

# The 2dF Galaxy Redshift Survey: the dependence of galaxy clustering on luminosity and spectral type

Peder Norberg,<sup>1\*</sup> Carlton M. Baugh,<sup>1</sup> Ed Hawkins,<sup>2</sup> Steve Maddox,<sup>2</sup> Darren Madgwick,<sup>3</sup> Ofer Lahav,<sup>3</sup> Shaun Cole,<sup>1</sup> Carlos S. Frenk,<sup>1</sup> Ivan Baldry,<sup>4</sup> Joss Bland-Hawthorn,<sup>5</sup> Terry Bridges,<sup>5</sup> Russell Cannon,<sup>5</sup> Matthew Colless,<sup>6</sup> Chris Collins,<sup>7</sup> Warrick Couch,<sup>8</sup> Gavin Dalton,<sup>9</sup> Roberto De Propris,<sup>8</sup> Simon P. Driver,<sup>10</sup> George Efstathiou,<sup>3</sup> Richard S. Ellis,<sup>11</sup> Karl Glazebrook,<sup>4</sup> Carole Jackson,<sup>6</sup> Ian Lewis,<sup>5</sup> Stuart Lumsden,<sup>12</sup> John A. Peacock,<sup>13</sup> Bruce A. Peterson,<sup>6</sup> Will Sutherland<sup>9,13</sup> and Keith Taylor<sup>5,11</sup>

<sup>1</sup>*Department of Physics, University of Durham, South Road, Durham DH1 3LE*

<sup>2</sup>*School of Physics & Astronomy, University of Nottingham, Nottingham NG7 2RD*

<sup>3</sup>*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*

<sup>4</sup>*Department of Physics & Astronomy, Johns Hopkins University, Baltimore, MD 21218-2686, USA*

<sup>5</sup>*Anglo-Australian Observatory, PO Box 296, Epping, NSW 2121, Australia*

<sup>6</sup>*Research School of Astronomy & Astrophysics, The Australian National University, Weston Creek, ACT 2611, Australia*

<sup>7</sup>*Astrophysics Research Institute, Liverpool John Moores University, 12 Quays House, Birkenhead L14 1LD*

<sup>8</sup>*Department of Astrophysics, University of New South Wales, Sydney, NSW 2052, Australia*

<sup>9</sup>*Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH*

<sup>10</sup>*School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Fife KY6 9SS*

<sup>11</sup>*Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA*

<sup>12</sup>*Department of Physics, University of Leeds, Woodhouse Lane, Leeds LS2 9JT*

<sup>13</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*

Accepted 2002 January 17. Received 2002 January 9; in original form 2001 December 3

## ABSTRACT

We investigate the dependence of galaxy clustering on luminosity and spectral type using the 2dF Galaxy Redshift Survey (2dFGRS). Spectral types are assigned using the principal-component analysis of Madgwick et al. We divide the sample into two broad spectral classes: galaxies with strong emission lines ('late types') and more quiescent galaxies ('early types'). We measure the clustering in real space, free from any distortion of the clustering pattern owing to peculiar velocities, for a series of volume-limited samples. The projected correlation functions of both spectral types are well described by a power law for transverse separations in the range  $2 < (\sigma/h^{-1} \text{ Mpc}) < 15$ , with a marginally steeper slope for early types than late types. Both early and late types have approximately the same dependence of clustering strength on luminosity, with the clustering amplitude increasing by a factor of  $\sim 2.5$  between  $L^*$  and  $4L^*$ . At all luminosities, however, the correlation function amplitude for the early types is  $\sim 50$  per cent higher than that of the late types. These results support the view that luminosity, and not type, is the dominant factor in determining how the clustering strength of the whole galaxy population varies with luminosity.

**Key words:** methods: numerical – methods: statistical – galaxies: formation – large-scale structure of Universe.

## 1 INTRODUCTION

One of the major goals of large redshift surveys such as the 2-degree Field Galaxy Redshift Survey (2dFGRS) is to make an

accurate measurement of the spatial distribution of galaxies. The unprecedented size of the 2dFGRS makes it possible to quantify how the clustering signal depends on intrinsic galaxy properties, such as luminosity or star formation rate.

The motivation behind such a programme is to characterize the galaxy population and to provide constraints upon theoretical

\*E-mail: Peder.Norberg@durham.ac.uk

models of structure formation. In the current paradigm, galaxies form in dark matter haloes that are built up in a hierarchical way through mergers or by the accretion of smaller objects. The clustering pattern of galaxies is therefore determined by two processes: the spatial distribution of dark matter haloes and the manner in which dark matter haloes are populated by galaxies (Benson et al. 2000b; Peacock & Smith 2000; Seljak 2000; Berlind & Weinberg 2002). The evolution of clumping in the dark matter has been studied extensively using  $N$ -body simulations of the growth of density fluctuations via gravitational instability (e.g. Jenkins et al. 1998, 2001). With the development of powerful theoretical tools that can follow the formation and evolution of galaxies in the hierarchical scenario, the issue of how galaxies are apportioned amongst dark matter haloes can be addressed, and detailed predictions of the clustering of galaxies are now possible (Kauffmann, Nusser & Steinmetz 1997; Kauffmann et al. 1999; Benson et al. 2000a,b; Somerville et al. 2001).

The first attempt to quantify the difference between the clustering of early- and late-type galaxies was made using a shallow angular survey, the Uppsala catalogue, with morphological types assigned from visual examination of the photographic plates (Davis & Geller 1976). Elliptical galaxies were found to have a higher-amplitude angular correlation function than spiral galaxies. In addition, the slope of the correlation function of ellipticals was steeper than that of spiral galaxies at small angular separations. More recently, the comparison of clustering for different types has been extended to three dimensions using redshift surveys. Again, similar conclusions have been reached in these studies, namely that ellipticals have a stronger clustering amplitude than spirals (Lahav & Saslaw 1992; Santiago & Strauss 1992; Iovino et al. 1993; Hermit et al. 1996; Loveday et al. 1995; Guzzo et al. 1997; Willmer et al. 1998).

The subjective process of visual classification can now be superseded by objective, automated algorithms to quantify the shape of a galaxy. One recent example of such a scheme can be found in Zehavi et al. (2002), who measured a ‘concentration parameter’ for 30 000 galaxy images from the Sloan Digital Sky Survey, derived from the radii of different isophotes. Again, based upon cuts in the distribution of concentration parameter, early types are found to be more clustered than late types.

In this paper, we employ a different method to classify galaxies, based upon a principal-component analysis (PCA) of galaxy spectra, which is better suited to the 2dFGRS data (Madgwick et al. 2002). This technique has a number of attractive features. First, the PCA approach is completely objective and reproducible. An equivalent analysis can, for example, be applied readily to spectra produced by theoretical models of galaxy formation or to spectra obtained in an independent redshift survey. Secondly, the PCA can be applied over the full magnitude range of the 2dFGRS, whenever the spectra have a sufficiently high signal-to-noise ratio (see Section 2.2). For the 2dFGRS, the image quality is adequate to permit a visual determination of morphological type only for galaxies brighter than  $b_J \approx 17$ , which comprise a mere 5 per cent of the spectroscopic sample.

Two previous clustering studies have used spectral information to select galaxy samples. Loveday, Tresse & Maddox (1999) grouped galaxies in the Stromlo-APM redshift survey into three classes based upon the equivalent width of either the  $H\alpha$  or  $O\text{II}$  lines, and found that galaxies with prominent emission lines display weaker clustering than more quiescent galaxies. Tegmark & Bromley (1999) measured the relative bias between different spectral classes in the Las Campanas redshift survey (Shectman

et al. 1996), using a classification derived from PCA analysis (Bromley et al. 1998), and also found that early spectral types are more strongly clustered than late spectral types. [See also Blanton (2000) for a revision of Tegmark & Bromley’s analysis, which takes into account the effect of errors in the survey selection function.]

Here, we use the 2dFGRS survey to measure the dependence of galaxy clustering jointly on luminosity and spectral type, adding an extra dimension to the analysis carried out by Norberg et al. (2001). Previously, a pioneering study of bivariate galaxy clustering, in terms of luminosity and morphological type, was undertaken using the Stromlo-APM redshift survey (Loveday et al. 1995). To place the analysis presented here in context, the samples that we consider cover a larger volume and, despite being volume-limited (see Section 2.4), typically contain over an order of magnitude more galaxies than those available to Loveday et al.

We give a brief overview of the 2dFGRS in Section 2, along with details of the spectral classification and an explanation of how the samples used in the clustering analysis were constructed. The estimation of the redshift-space correlation function and its real-space counterpart, the projected correlation function, are outlined in Section 3. A brief overview of the clustering of 2dFGRS galaxies in redshift space, selected by luminosity and spectral type, is given in Section 4; a more detailed analysis of the redshift-space clustering can be found in Hawkins et al. (2002, in preparation). We present the main results of the paper in Section 5 and give our conclusions in Section 6.

## 2 THE DATA

### 2.1 The 2dFGRS sample

Detailed descriptions of the construction of the 2dFGRS and its properties are given by Colless et al. (2001). In summary, galaxies are selected down to a magnitude limit of  $b_J \approx 19.45$  from the APM Galaxy Survey (Maddox et al. 1990a,b, 1996; Norberg 2002). The sample considered in this paper consists of over 160 000 redshifts measured prior to 2001 May. We focus our attention on the two large contiguous volumes of the survey, one centred on the Southern Galactic Pole (hereafter SGP) and the other close to the direction of the Northern Galactic Pole (NGP).

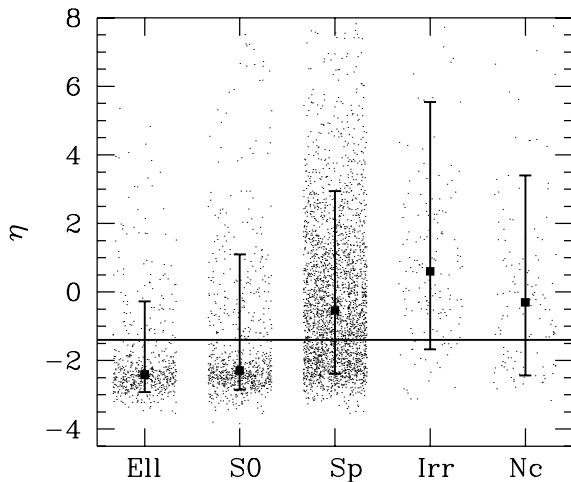
### 2.2 Spectral classification of 2dFGRS galaxies

The spectral properties of 2dFGRS galaxies are characterized using the principal-component analysis described by Madgwick et al. (2002). This analysis makes use of the spectral information in the rest-frame wavelength range 3700–6650 Å, thereby including all the major optical diagnostics between  $O\text{II}$  and  $H\alpha$ . For galaxies with  $z > 0.15$ , sky absorption bands contaminate the  $H\alpha$  line. Since this can affect the stability of the classification, we restrict our analysis to galaxies with  $z < 0.15$  following Madgwick et al. (2002).

The 2dFGRS spectra are classified by a single parameter,  $\eta$ , which is a linear combination of the first and second principal components. This combination has been chosen specifically to isolate the relative strength of emission and absorption lines present in the spectrum of each galaxy, thereby providing a diagnostic that is robust to the instrumental uncertainties that affect the calibration of the continuum. Physically, this parameter is related to the star formation rate in a galaxy, as is apparent from the tight correlation of  $\eta$  with the equivalent width of  $H\alpha$  in emission-line galaxies (Bland-Hawthorn et al. 2002, in preparation).

In this paper, we divide the 2dFGRS sample into two broad, distinct classes: galaxies with spectra for which the PCA returns  $\eta < -1.4$ , and which we refer to, for the sake of brevity, as early type, and galaxies with  $\eta > -1.4$ , which we call late type. The distribution of  $\eta$  for 2dFGRS spectra displays a shoulder feature at this value (see fig. 4 of Madgwick et al. 2002).

The spectral type of a galaxy, as given by the value of  $\eta$ , clearly depends upon its physical properties and is therefore a useful and effective way in which to label galaxies. Nevertheless, it is still instructive to see how well, if at all,  $\eta$  correlates with the morphological type assigned in a subjective fashion from a galaxy image. Madgwick et al. (2002) show that there is a reasonable correspondence between  $\eta$  and morphological type, using high signal-to-noise ratio spectra and photometry taken from Kennicutt (1992);  $\eta \approx -1.4$  approximately delineates the transition between early and late morphological types in  $b_j$ . We revisit the comparison between classifications based on spectral and morphological types in Fig. 1, this time using 2dFGRS spectra and UK Schmidt images. The horizontal axis shows the morphological type assigned to a subset of bright APM galaxies by Loveday (1996). Although there is a substantial amount of scatter in the  $\eta$  values of spectra that lie within a given morphological class, it is reassuring to see that the median  $\eta$  does correlate with the morphological class. Moreover, the median  $\eta$  values match up well with the broad division that we employ to separate early and late types. Galaxies denoted ‘early type’ on the basis of their morphology have a median  $\eta$  that is smaller than our fiducial value of  $\eta = -1.4$  and vice versa for late types. In practice, for the samples analysed in this paper, the correspondence between morphological type and spectral class will be better than suggested by Fig. 1. This is because the sample used in the comparison in Fig. 1 consists of nearby extended galaxies, and so the distribution of spectral types is distorted somewhat by aperture effects (see, e.g., Kochanek, Pahre & Falco 2002; Madgwick et al. 2002). This effect arises because the fibres



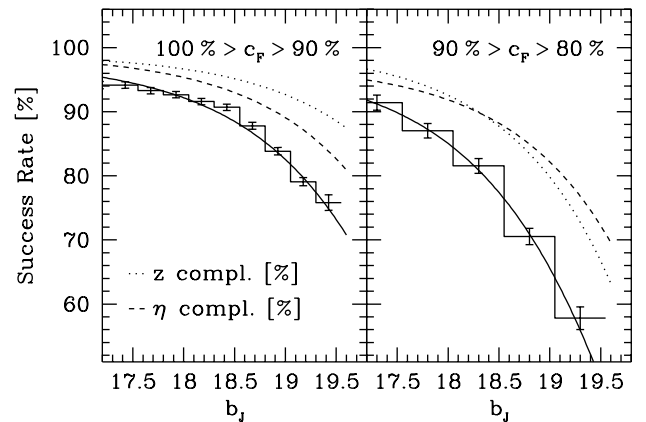
**Figure 1.** A comparison between the morphological classification of bright ( $b_j < 17.0$ ) APM galaxies by Loveday (1996) with the 2dFGRS spectral classification, as quantified by the continuous variable  $\eta$  (see text and Madgwick et al. 2002 for the definition). The morphological classification distinguishes between elliptical (Ell), lenticular (S0), spiral (Sp) and irregular (Irr) galaxies. All galaxies with both a morphological classification and a spectral classification are plotted. The non-classified (Nc) class contains objects for which morphological classification was attempted but for which Loveday was unable to assign a type. The squares show the median value of  $\eta$  for each morphological class defined by Loveday, and the error bars show the 10–90 percentile range.

used to collect the galaxy spectra are of finite size (subtending 2 arcsec on the sky). For this reason, when we measure the spectrum of a nearby galaxy it is possible that a disproportionate amount of light will be sampled from the bulge, thereby making the galaxy appear systematically redder or ‘earlier’ in type. We find that this effect is only significant for the most nearby galaxies ( $z < 0.05$ ) and should be completely negligible beyond  $z \sim 0.1$  (Madgwick et al. 2002).

### 2.3 Sample selection

In order to construct an optimal sample for the measurement of the two-point correlation function, we select regions with high completeness in terms of measured redshifts, using a redshift completeness mask for the 2dFGRS, similar to that described in Colless et al. (2001; see also Norberg et al. 2002). Such a mask is required because of the tiling strategy adopted to make the best use of the allocated telescope time, along with the fact that the survey is not yet finished. An additional consideration is the success rate with which spectral types have been assigned to galaxies, which depends upon the signal-to-noise ratio of the galaxy spectrum.

In Fig. 2, we show histograms of the spectral classification success rate for two different ranges of field completeness,  $c_F$ , which is defined as the ratio of the number of measured redshifts in a given 2dF field to the number of targets. The spectral classification success rate has two contributions. The first of these is the redshift completeness, shown by the dotted curve. This incompleteness arises because we do not always succeed in measuring a redshift for a targeted object. The redshift incompleteness is necessarily small for the high-completeness fields contributing to the histograms. The second contribution is the spectral classification completeness. Galaxies do not receive a spectral classification when a redshift is measured with  $z \leq 0.15$  (and is therefore within the redshift range over which the PCA can be carried out),



**Figure 2.** The histograms, plotted with Poisson error bars, show the success rate for assigning a spectral type to a targeted galaxy as a function of apparent magnitude. Two field completeness ( $c_F$ , defined in text) ranges are shown, as indicated by the values at the top of each panel. The redshift completeness, i.e. the fraction of targeted galaxies for which a redshift is measured, is shown by the dotted lines. The spectral classification completeness, i.e. the fraction of galaxies with measured redshifts below  $z = 0.15$  that have spectra of sufficient signal-to-noise ratio to be used in the PCA, is shown by the dashed lines. (The dotted and dashed curves show parametric fits to the inferred redshift completeness and spectral classification completeness, respectively.) The model for the spectral classification success rate, shown by the solid lines, is the product of the dotted and dashed lines in each panel, and is a good fit to the histogram in each case.

but the spectrum has too small a signal-to-noise ratio for the PCA to be applied successfully (typically signal-to-noise (S/N) ratio  $< 10$ ). The spectral classification success rate is given by the product of these two contributions. Our model for this effect, plotted as the solid curves in each panel of Fig. 2, is in good agreement with the success rate realized in the 2dFGRS, shown by the histograms.

Rather than weight the data to compensate for a spectral classification success rate below 100 per cent, we instead modulate the number of unclustered or random points laid down in each field in the clustering analysis to take into account the varying success rate. We have conducted a number of tests in which we varied the completeness thresholds used, adopted different weighting schemes using samples of higher completeness, and we have also compared our results with those from Norberg et al. (2001), whose samples are not subject to spectral classification incompleteness. The results of these tests confirm that our clustering measurements are robust to changes to the details of how we model the incompleteness; this is largely a result of our practice of restricting the analysis to high-completeness fields. Excluding areas below our relatively high sector completeness threshold (see Colless et al. 2001 for a definition), we estimate that the effective solid angle used in the SGP region is  $\sim 380 \text{ deg}^2$ , and in the NGP  $\sim 250 \text{ deg}^2$ .

#### 2.4 Constructing a volume-limited sample

We analyse a series of volume-limited samples drawn from the 2dFGRS, following the strategy adopted by Norberg et al. (2001). The chief advantage of this approach is simplicity; the radial distribution of galaxies is uniform apart from modulations in space density owing to clustering. Therefore, the complication of modelling the radial selection function in a flux-limited survey is avoided. This is particularly appealing for the current analysis, as separate selection functions would be required for each class of spectral type studied, since Madgwick et al. (2002) have demonstrated that galaxies with different spectral types have different luminosity functions.

The disadvantage of using volume-limited samples is that a large fraction of galaxies in the flux-limited catalogue do not satisfy the selection criteria. As Norberg et al. (2001) point out, a volume-limited sample specified by a range in absolute magnitude has both a lower ( $z_{\min}$ ) and an upper redshift cut ( $z_{\max}$ ), because the flux-limited catalogue has, in practice, bright and faint apparent magnitude limits. This seemingly profligate use of galaxy redshifts was a serious problem for previous generations of redshifts surveys. This is not, however, the case for the 2dFGRS, which contains sufficient galaxies to permit the construction of large volume-limited samples defined both by luminosity and spectral type. As we demonstrate in Section 5, the volume-limited samples we analyse are large enough, both in terms of volume and number of galaxies, to give extremely robust clustering measurements.

To construct a volume-limited sample, it is necessary to estimate the absolute magnitude that each galaxy would have at  $z = 0$ . This requires assumptions concerning the variation of galaxy luminosity with wavelength and with redshift, or equivalently, with cosmic time. We make use of the class-dependent  $k$ -corrections derived by Madgwick et al. (2002). The mean weighted  $k$ -corrections are given by the following expressions:

$$k(z) = 2.6z + 4.3z^2 \quad (\text{early types}) \quad (1)$$

$$k(z) = 1.5z + 2.1z^2 \quad (\text{late types}) \quad (2)$$

$$k(z) = 1.9z + 2.7z^2 \quad (\text{full sample}). \quad (3)$$

These  $k$ -corrections have the appeal that they are extracted directly in a self-consistent way from 2dFGRS spectra. However, no account is taken of evolution in the galaxy spectrum. The explicit inclusion of evolution could lead to the ambiguous situation whereby the spectral type of a galaxy changes with redshift. We have checked that our results are, in fact, insensitive to the precise choice of  $k$ -correction, comparing clustering results obtained with the spectral-type-dependent  $k$ -corrections given above with those obtained when a global  $k + e$ -correction (i.e. making an explicit attempt to account for galaxy evolution, albeit in an average sense) is applied (as in Norberg et al. 2001).

Since the  $k$ -corrections are class dependent, the  $z_{\min}$  and  $z_{\max}$  values corresponding to a given absolute magnitude range are also slightly class dependent. Hence, the volumes defining the samples for two different spectral classes for the same bin in absolute magnitude will not coincide exactly. In addition to this subtle class-dependent definition of the volumes, the values of  $z_{\min}$  and  $z_{\max}$  vary slightly with position on the sky. This is a result of revisions made to the map of galactic extinction (Schlegel, Finkbeiner & Davis 1998) and to the charge-coupled device (CCD) recalibration of APM plate zero-points since the definition of the original input catalogue.

Finally, throughout the paper, we adopt an  $\Omega_0 = 0.3$ ,  $\Lambda_0 = 0.7$  cosmology to convert the redshift into a comoving distance. The relative clustering strength of our samples is insensitive to this choice.

### 3 ESTIMATING THE TWO-POINT CORRELATION FUNCTION

The galaxy correlation function is estimated on a two-dimensional grid of pair separation parallel ( $\pi$ ) and perpendicular ( $\sigma$ ) to the line of sight. To estimate the mean density of galaxy pairs, a catalogue of randomly positioned points is generated with the same angular distribution and the same values of  $z_{\min}$  and  $z_{\max}$  as the data, taking into account the completeness of the survey as a function of position on the sky, as described in Section 2.3. The correlation function is estimated using

$$\xi_H = \frac{DDRR}{DR^2} - 1, \quad (4)$$

where  $DD$ ,  $DR$  and  $RR$  are the numbers of data–data, data–random and random–random pairs, respectively, in each bin (Hamilton 1993). This estimator does not require an explicit estimate of the mean galaxy density. We have also cross-checked our results using the estimator proposed by Landy & Szalay (1993):

$$\xi_{LS} = \frac{DD - 2DR + RR}{RR}, \quad (5)$$

where, this time  $DD$ ,  $DR$  and  $RR$  are the suitably normalized numbers of data–data, data–random and random–random pairs. We find that the two estimators give the same results over the range of pair separations in which we are interested.

The clustering pattern of galaxies is distorted when radial positions are inferred from redshifts, as expected in the gravitational instability theory of structure formation (e.g. Kaiser 1987; Cole, Fisher & Weinberg 1994). Clear evidence for this effect is seen in the shape of the two-point correlation function when plotted as  $\xi(\sigma, \pi)$ , as demonstrated clearly for galaxies in the 2dFGRS by Peacock et al. (2001) and for groups of galaxies in the Zwicky catalogue by Padilla et al. (2001). After giving a brief flavour of the clustering of 2dFGRS galaxies in redshift space in

Section 4, we focus our attention on clustering in real space in the remainder of the paper. The clustering signal in real space is inferred by integrating  $\xi(\sigma, \pi)$  in the  $\pi$ -direction (i.e. along the line of sight):

$$\frac{\Xi(\sigma)}{\sigma} = \frac{1}{\sigma} \int_{-\infty}^{\infty} \xi(\sigma, \pi) d\pi. \quad (6)$$

For the samples that we consider, the integral converges by pair separations of  $\pi \geq 50 h^{-1}$  Mpc. The projected correlation function can then be written as an integral over the spherically averaged real-space correlation function,  $\xi(r)$ ,

$$\frac{\Xi(\sigma)}{\sigma} = \frac{2}{\sigma} \int_{\sigma}^{\infty} \xi(r) \frac{r dr}{(r^2 - \sigma^2)^{1/2}}, \quad (7)$$

(Davis & Peebles 1983). If we assume that the real-space correlation function is a power law (which is a fair approximation for APM galaxies out to separations around  $r \sim 10 h^{-1}$  Mpc, see, e.g., Baugh 1996), then equation (7) can be written as

$$\frac{\Xi(\sigma)}{\sigma} = \left(\frac{r_0}{\sigma}\right)^{\gamma} \frac{\Gamma(1/2)\Gamma([\gamma - 1]/2)}{\Gamma(\gamma/2)} = \left(\frac{r_0}{\sigma}\right)^{\gamma} A(\gamma), \quad (8)$$

where  $\Gamma(x)$  is the usual Gamma function, and we have used  $\xi(r) = (r_0/r)^{\gamma}$ , where  $r_0$  is the real-space correlation length and  $\gamma$  is equal to the slope of the projected correlation function  $\Xi(\sigma)/\sigma$ . As we demonstrate in Section 4.2, the projected correlation function is well described by a power law.

We study a range of samples containing different numbers of galaxies and covering different volumes of the Universe. It is imperative to include sampling fluctuations when estimating the errors on the measured correlation function, to allow a meaningful comparison of the results obtained from different samples. This contribution to the errors has often been neglected in previous work. Following Norberg et al. (2001), we employ a sample of 22 mock 2dFGRS catalogues drawn from the  $\Lambda$  cold dark matter ( $\Lambda$ CDM) Hubble volume simulation (Evrard et al. 2002) to estimate the error bars on the measured correlation functions. The construction of these mock catalogues is explained in Baugh et al. (in preparation; see also Cole et al. 1998; Norberg et al. 2002). These catalogues have the same selection criteria and the same clustering amplitude as measured for galaxies in the flux-limited 2dFGRS. We have experimented with ensembles of mock catalogues constructed with different clustering strengths to ascertain how best to assign error bars when the measured clustering has a different amplitude from that of our fiducial sample of 22 2dFGRS mocks. We found that the error bars obtained directly by averaging over a test ensemble of mocks are reproduced most closely by using the scaled fractional rms scatter derived from the fiducial ensemble of 22 mocks, rather than by taking the absolute error.

## 4 CLUSTERING IN REDSHIFT SPACE

In this section we give a brief overview of the clustering of 2dFGRS galaxies in redshift space, for samples selected by luminosity and spectral type. First, in Section 4.1, we give a qualitative impression of the clustering differences by plotting the spatial distribution of galaxies in volume-limited samples. Then we quantify these differences by measuring the spherically averaged correlation function,  $\xi(s)$ . A more comprehensive analysis of the clustering of 2dFGRS galaxies in redshift space will be presented by Hawkins et al. (in preparation).

### 4.1 Spatial distribution of 2dFGRS galaxies

It is instructive to gain a visual impression of the spatial distribution of 2dFGRS galaxies before interpreting the measured correlation functions. In Fig. 3, we show the spatial distribution of galaxies in two ranges of absolute magnitude: in the upper panel we show a sample of faint galaxies ( $-18.0 \geq M_{b_j} - 5 \log_{10} h \geq -19.0$ ) and in the lower panel a sample of bright galaxies ( $-20.0 \geq M_{b_j} - 5 \log_{10} h \geq -21.0$ ). Within each panel, early- and late-type galaxies, as distinguished by their spectral types, are plotted with different symbols; the positions of early types are indicated by circles and the late types are marked by stars. For clarity, we show only a  $3^\circ$  slice in declination cut from the SGP region and we have sparsely sampled the galaxies, so that the space densities of the two spectral classes are equal. In order to expand the scale of the plot, the range of redshifts shown is restricted, taking a subset of the full volume-limited sample in each case. (Note also that the redshift ranges differ between the two panels.)

A hierarchy of structures is readily apparent in these plots, ranging from isolated objects, to groups of a handful of galaxies and on through to rich clusters containing over 100 members. It is interesting to see how structures are traced by galaxies in the different luminosity bins by comparing common structures between the two panels. For example, the prominent structure (possibly a supercluster of galaxies) seen at  $\alpha \approx 0^h$  and  $z \approx 0.061$  is clearly visible in both panels. The same is true for the overdensity seen at  $\alpha \approx 03^h 15^m$  at  $z \approx 0.068$ .

This is the first time that a large enough survey has been available, both in terms of the volume spanned and the number of measured redshifts, to allow a comparison of the clustering of galaxies of different spectral types and luminosities in representative volume-limited samples, without the complication of the strong radial gradient in number density seen in flux-limited samples.

It is apparent from a comparison of the distribution of the different spectral types in the upper panel of Fig. 3, that the faint early-type galaxies tend to be grouped into structures on small scales, whereas the faint late types are more spread out. One would therefore anticipate that the early types should have a stronger clustering amplitude than the late types, an expectation that is borne out in Section 4.2.

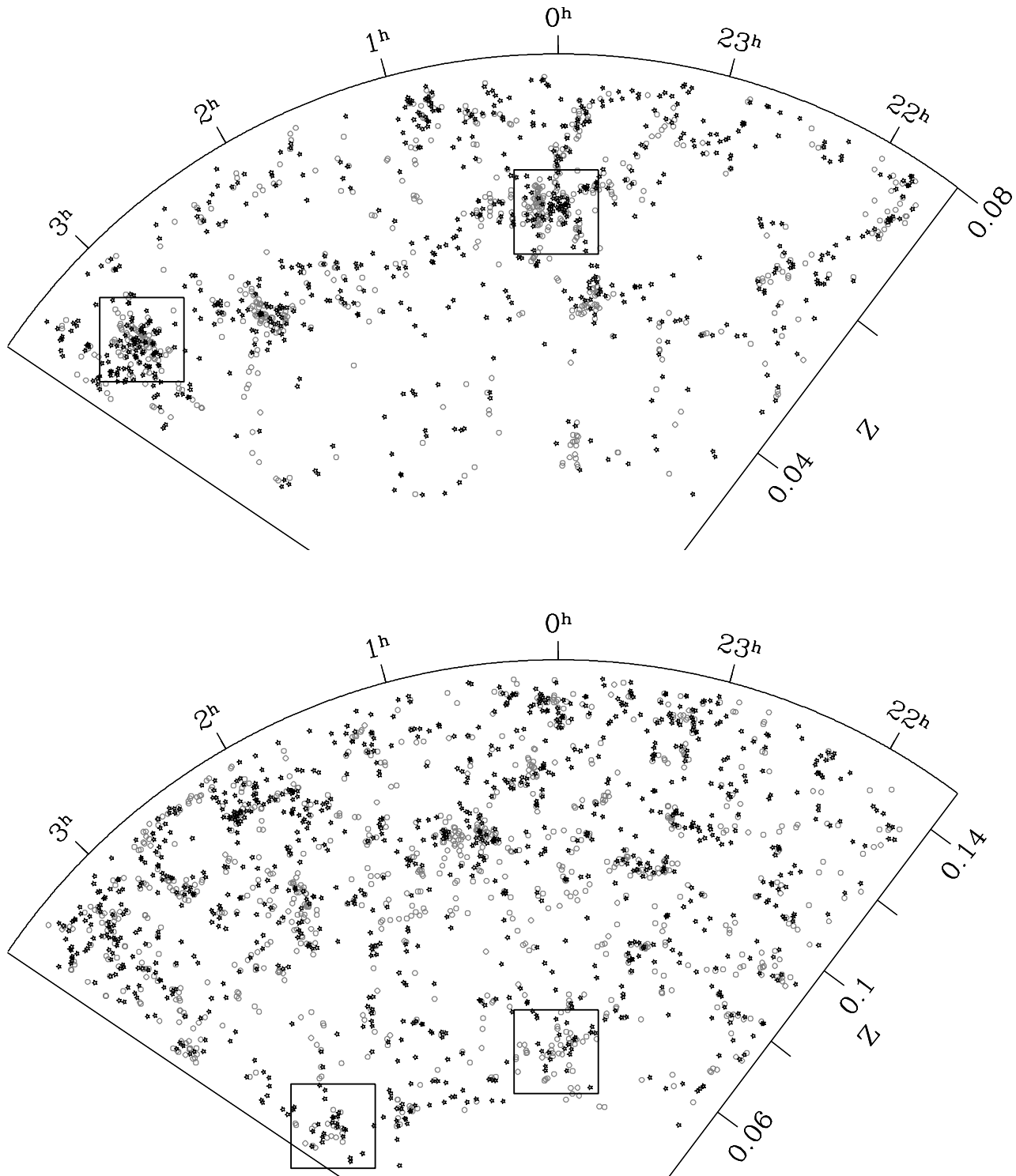
In the lower panel of Fig. 3, the distinction between the distribution of the spectral types is less apparent. This is partly caused by the greater importance of projection effects in the declination direction, as the cone extends to a greater redshift than in the upper panel of Fig. 3. However, close examination of the largest structures suggests that early types are more abundant in them than late types, again implying a stronger clustering amplitude.

### 4.2 $\xi(s)$ as a function of luminosity and spectral type

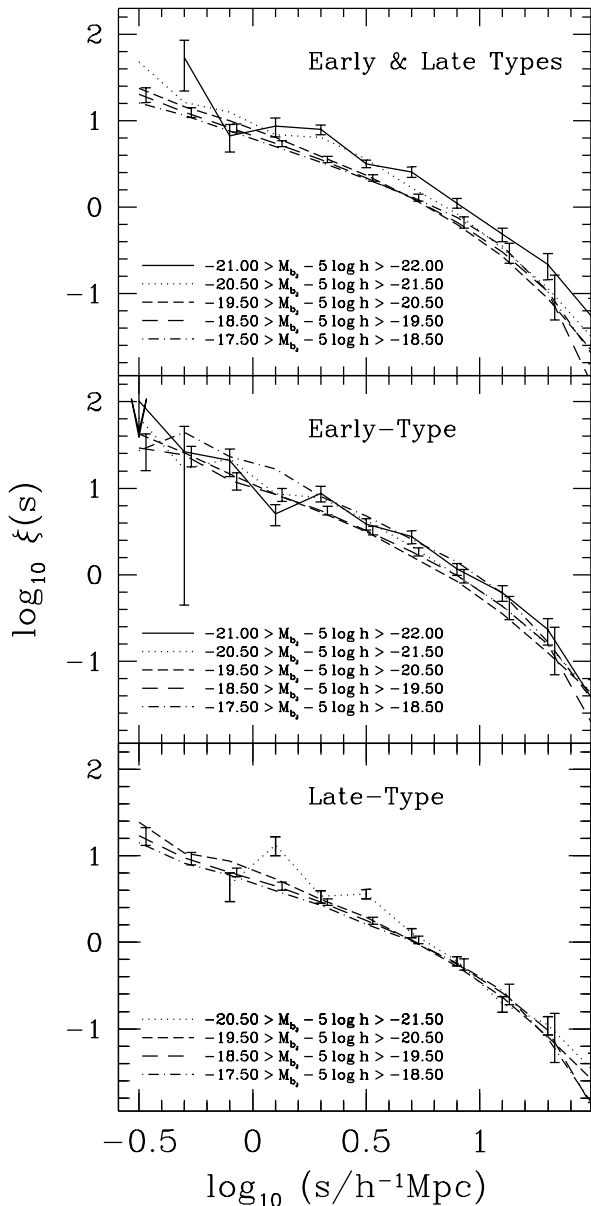
In Fig. 4, we show the spherically averaged redshift-space correlation function,  $\xi(s)$ , as a function of luminosity and spectral type. Results are shown for samples selected in bins of width 1 mag, as indicated by the legend in each panel. The top panel shows the correlation functions of all galaxies that have been assigned a spectral type, the middle panel shows results for galaxies classified as early types ( $\eta < -1.4$ ) and the bottom panel shows results for late types ( $\eta > -1.4$ ). Note that, at present, there are insufficient numbers of late-type galaxies to permit a reliable measurement of the correlation function for the brightest magnitude bin  $-21.0 \geq M_{b_j} - 5 \log_{10} h \geq -22.0$ .

Several deductions can be made immediately from Fig. 4. In all cases, the redshift-space correlation function is well described by a power law *only* over a fairly limited range of scales. The correlation functions of early-type galaxies are somewhat steeper than those of late types. However, the main difference is that the early-type galaxies have a stronger clustering amplitude than the

late-type galaxies. The correlation length, defined here as the pair separation for which  $\xi(s_0) = 1$ , varies for early types from  $s_0 = 7.1 \pm 0.7 h^{-1}$  Mpc for galaxies with absolute magnitudes around  $M_{b_j} - 5 \log_{10} h \sim -19.5$  to  $s_0 = 8.9 \pm 0.7 h^{-1}$  Mpc for the brightest sample with  $-21.0 \geq M_{b_j} - 5 \log_{10} h \geq -22.0$ . The faintest early types, with magnitudes  $-17.5 \geq M_{b_j} - 5 \log_{10} h \geq -18.5$ ,



**Figure 3.** The spatial distribution of 2dFGRS galaxies in the SGP. The panels show the redshift and right ascension of galaxies in a  $3^\circ$ -thick strip in declination for different magnitude ranges. To expand the scale of the plot, the redshift range shown has been restricted; note that the redshift scales are different in the two panels. Upper panel:  $-18.0 \geq M_{b_j} - 5 \log_{10} h \geq -19.0$ ; lower panel:  $-20.0 \geq M_{b_j} - 5 \log_{10} h \geq -21.0$ . Stars mark the locations of late-type galaxies and circles the positions of early-type galaxies. The boxes show the two structures that are referred to in the text.



**Figure 4.** The spherically averaged redshift-space correlation function of galaxies in disjoint absolute magnitude bins, as indicated by the key in each panel. The panels show the results for different samples: the top panel shows the correlation functions for all galaxies that have been assigned a spectral type, the middle panel shows the clustering of galaxies with  $\eta < -1.4$  and the lower panel shows  $\xi(s)$  for galaxies with  $\eta > -1.4$ . The error bars are obtained using 2dFGRS mock catalogues, as described in the text. For clarity, error bars are plotted only on the  $-18.5 \geq M_{b_1} - 5 \log_{10} h \geq -19.5$  sample curve and for the brightest sample in each panel.

display a clustering amplitude that is similar to that of the brightest early types. However, the measurement of the correlation function for this faint sample is relatively noisy, as the volume in which galaxies are selected is small compared with the volumes used for brighter samples. The late-type galaxies show, in contrast, little change in clustering amplitude with increasing luminosity, with a redshift-space correlation length of  $s_0 = 5.6 \pm 0.6 h^{-1} \text{Mpc}$ . Only a slight steepening of the redshift-space correlation function is apparent with increasing luminosity, until the brightest sample, which displays a modest increase in the redshift-space correlation length. The

correlation lengths of all our samples of early-type galaxies are larger than those of late-type galaxies.

## 5 CLUSTERING IN REAL SPACE

### 5.1 Robustness of clustering results

The approach adopted to study the dependence of galaxy clustering on luminosity relies upon being able to compare correlation functions measured in different volumes. It is important to ensure that there are no systematic effects, such as significant sampling fluctuations, that could undermine such an analysis. In Norberg et al. (2001), we demonstrated the robustness of this approach in two ways. First, we constructed a volume-limited sample defined using a broad magnitude range, that could be divided into co-spatial subsamples of galaxies in different luminosity bins, i.e. subsamples within the same volume and therefore subject to the same large-scale structure fluctuations. A clear increase in clustering amplitude was found for the brightest galaxies in the volume, establishing the dependence of clustering on galaxy luminosity (see fig. 1a of Norberg et al. 2001). Secondly, we demonstrated that measuring the correlation function of galaxies in a fixed luminosity bin, but using samples taken from different volumes, gave consistent results (see fig. 1b of Norberg et al. 2001).

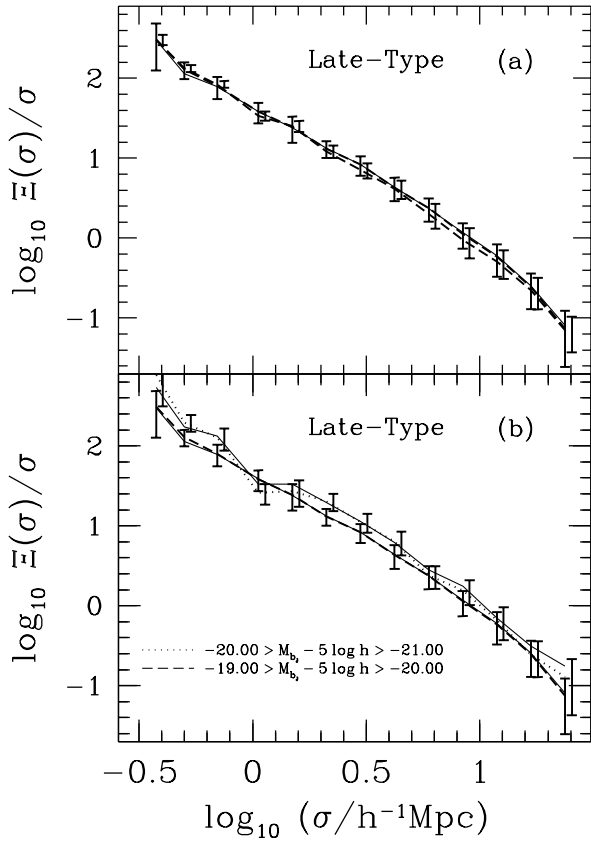
In this section we repeat these tests. The motivation for this exercise is that the samples considered here contain fewer galaxies than those used by Norberg et al. (2001), as only galaxies with  $z < 0.15$  are suitable for PCA spectral typing, and because the samples are more dilute as they have been selected on the basis of spectral type as well as luminosity. In Fig. 5(a) we plot the projected correlation function of late-type galaxies in a fixed absolute magnitude bin ( $-19.0 \geq M_{b_1} - 5 \log_{10} h \geq -20.0$ ), but measured for samples taken from volumes defined by different  $z_{\min}$  and  $z_{\max}$ . The clustering results are in excellent agreement with one another. In Fig. 5(b), we compare the projected correlation function of late-type galaxies in different absolute magnitude ranges but occupying the same volume. A clear difference in the clustering amplitude is seen. We have also performed these tests for early-type galaxies and arrived at similar conclusions.

As an additional test, we also show in Fig. 5 the correlation function measured in what we refer to as the *optimal sample* for a given magnitude bin. The optimal sample contains the maximum number of galaxies for the specified magnitude bin. The correlation functions of galaxies in optimal samples are shown by thin solid lines in both panels and are in excellent agreement with the other measurements shown.

### 5.2 Projected correlation function

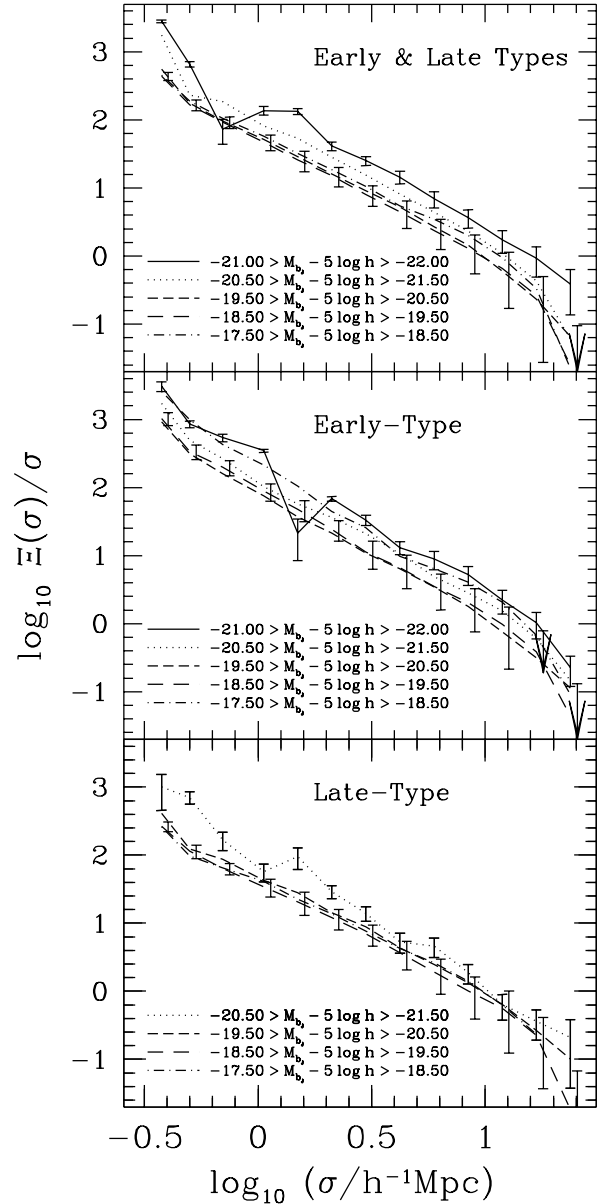
Fig. 6 shows how the real-space clustering of galaxies of different spectral type depends on luminosity. We use the optimal sample for each magnitude bin, i.e. the volume-limited sample with the maximum possible number of galaxies, the properties of which are listed in Tables 1 (early and late types together), 2 (early types only) and 3 (late types only).

The top panel of Fig. 6 confirms the results found by Norberg et al. (2001), namely that the clustering strength of the full sample increases slowly with increasing luminosity for galaxies fainter than  $M^*$ , and then shows a clear, strong increase for galaxies brighter than  $M^*$ . [We take  $M^*$  to be  $M_{b_1} - 5 \log_{10} h \approx -19.7$ , following Folkes et al. (1999) and Norberg et al. (2002).]



**Figure 5.** (a) The projected correlation function of late-type galaxies in a fixed absolute magnitude bin taken from different, almost independent volumes. We show the correlation function of galaxies with  $-19.0 \geq M_b - 5 \log_{10} h \geq -20.0$  taken from volumes defined by  $-18.0 \geq M_b - 5 \log_{10} h \geq -20.0$  and  $-19.0 \geq M_b - 5 \log_{10} h \geq -21.0$  (both shown by heavy dashed lines). The thin solid line shows the estimate from the optimal sample for the  $-19.0 \geq M_b - 5 \log_{10} h \geq -20.0$  mag bin. The different measurements are in almost perfect agreement. (b) The projected correlation function measured for late-type galaxies in two different absolute magnitude bins taken from the same volume. The volume is defined by the magnitude range  $-19.0 \geq M_b - 5 \log_{10} h \geq -21.0$ . Within a fixed volume, there is clear evidence for an increase (albeit small) in the clustering amplitude with luminosity. The two thin solid lines show estimates obtained from the corresponding optimal samples for the stated magnitude bins. In both panels the error bars come from the analysis of mock 2dFGRS catalogues.

Furthermore, the projected correlation functions are well described by a power law with a slope that is independent of luminosity. The middle panel of Fig. 6 shows the projected correlation function of early-type galaxies for different absolute magnitude ranges. The clustering amplitude displays a non-monotonic behaviour, with the faintest sample having almost the same clustering strength as the brightest sample. The significance of this result for the faintest galaxies will be discussed further in the next section. Early-type galaxies with  $M_b - 5 \log_{10} h \approx -19.5$ , display weaker clustering than the faint and bright samples. The bottom panel of Fig. 6 shows the real-space clustering of late-type galaxies as a function of luminosity. In this case, the trend of clustering strength with luminosity is much simpler. There is an increase in clustering amplitude with luminosity, and also some evidence that the projected correlation function of the brightest subset is steeper than that of the other late-type samples. In general, for the luminosity



**Figure 6.** The projected galaxy correlation function for samples of different spectral type, split into one absolute magnitude wide bins. The top panel shows the correlation function measured for all galaxies with an assigned spectral type. The middle panel shows correlation functions for early types and the bottom panel shows the results for late types. The absolute magnitude range of each sample is indicated in the legend on each panel. The error bars are derived from the 2dFGRS mock catalogues and, for clarity, are only plotted on the correlation functions of the  $-18.5 \geq M_b - 5 \log_{10} h \geq -19.5$  sample and of the brightest samples in each panel.

ranges for which a comparison can be made, the clustering strength of early-type galaxies is always stronger than that of late types.

The comparison of the correlation functions of different samples is made simpler if we divide the curves plotted in Fig. 6 by a fiducial correlation function. As a reference sample we choose all galaxies that have been assigned a spectral type, with absolute magnitudes in the range  $-19.5 \geq M_b - 5 \log_{10} h \geq -20.5$  (the short-dashed line in the top panel of Fig. 6). In Fig. 7, we plot the ratio of the correlation functions shown in the panels of Fig. 6, to the reference correlation function, with error bars obtained from the mock catalogues. The trends reported above for the variation of



**Table 1.** Properties of the combined NGP and SGP volume-limited samples for all galaxies with a spectral type. Column 1 lists the absolute magnitude range, which defines the volume-limited samples. Columns 2 and 3 list the median magnitude and the number of galaxies in the sample. Columns 4 and 5 give the redshift limits of the sample. Columns 6 and 7 list the best-fitting correlation length,  $r_0$ , and power-law slope,  $\gamma$ , of the real-space correlation function, and column 8 gives  $A(\gamma)$ , as defined by equation (8).

Mag. range $M_{b_j} - 5 \log_{10} h$	Median magnitude $M_{b_j} - 5 \log_{10} h$	$N_{\text{gal}}$	$z_{\text{min}}$	$z_{\text{max}}$	$r_0$ ( $h^{-1}$ Mpc)	$\gamma$	$A(\gamma)$
−17.5–18.5	−17.98	8510	0.0164	0.0724	$5.19 \pm 0.95$	$1.68 \pm 0.12$	4.14
−18.0–19.0	−18.46	13 795	0.0204	0.0886	$4.36 \pm 0.89$	$1.83 \pm 0.10$	3.58
−18.5–19.5	−18.93	19 207	0.0255	0.1077	$4.65 \pm 0.61$	$1.80 \pm 0.08$	3.68
−19.0–20.0	−19.40	24 675	0.0317	0.1302	$4.93 \pm 0.50$	$1.79 \pm 0.10$	3.71
−19.5–20.5	−19.85	22 555	0.0394	0.1500	$4.89 \pm 0.48$	$1.79 \pm 0.05$	3.71
−20.0–21.0	−20.30	10 399	0.0487	0.1500	$5.37 \pm 0.61$	$1.78 \pm 0.11$	3.75
−20.5–21.5	−20.74	3423	0.0602	0.1500	$6.57 \pm 0.83$	$1.83 \pm 0.23$	3.58
−21.0–22.0	−21.19	751	0.0739	0.1500	$8.47 \pm 0.97$	$1.80 \pm 0.29$	3.68

**Table 2.** Properties of the combined NGP and SGP volume-limited samples of early-type galaxies. See Table 1 for column definitions.

Mag. range $M_{b_j} - 5 \log_{10} h$	Median magnitude $M_{b_j} - 5 \log_{10} h$	$N_{\text{gal}}$	$z_{\text{min}}$	$z_{\text{max}}$	$r_0$ ( $h^{-1}$ Mpc)	$\gamma$	$A(\gamma)$
−17.5–18.5	−18.05	1 909	0.0163	0.0707	$8.33 \pm 1.82$	$1.87 \pm 0.23$	3.46
−18.0–19.0	−18.53	3 717	0.0203	0.0861	$6.28 \pm 1.46$	$1.98 \pm 0.11$	3.19
−18.5–19.5	−18.98	6 405	0.0253	0.1041	$5.92 \pm 1.00$	$1.83 \pm 0.10$	3.58
−19.0–20.0	−19.44	10 135	0.0314	0.1249	$5.71 \pm 0.57$	$1.87 \pm 0.09$	3.46
−19.5–20.5	−19.89	11 346	0.0388	0.1486	$5.66 \pm 0.56$	$1.87 \pm 0.09$	3.46
−20.0–21.0	−20.33	6 434	0.0480	0.1500	$6.10 \pm 0.72$	$1.80 \pm 0.12$	3.68
−20.5–21.5	−20.77	2 587	0.0590	0.1500	$7.60 \pm 1.02$	$1.87 \pm 0.26$	3.46
−21.0–22.0	−21.21	686	0.0722	0.1500	$9.74 \pm 1.16$	$1.95 \pm 0.37$	3.26

**Table 3.** Properties of the combined NGP and SGP volume-limited samples of late-type galaxies. See Table 1 for column definitions. Note that the brightest sample listed contains too few galaxies to permit a reliable measurement of the projected correlation function.

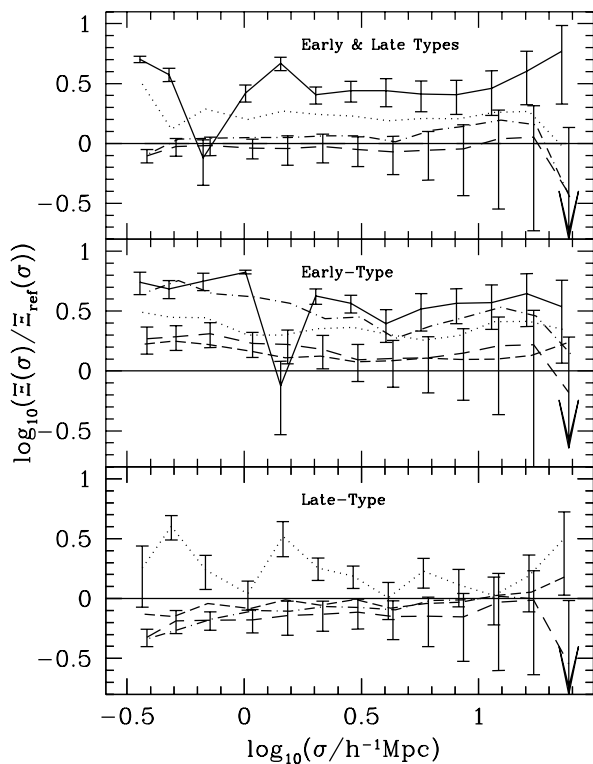
Mag. range $M_{b_j} - 5 \log_{10} h$	Median magnitude $M_{b_j} - 5 \log_{10} h$	$N_{\text{gal}}$	$z_{\text{min}}$	$z_{\text{max}}$	$r_0$ ( $h^{-1}$ Mpc)	$\gamma$	$A(\gamma)$
−17.5–18.5	−17.96	6 674	0.0164	0.0734	$4.27 \pm 0.81$	$1.65 \pm 0.12$	4.29
−18.0–19.0	−18.44	9 992	0.0205	0.0901	$3.71 \pm 0.77$	$1.76 \pm 0.11$	3.82
−18.5–19.5	−18.90	12 619	0.0256	0.1099	$4.17 \pm 0.64$	$1.79 \pm 0.10$	3.71
−19.0–20.0	−19.37	14 420	0.0319	0.1333	$4.45 \pm 0.47$	$1.76 \pm 0.09$	3.82
−19.5–20.5	−19.82	11 122	0.0397	0.1500	$4.59 \pm 0.44$	$1.76 \pm 0.07$	3.82
−20.0–21.0	−20.26	4 300	0.0492	0.1500	$5.52 \pm 0.88$	$1.87 \pm 0.13$	3.46
−20.5–21.5	−20.71	1 118	0.0608	0.1500	$6.33 \pm 1.01$	$2.01 \pm 0.29$	3.12
−21.0–22.0	−21.17	198	0.0749	0.1500	–	–	–

clustering strength with luminosity and spectral type are now clearly visible (see Fig. 7), particularly the difference in clustering amplitude between early types and late types. In the upper and lower panels, the ratios of correlation functions are essentially independent of  $\sigma$ , indicating that a single power-law slope is a good description over the range of scales plotted. The one exception is the brightest sample of late-type galaxies, which shows some evidence for a steeper power law. In the middle panel, the ratios for early-type galaxies show tentative evidence for a slight steepening of the correlation function at small pair separations,  $\sigma < 2 h^{-1}$  Mpc, which is most pronounced for the brightest sample.

### 5.3 Real-space correlation length

In the previous subsection, we demonstrated that the projected correlation functions of galaxies in the 2dFGRS have a power-law form with a slope that varies little as the sample selection is

changed, particularly for pair separations in the range  $2.0 \leq \sigma/(h^{-1} \text{ Mpc}) \leq 15.0$ . To summarize the trends in clustering strength found when varying the spectral type and luminosity of the sample, we fit a power law over this range of scales. We follow the approach, based on equation (8), used by Norberg et al. (2001), who performed a  $\chi^2$  minimization to extract the best-fitting values of the parameters in the power-law model for the real-space correlation function: the correlation length,  $r_0$ , and the power-law slope,  $\gamma$ . As pointed out by Norberg et al. (2001), a simple  $\chi^2$  approach does not give reliable estimates of the errors in the fitted parameters because of the correlation between the estimates of the correlation function at differing pair separations. We therefore use the mock 2dFGRS catalogues to estimate the errors in the fitted parameters in the following manner. The best-fitting values of  $r_0$  and  $\gamma$  are found for each mock individually, using the  $\chi^2$  analysis. The estimated error is then taken to be the fractional rms scatter in the fitted parameters over the ensemble of mock catalogues. The best-fitting parameters for each 2dFGRS sample are listed in

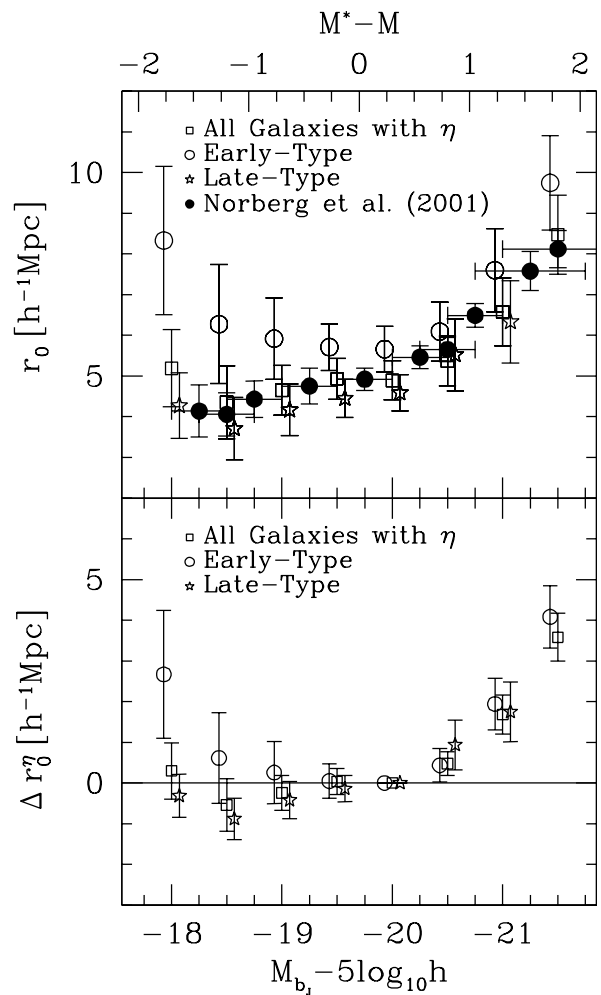


**Figure 7.** The ratio of the projected correlation function measured for galaxies selected by luminosity and spectral type to the projected correlation function of a reference sample. The reference sample consists of all galaxies with an assigned spectral type that lie within the magnitude range  $-19.5 \geq M_b - 5 \log_{10} h \geq -20.5$ . The top panel shows the ratios for different luminosity bins for all galaxies with a spectral type, the middle panel shows the ratios obtained for early types and the bottom panel shows the ratios for late types. The same line styles plotted in Fig. 6 are used to indicate different luminosities. The error bars, which take into account the correlation between the samples (but not between the bins), are from the mock 2dFGRS catalogues, and for clarity, are only plotted on two curves in each panel: the ratio for the sample with  $-18.5 \geq M_b - 5 \log_{10} h \geq -19.5$  and the ratio for the brightest sample in each panel.

Tables 1 (early and late types together), 2 (early types only) and 3 (late types only). The results for the correlation length are plotted in the top panel of Fig. 8.

The correlation lengths estimated for the full sample with assigned spectral types (shown by the open squares in Fig. 8) are in excellent agreement with the results of Norberg et al. (2001) (shown by the filled circles). The bright samples constructed by Norberg et al. (2001) have  $z_{\max}$  values in excess of the limit of  $z_{\max} = 0.15$  enforced upon the samples analysed in this paper by the PCA. The bright samples used in this paper therefore cover smaller volumes than those used by Norberg et al. and so the error bars are substantially larger. A cursory inspection of Fig. 8 would give the misleading impression that we find weaker evidence for an increase in correlation length with luminosity. It is important to examine this plot in conjunction with Tables 1–3, which reveals that there is significant overlap in the volumes defined by the four brightest magnitude slices, because of the common  $z_{\max} = 0.15$  limit.

In this case, the error bars inferred directly from the mocks do not take into account that our samples are correlated. The errors fully incorporate cosmic variance, i.e. the variance in clustering signal expected when sampling a given volume placed at different, independent locations in the Universe. The volumes containing the four brightest samples listed in Table 2 contain long-wavelength



**Figure 8.** Top panel: the real-space correlation length,  $r_0$ , as a function of the absolute magnitude of the sample, for galaxies of different spectral type. The stars show the  $r_0$  values fitted to the projected correlation function of late-type galaxies and the open circles show the  $r_0$  values for early types. The squares show  $r_0$  for the full sample with spectral types. The latter results are in excellent agreement with those obtained from the larger sample analysed by Norberg et al. (2001), which are plotted as filled circles. The horizontal bars on the filled circles show the magnitude range used to define the volume-limited samples. Lower panel: the difference  $\Delta r_0^\eta = r_0^\eta(M_b) - r_0^\eta(M_b^{\text{ref}})$  for each spectral type. This quantity gives the significance of the variation of the correlation length with luminosity for each spectral type with respect to a reference sample, chosen to be the sample defined by  $-19.5 \geq M_b - 5 \log_{10} h \geq -20.5$ . The error bars take into account the correlation between the two samples.

fluctuations in common and so clustering measurements from these different volumes are subject to a certain degree of coherency. The clearest way to show this is by calculating the difference between the correlation lengths fitted to two samples. The error in the difference, derived using the mock catalogues, has two components: the first comes from adding the individual errors in quadrature; the second is from the correlation of the samples. For correlated samples, this second term is negative and, therefore lowers the estimated error in the difference. This is precisely what is seen in the lower panel of Fig. 8, where we plot  $\Delta r_0^\eta = r_0^\eta(M_b) - r_0^\eta(M_b^{\text{ref}})$  for each spectral type as a function of absolute magnitude, with error bars taking into account the correlation of the samples. For all samples brighter than  $M^*$ , the increase in the clustering length with luminosity is clear.

There is a suggestion, in the top panel of Fig. 8, of a non-monotonic dependence of the correlation length on luminosity for early-type galaxies. The evidence for this behaviour is less apparent on the  $\Delta r_0^{\eta}$  panel, where a difference in  $r_0$  for the faintest sample is seen at less than the  $2\sigma$  level. These volumes are small compared with those defining the brighter samples. Furthermore, when analysing the faintest sample for SGP and NGP separately, our results are somewhat sensitive to the presence of single structures in each region (at  $\alpha \approx 0^h$  and  $z \approx 0.061$  for the SGP, as shown in Fig. 3). Thus we conclude that the upturn at faint magnitudes in the correlation length of early types is not significant.

The projected correlation function of early-type galaxies brighter than  $M^*$  is well fitted by a power-law real-space correlation function, with a virtually constant slope of  $\gamma \approx 1.87$  and a correlation length that increases with luminosity, from  $r_0 = 5.7 \pm 0.6 h^{-1}$  Mpc for  $M^*$  galaxies to  $r_0 = 9.7 \pm 1.2 h^{-1}$  Mpc for brighter galaxies ( $M_{bj} - 5 \log_{10} h \approx -21.2$ ). This represents an increase in clustering strength by a factor of 2.7, as seen in Fig. 7. The projected correlation functions of late-type galaxies are also consistent with a power law in real space, with an essentially constant slope. There is a very weak trend for  $\gamma$  to increase with luminosity, although at little more than the  $1\sigma$  level. Ignoring this effect, the fitted slope of the late-type correlation function is  $\gamma \approx 1.76$ . The correlation length increases with luminosity from a value of  $r_0 = 3.7 \pm 0.8 h^{-1}$  Mpc for faint galaxies ( $M_{bj} - 5 \log_{10} h \approx -18.4$ ) to  $r_0 = 6.3 \pm 1.0 h^{-1}$  Mpc for bright galaxies ( $M_{bj} - 5 \log_{10} h \approx -20.7$ ) – a factor of 2.5 increase in clustering strength. It should be possible to extend the analysis for late-type galaxies beyond  $M_{bj} - 5 \log_{10} h \approx -21$  when the 2dFGRS is complete.

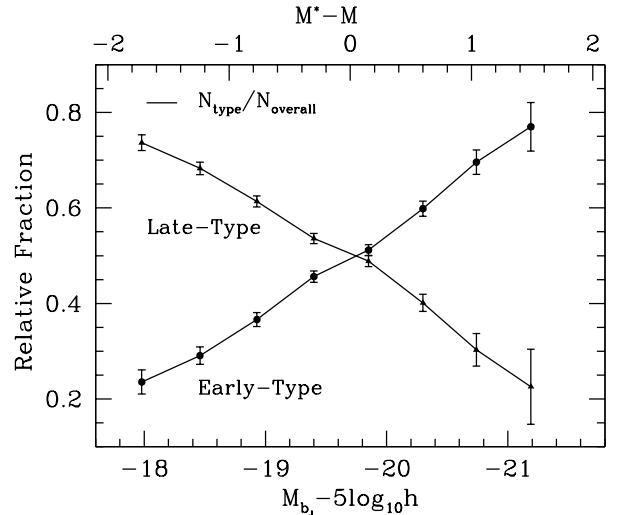
The top panel of Fig. 8 confirms our earlier conclusion that the clustering of early-type galaxies is stronger than that of late-type galaxies. At  $M^*$ , early types typically have a real-space clustering amplitude that is 1.5–1.7 times greater than that of late types.

## 6 DISCUSSION AND CONCLUSIONS

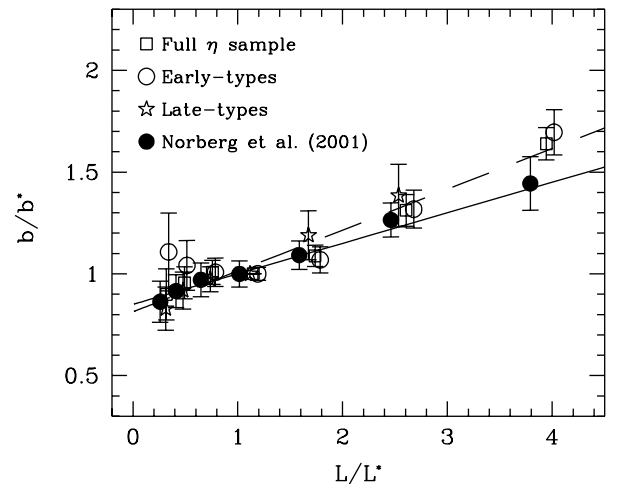
We have used the 2dFGRS to study the dependence of clustering on spectral type for samples spanning a factor of 20 in galaxy luminosity. The only previous attempt at a bivariate luminosity–morphology/spectral-type analysis of galaxy clustering was performed by Loveday et al. (1995) using the Stromlo-APM redshift survey. They were able to probe only a relatively narrow range in luminosity around  $L^*$ , which is more readily apparent if one considers the median magnitude of each of their magnitude bins (see fig. 3b of Norberg et al. 2001). The scatter between spectral and morphological types as illustrated in Fig. 1 precludes a more detailed comparison of our results with those of earlier studies, based on morphological classifications.

In Norberg et al. (2001), we used the 2dFGRS to make a precise measurement of the dependence of galaxy clustering on luminosity. The clustering amplitude was found to scale linearly with luminosity. One of the aims of the present paper is to identify the phenomena that drive this relation. In particular, there are two distinct hypotheses that we wish to test. The first is that there is a general trend for clustering strength to increase with luminosity, regardless of the spectral type of the galaxy. The second is that different types of galaxies have different clustering strengths, which may vary relatively little with luminosity, but a variation of the mix of galaxy types with luminosity results in a dependence of the clustering strength on luminosity.

Madgwick et al. (2002) estimated the luminosity function of 2dFGRS galaxies for different spectral classes, and found that, in going from early type to late types, the slope of the faint end of the luminosity function becomes steeper while the characteristic luminosity becomes fainter. Another representation of the variation of the luminosity function with spectral class is shown in Fig. 9, where we plot the fraction of early- and late-type galaxies in absolute magnitude bins. The plotted fractions are derived from the volume-limited samples listed in Tables 1–3. The mix of spectral types changes dramatically with luminosity; faint samples are dominated by late types, whereas early types are the most common galaxies in bright samples. Similar trends were found for galaxies



**Figure 9.** The fraction of galaxies in the two broad spectral classes, early type and late type, as a function of absolute magnitude. The fractions are derived from the volume-limited samples listed in Tables 1–3. The error bars show the Poisson errors on the fractions.



**Figure 10.** The relative bias (on the scale of  $r = 4.89 h^{-1}$  Mpc) of the different spectral classes as a function of luminosity, as indicated by the key. The definition of relative bias is given in the text. The reference sample covers the absolute magnitude range  $-19.5 \geq M_{bj} - 5 \log_{10} h \geq -20.5$ . The fiducial luminosity,  $L^*$ , is taken to be  $M_{bj} - 5 \log_{10} h = -19.7$ . The solid line shows the fit to the results of Norberg et al. (2001),  $b/b^* = 0.85 + 0.15L/L^*$ , whereas the dashed line shows the fit to the data analysed here. All the error bars plotted take into account the correlation between the various samples.

labelled by morphological type in the SSRS2 survey by Marzke et al. (1998).

We find that the change in the mix of spectral types with luminosity is not the main cause for the increase in the clustering strength of the full sample with luminosity. To support this assertion, we plot in Fig. 10 the variation of clustering strength with luminosity normalized, for *each* spectral class, to the clustering strength of a fiducial sample of  $M^*$  galaxies, i.e. the sample that covers the magnitude range  $-19.5 \geq M_{b_1} - 5 \log_{10} h \geq -20.5$ . For a galaxy sample with best-fitting correlation function parameters  $r_0^i$  and  $\gamma^i$ , we define the relative bias with respect to the  $M^*$  sample of the same type by

$$\frac{b^i}{b^*} \Big|_{\eta}(r) = \sqrt{\frac{(r_0^i)^{\gamma_i}}{r_0^{\gamma}} r^{\gamma-\gamma_i}}, \quad (9)$$

where  $r_0$  and  $\gamma$  are the best-fitting power-law parameters for the fiducial sample. In Fig. 10, we plot the relative bias evaluated at a fixed scale,  $r = 4.89 h^{-1}$  Mpc, which is the correlation length of the reference sample for all  $\eta$ -classified galaxies. A scale dependence in equation (9) arises if the slopes of the real-space correlation functions are different for the galaxy samples being compared. In practice, the term  $r^{\gamma-\gamma_i}$  is close to unity for the samples considered. The dashed line shows a fit to the bias relation defined by the open symbols. The solid line shows the effective bias relation obtained by Norberg et al. (2001), which is defined in a slightly different way to the effective bias computed here.

From Fig. 10, we see that the trend of increasing clustering strength with luminosity in both spectral classes is very similar for galaxies brighter than  $L > 0.5L^*$ . At the brightest luminosity, corresponding to  $\sim 4L^*$ , the clustering amplitude is a factor of 2–2.5 times greater than at  $L^*$ . This increase is much larger than the offset in the relative bias factors of early and late types at any given luminosity. We conclude that the change in correlation length with absolute magnitude found by Norberg et al. (2001) is primarily a luminosity effect rather than a reflection of the change in the mix of spectral types with luminosity.

Benson et al. (2001) showed that a dependence of clustering strength on luminosity is expected in hierarchical clustering cold dark matter universes because of the preferential formation of the brightest galaxies in the most massive, strongly clustered dark haloes. The close connection between the spectral characteristics of galaxies and their clustering properties discussed in this paper provides further evidence that the galaxy type is also related to the mass of the halo in which the galaxy forms.

## ACKNOWLEDGMENTS

The 2dFGRS is being carried out using the 2-degree field facility on the 3.9-m Anglo-Australian Telescope (AAT). We thank all those involved in the smooth running and continued success of the 2dF and the AAT. We thank the referee, Dr J. Loveday, for producing a speedy and helpful report. PN is supported by the Swiss National Science Foundation and an ORS award, and CMB acknowledges the receipt of a Royal Society University Research Fellowship. This work was supported in part by a PPARC rolling grant at Durham.

## REFERENCES

Baugh C. M., 1996, MNRAS, 280, 267  
 Benson A. J., Baugh C. M., Cole S., Frenk C. S., Lacey C. G., 2000a, MNRAS, 316, 107

Benson A. J., Cole S., Frenk C. S., Baugh C. M., Lacey C. G., 2000b, MNRAS, 311, 793  
 Benson A. J., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2001, MNRAS, 327, 1041  
 Berlind A. A., Weinberg D. H., 2002, ApJ, submitted (astro-ph/0109001)  
 Blanton M., 2000, ApJ, 544, 63  
 Bromley B. C., Press W. H., Lin H., Kirshner R. P., 1998, ApJ, 505, 25  
 Cole S., Fisher K. B., Weinberg D. H., 1994, MNRAS, 267, 785  
 Cole S., Hattton S., Weinberg D. H., Frenk C. S., 1998, MNRAS, 300, 945  
 Colless M. et al. (the 2dFGRS Team), 2001, MNRAS, 328, 1039  
 Davis M., Geller M. J., 1976, ApJ, 208, 13  
 Davis M., Peebles P. J. E., 1983, ApJ, 267, 465  
 Evrard A. et al. (the Virgo Consortium), 2002, ApJ, in press (astro-ph/0110246)  
 Folkes S. et al. (the 2dFGRS Team), 1999, MNRAS, 308, 459  
 Guzzo L., Strauss M. A., Fisher K. B., Giovanelli R., Haynes M. P., 1997, ApJ, 489, 37  
 Hamilton A. J. S., 1993, ApJ, 417, 19  
 Hermit S., Santiago B. X., Lahav O., Strauss M. A., Davis M., Dressler A., Huchra J. P., 1996, MNRAS, 283, 709  
 Iovino A., Giovanelli R., Haynes M., Chincarini G., Guzzo L., 1993, MNRAS, 265, 21  
 Jenkins A. et al., 1998, ApJ, 499, 20  
 Jenkins A., Frenk C. S., White S. D. M., Colberg J., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372  
 Kaiser N., 1987, MNRAS, 227, 1  
 Kauffmann G., Nusser A., Steinmetz M., 1997, MNRAS, 286, 795  
 Kauffmann G., Colberg J. M., Diaferio A., White S. D. M., 1999, MNRAS, 303, 188  
 Kennicutt R. C., 1992, ApJ, 388, 310  
 Kochanek C. S., Pahre M. A., Falco E. E., 2002, ApJ, submitted (astro-ph/0011458)  
 Landy S. D., Szalay A. S., 1993, ApJ, 412, 64  
 Lahav O., Saslaw W. C., 1992, ApJ, 396, 430  
 Loveday J., 1996, MNRAS, 278, 1025  
 Loveday J., Maddox S. J., Efstathiou G., Peterson B. A., 1995, ApJ, 442, 457  
 Loveday J., Tresse L., Maddox S. J., 1999, MNRAS, 310, 281  
 Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1990a, MNRAS, 243, 692  
 Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1990b, MNRAS, 246, 433  
 Maddox S. J., Efstathiou G., Sutherland W. J., Loveday J., 1996, MNRAS, 283, 1227  
 Madgwick D. S. et al. (the 2dFGRS Team), 2002, MNRAS, in press (astro-ph/0107197)  
 Marzke R. O., da Costa L. N., Pellegrini P. S., Willmer C. N. A., Geller M. J., 1998, ApJ, 503, 617  
 Norberg P. et al. (the 2dFGRS Team), 2001, MNRAS, 328, 67  
 Norberg P. et al. (the 2dFGRS Team), 2002, MNRAS, submitted (astro-ph/0111011)  
 Padilla N. D., Merchan M. E., Valotto C. A., Lambas D. G., Maia M. A. G., 2001, ApJ, 554, 873  
 Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144  
 Peacock J. A. et al. (the 2dFGRS team), 2001, Nat, 410, 169  
 Santiago B. X., Strauss M. A., 1992, ApJ, 387, 9  
 Shectman S. A., Landy S. D., Oemler A., Tucker D. L., Lin H., Kirshner R. P., Schechter P. L., 1996, ApJ, 470, 172  
 Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525  
 Seljak U., 2000, MNRAS, 318, 203  
 Somerville R. S., Lemson G., Sigad Y., Dekel A., Kauffmann G., White S. D. M., 2001, MNRAS, 320, 289  
 Tegmark M., Bromley B. C., 1999, ApJ, 518, L69  
 Willmer C. N. A., Da Costa L. N., Pellegrini P. S., 1998, AJ, 115, 869  
 Zehavi I. et al. (the SDSS Collaboration), 2002, ApJ, in press (astro-ph/0106476)

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.