqrt{s(1 - S/s)^3}$. In contrast, the simplest well-motivated approximation for the conditional distribution (equation 22 of Paranjape et al. 2012) is $(b - \langle \delta|B \rangle) / \sqrt{C_{\delta\delta|\Delta}}$. Our analysis provides a simple way to understand why that approximation works rather well: it boils down to fixing V

to its correct mean value and ignoring the fact that the constrained scatter around this mean is narrower than the unconstrained value, but this difference is vanishingly small as $S/s \ll 1$.

4.3 Assembly bias

Assembly bias effects associated with Markov Velocity models, while present, are particularly simple. To see why, consider walks which first crossed the barrier on scale S . Let Δ_0 denote the value of the field on large scales $S_0 < S$, and δ and v the values on some smaller scale $s > S$. Then $p(\delta, v|\Delta, \Delta_0)$ equals $p(\delta, v|\Delta)$ if the walk heights were Markov, but in general, and for our Markov Velocity walks in particular,

$$p(\delta, v|\Delta, \Delta_0) \neq p(\delta, v|\Delta). \quad (68)$$

The dependence of δ and v on the large scale Δ_0 is a manifestation of Assembly bias.

Now consider $p(\delta, v|\Delta, V, \Delta_0)$. We have already argued that, for Markov Velocity models,

$$p(\delta, v|\Delta, V, \Delta_0) = p(\delta, v|\Delta, V), \quad (69)$$

illustrating that, if Δ and V are specified, there is no correlation with the large scale S_0 . Therefore, although assembly bias is present (equation 68) – the analysis of the previous subsection showed that the mean values of $p(b|B, V)$ and $p(v|b, B, V)$ in $f_{\text{up}}(s|B, V)$ are shifted – these effects are particularly simple. In particular, if both B and V have been specified, there are *no* additional correlations with the larger scale environment.

Now, write the left-hand side of equation (69) as $p(\delta, v, \Delta, V, \Delta_0)/p(\Delta, V, \Delta_0)$, and the right-hand side as $p(\delta, v, \Delta, V)/p(\Delta, V)$. Then equation (69) implies that

$$\frac{p(\Delta_0|\delta, v, \Delta, V) p(\delta, v, \Delta, V)}{p(\Delta, V, \Delta_0)} = \frac{p(\delta, v, \Delta, V)}{p(\Delta, V)}, \quad (70)$$

that is

$$p(\Delta_0|\delta, v, \Delta, V) = \frac{p(\Delta, V, \Delta_0)}{p(\Delta, V)} = p(\Delta_0|\Delta, V), \quad (71)$$

which is explicitly independent of δ and v . Note that $p(\Delta_0|\Delta, V)$ is Gaussian, with mean

$$\langle \Delta_0|\Delta, V \rangle = \frac{\langle \Delta_0 \Delta \rangle}{S} \Delta + \frac{\langle \Delta_0 (V - \Delta/2S) \rangle}{\langle V^2 \rangle (1 - \gamma_s^2)} \left(V - \frac{\Delta}{2S} \right). \quad (72)$$

For scale-free $P(k)$ the term which multiplies $V - \Delta/2S = -S_0\psi(S_0, S)$, showing explicitly that if $V > \Delta/2S$ then the dependence on V acts to decrease the large scale Δ_0 .

To illustrate these effects, Figure 6 shows $\langle \Delta|\delta, v \rangle / \delta$ for a range of choices of v , for TopHat smoothing of $P(k) \propto k^{-2}$. In this case $\gamma^2 = 1/6$, so

$$\langle \Delta|\delta, v \rangle = \delta \frac{S}{4s} \left[5 - \frac{S^2}{s^2} - \left(1 - \frac{S^2}{s^2} \right) \left(\frac{v}{\delta/2s} - 1 \right) \right] \quad (73)$$

and

$$\langle \delta|\Delta, V \rangle = \Delta + S \frac{1 - (S/s)^2}{2} V, \quad (74)$$

where our notation is for $S < s$. The steeper curves in the figure are associated with larger v : clearly, large overdensity at $S/s \geq 1$ is associated with less dense environments

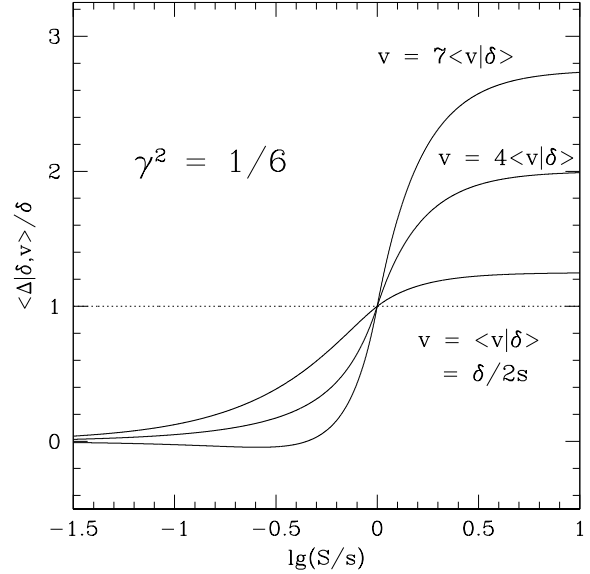


Figure 6. Assembly bias associated with TopHat smoothing of $P(k) \propto k^{-2}$. Curves show the mean walk height on scale S subject to the constraint that on scale s the walk has height δ and slope v . The three choices of $v = 1, 4, 7 \times \langle v|\delta \rangle$ show that steeper walks are associated with less dense large scale environments.

(i.e., smaller heights at $S/s \leq 1$. The other quantities which matter for assembly bias related quantities are

$$\begin{aligned} \langle v|\Delta, V \rangle &= (S/s)^3 V, & C_{vv|\Delta V} &= \frac{1 - (S/s)^5}{2s/3}, \\ C_{\delta\delta|\Delta V} &= s(1 - S/s)^3 \left(1 + \frac{9S}{8s} + \frac{3S^2}{8s^2} \right), \\ C_{\delta v|\Delta V} &= \frac{1}{2} - \frac{5S^3}{4s^3} + \frac{3S^5}{4s^5}. \end{aligned} \quad (75)$$

To see that these effects differ quantitatively for different Markov Velocity models, note that the simplest toy model ($\gamma = 1/2$) has

$$\begin{aligned} \langle \Delta|\delta, v \rangle &= \delta \frac{S}{2s} \left[3 - \frac{S}{s} - \left(1 - \frac{S}{s} \right) \left(\frac{v}{\delta/2s} - 1 \right) \right], \\ \langle \delta|\Delta, V \rangle &= \Delta + S(1 - S/s) V, \\ \langle v|\Delta, V \rangle &= (S/s)^2 V, & C_{vv|\Delta V} &= \frac{1 - (S/s)^3}{s}, \\ C_{\delta\delta|\Delta V} &= s(1 - S/s)^3, \\ C_{\delta v|\Delta V} &= \frac{(1 - S/s)^2}{2} (1 + 2S/s), \end{aligned} \quad (76)$$

and

$$\begin{aligned} \langle \delta|v, \Delta, V \rangle &= \Delta + S\psi V + \frac{C_{\delta v|\Delta V}}{C_{vv|\Delta V}} (v - \langle v|V \rangle), \\ C_{\delta\delta|v\Delta V} &= s \frac{3(1 - S/s)^4}{4(1 - S^3/s^3)}, \\ \langle v|\delta, \Delta, V \rangle &= \frac{(\delta - \Delta)}{2s} \frac{1 + 2S/s}{1 - S/s} - \frac{S/s}{2} V, \\ C_{vv|\delta\Delta V} &= (3/4) (1 - S/s)/s. \end{aligned} \quad (77)$$

In general, since $S\psi V \geq 0$, the dependence on V always acts to increase the effective value of B in $p(b|B, V)$. Hence,

if we define a characteristic scale as that s where $b - B - S(1 - S/s)V = \sqrt{s}(1 - S/s)^{3/2}$, then larger V means smaller s . That is, walks which upcross B with steeper slopes are associated with more massive progenitors at the earlier time when the barrier height was b . Since steeper slopes are also associated with smaller bias factors (see analysis in Musso et al. 2012, and Figure 6 here for an illustrative example), the analysis above allows one to quantify the assembly bias effect in this model.

4.4 Merger trees

Our equation (13) is a key feature of Markov Velocity models – one that is potentially extremely useful for fast generation of merger history trees of what is sometimes called the ‘main progenitor’ (Lacey & Cole 1993). This is because, in the Markov heights case, one makes independent picks from $\prod_i f(\delta_i|\delta_{i-1})$, where the δ_i need not be associated with closely spaced scales. This boils down to making independent picks from $\prod_i p(\delta_i|\delta_{i-1})$ for each i and then making a change of variables.

In principle, our Markov Velocity models allow us to work with $\prod_i f(\delta_i, v_i|\delta_{i-1}, v_{i-1})$. In the approximation where one replaces $f(\delta_i, v_i|\delta_{i-1}, v_{i-1})$ with $f_{\text{up}}(\delta_i, v_i|\delta_{i-1}, v_{i-1})$, we can view f_{up} of equation (67) as an integral over terms of the form $f_{\text{up}}(\delta, v|\Delta, V)$. The appearance of v and V mean that such an approach will return merger histories with more information than the traditional ones; v and V are expected to encode information about the halo concentration and large scale environment, so our approach leads naturally to merger histories which incorporate a form of assembly bias (c.f. Section 4.3). We are in the process of determining if the assembly bias associated with these Markov Velocity trees is realistic.

Finally, before closing, we note that the closely related quantity, $\prod_i p(\delta_i, v_i|\delta_{i-1}, v_{i-1})$ with $\delta_i = \delta_{i-1}$, also plays a key role in equation (A3) of Musso & Sheth (2014a), the formal exact expression for $f(s)$ (where scales i and $i-1$ may be rather far apart). Since each $p(\delta_i, v_i|\delta_{i-1}, v_{i-1}) = p(v_i|v_{i-1})p(\delta_i|v_i, \delta_{i-1}, v_{i-1})$ the product may simplify, so we are in the process of checking if the analysis of the previous section allows a fully analytic solution of the formal expression for $f(s)$ for some if not all Markov Velocity models.

5 DISCUSSION

Previous work on the first crossing distribution has shown the power of including the constraint that walks must cross upwards (Bond et al. 1991; Musso & Sheth 2012, 2014a,b). In particular, this has shown that studying the velocity structure of the walks – the continuum limit of the steps – is particularly fruitful when the steps are correlated. To explore this structure further, we first developed a toy model in which the correlation matrix of the steps is particularly simple (equation 9), and yet non-diagonal: even though this matrix has no zero entries, for this process the conditional distribution of any step depends only on the one just before it (equation 10). Therefore, this toy model exhibits the simplest level of complication one could have added to walks with completely uncorrelated steps.

We showed how to make fast Monte Carlo realizations

of such walks, providing an explicit expression for how one should think of the smoothing filter associated with the model (equation 16). We then used the Monte Carlo to obtain the first crossing distribution associated with the toy model, showing that it was rather similar to that for TopHat-smoothed Λ CDM walks (Figure 2). Along the way, we also used the toy model to illustrate how the distribution of walk heights – if it is known that the barrier was crossed on a larger smoothing scale – is modified by the correlations between steps (Figure A1).

We then showed that the toy model was a special case of a more general Markov Velocities model, in which it is not the heights of the walk, but the steps, which are Markovian. The Markov assumption allowed us to include the effects of correlations between steps rather efficiently. We did so by first writing down the Langevin equation which governs the process (equation 25), showing explicitly how the scale dependence of the velocity variance $\langle v^2(s) \rangle$ determines the process. We then solved the Langevin equation (equations 26–28), and studied the special case in which $\langle v^2(s) \rangle \propto 1/s$, arguing that such walks should be thought of as a family of scale-invariant models indexed by the constant of proportionality. The associated first crossing distributions interpolate smoothly between the case of walks with completely correlated and completely uncorrelated steps (Figure 4). Moreover, although Gaussian smoothing of scale-free power spectra produces walks that do not have Markov Velocities, their first crossing distributions are rather well approximated by those of scale-invariant Markov Velocity models having the same velocity variance structure (Figure 4).

In the more general case, $\langle v^2(s) \rangle$ may be a more complicated function of s ; e.g., for Λ CDM $P(k)$, $\gamma(s) = (4s\langle v^2 \rangle)^{-1/2}$ is quite well approximated by equation (49). We again provided explicit expressions for the effective smoothing filter of the white noise which affects the Langevin trajectories, arguing that equation (52) should provide a good approximation for Λ CDM-like $P(k)$. And we described our Markov Velocity Monte Carlo algorithm (equation 53) for generating Markov Velocity walks whose first crossing distribution closely approximates that of TopHat-smoothed Λ CDM (Figure 5).

Truncated Fourier smoothing kernels of the form given by our equation (54) will generically yield Markov velocity models whatever the underlying power spectrum (Section 3.5). For power-law $P(k)$, we provided an explicit mapping between the index γ of the Markov Velocity model, the shape of the smoothing filter, and the slope of the power-law $P(k)$ (equation 55). For the Λ CDM family, this mapping may be combined with equation (49) to get a feel for the smoothing window shape which will yield Markovian Velocities.

In this context, TopHat smoothing of $P(k) \propto k^{-n}$ with $n = -2$ is special: it has the same correlation structure as the Markov Velocity model obtained by smoothing with $W = 1 - (kR)^2$ (for $kR \leq 1$). Since $n = -2$ is close to the effective spectral index of the Λ CDM family of power spectra on the scales where $s \sim \delta_c^2$, this correspondence may provide an easy way to think of issues such as assembly bias (Sheth & Tormen 2004). This is because, in Markov Velocity models, if the walk height and slope are known on one scale, say S , the walk height on smaller scales (where $s \geq S$) depends only these two values, and not on the walk heights on

scales larger than S (equation 13). Therefore, at fixed mass, the assembly history in Markov Velocity models should be correlated with the larger scale environment because of the dependence on slope: assembly bias is present (equations 67 and 68 and related discussion, as well as Figure 6). However, at fixed mass *and* slope, there should be no further correlation between the formation history of a halo and its environment (equations 69 and 71). The slope of the walk associated with a protohalo patch is an indicator of the concentration of the final halo (e.g. Dalal et al. 2008); therefore, in Markov velocity models, there should be no correlations between formation history and environment if done at fixed halo mass *and* concentration. In this sense, assembly bias effects in Markov Velocity models are relatively simple.

The exact Markov Velocity nature of TopHat-smoothed $P(k) \propto k^{-2}$, and the fact that Markov Velocity smoothing of Λ CDM $P(k)$ yields a first crossing distribution that is in good agreement with that of TopHat smoothed Λ CDM (Figure 5) strongly suggest that Markov velocities are a useful approximation for future excursion set studies. However, the first crossing distribution (of a constant barrier) for TopHat smoothed Λ CDM walks is not in as good agreement with the actual mass fractions measured in numerical simulations of halo formation. These are rather well matched by the first crossing distribution associated with our simplest model (having $\gamma = 1/2$) which provides a good description of the mass fraction in haloes (Figure 3). Although one could explore how allowing the barrier height to depend on s might improve the agreement, we expect even this simplest constant barrier model to provide the basis of fast Monte Carlo merger history tree algorithms which include some of the assembly effects associated with correlated steps (Sections 4.3 and 4.4), and so represent a significant improvement on what is currently available.

ACKNOWLEDGEMENTS

The work of 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 NASA NNX11A125G.

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
 Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12, 0803.3453
 Epstein R. I., 1983, MNRAS, 205, 207, ADS
 Lacey C. G., Cole S., 1993, MNRAS, 262, 627, ADS
 Maggiore M., Riotto A., 2010, ApJ, 711, 907, 0903.1249
 Musso M., Paranjape A., Sheth R. K., 2012, MNRAS, 427, 3145, 1205.3401
 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
 Paranjape A., Lam T. Y., Sheth R. K., 2012, MNRAS, 420, 1429, 1105.1990
 Paranjape A., Sheth R. K., 2012, MNRAS, 426, 2789, 1206.3506
 Press W. H., Schechter P., 1974, ApJ, 187, 425, ADS
 Sheth R. K., 1998, MNRAS, 300, 1057, 9805319
 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., 1999, MNRAS, 308, 119, astro-ph/9901122, ADS
 Sheth R. K., Tormen G., 2002, MNRAS, 329, 61, astro-ph/0105113
 Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385, astro-ph/0402237

APPENDIX A: DISTRIBUTIONS AFTER FIRST CROSSING

In this appendix we use the Monte Carlo algorithm described in the main text to illustrate how correlations between steps modify one of the key quantities in the usual excursion set approach: The probability $p_{\times}(\delta)$ that a walk reaches δ at scale s having crossed the barrier at least once at larger scale. This can be written, exactly, as

$$p_{\times}(\delta, s) = \int_0^s dS \int_{B'}^{\infty} dV (V - B') p(\delta, B, V, \text{first } S), \quad (\text{A1})$$

from which one can obtain the first crossing rate as

$$f(s) = -\frac{d}{ds} \int_{-\infty}^{b(s)} d\delta [p(\delta, s) - p_{\times}(\delta, s)]. \quad (\text{A2})$$

For barriers of constant height, equation (A2) is the generalization to correlated steps of the symmetry argument used by Bond et al. (1991) for walks with uncorrelated steps (in which case $p_{\times}(\delta)$ is a Gaussian with mean $2b$ and variance S).

We can approximate $p_{\times}(\delta, s)$ by dropping from the term in the integrand of equation (A1) the requirement that the walk never crossed before S (this is consistent with the up-crossing approximation, $f \approx f_{\text{up}}$, of equation 3). We will call this approximation $p_{\times}^{\text{MS}}(\delta, s)$. Now, in our toy model, $p(B, V, \delta) = p(B, V) p(\delta - B - \psi V)$ so

$$p_{\times}^{\text{MS}}(\delta) = \int_0^s \frac{dS}{S} \frac{2s}{s-S} \int_0^{\infty} dx x \frac{e^{-x^2/2S}}{\sqrt{2\pi S}} \frac{e^{-(x-2B)^2/6S}}{\sqrt{6\pi S}} \times \frac{e^{-[(\delta-B)s/(s-S)-x]^2/2(s-S)}}{\sqrt{2\pi(s-S)}}, \quad (\text{A3})$$

where we have defined $x \equiv SV$. Figure A1 shows that this yields an excellent approximation to the Monte Carlo'd result. It also shows the special cases in which steps are completely correlated (a step function) or completely uncorrelated (the ‘mirror image’ of the tail of an error function).

Inserting $p_{\times}^{\text{MS}}$ of equation (A3) in place of $p_{\times}$ in equation (A2) gives

$$f(s) \simeq f_{\text{PS}}(s) + \int_0^s dS \int_0^{\infty} dV V \frac{dp(\delta \leq b, B, V + B')}{ds}, \quad (\text{A4})$$

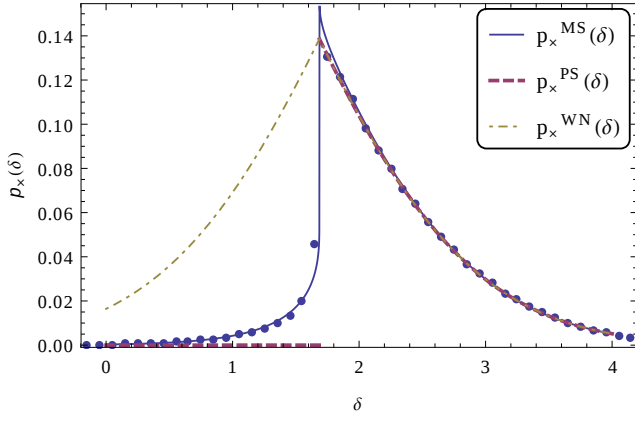


Figure A1. Probability distribution $p_{\times}(\delta)$ of the height δ on scale $s = 2$, of walks that crossed a constant barrier of height $\delta_c = 1.69$ at some $S < s$. Filled circles show this quantity measured in Monte Carlo simulations of walks having $\langle v_i, v_j \rangle = s_i/s_j^2$ for $s_i < s_j$, and solid line shows equation (A3). For comparison, the dashed line shows $p(\delta)\vartheta(\delta - \delta_c)$, which is exact for $\delta > \delta_c$ and corresponds to walks with completely correlated steps, whereas the dot-dashed line is for walks with uncorrelated steps.

where we have redefined $V \rightarrow V + B'$, and brought d/ds inside the integral over S (since its action on s in the integration limit gives zero). For our toy model, $p(\delta \leq b, B, V + B')$ is known exactly, so

$$sf(s) \approx sf_{\text{PS}}(s) + \int_0^1 \frac{dy}{y} \frac{1}{1-y} \int_0^\infty dw w^2 \times \frac{e^{-w^2/2y}}{\sqrt{2\pi y}} \frac{e^{-[w-2(b/\sqrt{s})]^2/6y}}{\sqrt{6\pi y}} \frac{e^{-w^2/2(1-y)}}{\sqrt{2\pi(1-y)}}. \quad (\text{A5})$$

Figure 2 shows that this approximation for $f(s)$ is substantially more accurate than $f_{\text{up}}(s)$ of equation (3); see Musso & Sheth (2014b) for why it is not quite as accurate as $f_{\text{BS}}(s)$.