The clustering and evolution of Hα emitters at $z \sim 1$ from HiZELS[★]

David Sobral,[1]† Philip N. Best,[1] James E. Geach,[2] Ian Smail,[2] Michele Cirasuolo,[3] Timothy Garn,[1] Gavin B. Dalton[4,5] and Jaron Kurk[6]

[1]*SUPA, Institute for Astronomy, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[2]*Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE*
[3]*Astronomy Technology Centre, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[4]*Astrophysics, Department of Physics, Keble Road, Oxford OX1 3RH*
[5]*Space Science and Technology, Rutherford Appleton Laboratory, HSIC, Didcot OX11 0QX*
[6]*Max-Planck-Institut für Extraterrestrische Physik, Postfach 1312, 85741 Garching, Germany*

ABSTRACT

The clustering properties of a well-defined sample of 734 Hα emitters at $z = 0.845 \pm 0.015$, obtained as part of the Hi-z Emission Line Survey, are investigated. The spatial correlation function of these Hα emitters is very well described by the power-law $\xi = (r/r_0)^{-1.8}$, with a real-space correlation, r_0, of $2.7 \pm 0.3\, h^{-1}$ Mpc. The correlation length r_0 increases strongly with Hα luminosity ($L_{\mathrm{H}\alpha}$), from $r_0 \sim 2\, h^{-1}$ Mpc for the most quiescent galaxies [star formation rates (SFRs) of $\sim 4\, \mathrm{M}_\odot\, \mathrm{yr}^{-1}$] up to $r_0 > 5\, h^{-1}$ Mpc for the brightest galaxies in Hα. The correlation length also increases with increasing rest-frame K-band (M_K) luminosity, but the r_0–$L_{\mathrm{H}\alpha}$ correlation maintains its full statistical significance at fixed M_K. At $z = 0.84$, star-forming galaxies classified as irregulars or mergers are much more clustered than discs and non-mergers, respectively; however, once the samples are matched in $L_{\mathrm{H}\alpha}$ and M_K, the differences vanish, implying that the clustering is independent of morphological type at $z \sim 1$ just as in the local Universe. The typical Hα emitters found at $z = 0.84$ reside in dark matter haloes of $\approx 10^{12}\, \mathrm{M}_\odot$, but those with the highest SFRs reside in more massive haloes of $\approx 10^{13}\, \mathrm{M}_\odot$. The results are compared with those of Hα surveys at different redshifts: although the break of the Hα luminosity function $L_{\mathrm{H}\alpha}^*$ evolves by a factor of ~ 30 from $z = 0.24$ to 2.23, if the Hα luminosities at each redshift are scaled by $L_{\mathrm{H}\alpha}^*(z)$ then the correlation lengths indicate that, independently of cosmic time, galaxies with the same $(L_{\mathrm{H}\alpha})/L_{\mathrm{H}\alpha}^*(z)$ are found in dark matter haloes of similar masses. This not only confirms that the star formation efficiency in high redshift haloes is higher than locally but also suggests a fundamental connection between the strong negative evolution of $L_{\mathrm{H}\alpha}^*$ since $z = 2.23$ and the quenching of star formation in galaxies residing within dark matter haloes significantly more massive than $10^{12}\, \mathrm{M}_\odot$ at any given epoch.

Key words: galaxies: evolution – galaxies: haloes – galaxies: high-redshift – cosmology: observations – large-scale structure of Universe.

1 INTRODUCTION

In a Universe dominated by cold dark matter (CDM), galaxies are found in dark matter haloes with a structure determined by universal scaling relations (e.g. Navarro, Frenk & White 1996). Since baryons trace the underlying distribution of dark matter, measurements of the clustering of baryonic matter can be used to extract typical dark matter halo masses (Mo & White 1996; Sheth, Mo & Tormen 2001) and to suggest links between populations found at different epochs.

Wide surveys of the nearby Universe [e.g. 2dF, Sloan Digital Sky Survey (SDSS); York et al. 2000; Colless et al. 2001] have now assembled extremely large samples of galaxies which can be explored to study their clustering properties in great detail. Using those data, several studies have found that the amplitude of the two-point correlation function rises continuously with galaxy luminosity (e.g. Norberg et al. 2001; Zehavi et al. 2005; Li et al. 2006). They also reveal that the most rapid increase occurs above the characteristic luminosity, L^*. Red galaxies are found to be more clustered than blue galaxies, but the correlation amplitude of the latter population

also increases continuously with blue or near-infrared (near-IR) luminosities (e.g. Zehavi et al. 2005). This seems to indicate that both star formation rate [SFR; traced by ultraviolet (UV)/blue light] and stellar mass (traced by near-IR light) are important for determining the clustering of different populations of galaxies. Other studies have used morphological classifications. Skibba et al. (2009), for example, took advantage of the largest sample of visually classified morphologies to date (from the SDSS) to clearly reveal that although early-type galaxies cluster more strongly than discs, once the analysis is done at a fixed colour and luminosity no significant difference is found; this confirms previous results (e.g. Beisbart & Kerscher 2000; Ball, Loveday & Brunner 2008) for the nearby Universe.

Understanding when these trends were created and how they evolved with cosmic time is a key input to galaxy formation models and to our general understanding of how galaxies formed and evolved. The first clustering studies of different populations of galaxies beyond the local Universe (e.g. Efstathiou et al. 1991; Fisher et al. 1994; Brainerd, Smail & Mould 1995; Le Févre et al. 1996), although pioneering, were mostly limited by the lack of information on the redshifts for their samples. Fortunately, those problems are now starting to be effectively tackled by larger and deeper surveys. Such surveys have recently led to robust clustering studies of specific populations of sources at moderate and high redshift such as active galactic nuclei (AGN; e.g. da Ângela et al. 2008), sub-mm galaxies (e.g. Blain et al. 2004; Weiß et al. 2009), luminous red and massive galaxies (e.g. Wake et al. 2008) or star-forming galaxies such as Hα emitters, Lyman-break galaxies or Lyman-α emitters (e.g. Geach et al. 2008, hereafter G08; McLure et al. 2009; Shioya et al. 2009). A dependence of the clustering on galaxy luminosity has also been identified beyond the local Universe (e.g. Giavalisco & Dickinson 2001). In particular, Kong et al. (2006), Hayashi et al. (2007) and, more recently, Hartley et al. (2008) found a clear dependence of the clustering amplitude on near-IR luminosity for $z \sim 1$–2 'BzK'-selected galaxies, indicating that already in the young Universe the most massive galaxies were much more clustered than the least massive ones. These results mean that although different studies have been combined to suggest links between the same population at different redshifts, or between different populations at different epochs, interpreting these requires a great deal of care, since luminosity limits typically increase significantly with redshift and are different for different populations.

Narrow-band Hα surveys are now a very effective way to obtain representative samples of star-forming galaxies. Since the Gallego et al. (1995) wide-area study in the local Universe, tremendous progress has been achieved, with recent narrow-band studies (e.g. Ly et al. 2007; Shioya et al. 2008; G08; Shim et al. 2009; Sobral et al. 2009a, hereafter S09), taking advantage of state-of-the-art wide-field cameras and obtaining large samples of typically hundreds of Hα emitters from $z = 0$ to $z = 2.23$, down to limiting SFRs of $\approx$1–$10\,M_\odot\,yr^{-1}$. In particular, the Hi-z Emission Line Survey (HiZELS; cf. G08 and S09) is obtaining and exploring unique samples of star-forming galaxies presenting Hα emission (and other major emission lines, e.g. Sobral et al. 2009b), redshifted into the J, H and K bands. By using a set of existing and custom-made narrow-band filters, HiZELS is surveying Hα emitters at $z = 0.84$, 1.47 and 2.23 over several square degree areas of extragalactic sky.

Narrow-band surveys have a great potential for determining the clustering properties of large samples of galaxies and their evolution with cosmic time. Such surveys probe remarkably thin redshift slices ($\Delta z \approx 0.02$), which not only provide an undeniable advantage over photometric surveys that can be significantly affected by sys-

tematic uncertainties but also allow the study of very well-defined cosmic epochs. Additionally, the selection function is well understood and easy to model in detail, a feature that contrasts deeply with that of current large high-redshift spectroscopic surveys. Narrow-band surveys also populate the redshift slices they probe with high completeness down to a known flux limit, in contrast to spectroscopic surveys which are usually very incomplete at any single redshift and present the typical pencil-beam distribution problems. Finally, narrow-band surveys can select equivalent populations at different redshifts, making it possible to really understand the evolution with cosmic time, avoiding the biases arising from comparing potentially different populations.

This paper presents a detailed clustering study of the largest sample of narrow-band-selected Hα emitters at $z \sim 1$ (cf. S09), obtained after conducting deep narrow-band imaging in the J band over $\sim$1.3 deg^2, as part of HiZELS. It is organized in the following way. Section 2 outlines the data, the samples and their on-sky distribution. Section 3 presents the angular two-point correlation function, carefully estimating the errors and other potential bias, along with the exact calculation of the real-space correlation. Section 4 quantifies the clear clustering dependences on the host galaxy properties. Section 5 presents the first detailed comparison between the clustering properties of Hα emitters from $z = 0.24$ up to 2.23. Finally, Section 6 presents the conclusions. An $H_0 = 100\,h\,km\,s^{-1}\,Mpc^{-1}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$ cosmology is used and, except where otherwise noted, magnitudes are presented in the Vega system and $h = 0.7$.

2 DATA AND SAMPLES

2.1 The sample of Hα emitters from HiZELS

This paper takes advantage of the HiZELS large sample of Hα emitters at $z = 0.845$ presented in S09 (the reader is referred to that paper for full details of how the sample was constructed). Briefly, the sample was obtained from data taken through a narrow-band J filter ($\lambda_{eff} = 1.211 \pm 0.015\,\mu m$) using the Wide Field Camera (WFCAM) on the United Kingdom Infrared Telescope (UKIRT), and reaches an effective flux limit of $8 \times 10^{-17}\,erg\,s^{-1}\,cm^{-2}$ over $\sim$0.7 deg^2 in the UKIDSS Ultra Deep Survey (UDS; Lawrence et al. 2007) field and $\sim$0.8 deg^2 in the Cosmological Evolution Survey (COSMOS; Scoville et al. 2007) field. The data allowed the selection of a total of 1517 line emitters which are clearly detected in NB$_J$ (signal-to-noise ratio >3) with a $J-$NB$_J$ colour excess significance of $\Sigma > 2.5$ and observed equivalent width EW > 50 Å.

As the detected emission may originate from several possible emission lines at different redshifts, photometric redshifts[1] were used to distinguish between them and select the complete sample of 743 Hα emitters over a comoving volume of 1.8×10^5 Mpc^3 at $z = 0.84$. Of these, 477 are found in COSMOS (0.76 deg^2) and 266 in UDS [0.54 deg^2; the reduction in area is driven by the overlap with the high-quality photometric catalogue used (cf. S09 for more details)]. The completeness and reliability of this sample were studied using the $\sim$10^4 available redshifts from zCOSMOS Data Release 2 (Lilly et al. 2009), spectroscopically confirming $\sim$100 Hα emitters within the photometric-redshift-selected sample; this allowed

---

[1] Photometric redshifts used in S09 for COSMOS present $\sigma(\Delta z) = 0.03$, where $\Delta z = (z_{phot} - z_{spec})/(1 + z_{spec})$; the fraction of outliers, defined as sources with $\Delta z > 3\sigma(\Delta z)$, is lower than 3 per cent while for UDS, the photometric redshifts have $\sigma(\Delta z) = 0.04$, with 2 per cent of outliers.
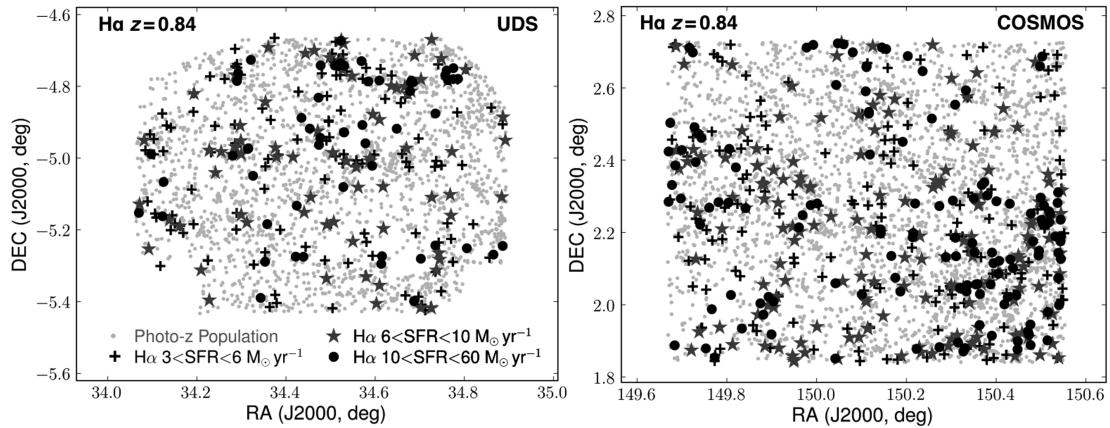
Figure 1. The on-sky distribution of the Hα emitters found at $z = 0.84$ in the UDS (left-hand panel) and COSMOS (right-hand panel) fields, together with a photometric-redshift-selected sample at the same redshift. Both panels/boxes cover the same angular/physical area, corresponding to $\approx 29 \times 29$ Mpc at $z = 0.84$. The Hα emitters are plotted in three different symbols corresponding to different SFRs (see key in the left-hand panel).

to estimate a >95 per cent reliability and >96 per cent completeness for the sample in COSMOS. Hα fluxes, Hα luminosities and Hα-based SFRs were obtained for all Hα candidates using HiZELS data – the details on how these were computed are fully detailed in S09.

Since the work presented in S09, improved photometric redshifts for COSMOS and UDS have been produced, incorporating a higher number of bands, deeper data and accounting for possible emission-line flux contamination of the broad-bands (cf. Ilbert et al. 2009; Cirasuolo et al. 2010; Cirasuolo et al. in preparation); these further confirm the robustness and completeness of the selection done in S09. Nevertheless, the Hα sample used for the analysis in this paper is slightly modified from that in S09 on the basis of the new photometric redshifts. In particular, the sample in UDS now contains 257 Hα emitters over 0.52 deg². The very small reduction in area simply results from the overlap with deep mid-IR data used for deriving the new improved photometric redshifts – this reduction places three sources from the initial sample outside the coverage. The remaining six sources excluded from the S09 sample were removed on the basis that the new photometric redshifts clearly place those sources at $z \approx 2.2$ (2) and 1.45 (4) (they are therefore identified as candidate [O II] 3727 and [O III] 5007/Hβ emitters, respectively). For COSMOS, the sample of Hα emitters is the same as in S09 (477 emitters over 0.76 deg²), since the new photometric redshifts do not change any of the classifications (noting that spectroscopic data had already been used to produce a cleaner sample in S09).

As the samples were obtained in two of the best studied square degree areas, a wealth of multiwavelength data are available. These include deep broad-band imaging from the UV to the IR – including deep *Spitzer* data. These data make it possible to compute rest-frame luminosities for the sample of Hα emitters at $z = 0.84$. Here, rest-frame B luminosities (M_B) are estimated by using i^+-band data (probing 4152.6 Å at $z = 0.84$) obtained with the Subaru telescope in both UDS and COSMOS. This is obtained by applying an aperture correction of -0.097 mag (to recover the total flux for magnitudes measured in 3-arcsec apertures) and a galactic extinction correction of -0.037 to i^+ magnitudes (Capak et al. 2007). It is assumed that all sources are at a luminosity distance of 5367 Mpc ($z = 0.845$); this yields a distance modulus of 43.649. Rest-frame K luminosities (M_K) are computed following Cirasuolo et al. (2010) by assuming

the same luminosity distance and interpolating using deep 3.6 and 4.5 μm *Spitzer* data.

Furthermore, the COSMOS field has been imaged by the Advanced Camera for Surveys/*Hubble Space Telescope*, and the sample of Hα emitters has been morphologically classified both automatically, with ZEST (Scarlata et al. 2007), and visually, as detailed in S09. The galaxies were classified into early types, discs and irregulars and, independently (visually only), into non-mergers, potential mergers and mergers (cf. S09). The sample has also been investigated for AGN contamination both in S09, using emission-line diagnostics, and in Garn et al. (2010) making use of a wide variety of techniques.

Fig. 1 presents the distribution of the Hα emitters at $z = 0.84$ for the UDS and COSMOS fields. The panels visually indicate how these emitters cluster across these two regions of the sky relative to a simple photometric-redshift-selected population at the same redshift, selected with $0.82 < z_{\rm photo} < 0.87$.

2.2 The random sample

Random catalogues are essential for robust clustering analysis and an over or underestimation of the clustering amplitude can easily be obtained if one fails to produce accurate random catalogues. These are produced by generating samples with 100 times more sources than the real data and by distributing those galaxies randomly over the geometry corresponding to the survey's field of view. This takes into account both the geometry of the fields and the removal of masked areas (due to bright stars/artefacts) – see discussion of masked regions in S09 for more details.

While the survey is fairly homogeneous in depth, there are some small variations from chip to chip and field to field reaching a maximum of ~0.2 mag (NB$_J$ of 21.5 to 21.7 mag). This corresponds to a maximum variation in luminosity [in log(L)] of ~0.1. The observed luminosity function presented in S09 shows that going deeper by 0.1 in log (L) increases the number count by ~20 per cent. This can have an effect on estimating the clustering, although simulations indicate that this is smaller than the measured errors. The final random catalogues were created reproducing source densities varying in accordance with the measured depths and the S09 number counts.

3 THE CLUSTERING PROPERTIES OF Hα EMITTERS

3.1 The two-point $\omega(\theta)$ correlation function at $z = 0.84$

In order to evaluate the two-point angular correlation function, the minimum variance estimator suggested by Landy & Szalay (1993) is used:

$$\omega(\theta) = 1 + \left(\frac{N_R}{N_D}\right)^2 \frac{DD(\theta)}{RR(\theta)} - 2\frac{N_R}{N_D}\frac{DR(\theta)}{RR(\theta)}, \qquad (1)$$

where $DD(\theta)$ is the number of pairs of real data galaxies within $(\theta, \theta + \delta\theta)$, $DR(\theta)$ is the number of data–random pairs and $RR(\theta)$ is the number of random–random pairs. N_R and N_D are the number of random and data galaxies in the survey. Errors are computed using the Poisson estimate (Landy & Szalay 1993):

$$\Delta\omega(\theta) = \frac{1 + \omega(\theta)}{\sqrt{DD(\theta)}}. \qquad (2)$$

The angular correlation function is computed for the entire sample of Hα emitters,[2] as well as for COSMOS and UDS separately, and it is found to be very well fitted by a power law with the form $A\theta^\beta$ with θ in arcsec. The power-law fit is obtained by determining $\omega(\theta)$ 2000 times using different random samples and a range of bin widths ($\Delta\log\theta = 0.1$–0.3, randomly picked for each determination) and by performing a χ^2 fit to each of these (over $5 < \theta < 600$ arcsec, corresponding to 38.2 kpc to 4.5 Mpc at $z = 0.845$; these correspond to angular separations for which fitting $\omega(\theta)$ with one single power law is appropriate; see details in Section 3.2.2). The results are shown in Fig. 2 and presented in Table 1; they imply $A = 14.1 \pm 3.9$ and $\beta = -0.79 \pm 0.06$ for the entire sample (the errors present the 1σ deviation from the average). This reveals a very good agreement with the fiducial value, $\beta = -0.8$.

Recent studies (e.g. Ouchi et al. 2005) have found a transition between small- and large-scale clustering (corresponding to one- and two-halo clustering, respectively) at high redshift, manifested as a clear deviation from the power law at small scales (e.g. smaller than ≈ 50 kpc for Lyman-break galaxies at $z \sim 4$). While there is no definitive evidence for the one-halo term for the Hα sample presented here (especially for separations larger than ~ 5 arcsec – corresponding to 38 kpc), for smaller separations the results point towards a departure from the power law, but only at $\approx 1\sigma$ level. Therefore, a larger sample is required to reliably determine $\omega(\theta)$ down to the smallest scales and constrain the contribution of the one-halo term.

The best power-law fit seen in Fig. 2 also reveals a good agreement both at small and larger scales between the COSMOS and UDS fields (see also Table 1); the UDS field presents a slightly higher clustering amplitude, but consistent within 1σ (fixing $\beta = -0.8$). Note that the Hα luminosity function derived in S09 for each field also revealed a good agreement between the two fields. These results are consistent with ~ 0.6–0.8 deg^2 fields being sufficient to overcome most of the effects of cosmic variance when conducting clustering analysis with narrow-band surveys at $z \sim 1$ – although the agreement could be caused by chance.

[2] Due to the finite area probed, the measured clustering amplitude will be underestimated by an amount C (known as the integral constraint; cf. Roche et al. 2002) which depends on the assumed true power law and the probed area. This is estimated as $C = 0.0023$ for the entire survey area and it only represents ≈ 0.01 per cent change in A, but it is still included, since this becomes significant for obtaining $\omega(\theta)$ at large separations; cf. Fig. 4.
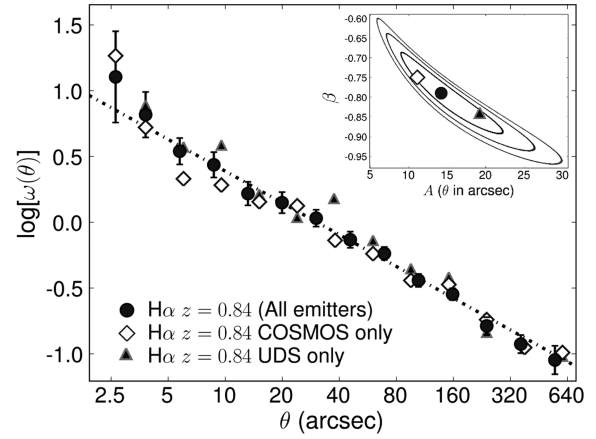
Figure 2. The two-point angular correlation function for the entire sample and for the COSMOS and UDS fields separately for angular separations up to 600 arcsec and the 1σ, 2σ and 3σ covariance contours when fitting both A and β for the entire sample compared to the results of each field. These results show a very good agreement between COSMOS and UDS, revealing that Hα emitters in completely different regions of the sky cluster with comparable amplitudes and power-law slopes. The results also show that there is no clear departure from the power law for this sample of Hα emitters down to 5 arcsec, implying that the two-halo contribution to $\omega(\theta)$ is dominant for scales larger than ~ 50 kpc, but there is tentative evidence of a significant one-halo contribution to $\omega(\theta)$ for smaller scales.

3.1.1 Robust error estimation and the effect of cosmic variance

The survey covers 1.3 deg^2 in two completely independent fields; this is a fundamental advantage over smaller surveys which only probe one single field, as it allows, in principle, a reliable estimate of possible errors due to cosmic variance. In order to achieve this, $\omega(\theta)$ is estimated over randomly picked square areas of different sizes (from 0.05 deg^2, corresponding to the area probed by one WFCAM chip, up to 0.5 deg^2) in COSMOS and UDS. The minimum number of emitters ranges from 15 to 65 within the smaller areas used (0.05 deg^2). Either 100 or 1000 randomly chosen regions are considered for each area (100 for 0.3–0.5 deg^2 and 1000 for smaller areas), and a power law is fitted to each $\omega(\theta)$ determination by fixing $\beta = -0.8$. For each area, the standard deviation on the amplitude A is used to quantify the uncertainty. The results are shown in Fig. 3 and demonstrate that the measured standard deviation is effectively reduced with area. The error in A (per cent) can be approximated by a power law of the form $20 \times \theta^{-0.35}$ with θ in deg^2; extrapolating this suggests an error slightly lower than ~ 20 per cent for the total area of the survey and also agrees with the small difference found between COSMOS and UDS (see Fig. 3). Furthermore, this suggests that one requires areas of ~ 7 deg^2 (the HiZELS target) or more to reduce the effects of cosmic variance to less than 10 per cent.

These estimates could still slightly underestimate the errors, mostly because at larger areas the number of approximately independent areas is strongly reduced. An alternative estimate of the clustering uncertainty can be derived using a bootstrap analysis. Since this often leads to a slight overestimation of the errors, it can be combined with the previous analysis to give an approximate range for the expected error. The analysis is done by naturally dividing the complete sample into 26 regions of ~ 0.05 deg^2 each (these are the minimum individual probed areas as these are covered by one WFCAM chip) and computing $\omega(\theta)$ with 26 randomly picked regions each time for 10 000 times. Fitting all realizations with a power law (fixing $\beta = -0.8$) results in a distribution with

Table 1. The power-law fitting parameters which best describe $\omega(\theta)$ for Hα emitters at $z = 0.84$, resulting from computing $\omega(\theta)$ 2000 times. These were obtained over $5 < \theta < 600$ arcsec, corresponding to 38 kpc to 4.5 Mpc at $z = 0.84$, avoiding both the possible one-halo clustering contribution (at the smallest scales) and the break of the Limber's approximation (at the largest scales). Note the degeneracy between A and β in Fig. 2. $A_{\beta=-0.8}$ is obtained by fixing $\beta = -0.8$, the fiducial value and in excellent agreement with what has been found for the entire sample with or without possible/likely AGN contamination. An error of 20 per cent is added in quadrature to random error in $\Delta A_{\beta=-0.8}$ to account for cosmic variance for the entire sample (or 25 per cent when considering just one subfield – COSMOS or UDS; cf. Section 3.1.1).

Sample ($z = 0.84$)	Number	A (θ in arcsec)	β	$A_{\beta=-0.8}$ (θ in arcsec)
All emitters	734	14.1 ± 3.9	-0.79 ± 0.06	14.2 ± 3.1
No AGN	660	13.6 ± 4.4	-0.78 ± 0.07	14.1 ± 3.0
No likely AGN	700	15.4 ± 4.0	-0.81 ± 0.06	14.6 ± 3.2
COSMOS	477	11.1 ± 3.0	-0.75 ± 0.08	13.8 ± 3.7
UDS	257	19.2 ± 8.9	-0.84 ± 0.15	18.4 ± 5.5
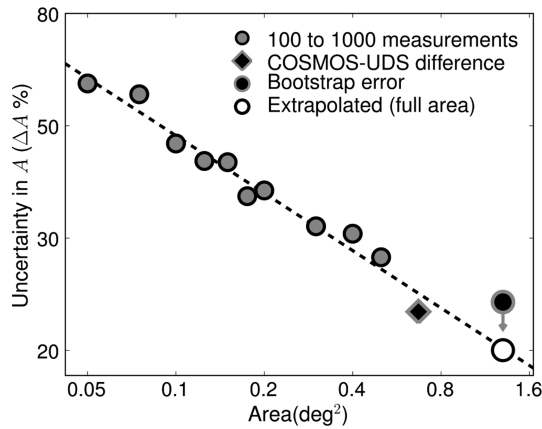
Figure 3. The uncertainty in A (per cent; $\Delta A = \sigma/A \times 100$, with $\beta = -0.8$) derived from 100–1000 $\omega(\theta)$ measurements on subsamples spanning different areas within the complete 1.3 deg^2 (up to 0.5 deg^2), which is considered to be mostly due to cosmic variance. This reveals that ΔA decreases with increasing area, and this trend is very well fitted by a simple power law ($20 \times \theta^{-0.35}$ per cent, with θ in deg^2, shown by the dashed line). The simple standard deviation between results from the COSMOS and UDS fields is also shown for comparison. The error estimated from the bootstrap analysis and that given by extrapolating the fitted power law are also shown for the full area.

a standard deviation of 23 per cent in the clustering amplitude, A. When compared and combined with the previous estimation ($\approx$18 per cent error), it suggests an error of $\sim$20 per cent in the clustering amplitude; this will be added in quadrature to the directly calculated errors in A.

3.1.2 AGN contamination

Some of the Hα emitters are likely to contain an AGN. By using emission-line ratio diagnostics for a small subset of emitters in zCOSMOS, S09 estimated a contamination of $\sim$15 per cent AGN in the sample. More recently, Garn et al. (2010) performed an extended search for AGN within the sample using several methods for identifying those sources, such as radio, X-rays and mid-IR colours. This resulted in the identification of a maximum of 74 AGN (40 in

COSMOS and 34 in UDS). From these, 34 are classified as likely AGN and the remaining 40 as possible AGN – this corresponds to an estimated $\sim$5–11 per cent AGN contamination within the sample.

The AGN may well have different clustering properties from the star-forming population. However, the nature of these potential AGN contaminants is very unclear, particularly the origin of the detected Hα emission. For example, they are mostly morphologically classified as irregulars and mergers and they span the entire luminosity range; it thus seems that although some of them might have their Hα emission powered by the AGN, they may well be undergoing significant star formation. Indeed, recently, Shi et al. (2009) presented a study of unambiguous AGN at $z \sim 1$, showing that at least half of the sample shows clear signatures of intense star formation.

In order to understand how AGN contaminants might affect the clustering measurements, the angular correlation function was calculated after removing all the potential AGN contaminants. This results in no statistical change in either A or β for the entire sample (see Table 1). Removing only the likely AGN leads to an identical result. However, the true AGN contaminants (and the possible bias they may introduce) might well have a larger effect on the results when studying the clustering as a function of host galaxy properties (presented in Section 4), since for certain host galaxy properties the AGN contamination may be significantly higher than in the sample as a whole.

To ensure that the analysis is really tackling the AGN contamination and to robustly determine the clustering properties of each sample being studied (and test trends), all the clustering properties in this paper (for the $z = 0.84$ sample and subsamples) are derived from a combination of measurements from samples with and without possible AGN contaminants. In practice, $\omega(\theta)$ is computed 500 times each using (i) all the emitters in the sample, (ii) removing likely AGN and (iii) removing all (likely plus possible) AGN. χ^2 fits are obtained for each realization of each sample and the total resulting distribution is used for the analysis. This also allows to carefully confirm that there is a good degree of consistency between the three different distributions and in no case is the difference between samples with and without AGN larger than the 1σ errors. This mostly results in estimating larger errors than those which would be obtained from either simply excluding all AGN or not dealing with AGN contamination and therefore represents a conservative approach.

3.2 Real-space correlation

3.2.1 First order approximation: Limber's equation

The real-space correlation, r_0, is a very useful description of the physical clustering of galaxies when the spatial correlation function is well described by $\xi = (r/r_0)^\gamma$. The inverse Limber transformation (Peebles 1980) can be easily used to obtain an approximation of the spatial correlation function[3] from the angular correlation function, provided the redshift distribution is known. For narrow-band surveys, the expected redshift distribution depends solely on the shape of the narrow-band filter profile; this can be well approximated by a Gaussian,[4] which, for the Hα emission line, corresponds to a redshift distribution centred at $z = 0.845$ with $\sigma = 0.0075$. This can be compared with the redshift distribution from Hα emitters confirmed by zCOSMOS (93 sources) which has a mean redshift of 0.844 and $\sigma = 0.0076$; the excellent agreement reveals that the real redshift distribution should be very close to the one assumed.

In order to calculate the real-space correlation length, r_0, and when performing the deprojection analysis, it is assumed that the redshifts are drawn from this Gaussian distribution and that the real-space correlation function is independent of redshift (cf. G08; Kovač et al. 2007) over the redshift range probed; this is a very good approximation, since the redshift distribution is extremely narrow. Contamination by sources which are not Hα emitters at $z = 0.84$ could have a significant effect on r_0, since the real redshift distribution will be different from the one assumed. Nevertheless, in S09 the sample was studied to find that only two out of 90 sources with spectroscopic redshifts that had been photometrically selected as Hα emitters were not real Hα emitters (one [S II] 6717 emitter at $z = 0.79$ and a [O III] 5007 emitter at $z = 1.42$, which were then removed from the sample). This conclusion is drawn from the analysis of limited spectroscopic data ($\sim$20 per cent of the HiZELS sample in COSMOS), but contamination at that level will only lead to underestimating r_0 by a maximum of 6 per cent[5] Throughout this paper, a 5 per cent correction is applied when computing r_0 to account for the expected contamination.

Computing r_0 for each realization of $\omega(\theta)$ fitted with a power law with $\beta = -0.8$ results in $r_0 = 2.5 \pm 0.3\,h^{-1}$ Mpc for the entire sample or $r_0 = 2.6 \pm 0.3\,h^{-1}$ Mpc when accounting for 5 per cent contamination. Note that the 20 per cent error in A due to cosmic variance results in an error of 11 per cent in r_0, and this is added in quadrature. However, as mentioned before, Limber's equation is only an approximation which can potentially result in significant errors for narrow-band surveys at high redshift. Section 3.2.2 describes how r_0 is robustly calculated by fully deprojecting $\omega(\theta)$.

3.2.2 Accurate determination of r_0 for narrow-band surveys

The errors introduced by the Limber's approximation are tackled by numerically integrating the exact equation connecting the spatial and angular correlation functions (following Simon 2007). This relation implies that spatial correlation functions described by $\xi = (r/r_0)^\gamma$ are projected as angular correlation functions with slopes $\beta = \gamma + 1$ for small scales and $\beta = \gamma$ for large angular separations when using a narrow filter – see full discussion in Simon (2007).

Here, it is assumed that the spatial correlation function is given by the power law $\xi = (r/r_0)^\gamma$ (as this is able to reproduce the observed $\omega(\theta)$ very well with $\gamma = -1.8$) and that the narrow-band filter profile is described by a Gaussian as detailed in Section 3.2.1. The following is then numerically integrated for the same angular separations as for the data, and a χ^2 fit is done for $\omega(\theta)$ around the value of r_0 obtained in the previous section:

$$\Delta\omega(\theta) = \psi^{-1} \int_0^{+\infty} \int_{s\sqrt{2\phi}}^{2s} \frac{2 f_S(s-\Delta) f_S(s+\Delta)}{R^{-\gamma-1} r_0^\gamma \Delta} dR ds, \qquad (3)$$

where $\psi = 1 + \cos\theta$, $\phi = 1 - \cos\theta$, $\Delta = \sqrt{(R^2 - 2s^2\phi)/(2\psi)}$, and f_S is the profile of the filter being used in comoving distance (assumed to be a Gaussian distributed value with an average of $2036.3\,h^{-1}$ Mpc and $\sigma = 14.0\,h^{-1}$ Mpc). The χ^2 fitting results in a much more robust estimate of r_0 and allows the use of $\omega(\theta)$ up to larger angular separations than 600 arcsec, a regime for which a single power law starts to become inadequate (as β changes from $\gamma + 1$ to γ; cf. Fig. 4). For the entire sample, however, this leads to little change: $r_0 = 2.7 \pm 0.3\,h^{-1}$ Mpc (this includes a 5 per cent correction for contamination, while the errors also include cosmic variance – this error estimation has been discussed in Section 3.1.1). Indeed, whilst Limber's equation breaks down for large galaxy separations, it is shown to do well as long as only smaller angular distances are
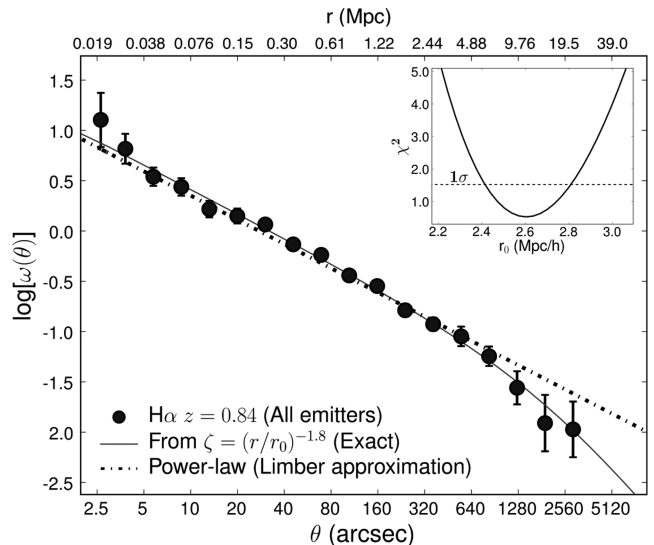
Figure 4. The two-point angular correlation function for the entire sample, compared with the simple power-law fit expected from the Limber's approximation. These are compared with the best (from a χ^2 fit – shown in the figure) exact angular correlation function obtained with the narrow-band filter profile for a spatial correlation function given by the power-law $\xi = (r/r_0)^\gamma$, with $\gamma = -1.8$ and $r_0 = 2.61 \pm 0.13\,h^{-1}$ Mpc (not corrected for contamination; errors directly from χ^2), revealing an excellent agreement with the data. This also reveals the regime for which the Limber's approximation breaks for this particular case ($\sim$600 arcsec for a $\sim$15 per cent difference; cf. Simon 2007 for a general analysis of the typical separation at which Limber's approximation breaks down).

---

[3] Limber's equation is an approximation that breaks down for large angular separations when one uses a narrow filter, but can still be used to obtain at least a first approximation of r_0 within $\approx$15 per cent for the NB$_J$ filter used; see Section 3.2.2 for details.

[4] A simpler way to model the narrow-band filter is to assume it is a top-hat; this results in a Hα redshift distribution of 0.845 ± 0.015. The two approaches produce results for different r_0 determinations which are consistent within 5 per cent.

[5] Assuming that contaminants will not cluster significantly, the contamination fraction of f will result in a maximum underestimation of A given by $(1 - f)^2$, and an underestimation of r_0 given by $\sim(1 - f)^{2/|\gamma|}$.

Table 2. The correlation length (fixing $\gamma = -1.8$ and corrected by 5 per cent for contamination) for Hα emitters at $z = 0.84$. Samples with fixed M_K have $-24.44 < M_K < -22.68$ (AB) and those with fixed $\log L_{H\alpha}$ have $41.76 < \log L_{H\alpha} < 42.22$ (erg s^{-1}). Hα luminosities are given in erg s^{-1}, M_K and M_B are AB rest-frame luminosities. Unclassified sources or with no data available were not considered. Errors include cosmic variance (11 per cent of r_0 for the entire sample and 14 per cent of r_0 for samples of each individual field).

Subsample (Hα $z = 0.84$)	Number	r_0 (h^{-1} Mpc)
All emitters	734	2.74 ± 0.29
Pure star-forming	660	2.61 ± 0.28
COSMOS	477	2.67 ± 0.37
UDS	257	3.13 ± 0.52
$41.95 < \log L_{H\alpha} < 43.18$	367	4.84 ± 0.58
$41.61 < \log L_{H\alpha} < 41.95$	367	2.52 ± 0.41
$42.40 < \log L_{H\alpha} < 42.80$	46	5.24 ± 1.01
$42.25 < \log L_{H\alpha} < 42.40$	55	5.03 ± 0.96
$42.10 < \log L_{H\alpha} < 42.25$	92	4.55 ± 0.72
$41.88 < \log L_{H\alpha} < 42.10$	243	2.50 ± 0.45
$41.75 < \log L_{H\alpha} < 41.88$	173	2.26 ± 0.42
$41.62 < \log L_{H\alpha} < 41.75$	125	2.13 ± 0.55
$42.20 < \log L_{H\alpha} < 42.44$ (fixed M_K)	56	4.20 ± 0.86
$41.98 < \log L_{H\alpha} < 42.20$ (fixed M_K)	117	3.87 ± 0.58
$41.83 < \log L_{H\alpha} < 41.98$ (fixed M_K)	125	2.84 ± 0.55
$41.71 < \log L_{H\alpha} < 41.83$ (fixed M_K)	117	1.70 ± 0.51
$-26.1 < M_K < -24.5$	82	3.74 ± 0.69
$-24.5 < M_K < -24.0$	116	4.24 ± 0.63
$-24.0 < M_K < -23.5$	145	3.08 ± 0.51
$-23.5 < M_K < -23.0$	145	2.72 ± 0.43
$-23.0 < M_K < -22.5$	117	2.65 ± 0.54
$-22.5 < M_K < -20.5$	116	2.31 ± 0.49
$-24.96 < M_K < -24.00$ (fixed $\log L_{H\alpha}$)	116	3.99 ± 0.56
$-24.00 < M_K < -23.40$ (fixed $\log L_{H\alpha}$)	108	2.52 ± 0.62
$-23.40 < M_K < -22.84$ (fixed $\log L_{H\alpha}$)	112	2.81 ± 0.54
$-22.84 < M_K < -21.81$ (fixed $\log L_{H\alpha}$)	111	2.48 ± 0.50
$-23.23 < M_B < -21.61$	96	4.47 ± 0.77
$-21.61 < M_B < -21.11$	186	2.97 ± 0.44
$-21.11 < M_B < -20.61$	246	2.25 ± 0.30
$-20.61 < M_B < -20.11$	145	3.92 ± 0.63
$-20.11 < M_B < -18.71$	59	2.54 ± 0.92
Discs (COSMOS)	363	2.52 ± 0.32
Irregulars (COSMOS)	68	5.12 ± 0.88
Discs (SFR and M_K match)	109	2.59 ± 0.45
Irregulars (SFR and M_K match)	46	2.96 ± 0.80
Non-mergers (COSMOS)	298	2.30 ± 0.31
Mergers (COSMOS)	111	3.75 ± 0.50
Non-mergers (SFR and M_K match)	128	2.35 ± 0.42
Mergers (SFR and M_K match)	100	2.87 ± 0.56
Bulge-dominated discs (COSMOS)	97	2.72 ± 0.54
Disc-dominated discs (COSMOS)	204	2.51 ± 0.39

considered. Note that real-space correlations for different samples (see Table 2) are computed in the following sections using the same procedure described here.

3.2.3 *The dependence of the redshift distribution on limiting line luminosity for narrow-band surveys*

Since the narrow-band filter profile is not a perfect top-hat, emitters with different Hα luminosities are detected over a slightly different range of redshifts. As the exact relation between $\xi(r)$ and $\omega(\theta)$

depends on the redshift distribution (given by the filter profile), assuming the same redshift distribution for different luminosity limits can lead to biases, especially when looking at a possible relationship between the clustering amplitude and Hα luminosity.

This potential bias is studied by performing the same simulations described in S09. Briefly, the Hα luminosity function presented in S09 is used to generate a fake population of emitters equally distributed over a wider range of redshifts than those probed by the narrow-band filter, and this is used to look at the recovered redshift distribution of sources as a function of measured luminosity. The redshift distribution is found to become continuously narrower with increasing Hα luminosity limit,[6] although all the distributions can be equally well fitted by a Gaussian. The variation can be written as a function of observed Hα flux as

$$\sigma = -\eta \times \left(\log F_{H\alpha(\text{limit})} - \log F_0 \right) + \sigma_0, \qquad (4)$$

where F_0 is a flux limit at which the redshift distribution is relatively well understood and σ_0 is the width of the redshift distribution at that flux limit; for the NB$_J$ filter, $\eta = 0.00117$, as derived from simulations.

Over the luminosity range of the sample (and for the assumed luminosity function and the real filter profile), neglecting this effect results in overestimating r_0 by a maximum of 8 per cent; this is only found when computing r_0 for a sample containing the brightest 5 per cent emitters (the $\approx$40 brightest emitters, for which $\sigma \approx 0.0065$). Nevertheless, narrow-band surveys which span a much wider luminosity range will be much more sensitive to this and line-luminosity trends could naturally arise from this bias.

This luminosity dependence is fully taken into account when computing r_0 for subsamples defined by different Hα luminosity/SFR limits. For the analysis of other subsamples, if the Hα luminosity distribution does not present any clear offset from that of the entire survey, it is assumed that the redshift distribution of those subsamples is very well approximated by the complete redshift distribution.

4 THE CLUSTERING DEPENDENCES ON HOST GALAXY PROPERTIES

As discussed in the Introduction, in the local Universe the clustering has been shown to depend on several host galaxy properties. At $z = 0.24$, Shioya et al. (2008) have shown that the clustering of Hα emitters seems to be stronger for the most luminous sources in Hα, but so far testing and quantifying that clustering dependence at higher redshifts for Hα emitters has not been possible, as surveys have lacked area and sample size. The HiZELS sample is large enough to achieve this.

The complete sample is divided into several subsamples in order to investigate the clustering as a function of fundamental observable host galaxy properties. These include (i) Hα luminosity corrected for extinction as in S09 (1 mag), (ii) rest-frame K luminosity (M_K), (iii) rest-frame B luminosity (M_B) and (iv) morphological class (discs, irregulars; non-mergers, mergers). The $\omega(\theta)$ correlation function is computed for each subsample (a single example is shown in Fig. 5, presenting $\omega(\theta)$ for the brightest and faintest halves

---

[6] Although intrinsically more luminous sources are detectable over a wider redshift range, their detection in the wings of the filter leads to an underestimation of their luminosity; galaxies are only measured to be luminous in Hα when the emission line is being detected near the peak of the filter profile.
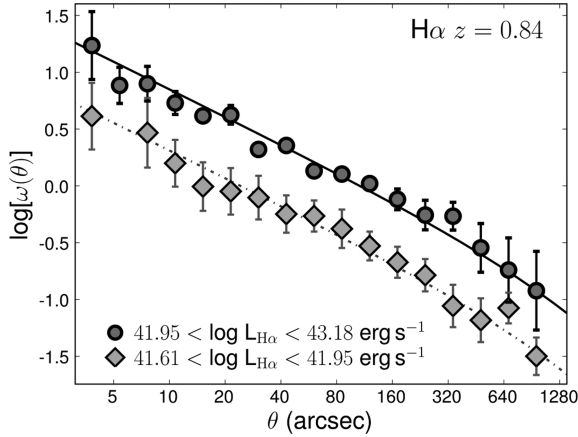
Figure 5. The two-point angular correlation function for subsamples obtained by splitting the sample in two halves based on Hα luminosity/SFR. These clearly show that the clustering amplitude is higher for galaxies with higher Hα luminosities, implying that the most active star-forming galaxies are more clustered than the least active. Lines show the best χ^2 fits to the exact relation between $\xi(r)$ and $\omega(\theta)$ by fixing $\gamma = -1.8$, which (after applying all corrections already detailed) result in $r_0 = 4.84 \pm 0.58 \, h^{-1}$ Mpc for the brightest emitters and $2.52 \pm 0.41 \, h^{-1}$ Mpc for the faintest half (in Hα luminosity) of the sample.

of the sample with respect to Hα luminosity). The approach detailed in the previous sections is then used by firstly computing r_0 using Limber's approximation power-law fit to $\omega(\theta)$ (fixing $\beta = -0.8$ and restricting the analysis to $5 < \theta < 600$ arcsec) and then using the exact relation between the spatial and angular correlation functions to improve the r_0 estimation. This analysis also accounts for the potential AGN contamination in the sample, following the procedures described at the end of Section 3.1.2 [computing $\omega(\theta)$ with and without the possible AGN contaminants and using the combined distribution]. The results obtained when splitting the sample into different subsamples are presented in Table 2, Figs 6 and 7. For subsamples obtained from the complete survey area, an error of 11 per cent in r_0 is added in quadrature to the 1σ errors (see
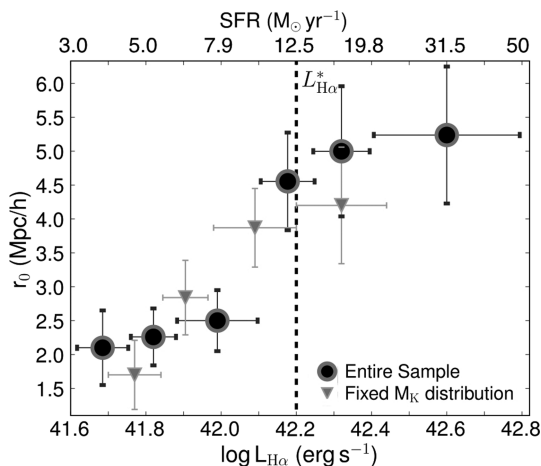
Figure 6. The dependence of the clustering length, r_0, with Hα luminosity at $z = 0.84$. This clearly reveals that galaxies presenting higher Hα luminosities/SFRs are more clustered, and that galaxies above the break in the Hα luminosity function ($L^*_{H\alpha}$) are much more clustered than those below. This dependence is not a result of the correlation between SFR (Hα luminosity) and stellar mass (M_K; see Fig. 8), as r_0 correlates as well with Hα luminosity for subsamples with the same M_K.

Section 3.1.1), while for subsamples derived based on morphology a 14 per cent error is added in quadrature to Δr_0 (for the COSMOS field only).

4.1 Hα luminosity/star formation rate

Following previous studies, such as Shioya et al. (2008), one can do a simple splitting of the sample in two halves with the same number of emitters (367 with $41.95 < \log L_{H\alpha} < 43.18$ erg s^{-1} and equal number with $41.61 < \log L_{H\alpha} < 41.95$ erg s^{-1}) and study the clustering properties of each of those samples. Fig. 5 presents the angular correlation function obtained for the brightest and faintest halves of the sample, revealing that the brightest Hα emitters are much more clustered than the faintest ones. This is also very clear when one compares the spatial correlations obtained for each sample: while the brightest emitters in Hα present $r_0 = 4.8 \pm 0.6 \, h^{-1}$ Mpc, for the faintest emitters one finds $2.5 \pm 0.4 \, h^{-1}$ Mpc. It should also be noted that because γ is fixed, the significant difference in r_0 for the two samples is not a result of the degeneracy between γ and r_0 (see Fig. 2 for a similar degeneracy between β and A), contrarily to studies such as Shioya et al. (2008), which allow for both γ and r_0 to vary; such note is valid throughout this paper.

With the large sample of Hα emitters obtained by HiZELS at $z = 0.84$, it is possible to perform a much more detailed investigation into the relation between r_0 and Hα luminosity by separating the emitters into a larger number of subsamples. Fig. 6 and Table 2 show that r_0 increases by a factor of almost 3 from the faintest Hα galaxies to the most active star-forming galaxies found above the break in the Hα luminosity function, $L^*_{H\alpha}$ ($10^{42.2}$ erg s^{-1}; S09).

Whilst the increase of r_0 with $L_{H\alpha}$ can be reasonably well described by a straight-line fit ($\chi^2 = 1.4$), there are hints that there might be a stronger increase in r_0 around $L^*_{H\alpha}$. This provides a description which is in line with previous studies but gives an unprecedented degree of detail of the relation between r_0 and Hα luminosity/SFR. Moreover, the robustness of these results is also increased by including the small correction for the fact that samples with different luminosity limits will present a different redshift distribution just from considering the filter profile.

Recently, Orsi et al. (2009) made predictions for the clustering properties of Hα emitters using two different versions of the GALFORM galaxy formation model. At $z = 0.84$ and for the flux limit of S09, Orsi et al. (2009) predict $r_0 \approx 4$–$5 \, h^{-1}$ Mpc. This shows that the disagreement between the data and the models found at $z = 0.24$ and 0.4 in Orsi et al. (2009) is also found at $z = 0.84$. The models predictions would be consistent with the data for flux limits two times higher due to the strong clustering dependence with Hα luminosity at $z = 0.84$; however, none of the models is able to fully reproduce this r_0–$L_{H\alpha}$ relation at the moment.

4.2 Rest-frame continuum luminosity

Hα emitters presenting the highest rest-frame K-band luminosities ($M_K < -24$) are strongly clustered, and a general trend of an increasing r_0 with K-band luminosity is found (see the left-hand panel of Fig. 7), similar to what has been found for other populations of galaxies at different redshifts. This suggests that this correlation must be valid regardless of the population being observed – at least for the range of luminosities probed – and seems to simply imply that galaxies with the highest stellar masses (as these are expected to correlate very closely with K luminosity, although the contribution from the thermally pulsing asymptotic giant branch phase of stellar
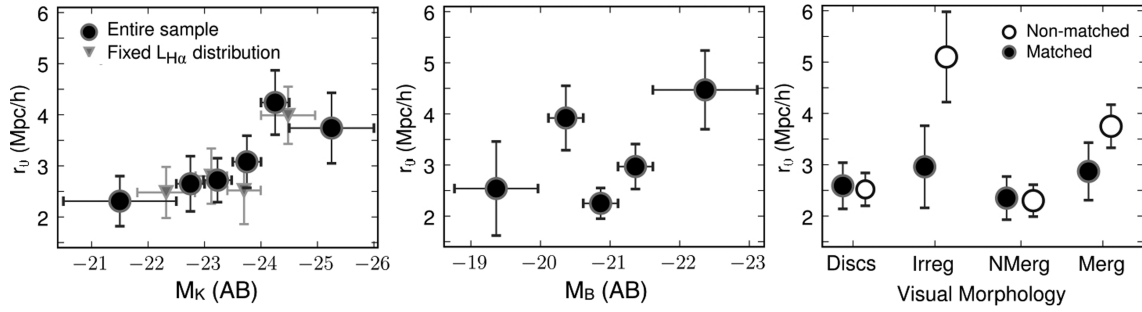
Figure 7. The real-space correlation length as a function of rest-frame K luminosity (left-hand panel), rest-frame B luminosity (middle panel) and visual morphology (right-hand panel). These results show that star-forming galaxies cluster more with increasing rest-frame K luminosity (which traces stellar mass), with this trend being maintained (but slightly weakened) even when one uses subsamples that have the same Hα luminosity distribution. No statistically significant correlation between r_0 with rest-frame B luminosity is found. Morphologically classified irregulars and mergers cluster more strongly than discs and non-mergers, but this is due to a different Hα luminosity and M_K distribution within each population; matching these distributions results in a similar r_0.

evolution can lead to some scatter in the M_K versus stellar mass relation; cf. Maraston et al. 2006) reside in the most massive haloes. However, it is also clear that the increase of r_0 with rest-frame K luminosity, whilst continuous, is not as pronounced as the increase with SFR.

In the local Universe, galaxies with higher B-band luminosities (M_B) are more clustered than galaxies with lower M_B. However, studying the clustering properties of Hα emitters at $z = 0.845$ as a function of rest-frame B luminosity reveals no statistically significant trend ($<1\sigma$) (see the middle panel of Fig. 7). These results are in reasonable agreement with recent studies (e.g. Meneux et al. 2009), which did not find any significant dependence of galaxy clustering on B luminosity for a similar redshift range.

4.3 Morphological class

By splitting the sample into morphological classes (only for the COSMOS field; cf. S09), one finds that galaxies classified as irregulars are more clustered than those classified as discs; mergers have a measured r_0 which lies between these populations, but significantly above the non-mergers (see Table 2 and the right-hand panel of Fig. 7 – open circles). However, S09 found that irregulars and mergers are typically brighter in Hα and M_K than discs (which completely dominate the faint end of the Hα luminosity function at $z \sim 1$). Therefore, in order to understand if the clustering really does depend on the morphological class or if this is simply a secondary effect driven by SFR and stellar mass dependencies, one needs to compare samples which are matched in log $L_{H\alpha}$ and M_K. This is done by using the distribution of Hα emitters in the log $L_{H\alpha}$–M_K plane and matching irregulars with discs and mergers with non-mergers. A Δlog $L_{H\alpha} < 0.02$ and $\Delta M_K < 0.02$ criteria is used: these values were chosen to ensure that the samples are very well matched in both Hα luminosity and M_K while still retaining the sample sizes required for the analysis. The population match used still results in a severe reduction of the larger disc and non-merger samples (along with a smaller reduction of the number of irregulars and mergers), but by matching these on the basis of M_K and log $L_{H\alpha}$ distribution, it is then possible to directly compare these populations on the basis of the morphological class only.

There is no significant difference in r_0 for the matched samples of irregulars and discs (see Fig. 7 – filled circles) within 1σ. The r_0 difference between mergers and non-mergers is also greatly reduced, although mergers are still found to be slightly more clustered than non-mergers at $\sim1\sigma$ level.

The sample of Hα emitters with disc morphologies has also been classified according to how much disc/bulge dominated each galaxy is (with ZEST; cf. Scarlata et al. 2007). By computing the correlation length for disc-dominated discs and bulge-dominated discs, one finds that they present reasonably the same r_0 (2.7 ± 0.5 h^{-1} Mpc for disc-dominated and 2.5 ± 0.4 h^{-1} Mpc for bulge-dominated galaxies). Therefore, the bulge–disc ratio of the studied star-forming disc galaxies does not have any significant effect on how these galaxies cluster at $z \sim 1$.

These results show that Hα luminosity and M_K (probing stellar mass) are the key host galaxy properties driving the clustering of star-forming galaxies at $z \sim 1$, and morphology is unimportant for the clustering of galaxies at high redshift, just as has been found in the local Universe (e.g. Skibba et al. 2009).

4.4 Hα luminosity versus rest-frame K luminosity

It has been shown that r_0 is very well correlated with both Hα luminosity (or SFR) and rest-frame K luminosity (tracing stellar mass). On the other hand, star-forming galaxies both in the local Universe and at high redshift reveal a correlation between both, with SFR typically being higher for galaxies with higher stellar masses (see Fig. 8). Thus, how much is the correlation between r_0 and Hα luminosity driven by different rest-frame K luminosity distributions and vice-versa?

In order to clearly test if both relations hold – or whether they are a result of a single strong relation between r_0 and either Hα luminosity of rest-frame K luminosity – new subsamples are derived. To do this, two subregions are defined in the 2D $L_{H\alpha}$–M_K space (see Fig. 8). To test the r_0–$L_{H\alpha}$ correlation at a fixed M_K, the region $-24.44 < M_K < -22.68$ (AB) and $41.71 < \log L_{H\alpha} < 42.44$ erg s^{-1} (corresponding to $4.1 < $ SFR $ < 21.8 \, M_\odot$ yr^{-1}) is considered (see Fig. 8). This region contains 409 Hα emitters, which can be divided into four subsamples on the basis of Hα luminosities with the same distribution of M_K (means of -23.6, -23.7, -23.6 and -23.7 with standard deviations of 0.5 in all instances, ordered from highest to lowest in respect to Hα luminosity). Similarly, in order to test the r_0–M_K correlation, the sample is restricted to Hα emitters within the region defined by $-24.96 < M_K < -21.81$ (AB) and $41.76 < \log L_{H\alpha} < 42.22$ erg s^{-1} (see Fig. 8) in which subsamples split on the basis of their M_K have the same $L_{H\alpha}$ distributions (means of 41.97, 41.95, 41.95 and 41.97 and standard deviations of 0.12 in all instances, ordered from highest to lowest in M_K) – cf. Table 2. The angular correlation functions of these matched subsamples are then computed and values of r_0 derived as fully described before.
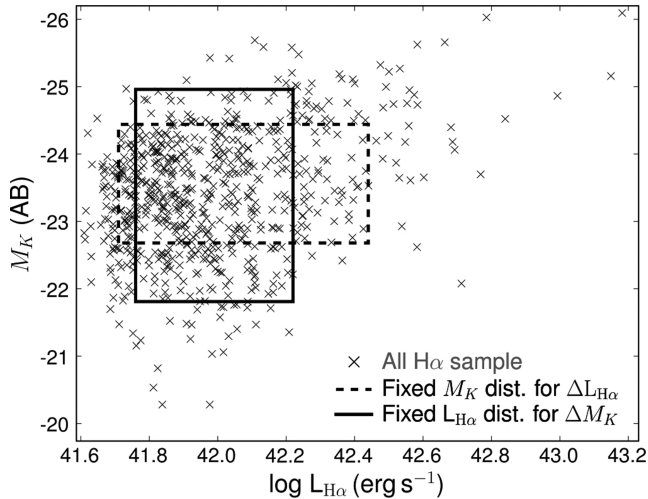
Figure 8. The complete 2D M_K–log $L_{H\alpha}$ space, showing the correlation between rest-frame K luminosity and Hα luminosity, compared to the regions used to select subsamples with the same M_K or $L_{H\alpha}$ distributions. This clearly shows that without using subregions of this 2D space, samples defined only on either $L_{H\alpha}$ or M_K will present different M_K and $L_{H\alpha}$ distributions, respectively, since these quantities are correlated.

The results are presented in Table 2 and in Figs 6 and 7. These reveal that both Hα luminosity and rest-frame K luminosity are relevant for the clustering of star-forming galaxies, since for a fixed distribution of one of these, a variation in the other one leads to a change in r_0. However, it should be noted that the correlation between r_0 and Hα luminosity (for a fixed M_K distribution) is maintained as highly statistically significant, while the statistical significance of the correlation between r_0 and M_K at a fixed $L_{H\alpha}$ distribution is slightly reduced.

5 THE CLUSTERING OF Hα EMITTERS ACROSS COSMIC TIME

The clustering properties of the Hα emitters at $z = 0.84$ can be compared with those of the HiZELS sample at $z = 2.23$ (G08) and with results derived from lower redshift surveys of Hα emitters at $z = 0.24$ and 0.4 (Nakajima et al. 2008; Shioya et al. 2008). This can then be interpreted in the context of the ΛCDM cosmological model.

5.1 The clustering of Hα emitters at $z = 0.24$

With the motivation of reliably comparing the clustering measurements at different redshifts, and since the Shioya et al. (2008) sample is publicly available, $\omega(\theta)$ is carefully recomputed for this sample.

Shioya et al. (2008) used galaxy colours to identify ~1000 candidate $z = 0.24$ Hα emitters within their narrow-band excess sample. For this analysis, the sample is restricted to 492 sources by limiting it to a very robust completeness limit ($L_{H\alpha} > 10^{39.8}$ erg s^{-1}, corresponding to SFR > 0.05 M$_\odot$ yr^{-1}). The robustness of this sample can be tested and improved taking advantage of a large set of spectroscopic data recently made available in the COSMOS field by the zCOSMOS project. There are 75 Hα candidates with a spectroscopic redshift available from zCOSMOS and 69 are identified as Hα emitters at $0.236 < z < 0.252$. The remaining six sources are identified as [S II] 6717 emitters at $0.229 < z < 0.231$ (contamination by S II emitters is even more significant at lower line fluxes, as

pointed out recently by Westra et al. 2010). With the robust luminosity cut used here, this corresponds to a contamination of 8 per cent, and by removing the six confirmed contaminants from the sample of Hα emitters, the contamination is estimated to be 7 per cent for the remaining sample of 486 emitters. This results in underestimating r_0 by a maximum of 8.4 per cent, and r_0 will be corrected by 7 per cent to account for contamination in the sample at $z = 0.24$; this is 80 per cent of the maximum correction, consistent with the approach used for $z = 0.84$.

Finally, in order to obtain r_0 (following the same procedures as for the $z = 0.84$ sample, except AGN contamination), two redshift distributions are considered: the one assumed in Shioya et al. (2008) (a top-hat characterized by $z = 0.242 \pm 0.009$) and a redshift distribution derived from the distribution of the 69 spectroscopically confirmed Hα emitters, which can be well described by a Gaussian with an average of $z = 0.245$ and a standard deviation of $\sigma = 0.006$. In practice, the values derived from both distributions agree well, but for consistency in determining r_0, the Gaussian distribution will be assumed.

The correlation length results in $r_0 = 1.8 \pm 0.2\, h^{-1}$ Mpc, with the error including random errors [from the standard deviation obtained after 1000 measurements of $\omega(\theta)$] and cosmic variance (assumed to affect r_0 by 11 per cent for this area). Only separations of $5 < \theta < 800$ arcsec are used when obtaining a power-law fit: angular separations lower than 5 arcsec are not used as $\omega(\theta)$ becomes steeper, deviating from the power law; this is simply interpreted as the transition between the two-halo and one-halo clustering, happening at ≈ 20 kpc. It should also be noted that the Limber equation works well for this case, with the differences between using such approximation and integrating the full relation between $\omega(\theta)$ and $\xi(r)$ being lower than 5 per cent for $\theta < 800$ arcsec.

5.2 The clustering of Hα emitters at $z = 2.23$

It has been mentioned that Limber's equation breaks down for large angular separations and for very narrow filters. Indeed, at $z = 2.23$ the comoving space width of the filter is extremely narrow, presenting $\sigma = 7.3\, h^{-1}$ Mpc and representing only ≈ 0.2 per cent of the total comoving distance to $z = 2.23$; this means that if the spatial correlation function of the Hα emitters is a power law $\xi = (r/r_0)^{-1.8}$, then $\omega(\theta)$, measured at $5 < \theta < 1000$ arcsec, cannot be well fitted by a power law with $\beta = -0.8$, since this regime is probing the exact change between $\beta = \gamma + 1$ and $\beta = \gamma$.

The correlation length r_0 is therefore recomputed. This is done by using the $\omega(\theta)$ data points presented in G08 and by doing a χ^2 fit to the obtained $\omega(\theta)$ from equation (3). This assumes that the spatial correlation function is a power law with $\gamma = -1.8$ (which is shown to reproduce the data very well), and only r_0 is allowed to vary. This procedure results in $r_0 = 2.6 \pm 0.5\, h^{-1}$ Mpc, with the error being derived directly from the χ^2 fit. Note that this implies a best fit $\omega(\theta)$ with $\beta \approx -1$ (for $\gamma = -1.8$) over the separations studied, agreeing well with the G08 best β fit to $\omega(\theta)$ of $\beta = -1$. The value of r_0 is notably lower than the value of $r_0 = 3.6 \pm 0.4\, h^{-1}$ Mpc derived by G08 from using the Limber approximation and their fitted values for A and β. A further correction is applied to account for a 15 per cent known contamination (the confirmed contaminants are all at different redshifts; this results in underestimating r_0 by 20 per cent) based on limited spectroscopic data available at the moment (Geach et al., in preparation); this finally results in $r_0 = 3.1 \pm 0.7\, h^{-1}$ Mpc (this also assumes a 14 per cent error in r_0 due to cosmic variance based on the area of the survey, but such uncertainty is likely to be underestimated).
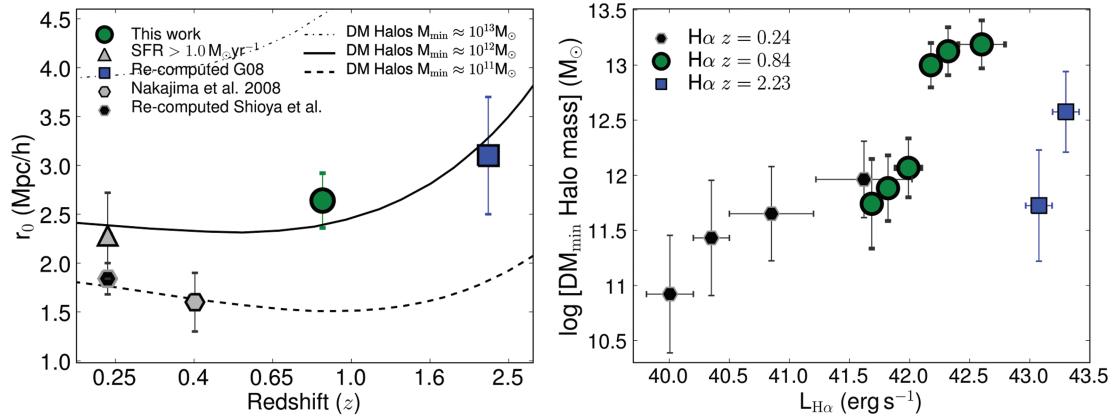
Figure 9. Left-hand panel: the clustering length (r_0) as a function of redshift for Hα emitters selected by narrow-band surveys. This reveals that the Hα emitters at $z = 0.84$ studied by HiZELS reside in typical dark matter haloes of $M_{min} \approx 10^{12}\,M_\odot$, consistent with being the progenitors of Milky Way type galaxies. The lower luminosity Hα emitters found in smaller volumes at $z = 0.24$ and 0.4 reside in less massive haloes; however, the most luminous Hα emitters at $z = 0.24$ cluster more strongly, and seem to reside in $M_{min} \approx 10^{12}\,M_\odot$, while Hα emitters at $z = 2.23$ reside in haloes just slightly less massive. Right-hand panel: the minimum mass of host dark matter haloes as a function of Hα luminosity; this reveals that more luminous emitters reside in more massive haloes, at any given cosmic time, but it also shows that the relation between halo mass and Hα luminosity/SFR evolves across cosmic time, with $M_{min} \approx 10^{12}\,M_\odot$ being much more effective (with respect to SFRs) at $z \sim 2$ than at $z < 1$.

5.3 The clustering evolution of Hα emitters since $z = 2.23$

Fig. 9 presents r_0 as a function of redshift for Hα emitters; this is the first combination of self-consistent clustering measurements for Hα emitters spanning more than half of the history of the Universe ($\approx 8\,\mathrm{Gyr}$) whilst probing four different well-defined epochs. For comparison, the $r_0(z)$ predictions for dark matter haloes with a fixed *minimum* mass of $M_{min} > 10^{11-13}\,M_\odot$ (Matarrese et al. 1997; Moscardini et al. 1998) are also shown. A ΛCDM cosmology is assumed, together with an evolving bias model $b(z)$ from Moscardini et al. (1998), and the values from that study are used for various fixed minimum mass haloes (cf. G08).

The results show that Hα emitters found in HiZELS at $z = 0.84$ (SFRs $> 3\,M_\odot\,\mathrm{yr}^{-1}$) reside in Milky Way type haloes of $M_{min} \approx 10^{12}\,M_\odot$ – the typical haloes where L^* galaxies in the local Universe reside. This contrasts with the low-Hα-luminosity galaxies for the samples at $z = 0.24$ and 0.4 (presenting SFRs $> 0.05\,M_\odot\,\mathrm{yr}^{-1}$). These seem to reside in much less massive haloes ($M_{min} \approx 10^{11}\,M_\odot$). These low-redshift emitters also present very low stellar masses and low luminosities in all available bands and are therefore likely to be small, young, dwarf-like galaxies, very different from the already much more massive and active star-forming galaxies found at $z = 0.84$.

Nevertheless, motivated by the correlation between r_0 and $L_{H\alpha}$ found both at $z = 0.24$ and 0.84, it is found that if one considers only the ≈ 5 per cent most luminous emitters (in Hα) at $z = 0.24$ (a rough match in $L_{H\alpha}$ to the $z = 0.84$ sample), then these are much more clustered than the complete sample, presenting $r_0 \approx 2.4\,h^{-1}$ Mpc; this is consistent with these emitters residing in dark matter haloes of $M_{min} \approx 10^{12}\,M_\odot$. These emitters also present stellar masses closer to those presented by the $z = 0.84$ Hα-selected population, suggesting that the brightest $z = 0.24$ Hα emitters and part of the sample at $z = 0.84$ might be related.

At higher redshift ($z = 2.23$), one finds Hα emitters residing in the dark matter haloes around $M_{min} \approx 10^{12}\,M_\odot$, and it is therefore possible that at least a fraction of the very active star-forming Hα emitters found at $z = 2.23$ (SFRs $>70\,M_\odot\,\mathrm{yr}^{-1}$) by G08 with HiZELS will turn into the typical Hα emitters at $z = 0.84$ with a strong $L_{H\alpha}^*$ decrease.

5.4 The dark matter host halo–$L_{H\alpha}^*$ relation

The right-hand panel of Fig. 9 presents how the minimum mass of the host dark matter halo changes with measured Hα luminosity (all luminosities are derived assuming a constant 1 mag of extinction as in S09). This shows that while the host halo mass increases with luminosity at any given redshift, there seems to be a different relation for each redshift/epoch. For a given dark matter halo mass, one finds that star formation is tremendously more effective at high-redshift than at lower redshift; this difference is especially pronounced from $z = 0.84$ to 2.23. On the other hand, G08 and S09 demonstrated that there is a clear evolution in the Hα luminosity function, showing that the characteristic luminosity, $L_{H\alpha}^*$, evolves by a factor of >20 from the local Universe to $z = 2.23$; the $L_{H\alpha}^*$ evolution is also most pronounced from $z = 0.84$ to 2.23. This suggests that there could be a relation between the host dark matter halo and $L_{H\alpha}^*$ found at different epochs.

In order to investigate this, the results from the right-hand panel of Fig. 9 are shown in Fig. 10 after scaling the measured luminosities by $L_{H\alpha}^*$ found at each individual redshift.[7] Remarkably, this scaling reveals a clear relation between the typical dark matter halo host of Hα emitters and the value of $L_{H\alpha}/L_{H\alpha}^*$, removing essentially all of the evolution seen in Fig. 9. Even though there is a significant evolution in the Hα emitters population from $z = 2.23$ to the present epoch, at $L_{H\alpha}^*$, Hα emitters seem to reside in $M_{min} \approx 10^{13}\,M_\odot$ dark matter haloes, independently of cosmic time. Galaxies below the luminosity function break reside in least massive haloes, while (at least for $z = 0.84$), Hα emitters above $L_{H\alpha}^*$ seem to reside in haloes just slightly more massive than $M \approx 10^{13}\,M_\odot$.

As these results suggest an epoch-independent constancy between $L_{H\alpha}^*$ and the minimum mass of the dark matter halo host of Hα emitters, it seems plausible to suggest a simple connection

---

[7] The best-fitting values for $L_{H\alpha}^*$ are taken from Shioya et al. (2008) at $z = 0.24$, from S09 at $z = 0.84$, and from the recent determination of Hayes, Schaerer & Ostlin (2010) at $z = 2.23$. The Hayes et al. measurement combines the results of G08 with deeper HAWK-I observations, and suggests a ~ 0.2 dex higher value of $L_{H\alpha}^*$ than G08 derived, but within the errors of that study.
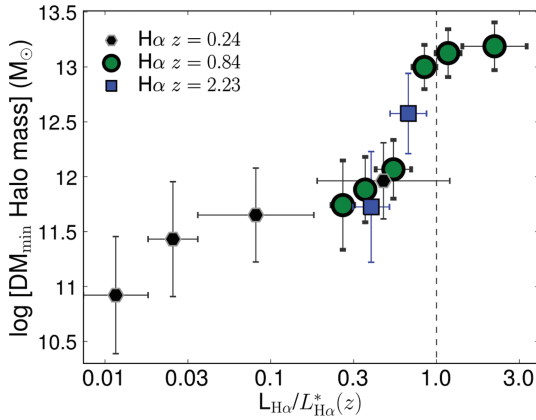
Figure 10. The minimum mass of host dark matter haloes as a function of $L_{H\alpha}/L_{H\alpha}^*(z)$ ($L_{H\alpha}^*(z)$ is the break in the Hα luminosity function, shown to evolve greatly from the local Universe up to $z = 2.23$; see S09 for more details). This reveals a good agreement between the data at three very distinct epochs of the history of the Universe. This shows that luminous Hα emitters reside in more massive haloes, but also that across eight billion years, emitters at the same fraction of $L_{H\alpha}^*$ seem to reside in haloes of typically the same mass.

between the two properties. Indeed, it is possible that the strong evolution in the break of the Hα luminosity function is being driven by the quenching of star formation for galaxies residing in haloes much more massive than $M \approx 10^{12}$ M$_\odot$, either because such haloes favour very intense and fast bursts of star formation, capable of using the majority of the gas or because such halo masses create the necessary conditions for AGN feedback to become important.

6 CONCLUSIONS

The clustering properties of Hα emitters at $z = 0.84$ have been fully detailed. These emitters are moderately clustered, with their angular correlation function $\omega(\theta)$ being very well fitted by a power-law $A\theta^{-\beta}$ with $A = 15.5 \pm 4.1$ (for θ in arcsec) and $\beta = -0.81 \pm 0.05$ out to 600 arcsec (4.5 Mpc). The exact relation between the spatial and angular correlation function is used to show that the Hα emitters at $z = 0.84$ present $r_0 = 2.7 \pm 0.3\,h^{-1}$ Mpc for a spatial correlation function given by $\xi = (r/r_0)^{-1.8}$, with the errors accounting for cosmic variance.

A strong dependence of the correlation length on Hα luminosity is found, with the most active star-forming galaxies presenting $5.2 \pm 1.0\,h^{-1}$ Mpc, while the lower Hα luminosity galaxies present a spatial correlation as low as $2.1 \pm 0.6\,h^{-1}$ Mpc. The correlation length also depends on rest-frame K luminosity (broadly tracing stellar mass) for star-forming galaxies at $z \sim 1$ but not on rest-frame B luminosity (or only very weakly).

The r_0 correlation with $L_{H\alpha}$ is fully maintained at a fixed rest-frame K luminosity, clearly revealing that it is not a simple result of the $L_{H\alpha}-M_K$ correlation; the r_0-M_K correlation is also maintained at a fixed $L_{H\alpha}$ distribution, but slightly weakened.

Irregular galaxies and mergers are found to cluster more strongly than discs and non-mergers, respectively. This, however, seems to be driven by the different Hα luminosity and M_K distributions of the distinct populations; once they are matched on the two properties, irregulars and discs present the same r_0 and mergers and non-mergers are consistent within $\approx 1\sigma$. Bulge-to-disc fraction is also shown not to be important to the clustering of star-forming

galaxies at $z \sim 1$, and morphology seems to be as unimportant for the clustering at high redshift as it is in the local Universe.

Hα emitters at $z = 0.84$ found by HiZELS reside in minimum dark matter haloes of $\approx 10^{12}$ M$_\odot$ similar to those of Milky Way type galaxies in the local Universe. Those are roughly the same as the haloes hosting the brightest Hα emitters at $z = 0.24$ and just slightly higher mass haloes than the hosts of Hα emitters at $z = 2.23$. Furthermore, the minimum dark matter halo mass hosting Hα emitters increases with Hα luminosity, and a $L_{H\alpha}^*$ scaling is able to combine observational results probing the last eight billion years of the age of the Universe in one single relation. This also suggests a connection between the strong $L_{H\alpha}^*$ evolution of the Hα luminosity function and star formation being truncated in galaxies residing within dark matter haloes with masses much higher than $\approx 10^{12}$ M$_\odot$.

The results presented in this study provide important details on the clustering and evolution of Hα-selected star-forming galaxies at $z = 0.84$, which are mostly discs, but with a significant population of irregular and mergers above $L_{H\alpha}^*$. Besides identifying and quantifying clustering relations with fundamental galaxy observables at $z \sim 1$ for the first time, the results clearly show that the highest SFRs at any epoch only occur in galaxies residing within massive haloes. Dark matter haloes of a given mass seem to be more effective at providing the conditions for intense starbursts at higher redshift. On the other hand, the fact that the brightest Hα emitters become rarer with cosmic time down to the present (seen by the strong evolution in the Hα luminosity function; e.g. G08 and S09) suggests that the most massive haloes not only provide the conditions and the environment required for the highest SFRs to take place, but they also seem to be the sites in which the quenching of star formation in galaxies occurs across cosmic time.

REFERENCES

Ball N. M., Loveday J., Brunner R. J., 2008, MNRAS, 383, 907
Beisbart C., Kerscher M., 2000, ApJ, 545, 6
Blain A. W., Chapman S. C., Smail I., Ivison R., 2004, ApJ, 611, 725
Brainerd T. G., Smail I., Mould J., 1995, MNRAS, 275, 781
Capak P. et al., 2007, ApJS, 172, 99
Cirasuolo M., McLure R. J., Dunlop J. S., Almaini O., Foucaud S., Simpson C., 2010, MNRAS, 401, 1166
Colless M. et al., 2001, MNRAS, 328, 1039
da Ângela J. et al., 2008, MNRAS, 383, 565
Efstathiou G., Bernstein G., Tyson J. A., Katz N., Guhathakurta P., 1991, ApJ, 380, L47
Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J., 1994, MNRAS, 266, 50

Gallego J., Zamorano J., Aragon-Salamanca A., Rego M., 1995, ApJ, 455, L1

Garn T. et al., 2010, MNRAS, 402, 2017

Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, MNRAS, 388, 1473 (G08)

Giavalisco M., Dickinson M., 2001, ApJ, 550, 177

Hartley W. G. et al., 2008, MNRAS, 391, 1301

Hayashi M., Shimasaku K., Motohara K., Yoshida M., Okamura S., Kashikawa N., 2007, ApJ, 660, 72

Hayes M., Schaerer D., Ostlin G., 2010, A&A, 509, L5

Ilbert O. et al., 2009, ApJ, 690, 1236

Kong X. et al., 2006, ApJ, 638, 72

Kovač K., Somerville R. S., Rhoads J. E., Malhotra S., Wang J., 2007, ApJ, 668, 15

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Lawrence A. et al., 2007, MNRAS, 379, 1599

Le Févre O., Hudon D., Lilly S. J., Crampton D., Hammer F., Tresse L., 1996, ApJ, 461, 534

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

Lilly S. J. et al., 2009, ApJS, 184, 218

Ly C. et al., 2007, ApJ, 657, 738

McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, MNRAS, 395, 2196

Maraston C., Daddi E., Renzini A., Cimatti A., Dickinson M., Papovich C., Pasquali A., Pirzkal N., 2006, ApJ, 652, 85

Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS, 286, 115

Meneux B. et al., 2009, A&A, 505, 463

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Moscardini L., Coles P., Lucchin F., Matarrese S., 1998, MNRAS, 299, 95

Nakajima A., Shioya Y., Nagao T., Saito T., Murayam T., Sasaki S. S., Yokouchi A., Taniguchi Y., 2008, PASJ, 60, 1249

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Norberg P. et al., 2001, MNRAS, 328, 64

Orsi A., Baugh C. M., Lacey C. G., Cimatti A., Wang Y., Zamorani G., 2009, preprint (arXiv:0911.0669)

Ouchi M. et al., 2005, ApJ, 635, L117

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

Roche N. D., Almaini O., Dunlop J., Ivison R. J., Willott C. J., 2002, MNRAS, 337, 1282

Scarlata C. et al., 2007, ApJS, 172, 406

Scoville N. et al., 2007, ApJS, 172, 1

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Shi Y., Rieke G. H., Ogle P., Jiang L., Diamond-Stanic A. M., 2009, ApJ, 703, 1107

Shim H., Colbert J., Teplitz H., Henry A., Malkan M., McCarthy P., Yan L., 2009, ApJ, 696, 785

Shioya Y. et al., 2008, ApJS, 175, 128

Shioya Y. et al., 2009, ApJ, 696, 546

Simon P., 2007, A&A, 473, 711

Skibba R. A. et al., 2009, MNRAS, 399, 966

Sobral D. et al. 2009a, MNRAS, 398, 75 (S09)

Sobral D. et al. 2009b, MNRAS, 398, L68

Wake D. A. et al., 2008, MNRAS, 387, 1045

Weiß A. et al., 2009, ApJ, 707, 1201

Westra E., Geller M. J., Kurtz M. J., Fabricant D. G., Dell'Antonio I., 2010, ApJ, 708, 534

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

Zehavi I. et al., 2005, ApJ, 630, 1

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