Biasing

[astro-ph/9712192
astroph/0206508
astroph/0112551]

Introduction

We know from observations that different objects (galaxies of different type, clusters) cluster with different amplitudes. As we shall see, at large enough scales one expects all of them to trace the shape of the two-point function or powerspectrum given by that of the dark matter, but the relative amplitudes could differ. The ratio of the ξ's or $P(k)$'s to those of the dark matter is known as the bias factor, e.g.

$$b^2(r) \equiv \frac{\xi_g(r)}{\xi(r)} \qquad b^2(k) \equiv \frac{P_g(k)}{P(k)}$$

where ξ_g and P_g correspond to galaxies, and ξ & P to dark matter. Note that the bias factor is a priori a scale-dependent quantity, only at large scales is expected to be a constant.

Since the definition of b involves the dark matter clustering, how do we know that we are getting $b \neq 1$ due to a wrong dark matter spectrum? Well, first we know different galaxies cluster differently, so not all of them can have $b=1$. But what about if we include "all" galaxies, wouldn't the "general" bias go to 1 if we include a representative ~~sample~~ of the galaxies which by gravity should sample dark matter? Here one can do the following experiment: take a sample of galaxies thought to be representative, then change the linear power spectrum shape until the non-linear spectrum agrees with that of the galaxies, by definition having $b=1$ according to the definition above. Doesn't this make bias go away? Not, necessarily! The problem is that we also measure higher-order correlation functions, if one calculates say the S_p parameters $(p=3,4,\ldots)$ for the dark matter with this new linear power spectrum, one must match the observed ones for galaxies. In the case where this has been ied (the APM survey), this experiment fails. That is, we know at small scales there must be bias.

Note the assumption here is that one is starting with Gaussian initial conditions, otherwise one could alter the S_p's by primordial non-Gaussianity. However, there is no evidence of primordial non-Gaussianity, so the procedure above seems a solid one.

[Show figures]

There are general results one can obtain for bias at large scales with some minimal assumptions, we first discuss these "local bias" arguments. When applied to higher-order statistics they provide a way of <u>measuring</u> bias without directly determining the dark matter power spectrum. Then we apply the same idea to dark matter halos, to look at their bias as a function of halo mass. This is important for galaxy bias as modern approach to galaxy bias is to go through the halos: one populates halos with galaxies and thus their bias is related to that of the halos and their "occupation numbers" as a function of halo mass. We'll discuss this approach next class.

Local Bias

The assumption one ~~that~~ ~~can~~ make is that at scales R large enough compared to those where non-gravitational physics takes place in forming galaxies, the <u>smoothed</u> galaxy density at some point $\bar{x}$ is a local function of the smoothed density field at the same point, i.e

$$\tilde{\delta}_g(\bar{x}) = \mathcal{F}(\tilde{\delta}(\bar{x})) \qquad \text{where} \qquad \tilde{A}(\bar{x}) = \int d^3y \, A(\bar{x}-\bar{y}) \, \underset{\substack{\uparrow \\ \boxed{\text{smoothing filter}}}}{W_R(\bar{y})}$$

Since for large enough R, $\tilde{\delta} \ll 1$, one can expand in Taylor series,

$$\tilde{\delta}_g = \sum_{k=0}^{\infty} \frac{b_k}{k!} \tilde{\delta}^k$$

where the linear term b_1 corresponds to the linear bias factor, and b_0 is chosen so that $\langle \tilde{\delta}_g \rangle = 0$. How does this change the statistical properties of clustering at large scales? By ~~adding~~ keeping only leading order contributions in $\tilde{\delta}$, one finds for moments of the galaxy density field,

$$\begin{cases} \sigma_g^2 = b_1^2 \sigma^2 \\ S_{g,3} = \dfrac{S_3 + 3c_2}{b_1} \\ S_{g,4} = \dfrac{S_4 + 12 c_2 S_3 + 4 c_3 + 12 c_2^2}{b_1^2} \\ \vdots \end{cases}$$

where $c_k \equiv b_k / b_1$. It is clear that for higher-order correlations linear bias is not enough information at large scales, since a non-linear bias (e.g. $b_2 \neq 0$) can generate non-Gaussianity. Since Fourier transforms are effectively a smoothing operation, one can derive similar results in Fourier space, in particular

$$\begin{cases} P_g(k) = b_1^2 P(k) \\ Q_g^{123} = \dfrac{1}{b_1} Q^{123} + \dfrac{b_2}{b_1^2} \\ \dots \end{cases}$$

the first is the analog of the variance, the second for the bispectrum when integrated over k's to give the skewness it reduces to $S_{g,3} = \frac{1}{b_1} S_3 + \frac{3 b_2}{b_1^2}$ as above. The expression for the bispectrum Q_g shows how one can actually measure bias parameters:

i) measure $P_g(k)$, since spectral index of P_g is the same as $P(k)$, one can compute Q^{123} (which, remember, is just a function of n and triangle shape, not on cosmology)

ii) Compare Q_g to Q as a function of triangle shape, then get $1/b_1$ and b_2/b_1^2

By measuring higher-order correlation functions one can constrain $b_3, \ldots$ etc as well. Note that measuring skewness does not give enough information since one measures S_3 at a given scale, this is a number, not a function, so one only constrains a linear combination of $1/b_1$ and b_2/b_1^2. In principle one could use S_3 @ different scales but getting $1/b_1$ and b_2/b_1^2 is nearly degenerate.

The argument above on the fact that at large scales ξ_g and ξ have the same shape relies on the fact that $\tilde\delta^2 \ll 1$ at the smoothing scale R where local bias operates. This can be generalized to $\tilde\delta^2 > 1$ as long as the assumption that the following is true:

$$\langle [\tilde\delta(\vec x_1)]^p \, [\tilde\delta(\vec x_2)]^q \rangle_c = C_{pq} \, (\sigma^2)^{p+q-2} \underbrace{\langle \tilde\delta(\vec x_1)\, \tilde\delta(\vec x_2)\rangle}_{\xi_{12}} + \mathcal{O}\left(\langle \tilde\delta(\vec x_1)\tilde\delta(\vec x_2)\rangle^2\right)$$

where C_{pq} are numbers independent of scale. This holds in the weakly non-linear regime, and simulations show to be true to a good approximation even when $\sigma^2 \gtrsim 1$. Then one can write

$$\xi_g(|\vec x_1 - \vec x_2|) = \sum_{p,q} \frac{b_p b_q}{p!\, q!} \left[C_{pq} (\sigma^2)^{p+q-2} \xi_{12} + \langle \tilde\delta^p(\vec x_1)\, \tilde\delta^q(\vec x_2) \rangle_{\text{unconnected}} \right] +$$

$$+ \mathcal{O}(\xi^2)$$

Since the unconnected piece can be written in terms of connected, one can rewrite

$$\xi_g(|\vec x_1 - \vec x_2|) = \left[\sum_{p,q} K_{pq} \frac{b_p b_q}{p!\, q!} (\sigma^2)^{p+q-2} \right] \xi_{12} + \mathcal{O}(\xi_{12}^2)$$

Thus at large scales $\xi_g \propto \xi_{12}$ even if $\sigma^2 > 1$.

etc that if $\sigma^2 \ll 1$, only $p=q=1$ contributes and one recovers $\xi_g = b_1^2\, \xi_{12}$ - Also, if one is probing scales $|\bar{x}_1-\bar{x}_2|$ such that $\xi \sim 1$, other terms $\mathcal{O}(\xi_{12}^2)$ and higher contribute and thus bias becomes scale dependent, which is good because this is what appears to be necessary to explain observations.

Note that since one cannot say much about ξ_g when $\xi \gtrsim 1$, the effect on the power spectrum can be slightly different than in ξ_g. For example, one can write in general

$$\xi_g = b_1^2\, \xi_{12} + F(r) \qquad r = |\bar{x}_1-\bar{x}_2|$$

where if $\xi_{12}(r_0) \equiv 1$, $F(r) = 0$ for $r > r_0$ (where the argument above using C_{pq}'s holds) and $F(r) \neq 0$ for $r < r_0$, i.e. $F(r)$ gives the contributions from scale-dependent bias at small scales. Then we have for the power spectrum,

$$P_g(k) = b_1^2\, P(k) + \int F(r)\, e^{-i\bar{k}\cdot\bar{r}}\, d^3r$$

$$P_g(k) = b_1^2\, P(k) + \int_0^{r_0} F(r)\, j_0(kr)\, d^3r$$

Now if we consider very large scales, where $kr_0 \ll 1$, then we can take

$j_0(kr) \simeq 1 \quad \Rightarrow \qquad P_g(k) = b_1^2 P(k) + c$

$$c \equiv \int_0^{r_0} F(r)\, d^3r$$

So, one gets a constant contribution to the power.

Halo bias

Now we use the ~~[illegible]~~ results of the previous class to derive the bias of dark matter halos as a function of mass. We saw that the fraction of mass at "time" δ_1 in halos of mass m_1 ($\sigma^2(m_1) = S_1$) that later ("δ_2") is in a halo of mass m_2 is proportional to ~~[illegible]~~ the conditional quantity,

$$f(S_1, \delta_1 | S_2, \delta_2) = \frac{\delta_1 - \delta_2}{(S_1 - S_2)^{3/2}} \frac{e^{-\frac{(\delta_1-\delta_2)^2}{2(S_1-S_2)}}}{\sqrt{2\pi}} \qquad \delta_2 < \delta_1 \quad ; \quad S_2 > S_1$$

Now we see how we can use this to figure out the clustering of dark matter halos. What we need to know is how $\delta_h(m)$, the density contrast of halos of mass m is related to δ, the dark matter density contrast.

Suppose we have a region of mass M and volume V, with density contrast δ, $\frac{M}{V} = \bar{\rho}(1+\delta)$. This δ could be somewhat nonlinear we are not going to assume is vanishing small, but it is not huge, so that it is not yet a halo. We shall assume the region is large enough so that M is so large that $\sigma^2(M) \ll \sigma^2(M_*) = \delta_c$, and we can take $\sigma^2(M) \to 0$. Now, this region in the future (δ_c) will become a halo. Then we can use the result above for the fraction of mass in halos of mass m (formerly m_1) that collapse at time δ_c (formerly δ_1), that are in a region of mass M (m_2) and density contrast δ:

$$f(m, \delta_c / M, \delta) \underset{\sigma^2(M)\to 0}{\simeq} \frac{\delta_c - \delta_{lin}(\delta)}{\sigma^3_{(m)}} \frac{e^{-\frac{[\delta_c - \delta_{lin}(\delta)]^2}{2\sigma^2(m)}}}{\sqrt{2\pi}}$$

remember this is for interval in $d\sigma^2$

the average number of such halos between m and $m+dm$ is:

$$N(m / M, \delta) = \frac{M}{m} f(m, \delta_c / M, \delta) \left| \frac{d\sigma^2}{dm} \right|$$

and their density contrast is thus:

$$1 + \delta_h = \frac{N(m / M, \delta)}{n(m)\, V}$$

..te that in the large scale limit indeed we have $\delta_h \to 0$,

since:

$$1+\delta_h = \frac{\frac{M}{m} f \left|\frac{d\sigma^2}{dm}\right|}{n(m)\, M} \bar{\rho}(1+\delta) \xrightarrow[\delta \to 0]{} \bar{\rho} \frac{\frac{\cancel{M}}{m} \frac{\nu P(\nu)}{\sigma^2} \left|\frac{d\sigma^2}{dm}\right|}{n(m)\, \cancel{M}}$$

$$\uparrow$$

$$\frac{M}{V} = \bar{\rho}(1+\delta)$$

$$= \frac{\frac{\bar{\rho}}{m} \nu P(\nu) \left|\frac{d \ln \sigma^2}{dm}\right|}{n(m)} = 1$$

$\Rightarrow \delta_h \to 0$ ✓

In other words as $\delta \to 0$ one is considering a large enough region that corresponds to $M = \infty$ and conditional mass function goes to unconditional mass function.
Now consider the one where δ is small enough that we can expand in Taylor series, then from the exact relationship

$$1+\delta_h = \frac{\frac{\bar{\rho}}{m} f(m, \delta_c | M, \delta) \left|\frac{d\sigma^2}{dm}\right|}{n(m)} (1+\delta)$$

We can expand for small δ inside f:

$$\frac{\frac{\bar{\rho}}{m} f(m, \delta_c | M, \delta) \left|\frac{d\sigma^2}{dm}\right|}{n(m)} = 1 + \delta_{lin}(\delta) \frac{\nu^2-1}{\delta_c} + \frac{\nu^2(\nu^2-3)}{2\delta_c^2} \delta_{lin}^2(\delta) + \dots$$

Then we use the spherical collapse model

$$\delta = \delta_{lin} + \frac{\nu_2}{2} \delta_{lin}^2 + \frac{\nu_3}{3!} \delta_{lin}^3 + \dots$$

which up to quadratic order can be inverted:

$$\delta_{lin}(\delta) \simeq \delta - \frac{\nu_2}{2} \delta^2 \dots \qquad \left(\nu_2 = \frac{34}{21}\right)$$

thus we have, up to δ^2 terms:

$$1+\delta_h \simeq (1+\delta)\left[1 + \left(\delta - \frac{\nu_2}{2}\delta^2\right) \frac{\nu^2-1}{\delta_c} + \frac{\nu^2(\nu^2-3)}{2\delta_c^2} \delta^2 + \dots\right]$$

$$\Rightarrow \quad \delta_h = b_1(m)\, \delta + \frac{b_2(m)}{2} \delta^2 + \dots$$

where:

$$b_1(m) \equiv 1 + \frac{\nu^2-1}{\delta_c}$$

$$b_2(m) \equiv 2\left(1-\frac{\nu_2}{2}\right)\frac{\nu^2-1}{\delta_c} + \frac{\nu^2(\nu^2-3)}{\delta_c^2}$$

these represent the linear and quadratic bias of halos, respectively. From b_1 one can see that for $M > M^*$ ($\nu > 1$) $b_1(M) > 1$ so these halos are biased, strongly so for $M \gg M^*$, whereas $M < M^*$ halos are anti biased. A similar behavior holds for $b_2(M)$: $b_2(M) < 0$ for $M < M^*$, $b_2(M) > 0$ for $M \gg M^*$. [Show figures]

Similarly, one could work out $b_3(m)$, etc. Note the following consistency relation

$$\frac{1}{\bar\rho}\int m\, n(m)\, dm\; b_i(m) = \begin{cases} 1 & i = 1 \\ 0 & i > 1 \end{cases}$$

e.g. for $i=1$ $\quad \frac{1}{\bar\rho}\int m\, n(m)\, dm\; b_1(m) = \int_{-\infty}^{\infty} P(\nu)\, d\nu \left(1+\frac{\nu^2-1}{\delta_c}\right) = 1$

$i=2$ $\quad \frac{1}{\bar\rho}\int m\, n(m)\, dm\; b_2(m) = \int_{-\infty}^{\infty} P(\nu)\, d\nu \left[2\left(1-\frac{\nu_2}{2}\right)\frac{\nu^2-1}{\delta_c} + \frac{\nu^4-3\nu^2}{\delta_c^2}\right] = 0$

both terms vanish, first one due to $\langle \nu^2 \rangle = 1$, second due to kurtosis of a Gaussian field being 3, $\langle \nu^4 \rangle = 3 = 3\langle \nu^2 \rangle$.