Three-Point Correlations of Abell Clusters

G. Tóth and J. Hollósi
Space Telescope Science Institute
Baltimore, MD, USA, and
Roland Eötvös University
Budapest, Hungary

A.S. Szalay
Roland Eötvös University
Budapest, Hungary, and
Johns Hopkins University
Baltimore, MD, USA

Abstract

We estimate the irreducible three-point correlations of Abell clusters using all distance class 5+6 clusters with latitude $|b^{II}| \geq 40°$. We find that these clusters satisfy a relation between the two- and three-point correlation functions:

$$\zeta(r,s,u) \approx Q\big(\xi(r)\,\xi(s) + \xi(s)\,\xi(u) + \xi(u)\,\xi(r)\big)$$

similar to that for galaxies. The value of Q has large uncertainties:

$$Q = 0.9 \pm 0.5\,,$$

with a strong discrepancy between the northern and southern hemispheres, $Q_{north} \approx 0.3$, while $Q_{south} \approx 1.7$ [1]. Higher order terms seem to be absent in ζ. Several error estimation methods are applied.

1 Definition of the three-point correlation function

Spatial:

Let V_1, V_2 and V_3 be three volume elements in space. Let us denote their separations by r_{12}, r_{23} and r_{31} (*Fig. 1a*). With the mean density of clusters being ρ the expected number of triplets in the volume elements is $n_{exp} = \rho^3 V_1 V_2 V_3$ for a uniform distribution. In case of a correlated distribution there is an excess of triplets, thus the number of them will be:

$$n = \rho^3 V_1 V_2 V_3 \big(1 + \xi(r_{12}) + \xi(r_{23}) + \xi(r_{31}) + \zeta(r_{12}, r_{23}, r_{31})\big)\ , \tag{1}$$

where ξ and ζ are the spatial two- and three-point correlation functions, respectively.

[1] Work in progress indicates that this discrepancy is less significant, when one takes into account that the distribution of triplets is non-Poissonian; strongly correlated.

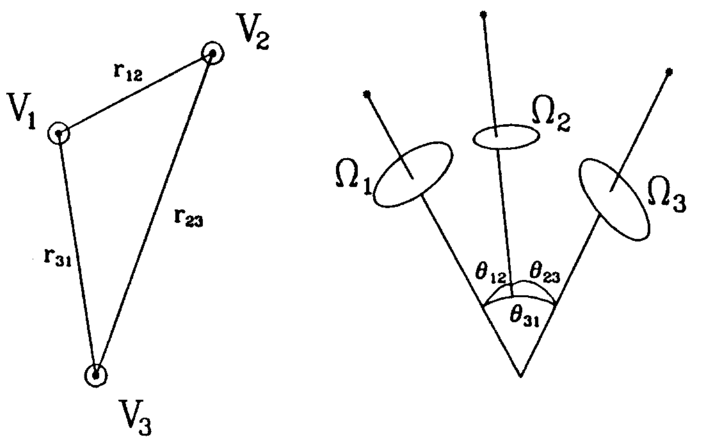

Fig. 1a, b. Definition of the three-point correlation functions.

Angular:

Let Ω_1, Ω_2 and Ω_3 be three solid angle elements on the surface of the unit sphere with separations θ_{12}, θ_{23} and θ_{31} (*Fig. 1b*). With the mean surface density of clusters being η, the expected number of triplets in the angle elements is $n_{exp} = \eta^3 \Omega_1 \Omega_2 \Omega_3$ for a uniform distribution. In case of a correlated distribution there is an excess of triplets, thus the number of them will be

$$n = \eta^3 \Omega_1 \Omega_2 \Omega_3 \left(1 + w(\theta_{12}) + w(\theta_{23}) + w(\theta_{31}) + z(\theta_{12}, \theta_{23}, \theta_{31})\right) , \quad (2)$$

where w and z are the angular two- and three-point correlation functions, respectively.

2 Motivation

Comparison with galaxies:

The distribution of galaxies were studied in detail by Groth and Peebles (1977). They determined the three- and even the four-point correlations so as to get quantitative results on the distribution of galaxies. An interesting relation was found between the spatial two- and three-point correlation functions:

$$\zeta_{gal}(r, s, u) \approx Q_{gal} \left(\xi_r \xi_s + \xi_s \xi_u + \xi_u \xi_r\right) , \quad (3)$$

where $Q_{gal} = 0.8 \ldots 1.3$ for various catalogues.

Comparison with models:

There are several models (both numerical and analytical) of clustering and they can be tested by comparing their correlation functions with the correlation function of an observed catalogue. A well known analytical model (Kaiser 1984, Bardeen *et al.* 1986, Politzer and Wise 1984) is the biasing of density fluctuations, where we assume that there were primordial Gaussian fluctuations of the mass density, and visible objects were formed only at places where the density reached a certain value ('biasing'). In a more general model (Szalay 1988) we have an arbitrary non-linear relation between luminosity and mass density. The three-point correlation function can be expanded in terms of the two-point correlation function, and the leading terms are:

$$\zeta(r,s,u) = Q\left(\xi_r\xi_s + \xi_s\xi_u + \xi_u\xi_r\right) + Q^3\xi_r\xi_s\xi_u + Q'\left(\xi_r^2\xi_s + \xi_r^2\xi_u + \ldots\right) . \quad (4)$$

Clusters are more likely to have preserved their initial distribution below the correlation length than galaxies, so the relation between the 2nd and 3rd coefficients can be checked for this.

3 Data

A magnetic tape of the Abell catalogue prepared by the Bulgarian Academy of Sciences is used. Our copy was obtained from UC Berkeley. We have discovered that somewhere in the copying processes an error occured: every 57th cluster is missing from the catalogue. Subsequently we found that this error is by no means specific to our tape. Several major institutes also had the faulty catalogue on their computer, thus we warn everybody to check his catalogue for the error. Abell's (1958) original paper was used to complete the data.

The catalogue contains 2712 clusters altogether (see *Fig. 1* in Hollósi and Efstathiou, *these proceedings*), the $D = 5+6$ (distance groups 5 and 6 and richness class $R \geq 1$) sample is used with declinations greater than $-27°$ and galactic latitudes $|b^{II}| \geq 40°$. Two subsamples are chosen for error estimation :

(HL)	High Latitude with $\|b^{II}\| \geq 40°$	(1323 clusters)
(NC)	North Cap with $b^{II} \geq +40°$	(844 clusters)
(SC)	South Cap with $b^{II} \leq -40°$	(479 clusters)

Within these geometrical boundaries only clusters of the statistical sample are present (the only exception is A915), thus no further geometrical constraints are necessary to restrict ourselves to the statistical sample. The areas of HL, NC and SC are 3.46, 2.24 and 1.22 sterad, respectively.

The extinction function is determined from the distribution of surface density $\eta(b^{II})$ from the data at different galactic latitudes. The clusters are counted in 50 equal area bins in the north galactic cap. The result is smoothed by a simple averaging process, and saved in a file to be used for generating random catalogues (*Fig. 2*). The same function is used for the south galactic cap.

4 How to Estimate the three–point correlation function

Triplet counts in space:

The usual way of calculating the three-point correlation function $\zeta(r,s,u)$, *e.g.* for galaxies, is to determine the number of triplets with given separations both in the data and random catalogues: $N_D(r,s,u)$ and $N_R(r,s,u)$. Assuming that we have already determined the two-point correlation function ξ, we can estimate ζ:

$$\zeta(r,s,u) = \frac{N_D(r,s,u)}{N_R(r,s,u)} - \xi_r - \xi_s - \xi_u - 1 \,. \tag{5}$$

Unfortunately only ≈ 300 Abell clusters have measured redshifts, thus we cannot apply this method to this subsample, because $N_D(r,s,u)$ would be too small.

From the angular three-point correlation function:

For two-point correlations the angular function $w(\theta)$ can be expressed in terms of the spatial function $\xi(r)$ and vice versa, so we can estimate $\xi(r)$ from $w(\theta)$ even if we do not know the redshifts.

For three-point correlations we can express the angular function z in terms of the spatial function ζ, but the inversion is unknown.

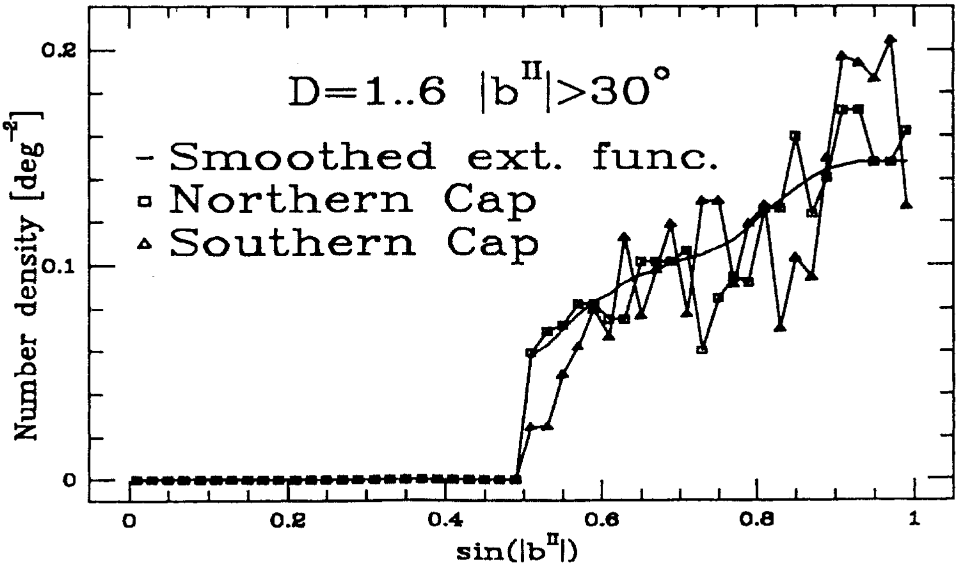

Fig. 2. Extinction function of the Abell clusters.

Expansion of the three-point correlation function:

Let us expand ζ into the symmetric terms of ξ up to the third order:

$$
\begin{aligned}
\zeta(r,s,u) = \quad & Q_0 \\
& +Q_1 \quad \cdot(\xi_r+\xi_s+\xi_u) \\
& +Q_{11} \quad \cdot(\xi_r\xi_s+\xi_s\xi_u+\xi_u\xi_r) \\
& +Q_2 \quad \cdot(\xi_r^2+\xi_s^2+\xi_u^2) \\
& +Q_{111} \cdot(\xi_r\xi_s\xi_u) \\
& +Q_{12} \quad \cdot(\xi_r^2\xi_s+\xi_r^2\xi_u\xi_s^2\xi_u+\xi_s^2\xi_r\xi_u^2\xi_r+\xi_u^2\xi_s) \\
& +Q_3 \quad \cdot(\xi_r^3+\xi_s^3+\xi_u^3)
\end{aligned}
\tag{6}
$$

Now the coefficients $Q_{\underline{k}}$ are to be determined instead of ζ, and this is possible, because the terms can be projected separately into angular terms, as we will see below. On the other hand we can easily answer the questions asked in Sect. 2 from the coefficients.

5 Projection of the expansion

Using the notation $\underline{k} \in \{0,1,11,2,111,12,3\}$ the expansion above can be written in a shorter form:

$$\zeta(r,s,u) = \sum_{\underline{k}} Q_{\underline{k}}\,\xi_{\underline{k}}(r,s,u)\ , \tag{7}$$

where $\xi_{\underline{k}}(r,s,u) = \xi_r^{k_1}\xi_s^{k_2}\xi_u^{k_3} + \text{symm.}$
Assuming that the selection function $P(r)$ is known, the expansion can be projected onto an equation for angular correlation functions:

$$z(a,b,c) = \sum_{\underline{k}} A_{\underline{k}} w_{\underline{k}}(a,b,c)\ . \tag{8}$$

The connection between $Q_{\underline{k}}\xi_{\underline{k}}$ and $A_{\underline{k}}w_{\underline{k}}$ is:

$$
\begin{aligned}
& A_{\underline{k}}w_{\underline{k}}(a,b,c) = \\
& \quad \frac{\rho^3}{\eta^3}\int_0^\infty dr_1\, r_1^2 P(r_1)\int_0^\infty dr_2\, r_2^2 P(r_2)\int_0^\infty dr_3\, r_3^2 P(r_3)\, Q_{\underline{k}}\xi_{\underline{k}}(r,s,u)\ ,
\end{aligned}
\tag{9}
$$

where

$$
\begin{aligned}
r^2 &= r_1^2 + r_2^2 - 2r_1r_2\cos a \\
s^2 &= r_2^2 + r_3^2 - 2r_2r_3\cos b \\
u^2 &= r_3^2 + r_1^2 - 2r_3r_1\cos c
\end{aligned}
\tag{10}
$$

Adopting $\xi(r) = (r/r_0)^{-\gamma}$ and $P(r) = 1$ for $d \le r \le D$ we can integrate analytically. The calculation is similar to the projection of $\xi(r)$ onto $w(\theta)$, but it is more complicated. The results are:

$$
\begin{aligned}
w_0 \quad & (a,b,c) = 1 \\
w_1 \quad & (a,b,c) = w_a + w_b + w_c \\
w_{11} \quad & (a,b,c) = w_a w_b + w_b w_c + w_c w_a \\
w_2 \quad & (a,b,c) = \tfrac{w_a^2}{a} + \tfrac{w_b^2}{b} + \tfrac{w_c^2}{c} \\
w_{111} \, & (a,b,c) = \tfrac{w_a w_b w_c}{a+b+c} \qquad \text{(for } \gamma = 2 \text{ only)} \\
w_{12} \quad & (a,b,c) = \tfrac{w_a^2 w_b + w_a^2 w_c}{a} + \tfrac{w_b^2 w_c + w_b^2 w_a}{b} + \tfrac{w_c^2 w_a + w_c^2 w_b}{c} \\
w_3 \quad & (a,b,c) = \tfrac{w_a^3}{a^2} + \tfrac{w_b^3}{b^2} + \tfrac{w_c^3}{c^2}
\end{aligned}
\tag{11}
$$

The angular coefficients are proportional to the spatial coefficients, but the ratios are different for the different terms due to the various moments of $P(r)$ involved. Let us denote the ratios by $R_{\underline{k}} = A_{\underline{k}}/Q_{\underline{k}}$. The numerical results for $\gamma = 2$ and for the distance limits $d = 250$ Mpc and $D = 600$ Mpc are:

$$
\begin{aligned}
R_0 \quad & = 1 \\
R_1 \quad & = 1 \\
R_{11} \quad & = 1.036 \\
R_2 \quad & = 5.9 \\
R_{111} \, & = 11.3 \\
R_{12} \quad & = 5.65 \\
R_3 \quad & = 63.4
\end{aligned}
$$

a, b and c are measured in degrees. Note that the spatial terms and the corresponding angular terms are similar, but all higher order terms are strongly amplified in the angular coefficients.

If we know *the angular two- and three-point correlation functions* we can determine the coefficients $A_{\underline{k}}$ from the angular expansion (Eqn. 7), and then the spatial coefficients $Q_{\underline{k}} = A_{\underline{k}}/R_{\underline{k}}$.

6 Estimation of angular correlations

Angular two-point correlation:

20 random catalogues were genarated with the proper areas and extinction functions for the HL, NC and SC samples. The number of pairs in 12 linear bins is counted up to 5° separation for the data samples and the random catalogues as well. The estimator of the angular correlation for each bin is

$$
w(\theta) = \frac{N_D(\theta)}{N_R(\theta)} - 1 \,, \tag{12}
$$

where $N_R(\theta)$ and $N_D(\theta)$ are the number of pairs in a bin of $[\theta - \frac{\Delta\theta}{2}, \theta + \frac{\Delta\theta}{2}]$ for the random and data catalogues, respectively. N_R can be estimated with an analytic formula, because for such small separations the non-linear effects of the boundary and the density gradient are negligible. Thus we can calculate the expected value of pair frequency analytically instead of generating several hundreds of random catalogues to eliminate the statistical scattering:

$$N_R(\theta) = \frac{1}{2} A \cdot \overline{\eta^2} \cdot 2\pi\theta\Delta\theta \ , \tag{13}$$

where A is the area of the sample, $\overline{\eta^2}$ is the mean squared surface density, $2\pi\theta\Delta\theta$ is the area where the second point can be if the first one is fixed. The formula is divided by 2, because otherwise all pairs would be counted twice. It is still necessary to determine the accurate coefficient of θ from fitting to the pair-frequency function of the random catalogues, because the presence of boundaries reduces the number of pairs.
The best power law fit for the HL sample is

$$w(\theta) = \frac{1}{\theta} \ . \tag{14}$$

The estimated angular correlation function is determined up to 30° and compared with the projected spatial correlation function and its analytic approximation (*Fig. 3*). The angular correlation functions of the different samples agree fairly well (*Fig. 4*).

Angular three-point correlation:

All triplets with separations up to 5° are counted and put into bins according to the length of the sides. Let $n_D(a,b,c)$ and $n_R(a,b,c)$ be the number of triplets in

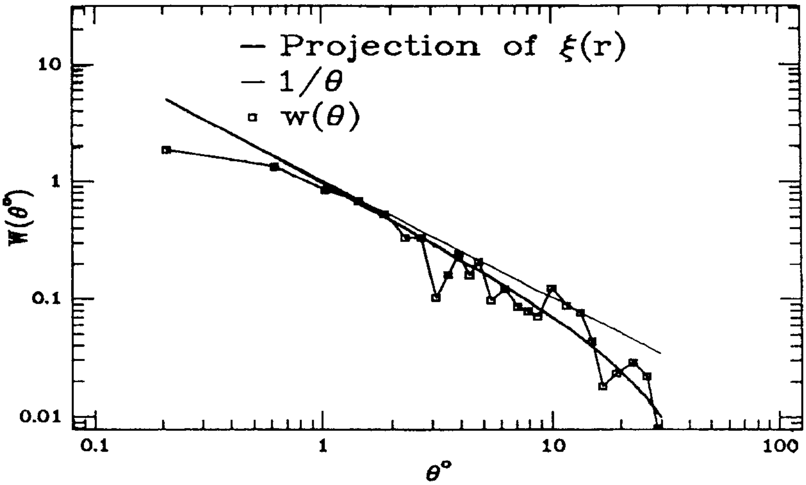

Fig. 3. Smooth curve: projection of $\xi(r) = (r/30\,\mathrm{Mpc})^{-2}$ analytically.
Straight line: best power law fit: $1/\theta$.
Points: angular two-point correlation function of the HL sample.

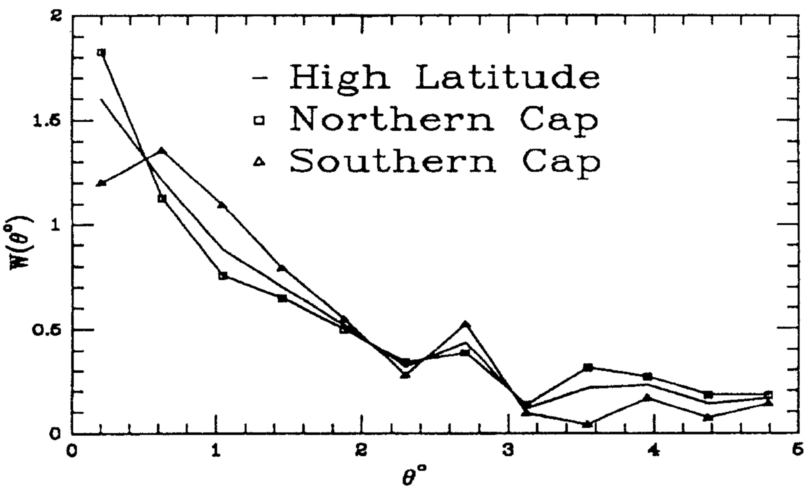

Fig. 4. Comparison of the angular two-point correlation functions of the various samples.

a bin of $V(a,b,c) = [a - \frac{\Delta}{2}, a + \frac{\Delta}{2}] \times [b - \frac{\Delta}{2}, b + \frac{\Delta}{2}] \times [c - \frac{\Delta}{2}, c + \frac{\Delta}{2}]$ for the data and random catalogues, respectively. Δ is the size of the bin. The estimator for the angular three-point correlation function is

$$z(a,b,c) \approx \frac{n_D(a,b,c)}{n_R(a,b,c)} - 1 - w_a - w_b - w_c \,. \tag{15}$$

This estimation, however, has some disadvantages. The equation is accurate for infinitesimal Δ only, but reducing Δ is limited because of the relatively small number of triplets in the data ($\approx$ 30000). There are also problems with generating enough random catalogues to get a good estimation for $n_R(a,b,c)$.

7 Transformation into triplet counts

Our purpose is to determine the expansion of the three-point correlation function and we do not need $z(a,b,c)$ itself. Thus a simple transformation will solve the problems mentioned above. Let $P(a,b,c)$ be the density distribution function of triplets:

$$P(a,b,c) = \lim_{\Delta \to 0} \frac{n(a,b,c)}{\Delta^3} \,. \tag{16}$$

The estimator equation for $z(a,b,c)$ can be written in the following exact form:

$$z(a,b,c) = \frac{P_D(a,b,c)}{P_R(a,b,c)} - 1 - w_a - w_b - w_c \,. \tag{17}$$

Now we will rearrange this equation and integrate over the volume of a bin; thus an exact equation can be obtained for the data triplet counts independently of the size of the bins:

$$P_D(a,b,c) \;=\; P_R(a,b,c)\big(1 + w_a + w_b + w_c + z(a,b,c)\big)$$

$$n_D(a,b,c) = \iiint\limits_{V(a,b,c)} da\,db\,dc P_R(a,b,c)\Big(w_0 + w_1 + \sum_{\underline{k}} A_{\underline{k}} w_{\underline{k}}(a,b,c)\Big)$$
$$n_D(a,b,c) = \sum_{\underline{k}} B_{\underline{k}} n_{\underline{k}}(a,b,c)\,, \tag{18}$$

where

$$B_0 = A_0 + 1\,, \quad B_1 = A_1 + 1\,, \quad B_{\underline{k}} = A_{\underline{k}} \quad (\underline{k} \neq 0,1)\,,$$
$$n_{\underline{k}} = \iiint\limits_{V(a,b,c)} da\,db\,dc P_R(a,b,c) w_{\underline{k}}(a,b,c)\,. \tag{19}$$

Having determined the $n_{\underline{k}}(a,b,c)$ functions from Eqn. 19 we can fit them to $n_D(a,b,c)$ with the $B_{\underline{k}}$ parameters in Eqn. 2. From $B_{\underline{k}}$ we can calculate $A_{\underline{k}}$ and $Q_{\underline{k}}$ easily. For $n_{\underline{k}}$ we still need $P_R(a,b,c)$, *the probability density of random triplets*, and we have to *integrate* in Eqn. 19.

8 Distribution of random triplets

Since we use small separations, the non-linear effects of the boundary and the density gradient are negligible, so uniform distribution is a good approximation for the random catalogues. For a uniform distribution of points the distribution of triplets can be calculated analytically in a similar way to that which we used for random pairs. The result is

$$P_R(a,b,c) = A\overline{\eta^3} \frac{8\pi abc}{\sqrt{2\left(a^2b^2 + b^2c^2 + c^2a^2\right) - a^4 - b^4 - c^4}}\,, \tag{20}$$

where η is the surface density and A is the area of the sample.

The accurate coefficient in this equation for a finite area might be slightly smaller due to the presence of the boundaries. The coefficient is corrected after the integration of Eqn. 19.

Since $P_R(a,b,c)$ is singular at $a + b = c$, the terms in Eqn. 19 must be integrated carefully. We used a special Monte Carlo integration.

9 Fit to the triplet data

We determine $n_D(a,b,c)$ in 239 bins with size $\Delta = 5^\circ/12$, the sides of the triplets are $a \leq b \leq c \leq 5^\circ$. A weighted least-squares method is used for fitting in Eqn. 18:

$$\chi^2 = \frac{1}{239} \sum_{a,b,c} \frac{1}{\sigma_D^2(a,b,c)} \Big(n_D(a,b,c) - \sum_{\underline{k}} B_{\underline{k}} n_{\underline{k}}(a,b,c)\Big)^2\,, \tag{21}$$

where $\sigma_D(a,b,c)$ is the scatter of the triplet count in the $V(a,b,c)$ bin. Poisson errors $\sigma_D^2(a,b,c) = n_D(a,b,c)$ were assumed.

10 The results

When we fit all terms, the third order terms have very small coefficients. Neglecting them and fitting only up to second order we get:

$$Q_0 = -0.048 \quad Q_1 = 0.045 \quad \mathbf{Q_{11} = 0.94} \quad Q_2 = -0.002 \qquad (\chi^2 = 1.52)$$

Only the third term seems to be significant, so let us fit this one only:

$$\mathbf{Q_{11} = 0.93 \qquad (\chi^2 = 1.53)}$$

Note that reducing the number of fitted terms did not increase χ^2 too much. If we had only Poisson errors, χ^2 should be equal to 1.

11 Error estimate

Fitting on subsamples:

Q_{11} (hereafter Q) is determined for the NC and SC subsamples as well:

$$Q_{NC} = 0.31 \qquad\qquad Q_{SC} = 1.74$$

The general behaviour of the coefficients for the NC and SC sample is similar to the coefficients for the HL sample (the third order terms are less significant than the lower order terms, and the third coefficient is outstanding), but the results are different.

Perturbation of the triplet distribution:

We test whether the above mentioned discrepancy is due to the Poisson errors or not. The triplet counts are perturbed with 1σ noise and Q is calculated for every perturbed data set:

$$Q_{HL} = 0.91\ldots0.94 \quad Q_{NC} = 0.31\ldots0.4 \quad Q_{SC} = 1.7\ldots1.75$$

The lack of overlapping shows clearly that the discrepancy cannot be explained by Poisson errors.

The relation χ^2 vs. Q:

Figure 5 shows well that $Q \approx 1$ gives the best compromise with $\chi^2 \approx 3$ for the NC and SC samples.

Examination of the deviation from the fitted function:

We tested whether the rest (after subtracting the fitted function) has a Gaussian distribution. The rest is calculated for each bin first, then divided by $1\sigma = \sqrt{n_D(a,b,c)}$. The statistics of these numbers is shown in *Fig. 6*. Although the curve is flatter than the Gaussian belonging to $\chi^2 = 1$ (smooth curve), it is not too far from a Gaussian.

11 Conclusions

Our main result is that there seems to be a connection between the spatial three-point and two-point correlation functions for clusters similar to that for galaxies:

$$\zeta(r,s,u) \approx Q\big(\xi(r)\,\xi(s) + \xi(s)\,\xi(u) + \xi(u)\,\xi(r)\big) \ , \tag{22}$$

where $Q = 0.9{\pm}0.5$. There is a strong discrepancy between the northern and southern galactic caps. It might be due to the non-Poisson distribution of triplets, or due to the differences between the hemispheres in the completeness of the catalogue. If we accept that $Q = 1$ and if there is no cubic term in ζ, we have to conclude that

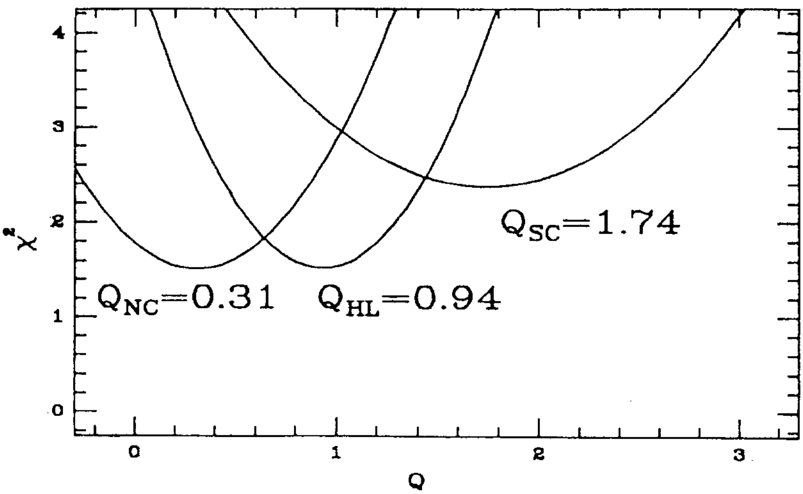

Fig. 5. Relation between χ^2 and Q for the three samples.

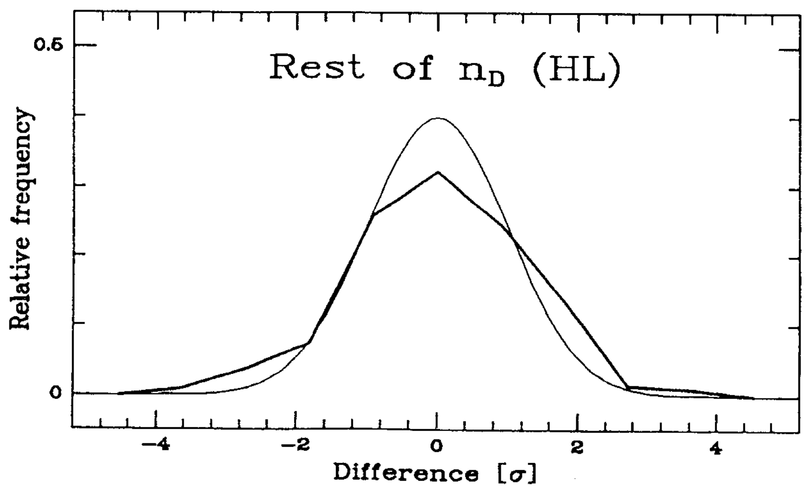

Fig. 6. Deviation of the triplet counts from the fit.

Gaussian biasing cannot be responsible for a major amplification of the correlations. These problems are under investigation (Hollósi *et al.* 1988).

Finally, we mention that we are planning to carry through the whole process for other cluster and galaxy catalogues, too.

References

Abell, G.O., 1958. *Astrophys. J. Suppl.*, **3**, 211.

Bardeen, J.R., Bond, J.R., Kaiser, N., Szalay, A.S., 1986. *Astrophys. J.*, **304**, 15.

Groth, E.J., Peebles, P.J.E., 1977. *Astrophys. J.*, **217**, 385.

Hollósi, J., Szalay, A.S., Tóth, G., 1988. In preparation.

Kaiser, N., 1984. *Astrophys. J.*, **284**, L9.

Peebles, P.J.E., 1980. *The Large Scale Structure of the Universe*, Princeton University Press, Princeton.

Politzer, D., Wise, M.B., 1984. *Astrophys. J.*, **285**, L1.

Struble, M.F., Rood, H.J., 1987. *Astrophys. J. Suppl.*, **63**, 543.

Szalay, A.S., 1988. *Astrophys. J.*, in press.