Mon. Not. R. astr. Soc. (1987) **227**, 921–931

Clustering of quasars from the ROE/ESO large-scale AQD survey

Roger G. Clowes *Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

Angela Iovino and Peter Shaver *European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-8046 Garching bei München, Federal Republic Germany*

Accepted 1987 March 20. Received 1987 March 18; in original form 1986 November 28

Summary. The new ROE/ESO large-scale AQD survey for quasars forms a connected area of ~200 deg^2 near the South Galactic Pole, and has resulted in the discovery of a total number of quasar candidates that is comparable to the number previously published from all other sources.

In this paper we describe a three-dimensional clustering analysis of ~1100 'high-probability' candidates occupying the assigned-redshift band of 1.8–2.4. The analysis is sensitive to very weak clustering – to a level of 7 per cent of our quasars occurring in pairs on scales $\sim 5h^{-1}$ Mpc – but none is found.

1 Introduction

One of the main difficulties of using quasars as cosmological indicators has, for many years, been the lack of large, complete samples. However, in the ROE/ESO survey for quasars, we have now successfully completed a very important stage in the assembly of such samples.

Briefly, seven low-dispersion, IIIa-J objective-prism plates from the UK Schmidt Telescope (UKST) have been measured with the COSMOS measuring machine (MacGillivray & Stobie 1984) at ROE, and the data processed with the software for Automated Quasar Detection (AQD – Clowes, Cooke & Beard 1984; Clowes 1986) to find the quasars. The seven plates form a connected area of ~200 deg^2 near the South Galactic Pole, and the total number of newly discovered quasar candidates is comparable to the number published from all other sources since the first quasars were discovered.

Objective-prism surveys remain one of the most efficient and neatest ways of finding quasars in large numbers over large areas of the sky, and the plates are particularly well exploited by automation of the discovery process: the ROE/ESO survey was carried out in only eight months. All varieties of survey, of course, have their own peculiar advantages (and disadvantages), but it is often overlooked that objective-prism surveys, in addition to the well-known selection by emission lines, also incorporate selection by the popular ultraviolet-excess method; and, moreover, with AQD, the ultraviolet excess refers to the continuum only, and so avoids biases from

emission lines. All of our candidates were selected by one or both of emission lines and ultraviolet excess.

With such a large number of new candidates the ROE/ESO survey ideally needs a long programme of follow-up spectroscopy both for confirmation and for obtaining accurate redshifts. In this paper on the clustering properties, however, we circumvent the present absence of such spectroscopy by restricting our samples to the 'high-probability' candidates from the five best plates.

The ability to grade the candidates in order of probability of being quasars is a very useful feature (it has something in common with the expert-system approach used in machine intelligence). It means that we can produce high-probability subsets with surface densities of $\sim 10\,\mathrm{deg}^{-2}$ – more than would be found in a typical visual search – with minimal need for access to large telescopes. There are, of course, many further quasars in the lower probability grades.

2 The quasars

The procedure used for the production of the high-probability subsets differs from that in Clowes (1986): a scheme for awarding points according to quasar-like properties is now used, and the sum of the points for a given object should be directly related to the probability of its being a quasar. One consequence of this scheme is that a high-probability candidate must have an emission line – so a redshift can always be assigned.

First, we produced the high-probability subsets in the following five fields.

ESO/SERC field	UKST plate	Plate centre (1950) (h m s)	(° ′ ″)	N1	N2	N3
295	UJ6536P	00 52 00	−40 00 00	399	328	284
296	UJ5406P	01 18 00	−40 00 00	219	213	171
297	UJ4514P	01 44 00	−40 00 00	407	380	308
351	UJ6528P	00 48 00	−35 00 00	275	261	208
411	UJ7307P	00 46 00	−30 00 00	216	205	168
Totals				1516	1387	1139

N1=number in the high-probability subset.
N2=number after rejection of obvious contaminants.
N3=number with assigned redshift between 1.8 and 2.4

Obvious contaminants were rejected visually; these are usually stars or surviving overlaps which, for special reasons, have slipped through the grading process.

Our objects cover $\sim 28.3\,\mathrm{deg}^2$ of each field. The precise limiting magnitudes are unknown because of the present lack of external data for calibration, but they are $B \sim 20.5$ for the best plates; consequently, no significance should be attached to the variations in surface density from field to field.

Redshifts were assigned on the assumption that the typically single lines are Lyα, and the objects then restricted to the range 1.8 to 2.4. The reason for the restriction is that the clustering analysis requires a comoving cuboid of points, which can only be defined for a small range of redshifts if there is to be no serious distortion of scales. A useful consequence is the avoidance of emission lines nearing the emulsion cut-off, which will have increasingly inaccurate wavelengths. The range 1.8 to 2.4 contains a large fraction of the objects.

The objective-prism redshifts have been compared with redshifts from published slit-spectroscopy. This shows that the objective-prism redshifts are systematically too small by ~0.04 at z~2.1; however, this offset is small, is irrelevant to the separations of pairs of quasars, and has been ignored. More importantly, the comparison yields a probability distribution for apparent separations that will result from a particular intrinsic separation: thus for an intrinsic separation of $\sim 5h^{-1}$ Mpc ($\Delta z=0.003$) the probability that the apparent separation will be $0h^{-1}$ Mpc is ~0.2 and the probability that it will be $\leqslant 33h^{-1}$ Mpc ($\Delta z=0.02$) is ~0.65. For comparison, the mean nearest-neighbour separation is $\sim 50h^{-1}$ Mpc. This probability distribution is used in the simulations which will be described in the next section.

3 The clustering analysis – 3D power spectrum analysis

We use the method of power spectrum analysis (PSA) to test for clustering. The method was developed by Webster (1976) with the particular aim of distinguishing a population that is weakly clustered from one that is randomly distributed.

A brief summary of the theory of PSA is given by Clowes (1986). Three-dimensional PSA is a more sensitive test than 2D PSA because, as Webster (1982) points out, the number of PSA statistics having $|\mathbf{k}|$ between k and $k+dk$, where $\mathbf{k}$ is the propagation vector, is then proportional to k^2 rather than k. By excluding the purely 1D and 2D propagation vectors (as determined by the special directions of the coordinate axes) from the 3D PSA the effects of any 1D and 2D artefacts in the data will be avoided. To obtain the third coordinate for the 3D PSA the cosmological hypothesis of redshifts is assumed and a particular cosmological model is adopted $-q_0=0$, $H_0=100h$ km s^{-1} Mpc^{-1}.

The way to define the comoving cuboid required by 3D PSA is given in Clowes (1986); its present-epoch volume is $\sim 390\times 390\times 1000h^{-3}$ Mpc^3 for each of the fields, the largest dimension being in the redshift direction.

The success of 3D PSA in detecting weak clustering is shown by random simulations of our fields. A level of clustering in which only 12 per cent of objects form pairs, in excess of random, of intrinsic separation $\sim 5h^{-1}$ Mpc, yields strong detections in all cases. Conversely, the detection of clustering in an intrinsically clustered population would still be possible with a very large proportion of randomly distributed contaminants.

In such a simulation 88 per cent of objects are distributed randomly and 12 per cent are forced to form randomly distributed pairs of intrinsic separation $\sim 5h^{-1}$ Mpc, with apparent separations generated according to the corresponding probability distribution. All following discussions of the sensitivity of the clustering analysis are in these terms.

At this point it is useful to have an estimate of what fraction of the objects are definitely quasars. Visual prejudice applied to a film copy of each plate and 1D plots of the objective-prism spectra suggests that the mean fraction corresponding to convincing quasars is ~80 per cent. This value is further supported by recent slit-spectroscopy of a small but representative selection.

Thus, we expect that our final samples for 3D PSA will contain a very low fraction of non-quasar contaminants, but we should also estimate in what fraction the assumption of Lyα is correct, and to what extent misidentified lines actually matter. Assumed Lyα could really be C IV λ1549, C III] λ1909, and Mg II λ2798, in which case the range 1.8–2.4 is really 1.20–1.67, 0.78–1.17, and 0.22–0.48 respectively. Hazard's (1985) histogram of redshifts is not ideal but gives the most relevant data: it suggests that the assumption will be correct in 30/43 or ~70 per cent of cases. Additional support for this value is given by comparison with existing spectroscopic observations (correct in 34/48 or ~71 per cent of cases).

A line-misidentification rate of 30 per cent will affect the detection of clustering. We have confirmed this by further simulations of our fields in which objects are paired, in excess of

random, at the 12 per cent level as before, but, instead of assuming all of the objects to be genuinely in the redshift range of 1.8 to 2.4, 30 per cent were assumed to be put there by line-misidentification. The significance of the detection of clustering generally falls and sometimes becomes indistinct; 12 per cent can therefore be adopted as a realistic threshold of detection. Of course, misidentified lines must lead to confusion of the scales of any clustering, but this is not apparent from the simulations.

The alternative to simulating the effects of the misidentified lines would be to identify the affected quasars and remove them. We will assume that equivalent widths are normally distributed with the parameters given by Wilkes (1986). Then essentially all the affected quasars could, in principle, be removed by setting a lower limit to the observed equivalent widths of ~180 Å; in practice, because of the errors on equivalent widths measured from objective-prism plates, the limit would need to be set at ~250 Å. In this way we would remove the affected quasars (~30 per cent of the total) and would simultaneously lose nearly 70 per cent (~50 per cent of the total) of the Lyα quasars, leaving only ~20 per cent of the total number of quasars.

4 Results of the 3D PSA

If we assume that we can detect clustering in a single field if only 12 per cent of objects are paired, and that 80 per cent of the objects are actually quasars, the other 20 per cent being randomly distributed contaminants, then we can detect clustering if ~15 per cent of all our quasars are paired. Furthermore, if we then combine the statistics from the five fields we can detect clustering if only ~7 per cent of all our quasars are paired.

Recall that 7 per cent in pairs refers to 7 per cent in pairs of intrinsic separation $\sim 5h^{-1}$ Mpc: that is, we are able to *detect* small-scale clustering at this level even though the actual scale of the clustering will be indeterminate.

Figs 1–5 show, for each of the five fields, the histograms of the 3D PSA x, y, z coordinates. Note that the origin of each set of x, y coordinates is the field centre, and that the x, y extents from the origin differ from field to field, although the total x, y extents are the same for all fields. The origin of each set of z coordinates corresponds to an assigned redshift of 2.1, and all the z extents are limited by the redshifts 1.8 and 2.4.

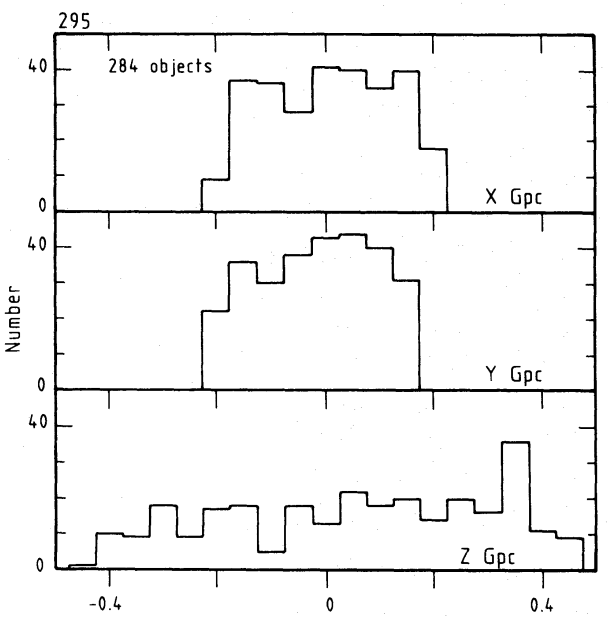

Figure 1. Histograms, for a bin size of 0.05 Gpc, of the 3D PSA x, y, z coordinates for field 295.

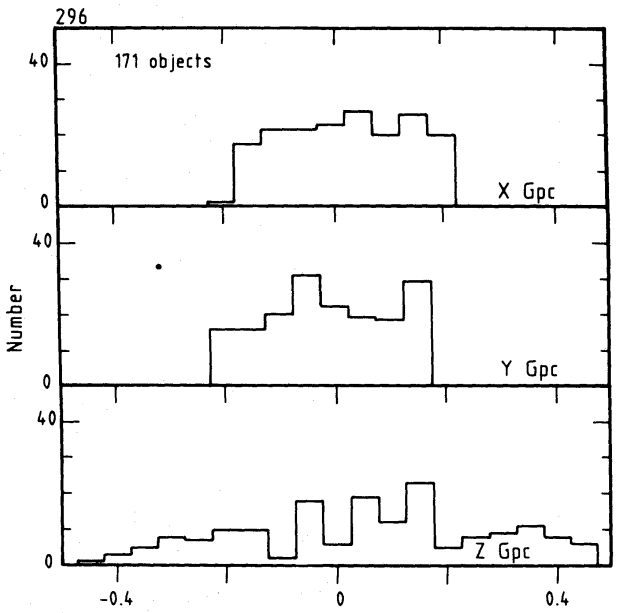

Figure 2. As Fig. 1, but for field 296.

Inspection of the histograms for gross features reveals nothing in x, y – the 'vignetting' losses are actually artefacts of the absolute limits of the data and the binning. The z histograms suggest a gross feature for field 296, in which there is a central concentration, and for field 411, in which there is a concentration at large z.

The appearance of small-scale features – peaks at $z=-0.15$, -0.05, 0.05, 0.15, 0.35 – is an artefact of the digitization of the locations of line peaks (i.e. pixels) and the binning. However, fields 295 and 297 seem to possess some excessively prominent peaks. Recall that the exclusion of the purely 1D and 2D terms from the 3D PSA avoids the effects of gross features, unless they coincide in some objects.

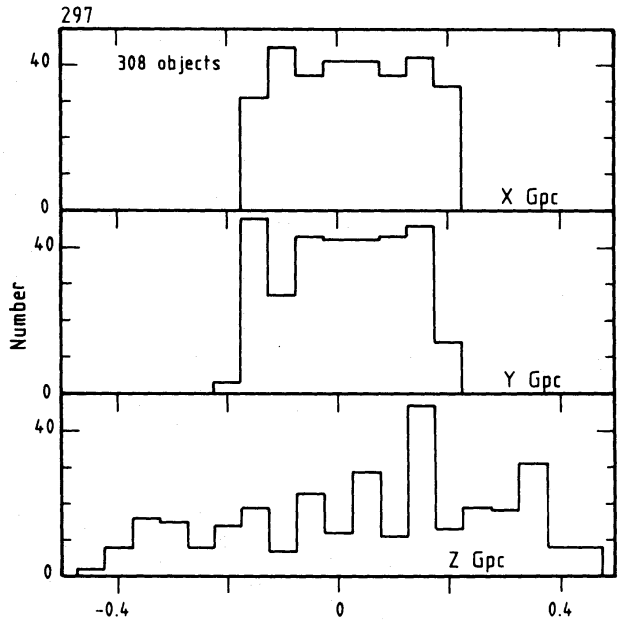

Figure 3. As Fig. 1, but for field 297.

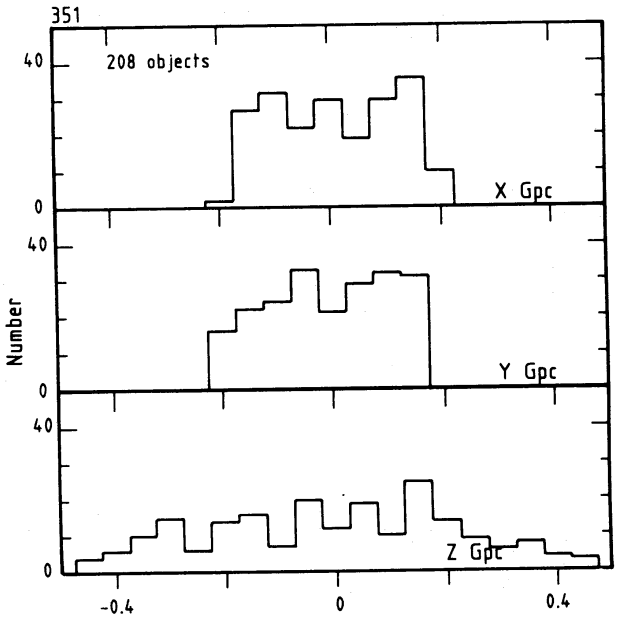

Figure 4. As Fig. 1, but for field 351.

Figs 6–10 show, for each of the five fields, the results of the 3D PSA, with the purely 1D and 2D terms excluded. The plots are of Q' against $1/\lambda$ for a bin size of 2.0 Gpc^{-1}, and the error bars are $\pm\sigma$ (erratum: the stated bin size in Clowes 1986 was accidentally transposed to 0.2 Gpc^{-1}). Recall from Webster (1976) and Clowes (1986) that if the plot of Q' against $1/\lambda$ suggests clustering on a scale of λ_c then the PSA statistics are combined for $\lambda>\lambda_c$ ($k<k_c$), giving Q, to establish the significance of the clustering and to estimate the mean number of objects per cluster. For a random distribution the expectation values of Q' and Q are 1.

Fig. 11 similarly shows a plot of Q' against $1/\lambda$, but this is for a Q' obtained by combining the PSA statistics from the five fields, thus giving the most sensitive test of clustering.

The requirement for a possible detection of clustering on a scale λ_c is taken to be a value of Q

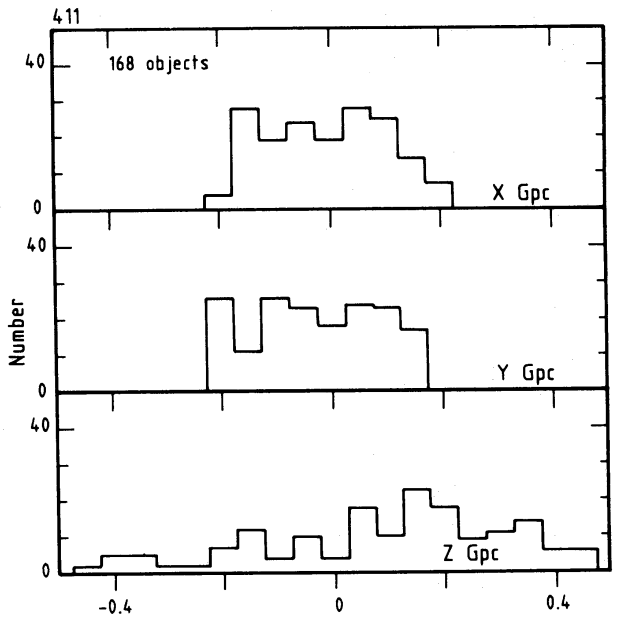

Figure 5. As Fig. 1, but for field 411.

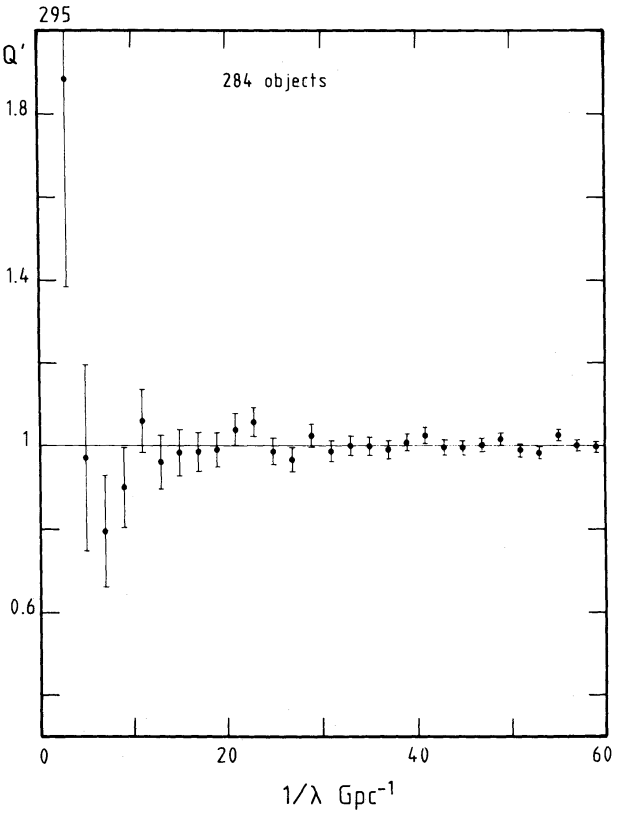

Figure 6. The 3D PSA plot of Q' against $1/\lambda$, with 1D and 2D terms excluded, for field 295. The bin size is 2.0 Gpc^{-1} and the error bars are $\pm\sigma$.

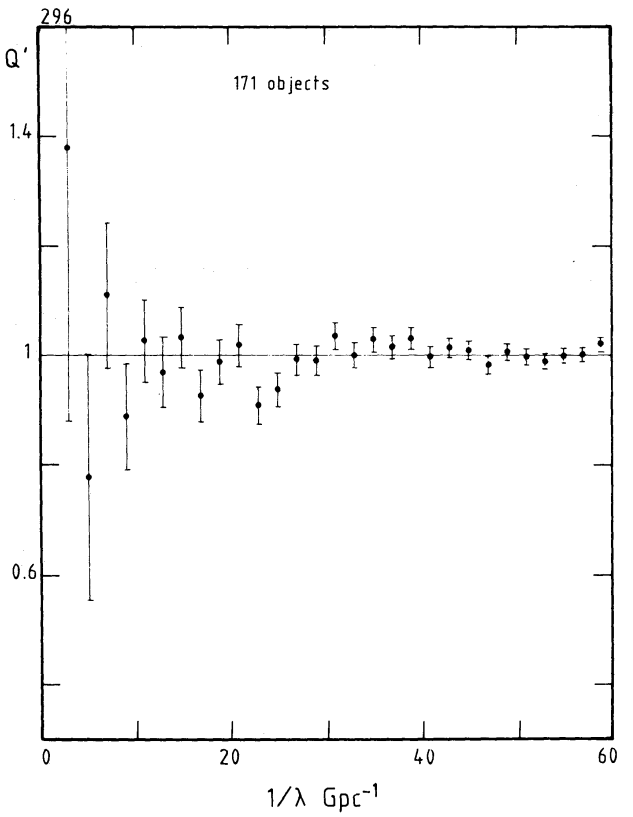

Figure 7. As Fig. 6, but for field 296.

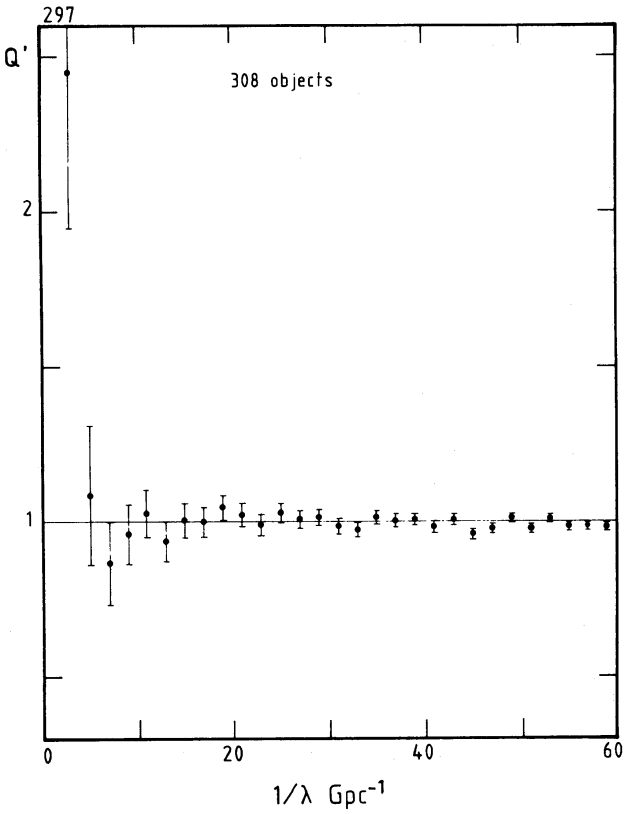

Figure 8. As Fig. 6, but for field 297.

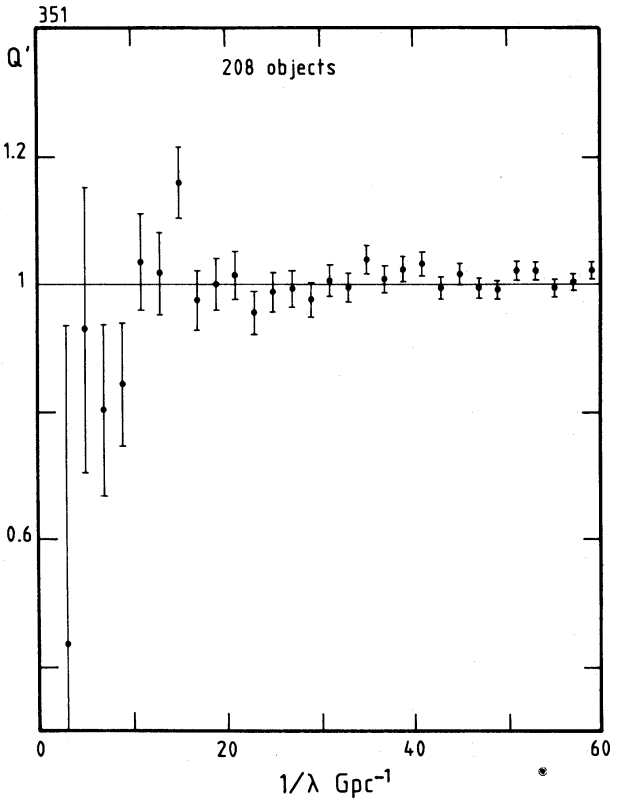

Figure 9. As Fig. 6, but for field 351.

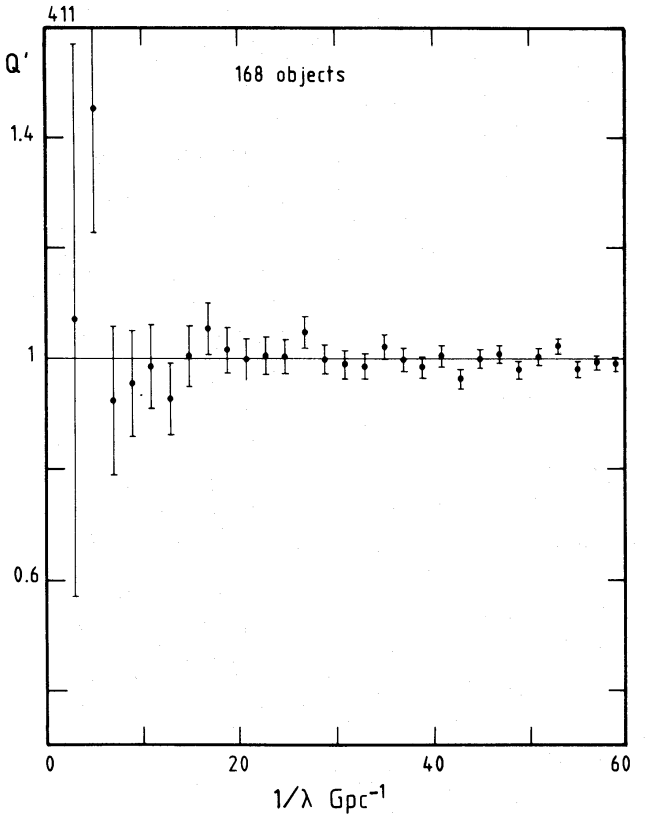

Figure 10. As Fig. 6, but for field 411.

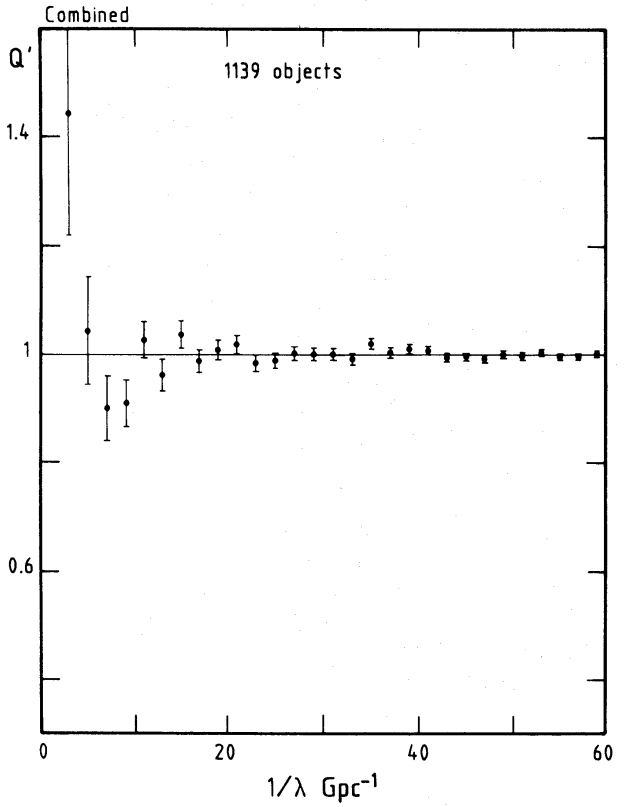

Figure 11. As Fig. 6, but with the statistics from the five fields combined.

that differs from 1 by at least 2.5σ. There are only two possible detections from the five fields (Figs 6–10): in field 296 at $1/\lambda_c \sim 25$ ($Q=0.962\pm0.015$), and in field 297 at $1/\lambda_c \sim 3$ ($Q=2.447\pm0.500$). The result in field 296 indicates anticlustering (i.e. a tendency towards the uniformity of a lattice structure), but Q differs from 1 by only 2.53σ, which is scarcely significant. The result in field 297 indicates clustering and is more significant – 2.89σ – but is contained by only one bin and occurs for the largest scale; a very plausible explanation is a 3D artefact.

More importantly, Fig. 11, for the combined statistics, shows no indication of clustering.

5 Comments

The results of this clustering analysis are that no single field shows any convincing evidence for physical clustering of the quasars, and that, in combination, the fields show convincing evidence for no clustering. If quasars are, in fact, physically clustered as pairs on the scale of $\sim 5h^{-1}$ Mpc (see, for example, Oort 1983) then we can deduce that the fraction of our quasars participating in pairs is less than ~7 per cent.

There have been other clustering analyses which cover the same small ($\sim 10h^{-1}$ Mpc) to intermediate ($\sim 100h^{-1}$ Mpc) scales that are considered here, and they have divided into results of no clustering (e.g. Osmer 1981; Webster 1982; Clowes 1986; Crampton, Cowley & Hartwick 1987) and results of possible clustering on small scales but not on intermediate scales (Shaver 1984; Boyle 1986). The analysis of clustering on large scales ($\gtrsim 1000h^{-1}$ Mpc) is more difficult because plate-to-plate artefacts are so important, but there are some indications that the random distribution at intermediate scales is not always extended to large scales (Clowes & Cannon, work in progress).

Our detection thresholds of ~15 per cent paired quasars for single fields and ~7 per cent for the combination of fields refer to the scales at which clustering is claimed by Shaver (1984) and Boyle (1986), and they incorporate allowances for non-quasar contaminants, misidentified lines and redshift errors. Thus, we are confident that the fraction of quasars in our samples that can be paired is less than ~7 per cent, and must now compare this result with the results of Shaver (1984) and Boyle (1986).

Shaver's (1984) approach is to use the published positions and redshifts of quasars that are collected in a large, inevitably heterogeneous, compilation (Véron-Cetty & Véron 1984). The heterogeneity makes comparison very difficult: the result could be real, and possibly attributable to low-redshift quasars only (Kruszewski 1986); it could, despite the allowances made, be caused by the selection effects of observation and publication that are inherent in a compilation.

Comparison with Boyle's (1986) result is rather easier. The quasars were discovered in an automated ultraviolet-excess survey, and the 3D clustering analysis, using the autocorrelation method, of 169 quasars had the advantage of spectroscopic redshifts. The clustering result implies that ~8 per cent of quasars are paired, in excess of random, on scales $\gtrsim 10h^{-1}$ Mpc ($q_0=0.5$), but it is also so noisy that a random distribution is not definitely excluded. A fraction of ~8 per cent would transform to ~1–2 per cent for our brighter quasars, and be below the detection threshold. (This transformation occurs because the fraction in pairs must decrease as the magnitude limit is set brighter.) The mean redshift of Boyle's quasars is ~1.5 compared with ~2.1 for ours, which could also be relevant.

Fewer than ~7 per cent of our quasars can be paired. To detect clustering still weaker than ~7 per cent the sensitivity of the analysis must be increased still further. In principle, this can be done quite easily by substituting spectroscopic redshifts for objective-prism redshifts. A quasar redshift survey would, in just one of our fields, improve the sensitivity to better than 3 per cent for pairs of maximum separation $10h^{-1}$ Mpc. There would also be the opportunity to investigate variations of clustering with epoch.

Acknowledgments

The staff of the COSMOS Unit at ROE and of the UK Schmidt Telescope Unit at ROE and Coonabarabran are thanked for their help. Most of the data processing used the STARLINK facilities at ROE.

References

Boyle, B. J., 1986. *PhD thesis,* University of Durham.
Clowes, R. G., 1986. *Mon. Not. R. astr. Soc.,* **218,** 139.
Clowes, R. G., Cooke, J. A. & Beard, S. M., 1984. *Mon. Not. R. astr. Soc.,* **207,** 99.
Crampton, D., Cowley, A. P. & Hartwick, F. D. A., 1987. *Astrophys. J.,* **314,** 129.
Hazard, C., 1985. In: *Active Galactic Nuclei,* ed. Dyson,. J. E., Manchester University Press.
Kruszewski, A., 1986. Preprint.
MacGillivray, H. T. & Stobie, R. S., 1984. *Vistas Astr.,* **27,** 433.
Oort, J. H., 1983. In: *Quasars and Gravitational Lenses,* 24th Liège Astrophysical Colloq., p. 301, Université de Liège.
Osmer, P. S., 1981. *Astrophys. J.,* **247,** 762.
Shaver, P. A., 1984. *Astr. Astrophys.,* **136,** L9.
Véron-Cetty, M.-P. & Véron, P., 1984. *ESO Scientific Report No. 1.*
Webster, A., 1976. *Mon. Not. R. astr. Soc.,* **175,** 61.
Webster, A., 1982. *Mon. Not. R. astr. Soc.,* **199,** 683.
Wilkes, B. J., 1986. *Mon. Not. R. astr. Soc.,* **218,** 331.