# Notations

$\rho(\bar{x})$ – density field at $\bar{x}$ (comoving coord.)

$\delta(\bar{x}) = \dfrac{\rho(\bar{x})}{\bar{\rho}} - 1$ = density contrast

Note: $\rho(\bar{x}) \geq 0$ and hence $\delta(\bar{x}) \geq -1$
$\Rightarrow$ both are physical quantities

In contrast to the extrapolated density contrast, where

$$\delta_e(z=0) = \underbrace{\frac{D(z=0)}{D(z=z_i)}}_{\text{growth factor}} \delta_e(z=z_i)$$

$\delta_e$ are not bounded below by $-1$ !

Homogeneity: $\langle \delta(\bar{x}) \rangle = \langle \delta \rangle$
$\langle \delta(\bar{x}+\bar{r})\,\delta(\bar{x}) \rangle = \langle \delta(\bar{r}) \rangle$

Smoothing :

The observed quantities are usually smoothed by some window functions

$$\delta(\bar{x}\,;\,W_R) = \int d^3y\ \delta(\bar{y})\ W_R(|\bar{x}-\bar{y}|)$$

where we assume the filter function $W_R$ is normalized.

Applying filtering is easier in Fourier space (convolution → multiplication)

$$\delta_R(\bar{k}) = \tilde{W}_R(\bar{k})\ \delta(\bar{k})$$

Common filters

① Real-space

$$W_R(r) = \begin{cases} \frac{1}{V_R} & \text{if} \quad r \le R \\ 0 & \text{otherwise} \end{cases}$$

$$\tilde{W}_R(k) = \frac{3}{(kR)^3}\left[\sin(kR) - kR\ \cos(kR)\right]$$

② Gaussian

$$W_G(x;R) = \frac{1}{(2\pi)^{3/2} R^3} \exp\left[-\frac{x^2}{2R}\right]$$

$$\widetilde{W}_G(k;R) = \exp\left[-(kR)^2/2\right]$$

③ Sharp-k

$$\widetilde{W}_{SK}(k;k_0) = \begin{cases} 1 & k \le k_0 \\ 0 & \text{otherwise} \end{cases}$$

$$W_{SK}(x;k_0) \propto \frac{\sin(k_0 x) - k_0 x \cos k_0 x}{(k_0 x)^3}$$

## Power spectrum & correlation function

$$P(k) \propto \langle |\delta_k|^2 \rangle$$

$$\xi(r) = \langle \delta(\bar{x}+\bar{r})\, \delta(\bar{x}) \rangle$$

"Smoothed" version

$$\xi(r; R_1, R_2) = \langle \delta_{R_1}(\bar{x}+\bar{r})\, \delta_{R_2}(\bar{x}) \rangle$$

## Normalization

Define the variance of the smoothed density at zero lag

$$\sigma^2(R) = \xi(r=0; R, R) = \langle \delta_R^2 \rangle$$

$$= \int \frac{dk}{k} \frac{k^3 P(k)}{2\pi^2} W^2(kR)$$

$\sigma_8$ — normalization of linear power spectrum with (real-space) top hat filter and $R = 8$ Mpc/h

## Variance, smoothing scale, mass

Consider $s(R) \equiv \sigma^2(R)$, $s(R)$ describes the fluctuation of the density field $\delta_R$ when the smoothing scale $R$ changes.

For the power spectrum in cosmology $s(R)$ is always a decreasing function.

In particular for power-law $P(k) \propto k^{n_s}$

$$s(R) = \sigma^2(R) = A\int \frac{dk}{k} \frac{k^3 P(k)}{2\pi^2} W^2(kR)$$

Tophat: $$s(R) = A\int \frac{dk}{k} \frac{k^{n_s+3}}{2\pi^2} \frac{\sin kR - kR\cos kR}{(kR)^3}$$

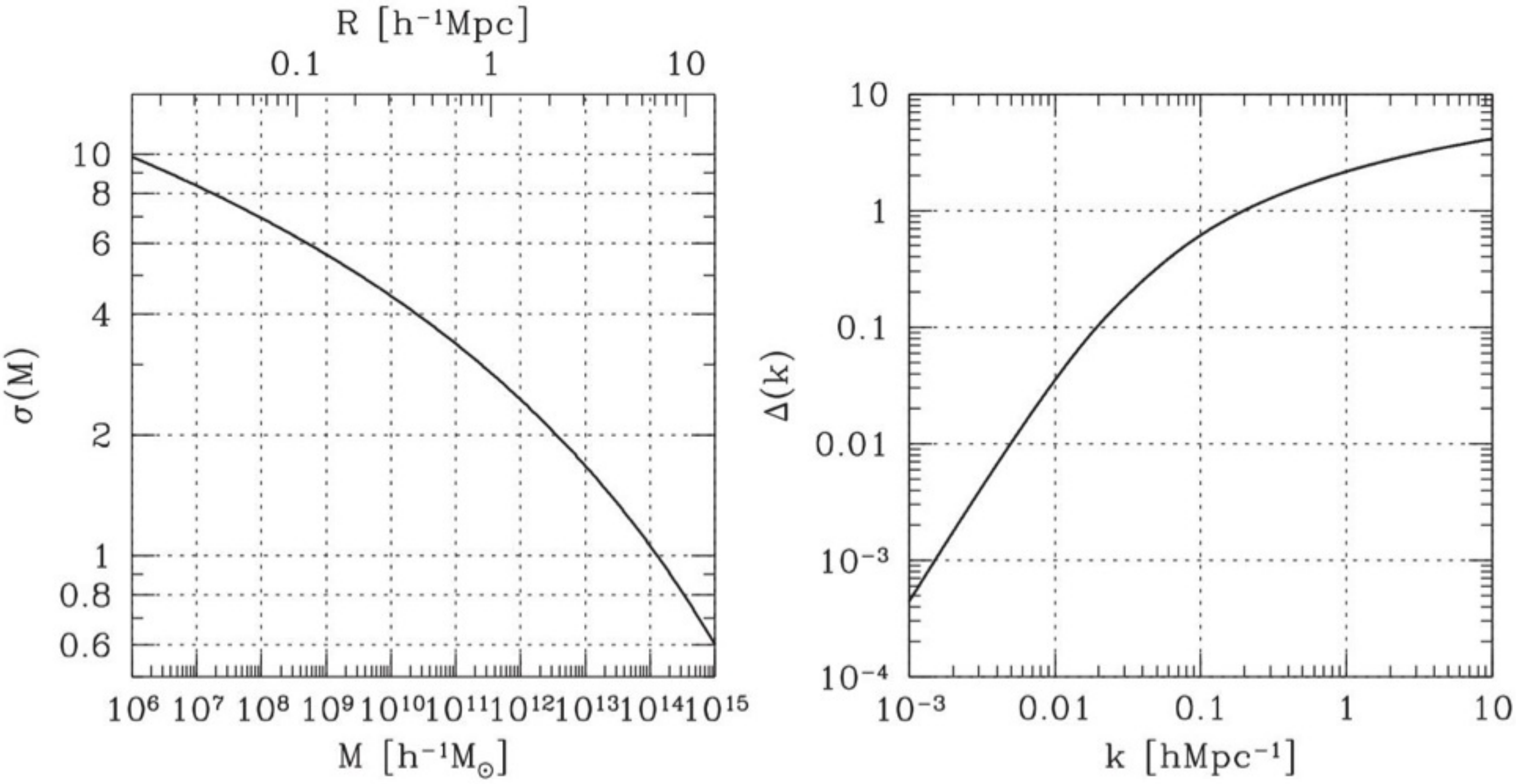
$$\propto R^{-(n_s+3)}$$

$$\propto M^{-\frac{(n_s+3)}{3}}$$

The constant of proportionality is set by $\sigma_8$.

Three equivalent quantities:

$$S \longleftrightarrow R \longleftrightarrow M$$




Zentner 2007

Review on excursion set

# 1. Halo Abundance

Number density of rare objects would tell us something about cosmology.

Dark matter halos : most massive bound systems

Building blocks for mock galaxy catalog and analytic models.

Postulate : halos preferably form at high density regions

Spherical collapse (Gunn & Gott 1972)

The virialized objects would have a density $\sim 200$ the background.
The linear extrapolated density contrast to $z=0$ is $\delta_c \approx 1.68$

Abundance : "count" the regions in the linear density field that have $\delta_R \geq \delta_c$

Note : The comparsion is in Lagrangian (linear extrapolation).

$\delta_R$ — a Gaussian random variable $\left( \begin{array}{l} \text{mean} = 0 \\ \text{variance} = S(R) \end{array} \right)$

# Press - Schechter (1974)

Consider

$$\mathbb{P}(s) = \int_{\delta_c}^{\infty} p(\delta ; s)\, d\delta$$

$$= \frac{1}{2} \operatorname{erfc}\left(\frac{\nu}{\sqrt{2}}\right) \quad , \text{where } \nu = \frac{\delta_c}{\sqrt{s}}$$

$$= 1 - \text{cumulative probability at } s$$

Another view, $\mathbb{P}$ is the mass fraction in halos more massive than $M(s)$.

We already know as $M \to 0$ (or $R \to 0$), $s \to \infty$ (hence $\nu \to 0$).

$\Rightarrow$ Total mass fraction in halo $= \frac{1}{2}$

Press-Schechter then multiplied the above factor by 2 $\rightarrow$ all mass are bound to some halo (to the zero mass halo limit).

Converting from $\mathbb{P}(s)$ to $\frac{dn}{dM}$

number density $\rightarrow$ $$\frac{dn}{dM} dM = 2 \boxed{\frac{\bar{\rho}}{M} \left| \frac{d\mathbb{P}}{dM} \right|} dM$$

mass fraction in mass bin $(M, M+dM)$

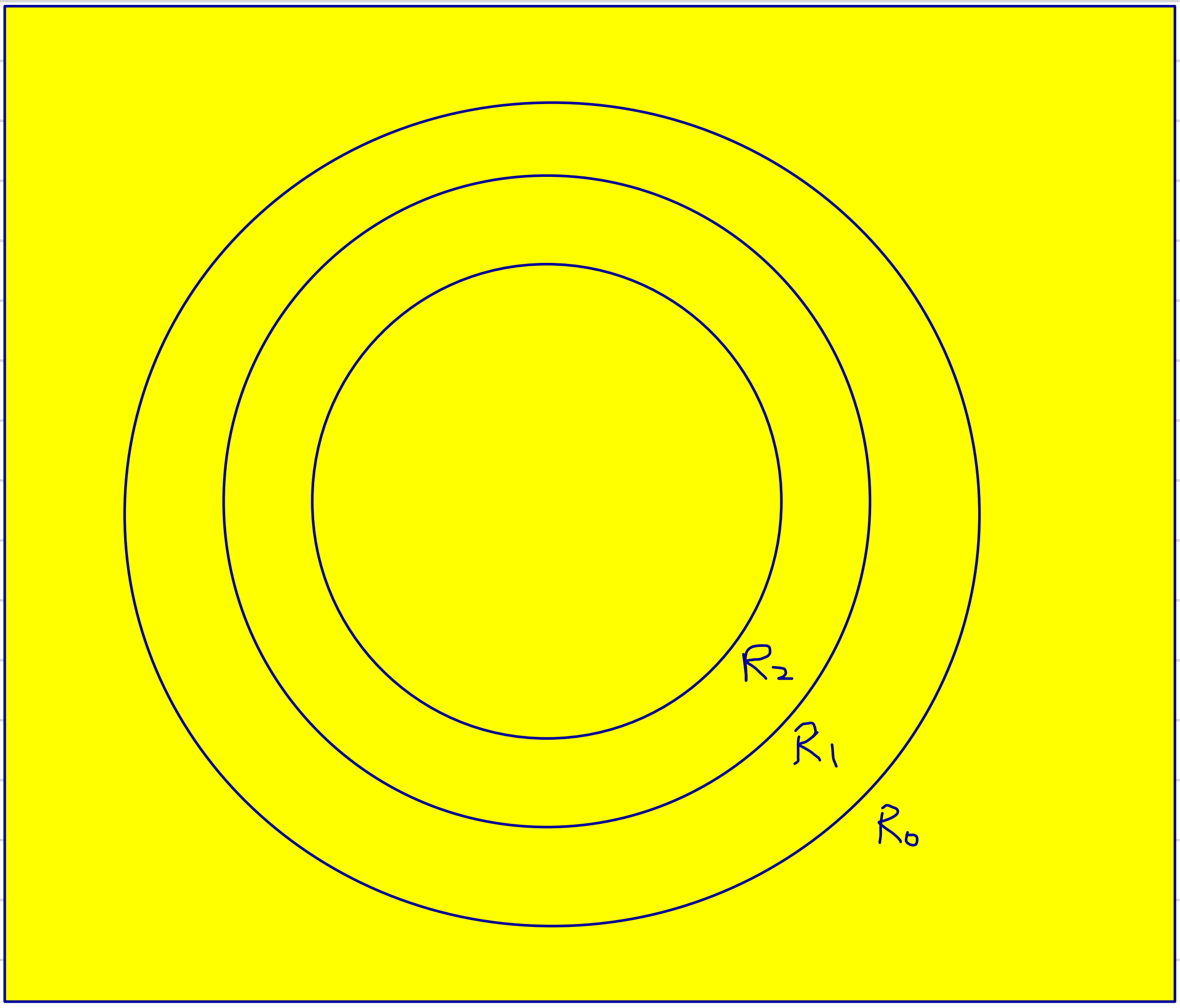
$$= 2 \frac{\bar{\rho}}{M^2} \frac{d\mathbb{P}}{d\ln s} \left| \frac{d\ln s}{d\ln M} \right| dM$$

$$= \frac{1}{\sqrt{2\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sqrt{s}} e^{-\delta_c^2/2s} \left| \frac{d\ln s}{d\ln M} \right| dM$$

Excursion set (or extended Press-Schechter)

Bond et. al. 1991

- Look for the biggest smoothing scale (or most massive) regions that satisfy the halo formulation criteria.




At each $\bar{x}$, form a series of concentric smoothing sphere $R_0 > R_1 > R_2 \cdots$, measure the corresponding $\{\delta_0, \delta_1, \delta_2, \ldots\}$ and find the "first" scale that is $\geq \delta_c$.

Once such $\delta_{R_{first}} \geq \delta_c$ is found, sequent smoothing is no longer relevant

→ The collapse of the region $R_{first}$ will crash any internal structure formation.

By considering only the biggest smoothing scale that has $\delta_R \geq \delta_c$ → avoid double counting.

Instead of thinking we progressively reduce the smoothing scale until $\delta_c$ is hit, an alternating view using $S$ is more common.

Recall that as $R$ and $s$ (as well as $M$) are equivalent.

As $R \Rightarrow \infty$, $S \Rightarrow 0$ and decreasing $R$ means increasing $S$.

$\delta_\ell(R)$ is a random variable (mean $= 0$, variance $= S$)
(Gaussian)

→ A random walk with an absorbing barrier.


We start from the origin where $(s, \delta) = (0, 0)$.
Increasing $s$ (smaller $R$) the random walks go up and down with increasing fluctuation.

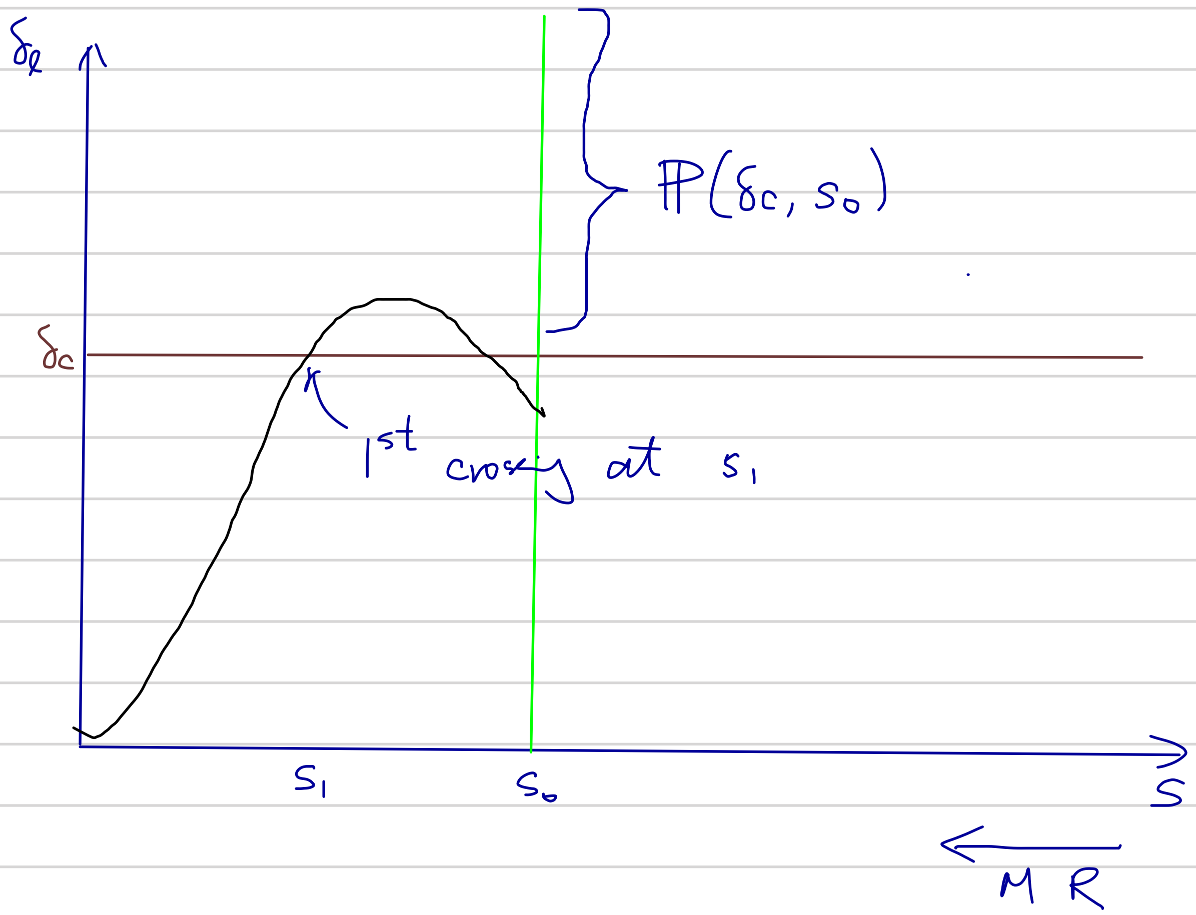
Follow each trajectory until the walk hits the absorbing barrier.

So what we want to know is the first upcrossing probability of random walk across the $\delta_c$ – barrier.

Approach 1 : Analytic

Approach 2 : monte-carlo simulation.

Explanation of the factor "2" in Press-Schechter




Consider a random walk first crosses $\delta_c$ at $s_1$, (for a Markovian walk) it is equally likely to go up or down.

Hence the survival prob. (never cross barrier):

$$\Pi(\delta, s, \delta_c) = p_g(\delta, s) - p_g(2\delta_c - \delta, s).$$

$$f(\delta_c, s) = -\frac{\partial}{\partial s} \int_{-\infty}^{\delta_c} \Pi(\delta, s, \delta_c)\, d\delta$$

$$\Pi(\delta_n, s_n, \delta_c) = \int_{-\infty}^{\infty} d\delta_{n-1}\ \Pi(\delta_{n-1}, s_{n-1}, \delta_c)\ p(\delta_n, s_n \mid \delta_{n-1}, s_{n-1})$$

where conditional prob.

$p(\delta+\Delta\delta, s+\Delta s \mid \delta, s)$ is Gaussian distributed with

$$\mu = \langle \Delta\delta \mid \delta \rangle = \frac{\langle \Delta\delta\, \delta \rangle}{\langle \delta^2 \rangle}\, \delta$$

$$\sigma^2 = \langle \Delta\delta^2 \rangle - \frac{\langle \Delta\delta\, \delta \rangle^2}{\langle \delta^2 \rangle}$$

These are general

Fokker-Planck eqn.

$$\frac{\partial \Pi}{\partial s} = \lim_{\Delta s \to 0} \frac{1}{\Delta s} \left[ \frac{1}{2} \frac{\partial^2 \langle \Delta\delta^2 \mid \delta \rangle \Pi}{\partial \delta^2} - \frac{\partial \langle \Delta\delta \mid \delta \rangle \Pi}{\partial \delta} \right]$$

$$= \frac{1}{2} \frac{\partial^2 \Pi}{\partial \delta^2}$$

# First Upcrossing

What we are trying to get is the first upcrossing probability across some barriers — constant $\delta_c$.

Another interpretation is $\delta_c$ is an absorbing barrier > once random walks touch it they are absorbed.

$$\mathbb{P}(\delta_c, S_0) = \int_0^{S_0} dS_1\, f(\delta_c, S_1)\; p(\delta \geq \delta_c, S_0 \mid \delta_c, S_1, \text{first})$$

$$= \int_0^{S_0} dS_1\, f(\delta_c, S_1)\, \mathbb{P}(\delta_c, S_0 \mid \delta_c, S_1, \text{first})$$

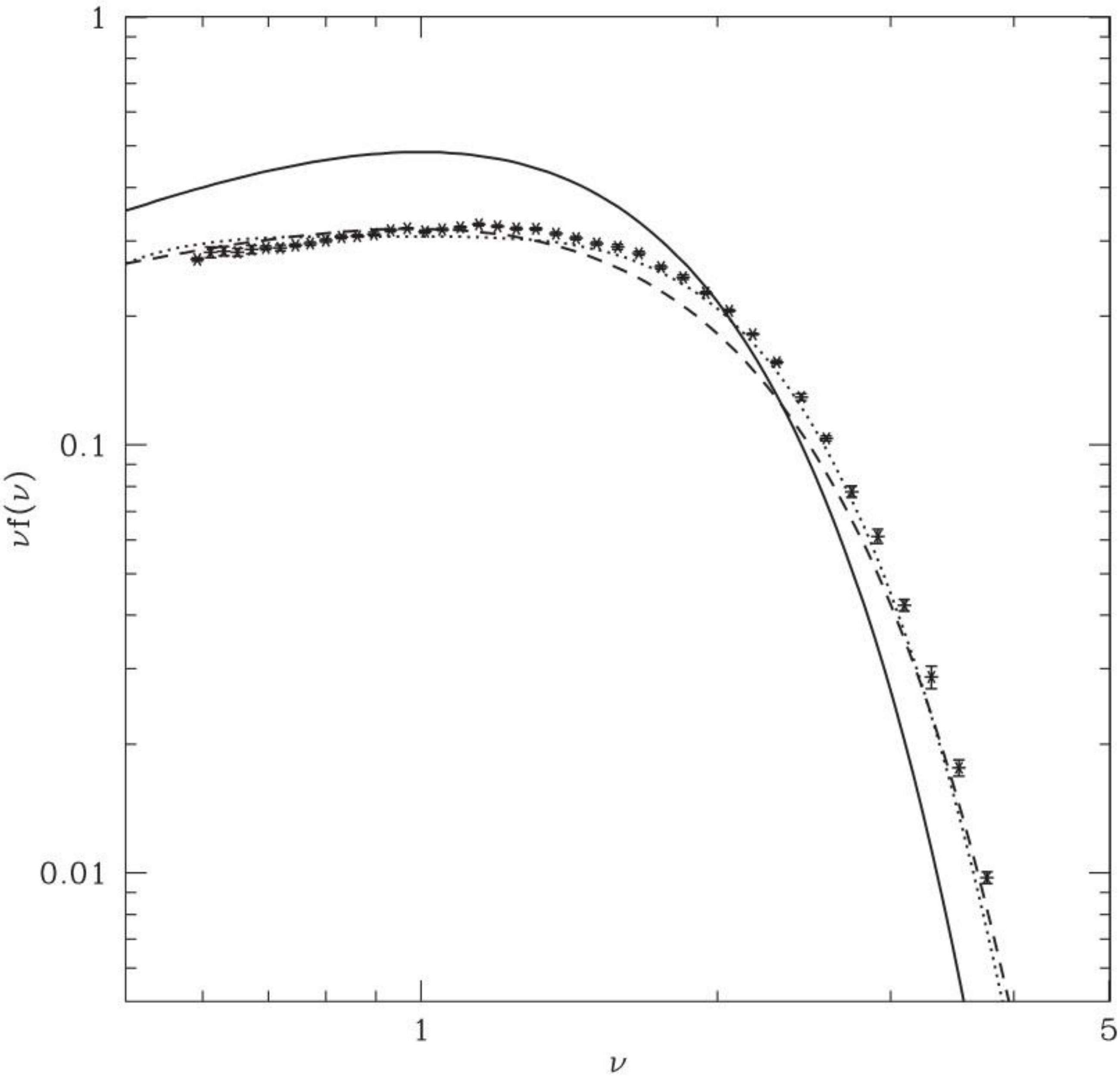
$$\mathbb{P}(\delta_c, S_0) = \int_0^{S_0} dS_1\; f(\delta_c, S_1)\; \mathbb{P}(\delta_c, S_0 \mid \delta_c, S_1, \text{first})$$

For sharp-k filter,

$$\mathbb{P}(\delta_c, S_0 \mid \delta_c, S_1, \text{first}) = \mathbb{P}(\delta_c, S_0 \mid \delta_c, S_1) = \frac{1}{2}$$

$$\Rightarrow \quad \mathbb{P}(\delta_c, S_0) = \frac{1}{2}\int_0^{S_0} dS_1\; f(\delta_c, S_1)$$

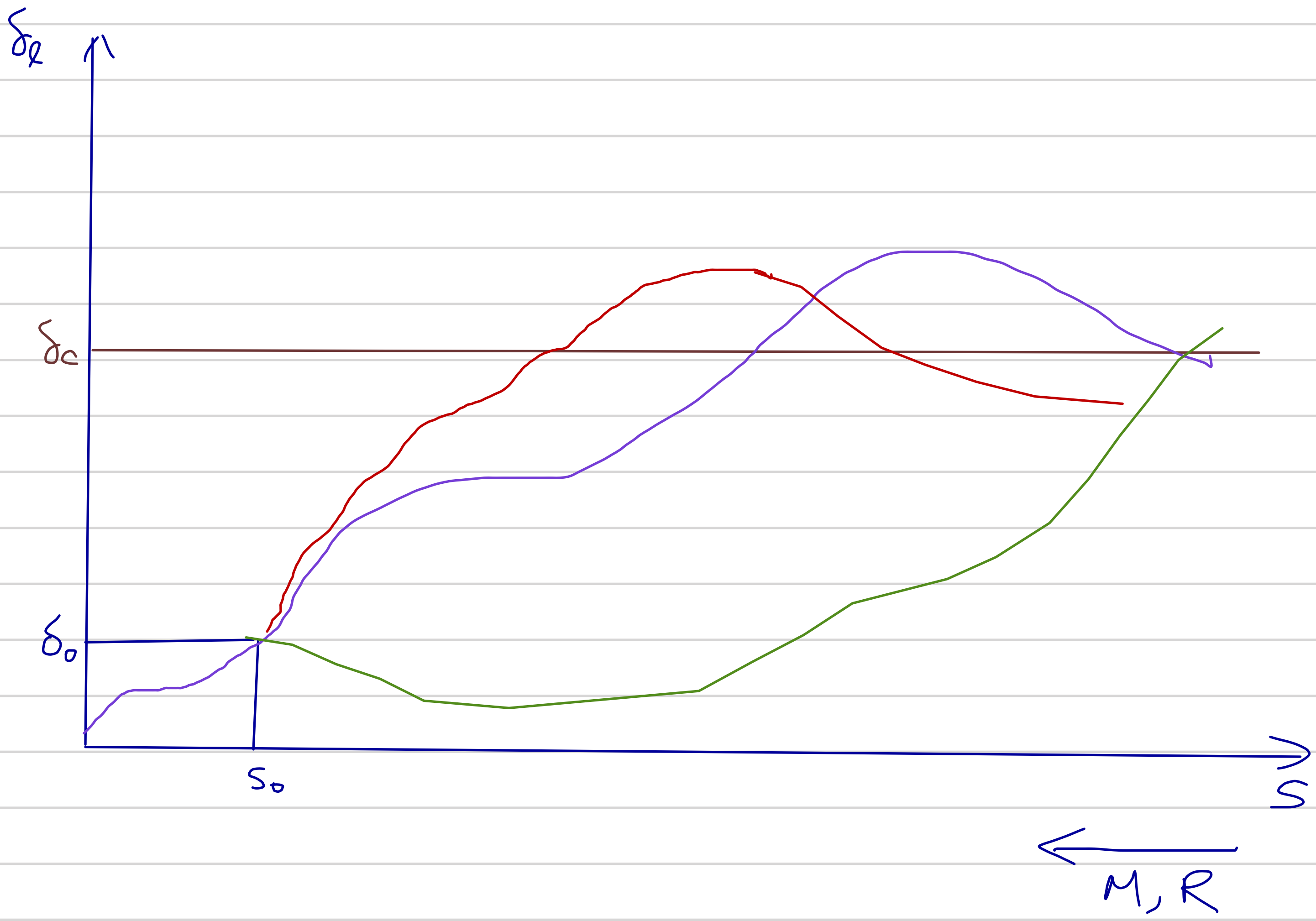
or $$f(\delta_c, S_0) = 2\frac{\partial}{\partial S_0}\mathbb{P}(\delta_c, S_0) = \frac{\delta_c}{S}\frac{e^{-\delta_c^2/2S}}{\sqrt{2\pi S}}$$




Zentner 2007

## 2. Halo Bias

Bias : a response of the halo number density to the underlying matter density field




In the random walk picture, we essentially shift the origin.

The first crossing probability

$$f(\delta_c, s \mid \delta_0, s_0) = f(\delta_c - \delta_0, s - s_0)$$

Markovian

$$f(\delta_c, s \mid \delta_0, s_0) = \frac{\delta_c - \delta_0}{(s - s_0)^{3/2}} \, \frac{e^{-(\delta_c-\delta_0)^2/2(s-s_0)}}{\sqrt{2\pi}}$$

Relative change $(s_0 \Rightarrow 0)$ :

$$\frac{f(\delta_c, s \mid \delta_0, s_0)}{f(\delta_c, s)} - 1 = \frac{\delta_c - \delta_0}{\delta_c} \exp\left[ - \frac{(\delta_c - \delta_0)^2 - \delta_c^2}{2s} \right] - 1$$

$$= \sum_{k>0} b_k^L \, \delta_0^k$$

where $b_0^L = 0$ , $b_1^L = \dfrac{\nu - 1}{\delta_c}$

*** These $b_k^L$ are Lagrangian bias ***

To obtain Eulerian bias :

① Convert $\delta_0$ to $\delta$

② Take into account the change of the comoving volume corresponding to $s_0$ .

$\delta_0 \Longleftrightarrow \delta$ : either PT or spherical collapse

$$V = \frac{V_0}{(1+\delta_0)^3}$$

$$\Rightarrow \quad b_1^E = 1 + b_1^L$$

## 3 Progenitor mass function (Lacey & Cole 1993)

So far we consider halos being collapsed at one single redshift.

The linear extrapolated density contrast to $z_1$ is

$$\delta(z_1) = \frac{D(z_1)}{D(z_0)} \delta(z_0)$$

Take $z_0 = 0$ and $D(z_0) = 1$, $\delta(z_1) = D(z_1)\,\delta(z_0)$
where $D(z_1) < D(z_0)$ for $z_1 > z_0$.

Effectively the amplitude of the variance at all smoothing scales are reduced by the same amount.

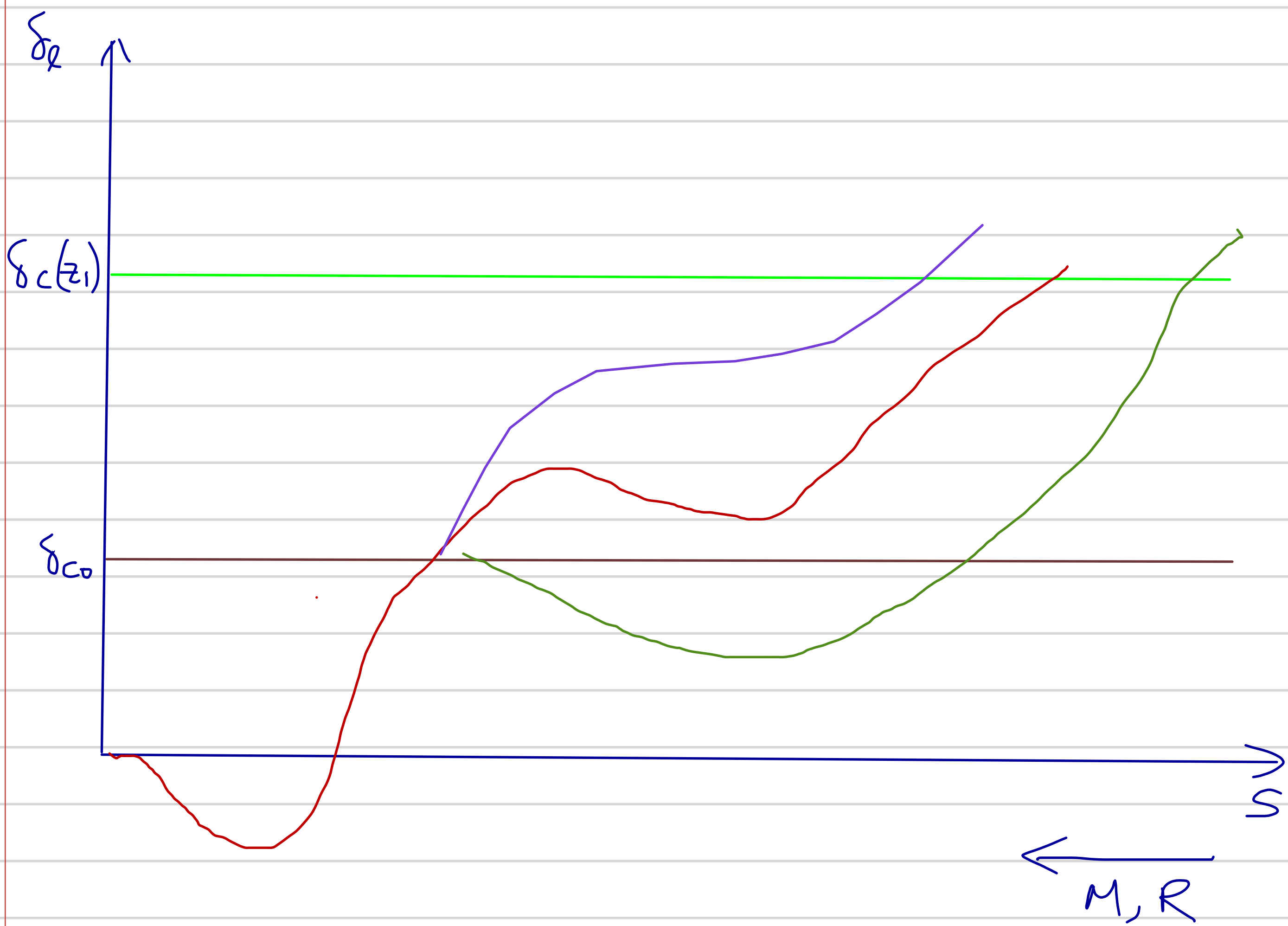
In first crossing probability, all we care about is the relative height of the walk compared to that of the barrier.

$\Rightarrow$ lowering the amplitude of the walk
$=$ increasing the height of the barrier.

$\delta_{c0} \simeq 1.68$ at $z_0 = 0$

$$\delta_c(z) = \frac{\delta_{c0}}{D(z)}$$

linear growth factor

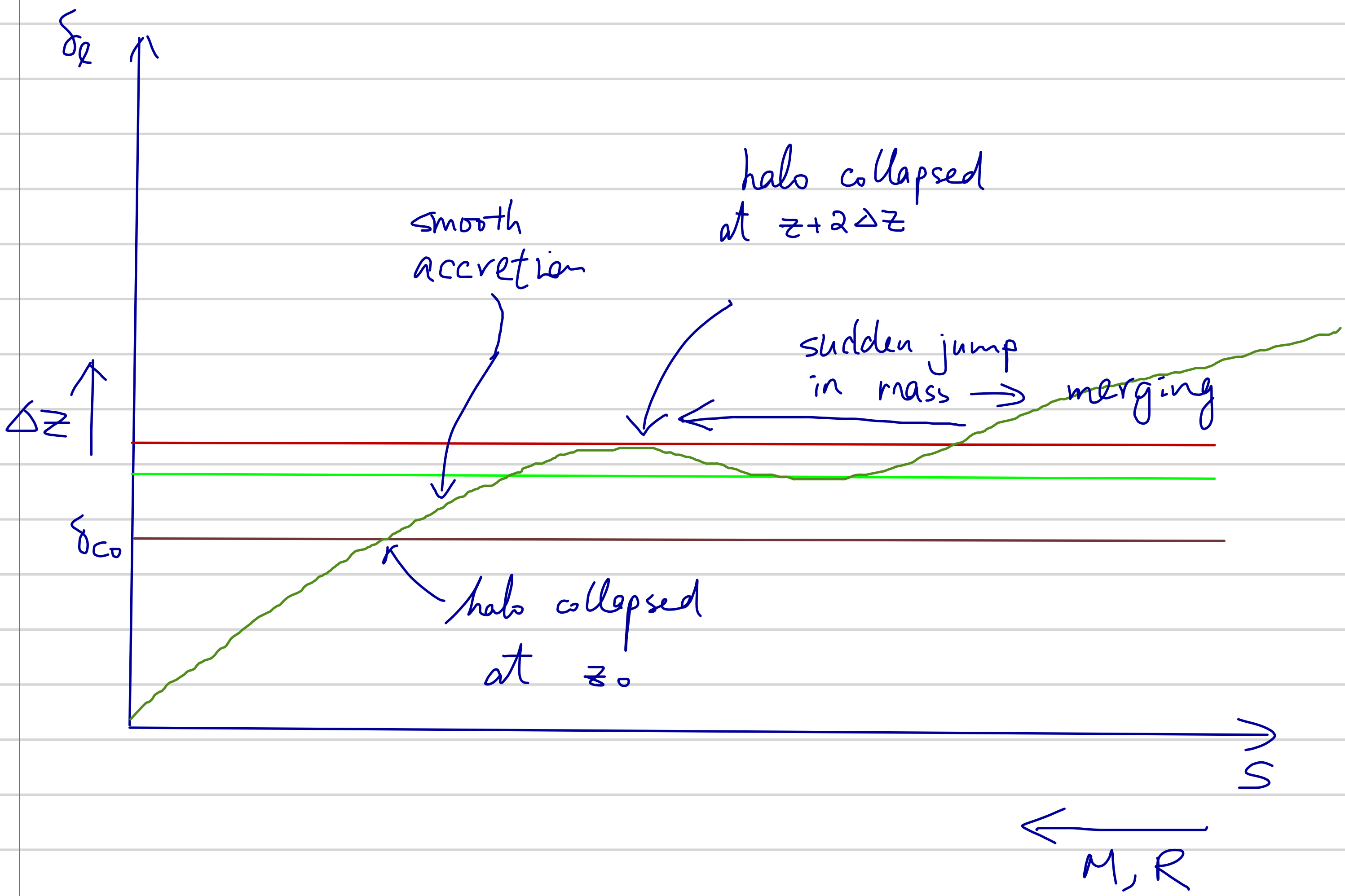



$$f\left(\delta_c(z_1), s_1 \,\middle|\, \delta_{c0}, s_0\right)$$

= mass fraction of halo $M(s)$ which was halo of mass $M_1(s_1)$ at $z_1$

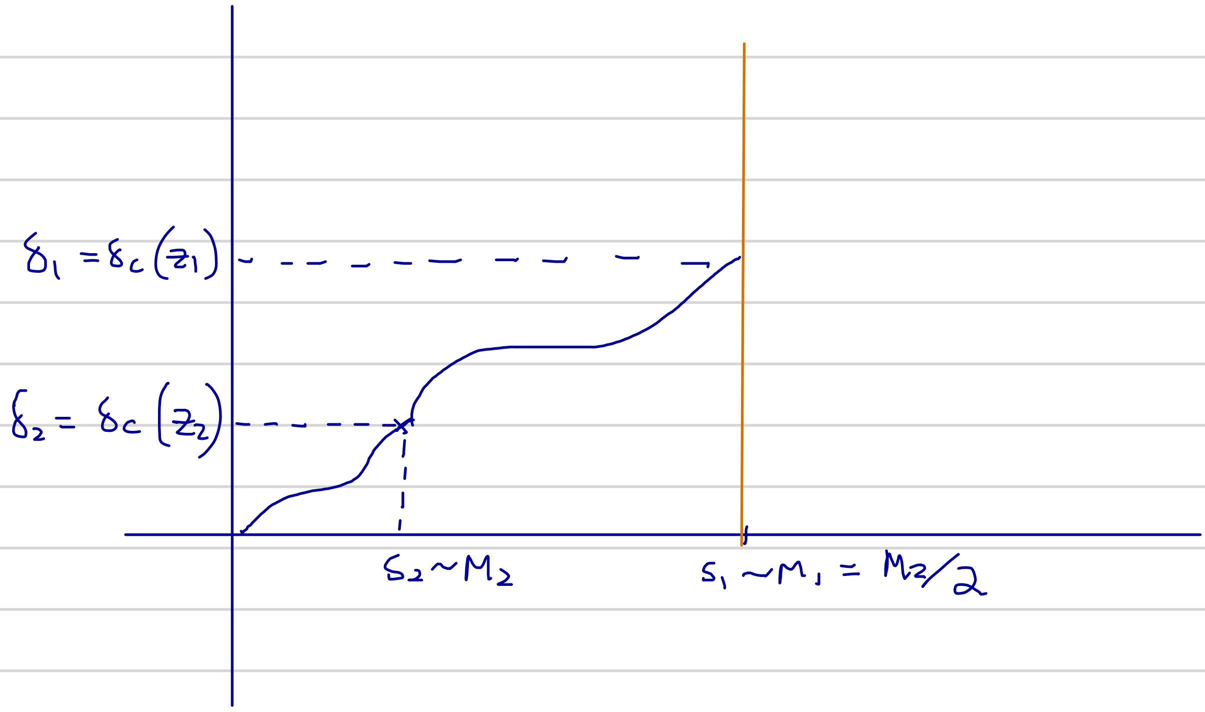
$$\frac{dn(M_2|M_1)}{dM_2} = \frac{M_1 \, f\left(\delta_c(z_2), s_2 \,\middle|\, \delta_c(z_1), s_1\right)}{M_2} \left|\frac{dS_2}{dM_2}\right| dM_2$$

4. Merging history (Lacey & Cole 1993)



## Formation time (Lacey & Cole 1993)

Definition: Formation time $t_f$ = parent halo has $> 50\%$ of the mass




Formation time at $z_1 \sim f(\delta_1, s_1 \mid \delta_2, s_2)$

$$P(z_f > z_1 \mid M_2, z_2) = \int_{s_2}^{s_1} ds_1' \, f(\delta_1, s_1' \mid \delta_2, s_2)$$

But it is not the distribution of formation time

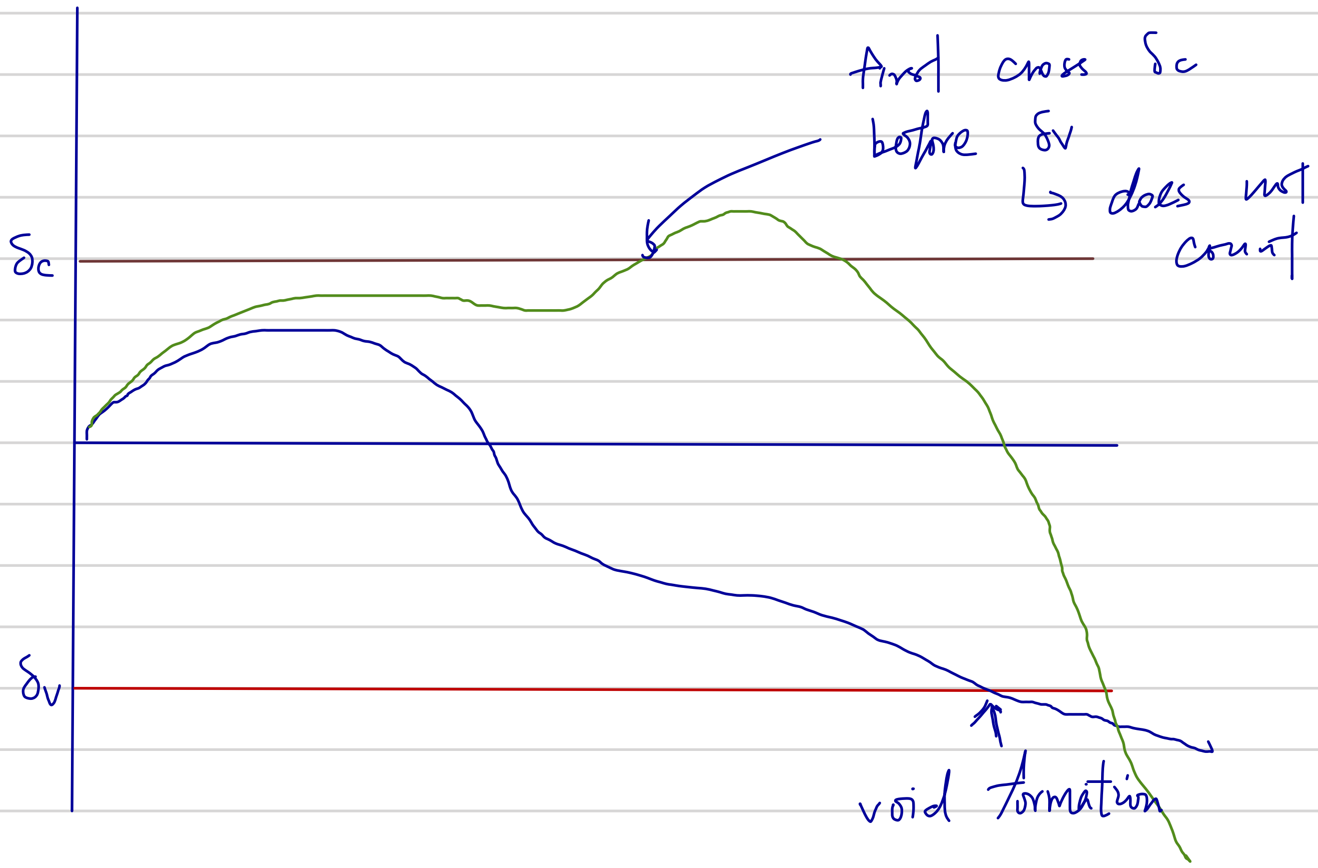
$f(\delta_1, s_1 \mid \delta_2, s_2)$ measures the mass distribution of halos at $z_1$ that form $M_2$ halo at $z_2$.

Exercise: Why? And how to correct that?

## 5. Void Abundance

Similar to halo abundance, the excursion set formalism would be used to compute void abundance.

Spherical collapse predicts another barrier $\delta_v = -2.7$




Void-in-cloud: void cannot reside in halo

2-barrier first crossing:

first crossing probability of the $\delta_v$-barrier without crossing the $\delta_c$-barrier at smaller s.

$$f(s,\delta_v,\delta_c) \simeq f_0(s,\delta_v)\exp\left(-\frac{|\delta_v|}{\delta_c}\frac{\mathcal{D}^2}{4\nu^2} - 2\frac{\mathcal{D}^4}{\nu^4}\right)$$

(Sheth & van de Weygaert 2004) $\mathcal{D} = \dfrac{|\delta_v|}{\delta_c + |\delta_v|}$

# 6. Count-in-cell

So far we consider "collapsed" objects (including voids). The excursion set formalism can also study the distribution of matter.

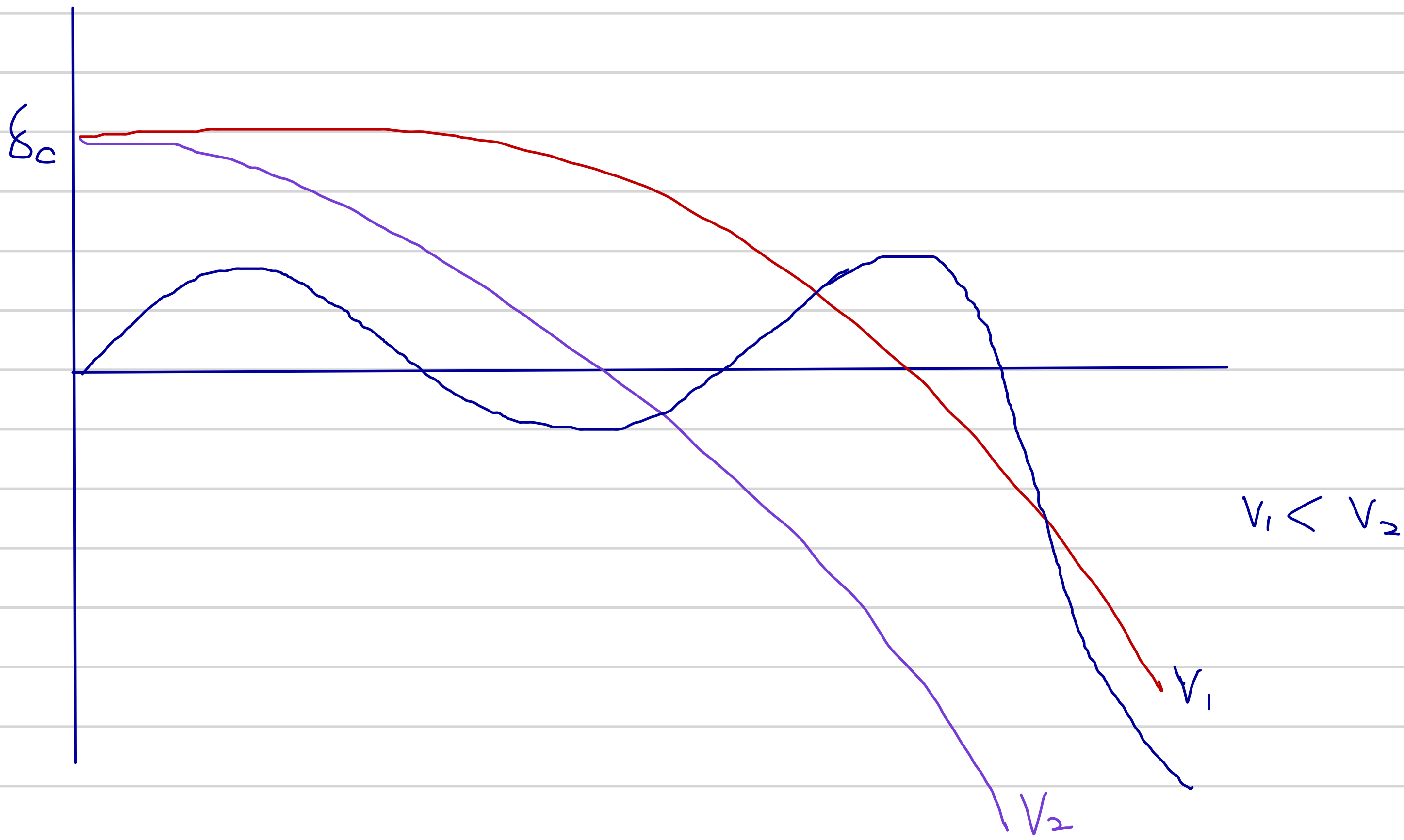
Consider the following approximation for the spherical collapse model

$$\rho_{NL} = 1 + \delta_{NL} = \left(1 - \frac{\delta_\ell}{\delta_c}\right)^{-\delta_c}$$

Assume we are interested in the distribution of mass for some fixed volume $V$. Rewrite the approximation :

$$\frac{M(s)}{\bar{\rho} V} = \left(1 - \frac{\delta_\ell}{\delta_c}\right)^{-\delta_c} \longrightarrow \text{a curve in } (S, \delta_\ell)$$

As $V \to 0$ (halo limit), $\delta_\ell \to \delta_c$




When the random walk touches the barrier the first time at $s_1$ ($\sim M_1$), we can interpret that as the mass $M_1$ is contained in an eulerian volume $V_1$.

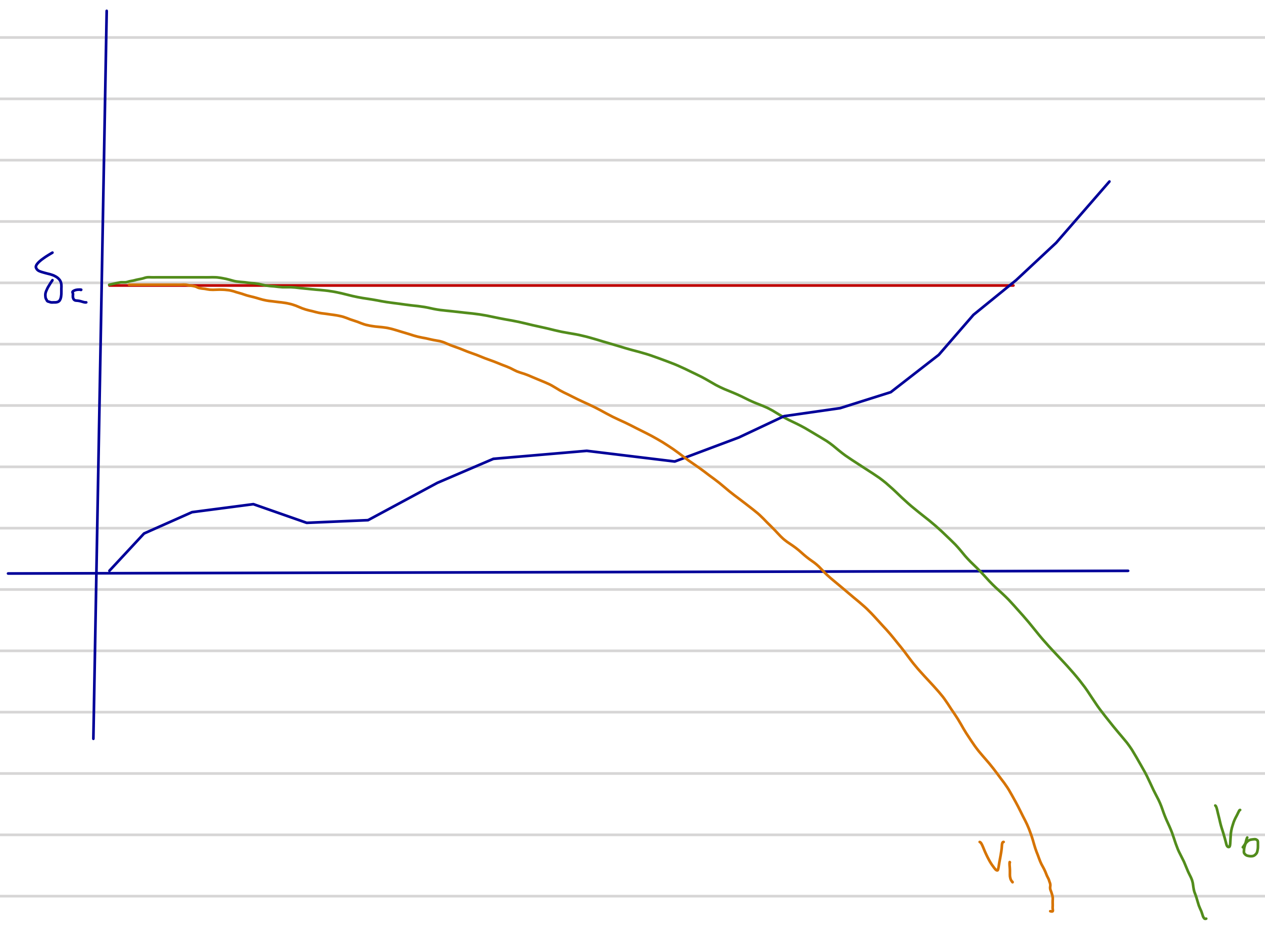
Suppose at $s_2$ ($> s_1$) the walk crosses the same barrier again, so mass $M_2$ ($< M_1$) is contained in the same eulerian volume $V_1$.

Such mass assignment preserves the ordering in Lagrangian space, i.e. no shell crossing.

Hierarchy of barriers of different eulerian volumes also conserve this ordering.

$$\rho_{NL}\, p(\rho_{NL}; V) = f(s(\rho_{NL}); V)$$

# 7. Density profile around halo




The first crossing prob. across segment of eulerian barrier given that the walk first upcrossed $\delta_c$-barrier at some $s \Rightarrow$ mass contained in various eulerian sphere around halo.

# Random walk generation

At each smoothing scale $s$, $\delta_R$ at different position $\vec{x}$ is a random variable whose mean is zero and variance is $s$.

To generate random walks for Monte-carlo simulation,

discretize $s$ into $n$ small steps of size $\Delta s$.

Generate a gaussian random number at each step with zero mean and variance $= \Delta s$

Height of random walk at step ith

$$\delta_i = \delta(s = i\Delta s) = \sum_{j\leq i} g_j$$

$$\langle \delta_i \rangle = \sum_{j\leq i} \langle g_j \rangle = 0$$

$$\langle \delta_i^2 \rangle = \sum_{j\leq i} \sum_{k\leq i} \langle g_j g_k \rangle = \sum_{j\leq i} \sum_{k\leq i} \Delta s \, \delta_{jk}$$

$$= \sum_{j\leq i} \Delta s$$

$$= i \Delta s$$

$$\langle \delta_i \delta_\ell \rangle = \sum_{j\le i} \sum_{k\le \ell} \langle g_j g_k \rangle = \sum_{j\le i} \sum_{k\le \ell} \Delta_s \delta_{jk}$$

$$= \sum_{j\le \min(i,\ell)} \Delta_s$$

$$= \min(i,\ell)\, \Delta s$$

These are valid for sharp-k filter (and assume the density field is Gaussian distributed)

$$\langle \delta_{R_1} \delta_{R_2} \rangle = \int \frac{dk}{k} \Delta(k)\, \underbrace{W(kR_1)}_{\theta(1-kR_1)}\, \underbrace{W(kR_2)}_{\theta(1-kR_2)}$$

$$= \int_0^{\min\left(\frac{1}{R_1},\frac{1}{R_2}\right)} \frac{dk}{k} \Delta(k)$$

$$= s\left(\min\left(\frac{1}{R_1}, \frac{1}{R_2}\right)\right)$$

$$= \min\left(s(R_1),\ s(R_2)\right)$$

For other filters, the computation is more expensive since steps are correlated.

$$\langle \delta_{R_1} \delta_{R_2} \rangle = \int \frac{dk}{k} \Delta(k)\, W(kR_1)\, W(kR_2)$$

And $W(kR)$ is no longer a step function.

For example the Gaussian filter

$$W(kR_1)\, W(kR_2) = \exp\left[ - \frac{k^2 (R_1^2 + R_2^2)}{2} \right]$$

To generate random walks, consider a random variate $Y_i$ that is Gaussian distributed with zero mean and variance

$$\text{variance} = \frac{k_i^3 P(k_i)}{2\pi^2}$$

Then $\delta_i = \sum_i \Delta \ln k \;\, Y_i \;\, W(k_i R)$

and we have the correct statistics

$$\langle \delta_i \rangle = 0$$

$$\langle \delta_i^2 \rangle = \sum_{i,j} \Delta \ln k \, \langle Y_i Y_j \rangle \, W(k_i R)\, W(k_j R)$$

$$= \sum_{i,j} \Delta \ln k \, \frac{k_i^3 P(k_i)}{2\pi^2} \delta_{ij} \, W(k_i R)\, W(k_j R)$$

$$= \sum_i \Delta \ln k \, \frac{k_i^3 P(k_i)}{2\pi^2} \, W^2(k_i R)$$

# Short summary of excursion set formalism

① Define some barriers in $(\delta, s)$ plane

② Calculate the first crossing probability

③ Relate this first crossing probability to number density

( mass conservation , $s \to$ mass )

## 1. Barrier

The halo mass function derived from a constant barrier $\delta_c$ only matches numerical measurements qualitatively.

Sheth-Tormen (1999, 2002) and Sheth, Mo & Tormen (2001) arrives at a better solution by using a moving barrier motivated by ellipsoidal collapse.

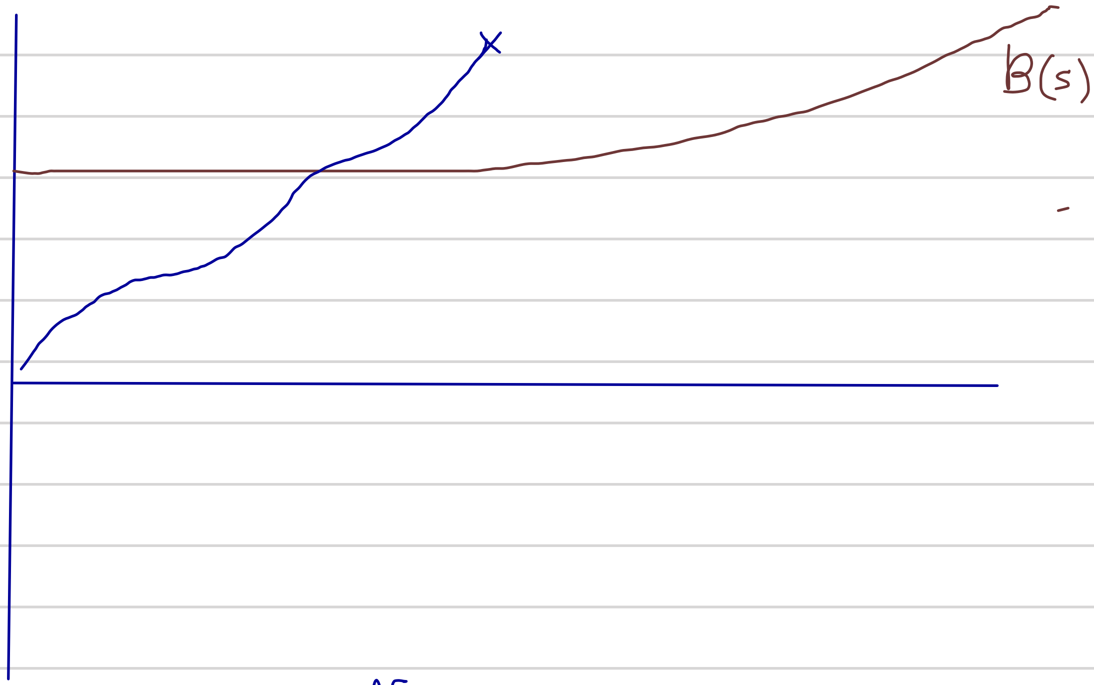
$$B(s) = \sqrt{a}\,\delta_c \left[ 1 + \beta \left( \frac{\sqrt{s}}{\sqrt{a}\,\delta_c} \right)^{2\gamma} \right]$$

$a = 0.7$ , $\beta = 0.4$ , $\gamma = 0.6$

The approximation of the first crossing probability

$$s\,f(s,b) \simeq \frac{b(0)}{\sqrt{2\pi s}} \exp\left[ -\frac{b^2}{2s} \right] \left[ 1 + 0.067 \frac{s^\gamma}{(a\,\delta_c^2)^\gamma} \right]$$

First crossing probability for general barrier (see also Zhang & Hui 2006)




$$p(\delta, s) = \int_0^s ds_1 \; f(b, s_1) \; p(\delta, s \mid b, s_1) \quad , \; \delta > B(s)$$

$$\Rightarrow \mathbb{P}(B(s), s) = \int_B^\infty d\delta' \; p(\delta', s) = \int_0^s ds_1 \; f(b, s_1) \int_B^\infty d\delta' \; p(\delta', s \mid b, s_1)$$

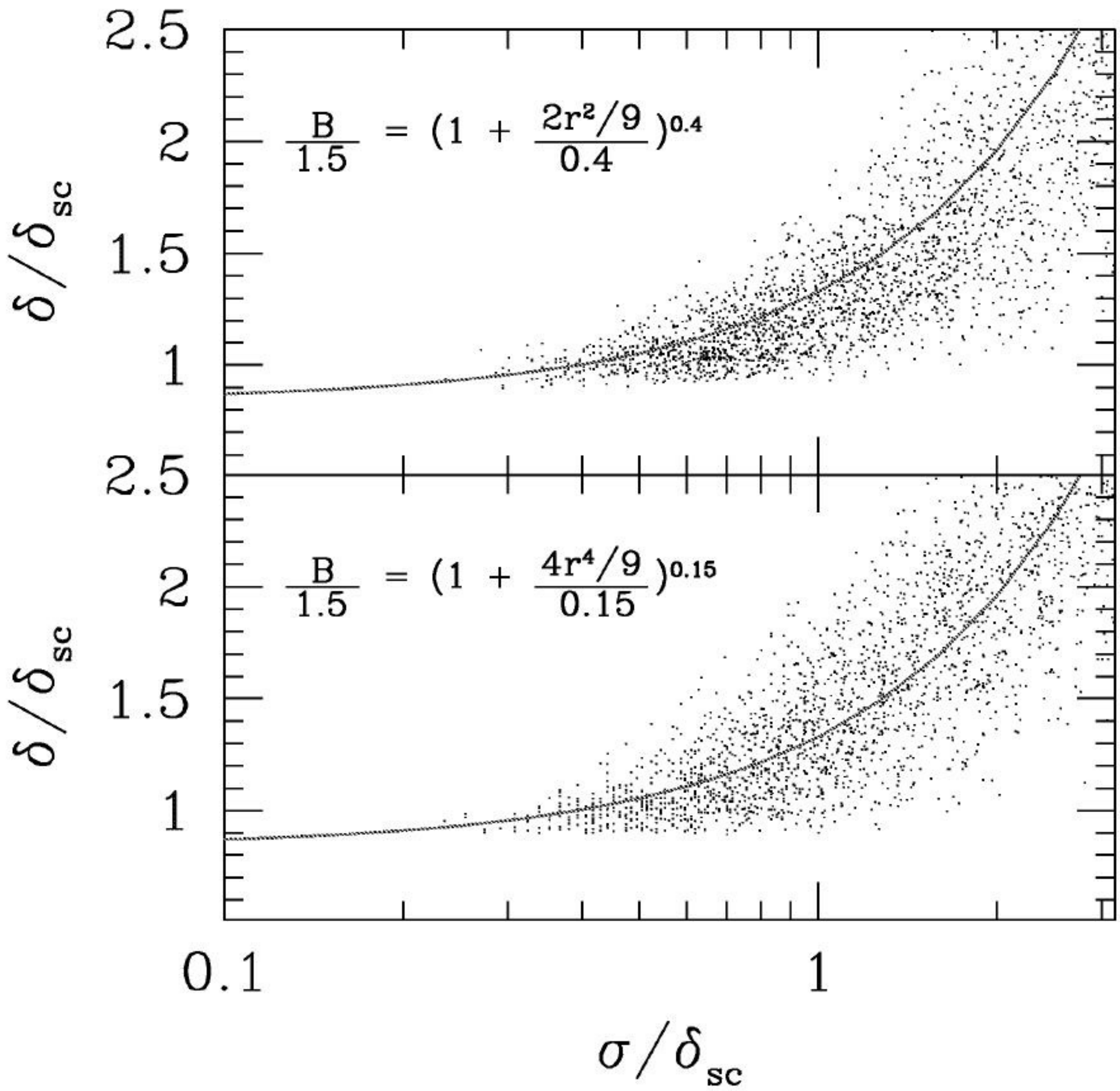
$$\Rightarrow \frac{\partial \mathbb{P}}{\partial s} = \frac{f(B,s)}{2} + \int_0^s ds_1 \; f(b, s_1) \frac{\partial}{\partial s} \mathbb{P}(B, s \mid b, s_1)$$

$$\frac{B}{S} \frac{e^{-B^2/2s}}{\sqrt{2\pi s}} = f(B,s) + \int_0^s ds_1 \; f(b, s_1) \frac{e^{-(B-b)^2/2(s-s_1)}}{\sqrt{2\pi (s-s_1)}} \frac{(B-b)}{s-s_1}$$

Taylor expand $b(s_1)$ around $B(s)$

$$s f(B,s) = \left[ B - s \frac{\partial B}{\partial s} \right] \frac{e^{-B^2/2s}}{\sqrt{2\pi s}} - \sum_{i=2}^{\infty} \frac{s^i}{i!} \frac{\partial^i b}{\partial s^i} \int_0^s ds_1 \; f(b, s_1)$$

$$\frac{e^{-(B-b)^2/2(s-s_1)}}{\sqrt{2\pi (s-s_1)}} \left( \frac{s_1}{s} - 1 \right)^{i-1}$$



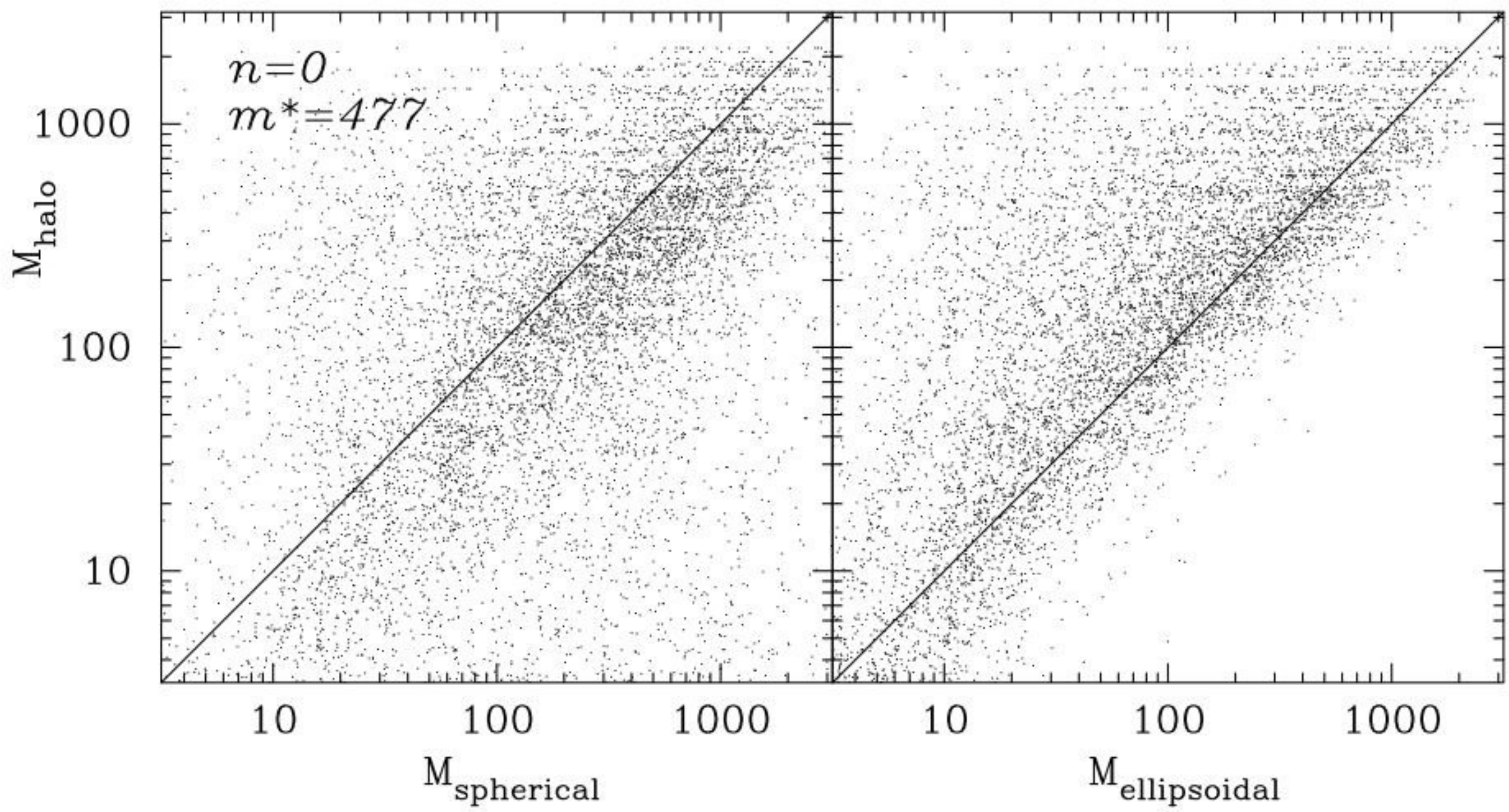

Sheth - Tormen (2002)

Indication of a moving barrier

① Require higher density for less massive halos

② At massive end, $\delta/\delta_c < 1$.

(see recent work in Robertson et. al. 2009)




Sheth, Mo & Tormen 2001

Scatter between $M_{halo}$ and theoretical predictions
(random particles)

## Scaling of barrier

One may notice the moving barrier contains a prefactor $\sqrt{a}$ which rescales the height of the barrier. Where does it come from?

### Stochastic barrier (see Maggiore & Riotto 2009, Corasaniti & Achitouv 2011)

In addition to a random walk in $\delta_\ell$, suppose the barrier itself is also stochastic.

$B(s) = \delta_c + Y_s$ , where $Y_s$ is Gaussian random variate with $\langle Y_s \rangle = 0$ , $\langle Y_s^2 \rangle = D_B S$

Assume $Y_s$ and $\delta_R$ are uncorrelated, then the resulting first crossing probability will have

$$\delta_c \Rightarrow \delta_c / \sqrt{1+D_B}$$

Reason: We want to find $\delta_s \geq B(s) = \delta_c + Y_s$

$$\delta_s - Y_s \geq \delta_c$$

uncorrelated Gaussian $X_s$ with $\langle X_s \rangle = 0$ , $\langle X_s^2 \rangle = S + D_B S = (1+D_B) S$

Rescaling variance by $1+D_B$ is equivalent to rescaling $\delta_c$ by $\frac{1}{\sqrt{1+D_B}}$

## Distribution of deterministic barrier

The scattering in the $(\delta, s)$ plot can also be explained by a distribution of barrier.

For each random walk, the barrier is chosen from a distribution $p(\delta_c)$.

→ it is not the same as the stochastic barrier discussed above.

Note: This concept of distribution of barrier is particularly interesting since it is used in studies of modified gravity models to evaluate the halo/void abundances.

In such MG models, the critical density depends on the environment density (due to some screening mechanisms to restore gravity to GR).

For example, consider

$$B(s) = \delta_c + \beta \sqrt{s}, \quad \beta - \text{Gaussian distributed}$$

$$s f(s) = \int d\beta \, P_g(\beta) \, s f(s|\beta)$$

$$\simeq \left[ 1 - \frac{\Sigma_\beta^2}{4(1+\Sigma_\beta^2)} \right] \frac{\delta_c}{\sqrt{s(1+\Sigma_\beta^2)}} \frac{e^{-\delta_c^2/2s(1+\Sigma_\beta^2)}}{\sqrt{2\pi}}$$

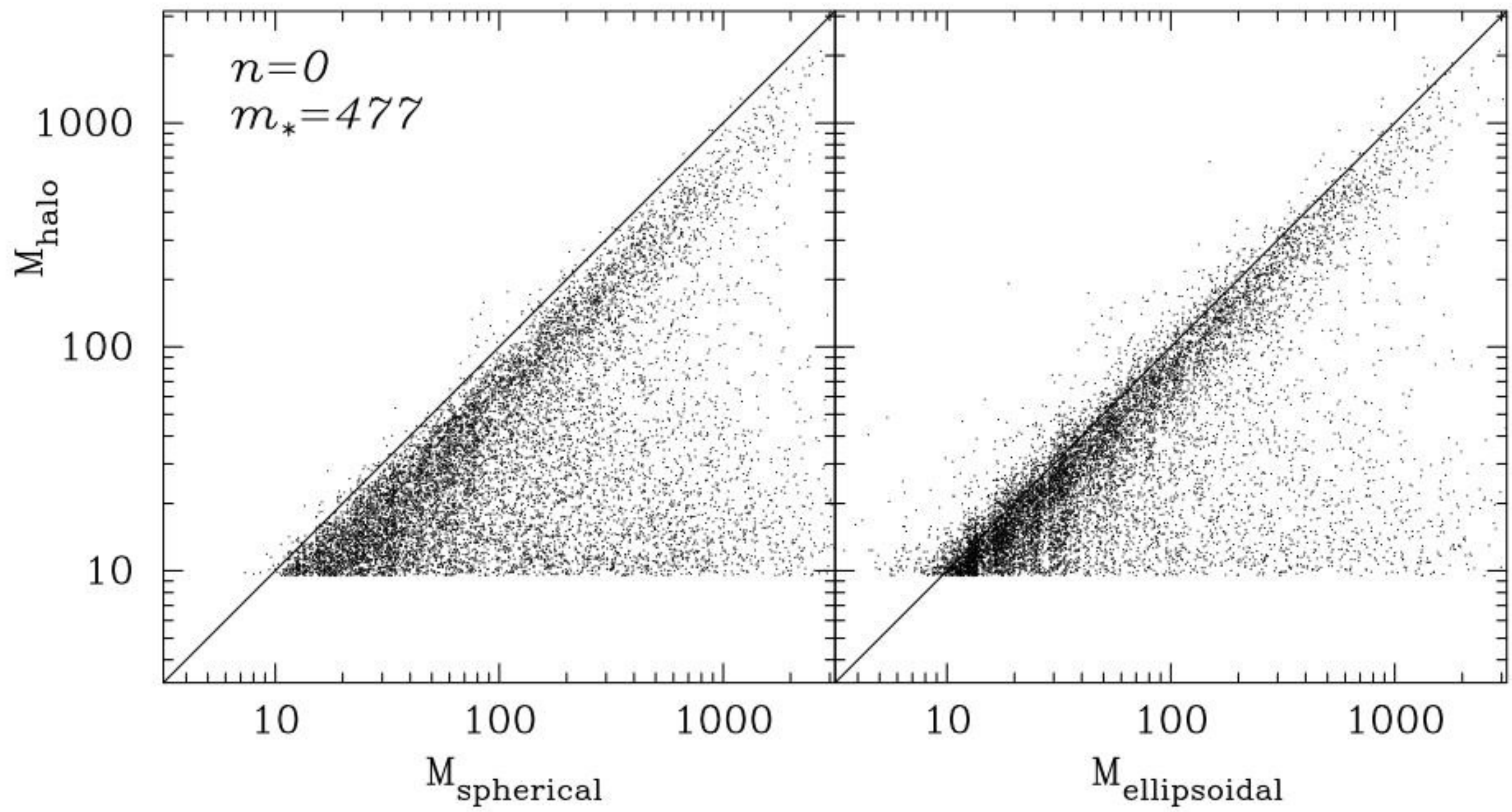
→ give the functional form of ST-mass function.

## Excursion set on special locations

The whole formalism conserves the ordering of concentric shells (i.e. no shell crossing)

Hence the excursion set is looking at the Lagrangian position of the center of the halo.

(see Bond & Myers 1996 for peak-patch algorithm)




Sheth, Mo & Tormen 2001

## 2. First crossing probability

So far we focus on solving the first crossing probability with the "easy" random walk.

Random walk (assume Gaussian initial condition) with sharp-k filter is uncorrelated since each step is independent of the previous steps.

Problem: We define halos with real-space tophat
+ no well defined mass for sharp-k filter

Try to solve the first crossing problem with correlated steps random walk.

Note: The first crossing probability for correlated steps underpredicts the halo abundance
— if we assume the same mass fraction mapping.

Suppose $\{\delta_0, \delta_1, \dots, \delta_n\}$ where $\delta_k = \sum_j W_{kj} g_j$

So we have a system of linear inequality

$$W_{01} g_1 + W_{02} g_2 + \dots + W_{0m} g_m = \delta_0 < \delta_c$$

$$W_{11} g_1 + W_{12} g_2 + \dots + W_{1m} g_m = \delta_1 < \delta_c$$

$$\vdots$$

$$W_{k1} g_1 + W_{k2} g_2 + \dots + W_{km} g_m = \delta_k > \delta_c$$

Method 1 : Brute force Monte-carlo

Each step in the random walk requires a weighted summation of many Gaussian variables.

Possible with modern computation power.

How about extension to calculate halo bias, etc? Need to generate great number of walks to get the statistics.

Method 2: Approximation formula

Several approximations are available

Peacock-Heavens (1991) (or its extension in Paranjape et. al. 2012)

Field theory approaches (see the series by Maggiore & Riotto 2009)

PH : An expansion around completely correlated steps.
MR : An expansion around uncorrelated steps.

# One step beyond approach (Musso & Sheth 2012, 2013)

Original statement: $\delta(S) < B(S)$ for all $S < s$

Assumption: Replace by $\delta(s-\Delta s) < B(s-\Delta s)$

i.e. Random walks are below the barrier at $s-\Delta s$, looking at one preceding step instead of the whole history.

Next expand $\delta(s-\Delta s)$ & $B(s-\Delta s)$ in Taylor series

$$\Rightarrow \delta(s) - \Delta s \frac{\partial \delta}{\partial s} < B(s) - \Delta s \frac{\partial B}{\partial s}$$

Hence $B(s) \leq \delta(s) \leq B(s) + \Delta s (\delta' - B')$

and $\delta' \geq B'$

First crossing prob. $f(s)\,ds = \lim_{\Delta s \to 0} \int_{B'}^{\infty} d\delta' \int_{B(s)}^{B(s)+\Delta s(\delta'-B')} d\delta \; p(\delta, \delta')$

$$f(s)\,ds = \Delta s \; p(B, s) \int_{B'}^{\infty} d\delta' \; p(\delta' | B) \, (\delta' - B')$$

$p(B, s)$ — Gaussian distribution

$p(\delta' | B)$ — Gaussian $\left( \langle \delta' | B \rangle = \gamma B \sqrt{\frac{\langle \delta'^2 \rangle}{\langle \delta^2 \rangle}} = \frac{B}{2s}, \right.$
$\left. \langle \delta'^2 \rangle (1-\gamma^2) \right)$

$$\gamma^2 = \frac{\langle \delta \delta' \rangle^2}{\langle \delta^2 \rangle \langle \delta'^2 \rangle}$$

$$s f(s) = \frac{e^{-\beta^2/2}}{2\sqrt{2\pi}} \beta_* \left[ \frac{1 + \mathrm{erf}\left(\Gamma \beta_* / \sqrt{2}\right)}{2} + \frac{e^{-\Gamma^2 \beta_*^2/2}}{\sqrt{2\pi}\,\Gamma \beta_*} \right]$$

$$\Gamma^2 = \frac{\gamma^2}{1-\gamma^2}, \quad \beta = \frac{B(s)}{\sqrt{s}}, \quad \beta_* = -\beta \frac{\partial \ln \beta}{\partial \ln \sqrt{s}}$$

This treatment works very well ( $s \sim$ a few times $\delta_c$ )

Dependence on $\Gamma$ breaks universality

① For small $s$, multiple crossings are less common for correlated steps.

② Simple improvement can include multiple crossings

$$f(s) = \int_0^\infty d\delta' \, \delta' \, p(B, \delta')$$

$$- \int_0^s ds_1 \int_0^\infty d\delta_1' \, \delta_1' \int_0^\infty d\delta' \delta' \, p(B_1, \delta_1', B, \delta')$$

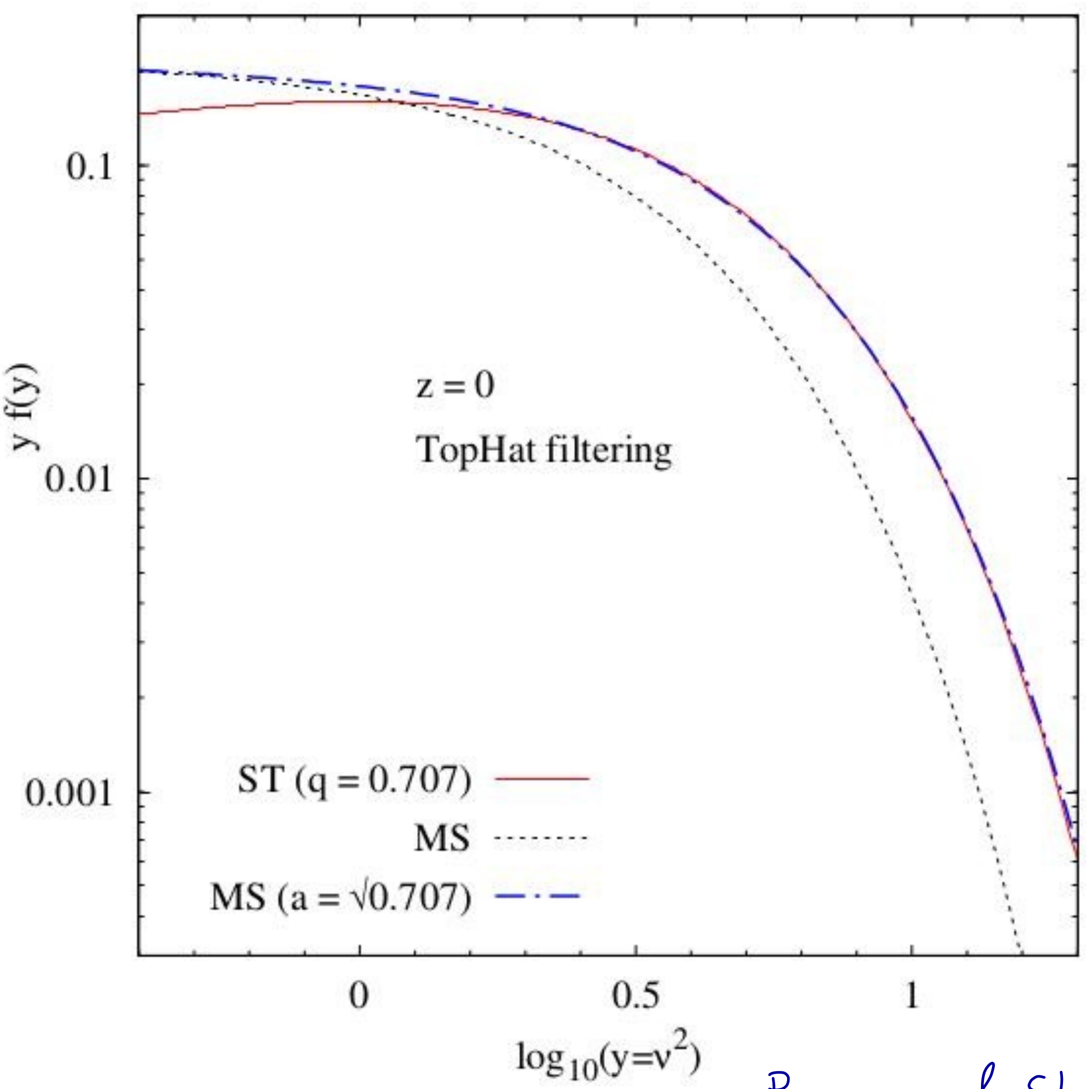
$$+ \ldots$$

Note: For first crossing probability with correlated steps in modified gravity, the one-step approach works for $\Delta \geq 1$ ( without the correction term).

P.S. In MG models, the calculation actually involves the upcrossing $p(B, \Delta', b, \delta')$

Note 2: Although this one-step approach is a very good approximation for the first upcrossing probability, applying it with constant $\delta_c$ as well as the mass fraction does not match the measured mass function.

$\nu \Rightarrow \sqrt{a}\,\nu$, $a \simeq 0.7$




Paranjape & Sheth 2012

## 3. Going from first crossing probability to abundance

Excursion set formalism discussed so far does NOT consider correlation between walks.

⇒ Any points can be the center of mass of halos

It does not seem to be the case.
(Bond & Myers 1996, White 1996, Sheth, Mo & Tormen 2001, Ludlow & Porciani 2011)

Peak-Patch Algorithm (Bond & Myers 1996)

1. Identify peaks in initial condition
2. Perform excursion set on those peaks
3. Use mass ranking to select overlapping proto-halo patches.

⇒ These steps take into account the non-local effect.

Excursion set peak EPS (Paranjape & Sheth 2012, Paranjape, Sheth & Desjacques 2013)

Number density of peak (smoothed with some R)

$$N_{pk}(\nu) = \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{G_0(\gamma\nu, 1-\gamma^2)}{(2\pi R_*^2)^{3/2}} \quad \text{(BBKS 1986)}$$

Gaussian filter volume

$$R_* = \frac{\sqrt{3}\,\sigma_1}{\sigma_2}, \quad \gamma = \frac{\sigma_1^2}{\sigma_0 \sigma_2}$$

Note: Peaks require the first and second orders of the spatial derivatives (encoded in $G_0$)

Problem: ① Mass assignment
We have fixed smoothing scale but we expect high-$\nu$ peaks are associated with more massive halos.

② cloud-in-cloud problem

Consider applying the one-step beyond approach on peak

Random positions

$$f(s) = \int_0^\infty dv \; v \, p(v, \delta_c)$$

Normalized variables

$$\nu = \frac{\delta_c}{\sigma_0} \quad , \quad x = \frac{v}{\sqrt{\langle v^2 \rangle}}$$

$$s f(s) = \frac{e^{-\nu^2/2}}{2\gamma\sqrt{2\pi}} \int_0^\infty dx \; x \; p_g(x; \gamma\nu, 1-\gamma^2)$$

$$f(\nu) = \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{\langle x \mid \gamma\nu, 1-\gamma^2 \rangle}{\gamma\nu}$$

$$\int_0^\infty dx \; x \; p_g(x; \gamma\nu, 1-\gamma^2)$$

Peak

Peak height : $\nu$

Need something to $\frac{d\nu}{ds}$

But for a Gaussian window function

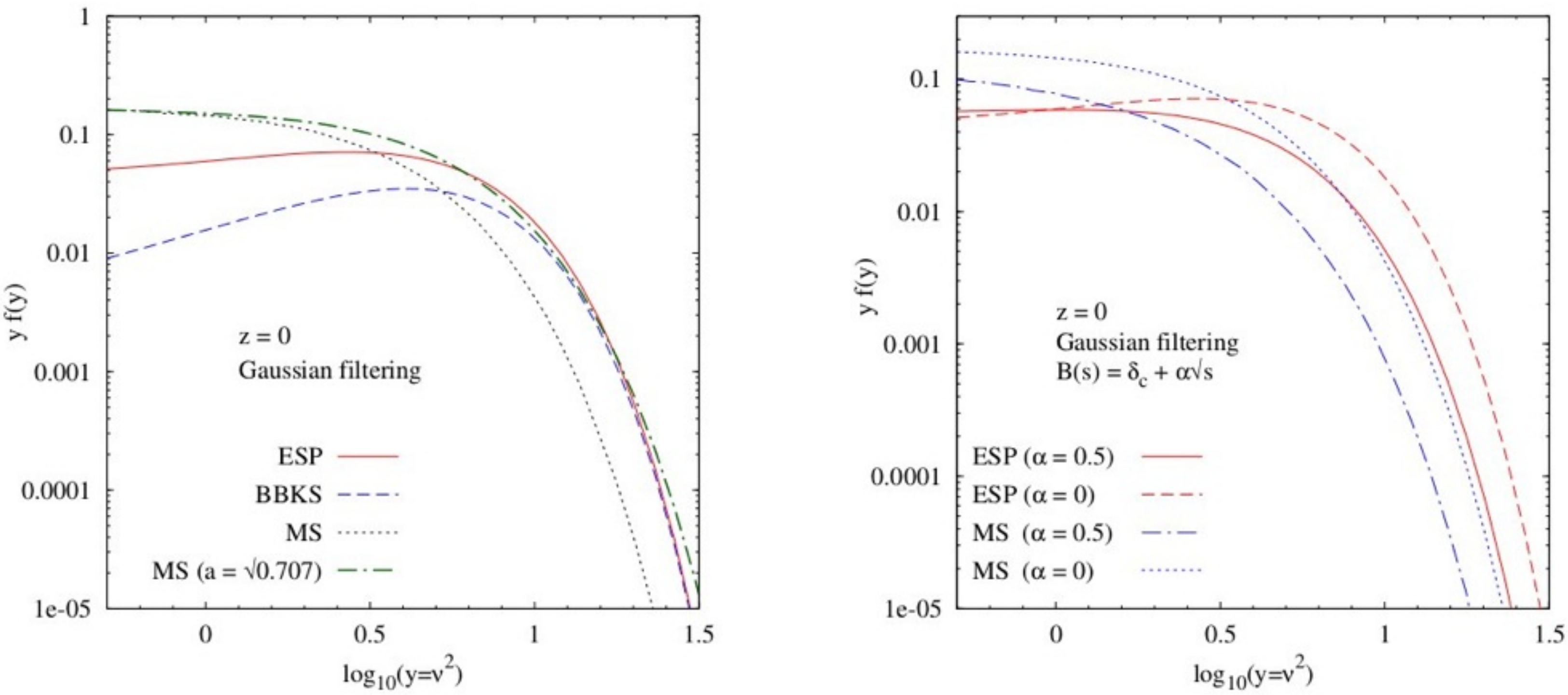
$$\frac{d\delta_R}{ds} \sim \nabla^2 \delta_R \sim x \; (\text{curvature around peak})$$

$$N_{ESP}(\nu) \sim \int_0^\infty dx \; x \; N_{pk}(x, \nu)$$

$$f_{ESP}(\nu) = \frac{e^{-\nu^2/2}}{\sqrt{2\pi}} \frac{V}{V_*} G_0(\gamma\nu, 1-\gamma^2) \frac{\langle x \mid \gamma\nu, 1-\gamma^2 \rangle}{\gamma\nu}$$

$$\frac{G_1(\gamma\nu, 1-\gamma^2)}{G_0(\gamma\nu, 1-\gamma^2)} \quad , \quad G_k(\mu, \sigma^2) = \int_0^\infty dx \; x^k \; F(x) \; p_g(x; \mu, \sigma^2)$$

Same formula as in Appel & Jones (1990)




Paranjape & Sheth 2012

① No rescaling is necessary

② Extension to tophat filter is possible
(see Paranjape, Sheth & Desjacques 2013)

③ EPS does the first 2 steps in the peak-patch algorithm. How about the 3rd step?

# Mapping from Lagrangian space to Eulerian

So far we always assume the number density we found in Lagrangian space = eulerian number density

The 3rd step in the peak-patch algorithm indicates such treatment neglects overlapping Lagrangian proto-halo regions.

i) Change number density
ii) Remapping of mass

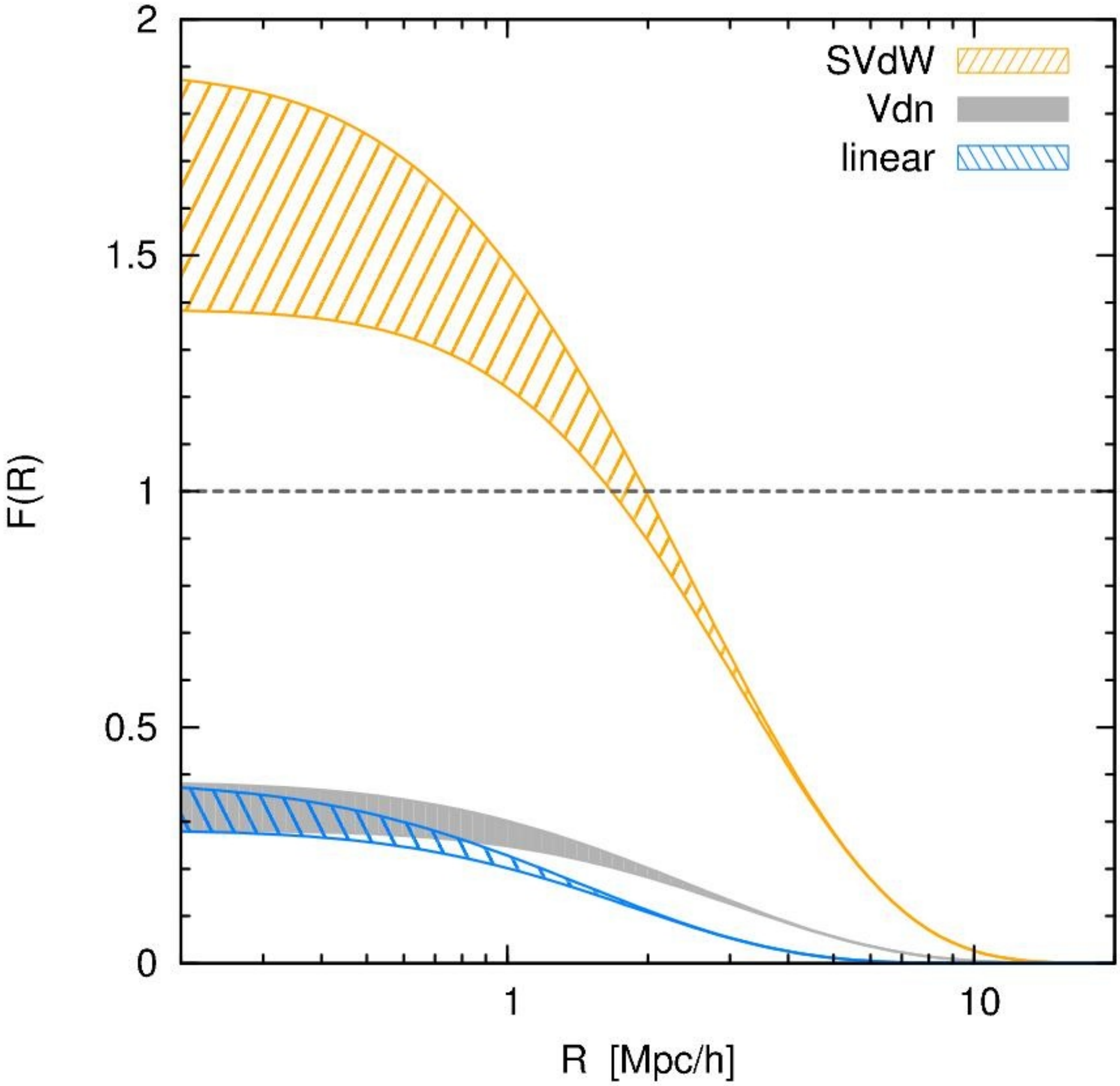
These changes are particularly important for voids — expanding compared to background.

From spherical collapse, voids will have a density contrast of $-0.8$ $(V = 5V_0)$.

We can estimate the volume fraction occupied by voids larger than $V$

$$F(R) = \int_V^{\infty} dm \quad n(m)\ V(m)$$

$$n(m) = \frac{\bar{\rho}}{m^2} s f(s, \delta_v, \delta_c) \left| \frac{d\ln s}{d\ln m} \right| \quad , \quad V(m) = 5V_0(m)$$




Jennings, Li & Hu (2013)

$\Rightarrow$ unphysical result as volume fraction $> 1$

## Volume conserving mapping (Jenning, Li & Hu 2013)

Instead of conserving number density, keep volume fraction conserved

$$V\, n_v(V)\, dV = V_0\, n_L(V_0)\, dV_0$$

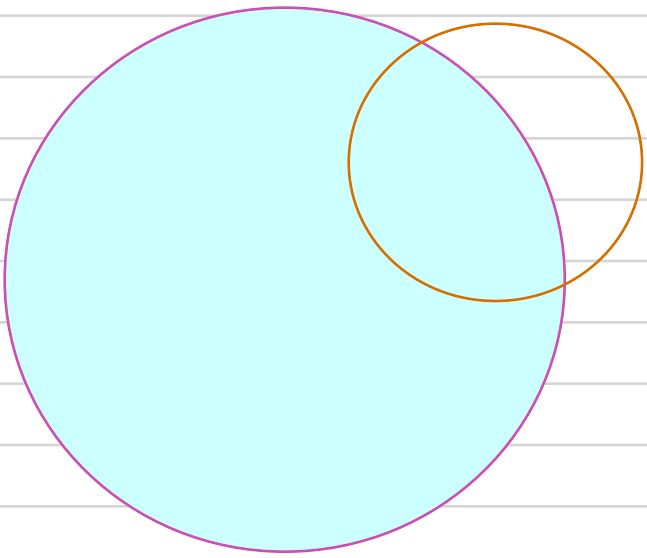
Since $V_0^2\, n_L(V_0)\, d\ln V_0 = \frac{1}{\bar{e}}\, s\, f(s, \delta_v, \delta_c)\, d\ln s$

$$n_v(V)\, dV = \frac{1}{\bar{e} V}\, s f(s, \delta_v, \delta_c)\, d\ln s$$

Note: ① Only a first order estimate

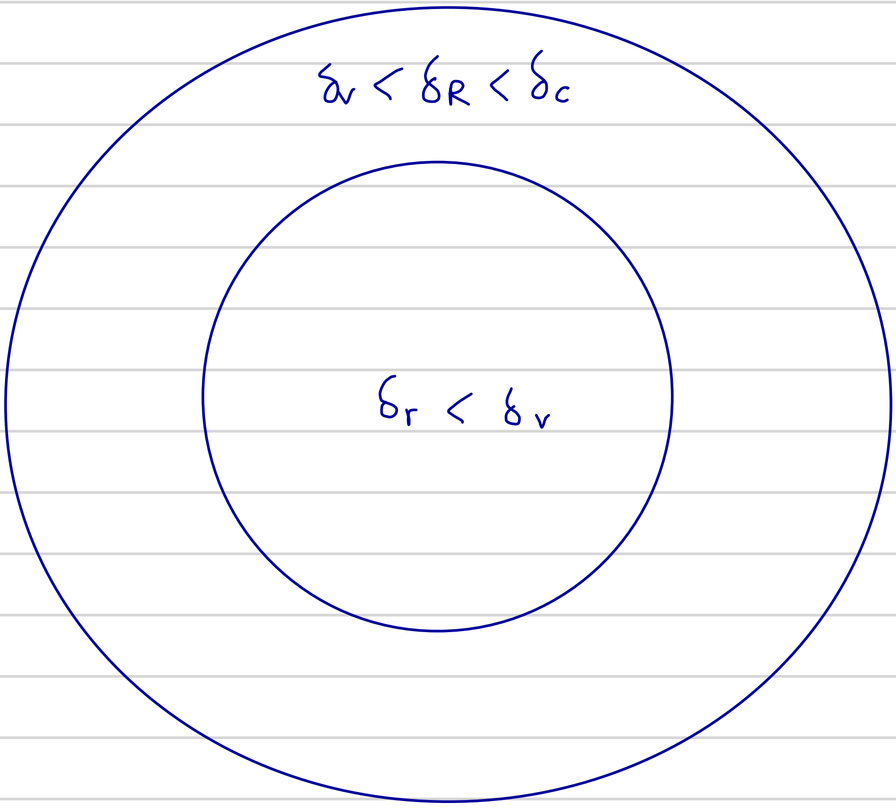
— merging of voids of various sizes

— Total volume conservation after merging



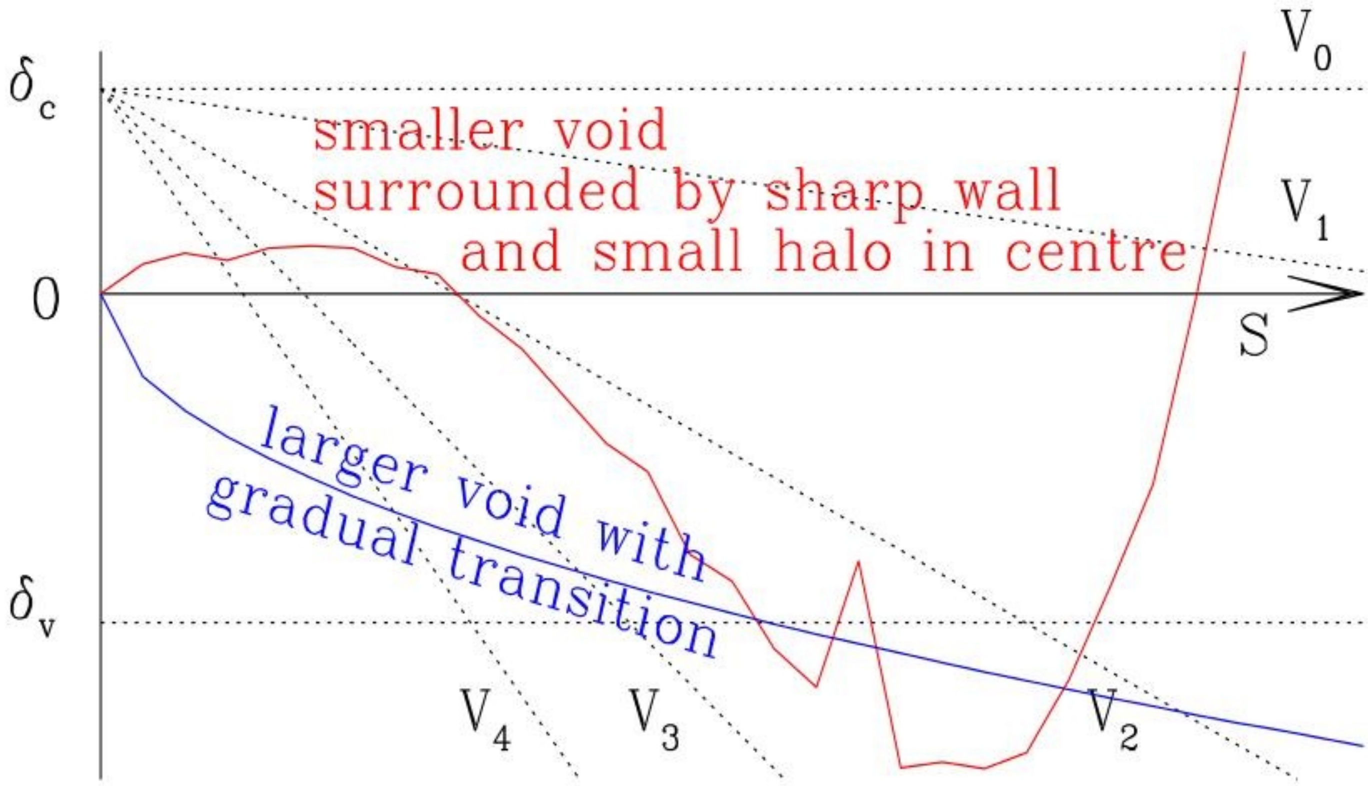
## Improve version of void excursion set
(Paranjape, Lam & Sheth 2012)

The 2-barriers excursion set solves the void-in-cloud problem.
However consider the following situation




The comoving size exterior shell will expand (if $\delta_R < 0$) or shrink (if $\delta_R > 0$)

$\Rightarrow$ it restricts the growth of the inner region which is supposed to expand 5 times.




Paranjape, Lam & Sheth 2012

- Void : $\rho(V) = 0.2\,\bar{\rho}$
- Find the biggest eulerian volume satisfies $\rho \leqslant 0.2\,\bar{\rho}$

$V_4 > V_3 > V_2 > \ldots$ all are eulerian volume

- non-smooth mass at $V_2 \Rightarrow$ wall around void
- void volume fraction $\sim 1.17$

## Alternative approaches

① Using Zeldovich approximation $\{\lambda_1, \lambda_2, \lambda_3\}$ to do 3D random walks

- mass assignment (crossing at different scales)
- normalization (regions that never satisfy some pre-defined conditions)

② Probability conditional formalism (Firmani & Avila-Reese 2013)

→ using the concept of isolated regions (Jedamzik 1995) and conditional probability to rediscover the first upcrossing picture

- use Tinker mass function to calibrate the unconditational cumulative probability
- assume markovian and infer conditional probability
- a mass rescaling of 30% to progenitors
- fitting mass function, progenitor mass function and merging rates

In fact we have many freedoms in the model

① Barrier ( constant vs moving ; stochastic vs deterministic vs ensembles )

② First crossing probability ( correlated vs uncorrelated )

③ Mapping from Lagrangian to Eulerian

( random pt vs peak ; number density conservation vs volume fraction vs total mass )

④ Mass rescaling ?

Obviously mass function alone cannot constrain all these "free" parameters.

Additional information : bias, conditional mass function, progenitor mass function, etc.

# Halo model

Use halo as building block

Missing ingredient $\Rightarrow$ how stuff distributed in halo ?

NFW profile :

$$\rho(r|m) = \frac{\rho_s}{\left(\frac{r}{r_s}\right)\left(1+\frac{r}{r_s}\right)^2}$$

$$m = \int_0^{r_{vir.}} dr\, 4\pi r^2 \rho(r|m) = 4\pi \rho_s r_s^3 \left[\ln(1+c) - \frac{c}{1+c}\right]$$

where $c$ = concentrat^n^ $= \frac{r_{vir}}{r_s}$

Lognormal distribution of concentration :

$$p(c|m,z)\, dc = \frac{d\ln c}{\sqrt{2\pi \sigma_c^2}} \exp\left[-\frac{\ln^2[c/\bar{c}]}{2\sigma_c^2}\right]$$

$$\bar{c} = \frac{9}{1+z}\left[\frac{m}{m_*(z)}\right]^{-0.13}, \qquad \sigma_c \simeq 0.25$$

2-pt correlation function of dark matter

$$\rho(x) = \sum_i \rho(x - x_i | m_i)$$

Summing up all contributions from each halo $(x_i, m_i)$

Using a normalized profile

$$\rho(x) = \sum_i m_i \, u(x - x_i | m_i)$$

$$= \sum_i \int dm \, d^3y \, \delta(m - m_i) \, \delta^3(\bar{y} - \bar{x}_i) \, m \, u(\bar{x} - \bar{y} | m)$$

We have

$$n(m) = \left\langle \sum_i \delta(m - m_i) \, \delta^3(\bar{y} - \bar{x}_i) \right\rangle$$

and $$\bar{\rho} = \langle \rho(\bar{x}) \rangle = \int dm \, n(m) \, m$$

2-pt CF

$$\bar{\rho}^2 [1 + \xi(r)] = \langle \rho(x) \, \rho(x + r) \rangle$$

$$= \int dm \, m \int dm' \int d^3y \, u(\bar{x} - \bar{y} | m) \int d^3z \, u(\bar{x} + \bar{r} - \bar{z} | m')$$

$$\left\langle \sum_{i,j} \delta(m - m_i) \, \delta(m' - m_j) \, \delta^3(\bar{y} - \bar{x}_i) \, \delta^3(z - \bar{x}_j) \right\rangle$$

For $i=j$ : 1-halo contribution

$$\langle \quad \rangle_{1h} = \int dm\ m^2 n(m) \int d^3y\ u(\bar{x}-\bar{y}|m)\, u(\bar{x}+\bar{r}-\bar{y}|m)$$

$$= \int dm\ m^2 n(m) \int d^3y\ u(y|m)\, u(r+y|m)$$

For $i \neq j$ : 2-halo term

$$\langle \quad \rangle_{2h} = \int dm\ m \int dm'\ m' \int d^3y\ u(x-y|m) \int d^3z\ u(x+r-z|m')$$

$$n(m)\, n(m') \left[1 + \xi_{hh}(r, m, m')\right]$$

$$= \bar{\rho}^2 + \int dm\ m\, n(m) \int dm'\ n(m')\, m' \int d^3y\ u(y|m) \int d^3z\ u(z|m')$$

$$\xi_{hh}(y+z-r, m, m')$$

$$\Rightarrow \xi(r) = \xi_{1h} + \xi_{2h}$$

$$\xi_{1h} = \int dm\ n(m) \frac{m^2}{\bar{\rho}^2} \int d^3y\ u(y|m)\, u(y+r|m)$$

$$\xi_{2h} = \frac{1}{\bar{\rho}^2}$$

In Fourier sp. (and $\xi_{hh} = b(m)\, b(m')\, \xi_{lin}(r)$ )

$$P_{1h}(k) = \int dm\; n(m)\, \frac{m^2}{\bar{\rho}^2}\, |u(k|m)|^2$$

$$P_{2h}(k) = P_{lin}(k) \left[ \int dm\; n(m)\, \frac{m}{\bar{\rho}}\, b(m)\; u(k|m) \right]^2$$

$$u(k|m) = \left[ \ln(1+c) - \frac{c}{1+c} \right]^{-1} \Big\{ \sin(kr_s) \left[ \mathrm{Si}\big( (1+c)\, kr_s \big) - \mathrm{Si}(kr_s) \right] - \frac{\sin(c\, kr_s)}{(1+c)\, kr_s} + \cos(kr_s) \left[ \mathrm{Ci}\big( (1+c)\, kr_s \big) - \mathrm{Ci}(kr_s) \right] \Big\}$$

# Halo occupation distribution (Scoccimarro et al 2001, Berlind & Weinberg 2002)

Empirical approach to plant galaxies into halos

1. Some minimum "cutoff" below which no galaxies reside
2. The first galaxy is the center galaxy (usually the BCG)
3. Others (satellites) follow matter distribution

Number density of galaxies

$$n_g = \int dm \, n(m) \, \underbrace{f_{cen}(m)}_{\text{average number of central in m-halo}} \Big[ \overbrace{1}^{\text{central}} + \overbrace{\langle N_s | m \rangle}^{\text{\# density of satellites in m-halo}} \Big]$$

Correlation function of galaxy

$$n_g^2 \left(1 + \xi_{gg}(r)\right) = \langle n_g(x)\, n_g(x+r) \rangle$$

$\langle \ \rangle_{1h}$ — center - satellite; satellite - satellite

$$\langle \ \rangle_{1h,cs} = \int dm\, n(m)\, f_{cen}(m) \langle N_s | m \rangle\, u(r|m)$$

$$\langle \ \rangle_{1h,ss} = \int dm\, n(m)\, f_{cen}(m) \frac{\langle N_s(N_s-1) | m \rangle}{2} \int d^3y\, u(y|m)\, u(y+r|m)$$

# of distinct satellite pairs

$$\langle\ \rangle_{2h} \rightarrow \begin{cases} \text{center-center} \\ \text{center - satellite} \\ \text{satellite - center} \\ \text{satellite - satellite} \end{cases}$$

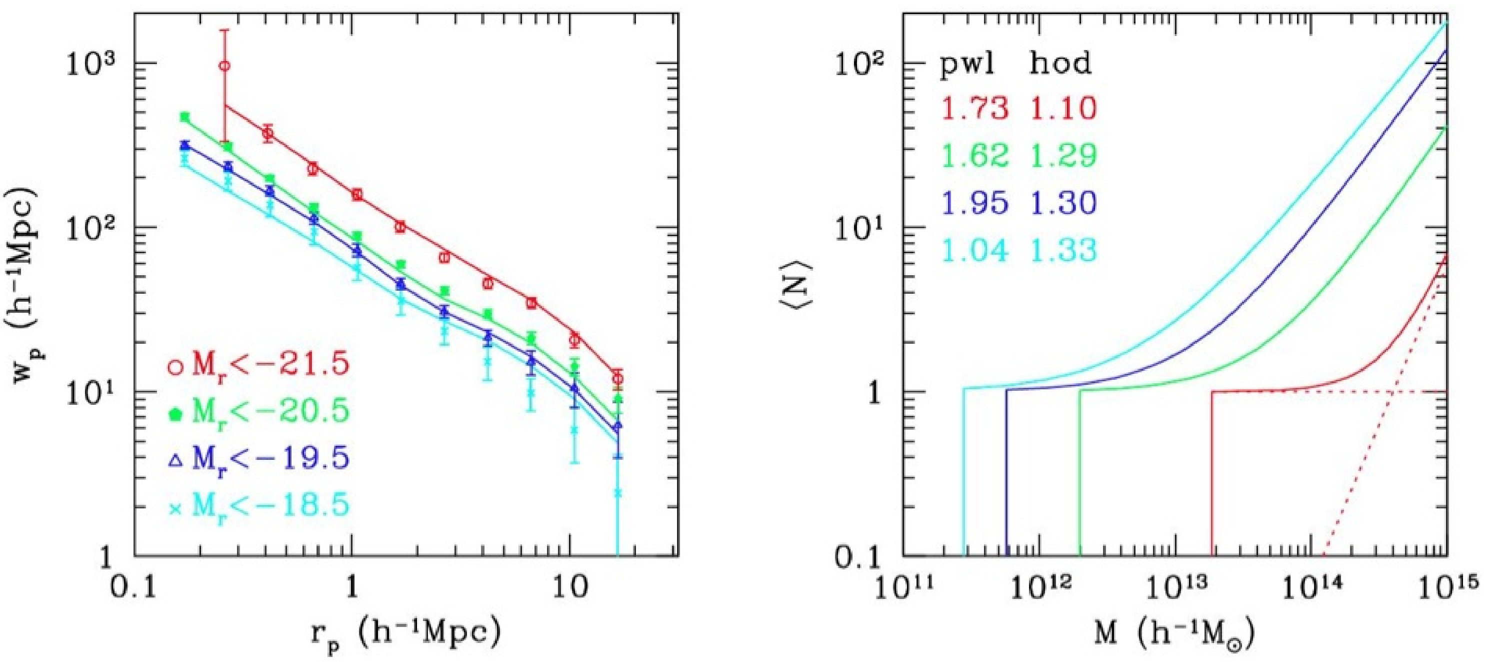
$$\langle\ \rangle_{2h,cc} = \int dm\, n(m)\, f_{cen}(m) \int dm'\, f_{cen}(m') \left[1+\xi_{hh}(r,m,m')\right]$$

$$\langle\ \rangle_{2h,cs} = \int dm\, n(m)\, f_{cen}(m) \int dm'\, f_{cen}(m')\, \langle N_s|m'\rangle \int d^3y\, u(y|m) \left[1+\xi_{hh}(y+r,m,m')\right]$$

$$\langle\ \rangle_{2h,ss} = \int dm\, n(m)\, f_{cen}(m) \int dm'\, f_{cen}(m')\, \langle N_s|m\rangle \langle N_s|m'\rangle \int d^3y\, u(y|m) \int d^3z\, u(z|m) \left[1+\xi_{hh}(y+z-r,m,m')\right]$$

For $\xi_{hh} \simeq b_m b_{m'} \xi_{lin}(r)$

$$P_{gg}^{2h} = P_{lin}(k) \left[ \int dm\, n(m) \frac{f_{cen}(m)}{n_g} b(m) \left(1+\langle N_s|m\rangle u(k|m)\right)\right]$$




Zehavi et al. 2005