# Galaxy ecology: groups and low-density environments in the SDSS and 2dFGRS

Michael Balogh,[1]★ Vince Eke,[1] Chris Miller,[2] Ian Lewis,[3] Richard Bower,[1] Warrick Couch,[4] Robert Nichol,[2] Joss Bland-Hawthorn,[5] Ivan K. Baldry,[6] Carlton Baugh,[1] Terry Bridges,[5] Russell Cannon,[5] Shaun Cole,[1] Matthew Colless,[7] Chris Collins,[8] Nicholas Cross,[7,9] Gavin Dalton,[4] Roberto De Propris,[4] Simon P. Driver,[9] George Efstathiou,[10] Richard S. Ellis,[11] Carlos S. Frenk,[1] Karl Glazebrook,[6] Percy Gomez,[2] Alex Gray,[12] Edward Hawkins,[13] Carole Jackson,[7] Ofer Lahav,[10] Stuart Lumsden,[14] Steve Maddox,[13] Darren Madgwick,[10] Peder Norberg,[15] John A. Peacock,[16] Will Percival,[16] Bruce A. Peterson,[7] Will Sutherland[16] and Keith Taylor[11]

[1]*Department of Physics, University of Durham, South Road, Durham DH1 3LE*
[2]*Department of Physics, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA*
[3]*Astrophysics, Nuclear and Astrophysics Laboratory, Keble Road, Oxford OX1 3RH*
[4]*School of Physics, University of New South Wales, Sydney 2052, Australia*
[5]*Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710, Australia*
[6]*Department of Physics and Astronomy, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218-2686, USA*
[7]*Research School of Astronomy and Astrophysics, The Australian National University, Weston Creek, ACT 2611, Australia*
[8]*Astrophysics Research Institute, Liverpool John Moores University, Twelve Quays House, Egerton Wharf, Birkenhead L14 1LD*
[9]*School of Physics and Astronomy, University of St Andrews, North Haugh, St Andrews, Fife KY16 9SS*
[10]*Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA*
[11]*California Institute of Technology, Pasadena, CA 91125-2400, USA*
[12]*Department of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, USA*
[13]*School of Physics and Astronomy, University of Nottingham, University Park, Nottingham NG7 2RD*
[14]*Department of Physics and Astronomy, E. C. Stoner Building, Leeds LS2 9JT*
[15]*ETHZ Institut für Astronomie, ETH Hönggerberg, CH-8093 Zürich, Switzerland*
[16]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ*

**ABSTRACT**

We analyse the observed correlation between galaxy environment and H$\alpha$ emission-line strength, using volume-limited samples and group catalogues of 24 968 galaxies at $0.05 < z < 0.095$, drawn from the 2dF Galaxy Redshift Survey ($M_{b_J} < -19.5$) and the Sloan Digital Sky Survey ($M_r < -20.6$). We characterize the environment by: (1) $\Sigma_5$, the surface number density of galaxies determined by the projected distance to the fifth nearest neighbour; and (2) $\rho_{1.1}$ and $\rho_{5.5}$, three-dimensional density estimates obtained by convolving the galaxy distribution with Gaussian kernels of dispersion 1.1 and 5.5 Mpc, respectively. We find that star-forming and quiescent galaxies form two distinct populations, as characterized by their H$\alpha$ equivalent width, $W_0(H\alpha)$. The relative numbers of star-forming and quiescent galaxies vary strongly and continuously with local density. However, the distribution of $W_0(H\alpha)$ amongst the star-forming population is independent of environment. The fraction of star-forming galaxies shows strong sensitivity to the density on large scales, $\rho_{5.5}$, which is likely independent of the trend with local density, $\rho_{1.1}$. We use two differently selected group catalogues to demonstrate that the correlation with galaxy density is approximately independent of group velocity dispersion, for $\sigma = 200$–$1000$ km s$^{-1}$. Even in the lowest-density environments, no more than

★E-mail: M.L.Balogh@Durham.ac.uk

~70 per cent of galaxies show significant Hα emission. Based on these results, we conclude that the present-day correlation between star formation rate and environment is a result of short-time-scale mechanisms that take place preferentially at high redshift, such as starbursts induced by galaxy–galaxy interactions.

**Key words:** galaxies: clusters: general – galaxies: evolution – galaxies: interactions.

## 1 INTRODUCTION

The galaxy population today can be effectively described as a combination of two distinct types. The first are red, morphologically early-type galaxies with little or no current star formation; the remainder are blue, late-type galaxies with active star formation. This segregation has been known for a long time; however, the superb data from the Sloan Digital Sky Survey has revealed how surprisingly distinct these two populations are, at least in terms of their colours (Strateva et al. 2001; Baldry et al. 2004). This segregation is known to be strongly mass-dependent, with the most massive galaxies being predominantly red, early types (Kauffmann et al. 2003). It is currently unclear whether galaxy evolution at the present day consists of galaxies changing from one type to another in a short time, or of evolution in properties within a given class.

Direct evidence of galaxy evolution comes from observations of galaxies at different redshifts; this shows that, in the past, the average star formation rate (SFR) was much higher (Lilly et al. 1996; Madau et al. 1996; Wilson et al. 2002), and the typical star-forming galaxy was more luminous (Cowie, Songaila & Barger 1999). Given the near-constancy of the SFR in the Milky Way (e.g. Rocha-Pinto et al. 2000), it seems likely that this downsizing effect is the main characteristic of the decline in global star formation: the most massive galaxies have recently stopped forming stars altogether, while less massive galaxies continue unhindered.

A more indirect form of evolution is observed as the change in galaxy populations as a function of their environment at a given epoch. In particular, galaxies in dense environments (i.e. clusters) tend to have early-type morphologies (e.g. Dressler 1980; Domínguez, Muriel & Lambas 2001; Treu et al. 2003) and low SFRs (Balogh et al. 1997, 1998; Poggianti et al. 1999; Balogh et al. 2000). One interpretation of this trend has been that the cluster environment causes galaxies to transform their properties as they move, pulled by gravity, from low-density regions into the cluster centre. However, recent work (e.g. O'Hely 2000; Kodama et al. 2001; Pimbblet et al. 2002), especially that based on the 2dF Galaxy Redshift Survey (2dFGRS, Lewis et al. 2002b, hereafter Paper I; Mateus & Sodré 2003) and Sloan Digital Sky Survey (SDSS, Gomez et al. 2003, hereafter Paper II), has shown that the correlation between galaxy type and local density extends to very low densities, well beyond the region where the cluster is expected to have much influence. These works showed that there is a smooth dependence of SFR on local galaxy density, and identified a 'critical' surface density of 1 Mpc$^{-2}$, where SFR correlations with environment first occur. This critical density is quite low, and corresponds to regions well outside the virialized cluster region; this suggests that galaxies may be preprocessed in groups, before they end up in clusters. This had been anticipated (e.g. Zabludoff & Mulchaey 1998, 2000), since galaxy–galaxy interactions are known to induce star formation (e.g. Barton, Geller & Kenyon 2000; Lambas et al. 2003), and such interactions should be common in groups. There is evidence that these interactions lead to the buildup of elliptical galaxies with hot X-ray haloes that will later be incorporated into clusters

(e.g. Ponman & Bertram 1993; Mulchaey & Zabludoff 1999; Smith et al. 2003). Groups are also expected to be important because they are the first level of the supergalactic hierarchy. Models of galaxy formation assume the rate at which gas falls on to a galaxy depends on its environment (e.g. Cole et al. 2000; Hernquist & Springel 2003). In these models, satellite galaxies do not have their own supply of hot gas to replenish gas in the disc used to form stars. Thus, star formation will gradually decline in any galaxy that is part of a larger halo (Balogh, Navarro & Morris 2000; Diaferio et al. 2001); groups will be the first environment to demonstrate this effect.

### 1.1 Previous work

Our understanding of galaxies in groups and lower-density environments has been hindered by the difficulties in obtaining large, unbiased samples. Most of the work has been restricted to compact groups (e.g. Hickson 1982; Rubin, Hunter & Ford 1991; Iglesias-Páramo & Vílchez 1999; Coziol, Iovino & de Carvalho 2000; Nishiura et al. 2000; Verdes-Montenegro et al. 2001; de la Rosa, de Carvalho & Zepf 2001; Kelm & Focardi 2003), which are most likely to be physically bound systems. However, these may be a special class of group, not representative of a typical stage of the hierarchy through which most galaxies evolve.

Postman & Geller (1984) analysed the morphology–density relation in groups selected from the CfA Redshift Survey (Geller & Huchra 1983) and demonstrated that the original relation found in clusters by Dressler (1980) extends continuously to lower densities. Below a density of ~1 Mpc$^{-3}$ ($M_{B(0)} < -17.5$), no further correlation is seen. Similar trends have since been confirmed in numerous other studies (e.g. Zabludoff & Mulchaey 1998; Hashimoto & Oemler 1999; Tran et al. 2001; Treu et al. 2003). Interestingly, Postman & Geller (1984) found that the trend with local density is the same in both rich clusters and poor groups; the only difference is that poor groups typically sample lower-density regions and thus have populations more dominated by late-type galaxies. It has been argued based on this evidence that galaxy–galaxy interactions within the group environment are the mechanism responsible for the high fractions of early-type galaxies in clusters (e.g. Zabludoff & Mulchaey 1998, 2000; Hashimoto & Oemler 2000).

Less work has been done on the stellar populations of group galaxies, though there is good evidence that they are intermediate between those of the field and rich clusters (Allington-Smith et al. 1993; Hashimoto et al. 1998; Tran et al. 2001; Carlberg et al. 2001; Girardi et al. 2002). More recently, analysis of SFRs in galaxy groups selected from the partially completed 2dFGRS survey has been done by Merchán & Zandivarez (2002). Based on this group catalogue, Martínez et al. (2002) found that even the lowest-mass groups, with masses ~$10^{13}$ M$_\odot$, show reduced total SFRs relative to the field. Domínguez et al. (2002) claim that the spectral-type dependence on local density is only observed in groups more massive than

$3 \times 10^{13}$ M$_\odot$; however, this claim appears to arise mostly from the fact that their low-mass groups do not sample densities as high as found in the higher-mass groups.

## 1.2 The purpose of this paper

We analyse the local correlation between star formation activity and environment using data obtained from the two largest galaxy surveys ever conducted: the 2dFGRS (Colless et al. 2001, 2003) and the SDSS (York et al. 2000). In particular, we have selected a sample of galaxy groups from each survey, but in significantly different ways. Since the definition of a group is partly subjective, and a variety of algorithms have been developed to find groups in redshift surveys, comparing results from the two catalogues allows us to investigate the sensitivity of our results to the way the group catalogue is constructed.

The purpose of this paper is to establish how star formation in the galaxy population, as characterized by the distribution of H$\alpha$ emission, depends on environment. Specifically, we will investigate whether the most important variable is (1) velocity dispersion of the embedding group or cluster, (2) local galaxy density, on scales $\lesssim 1$ Mpc, or (3) large-scale structure, as parametrized by the density on $\sim 5$ Mpc scales. We will show that the number of actively star-forming galaxies depends on both the local and large-scale density, but that the properties of the star-forming galaxies themselves do not.

A summary of the paper follows. In Section 2 we describe the details of the two data sets analysed herein, focusing on the construction of group catalogues, measurement of emission-line strengths and sample selection. Our results are presented in Section 3. We discuss the physical implications of our results in Section 4, and draw some conclusions in Section 5. Throughout the paper we use a cosmology of $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 100\,h$ km s$^{-1}$ Mpc, with $h = 0.7$. All distances are proper lengths in units of megaparsecs.

## 2 DATA

We will use local data gleaned from the two largest redshift surveys available: the 2dFGRS (Colless et al. 2001, 2003) and the SDSS (York et al. 2000). Below, we summarize the data in each survey, and describe our group catalogues, emission-line measurements and sample selection criteria.

## 2.1 The 2dF Galaxy Redshift Survey

### 2.1.1 Summary of the data

The 2dFGRS has obtained over 220 000 spectra of galaxies selected in the photographic $b_J$ band, from the APM galaxy catalogue. The targeted galaxies are located in two contiguous declination strips, plus 99 randomly located fields. One strip is in the Southern Galactic hemisphere and covers approximately $80 \times 15$ deg$^2$ centred close to the South Galactic Pole. The other strip is in the Northern Galactic hemisphere and covers $75 \times 10$ deg$^2$. The 99 random fields are located over the entire region of the APM galaxy catalogue in the Southern Galactic hemisphere outside of the main survey strip. We only use the contiguous fields for this work. Full details of the survey strategy are given in Colless et al. (2001).

The survey spectra were obtained through $\sim 2$ arcsec fibres, and cover the wavelength range 3600–8000 Å at 9-Å resolution. Only the wavelength range of 3600–7700 Å is used during the line fitting procedure because of poor signal-to-noise ratio and strong sky

emission in the red part of the spectrum. The wide wavelength range is made possible by the use of an atmospheric dispersion compensator (ADC) within the 2dF instrument (Lewis et al. 2002a). The accuracy of each individual redshift is $\sim 85$ km s$^{-1}$ (Colless et al. 2001).

### 2.1.2 The group catalogue

The group catalogue is based on a friends-of-friends percolation algorithm, which links neighbouring galaxies together if they lie within a specified linking length of each other (Eke et al. 2003). The linking length is scaled with redshift in order to obtain groups of a constant overdensity in the magnitude-limited survey. This is done by scaling the length according to $n(z)^{-1/3}$, where $n(z)$ is the mean galaxy density at redshift $z$. Furthermore, the linking length along the line of sight ($\ell_\parallel$) is allowed to be larger than that in the plane of the sky ($\ell_\perp$), to account for the effects of peculiar velocities. The algorithm is extensively tested on mock catalogues derived from numerical simulations with $\sigma_8 = 0.9$, so the completeness and contamination of the catalogue are understood. For details we refer the reader to (Eke et al. 2003); we briefly review the salient points here. The mock galaxy catalogues are created using the semi-analytic models of Cole et al. (2000), with the 2dFGRS selection function. Groups identified from this mock catalogue are then compared with the corresponding 'true' group, identified as galaxies populating dark matter haloes identified using a friends-of-friends algorithm with a linking length of $b = 0.2$ times the mean particle separation. The parameters of the group-finding algorithm are tuned to provide the best match between the median properties (size and mass) of the observed and true groups. These best-fitting parameters are $b = \ell_\perp n(z)^{1/3} = 0.13$, $L_{\perp,\max} = 2\,h^{-1}$ Mpc and $R = 11$, where $L_{\perp,\max}$ is the maximum linking length permitted across the line of sight and $R = \ell_\parallel / \ell_\perp$. These values are insensitive to $\sigma_8$ for reasonable values of the normalization. Small, parametrized perturbations of $b$ and $R$ are allowed to remove small differences in the recovered properties that correlate with halo mass. The median recovered velocity dispersions are accurate to better than $\sim 10$ per cent, independent of halo mass, when compared with the velocity dispersion of the parent dark matter halo; however, the scatter in this accuracy is large. The catalogue is highly complete, recovering $>95$ per cent of haloes with dark matter mass $M \lesssim 4 \times 10^{14}$ M$_\odot$; the price to be paid is some contamination from unphysical systems. For the most massive haloes, the algorithm is susceptible to a small amount of fragmentation; approximately 10 per cent of haloes with $M \gtrsim 10^{14}$ M$_\odot$ are fragmented into more than one group with mass at least 20 per cent that of the parent dark matter halo.

For each group, we calculate the velocity dispersion using the gapper estimate of Wainer & Thissen (1976), as discussed in Beers, Flynn & Gebhardt (1990), which is insensitive to outliers. The group centre is computed by iteratively rejecting the most distant galaxy until only two galaxies remain; the centre is taken to be the position of the brighter of these two galaxies.

### 2.1.3 H$\alpha$ as a star formation tracer in the 2dFGRS

The SFR is directly related to the H$\alpha$ emission luminosity (e.g. Kennicutt 1983), and we will use this emission line as our tracer of star formation. To provide a reliable description of SFR, however, H$\alpha$ luminosities need to be corrected for underlying absorption, dust extinction and aperture effects, all of which are important (e.g. Charlot & Longhetti 2001; Hopkins et al. 2001, 2003; Afonso et al.

2003). When these corrections are made, there is good agreement between Hα-derived SFRs, and those derived from the far-infrared, radio or ultraviolet continuum (Hopkins et al. 2003). In this paper, however, we will focus only on the rest-frame equivalent width of the Hα line, $W_0(H\alpha)$, corrected for underlying absorption (see below). Uniform dust extinction will not affect the $W_0(H\alpha)$, though selective extinction around massive stars will (Charlot & Longhetti 2001). Aperture corrections to the *flux* can be substantial, typically ranging from a factor of ~2 to ~6 (Hopkins et al. 2003); however, the effect of this missing flux on $W_0(H\alpha)$ depends on the spatial distribution of the Hα emission. As long as the size distribution of galaxies is independent of environment, neglecting this correction does not affect the results of this paper. We will demonstrate that this assumption is justified, in Appendix A.

As described in Paper I, all of the measurements of line equivalent widths have been performed using a completely automatic procedure. In summary, up to 20 individual absorption and emission lines are fitted simultaneously with Gaussian profiles. The Hα emission line is accurately deblended from the adjacent [N II]λ6548 and [N II]λ6583 lines; the [N II] lines are constrained to be in emission, while the Hα line may be either emission or absorption (but not both); measurements are made whether or not the line is detected in emission. We will add 1 Å to the Hα equivalent widths to account approximately for the effects of stellar absorption (Hopkins et al. 2003). This correction is not important for galaxies with significant emission, which are of interest here, but ensures that the mean $W_0(H\alpha)$ is never much less than zero. Because of the uncertain flux calibration of the 2dFGRS spectra, we do not derive Hα luminosities, or SFRs, but restrict the analysis to the observable quantity $W_0(H\alpha)$.

We also wish to exclude from the analysis galaxies in which most of the emission comes from an active galactic nuclei (AGN). Since the flux calibration is not reliable over long wavelengths, we just use the criterion [N II]λ6583/Hα > 0.63 to identify AGN-dominated spectra, when the equivalent widths of both lines are greater than 2 Å (e.g. Miller et al. 2003). Since the AGN fraction does not appear to correlate with environment, however (Miller et al. 2003), the correction does not affect any of the conclusions in this paper.

### 2.1.4 Sample selection

We select a volume-limited subset of the contiguous fields, where the lower redshift limit, $z > 0.05$, is chosen to reduce our sensitivity to aperture effects. The upper redshift limit, $z < 0.095$, is chosen because measurement of Hα emission becomes difficult when it is redshifted beyond ~7200 Å and night sky emission lines are strong.[1] At this redshift, the magnitude limit of the survey ($b_J = 19.45$) corresponds to a *k*-corrected luminosity of $M_b \approx -19$. However, for consistency with the SDSS sample (see Appendix B), we limit this sample to $M_b = -19.5$, using the average *k*-correction of Norberg et al. (2002). Finally, we exclude galaxies that are within 2 Mpc of a survey boundary; more stringent cuts are made when computing densities on various scales, as appropriate.

Our final, volume-limited sample contains 20 154 galaxies. For computations of $W_0(H\alpha)$, we exclude galaxies in which the continuum was negative, or a Gaussian was a poor fit to the line (see Paper I). Furthermore, we restrict our analysis to data taken after 1999 August, since the extreme ends of spectra obtained earlier are

---

[1] In Paper I we used an upper limit of $z = 0.1$; the small change here is made for consistency with the SDSS sample as presented in Paper II and this paper.
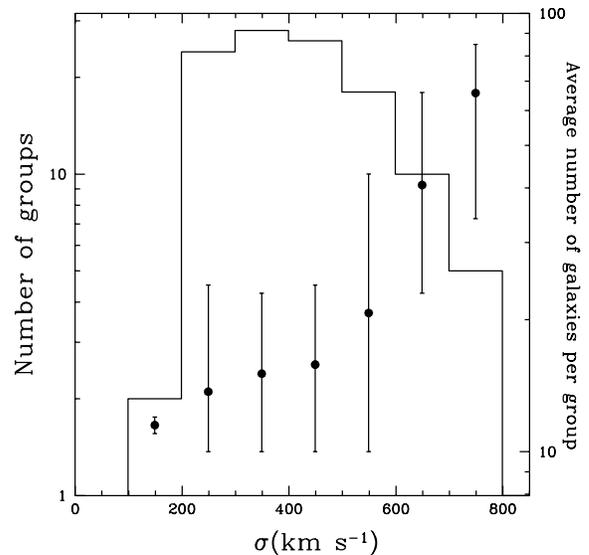


**Figure 1.** The histogram shows the abundance of groups with at least 10 members in the 2dFGRS sample as a function of velocity dispersion, corresponding to the left axis. The solid points are the average number of galaxies (brighter than $M_b = -19$) per group, with values corresponding to the right axis. The error bars span the range from the 10th to the 90th percentile of the distribution.

severely affected by problems with the ADC (Lewis et al. 2002a). These restrictions reduce the usable sample to 12 683 galaxies. Of these, 846 (6.7 per cent) are identified as AGN (see Section 2.1.3) and excluded from the sample. This leaves 11 837 galaxies, of which 7012 (59 per cent) are in the friends-of-friends group catalogue, and 1577 (13 per cent) are in groups with at least 10 members above the luminosity limit, which we use for our analysis. This latter requirement is made so that meaningful estimates of local density can be made. We use the friends-of-friends linking algorithm to designate group membership, rather than selecting all galaxies within some specified radius of a chosen centre. No further selection of the clusters is made, and some of the systems are therefore dynamically unrelaxed groups, for which the velocity distribution is not a reliable indicator of dynamical mass. The velocity dispersion distribution of groups with more than 10 bright members is shown in Fig. 1. As described in Section 2.1.2, this is designed to be a very highly complete catalogue that will still have some contamination from unphysical systems, though this contamination is substantially reduced by our selection of groups with at least 10 members (Eke et al. 2003).

## 2.2 The first SDSS data release

### 2.2.1 Summary of the data

The SDSS (http://www.sdss.org) is a joint, five-passband (*u, g, r, i, z*), imaging and medium-resolution ($R \simeq 1800$) spectroscopic survey of the Northern Galactic hemisphere (see York et al. 2000 for details). In 2003 May, the SDSS publicly released the first official set of data, named DR1, which comprises 186 240 spectra of galaxies, stars and quasi-stellar objects (QSOs) over 1360 deg$^2$ of sky. This release is fully described in Abazajian et al. (2003), and the reader is referred to Strauss et al. (2002) for a detailed description of the spectroscopic target selection for the SDSS main galaxy survey.

The spectra are obtained from 3 arcsec diameter fibres, larger than those of the 2dFGRS. The spectrographs produce data covering 3800–9200 Å, with the beam split at 6150 Å by a dichroic. The spectral resolution at $\lambda \sim 5000$ Å is $\sim 2.5$ Å, and redshift uncertainties are $\sim 30$ km s$^{-1}$.

### 2.2.2 The group catalogue

Groups are selected from the SDSS in a fundamentally different way from 2dFGRS groups (cf. Section 2.1.2), which allows us to test the sensitivity of our results to the group-finding algorithm. The SDSS algorithm (Nichol et al., in preparation) is a semi-parametric, high-dimensional technique developed to find galaxies clustered in both position and colour. The premise is that galaxies within clusters evolved similarly, and thus galaxy clusters contain subsets of galaxies that have similar spectral energy distributions (SEDs). Note that this is not a red-sequence finder (e.g. Gladders & Yee 2000), and will detect groups of blue galaxies if they have similar colours. Galaxy colours are used as a proxy for the shape of the SED since they cover a larger wavelength range than the spectra and because the dimensionality of the problem is reduced to a manageable seven-dimensional space (two spatial positions, redshift, $u - g$, $g - r$, $r - i$ and $i - z$). In practice, we expect clustering in colour space to be a signal from the red sequence (e.g. Gladders & Yee 2000). Galaxy overdensities in this 7D space are found by comparing to random locations in the SDSS DR1 survey. By design, the algorithm is extremely pure (i.e. few false positives). This high purity is the result of the lack of projection in 7D space, and the use of the false discovery rate statistical method in choosing a threshold above which galaxies are considered clustered (see Miller et al. 2001).

The algorithm has been tested against mock catalogues (Wechsler et al., in preparation), in which galaxies are assigned $r$-band magnitudes and placed on to dark matter particles in such a way that the luminosity-dependent clustering of the SDSS seen in Blanton et al. (2003a) is matched. The remaining magnitudes are then added in so that the correlation between colour and local density, as observed in the SDSS, is also matched. The same cluster-finding algorithm used on the real data is then run on the mock catalogues. Preliminary tests show that the cluster catalogue is $>90$ per cent complete for clusters with mass $M > 2.5 \times 10^{14}$ M$_\odot$; the completeness decreases towards lower masses. The catalogue is 100 per cent pure for systems with $M > 5 \times 10^{14}$ M$_\odot$ and always greater than 90 per cent pure (Nichol et al., in preparation). We exclude groups in which the velocity distribution (measured using all galaxies, not only those identified as clustered in colour space) is significantly different from a Gaussian, so that we preferentially select dynamically isolated, relaxed systems. The final catalogue contains 104 bona fide clusters with reliable velocity dispersions that are likely indicative of system mass. This catalogue is a good complement to the 2dFGRS catalogue, as it is a highly pure catalogue (i.e. with little contamination), at the expense of being incomplete for low-mass groups. All galaxies within 1000 km s$^{-1}$ and twice the virial radius (defined by the galaxy number overdensity relative to the field) of the cluster are considered group members; the exact radius chosen is not important because we present our data as a function of local density, which correlates well with radius within the virialized region.

### 2.2.3 Hα as a star formation tracer in the SDSS

We will again use the Hα equivalent width as a tracer of star formation, as we have done for the 2dFGRS sample (see Section 2.1.3).

Emission and absorption lines are measured automatically from the spectra by fitting multiple Gaussians where required. Again a measurement is provided for every line, whether or not it was detected in emission, and we make a 1 Å correction for underlying stellar absorption (Hopkins et al. 2003). The effects of converting $W_0(\text{H}\alpha)$ to SFR using various conversions (e.g. Kennicutt 1998; Hopkins et al. 2001; Charlot & Longhetti 2001) were discussed in Paper II and will not be considered here.

For the SDSS spectra, we can more effectively exclude AGN (relative to the 2dFGRS), since the spectra are accurately flux-calibrated. Here, we will consider the [N II]$\lambda$6583/Hα and [O III]/Hβ ratios, using the classification of Miller et al. (2003). Where possible, all four lines are used to identify AGN; if only one of the ratios is available, then only that one is used. Most (87 per cent) of the AGN are identified based on the [N II]/Hα ratio alone.

### 2.2.4 Sample selection

We will use the same volume-limited sample as in Paper II. The main criteria for selection are $0.05 < z < 0.095$ and $M_r < -20.6$ ($k$-corrected, for $H_0 = 70$ km s$^{-1}$ Mpc). As for the 2dFGRS data, the lower redshift limit is imposed to minimize aperture bias due to large nearby galaxies (see Appendix A for more detail). The upper redshift limit is that where our luminosity limit equals the magnitude limit of the SDSS ($r = 17.7$; Strauss et al. 2002). From a volume-limited sample of 19 287 galaxies, we exclude 4237 that are within 2 Mpc of a survey boundary, and another 64 galaxies for which no reliable measurement of Hα is available, leaving 14 986 galaxies. Finally, removing a $\sim 12$ per cent AGN contribution, we have 13 131 galaxies in the final sample. Of these, 1939 (15 per cent) are associated with a group in the catalogue.

In Fig. 2 we show the velocity dispersion distribution for the selected subset of SDSS groups. The distribution is different from that of the 2dFGRS groups (Fig. 1), probably due to the fact that the former catalogue is incomplete at the lowest velocity dispersions, while the latter has more contamination. Furthermore, the SDSS catalogue includes a few clusters with very high velocity dispersions,
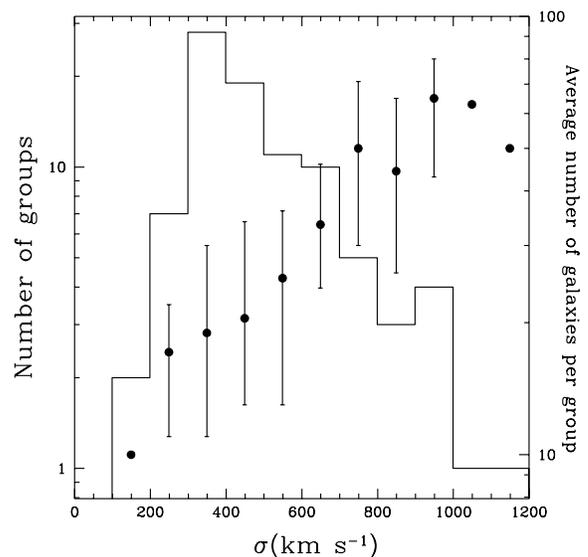


**Figure 2.** As Fig. 1, but for groups in the SDSS sample, restricted to groups with at least 10 member galaxies brighter than $M_r = -20.6$.

$\sigma > 800$ km s$^{-1}$, while such large systems tend to be fragmented in the 2dFGRS catalogue.

## 2.3 Homogeneity of the samples

In using the 2dFGRS and SDSS data together, we need to ensure that the samples are comparable in terms of selection and derived measurements. This detailed comparison is made in Appendix B. The only difference relevant to our results is that the $b_J$ selection of the 2dFGRS results in a sample that is more biased towards galaxies with large $W_0(H\alpha)$ than the $r$-selected SDSS. Because of this difference, we will show most of our results for the two surveys separately. The final combined sample contains 24 968 galaxies and 204 groups with at least 10 members. Approximately 14 per cent of all galaxies in the sample are associated with a group.

## 2.4 Density estimates

The local galaxy number density around a given galaxy is not in itself a well-defined quantity. As can be anticipated from the shape of the correlation function (e.g. Baugh et al. 1999), the density of neighbouring galaxies tends to increase as you probe closer to the target galaxy (see Appendix C, Section C2). We are therefore faced with two choices when characterizing the density around a galaxy. One is to measure the density within a fixed distance scale. This ensures that both high- and low-density regions are measured at the same scale; however, the measurement is noisy and of limited dynamic range due to the arbitrarily small, and finite, number of galaxies within the chosen distance. On the other hand, we can use a systematically larger scale in lower-density regions. This improves sensitivity and precision at low densities, but may be difficult to interpret because high- and low-density environments are measured at different physical scales. Without prior knowledge of the nature of the relevant physical effects, we cannot say which is the more meaningful estimate. For example, if we believe galaxy–galaxy interactions are important, then perhaps it is only the distance to the nearest neighbour that is relevant (Lambas et al. 2003). It is easy to imagine scenarios where either the distance to the *N*th neighbour, or density measured on a physical scale, is likely to be more meaningful.

In this paper, we will adopt two density estimators, which are described in Appendix C. The first is a traditional projected density estimate, $\Sigma_5$, which is based on the projected distance to the fifth nearest neighbour within $\pm 1000$ km s$^{-1}$. Inasmuch as we are willing to believe that galaxies can be found in discrete, relatively isolated groups, we must be cautious about interpreting our density estimate in groups where the number of members is less than five. For a group with only four members, the fifth nearest neighbour will clearly not lie in the same group and, therefore, the density computed from this distance may not be what is wanted. To ensure that we maintain our intuitive idea of a local environment within groups, we only consider groups that have at least 10 galaxies brighter than our luminosity limit. Furthermore, we exclude any galaxies in which the fifth nearest neighbour is more distant than the closest survey boundary, or within 1000 km s$^{-1}$ of our redshift limits, to ensure accurate measurements. This may bias us against finding low-density regions, but does not affect the observed trends of galaxy property with $\Sigma_5$ (e.g. Miller et al. 2003).

The second estimator is a three-dimensional density $\rho_\theta$, obtained by convolving the galaxy field with a Gaussian of dispersion $\theta$. In particular, we will consider the density measured on scales $\theta =$

1.1 and 5.5 Mpc (see Appendix C). These measurements underestimate the density in clusters with large velocity dispersions, but are particularly useful for probing low-density regimes where peculiar velocities are small. The main disadvantage is that the signal-to-noise ratio varies with density, and is low at low densities, when there are few galaxies within the aperture. In our analysis we will therefore indicate the density at which there is less than one galaxy within the Gaussian dispersion $\theta$, and will exclude any galaxies that are located less than $2\theta$ from a survey boundary, to avoid biasing the density estimate.

## 3 RESULTS

### 3.1 Dependence of $W_0(H\alpha)$ on environment

In Papers I and II we showed that there is a strong correlation between the $W_0(H\alpha)$ distribution and local galaxy density. In this section we will first define precisely what property of the galaxy distribution correlates with environment; we then proceed to explore how this correlation depends on the definition of local environment.

#### 3.1.1 Bimodality in $W_0(H\alpha)$ distribution

In Fig. 3 we present the correlation between $W_0(H\alpha)$ and $\Sigma_5$ for the 2dFGRS and SDSS samples. On the top axis we show the distance to the fifth nearest neighbour corresponding to $\Sigma_5$, which shows that at the lowest densities we are measuring the galaxy distribution on $\gtrsim 3$ Mpc scales, while in the densest regions the measurement is made at $\lesssim 300$ kpc. Recall (Section 2.4) that we exclude any galaxy for which the fifth nearest neighbour is more distant than the
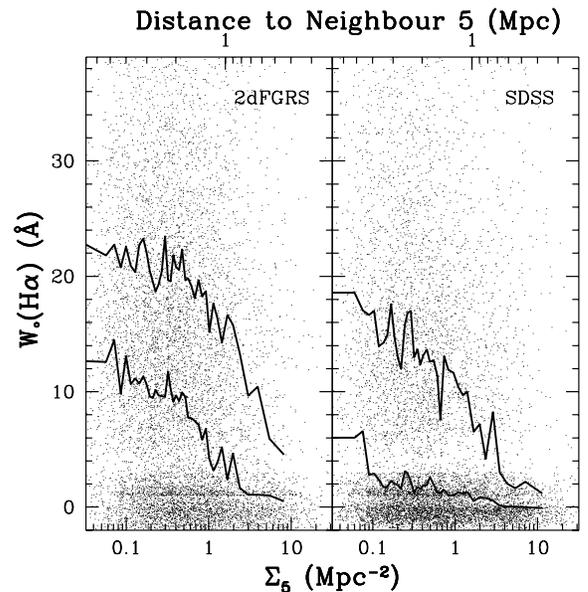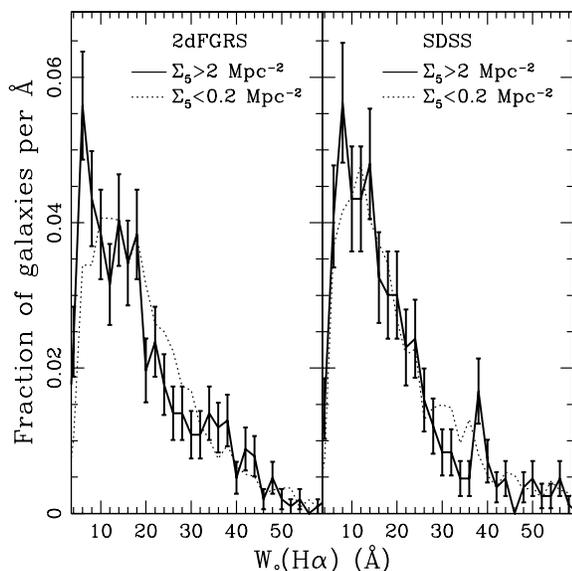


**Figure 3.** The dependence of $W_0(H\alpha)$ on projected, local density for the 2dFGRS (left panel) and SDSS (right panel) samples. The top axis shows the distance to the fifth nearest neighbour, from which $\Sigma_5$ is computed. The solid lines show the median and 75th percentile of $W_0(H\alpha)$, in bins of 100 galaxies. Although we only plot the data for $W_0(H\alpha) < 40$ Å, they extend to $W_0(H\alpha) \sim 100$ Å; see Fig. B2 for the full $W_0(H\alpha)$ distributions as histograms.
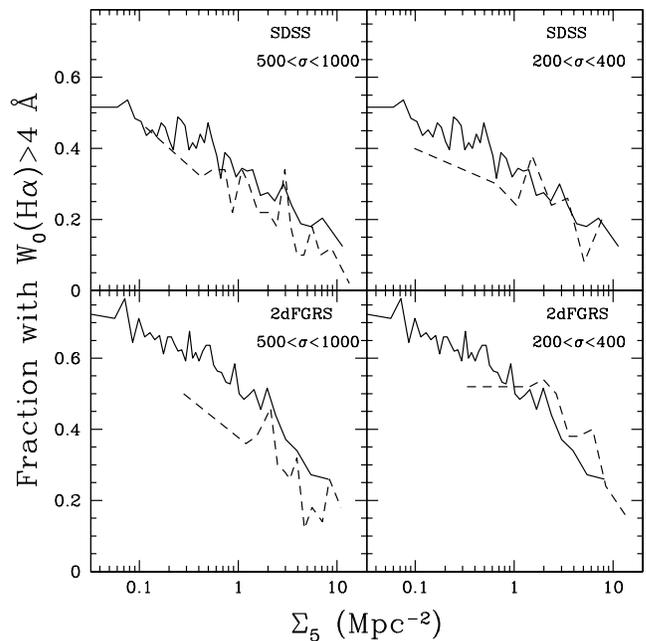
nearest survey boundary, or within 1000 km s$^{-1}$ of our redshift limits. The solid lines show the median and 75th percentile of the $W_0(H\alpha)$ distribution. In both cases, we reproduce the results of Papers I and II; the distribution changes distinctly at a density $\Sigma_5 \sim 2$ Mpc$^{-2}$. This is characterized by the near-total lack of galaxies with large $W_0(H\alpha)$ at densities greater than this value. At lower densities, there remains a correlation between $W_0(H\alpha)$ and $\Sigma_5$; however, it is weak, with the average $W_0(H\alpha)$ increasing by only $\sim$25 per cent over an order of magnitude in local density. The difference in normalization between the SDSS and 2dFGRS surveys is due to the different selection criteria for the spectroscopic sample, as discussed in Appendix B.

The $W_0(H\alpha)$ distribution shown in Fig. 3 reveals the presence of two distinct galaxy populations: those with significant, on-going star formation covering a broad range in H$\alpha$ strength (from $\sim$4 to >40 Å), and those with no on-going star formation, which form a conspicuous horizontal ridge line about $W_0(H\alpha) = 0$. [This is seen more clearly in Fig. B2, where we show the $W_0(H\alpha)$ distributions as histograms.] This recalls the bimodality observed in the colour distribution (Strateva et al. 2001; Blanton et al. 2003b); Baldry et al. (2004) find that the red peak distribution can be explained as the result of mergers between galaxies in the blue peak. We will therefore focus our analysis on the star-forming population, and its variation with environment.

In Fig. 4 we show the $W_0(H\alpha)$ distribution only for those galaxies with $W_0(H\alpha) > 4$ Å, in environments with the highest and lowest $\Sigma_5$ densities. The distributions are very similar; a Kolmogorov–Smirnov test cannot reject the hypothesis that the low- and high-density distributions are drawn from the same population with more than 1$\sigma$ confidence. Any difference, no matter how significant, must be small: the mean $W_0(H\alpha)$ is $\sim$21.5 and $\sim$20 Å for the low- and high-density populations, respectively, in both the 2dFGRS and SDSS samples. Thus, the observed trend of mean $W_0(H\alpha)$ with density (Fig. 3) is due almost entirely to the relative proportion of



**Figure 5.** The fraction of galaxies with $W_0(H\alpha) > 4$ Å as a function of local density for the 2dFGRS (bottom panels) and SDSS (top panels) samples. The solid lines represent the full galaxy sample, in bins each containing 250 galaxies. The dashed lines are restricted to galaxies that lie in groups or clusters with the indicated velocity dispersion, in bins of 50 galaxies each. Poisson-distributed uncertainties are typically $\sim$0.1 on the dashed lines and $\sim$0.05 on the solid lines.

galaxies with $W_0(H\alpha) > 4$ Å; there is at most weak sensitivity to environment within the actively star-forming population alone.[2]

### 3.1.2 The abundance of star-forming galaxies

Motivated by the results of the previous subsection, we show, in Fig. 5, how the fraction of galaxies with $W_0(H\alpha) > 4$ Å depends on $\Sigma_5$. It is evident that the fraction decreases steadily with increasing density; the break at $\Sigma_5 \sim 2$ Mpc$^{-2}$ is still there, but less apparent than in Fig. 3.

Since the identification of a characteristic density can have important implications, it is useful to understand why it appears so strongly in Fig. 3 (and Papers I and II) but is much weaker in Fig. 5. Upon close examination, a change in slope is most evident in the median of the 2dFGRS, at $\Sigma_5 \sim 1$ Mpc$^{-2}$, and in the 75th percentile of the SDSS, at $\Sigma_5 \sim 3$ Mpc$^{-2}$. From Fig. 5 we see that these two densities correspond to the point at which the fraction of galaxies with $W_0(H\alpha) > 4$ Å drops below 50 (2dFGRS) or 25 (SDSS) per cent; at higher densities the median and 75th percentile, respectively, are tracing the non-star-forming population with $W_0(H\alpha) < 4$ Å, and the trend with $\Sigma_5$ largely disappears.

It is worth stressing that the trend with $\Sigma_5$ is fairly insensitive to projection effects, despite the fact that we are projecting over a $\pm$14 Mpc cylinder, and despite the non-uniform density of the background, as we demonstrate in Appendix C (Section C1.1). To



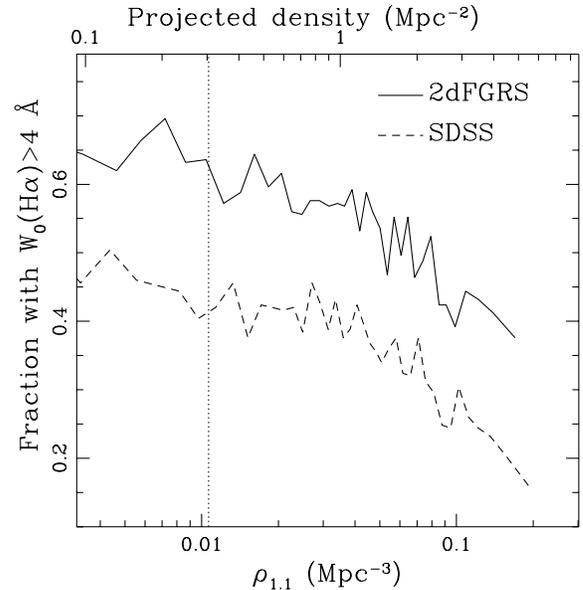**Figure 4.** The $W_0(H\alpha)$ distribution for galaxies in the 2dFGRS (left panel) and SDSS (right panel) with $W_0(H\alpha) > 4$ Å, in low-density environments ($\Sigma_5 < 0.2$ Mpc$^{-2}$, dotted line) and high-density environments ($\Sigma_5 > 2$ Mpc$^{-2}$, solid line). We show Poisson-distributed error bars on the high-density subsample, which is the smaller of the two.

---

[2] We will not deal directly with the distribution of $W_0(H\alpha)$ in galaxies with $W_0(H\alpha) < 4$ Å, since this distribution is most likely dominated by measurement uncertainties, including systematic effects like stellar absorption, rather than star formation activity.

summarize the results there, at $\Sigma_5 \sim 0.1\,\mathrm{Mpc}^{-2}$, where the projected contamination is $\sim 50$ per cent, we only overestimate the emission-line fraction by about 5 per cent. This is because the emission-line fraction varies weakly with $\Sigma_5$ at low densities, so there is little physical difference between the properties of the target galaxies and the projected galaxies. At high $\Sigma_5$, the projected fraction drops strongly, both because the contrast with the field is increasing, and because the radius of the projected cylinder is decreasing. Thus, the observed emission-line fraction remains within $\sim 5$ per cent of the true value, despite the fact that the contrast between the target and projected galaxy populations becomes large. The observed trend with $\Sigma_5$ is, therefore, a physical one.

Fig. 5 also shows the correlation for galaxies in groups of different velocity dispersion. We note that, although velocity dispersion gives a good indication of the group mass on average, there is a significant scatter arising because our groups are not all relaxed, spherical systems. Therefore, the velocity dispersion in some cases will be a better indicator of substructure or dynamic state than of mass. The correlation of emission-line fraction with density is present in both high- and low-velocity dispersion groups, selected from either the SDSS or 2dFGRS. There is marginal evidence (particularly in the 2dFGRS sample) that galaxies in the highest-velocity dispersion clusters have a low fraction of galaxies with $W_0(\mathrm{H}\alpha) > 4\,\text{Å}$ for their local density, relative to galaxies in lower-dispersion groups. Projection will have a more complex effect on this result, since the physical size of the group depends on $\sigma$; in particular, we expect the overestimation of both $\Sigma_5$ and $W_0(\mathrm{H}\alpha)$ to be greater in high-$\sigma$ clusters than in low-$\sigma$ clusters. This means the high-$\sigma$ cluster $W_0(\mathrm{H}\alpha)$ distribution should exceed that of low-$\sigma$ groups at a fixed $\Sigma_5$; this is opposite to what we find, and cannot therefore be the explanation for the small observed difference. Finally we note the remarkable similarity between the results for $\sigma < 400\,\mathrm{km\,s}^{-1}$ groups in both samples. Since the 2dFGRS catalogue is highly complete, and the SDSS catalogue is highly pure, the persistence of a strong $W_0(\mathrm{H}\alpha)$–$\Sigma_5$ relation in groups from both catalogues is not likely to be due to a selection effect.

We now consider the three-dimensional Gaussian kernel density estimator with a 1.1 Mpc standard deviation, $\rho_{1.1}$ (the motivation for the choice of filtering size is given in Appendix C). The $\rho_{1.1}$ density measure allows us to find galaxies in the very lowest-density environments, which are not easily measured from $\Sigma_5$ since that quantity projects the galaxy distribution over a 2000 km s$^{-1}$ cylinder.[3] In our analysis, we exclude galaxies that are within 2.2 Mpc of a survey boundary, for which $\rho_{1.1}$ cannot be reliably determined; we also note that $\rho_{1.1}$ becomes noisy when there are no galaxies within the filtering scale 1.1 Mpc, $\rho_{1.1} < 0.01\,\mathrm{Mpc}^{-3}$.

Fig. 6 shows how the fraction of galaxies with $W_0(\mathrm{H}\alpha) > 4\,\text{Å}$ depends on $\rho_{1.1}$. To compare this approximately with the corresponding $\Sigma_5$ value, we compute the projected density within a cylinder of length $2 \times 1000$ km s$^{-1}$ along the line of sight, to obtain an equivalent surface density, shown along the top axis. The emission-line fraction shows a strong dependence on $\rho_{1.1}$, and gets significantly steeper at $\rho_{1.1} \gtrsim 0.05\,\mathrm{Mpc}^{-3}$. However, the correlation is still present at the lowest densities observed. It is not clear whether the change in slope at $\rho_{1.1} \gtrsim 0.05\,\mathrm{Mpc}^{-3}$ highlights an interesting physical scale, or if $\rho_{1.1}$ simply loses sensitivity to the underlying density distribution at low densities, an effect that would be exaggerated by the logarithmic scale. Furthermore, the highest

---

[3] This projection corresponds to $\sim 25$ Mpc, much larger than the linear size of typical void regions (Benson et al. 2003).



**Figure 6.** The fraction of galaxies with $W_0(\mathrm{H}\alpha) > 4\,\text{Å}$ in the SDSS (dashed line) and 2dFGRS (solid line), as a function of three-dimensional density $\rho_{1.1}$, estimated with a Gaussian kernel with 1.1 Mpc standard deviation. Each bin contains 250 galaxies. The vertical, dotted line corresponds to the density where no galaxy lies within the Gaussian filtering scale $\theta$, making the density estimate particularly noisy. On the top axis we show the estimated two-dimensional projected density within a cylinder of length 2000 km s$^{-1}$, for comparison with $\Sigma_5$.

densities (which tend to correspond to clusters with high velocity dispersions) will generally be underestimated as a result of large peculiar velocities, and this may be partly responsible for the change in slope.

## 3.2 Local or global density?

We have shown that the fraction of star-forming galaxies depends strongly on the local density, measured on scales $\lesssim 1$ Mpc, and that this correlation is approximately independent of the velocity dispersion of the embedding cluster or group. However, velocity dispersion may not be the best measurement of large-scale structure; in this subsection, we will compare densities measured on 1.1 and 5.5 Mpc scales, to determine whether $W_0(\mathrm{H}\alpha)$ shows any independent sensitivity to structure on large scales.

In Fig. 7 we compare the Gaussian-filtered densities $\rho_{1.1}$ and $\rho_{5.5}$ for galaxies in the SDSS and 2dFGRS samples. Galaxies are only considered if they are at least 11 Mpc from a survey boundary. In Appendix C (Section C2) we show that these density estimators are good tracers of intuitively dense and low-density regions, despite the complications of peculiar velocities in clusters (Fig. C4). Note that the 2dFGRS data extend to larger values of $\rho_{5.5}$ than the SDSS; these data arise from a large supercluster region at $z \sim 0.0839$ and $\alpha \sim 195.5$, $\delta \sim -2.9$ (J2000 degrees). Fig. 7 shows a clear separation between galaxies in groups (200 km s$^{-1} < \sigma < 400$ km s$^{-1}$) and clusters ($\sigma > 600$ km s$^{-1}$). Galaxies in both environments span a similar range in local environment, characterized by $\rho_{1.1}$. However, galaxies in the high-velocity-dispersion clusters lie at higher densities on 5.5 Mpc scales, as expected since they are physically larger systems. This is particularly true for the SDSS sample, where the groups are selected to have approximately Gaussian distributed velocities, so that there is likely to be less scatter in the relation
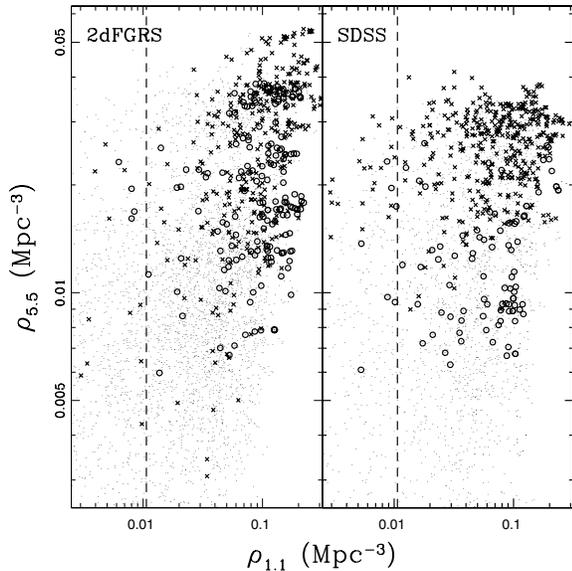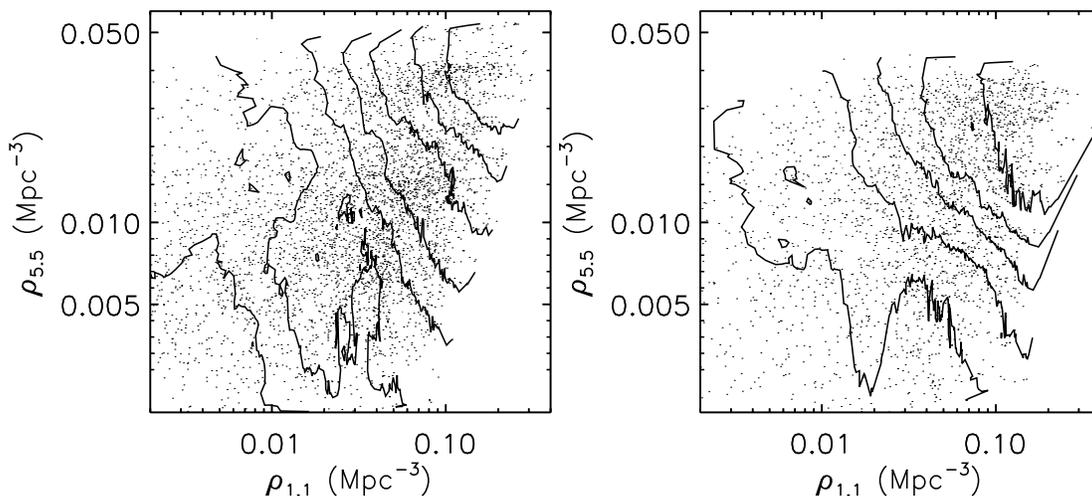
**Figure 7.** Galaxies in the 2dFGRS (left panel) and SDSS (right panel) are shown as a function of density computed on 1.1 and 5.5 Mpc scales, using a Gaussian filtering kernel (small dots). Galaxies in groups with $\sigma = 200$–$400\,\mathrm{km\,s^{-1}}$ are shown as open circles, while those with $\sigma > 600\,\mathrm{km\,s^{-1}}$ are represented as crosses. Only galaxies that are at least 11 Mpc from a survey boundary are considered. The dashed line shows the minimum reliable density; below this there are fewer than one galaxy within the filter size $\theta$ (the corresponding limit on $\rho_{5.5}$ is off the scale of the plot).

between $\sigma$ and virial mass. Note that the higher velocity dispersions of clusters mean that this pseudo-three-dimensional measurement will *underestimate* the density, more so than in groups. Thus the real difference between the clusters and groups is even larger than shown here.

In Fig. 8, we illustrate how the fraction of emission-line galaxies (measured from the nearest 500 galaxies in this plane) depends on these two very different density scales. Interestingly, the contours are not parallel with either axis, which indicates that the population

composition shows a dependence on large ($>5$ Mpc) scales, in addition to the more local density measured at 1.1 Mpc. That is, the fraction of emission-line galaxies is lower in regions that are overdense on 5.5 Mpc scales, even when the local (1.1 Mpc) overdensity is the same. We have checked that this is not an artefact of the smoothing or the correlation between $\rho_{1.1}$ and $\rho_{5.5}$, in the following way. For each galaxy we assign a value of $\rho_{5.5}$, chosen at random from among galaxies that have similar (within 10 per cent) values of $\rho_{1.1}$. This preserves the correlations between $\rho_{1.1}$ and $\rho_{5.5}$, as well as between $\rho_{1.1}$ and $W_0(\mathrm{H}\alpha)$, but removes any residual correlation between $\rho_{5.5}$ and $W_0(\mathrm{H}\alpha)$. In this case, the contours lie nearly parallel to the $\rho_{5.5}$ axis, confirming that the correlation we see in Fig. 8 is real. There is some evidence that, at high densities, the fraction of H$\alpha$-emitting galaxies is mostly dependent on $\rho_{1.1}$, and the dependence on larger scales becomes more important at lower densities. The dependence on large-scale densities is evidently stronger than the dependence on velocity dispersion (Fig. 5), despite the correlation between them (Fig. 7).

A cautionary note needs to be added, however. Since $\rho_{1.1}$ and $\rho_{5.5}$ are intrinsically correlated, errors on these measurements can give rise to the trends shown in Fig. 8, even if the galaxy population only depends on one parameter. However, a comparison with mock catalogues strongly suggests that the observed dependence on both density scales is real (Balogh et al., in preparation).

## 4 DISCUSSION

### 4.1 Overview

Thanks to the unique size and homogeneity of the 2dFGRS and SDSS data sets, we have been able to trace the H$\alpha$ distribution of galaxies over the full range of environments at the present day. This analysis has revealed three new, important clues about the process of galaxy formation:

(1) The change in $W_0(\mathrm{H}\alpha)$ distribution as a function of environment is predominantly due to a change in the relative number of star-forming and non-star-forming galaxies; the $W_0(\mathrm{H}\alpha)$ distribution of



**Figure 8.** Galaxies in the 2dFGRS (left panel) and SDSS (right panel) are shown as a function of density computed on 1.1 and 5.5 Mpc scales, using a Gaussian filtering kernel. Only galaxies that are at least 11 Mpc from a survey boundary are considered. The contours trace the fraction of galaxies with $W_0(\mathrm{H}\alpha) > 4$ Å, computed for the nearest 500 galaxies at each point in this plane. The contours are spaced in steps of 0.05, and increase towards lower densities. For the SDSS sample the contours span fractions 0.25–0.45, while for the 2dFGRS they span 0.35–0.65.

actively star-forming galaxies does not appear to depend strongly on environment. This is a surprising result that challenges theories in which the environment at the present day induces a transformation in galaxy properties, as we discuss in detail in Section 4.2.

(2) The fraction of star-forming galaxies increases continuously with decreasing density. However, even at the lowest densities there is a substantial fraction of galaxies with negligible Hα emission; we will discuss the consequences of this in Section 4.3. There is evidence that the correlation with environment becomes steeper at $\Sigma_5 \gtrsim 1\ \mathrm{Mpc}^{-2}$, an effect that is also seen in the morphology–density relation (Postman & Geller 1984; Treu et al. 2003). This may represent an interesting scale at which different physics becomes important; for example, as noted by Postman & Geller (1984), at this density the typical dynamical time of the group is longer than a Hubble time, so galaxy–galaxy interactions may start to play a role (see also Zabludoff & Mulchaey 1998, 2000).

(3) The fraction of galaxies with Hα emission may depend not only on the local density, but also on the density on much larger scales, ∼5.5 Mpc. If this trend is real, it means that the galaxy population must be only *indirectly* related to its present-day environment. We discuss the implications of this in Section 4.4.

These three clues allow us to make advances in our understanding of the evolution of star formation in the Universe (Section 4.5), and of any additional physics that may take place in dense environments (Section 4.6).

## 4.2 Star-forming galaxies in dense environments

If galaxies at the present day are evolving as they move from low-density regions to high-density regions, we should see a signature of this transformation which depends on the relevant time-scale. In particular, if all galaxies experience a gradual ($\gtrsim 2$ Gyr) decline in star formation when they become bound to groups or clusters (Balogh, Navarro & Morris 2000; Bekki, Couch & Shioya 2002), we expect star-forming galaxies in dense regions to show systematically lower $W_0(\mathrm{H}\alpha)$. Fig. 4 shows that this is not the case, implying that such a gradual decline is not a common phenomenon at the present time.

On the other hand, if the SFR declines rapidly to zero, this might not be reflected in Fig. 4, since galaxies will quickly move from the star-forming distribution to the quiescent one with $W_0(\mathrm{H}\alpha) < 4$ Å. Instead, we need to look at a longer-lived tracer of star formation (i.e. one that changes slowly after star formation ceases) to observe this effect. For example, the Hδ absorption line, which remains strong for ∼1 Gyr after star formation ceases, is a common diagnostic of recent activity. The extreme rarity (<0.1 per cent) of bright galaxies with strong Hδ but negligible star formation at low redshift provides evidence against the possibility that SFR has recently declined on a short time-scale for a significant number of galaxies (Zabludoff et al. 1996; Goto et al. 2003); however, this depends on an accurate quantification of the duration of the enhanced Hδ phase, which is systematically uncertain to within at least a factor of 2. In particular, if star formation is truncated in normal spirals, without a preceding burst, the Hδ line never gets very strong, and is only enhanced for a short time (e.g. Balogh et al. 1999).

In a similar way, we can use the accurate colours of the SDSS, which do not suffer from aperture bias, to trace the longer-lived stellar population. In particular, we will consider the rest-frame Petrosian $(g - r)$ colour, denoted $(g - r)_0$, and use the $k$-corrections of Blanton et al. (2003a). To isolate the population of galaxies with recent star formation, we select blue galaxies as those with $(g - r)_0$
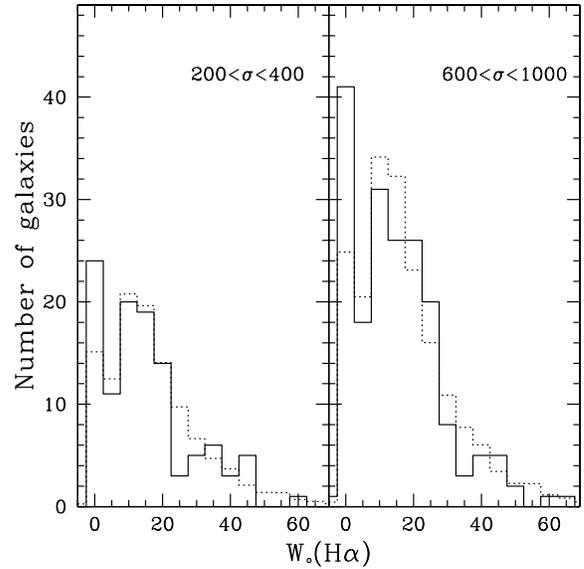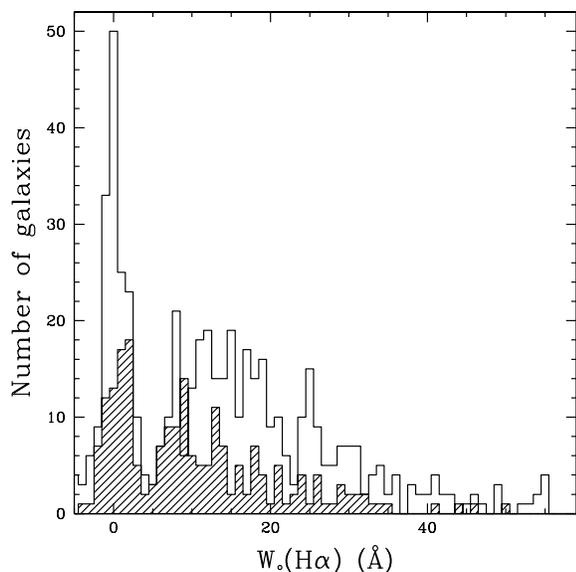


**Figure 9.** The $W_0(\mathrm{H}\alpha)$ distribution for blue galaxies ($(g - r)_0 < 0.7$) in the SDSS volume-limited sample. The solid line shows the distribution for galaxies in groups (left panel) and clusters (right panel), while the dotted histogram shows the distribution for all blue galaxies in the SDSS sample, renormalized to include the same area as the solid lines.

< 0.7, which will include any galaxy that has formed stars within the last ∼300 Myr, using the latest Bruzual & Charlot (2003) stellar population models.[4] Fig. 9 shows the $W_0(\mathrm{H}\alpha)$ distribution for all such blue galaxies, compared with that of blue galaxies in groups ($\sigma = 200$–$400$ km s$^{-1}$) and clusters ($\sigma > 600$ km s$^{-1}$). All three distributions are consistent with being drawn from the same parent population. If cluster galaxies had slowly declining SFRs, they would remain blue, but with reduced $W_0(\mathrm{H}\alpha)$ (e.g. Shioya et al. 2002), contrary to our observations. This reinforces our conclusions, drawn from Fig. 4, that star-forming galaxies in dense environments have normal SFRs, and are not being inhibited by their environment at the present day.

However, we are still not able to rule out the possibility that the environment has caused the SFR to decline *rapidly* in a substantial number of cluster galaxies, as we now demonstrate. There are 617 galaxies with $W_0(\mathrm{H}\alpha) < 4$ Å in the $\sigma > 600$ km s$^{-1}$ clusters. If the clusters have been constantly, and uniformly, truncating star formation in active galaxies over the last ∼10 Gyr, we would only need to find ∼20 blue galaxies with $W_0(\mathrm{H}\alpha) < 4$ Å in clusters with $\sigma > 600$ km s$^{-1}$ at the present day to account for these currently inactive galaxies. This number is not inconsistent with Fig. 9; in fact there is perhaps evidence for such a population, though this may equally well be contamination from the tail of the distinct, red galaxy population at $(g - r)_0 > 0.7$. By the same argument, our results are consistent with a model in which 20 per cent of the cluster members not presently forming stars had their star formation truncated some time within the last 2 Gyr.

Thus, the $(g - r)_0$ colour is too sensitive to recent star formation for us to put strong constraints on the number of cluster galaxies that have had their star formation truncated in the last few gigayears. Galaxy morphology may provide a better indicator, since an observable disc structure should persist for $\gtrsim 1$ Gyr after star formation

---

[4] Based on a model with a Salpeter initial mass function, and a SFR that declines exponentially ($\tau = 4$ Gyr) for 10 Gyr.

**Figure 10.** The $W_0(H\alpha)$ distribution for all SDSS and 2dFGRS galaxies in low-density environments ($\rho_{1.1} < 0.01$ Mpc$^{-3}$ and $\rho_{5.5} < 0.005$ Mpc$^{-3}$, unassociated with a group or cluster). The shaded histogram shows the distribution for bright ($>L^*$) galaxies, and the solid line represents the fainter population.

ceases (e.g. Bekki et al. 2002). Past analysis (Balogh et al. 1998; Hashimoto et al. 1998; Paper I; Paper II; Girardi et al. 2003) has indeed suggested that morphology and SFR are partly independent. Furthermore, the existence in clusters of H I-deficient spirals (e.g. Solanes et al. 2001), red spiral galaxies with little star formation (e.g. Poggianti et al. 1999; Balogh et al. 2002a) and spiral galaxies with unusually smooth structure (McIntosh, Rix & Caldwell 2003) all argue for some sort of transformation of spiral galaxies to be taking place in clusters. However, many of these results are based on a fairly coarse morphological binning; therefore, we reserve drawing firm conclusions until reliable, automated morphological measurements are available for the present samples (e.g. Liske et al. 2003; Blanton et al. 2003b; Kelly & McKay 2004).

### 4.3 Isolated galaxies

Although the fraction of emission-line galaxies continually increases with decreasing density, it never gets much larger than ~70 per cent, even in the lowest-density environments studied here. In these empty regions of the Universe, environment is not likely to have played a large role in galaxy evolution; therefore many galaxies must have ceased their star formation activity for reasons independent of their surroundings (but see below). In Fig. 10 we show the $W_0(H\alpha)$ distribution for galaxies in the lowest-density environments, with $\rho_{1.1} < 0.01$ Mpc$^{-3}$ and $\rho_{5.5} < 0.005$ Mpc$^{-3}$, and unassociated with any catalogued group or cluster. We combine all galaxies from both the 2dFGRS and SDSS samples, and consider only those galaxies which lie at least 11 Mpc from a survey boundary. Even in these extremely sparse regions of space, only ~70 per cent of galaxies are forming a significant number of stars, with $W_0(H\alpha) > 4$ Å, and this fraction is similar for both faint and bright galaxies[5] (shown separately in Fig. 10). A possible interpretation is that the ~30 per cent of bright, isolated galaxies with no sign of star

---

[5] Note that only the brightest isolated galaxies are sure to have no companions of comparable brightness, since fainter galaxies could have neighbours that are just below our luminosity limit.

formation are fossil groups (Ponman & Bertram 1993; Mulchaey & Zabludoff 1999; Smith et al. 2003; Jones et al. 2003). Although we await morphological confirmation, we note that these galaxies do have the colours typical of elliptical galaxies, $(g - r)_0 \sim 0.75 \pm 0.05$. In this case, their isolation may be a misleading representation of environment; in fact, they may represent the most dense environments, where all bright, surrounding galaxies have merged into one.

### 4.4 Dependence on large-scale structure

We showed in Section 3.2 that, although the galaxy population depends strongly on its local environment, it may also be sensitive to the density on $\gtrsim 5$ Mpc scales, especially at low densities. That is, fewer galaxies in supercluster-like environments have significant $H\alpha$ emission, relative to galaxies in environments with similar local densities. We currently treat this conclusion as tentative, because a similar trend could be caused by uncertainties in the correlated densities $\rho_{1.1}$ and $\rho_{5.5}$; however, comparison with mock catalogues strongly suggests that this is not the case (Balogh et al., in preparation). A similar dependence on large-scale structure has previously been observed in the morphologies of galaxies in clusters at $z \sim 0.2$ (Balogh et al. 2002c). Using high-resolution *Hubble Space Telescope* imaging to decompose the disc and bulge components of cluster members, Balogh et al. (2002c) found evidence that the bulges of galaxies at a fixed local density are systematically brighter in massive clusters (as traced by their X-ray emission) than in low-mass clusters. On the other hand, the disc luminosity function does not show this dependence. It was suggested that bulge components grow preferentially within large-scale overdensities, perhaps due to richer merger histories, but that the disc properties are not sensitive to this structure.

However, our results show that the galaxy population depends more strongly on the large-scale density than on the mass of the embedding halo, as evidenced by the lack of correlation with cluster velocity dispersion, at a fixed $\Sigma_5$ (Fig. 5). This, together with the results of Section 4.2, suggests that galaxy properties are only indirectly related to their environment at the present day. For example, the early-type population that dominates clusters today likely arises from the fact that galaxies in dense regions form earlier, and experience more galaxy–galaxy interactions throughout their longer lifetime, than galaxies in underdense regions (Zabludoff & Mulchaey 1998; Blanton et al. 1999; Balland, Devriendt & Silk). Another possibility is that galaxy evolution is strongly affected by tidal forces, which arise from structure on all scales (e.g. Moss & Whittle 2000; Gnedin 2003b).

### 4.5 Implications for the global SFR evolution

It has been proposed that changes in the typical environment, due to the hierarchical growth of structure, drive the evolution in global SFR (Balogh & Bower 2002; Bower & Balogh 2003). This hypothesis maintains that the correlation between SFR and local density remains unchanged, but galaxies at higher redshift typically lie in lower-density environments. However, this interpretation is inconsistent with the present data, as we demonstrate here. The fraction of star-forming galaxies [$W_0(H\alpha) > 4$ Å] is ~0.4 in the SDSS and ~0.57 in the 2dFGRS. This is close to the maximum fraction achieved in any environment at $z = 0$ (e.g. Figs 5 and 6; also Section 4.3). On the other hand, the average SFR at $z \sim 0.4$ is expected to be about 75 per cent higher than at the present day. This can be inferred from the increase in the ultraviolet luminosity density of

the Universe (e.g. Wilson et al. 2002); a similar result is obtained by comparing the average [O II] equivalent width at $z \sim 0.3$ from the CNOC surveys with that at $z = 0$ (Paper II). To achieve such a dramatic increase in the mean SFR, either the fraction of star-forming galaxies at $\sim 0.4$ is much higher than anywhere at the present day, or the typical star-forming galaxy has a higher SFR. It is interesting that Treu et al. (2003) find the fraction of morphologically classified early-type galaxies in low-density regions has not evolved substantially between $z = 0.4$ and $z = 0$, suggesting that evolution is of the latter type. Either way, the galaxy population at $z = 0.4$ must be different from that in any environment at $z = 0$.

A direct test of evolution in the SFR–density correlation is not yet possible, as intermediate- and high-redshift cluster data do not generally extend far enough from the cluster centre (Balogh et al. 1997; Couch et al. 2001; Balogh et al. 2002a,b). Kodama et al. (2001), however, have analysed multicolour *Subaru* data, which show that the red sequence in the $z \sim 0.4$ cluster Cl0939 first becomes apparent in regions where the local projected density exceeds 30 Mpc$^{-2}$ (see the Erratum of Kodama et al. 2003). Their photometric observations are much deeper than ours, and are complete to $\sim 0.02 L_V^*$. Correcting for this difference, their measurement of the critical density corresponds to $\Sigma_5 \sim 4 \pm 2$ Mpc$^{-2}$ in our units. This is similar to the result found in Papers I and II, and seen in Fig. 3, which suggests that any evolution in the SFR–density relation has been small. However, any serious interpretation of this comparison should wait until comparable, spectroscopic measurements can be made of the galaxy population composition as a continuous function of density.

### 4.6 The physics of galaxy evolution

Based on the results presented here, it seems unlikely that a substantial fraction of the star-forming galaxy population is today undergoing a physical transformation induced by its environment. Instead, the observed trends with density at $z \sim 0$ are probably only related indirectly to their environment, and the physics that determines the final composition of galaxy groups and clusters probably took place at a much earlier time. Similar conclusions were reached by Zabludoff & Mulchaey (1998), based on the high fractions of early-type galaxies in nearby, X-ray groups. Observations of the local Universe alone can only put weak constraints on what the relevant physics might be, except to say that it is unlikely to be dominated by ram-pressure stripping (e.g. Gunn & Gott 1972; Quilis, Moore & Bower 2000). This mechanism has the distinct advantage that there is good observational evidence for such activity (e.g. Veilleux et al. 1999; Vollmer et al. 2000; Gavazzi et al. 2002; van Gorkom 2003); however, it seems unlikely that it could play any role in the lowest-density environments.

Similarly, it may be difficult to interpret our results in the context of models in which SFR declines slowly in galaxies that are accreted into groups and clusters (Balogh et al. 2000; Diaferio et al. 2001; Bekki et al. 2002), for reasons discussed in Section 4.2. In particular, galaxy groups are expected to represent the first level of the hierarchy in which heating and stripping of the hot halo gas is likely to influence star formation (e.g. Blanton et al. 1999; Cole et al. 2000; Diaferio et al. 2001; Hernquist & Springel 2003), yet blue galaxies within such groups appear to have a normal $W_0(\text{H}\alpha)$ distribution, suggesting normal, recent star formation histories. However, the predictions of hierarchical models are not straightforward, because local density, halo mass and formation time are all correlated in a non-trivial way. It is possible that including the effects of infall (Ellingson et al. 2001) and projection effects (Diaferio et al. 2001), in addition to a slow-decay model (Balogh et al. 2000), conspire to

keep the shape of the $W_0(\text{H}\alpha)$ distribution constant. We will therefore leave a detailed comparison with theory for a subsequent paper (Balogh et al., in preparation).

A more viable explanation is perhaps suggested by observations of close galaxy pairs, which are the only environmentally selected population to show *enhancements* of star formation (Barton et al. 2000; Lambas et al. 2003). These bursts likely lead to the rapid consumption of cold gas, and the eventual formation of gas-poor elliptical galaxies with little star formation (e.g. Ponman & Bertram 1993; Mulchaey & Zabludoff 1999; Smith et al. 2003). Such galaxy–galaxy interactions must typically occur long before a galaxy is bound to a virialized group or cluster, since the SFRs in those environments are so low (e.g. Zabludoff & Mulchaey 1998). In fact, there is evidence for interaction-induced star formation in the unvirialized regions of clusters (e.g. Caldwell & Rose 1997; Moss & Whittle 2000), and for a correlation between morphological distortions and environment (Hashimoto & Oemler 2000). The observed correlation with local density then arises because galaxies in dense regions have typically had more interactions, over a longer period of time, than those in low-density regions (e.g. Balland et al. 2003; Gnedin 2003a,b). Since the interaction rate may increase substantially with redshift (Patton et al. 2002; Conselice et al. 2003), we would expect most of the environmentally induced galaxy transformation to have taken place at higher redshift; this would be consistent with the strong evolution observed in the fraction of post-starburst galaxies (Poggianti et al. 2003).

We might therefore be surprised that the correlation between emission-line fraction and local density is so similar in the groups in the present study, relative to other environments, since the low velocity dispersions and high densities in groups should encourage interactions (e.g. Aarseth & Fall 1980; Barnes 1985; Merritt 1985). Our groups are very heterogeneous, and not all are necessarily bound systems. It is possible that interactions are important in specific types of group, such as compact groups. We will leave such investigations for future work.

Finally, we note that there is no observed correlation between the fraction of AGN and environment at the present day (Miller et al. 2003). This is puzzling, especially in the context of models where star formation and AGN activity are both linked to the availability of cold gas (Kauffmann & Haehnelt 2000; Granato et al. 2003; Di Matteo et al. 2003). It appears that star formation activity must stop (or become undetectable) when there is still some cold gas left. One possibility is that the gas density drops below a certain threshold necessary for star formation (e.g. Kennicutt 1989; Martin & Kennicutt 2001; Wong & Blitz 2002; Boissier et al. 2003; Gnedin 2003b), but can still flow to the centre and fuel an AGN.

## 5  CONCLUSIONS

We have made a joint analysis of 24 968 galaxies selected from the SDSS and 2dFGRS. The distribution of $W_0(\text{H}\alpha)$ among these galaxies is bimodal, consisting of an active population of galaxies with $W_0(\text{H}\alpha) > 4$ Å, and a quiescent population with no significant star formation at the present day. We have investigated how this distribution depends on environment, as characterized by:

(i) the projected surface number density of galaxies $\Sigma_5$, determined from the distance to the fifth nearest neighbour;

(ii) the three-dimensional number density of galaxies measured on 1.1 and 5.5 Mpc scales, $\rho_{1.1}$ and $\rho_{5.5}$; and

(iii) the velocity dispersion of the embedding structure, as determined from the group catalogues of Eke et al. (2003) and Nichol et al. (in preparation).

We use these different measures of environment to establish the scales and structures on which the present-day galaxy population depends. Our findings are summarized as follows:

(1) The distribution of H$\alpha$ line strength for the star-forming population, selected on $W_0$(H$\alpha$) or $(g - r)$ colour, does not itself depend strongly on environment. Thus, it is unlikely that SFRs are gradually decreasing in a substantial number of star-forming galaxies in or near dense environments today.

(2) The fraction of galaxies with $W_0$(H$\alpha$) $>$ 4 Å decreases steadily with increasing local density. There is evidence that it decreases more strongly at densities exceeding $\Sigma_5 \sim 1$ Mpc$^{-2}$, or $\rho_{1.1} \gtrsim 0.05$ Mpc$^{-3}$. The persistence of this correlation at low densities means that ram-pressure stripping, at any redshift, cannot be the only physical mechanism at work.

(3) The fraction of galaxies brighter than $M* + 1$ with $W_0$(H$\alpha$) $>$ 4 Å is never more than $\sim$70 per cent, even in the least dense environments explored here. We have shown that this means the recent decline in globally averaged star formation rate cannot be wholly due to the growth of large-scale structure.

(4) The emission-line fraction of a galaxy population appears to depend both on the local environment (on $\sim$1 Mpc scales) *and* on the large-scale structure as parametrized by the density $\sim$5.5 Mpc scales. There is little further dependence on the velocity dispersion of the group or cluster in which the galaxy is embedded. This result suggests that the composition of the galaxy population today is likely related *indirectly* to its present environment (see also Zabludoff & Mulchaey 1998, 2000).

The most likely physical explanations for the correlation between $W_0$(H$\alpha$) distribution and environment at $z = 0$ are those which are effective over a large range of environment, affect the SFR on short ($<$1 Gyr) time-scales, and were much more effective in the past. As suggested by previous authors (e.g. Zabludoff et al. 1996; Mulchaey & Zabludoff 1998; Hashimoto et al. 1998; Hashimoto & Oemler 2000), one good candidate is starbursts induced by galaxy interactions, since close pairs of galaxies are the only environment known to provoke a physical transformation directly. Such interactions will likely be more common at high redshift, and will have had more time to influence galaxies that end up in high-density environments.

## ACKNOWLEDGMENTS

## REFERENCES

Aarseth S. J., Fall S. M., 1980, ApJ, 236, 43
Abazajian K. et al., 2003, AJ, 126, 2081
Afonso J., Hopkins A., Mobasher B., Almeida C., 2003, ApJ, 597, 269

Allington-Smith J. R., Ellis R., Zirbel E. L., Oemler A. J., 1993, ApJ, 404, 521
Baldry I. K., Glazebrook K., Brinkmann J., Ivezić, Z., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, in press (astro-ph/0309710)
Balland C., Devriendt J. E. G., Silk J., 2003, MNRAS, 343, 107
Balogh M. L., Bower R. G., 2002, in Avila-Reese C. F. V., Firmani C., Allen C., eds, Rev. Mex. Astron. Astrofis. Ser. Conf. Vol. 17, Galaxy Evolution: Theory and Observations. UNAM, Mexico, p. 220
Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1997, ApJ, 488, L75
Balogh M. L., Schade D., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1998, ApJ, 504, L75
Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1999, ApJ, 527, 54
Balogh M. L., Navarro J. F., Morris S. L., 2000, ApJ, 540, 113
Balogh M., Bower R. G. et al., 2002a, MNRAS, 337, 256
Balogh M. L., Couch W. J., Smail I., Bower R. G., Glazebrook, K., 2002b, MNRAS, 335, 10
Balogh M. L., Smail I., Bower R. G. et al., 2002, ApJ, 566, 123
Barnes J., 1985, MNRAS, 215, 517
Barton E. J., Geller M. J., Kenyon S. J., 2000, ApJ, 530, 660
Baugh C. M., Benson A. J., Cole S., Frenk C. S., Lacey C. G., 1999, MNRAS, 305, L21
Beers T. C., Flynn K., Gebhardt K., 1990, AJ, 100, 32
Bekki K., Couch W. J., Shioya Y., 2002, ApJ, 577, 651
Benson A. J., Hoyle F., Torres F., Vogeley M. S., 2003, MNRAS, 340, 160
Blanton M., Cen R., Ostriker J. P., Strauss M. A., 1999, ApJ, 522, 590
Blanton M. R., Hogg D. W. et al., 2003a, ApJ, 592, 819
Blanton M. R. et al., 2003b, ApJ, 594, 186
Boissier S., Prantzos N., Boselli A., Gavazzi G., 2003, MNRAS, 346, 1215
Bower R. G., Balogh M. L., 2003, in Mulchaey A. D. J. S., Oemler A., eds, Carnegie Obs. Astrophys. Ser. Vol. 3, Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution. astro-ph/0306342
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Caldwell N., Rose J. A., 1997, AJ, 113, 492
Carlberg R. G., Yee H. K. C., Morris S. L. et al., 2001, ApJ, 563, 736
Charlot S., Longhetti M., 2001, MNRAS, 323, 887
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Colless M. et al., 2001, MNRAS, 328, 1039
Colless M. et al., 2003, astro-ph/0306581
Conselice C. J., Bershady M. A., Dickinson M., Papovich C., 2003, AJ, 126, 1183
Couch W. J., Balogh M. L., Bower R. G., Smail I., Glazebrook K., Taylor M., 2001, ApJ, 549, 820
Cowie L. L., Songaila A., Barger A. J., 1999, AJ, 118, 603
Coziol R., Iovino A., de Carvalho R. R., 2000, AJ, 120, 47
de la Rosa I. G., de Carvalho R. R., Zepf S. E., 2001, AJ, 122, 93
Diaferio A., Kauffmann G., Balogh M. L., White S. D. M., Schade D., Ellingson E., 2001, MNRAS, 323, 999
Di Matteo T., Croft R. A. C., Springel V., Hernquist L., 2003, ApJ, 593, 56
Domínguez M., Muriel H., Lambas D. G., 2001, AJ, 121, 1266
Domínguez M. J., Zandivarez A. A., Martínez H. J., Merchán M. E., Muriel H., Lambas D. G., 2002, MNRAS, 335, 825
Dressler A., 1980, ApJ, 236, 351
Eke V. R., Baugh C., Cole S., Frenk C. S., Norberg P., Peacock J. A. et al., 2003, MNRAS, in press
Ellingson E., Lin H., Yee H. K. C., Carlberg R. G., 2001, ApJ, 547, 609
Fukugita M., Shimasaku K., Ichikawa T., 1995, PASP, 107, 945
Gavazzi G., Boselli A., Pedotti P., Gallazzi A., Carrasco L., 2002, A&A, 396, 449
Geller M. J., Huchra J. P., 1983, ApJS, 52, 61
Girardi M., Giuricin G., Mardirossian F., Mezzetti M., Boschin W., 1998, ApJ, 505, 74
Girardi M., Manzato P., Mezzetti M., Giuricin G., Limboz F., 2002, ApJ, 569, 720
Girardi M., Rigoni E., Mardirossian F., Mezzetti M., 2003, A&A, 406, 403

Gladders M. D., Yee H. K. C., 2000, AJ, 120, 2148

Gnedin O. Y., 2003a, ApJ, 582, 141

Gnedin O. Y., 2003b, ApJ, 589, 752

Gomez P. L., Nichol R. C., Miller C. J., Balogh M. L. et al., 2003, ApJ, 584, 210 (Paper II)

Goto T., Nichol R. C. et al., 2003, PASJ, 55, 771

Granato G. L., De Zotti G., Silva L., Bressan A., Danese L., 2003, ApJ, in press (astro-ph/0307202)

Gray A., Moore A. W., 2003a, in Barbara D., Kanath C., eds, Proc. Third SIAM International Conference on Data Mining. San Francisco

Gray A., Moore A. W., 2003b, in Proc. ASA Joint Statistical Meeting, Very Fast Multivariate Kernel Density Estimation via Computational Geometry. Am. Statistical Association, San Francisco

Gunn J. E., Gott J. R., 1972, ApJ, 176, 1

Hashimoto Y., Oemler A. J., 1999, ApJ, 510, 609

Hashimoto Y., Oemler A. J., 2000, ApJ, 530, 652

Hashimoto Y., Oemler A., Lin H., Tucker D. L., 1998, ApJ, 499, 589

Hernquist L., Springel V., 2003, MNRAS, 341, 1253

Hickson P., 1982, ApJ, 255, 382

Hopkins A. M., Connolly A. J., Haarsma D. B., Cram L. E., 2001, AJ, 122, 288

Hopkins A. M., Miller C. J., Nichol R. C. et al., 2003, ApJ, in press (astro-ph/0306621)

Iglesias-Páramo J., Vílchez J. M., 1999, ApJ, 518, 94

Jones L. R., Ponman T. J., Horton A., Babul A., Ebeling H., Burke D. J., 2003, MNRAS, 343, 627

Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576

Kauffmann G., Heckman T. M., White S. D. M. et al., 2003, MNRAS, 341, 54

Kelly B. C., McKay T. A., 2004, AJ, in press (astro-ph/0307395)

Kelm B., Focardi P., 2003, A&A, submitted (astro-ph/0306414)

Kennicutt R. C., 1983, ApJ, 272, 54

Kennicutt R. C., 1989, ApJ, 337, 761

Kennicutt R. C., 1998, ARA&A, 36, 189

Kodama T., Smail I., Nakata F., Okamura S., Bower R. G., 2001, ApJ, 562, L9

Kodama T., Smail I., Nakata F., Okamura S., Bower R. G., 2003, ApJ, 591, L169

Lambas D. G., Tissera P. B., Sol Alonso M., Coldwell G., 2003, MNRAS, 346, 1189

Lewis I. J. et al., 2002a, MNRAS, 333, 279

Lewis I., Balogh M., De Propris R., Couch W., Bower R. et al., 2002b, MNRAS, 334, 673 (Paper I)

Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJ, 460, L1

Liske J., Lemon D. J., Driver S. P., Cross N. J. G., Couch W. J., 2003, MNRAS, 344, 307

McIntosh D., Rix H.-W., Caldwell N., 2003, ApJ, submitted (astro-ph/0212427)

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter A., 1996, MNRAS, 283, 1388

Martin C. L., Kennicutt R. C., 2001, ApJ, 555, 301

Martínez H. J., Zandivarez A., Domínguez M., Merchán M. E., Lambas D. G., 2002, MNRAS, 333, L31

Mateus A. J., Sodré L. J., 2003, MNRAS, submitted (astro-ph/0307349)

Merchán M., Zandivarez A., 2002, MNRAS, 335, 216

Merritt D., 1985, ApJ, 289, 18

Miller C. J. et al., 2001, AJ, 122, 3492

Miller C. J., Nichol R. C., Gómez P. L., Hopkins A. M., Bernardi M., 2003, ApJ, 597, 142

Moss C., Whittle M., 1993, ApJ, 407, L17

Moss C., Whittle M., 2000, MNRAS, 317, 667

Mulchaey J. S., Zabludoff A. I., 1998, ApJ, 496, 73

Mulchaey J. S., Zabludoff A. I., 1999, ApJ, 514, 133

Nishiura S., Shimada M., Ohyama Y., Murayama T., Taniguchi Y., 2000, AJ, 120, 1691

Norberg P. et al., 2002, MNRAS, 336, 907

O'Hely E., 2000, PhD thesis, Univ. New South Wales

Patton D. R., Pritchet C. J., Carlberg R. G. et al., 2002, ApJ, 565, 208

Pimbblet K. A., Smail I., Kodama T., Couch W. J., Edge A. C., Zabludoff A. I., O'Hely E., 2002, MNRAS, 331, 333

Poggianti B. M., Smail I., Dressler A., Couch W. J., Barger A. J., Butcher H., Ellis R. S., Oemler A. J., 1999, ApJ, 518, 576

Poggianti B. M., Bridges T. J., Komiyama Y., Yagi M., Carter D., Mobasher B., Okamura S., Kashikawa N., 2003, ApJ, in press (astro-ph/0309449)

Ponman T. J., Bertram D., 1993, Nat, 363, 51

Postman M., Geller M. J., 1984, ApJ, 281, 95

Quilis V., Moore B., Bower R., 2000, Sci, 288, 1617

Rocha-Pinto H. J., Scalo J., Maciel W. J., Flynn C., 2000, A&A, 358, 869

Rubin V. C., Hunter D. A., Ford W. K. J., 1991, ApJS, 76, 153

Shioya Y., Bekki K., Couch W. J., De Propris R., 2002, ApJ, 565, 223

Silverman B. W., 1986, Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York, pp. 61–66

Smith B. J., Nowak M., Donahue M., Stocke J., 2003, AJ, 126, 1763

Solanes J., Manrique A., García-Gómez C., González-Casado G., Giovanelli R., Haynes M. P., 2001, ApJ, 548, 97

Strateva I. et al., 2001, AJ, 122, 1861

Strauss M. A. et al., 2002, AJ, 124, 1810

Tran K. H., Simard L., Zabludoff A. I., Mulchaey J. S., 2001, ApJ, 549, 172

Treu T., Ellis R. S., Kneib J., Dressler A., Smail I., Czoske O., Oemler A., Natarajan P., 2003, ApJ, 591, 53

van Gorkom J. H., 2003, in Mulchaey A. D. J. S., Oemler A., eds, Carnegie Obs. Astrophys. Ser. Vol. 3, Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution. astro-ph/0308209

Veilleux S., Bland-Hawthorn J., Cecil G., Tully R. B., Miller S. T., 1999, ApJ, 520, 111

Verdes-Montenegro L., Yun M. S., Williams B. A., Huchtmeier W. K., Del Olmo A., Perea J., 2001, A&A, 377, 812

Vollmer B., Marcelin M., Amram P., Balkowski C., Cayatte V., Garrido O., 2000, A&A, 364, 532

Wainer H., Thissen D., 1976, Psychometrika, 41, 9

Wasserman L. et al., 2001, astro-ph/0112050

Wilson G., Cowie L. L., Barger A. J., Burke D. J., 2002, AJ, 124, 1258

Wong T., Blitz L., 2002, ApJ, 569, 157

York D. G. et al., 2000, AJ, 120, 1579

Zabludoff A. I., Mulchaey J. S., 1998, ApJ, 496, 39

Zabludoff A. I., Mulchaey J. S., 2000, ApJ, 539, 136

Zabludoff A. I., Zaritsky D., Lin H., Tucker D., Hashimoto Y., Shectman S. A., Oemler A., Kirshner R. P., 1996, ApJ, 466, 104

## APPENDIX A: APERTURE EFFECTS

Since the spectra are obtained from 2 arcsec (2dFGRS) or 3 arcsec (SDSS) fibres, they do not represent the light from the whole galaxy. In particular, our measurements of Hα flux will underestimate the flux from the galaxy as a whole; we refer to this as the aperture bias. Aperture bias will not impact on our results if the average galaxy size does not correlate strongly with environment. We test this in Fig. A1, where we plot the SDSS galaxy Petrosian radius as a function of $\Sigma_5$ and of group velocity dispersion. Note that Hopkins et al. (2003) find that the amount of emission-line flux lost due to aperture bias in the SDSS is a factor of ∼2 for galaxies with radius ∼3 arcsec, increasing to a factor of ∼10 for galaxies with radius >12 arcsec. However, there is no significant trend of galaxy size with environment. Therefore, although the absolute values of $W_0(Hα)$ may be affected[6] by aperture bias, the relative trends as a function of environment are secure. This conclusion requires the further assumption that the spatial distribution of star formation in a galaxy depends only on its SFR, and not its environment. For

---

[6] Note that $W_0(Hα)$ is not necessarily underestimated, as it is a relative quantity. The effect of aperture bias depends on the spatial distribution of star formation; if it is uniform, $W_0(Hα)$ is not affected by this bias.
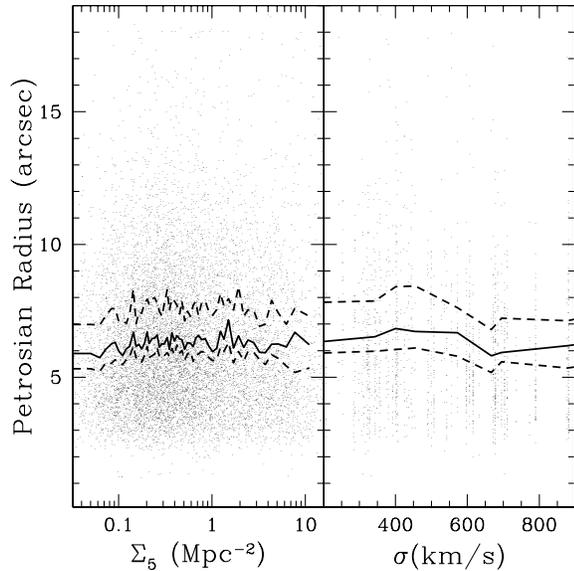
**Figure A1.** The Petrosian radius is shown as a function of local, projected density $\Sigma_5$ (left panel) and of group velocity dispersion (right panel), for galaxies in the SDSS sample. The dashed lines show the median and 75th percentile of the distribution, in running bins each containing 500 galaxies; the solid line represents the mean.

example, if star formation in cluster galaxies is more concentrated than for galaxies in the field with similar SFRs (e.g. Moss & Whittle 1993, 2000), the aperture corrections for cluster galaxies should be smaller.

## APPENDIX B: DETAILED COMPARISON OF THE CATALOGUES

Here we compare the 2dFGRS and SDSS samples in detail. The most important difference between them for our purposes is that the 2dFGRS is selected from $b_J$ photometric magnitudes, while the SDSS is selected from digital $r$ photometry. The SDSS survey is somewhat shallower, and our volume-limited sample is complete to $M_r = -20.6$. From the tables of Fukugita, Shimasaku & Ichikawa (1995), an Sab galaxy at $z = 0.06$ has a colour $(b_J - r) = 1.13$, so we adopt a magnitude limit of $M_b = -19.5$ for the 2dFGRS sample; the survey is actually complete to 0.5 magnitudes fainter than this.

In the area of sky covering approximately $155 < \alpha < 220$ and $-3 < \delta < 1.5$ degrees, the 2dFGRS and SDSS surveys overlap. We can use this overlap region to compare the galaxy and group catalogues directly. In our volume-limited subsample, we find 1029 galaxies in common between the two surveys (within 2 arcsec and 100 km s$^{-1}$ of one another). The distribution of the difference between the absolute 2dFGRS $b_J$ magnitude and the absolute SDSS $r$ magnitude is shown in Fig. B1. The average colour is $M_{b_J} - M_r \sim 0.95$, similar to the colour we expected (see above). Thus we can be confident that our magnitude limits are matched as closely as possible. However, the $1\sigma$ standard deviation in colour is $\sim 0.35$; thus there will be large numbers of galaxies at the faint limit of one catalogue that will not be included in the other.

The consequences of this are shown in Fig. B2, where we compare the distributions of $W_0(H\alpha)$ in the two samples. Recall that an absorption correction of 1 Å has been applied to all measurements (see Section 2). The two distributions are very similar; however, there are relatively more galaxies with substantial $H\alpha$ emission in
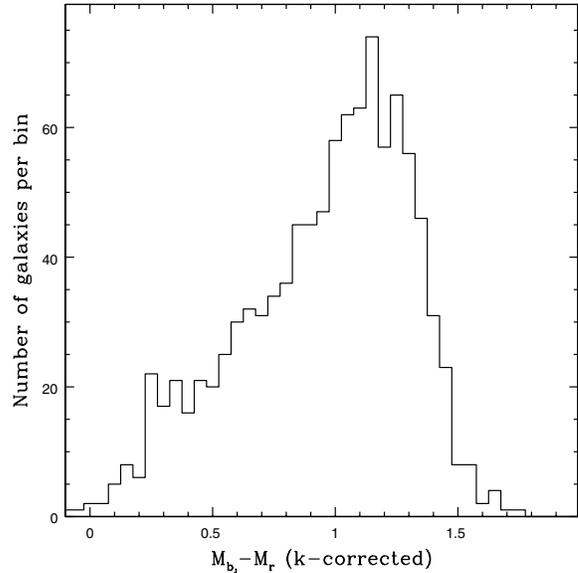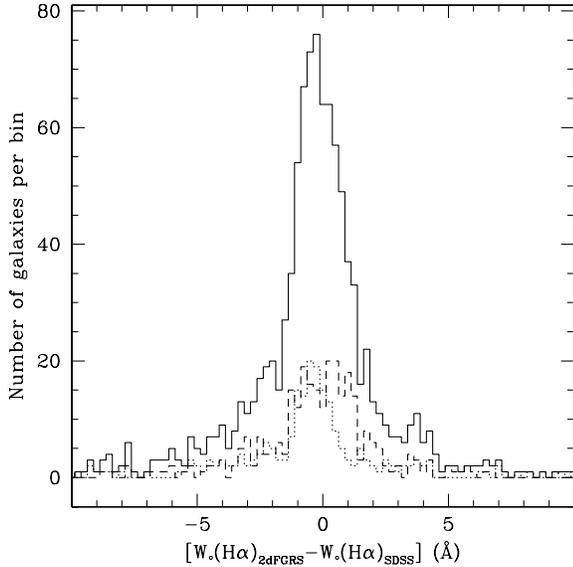


**Figure B1.** The distribution of the differences between the absolute magnitudes of 1460 galaxies in our sample that are identified in both SDSS and 2dFGRS catalogues.
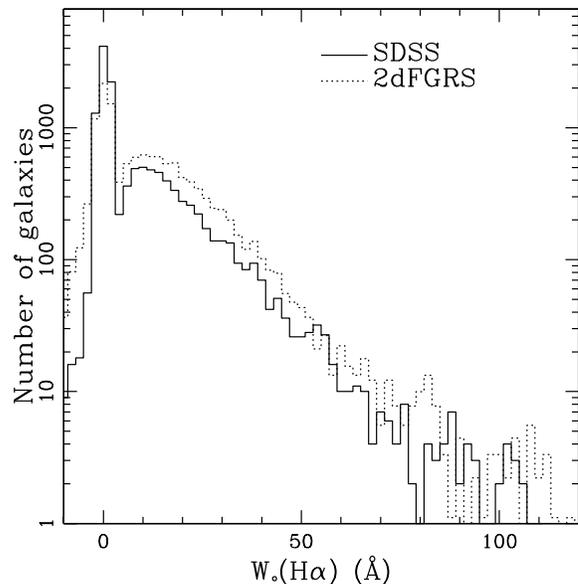


**Figure B2.** The distribution of $H\alpha$ equivalent width in the two samples. A 1 Å correction for underlying stellar absorption has been made in both cases. The 2dFGRS distribution has been renormalized to match the number of galaxies in the SDSS, to facilitate the comparison.

the 2dFGRS sample. The mean equivalent width is 8.1 Å in the SDSS, and 11.9 Å in the 2dFGRS; a similar difference is seen in the 75th percentiles, which are also $\sim 30$ per cent larger in the 2dF-GRS sample. The larger errors on the 2dFGRS spectra also result in a larger population of galaxies with $W_0(H\alpha) < 0$ Å, at the expense of galaxies in the peak (with $W_0(H\alpha) = 0$ Å).

The difference in Fig. B2 is due entirely to the blue selection of the 2dFGRS. In Fig. B3 we directly compare the $W_0(H\alpha)$ for galaxies in common between the two surveys. Most (67 per cent) of the measurements are within 2 Å of one another. In the mean, the SDSS measurement of $W_0(H\alpha)$ is 0.6 Å larger than that of the

**Figure B3.** The difference between $W_0(H\alpha)$ measured in the 2dFGRS and SDSS samples, for 1029 galaxies in common. Thirty-three galaxies lie outside the bounds of this diagram. The dotted and dashed lines show the same distribution, but restricted to galaxies at $z < 0.075$ and $z > 0.085$, respectively.
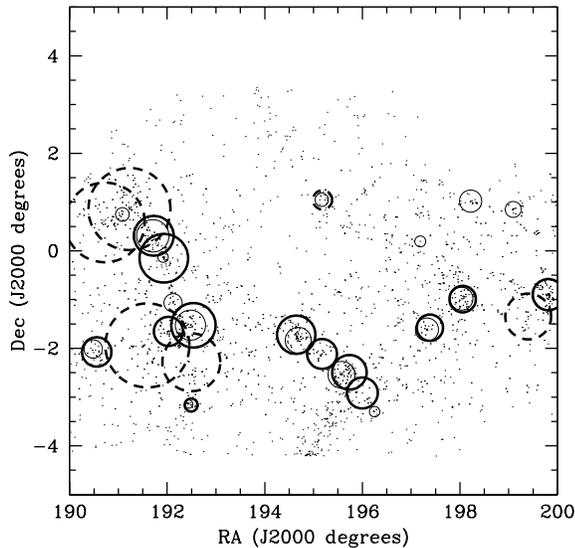


**Figure B4.** Galaxies are shown in a 25 deg$^2$ region of space where the 2dFGRS and SDSS overlap. Galaxy groups in the 2dFGRS sample, with $\sigma > 200$ km s$^{-1}$ and at least 10 bright members, are shown as thin/feint circles, with a radius corresponding to the virial radius (proportional to velocity dispersion). The thick circles are SDSS groups; dashed circles have significant substructure and poorly determined velocity dispersions.

2dFGRS. This difference is too small (and in the wrong sense) to account for the difference in Fig. B2. We also show the comparison for galaxies at $z < 0.075$ and $z > 0.085$; there is no discernible difference between them. Since the 2dFGRS and SDSS use different size fibres, this gives added confidence that aperture effects do not have a large influence on our results.

We can also make a direct comparison between the group catalogues. In Fig. B4 we show the positions of groups in both surveys, within a $10 \times 7$ deg$^2$ region. In both cases we show all groups within

$0.05 < z < 0.095$ and with $\sigma > 200$ km s$^{-1}$ and at least 10 members above our luminosity limit. For the SDSS we show all the groups satisfying these criteria; however, in this paper we have only used those with well-determined velocity dispersions, shown as the solid circles. Within the full overlap region there are 29 such 2dFGRS groups, compared with 24 SDSS groups; 16 of the latter have well-determined velocity dispersions. The radii of the circles represent the cluster virial radius, estimated from the velocity dispersion as $R_{vir} = 3.5\,\sigma\,(1 + z)^{-1.5}$ (Girardi et al. 1998). The correspondence between the two group catalogues is remarkably good.

## APPENDIX C: DENSITY ESTIMATORS

In this appendix, we consider two different density estimators in turn. The first is a projected surface density, based on counting the number of nearest neighbours. The second is a three-dimensional, fixed-scale estimator derived by kernel density estimation.

### C1 Nearest-neighbour approach

The most common method is to measure the distance to the $N$th nearest neighbour, and measure the density within that distance. Dressler's (1980) original prescription called for taking the area to be a box enclosing the 10th nearest neighbour, before background subtraction. Others have employed similar methods, but by using a spherical area and first subtracting foreground and background galaxies (e.g. Papers I and II). This latter difference means that densities are typically measured in a larger volume, since the 10th nearest neighbour is further away after background subtraction. In general, we will define $d_N$ as the distance to the $N$th nearest neighbour after background subtraction; then the projected density is given by $\Sigma_N = N/(\pi\,d^2_N)$.

Although this is a two-dimensional estimate, we wish to use the redshift information to remove the foreground and background. In some cases, this has been done by removing galaxies more than $3\sigma$ from the cluster (e.g. Paper I; Balogh et al. 2002a), where $\sigma$ is the cluster velocity dispersion. However, this is only reliable if $\sigma$ is well known, and the velocity distribution is Gaussian. We will therefore choose a fixed velocity interval of $\pm 1000$ km s$^{-1}$ within which to compute the local density. This allows us to include most of the galaxies in systems with large velocity dispersions, while still keeping contamination low. We choose $N = 5$ to be similar to that defined by Dressler (1980), who used $N = 10$ *before* background subtraction.

### C1.1 Projection effects

Our density estimator $\Sigma_5$ is a projected quantity, including all galaxies within $\pm 1000$ km s$^{-1}$ of the target galaxy, or $\pm 14$ Mpc for $H_0 = 70$ km s$^{-1}$ Mpc. We will use the 2dFGRS as an example (an analogous calculation can be done for the SDSS) and estimate the contamination at a given measured density $\Sigma_5$ by assuming that background galaxies are distributed at the mean galaxy density. From the galaxy luminosity function of Norberg et al. (2002), we calculate that the number of galaxies within a projected cylinder of radius $d$ is $\sim 0.17d^2$. If we take $d_5$ to be the distance to the fifth nearest neighbour, then the fraction of projected galaxies within $d_5$ is given by $f_{proj} = 0.17/(\pi\,\Sigma_5)$. This varies from 100 per cent at $\Sigma_5 = 0.054$ Mpc$^{-2}$ (which therefore represents the field density) to 5 per cent by $\Sigma_5 = 1$ Mpc$^{-2}$. In environments typical of the field, the
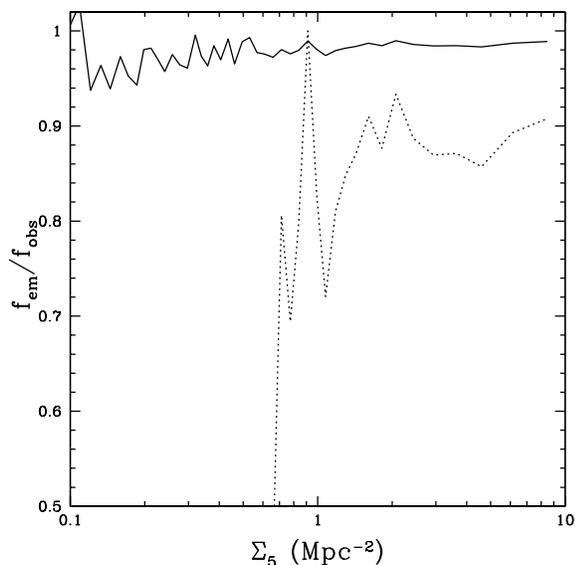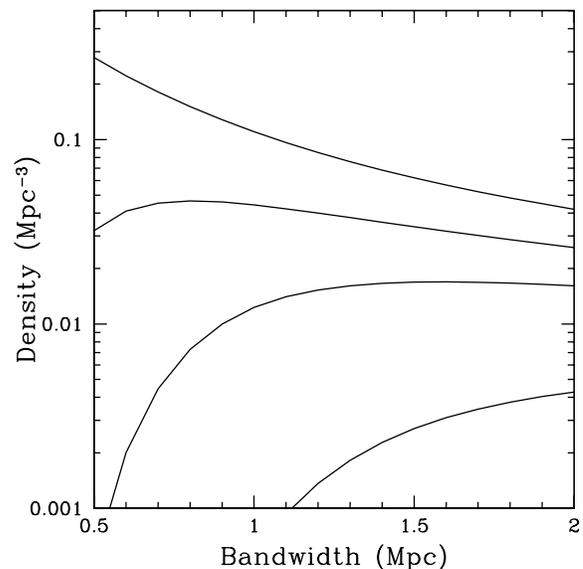
**Figure C1.** The ratio between the true fraction of emission-line galaxies $[W_0(H\alpha) > 4$ Å$]$ and the observed fraction, which is contaminated by projection along the line of sight, for the 2dFGRS. The solid line represents the case where the projected galaxies are at the average density of the Universe, while the dotted line is calculated assuming this density is 10 times larger.



**Figure C2.** The dependence of three-dimensional density estimate on bandwidth, for the combined sample. Here, the bandwidth is the standard deviation of the kernel. The sample is divided into four based on the quartiles of the density at $\theta = 2$ Mpc. Thus, the top curve shows the average density of the 25 per cent of galaxies that have the greatest density at $\theta = 2$ Mpc, while the bottom curve corresponds to the 25 per cent with the lowest density at that point.

fraction of emission-line galaxies is $\sim 70$ per cent (Fig. 5). Therefore, the true, unprojected emission-line fraction, $f_{em}$, is related to the observed fraction, $f_{obs}$, by $f_{em} = (f_{obs} - 0.7f_{proj}) / (1 - f_{proj})$. In Fig. C1 we show how the ratio $f_{em}/f_{obs}$ depends on $\Sigma_5$. The correction is small, $\lesssim 5$ per cent at all values of $\Sigma_5$. We also show how this ratio changes if we assume that the background is 10 times more dense than the average, which might be the case in the vicinity of clusters and groups. However, in this case we should adopt a smaller intrinsic emission-line fraction, because the projected galaxies are themselves in a more dense environment; based on Fig. 5 we take this to be 60 per cent. Even in this case, the correction required to account for projection is $< 25$ per cent for $\Sigma_5 > 1$ Mpc$^{-2}$. In reality, such a dense field projection is probably only reasonable at the highest $\Sigma_5$, where the correction is $\lesssim 15$ per cent. We conclude that projection will only have a small effect on the trends observed in Fig. 5 and similar figures.
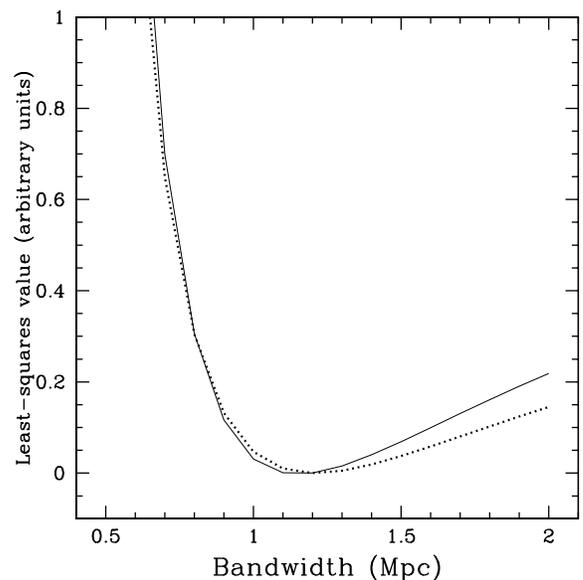
## C2 Fixed scale estimates: kernel density estimation

In this section, we will consider density estimates within a fixed distance of a galaxy. We will use a Gaussian filtering kernel which has some weight in the wings, so that more weight is given to galaxies that are closer, while there is still sensitivity to more distant structures. A rigorous statistical analysis shows that the density estimate, for any data distribution, is sensitive mostly to the bandwidth $\theta$ of the smoothing kernel, and much less sensitive to the shape of the kernel (Silverman 1986). The strong bandwidth dependence is shown explicitly in Fig. C2, where we plot the average density as a function of $\theta$. Four curves are shown, corresponding to the four quartiles of the density distribution evaluated at $\theta = 2$ Mpc.

In general, smaller bandwidths result in greater densities. Thus, what do we choose for the optimal bandwidth? In abstract terms, the problem is to find an estimate $\hat{f}$ of the underlying density distribution function $f$ that minimizes the average value of the mean
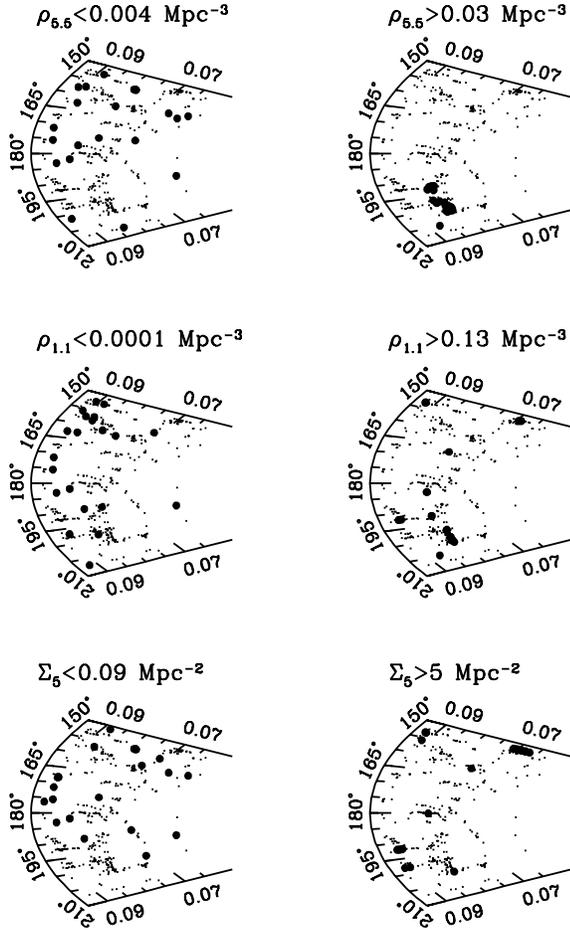


**Figure C3.** The least-squares cross-validation score (arbitrarily scaled), as a function of bandwidth, for the SDSS and 2dFGRS samples (solid and dotted lines, respectively). Both curves show a clear minimum near $\sim 1.1$ Mpc.

square error:

$$\text{MSE}(f, \hat{f}) = \left\langle \int [f(x) - \hat{f}(x)]^2 \, dx \right\rangle. \tag{C1}$$

We want to use the actual data to find the optimal bandwidth, $\theta$, which minimizes this error. Since $\int f^2(x) \, dx$ does not depend on $\theta$, this corresponds to minimizing the function:

$$J(\theta) = \int \hat{f}^2(x) \, dz - 2 \int \hat{f}(x) f(x) \, dx. \tag{C2}$$

$\rho_{5.5} < 0.004 \ \mathrm{Mpc}^{-3}$

$\rho_{5.5} > 0.03 \ \mathrm{Mpc}^{-3}$

$\rho_{1.1} < 0.0001 \ \mathrm{Mpc}^{-3}$

$\rho_{1.1} > 0.13 \ \mathrm{Mpc}^{-3}$

$\Sigma_5 < 0.09 \ \mathrm{Mpc}^{-2}$

$\Sigma_5 > 5 \ \mathrm{Mpc}^{-2}$

**Figure C4.** Pie diagrams (with redshift along the radial direction and right ascension as the angle) for galaxies in the 2dFGRS within 0.5° of zero declination, and within 11 Mpc of a survey boundary. The filled circles show galaxies in different density ranges, as indicated above each panel.

The method of least-squares cross-validation uses the data $X_i$ to obtain an unbiased estimate of this function:

$$\hat{J}(\theta) \approx \frac{1}{\theta n^2} \sum_i \sum_j K^* \left( \frac{X_i - X_j}{\theta} \right) + \frac{2}{n\theta} K(0), \qquad (C3)$$

where $K(x)$ is the kernel function, and $K^* = K(x) * K(x) - 2K(x)$ (Wasserman et al. 2001). This function can be computed using the fast Fourier transform (Silverman 1986), but this method is relatively slow, and can be inaccurate due to a number of necessary approximations. We have implemented a new, substantially improved algorithm based on adaptive computational geometry and a hierarchical finite-difference approximation (Gray & Moore 2003a,b).

We performed a cross-validation analysis of our volume-limited, combined sample. We use a Gaussian kernel for the cross-validation, but recall that the optimal value of $\theta$ is insensitive to the exact kernel choice. Here, we are computing three-dimensional densities $\rho_\theta$, assuming the redshifts indicate line-of-sight position in a $\Lambda$CDM model, and ignoring the effects of redshift distortions. Fig. C3 shows the least-squares value as a function of bandwidth. We see that the optimal bandwidth is $\sim 1.1$ Mpc, which we will therefore use as our best estimate of the three-dimensional local density. To measure the sensitivity of galaxy properties to larger scales, though, we will also consider densities measured on 5.5 Mpc scales, $\rho_{5.5}$.

Note that the density estimate at a given galaxy may be arbitrarily low, since the density at that point is measured from the surrounding data, not including the point itself. Densities become significantly underestimated if the width of the Gaussian, $\theta$, is comparable to the distance to the nearest boundary. We therefore only consider galaxies farther than $2\theta$ from a survey boundary when using these estimators.

To see more directly what the density estimates on different scales are measuring, we show, in Fig. C4, pie diagrams for a section of the 2dFGRS. In each diagram, we show all galaxies that lie at least 11 Mpc from the survey boundary.[7] Marked as thick, black circles are galaxies in different density environments, as labelled. For each of the three density estimators, we show galaxies in the highest and lowest $\sim 6$ per cent density environments. High values of both $\Sigma_5$ and $\rho_{1.1}$ are good at finding relatively small, but dense clumps. At low densities, both estimators are good at finding isolated galaxies; however, $\Sigma_5$ is unable to distinguish truly isolated galaxies from those in low-density filaments in the plane of the sky, while $\rho_{1.1}$ measures as low-density some cluster galaxies that are elongated along the line of sight. In contrast with these, $\rho_{5.5}$ is sensitive to much larger scales. The densest galaxies, as measured on this scale, are only those in the massive supercluster, while the lowest-density environments are true voids, far from filaments or clusters.

[7] This motivates our choice to show the 2dFGRS data rather than the SDSS data. The SDSS DR1 catalogue is much more patchy, and there are few large, contiguous regions where all galaxies are at least 11 Mpc from a boundary.

This paper has been typeset from a TEX/LATEX file prepared by the author.