Clustering of galaxies

- Notions: peculiar velocities, redshift space
- Correlation function: definitions and methods of estimation
- Power Spectrum
- Effects: redshift distortions
- Biases and biasing
- Observations

Correlation function: definition

This is usually quantified using the ***2-point correlation function, ξ (r),*** defined as an "excess probability" of finding another galaxy at a distance ***r*** from some galaxy, relative to a uniform random distribution; averaged over the entire set:

$$dN(r) = \rho_0 \left(1 + \xi(r)\right) dV_1 dV_2$$

Correlation function is often approximated with a power law:

$$\xi(r) = \left(r / r_0\right)^{-\gamma}$$

Parameter r_0 is called the correlation length

Estimators of the correlation function

- Simplest estimator: count the number of data-data pairs, $\langle DD \rangle$, and the equivalent number in a randomly generated (Poissonian) catalog, $\langle RR \rangle$:

$$\xi(r)_{est} = \frac{\langle DD \rangle}{\langle RR \rangle} - 1$$

- A better (Landy-Szalay) estimator is:

$$\xi(r)_{est} = \frac{\langle DD \rangle - 2\langle RD \rangle + \langle RR \rangle}{\langle RR \rangle}$$

where $\langle RD \rangle$ is the number of data-random pairs

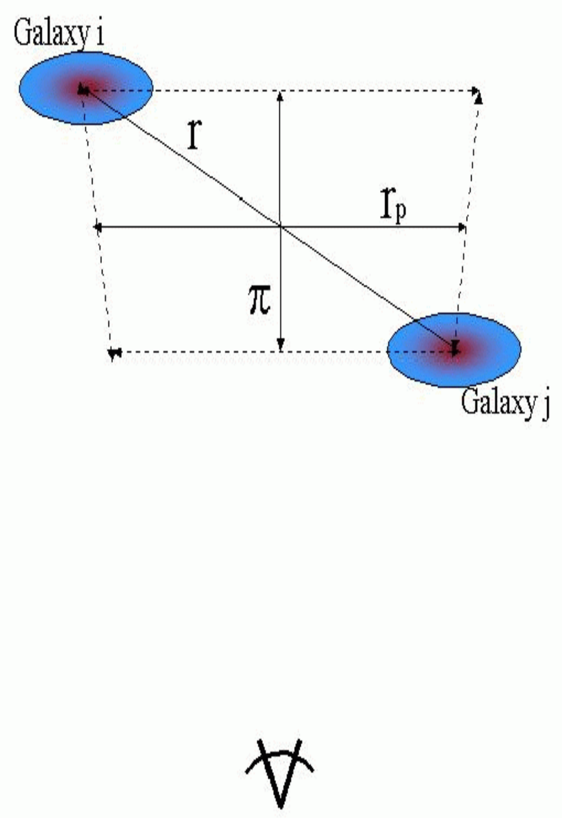
- This takes care of the edge effects, where one has to account for the missing data outside the region sampled, which can have fairly irregular boundaries

Angular and 3D correlation functions

$$w(r_p) = 2\int_{r_p}^{\infty} \xi(r)\left(r^2 - r_p^2\right)^{-1/2} dr$$

r_p: projected distance between pairs of galaxies,

π: distance parallel to the line of sight



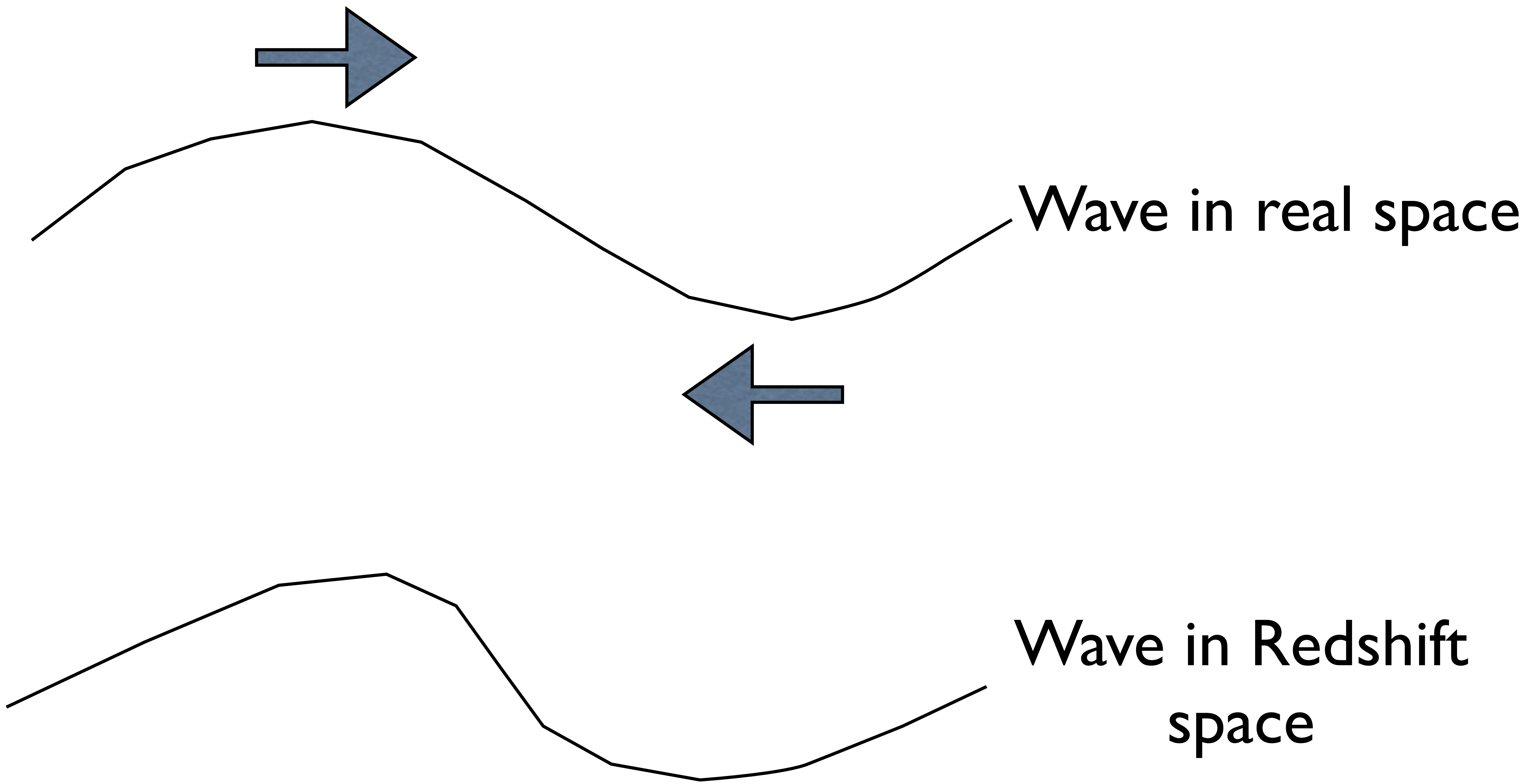
$$w(r_p) = \int_{-\delta\pi}^{+\delta\pi} \xi(r_p, \pi)\, d\pi$$

Inverting angular correlation function

$$w_p(r_p) = 2\int_0^\infty dy\, \xi\left[(r_p^2 + y^2)^{1/2}\right] = 2\int_{r_p}^\infty r\, dr\, \xi(r)\big(r^2 - r_p^2\big)^{-1/2} \tag{3}$$

$$\xi(r) = -\frac{1}{\pi}\int_r^\infty w_p'(r_p)(r_p^2 - r^2)^{-1/2} dr_p$$

Redshift distortions: long waves



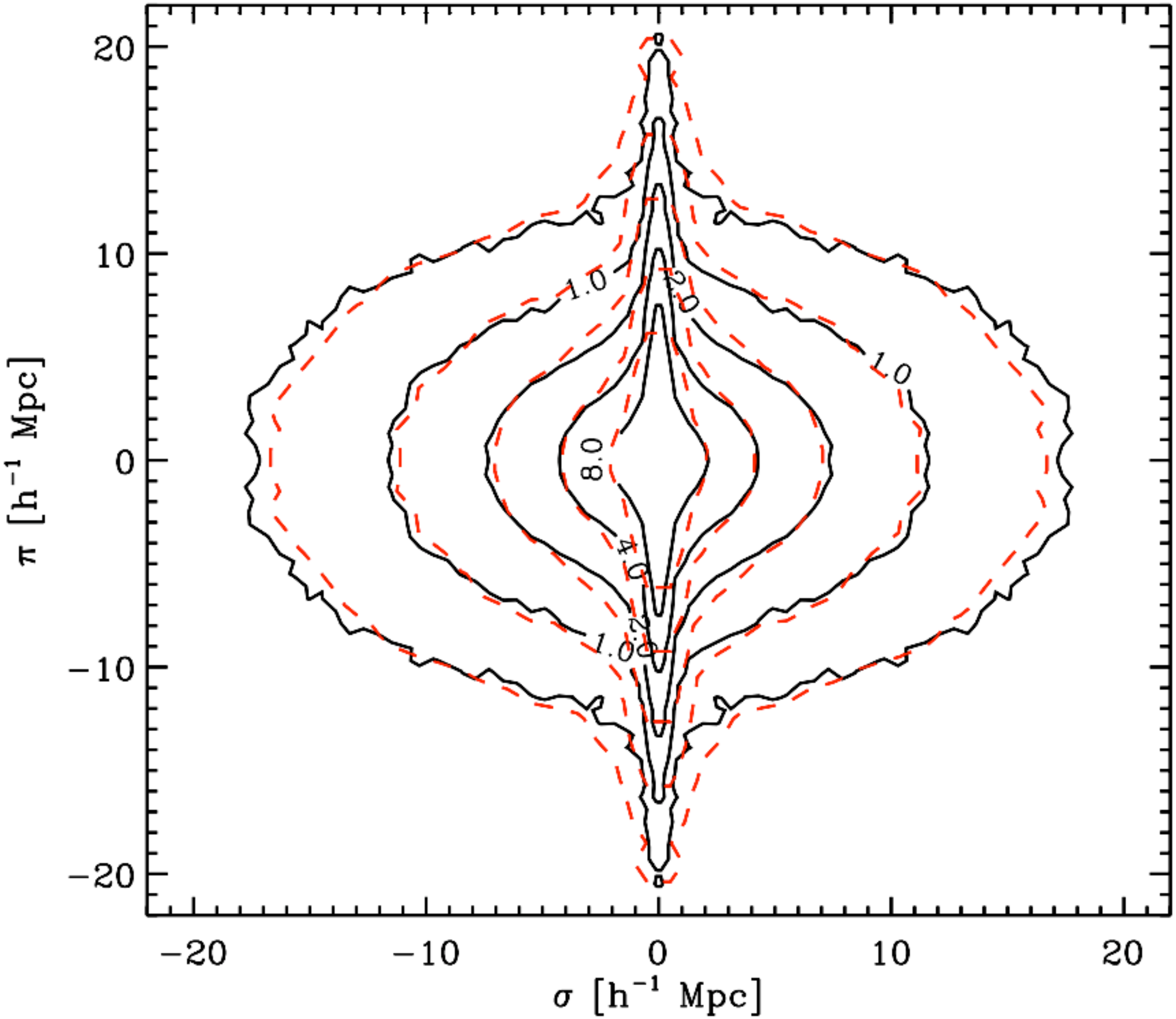
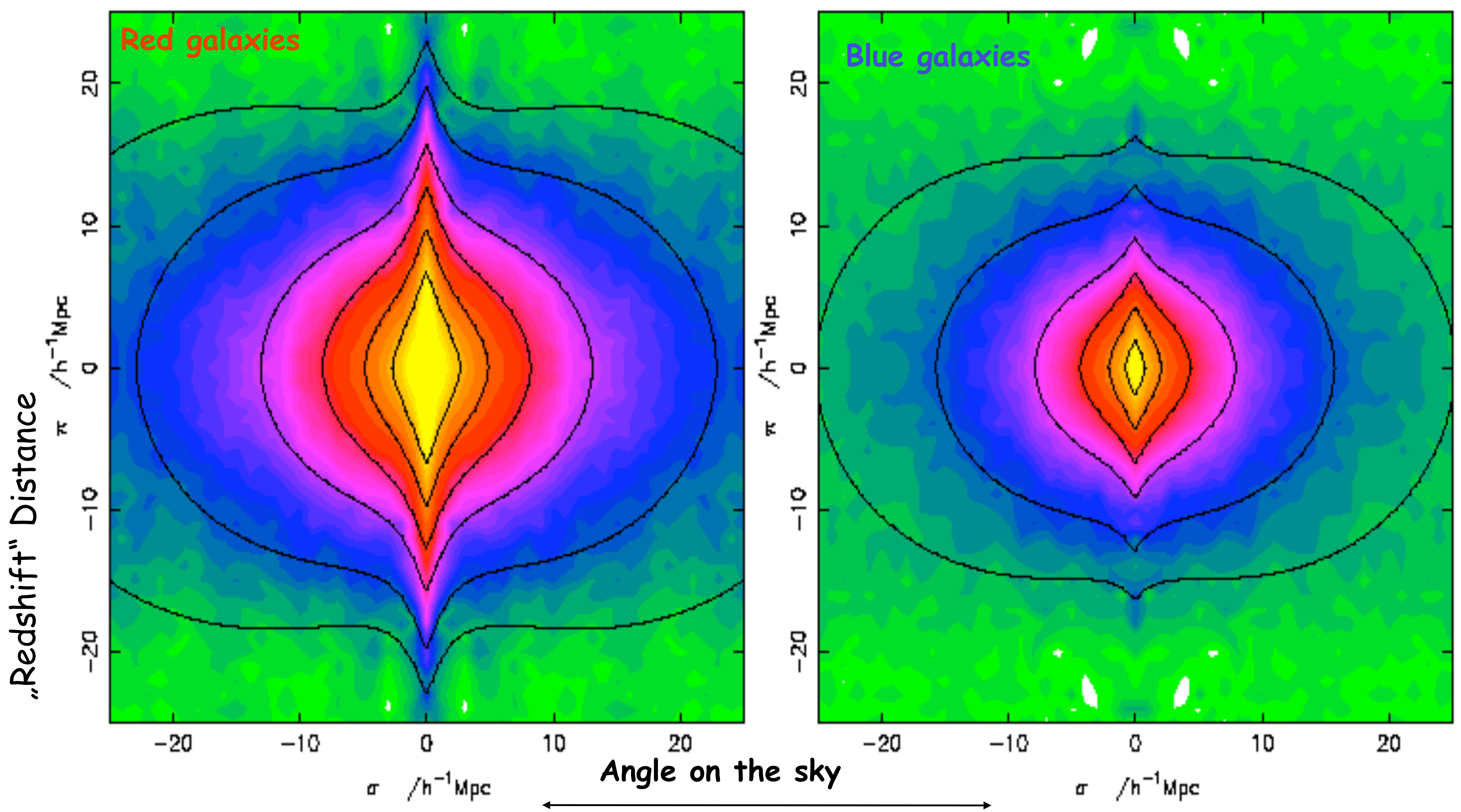


Figure 7. Contours of the two-dimensional correlation function $\xi(\sigma, \pi)$ estimated from the two-year BOSS-CMASS North galaxy sample (dashed line) at $0.4 < z < 0.7$ and for our MultiDark halo catalog constructed using the HAM technique at $z = 0.53$.

Redshift distortions: 'finger-of-god' effect on small scales



Angular correlation function $w_p(r_p)$

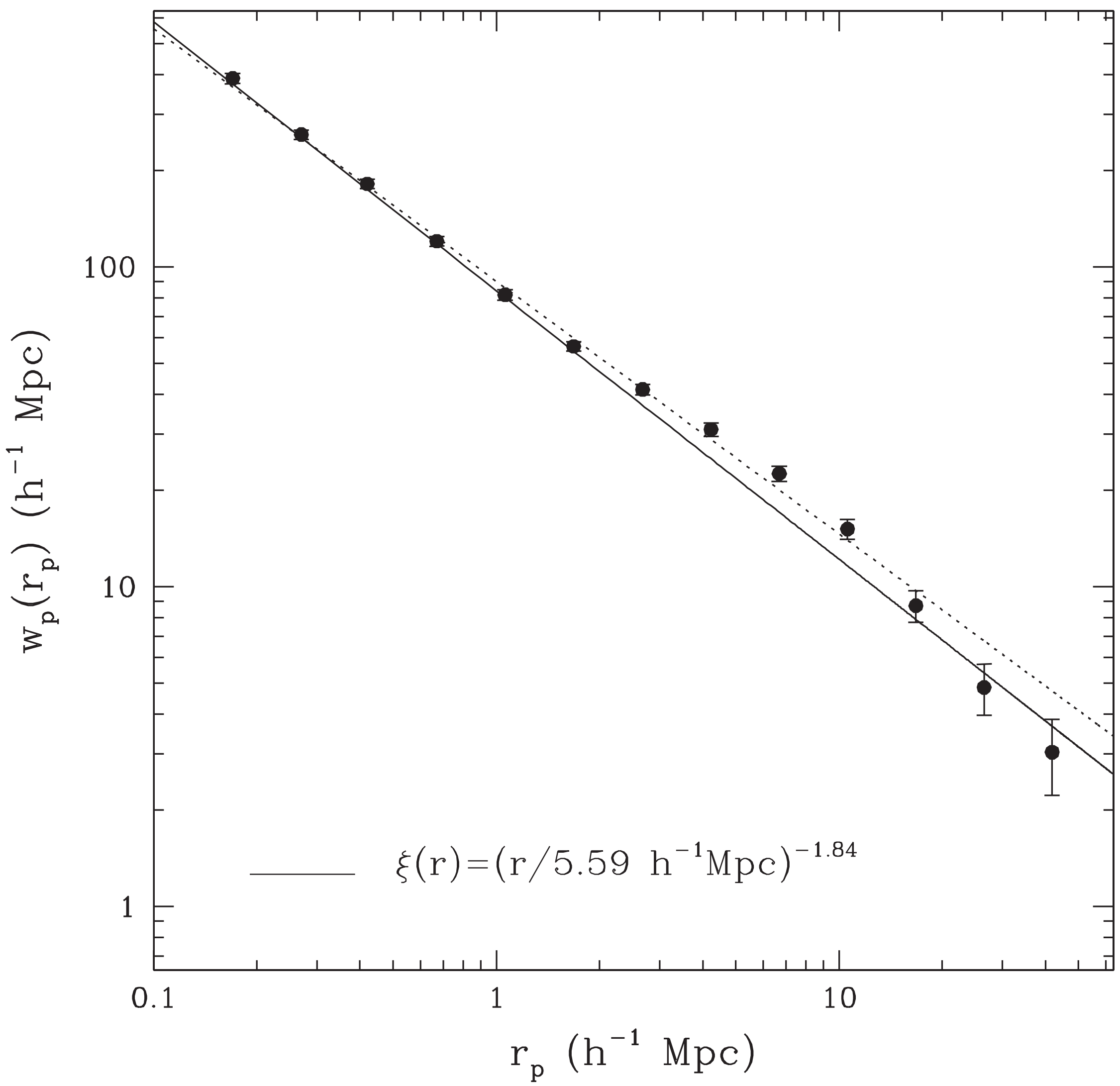


FIG. 6.—Projected galaxy correlation function $w_p(r_p)$ for the flux-limited galaxy sample. The solid line shows a power-law fit to the data points, using the full covariance matrix, which corresponds to a real-space correlation function $\xi(r) = (r/5.59\ h^{-1}\ \mathrm{Mpc})^{-1.84}$. The dotted line shows the fit when using only the diagonal error elements, corresponding to $\xi(r) = (r/5.94\ h^{-1}\ \mathrm{Mpc})^{-1.79}$. The fits are performed for $r_p < 20\ h^{-1}$ Mpc.

If only 2-D positions on the sky are known, then use angular separation $\boldsymbol{\theta}$ instead of distance $\boldsymbol{r}$:

$$w(\theta) = (\theta/\theta_0)^{-\beta}, \quad \beta = \gamma - 1$$

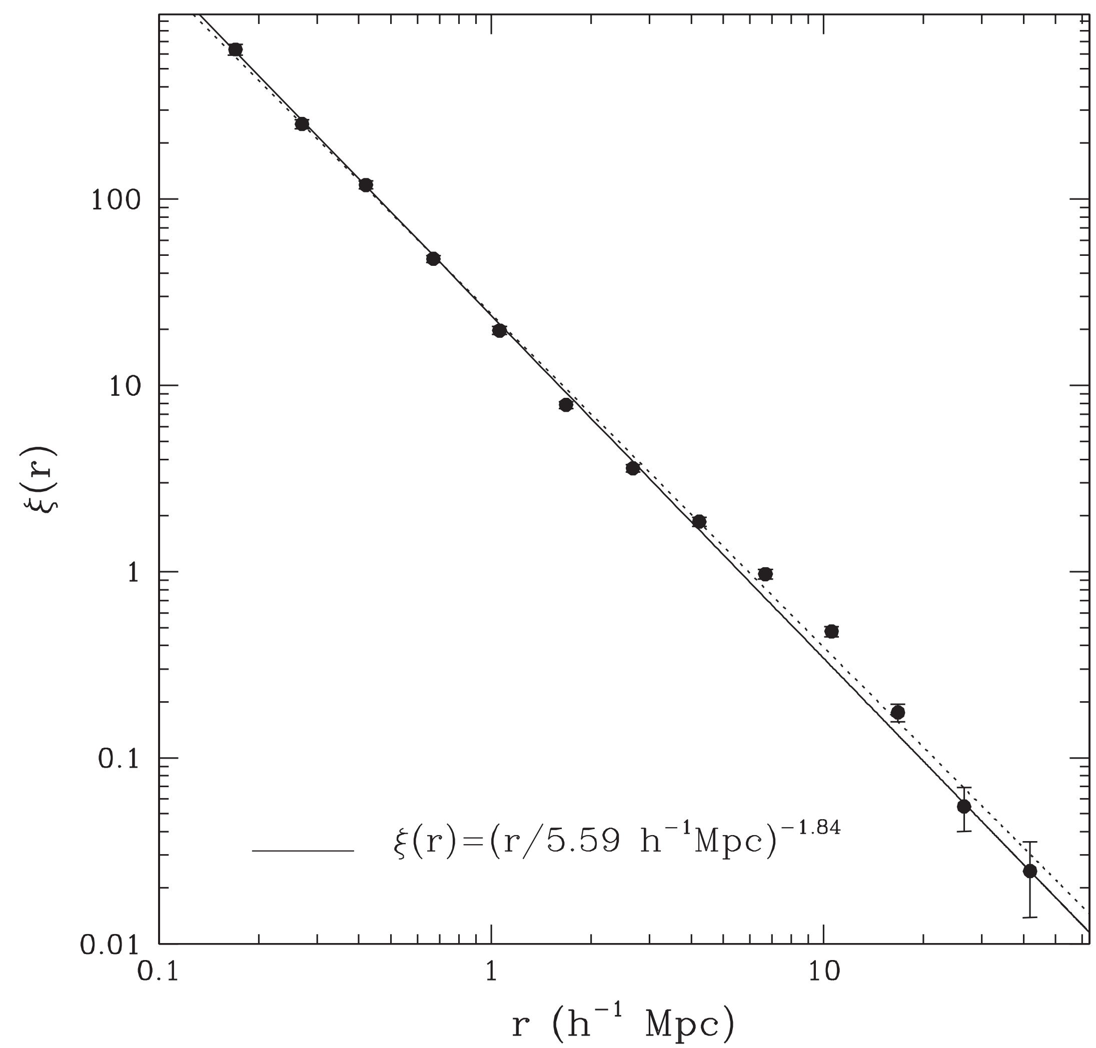


FIG. 7.—Real-space correlation function $\xi(r)$ for the flux-limited galaxy sample, obtained from $w_p(r_p)$ as discussed in the text. The solid and dotted lines show the corresponding power-law fits obtained by fitting $w_p(r_p)$ using the full covariance matrix or just the diagonal elements, respectively.

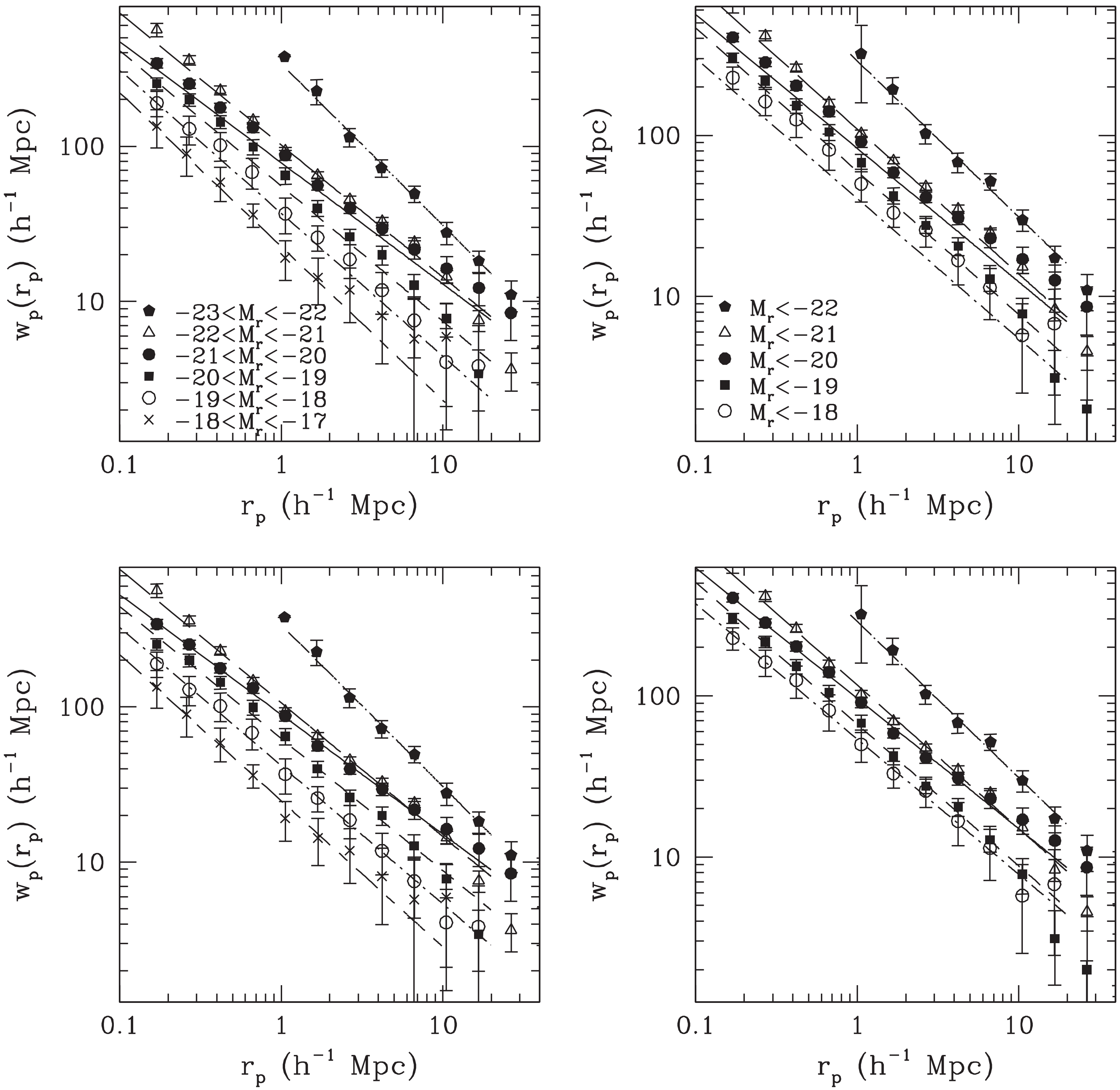
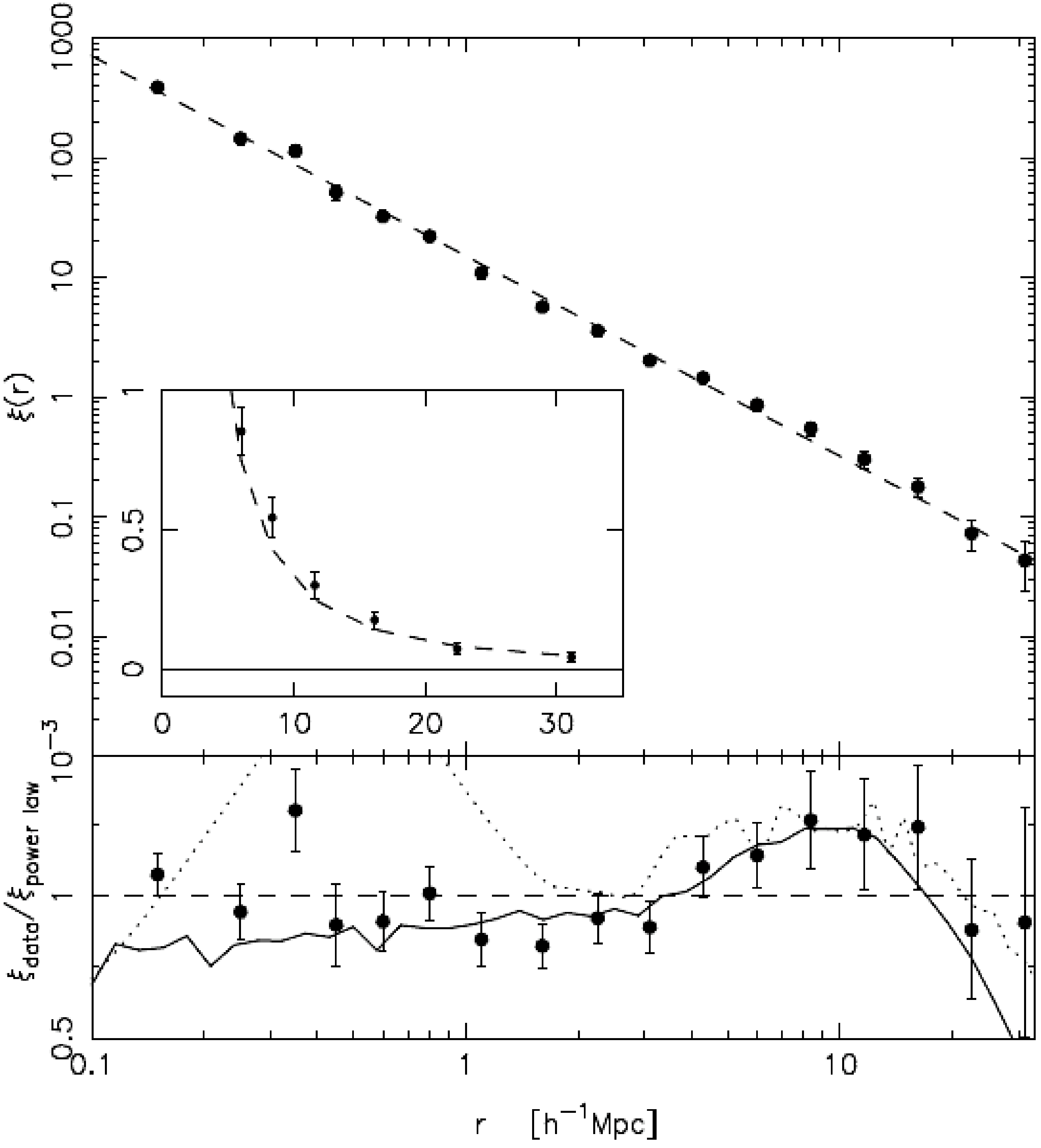


FIG. 8.—*Top left*: Projected galaxy correlation functions $w_p(r_p)$ for volume-limited samples with the indicated absolute magnitude and redshift ranges. Lines show power-law fits to each set of data points, using the full covariance matrix. *Top right*: Same as top left, but now the samples contain all galaxies brighter than the indicated absolute magnitude; i.e., they are defined by luminosity thresholds rather than luminosity ranges. *Bottom panels*: Same as the top panels, but now with power-law fits that use only the diagonal elements of the covariance matrix. [*See the electronic edition of the Journal for a color version of this figure.*]

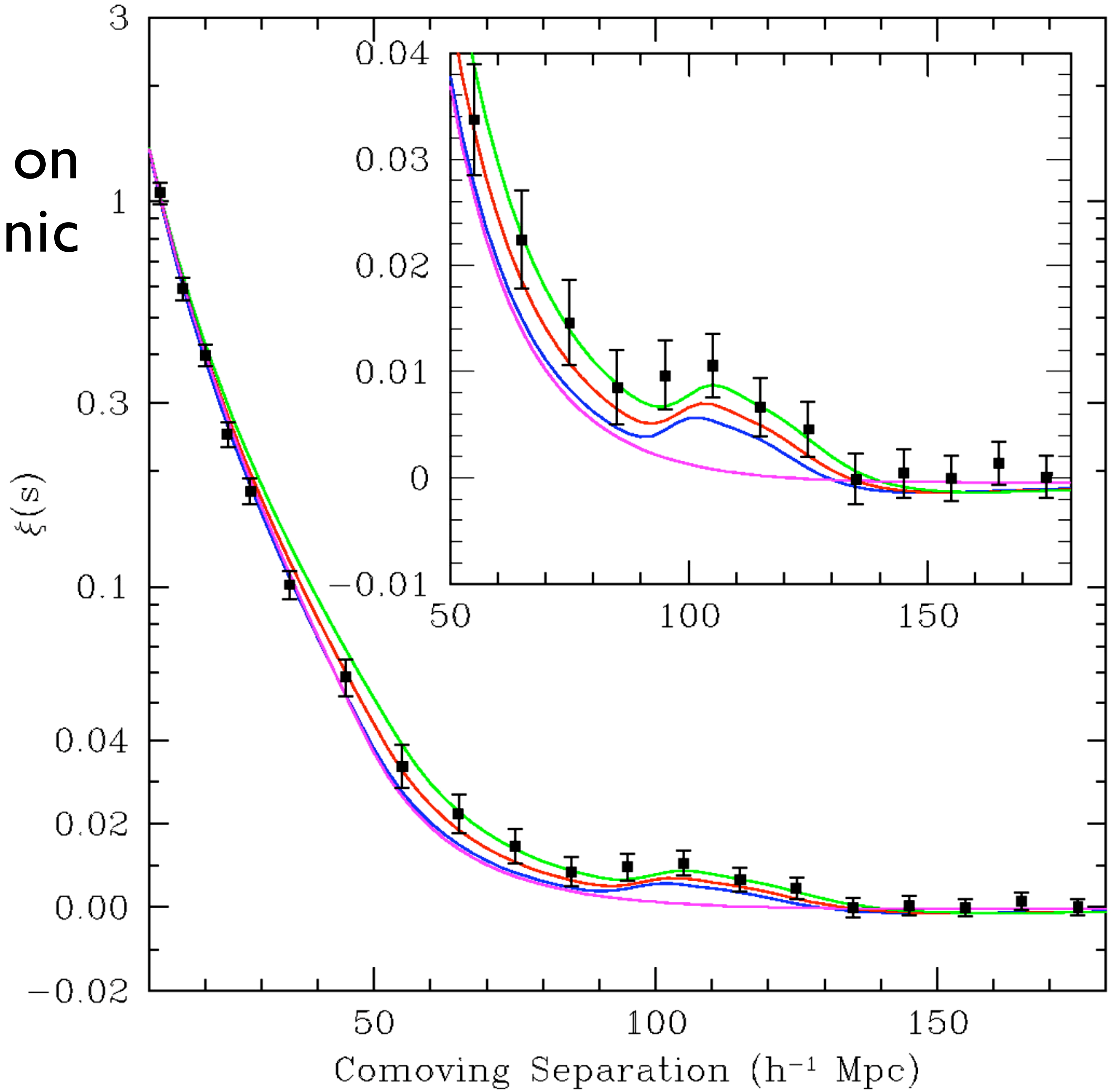
Galaxy Correlation Function

As measured by the 2dF redshift survey



Deviations from the power law:

Correlation function on large scales: baryonic oscillations

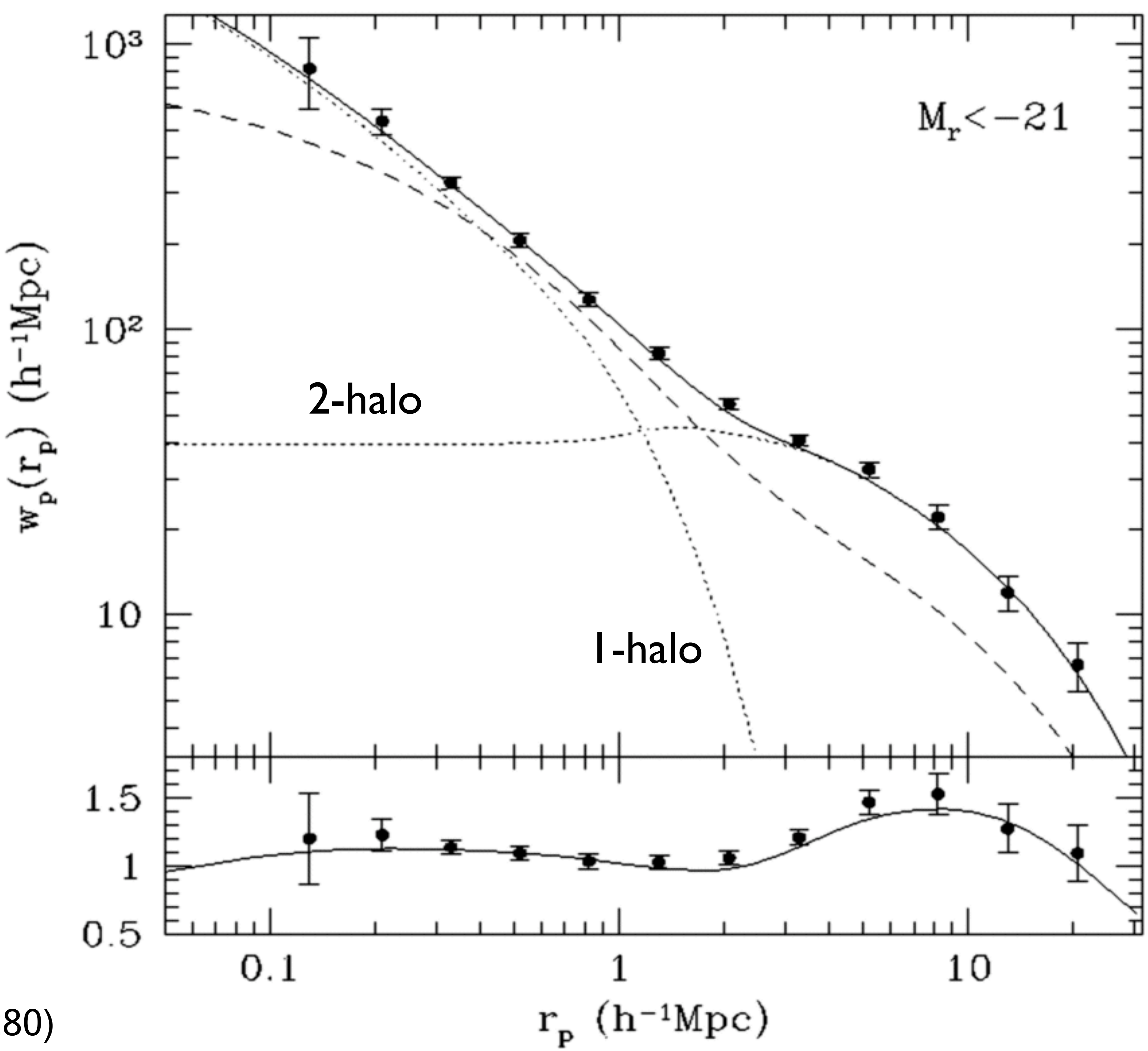


SDSS (Eisenstein et al.)

Angular correlation function: SDSS results

Two contributions:

- number-density profile of galaxies inside the same halo
- clustering of halos



Zehavi et al. (astro-ph/0301280)

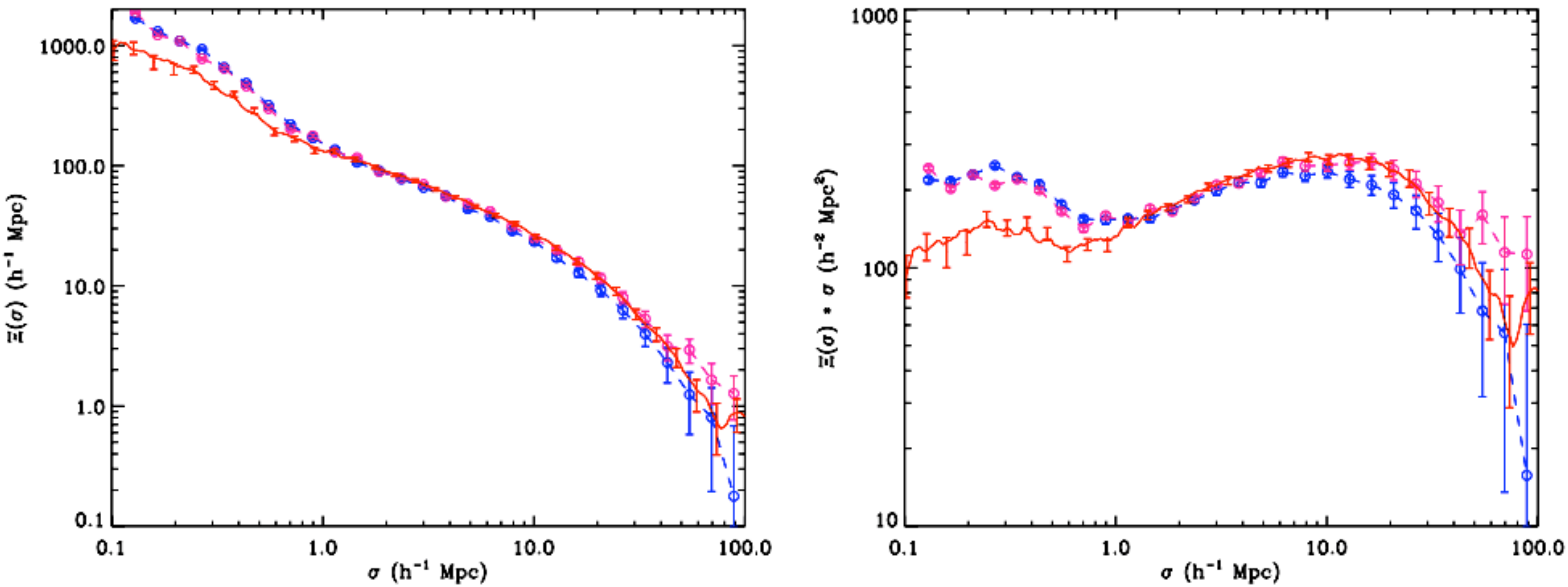


Figure 8. *Left panel:* Projected correlation function for the $0.4 < z < 0.7$ two-year BOSS-CMASS North and South galaxy samples (blue and magenta open circles respectively) and the MultiDark catalog selected with the HAM procedure at $z = 0.53$ (solid line). Error bars for MultiDark give an estimate of the cosmic variance magnitude. BOSS-CMASS error bars were estimated using an ensemble of 600 PTHalos mock galaxies. The transition between the 1st and 2nd halo terms can be seen at $\sim 1\ h^{-1}$ Mpc. Flattening of the signal at intermediate scales and bending at large scales are also evident features. *Right panel:* Detailed differences between our ΛCDM model and BOSS clustering measures is better seen when plotting the quantity $\Xi(\sigma)\,\sigma$ as a function of projected distance (see text).

$$\Xi(\sigma) = 2\int_0^\infty \xi(\sigma,\pi)\,\mathrm{d}\pi. \tag{2}$$

In practice, we integrate out to $\pi_{\rm max} = 200\,h^{-1}\,$Mpc .

We compute the full correlation functions $\xi(\sigma,\pi)$ using the Landy & Szalay (1993) estimator

$$\xi(\sigma,\pi) = \frac{\mathrm{DD} - 2\mathrm{DR} + \mathrm{RR}}{\mathrm{RR}} \tag{3}$$

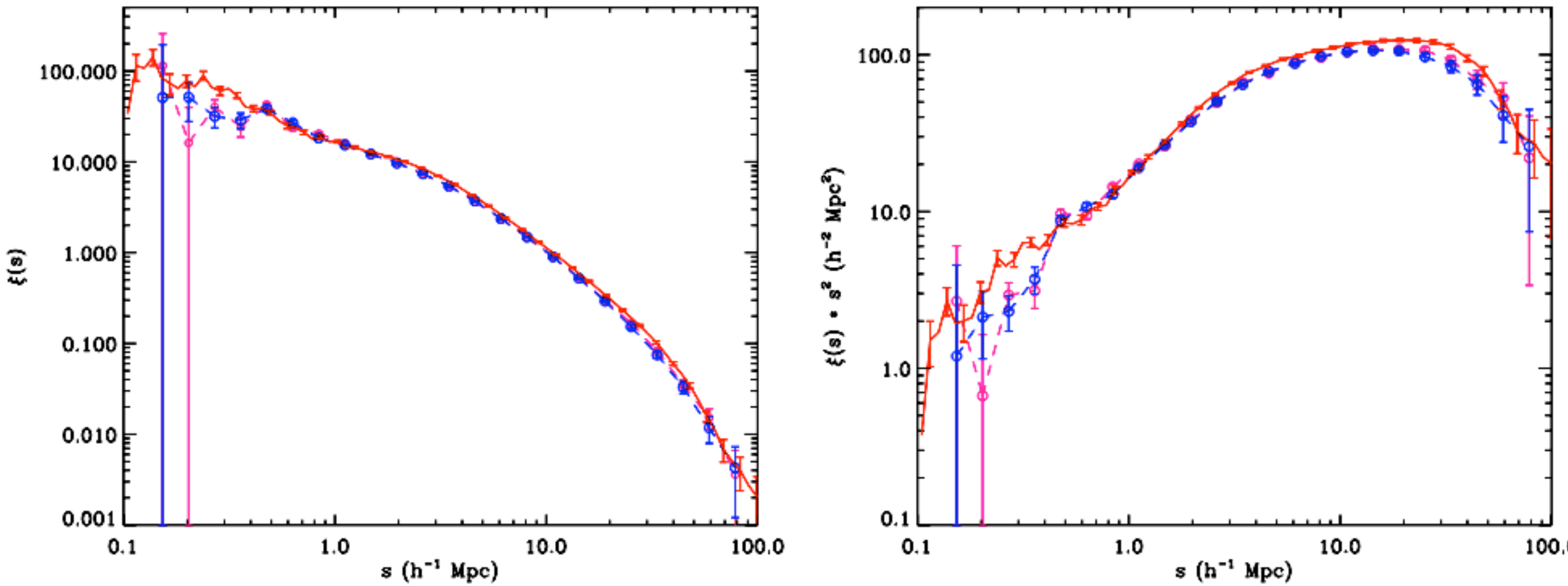


Figure 9. *Left panel:* Redshift-space correlation function both for the tow-year BOSS-CMASS North and South galaxy samples at $0.4 < z < 0.7$ (blue and magenta open circles respectively) and the MultiDark catalog selected with the HAM procedure at $z = 0.53$ (solid line). Error bars are obtained in tha same way as in Fig. 8. *Right panel:* Shown is the quantity $\xi(s)\,s^2$ which better reflects the differences between our ΛCDM model and BOSS clustering measures.

Power Spectrum

δ(k) is the Fourier amplitude

$$P(\mathbf{k}) = \left|\delta(\mathbf{k})\right|^2$$

- Naïve estimator for a discrete density field is

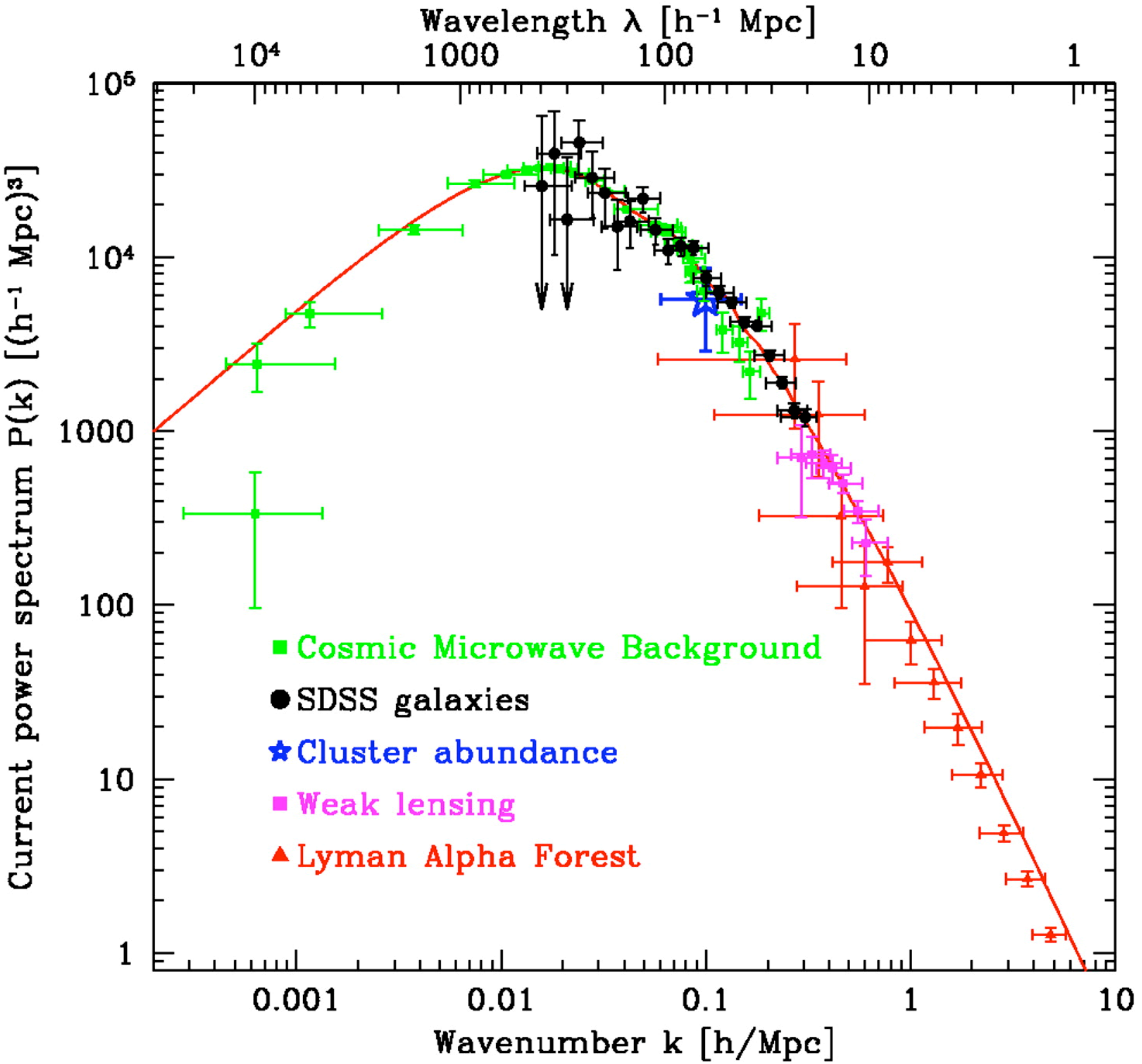
$$\hat{f}(\mathbf{k}) = \frac{1}{N}\sum_n e^{i\mathbf{k}\mathbf{r}_n}$$

- We need to take into account (1) selection function ϕ(r) and shot noise w(k)

$$\hat{f}(\mathbf{k}) = \sum_n \phi(\mathbf{r}_n)e^{i\mathbf{k}\mathbf{r}_n} - w(\mathbf{k})$$

$$\phi(\mathbf{r}) = \frac{\overline{n}(r)}{1+\overline{n}(r)P(k)}$$

The Observed Power Spectrum



(Tegmark et al.)

Baryonic acoustic oscillations: Power spectrum

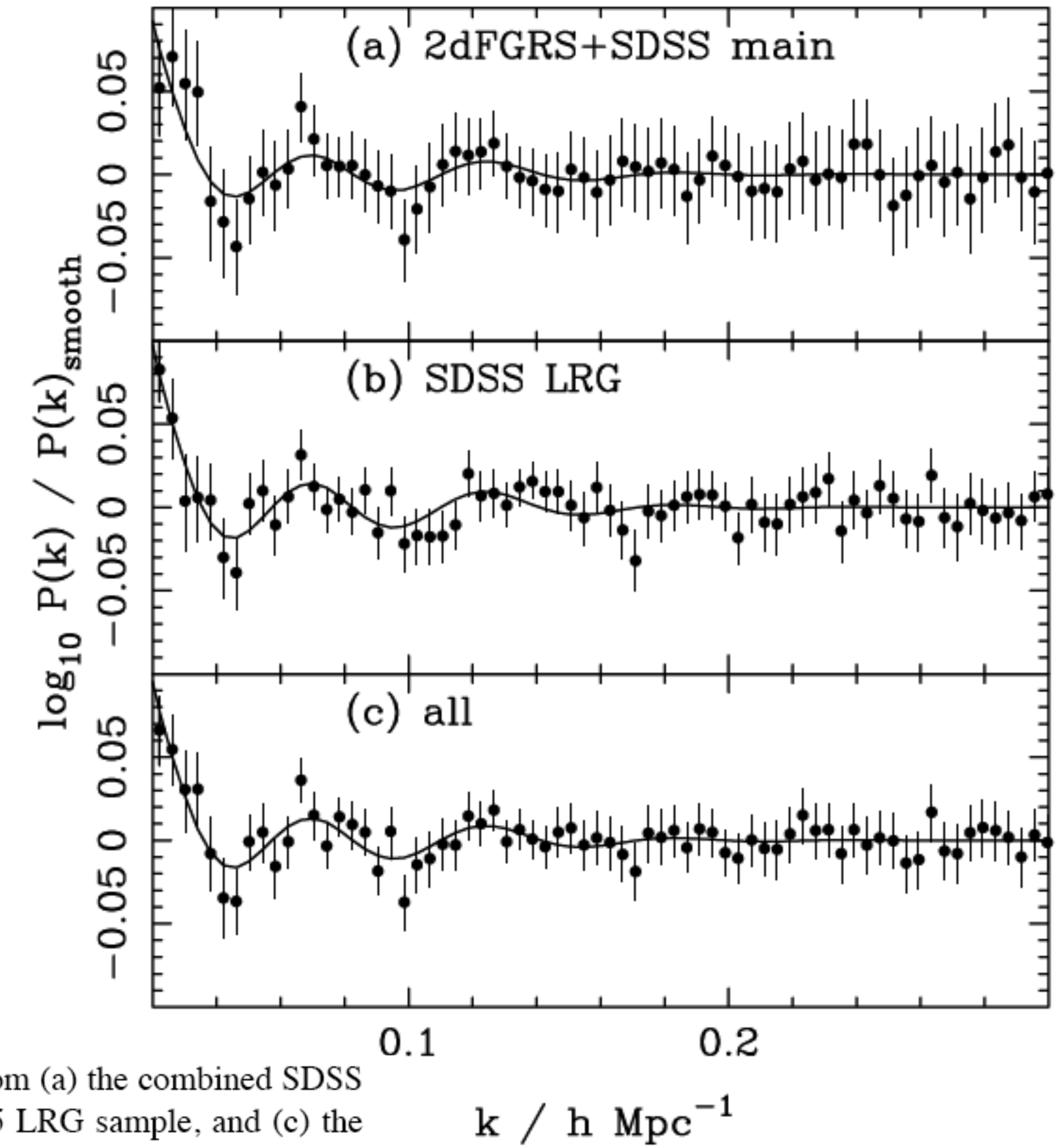


Figure 2. BAOs in power spectra calculated from (a) the combined SDSS and 2dFGRS main galaxies, (b) the SDSS DR5 LRG sample, and (c) the combination of these two samples (solid symbols with 1σ errors). The data are correlated and the errors are calculated from the diagonal terms in the covariance matrix. A standard ΛCDM distance–redshift relation was assumed to calculate the power spectra with $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$. The power spec-

Percival etal 2007