Title: Not so random walks in cosmology

Date: Sep 24, 2013 11:00 AM

URL: http://pirsa.org/13090063

Abstract: I'll discuss a number of insights into the process of nonlinear structure formation which come from the study of random walks crossing a suitably chosen barrier. These derive from a number of new results about walks with correlated steps, and include a unified framework for the peaks and excursion set frameworks for estimating halo abundances, evolution and clustering, as well as nonlinear, nonlocal and stochastic halo bias, all of which matter for the next generation of large scale structure datasets.

Not so random walks in cosmology

Can theory step up to precision cosmology?

Ravi K Sheth (ICTP/Penn),
Marcello Musso (CP3-IRMP),
Aseem Paranjape (ETH Zurich)

- Solving Press-Schechter/Excursion Set
  - The fundamental ansatz
  - The importance of stepping up
  - One step beyond
- Failure of fundamental ansatz
  - Stochasticity from other variables
  - Averaging over all vs special positions
- Special positions (e.g. peaks)
  - Lagrangian and Eulerian bias

Fundamental ansatz

dn/dm = comoving density of m-halos

$$\frac{m}{\rho}\frac{dn}{dm} = \text{mass fraction in } m-\text{halos}$$
$$= \text{fraction of positions in field which when smoothed on scale m are predicted to collapse into halos of mass m}$$
$$= f(s)\,ds$$

where s is variance in density field when smoothed on scale m.
s(m) is monotonic function of 1/m.

Typically written as 'concentric spheres'.

Press-Schechter: Want $\delta \geq \delta_c$

Bond, Cole, Efstathiou, Kaiser:
$\delta(s) \geq \delta_c$ and $\delta(S) \leq \delta_c$ for all $S \leq s$:

$$f(s)\Delta s = \int_{-\infty}^{\delta_c} d\delta_1 \cdots \int_{-\infty}^{\delta_c} d\delta_{n-1} \int_{\delta_c}^{\infty} d\delta_n \, p(\delta_1, \ldots, \delta_n)$$

Since $s = n\Delta s$ this requires n-point
distribution in limit as $n \to \infty$ and $\Delta s \to 0$.
(Best solved by Monte-Carlo methods.)

Musso-Sheth: $\delta \geq \delta_c$ while 'stepping up'
$\delta(S) \geq \delta_c$ and $\delta(S - \Delta S) \leq \delta_c$
$\delta(S) \geq \delta_c$ and $\delta(S) - \Delta S d\delta/dS \leq \delta_c$
$\delta(S) \geq \delta_c$ and $\delta(S) \leq \delta_c + \Delta S\, v$
so

$$f(s)\,\Delta s = \int_0^{\infty} dv \int_{\delta_c}^{\delta_c + \Delta S\, v} db\, p(b, v)$$

$$= \int_0^{\infty} dv\, \Delta S\, v\, p(\delta_c, v)$$

making

$$f(s) = p(\delta_c|s) \int_0^{\infty} dv\, v\, p(v|\delta_c)$$

Requires only 2-point statistics.
Logic general; applies to very NG fields also

One step beyond:

Start from exact statement:

$$p(\geq b|s) = \int_0^s dS\, f(S)\, p(\geq b, s|\text{first at } S)$$

Approximate as:

$$p(\geq b|s) \approx \int_0^s dS\, f(S)\, p(\geq b, s|B, S)$$

Completely correlated: $p(\geq \delta_c, s|\delta_c, S) = 1$
(what Press-Schechter really means)

Completely uncorrelated:
$p(\geq \delta_c, s|\delta_c, S) = 1/2$
(Bond, Cole, Efstathiou, Kaiser)

Next simplest approximation (step up):

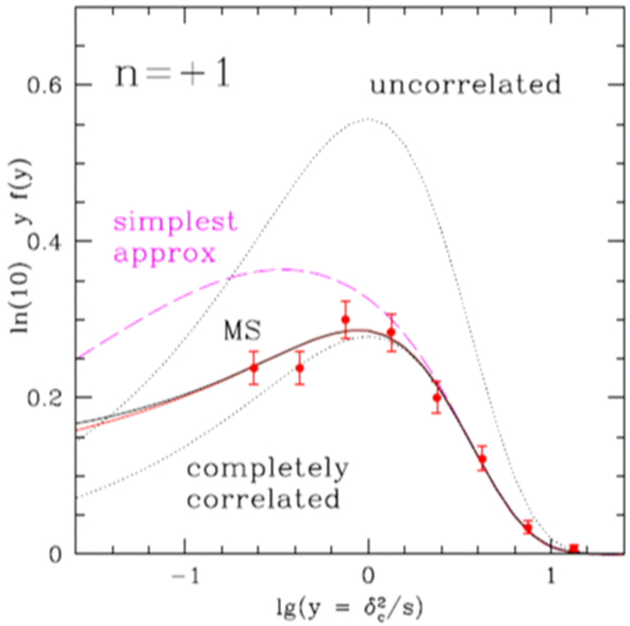
$$p(\geq b|s) = \int_0^s dS\, f(S)\, p(\geq b, s|\text{first at } S)$$
$$\approx \int_0^s dS\, f(S)\, p(\geq b|\text{up at } S)$$

where

$$p(\geq b|\text{up at } S) = \frac{\int_0^\infty dV\, V\, p(\geq b, V|B)}{\int_0^\infty dV\, V\, p(V|B)}$$

Requires only 3-point statistics.

Works for all smoothing filters and (monotonic) barriers.


n=+1
uncorrelated
simplest
approx
MS
completely
correlated
ln(10) y f(y)
lg(y = δ_c^2/s)
0.6
0.4
0.2
0
−1
0
1

One step beyond:

Start from exact statement:

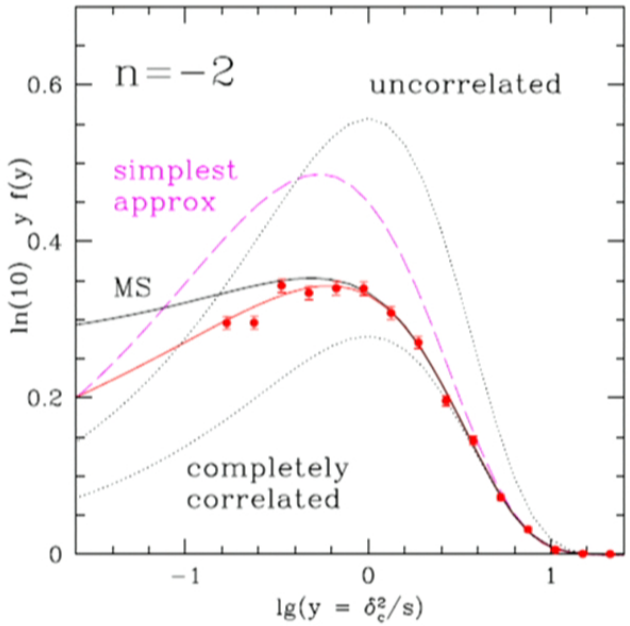
$$p(\geq b|s) = \int_0^s dS\, f(S)\, p(\geq b, s|\text{first at } S)$$

Approximate as:

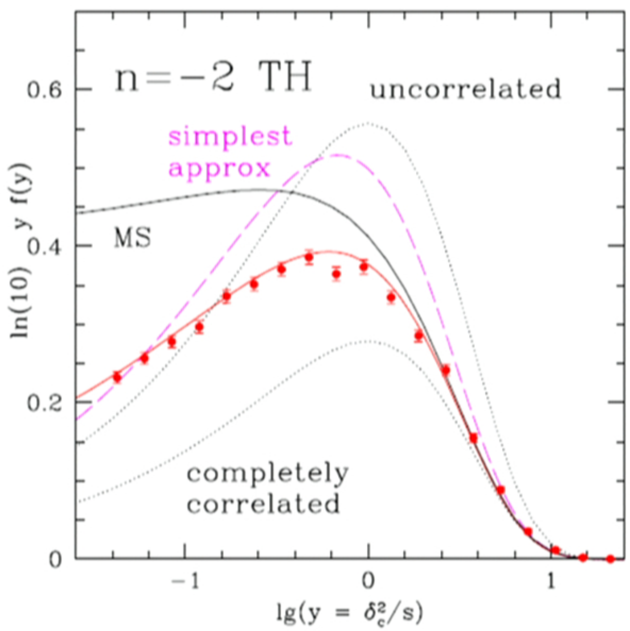
$$p(\geq b|s) \approx \int_0^s dS\, f(S)\, p(\geq b, s|B, S)$$

Completely correlated: $p(\geq \delta_c, s|\delta_c, S) = 1$
(what Press-Schechter really means)

Completely uncorrelated:
$p(\geq \delta_c, s|\delta_c, S) = 1/2$
(Bond, Cole, Efstathiou, Kaiser)



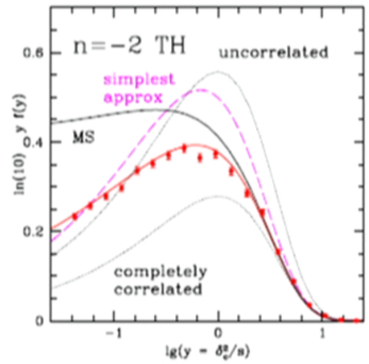
n = −2
uncorrelated
simplest
approx
MS
completely
correlated
ln(10) y f(y)
lg(y = δ_c^2/s)
0.6
0.4
0.2
0
−1
0
1


n=−2 TH
uncorrelated
simplest
approx
MS
completely
correlated
ln(10) y f(y)
lg(y = δ_c^2/s)
0.6
0.4
0.2
0
−1
0
1

Note that n=-2 TopHat has

Markov velocities

(rather than heights):



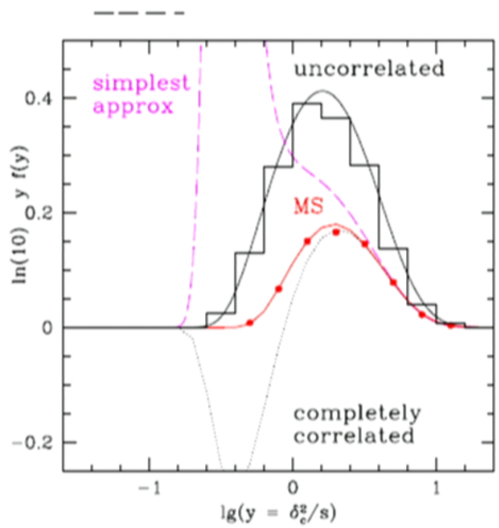
$p(\delta_n|\delta_k,\ldots,\delta_1) = p(\delta_n - \delta_k - \psi_{nk} s_k v_k)$

Previous work with uncorrelated steps had
$p(\delta_n|\delta_k,\ldots,\delta_1) = p(\delta_n - \delta_k)$
so Markov velocities is next most
complicated (Ising-like) model
(Musso, Sheth 2013).

Back-substitution also works for

moving barriers

(used to model filaments and sheets in cosmic web, reionization bubbles)



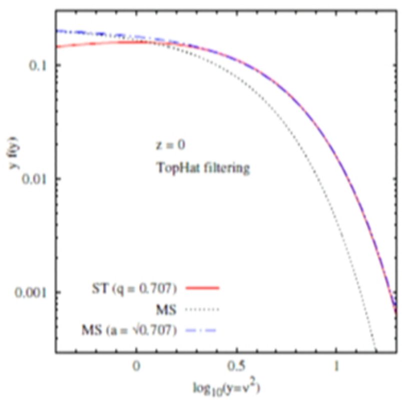
Moving barrier: $b = \delta_c[1 + (s/\delta_c)^2/4]$

In fact, MS is almost always a good enough approximation.

Problem

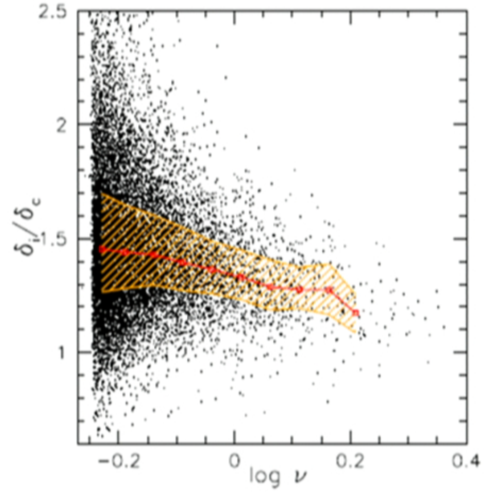
$f(s) \propto \exp(-\delta_c^2/2s)$ at $s \ll 1$ (large m)

This generically underestimates halo counts in regime of most interest to cosmology.



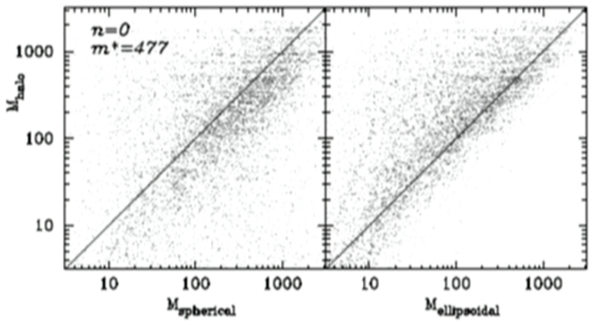
Decreasing $\delta_c \to \sqrt{a}\delta_c$ is an 'effective' fix, and is employed in all fitting formulae.

Cheat, because physics of collapse has
$\delta \geq \delta_c$.



(Sheth, Mo, Tormen 2001; Despali, Tormen, Sheth 2013)

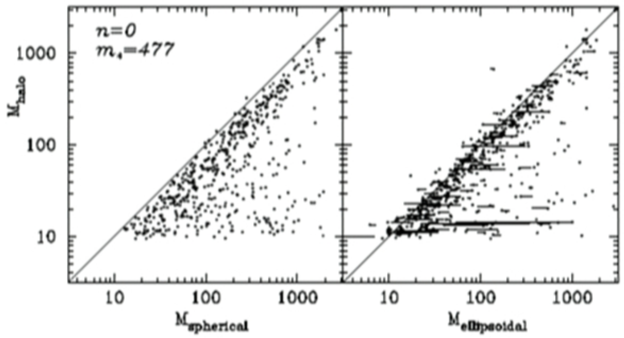
Moreover, predicted mass wrong for almost all walks (White 1996) ...



... even if one uses a better model for the physics of collapse (Sheth, Mo, Tormen 2001)!

Collapse model + uncorrelated steps + uncorrelated walks

Walk associated with center of mass particle usually predicts mass much better.



(Sheth, Mo & Tormen 2001)

Ansatz: Halos form around local maximum in predicted mass.

- if predicted mass is m, all other walks within R_m should predict smaller masses;
- walk which predicts m should be further than R_M from all walks for which predicted mass is $M>m$.

Walk associated with center of mass particle usually predicts mass much better (Sheth, Mo, Tormen 2001)

Suggests first of two fixes:

Collapse model + (un)correlated steps + uncorrelated walks

Accounting for correlated walks is tough. Easier to average over special positions in space: e.g. Peaks
Attractive because simulations show good correspondence between peaks and halos (Ludlow, Porciani 2011; Hahn, Paranjape 2013).

I. Average only over special positions

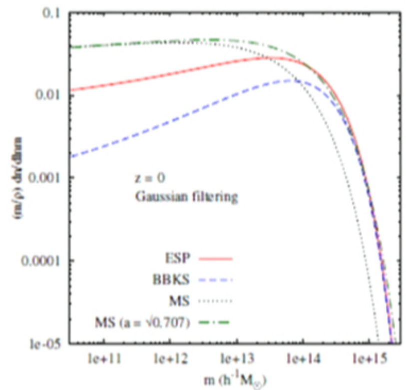
Can write Excursion Set Peaks model by noting that distribution of slopes v for peaks (i.e. curvature around a peak) is different from that for random positions (of same height):

$$f(s) = p(b|s) \int_{b'}^{\infty} dv\, (v - b')\, p(v|b)\, C_{pk}(v)$$

C_{pk} known from BBKS (1986).

In effect, this yields peaks theory for arbitrary barriers. (Musso, Sheth 2012)

Including this extra factor
(from peak curvatures) makes
$f_{pk}(s) \propto (\delta_c^2/s)\, f(s)$ at $s \ll 1$ (large m)



Provides much better match to halo counts
(Paranjape, Sheth 2012; Paranjape et al. 2013).

Physics 'easy', statistics tough

II. Model the error: Stochasticity

(Castorina, Sheth 2013; Achitouv et al 2013)

Assume

$$\delta \geq \delta_c + q$$

If q zero-mean Gaussian, with variance $\beta^2 s$, then useful to rewrite as

$$\delta - q \geq \delta_c$$

Since distribution of sum of two Gaussians is Gaussian, we have the same (solved!) problem as before, with all $s \to (1+\beta^2)s$. Therefore $\delta_c^2/s \to \delta_c^2/(1+\beta^2)/s$, which looks like rescaling of δ_c as required.

But what is q?

Statistics 'easy', physics obsure

Other things than initial overdensity may also matter (e.g. external shear, alignment of initial shape and shear, etc.)

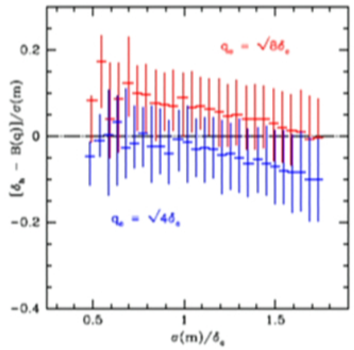
These lead to models with more than one walk (Sheth, Tormen 2002), sometimes called stochastic barrier models (Castorina, Sheth 2013), which generically exhibit 'nonlocal' stochastic bias (Sheth et al. 2012).

All of these can be treated with the same formalism.

Assume

$$\delta \geq \delta_c \,(1 + \sqrt{q^2/q_c^2})$$

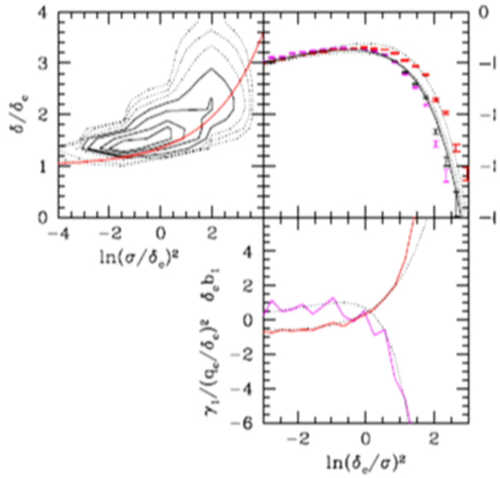
where q is related to (traceless) tidal field.



This is in reasonable agreement with protohalos in simulations.

This model for the physics always has $\delta \geq \delta_c$.

First crossing distribution can be found analytically, even if walks are required to start from nonzero δ_0 and q_0.



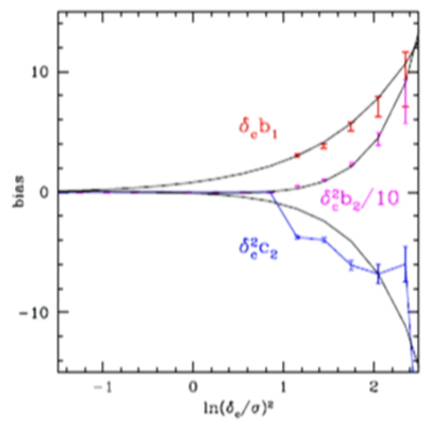
(Sheth, Chan, Scoccimarro 2012)

This yields analytic expressions for ...

Nonlocal, stochastic bias

$$1 + \delta_h^L(m|\delta_0, q_0^2) \equiv \frac{\langle N_m|\delta_0, q_0^2\rangle}{n(m)V_0}$$

$$= 1 + b_1^L\,\delta_0 + b_2^L\,\frac{\delta_0^2}{2} + c_2^L\,\frac{q_0^2}{2} + \dots$$



In initial conditions, expect nonlocality matters most for most massive halos.

Account for additional nonlocality from contribution of tidal term to nonlinear evolution $\delta(\delta_0, q_0)$ to get Eulerian bias.

$$1 + \delta_h^E(\delta, q^2) = (1+\delta)(1+\delta_h^L)$$

$$\begin{aligned}
&= (1+\delta)\left(1 + b_1^L\,\delta_0 + b_2^L\,\frac{\delta_0^2}{2} + c_2^L\,\frac{q_0^2}{2} + \dots\right) \\
&= 1 + b_1^L\,\delta_0 + b_2^L\,\frac{\delta_0^2}{2} + c_2^L\,\frac{q_0^2}{2} + \delta + b_1^L\,\delta_0\delta \\
&= 1 + \delta\,(b_1^L + 1) + \frac{\delta^2}{2}(8b_1^L/21 + b_2^L) \\
&\quad + \frac{q_0^2}{2}(c_2^L - 8b_1^L/21).
\end{aligned}$$

Since, to lowest order, $q_0^2 = q^2$, the Eulerian bias factors are

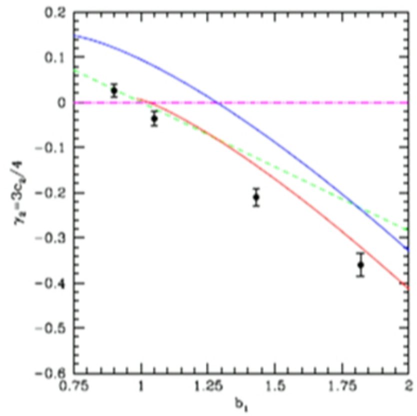
$$b_1 = 1 + b_1^L, \qquad b_2 = b_2^L + \frac{8}{21}\,b_1^L,$$

and

$$c_2 = c_2^L - \frac{8}{21}\,b_1^L.$$

Tidal term gives rise to new observable: quadrupolar signature in bias.

Matters for bispectrum. Clipping the field, Gaussianization, etc, are local operations, so will get this wrong.



Some evidence for this expected departure from local Lagrangian bias seen in simulations (Sheth et al. 2013), on mass scales which are relevant to LRGs.

Adding peaks constraint makes the bias factors k-dependent.

Systematic expansions for peaks are in Paranjape et al. (2013) and Desjacques (2013).

Stay tuned!

Summary

Its always good to step up; first passage problem with correlated steps 'solved' using only 3 variables.

Markov velocities are the next most natural (Ising-model like) generalization of the usual Markov heights model; good approximation for CDM-like P(k)!

Self-consistent model must only average over special subset of walks. Peaks are a good choice, for which closed form expressions are now available.

Must incorporate stochasticity in halo formation from tidal field and (mis-alignment with!) proto-halo shapes.

Tidal field leaves signature on halo abundances, clustering, especially in higher-order statistics of highly clustered objects (typically high-mass halos).