Cosmic Density Field

June 12, 2018

1 Large-Scale Structure

1.1 Two-Point Correlation Functions

[1] One of the most important statistics used to characterize the spatial distribution of galaxies is the two-point correlation function, defined as the excess number of galaxy pairs of a given separation, r, relative to that expected for a random distribution:

$$\xi(r) = \frac{DD(r)\Delta r}{RR(r)\Delta r} - 1 \ . \tag{1}$$

Here $DD(r)\Delta r$ is the number of galaxy pairs with separations in the range $r \pm \Delta r/2$, and $RR(r)\Delta r$ is the number that would be expected if galaxies were randomly distributed in space. Galaxies are said to be positively correlated on scale r if $\xi(r) > 0$, to be anticorrelated if $\xi(r) < 0$, and to be uncorrelated if $\xi(r) = 0$. Since it is relatively straightforward to measure, the two-point correlation function of galaxies has been estimated from various samples. In many cases, red-shifts are used as distances and the corresponding correlation

function is called the correlation function in redshift space. Because of peculiar velocities, this redshift-space correlation is different from that in real space. The latter can be estimated from the projected two-point correlation function, in which galaxy pairs are defined by their separations projected onto the plane perpendicular to the line-of-sight so that it is not affected by using redshift as distance. On scales smaller than about $10h^{-1}$ Mpc the real-space correlation function can well be described by a power law,[1]

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma} , \qquad (2)$$

with $\gamma \sim 1.8$ and with a correlation length $r_0 \approx 5h^{-1}$ Mpc. This shows that galaxies are strongly clustered on scales $\lesssim 5h^{-1}$ Mpc, and the clustering strength becomes weak on scales much larger than $\sim 10h^{-1}$ Mpc. The exact values of γ and r_0 are found to depend significantly on the properties of the galaxies. In particular the correlation length, r_0, defined by $\xi(r_0) = 1$, is found to depend on both galaxy luminosity and color in the sense that brighter and redder galaxies are more strongly clustered than their fainter and bluer counterparts.

One can apply exactly the same correlation function analysis to groups and clusters of galaxies. This shows that their two-point correlation functions has a logarithmic slope, γ, that is similar to that of galaxies, but a correlation length, r_0, which increases strongly with the richness of the systems in question, from about $5h^{-1}$ Mpc for poor groups to about $20h^{-1}$ Mpc for rich clusters.

Another way to describe the clustering strength of a certain population of objects is to calculate the variance of the number counts within randomly placed spheres of given

---

[1]because of the definition of the two-point correlation function, $\xi(r)$ has to become negative on large scales. A power law can only fit the data up to a finite scale.

radius r:

$$\sigma^2(r) \equiv \frac{1}{(\bar{n}V)^2} \sum_{i=1}^{M} (N_i - \bar{n}V)^2 \ , \tag{3}$$

where $\bar{n}$ is the mean number density of objects, $V = 4\pi r^3/3$, and $N_i(i = 1, \cdots, M)$ are the number counts of objects in M randomly placed spheres. For optically selected galaxies with a luminosity of the order of L^* one finds that $\sigma \sim 1$ on a scale of $r = 8h^{-1}$ Mpc and decreases to $\sigma \sim 0.1$ on a scale of $r = 30h^{-1}$ Mpc. This confirms that the galaxy distribution is strongly inhomogeneous on scales of $\lesssim 8h^{-1}$ Mpc, but starts to approach homogeneity on significantly larger scales.

Since galaxies, groups and clusters all contain large amounts of matter, we expect their spatial distribution to be related to the mass distribution in the Universe to some degree. However, the fact that different objects have different clustering strengths makes one wonder if any of them are actually fair tracers of the matter distribution. The spatial distribution of luminous objects, such as galaxies, groups and clusters, depends not only on the matter distribution in the Universe, but also on how they form in the matter density field. Therefore, without a detailed understanding of galaxy formation, it is unclear which, if any, population of galaxies accurately traces the matter distribution. It is therefore very important to have independent means to probe the matter density field.

2   Statistical Properties

[1]

3 Gravitational Collapse of the Cosmic Density Field

3.1 Hierarchical Clustering

[1] In the linear regime the density perturbations $\delta(\boldsymbol{x}, t)$ grow with time as $\delta(\boldsymbol{x}, t) \propto D(t)$, and so the variance $\sigma^2(r, t) \propto D^2(t)$. For a power-law spectrum $P(k) \propto k^n$, $\sigma^2(r) \propto r^{-(n+3)}$

$$\sigma^2(r, t) = \left[\frac{r}{r^*(t)} \right]^{-(n+3)} = \left[\frac{M}{M^*(t)} \right]^{-(n+3)/3}, \tag{4}$$

3.2 Results from Numerical Simulations

[1] The initial density field contains perturbations over a wide range of scales, and nonlinear evolution couples structures of different scale. Ro follow the evolution of the cosmic density field in detail, one must use N-body simulations. The first step is to generate an initial density field with precisely the desired statistical properties. One can start with an array of Fourier modes, each characterized by its wavevector $\boldsymbol{k}$, and assign each a random amplitude $|\delta_{\boldsymbol{k}}|$ and a random phase $\varphi_{\boldsymbol{k}}$ according to the distribution function. The linear overdensity field, $\delta(\boldsymbol{x})$, can then be obtained using fast Fourier transforms (FFTs), and set up in a simulation box. Because both $|\delta_{\boldsymbol{k}}|$ and $\varphi_{\boldsymbol{k}}$ are random variables, the field $\delta(\boldsymbol{x})$ from a particular set of $(|\delta_{\boldsymbol{k}}|, \varphi_{\boldsymbol{k}})$ represents only one specific realization of the model in consideration. The perturbation fields will differ from one realization to another, but should all be equivalent in a statistical sense. Because the density field must be sampled in a finite simulation box, there are also artificial effects due to the finite number of Fourier modes that is sampled. The realization to realization variance caused by these effects of

finite volume is similar to the uncertainty in inferences from observational surveys due to their finite volume and is usually referred to as cosmic variance.

4 Large-Scale Mass Distribution

4.1 Correlation Functions

[1] The properties of the cosmic density field can be described by the moments of the distribution function $\mathscr{P}_x$;

two-point correlation function of the density perturbation field

$$\xi(x) = \langle \delta_1 \delta_2 \rangle \ , \tag{5}$$

where $x = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$ is the separation in comoving units. $\xi(x)$ depends only on the amplitude x of $\boldsymbol{x}$: the density perturbation field, $\delta(\boldsymbol{x})$, is a homogeneous and isotropic random field. Writing both δ_1 and δ_2 in terms of their Fourier transforms,

$$\xi(x) = \frac{1}{V_u} \sum_k P(k) e^{i\boldsymbol{k} \cdot \boldsymbol{x}} = \frac{1}{(2\pi)^3} \int P(k) e^{i\boldsymbol{k} \cdot \boldsymbol{x}} \mathrm{d}^3 \boldsymbol{k} \tag{6}$$

where

$$P(k) = V_u \langle |\delta_{\boldsymbol{k}}|^2 \rangle \tag{7}$$

is the power spectrum. The two-point correlation function and power spectrum form a Fourier transform pair. Carrying out the integration over the angle between $\boldsymbol{k}$ and $\boldsymbol{x}$,

$$\xi(x) = \frac{1}{2\pi^2} \int_0^\infty k^3 P(k) \frac{\sin kx}{kx} \frac{\mathrm{d}k}{k} = \int_0^\infty \Delta^2(k) \frac{\sin kx}{kx} \frac{\mathrm{d}k}{k} \ , \tag{8}$$

where

$$\Delta^2(k) \equiv \frac{k^3}{2\pi^2} P(k) \ . \tag{9}$$

5

The definition of the power spectrum gives

$$P(k) = 4\pi \int_0^\infty \xi(x) \frac{\sin kx}{kx} x^2 \mathrm{d}x \tag{10}$$

The volume average of $\xi(x)$ is

$$\bar{\xi}(x) = \frac{3}{x^3} \int_0^x \xi(x') x'^2 \mathrm{d}x' \;, \tag{11}$$

which is related to the power spectrum by

$$\bar{\xi}(x) = \int_0^\infty \Delta^2(k) \left[\frac{3(\sin kx - kx \cos kx)}{(kx)^3} \right] \frac{\mathrm{d}k}{k} \;. \tag{12}$$

The window function (the term in brackets) involved here dies off faster with increasing k than the term $(\sin kx / kx)$ in the expression of $\xi(x)$, so $\bar{\xi}(x)$ gives a cleaner measure of the power spectrum at $k \sim \frac{1}{x}$ than does $\xi(x)$.

$$J_3(x) = \int_0^x \xi(x') x'^2 \mathrm{d}x' = \frac{1}{3} x^3 \bar{\xi}(x) \;.$$

If $P(k) \approx k^n$ as $k \to 0$, then for $n > -3$ the main contribution to the integral of $\xi(x)$ is from $k \lesssim \frac{1}{x}$. Since the window function in $\xi(x)$ is about unity for $k \lesssim \frac{1}{x}$,

$$\bar{\xi}(x) \simeq \int_0^{1/x} \Delta^2(k) \frac{\mathrm{d}k}{k} \propto x^{-(n+2)} \;, \tag{13}$$

and $J_3(x) \propto x^{-n}$. For $n > 0$,

$$\int_0^\infty \xi(x') x'^2 \mathrm{d}x' = 0 \;.$$

Because $\xi(0) = \langle \delta^2(\boldsymbol{x}) \rangle > 0$, this constraint means that $\xi(x)$ must pass through zero at some large x.

The *l*-point correlation function is defined as

$$\xi^{(l)}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_l) \equiv \langle \delta_1 \delta_2 \cdots \delta_l \rangle \;. \tag{14}$$

The connected l-point function is obtained by subtracting from $\xi^{(l)}$ all the disconnected terms arising from lower-order correlations. The connected three- and four-point correlation functions are

$$
\begin{aligned}
\zeta(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3) & \equiv \langle \delta_1 \delta_2 \delta_3 \rangle - \langle \delta_1 \rangle \langle \delta_2 \delta_3 \rangle (\text{three terms}) - \langle \delta_1 \rangle \langle \delta_2 \rangle \langle \delta_3 \rangle \\
& = \langle \delta_1 \delta_2 \delta_3 \rangle \;,
\end{aligned}
$$

$$
\begin{aligned}
\eta(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \boldsymbol{x}_4) & \equiv \langle \delta_1 \delta_2 \delta_3 \delta_4 \rangle - \langle \delta_1 \rangle \langle \delta_2 \delta_3 \delta_4 \rangle (\text{four terms}) \\
& \quad - \langle \delta_1 \delta_2 \rangle \langle \delta_3 \delta_4 \rangle (\text{three terms}) - \langle \delta_1 \rangle \langle \delta_2 \rangle \langle \delta_3 \rangle \langle \delta_4 \rangle \\
& = \langle \delta_1 \delta_2 \delta_3 \delta_4 \rangle - \xi(x_{12}) \xi(x_{34}) (\text{three terms}) \;.
\end{aligned}
$$

For a Gaussian random field, $\zeta = \eta = 0$, and so are all connected higher-order correlation functions. The connected high-order correlation functions can be used to test whether a density perturbation field is Gaussian.

4.2   Particle Sampling and Bias

[1] A continuous density field with particles is represented by dividing the space into infinitesimal cells, and sample the density field in such a way that the number of particles in a particular cell has a Poisson distribution with a mean proportional to the mean density of the cell. Suppose that the volume of each cell, ΔV, is chosen so small that the probability for a cell to contain more than one particle is zero. The occupancy of an arbitrary cell ($\mathcal{N}_i$) has the following properties:

$$
\mathcal{N}_i = \mathcal{N}_i^2 = \mathcal{N}_i^3 = \cdots \;. \tag{15}
$$

To sample a given density field $\rho(\boldsymbol{x})$ by a Poisson process, just need to specify $p^{(1)}(\boldsymbol{x})$, the probability for a cell located at $\boldsymbol{x}$ to contain one particle. This probability is equal to

7

the average of $\mathcal{N}$ at $\boldsymbol{x}$

$$p^{(1)}(\boldsymbol{x}) = \langle \mathcal{N}(\boldsymbol{x}) \rangle_{\mathrm{P}} = [1 + \delta(\boldsymbol{x})] \bar{n} \Delta V \; , \qquad (16)$$

where $\bar{n} = \bar{\rho}/m$ (m being the mass of a particle) is the mean number density of particles, and $\langle \cdots \rangle_{\mathrm{P}}$ is the average over the Poisson distribution.

5 Galaxy Clustering

[1] To study the spatial clustering of galaxies, one usually starts with a galaxy sample for which sky positions and redshifts are listed for all members. Such a sample typically provides a non-uniform sampling of the true galaxy distribution in redshift space (and thus a distorted image of their distribution in real space) throughout a finite volume defined by the survey boundaries. The observational criteria which define real samples typically induce strong selection effects. For example, only the intrinsically brightest galaxies may be included at large distances whereas relatively faint systems may be included in the foreground. Such effects must be carefully taken into account when using real samples to study galaxy clustering.

5.1 Correlation Analyses

[1] The two-point correlation function of galaxies can be estimated from a redshift survey as

$$\xi(r) = \frac{DD(r)}{RR(r)} - 1 \; , \qquad (17)$$

where $DD(r)\Delta r$ is the observed number of galaxy-galaxy pairs with separations in the range $r \pm \Delta r/2$, and $RR(r)\Delta r$ is the expected number of such pairs in a 'random' (i.e.

uniform, Poisson distributed) sample with the same number of objects filling the same volume and with the same selection function as the real sample. $\xi(r)$ is zero if galaxies are distributed as a uniform Poisson process; any deviation from zero indicates spatial clustering. This definition is motivated by the idea that galaxies can be treated as a Poisson sampling of a smooth underlying density field.

If the number density of galaxies at each spatial location is proportional to the underlying density field, then $\xi(r)$ as defined above will be the same as the two-point correlation function of the density field. Other estimators of $\xi(r)$ are possible. Three that are frequently used can be written symbolically as

$$\xi(r) = \frac{DD}{DR} - 1 \ , \ \xi(r) = \frac{DD \cdot RR}{DR^2} - 1 \ , \ \xi(r) = \frac{DD - 2DR. + RR}{RR} \ , \quad (18)$$

where $DR(r)$ is the number of cross-pairs between the real and random samples. For real applications, the last two estimators are usually preferred because they are less affected by sample boundaries. It is often advantageous to construct random samples that contain many more particles than there are galaxies in the real sample. In this case, RR has to be multiplied by $(N_g/N_r)^2$ and DR by N_g/N_r, where N_g and N_r are the number of galaxies in the real and the random samples, respectively.

Assume that the fluctuation in the number of independent pairs in a given bin of r has a Poisson distribution: $\delta N_{\text{pair}} = N_{\text{pair}}^{1/2}$. The error in ξ is

$$\delta \xi(r) = \frac{1 + \xi(r)}{\sqrt{N_{\text{pair}}(r)}} \ . \quad (19)$$

However, this method significantly underestimate the true uncertainty in the data. Another class of 'internal' error estimators is based on dividing the data sample into a set of N similar subsamples in space. The simplest one, often referred to as the subsample

method, is to use the variance among all the subsamples, under the assumption that each of them is an independent realization of the underlying distribution. The covariance matrix is

$$C(\xi_i, \xi_j) = \frac{1}{N} \sum_{k=1}^{N} (\xi_i^k - \bar{\xi}_i)(\xi_j^k - \bar{\xi}_j) \; , \tag{20}$$

where ξ_i^k denotes the measurement of ξ at separation r_i from the kth subsample, and $\bar{\xi}$ is the expectation value estimated independently from the N subsamples. The problem with this method is that in real applications the subsamples, each with a volume much smaller than the total sample, may not be mutually independent because of the existence of long-range modes in the density fluctuations. Two other commonly used estimators in this class are the bootstrap method and jackknife method. In the bootstrap method, one forms a set of $N_{\rm rs}$ resamplings of the original sample, each containing N galaxies (including duplicates) randomly picked from the original N galaxies with replacement (i.e. a galaxy is retained in the stack even if it has already been picked). Thus, although each resample consists of the same number of galaxies as the original sample, it will include some of the galaxies more than once, while others may not be included at all. The covariance matrix is estimated from

$$C(\xi_i, \xi_j) = \frac{1}{N_{\rm rs}} \sum_{k=1}^{N_{\rm rs}} (\xi_i^k - \bar{\xi}_i)(\xi_j^k - \bar{\xi}_j) \; , \tag{21}$$

where $\bar{\xi}$ is the mean obtained from the $N_{\rm rs}$ resamplings. In practice for large samples, one chooses $N_{\rm rs} \ll N$ but still large enough to provide a good measurement of the covariance matrix. In the jackknife method, one forms a set of N 'copies' of the original sample, each time leaving out one of the N galaxies. The covariance matrix is then estimated from

$$C(\xi_i, \xi_j) = \frac{N-1}{N} \sum_{k=1}^{N} (\xi_i^k - \bar{\xi}_i)(\xi_j^k - \bar{\xi}_j) \; , \tag{22}$$

where ξ_i^k is the measurement of ξ at separation r_i in the kth 'copy' and ξ is the mean of the N copies. For large samples the jackknife approach as outlined above is impractical, since excluding a single galaxy has almost no effect. An alternative is to divide the full sample into $N_{\rm rs}$ disjoint subsamples, each containing $N/N_{\rm rs}$ galaxies. One can then proceed as before, estimating a correlation function for $N_{\rm rs}$ jackknife copies of the original sample obtained by leaving out one subsample at a time.

To take care of redshift distortions, first estimate the two-dimensional function $\xi(r_\pi, r_p)$, and then obtain the real-space correlation function from the projected function $w(r_p)$. For optically selected galaxies with $L \sim L^*$, the observed correlation function at $r \lesssim 20h^{-1}$ Mpc can be approximated by a power law,

$$\xi(r) = \left(\frac{r}{r_0}\right)^{-\gamma} \text{, with } r_0 \approx 5h^{-1}\text{Mpc and } \gamma \approx 1.7 \tag{23}$$

This indicates that the galaxy distribution becomes highly nonlinear on scales $\lesssim 5h^{-1}$ Mpc. Bright, red galaxies are more strongly clustered than faint, blue ones.

5.2   Power Spectrum Analysis

[1] Although the power spectrum is just the Fourier transform of the two-point correlation function, it is more advantageous to work with the power spectrum than with the two-point correlation function when studying galaxy clustering on large scales. The reason for this is that on large scales, where the density field is still in the linear regime, different Fourier modes evolve independently while the amplitude of the two-point correlation function is affected by many different modes. Consequently, power-spectrum amplitudes on large scales are less affected by small-scale structure than correlation function estimates, making observational results much easier to interpret.

Given a galaxy sample, it is straightforward to measure the power spectrum. Suppose that the volume of the sample is V_s, which is described by a window function $W(\boldsymbol{x})$ such that $W(\boldsymbol{x})$ is equal to 1 if $x \in V_s$ and to 0 otherwise. Suppose that this volume is contained in a large box V_u on which the Universe is assumed to be periodic. We divide V_u into small cells of volume ΔV so that the galaxy occupation number in each cell, $\mathcal{N}_i$, is either 1 or 0. The observed galaxy density field can then be written as

$$n_o(\boldsymbol{x}) = \sum_j \mathcal{N}_j \delta^{\mathrm{D}}(\boldsymbol{x} - \boldsymbol{x}_j) W_j \;, \tag{24}$$

where $W_j = W(\boldsymbol{x}_j)$ and $\delta^{\mathrm{D}}(\boldsymbol{x})$ is the Dirac delta function. For a magnitude-limited sample, this density must be multiplied by the reciprocal of the selection function to correct for selection effects. The corrected density field is

$$n_c(\boldsymbol{x}) = A \sum_j \mathcal{N}_j S_j^{-1} \delta^{\mathrm{D}}(\boldsymbol{x} - \boldsymbol{x}_j) W_j \;, \tag{25}$$

where $S_j = S(\boldsymbol{x}_j)$, and the prefactor $A = \overline{n} V_s / \sum_j \mathcal{N}_j S_j^{-1} W_j$ is included so that $\int n_c(\boldsymbol{x}) \mathrm{d}^3 \boldsymbol{x} = \overline{n} V_s$, with $\overline{n}$ the mean number density of galaxies. Note that A is equal to 1 for the top-hat window considered here and is included in the equation to encompass other possible choices.

$$\tilde{n}_c(\boldsymbol{k}) = \frac{A}{V_u} \sum_j \mathcal{N}_j S_j^{-1} W_j \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{x}_j} \;. \tag{26}$$

The average of this is

$$\langle \tilde{n}_c(\boldsymbol{k}) \rangle = \overline{n} \tilde{W}(\boldsymbol{k}) \;, \quad \text{where} \quad \tilde{W}(\boldsymbol{k}) = \frac{A}{V_u} \int W(\boldsymbol{x}) \mathrm{e}^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \mathrm{d}^3 \boldsymbol{x} \;, \tag{27}$$

is the Fourier transform of the window function. $\langle \mathcal{N}_j \rangle / S_j = \overline{n} \Delta V$. Using $\mathcal{N}_i^2 = \mathcal{N}_i$ and

$$\langle \mathcal{N}_i \mathcal{N}_j \rangle = \overline{n}^2 S_i S_j [1 + \xi(\boldsymbol{x}_i - \boldsymbol{x}_j)],$$

$$\langle \tilde{n}_c(\boldsymbol{k}) \tilde{n}_c^*(\boldsymbol{k}) \rangle = \frac{A^2}{V_u^2} \sum_{i,j} \langle \mathcal{N}_i \mathcal{N}_j \rangle S_i^{-1} S_j^{-1} W_i W_j \mathrm{e}^{i\boldsymbol{k} \cdot (\boldsymbol{x} - \boldsymbol{x}_j)}$$

$$= \frac{A^2}{V_u^2} \sum_j \mathcal{N}_i (W_j/S_j)^2 + \overline{n}^2 |\tilde{W}(\boldsymbol{k})|^2 + \frac{\overline{n}^2}{V_u} \sum_{\boldsymbol{k}'} |\tilde{W}(\boldsymbol{k}')|^2 P(\boldsymbol{k} - \boldsymbol{k}') , \quad (28)$$

where $P(k) = \int \xi(\boldsymbol{x}) \mathrm{e}^{-i\boldsymbol{k} \cdot \boldsymbol{x}} \mathrm{d}^3 \boldsymbol{x}$ is the power spectrum of the galaxy density field. If the survey volume is large in comparison to the wavelengths in consideration, $\tilde{W}(\boldsymbol{k}')$ is sharply peaked at $\boldsymbol{k}' = 0$, and we can replace $P(\boldsymbol{k} - \boldsymbol{k}')$ by $P(\boldsymbol{k})$ in the above equation and pull it out of the summation.

$$P(\boldsymbol{k}) \approx \frac{V_u^2 \langle |\tilde{n}_c(\boldsymbol{k}) - \langle \tilde{n}_c(\boldsymbol{k}) \rangle |^2 \rangle - N_{\mathrm{eff}}}{\overline{n}^2 V_u \sum_{\boldsymbol{k}'} |\tilde{W}(\boldsymbol{k}')|^2} , \quad (29)$$

$$N_{\mathrm{eff}} \equiv \frac{(\overline{n} V_s)^2 \sum_j \mathcal{N}_j (W_j S_j^{-1})^2}{(\sum_j \mathcal{N}_j W_j S_j^{-1})^2} . \quad (30)$$

The summations over j in N_{eff} can be replaced by those over all the galaxies in the sample with $\mathcal{N}_j$ set to 1. In a volume-limited sample, $S_i = 1$ and so $N_{\mathrm{eff}} = \overline{n} V_s$. It shows that the true power spectrum is equal to the power spectrum of the corrected number density field, $n_c(\boldsymbol{x})$, with a subtraction of the shot noise, N_{eff}, and a deconvolution with the window function. The final power spectrum is usually binned in shells in k-space:

$$P(k) = \frac{1}{V_k} \int_{V_k} P(\boldsymbol{k}') \mathrm{d}^3 \boldsymbol{k}' , \quad (31)$$

where V_k is the volume of the shell in $\boldsymbol{k}$-space.

Since the derivation is independent of the assumption that $W(\boldsymbol{x})$ is a top-hat, it is valid for any form of $W(\boldsymbol{x})$. As for the correlation functions, assign different weights to different regions to reduce the variance in the estimate. This can be done by using a window function that is inhomogeneous in space.

5.3   Angular Correlation Function and Power Spectrum

[1] For a two-dimensional sample, estimate the angular correlation function of galaxies, $w(\vartheta)$, from the definition:

$$\langle \nu(\hat{\boldsymbol{r}}_1)\nu(\hat{\boldsymbol{r}}_2)\rangle \mathrm{d}\omega_1 \mathrm{d}\omega_2 = \bar{\nu}^2[1 + w(\vartheta)]\mathrm{d}\omega_1 \mathrm{d}\omega_2 \ , \tag{32}$$

where $\nu(\hat{\boldsymbol{r}})\mathrm{d}\omega$ is the number of galaxies within a solid angle $\mathrm{d}\omega$ in the direction $\hat{\boldsymbol{r}}$, and $\bar{\nu}$ is the mean surface density. Suppose that the volume density of galaxies (assumed to be in comoving units) at redshift z and in the direction $\hat{\boldsymbol{r}}$ is $n(\hat{\boldsymbol{r}}, z)$. The surface density, which is the projection of the three-dimensional distribution on the sky, can be written as

$$\nu(\hat{\boldsymbol{r}}) = \int_0^\infty n(\hat{\boldsymbol{r}}, z)S(z)\frac{\mathrm{d}^2 V}{\mathrm{d}z\mathrm{d}\omega}\mathrm{d}z \ , \tag{33}$$

where $S(z)$ is the selection function of the survey, and $\mathrm{d}^2 V$ is the comoving volume element corresponding to $\mathrm{d}z$ and $\mathrm{d}\omega$ at redshift z.

$$\bar{\nu} = \int_0^\infty \overline{n}(z)S(z)\frac{\mathrm{d}^2 V}{\mathrm{d}z\mathrm{d}\omega}\mathrm{d}z \ , \tag{34}$$

$$w(\vartheta) = \frac{\overline{n}^2}{\bar{\nu}^2}\int_0^\infty \overline{n}(z_1)S(z_1)\overline{n}(z_2)S(z_2)\xi(r_{12}, z)\frac{\mathrm{d}^2 V_1}{\mathrm{d}z_1\mathrm{d}\omega_1}\frac{\mathrm{d}^2 V_2}{\mathrm{d}z_2\mathrm{d}\omega_2}\mathrm{d}z_1\mathrm{d}z_2 \ , \tag{35}$$

where $\xi(r_{12}, z)$ is the spatial correlation function at redshift $z \equiv (z_1 + z_2)/2$, r_{12} is the separation between $(\hat{\boldsymbol{r}}_1, z_1)$ and $(\hat{\boldsymbol{r}}_2, z_2)$, and $\cos \vartheta = \hat{\boldsymbol{r}}_1 \cdot \hat{\boldsymbol{r}}_2$. This is a general relation between the angular and spatial correlation functions, taking into account possible evolutions in $\overline{n}$ and ξ with redshift, as well as cosmological effects. If the survey is not very

14

deep, so that the evolutionary and cosmological effects are all negligible,

$$\bar{\nu} = \int_0^\infty S(r) r^2 \mathrm{d}r \ , \tag{36}$$

$$w(\vartheta) = \frac{1}{\bar{\nu}^2} \int_0^\infty S(r_1) S(r_2) \xi(r_{12}) r_1^2 r_2^2 \mathrm{d}r_1 \mathrm{d}r_2 \ , \tag{37}$$

where $r_{12}^2 = r_1^2 + r_2^2 2 r_1 r_2 \cos \vartheta$. In the small-angle limit where ϑ is small and the mean distance of the pair, $x \equiv (r_1 + r_2)/2$, is much larger than the value of $y \equiv r_1 - r_2$, $r_{12}^2 \approx y^2 + x^2 \vartheta^2$ and

$$w(\vartheta) = \int_0^\infty x^4 S^2(x) \mathrm{d}x \int_{-\infty}^\infty \mathrm{d}y \xi \left[(y^2 + x^2 \vartheta^2)^{1/2} \right] \bigg/ \left[\int_0^\infty S(x) x^2 \mathrm{d}x \right]^2 \ . \tag{38}$$

If $\xi(r)$ is a power law,

$$\xi(r) = A/r^\gamma \ , \tag{39}$$

the angular correlation function is also a power law,

$$w(\vartheta) = B/\vartheta^{\gamma-1} \ , \tag{40}$$

with the amplitude B related to A by

$$B = A \sqrt{\pi} \frac{\Gamma[(\gamma-1)/2]}{\Gamma(\gamma/2)} \int_0^\infty x^{5-\gamma} S^2(x) \mathrm{d}x \bigg/ \left[\int_0^\infty S(x) x^2 \mathrm{d}x \right]^2 \ . \tag{41}$$

For given γ and A, the amplitude of the angular correlation function depends only on the shape of the selection function. For a magnitude-limited sample, the selection function can be written as $S(x) = f(x/d^*)$, where $d^* \approx 10^{0.2 m_{\mathrm{lim}}}$ (with m_{lim} the magnitude limit) is the characteristic depth of the sample, and f is a universal function (under the assumption that cosmological effects and K corrections are small). This form of $S(x)$ implies that for a given $\xi(r)$ the angular correlation function scales with the sample depth as

$$w(\vartheta, d^*) = \mathscr{W}(\vartheta d^*)/d^* \ , \tag{42}$$

15

where $\mathscr{W}$ is a scaling function.

The power spectrum can also be estimated for a two-dimensional sky survey of galaxies. Dividing the sky into small cells, $j = 1, 2, \cdots$, so that the galaxy occupation number $\mathscr{N}_j$ is either 1 or 0, we can write the observed surface density field as

$$\nu_o(\hat{\boldsymbol{r}}) = \sum_j \mathscr{N}_j \delta^{(2)}(\hat{\boldsymbol{r}} - \hat{\boldsymbol{r}}_j) W_j \ , \tag{43}$$

where $W_j \equiv W(\hat{\boldsymbol{r}}_j)$ and $W(\hat{\boldsymbol{r}})$ is a window function specifying the sky coverage of the survey. Expanding $\nu_o(\hat{\boldsymbol{r}})$ in spherical harmonics,

$$\nu_o(\hat{\boldsymbol{r}}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{\boldsymbol{r}}) \ , \tag{44}$$

where $Y_{\ell m}(\hat{\boldsymbol{r}})$ is the spherical harmonic function calculated at the direction $\hat{\boldsymbol{r}} = (\vartheta, \varphi)$,

$$a_{\ell m} = \sum_j \mathscr{N}_j W_j Y_{\ell m}^*(\hat{\boldsymbol{r}}_j) \ . \tag{45}$$

The angular power spectrum is

$$C_\ell \approx \frac{\langle |a_{\ell m} - \langle a_{\ell m} \rangle|^2 \rangle}{\bar{\nu}^2 \langle J_{\ell m} \rangle} - \frac{1}{\bar{\nu}} \ , \tag{46}$$

where

$$\langle a_{\ell m} \rangle = \int W(\vartheta, \varphi) Y_{\ell m}(\vartheta, \varphi) \mathrm{d}\omega \ ,$$

$$\langle J_{\ell m} \rangle = \int W(\vartheta, \varphi) |Y_{\ell m}(\vartheta, \varphi)|^2 \mathrm{d}\omega \ ,$$

and the approximation assumes that the size of the window is much larger than the angular scale corresponding to mode ℓ. The angular power spectrum C_ℓ is related to the angular correlation function by

$$w(\vartheta) = \frac{1}{4\pi} \sum_\ell (2\ell + 1) C_\ell \mathscr{P}_\ell(\cos \vartheta) \ , \tag{47}$$

where P_ℓ is the Legendre function. Using the relation between the angular and spatial correlation functions, one can relate C_ℓ to the spatial power spectrum $P(k)$:

$$C_\ell = \frac{2}{\pi(2\ell+1)} \int \mathrm{d}k k^2 P(k) \left[\int S(r) j_\ell(kr) r^2 \mathrm{d}r \right]^2 \Big/ \left[\int S(r) r^2 \mathrm{d}r \right]^2 \tag{48}$$

where j_ℓ is the spherical Bessel function.

[2] The simplest statistic is the two-point function: in real space it is $w(\theta)$ the angular correlation function. In Fourier space, the relevant function is the Fourier transform of w, $P_2(l)$, the two-dimensional power spectrum.

A given galaxy is at comoving distance $\chi(z)$ away from us. The z-axis is typically chosen so that it points to the center of the distribution of galaxies. In the plane perpendicular to this axis, a galaxy's position is determined by the two-dimensional vector $\boldsymbol{\theta} = (\theta_1, \theta_2)$. Therefore, the three-dimensional position vector $\boldsymbol{x}$ has components

$$\boldsymbol{x}(\chi(z), \boldsymbol{\theta}) = \chi(z)(\theta_1, \theta_2, l) . \tag{49}$$

The assumption that all galaxies are located near the z-axis clearly breaks down if the survey measures structure on very large angular scales.

We measure all galaxies along the line of sight, effectively integrating over $\chi(z)$. Therefore an overdensity at angular position $\boldsymbol{\theta}$ is

$$\delta_2(\boldsymbol{\theta}) = \int_0^{\chi_\infty} \mathrm{d}\chi W(\chi) \delta(\boldsymbol{x}(\chi, \boldsymbol{\theta})) , \tag{50}$$

where the subscript 2 denotes the fact that δ on the left is the angular - or two-dimensional - overdensity, while δ on the right is the full three-dimensional overdensity. The upper limit on the χ integral corresponds to $z \to \infty$, equal to $\chi_\infty = 2/H_0$ in a flat, matter-dominated universe. The selection function $W(\chi)$ encodes the information: it is the

probability of observing a galaxy a comoving distance χ from us. Galaxies at large distances are too faint to be included in a survey, whereas there are relatively few galaxies at very low redshift simply because the volume is small. Since it is a probability, the selection function is normalized so that $\int_0^{\chi_\infty} \mathrm{d}\chi W(\chi) = 1$.

The 2D vector conjugate to $\boldsymbol{\theta}$ will be $\boldsymbol{l}$, so that the Fourier transform of $\delta_2(\boldsymbol{\theta})$ is

$$\tilde{\delta}_2(\boldsymbol{l}) = \int \mathrm{d}^2\theta \mathrm{e}^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} \delta_2(\boldsymbol{\theta}) \ . \tag{51}$$

The two-dimensional power spectrum is defined as

$$\langle \tilde{\delta}_2(\boldsymbol{l}) \tilde{\delta}_2^*(\boldsymbol{l}') \rangle = (2\pi)^2 \delta^2(\boldsymbol{l} - \boldsymbol{l}') P_2(l) \ , \tag{52}$$

The 2D power spectrum is

$$\begin{aligned} P_2(l) &= \frac{1}{(2\pi)^2} \int \mathrm{d}^2 l' \langle \delta_2(\boldsymbol{l}) \delta_2^*(\boldsymbol{l}') \rangle \ , \\ &= \frac{1}{(2\pi)^2} \int \mathrm{d}^2 l' \int \mathrm{d}^2\theta \int \mathrm{d}^2\theta' \mathrm{e}^{-i\boldsymbol{l}\cdot\boldsymbol{\theta}} \mathrm{e}^{-i\boldsymbol{l}'\cdot\boldsymbol{\theta}'} \int_0^{\chi_\infty} \mathrm{d}\chi W(\chi) \int_0^{\chi_\infty} \mathrm{d}\chi' W(\chi') \\ &\langle \delta(\boldsymbol{x}(\chi, \boldsymbol{\theta})) \delta(\boldsymbol{x}'(\chi', \boldsymbol{\theta}')) \rangle \ . \end{aligned}$$

The integral over $\boldsymbol{l}'$ gives $(2\pi)^2$ times a Dirac delta function in $\boldsymbol{\theta}'$ and the brackets give the 3D correlation function,

$$\xi(\boldsymbol{x} - \boldsymbol{x}') \equiv \langle \delta(\boldsymbol{x}) \delta(\boldsymbol{x}') \rangle = \int \frac{\mathrm{d}^3 k}{(2\pi)^2} P(k) \mathrm{e}^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x}')} \ . \tag{53}$$

The average here $\langle \cdots \rangle$ is over all realizations of the density field. At very small distance, we expect galaxies to be clustered strongly as a result of gravity, so ξ is positive. As the distance gets larger, correlations die off and ξ gets smaller and eventually goes negative. The second line follows since the correlation function is the Fourier transform of the power spectrum.

$$P_2(l) = \int d^2\theta e^{-il\cdot\theta} \int_0^{\chi_\infty} d\chi W(\chi) \int_0^{\chi_\infty} d\chi' W(\chi') \int \frac{d^3k}{(2\pi)^3} P(k) e^{ik\cdot[x(\chi,\theta)-x(\chi',0)]} , \qquad (54)$$

$$= \int_0^{\chi_\infty} d\chi W(\chi) \int_0^{\chi_\infty} d\chi' W(\chi') \int_{-\infty}^{\infty} \frac{dk_3}{2\pi} P\left(\sqrt{k_3^2 + l^2/\chi^2}\right) e^{ik_3[\chi-\chi']} . \qquad (55)$$

The argument of the exponential is $i[k_1\chi\theta_1 + k_2\chi\theta_2 + k_3(\chi - \chi')]$, so the integral over angles $\boldsymbol{\theta}$ gives Dirac delta functions setting $l_1 = \chi k_1$ and $l_2 = \chi k_2$.

The only 3D Fourier modes that contribute to the integral are those with k_3 very small, much smaller than l/χ. We first need to estimate l, the variable conjugate to θ. Roughly, l^{-1} is of order the angular scales probed by the survey. Since we are working in the small angle approximation, $l/\chi \sim 1/(\chi\theta) \gg 1/\chi$.

The modes with longitudinal wavenumber k_3 much greater than χ^{-1} do not give rise to angular correlations because of cancellations along the line of sight. Only modes with k_3 smaller than χ^{-1} lead to angular correlations. Therefore, the relevant transverse wavenumbers l/χ are much larger than the relevant longitudinal wavenumbers, and the argument of the 3D power spectrum can be set to l/χ. With this approximation, the k_3 integral gives a Dirac delta function in $\chi - \chi'$ so

$$P_2(l) = \int_0^{\chi_\infty} d\chi \frac{W^2(\chi)}{\chi^2} P(l/\chi) . \qquad (56)$$

This is an expression for the 2D power spectrum as an integral over the line of sight. Change dummy variables from $\chi \to k \equiv l/\chi$ to rewrite the integral as

$$P_2(l) = \frac{1}{l} \int_0^{\infty} dk P(k) W^2(l/k) . \qquad (57)$$

The angular correlation function is the Fourier transform of the 2D power spectrum,

$$w(\theta) = \int \frac{d^2l}{(2\pi)^2} e^{il\cdot\theta} P_2(l) . \qquad (58)$$

Since P_2 depends only on the magnitude of $\boldsymbol{l}$, the angular part of the integration over l is $\int_0^{2\pi} \mathrm{d}\phi \mathrm{e}^{il\theta \cos\phi}$, which is proportional to $J_0(l\theta)$, the Bessel function of order zero.

$$w(\theta) = \int_0^\infty \frac{\mathrm{d}l}{2\pi} l P_2(l) J_0(l\theta) = \int_0^\infty \mathrm{d}k k P(k) F(k,\theta) \ , \qquad (59)$$

where the second line follows from changing the order of integration. Here the kernel for the angular correlation function is

$$F(k,\theta) \equiv \frac{1}{k} \int_0^\infty \frac{\mathrm{d}l}{(2\pi)} J_0(l\theta) W^2(l/k) \ . \qquad (60)$$

The kernel is a function of $k\theta$. The kernel is constant at small $k\theta$ and then begins damped oscillations. The contribution from small k though is suppressed because the integral is over the kernel weighted by $kP(k)$, and the latter goes to zero as $k \to 0$. The modes that contribute most to $w(\theta)$ are typically those with wavenumbers of order the first turnover in the kernel, $k\theta \sim 0.2h \ \mathrm{Mpc}^{-1}$ degrees for APM and a factor of 3 smaller for the deeper Sloan Digital Sky Survey (SDSS). This means that the angular correlation function at $5°$ in APM is most sensitive to power at $k = 0.04h \ \mathrm{Mpc}^{-1}$. The wavenumbers contributing to $w(\theta)$ in a deeper angular survey are smaller. This makes sense: the same angle probes larger physical scales in a deeper survey.

[3]

References

[1] H. Mo, F. C. van den Bosch, and S. White. *Galaxy Formation and Evolution*. May 2010.

[2] S. Dodelson. *Modern cosmology*. 2003.

[3] S. Weinberg. *Cosmology.* Oxford University Press, 2008.