A large narrow-band Hα survey at z ~ 0.2: the bright end of the luminosity function, cosmic variance and clustering across cosmic time

Stroe, A.; Serrano Goncalves Sobral, D.R.

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **453,** 242–258 (2015) doi:10.1093/mnras/stv1555

A large narrow-band Hα survey at $z \sim 0.2$: the bright end of the luminosity function, cosmic variance and clustering across cosmic time

Andra Stroe[1]★ and David Sobral[1,2,3]†

[1]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[2]*Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, Observatório Astronómico de Lisboa, Tapada da Ajuda, P-1359-018, Lisbon, Portugal*
[3]*Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, P-1748-016, Lisbon, Portugal*

Accepted 2015 July 9. Received 2015 July 7; in original form 2015 June 3

ABSTRACT

We have carried out the largest ($>3.5 \times 10^5$ Mpc^3, 26 deg^2) Hα narrow-band survey to date at $z \sim 0.2$ in the SA22, W2 and XMMLSS extragalactic fields. Our survey covers a large enough volume to overcome cosmic variance and to sample bright and rare Hα emitters up to an observed luminosity of $\sim 10^{42.4}$ erg s^{-1}, equivalent to $\sim 11\,M_\odot$ yr^{-1}. Using our sample of 220 sources brighter than $>10^{41.4}$ erg s^{-1} ($>1\,M_\odot$ yr^{-1}), we derive Hα luminosity functions, which are well described by a Schechter function with $\phi^\star = 10^{-2.85\,\pm\,0.03}$ Mpc^{-3} and $L^*_{\rm H\alpha} = 10^{41.71\pm0.02}$ erg s^{-1} (with a fixed faint end slope $\alpha = -1.35$). We find that surveys probing smaller volumes ($\sim 3 \times 10^4$ Mpc^3) are heavily affected by cosmic variance, which can lead to errors of over 100 per cent in the characteristic density and luminosity of the Hα luminosity function. We derive a star formation rate density of $\rho_{\rm SFRD} = 0.0094 \pm 0.0008\,M_\odot$ yr^{-1}, in agreement with the redshift-dependent Hα parametrization from Sobral et al. The two-point correlation function is described by a single power law $\omega(\theta) = (0.159 \pm 0.012)\theta^{(-0.75\,\pm\,0.05)}$, corresponding to a clustering length of $r_0 = 3.3 \pm 0.8$ Mpc h^{-1}. We find that the most luminous Hα emitters at $z \sim 0.2$ are more strongly clustered than the relatively fainter ones. The $L^*_{\rm H\alpha}$ Hα emitters at $z \sim 0.2$ in our sample reside in $\sim 10^{12.5-13.5}\,M_\odot$ dark matter haloes. This implies that the most star-forming galaxies always reside in relatively massive haloes or group-like environments and that the typical host halo mass of star-forming galaxies is independent of redshift if scaled by $L_{\rm H\alpha}/L^*_{\rm H\alpha}(z)$, as proposed by Sobral et al.

Key words: galaxies: evolution – galaxies: formation – galaxies: luminosity function, mass function – large-scale structure of Universe.

1 INTRODUCTION

The star formation (SF) activity in the Universe was significantly higher in the past, reaching a peak ∼10–11 Gyr ago ($z \sim 2$–3; e.g. Lilly et al. 1996; Bouwens et al. 2011; Karim et al. 2011; Gunawardhana et al. 2013; Sobral et al. 2013; Bouwens et al. 2015), and with the typical star formation rate (SFR) of galaxies (SFR★) at $z \sim 2$ being a factor of ∼10 times higher than at $z = 0$ (Sobral et al. 2014). However, the understanding of how, and through which physical mechanisms, the typical SFRs of galaxies have declined over the last 11 Gyr is still poor.

In order to study SF across cosmic time, a number of tracers can be used. Ultraviolet (UV) data can be used to trace radiation coming from massive, short-lived stars. Dust heated by the UV radiation emits in the far-infrared (FIR). The radiation from massive stars also ionizes the surrounding gas and leads to numerous recombination lines, such as Hα (6563 Å) and [O II] (3727 Å). Radio observations can be used to trace emission from supernova remnants. However, it is not trivial to combine these SF indicators, given that they trace different phases of SF (averaged on short, ∼10 Myr, or long, ∼100 Myr, time-scales, dust obscured, etc.), with different selection functions. Some selections are significantly biased: UV-selected samples miss dusty/metal-enriched star-forming galaxies, while the FIR exclusively selects dusty star-forming regions. Therefore, one of the main challenges in obtaining a complete picture of the SF evolution is the direct comparison of equally selected large samples of SF galaxies at a range of redshifts. Samples at high redshift tend to be obtained with a completely different selection than those at lower redshift, which can result in misinterpreted evolutionary trends, which are more likely connected with the different selections at different redshifts than the actual evolution of galaxies across time (e.g. Stott et al. 2013).

★E-mail: astroe@strw.leidenuniv.nl
†VENI/IF Fellow.

An effective way of overcoming such limitations is by using a single technique and a single SF indicator up to the peak of the SF activity. This can be achieved by tracing the Hα emission line, which is one of the most sensitive and well-calibrated SF traces and also benefits from low intrinsic dust extinction within the host galaxy (when compared to, for example, UV). Hα surveys performed using the narrow-band (NB) technique can provide clean, large and complete samples SF galaxies (see Oteo et al. 2015).

A successful example of the NB technique put into practice is the High Redshift Emission Line Survey (HiZELS; Geach et al. 2008; Best et al. 2013; Sobral et al. 2013), but also see the pioneering works of Bunker et al. (1995), Moorwood et al. (2000), Kurk et al. (2004), Ly et al. (2007) and Shioya et al. (2008). At $z \sim 1$–2, the volumes probed by HiZELS over a number of different fields ($\sim$5–10 deg^2) virtually overcome cosmic variance (Sobral et al. 2015b). However, at $z < 0.4$, the volumes probed over 1–2 deg^2 areas are only a minor fraction of those at high redshift. Indeed, the samples at low redshift are greatly limited by cosmic variance, and even the widest surveys, for example, the Cosmological Evolution Survey (COSMOS; see Shioya et al. 2008) struggle to reach the characteristic Hα luminosity ($L^*_{\rm H\alpha}$). An additional limitation is saturation, which means missing the luminous population of Hα emitters (with $>$1–3 M$_\odot$ yr^{-1}; for a discussion of this effect, see Stroe et al. 2014). This can lead to an underestimation of the Hα luminosity function (LF) bright end and an exaggeration of the evolution of $L^*_{\rm H\alpha}$ from high to low redshift.

The combination of all these issues and the different selection techniques applied by each study makes it extremely hard to compare fairly between $z < 0.4$ and $z > 1$ samples when based on the same surveys. While it is possible to use other samples at lower redshift (e.g. spectroscopic selection; Gunawardhana et al. 2013), the importance of using the same selection in order to obtain clean and clear evolutionary trends cannot be stressed enough: without the guarantee of a unique selection, any evolutionary trends become hard/impossible to understand and interpret, limiting our understanding.

In order to overcome the current shortcomings, we clearly require a large Hα survey at lower redshifts, which can be directly matched to higher redshift. In this paper, we present a large survey at $z \sim 0.2$, covering a similar comoving volume (3.5×10^5 Mpc^3, spread over three independent fields to overcome cosmic variance) and complete down to similar luminosity limits relative to $L^*_{\rm H\alpha}$ as surveys at $z > 1$. The structure of the paper is as follows. In Section 2, we present the observations and the reduction of the NB data, while in Section 3, we show the selection of the Hα emitters. Section 4 deals with the $z \sim 0.2$ Hα LF. Section 5 deals with the clustering of bright Hα sources, and we present the implications of our results for the cosmic SF evolution. We present concluding remarks in Section 6.

At the two redshifts probed, $z \sim 0.19$ and 0.22, 1 arcsec covers a physical scale of 3.2 and 3.6 kpc, respectively. The luminosity distance is $d_{\rm L} \approx 940$ Mpc at $z \sim 0.19$ and $\approx$1110 Mpc at $z = 0.22$. All coordinates are in the J2000 coordinate system. We use the Chabrier (2003) initial mass function (IMF) throughout the paper, and results from other studies are also converted to this IMF.

2 OBSERVATIONS AND DATA REDUCTION

We obtain NB data tracing Hα at $z \sim 0.19$ and $\sim$0.22 in three well-studied extragalactic fields located at high Galactic latitude. W2 is part of the Canada–France–Hawaii Telescope Legacy Survey (CFHTLS) 155 deg^2, wide and shallow survey (Gwyn 2012), aimed at studying the large-scale structure and matter distribution using weak lensing and galaxy distribution. SA22 is part of the W4 field in the CFHTLS and multiwavelength data have been compiled by Matthee et al. (2014) and Sobral et al. (2015b). The XMM Large-Scale Structure Survey (XMMLSS; Pierre et al. 2004) is aimed at mapping large-scale structures through clusters and groups of galaxies.

2.1 Narrow-band Hα observations

We obtained NB data using the NOVA782HA and NOVA804HA (Stroe et al. 2014, 2015; Sobral et al. 2015a) filters on the Wide Field Camera (WFC)[1] mounted on the Isaac Newton Telescope (IN, I13BN008, PI Sobral).[2] For brevity, we label the filters NB1 (NOVA782HA) and NB2 (NOVA804HA). Given that the central wavelengths of the filters are 7852.4 and 8036.15 Å, with a full width at half-maximum (FWHM) of 110 Å, the two filters trace Hα emission in the $z = 0.1865$–0.2025 and $z = 0.2170$–0.2330 redshift ranges. Note that given the large field of view of the WFC, a slight blueshift in the filter central wavelength is expected at large off-axis distances. However, given the WFC focal ratio (f/3.29), this effect is expected to be very low (a few per cent; Bland-Hawthorn et al. 2001). Sobral et al. (2015a) and Stroe et al. (2015) characterized the filters with spectroscopy from the Keck and William Herschel Telescopes, with sources located both towards and away from the pointing centre, and found that the redshift distribution of Hα emitters matches that expected from the filter profile, without any noticeable offset.

Observations were conducted on five bright nights, between 2013 October 22 and 26, under $\sim$1 arcsec seeing conditions. A five-position dither pattern was employed for the individual exposures (of 600 s each) to cover the spacings between the four WFC CCDs. Forty-nine individual pointings (of $\sim$0.3 deg^2 each with WFC) split between the three fields (SA22, W2 and XMMLSS) cover an area of almost 13 deg^2 at each of the two redshifts (thus an effective area of $\sim$26 deg^2 combined), tracing a total comoving volume of about 3.63×10^5 Mpc^3. The overlap with the multiwavelength data extends to about 10 deg^2 per redshift.

2.2 Narrow-band data reduction

We reduce the data using the PYTHON-based pipeline described in Stroe et al. (2014). In short, we median combine the sky flats and biases and use the stacks to correct the science data. After detecting sources using the SEXTRACTOR package (Bertin & Arnouts 1996) and masking them in each science exposure, we median combine the exposures to obtain a 'super-flat'. We divide the data through the 'super-flat' to correct for 'fringing'. We then use SCAMP (Bertin 2006) to find astrometric solutions for the science exposures. The exposures are normalized to the same zero-point (ZP) by comparison to the red magnitude in the fourth United States Naval Observatory (USNO) Catalogue (UCAC4; Zacharias et al. 2013). We combine the processed data into final stacked images using SWARP (Bertin et al. 2002). We photometrically calibrate our data against the i-band magnitude from the Sloan Digital Sky Survey (SDSS) Data Release 9 (DR9; Ahn et al. 2012), which covers all our fields (SA22, W2 and XMMLSS). We extract magnitudes within 5 arcsec apertures using SEXTRACTOR (Bertin & Arnouts 1996). This corresponds to a physical diameter of $\sim$18 kpc at $\sim$0.2 redshift.

[1] http://www.ing.iac.es/engineering/detectors/ultra_wfc.htm
[2] http://www.ing.iac.es/Astronomy/telescopes/int/

We calculate 3σ limiting magnitudes using the rms noise reported by SEXTRACTOR (see Table 2). The depth of the observations varies across the pointings and even between the different chips of the WFC. Hence, we calculate the rms noise individually for each CCDs, for each pointing, across the three fields.

We apply the NB technique to select line emitters, using a NB filter tracing line emission within a narrow range in redshift, in combination with another NB or broad-band (BB) filter used for the estimation of the continuum emission underlying the emission line. We use two NB filters to trace Hα emission in two redshift ranges (0.1865–0.2025 and 0.2170–0.2330). For each NB filter, we use the other NB filter to estimate the continuum BB emission. In this way, for line emitters, one NB filter captures the BB emission as well as the line emission, while the other NB filter only captures the stellar continuum emission. Our method is similar to that of Dale et al. (2010), who use twin NB filters for continuum subtraction. In the following, we use labels according to the filter that was used as the NB filter in that particular case. Therefore, when we use the label NB1, we refer to line emitters in the 0.1865–0.2025 redshift range, while NB2 refers to the 0.2170–0.2330 range. The details of the selection method are given in Section 3.2.

2.3 Optical and IR data

In our analysis, we use the rich multiwavelength optical and infrared (IR) data available for the SA22, W2 and XMMLSS fields.

All three fields are part of the CFHTLS wide and shallow surveys (SA22, W2 and XMMLSS are in fields W4, W2 and W1). We make use of the g, r, i and z photometry (Erben et al. 2013) and photometric redshifts (Ilbert et al. 2006) available through the CFHTLS T0007 release.

We also employ near-IR data in the J and K_s filters, down to $\sim$21.2 and $\sim$20.0 mag, respectively, obtained as part of the Visible and Infrared Survey Telescope for Astronomy (VISTA) Hemisphere Survey (VHS; McMahon and the VHS Collaboration, in preparation). Where available, in the XMMLSS field, we preferentially use data from the VISTA Deep Extragalactic Observations (VIDEO) Survey (Jarvis et al. 2013), which is about 3.5 mag deeper than VISTA. We also make use of the IR photometric data taken in the SA22 field as part of the second data release of the UKIDSS Deep Extragalactic Survey (Warren et al. 2007), which reaches 23.4 and 22.8 mag in the J and K_s bands, respectively, with a catalogue from Sobral et al. (2015b).

We make use of the photometric and spectroscopic redshift compilation in the UKIDSS Ultra Deep Survey (part of XMMLSS) available as part of their eighth data release,[3] as well as other publicly available spectroscopy in the XMMLSS field (Garcet et al. 2007; Polletta et al. 2007; Tajer et al. 2007; Melnyk et al. 2013).

3 METHODS AND SELECTING THE Hα SAMPLES

Once sources are detected in the NB images, we cross-match the NB catalogues with the optical and IR catalogues presented in Section 2.3, using a 1-arcsec positional tolerance. Note that because the BB catalogues are deeper than our data by at least 2 mag, we have 100 per cent optical and IR coverage in the areas we have field-of-view overlap with all the multiwavelength data. We use each NB catalogue as a base catalogue for the cross-match.

[3] http://www.nottingham.ac.uk/astronomy/UDS/data/data.html

3.1 Star removal

As explained in Section 2.2, we use the two NB filters to trace Hα emission at two redshifts ranges (0.1865–0.2025 and 0.2170–0.2330). However, given the wavelength coverage of the two adjacent filters, our sample of line emitters is contaminated by stars (see also Stroe et al. 2014). Stars could mimic having an emission line if they have an extremely red or a broad absorption feature, which would lead to a strong colour between the two NB filters. We expect the line emitters selected in the NB2 filter to be particularly contaminated with a population of (L, M) dwarf stars (Kirkpatrick, Henry & McCarthy 1991; Kirkpatrick et al. 1999). They will be selected as having excess in NB2 because their continuum has a broad absorption feature falling within the NB1 filter, leading to an underestimation of the continuum emission. The extremely red BB colours of these sources are also consistent with them being red dwarfs.

We exclude stars using a colour–colour selection criterion using optical and IR colours based on Sobral et al. (2012), keeping in mind the distribution of sources in the colour–colour diagram. This is illustrated in Fig. 1.

Red stars are selected using

$$(g-r) > 2(J-Ks)+1, \quad (g-r) > 0.8, \quad (J-Ks) > -0.7. \tag{1}$$

We select dwarf stars via

$$(g-r) > (7/3(r-i)-2/3), \qquad (g-r) > 1.0. \tag{2}$$

Optically blue stars and dwarf stars with absorption features are selected by

$$(g-r) > 2(J-Ks)+1, \qquad (g-r) < 0.8. \tag{3}$$

We additionally use the 'StarGal' parameter in the CFHTLS photometric redshift catalogue to select stars (Ilbert et al. 2006), which categorizes sources as point-like or extended objects.

Thus, in summary, we label sources as stars if (i) the source passes the red star selection criterion (equation 1), (ii) the source passes the blue star selection criterion (equation 3), (iii) the source passes the dwarf star selection criterion (equation 2) or (iv) the source is classified as a star by the CFHTLS 'StarGal' parameter.

About 60–80 per cent of the sources mimicking emission lines are marked as stars. Spectroscopic observations using NB1 and NB2 (e.g. Stroe et al. 2015; Sobral et al. 2015a) confirm the presence of such stars. All the sources masked as stars are removed from catalogues such that they are not selected as line emitters.

3.2 Selection of line emitters

We use the formalism developed by Bunker et al. (1995), which is widely used in the literature (e.g. Shioya et al. 2008; Sobral et al. 2009; Stroe et al. 2014) to select large numbers of line emitters. We refer the interested readers to those papers for the details of the selection criteria.

We select line emitters separately in each field and each NB filter. For brevity, in the following equations, we label with NB the filter used to select emitters, while we call the other NB filter, used to quantify the continuum emission, the BB filter. Note that we have attempted the selection of line emitters using the i-band filter, following Stroe et al. (2014) and Stroe et al. (2015). However, the relatively deep CFHTLS data become saturated at 17–18 mag and would prevent the selection of bright line emitters. Therefore, the

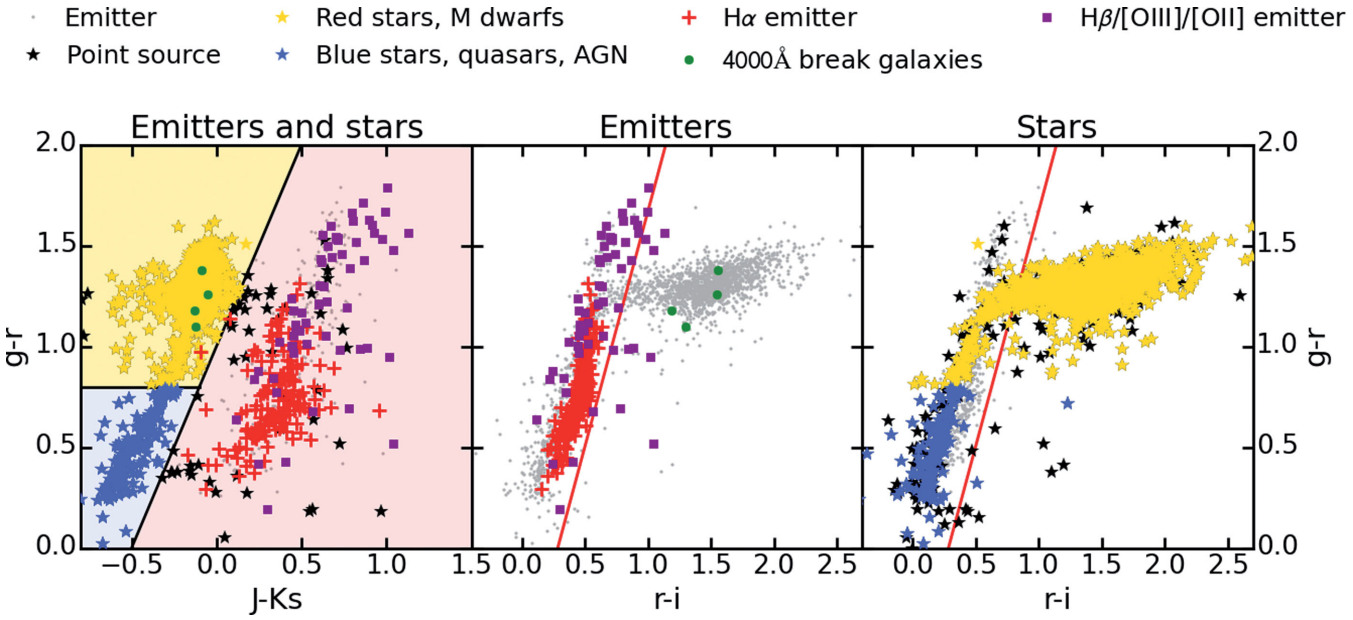

Figure 1. Colour–colour plots for the SA22, W2 and XMMLSS fields, mainly used to remove stars. The first plot shows $g - r$ versus $J - K_s$ while the middle and last plots show $g - r$ versus $r - i$. We first separate stars and emitters using the $g - r$ versus $J - K_s$, and the apply an extra cut using the optical colours to further remove stars with absorption features in one of the filters. The solid red and black lines display the colour cuts used to select point-like objects. Hα emitters are plotted in red crosses, while point-like sources are plotted as stars. 4000 Å break galaxies are plotted as green crosses and high-redshift sources as purple crosses.

use of each NB filter for continuum subtraction of the other represents the optimal strategy, enabling the selection of line emitters up to 10 mag. Using much deeper BB i data would allow us to probe down to fainter emitters, but our aim in this paper is to study the bright population. By comparison, the widest Hα survey at $z \sim 0.2$ to date, performed by Shioya et al. (2008), can only probe sources as bright as ∼18 mag, but excels at the faint end (going down to 24 mag).

We select emitters in each NB filter based on their excess emission compared to the BB emission (quantified using the other NB filter). We first correct for any systematic colour offset between the two NB filters. Colour is defined here as the difference in magnitude between the filter used as NB and the filter used to measure BB. We estimate a median offset of this colour, based on the scatter in the colours at non-saturated, but still bright NB magnitudes. We then apply this correction to the colour and the NB magnitude. However, because the filters are close in wavelength, this correction is small (0.02 and 0.03 mag, for NB1 and NB2, respectively).

The excess emission is then quantified through the colour excess significance Σ, which is used to separate sources with real colour excess, compared to excess caused by random scatter (Sobral et al. 2009; Sobral et al. 2012):

$$\Sigma = \frac{10^{-0.4(m_{\rm BB}-m_{\rm NB})}}{10^{-0.4(ZP_{\rm AB}-m_{\rm NB})}\sqrt{\pi r^2(\sigma^2_{\rm NB}+\sigma^2_{\rm BB})}}, \quad (4)$$

Here, $ZP_{\rm AB}$ is the magnitude system zero-point, $m_{\rm NB}$ and $m_{\rm BB}$ are the NB and BB magnitudes (where NB is the filter used for detection of line emitters and BB is the other NB filter used for quantifying the continuum emission), r is the radius of the aperture in pixels and $\sigma_{\rm NB}$ and $\sigma_{\rm BB}$ are the rms noise levels.

The NB or BB flux $f_{\rm NB,\,BB}$ is calculated as

$$f_{\rm NB,BB} = \frac{c}{\lambda^2_{\rm NB,BB}} 10^{-0.4(m_{\rm NB,BB}-ZP_{\rm AB})}, \quad (5)$$

where c is the speed of light, $\lambda_{\rm NB}$ and $\lambda_{\rm BB}$ are the central wavelengths of the two NB filters and $ZP_{\rm AB} = 48.574$ is the zero-point of the AB magnitude system. The line flux is

$$F_{\rm line} = \Delta\lambda_{\rm NB}(f_{\rm NB} - f_{\rm BB}). \quad (6)$$

Note that the two filters are independent, and hence there is no overlap in wavelength between NB1 and NB2. Therefore, if one filter captures line emission on top of the continuum, automatically the other NB filter picks up only continuum emission. Therefore, the line flux formula accounts for the fact the filter used as BB does not contain any line emission.

We use the Σ parameter in conjunction with an equivalent width (EW) cut, which ensures that we select only sources that have a ratio of the line to continuum flux larger than the scatter at bright magnitudes. The observed EW is defined as

$$\mathrm{EW} = \Delta\lambda_{\rm NB}\frac{f_{\rm NB} - f_{\rm BB}}{f_{\rm BB}}, \quad (7)$$

where $\Delta\lambda_{\rm NB} = 100$ Å is the FWHM of the NB filters, while $f_{\rm NB}$ and $f_{\rm BB}$ are the NB and continuum fluxes. Note that this formula is a simplified version of those presented in, for example, Bunker et al. (1995) and Sobral et al. (2009), because we do not expect our BB filter to contain any emission-line flux.

In the rest frame of the sources, the intrinsic EW_0 is

$$\mathrm{EW}_0 = \mathrm{EW}/(1+z). \quad (8)$$

In conclusion, we select sources as emitters if (i) their colour significance Σ is higher than 3 and (ii) their EW is higher than 3σ, where σ is the colour excess (BB − NB) scatter at bright, but not saturated magnitudes. The $\Sigma = 3$ colour significance and the 3σ excess depend on the depth of the observations in each field (see Fig. 2). We choose not to impose a single, common cut, in order to follow the natural depth of the data, rather than cutting the sample at

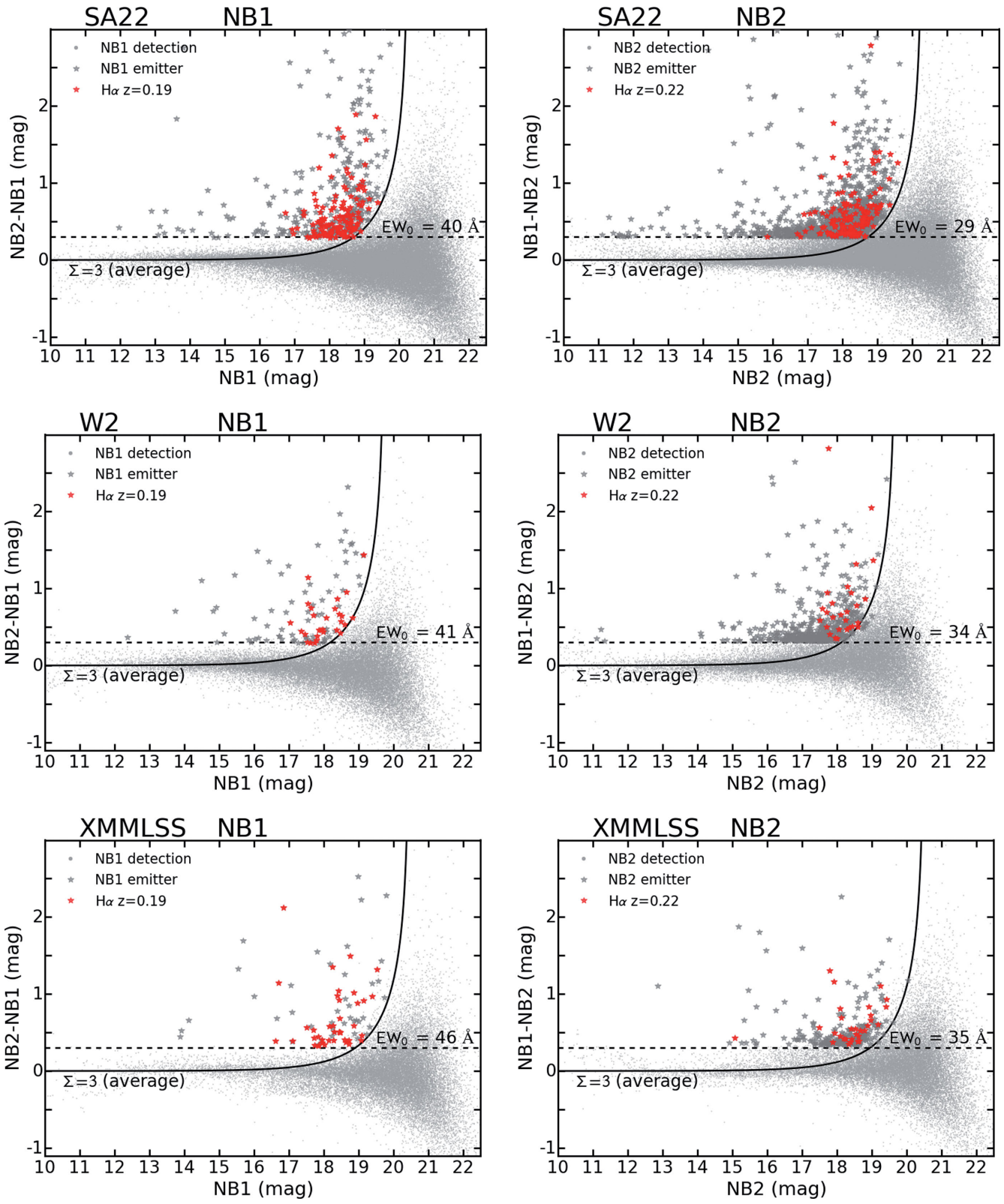

Figure 2. Colour–magnitude diagrams showing the excess as a function of NB magnitude. The selection is performed separately for each CCD/pointing, field and NB filter, using the other NB filter for continuum estimation. Each panel is labelled with the corresponding field and the filter that is used as the NB. The curves show average 3Σ colour significances for the average depth, as the rms value varies between the pointings and CCDs. The horizontal dashed, black lines represent the intrinsic EW cuts. Note that we correct for incompleteness arising from our slightly different EW and colour significance cuts.

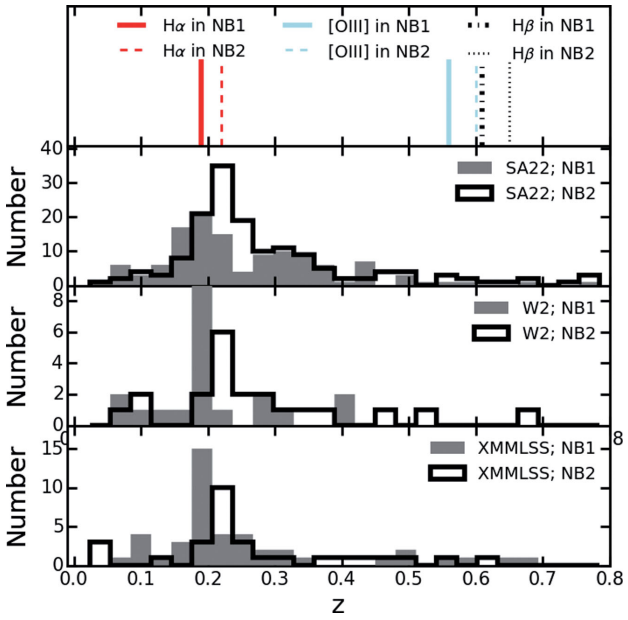

Figure 3. Photometric redshift distribution of line emitters for each field. Note the quality of the photometric redshifts varies between the fields. The top panel shows the main lines we expect to capture with our two NB filters. The distribution contains clear peaks around $z \sim 0.2$, indicating our sample is dominated by Hα emitters, with little contamination from higher redshift emitters.

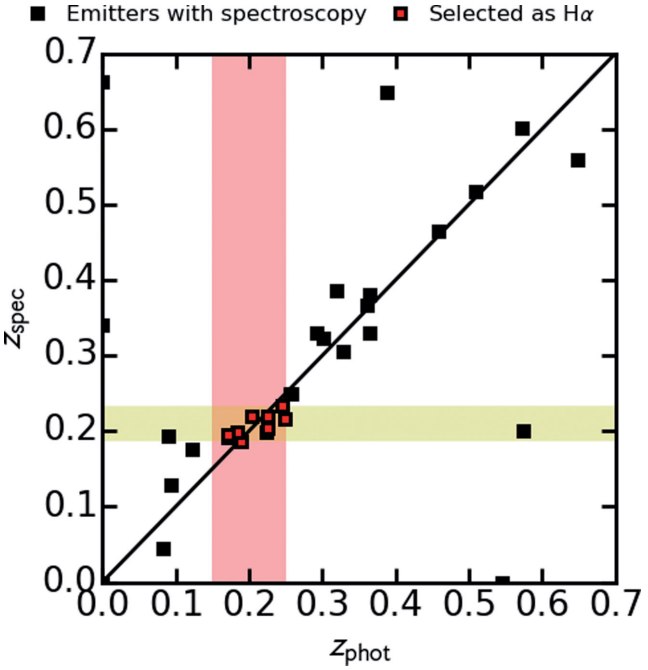

Figure 4. Photometric versus spectroscopic redshift for sources selected as emitters. The vertical shaded red area indicates sources that, based on their photometric redshift, were selected as Hα. The horizontal yellow shaded area indicates the redshift range captured by the filters.

excessively high EW and Σ. However, we note that we fully correct for the sources missed by our cuts, as explained in Section 4.1.

3.3 Selection of Hα candidates

The line emitter population is made of Hα emitters at $z \sim 0.2$, as well as higher-redshift line emitters: Hβ ($\lambda_{rest} = 4861$ Å), [O III]λλ4959, 5007 emitters at $z \sim 0.61$–0.65 and [O II] ($\lambda_{rest} = 3727$ Å) emitters at $z \sim 1.15$ (see Fig. 3). Our sample could be contaminated by a population of 4000 Å break galaxies at $z \sim 0.8$. As shown in Stroe et al. (2014), at ∼8000 Å and lower line fluxes, the line emitter population is dominated by [O II]λ3727 emitters and $z \sim 0.8$ 4000 Å break galaxies. However, at high fluxes, the number of Hα and Hβ/[O III] steeply rises, each amounting to about 50 per cent of the line emitter population. Therefore, given the shallow depth of our survey, we are strongly biased against detecting high-redshift ($z > 0.6$) sources. We expect the Hα emitters to amount to about half of the emitter population. Fig. 3, which presents the photometric redshift distribution of the line emitters, confirms these findings. The steps we undertake to robustly separate the Hα emitters from the other sources are described in the following paragraphs.

We first visually inspected all line emitter candidates to flag any spurious sources coming from noisy edge regions of the chips or from false detections within the haloes of bright sources.

Hα emitters are selected in the following way.

(i) The photometric or spectroscopic redshift of the source does not lie in the expected ranges for Hβ/[O III]/[O II] emitters ($0.37 < z < 0.7$ and $0.9 < z < 1.2$) and 4000 Å break galaxies ($0.7 < z < 0.9$).

(ii) The photometric or spectroscopic redshift of the source lies in the $0.15 < z < 0.25$ range.

Fig. 1 displays the colour–colour distribution of line emitters and the cut employed to separate the source types, and highlights the location of the Hα emitters. All three fields and both filters are shown in the same plot. Separating the data per field and filter results in colour–colour diagrams that are consistent with Fig. 1, indicating there are no systematic differences between the populations selected with the two NB filters. The number of Hα emitters selected in each field can be found in Table 3, amounting to a total of 220 Hα emitters. This amounts to almost 40 per cent of the total number of emitters, as expected and as explained in Section 3.3.

3.3.1 Purity of the Hα sample

We compare the spectroscopic and photometric redshifts in order to study the purity of the Hα sample (Fig. 4). We find that the photometric redshifts are within 0.05 of the spectroscopic redshifts. From the sources spectroscopically confirmed to be at lower or higher redshift, none makes it into the Hα catalogue, implying a very low contamination. Note that the range we used for selecting sources as Hα from photometric redshifts is 0.15–0.25, which is large enough to capture Hα emitters in both filters, while minimizing contamination. Out of 12 spectroscopically confirmed emitters, we miss two sources, implying completeness higher than 80 per cent. However, the spectroscopy is limited and the low number statistics could lead to an overestimation or underestimation of the completeness and contamination. Future spectroscopic observations will allow us to further investigate this.

4 Hα LUMINOSITY FUNCTION AND STAR-FORMATION RATE DENSITY

We use the sample of 220 Hα sources to build LFs.

Our filters are sensitive not only to Hα, but also to the adjacent [NII] double (6450 and 6585 Å) forbidden line. We subtract the [NII] contribution from the line fluxes using the method from Sobral et al.

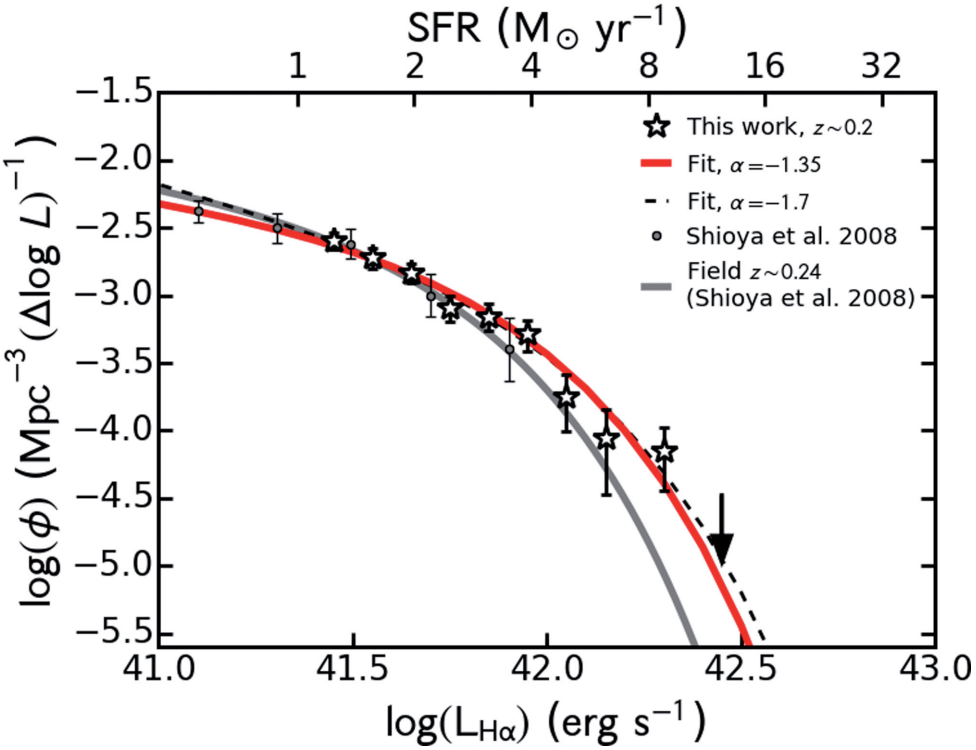

Figure 5. Hα LF at $z \sim 0.2$ from our study and the best-fitting Schechter function. $L_{H\alpha}$ is not corrected for intrinsic dust attenuation. For comparison, the results from Shioya et al. (2008) are also shown. Note the excellent agreement between the data in the overlapping luminosity range. However, our data probe brighter luminosities, enabling the first determination of the bright end of the Hα LF at $z \sim 0.2$.

(2012) to obtain Hα fluxes ($F_{H\alpha}$), which has been spectroscopically confirmed by Sobral et al. (2015b). The average [NII] contribution is about 30 per cent of the total line flux.

After we obtain pure Hα fluxes $F_{H\alpha}$, we calculate the Hα luminosity $L_{H\alpha}$,

$$L_{H\alpha} = 4\pi d_L^2(z) F_{H\alpha}, \quad (9)$$

where $d_L(z)$ is the luminosity distance (940 Mpc for the NB1 filter and 1110 Mpc for NB2).

4.1 Completeness, volume and filter profile corrections

We use the method of Sobral et al. (2012) to correct for the incompleteness arising from missing sources with faint Hα fluxes and/or low EW. We select random samples of sources passing the selection criteria for being located at the redshifts traced by the two filters, but which are not selected as Hα emitters. Fake Hα emission lines are added to these sources, which are then passed through the Hα selection criteria (EW and Σ) described at the end of Section 3.3.

Because of the different depth between the pointings and between the four CCD chips, we independently study the recovery rate as a function of the Hα flux for each chip, pointing, filter and field. The results of the completeness study can be found in Fig. A1 of Appendix A. Our results are corrected for the effects of incompleteness, especially the Hα LF (see Sections 4.3 and 4.4 and, for example, Figs 5–7).

The volumes probed in each field and at each redshift, assuming that the filters have a perfect top-hat shape, are listed in Table 1. The total comoving volume probed is 3.63×10^5 Mpc^3, by far the largest volume ever surveyed in Hα at $z \sim 0.2$. However, because the filter transmission does not follow perfectly an idealized top hat, we follow the method of Sobral et al. (2009) and Sobral et al. (2012) and correct the volumes to account for sources missed at the edges of the filter.

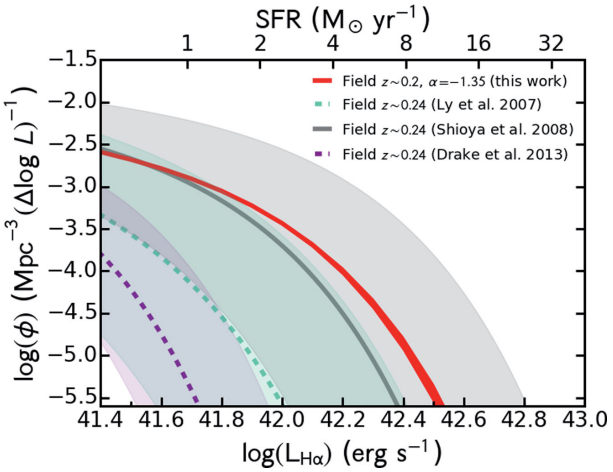

Figure 6. A range of LFs at $z \sim 0.2$, from the current work and the works of Ly et al. (2007), Shioya et al. (2008) and Drake et al. (2013). In shaded areas, we overplot the ranges allowed by the 1σ error bars of the LF parameters. The works of Ly et al. (2007), Shioya et al. (2008) and Drake et al. (2013) explore the faint-end part of the luminosity. The shaded areas indicate the 1σ uncertainties of the Schechter function parameters. Our measurements are consistent with previous work, but significantly improve the previously unexplored bright end, while our measurement error is given by cosmic variance, as shown in Section 4.6. However, the other measurements do not include the error given by cosmic variance, which would add an error of about 100–200 per cent in the parameters.

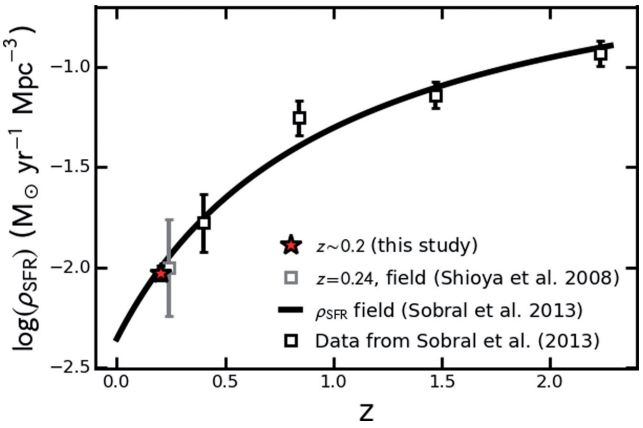

Figure 7. Evolution of the SFRD from $z \sim 2.23$ to $z \sim 0.2$. Our measurement at $z \sim 0.2$ confirms the previously discovered decline in SFRD, which can be simply parametrized as a function of redshift (Sobral et al. 2013).

Table 1. Area and volumes covered by the narrow-band observations. Only the common area between the two filters is listed. The same area is used to calculate the comoving volume.

Field	No pointings	Area (deg^2)	z	Volume (10^4 Mpc^3)
SA22	24	6.1	0.19	7.5
			0.22	9.8
W2	12	3.6	0.19	4.4
			0.22	5.7
XMMLSS	13	3.1	0.19	3.9
			0.22	5.0
Total	49 × 2	12.8 × 2		36.3

Table 2. Typical 3σ limiting magnitudes for the three fields (including the standard spread in values), for each filter. The depth for each pointing (and within each CCD of out of the four WFC CCDs) varies across the fields over the ranges reported in the third column.

Field	Filter	3σ (mag)
SA22	NB1	$17.5^{+0.4}_{-0.3}$
	NB2	$17.4^{+0.4}_{-0.3}$
W2	NB1	$16.8^{+1.5}_{-0.6}$
	NB2	$16.7^{+0.7}_{-0.4}$
XMMLSS	NB1	$17.7^{+0.4}_{-0.3}$
	NB2	$17.5^{+0.5}_{-0.3}$

4.2 Survey limits

A 50 per cent completeness (see Fig. A1) translates to average limiting Hα luminosities of $10^{41.1-41.6}$ erg s^{-1} for our survey. This is equivalent to limiting SFRs of 0.5–1.8 M$_\odot$ yr^{-1}, with no intrinsic dust extinction applied. If 1 mag of dust extinction is applied, then this is equivalent to 0.2–0.8 SFR* (see equation 11).

The maximum observed Hα luminosity that our survey probes is $\sim 10^{42.4}$ erg s^{-1}, equivalent to SFRs of 11 M$_\odot$ yr^{-1} (or $\gtrsim$ 27 M$_\odot$ yr^{-1} if 1 mag of dust extinction is applied). By comparison, the widest Hα survey at a similar redshift, performed by Shioya et al. (2008), reaches $\sim 10^{41.9}$ erg s^{-1}, or 3.5 M$_\odot$ yr^{-1} (8.7 M$_\odot$ yr^{-1} with dust extinction). This means that our survey probes galaxies more than three times more star-forming than previous surveys.

Table 3. Number of line emitters and Hα emitters selected in each field and filter. We also list the average limiting observed Hα luminosity at 50 per cent completeness and the equivalent SFR (using equation 11).

Field	Filter	Emitters	Hα emitters	log ($L_{\rm H\alpha}$) (erg s^{-1})	SFR (M$_\odot$ yr^{-1})
SA22	NB1	153	59	41.4	1.1
	NB2	238	91	41.4	1.1
W2	NB1	33	13	41.4	1.1
	NB2	55	15	41.6	1.7
XMMLSS	NB1	51	23	41.1	0.5
	NB2	50	19	41.4	1.1
Total	Both	576	220		

4.3 Hα luminosity function

Using our final sample of Hα emitters, we build LFs that characterize the density of sources at any given Hα luminosity. To do so, we bin sources based on their luminosity (corrected for the [NII] contribution – Section 4 – but not for intrinsic dust extinction), by adding their associated inverse comoving volume, corrected for the real filter profile and incompleteness (as shown in Section 4.1).

We fit the binned data with a Schechter (1976) parametrization:

$$\phi(L_{\rm H\alpha}){\rm d}L_{\rm H\alpha} = \phi^* \left(\frac{L_{\rm H\alpha}}{L^*_{\rm H\alpha}}\right)^{\alpha} {\rm e}^{-(L_{\rm H\alpha}/L^*_{\rm H\alpha})} {\rm d}\left(\frac{L_{\rm H\alpha}}{L^*_{\rm H\alpha}}\right). \quad (10)$$

Here, $L^*_{\rm H\alpha}$ is the characteristic Hα luminosity, ϕ^* is the characteristic density of Hα emitters and α is the faint-end slope of the LF. Because our data are not deep enough to properly constrain the faint-end slope of the LF (see Table 3), we fix α to two values previously derived in the literature using deep data: −1.35 from Shioya et al. (2008) and −1.7 from Ly et al. (2007). In fitting the LFs, we assume Poissonian errors.

Our best-fitting Hα LF is described by a typical luminosity $\log(L^*_{\rm H\alpha}) = 10^{(41.71\pm0.02)}$ erg s^{-1} and a characteristic density $\log(\phi^*) = 10^{(-2.85\pm0.03)}$ Mpc^{-3} (see Table 4 and Fig. 5). Our data sample really well the bright end of the LF, which enables us to place tight constraints on ϕ^* and $L^*_{\rm H\alpha}$ (errors lower than 15 per cent). However, we lack depth (lowest bin at $\sim 10^{41.4}$ erg s^{-1}), so we fix the faint-end slope to −1.35, as obtained by Shioya et al. (2008) from the previously widest Hα survey, which benefits from high-quality, deep data reaching luminosities of $10^{39.3}$ erg s^{-1}, but is limited at the bright end. Therefore, the two surveys are highly complementary. Within the overlapping regions with data from both Shioya et al. (2008) and our survey, the measurements are in excellent agreement. However, our LF, constrained up to $L_{\rm H\alpha} = 10^{42.5}$ erg s^{-1}, indicates a slightly larger value of $L^*_{\rm H\alpha}$, but still consistent with Shioya et al. (2008) within their large error bars (see Fig. 6). Note that their uncertainties do not include the error from cosmic variance, which can result in 100–200 errors in the parameters of the LF (see Section 4.6). Any discrepancy between the results can be explained by cosmic variance, given that Shioya's volume is ∼10 times smaller than ours and probes a single field. The differences between the ϕ^* results could also be explained by the different colour–colour

Table 4. Best-fitting LF at $z \sim 0.2$ obtained from combining data in the three fields (SA22, W2 and XMMLSS) and two NB filters. Because our data are not very deep, but probe the bright end really well, we fix the faint-end slope α at two values. For comparison, we also list the results and volumes probed from other studies at a similar redshift. Note that none of the $L^*_{H\alpha}$ is corrected for Hα extinction.

Source	z	V (10^4 Mpc^3)	α	$\log \phi^\star$ (Mpc^{-3})	$\log L^*_{H\alpha}$ (erg s^{-1})
This study	∼0.2	36.3	−1.35	-2.85 ± 0.03	41.71 ± 0.02
			−1.70	-3.06 ± 0.04	41.83 ± 0.03
Shioya et al. (2008)	∼0.24	3.1	$-1.35^{+0.11}_{-0.13}$	$-2.65^{+0.27}_{-0.38}$	$41.54^{+0.38}_{-0.29}$
Ly et al. (2007)	∼0.24	0.5	-1.70 ± 0.10	-2.98 ± 0.40	41.25 ± 0.34
Drake et al. (2013)	∼0.25	1.2	$-1.03^{+0.17}_{-0.15}$	$-2.53^{+0.17}_{-0.21}$	$40.83^{+0.19}_{-0.16}$

methods used to separate the Hα emitters from higher-redshift line emitters.

The discrepancy with other studies is much larger however (see Fig. 6). Compared to our results, Ly et al. (2007) slightly overestimate ϕ^* (not significant) and underestimate $L^*_{H\alpha}$ (at the 2σ level). Drake et al. (2013) obtain an $L^*_{H\alpha}$ that is highly underestimated ($10^{40.83}$ erg s^{-1}). The difference to our value is significant at the 11σ level. This is entirely driven by Drake's small volume (∼30 times smaller than ours) and the long exposures they were using in their study, which prevented the study of sources brighter than 20 mag in the NB filter. Given the large variations in the LF parameters from cosmic variance, we expect all these results to be consistent with our measurement, once the cosmic variance error is folded in (see Section 4.6).

4.4 Star formation rate density

We can calculate the star formation rate density (SFRD) at $z \sim 0.2$ by integrating the LF and converting Hα luminosity to SFR. We use the $L_{H\alpha}$ to SFR conversion from Kennicutt (1998), corrected for the Chabrier (2003) IMF:

$$SFR(\mathrm{M_\odot yr^{-1}}) = 4.4 \times 10^{-42} L_{H\alpha}(\mathrm{erg\,s^{-1}}). \quad (11)$$

The luminosity density is obtained by integrating the Hα LF,

$$\rho_{L_{H\alpha}} = \int_0^\infty \phi(L_{H\alpha}) L_{H\alpha} \mathrm{d}L_{H\alpha} \quad (12)$$

$$= \Gamma(\alpha + 2)\phi^* L^*_{H\alpha}, \quad (13)$$

where $\Gamma(n) = (n-1)!$ is the Gamma function. By converting from luminosity to SFR through equation (11), the SFRD ρ_{SFR} is

$$\rho_{SFR} = \Gamma(\alpha + 2)\phi^* L^*_{H\alpha} 10^{0.4 A_{H\alpha}} (1 - f_{AGN}), \quad (14)$$

where $A_{H\alpha}$ is the intrinsic Hα dust extinction, which we assume to be 1 mag, and $f_{AGN} = 0.15$ is the fraction of the Hα luminosity expected to be due to contributions from broad-line and narrow-line AGN emission (e.g. Garn & Best 2010; Sobral et al. 2015a).

Our measurement of the SFRD is $\rho_{SFRD} = 0.0094 \pm 0.0008\,\mathrm{M_\odot\,yr^{-1}\,Mpc^{-3}}$, which matches with the value of Shioya et al. (2008) ($0.010 \pm 0.006\,\mathrm{M_\odot\,yr^{-1}\,Mpc^{-3}}$). Sobral et al. (2013) derive a redshift-dependent parametrization of the SFRD [$\rho_{SFRD} = -2.1/(1+z) + \log 10(4.4/7.9)$, corrected for the Chabrier IMF] based on their measurements and results from Ly et al. (2007) at $z \sim 0.08$ and Shioya et al. (2008) at $z \sim 0.24$ (see Fig. 7). Our measurement perfectly agrees with the parametrization, which predicts a value of 0.01 at $z \sim 0.2$.

4.5 Distribution of Hα emitters

Fig. 8 shows the distribution of the Hα emitters in the three fields at the two redshifts, as selected in Section 3.3. Note the high degree of cosmic variance within and between the field and at the adjacent redshifts.

On average, down to a limiting Hα luminosity of $10^{41.4}$ erg s^{-1} or $SFR \sim 1\,\mathrm{M_\odot\,yr^{-1}}$, we find ∼ two H$\alpha$ emitters per deg^2 (or ∼ three per Mpc^3). However, there are large areas with no emitters, while parts of the W2 and XMMLSS fields have densities of up to 20 sources per deg^2. The 'Sausage' massive, young post-merger galaxy cluster Stroe et al. (2014, 2015), where Hα emitters were selected with the NB1 filter, was found to be extremely dense in star-forming galaxies and AGNs, compared to blank fields. Down to the faintest Hα luminosities as our current data survey ($10^{41.1}$ erg s^{-1}), the density is ∼140 emitters per deg^2, about 70 times above the average we find over an area of 20 deg^2. Assuming Poissonian noise, the 'Sausage' cluster overdensity is significant at the $>11\sigma$ level.

The older 'Toothbrush' galaxy cluster merger, where the two subclusters collided about 2 Gyr ago, behaves differently. The density is about ∼16 emitters per deg^2, densities similar to the densest parts of our wide, shallow Hα survey. Our results thus corroborate the conclusions from Stroe et al. (2014) and Stroe et al. (2015).

4.6 Quantifying cosmic variance

One of our goals is to understand the impact of cosmic variance and low number statistics on the determination of the LF parameters, especially motivated by the differences in LF found with the previous studies of Ly et al. (2007) and Drake et al. (2013). We generate random subsamples of Hα emitters, probing a range of volumes. We perform 1000 realizations starting from the smallest volumes for which we can fit a LF, up to the entire volume of our survey. We perform this experiment using Hα emitters in each NB filter and we also combine all the data together, following Sobral et al. (2015b).

The number of sources for each realization is plotted in Fig. 10. As expected, the average number of sources increases with the volume surveyed. We calculate the standard deviation of the spread in number of sources at each volume and compare that to the Poissonian error. In the calculation of the Poissonian error, we take into account the fact that the sources are divided into bins. At very low volumes, the relative Poissonian error dominates over the spread in the number of sources, which is caused by cosmic variance. Given the depth of our survey, at the very small volumes ($<2 \times 10^4$ Mpc^3) the Poissonian error essentially goes to infinity. Overall, the total relative error, calculated as the sum in quadrature of the Poissonian and cosmic variance error, goes down with increasing volume.

Naturally, when surveying a smaller volume, the number of Hα sources is proportionally smaller. Therefore, we adapt the number

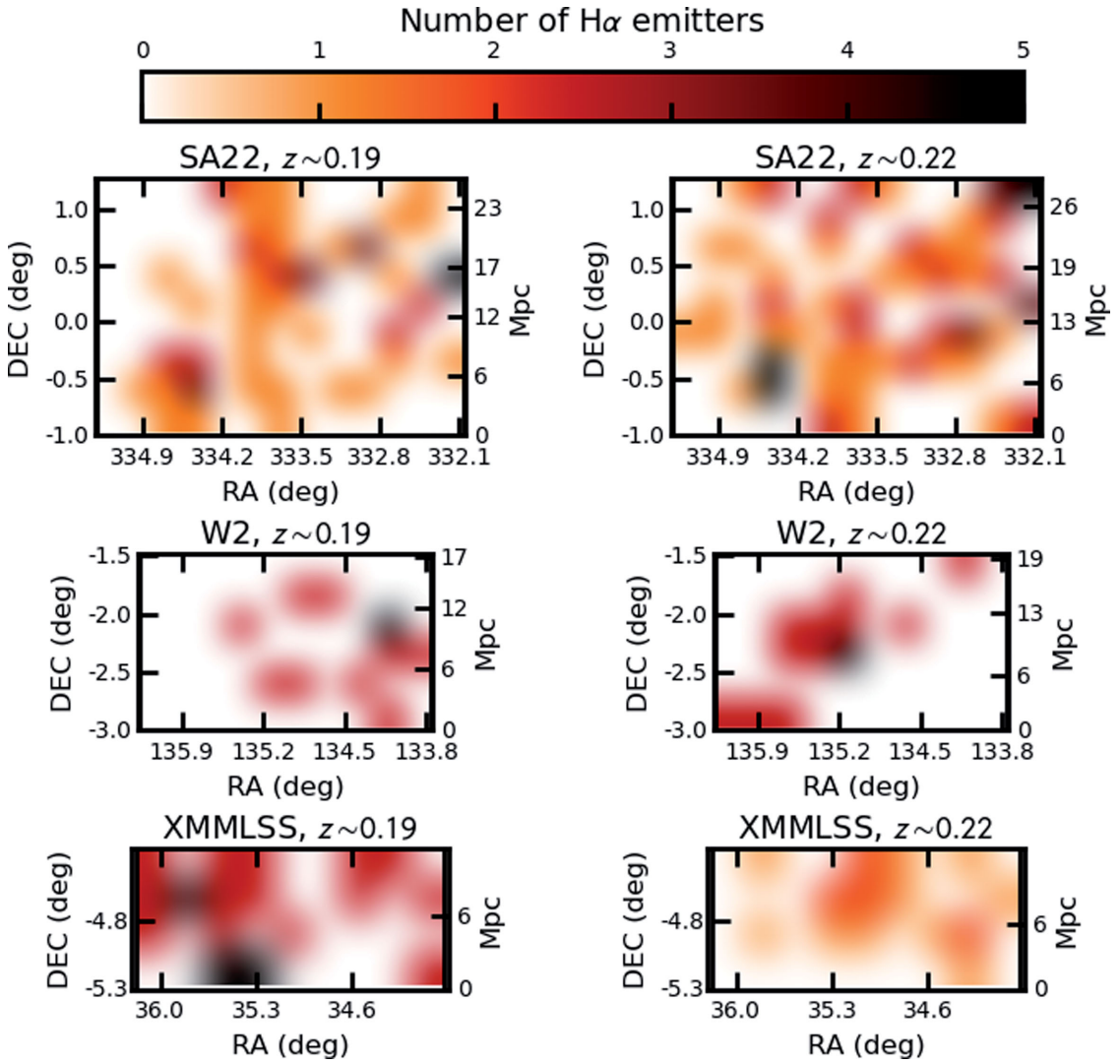

Figure 8. Smoothed sky distribution of the Hα emitters. Note the amount of cosmic variance within the fields. On average, two emitters are found per deg^2, but the values vary between zero and five sources per deg^2.

Table 5. Bin width $\Delta \log L_{\mathrm{H}\alpha}$, starting bin $\log L_{\mathrm{H}\alpha,\,\mathrm{min}}$ and number of bins (N) chosen for studying the LF, depending on the volume V probed.

V range	$\Delta \log L_{\mathrm{H}\alpha}$	$\log L_{\mathrm{H}\alpha,\,\mathrm{min}}$	N_{bins}
$<2 \times 10^4$ Mpc^3	0.3	41.5	4
$2 \times 10^4 - 9 \times 10^5$ Mpc^3	0.2	41.5	4
$9 \times 10^5 - 18 \times 10^5$ Mpc^3	0.15	41.4	5
$18 \times 10^5 - 27 \times 10^4$ Mpc^3	0.15	41.4	6
$>27 \times 10^5$ Mpc^3	0.1	41.4	8

of bins (N), the bin width $\Delta \log L_{\mathrm{H}\alpha}$ and the starting bin $\log L_{\mathrm{H}\alpha}$, depending on the volume V probed, as detailed in Table 5.

The results from the different realizations of the LF calculated from Hα emitters extracted over a range of volumes can be found in Fig. 9. At small volumes ($<4 \times 10^4$ Mpc^3), the random realizations of the LF give wildly different results, with values spanning 4–5 dex. This is driven by two main factors: low number statistics and cosmic variance. The low number of Hα emitters in small volumes imposes wide and few $L_{\mathrm{H}\alpha}$ bins to gain enough number statistics. With few bins, the LF function is barely constrained. Additionally, small volumes do not fully sample the LF at the brightest $L_{\mathrm{H}\alpha}$, where Hα emitters are rare. Therefore, when the volumes are small, cosmic variance is significant. However, with the increase of the probed volume, we can much better constrain the ϕ^* and $L^*_{\mathrm{H}\alpha}$ parameters, by overcoming both Poissonian errors and cosmic variance. This is exemplified in Fig. 11. The standard deviation of the $L^*_{\mathrm{H}\alpha}$ and ϕ^* parameters at each volume size becomes smaller with increasing volume. Note, however, that the values of $L^*_{\mathrm{H}\alpha}$ and ϕ^* are highly correlated (Fig. B1).

As shown in this section, cosmic variance can fully explain the differences found in the literature regarding the Hα LF at $z \sim 0.2$. By accounting for cosmic variance, our LF results can be reconciled with those of Drake et al. (2013) and Ly et al. (2007). Our results indicate that at $z \sim 0.2$, volumes of at least 10^5 Mpc^3 are required to overcome cosmic variance.

5 CLUSTERING OF Hα EMITTERS

To study the clustering of our sample of 220 bright Hα emitters at $z \sim 0.2$, we start by generating a random catalogue with one million sources. The random catalogue sources follow the geometry of the

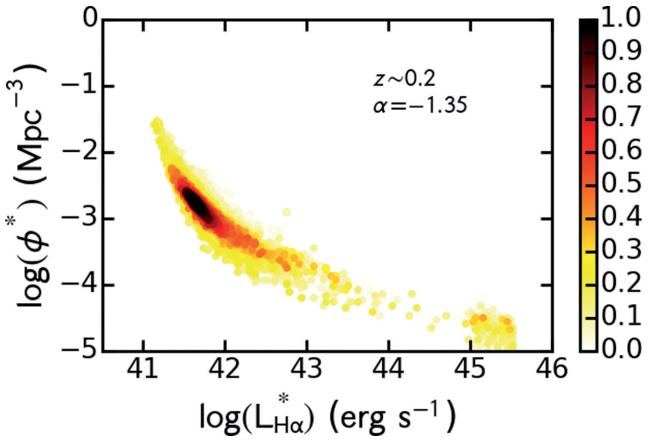

Figure 9. Values of LF Schechter parameters ϕ^* and $L^*_{H\alpha}$, when we fix $\alpha = -1.35$. For fitting the LF, we create 1000 random subsamples of Hα emitters, at a range of probed volumes. The data points are colour-coded with the comoving volume probed in units of 1000 Mpc^3. Note how, at small volumes, the scatter of the values is extremely large (up to 4–5 dex), while at large volumes the values for ϕ^* and $L^*_{H\alpha}$ converge. We obtain similar results with a different value of α or when we use the data for the two filters separately (see Fig. B1)

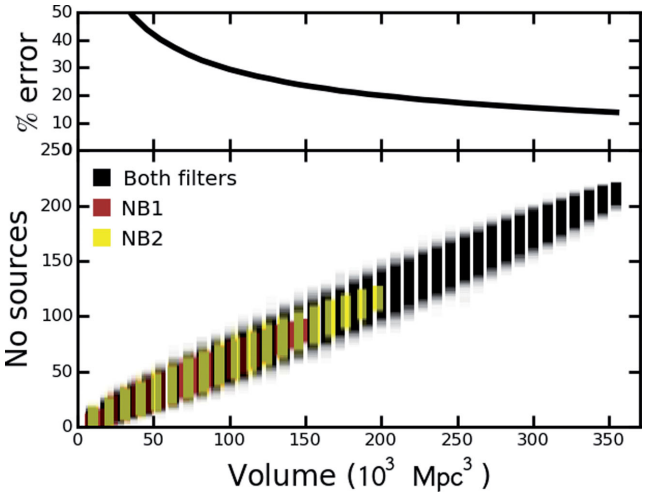

Figure 10. Distribution of the number of Hα emitters randomly selected within a range of volumes. As expected the larger the volume, the larger the number of sources, with a spread at each volume size caused by cosmic variance. The Poissonian error relative to the mean number of sources does not dominate over the spread caused by cosmic variance, except where cosmic variance is minimized through the sampling of a large volumes.

actual observed fields and masked areas (due to saturated stars) and their number in each CCD of each pointing is normalized according to the depth attained (and hence the density of sources in that area).

We follow the method described in detail in Sobral et al. (2010), which evaluates the two-point angular correlation function minimum variance estimator proposed by Landy & Szalay (1993):

$$\omega(\theta) = 1 + \left(\frac{N_R}{N_D}\right)^2 \frac{DD(\theta)}{RR(\theta)} - 2\frac{N_R}{N_D}\frac{DR(\theta)}{RR(\theta)}, \quad (15)$$

where θ is the angle on the sky, N_R and N_D are the numbers of sources in the random and real catalogues of Hα sources, respectively, DD(θ), RR(θ) and DR(θ) are the number pairs of sources located at distances between θ and $\theta + \delta\theta$ in the real data, random data and between real and random data, respectively.

Errors on $\omega(\theta)$ are then (Landy & Szalay 1993)

$$\Delta\omega(\theta) = \frac{1 + \omega(\theta)}{\sqrt{DD(\theta)}}. \quad (16)$$

We determine $\omega(\theta)$ using 1000 different randomly selected subsamples of sources selected from the randomly generated catalogue. We perform our analysis separately on emitters selected in each filter, but combine the data for the SA22, W2 and XMMLSS fields. We use the full luminosity range ($L_{H\alpha} = 10^{41.0-42.4}$ erg s^{-1}) of the Hα emitters, as well as splitting the sample into two roughly equal halves: a faint sample with luminosities in the range $10^{41.0-41.55}$ erg s^{-1} and a bright sample with luminosities $10^{41.55-42.40}$ erg s^{-1}. We bin the data using a range of angular scale bins (with different starting bin θ_{min}, bin width $\delta\theta$ and maximum bin θ_{max}).

The results are presented in Fig. 12 and Table 6. The two-point correlation function for the samples is well described by a single power law. The results for the two filters are considered separately and when combined give fully consistent results within the error bars.

Note we studied only the range $0.^\circ02 < \theta < 3.^\circ0$, where there was enough signal. At scales smaller than $<0.^\circ02$, a flattening of $\omega(\theta)$ occurs, perhaps caused by bright Hα emitters not being able to reside in a single halo. Additionally, because our survey is not very deep, we do not probe the regime where satellites are expected. Therefore, we cannot evaluate the departure of the two-point correlation function from a single power law, which is caused by the transition from the large-scale – that is, two galaxies residing in a separate dark matter (DM) halo – to the small-scale clustering regime (galaxies sharing a single halo; e.g. Ouchi et al. 2005).

Previous research indicates that bright Hα galaxies as well as Lyman break galaxies are more clustered than faint ones. Shioya et al. (2008) found that the two-point correlation function for faint Hα emitters ($L_{H\alpha} < 10^{40.54}$ erg s^{-1}) at $z \sim 0.24$ follows the relationship $\omega(\theta) = (0.011 \pm 0.002)\theta^{(-0.84 \pm 0.05)}$, while brighter emitters with $10^{40.54} < L_{H\alpha} \lesssim 10^{41.5}$ erg s^{-1} follow the relationship $\omega(\theta) = (0.019 \pm 0.004)\theta^{(-1.08 \pm 0.05)}$. The amplitude of the two-point correlation function for our faint sample is 0.208 ± 0.035, while for the bright sample it is slightly larger: 0.295 ± 0.026. The relation is also steeper for the bright sample than for the faint sample. Our results therefore support and extend the claim that brighter (and hence more star-forming galaxies) are more clustered than faint ones to very high luminosities beyond $10^{41.0}$ erg s^{-1} up to $10^{42.4}$ erg s^{-1} ($L/L^*_{H\alpha} \sim 5.0$).

We use the inverse Limber transformation and the redshift distribution of the NB filters to translate the two-point correlation function into a three-dimensional spatial correlation (Peebles 1980), assuming the latter is well described by $\epsilon = (r/r_0)^\gamma$, where r_0 is the real-space correlation length of the Hα emitters. Following the method of Sobral et al. (2010), we assume that the two filters have a perfect top-hat shape. We compute r_0 for each realization of $\omega(\theta)$ in each filter, by fixing $\beta = -0.8$. We finally combine the data for the two filters. The dependence of r_0 on redshift is shown in Fig. 13.

For the full sample, we obtain a correlation length $r_0 = 3.3$ Mpc h^{-1} with a standard deviation 0.8 Mpc h^{-1}. We obtain $r_0 = 3.5 \pm 1.1$ Mpc h^{-1} for our fainter Hα sample and 5.0 ± 1.5 Mpc h^{-1} for the brighter sample. Our measurements are larger than those of Sobral et al. (2010) at $z \sim 0.24$ (based on the sample from Shioya et al. 2008), who find a value of 1.8 ± 0.2 Mpc h^{-1} for their sample with $10^{39.4} < L_{H\alpha} < 10^{41.5}$ erg s^{-1}. As expected, fainter Hα galaxies have smaller correlation lengths than brighter ones (Norberg et al. 2001; Shioya et al. 2008; Sobral et al. 2010).

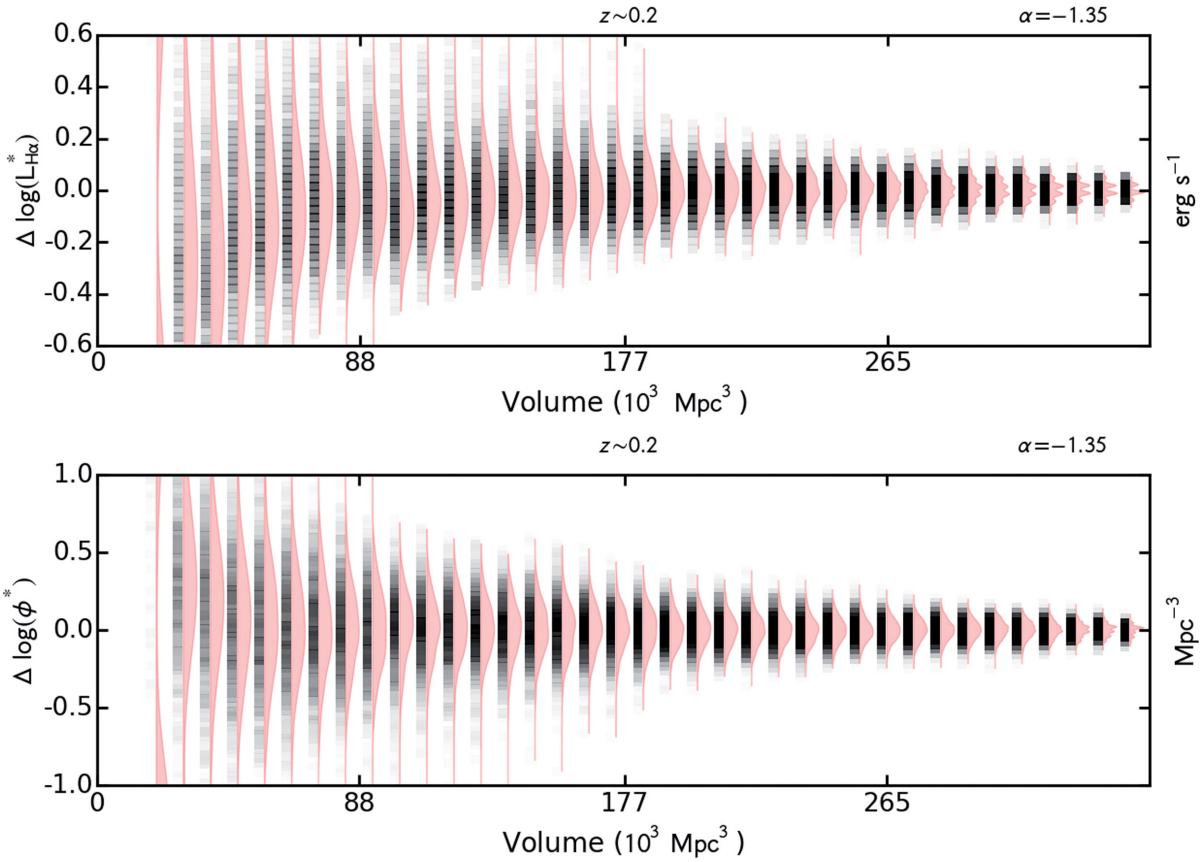

Figure 11. Error distribution of the characteristic Hα luminosity $L^*_{\rm H\alpha}$ and number density ϕ^*, as function of the volume probed. The error is calculated as fitted value minus the mean of the distribution at each volume. The results are obtained when combining data from both NB filters, with the faint-end slope fixed to -1.35 (see Figs B2, B3 and B4 for results for other α and for the two filters independently). At each volume, 1000 realizations are performed, based on random samples of sources. Each figure shows the values obtained from the LF fitting in grey-black stripe. Darker colours mean more of the realizations found at that particular $L^*_{\rm H\alpha}$ or ϕ^* value. The violin plot next to each stripe encodes the $L^*_{\rm H\alpha}/\phi^*$ histogram. The top panel shows the standard deviation σ of the $L^*_{\rm H\alpha}$ values at each volume size. Note that spread of values drops, the larger the volume probed, indicating a convergence in the values of $L^*_{\rm H\alpha}$ and ϕ^*.

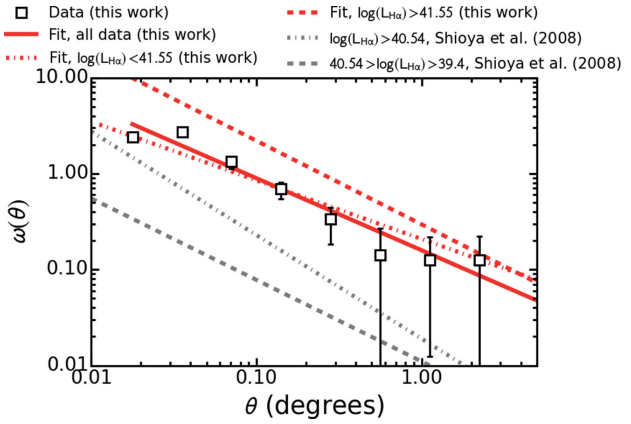

Figure 12. Angular two-point correlation function for bright Hα emitters ($L_{\rm H\alpha} \gtrsim 10^{41.0}$ erg s^{-1}) at $z \sim 0.2$. The best-fitting power-law relation is $\omega(\theta) = (0.109 \pm 0.005)\theta^{(-0.79 \pm 0.04)}$. For comparison, we plot the results for fainter emitters ($L_{\rm H\alpha} \lesssim 10^{41.5}$ erg s^{-1}) from Shioya et al. (2008). We find that more luminous Hα emitters are more clustered.

The correlation length also depends on redshift, but the evolution is driven by the typical luminosity: at high redshift, Hα emitters are, on average, brighter and have larger r_0 than lower-redshift sources.

Similar results are found by Hartley et al. (2010), who select galaxies using K-band luminosity as a proxy for stellar mass. The authors find that red galaxies, likely mostly ellipticals, are more clustered than the blue galaxies. Selecting star-forming galaxies based on colours, they find that r_0 drops with redshift. However, no dependence of r_0 on BB luminosity was found. By contrast, Bielby et al. (2014) use a mass selected sample and find that higher-mass galaxies tend to have larger clustering lengths. Additionally, they find that the clustering strength increases with stellar mass. Stellar mass correlates well with SFR (e.g. at $z \sim 0.2$; Stroe et al. 2015), which can then be translated to an equivalent Hα luminosity though equation (11). The results from Bielby et al. (2014) might indicate that more star-forming, more luminous galaxies have larger r_0, which is consistent with our findings. Note, however, that Sobral et al. (2010) controlled for both Hα luminosity and mass (K-band luminosity) and found that both are important for the evolution of r_0: r_0 increases with both higher $L_{\rm H\alpha}$ and K-band luminosity.

The clustering of the Hα emitters depends on the clustering of their host DM haloes. The bias parameter $b(z)$ describes how the matter distribution traces the DM distribution, as a function of redshift. In the bias model of Matarrese et al. (1997), the physical parameters of galaxies are determined by their host DM halo mass. In such a model, $b(z)$ depends on the minimum mass of the DM halo. Fig. 13 also contains r_0 predictions for DM haloes with fixed minimum mass of $M_{\rm min} = 10^{11-13}$ M$_\odot$, as calculated by Geach et al.

Table 6. Two-point correlation function for Hα emitters at $z \sim 0.2$, best fit as a single power law of the form $\omega(\theta) = A\theta^{\beta}$. Note that the filters and redshift distribution are different for Shioya et al. (2008) than for our study, so the amplitudes cannot be directly compared.

Source	$\log(L_{\rm H\alpha})$ (erg s^{-1})	A	β
This study	41.00–42.40	0.159 ± 0.012	-0.75 ± 0.05
Faint	41.00–41.55	0.208 ± 0.035	-0.61 ± 0.07
Bright	41.55–42.40	0.295 ± 0.026	-0.87 ± 0.06
Shioya et al. (2008)	40.54–41.50	0.019 ± 0.004	-1.08 ± 0.05
Shioya et al. (2008)	39.40–40.54	0.011 ± 0.002	-0.85 ± 0.05

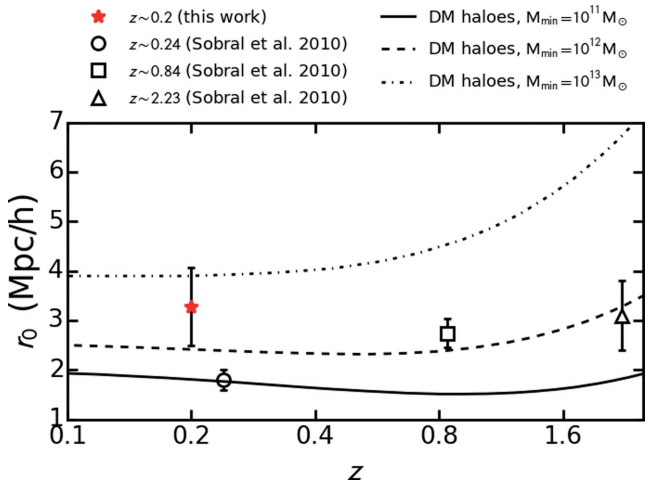

Figure 13. Dependence of the clustering length r_0 on redshift, using a consistent set of Hα emitters selected through NB surveys. For comparison, we also show data from Sobral et al. (2010). The plot suggests that typical ($L^*_{\rm H\alpha}$) emitters have very similar r_0 across cosmic time. At $z \sim 0.2$, there is a sharp increase in the typical DM halo mass with luminosity of the Hα sample. Note however, as shown in Fig. 14, that once corrected for the redshift evolution of the characteristic luminosity, $L_{\rm H\alpha}$ sets the position of galaxies in relation to the DM halo host.

(2008) assuming a ΛCDM cosmology and an evolving bias model from Matarrese et al. (1997) and Moscardini et al. (1998). Note, however, that the r_0 prediction is highly dependent on the model (see e.g. Hartley et al. 2010). Thus, we note that while the trends are valid, the normalization of $M_{\rm min}$ could be higher than that used here, leading to lower masses than derived here.

The emitters from Shioya et al. (2008), probing fainter Hα regimes with $L_{\rm H\alpha} < 10^{41.5}$ erg s^{-1}, reside in DM haloes of 10^{11} M$_\odot$ mass. These are most likely dwarf galaxies. By contrast, our faint sample is hosted by DM haloes of about $10^{12.5}$ M$_\odot$ mass, about the mass of the Milky Way. The bright Hα emitters are hosted by $\sim 10^{13-13.5}$ M$_\odot$ DM haloes, which are most probably already galaxy groups.

Fig. 14 shows how the DM halo minimum mass varies as a function of Hα luminosity and the luminosity scaled by the characteristic luminosity at that redshift [$L_{\rm H\alpha}/L^*_{\rm H\alpha}(z)$]. By comparing our results, with the results from Sobral et al. (2010) (based on data from Shioya et al. 2008), we find a linear correlation between the host minimum DM halo mass and luminosity (in log–log space; see Fig. 14). This indicates that more luminous, more star-forming galaxies reside in more massive DM haloes.

Accounting for the evolution of the characteristic luminosity with redshift, we find that more luminous emitters reside in more massive DM haloes, irrespective of redshift. Such a comparison between $z < 0.4$ and $z > 0.4$ samples has previously been difficult because of the different $L_{\rm H\alpha}/L^*_{\rm H\alpha}(z)$ ranges probed in the different redshift ranges. With our measurements, we probe beyond $L^*_{\rm H\alpha}$ at $z \sim 0.2$ for the first time, to be fully comparable with samples up to $z \sim 2.23$. Our measurements therefore confirm the results from Sobral et al. (2010) and Geach et al. (2012), who find that $L^*_{\rm H\alpha}$ galaxies reside in 10^{12-13} M$_\odot$, Milky Way size DM haloes, at all redshifts. The results indicate that the position of star-forming galaxies within the Hα LF is dictated by the host DM halo mass, at all cosmic times since $\sim$2.3.

6 CONCLUSIONS

In order to constrain the evolution of the star-forming galaxies across cosmic time, large samples of sources are necessary. Such samples are available at high redshifts ($z > 0.8$) through NB-selected Hα emitter samples, which probe large volumes ($>10^5$ Mpc^3) and overcome cosmic variance. However, at low redshifts ($z < 0.8$), large areas (>15 deg^2) need to be surveyed in order to match the volumes at high redshift. By carrying out the largest survey of Hα emitters at $z \sim 0.2$, we produce a LF describing typical galaxies within representative volumes of the Universe. With our large sample of bright emitters, we study their distribution and clustering, and place it in the context of the evolution of the SFRD throughout cosmic history. Our main results are as follows.

(i) The Hα LF at $z \sim 0.2$ is well described by a Schechter function with $\log(\phi^*) = -2.85 \pm 0.03$ (Mpc^{-3}) and $\log(L^*_{\rm H\alpha}) = 41.71 \pm 0.02$ (erg s^{-1}). We find that previous studies, probing far smaller volumes, underestimate the characteristic luminosity $L^*_{\rm H\alpha}$, but are reconciled with our results if cosmic variance uncertainties are taken into account. For volumes typically probed in previous Hα works at $z \sim 0.2$ of $<5 \times 10^4$ Mpc^3, cosmic variance can account to more than 50 per cent variance in the LF parameters.

(ii) By assuming a 15 per cent AGN fraction, we derive a SFRD of $\rho_{\rm SFRD} = 0.0094 \pm 0.0008$ M$_\odot$ yr^{-1} Mpc^{-1}.

(iii) We find significant cosmic variance in the distribution of the Hα emitters, but on average 1–4 bright ($L_{\rm H\alpha} > 10^{41.1}$ erg s^{-1}) Hα emitters are found per deg^2.

(iv) We study the clustering of the Hα emitters. The two-point correlation function is well fit by a single power law $\omega(\theta) = (0.159 \pm 0.012)\theta^{(-0.75 \pm 0.05)}$, with a spatial clustering length $r_0 = 5.0 \pm 1.1$ Mpc h^{-1} for the bright sample ($10^{41.0-41.55}$ erg s^{-1}) and $r_0 = 3.5 \pm 1.1$ Mpc h^{-1} for the faint sample ($10^{41.55-42.40}$ erg s^{-1}). Our results confirm that luminous, strongly star-forming galaxies are more clustered than those that are weakly star-forming.

(v) We find that, at $z \sim 0.2$, the higher the SFR, the more massive the DM halo host is. When accounting for the redshift dependence

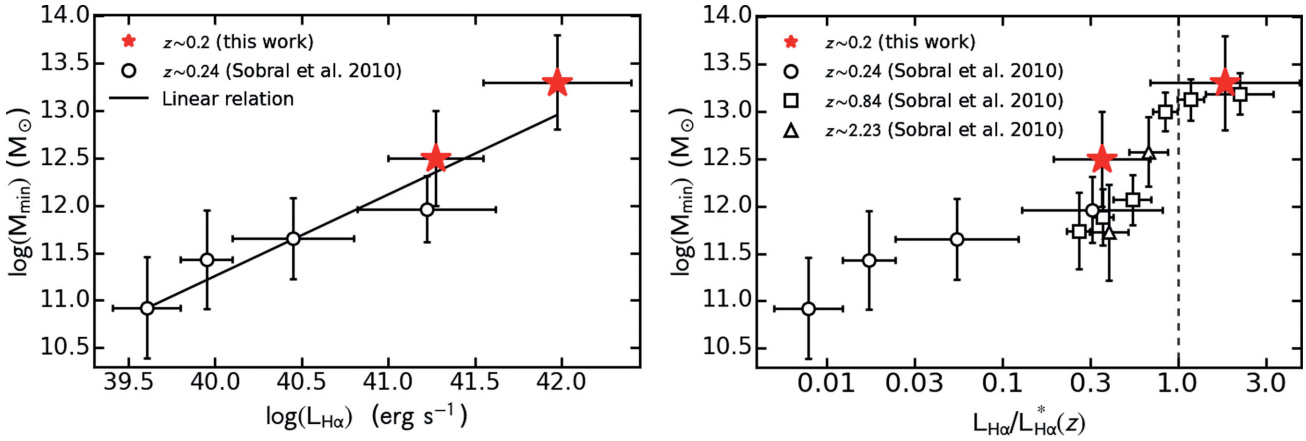

Figure 14. Minimum DM halo mass ($M_{\rm DM}$) as a function of luminosity ($L_{\rm H\alpha}$, left) and luminosity scaled by the characteristic luminosity at the respective redshift ($L_{\rm H\alpha}/L^*_{\rm H\alpha}(z)$, right). The data from Sobral et al. (2010), split per luminosity bin, are shown for comparison. The ∼0.2 points are renormalized using the $L^*_{\rm H\alpha}$ derived in this paper. All luminosities are not corrected for intrinsic dust extinction. Note the relation between the Hα luminosity and host mass. When scaled for the typical luminosity, a clear relation between DM halo mass and luminosity is observed from $z \sim 2.23$ to $z \sim 0.2$.

of the characteristic Hα luminosity, there is no redshift dependence of the host mass, but a strong dependence on $L_{\rm H\alpha}/L^*_{\rm H\alpha}(z)$.

ACKNOWLEDGEMENTS

We thank the referee for comments that have improved the clarity and interpretation of our results. This work is based on observations made with the Isaac Newton Telescope (proposal I13BN008) operated on the island of La Palma by the Isaac Newton Group in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. It is also based on observations obtained with MegaPrime/MegaCam, a joint project of Canada–France–Hawaii Telescope (CFHT) and CEA/IRFU, at the CFHT, which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France and the University of Hawaii. This work is based in part on data products produced at Terapix available at the Canadian Astronomy Data Centre as part of the CFHTLS, a collaborative project of NRC and CNRS. Based on observations obtained as part of the VISTA Hemisphere Survey, ESO Progamme, 179.A-2010 (PI: McMahon). Based on data products from observations made with ESO telescopes at the La Silla or Paranal Observatories under ESO programme ID 179.A-2006. The UKIDSS project is defined in Lawrence et al. (2007). UKIDSS uses the UKIRT Wide Field Camera (WFCAM; Casali et al. 2007). The photometric system is described in Hewett et al. (2006), and the calibration is described in Hodgkin et al. (2009). The pipeline processing and science archive are described in Irwin et al. (in preparation) and Hambly et al. (2008). This research has made use of the NASA/IPAC Extragalactic Data base (NED), which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration (NASA). This research has made use of NASA's Astrophysics Data System. AS acknowledges financial support from a Netherlands Organization for Scientific Research (NWO) top subsidy (614.001.006). DS acknowledges financial support from the NWO through a Veni fellowship, from FCT through a FCT Investigator Starting Grant and Start-up Grant (IF/01154/2012/CP0189/CT0010) and from FCT grant PEst-OE/FIS/UI2751/2014.

REFERENCES

Ahn C. P. et al., 2012, ApJS, 203, 21
Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Solano E., eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac., San Francisco, p. 112
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, ASP Conf. Ser. Vol. 281, Astronomical Data Analysis Software and Systems XI. Astron. Soc. Pac., San Francisco, p. 228
Best P. et al., 2013, in Adamson A., Davies J., Robson I., eds, Astrophysics and Space Science Proc. Vol. 37, Thirty Years of Astronomical Discovery with UKIRT. Springer-Verlag, Berlin, p. 235 (arXiv:1003.5183)
Bielby R. M. et al., 2014, A&A, 568, A24
Bland-Hawthorn J., van Breugel W., Gillingham P. R., Baldry I. K., Jones D. H., 2001, ApJ, 563, 611
Bouwens R. J. et al., 2011, ApJ, 737, 90
Bouwens R. J. et al., 2015, ApJ, 803, 34
Bunker A. J., Warren S. J., Hewett P. C., Clements D. L., 1995, MNRAS, 273, 513
Casali M. et al., 2007, A&A, 467, 777
Chabrier G., 2003, PASP, 115, 763
Dale D. A. et al., 2010, ApJ, 712, L189
Drake A. B. et al., 2013, MNRAS, 433, 796
Erben T. et al., 2013, MNRAS, 433, 2545
Garcet O. et al., 2007, A&A, 474, 473
Garn T., Best P. N., 2010, MNRAS, 409, 421
Geach J. E., Smail I., Best P. N., Kurk J., Casali M., Ivison R. J., Coppin K., 2008, MNRAS, 388, 1473
Geach J. E., Sobral D., Hickox R. C., Wake D. A., Smail I., Best P. N., Baugh C. M., Stott J. P., 2012, MNRAS, 426, 679
Gunawardhana M. L. P. et al., 2013, MNRAS, 433, 2764
Gwyn S. D. J., 2012, AJ, 143, 38
Hambly N. C. et al., 2008, MNRAS, 384, 637
Hartley W. G. et al., 2010, MNRAS, 407, 1212
Hewett P. C., Warren S. J., Leggett S. K., Hodgkin S. T., 2006, MNRAS, 367, 454
Hodgkin S. T., Irwin M. J., Hewett P. C., Warren S. J., 2009, MNRAS, 394, 675
Ilbert O. et al., 2006, A&A, 457, 841
Jarvis M. J. et al., 2013, MNRAS, 428, 1281
Karim A. et al., 2011, ApJ, 730, 61
Kennicutt R. C., Jr, 1998, ARA&A, 36, 189
Kirkpatrick J. D., Henry T. J., McCarthy D. W. Jr, 1991, ApJS, 77, 417

Kirkpatrick J. D. et al., 1999, ApJ, 519, 802
Kurk J. D., Pentericci L., Röttgering H. J. A., Miley G. K., 2004, A&A, 428, 793
Landy S. D., Szalay A. S., 1993, ApJ, 412, 64
Lawrence A. et al., 2007, MNRAS, 379, 1599
Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJ, 460, L1
Ly C. et al., 2007, ApJ, 657, 738
Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS, 286, 115
Matthee J. J. A. et al., 2014, MNRAS, 440, 2375
Melnyk O. et al., 2013, A&A, 557, A81
Moorwood A. F. M., van der Werf P. P., Cuby J. G., Oliva E., 2000, A&A, 362, 9
Moscardini L., Coles P., Lucchin F., Matarrese S., 1998, MNRAS, 299, 95
Norberg P. et al., 2001, MNRAS, 328, 64
Oteo I., Sobral D., Ivison R. J., Smail I., Best P. N., Cepa J., Pérez-García A. M., 2015, MNRAS, in press (arXiv:1506.02670)
Ouchi M. et al., 2005, ApJ, 620, L1
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton University Press, Princeton, NJ
Pierre M. et al., 2004, JCAP, 9, 011
Polletta M. et al., 2007, ApJ, 663, 81
Schechter P., 1976, ApJ, 203, 297
Shioya Y. et al., 2008, ApJS, 175, 128
Sobral D. et al., 2009, MNRAS, 398, 75
Sobral D., Best P. N., Geach J. E., Smail I., Cirasuolo M., Garn T., Dalton G. B., Kurk J., 2010, MNRAS, 404, 1551
Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, MNRAS, 420, 1926
Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, MNRAS, 428, 1128
Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 437, 3516
Sobral D., Stroe A., Dawson W. A., Wittman D., Jee M. J., Röttgering H., van Weeren R. J., Brüggen M., 2015a, MNRAS, 450, 630
Sobral D. et al., 2015b, MNRAS, 451, 2303
Stott J. P., Sobral D., Smail I., Bower R., Best P. N., Geach J. E., 2013, MNRAS, 430, 1158
Stroe A., Sobral D., Röttgering H. J. A., van Weeren R. J., 2014, MNRAS, 438, 1377
Stroe A. et al., 2015, MNRAS, 450, 646
Tajer M. et al., 2007, A&A, 467, 73
Warren S. J. et al., 2007, preprint (arXiv:astro-ph/0703037)
Zacharias N., Finch C. T., Girard T. M., Henden A., Bartlett J. L., Monet D. G., Zacharias M. I., 2013, AJ, 145, 44

APPENDIX A: SURVEY COMPLETENESS

The method for studying the completeness is detailed in Section 4.1. The dependence of the completeness on line flux is shown in Fig. A1.

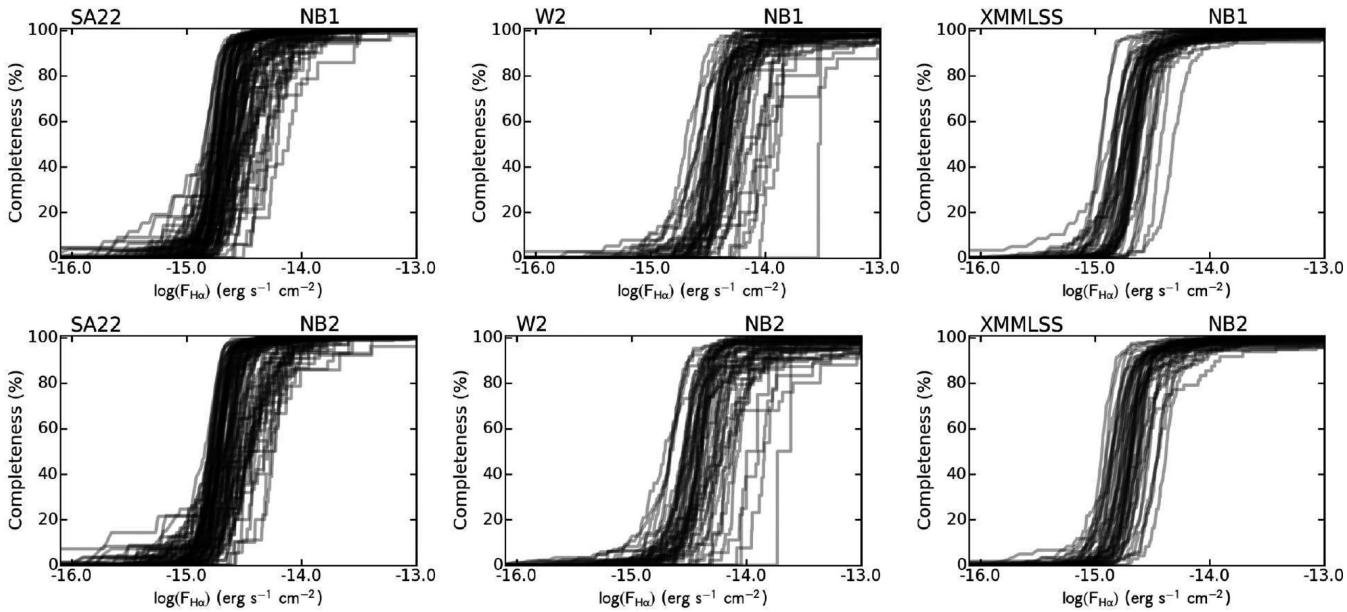

Figure A1. Survey completeness as a function of Hα flux, plotted separately for each field and NB filter used to select Hα candidates. Each curve is associated with the the completeness study for a different CCD chip within each pointing. The darker the colour, the more completeness curves fall within that region. Note the XMMLSS field is significantly more complete than the W2 field.

APPENDIX B: SURVEY COMPLETENESS

The results of the resampling of the LF at $z \sim 0.2$ with different binnings is presented for a range of data selections. The faint-end slope is fixed at -1.35 and -1.7 and ϕ and L are fit using data selected from the two NB filters independently and combined. The results are shown in Figs B1–B4.

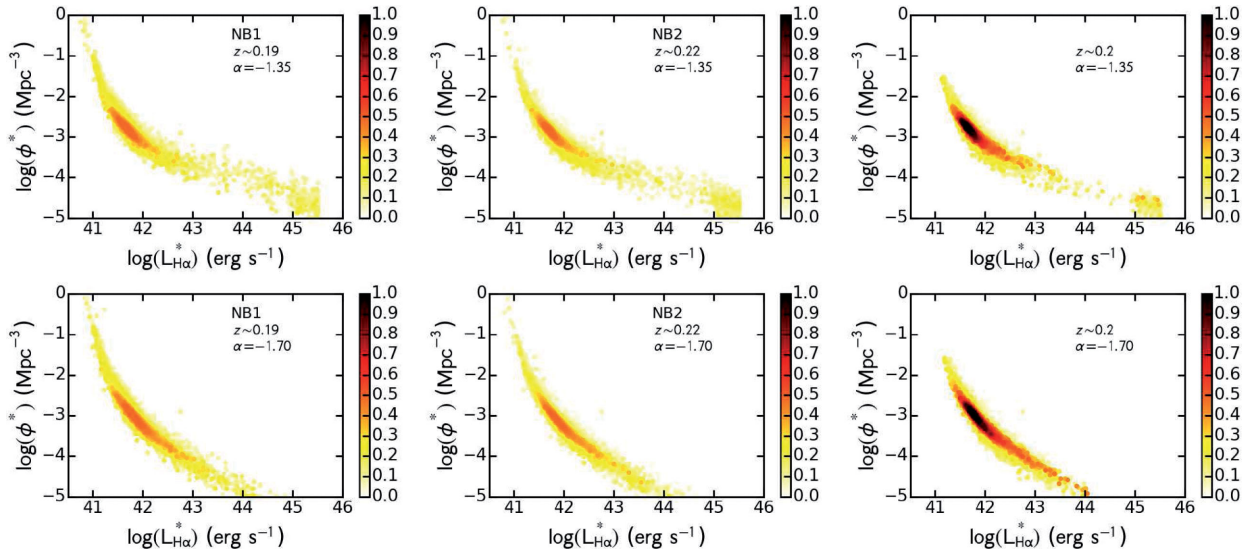

Figure B1. As for Fig. 9, but with different values of α and when using the data for the two filters separately or together.

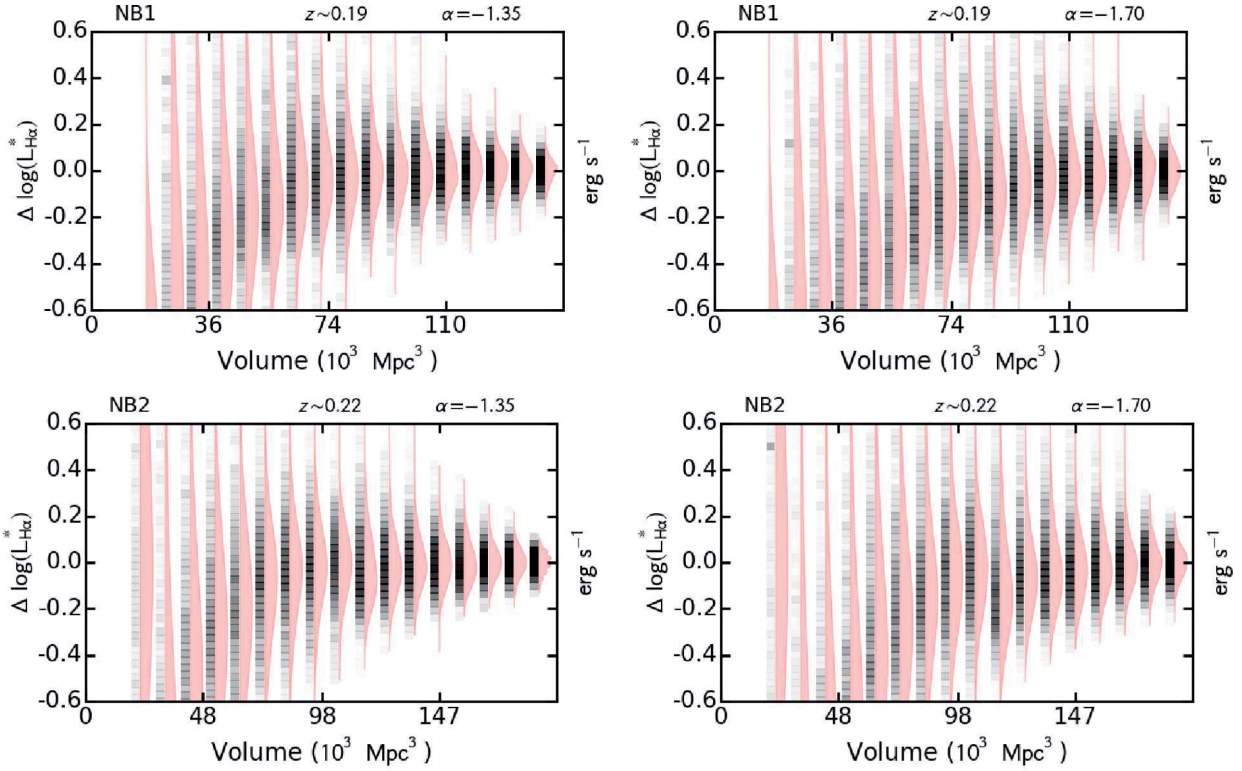

Figure B2. As Fig. 11, but for data samples from the two NB filters independently. Note that similar results are found for the two filters, even when considered separately.

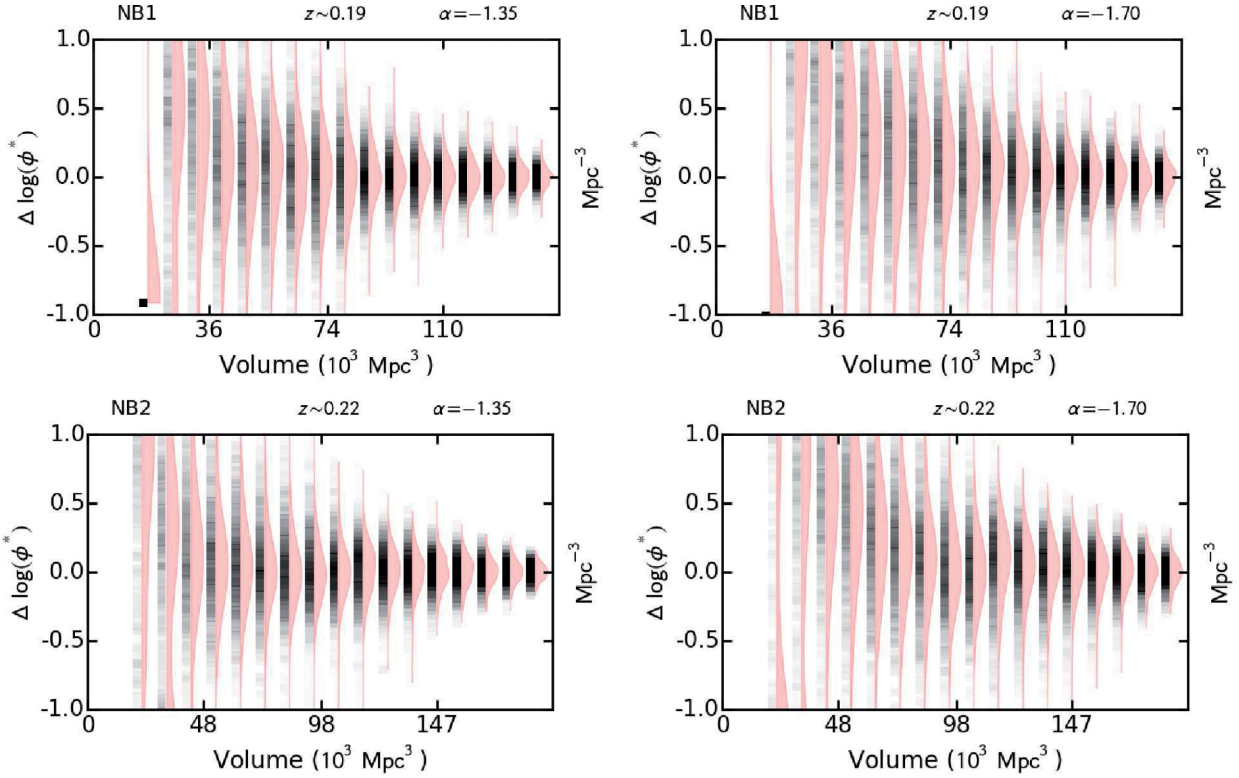

Figure B3. As Fig. 11, but for data samples from the two NB filters independently. Note that similar results are found for the two filters, even when considered separately.

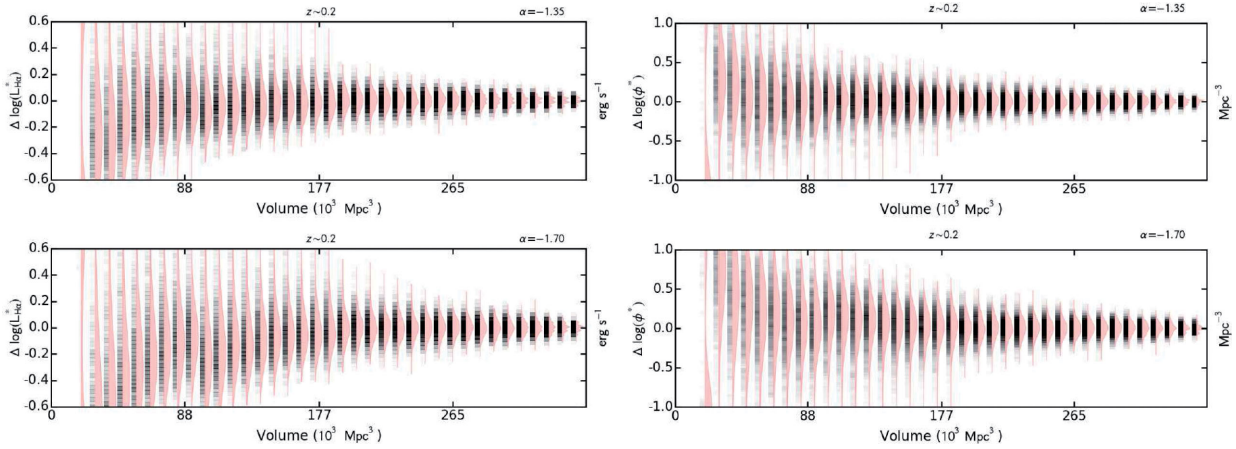

Figure B4. As Fig. 11, but for different α values.

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