Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY
MNRAS **451,** 2303–2323 (2015) doi:10.1093/mnras/stv1076

CF-HiZELS, an ∼10 deg^2 emission-line survey with spectroscopic follow-up: Hα, [O III] + Hβ and [O II] luminosity functions at $z = 0.8$, 1.4 and 2.2

D. Sobral,[1,2,3★†] J. Matthee,[3†] P. N. Best,[4] I. Smail,[5] A. A. Khostovan,[6] B. Milvang-Jensen,[7] J.-W. Kim,[8] J. Stott,[5] J. Calhau,[1,2] H. Nayyeri[6,9] and B. Mobasher[6]

[1]*Instituto de Astrofísica e Ciências do Espaço, Universidade de Lisboa, OAL, Tapada da Ajuda, P-1349-018 Lisboa, Portugal*
[2]*Departamento de Física, Faculdade de Ciências, Universidade de Lisboa, Edifício C8, Campo Grande, P-1749-016 Lisbon, Portugal*
[3]*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[4]*SUPA, Institute for Astronomy, Royal Observatory of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, UK*
[5]*Institute of Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK*
[6]*Department of Physics and Astronomy, University of California, 900 University Ave, Riverside, CA 92521, USA*
[7]*Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen Ø, Denmark*
[8]*Center for the Exploration of the Origin of the Universe, Department of Physics and Astronomy, Seoul National University, Seoul 151-742, Korea*
[9]*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*

Accepted 2015 May 11. Received 2015 May 1; in original form 2015 February 21

ABSTRACT

We present results from the largest contiguous narrow-band survey in the near-infrared. We have used the wide-field infrared camera/Canada–France–Hawaii Telescope and the lowOH2 filter (1.187 ± 0.005 μm) to survey ≈10 deg^2 of contiguous extragalactic sky in the SA22 field. A total of ∼6000 candidate emission-line galaxies are found. We use deep *ugrizJK* data to obtain robust photometric redshifts. We combine our data with the High-redshift(Z) Emission Line Survey (HiZELS), explore spectroscopic surveys (VVDS, VIPERS) and obtain our own spectroscopic follow-up with KMOS, FMOS and MOSFIRE to derive large samples of high-redshift emission-line selected galaxies: 3471 Hα emitters at $z = 0.8$, 1343 [O III] + Hβ emitters at $z = 1.4$ and 572 [O II] emitters at $z = 2.2$. We probe comoving volumes of $>10^6$ Mpc^3 and find significant overdensities, including an 8.5σ (spectroscopically confirmed) overdensity of Hα emitters at $z = 0.81$. We derive Hα, [O III] + Hβ and [O II] luminosity functions at $z = 0.8$, 1.4, 2.2, respectively, and present implications for future surveys such as *Euclid*. Our uniquely large volumes/areas allow us to subdivide the samples in thousands of randomized combinations of areas and provide a robust empirical measurement of sample/cosmic variance. We show that surveys for star-forming/emission-line galaxies at a depth similar to ours can only overcome cosmic-variance (errors <10 per cent) if they are based on volumes $>5 \times 10^5$ Mpc^3; errors on L^* and ϕ^* due to sample (cosmic) variance on surveys probing ∼10^4 and ∼10^5 Mpc^3 are typically very high: ∼300 and ∼40–60 per cent, respectively.

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

1 INTRODUCTION

Determining and understanding the star formation history of the Universe is of fundamental importance to improve our understanding of galaxy formation and evolution. Surveys measuring the star formation rate (SFR) density ($\rho_{\rm SFR}$) as a function of redshift/cosmic time suggest that $\rho_{\rm SFR}$ rises up to $z \sim 1$–2 (e.g. Lilly et al. 1996; Hopkins & Beacom 2006; Karim et al. 2011; Sobral et al. 2013a) and thus reveal that the 'epoch' of galaxy formation occurs at $z > 1$. Most interesting, recent studies are also showing that such behaviour happens for star-forming galaxies at all masses (e.g. Karim et al. 2011; Sobral et al. 2014).

★FCT-IF/Veni Fellow.
†E-mail: sobral@strw.leidenuniv.nl (DS); matthee@strw.leidenuniv.nl (JM)

There are several star formation indicators that one can use to identify and study star-forming galaxies in cosmological volumes. The most direct tracer of recent star formation is the far-ultraviolet (UV) light, coming directly from very massive, short-lived stars. Recombination lines resulting from the strong ionizing radiation coming from such stars are also excellent tracers of recent star formation, in particularly Hα, but forbidden lines such as [O II]λ3727 can also be used. Studies have also used the [O III]λ5007 and other emission lines for this purpose (e.g. Ly et al. 2007; Drake et al. 2013), although with many caveats such as a potential large AGN contamination. Alternative methods for selecting star-forming galaxies are surveying in the far-infrared (FIR) to detect the blackbody re-radiation of dust-absorbed UV light from young, massive stars, or surveying in the radio for emission of supernova remnants. There is of course a significant difference between using a star formation indicator to both identify and study star-forming galaxies (which can identify and study star-forming selected galaxies down to some limit), and using such indicators to just measure SFRs from samples selected by some other means, with more complicated and potentially biased selection functions.

Hα stands out as a sensitive star formation indicator, very well calibrated and not strongly affected by dust extinction in typical star-forming galaxies, unlike the UV or bluer emission lines (typically $A_{\rm H\alpha} = 1$ mag; e.g. Garn et al. 2010; Sobral et al. 2012; Ibar et al. 2013; Stott et al. 2013a). It is also much more sensitive than even the most sensitive FIR or radio surveys, and, with the depths that current instrumentation now allow, Hα surveys are able to identify both relatively dust-free and dusty star-forming galaxies, and thus are ideal for an approximately complete selection of star-forming galaxies (see e.g. Oteo et al. 2015). When combined with measurements in the FIR, Hα becomes an even better SFR selector and indicator (e.g. Kennicutt et al. 2009; Garn et al. 2010; Ibar et al. 2013). Another potential advantage of Hα is the possibility to perform surveys with the narrow-band (NB) technique, taking advantage of wide-field optical and near-infrared cameras. NB surveys can probe large areas and, when combined with another NB filter at a close-by wavelength, or a broad-band (BB) filter to estimate the continuum, they allow for an effective way to obtain clean, complete samples of emission-line selected galaxies (e.g. Bunker et al. 1995; Ly et al. 2007, 2011; Geach et al. 2008; Koyama et al. 2011, 2013; Tadaki et al. 2011; Lee et al. 2012; Sobral et al. 2012, 2013a; An et al. 2014; Stroe et al. 2014, 2015).

The High-redshift(Z) Emission Line Survey (HiZELS; Geach et al. 2008; Sobral et al. 2009b, 2012, 2013a) has exploited various NB filters in the z, J, H and K bands to undertake deep, wide surveys for line emitters, with a particular strong emphasis on exploring Hα emitters across redshift (e.g. Sobral et al. 2010, 2014; Geach et al. 2012; Swinbank et al. 2012a,b; Stott et al. 2013a,b). HiZELS results in a homogeneous selection of Hα emitters at $z = 0.40$, 0.84, 1.47 and 2.23, over a few $\sim$1 deg^2 areas, detecting around 500–1000 emitters at each redshift down to limiting Hα SFRs of $\sim$3–10 M$_\odot$ yr^{-1} (cf. Sobral et al. 2013a). HiZELS allowed for the first fully self-consistent measurement of the evolution of the Hα luminosity function (LF) from $z = 0$ to $\sim$2, revealing that $\rho_{\rm SFR}$ rises up to $z \sim 2$, using one single star formation indicator (Sobral et al. 2013a). The results also show that the Hα star formation history can fully reproduce the evolution of the stellar mass density since $z = 2.23$ and that the typical/characteristic SFR of star-forming galaxies (SFR$^*(z)$) has decreased by a factor $\sim$13 since that time (Sobral et al. 2014).

Results from NB and slitless grism surveys are generally in good agreement, but some disagreement by factors of a few have been reported (Ly et al. 2007, 2011; Lee et al. 2012; Sobral et al. 2012; Colbert et al. 2013; Drake et al. 2013; Stroe et al. 2014). It is expected that such discrepancies are mostly caused by sample variance (cosmic variance) due to the relatively small areas studied. By probing large enough areas, one can overcome the effect of cosmic variance (e.g. Sobral et al. 2010), and such data can be used to empirically measure the effects of sample variance, instead of having to rely on rather uncertain theoretical/model predictions. Minimizing and understanding these discrepancies is not only important for our understanding of galaxy formation and evolution, but also as a forecasting tool for upcoming surveys, such as *Euclid* and *WFIRST* (e.g. Geach et al. 2010; Colbert et al. 2013; Wang, Chuang & Hirata 2013). In particular, determining the bright end of the LF with sufficient accuracy, and minimizing cosmic variance, can have a significant impact on estimates of the number of emitters to be recovered with very wide space surveys with e.g. *Euclid*.

NB surveys are also sensitive to many other emission lines besides Hα that can be very useful to (potentially) extend measurements to higher redshift, but also to provide alternative/comparative views. These can be used to search for Ly α (e.g. Sobral et al. 2009a; Matthee et al. 2014), but also to trace the evolution of the LF of e.g. oxygen lines, such as [O II] (e.g. Ly et al. 2007; Sobral et al. 2012). So far studies have mostly just covered relatively small areas (e.g. Ly et al. 2007; Drake et al. 2013), probing down to low luminosities, but at the expense of being strongly affected by cosmic variance, particularly when based on a single field. Significant progress requires probing volumes $\sim$10 $\times$ larger.

In order to address the current shortcomings/limitations, we have undertaken by far the largest NB survey for high-redshift emission-line galaxies. Observations were made using the LowOH2 NB filter on the Canada–France–Hawaii Telescope (CFHT)/Wide-field InfraRed CAMera (WIRCam) over an $\sim$10 deg^2 area in the SA22 field and we also combine our new data with similar data in the Cosmic Evolution Survey (COSMOS) and the Ultra Deep Survey (UDS) from Sobral et al. (2013a). Our study provides a major increase in the sample sizes and statistics, but also probes a new part of the parameter space. The surveyed area corresponds to an area coverage which is $\sim$5–10 times larger than the largest previous emission-line surveys at $z \sim 1$ (e.g. Sobral et al. 2009b, 2013a; Ly et al. 2011). We also undertake a significant spectroscopic follow-up to improve the robustness of our results. Finally, we present the first [O II] LF at $z = 2.2$.

The paper is organized in the following way. Section 2 outlines the details of the observations, and describes the data reduction, photometric calibration, source extraction and survey limits. Section 3 presents the NB selection criteria, the final sample of emitters and the photometric and spectroscopic redshift analysis, including the follow-up observations with MOSFIRE and FMOS. Section 3 also presents the selection of different line emitters within the full sample, by using colour–colour selections and photometric redshifts. Section 4 presents the methods and corrections applied to the data in order to derive LFs. Results are presented in Section 5; the Hα, Hβ + [O III], and [O II] LFs, their evolution to $z \sim 2$, implications for *Euclid*, and a quantification of sample (cosmic) variance over the full 10 deg^2, which includes the HiZELS data. Finally, Section 6 outlines the conclusions. A $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_{\rm M} = 0.3$ and $\Omega_\Lambda = 0.7$ cosmology is used. We use a Salpeter initial mass function and all magnitudes are in the AB system, unless noted otherwise.

Table 1. Observation log for the NB observations conducted with the lowOH2 filter on CFHT/WIRCam. A total of 80 pointings, numbered from 0 to 79 (Field ID), were obtained with WIRCam, to cover a total area of approximately 10 deg². The seeing in all observations was in the range 0.5–0.7 arcsec and conditions were photometric.

Field ID	RA (J2000)	Dec. (J2000)	Int. time (ks pixel^{-1})	Dates of observations	3σ limit (AB, 2 arcsec)
0–28	22 18 00 to 22 22 00	−00 04 00 to +00 06 24	1.0	20–30 Sept, 1–18 Oct, 6 Dec 2011, 4 Oct–3 Nov 2012	22.7
29–53	22 14 00 to 22 18 00	−00 04 00 to +00 06 24	1.0	4–31 Oct, 1–3 Nov 2012	22.7
54–79	22 10 00 to 22 14 00	−00 04 00 to +00 06 24	1.0	20 Sept–6 Dec 2011, 4 Oct–3 Nov 2012	22.7

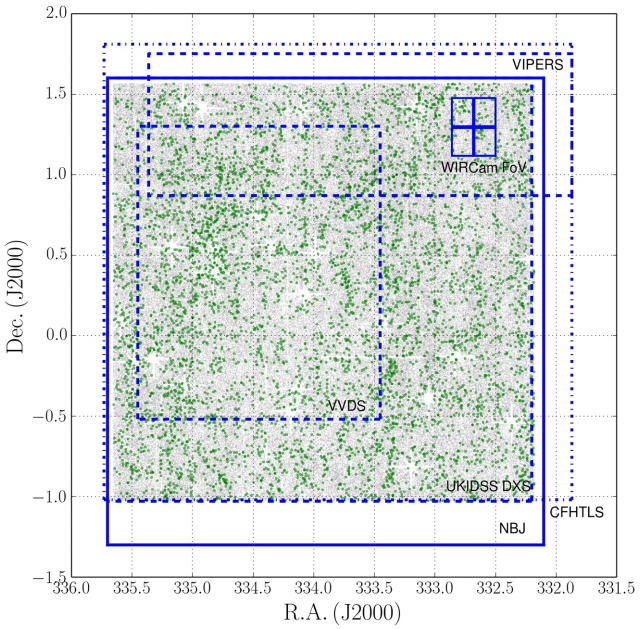

Figure 1. Surveyed area in the SA22 with the LowOH2 NB filter (NB$_J$) and comparison with other surveys and with the individual WIRCam footprint/field of view. For Hα emitters at $z = 0.81$, the surveyed area roughly corresponds to a box with ∼60 × 95 Mpc (physical). The figure also shows (in grey) all NB detections matched to J and in green the location of all line emitters, irrespectively of redshift (see Section 3). The overlapping regions with CFHTLS W4 (*ugriz*), UKIDSS DXS (*JK*), VVDS and VIPERS (z_{spect}) are also shown.

2 OBSERVATIONS AND DATA REDUCTION

2.1 Observations

Observations were made with the CFHT using the WIRCam (Puget et al. 2004) during three semesters in queued service mode (Program IDs 2011B/029, 2012A019 and 2012B/016).[1] Data were obtained using the LowOH2 NB filter (central wavelength of 1187 nm[2] and a full width at half-maximum, $\Delta\lambda = 10$ nm, also referred in this paper as NB$_J$), and cover a contiguous region of ∼10 deg² (∼3 × 3 deg²) in the SA22 field (centred at 22:16:00, +00:18:00; see Table 1). The SA22 field was chosen for this very wide NB survey as it presents excellent ancillary data over ∼9.2 deg² area (see Fig. 1 for overlap with other surveys).

Observations were conducted from 2011 September 20 to 2011 December 6, and were concluded between 2012 October 4 and 2012 November 3. The seeing over the entire data set is excellent, in the range 0.5–0.7 arcsec and observations were done in clear conditions. WIRCam's standard 'paw-print' configuration of four 2048 × 2048 cryogenically cooled HgCdTe arrays (0.3045 arcsec pixel^{-1}), offset by ∼1 arcmin from each other, probe a region of about 0.11 deg² (21 × 21 arcmin field of view at prime focus, see Fig. 1) at any given time. In order to cover the full SA22 field, we obtained 80 different pointings. Each pointing was obtained using a dither pattern which resulted in observations being obtained over 10 different positions (individual exposures of 100 s) per pointing to fill in the detector gaps and minimize the effect of bad pixels and cosmic rays. The final exposure time per pixel is 1.0 ks. A summary of the observations is given in Table 1.

2.2 Data reduction

The NB data were reduced with a dedicated pipeline using PYTHON and its PYFITS and IMAGE modules, based on PfHiZELS (Sobral et al. 2009b, 2013a). Briefly, we start by median combining the dark frames to produce master darks and then use them to dark subtract the individual science frames. We obtain first-pass flat-fields by median combining jittered science frames, and use those to flatten the data. We then run SExtractor (Bertin & Arnouts 1996) on the first-pass flattened frames to produce individual masks. We use those to mask out all individual sources, and, excluding each frame that is being flattened, we produce a final flat-field for that frame and flatten the frame. We then use SCAMP[3] (Bertin 2006) to fit a World Coordinate System by matching sources detected in individual reduced science images to those in 2MASS (Skrutskie et al. 2006) catalogue. We also use SCAMP to correct for distortions across the field of view by fitting a third-order polynomial. Frames are also normalized individually to the same zero-point (ZP) by computing the ratio between the expected flux/magnitude from 2MASS and that found in our data (see more details in Section 2.3). For both steps, we are able to use on average ∼75 stars per individual frame, resulting in astrometric solutions with typical rms of ∼0.15 arcsec. Finally, the individual reduced frames from the four detectors are median combined using SWARP[4] (Bertin 2010), to obtain stacked reduced data for the entire field. For the BB J (and K) comparison data, we use UKIDSS-DXS-DR10[5] (Lawrence et al. 2007).

2.3 Source extraction and survey limits

We obtain the magnitude ZP by comparing the magnitudes of the sources in the 2MASS catalogue to those in our data set, excluding

[1] FP7/2007–2013: the research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007–2013) under grant agreement number RG226604 (OPTICON).
[2] We fully confirm the central wavelength by exploring all spectroscopic redshifts in our study.

[3] http://www.astromatic.net/software/scamp
[4] http://www.astromatic.net/software/swarp
[5] http://surveys.roe.ac.uk/wsa/

the faintest ($J > 17$, low S/N in 2MASS) and the brightest ($J < 12$, saturated in our data) sources. In order to simplify the analysis, and once accurate ZPs are determined for each stacked image, we scale all images such that ZPs are set to 25, including the BB images (J, UKIDSS). Sources were extracted using SExtractor (Bertin & Arnouts 1996) on both NB and BB J images, using 2 arcsec apertures. The 3σ AB-magnitude limit for the survey is 22.7 (5σ: 22.15; cf. Matthee et al. 2014), corresponding to an emission-line flux limit of 8×10^{-17} erg s^{-1} cm^{-2}. This limit is computed by measuring the average background rms of the NB images in one million empty 2 arcsec diameter apertures and for point-like sources (no aperture correction is applied as it is very small given the excellent seeing, assuming sources are point sources); this is the aperture we use throughout the paper for all measurements and corresponds to about ∼16 kpc physical diameter for the three lines investigated here. We note that because we use random aperture measurements, the rms that we measure already accounts for correlations in the noise.

3 SELECTION OF EMITTERS

Down to the 3σ limit, we detect 346 244 sources in our NB images. Once catalogues with sources in the NB and in the BB are made, they are matched using a sky algorithm with a maximum separation of 1 arcsec. NB sources with no matching BB source are likely to be spurious, but they are kept in the catalogue and assigned a J upper limit.

3.1 Emission-line candidates

In order to robustly select sources that show a real colour excess of the NB over the BB, instead of just random scatter or uncertainty in the measurements, two criteria are used. First, the parameter Σ (Bunker et al. 1995) is used to quantify the real excess compared to an excess due to random scatter. This means that the difference between counts in the NB and the BB must be higher than the total error times Σ:

$$c_{\mathrm{NB}_J} - c_{\mathrm{BB}_J} > \Sigma\delta. \quad (1)$$

Here, c_{NB_J} and c_{BB_J} are the counts in the NB and BB, respectively, while δ is the total photometric error, the combination of the errors in both bands:

$$\delta = \sqrt{\pi r_{\mathrm{ap}}^2 \left(\sigma_{\mathrm{NB}_J}^2 + \sigma_{\mathrm{BB}_J}^2\right)}, \quad (2)$$

where r_{ap} is the radius of the apertures in pixels and σ the rms per pixel in each band. These two equations can be combined into the following equation for Σ (Sobral et al. 2013a),

$$\Sigma = \frac{1 - 10^{-0.4(\mathrm{BB}-\mathrm{NB}_J)}}{10^{-0.4(\mathrm{ZP}-\mathrm{NB}_J)}\sqrt{\pi r_{\mathrm{ap}}^2 \left(\sigma_{\mathrm{NB}_J}^2 + \sigma_{\mathrm{BB}_J}^2\right)}}. \quad (3)$$

ZP is the zero-point of the NB_J (which is the same as the BB_J, because they are both scaled to ZP = 25 in our analysis). We note that the central wavelength of the NB is not perfectly in the centre of the BB (J), but rather at the blue end of the filter, similarly to other NB filters in the J band (see e.g. Sobral et al. 2013a). Here, we correct for this effect using CFHTLS[6] z-band, which is the closest band on the blue side of J. Our colour correction (empirical) is given by

$$BB_J - NB_J = (BB_J - \mathrm{NB}_J)_0 + 0.04(z' - BB_J) - 0.05. \quad (4)$$

[6] Canada–France–Hawaii Telescope Legacy Survey.

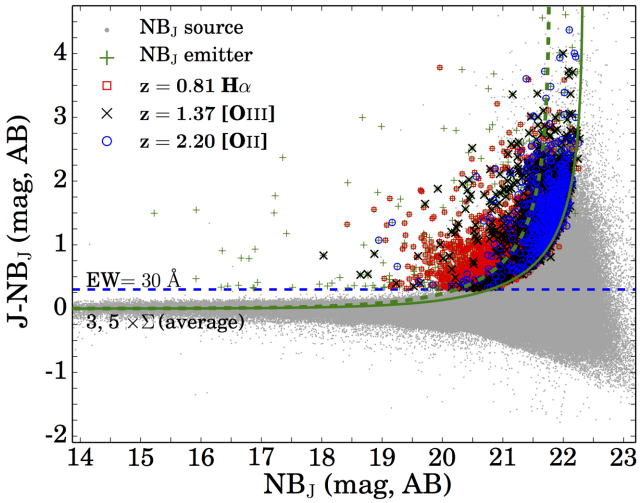

Figure 2. Colour–magnitude diagram for our NB_J detected sources. The dotted horizontal line is for an observed EW of 30 Å, which corresponds to J-$\mathrm{NB}_J = 0.3$. The 3σ and 5σ selection curves are shown for the average depth of the survey for reference (but note that some small variations exist across the survey). Candidate line emitters that passed the EW, Σ and visual filtering (to exclude spurious sources or artefacts) are shown in colour, while all other non-excess sources are shown in grey. Emitters classed as Hα ($z = 0.81$), [O III] (+Hβ) at $z = 1.37$ and [O II] at $z = 2.2$, or unclassified, are plotted in different colours as indicated by the key. Note that, as expected, higher redshift emitters are always preferentially found at fainter and fainter magnitudes. Thus, for example, a survey a magnitude shallower would recover little to no [O II] emitters at $z = 2.2$.

Table 2. Redshifts at which the LowOH2 filter ($\lambda_c = 1187$ nm, $\Delta\lambda = 10$ nm), targets strong emission lines and corresponding probed comoving volumes per square degree. There are, of course, many other rarer emission lines which the filter can obtain – see Section 3.3.

Emission line	λ_0 (nm)	z	Volume 10^6 (Mpc^3 deg^{-2})
Hα	656.3	0.81 ± 0.01	0.11
[O III]	500.7	1.37 ± 0.01	0.23
[O III]	495.9	1.39 ± 0.01	0.23
Hβ	486.1	1.44 ± 0.01	0.24
[O II]	372.7	2.18 ± 0.02	0.36
Ly α*	121.6	8.76 ± 0.04	0.52

*See Matthee et al. (2014) for full details on the search for Ly α at $z = 8.8$.

For sources with no z band ($<3\sigma$) available (2 per cent, either because they are too faint in z or because they are masked in CFHTLS), we apply the average correction obtained for all the sources which have reliable z detections (-0.02). Thus, we note that this is, on average, a very small correction.

We classify as potential emitters the sources that have $\Sigma > 3$ (see Fig. 2). Table 2 indicates the major emission lines expected. The second criterion for an excess source to be an emitter is that the emission line must have an observed-frame equivalent width (EW, the ratio of the line flux and the continuum flux densities) higher than the scatter at bright magnitudes. This step avoids selecting sources with highly non-uniform continua (with e.g. strong features). We compute EWs by using

$$\mathrm{EW} = \Delta\lambda_{\mathrm{NB}_J} \frac{f_{\mathrm{NB}_J} - f_{\mathrm{BB}_J}}{f_{\mathrm{BB}_J} - f_{\mathrm{NB}_J}(\Delta\lambda_{\mathrm{NB}_J}/\Delta\lambda_{\mathrm{BB}_J})}, \quad (5)$$

where $\Delta\lambda_{\mathrm{NB}_J}$ and $\Delta\lambda_{\mathrm{BB}_J}$ are the widths of the filters and f_{NB_J} and f_{BB_J} are the flux densities for the NB and BB, respectively. In order to identify a source as a potential line emitter, we require it to have EW (observed) higher than 30 Å, corresponding to an excess of $J - \mathrm{NB}_J > 0.3$. Note that this will correspond to different rest-frame EWs depending on the line/redshift being looked at. For $z = 0.81$, the rest-frame EW limit (Hα + [N II]) is ∼17 Å, while for [O III] + Hβ, at $z \sim 1.4$, and [O II], at $z = 2.2$, it is 12.5 and 9 Å, respectively. These are all relatively low EWs for sources at $z \sim 1$–2 (e.g. Fumagalli et al. 2012; Sobral et al. 2014), and thus the sample completeness will be high (note that we take this into account when obtaining completeness corrections).

Fluxes of emission lines are calculated as follows:

$$F_{\mathrm{line}} = \Delta\lambda_{NB_J} \frac{f_{NB_J} - f_{BB_J}}{1 - (\Delta\lambda_{NB_J}/\Delta\lambda_{BB_J})}. \quad (6)$$

Throughout these calculations, the conversion from magnitudes to flux densities (f_ν) is

$$f_\nu = 10^{23} \times 10^{-0.4(m_{\mathrm{AB}}+48.6)}, \quad (7)$$

where the flux density is given in Jansky and m_{AB} is the magnitude in AB.

Using our selection criteria, out of the 346 244 NB sources individually detected, 8599 emitters were selected as potential line emitters. However, many of these are still likely to be artefacts and/or sources in very noisy regions. This is because these numbers do not include any masking for bright stars and their haloes/nor for other artefacts (see e.g. Geach et al. 2008; Sobral et al. 2009b). We therefore clean our list of potential emitters by visually inspecting all candidates before flagging them as final emitters and produce a final mask. After visually checking all the candidates (obtained before masking), 2284 are marked as spurious or artefacts (fully consistent with e.g. Geach et al. 2008; Sobral et al. 2013a), with the vast majority being either fake sources/artefacts caused by the many bright ($J < 12$) stars in our full coverage, or real sources which have their NB flux boosted (relative to J) due to enhanced stripes, again caused by bright stars. We note that the vast majority of these sources could be masked in an automated way, but due to the very wide area of the survey and the different artefacts caused by stars with different bright magnitudes, our method (visually checking every single potential emitter) assured that no region would be overlooked. The remaining sources are excluded because they are detected in low S/N or noisy regions. This leads to a sample of 6315 potential emitters (see Fig. 2), covering an effective area (after masking) of 7.6 deg².

Following Sobral et al. (2012), we then exclude potential stars within the sample by using an optical versus a near-IR colour. In our case, we use $g - z$ versus $J - K$ and identify blue sources in $J - K$ which are very red in $g - z$.[7] This results in identifying 339 potential stars, which we remove from the sample, resulting in a final sample of robust line emitters of 5976 line emitters. This density of emitters is in very good agreement with similar surveys of smaller areas (Sobral et al. 2013a).

We present the catalogue of emitters containing the final 6315 excess sources, and identify candidate stars with the flag −1 in Appendix A. The catalogue contains photometry in NB J (from this work, 2 arcsec) and J (from UKIDSS, obtained by us, also with 2 arcsec apertures), line flux, EWs and Σ. Table 3 summarizes the catalogue which is available online together with the paper.

Table 3. Summary of the information in the SA22 catalogue. It first shows the number of NB_J detections, then the number of sources selected as emitters in Section 3.1. It shows the number of candidate emitters after removing spurious sources, the number of candidate stars and the final number of robust emitters. The remaining columns present the number of Hα, [O III]/Hβ and [O II] candidates. When available, the number of spectroscopic redshifts are shown as well.

Sample	No. of sources	z-spec
NB_J detections	346 244	16 964
Candidate emitters (after visual check)	6315	541
Stars	339	30
Robust emitters	5976	511
Hα ($z = 0.81$)	2834	295
[O III] + Hβ ($z = 1.37/1.44$)	1056	46
[O II] ($z = 2.18$)	488	19
[S II] ($z = 0.76$)	–	37
$z < 0.7$ ($z > 3$ or unidentified)	1638	82 (26)

The data provides the opportunity to identify large statistical samples of emission-line galaxies, mainly Hα at $z = 0.81$, but also [O III] + Hβ at $z \sim 1.3$–1.4, [O II] at $z = 2.18$, and Lyman-α candidates at $z = 8.8$ (see Matthee et al. 2014). Redshifts and probed volumes for the strongest/main emission lines to which our survey is sensitive are found in Table 2.

3.2 Photometry and photometric redshifts

We combine deep (∼23 AB) J and K data (UKIDSS DXS DR10) from UKIRT/WFCAM (Lawrence et al. 2007) with CFHTLS data in *ugriz* (limit ∼26 AB), to produce a photometric catalogue down to the 5σ depth of the J-band data (J-selected) using 2 arcsec apertures. We then match the catalogue with our catalogue of NB_J emitters.[8] By using the J-selected catalogue, we measure point spread function matched magnitudes in *ugrizK* (2 arcsec apertures), and compile a *ugrizJK* catalogue. We use the photometric catalogue to distinguish between different line emitters by using colour–colour diagnostics, but also by deriving and exploring photometric redshifts.

We compute photometric redshifts by using EAZY (Brammer, van Dokkum & Coppi 2008), which contains a wealth of templates with the main emission lines included. We use *ugrizJK* photometry, but add our NB photometry as well (we also run EAZY without the NB to control for this addition). In total, we obtain photometric redshifts for 5953 emitters (>99 per cent). The photometric redshift distribution of our sample of emitters is shown in Fig. 3. We find that the EAZY photometric redshift distribution has clear peaks at the redshifts of our expected strong emission lines (e.g. Hα at $z = 0.8$, Hβ/[O III] at $z = 1.4$; see Fig. 3). We also find a tentative peak around $z \sim 0.25$, which may be driven by [S III] 9530.6 Å emitters (e.g. Milvang-Jensen et al. 2013) – we confirm some spectroscopically (see Fig. 3). We note that the peaks corresponding to the redshifts of strong emission lines are clearly enhanced by including the NB data. These, probably more accurate, photometric redshifts are used additionally to colour–colour selection and spectroscopic

[7] We identify stars as sources with $(g - z) > 7 \times (J - K) + 2 \wedge (g - z > 1.4) \wedge (J - K) > -0.5$. We check that this procedure eliminates the bulk of spectroscopic stars in both SA22 and COSMOS NB_J samples without excluding real line emitters.

[8] For those without a J band >5σ detection (5 per cent), we measure their J magnitude centred on the NB_J position and, if undetected above 2σ (8 sources, 0.13 per cent), we assign the 2σ limit.

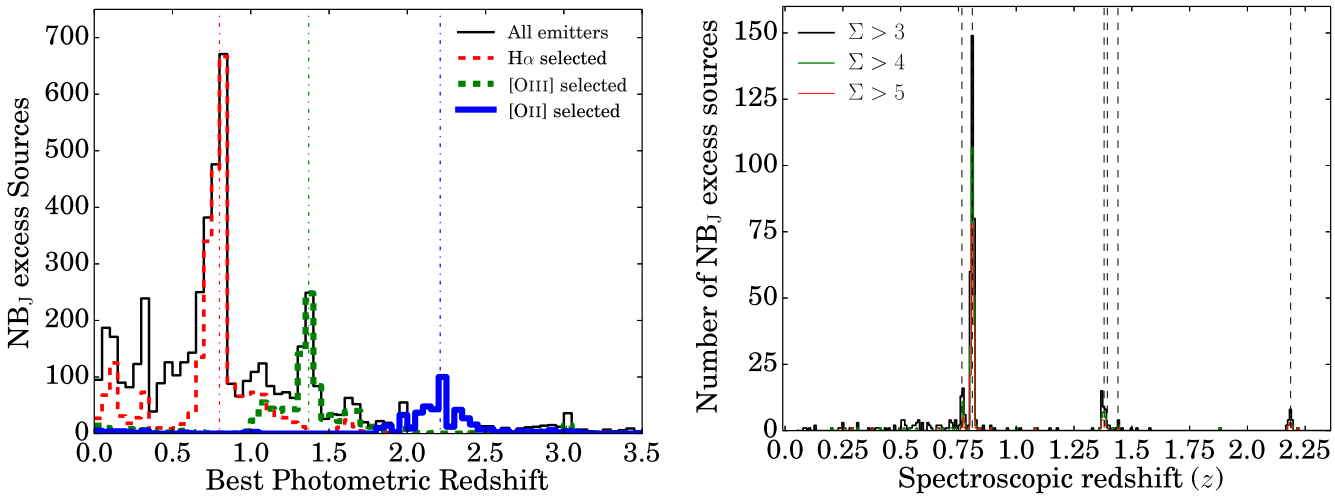

Figure 3. Left: photometric redshift distributions obtained with EAZY (including the NB data) for our NB$_J$ selected line emitters in the SA22 field and comparison with the final samples of the different emitters. This shows that the sample of line emitters is dominated by Hα emitters at $z = 0.8$, followed by Hβ + [O III] emitters ($z \sim 1.4$), while [O II] emitters at $z \sim 2.2$ represent a lower fraction. Right: the distribution of spectroscopic redshifts of the full sample of line emitters, including our follow-up observations with FMOS on Subaru and MOSFIRE on Keck. This clearly shows that the sample of spectroscopically confirmed line emitters is dominated by Hα emitters, followed by [O III], [O II], Hβ and [S II] emitters. Other much rarer lines are found, including Pa γ 10 870, N I 10 406, C I 9853/9827, [S III] 9533, He II 8237, O I 7774 O II 7330/7320, [Ar III] 7135, [O I] 6363/6300, He I 5876, [S III] 6311+He II 6311. Note that the sample with available spectroscopic redshifts is still dominated by optical follow-up (and observed-optical selection), and thus is highly biased against the higher redshift line emitters. Dashed lines indicate the redshifts of the major emission lines studied in this paper.

redshifts to distinguish between various emission lines, as outlined in Section 3.5. Furthermore, in Fig. 5 we show how our EAZY photometric redshifts (with the inclusion of the NB) compare with all the spectroscopic redshifts available for the sample, revealing good agreement (the spectroscopic redshift distribution of the sample with spectroscopic redshifts is shown in Fig. 3).

3.3 Spectroscopic redshifts: literature compilation

Spectroscopic redshifts are available from the VIMOS-VLT Deep Survey (VVDS),[9] which covered 4 deg^2 in the SA22 field. Additional spectra are available from VIPERS (Garilli et al. 2014). In total, 289 of our candidate line emitters have a spectroscopic redshift from either VVDS or VIPERS. However, and particularly for redshifts above 1 (but even for $z \sim 0.8$), it becomes significantly more difficult to identify emission lines in the VVDS survey (the survey selection is $I < 22.5$ and there are few and weak lines in the optical), as the strongest lines are shifted to the (near)-infrared, or still at bluer wavelengths (like Lyman α). Thus, targeted spectroscopic follow-up (directly in the observed NIR) is ideal to increase the fraction of spectroscopically confirmed line emitters, and allow a more representative evaluation of the range of line emitters within the sample that is not biased towards rest-frame UV bright galaxies.

3.4 Follow-up observations with MOSFIRE and FMOS

In order to significantly complement spectroscopic redshifts available from the literature, we follow-up a significant fraction of the sources. We obtained spectra for some of our sources using KMOS (Sharples et al. 2013), and these are presented in Sobral et al. (2013b) and Stott et al. (2013b). We also followed up some of the bright line emitters with the William Herschel Telescope (WHT) and the New Technology Telescope (NTT); these are presented in Sobral et al. (2015).

[9] http://cesam.oamp.fr/vvdsproject/index.html; Le Fèvre et al. (2013).

Furthermore, we have observed large samples of our emitters using FMOS (Kimura et al. 2010) on Subaru and MOSFIRE (McLean et al. 2010, 2012) on Keck. Both instruments are shown to be tremendously efficient for our targets. This is because not only do we know where one of the main emission lines should be found (within a low OH emission window, thus maximizing the S/N), but we also know that the vast majority have fluxes high enough to be detected in modest exposure times.

3.4.1 FMOS observations

FMOS observations were taken on 2014 June 15, under clear conditions and good seeing (0.7–0.9 arcsec). We observed sources within the SA22 field in two different configurations, centred on 22 19 57.62 +00 19 35.28 (P1) and 22 19 05.06 +00 52 34.19 (P2). We selected a total of 128 sources from our candidate line emitters in SA22 for P1 and another set of 128 sources from our candidate line emitters in SA22 for P2. The remaining fibres (72) per configuration targeted fillers, calibration stars or were disabled. We used the *J*-Long ($R = 2200$) setting (high-resolution mode). Individual exposure times were 900 s in each nodding position and we obtained two of each position for a total exposure time of 3.6 ks per pointing, corresponding to 1.8 ks of on-target exposure time due to cross-beam switching.

Apart from the raw science frames, we obtained (per configuration) dome-flats and Th-Ar spectral calibration arcs which were used for the reduction. In order to reduce the data, we used the Subaru FMOS reduction pipeline, FMOS Image-Based REduction Package (FIBRE-PAC; Iwamuro et al. 2012). FIBRE-PAC is a combination of IRAF tasks and C programs using the CFITSIO library (Pence 1999). Details are found in Iwamuro et al. (2012). Briefly, data are flat-fielded and bad pixels removed to begin with. After that, corrections are applied to fix spatial and spectral distortions present and then wavelength calibration is performed. An initial background subtraction is then achieved using the ABAB nodding

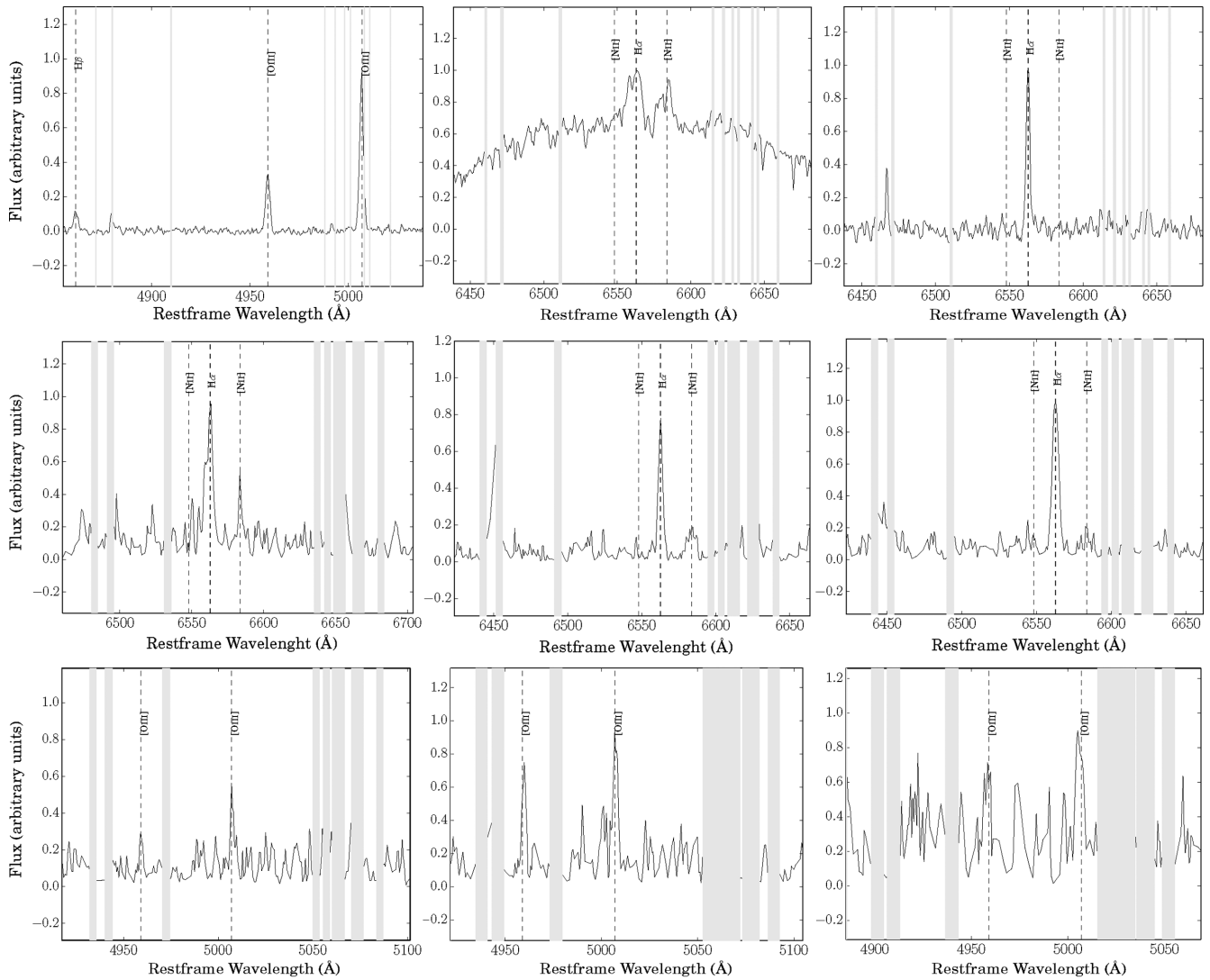

Figure 4. Top: some examples of our MOSFIRE spectra. Top left: one of the confirmed [O III] emitters with a strong detection of Hβ, where the emission line being measured and detected by our NB filter is actually [O III] 4959. Top middle: a very luminous broad-line AGN with extremely broad Hα line, implying a very massive black hole of $\sim 10^9$ M$_\odot$. Top right: a high EW Hα emitter, representative of the less massive population of Hα emitters. The locations of the strongest OH lines are indicated as shaded lines, but we note that other weaker OH lines are also present. No smoothing has been applied. Middle: representative examples of spectroscopically confirmed Hα emitters in our sample within our FMOS data set. No smoothing has been used. The strongest OH emission lines are represented with shaded regions (but weaker OH lines are not indicated). Bottom: representative examples of spectroscopically confirmed [O III] emitters in our sample within our FMOS data set. No smoothing has been used.

pattern of the telescope to perform an A-B sky subtraction. Further bad pixel, detector cross talk, bias difference, distortion, residual background and sky corrections are applied. The 2D spectra are combined, which in cross beam switching mode means inverting and adding the negative B spectra to the A spectra. The final step is an initial flux calibration, which we cross check with our NB estimated fluxes (see Stott et al. 2013b, for more details).

We obtained redshifts by first identifying the emission line within the NB$_J$ filter profile wavelength range, and then fitting with various combinations. For most spectroscopic redshift determinations there were at least two lines, but whenever only one line was present, all solutions were evaluated, and a lower confidence flag was assigned (in almost all cases, such sources were classified as potential Hα emitters).

We obtain redshifts for all NB$_J$ emitters with a usable spectrum above a flux of 1.5×10^{-16} erg s^{-1} cm^{-2}, but we also obtain redshifts for more than half of sources with fluxes down to 1.0×10^{-16} erg s^{-1} cm^{-2}. In total, out of the 241 usable spectra, we detect a strong emission line and obtain redshifts for 185 sources. For 135 sources, we are able to obtain a robust spectroscopic redshift, based on at least two or more emission lines. Sources without a redshift determination are all lower flux sources, requiring higher exposure times in order to be significantly detected (note that we integrated for only 1.8 ks per source). We show examples of FMOS spectra in Fig. 4, which shows examples of Hα and [O III] emitters.

3.4.2 MOSFIRE observations

MOSFIRE observations were obtained on 2014 November 5, under clear conditions. The seeing was 0.8 arcsec. We observed two masks, both with 0.7 arcsec slits. The first one was centred on

22:13:24.89, −00 12 16.31 (p1), while the second one was centred on 22 13 30.27, −00 53 21.24 (p3). Individual exposures were 120 s in each A and B positions, and we repeated each twice, for a total exposure time of 480 s pixel^{-1} for each of the masks. We reduced the data using the MOSFIRE team data reduction pipeline (DRP).[10] The DRP produces flat-fielded and wavelength-calibrated combined 2D spectra for the individual objects. The spectra are wavelength-calibrated using the sky lines and then. In practice, DRP follows very similar steps to those used to reduce FMOS data.

For mask p1, we were able to observe 17 of our candidate line emitters, obtaining robust redshifts for all 17 sources. For mask p3, we targeted eight of our line emitter candidates, and obtained a robust spectroscopic redshift for seven out of the eight. For the missing source, the S/N was too low to detect the emission lines. We note that for most sources the S/N obtained with MOSFIRE are at least comparable, and in most cases much higher than FMOS, despite MOSFIRE data having an exposure time of only ∼13 per cent that of FMOS. However, FMOS has a larger multiplexing, allowing us to target about 7–10 times more sources (taking into account cross beam switching), and thus it is competitive with MOSFIRE for our targets, i.e. for sources spread over relatively wide areas and that have emission lines which avoid strong OH lines.

In total, we obtained redshifts for 24 sources with MOSFIRE. We show examples of MOSFIRE spectra in Fig. 4. Fig. 3 shows the full spectroscopic redshift distribution.

Together, the FMOS and MOSFIRE data sets significantly add to the number of spectroscopic redshifts in the sample of line emitters, not only allowing for a spectroscopically confirmed sample of just over 300 Hα emitters at $z = 0.81$, but also allowing us to directly investigate the contamination by the adjacent [N II] line and contributing to a robust correction (see Section 4.1).

Finally, by compiling all the 511 spectroscopic redshifts, we test the accuracy of our photometric redshifts for line emitters, which can be seen in Fig. 5. Fig. 5 shows that the redshifts overall agree well, but that there is some scatter, particularly for $z > 1$. We use our results, and our full spectroscopic sample, to optimize our selection of different line emitters and to estimate the completeness and contamination of each sample.

3.5 Testing SA22 photometric redshift selection using COSMOS

In order to further check the validity/accuracy of using our EAZY photometric redshifts (*pz*) based on *ugrizJK* and our NB for the bulk of our sample, we make use of similar data available in the COSMOS field (Sobral et al. 2013a). As the COSMOS field (Capak et al. 2007; Scoville et al. 2007; McCracken et al. 2012) has been widely studied, there is data available in 30 bands (see e.g. Ilbert et al. 2009, 2010, 2013), making photometric redshifts much more reliable, while there are also relatively more spectroscopic redshifts available in the literature. Thus, in order to have a fully comparable sample, we compute photometric redshifts in COSMOS by following the same method as for our SA22 sample, i.e. by using *ugrizJK* and NB$_J$. This allows us to directly compare photometric redshifts for a sample of similar line emitters and investigate any biases/incompleteness due to the limited availability of photometric bands.

For the 700 sources selected as line-emitters in COSMOS by Sobral et al. (2013a) in their NB$_J$ band, 76 had spectroscopic redshifts (Lilly et al. 2009). As a first step, we derive photometric redshifts

[10] https://code.google.com/p/mosfire/

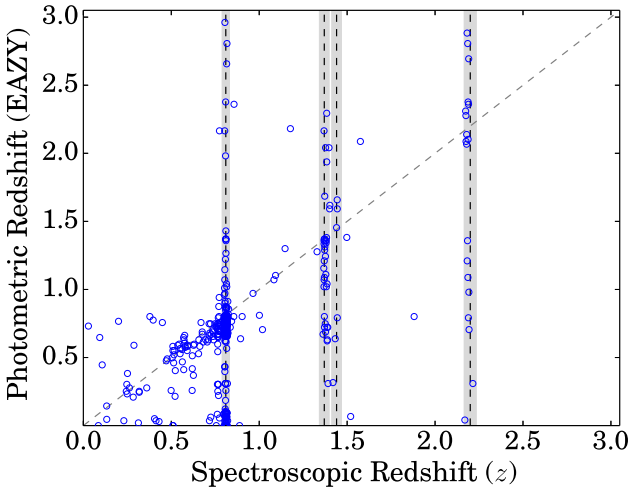

Figure 5. Comparison between photometric and spectroscopic redshifts for NB$_J$ selected line emitters for our SA22 NB survey. Redshifts for the emission lines that we select are shown as vertical lines. Many other lines are not labelled, but found, e.g. from low to high redshift: Pa γ 10 870, N I 10 406, C I 9853/9827, [S III] 9533, He II 8237, O I 7774 O II 7330/7320, [Ar III] 7135, [S II] 6731/6716, Hα 6563, [O I] 6363/6300, He I 5876, [S III] 6311+He II 6311, [O III] 5007/4959, Hβ 4861, [O II] 3727.

using *ugrizJK* in COSMOS without the inclusion of the NB$_J$ filter and compare them with those derived with the inclusion of the NB. We find that both photometric redshift sets are in good agreement for most redshifts, except for the redshifts of strong emission lines such as Hα at $z = 0.8$; for these photometric redshifts clearly benefit from the addition of the NB and are always more accurate at recovering spectroscopic redshifts, resulting in typical photometric redshift uncertainties of ≈0.05. We compare the photometric redshifts obtained in COSMOS, with all bands, from Ilbert et al. (2009), with those we derive with EAZY using the restricted set of bands including NB. Our results are shown in Fig. 6. We find good

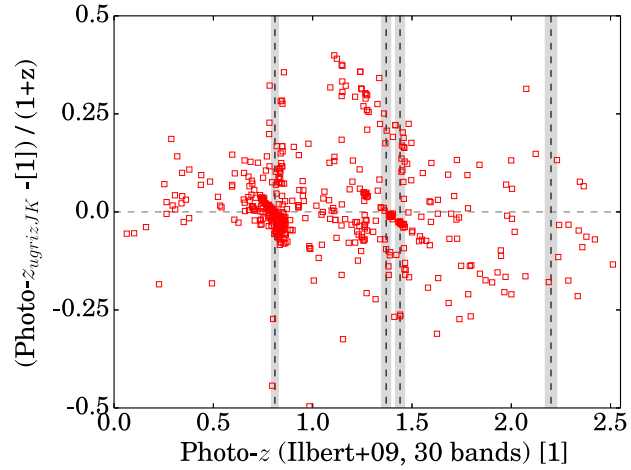

Figure 6. A comparison between photometric redshifts for the COSMOS sample of NB$_J$ line emitters (from Sobral et al. 2013a), derived by us with EAZY using *ugrizJK* and NB magnitudes and those derived by Ilbert et al. (2009) with 30 bands from FUV to IRAC 8 μm. We find a good correlation between the two estimates, with a relatively small scatter below $z < 1$. The scatter increases at higher redshift, particularly due to the lack of IRAC bands, but our photometric redshift ranges used for selection, together with our colour–colour selections are able to tackle that.

agreement between both. We find that for the bright ($NB_J < 21.2$, AB) subset of the sample, the EAZY photometric redshifts are consistent to ≈0.02. For fainter sources, the EAZY photometric redshifts can sometimes differ significantly from those using all bands, due to larger error bars for the faint sources, but also because the lack of bands redder than *K* and the lack of *H* and UV bands.

Finally, using the data and computed photometric redshifts in COSMOS we check our selection of emitters that we will apply for SA22. We apply the same selection criteria (see Section 3.6) in COSMOS, using both the Ilbert et al. (2009) photo-*z*s based on 30-bands and our photometric redshifts with *ugrizJK* + NB. We find that our SA22-like selection criteria (photo-*z* with just *ugrizJK* + NB and colour–colour selections; see Section 3.6) applied to COSMOS recovers 90 per cent of the Hα emitters (97 per cent of the spectroscopically confirmed Hα emitters) that are found with much more accurate photometric redshifts. Our photometric redshifts tend to sometimes favour higher redshift solutions compared to Ilbert et al. (2009), which leads to one spectroscopic Hα emitter being selected as [O II] emitter (but this is a small effect). However, apart from this, the selection is very clean and leads to the conclusion that for our emitters, *ugrizJK* + NB photometric redshifts + colour–colour selections are sufficient and comparable to the best photometric redshifts in COSMOS. We check that similar conclusions are reached for Hβ + [O III] and [O II] emitters. Therefore, our samples selected from SA22 and COSMOS (all selected with the same selection criteria) are unlikely to have strong differences and should have very comparable completeness (∼90 per cent) and contamination (∼10 per cent) fractions.

3.6 Selection of Hα, [O III]/Hβ & [O II] emitters

The selection of Hα, [O III]/Hβ & [O II] emitters at $z = 0.81$, 1.4, 2.2, respectively, is done following Sobral et al. (2013a), by using a combination of photometric redshifts (and spectroscopic redshifts, when available) and colour–colour selections optimized for star-forming galaxies at the redshifts of interest ($z \sim 0.8$, 1.4, 2.2). We apply these selections not only to the sample of emitters in SA22 presented here, but also to COSMOS and UDS (from Sobral et al. 2013a), in order to obtain larger samples over a larger number of independent volumes. Spectroscopic redshifts are also used to evaluate the completeness and contamination of the sample, although due to the selection function of most of the literature spectroscopic redshifts they are only really useful for sources up to $z \sim 0.8$. The selection criteria applied here are the same as Sobral et al. (2013a) for Hα at $z \sim 0.8$, while for [O III] + Hβ at $z \sim 1.4$ we use their criteria to select Hα sources at $z \sim 1.47$, and for [O II] emitters at $z \sim 2.2$ we use their criteria for $z \sim 2.2$ Hα emitters. We check with COSMOS and UDS that these criteria work well (see Khostovan et al. 2015, for more details).

The following sections describe the specific selection criteria used to identify the different line emitters, but briefly: (i) we apply a photometric redshift range, centred on the expected redshift of the line, and that takes into account typical errors (motivated by the estimated uncertainties in the photometric redshifts and on e.g. Figs 5 and 6) and the proximity of any other strong emission line, (ii) we use colour–colour selections (see e.g. Fig. 7), to increase completeness and (iii) we explicitly remove any spectroscopically confirmed contaminant and include any spectroscopic confirmed source in the sample. Finally, if a source is classified as e.g. Hα it can no longer be selected as a higher redshift emitter and, if classified as [O III] + Hβ it will not be able to be selected as an [O II] emitter. We use our spectroscopic redshifts in order to test different photometric redshift cuts and study the completeness and contamination of our samples. We find that we need to use a wider cut in photometric redshifts than Sobral et al. (2013a) in order to maximize the completeness (95 per cent for Hα emitters), while maintaining the contamination at a low level of ∼10–15 per cent. This is a simple consequence of our photometric redshifts in SA22 having a larger uncertainty.

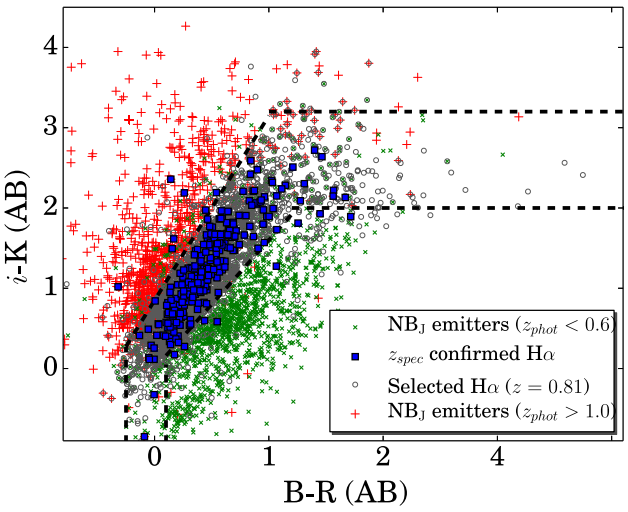

Figure 7. $B - R$ versus $i - K$ colour–colour separation (Sobral et al. 2009b) that we use (together with photometric redshifts) to identify Hα emitters within the full sample emitters. Here, only NB line emitters are plotted. Note how the selection is able to distinguish between $z \sim 0.8$ sources, those at $z > 1$ and those at lower redshifts ($z < 0.6$), but also that only ∼5 per cent of all spectroscopically confirmed Hα emitters do not fall on our *BRiK* colour–colour selection, thus implying high completeness.

3.6.1 Hα emitters at z = 0.81

Hα emitters at $z \sim 0.8$ can be distinguished from lower and higher redshift emitters very efficiently by using (i) our photometric redshifts and (ii) the *BRiK* (Sobral et al. 2009b) colour–colour diagram (see Fig. 7)[11], in combination with photometric redshifts and spectroscopic redshifts (Sobral et al. 2009b). We classify a source as Hα if (i) its photometric redshift is within $z_{\rm phot} = 0.70$–0.95 (1523 sources) or (ii) if it satisfies the *BRiK* colour–colour criteria (Fig. 7) and does not have $1.3 < z_{\rm phot} < 1.5$ (likely [O III]) or $2.0 < z_{\rm phot} < 2.4$ (likely [O II]): this step adds 1330 extra sources (a further 200 sources that satisfy the *BRiK* selection are not included due to their photo-*z*s clearly pointing towards higher redshift line emitters). We use the *BRiK* colour–colour selection to guarantee high completeness even for the faintest sources, where photometric redshifts become unreliable and/or unavailable. We recover 286 spectroscopically confirmed Hα emitters (see Fig. 4 for a few typical examples), remove 34 sources spectroscopically confirmed to be other emission lines either than Hα ([S II] emitters are clearly the largest contaminant, but there are also a few low-redshift emission lines for which VVDS and VIPERS is extremely complete), and introduce 15 spectroscopically confirmed Hα emitters which

[11] For SA22, we estimate *B*-band photometry by using *g* and mimicking the necessary empirical correction extracted from the COSMOS and UDS (Sobral et al. 2013a), for which *u*, *g* and *B* data is available. The correction is: $(B - r)_{\rm cor} = (g - r) \times 1.1657$–$0.223$.

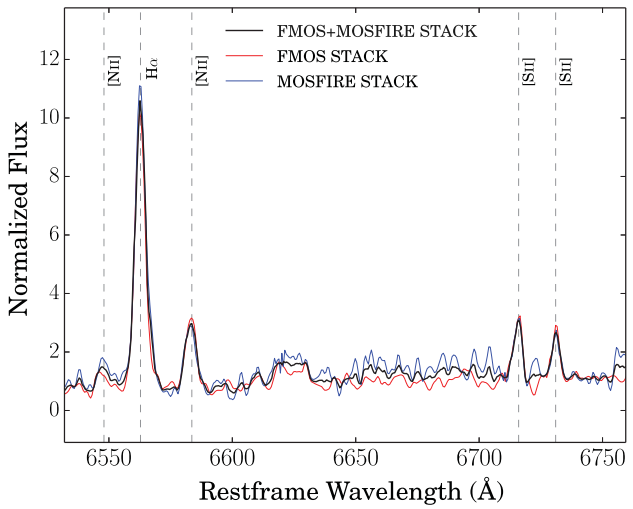

Figure 8. Stacked spectra of $z = 0.81$ Hα emitters obtained with FMOS and MOSFIRE. As a whole, our Hα emitters have a metallicity $12 + \log_{10}(\mathrm{O/H}) = 8.56 \pm 0.05$, slightly subsolar (but consistent with solar metallicity $12 + \log_{10}(\mathrm{O/H}) = 8.66 \pm 0.05$). We also make significant detections of the [S II] doublet, with a line ratio of $\mathrm{I([S\,II]_{6716})/I([S\,II]_{6731}) = 1.33 \pm 0.08}$, implying an electron density of 40–200 cm^{-3}. The $\mathrm{[S\,II]_{6716}/I(H\alpha)}$=0.14 ± 0.02 ratio also implies ionization potential of $\log_{10}(U) = -3.9 \pm 0.5$ cm^{-3}.

did not satisfy the photometric selection criteria. Comparing this number with the 286 emitters spectroscopically confirmed to be Hα which our selection successfully recovers indicates ∼95 per cent completeness. With the current spectroscopic follow-up, we estimate contamination on the level of ∼10 per cent, being dominated by [S II] emitters at $z = 0.76$. We note that this is a consequence of photometric redshifts and colour–colour selections not being able to completely distinguish between Hα and [S II] emitters as they are very close in redshift.

In total, we select 2834 Hα emitters in SA22 (photometric redshift distribution of the final sample is shown in Fig. 3). Out of all Hα emitters in SA22, 295 are spectroscopically confirmed (see Fig. 4 for examples from FMOS, and Fig. 8 for the composite spectra), from VVDS (75), VIPERS (74), KMOS (20; Sobral et al. 2013b; Stott et al. 2014), 10 sources followed-up with NTT and WHT, 16 with MOSFIRE and 100 with FMOS.

3.6.2 *[O III] and Hβ emitters at z ∼ 1.4*

In order to select [O III] + Hβ emitters[12], we use the same selection as for Hα emitters at $z = 1.47$ in Sobral et al. (2013a). Our selection criteria are the following: (i) we disregard all candidate Hα emitters obtained in the previous subsection (rejection of 2834 sources), (ii) we use our photometric redshifts and apply the selection $1.0 < z_{\mathrm{phot}} < 1.8$ (911 sources) and (iii) we use our *BzK* selection to identify $z \sim 1.4$ sources from which we remove sources likely to lie at $z > 2$ on the basis of their *izK* colours (see Sobral et al. 2013a) and a further 41 sources with $2.0 < z_{\mathrm{phot}} < 2.4$ which are likely to be [O II] emitters; this adds 141 sources to the sample. Just like for Hα, we also use spectroscopic redshift information to remove eight non-[O III] + Hβ emitters in the sample and add 12 spectroscopically confirmed [O III] + Hβ emitters. Based on the limited spectroscopic information, we estimate ∼80 per cent completeness and ∼15 per cent contamination. We have a total of 46 spectroscopically confirmed sources; in Fig. 4 we show some examples. By following these selection criteria, we select 1056 [O III] + Hβ emitters in SA22 (photometric redshift distribution of the final sample is shown in Fig. 3). We note that our limited spectroscopic follow-up already allows us to constrain the fraction of Hβ and [O III] emitters within this sample. Within the full sample of spectroscopically confirmed [O III] + Hβ emitters, we find that ∼16 per cent are Hβ emitters (Hβ line detected by the NB filter), and that Hβ emitters have lower luminosities than [O III] emitters, likely contributing more towards the faint-end of the LF, but very little at the bright end. We discuss the implications in more detail in Section 4.2.

3.6.3 *[O II] emitters at z = 2.2*

We select [O II] emitters at $z = 2.2$ with a similar selection as for Hα emitters at $z = 2.23$ in Sobral et al. (2013a). Our selection is as follows: (i) we disregard all candidate Hα emitters and [O III] + Hβ emitters obtained in the previous subsections (rejection of 3888 sources); (ii) we use our photometric redshifts, applying the selection $1.8 < z_{\mathrm{phot}} < 2.7$ (424 sources); and (iii) we use a *BzK* cut, to isolate $z > 1.5$ galaxies (Sobral et al. 2013a), adding 69 new sources to the sample. We also use spectroscopic redshift information to remove seven non-[O II] emitters in the sample and introduce two spectroscopically confirmed sources. We obtain a final sample of 488 [O II] emitters in SA22, with 19 sources being spectroscopically confirmed. Our spectroscopic sample is extremely limited for our [O II] sample, but based on the available redshifts, we estimate a ∼90 per cent completeness and a <25 per cent contamination. We note that we do not apply the Lyman-break selection in order to exclude $z > 3$ emitters as in Sobral et al. (2013a), because no major emission lines are found at slightly higher redshift – but the reader is referred to Matthee et al. (2014) for a full discussion of emitters at much higher redshift.

3.6.4 *Final samples: SA22*

Using the data and steps mentioned in previous chapters, we add an identifier in the range 0–3 to classify between other lines/unclassified (0), Hα (1), [O III] + Hβ (2) and [O II] (3) to the catalogue of emitters containing the final 5976 robust excess sources. Table 3 presents the final numbers per sample, including the number of spectroscopically confirmed.

3.6.5 *Samples in COSMOS and UDS*

We supplement our sample using the HiZELS catalogue of line emitters available from Sobral et al. (2013a), which has been derived in a very similar way to ours, and with almost identical data quality and depth. For Hα, we use the samples presented by Sobral et al. (2013a): 425 Hα emitters in COSMOS and 212 in UDS.[13]

[12] It is not possible to completely distinguish, with photometric redshifts and/or colour–colour selections between [O III] and Hβ emitters due to a very similar redshift, although we expect [O III] emitters to dominate the sample – see Section 4.2.

[13] By applying exactly the same selection as in SA22, we would obtain two extra sources in COSMOS, and three extra sources (and less six) in UDS relative to Sobral et al. (2013a), but we chose to use the sample presented in Sobral et al. (2013a) for consistency. These minor differences in the sample make no difference in the results, particularly as they are found at the fainter end.

Table 4. A summary of our final samples of Hα, [O III] + Hβ and [O II] emitters at $z \sim 0.8$, ∼1.4 and ∼2.2, respectively.

Samples Hα ($z \sim 0.8$)	No. of sources (#)	Volume (10^6 Mpc^3)	Depth ($\log_{10}L$) (erg s^{-1})
SA22	2834	0.9	41.35
COSMOS	425	0.1	41.25
UDS	212	0.08	41.30
Full sample	3471	1.1	41.35
[O III] + Hβ ($z \sim 1.4$)	(#)	(10^6 Mpc^3)	(erg s^{-1})
SA22	1056	2.48	42.0
COSMOS	159	0.33	41.90
UDS	128	0.24	41.95
Full sample	1343	3.05	42.0
[O II] ($z \sim 2.2$)	(#)	(10^6 Mpc^3)	(erg s^{-1})
SA22	488	2.62	42.6
COSMOS	41	0.29	42.50
UDS	43	0.22	42.55
Full sample	572	3.12	42.6

In order to select [O III] + Hβ and [O II] emitters, we apply exactly the same criteria as in SA22. In COSMOS, we find 159 [O III] + Hβ and 41 [O II] emitters. In UDS, we find 128 [O III] + Hβ emitters at $z = 1.4$ and 43 [O II] emitters at $z = 2.2$. We note that the selection criteria (optimized for the COSMOS and UDS fields) in Khostovan et al. (2015) to select [O III] + Hβ and [O II] emitters with the HIZELS NB$_J$ sample is slightly different than ours in the photo-z selection (slightly more restrictive due to photo-zs in COSMOS and UDS being better than in SA22), but although it results in some very minor differences in the source numbers (maximum ∼15 per cent but typically within 5–10 per cent), the samples are essentially the same (>90 per cent of the samples are the same) and lead to the same results. Here, we chose to use the same selection criteria as for SA22 to have fully consistent samples across fields, but we note that even if we applied Khostovan et al. (2015) selection criteria our results would not change.

We also note that the HiZELS data (Sobral et al. 2009b, 2013a) are slightly deeper (by ∼0.1 dex in luminosity) than our SA22 survey and that the NB filter used by Sobral et al. (2013a) is slightly wider, and thus naturally recovers a higher number of emitters per deg^2 if that is not taken into account.

We provide a summary of the sample in the three fields in Table 4. The final samples (SA22, COSMOS and UDS) are by far the largest ever assembled, yielding 3471 Hα emitters at $z \sim 0.8$ (∼400 spectroscopically confirmed),1343 [O III] + Hβ at $z \sim 1.4$ and 572 [O II] emitters at $z \sim 2.2$. Some relatively significant variations in source densities are found across fields and within fields. This is evaluated in Section 5, where we present how important is cosmic variance on different scales for each one of the emission lines.

3.7 AGN fraction for line emitters

We take advantage of the similar samples we selected in COSMOS and UDS to estimate the fraction of potential AGNs in our SA22 sample (where no *Chandra* data and no IRAC data are available). We start by using C-COSMOS (Elvis et al. 2009) to find that, in agreement with Garn et al. (2010), only ∼1 per cent of Hα sources are detected in the X-rays (five sources, X-ray luminosities of $10^{42.8 \pm 0.11}$ erg s^{-1}). For our [O III] + Hβ and [O II] samples in COSMOS, we only find one X-ray match per sample (both sources with X-ray luminosities of $\sim 10^{43.4}$ erg s^{-1}), thus resulting in a very small fraction (<1 per cent for [O III] + Hβ at $z \sim 1.4$) and ∼2–3 per cent for [O II] at $z = 2.23$. It is therefore clear that the fraction of X-ray detected AGNs in our samples are only at the level of ∼1 per cent. All X-ray AGN within our sample are found to have intermediate line luminosities (so they are not the highest luminosity line emitters, but none is a weak line emitter).

We also use deep IRAC data, in both COSMOS and UDS, to look for sources dominated by signatures of the stellar bump (star-forming dominated), thus showing a blue colour beyond ∼1.6 μm rest frame and sources dominated by signatures of a red power law (AGN-dominated), showing red colours even beyond ∼1.6 μm rest frame (see e.g. Garn et al. 2010). Given the typical errors on IRAC photometry, particularly for redder bands, and motivated by the colours of our X-ray AGN (in order to recover the majority of them with the appropriate IRAC colours given the redshift of the sources), we apply the following cuts to select potential AGN, which should be mostly interpreted as an upper limit to the AGN contamination. For Hα ($z = 0.8$): [3.6]–[4.5]>0.0 (10 per cent potential AGN); for [O III] + Hβ ($z \sim 1.4$): [4.5]–[5.8]>0.1 (18 per cent potential AGN) and for [O II] ($z = 2.2$): [5.8]–[8.0]>0.2 (23 per cent potential AGN). The contamination for these relatively 'typical' sources can be well approximated by a constant contamination across luminosities at least up for the luminosities probed by the UDS and COSMOS samples.

In order to constrain the fraction of AGN among the most luminous Hα emitters, we use the results from Sobral et al. (2015). Sobral et al. (2015) present the results of the spectroscopic follow-up of very luminous Hα emitters, allowing us to constrain the fraction of AGN for Hα luminosities $L > L^*$ (up to $L \sim 50L^*$). Their results show that while the AGN fraction is relatively constant up to $L \sim L^*$ (around 10 per cent), it strongly correlates with L/L^* for higher luminosities, reaching ∼100 per cent by $L \sim 50L^*$. Therefore, for Hα emitters, we use results from X-rays and from IRAC, which allow us to have a global AGN fraction for typical luminosities, and the results from Sobral et al. (2015), which allow us to estimate the AGN fraction at the highest luminosities (for Hα emitters). We thus assume that our Hα sample ($z = 0.8$) will likely be ∼10 per cent contaminated by AGN up to $L \sim L^*$, and we use the AGN fraction as a function of L/L^* found by Sobral et al. (2015), which is given by $0.59 \times \log_{10}(L/L^*) + 0.112$.

For our sample of [O III] + Hβ and [O II] emitters, we have no information regarding the AGN fraction of the highest luminosity emitters. We thus assume a constant AGN fraction when converting luminosity densities to star formation rate densities (SFRD). For our sample of [O III] + Hβ, we assume we will be ∼15 per cent contaminated by AGNs (Stott et al. 2013b, find that for Hα emitters at $z = 1.4$ only 10 per cent are AGN) and for our [O II] sample at $z = 2.2$ we use a correction of ∼20 per cent.

4 METHODS: COMPLETENESS AND CORRECTIONS

4.1 Flux corrections: [N II]/Hα correction

When computing line fluxes and EWs for the Hα emitters (Hα line), one must note that the adjacent [N II] lines at 6548 and 6583 Å will also contribute to both quantities, increasing them both (see e.g. Fig. 8). [N II] 6583 Å is clearly the strongest and the most important to take into account (see Fig. 8); we refer to it as simply [N II] for

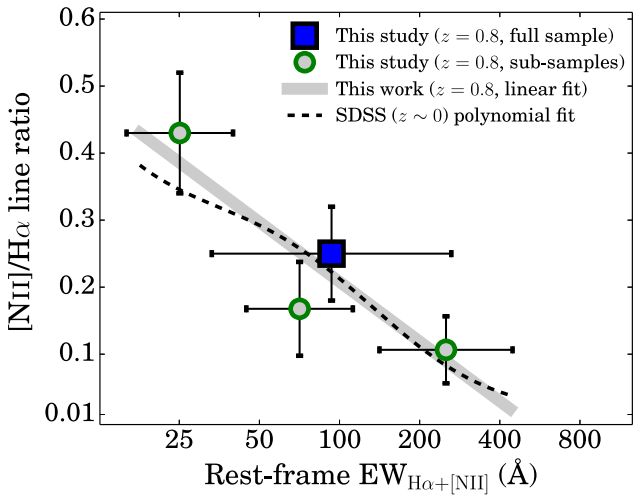

Figure 9. The observed anticorrelation between the [N II]6583 Å/Hα line ratio and the rest-frame EW(Hα + [N II]6583 Å) at $z = 0.81$ in this study and a comparison with the polynomial fit to SDSS at low redshift. Given the range of EWs and line ratios probed, we find that while the SDSS fit is a good fit to the data at $z = 0.81$, a linear fit does equally well, and is a much simpler way to robustly estimate [N II]/Hα. The relation is $f([\mathrm{N\,II}]/\mathrm{H}\,\alpha) = -0.296 \times \log_{10}(\mathrm{EW}_{\mathrm{H}\alpha+[\mathrm{N\,II}]}) + 0.8$. We show three bins which split the sample in three relatively evenly in parameter space and that do not overlap, together with the stack for the full sample, and three other stacks which split the sample in terms of number of Hα emitters.

the remaining of the paper. One way to correct for this is to use the Sloan Digital Sky Survey (SDSS) relation between $F_{[\mathrm{N\,II}]}/F_{\mathrm{H}\alpha}$ and the total measured EW (Hα + [N II]) (Villar et al. 2008; Sobral et al. 2012). Given the representative spectroscopic follow-up of our Hα emitters, we can directly test whether the polynomial correction presented in Sobral et al. (2012, based on SDSS) is appropriate for $z \sim 1$.

We use our FMOS and MOSFIRE data and stack as a function of rest-frame EW(Hα + [N II], measured from the NB data as that is what we want to evaluate). We recover a clear anticorrelation between $\mathrm{F}_{[\mathrm{N\,II}]}/\mathrm{F}_{\mathrm{H}\alpha}$ and rest-frame EW(Hα + [N II]), as shown in Fig. 9. We find that the trend is fully consistent with SDSS of decreasing $F_{[\mathrm{N\,II}]}/F_{\mathrm{H}\alpha}$ as a function of EW(Hα + [N II]). We note, as shown in Fig. 9, that a linear relation is an even simpler correction which is extremely similar to the SDSS relation. The linear relation that we derive is

$$f([\mathrm{N\,II}]/\mathrm{H}\,\alpha) = -0.296 \times \log_{10}(\mathrm{EW}_{\mathrm{H}\alpha+[\mathrm{N\,II}]}) + 0.8 \quad (8)$$

and is valid for rest-frame EW(Hα + [N II]) from ∼15 to ∼600 Å (see Fig. 9).

We note that if a single correction is applied (for similar surveys), then the value to be applied for $\mathrm{F}_{[\mathrm{N\,II}]}/\mathrm{F}_{\mathrm{H}\alpha}$ should be 25 per cent, as the $\mathrm{F}_{[\mathrm{N\,II}]}/\mathrm{F}_{\mathrm{H}\alpha}$ for the full stack (median stack, see Fig. 8) is 0.25 ± 0.05. This suggests (in agreement with e.g. Swinbank et al. 2012b; Sobral et al. 2013b; Stott et al. 2013b), using the Pettini & Pagel (2004) calibration, that as a whole, our Hα emitters have a metallicity $12 + \log_{10}(\mathrm{O/H}) = 8.56 \pm 0.05$, slightly subsolar (solar metallicity $12 + \log_{10}(\mathrm{O/H}) = 8.66 \pm 0.05$).

In our stacked spectra, we also make significant detections of the [S II] doublet. We find a median line ratio of $\mathrm{I}([\mathrm{S\,II}]_{6716})/\mathrm{I}([\mathrm{S\,II}]_{6731}) = 1.33 \pm 0.08$, which implies an electron density of 40–200 cm^{-3} (Osterbrock 1989). We also find that $[\mathrm{S\,II}]_{6716}/\mathrm{I}(\mathrm{H}\alpha) = 0.14 \pm 0.02$, which implies an ionization potential of $\log_{10}(U) = -3.9 \pm 0.5$ cm^{-3} (Osterbrock 1989). However, we emphasize that the stacked spectra provide only information on the median value, and no indication on the range of values.

4.2 The relative contributions from [O III]$_{5007}$, [O III]$_{4959}$ and Hβ to the sample of [O III] + Hβ emitters at $z \sim 1.4$

For [O III] + Hβ, we start by noting that based on our spectroscopic follow up, ∼16 per cent of the [O III] + Hβ emitters turned out to be Hβ and the rest to be [O III]. We also find that spectroscopically confirmed Hβ emitters are found to have lower luminosities than spectroscopically confirmed [O III] emitters. This, combined with the low fraction within the sample, would be a relatively good motivation towards ignoring such emitters (this is done by most studies; e.g. Ly et al. 2007). However, attempting to further split the Hβ and [O III] LFs, we take the following approach. We use the Sobral et al. (2013a) $z = 1.47$ Hα LF, and, with a simple assumption of $A_{\mathrm{H}\alpha} = 1$ mag and case B recombination, we predict the Hβ LF at $z \sim 1.44$ (also taking into account what flux we would recover with our NB filter). We find that the number densities of Hβ emitters are significantly below the number densities of [O III] + Hβ emitters. Given our depth, this simple prediction is very much in line with the 16 per cent fraction of Hβ found (and with Hβ emitters being preferentially found at fainter luminosities).

Due to the separation in wavelength, it is not possible for both Hβ and [O III] to contribute to the NB flux measured. We note, none the less, that there is a narrow range of redshifts where both lines from the [O III] doublet are detected (see one example of an emitter in Fig. 4 where both [O III] lines contribute to the NB flux) at opposite wings of the filter. Indeed, the [O III]4959 emission line is found to be even more important within the sample than Hβ and with important consequences for deriving LFs. While we find that [O III]5007 (only) dominates the sample of spectroscopically confirmed [O III] + Hβ emitters, representing ∼50 per cent of the full sample, [O III]4959 (only) represents 27 per cent (see e.g. Fig. 4), almost twice as common as Hβ and 7 per cent of the sources are actually detections of both [O III] lines at opposite wings of the filter (e.g. Fig. 4). Within the spectroscopic sample of [O III] + Hβ emitters, we study the fraction of each emitter as a function of luminosity and show the results in Fig. 10. Our analysis of the spectroscopically confirmed emitters shows that [O III]5007 emitters are distributed in a very Schechter like form, at all luminosities, Hβ are only found at lower luminosities, but [O III]4959 emitters, and simultaneous detections of [O III]5007 and [O III]4959 lines, are found both at the faintest luminosities (in the NB) and at the brightest luminosities, thus significantly boosting number counts at both the faintest luminosities (likely steepening the faint-end slope) and at the brightest luminosities (likely contributing to a non-Schechter form).

Our results are very important in order to interpret the results from the LF in Section 5.1.2 and for other similar studies (e.g. Ly et al. 2007; Khostovan et al. 2015). Thus, while we present a [O III] + Hβ LF, we caution that in order to properly derive LFs for each of the individual lines contributing to it, a much more detailed spectroscopic follow-up is needed, and that simply assuming that similar samples will be dominated by [O III] 5007 (ignoring, for example, the [O III] 4959 line, or simultaneous measurements of the doublet), will lead to strong systematic errors/biases. We therefore refer to these emitters as [O III] + Hβ throughout the paper. We also take our results into account when estimating the total volume probed by our survey for [O III] + Hβ emitters. Given that [O III] 5007 dominates

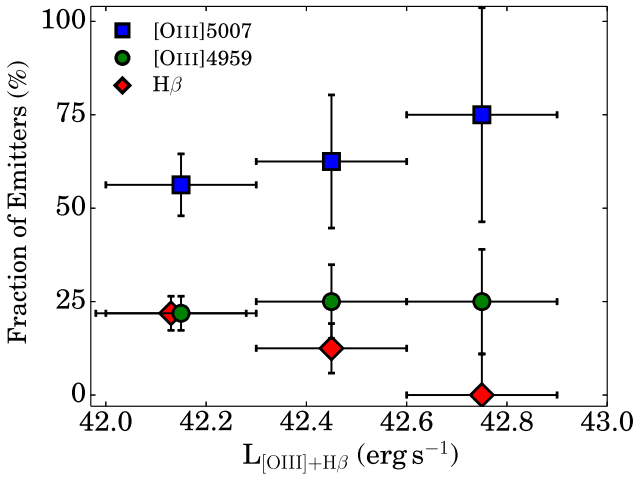

Figure 10. The fraction (within the spectroscopic sample) of [O III]5007, [O III]4959 and Hβ emitters as a function of [O III] + Hβ luminosity within the [O III] + Hβ sample of emitters at $z \sim 1.4$. We find that [O III]5007 emitters dominate at all luminosities, and that Hβ emitters only start to have a small contribution for lower luminosities.

the sample, we use the full volume probed with that emission line, but we correct for the expected addition of extra emitters (picked up over extra volumes). Thus, we also add to the total volume probed (because we are also sensitive to the other lines, and following our spectroscopic results) 16 per cent of full Hβ volume probed by us and 25 per cent of the volume probed for [O III] 4959 emitters.

4.3 Luminosity calculations and extinction corrections

In order to calculate LFs for our samples, line fluxes are converted to luminosities by applying

$$L_{\rm line} = 4\pi D_{\rm L}^2 {\rm F}_{\rm line}, \quad (9)$$

where $D_{\rm L}$ is the luminosity distance. We use $D_{\rm L}$=5367 Mpc for Hα emitters at $z = 0.81$, $D_{\rm L} = 9752.7$ Mpc for [O III]/Hβ emitters at $z = 1.4$ and $D_{\rm L}$=17 746.5 Mpc for [O II] emitters at $z = 2.2$.

We note that all our LFs are observed LFs, not dust-corrected. We only apply extinction corrections when converting luminosity density to SFRD. We also note that for the [O II] sample, we do not need to apply or investigate any correction, as no other line (close enough in rest-frame wavelength) is expected to contribute to the flux measured.

4.4 Completeness corrections

Fainter sources and those with weak emission lines might be missed and thus not included in the sample; this will result in the underestimation of the number of emitters, especially at lower luminosities. In order to account for that, we follow Sobral et al. (2013a) to estimate completeness corrections per subfield per emission line. Very briefly, we use sources which have not been selected as line emitters ($\Sigma < 3$ or EW<30 Å) and that have a photo-z within $\pm$0.5 of the appropriate line of interest (thus, we use different samples to estimate the completeness of different lines; see e.g. Sobral et al. 2012). We then add emission-line flux to all those sources, and study the recovery fraction as a function of input flux. We do these simulations in a subfield by subfield basis. We then apply those corrections in order to obtain our completeness-corrected LFs.

4.5 Filter profile corrections

The NB filter transmission function is not a perfect top-hat, so the real volume surveyed is a function of intrinsic luminosity. For example, luminous line emitters will be detectable over a larger volume than the fainter ones, as they can be detected in the wings of the filters (although they will be detected as fainter sources in these cases). Low-luminosity sources, however, will only be detectable in the central regions of the filter, leading to a smaller effective volume.

In order to correct for this when deriving the LFs, we follow Sobral et al. (2012). First, we compute the LF assuming a top-hat NB filter. We then generate a set of 10^{10} line emitters with a flux distribution given by the measured LF, but spread evenly over the redshift range being studied (assuming no cosmic structure variation or evolution of the LF over this narrow redshift range). We fold the fake line emitters through the top-hat filter model to confirm that we recover the input LF perfectly. Next, we fold the fake line emitters through the real NB profiles – their measured flux is not only a function of their real flux, but also of the transmission of the NB filter for their redshift. The simulations show that the number of brighter sources is underestimated relative to the fainter sources. A mean correction factor between the input LF and the one recovered (as a function of luminosity) was then used to correct each bin.

5 RESULTS

5.1 Luminosity functions

The estimate of the source density in a luminosity bin of width $\Delta(\log L)$ centred on $\log L_{\rm c}$ is given by the sum of the inverse volumes of all the sources in that bin, after correcting for completeness. The volume probed is calculated taking into account the survey area and the NB filter width, followed by applying the appropriate real filter profile corrections obtained in Section 4.5.

The LFs presented here are fitted with Schechter functions defined by the three parameters: α, ϕ^* and L^* (see Table 5). For the Hα LF, we can still get a reasonable constraint on α, but the data are too shallow to do the same for [O III] + Hβ and [O II], and thus for those we concentrate on just fitting for ϕ^* and L^* (and fixing α to the values commonly used in the literature for comparison). The final LFs are presented in Figs 11 and 12, and in Tables B1, B2 and B3. We note that we do not apply any dust correction prior to fitting the LFs – dust corrections are only applied when converting luminosity densities to estimates of the SFRD. The best-fitting Schechter functions are presented in Table 5.

5.1.1 Hα LF at z = 0.81

We present our Hα LF at $z = 0.81$ in Fig. 11 and Table B1. We confirm a strong evolution in the Hα LF from $z = 0$ to 0.8. We also find good agreement with both Sobral et al. (2013a) and Ly et al. (2011), as can be seen by directly comparing the data points and the best-fitting LF from e.g. Sobral et al. (2013a). However, as Fig. 11 clearly shows, we are able to not only extend the LF to higher luminosities, and to constrain the number densities of the most luminous Hα emitters with unprecedented accuracy, but also determine it with 3× more bins while reducing the Poissonian errors significantly at every bin. While the agreement is relatively good at all luminosities, we find a slightly lower L^* or slightly lower ϕ^*, in line with e.g. Colbert et al. (2013). Our results in Section 5.3 also show that we overcome cosmic variance, with the errors becoming a

Table 5. The LF and SFRD evolution for $z = 0.8$, 1.4, 2.2. L^*, ϕ^* and α values are derived without any correction for dust extinction. The measurements corrected for 1 mag extinction at Hα for Hα and [O III] + Hβ emitters (as our SFR calibration is based on observed Hα and [O III] line fluxes, thus the correction for Hα emitters is the most appropriate), but follow Hayashi et al. (2013) for [O II] selected emitters ($A_{\rm H\,\alpha} = 0.2$). Columns present the redshift, break of the LF, $L^*_{\rm line}$, normalization ($\phi^*_{\rm line}$) and faint-end slope (α) of the LFs. The two right-hand columns present the SFRD at each redshift based on integrating the LF down to each observational limit (41.4 for Hα, 42.1 for [O III] and 42.6 for [O II] in log L) and for a full integration, and include dust corrections. The two columns immediately to the left present similar measurements but for luminosity densities. SFRD include a correction for AGN contamination of 10 per cent at $z = 0.8$ for Hα (see Garn et al. 2010), up to L^*, and a correction given by $0.59 \times \log_{10}(L/L^*) + 0.112$ for $L > L^*$ (Sobral et al. 2015). For [O III] + Hβ at $z \sim 1.4$, we assume an AGN contamination of 15 per cent, while we assume a contamination of 20 per cent at $z = 2.2$ for [O II] emitters. We only fit α for the Hα LF and fix it for all the other LFs. L^* and ϕ^* are obtained by fixing $\alpha = -1.6$ (Hα), and the 1σ errors on L^* and ϕ^* are derived from such fits (with fixed α).

(z) Em. line ($z \pm 0.02$)	L^* (erg s^{-1})	ϕ^* (Mpc^{-3})	α	log ρ_L obs (erg s^{-1} Mpc^{-3})	log ρ_L (erg s^{-1} Mpc^{-3})	$\rho_{\rm SFR}$ obs (M$_\odot$ yr^{-1} Mpc^{-3})	$\rho_{\rm SFR}$ All (M$_\odot$ yr^{-1} Mpc^{-3})
($z = 0.81$) Hα	$41.72^{+0.03}_{-0.02}$	$-2.31^{+0.04}_{-0.05}$	$-1.6^{+0.2}_{-0.2}$	$38.17^{+0.03}_{-0.03}$	$40.14^{+0.01}_{-0.01}$	$0.022^{+0.002}_{-0.002}$	$0.086^{+0.003}_{-0.003}$
($z = 1.37$) [O III]	$42.10^{+0.05}_{-0.04}$	$-2.71^{+0.08}_{-0.09}$	-1.6 (Fixed)	$40.15^{+0.06}_{-0.02}$	$40.36^{+0.05}_{-0.04}$	$0.09^{+0.01}_{-0.01}$	$0.14^{+0.04}_{-0.04}$
($z = 2.20$) [O II]	$42.23^{+0.05}_{-0.04}$	$-2.23^{+0.14}_{-0.14}$	-0.9 (Fixed)	$39.01^{+0.01}_{-0.01}$	$40.02^{+0.10}_{-0.10}$	$0.11^{+0.02}_{-0.02}$	$0.26^{+0.03}_{-0.04}$

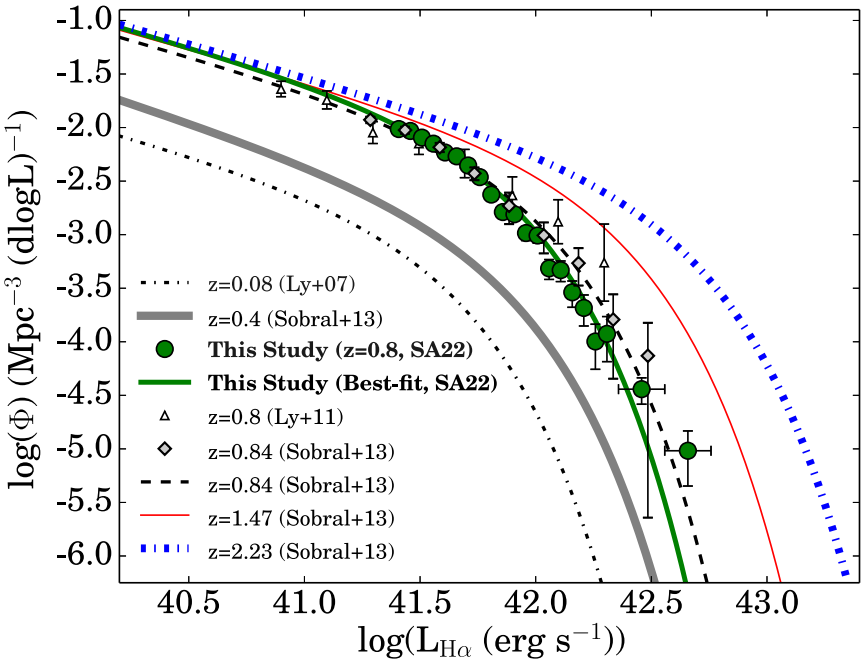

Figure 11. Observed Hα LF at $z = 0.81$ from our survey, comparison with previous surveys at a similar redshift, and evolution of the LF (none of the LFs have been corrected for dust extinction). We find a very good agreement between our results and those in the literature (well within the errors excepted from cosmic variance), but are able to extend the Hα LF to much higher luminosities, as well as reducing the Poissonian errors significantly. The errors due to cosmic variance are <10 per cent at most. Our results agree with the strong L^* evolution with increasing redshift.

few times smaller than all previous surveys, and that the differences with respect to Sobral et al. (2013a) and Ly et al. (2011) are fully explained by cosmic variance. Our results also confirm that the faint-end slope is relatively steep (we find $\alpha = -1.6 \pm 0.2$), in excellent agreement with both Sobral et al. (2013a) and Ly et al. (2011) and that L^* evolves significantly from $z \sim 0$ to $\sim$0.8. We also note that our spectroscopically confirmed subsample (301 spectroscopically confirmed Hα emitters in the SA22 sample) fully supports our results. Moreover, we find that the density of the most luminous emitters is still consistent with Sobral et al. (2013a), although, as shown in Sobral et al. (2015), the most luminous Hα emitters at $z \sim 0.8$ have a significant fraction of AGN, particularly broad-line AGN.

5.1.2 [O III] + Hβ LF at z ∼ 1.4

We present the observed (non-dust-corrected) [O III] + Hβ LF at $z \sim 1.4$ in Fig. 12 and in Table B2. We have fixed $\alpha = -1.6$ for the best fit. We find a clear luminosity evolution from $z \sim 0$ to $\sim$1.4 when compared to other LFs from the literature, in line with what is seen for Hα. Interestingly, if one uses empirical conversions from [O III] to Hα (e.g. Ly et al. 2007), [O III]/H$\alpha \approx 1.05$, and plot the Hα LF at $z \sim 1.4$, we find that it resembles the [O III] + Hβ LF surprisingly well. In other words, [O III] + Hβ emitters at $z \sim 1.4$ have similar number densities to Hα emitters at the same redshift. However, as noted in Section 4.2, the [O III] + Hβ LF is far from being simple/easy to interpret (as it is a composite of a few different lines, unlike Hα).

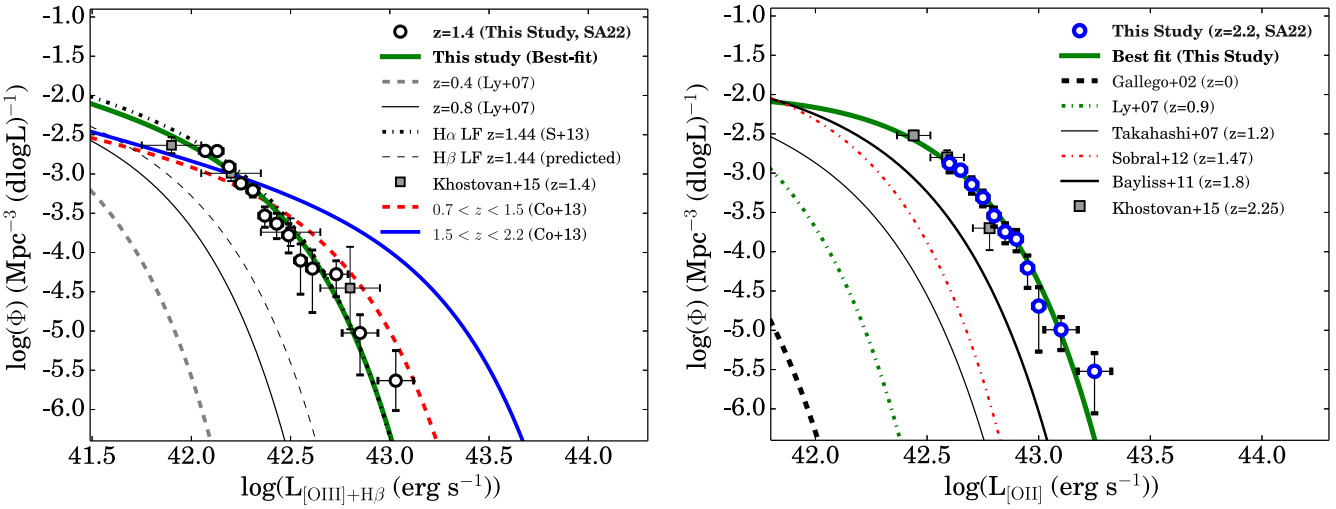

Figure 12. Left: observed [O III] + Hβ LF at $z \sim 1.4$ from our SA22 survey and comparison with other studies/surveys (including the LF from UDS+COSMOS derived by Khostovan et al. 2015). Our results clearly show that the [O III] + Hβ LF is evolving very strongly with redshift, mostly due to a strong increase in L^*, similarly to the behaviour of the Hα LF. Our best fit is very similar to the observed Hα LF at the same redshift (also shown), and, at the faint end, reveals a higher number of [O III] + Hβ emitters than in Colbert et al. (2013) – these are likely explained by a combination of Hβ emitters (we also show our predicted Hβ LF based on the Hα LF from Sobral et al. 2013a), but also due to the prevalence of the [O III] 4959 line (see Section 4.2). The deviations from a Schechter function are likely explained by the mix of different lines contributing to the final LF, as the observed LF is likely a combination of three different functions (for the three different emission lines). For higher luminosities, our data is in good agreement with the results from Colbert et al. (2013) and our slightly lower number densities are likely a result of accounting from the contribution of both [O III] 4959 and Hβ for the final value probed. Most importantly, we compare our results with Khostovan et al. (2015), who present the [O III] + Hβ LF from HiZELS in COSMOS+UDS, and find a very good agreement. Right: [O II] LF at $z = 2.2$ from our SA22 survey and comparison with other lower redshift surveys such as Gallego et al. (2002), Takahashi et al. (2007) and Bayliss et al. (2011). Similarly to the evolution of the Hα and [O III] + Hβ LFs, the [O II] LF shows a strong L^* evolution. Such increase in L^* with increasing redshift is even stronger by a few times than what is seen for other lines. On the other hand, there is little to no evidence of evolution in ϕ^*. We also compare our results from SA22 with those from UDS and COSMOS (HiZELS) presented by Khostovan et al. (2015) and find a very good agreement, at both the bright and faint ends.

Comparing with the largest *Hubble Space Telescope* slitless grism survey (Colbert et al. 2013), we find very good agreement at moderate to high luminosities, but we recover more line emitters at the faintest luminosities, as discussed earlier in Section 4.2. Based on our spectroscopic data, our [O III] + Hβ LF is likely a sum of three different LFs, all contributing in varying forms across the observed luminosities, thus showing a non-Schechter like form: [O III] 4959, [O III] 5007, [O III] 4959 + [O III] 5007 (where both lines are both being detected and measured by the NB filter) and Hβ.

5.1.3 [O II] LF at $z = 2.2$

We present the observed (non-dust-corrected) [O II] LF at $z = 2.2$ in Fig. 12 and in Table B3. This is the first time the [O II] LF has been measured at $z = 2.2$, and also constitutes by far the largest sample of [O II] emitters at $z \sim 2$. Since we do not constrain the faint-end slope, we fix it to the value measured at $z \sim 1.5$ by Ly et al. (2007) and Sobral et al. (2012), $\alpha = -0.9$.

We compare our results with results from the literature at lower redshifts: Gallego et al. (2002), Takahashi et al. (2007), Ly et al. (2007), Sobral et al. (2012) and Bayliss et al. (2011). Similarly to what has been found with Hα and [O III] + Hβ, we find a significant L^* evolution in the [O II] LF. However, for [O II] we find an even stronger/faster L^* evolution as L^* becomes ∼100 × brighter from the local Universe to $z = 2.2$. We also show data from a recent study by Khostovan et al. (2015), based on the HiZELS (COSMOS+UDS) data. We find excellent agreement when comparing to Khostovan et al. (2015), which probes to slightly lower luminosities, while we probe a much larger volume. Indeed, our fit to SA22 fits the results from Khostovan et al. (2015) well. However, even when combined, our data sets are still unable to constrain the faint-end slope, and thus it is possible that the faint-end slope is steeper than $\alpha = -0.9$.

5.2 Predictions for *Euclid* and *WFIRST*

We find that approximately ∼50 per cent of the emitters in our sample (for our full sample down to our flux limit) are likely Hα emitters at $z = 0.81$, while [O III] + Hβ account for ∼18 per cent and [O II] for about 9 per cent. The remaining ∼23 per cent consist of rarer emission lines (mostly at $z < 0.8$, but also at $z > 3$; see Matthee et al. 2014) and extra stars (with strong spectral features). If our flux limit was similar to the planned *Euclid* wide survey (Laureijs et al. 2012) of $>3 \times 10^{-16}$ erg s^{-1}, we would obtain a lower fraction (but still significant) of line emitters which are not Hα, [O III] + Hβ or [O II]: 18 per cent. Our *Euclid*-like sample is dominated by Hα emitters (67 per cent), followed by [O III] + Hβ emitters (12 per cent), but still with some [O II] emitters (3 per cent). The raw (observed) density of Hα emitters with $>3 \times 10^{-16}$ erg s^{-1} at $z \sim 0.8$ ($\lambda \sim 1.2$ μm) is 8.66×10^{-5} Mpc^{-3}, while at the same wavelength the observed number density of [O III] + Hβ emitters is about five times lower (1.6×10^{-5} Mpc^{-3}), and the number density of [O II] emitters (0.3×10^{-5} Mpc^{-3}) is almost 30 times lower than Hα at $\lambda \sim 1.2$ μm.

Using spectroscopic redshifts from the literature and from our own follow-up, we can have an even more robust quantification of the range of line emitters in the full sample. Fig. 3 shows the full distribution of spectroscopic redshifts. Fig. 3 also shows the different spectroscopic redshift subsamples obtained when restricting our sample to a much higher NB colour significance (Σ). For $\Sigma > 3$ (our full sample), and based solely on spectroscopic redshifts, we find 59 per cent Hα emitters, 9 per cent [O III] + Hβ emitters, 7 per cent

[S II] emitters, 4 per cent [O II] emitters, while 16 per cent are other $z < 1.0$ (including [N II]) and the remaining 5 per cent are other $z > 1$ emitters. Subsamples with higher Σ will be samples with higher flux limits, thus changing the distribution of emitters within such subsamples. For our sample, $\Sigma > 4$ corresponds to a cut in flux of 1.3×10^{-16} erg s^{-1} cm^{-2}, while the cut at $\Sigma > 5$ corresponds to cutting the flux down to a flux limit of 1.55×10^{-16} erg s^{-1} cm^{-2}. For $\Sigma > 4$ ($\Sigma > 5$), based on spectroscopic redshift, we find 74 per cent (81 per cent) Hα emitters, 9 per cent (7 per cent) [O III] + Hβ emitters, 6 per cent (4 per cent) [S II] emitters and 3 per cent (3 per cent) [O II] emitters, while 6 per cent (3 per cent) are other $z < 1.0$ (including [N II]) and the remaining 5 per cent (2 per cent) are other $z > 1$ emitters. It is therefore clear that a selection with a higher flux limit will be more and more dominated by Hα emitters, and thus our spectroscopic sample shows that both *Euclid* and *WFIRST* can expect ∼80 per cent of all line emitters to be Hα.

5.3 Luminosity density and cosmic SFRD

By using the largest samples ever assembled and empirically computed uncertainties due to cosmic variance, we integrate our observed LFs. We provide two different measurements for each emission line: (i) the full analytical integral and (ii) the numerical integration down to the observational limit. We provide the observed luminosity density values in Table 5.

Prior to converting our observed luminosity densities for each line to a SFRD ($\rho_{\rm SFR}$), we apply two corrections. First, we do a simple correction for dust extinction. We assume 1 mag at Hα ($A_{\rm H\alpha} = 1$). Dust corrections for [O III] + Hβ and [O II] are discussed below. Secondly, we assume a 10–20 per cent AGN contribution to our samples, following our results in Section 3.7.

In order to convert luminosity densities to SFRD we use, for Hα (Kennicutt 1998),

$$\mathrm{SFR}(\mathrm{M}_\odot\ \mathrm{yr}^{-1}) = 7.9 \times 10^{-42}\ L_{\rm H\alpha}(\mathrm{erg\,s}^{-1}). \quad (10)$$

For [O III] + Hβ, as seen before, directly interpreting as a star formation indicator has a significant number of caveats. However, cautioning the reader that any values need to be interpreted with caution and are mostly indicative, we use the relation derived from Osterbrock & Ferland (2006), in good agreement with the empirical calibration by Ly et al. (2007), who found typical line ratios of ∼1,

$$\mathrm{SFR}(\mathrm{M}_\odot\ \mathrm{yr}^{-1}) = 7.35 \times 10^{-42}\ L_{\rm [O\,III]+H\,\beta}(\mathrm{erg\,s}^{-1}). \quad (11)$$

For [O II], we use the standard Kennicutt (1998) calibration of [O II] as a star formation tracer (calibrated using Hα):

$$\mathrm{SFR}(\mathrm{M}_\odot\ \mathrm{yr}^{-1}) = 1.4 \times 10^{-41}\ L_{\rm [O\,II]}\ (\mathrm{erg\,s}^{-1}). \quad (12)$$

For [O II], in order to correct for dust extinction, we use the results from Hayashi et al. (2013) that present a large study of [O II] selected emitters at $z \sim 1.47$ and use Hα measurements for all of them (including stacking in Hα for the faintest [O II] emitters) to calibrate [O II]. Hayashi et al. (2013) shows that [O II] selected samples at much lower luminosities than ours have typical dust extinctions of $A_{\rm H\alpha} \approx 0.4$, that their median extinction drops with increasing [O II] luminosity, and that assuming $A_{\rm H\alpha} = 1.0$ for such emitters results in a significant overestimation of the real [O II] luminosity. As a simple test of this conclusion (and its validity at $z = 2.2$), we start by assuming $A_{\rm H\alpha} = 1.0$ for our sample of [O II] emitters at $z = 2.2$, and find that we obtain $\rho_{\rm SFR} = 1.0 \pm 0.3$ M$_\odot$ yr^{-1} Mpc^{-3}, about four to five times higher than that given by Hα (Sobral et al. 2013a). This suggests that the results presented by Hayashi et al. (2013) are also valid for $z = 2.2$. Indeed, if we use their results and instead extinction-correct the [O II] luminosities assuming $A_{\rm H\alpha} \approx 0.2$ (appropriate for our luminosity; this corresponds to $A_{\rm [O\,II]} \approx 0.4$), we obtain a good agreement with the value from Hα at the same redshift. However, we note that the uncertain faint-end slope of the [O II] LF at high redshift also plays a role. At lower redshift, it is found to be relatively shallow (α approximately in the range from −0.8 to −1.2), and thus here we fix it to $\alpha = -0.9$, as this is the value found by the best studies at $z \sim 1.5$ (Ly et al. 2007; Sobral et al. 2012).

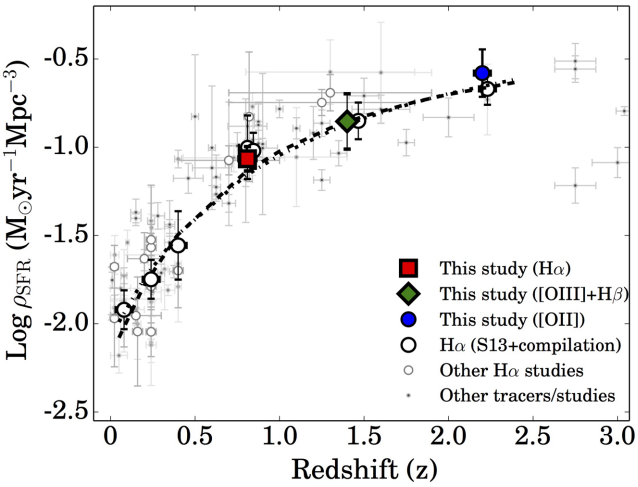

Figure 13. Star formation history of the Universe using our CF-HiZELS SA22 data: $z = 0.81$ (Hα), $z = 1.4$ ([O III] + Hβ) and $z = 2.2$ ([O II]). Our results agree very well with the Hα star formation history of the Universe (dashed line is the parametrization from Sobral et al. 2013a).

While e.g. Ly et al. (2007) explored a double-blind [O III]-Hα survey (and found typical line ratios around ∼1), no systematic investigation was performed in order to investigate the typical dust properties of [O III] + Hβ emitters at $z > 1$. However, as we have discussed, [O II]-selected emitters are significantly less dusty than Hα selected line emitters (as an [O II]-selection preferentially picks up dust-free sources and is biased against dusty sources). This is likely to be the case for [O III] + Hβ emitters, although we expect them to present typical dust extinction properties which are more similar to Hα-selected line emitters. However, given that we are using an SFR calibration based directly on observed line fluxes which used the Hα line (Ly et al. 2007), we will use the correction appropriate for Hα-selected emitters, $A_{\rm H\alpha} = 1$, corresponding to about 1.5 mag at [O III] + Hβ. We obtain $\rho_{\rm SFR} = 0.14 \pm 0.03$ M$_\odot$ yr^{-1} Mpc^{-3}.

Our results from Hα, [O III] and [O II] can be found in both Table 5 and in Fig. 13. We find a very good agreement with the overall star formation history of the Universe (e.g. Karim et al. 2011; Sobral et al. 2013a), indicating that our simple assumptions for dust and AGN correction (or the combination of both) provide decent average estimates for the full population of galaxies (although it obviously fails on a source by source basis, and it will have strong dependences with mass, luminosity/etc – see e.g. Ibar et al. 2013; Sobral et al. 2014).

Our results confirm the rise of the star formation history up to $z = 2.2$, but also indicate how important it is to apply appropriate dust extinction corrections (Hayashi et al. 2013) that take into account that [O II] emitters are typically almost dust free. While there are significant uncertainties in the use of [O II] as a star

formation indicator for higher redshifts where Hα is not available, if a better understanding of the typical dust properties is obtained, and particularly if the faint-end slope is constrained, it should provide at least a competitive view when compared to e.g. the UV.

5.4 Overdensities

We explore our very wide area coverage over different fields to look for significant overdensities. In order to do this, we take two approaches: we smooth the distribution of sources, but also compute the tenth nearest neighbour densities. As can be seen in Fig. 14, there appears to be a significant large-scale overdensity of Hα emitters which contains ∼ 300 candidate $z = 0.81$ Hα line emitters within an ∼ 20 arcmin field (Fig. 14). This includes a region where the number density of Hα emitters is ∼10 times higher than the general field. This strong overdensity of Hα emitters in the SA22 field has now been confirmed with the new KMOS instrument (Sobral et al. 2013b). This ∼8σ overdensity of Hα emitters is found within a volume of 3000 Mpc^3 (comoving).

We note that this finding could be important to also interpret any overdensity of star-forming galaxies detected at this or at even higher redshift. Without a careful analysis of the full galaxy population, one could be misled to conclude that such structure could be a cluster and/or a protocluster. However, following Sobral et al. (2011), Sobral et al. (2013b) find that the structure presents galaxy number densities which are more typical of groups, or intermediate densities, but not clusters. In practice, the structure is likely a dense cosmic web structure or filamentary/wall-like as a group. Darvish et al. (2014) find that the fraction of star-forming galaxies (Hα emitters) is much higher in filaments than in the general field or in clusters, making these easily detectable with wide Hα surveys.

We also search for overdensities within the samples of [O III] and [O II] emitters, but find only mild overdensities (<3σ). Nevertheless, these density fluctuations still result in significant cosmic variance, if only different parts of the fields would be investigated.

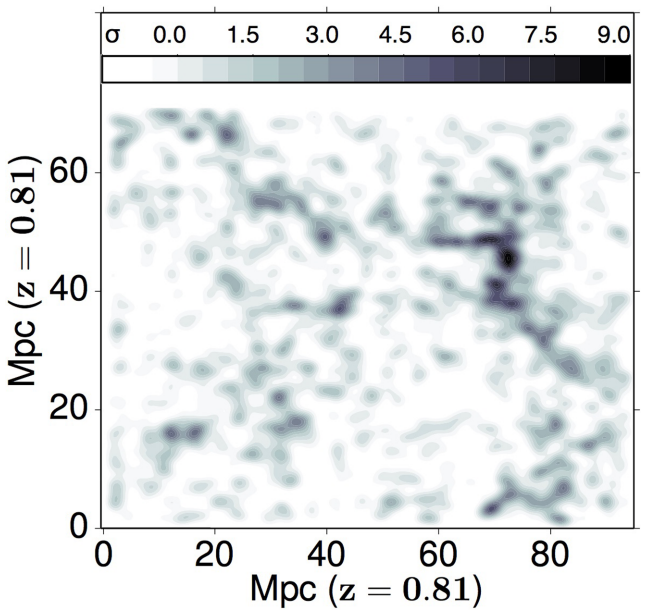

Figure 14. Density of Hα emitters at $z = 0.81$ in our SA22 survey. The angular scale is converted to a physical scale. The smoothed distribution highlights the large 7σ–9σ overdensity, which is approximately 15 Mpc across. It also shows the filamentary structure of the large-scale structure in the Universe.

5.5 Cosmic variance and cosmic convergence

We explore the large samples spread over large, multiple areas (SA22, COSMOS, UDS) to empirically quantify cosmic variance affecting NB surveys. In order to do this, we follow a similar procedure to Sobral et al. (2011), and divide the full sample in several areas corresponding to individual pointing/cameras, roughly corresponding to 0.02–0.05 deg^2. We then repeat the determination of the LFs by randomly selecting these areas. We start by computing them for the smallest, contiguous 0.02 deg^2 areas and then go up by steps of 0.02, sampling these in random combinations within SA22, COSMOS and UDS. We first investigate how α, ϕ^* and L^* are affected by sample variance, but find that, due to the depth of our data, α is in general completely unconstrained. Thus, for the following analysis, we fix α and concentrate on studying ϕ^* and L^* only. For each emission line, we obtain a total of a million realizations of the LF, spread over the minimum to the maximum total area. Quoted errors (in per cent) are computed as the ratio between the standard deviation of a parameter for each bin in area (from all LF realizations with a given total area) and the actual parameter value.

We note that our results can have a relatively general application for data sets with similar depths, but that for data sets that are e.g. significantly shallower, the errors measured with our method would become much larger, as they would be dominated by e.g. Poissonion errors and, in such cases (e.g. if depth is shallower than L^*), ϕ^* and L^* become unconstrained, and thus the results of our study no longer apply (in order to lower the errors in the parameters, deeper data is more efficient than probing larger areas). We also note that for [O III] + Hβ and [O II] LFs, due to the depth of our data, we only probe down to ∼L^* and thus, for the lowest volumes, our errors become completely Poissonion-dominated, resulting in very large variance which is due to low number statistics, not cosmic variance. Thus, we do not show the results for areas which are so small that low number statistics completely dominate.

The results are presented in Fig. 15. As expected, we find that cosmic/sample variance is very strong for small volumes (<10^5 Mpc^3), and strongly reduced with increasing area/volume. For the Hα LF, for example, a total area of ≈1 deg^2 (≈10^5 Mpc^3) still results in significant scatter/significant errors due to sample/cosmic variance, where L^* is clearly the most affected parameter (up to >100 per cent errors), followed by ϕ^* (∼40–50 per cent errors).

With our total area of ∼9 deg^2 the errors on ϕ^* are reduced to ∼10–15 per cent, while the errors on L^* drop to ∼8 per cent and the error on SFRD drops to close to 5 per cent. We therefore conclude that only by probing volumes comparable or larger than 10^6 Mpc^3 will the sample variance fully drop below 10 per cent. Our results are also very useful to interpret results from similar previous, current and future surveys of star-forming galaxies, particularly to interpret differences in number counts and LFs, without having to rely on simulations and on other indirect methods/arguments.

We find similar results for both [O II] and [O III] + Hβ. We note that we use our results to obtain a better estimate of our errors by adding (in quadrature) the expected errors due to cosmic variance, based on the total volume probed by each LF/measurement. This is a small contribution to the total error budget given that our survey probes a very large volume, but would be a significant amount to surveys probing small volumes.

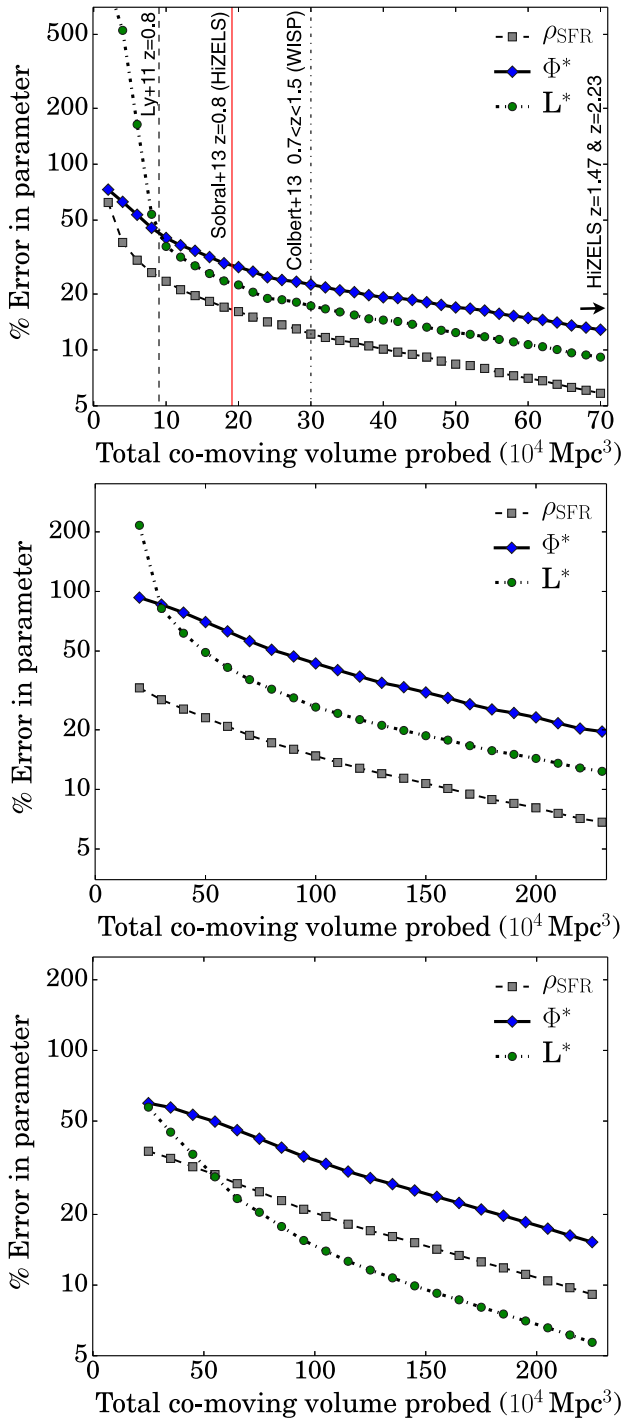

Figure 15. Top: the error/variance on ϕ^*, L^* and SFRD (ρ_{SFR}) as function of total probed comoving volume, based on a million realizations, each sampling a random combination of UDS, COSMOS and SA22 for Hα LF at $z \sim 0.8$. The results show that the variance on all parameters/quantities decreases with total comoving volume probed, but that it is significant for most volumes usually used in the literature. Our results are therefore expected to have an error <10 per cent. Middle: the error/variance on ϕ^*, L^* and SFRD (ρ_{SFR}) as function of total probed comoving volume, based on a million realizations, each sampling a random combination of UDS, COSMOS and SA22 for [O III] + Hβ LF at $z \sim 1.4$. Bottom: the error/variance on ϕ^*, L^* and SFRD (ρ_{SFR}) as function of total probed comoving volume, based on a million realizations, each sampling a random combination of UDS, COSMOS and SA22 for [O II] LF at $z \sim 2.2$.

6 CONCLUSIONS

We presented results from the largest contiguous NB survey in the near-infrared (*J* band). We surveyed ≈10 deg^2 of contiguous extragalactic sky and found a total of 5976 robust candidate emission-line galaxies ($\Sigma > 3$). By using deep CFHTLS *ugriz* and UKIDSS DXS *J* and *K*, a large sample of available spectroscopic redshifts from VVDS and VIPERS from the literature, and by obtaining new spectroscopic follow-up observations with MOSFIRE and FMOS, and by combining our sample with HiZELS, we derive the largest samples of emission-line selected galaxies. We find by far the largest sample of Hα emitters at $z \sim 0.8$: 3471 sources (∼400 spectroscopically confirmed), obtaining the most accurate measurement of the SFRD at that cosmic epoch. We also present the largest sample (1341) of [O III]/Hβ emitters at $z \sim 1.4$ and present the first large sample (572 sources) of [O II] emitters at the peak of the star formation history ($z = 2.2$). Our main conclusions are as follows.

(i) Our large spectroscopic sample from FMOS and MOSFIRE allows us to confirm that the [N II]/Hα correction as a function of EW(Hα + [N II]) from SDSS is applicable to $z = 0.81$ with no evolution. We none the less provide a simpler linear fit, based on our data, that can be used to correct similar data sets: $f([\mathrm{N\,II}]/\mathrm{H}\,\alpha) = -0.296 \times \log_{10}(\mathrm{EW}_{\mathrm{H}\,\alpha+[\mathrm{N\,II}]}) + 0.8$. Our spectroscopic H$\alpha$ sample also shows that our Hα emitters have a metallicity of $12 + \log_{10}(\mathrm{O/H}) = 8.56 \pm 0.05$, slightly subsolar. We also make significant detections of the [S II] doublet, with a line ratio of I([S II]6716)/I([S II]6731) = 1.33 ± 0.08, implying an electron density of 40–200 cm^{-3}. The [S II]6716/I(Hα)=0.14 ± 0.02 ratio also implies ionization potential of $\log_{10} U = -3.9 \pm 0.5$ cm^{-2}.

(ii) We obtain Hα, [O III] + Hβ and [O II] LFs at $z = 0.8$, 1.4, 2.2 with the largest statistical samples, reaching up to the highest luminosities. We find a strong luminosity evolution (L^*) in the LF of all the lines with increasing redshift up to at least $z \sim 2.2$, with a less significant (but present) ϕ^* evolution. This is consistent with the evolution seen across redshift presented by Khostovan et al. (2015).

(iii) We show that the [O III] + Hβ LF at $z = 1.4$ is very hard to interpret in general, particularly due to the complicated contribution from the two different [O III] lines, and, to a lesser extent, to the contribution of Hβ. While we find that [O III] 5007 (only) dominates the sample of spectroscopically confirmed [O III] + Hβ emitters, representing ∼50 per cent of the full sample, [O III] 4959 (only) represents 27 per cent, almost twice as common as Hβ. We find that 7 per cent of the [O III] + Hβ sources are actually detections of both [O III] lines at opposite wings of the filter.

(iv) We present the first [O II] LF at $z = 2.2$, and find a very strong evolution from $z \sim 0$ to ∼2.2, much stronger than that seen for either Hα or [O III] + Hβ. By correcting for dust extinction using Hayashi et al. (2013), the SFRD based on [O II] is in excellent agreement with Hα (Sobral et al. 2013a). If 1 mag of extinction at Hα was used for [O II] emitters instead, the SFRD would have been overestimated by a factor of 2–3. Thus, if the Hayashi et al. (2013) calibration is used (and if it remains valid for even higher redshift), [O II] may be a reasonably good way to measure SFRD beyond $z \sim 2.2$, as Hα gets redshifted out of the *K* band.

(v) We find a reasonable good agreement between the SFRD from [O III] and that from Hα at a similar redshift. We none the less caution that without a detailed investigation into the nature of [O III] + Hβ emitters (and without robustly separating them from Hβ emitters, interesting on their own), using [O III] + Hβ as a star formation indicator at high redshift is highly unreliable.

(vi) For the planned *Euclid* wide survey (Laureijs et al. 2012) flux limit of $>3 \times 10^{-16}$ erg s^{-1}, our sample is dominated by Hα emitters (67 per cent), followed by [O III] + Hβ emitters (12 per cent), but still with some [O II] emitters (3 per cent); the remaining 18 per cent are rarer emitters spread over a wide range of redshifts. The raw (observed) density of Hα emitters with $>3 \times 10^{-16}$ erg s^{-1} at $z \sim 0.8$ ($\lambda \sim 1.2$ μm) is 8.66×10^{-5} Mpc^{-3}, while at the same wavelength the observed number density of [O III] + Hβ emitters is about five times lower (1.6×10^{-5} Mpc^{-3}), and the number density of [O II] emitters (0.3×10^{-5} Mpc^{-3}) is almost 30 times lower than Hα at $\lambda \sim 1.2$ μm. Our fully spectroscopically confirmed sample confirms these numbers, predicting that ∼70–80 per cent of all line emitters found at $\lambda \sim 1.2$ μm will be Hα.

(vii) We find significant overdensities, out of which the strongest one is found at $z = 0.8$ and traced by Hα emitters. It is an 8.5σ (confirmed with KMOS) overdensity of Hα emitters, where the number density is a factor ∼10 × higher than the average; this is consistent with group-like densities, and most likely a rich filamentary structure, similar to what has been studied in Darvish et al. (2014). Only mild overdensities are found for [O III] + Hβ and [O II] emitters, although such overdensities would be much harder to find given the relatively shallow data when compared to Hα.

(viii) We take advantage of the large volumes/area and multiple fields to subdivide the samples in randomized areas and provide a robust empirical measurement of sample/cosmic variance for the different lines/redshifts. We find that surveys for star-forming/emission-line galaxies can only overcome cosmic variance (errors <10 per cent) if they are based on volumes $>5 \times 10^5$ Mpc^3. In other words, multiple/different volumes adding up to $<5 \times 10^5$ Mpc^3 show variance which results in errors being >10 per cent.

(ix) Errors due to sample (cosmic) variance on surveys probing ∼10^4 and ∼10^5 Mpc^3 are typically very high: ∼300 and ∼40–60 per cent, respectively. Focusing only on L^* and ϕ^*, the latter is the most affected parameter for large volumes, while L^* is completely undetermined for volumes $<10^5$ Mpc^3.

ACKNOWLEDGEMENTS

The authors wish to thank the anonymous reviewer for many helpful comments and suggestions which greatly improved the clarity and quality of this work. DS acknowledges financial support from the Netherlands Organization for Scientific research (NWO) through a Veni fellowship, from FCT through an FCT Investigator Starting Grant and Start-up Grant (IF/01154/2012/CP0189/CT0010), from FCT grant PEst-OE/FIS/UI2751/2014, and from LSF and LKBF. JM acknowledges the award of a Huygens PhD fellowship. PNB is grateful for support from STFC. IRS acknowledges support from STFC, a Leverhulme Fellowship, the ERC Advanced Investigator programme DUSTYGAL and a Royal Society/Wolfson Merit Award. BMJ acknowledges support from the ERC-StG grant EGGS-278202. The Dark Cosmology Centre is funded by the DNRF. The Dark Cosmology Centre is funded by the DNRF. JWK acknowledges support from the National Research Foundation of Korea (NRF) grant, no. 2008-0060544, funded by the Korea government (MSIP). JPS acknowledges support from STFC (ST/I001573/1). JC acknowledges support from the FCT-IF grant IF/01154/2012/CP0189/CT0010. The work was only possible due to OPTICON/FP7 and the invaluable access that it granted to the CFHT telescope. We would also like to acknowledge the excellent work done by CFHT staff in conducting the observations in service mode, and on delivering truly excellent data. We are also tremendously thankful to Kentaro Aoki for the incredible support while observing at Subaru with FMOS, and also to the Keck staff for the help with the observations with MOSFIRE. This work is based on observations obtained with WIRCam on the CFHT, OPTICON programme 2011B/029, 2012A019 and 2012B/016. Based on observations made with ESO telescopes at the La Silla Paranal Observatory under programmes IDs 60.A-9460 (data can be accessed through the ESO data archive), 087.A-0337 and 089.A-0965. Based on observations done with FMOS on Subaru under programme S14A-084, and on MOSFIRE/Keck observations under programme U066M. Part of the data on which this analysis is based are available from Sobral et al. (2013a). Dedicated to the memory of C. M. Sobral (1953-2014).

REFERENCES

An F. X. et al., 2014, ApJ, 784, 152
Bayliss K. D., McMahon R. G., Venemans B. P., Ryan-Weber E. V., Lewis J. R., 2011, MNRAS, 413, 2883
Bertin E., 2006, in Gabriel C., Arviset C., Ponz D., Enrique S., eds, ASP Conf. Ser. Vol. 351, Astronomical Data Analysis Software and Systems XV. Astron. Soc. Pac., San Francisco, p. 112
Bertin E., 2010, Astrophysics Source Code Library, record ascl:1010.068
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Bunker A. J., Warren S. J., Hewett P. C., Clements D. L., 1995, MNRAS, 273, 513
Capak P. et al., 2007, ApJS, 172, 99
Colbert J. W. et al., 2013, ApJ, 779, 34
Darvish B., Sobral D., Mobasher B., Scoville N. Z., Best P., Sales L. V., Smail I., 2014, ApJ, 796, 51
Drake A. B. et al., 2013, MNRAS, 433, 796
Elvis M. et al., 2009, ApJS, 184, 158
Fumagalli M. et al., 2012, ApJ, 757, L22
Gallego J., García-Dabó C. E., Zamorano J., Aragón-Salamanca A., Rego M., 2002, ApJ, 570, L1
Garilli B. et al., 2014, A&A, 562, A23
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
Geach J. E. et al., 2010, MNRAS, 402, 1330
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
Hayashi M., Sobral D., Best P. N., Smail I., Kodama T., 2013, MNRAS, 430, 1042
Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142
Ibar E. et al., 2013, MNRAS, 434, 3218
Ilbert O. et al., 2009, ApJ, 690, 1236
Ilbert O. et al., 2010, ApJ, 709, 644
Ilbert O. et al., 2013, A&A, 556, A55
Iwamuro F. et al., 2012, PASJ, 64, 59
Karim A. et al., 2011, ApJ, 730, 61
Kennicutt R. C., Jr, 1998, ARA&A, 36, 189
Kennicutt R. C., Jr, et al., 2009, ApJ, 703, 1672
Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, preprint (arXiv:1503.00004)
Kimura M. et al., 2010, PASJ, 62, 1135
Koyama Y., Kodama T., Nakata F., Shimasaku K., Okamura S., 2011, ApJ, 734, 66
Koyama Y. et al., 2013, MNRAS, 434, 423
Laureijs R. et al., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann J. M., eds, Proc. SPIE Conf. Ser. Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 84420T
Lawrence A. et al., 2007, MNRAS, 379, 1599
Le Fèvre O. et al., 2013, A&A, 559, A14

Lee J. C. et al., 2012, PASP, 124, 782
Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJ, 460, L1
Lilly S. J. et al., 2009, ApJS, 184, 218
Ly C. et al., 2007, ApJ, 657, 738
Ly C., Lee J. C., Dale D. A., Momcheva I., Salim S., Staudaher S., Moore C. A., Finn R., 2011, ApJ, 726, 109
McCracken H. J. et al., 2012, A&A, 544, A156
McLean I. S. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 7735, Ground-Based and Airborne Instrumentation for Astronomy III. SPIE, Bellingham, p. 1
McLean I. S. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-Based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460J
Matthee J. J. A. et al., 2014, MNRAS, 440, 2375
Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, preprint (arXiv:1502.07355)
Milvang-Jensen B. et al., 2013, A&A, 560, A94
Osterbrock D. E., 1989, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Mill Valley, CA
Osterbrock D. E., Ferland G. J., 2006, Astrophysics of Gaseous Nebulae and Active Galactic Nuclei. University Science Books, Mill Valley, CA
Oteo I., Sobral D., Ivison R., Smail I., Best P. N., Cepa N., Perez G., 2015, MNRAS, submitted
Pence W., 1999, in Mehringer D. M., Plante R. L., Roberts D. A., eds, ASP Conf. Ser. Vol. 172, Astronomical Data Analysis Software and Systems VIII. Astron. Soc. Pac., San Francisco, p. 487
Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59
Puget P. et al., 2004, in Moorwood A. F. M., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 5492, Ground-based Instrumentation for Astronomy. SPIE, Bellingham, p. 978
Scoville N. et al., 2007, ApJS, 172, 1
Sharples R. et al., 2013, The Messenger, 151, 21
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Sobral D. et al., 2009a, MNRAS, 398, L68
Sobral D. et al., 2009b, 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., Smail I., Geach J. E., Cirasuolo M., Garn T., Dalton G. B., 2011, MNRAS, 411, 675
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., 2013a, MNRAS, 428, 1128
Sobral D. et al., 2013b, ApJ, 779, 139
Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 437, 3516
Sobral D., Kohn S., Best P. N., Smail I., Stott J., Calhau J., Matthee J., 2015, MNRAS, submitted
Stott J., Sobral D., Smail I., Bower R., Best P. N., Geach J. E., 2013a, MNRAS, 430, 1158
Stott J. et al., 2013b, MNRAS, 436, 1130
Stott J. P. et al., 2014, MNRAS, 443, 2695
Stroe A., Sobral D., Röttgering H. J. A., van Weeren R. J., 2014, MNRAS, 438, 1377
Stroe A. et al., 2015, MNRAS, 450, 646
Swinbank A. M., Sobral D., Smail I., Geach J. E., Best P. N., McCarthy I. G., Crain R. A., Theuns T., 2012a, MNRAS, 426, 935
Swinbank A. M., Smail I., Sobral D., Theuns T., Best P. N., Geach J. E., 2012b, ApJ, 760, 130
Tadaki K.-I., Kodama T., Koyama Y., Hayashi M., Tanaka I., Tokoku C., 2011, PASJ, 63, 437
Takahashi M. I. et al., 2007, ApJS, 172, 456
Villar V., Gallego J., Pérez-González P. G., Pascual S., Noeske K., Koo D. C., Barro G., Zamorano J., 2008, ApJ, 677, 169
Wang Y., Chuang C.-H., Hirata C. M., 2013, MNRAS, 430, 2446

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Table A1. The catalogue of all $\Sigma > 3$ NB sources selected in the SA22 field from this paper (http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stv1076/-/DC1).

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.

APPENDIX A: CATALOGUE OF CF-HIZELS NB EMITTERS

The catalogue of NB emitters in SA22 is presented in Table A1. It contains IDs, right ascension (RA, J2000), declination (Dec., J2000), NB magnitude (NB, AB), BB magnitude (*J*, AB), the significance of the NB excess (Σ, estimated in 2 arcsec apertures), estimated flux ($\log_{10}$), estimated observed EW (Å), and a flag for those that are classified as Hα at $z = 0.8$ (1), [O III] or Hβ at $z \sim 1.4$ (2) and [O II] at $z = 2.20$ (3). Unclassified sources as flagged with 0 and candidate stars are identified with flag -1. We note that there are a few (15) sources with some extreme EWs (observed EWs $>10^4$ Å [observed]) and 17 sources with $\Sigma > 100$, but that these may be supernovae and/or strongly variable sources (see e.g. Matthee et al. 2015), apart from real extremely rare sources. Note that only the online version contains the full catalogue – here only three entries of the table are shown as examples of the entire catalogue.

Table A1. Example entries form the catalogue of all $\Sigma > 3$ NB sources selected in the SA22 field from this paper. The full catalogue is available online.

ID	RA (J2000)	Dec. (J2000)	NB_J (AB)	J (AB)	Σ	log flux (erg s^{-1} cm^{-2})	EW_{obs} (Å)	Class. flag
CFHIZELS-SM14-99905	22 09 11.64	+01 23 13.25	20.43 ± 0.06	21.40 ± 0.06	10.1	−15.48	170.8	1
CFHIZELS-SM14-98588	22 19 34.55	+00 24 51.36	20.41 ± 0.06	22.24 ± 0.11	13.2	−15.33	657.2	2
CFHIZELS-SM14-97547	22 11 31.53	−00 37 16.69	20.40 ± 0.07	20.82 ± 0.05	5.1	−15.72	52.4	3

APPENDIX B: LUMINOSITY FUNCTIONS

Table B1. Hα LF at $z = 0.81$ from our SA22 survey. Note that $L_{\mathrm{H}\alpha}$ has been corrected for [N II] contamination, but not for dust extinction. The total volume per bin is 8.6×10^5 Mpc^3, assuming a top-hat filter, but ϕ corr has been corrected for both incompleteness and the fact that the filter profile is not a perfect top hat.

log $L_{\mathrm{H}\alpha}$ ($z = 0.81$)	Sources (#)	ϕ obs (Mpc^{-3})	ϕ corr (Mpc^{-3})
41.40 ± 0.025	445 ± 21	-1.98 ± 0.02	-2.01 ± 0.02
41.45 ± 0.025	433 ± 20	-1.99 ± 0.02	-2.03 ± 0.02
41.50 ± 0.025	370 ± 19	-2.06 ± 0.02	-2.09 ± 0.02
41.55 ± 0.025	322 ± 17	-2.12 ± 0.02	-2.15 ± 0.03
41.60 ± 0.025	268 ± 16	-2.20 ± 0.03	-2.23 ± 0.03
41.65 ± 0.025	248 ± 15	-2.24 ± 0.03	-2.27 ± 0.03
41.70 ± 0.025	201 ± 14	-2.33 ± 0.03	-2.35 ± 0.03
41.75 ± 0.025	159 ± 12	-2.43 ± 0.04	-2.46 ± 0.04
41.80 ± 0.025	109 ± 10	-2.59 ± 0.04	-2.63 ± 0.05
41.85 ± 0.025	76 ± 8	-2.75 ± 0.05	-2.79 ± 0.06
41.90 ± 0.025	73 ± 8	-2.77 ± 0.05	-2.81 ± 0.06
41.95 ± 0.025	47 ± 6	-2.95 ± 0.07	-2.98 ± 0.07
42.00 ± 0.025	46 ± 6	-2.96 ± 0.07	-3.01 ± 0.07
42.05 ± 0.025	23 ± 4	-3.26 ± 0.10	-3.31 ± 0.10
42.10 ± 0.025	21 ± 4	-3.29 ± 0.10	-3.33 ± 0.11
42.15 ± 0.025	13 ± 3	-3.50 ± 0.14	-3.54 ± 0.14
42.20 ± 0.025	9 ± 3	-3.64 ± 0.17	-3.68 ± 0.17
42.25 ± 0.025	5 ± 2	-3.93 ± 0.26	-4.00 ± 0.26
42.30 ± 0.025	5 ± 2	-3.93 ± 0.26	-3.93 ± 0.26
42.45 ± 0.025	13 ± 3	-4.44 ± 0.14	-4.44 ± 0.14
42.65 ± 0.025	3 ± 1	-5.08 ± 0.33	-5.02 ± 0.33

Table B2. [O III] + Hβ LF – this has not been corrected for extinction. The total volume per bin is 24.8×10^5 Mpc^3, assuming a top-hat filter, but ϕ corr has been corrected for both incompleteness and the fact that the filter profile is not a perfect top-hat.

log $L_{[\mathrm{O\,III}]}$ ($z = 1.4$)	Sources (#)	ϕ obs (Mpc^{-3})	ϕ corr (Mpc^{-3})
42.07 ± 0.03	285 ± 16	-2.72 ± 0.03	-2.71 ± 0.04
42.13 ± 0.03	283 ± 16	-2.72 ± 0.03	-2.71 ± 0.04
42.19 ± 0.03	187 ± 13	-2.90 ± 0.03	-2.91 ± 0.04
42.25 ± 0.03	104 ± 10	-3.16 ± 0.04	-3.12 ± 0.06
42.31 ± 0.03	74 ± 8	-3.30 ± 0.05	-3.21 ± 0.08
42.37 ± 0.03	32 ± 5	-3.67 ± 0.08	-3.53 ± 0.15
42.43 ± 0.03	23 ± 4	-3.81 ± 0.10	-3.63 ± 0.19
42.49 ± 0.03	17 ± 4	-3.94 ± 0.12	-3.78 ± 0.23
42.55 ± 0.03	8 ± 2	-4.27 ± 0.19	-4.10 ± 0.43
42.61 ± 0.03	6 ± 2	-4.39 ± 0.23	-4.21 ± 0.56
42.73 ± 0.03	14 ± 3	-4.03 ± 0.14	-4.28 ± 0.29
42.85 ± 0.09	4 ± 2	-4.57 ± 0.30	-5.03 ± 0.53
43.03 ± 0.09	1 ± 1	-5.17 ± 0.38	-5.63 ± 0.38

Table B3. [O II] LF not corrected for dust extinction. The total volume per bin is 26.2×10^5 Mpc^3, assuming a top-hat filter, but ϕ corr has been corrected for both incompleteness and the fact that the filter profile is not a perfect top-hat.

log $L_{[\mathrm{O\,II}]}$ ($z = 2.2$)	Sources (#)	ϕ obs (Mpc^{-3})	ϕ corr (Mpc^{-3})
42.60 ± 0.025	141 ± 11	-2.97 ± 0.04	-2.88 ± 0.11
42.65 ± 0.025	118 ± 10	-3.04 ± 0.04	-2.96 ± 0.06
42.70 ± 0.025	73 ± 8	-3.25 ± 0.05	-3.14 ± 0.12
42.75 ± 0.025	53 ± 7	-3.39 ± 0.06	-3.31 ± 0.11
42.80 ± 0.025	30 ± 5	-3.64 ± 0.09	-3.54 ± 0.14
42.85 ± 0.025	18 ± 4	-3.86 ± 0.12	-3.75 ± 0.15
42.90 ± 0.025	15 ± 3	-3.94 ± 0.13	-3.84 ± 0.16
42.95 ± 0.025	6 ± 2	-4.34 ± 0.23	-4.21 ± 0.25
43.00 ± 0.025	2 ± 1	-4.82 ± 0.53	-4.69 ± 0.58
43.10 ± 0.075	5 ± 2	-4.99 ± 0.26	-4.99 ± 0.26
43.25 ± 0.075	2 ± 1	-5.52 ± 0.53	-5.52 ± 0.53

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