

XVI Large-scale Clustering Statistics: Bispectrum, Skewness and PDF

Last class we discussed solutions beyond linear perturbation theory (PT), now we look at what can be used for. The simplest application is the calculation of the development of non-Gaussianity due to non-linearity in the equations of motion. Non-Gaussianity is characterized, to lowest order, by the appearance of a three-point function (which vanishes for a Gaussian field) or its Fourier transform, the bispectrum. An average of these quantities determines the skewness, which measures the asymmetry in evolution between underdense and overdense regions - let's consider this in more detail.

We write the PT solution for the density,

$$\delta(\vec{k}) = \delta_1(\vec{k}) + \sum_{n=2}^{\infty} \int [d\vec{k}_0] F_n(t_1, \vec{k}_n) \delta_1(\vec{k}_1) \dots \delta_1(\vec{k}_n) = \sum_{n=1}^{\infty} \delta_n(\vec{k})$$

where we include the growth factor inside each $\delta_1(\vec{k})$, and we work in the approximation that the kernels are independent of time, i.e. we assume $f^2 = -\dot{\sigma}_m$ at all times. Here $[d\vec{k}_0] \equiv \delta_D(\vec{k} - \sum_{i=1}^n \vec{k}_i)$ and a similar expression holds for the velocity divergence replacing F_n by G_n .

Now consider schematically the calculation of the third moment of the density

$$\begin{aligned} \langle \delta^3 \rangle &= \langle (\delta_1 + \delta_2 + \dots)^3 \rangle \\ &= \underbrace{\langle \delta_1^3 \rangle}_{\mathcal{O}(\sigma^3)} + 3 \underbrace{\langle \delta_1^2 \delta_2 \rangle}_{\mathcal{O}(\sigma^4)} + 3 \underbrace{(\langle \delta_1 \delta_2^2 \rangle + \langle \delta_1^2 \delta_3 \rangle)}_{\mathcal{O}(\sigma^5)} + \underbrace{\langle \delta_2^3 \rangle + 6 \langle \delta_1 \delta_2 \delta_3 \rangle + 3 \langle \delta_1^2 \delta_4 \rangle}_{\mathcal{O}(\sigma^6)} + \dots \end{aligned}$$

where we have organized different terms by the number of powers in the linear density field. Now, if we assume Gaussian initial conditions, the expectation value of odd numbers of δ_1 's is zero, then

$$\langle \delta^3 \rangle_{GIC} = 3 \langle \delta_1^2 \delta_2 \rangle + [\langle \delta_2^3 \rangle + 6 \langle \delta_1 \delta_2 \delta_3 \rangle + 3 \langle \delta_1^2 \delta_4 \rangle] + \dots$$

The first term scales as σ^4 where σ^2 is the linear variance of the density field, the second term scales as σ^6 , so it will be much smaller than the first at large scales where $\sigma^2 \ll 1$.

then, in the large scale limit we have

$$\langle \delta^3 \rangle = 3 \langle \delta_1^2 \delta_2 \rangle$$

Now, back to the details. We can calculate the three-point function in Fourier space or "bispectrum", by

$$\langle \delta(k_1) \delta(k_2) \delta(k_3) \rangle = B_{123} \delta_D(k_1 + k_2 + k_3)$$

$$\Rightarrow \delta_D(k_{123}) B_{123} = 3 \langle \delta_1(k_1) \delta_1(k_2) \delta_2(k_3) \rangle + 2 \text{ cyc.}$$

↑
the 2 other cyclic permutations

let's calculate one term:

$$\langle \delta_1(k_1) \delta_1(k_2) \delta_2(k_3) \rangle = \int d^3q_1 d^3q_2 \delta_D(k_3 - q_1 - q_2) F_2(q_1, q_2) \underbrace{\langle \delta_1(k_1) \delta_1(k_2) \delta_1(q_1) \delta_1(q_2) \rangle}_{\substack{\delta_1's \text{ are Gaussian, so} \\ \text{this is:}}}$$

Now, the first term does not contribute because it sets $q_1 + q_2 = 0 \Rightarrow F_2 = 0$ (remember this is because $\langle \delta_2(k) \rangle = 0$). Then we have

$$\begin{aligned} & P(k_1) \delta_D(k_2) P(q_1) \delta_D(q_2) \\ & + P(k_1) P(k_2) \delta_D(k_1 + q_1) \delta_D(k_2 + q_2) \\ & + P(k_1) P(k_2) \delta_D(k_1 + q_2) \delta_D(k_2 + q_1) \end{aligned}$$

$$= \int \delta_D(k_3 - q_1 - q_2) 2F_2(q_1, q_2) P(k_1) P(k_2) d^3q_1 d^3q_2 \delta_D(k_1 + q_1) \delta_D(k_2 + q_2)$$

where we introduce a factor of 2 since last two terms give equal contribution. Then we have

$$= \delta_D(k_1 + k_2 + k_3) 2F_2(k_1, k_2) P(k_1) P(k_2)$$

Then, we have for the bispectrum,

$$\begin{aligned} B_{123} &= 2F_2(k_1, k_2) P(k_1) P(k_2) + 2F_2(k_1, k_3) P(k_1) P(k_3) + 2F_2(k_2, k_3) P(k_2) P(k_3) \\ &= 2F_2(k_1, k_2) P(k_1) P(k_2) + \text{cyc.} \end{aligned}$$

Note here that all time dependence in the bispectrum is through the power spectrum. Also note that the quadratic nonlinearities here induce a scaling

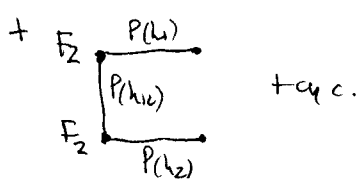
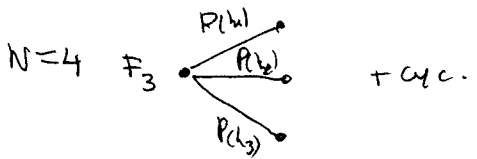
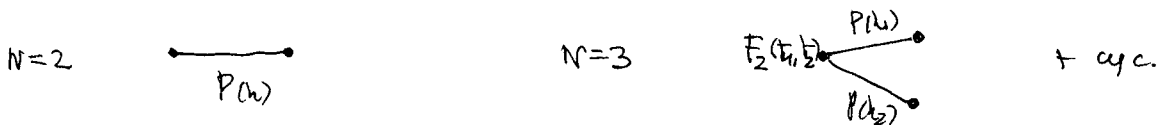
$$B \sim p^2$$

This in fact can be generalized to all higher-order correlation functions. For the N -point spectrum T_N we have

$$T_N \sim p^{N-1}$$

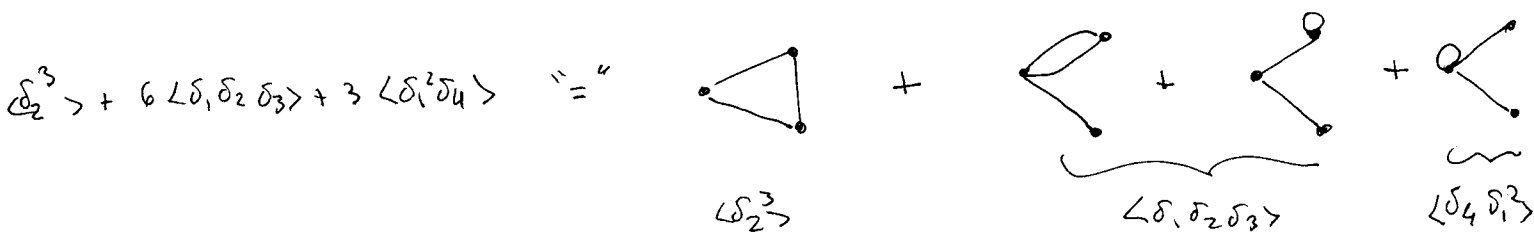
which corresponds to terms like $\langle \delta_{N-1} \underbrace{\delta_1 \dots \delta_1}_{N-1} \rangle_c$ in $\langle \delta^N \rangle$

The "c" here means we take connected part, i.e. get rid of Gaussian contribution to even-order correlation functions. Such a structure of scalings is related to a tree with N points, which requires $N-1$ links to make it connected, e.g.



etc.... Note that we defined vertices of different types, a vertex with n lines going out of it gets a factor of F_n (with $F_1 \equiv 1$). And each link carries a Power spectrum of the particular wavevector carried by that line. At each vertex "momentum conservation" is imposed, i.e. $k_1 + \dots + k_n = 0$.

Note that the next-to leading terms of the same order organized themselves into 1-loop diagrams in this language,



and similarly for the power spectrum,

$$\langle \delta^2 \rangle = \underbrace{\langle \delta_1^2 \rangle}_{O(\delta^2)} + \underbrace{2\langle \delta_1 \delta_2 \rangle}_{O(\delta^3)} + \underbrace{[\langle \delta_2^2 \rangle + 2\langle \delta_1 \delta_3 \rangle]}_{O(\delta^4)} + \dots \quad (4)$$

For Gaussian initial conditions the second term vanishes and one has,

$$P(\omega) = \underbrace{\text{---}}_{\langle \delta_1^2 \rangle} + \left[\underbrace{\text{---}}_{\langle \delta_2^2 \rangle} + \underbrace{\text{---}}_{\langle \delta_1 \delta_3 \rangle} \right] + \dots$$

The first being a tree diagram, the second contributions 1-loop diagrams - they correspond to the non-linear corrections to the power spectrum - We will discuss this next class -

OK, let's go back to the bispectrum and interpret physically what we get - Since $B \sim P^2$ it is convenient to define the redshifted bispectrum

$$Q_{123} \equiv \frac{B_{123}}{P_1 P_2 + P_2 P_3 + P_3 P_1}$$

i.e., we divide by the symmetrized combinations of P^2 - This definition removes the main dependencies on B_{123} , in fact:

- i) Q_{123} is indep of time (since linear growth factor cancels)
 - ii) " " " " cosmological parameters (" ")
 - iii) " " " " overall scale, since F_2 depends on ratios of k 's -
- for a power-law power spectrum, $P(k) \sim k^{-n}$ -

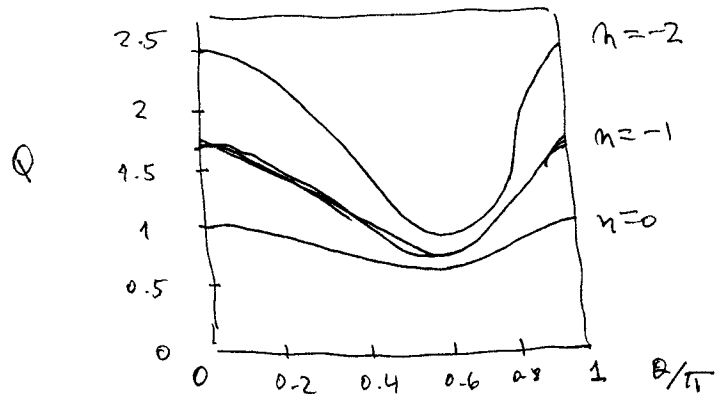
The first two properties remain true for a CDM spectrum, the last one does not because the spectral index depends on scale in CDM - So the bottom line is that Q_{123} only depends on the shape of the triangle and the spectral index at the range of scales involved in the k 's that correspond to the triangle sides -

One case that is particularly simple is that of an equilateral triangle, $k_1 \cdot k_2 = -1/2$ and $|k_i| = k$, thus

$$Q_{eq} = \frac{4}{7}$$

This is independent of the power spectrum.

For other triangles the dependence is more complicated, but to give an example consider $k_2 = 2k_1$ and plot things as a function of θ where $\hat{k}_1 \cdot \hat{k}_2 = \cos\theta$,



Q is maximal for colinear triangles and minimum for triangles close to equilateral. This reflects the dependence on the tidal field and on the transformation from a given mass element to Eulerian space. In summary it reflects the shapes of structures generated by gravitational instability, in fact in order to talk about shapes you need at least 3 points, so in this case the bispectrum is the lowest order statistic sensitive to shapes of structures generated by gravity. Note that for a more negative spectral index the bispectrum is more anisotropic, this is what we expect by just looking at the distribution of matter.

An averaged version of the bispectrum is to compute the skewness. Let's first calculate the skewness at a given point in space. We have

$$S_3 \equiv \frac{\langle \delta^3 \rangle}{\langle \delta^2 \rangle^2}$$

i.e. the same definition as for Q_{123} . Similarly, for higher-order moments one defines

$$S_N \equiv \frac{\langle \delta^N \rangle_c}{\langle \delta^2 \rangle^{N-1}}$$

again, to scale out the dependence on time, scale and cosmological parameters in the large-scale limit.

$$\langle \delta^3(\bar{x}) \rangle = 3 \langle \delta_2(\bar{x}) \delta_1(\bar{x})^2 \rangle$$

$$= 3 \left\langle \int e^{i \vec{k}_3 \cdot \bar{x}} d^3 k_3 \delta_2(\vec{k}_3) \int e^{i \vec{k}_1 \cdot \bar{x}} \delta_1(\vec{k}_1) d^3 k_1 \int e^{i \vec{k}_2 \cdot \bar{x}} \delta_1(\vec{k}_2) d^3 k_2 \right\rangle$$

$$= \int e^{i(\vec{k}_3 + \vec{k}_1 + \vec{k}_2) \cdot \bar{x}} B(\vec{k}_1, \vec{k}_2, \vec{k}_3) \delta_D(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) d^3 k_1 d^3 k_2 d^3 k_3$$

$$= \int B(\vec{k}_1, \vec{k}_2) d^3 k_1 d^3 k_2 = 6 \int F_2(\vec{k}_1, \vec{k}_2) P(k_1) P(k_2) d^3 k_1 d^3 k_2$$

Now, we write as last class

$$F_2(k_1, k_2) = \frac{\nu_2}{2} + k_1 \cdot k_2 \left(\frac{1}{k_2} + \frac{1}{k_1} \right) + \frac{2}{7} \left[(k_1 \cdot k_2)^2 \frac{1}{3} \right]$$

$$\nu_2 = \frac{34}{21}$$

$$\Rightarrow \langle \delta^3(\bar{x}) \rangle = 3 \nu_2 \left[\int P(k) d^3 k \right]^2 = 3 \nu_2 * \langle \delta^2(\bar{x}) \rangle^2$$

$$\Rightarrow S_3 = 3 \nu_2 = \frac{34}{7}$$

We see here that, i) only spherical dynamics enters

ii) does not depend on power spectrum or cosmological parameters -

iii) by translation invariance does not depend on \bar{x} .

Now, in practice one does not observe density field at "a point", but rather smoothed over some scale R . The most typical filter one smooths with is a top-hat filter, which in real space corresponds to counting things inside a sphere of radius R :

$$W_{TH}(\bar{x}) = \frac{\Theta(|\bar{x}| - R)}{\frac{4\pi}{3} R^3} \quad \text{where} \quad \Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases}$$

We can then write

$$\delta_R(\bar{x}) = \int \delta(\vec{y}) W_{TH}(\bar{x} - \vec{y}) d^3 y$$

or, in Fourier space,

$$\delta_R(\vec{k}) = \delta(\vec{k}) W_{TH}(kR)$$

where the Fourier transform is given by

$$W_{TH}(x) = \frac{3}{x^3} (\sin x - x \cos x) = 3 \sqrt{\frac{\pi}{2}} \frac{J_{3/2}(x)}{x^{3/2}}$$

$$\Rightarrow \langle \delta_R^2(\vec{x}) \rangle = \int P(k) W_{TH}^2(kR) d^3k = \sigma_R^2$$

this is used to specify the normalization of the power spectrum, e.g. at 8 Mpc/h it is called σ_8 ,

$$\sigma_8^2 = \int P(k) W_{TH}^2(k \cdot 8 \text{ Mpc}/h) d^3k$$

For the third moment we have,

$$\begin{aligned} \langle \delta_R^3(\vec{x}) \rangle &= \int B(k_1+k_2+k_3) \delta_D(k_{123}) W(k_1R) W(k_2R) W(k_3R) d^3k_1 d^3k_2 d^3k_3 \\ &= 6 \int F_2(k_1, k_2) P(k_1) P(k_2) W(k_1R) W(k_2R) W(k_1+k_2)R) d^3k_1 d^3k_2 \end{aligned}$$

Now we see that the integration over angles is a bit more complicated because of the angular dependence inside $W(k_1+k_2)R$ - But fortunately, this can be taken care of by the summation theorem for Bessel functions which imply:

$$\int \frac{d\Omega_{12}}{4\pi} W(k_1+k_2) [1 - (\hat{k}_1 \cdot \hat{k}_2)^2] = \frac{2}{3} W(k_1) W(k_2)$$

Note this is the same as doing $W \rightarrow W(k_1) W(k_2)$ and doing angle average.

Also,

$$\int \frac{d\Omega_{12}}{4\pi} W(k_1+k_2) \left(1 + \hat{k}_1 \cdot \hat{k}_2 \frac{k_2}{k_1} \right) = W(k_1) \left[W(k_2) + \frac{1}{3} k_2 W'(k_2) \right]$$

Then we have

$$\int F_2(k_1, k_2) W(k_1+k_2) W(k_1R) W(k_2R) \frac{d\Omega_{12}}{4\pi}$$

$$= \underbrace{-\frac{2}{7} \times \frac{2}{3} W_1 W_2 + W_1 W_2}_{\frac{17}{21} W_1 W_2} + \frac{1}{6} (W_1 k_2 R W_2' + W_2 k_1 R W_1')$$

$$\frac{17}{21} W_1 W_2 = \frac{\sqrt{2}}{2} W_1 W_2$$

contribution from "kinetic" L to E term

using that

$$\frac{d\sigma^2}{d\ln R} = 2 \int P(k) W(kR) W'(kR) kR d^3k$$

We have:

$$\begin{aligned} \langle \delta_R^3 \rangle &= 3V_2 \langle \delta_R^2 \rangle^2 + \frac{1}{3} \langle \delta_R^2 \rangle^2 \frac{d \ln \sigma^2}{d \ln R} \\ &= 3V_2 \langle \delta_R^2 \rangle^2 + \langle \delta_R^2 \rangle^2 \frac{d \ln \sigma^2}{d \ln R} \end{aligned}$$

$$\Rightarrow S_3 = \frac{34}{7} + \frac{d \ln \sigma^2(n)}{d \ln R}$$

We see that smoothing has introduced a correction that depends on spectral index at the smoothing scale.

For a power-law spectrum $P(k) \sim k^{-n}$

$$\sigma^2(k) \sim R^{-(n+3)} \quad \Rightarrow \quad \frac{d \ln \sigma^2}{d \ln R} = -(n+3)$$

$$\Rightarrow S_3 = \frac{34}{7} - (n+3)$$

By doing a similar calculation for all higher-order cumulants (that is, knowing γ_n from spherical collapse and incorporating smoothing through $L \rightarrow E$ mapping) one can sum up the cumulant generating function and compute the PDF. Show results.