



## The MUSE-Wide survey: A first catalogue of 831 emission line galaxies

Edmund Christian Herenz, Tanya Urrutia, Lutz Wisotzki, Josephine Kerutt, Rikke Saust, Maria Werhahn, Kasper Borello Schmidt, Joseph Caruana, Catrina Diener, Roland Bacon, et al.

### ► To cite this version:

Edmund Christian Herenz, Tanya Urrutia, Lutz Wisotzki, Josephine Kerutt, Rikke Saust, et al.. The MUSE-Wide survey: A first catalogue of 831 emission line galaxies. *Astronomy and Astrophysics - A&A*, 2017, 606, 10.1051/0004-6361/201731055 . insu-03710614

**HAL Id: insu-03710614**

**<https://insu.hal.science/insu-03710614>**

Submitted on 1 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The MUSE-Wide survey: A first catalogue of 831 emission line galaxies<sup>★,★★</sup>

Edmund Christian Herenz<sup>1,2</sup>, Tanya Urrutia<sup>1</sup>, Lutz Wisotzki<sup>1</sup>, Josephine Kerutt<sup>1</sup>, Rikke Saust<sup>1</sup>, Maria Werhahn<sup>1</sup>,  
Kasper Borello Schmidt<sup>1</sup>, Joseph Caruana<sup>1,3,4</sup>, Catrina Diener<sup>1,5</sup>, Roland Bacon<sup>6</sup>, Jarle Brinchmann<sup>7</sup>,  
Joop Schaye<sup>7</sup>, Michael Maseda<sup>7</sup>, and Peter M. Weilbacher<sup>1</sup>

<sup>1</sup> Leibniz-Institut für Astrophysik Potsdam (AIP), An der Sternware 16, 14482 Potsdam, Germany

<sup>2</sup> Department of Astronomy, Stockholm University, AlbaNova University Centre, 106 91, Stockholm, Sweden  
e-mail: [christian.herenz@astro.su.se](mailto:christian.herenz@astro.su.se)

<sup>3</sup> Department of Physics, University of Malta, Msida MSD 2080, Malta

<sup>4</sup> Institute of Space Sciences & Astronomy, University of Malta, Msida MSD 2080, Malta

<sup>5</sup> Institute of Astronomy, Madingley Road, Cambridge, CB3 0HA, UK

<sup>6</sup> CRAL, Observatoire de Lyon, CNRS, Université Lyon 1, 9 Avenue Charles André, 69561 Saint Genis-Laval Cedex, France

<sup>7</sup> Leiden Observatory, Leiden University, PO Box 9513, 2300 RA Leiden, The Netherlands

Received 26 April 2017 / Accepted 19 May 2017

## ABSTRACT

We present a first instalment of the MUSE-Wide survey, covering an area of 22.2 arcmin<sup>2</sup> (corresponding to ~20% of the final survey) in the CANDELS/Deep area of the *Chandra* Deep Field South. We use the MUSE integral field spectrograph at the ESO VLT to conduct a full-area spectroscopic mapping at a depth of 1 h exposure time per 1 arcmin<sup>2</sup> pointing. We searched for compact emission line objects using our newly developed LSDCat software based on a 3D matched filtering approach, followed by interactive classification and redshift measurement of the sources. Our catalogue contains 831 distinct emission line galaxies with redshifts ranging from 0.04 to 6. Roughly one third (237) of the emission line sources are Lyman  $\alpha$  emitting galaxies with  $3 < z < 6$ , only four of which had previously measured spectroscopic redshifts. At lower redshifts 351 galaxies are detected primarily by their [O II] emission line ( $0.3 \lesssim z \lesssim 1.5$ ), 189 by their [O III] line ( $0.21 \lesssim z \lesssim 0.85$ ), and 46 by their H $\alpha$  line ( $0.04 \lesssim z \lesssim 0.42$ ). Comparing our spectroscopic redshifts to photometric redshift estimates from the literature, we find excellent agreement for  $z < 1.5$  with a median  $\Delta z$  of only  $\sim 4 \times 10^{-4}$  and an outlier rate of 6%, however a significant systematic offset of  $\Delta z = 0.26$  and an outlier rate of 23% for Ly $\alpha$  emitters at  $z > 3$ . Together with the catalogue we also release 1D PSF-weighted extracted spectra and small 3D datacubes centred on each of the 831 sources.

**Key words.** galaxies: high-redshift – techniques: imaging spectroscopy – catalogs – surveys

## 1. Introduction

Most spectroscopic samples of high-redshift galaxies are based on a photometric pre-selection of targets (e.g. [Noll et al. 2004](#); [Vanzella et al. 2005, 2006, 2008](#); [Popesso et al. 2009](#); [Balestra et al. 2010](#); [Mallery et al. 2012](#); [Le Fèvre et al. 2013, 2015](#)). These surveys have very successfully maximised their spectroscopic success rates, i.e. the fraction of galaxies with scientifically usable spectra among all targeted objects, by employing photometric redshift priors. However, an inevitable conceptual drawback of this pre-selection approach is that the selection process itself will leave its imprint on the final sample properties. Moreover, multi-object spectrographs have only limited freedom in choosing targets for simultaneous observation, and even state-of-the-art surveys hardly ever obtain target sampling rates above 50%, more often well below this level. Finally, aperture effects in the preconfigured slit mask can lead to significant flux losses

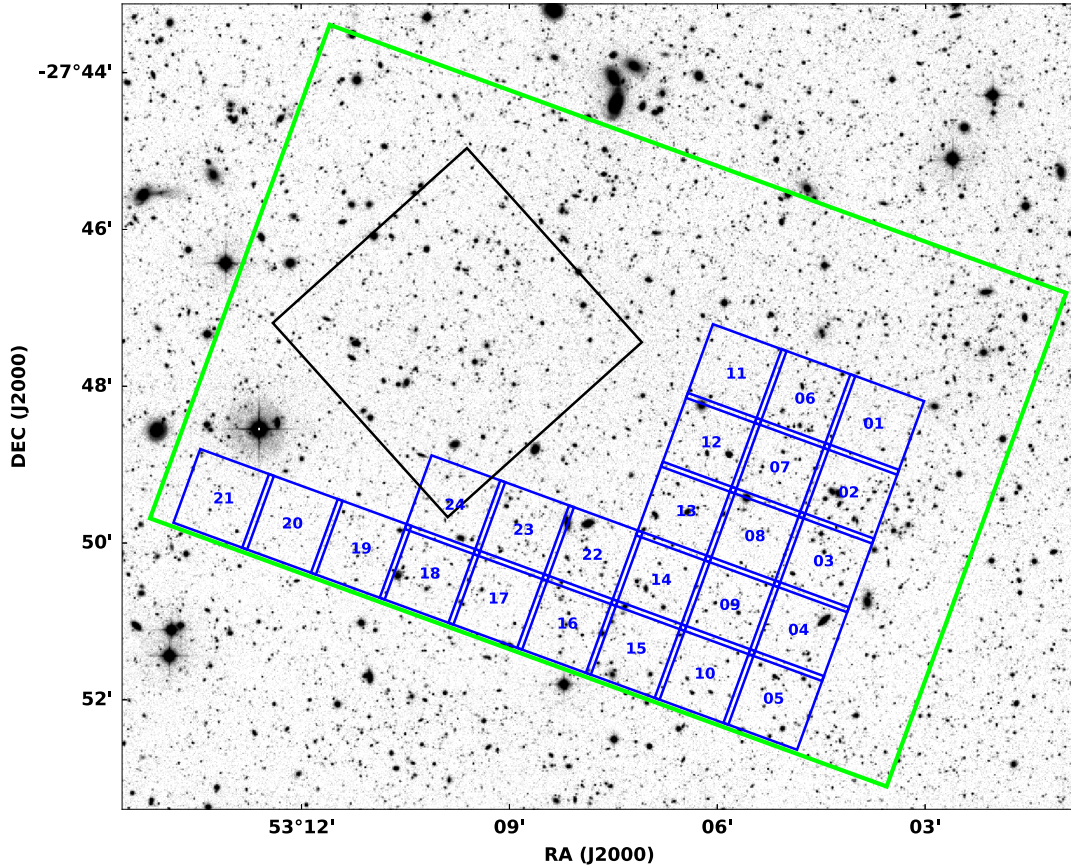
especially from complex objects and/or extended emission line regions.

Integral field spectroscopy (IFS) provides an alternative approach that circumvents many of these problems. Contiguous areas in the sky can be mapped instead of targeting individual objects, providing spectral information of everything within the field of view (FoV) and within the sensitivity limits of the observation (see [van Breukelen et al. 2005](#); and [Adams et al. 2011](#); for pioneering implementations). The new panoramic IFS instrument MUSE (Multi Unit Spectroscopic Explorer) at the ESO Very Large Telescope is a particularly powerful machine specifically designed to perform blind surveys for extremely faint high-redshift galaxies ([Bacon et al. 2009, 2014](#); [Caillier et al. 2014](#)). The discovery potential of MUSE was strikingly demonstrated by a 27 h integration in the *Hubble* Deep Field South ([Bacon et al. 2015](#)), where 189 redshifts could be measured inside a single MUSE field of 1 arcmin<sup>2</sup>. While surveying the sky with MUSE overcomes the above mentioned limitations, this obviously happens at the expense of covering only a very small area at a time.

Here we present results from the MUSE-Wide survey, a GTO programme complementing the ongoing ultra-deep pencil-beam MUSE surveys ([Bacon et al. 2017](#)). MUSE-Wide trades depth for survey area by covering many fields with relatively shallow

\* Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programme 094.A-0205.

\*\* Data products are available via <http://muse-vlt.eu/science/> and at the CDS via anonymous ftp to [cdsarc.u-strasbg.fr](http://cdsarc.u-strasbg.fr) (130.79.128.5) or via <http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/606/A12>.



**Fig. 1.** Footprint of the first 24  $1' \times 1'$  pointings of the MUSE-Wide survey in the CANDELS Deep region of GOODS-South (overlaid over the *R*-Band Image from GaBoDS – Erben et al. 2005; Hildebrandt et al. 2006). The black square indicates the region of the *Hubble* Ultra Deep Field and the green rectangle outlines the CANDELS Deep region.

exposures (1 h per field). Nevertheless, the obtained depth already suffices to obtain source densities (i.e. multiplex factors) of several tens of objects per  $\text{arcmin}^2$  with useful spectra, at a target sampling rate of essentially 100% and with all the benefits of a powerful integral field unit (IFU). In particular, MUSE features an excellent spatial sampling at  $0''.2 \times 0''.2$  spatial pixels and a spectral resolution of  $2.5 \text{ \AA}$  over one octave in wavelengths from  $4750 \text{ \AA}$  to  $9350 \text{ \AA}$ .

A full description of the MUSE-Wide survey strategy will be present in a forthcoming dedicated publication, accompanying a first general data release (Urrutia et al., in prep.). Here we only summarise the currently available material that forms the basis of the present publication. The MUSE-Wide survey focuses on areas with extremely deep HST imaging, with special emphasis on the GOODS-South (Giavalisco et al. 2004) and CANDELS-Deep/CDFS (Grogin et al. 2011; Koekemoer et al. 2011) regions, which in addition to the HST coverage also contains a plethora of multiwavelength data.

This paper describes the outcome of a blind search for emission line objects in 24 MUSE-Wide fields in the CDFS region, covering a footprint of  $22.2 \text{ arcmin}^2$  and yielding a total of 831 galaxies, of which more than half had no spectroscopic redshift until now. Besides the catalogue we also publish object-specific data products suitable for further investigations.

The structure of the paper is as follows: in Sect. 2 we outline the observations and data reduction. We then describe in Sect. 3 how we detect, parameterise and classify emission line sources in our datacubes, including details of the redshift determination procedure. In Sect. 4 we present and describe the

source catalogue and data products. Section 5 is dedicated to a global characterisation of the obtained sample. We present our conclusions and outlook in Sect. 6.

## 2. Observations and data reduction

Our current dataset is based on the analysis of 24 adjacent  $1' \times 1'$  MUSE pointings in the CANDELS Deep region of the GOODS-South field obtained during the first semester of MUSE guaranteed time observations. Observations were carried out in grey and dark time under photometric and clear conditions from September to December 2014 (ESO programme 094.A-0205, PI: Lutz Wisotzki). In Fig. 1 we show the footprint of the survey area. We matched the position angle of our pointings to the  $70^\circ$  position angle (east of north) of the CANDELS Deep region (indicated by the green box in Fig. 1). In Table 1 we provide a log of our observations. Standard star exposures were taken at the beginning and at the end of each night.

We integrated 1 h on each pointing. Each integration was split into 4 exposures of 15 min. In between exposures small dither offsets (typically smaller than  $1''$ ) were applied and the spectrograph was rotated by 90 degrees. This procedure, which is recommended in the MUSE User’s manual<sup>1</sup>, ensures that patterns of the 24 individual spectrographs and their image slicers are averaged out. With the exception of pointings 08, 09, 12, and 16 all exposures for a pointing were obtained in immediate succession. Pointing 08 was split into two exposure sequences

<sup>1</sup> <http://www.eso.org/sci/facilities/paranal/instruments/muse/doc.html>



**Table 1.** Log of observations.

| Pointing             | Pointing centre                                    |                  | Date          | AG seeing<br>[""] | Airmass | Conditions                |
|----------------------|--|------------------|---------------|-------------------|---------|---------------------------|
|                      | $\alpha_{J2000}$                                   | $\delta_{J2000}$ |               |                   |         |                           |
| MUSE-candels-cdfs-01 | 03 <sup>h</sup> 32 <sup>m</sup> 15.04 <sup>s</sup> | −27°48′29.4″     | 2014-10-20    | 0.86              | 1.087   | photometric / grey        |
| MUSE-candels-cdfs-02 | 03 <sup>h</sup> 32 <sup>m</sup> 16.48 <sup>s</sup> | −27°49′21.5″     | 2014-09-20    | 1.05              | 1.076   | clear / grey              |
| MUSE-candels-cdfs-03 | 03 <sup>h</sup> 32 <sup>m</sup> 17.80 <sup>s</sup> | −27°50′13.6″     | 2014-11-17    | 0.93              | 1.061   | clear / grey              |
| MUSE-candels-cdfs-04 | 03 <sup>h</sup> 32 <sup>m</sup> 19.67 <sup>s</sup> | −27°51′07.1″     | 2014-11-17    | 0.76              | 1.041   | clear / grey              |
| MUSE-candels-cdfs-05 | 03 <sup>h</sup> 32 <sup>m</sup> 20.70 <sup>s</sup> | −27°51′59.8″     | 2014-11-19    | 1.03              | 1.017   | clear / dark              |
| MUSE-candels-cdfs-06 | 03 <sup>h</sup> 32 <sup>m</sup> 18.91 <sup>s</sup> | −27°48′10.7″     | 2014-11-18    | 0.84              | 1.021   | clear / dark              |
| MUSE-candels-cdfs-07 | 03 <sup>h</sup> 32 <sup>m</sup> 20.36 <sup>s</sup> | −27°49′02.6″     | 2014-11-19    | 0.92              | 1.026   | photometric / dark        |
| MUSE-candels-cdfs-08 | 03 <sup>h</sup> 32 <sup>m</sup> 21.89 <sup>s</sup> | −27°49′55.3″     | 2014-11-19+20 | 1.00              | 1.034   | photometric / dark        |
| MUSE-candels-cdfs-09 | 03 <sup>h</sup> 32 <sup>m</sup> 23.25 <sup>s</sup> | −27°50′47.9″     | 2014-11-26    | 0.87              | 1.044   | photometric / dark + grey |
| MUSE-candels-cdfs-10 | 03 <sup>h</sup> 32 <sup>m</sup> 24.68 <sup>s</sup> | −27°51′40.8″     | 2014-11-27    | 0.90              | 1.071   | photometric / grey        |
| MUSE-candels-cdfs-11 | 03 <sup>h</sup> 32 <sup>m</sup> 22.91 <sup>s</sup> | −27°47′50.9″     | 2014-11-28    | 0.95              | 1.104   | photometric / grey        |
| MUSE-candels-cdfs-12 | 03 <sup>h</sup> 32 <sup>m</sup> 24.35 <sup>s</sup> | −27°48′43.2″     | 2014-11-27    | 1.02              | 1.144   | photometric / grey        |
| MUSE-candels-cdfs-13 | 03 <sup>h</sup> 32 <sup>m</sup> 25.88 <sup>s</sup> | −27°49′36.2″     | 2014-11-27    | 1.07              | 1.193   | photometric / dark        |
| MUSE-candels-cdfs-14 | 03 <sup>h</sup> 32 <sup>m</sup> 27.21 <sup>s</sup> | −27°50′28.9″     | 2014-11-28    | 0.88              | 1.229   | photometric / grey        |
| MUSE-candels-cdfs-15 | 03 <sup>h</sup> 32 <sup>m</sup> 28.66 <sup>s</sup> | −27°51′21.5″     | 2014-12-25    | 0.83              | 1.241   | photometric / grey        |
| MUSE-candels-cdfs-16 | 03 <sup>h</sup> 32 <sup>m</sup> 32.62 <sup>s</sup> | −27°51′02.1″     | 2014-11-28    | 0.83              | 1.228   | photometric / grey        |
| MUSE-candels-cdfs-17 | 03 <sup>h</sup> 32 <sup>m</sup> 36.60 <sup>s</sup> | −27°50′43.8″     | 2014-12-23    | 0.80              | 1.191   | clear / dark              |
| MUSE-candels-cdfs-18 | 03 <sup>h</sup> 32 <sup>m</sup> 40.58 <sup>s</sup> | −27°50′24.3″     | 2014-12-21    | 0.89              | 1.191   | photometric / dark        |
| MUSE-candels-cdfs-19 | 03 <sup>h</sup> 32 <sup>m</sup> 44.60 <sup>s</sup> | −27°50′04.6″     | 2014-12-21    | 0.82              | 1.223   | photometric / dark        |
| MUSE-candels-cdfs-20 | 03 <sup>h</sup> 32 <sup>m</sup> 48.61 <sup>s</sup> | −27°49′46.0″     | 2014-12-23    | 0.82              | 1.288   | clear / dark              |
| MUSE-candels-cdfs-21 | 03 <sup>h</sup> 32 <sup>m</sup> 52.54 <sup>s</sup> | −27°49′26.1″     | 2014-12-23    | 0.72              | 1.388   | clear / dark              |
| MUSE-candels-cdfs-22 | 03 <sup>h</sup> 32 <sup>m</sup> 31.19 <sup>s</sup> | −27°50′09.8″     | 2014-12-22    | 0.79              | 1.338   | clear / dark              |
| MUSE-candels-cdfs-23 | 03 <sup>h</sup> 32 <sup>m</sup> 35.23 <sup>s</sup> | −27°49′50.3″     | 2014-12-24    | 0.86              | 1.266   | photometric / dark        |
| MUSE-candels-cdfs-24 | 03 <sup>h</sup> 32 <sup>m</sup> 39.14 <sup>s</sup> | −27°49′31.5″     | 2014-12-26    | 0.81              | 1.168   | photometric / grey        |

**Notes.** The integration time for each exposure was 3600 s. Autoguider seeing (AG seeing) and airmass values refer to the average over all four individual exposures per pointing.

in two subsequent nights, while pointings 09, 12, and 16 were taken at different times during a night. Adjacent pointings have an overlap of 4″. Taking this overlap and the exact geometry of the MUSE FoV into account, the total area exposed with MUSE in the 24 pointings taken in the first cycle of the MUSE-Wide survey is 22.2 arcmin<sup>2</sup>.

For the reduction of the individual pointings we used version 1.0 of the MUSE data reduction system<sup>2</sup> (Weilbacher et al. 2014, and in prep.), in combination with custom developed python<sup>3</sup> routines and the ZAP tool presented in Soto et al. (2016). An in-depth description and validation of the data reduction procedure will be given in the publication complementing the full data release (Urrutia et al., in prep.), in the following we only provide a brief overview.

We used the set of calibration exposures taken closest in time to the actual observations to create master biases, master flats, dispersion solutions, and trace tables. For the illumination correction, we always chose the illumination frames that were taken before the science observation. Using the standard-star exposures we constructed response curves for flux-calibration. We applied these calibration products with the pipeline routine `muse_scibasic` to all 24 spectrographs CCD images belonging to one science exposure. The result of this process are so-called pixel tables for each exposure containing calibrated flux values, errors, wavelengths, and information on their location on the sky.

It is known that the current version of the MUSE pipeline sky-subtraction routine (Streicher et al. 2011) leaves significant systematic residuals that hamper the detection of faint object signals (Soto et al. 2016). For this reason we developed and applied our own sky-subtraction routine that works on the pixel tables. We found, that this procedure, in combination with the ZAP tool by Soto et al. (2016), provided a more optimal result compared to using only ZAP. Our procedure will be detailed in Urrutia et al., (in prep.). Finally, we applied the self-calibration method described in Sect. 3.1 of Bacon et al. (2015) to remove systematic mean zero-flux level offsets between slices<sup>4</sup>.

For each exposure we then used the pipeline routine `muse_scipost` that resamples the pixtables into datacubes and propagates errors into a corresponding variance datacube. The datacubes were corrected for differential atmospheric refraction using the formula by Filippenko (1982). Remaining sky-subtraction residuals after the application of our own routine were then purged from the cubes using the ZAP-software of Soto et al. (2016). We then created white-light images from these datacubes by summation over the spectral axis. By performing 2D Gaussian fits to compact objects within those images (compact galaxies, or, when available, stars) we determined a reference registration for each exposure datacube. Using these determined offsets for the individual cubes, we ran `muse_scipost` again, to resample all exposures onto a common

<sup>2</sup> Available from ESO via <http://www.eso.org/sci/software/pipelines/muse/muse-pipe-recipes.html>  
<sup>3</sup> <http://www.python.org>

<sup>4</sup> The self-calibration procedure is part of the *Muse Python Data Analysis Framework* (MPDAF, Bacon et al. 2016; see also Conseil et al. 2016). MPDAF is available at <http://mpdaf.readthedocs.io/en/latest> and <http://ascl.net/1611.003>

world-coordinate system grid. We created a combined datacube for each pointing by averaging the corresponding four exposures with rejection of  $3\sigma$  outliers.

In an additional postprocessing step we adjusted the zero level for all layers. This was needed since none of the preceding reduction steps (including the sky subtraction) actually enforces a spatially coherent zero background level. We obtained this zero level correction by masking out all continuum sources in a wavelength-collapsed white-light image of the field and then computing the average of the remaining unmasked spaxels in each layer. While these corrections were found to be small except very close to strong night sky lines, with median values around  $2 \times 10^{-22} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ , as systematic offsets they are nevertheless potentially significant for large aperture integrations. We subtracted the corrections layer by layer from the combined cubes and thus obtained the final flux datacube.

After the procedure described above we have 24 datacubes that contain  $\approx 3.5 \times 10^8$  exposed volume-pixels (so called voxels). The spatial sampling is  $0.2'' \times 0.2''$ , the spectral sampling is  $1.25 \text{ \AA}$ , and the wavelength range extends from  $4750 \text{ \AA}$  to  $9350 \text{ \AA}$ . The wavelength axis is given in air wavelengths in the barycentric reference frame. For each cube the astrometry is such that the east-west and south-north axes are parallel to the spatial coordinate axes. Complementary to the flux datacubes we produced a final variance datacube by propagating the individual variance values. We also created exposure-map cubes, where we track the number of individual single exposures that went into a voxel of the final datacube. The datacube, variance cube and exposure-map cubes are stored in separate header and data units (HDUs) of a single FITS file (Pence et al. 2010) taking 5 GB of disk space for each pointing.

### 3. Emission line source detection and classification

#### 3.1. Detection and parameterisation of emission line source candidates

On each of the 24 datacubes we performed the following tasks to build a catalogue of emission line source candidates:

1. Empirical estimation and correction of the MUSE pipeline propagated variance cubes.
2. Removal of source continua from the datacube.
3. Cross-correlation of the datacube with a 3D matched filter for compact emission line sources.
4. Thresholding and cataloguing of emission line source candidates.
5. Position, size, and flux measurements of the emission line source candidates.

For tasks 3 to 5 we have developed the emission *Line Source Detection and Cataloguing Tool* LSDCat<sup>5</sup>. In the following subsections we briefly describe the above steps. For an in-depth description of LSDCat we refer to Herenz & Wisotzki (2017, in the following HW17).

##### 3.1.1. Empirical correction of the MUSE pipeline propagated noise

Both for emission line source detection and flux measurements we require an accurate characterisation of the noise in our datacubes. However, we expect that the variance cubes provided by

the pipeline underestimate the true variance. First, because the pipeline propagates variances from the CCD level through the various reduction steps, thus these formal variance values do not account for hidden (non-Gaussian) systematics such as imperfect flat-fielding or imperfect sky-subtraction. Second, in the resampling process carried out by the MUSE pipeline co-variance terms are neglected.

In order to quantify and correct the formal variances produced by the pipeline, we determined an empirical variance estimate from the flux datacubes by analysing the pixel statistics for each datacube layer in regions of blank sky. However, direct pixel-to-pixel noise statistics as such would also be biased towards lower values since also here covariance terms introduced in the resampling process would be neglected. To overcome this issue we placed circular apertures with a  $2''$  diameter (10 pixels) at random positions in each pointing. By prohibiting these apertures to overlap with  $m_{F814} < 25$  mag sources from the Guo et al. (2013) CANDELS GOODS-S photometric catalogue we ensured that blank sky was sampled. The width of the distribution of average pixel values in those apertures, characterised by its standard deviation, was then used as an estimate of the noise for each spectral layer. Due to the small FoV of MUSE we were limited to use only 100 apertures. Hence, we estimated the width of the distribution for each layer via  $\sigma_{\text{emp}} = (q_{75} - q_{25}) / (2 \sqrt{2} \text{erf}^{-1}(1/2)) \approx 0.7413 (q_{75} - q_{25})$ , where  $\text{erf}^{-1}$  is the inverse error function and  $q_{75} - q_{25}$  is the interquartile range of the distribution (e.g., Sect. 3.2.2 in Ivezić et al. 2014). We then compared the so obtained empirical noise values  $\sigma_{\text{emp}}^2$  to an average value  $\sigma_{\text{pipe}}^2$  of the pipeline propagated variance in each layer – this process is illustrated in Fig. 2. As can be seen the ratio  $\sigma_{\text{emp}}/\sigma_{\text{pipe}}$  is greater than one in almost all datacube layers. For all cubes we found typically  $\sigma_{\text{emp}} = 1.15 - 1.20 \sigma_{\text{pipe}}$ . This is consistent with an empirical noise estimate made for MUSE datacubes obtained with a similar observing strategy (Borisova et al. 2016, their Sect. 3.3). However, as can be also seen in Fig. 2 there is a small number of layers where the ratio  $\sigma_{\text{emp}}/\sigma_{\text{pipe}}$  is less than one. The affected layers are in the cores of sky emission lines where they suffer from over-subtraction of the high-frequency noise by the ZAP routine.

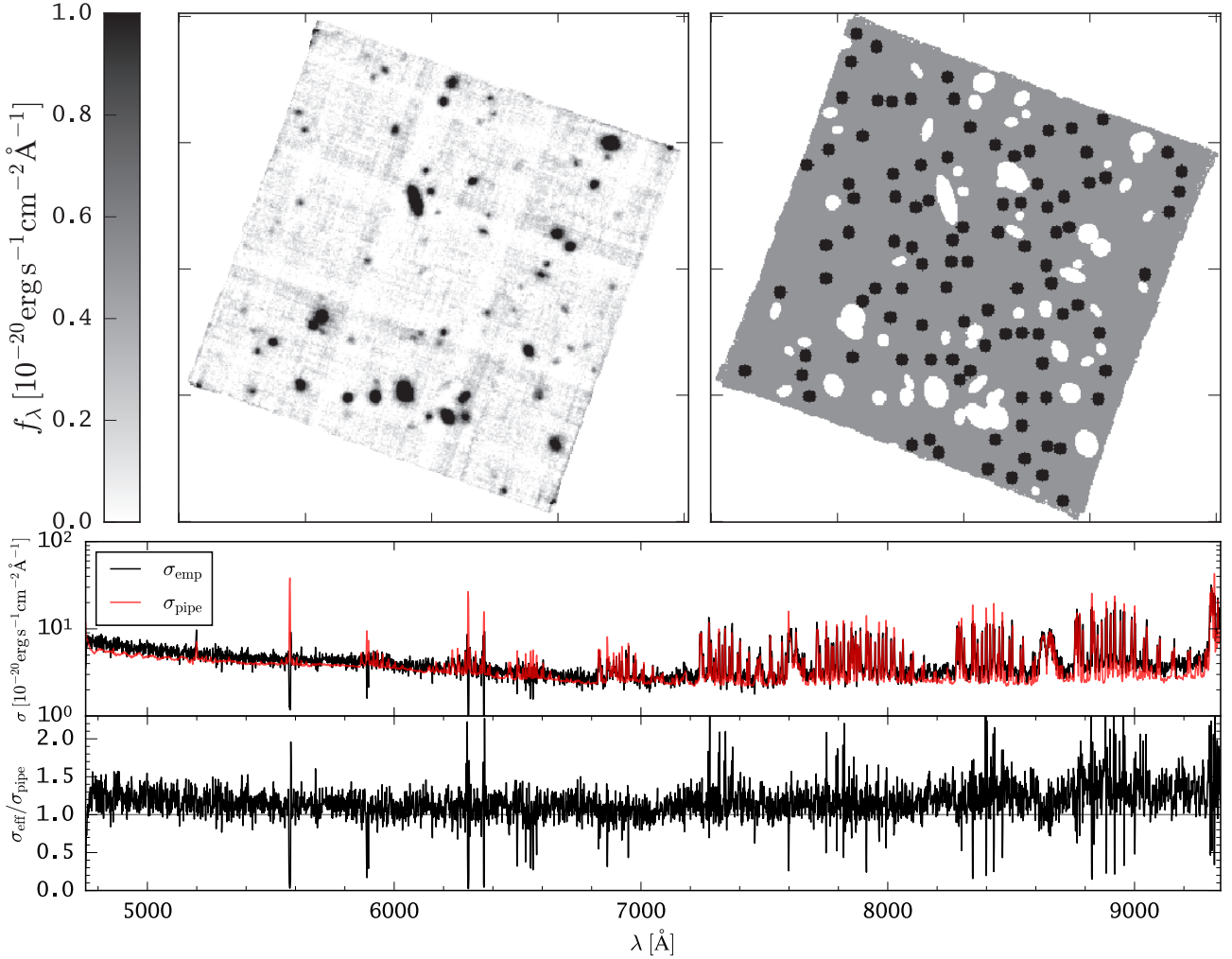
In order to correct the pipeline propagated variance estimate, we replace the values in the variance cubes with the empirical noise estimate  $\sigma_{\text{emp}}^2$ . In layers where  $\sigma_{\text{emp}}/\sigma_{\text{pipe}} < 1$  we use the average value from the pipeline. Due to the dither- and rotation pattern typically  $\lesssim 10\%$  of voxels of the datacube have not contributions from all four exposure. These voxels are mostly on the border of the FoV. To adjust the noise estimates for those voxels we rescale their  $\sigma_{\text{emp}}^2$  with  $4/N_{\text{exp}}$ , where  $N_{\text{exp}}$  is the number of exposures contributing to that voxel.

There are two caveats with this empirical noise estimate. First, due to the small number of apertures our noise estimate is itself noisy. This can be seen when comparing the smooth  $\sigma_{\text{pipe}}$ -curve to the noisier  $\sigma_{\text{emp}}$ -curve in Fig. 2. Second, our estimate does not take correlated noise in the spectral direction into account. We will address these shortcomings of the described empirical noise estimate in future releases of the MUSE-Wide data (Urrutia et al., in prep.). Nevertheless, compared to the MUSE pipeline propagated variances that underestimate the true variance our empirical estimation approach is an improvement.

##### 3.1.2. Removal of source continua

The source detection algorithm in LSDCat searches for emission line signals and implicitly assumes that significant source

<sup>5</sup> LSDCat is available via the Astrophysics Source Code Library: <http://www.ascl.net/1612.002> (Herenz & Wisotzki 2016).



**Fig. 2.** Illustration of the empirical noise calculation procedure. Shown is the case in the MUSE-Wide pointing MUSE-candels-cdfs-06. The *top left panel* shows a white-light image created by summing over all spectral layers and subsequent division by the spectral range. In the *top right panel* we show the 100 random 2'' diameter apertures (black) and the avoided regions (white) because of the presence of continuum bright objects (sources with  $m_{F814} < 25$  mag in Guo et al. 2013). In the *bottom panels* we compare the width of the distribution of the flux values extracted in the 100 apertures for each spectral layer  $\sigma_{\text{emp}}$  (black curve) normalised to one spectral pixel to the corresponding aperture average from the pipeline produced variance cube  $\sigma_{\text{pipe}}$  (red curve).

continua are subtracted from the datacube. We achieved this by subtracting a 151 pixel wide running median in the spectral direction from the datacube. The full width of the median filter was chosen to be 151 spectral layers ( $188.75 \text{ \AA}$ ), which is much broader than the width of the targeted emission lines and narrow enough to robustly subtract slowly varying continua. The remaining residuals in the datacube are either real emission lines or residual features from continua varying at high frequencies (e.g. cold stars). We demonstrate the effectiveness of our median-filter subtraction method in Fig. 3.

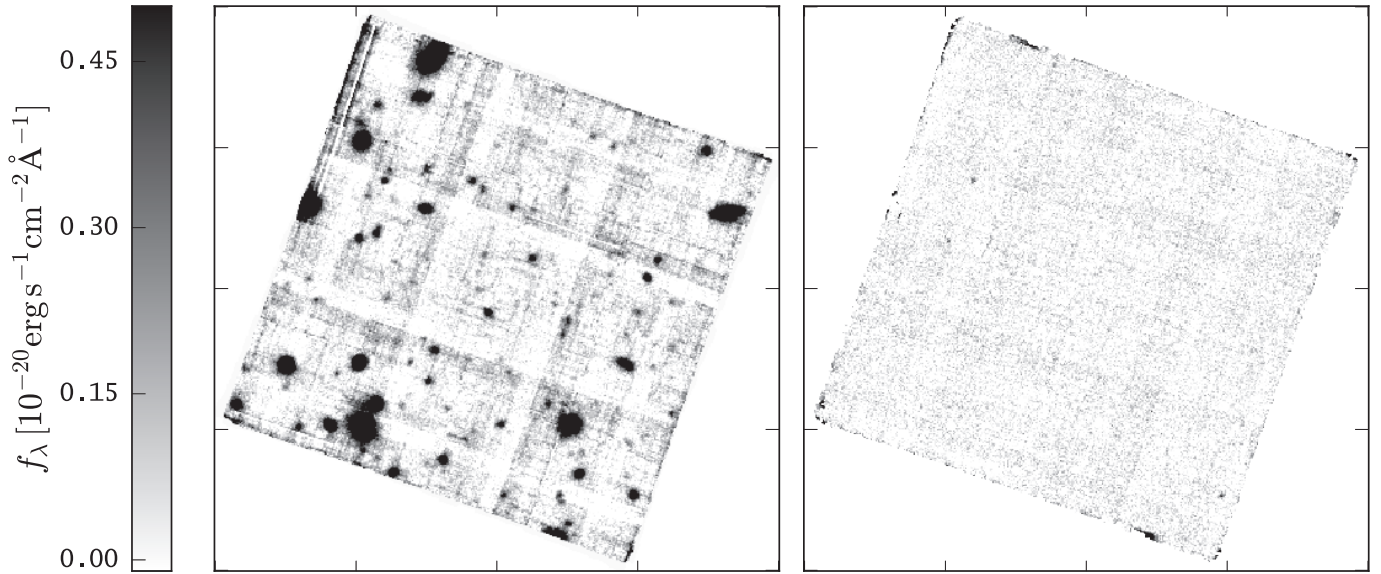
### 3.1.3. Cross-correlation with matched filter

The detection algorithm of LSDCat is based on matched filtering, an operation that maximises the signal-to-noise ratio (S/N) of emission lines within the datacube (e.g. Schwartz & Shaw 1975; Das 1991; Bertin 2001; Zackay & Ofek 2017; Vio & Andreani 2016). To this aim the algorithm transforms the input datacube by convolving it with a three-dimensional template that matches the expected signal of an emission line in the datacube. The 3D convolution is performed as two successive convolutions, a 2D

convolution in each spectral layer and a 1D convolution in the spectral direction for each spaxel.

As the template for the 2D convolution in each spectral layer we use the circular Gaussian profile option of LSDCat. For the matched filtering this template provides a reasonably good approximation of the seeing induced point spread function (PSF). The width of the PSF, typically given as the full width at half maximum (FWHM), depends on wavelength (e.g., Hickson 2014). The input-parameters for LSDCat describing this width and its wavelength dependence need to be supplied as the coefficients  $p_0$  and  $p_1$  of the linear function  $FWHM(\lambda)'' = p_0 + p_1(\lambda - 7000 \text{ \AA})$  that provides an acceptable approximation over the wavelength range under consideration here. In principle, the determination of the  $FWHM(\lambda)$  dependence, and thus the optimal coefficients in  $p_0$  and  $p_1$ , can be achieved by fitting a 2D Gaussian function to a reasonably bright star in each spectral layer of the datacube. However, by choice the CANDELS fields are devoid of bright stars. Hence, for numerous of our pointings selecting objects from the Guo et al. (2013) CANDELS GOODS-S photometric catalogue with  $CLASS\_STAR > 0.95$ , i.e. objects that are likely stars, results only in objects with

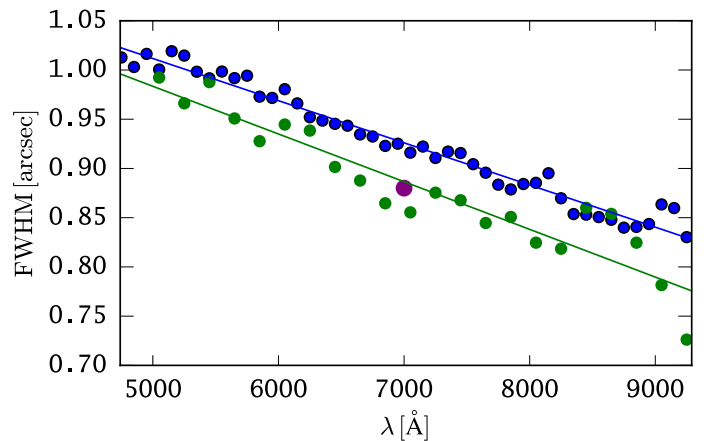




**Fig. 3.** Example illustrating the effectiveness of subtracting a running median in the spectral direction from the datacube to remove signal from bright continuum sources. Shown is the MUSE-Wide pointing MUSE-candels-cdfs-15. *Left:* white-light image created by summing all spectral layers and subsequent division by the spectral range. *Right:* white-light image of the median filter subtracted version of the same cube illustrating the effectiveness of this continuum subtraction method.

$m_{F814} > 23$  mag. For these objects even binning of 100 spectral layers of the datacube, does not result in a S/N sufficient for reliable 2D Gaussian fits. However, for stars brighter than  $m_{F814} \lesssim 22$  mag such fits do converge. For 13 of the 24 pointings this direct point-source fit could be employed. To get the FWHM values in the binned layers we used the PAMPELMUSE software (Kamann 2013; Kamann et al. 2013). For the remaining fields we devised a minimisation scheme utilising compact galaxies. For this we visually selected compact galaxies within the FoV of a pointing from the CANDELS F814W image. We then convolved these galaxies with 2D Gaussians of different FWHMs and resampled these convolved images to the spatial resolution of MUSE. In a series of  $\sim 100$ – $200$  binned layers we then determined the  $\chi^2$  of the differences between the convolved and resampled images to the real data. Finally, we fitted the sequence of FWHM values at different wavelengths with a linear function to obtain  $p_0$  and  $p_1$ . For the 13 pointings that have a star in the field we found that the two methods agree within 10% on the derived  $FWHM(\lambda)$  dependence. This is shown for one example in Fig. 4. We list in Table 2 all  $p_0$  and  $p_1$  coefficients that we used as input for the spatial filtering procedure in LSDCat. For pointings with both the stellar- and compact object-based estimates of the PSF available, we used the coefficients resulting in  $FWHM(7000 \text{ \AA})$  being closest to the autoguider seeing value given in Table 1, which we consider to be the best seeing estimate at this wavelength.

For the spectral convolution LSDCat uses a 1D Gaussian template. Its width needs to be specified as the velocity FWHM –  $v_{FWHM}$  – in  $\text{km s}^{-1}$ . We fixed this parameter in our emission line search to  $v_{FWHM} = 250 \text{ km s}^{-1}$ . We found this to be the best single value for achieving the highest S/N for the majority of LAEs in MUSE datacubes (HW17). Moreover, taking the instrumental resolution of MUSE into account, this value is also consistent with the expectations from the distributions of LAE FWHMs in the literature (e.g., Dawson et al. 2007; Mallery et al. 2012; Yamada et al. 2012). Although the width of the spectral filter is optimised for LAEs, we emphasise that the resulting S/N of an Gaussian emission line decreases only as



**Fig. 4.** Illustration of the determination of the wavelength dependence of the FWHM of the seeing PSF. As an example we show the results for the MUSE-Wide pointing MUSE-candels-cdfs-14. Blue points show the FWHM values obtained from fitting a 2D Gaussian to images of a star within the datacube. The used images were created by summing over  $100 \text{ \AA}$  along the spectral axis. Green points show the results from minimising the  $\chi^2$  difference between MUSE images of several compact galaxies within a pointing to 2D Gaussian-convolved and to MUSE resolution resampled HST images of those galaxies. For each image the FWHM of the 2D Gaussian kernel minimising  $\chi^2$  is displayed. Here the used images are created by summing over  $200 \text{ \AA}$  along the spectral axis. The green and the blue lines are the linear fits  $FWHM(\lambda) = p_0 + p_1(\lambda - 7000 \text{ \AA})$  to the individual data points of the FWHM determination using a star or several compact galaxies, respectively. The purple point at  $7000 \text{ \AA}$  is the mean value inferred from the VLT auto guider probe (AG seeing) averaged over all four exposures. As the green points and line are closer to the AG seeing value, we chose this fit to describe the  $FWHM(\lambda)$  dependence.

$\sqrt{2k/(k^2 + 1)}$  if instead of the optimal  $v_{FWHM, \text{correct}}$  a different  $v_{FWHM, \text{incorrect}}$  is chosen, where  $k = v_{FWHM, \text{incorrect}}/v_{FWHM, \text{correct}}$  (HW17; see also Zackay & Ofek 2017). Moreover, for the same reason the resulting S/N is also quite robust against moderate

**Table 2.** Coefficients  $p_0$  and  $p_1$  of  $FWHM(\lambda) = p_0 + p_1(\lambda - 7000 \text{ \AA})$  that describe the wavelength dependence of the PSF  $FWHM$  in each of the datacubes.

| Pointing<br>No. | $p_0$<br>[ $''$ ] | $p_1$<br>[ $10^{-5} ''/\text{\AA}$ ] | Method                               |
|-----------------|-------------------|--------------------------------------|--------------------------------------|
| 01              | 0.836             | -4.429                               | fit to ID 10548, $m_{F814W} = 22.18$ |
| 02              | 0.940             | -3.182                               | compact galaxies, no star            |
| 03              | 0.944             | -4.460                               | fit to ID 8374, $m_{F814W} = 19.39$  |
| 04              | 0.747             | -4.219                               | compact galaxies, no star            |
| 05              | 1.026             | -3.003                               | compact galaxies, no star            |
| 06              | 0.835             | -4.332                               | compact galaxies, no star            |
| 07              | 0.935             | -3.966                               | compact galaxies, no star            |
| 08              | 0.991             | -5.007                               | compact galaxies                     |
| 09              | 0.833             | -8.069                               | fit to ID 68879, $m_{F814W} = 19.73$ |
| 10              | 0.890             | -3.051                               | compact galaxies                     |
| 11              | 0.989             | -3.771                               | compact galaxies                     |
| 12              | 1.020             | -4.123                               | compact galaxies, no star            |
| 13              | 1.063             | -5.285                               | compact galaxies                     |
| 14              | 0.884             | -4.844                               | compact galaxies                     |
| 15              | 0.702             | -4.441                               | fit to ID 5744, $m_{F814W} = 20.3$   |
| 16              | 0.859             | -3.784                               | fit to ID 6475, $m_{F814W} = 18.86$  |
| 17              | 0.780             | -3.534                               | compact galaxies                     |
| 18              | 0.929             | -3.479                               | compact galaxies, no star            |
| 19              | 0.814             | -3.524                               | compact galaxies, no star            |
| 20              | 0.713             | -5.196                               | compact galaxies, no star            |
| 21              | 0.836             | -4.255                               | fit to ID 9801, $m_{F814W} = 21.92$  |
| 22              | 0.788             | -3.253                               | fit to ID 7813, $m_{F814W} = 20.23$  |
| 23              | 0.777             | -3.019                               | compact galaxies, no star            |
| 24              | 0.728             | -4.232                               | compact galaxies, no star            |

**Notes.** These coefficients are used as input parameters in LSDCat for the spatial filtering. In the method column we give the ID and  $F814$  magnitude from the Guo et al. (2013) catalogue when the polynomial coefficients were derived using this star. Otherwise we indicate with “compact galaxies”, that the minimisation utilising compact galaxies provided a  $p_0$  value closer to the AG Seeing and we therefore used  $p_0$  and  $p_1$  from this method. With “compact galaxies, no star” we indicate that no sufficiently bright star was present within the datacube, and we thus had to rely on the minimisation scheme utilising compact galaxies.

shape mismatches between the Gaussian template and the real emission line profiles.

Equipped with a set of carefully vetted parameters for the cross-correlation with the matched filter, we then applied the relevant LSDCat routines `lsd_cat_spatial.py` and `lsd_cat_spectral.py` to our 24 datacubes. These routines also propagate our empirically estimated variances accordingly. As a result, we obtained 24 new datacubes (S/N-cubes) that contained in each voxel the detection significance of an emission line being present at this position in terms of S/N.

### 3.1.4. Thresholding and cataloguing of emission line source candidates

LSDCat collects emission line candidates by thresholding the S/N-cubes from the matched-filtering procedure. This task is performed by the routine `lsd_cat.py`, which collects all detections in the form of a catalogue containing their peak S/N value, the 3D coordinate of the peak, and the numbers of voxels constituting the detection cluster.

In theory, i.e. for perfect data without artifacts, the voxel values of the matched-filter processed cubes are directly related to the probability of rejecting the null hypothesis of no source

signal being present at a voxel position. Under this assumption the choice of the S/N threshold can directly be related to the number of expected false detections in a datacube from its total number of voxels. In practice, however, we have to face the difficulties of possible unknown systematics in the data and the limitations of our empirical noise estimate (Sect. 3.1.1). Moreover, as we describe in Sect. 3.2, our source classification scheme is semi-automatic and requires a careful visual inspection of most sources. This necessitates a low ratio of spurious to real detections.

By successively lowering the detection threshold and visual inspection of the detected sources, we found that the number of unclassifiable (likely spurious) sources increased rapidly when we lowered the S/N threshold below eight. This value represents the *point of diminishing returns* that we adopted for our emission line search.

### 3.2. Classification and cleaning

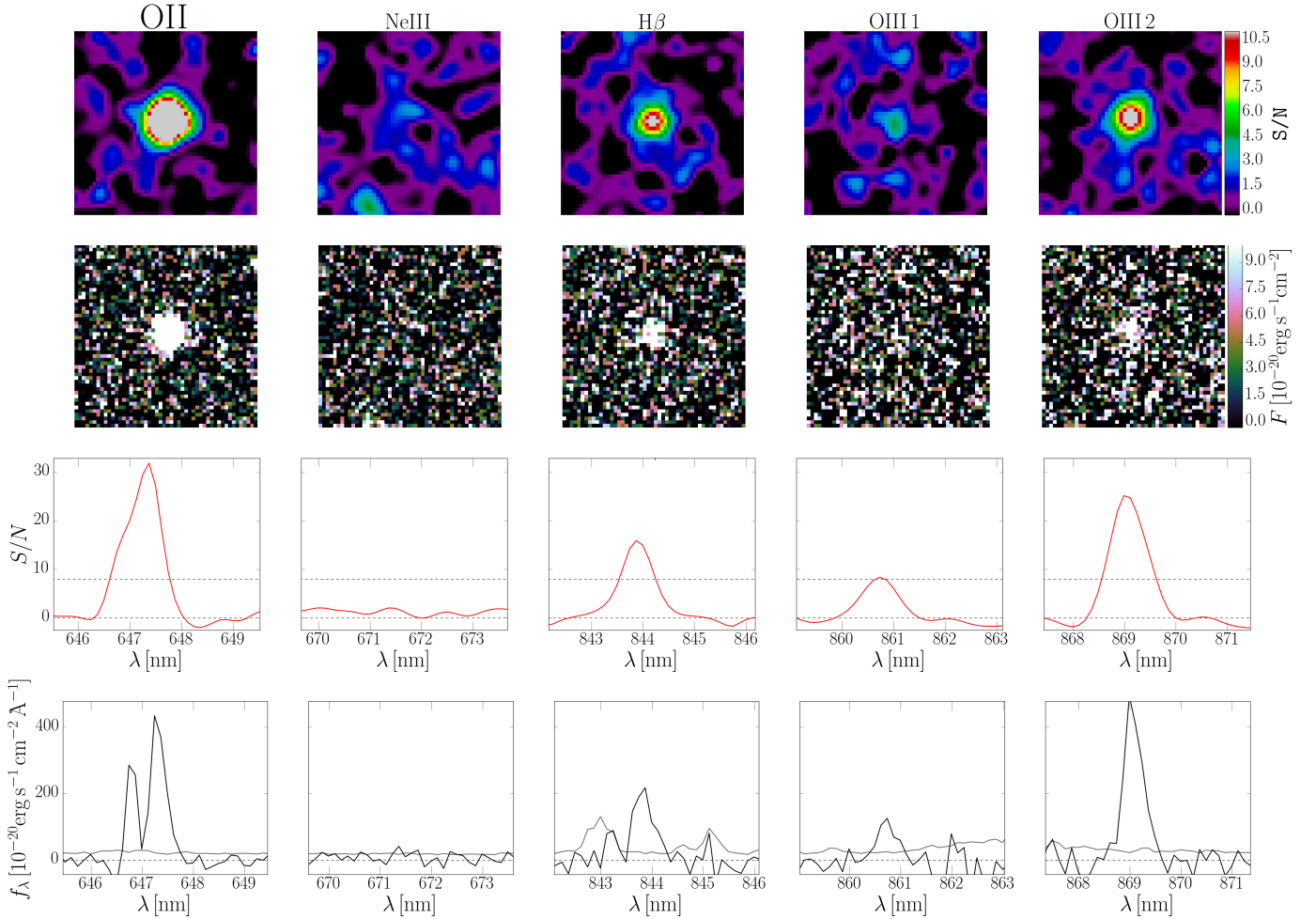
With the S/N threshold of eight LSDCat provided us with a catalogue of 2603 line detections over all 24 fields. The process of identifying those individual detections, grouping them into to objects, as well as purging spurious detections from the initial catalogue was the obvious next step.

LSDCat groups multiple line detections together if they are within a search radius of  $0.8''$  on the sky. Our initial catalogue contained 374 groups consisting of two or more detections and 642 spatially isolated single detections. We inspected those emission line groups and single line detections with an interactive graphical tool QtClassify developed especially for this task. The functionality and appearance of QtClassify are detailed in Appendix A. With this tool all groups and single line detections were inspected independently by three investigators (ECH, JK, and TU). Afterwards these individual classifications were consolidated into a final classification for each object. During consolidation a final quality and confidence value were assigned to each detection. While our quality value indicates the amount of objective information within the datacube and the cross-correlated S/N datacube that aided the decision process for a particular classification, our confidence value is a more subjective measure of “belief” towards the final classification.

In detail we assign quality flags A, B, and C according to the following criteria:

- **Quality A:** multiple emission lines were detected at the same location on the sky and it was possible to anchor a unique redshift solution for the object. These unambiguous identifications were assigned automatically in QtClassify but were confirmed by visual cross-checks. We show an example of a quality A classified object in Fig. 5. There are 288 “Quality A” objects in the final catalogue (34.7%).
- **Quality B:** only one emission line was detected above the detection threshold. However, one or more other lines below the detection threshold could also be seen. These, mostly unambiguous, identifications were assigned manually in QtClassify. We show an example of a quality B classified object in Fig. 6. There are 117 “Quality B” objects in the final catalogue (14%).
- **Quality C:** only one emission line was detected and no secondary lines are visible. The identification was based on the visual appearance of the line (e.g. double peak matches [O II]  $\lambda\lambda 3726, 3728$  profile, characteristic Ly $\alpha$  profile shape), and the appearance of the object in the HST CANDELS images. Prior information on the photometric redshifts was not





**Fig. 5.** Example of a quality A (confidence 3) object (MUSE-Wide ID 108025145 at  $z = 0.74$ , strongest line in S/N is O II). Multiple emission lines from this galaxy are detected above the detection threshold  $S/N_{\text{thresh}} = 8$ . *First row:* individual layers from the S/N cube after matched filtering with LSDCat. The position of the layers is chosen to match the classified redshift of the object. *Second row:* pseudo narrow-band images created by integration over 5 Å around the expected position of the emission lines in the continuum subtracted datacube. *Third row:* corresponding segments of the S/N cube with the highest S/N peak. The dotted line indicates the detection threshold  $S/N_{\text{thresh}} = 8$ . *Fourth row:* corresponding segments of an aperture extracted spectrum (aperture radius =  $0.7'' = 3.5$  spectral pixels) from the flux datacube. The grey line shows our empirically determined noise estimate.

included. We show an example of a quality C classification in Fig. 7. There are 426 “Quality C” objects in the final catalogue (51.3%). For these objects the confidence measure indicates how sure the investigators are on a certain classification.

The confidence values 3, 2, and 1 indicate the following:

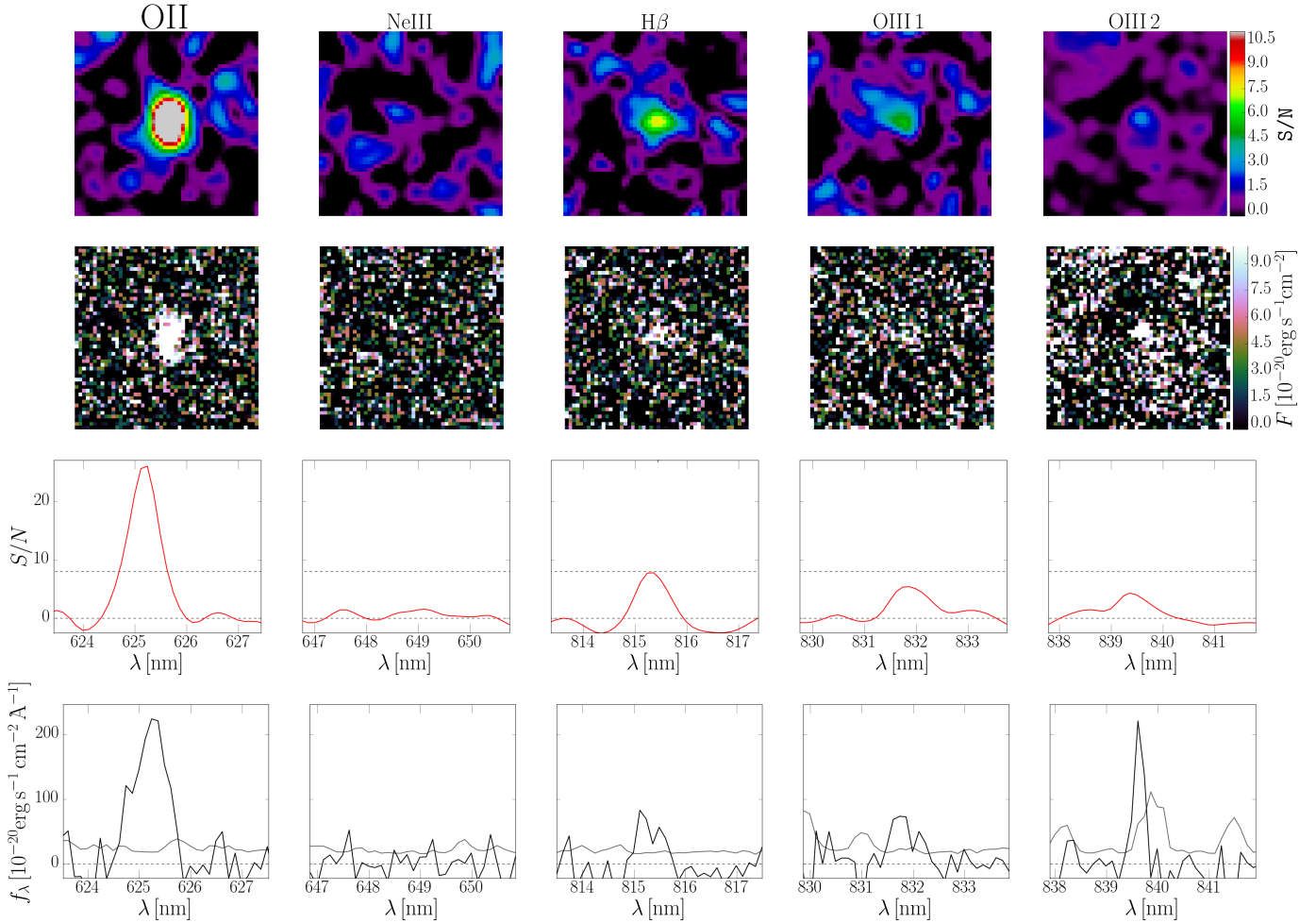
- *Confidence 3:* we are certain of the classification. All quality A and most quality B identifications have confidence 3. For single line detections these objects are characterised by a very well resolved [O II] double-peak or a characteristic Ly $\alpha$  profile. In total we have 578 objects marked with a confidence value 3 in the final catalogue (70%).
- *Confidence 2:* we are quite sure of the classification, but not with the same degree of certainty as for confidence 3. Often the reason was simply a somewhat lower S/N of the detection, or a remaining (however slight) possibility that another line might mimic the appearance of the classified line. Such cases were discussed among the classifiers in the consolidation process. Often the line profile had to be evaluated in detail and examined in apertures of different sizes in

QtClassify. In some cases also CANDELS images and photometry were inspected in the consolidation process. However, we did not consult any photometric redshift catalogues. 210 objects in the final catalogue are marked with confidence 2 (25.3%).

- *Confidence 1:* we are unsure regarding the classification but are certain that the detection is not spurious. Often the classifiers initially disagreed on the classification. Inspection of the line profile and HST imaging data did not resolve the uncertainty. The final classification represents our “best guess”. These lines usually show the lowest S/Ns. Only 43 objects, i.e. 5.2% of all objects in our catalogue, are in this category.

To illustrate our confidence measure we show several line profiles of emission lines that were classified either as Ly $\alpha$  or [O II] in Fig. 8.

During the consolidation session we also flagged spurious detections (e.g., telluric line residuals, or detections caused by increased noise near the FoV borders) and detections caused by residuals due to imperfect continuum subtraction. Only 74 of the 2603 line detections in the initial catalogue were flagged as spurious. Moreover 636 line detections were caused by contin-



**Fig. 6.** Similar to Fig. 5, but for a quality B (confidence 3) object (MUSE-Wide ID 107021114, strongest line in S/N is O II). Only one emission line is detected above the detection threshold, but other lines are clearly visible in the LSDCat generated S/N cube and in the flux datacube. In this object the [O III]  $\lambda 5007$  line falls on a sky emission line (as can be seen by the increased noise level in the *bottom right panel*), so its S/N is lower compared to the intrinsically weaker [O III]  $\lambda 4959$  line.

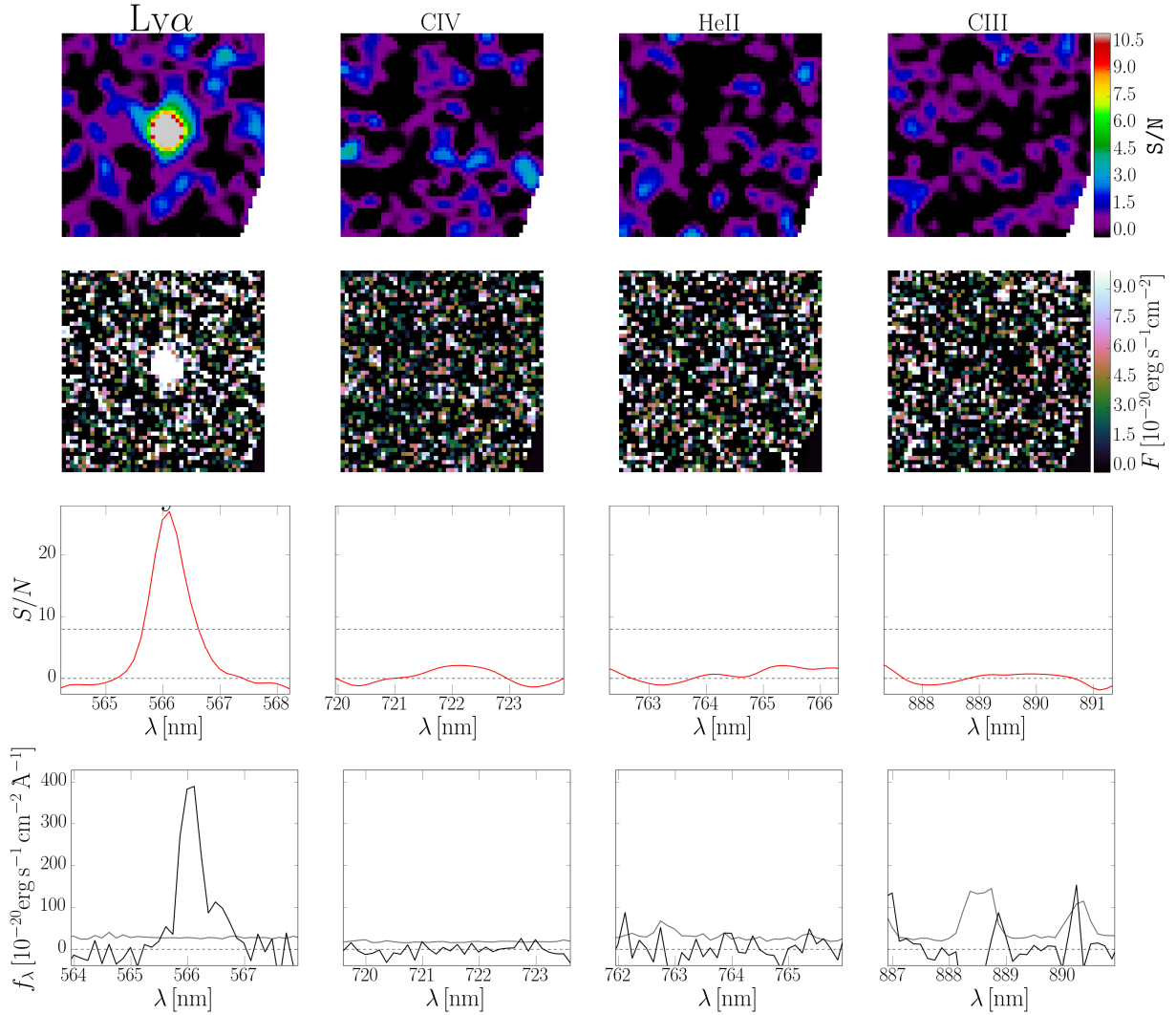
uum residuals, most of them were bundled up in very few objects, cold stars or bright early-type galaxies for which the median filter subtraction does a poor job of removing the continua, with 30 or more detections each. These detections were removed from the final emission line table. Moreover, due to the overlap of the pointings, some sources were detected twice in adjacent pointings. In the final catalogue we tabulate for such sources only the quantities determined for the detections in the pointing where the source is located farthest away from the edge, where the measurements are less affected by possible edge effects. Furthermore, some low- $z$  galaxies tended to fragment into detections of, e.g., multiple H II regions. Such fragmented detections were manually merged into single objects. The removal of double-detections and manual cleaning of fragmented objects removed 241 emission line detections from the initial catalogue. Finally, it turned out that eight of the inspected line groups (i.e. multiple detections within a  $0.8''$  radius) are not one object, but are a by chance superposition of two objects at different redshifts.

By construction, all objects with quality flag A have a confidence value of 3 in the final catalogue. Most of the objects with quality B also have a confidence value of 3 (107), but for a few objects (10) the S/N of the additional lines was very low so that we assigned them with a confidence value of 2. None of the qual-

ity B objects has a confidence value 1. In the quality C class 183 of the 426 single line detections got assigned a confidence value of 3, 200 of them got assigned a confidence value of 2, and for 43 objects we were unsure regarding the final classification (confidence value 1).

After the consolidation and cleaning steps of the initial catalogues described above we arrive at a final catalogue of 1652 emission lines from 831 emission line galaxies.

LSDCat determines the positions, spatial extents, and fluxes of detected emission lines with the routine `lsd_cat_measure.py`. As source positions in our catalogue we primarily use the first central moments that are determined in a pseudo narrow-band image, generated from summing over several layers in the matched-filtered version of the datacube. The layers used in this summation are given by the spectral coordinates of voxels that are above a certain analysis threshold  $S/N_{\text{ana}}$  in the S/N-cube. After visual inspection of the line profiles, we found that setting  $S/N_{\text{ana}} = 3$  delivers a band that is optimally suited for almost all emission lines. However, currently LSDCat does not offer an algorithm to deblend close-by neighbouring sources. For those sources the first central moments can be ambiguous. In such cases we tabulate as primary coordinate the S/N peak position introduced in Sect. 3.1.4. These cases are identified by searching in the



**Fig. 7.** Similar to Fig. 5, but for a quality C (confidence 3) object (MUSE-Wide ID 104015052, only one detected line). No other lines were found in the datacube. Also no veto lines were found if we would assume the detected line is an [O II] emission line. Based on the characteristic profile of the emission line we confidently classified it as Ly $\alpha$ .

LSDCat output catalogues for detections where the S/N-peak position differs significantly from the first-moment coordinate ( $\geq 0.5''$ ; cf. Sect. 5.2, where we comment on the astrometric precision of the catalogue). More details on the available coordinates per emission line are given in Sect. 4 where we describe the contents of the final source catalogue.

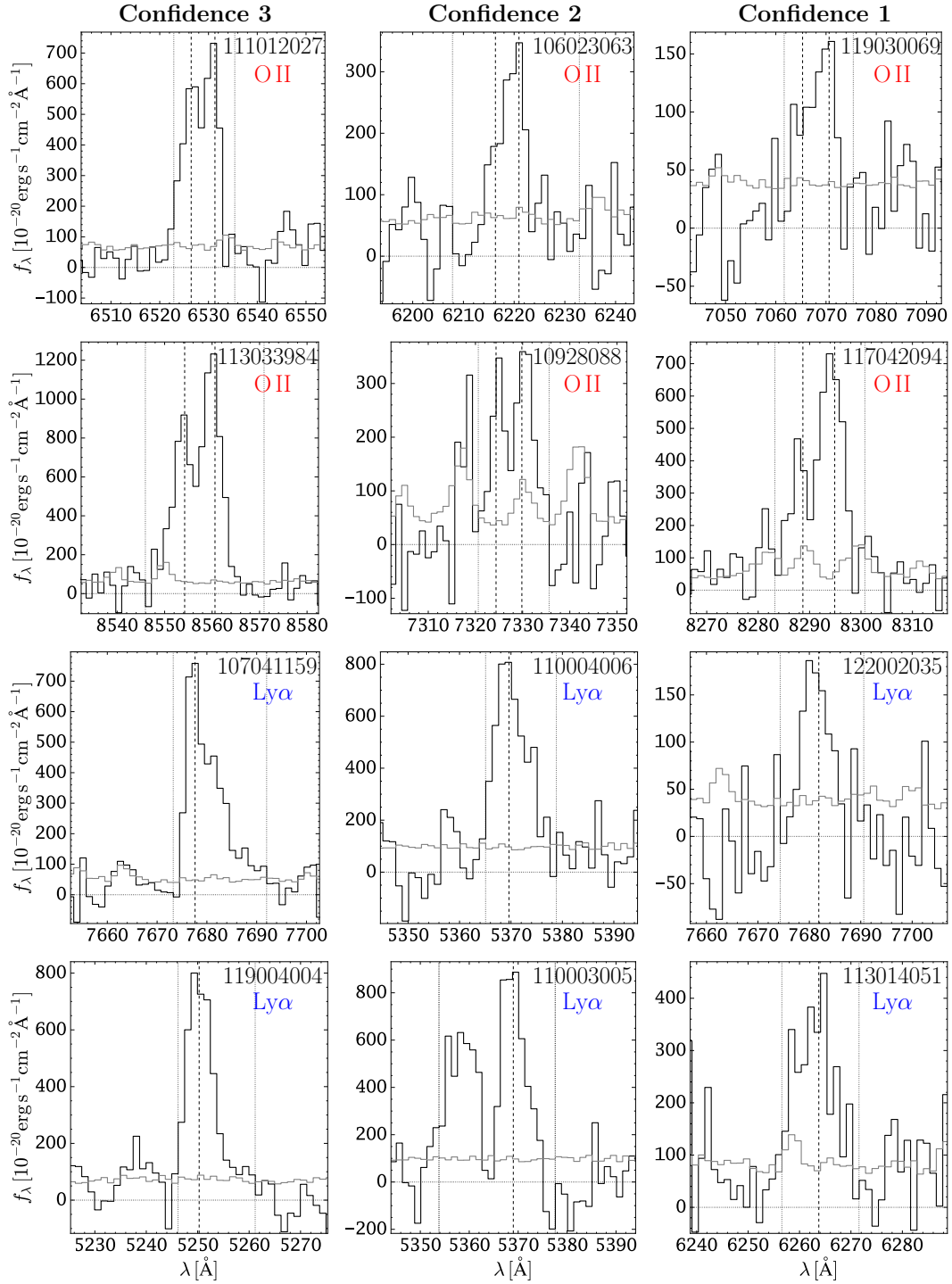
LSDCat also measures the spatial extents of our detections by calculating the characteristic light distribution weighted radius introduced by Kron (1980). To calculate the Kron radius  $R_{\text{Kron}}$  the LSDCat algorithm follows closely the SExtractor implementation (Bertin & Arnouts 1996). The calculation is performed on the same pseudo narrow-band images that were used for the determination of the centroids above. Since for low-S/N detections the Kron radius can become erroneously small, we limit the boundary to  $R_{\text{Kron}}^{\text{min}} = 0.6''$  in such cases. This ensures that the smallest aperture diameter for the flux measurement, described below, is always larger than the FWHM of the seeing disk.

Finally, we used LSDCat to measure the fluxes of all emission lines. To do this, first, the algorithm creates pure line emission images by summing up layers containing only the emission line signal. As above, the bandwidth of these images is given

by the spectral coordinates of the voxels above  $S/N_{\text{ana}} = 3$  in the S/N-cube. We then integrate the flux in these images within  $k \times R_{\text{Kron}}$  apertures, with  $k = 1, k = 2, k = 3$ , and  $k = 4$ . The  $k = 3$  aperture is expected to contain  $>95\%$  of the total flux for compact sources whose light-profile is mainly determined by PSF broadening (e.g. Graham & Driver 2005). Moreover, we show in the LSDCat publication (HW17) that automatically determined fluxes based on the LSDCat  $k = 3$  aperture compare well with manually determined fluxes based on a curve-of-growth method.

We caution that for some double peaked Ly $\alpha$  emission lines in our catalogue we find that the spectral width determined by LSDCat does not always encompass the weaker bluer peak of the profile in its entirety (e.g., object 110003005 in Fig. 8). In a few cases it even misses the blue peak completely (e.g., objects 107041159, 11004006, and 119004004 in Fig. 8). In particular, 87 of our 237 Ly $\alpha$  emitters show a blue peak and in 53 of those the blue peak is not or only partially included in the automatically determined flux integration bandwidth. Fitting the blue and red peak simultaneously in the 1D extracted spectra (described in Sect. 4.3), we find that the average flux loss for those 53 objects is 10%. Rather than manually changing the spectral integration width for those sources we opt for providing a homogenised set





**Fig. 8.** Six representative [O II] and Ly $\alpha$  emission line profiles from our sample. The first, second, and third column include objects with confidence flag 3, 2, and 1, respectively. In each panel we show extracted spectra from the datacube in black and their corresponding variances in grey. The dashed vertical lines mark the observed wavelengths of the [O II] or Ly $\alpha$  emission according to our determined redshift for that object (see Sect. 3.3). The dotted vertical lines mark the window used for flux integration by LSDCat (see Sect. 3.2). In each panel the object’s MUSE-Wide ID (cf. Sect. 4.1) is indicated in the top right corner.

of automatically determined flux measurements. We will address the accurate measurement of LAE fluxes in a forthcoming publication. Moreover, the fluxes in automatically determined  $3R_{Kron}$  apertures are also not robust for galaxies that are exceptionally extended or have close by companions. For such objects emission line flux ratios will likely be distorted. Another caveat of the provided fluxes is, that for all except eight [O II] detections

the  $\lambda\lambda 3726, 2729$  Å doublet is detected as a single line, thus the tabulated flux is integrated over both lines. We encourage users interested in more accurate emission line flux measurements to exploit the information contained in the 3D source datacubes that we provide with this catalogue (Sect. 4.4).

### 3.3. Redshift measurements

The line detection and identification gives us an approximate measure of the redshift which we improved as described in the following. To accurately measure redshifts of the 831 emission line galaxies in our sample we use 1D spectra that we extracted for each of our objects. These spectra are released with the catalogue and the extraction is described in Sect. 4.3. Depending on whether the source is a high- $z$  LAE or a low- $z$  galaxy detected by its rest-frame optical emission line we employed different emission line fitting strategies. We describe the fitting method in the following two subsections. All our redshifts are vacuum redshifts within the barycentric reference frame.

#### 3.3.1. Determining Lyman $\alpha$ galaxy redshifts

To determine the redshifts of the LAEs we fitted the Ly $\alpha$  line profiles with the formula

$$f(\lambda) = A \times \exp \left\{ -\frac{(\lambda - \lambda_0)^2}{2 \times (a_{\text{asym}}(\lambda - \lambda_0) + d)^2} \right\} \quad (1)$$

introduced by Shibuya et al. (2014). Equation (1) describes an asymmetric Gaussian profile used to fit Ly $\alpha$  profiles. Shibuya et al. (2014) argue that Eq. (1) provides a more robust peak wavelength for the typical LAE profiles than a simple Gaussian. The free parameters  $A$ ,  $\lambda_0$ ,  $a_{\text{asym}}$ , and  $d$  in our fit to Eq. (1) are the amplitude, the peak wavelength, the asymmetry parameter, and the typical width of the line, respectively. It is known and commonly attributed to the complex Ly $\alpha$  radiative transfer physics that Ly $\alpha$  redshift measurements are systematically offset by  $\sim 100$ – $200$  km s $^{-1}$  with respect to the systemic redshift determined from rest-frame optical emission lines (e.g. McLinden et al. 2011; Rakic et al. 2011; Chonis et al. 2013; Erb et al. 2014; Song et al. 2014; Hashimoto et al. 2015; Trainor et al. 2015). No correction for such offsets was applied here.

In practical terms, the fitting was performed in a window around the peak with flux values being greater than 10% of the peak flux. For the 87 Ly $\alpha$  emitters in our sample that show a double peaked profile we restricted the fit to the region of the red peak. For 11 objects we had to adjust the window size manually to avoid strong sky-subtraction residuals. We accounted for a possible continuum by subtracting a running median from the 1D spectrum. Finally, we determined the error on each redshift by repeating the fitting procedure 100 times on random realisations of the spectra generated by perturbing each spectral pixel according to the noise statistics of that pixel from the associated error spectrum.

#### 3.3.2. $z \lesssim 1.5$ galaxies

Emission line galaxies at  $z \lesssim 1.5$  are detected in MUSE datacubes by the typical strong rest-frame optical emission lines of star-forming galaxies (e.g. Kennicutt 1992): [O II]  $\lambda\lambda 3276, 3278$  (detected in 472 objects), [O III]  $\lambda\lambda 4958, 5006$  (detected in 310 objects), H $\beta$  (detected in 184 objects), and H $\alpha$   $\lambda 6563$  (detected in 73 objects). In Table 5 we list the air- and vacuum wavelengths of these transitions.

To determine redshifts we fitted 1D Gaussian profiles to those emission lines. The [O II]  $\lambda\lambda 3276, 3278$  doublet was fitted by a double component Gaussian with fixed separation and free intensity ratio, while all other lines were fitted with single components. The continuum was subtracted with a 151 pixel wide

**Table 3.** Columns of the object table.

| Column name | Short description   | Example entry <sup>a</sup> |
|-------------|---|----------------------------|
| UNIQUE_ID   | Unique MUSE-Wide object ID                                | 101001006                  |
| RA          | Right ascension $\alpha_{J2000}$ [deg]                    | 53.060185...               |
| DEC         | Declination $\delta_{J2000}$ [deg]                        | −27.813464...              |
| Z           | Redshift  | 0.310564...                |
| Z_ERR       | Error on the redshift                                     | 0.000018...                |
| LEAD_LINE   | Highest S/N detected line <sup>b</sup>                    | Ha                         |
| SN          | S/N of the LEAD_LINE                                      | 76.106950...               |
| QUALITY     | Quality flag (Sect. 3.2)                                  | a                          |
| CONFIDENCE  | Confidence value (Sect. 3.2)                              | 3                          |
| OTHER_LINES | Other detected lines <sup>b</sup>                         | 02, Hg, Hb...              |
| GUO_ID      | Associated source in<br>Guo et al. (2013) catalogue       | 10720                      |
| GUO_SEP     | Angular separation to<br>Guo et al. (2013) source ["]     | 0.37                       |
| SKELTON_ID  | Associated source in<br>Skelton et al. (2014) catalogue   | 20736                      |
| SKELTON_SEP | Angular separation to<br>Skelton et al. (2014) source ["] | 0.25                       |

<sup>(a)</sup> Entries containing ... are truncated compared to the original format of the table. <sup>(b)</sup> Emission line identifiers are explained in Table 4.

running median in spectral direction. For objects having several detected emission lines we computed the S/N-weighted mean redshift of all emission line fits. The error on the redshift was determined by repeating the fitting procedure 100 times on realisations of the spectra generated by perturbing each pixel according to the noise statistics from the error spectrum.

## 4. Source catalogue, spectra, and datacubes

With this publication we provide the following data products:

- A catalogue of all 831 detected emission line objects.
- A table of all 1652 detected emission lines in those objects.
- 1D PSF-weighted extracted spectra of the emission line objects.
- 3D datacubes of the emission line objects.

The tabular data is available in its entirety at the CDS. Moreover, this data is also available via the MUSE-Science website<sup>6</sup> and via the CDS, where also the 1D spectra and 3D datacubes are stored. In the following subsections we describe these data products in detail.

### 4.1. Object table

In Table 3 we present the columns of the catalogue that contains all 831 detected emission line galaxies in the first 24 pointings (22.2 arcmin<sup>2</sup>) of the MUSE-Wide survey. The details of the 14 columns are given below:

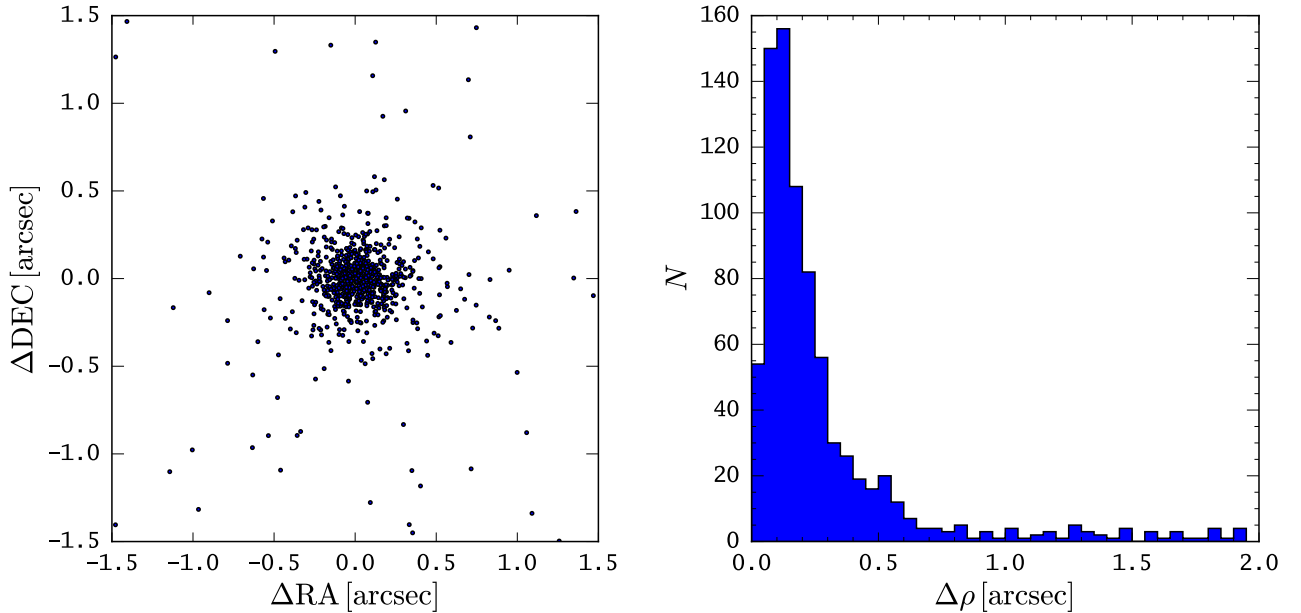
- UNIQUE\_ID contains a unique MUSE-Wide ID. This ID is composed of nine digits divided into four groups of the format ABBCCDDDD. Here A designates the MUSE-Wide survey-area (1≡ECDF-S CANDELS/Deep, 2≡COSMOS CANDELS (not used in the catalogue described here), or other numbers for future MUSE-Wide regions), BB indicates the

<sup>6</sup> <http://muse-vlt.eu/science>

**Table 4.** Columns of the emission line table.

| Column                          | Unit(s)                             | Description  |
|---------------------------------|-------------------------------------|--|
| UNIQUE_ID                       | —                                   | Unique MUSE-Wide object ID   |
| POINTING_ID                     | —                                   | Pointing Number (see Fig. 1)   |
| OBJ_ID                          | —                                   | Object ID – only unique per pointing   |
| RID                             | —                                   | Running ID – only unique per pointing  |
| IDENT                           | —                                   | Line identification  |
| COMMENT                         | —                                   | Free-form comment added during classification and cleaning (Sect. 3.2)                             |
| SN                              | —                                   | Detection significance (Sect. 3.1.3)   |
| {RA, DEC, LAMBDA}_SN            | deg / Å                             | 3D S/N-weighted position   |
| {RA, DEC, LAMBDA}_PEAK_SN       | deg / Å                             | S/N-peak position  |
| {RA, DEC}_1MOM                  | deg                                 | First central moment coordinate, determined in optimal narrow band image (Sect. 3.2)               |
| LAMBDA_NB_{MIN, MAX}            | Å                                   | Minimum- and maximum wavelength of optimal narrow band image. Used for flux integration            |
| R_KRON                          | arcsec                              | Kron-Radius (see Sect. 3.2)  |
| F_KRON, F_{2, 3, 4}KRON         | $10^{-20}$ erg s $^{-1}$ cm $^{-2}$ | Flux extracted in ( $k\times$ ) R_KRON aperture within narrow band defined by LAMBDA_NB_{MIN, MAX} |
| F_KRON_ERR, F_{2, 3, 4}KRON_ERR | $10^{-20}$ erg s $^{-1}$ cm $^{-2}$ | Error on the extracted flux  |
| BORDER_FLAG                     | —                                   | Flag indicating whether $3\times R_{\text{Kron}}$ overlaps with FoV border                         |

**Notes.** Comma separated list within curly braces in the first column indicate the set of similar columns. Wavelengths are vacuum wavelengths.



**Fig. 9.** *Left panel:* relative differences in right ascension and declination between object coordinates in the MUSE-Wide catalogue and objects with the closest on-sky separation in the 3D-HST catalogue (Skelton et al. 2014). *Right panel:* angular separation  $\Delta\rho$  for between MUSE-Wide objects and most closely separated 3D-HST sources. 83% of the objects from our MUSE-Wide catalogue have cross-matches within  $0.5''$  in the Skelton et al. (2014) catalogue.

pointing number (here 01–24, see Fig. 1), CCC refers to the per-pointing object ID, and DDD to the running ID of the strongest line. These last two identifiers relate to the emission-line table explained in Sect. 4.2.

- RA and Dec contain the position of the galaxy in right-ascension  $\alpha_{J2000}$  and declination  $\delta_{J2000}$ . For most sources this position is given as first central moments determined in an adaptive narrow-band image (Sect. 3.2) of the lead line (see column LEAD\_LINE description below). However, for 40 sources this position differed by more than  $0.5''$  to the peak S/N position found in the initial thresholding step

(Sect. 3.1.4). Visual inspection revealed that these cases are often affected by blends with neighbouring sources. Since the peak S/N-coordinate is less affected by blending, for those 40 cases we replace the first central moment coordinate with the peak S/N coordinate. We also provide more positional parameters per emission line detection in Sect. 4.2.

- Z contains the redshift  $z$  for each galaxy and column Z\_ERR contains the error on this quantity as explained in Sect. 3.3.
- LEAD\_LINE contains the lead line for each galaxy. The lead line is defined as a galaxy’s emission line that has the highest S/N after the matched-filtering process (Sect. 3.1.3). The



**Table 5.** Detected emission lines in the catalogue.

| Transition(s)                      | IDENT <sup>a</sup> | Air wavelength(s) <sup>b</sup><br>[Å] | Vacuum wavelength(s) <sup>b</sup><br>[Å] | Detections | Comments <sup>c</sup> |
|------------------------------------|--------------------|---------------------------------------|--|------------|-----------------------|
| O VI $\lambda\lambda 1032, 1038$   | O6_2               | ...                                   | 1031.93, 1037.62                         | 2          | (1)                   |
| Ly $\alpha$ $\lambda 1216$         | Lya                | ...                                   | 1215.670                                 | 242        | (2), (3)              |
| N V $\lambda\lambda 1238, 1242$    | N5                 | ...                                   | 1238.821, 1242.804                       | 1          | (4)                   |
| C IV $\lambda\lambda 1549, 1550$   | C4                 | ...                                   | 1548.203, 1550.777                       | 3          | (3), (4)              |
| Mg II $\lambda 2796$               | Mg2                | 2795.528                              | 2796.290                                 | 1          |                       |
| [O II] $\lambda\lambda 3726, 3729$ | O2                 | 3726.032, 3728.815                    | 3727.048, 3729.832                       | 480        | (3), (5)              |
| [Ne III] $\lambda 3869$            | Ne3                | 3869.06                               | 3870.115                                 | 49         |                       |
| [Ne III] $\lambda 3968$            | Ne32               | 3967.79                               | 3968.872                                 | 3          | (6)                   |
| H $\zeta$ $\lambda 3889$           | Hzet               | 3889.049                              | 3890.109                                 | 2          |                       |
| H $\epsilon$ $\lambda 3970$        | Heps               | 3970.072                              | 3971.154                                 | 1          |                       |
| H $\delta$ $\lambda 4102$          | Hd                 | 4101.734                              | 4102.852                                 | 18         |                       |
| H $\gamma$ $\lambda 4340$          | Hg                 | 4340.464                              | 4341.647                                 | 74         |                       |
| [O III] $\lambda 4363$             | O3_3               | 4363.210                              | 4364.400                                 | 2          |                       |
| H $\beta$ $\lambda 4861$           | Hb                 | 4861.325                              | 4862.650                                 | 183        | (3)                   |
| [O III] $\lambda 4959$             | O3_1               | 4958.911                              | 4960.263                                 | 135        | (3)                   |
| [O III] $\lambda 5007$             | O3_2               | 5006.843                              | 5008.208                                 | 304        | (3)                   |
| He I $\lambda 5876$                | He1                | 5875.615                              | 5877.217                                 | 6          |                       |
| [O I] $\lambda 6300$               | O1                 | 6300.304                              | 6302.022                                 | 3          |                       |
| [N II] $\lambda 6549$              | N2_1               | 6548.04                               | 6549.825                                 | 4          |                       |
| H $\alpha$ $\lambda 6563$          | Ha                 | 6562.80                               | 6564.589                                 | 73         | (3)                   |
| [N II] $\lambda 6583$              | N2_2               | 6583.46                               | 6585.255                                 | 19         |                       |
| [S II] $\lambda 6717$              | S2_1               | 6716.44                               | 6718.271                                 | 30         |                       |
| [S II] $\lambda 6730$              | S2_2               | 6730.81                               | 6732.645                                 | 17         |                       |

**Notes.** <sup>(a)</sup> Emission line identifier code used in the IDENT column of the emission line table described in Sect. 4.2. <sup>(b)</sup> Wavelengths are from the “Atomic Line List” compiled by P. van Hoof: <http://www.pa.uky.edu/~peter/atomic/>. Wavelengths longward of 2000 Å were queried in air wavelengths and converted to vacuum wavelengths with the formula used in the Vienna atomic line database (Ryabchikova et al. 2015): <http://www.astro.uu.se/valdwiki/Air-to-vacuum%20conversion>. <sup>(c)</sup> (1): Doublet in single object resolved into two individual line detections. (2): 237 LAEs, with five having the double peak profile resolved into two individual line detections. (3): Used for redshift determination. (4): Doublet unresolved, i.e. always detected as single line. (5): In all except eight detections (103035109, 105019069, 109021072, 109023081, 114015092, 115046179, 124009024, and 124040076) the [O II]  $\lambda\lambda 3726, 3729$  doublet is unresolved. (6): Blend with H $\epsilon$ .

lead line is therefore not necessarily the line with the highest flux. Potential identifiers encountered in the LEAD\_LINE column are tabulated in Table 5.

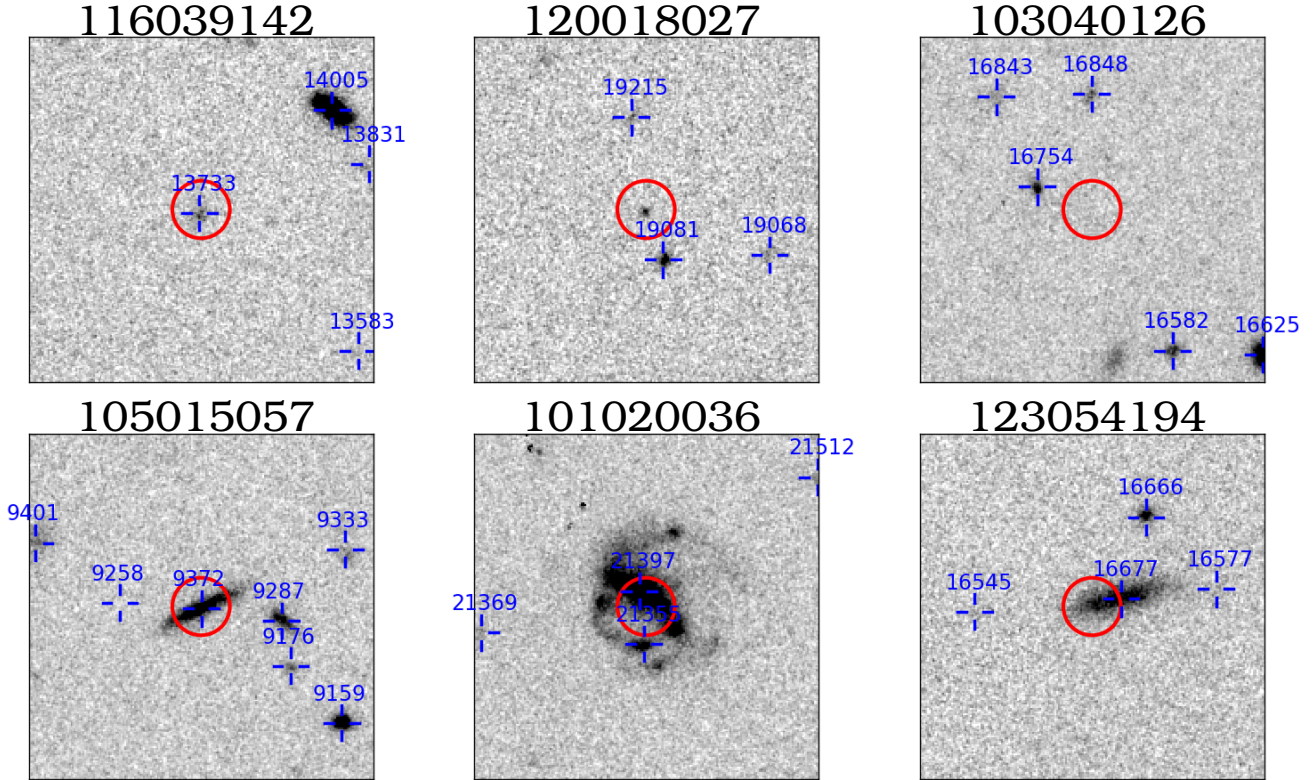
- SN contains the S/N of the lead line.
- QUALITY and CONFIDENCE tabulate the quality and confidence values indicating the robustness of the object classification and line identification as described Sect. 3.2.
- OTHER\_LINES we provide a comma-separated list of all other emission lines that were detected above the detection threshold, but at a lower S/N than the lead line. This field is empty in case of a single line detection. Potential identifiers encountered in the OTHER\_LINES column are tabulated in Table 5 (see also Sect. 4.2 below).

In the last four columns of the object-table we provide cross-matches (and angular separations) to existing catalogues based on source detection in HST imaging of the CANDELS-GOODS-S field.

- GUO\_ID contains the object identifier of the cross-matched source from the CANDELS GOODS-S catalogue (Guo et al. 2013). Its value is set to 0 for objects where no counterpart could be assigned.
- GUO\_SEP gives the angular separation between the MUSE-Wide position and Guo et al. (2013) source.
- SKELTON\_ID contains the in object identifier of the cross-matched source from the 3D-HST/CANDELS photometric catalogue (Skelton et al. 2014). It is set to 0 for objects where no counterpart could be assigned.

- SKELTON\_SEP gives the angular separation between the MUSE-Wide position and the Skelton et al. (2014) source.

These cross-matches were obtained by first selecting for each object in our catalogue the corresponding object with the smallest on-sky separation in the photometric catalogues. In Fig. 9 we show the relative distances in right ascension and declination with respect to the Skelton et al. (2014) catalogue, as well as histogram of the angular separations  $\Delta\rho$ . Both panels of the figure look almost identical if we compare to the positions from the Guo et al. (2013) catalogue instead. The isotropic distribution around  $(\Delta RA, \Delta Dec) = (0, 0)$  indicates that there is no systematic off-set in the positions reported in our catalogue. The median  $\Delta\rho$  of this catalogue cross-match is 0.17'' (0.19'') and 83% (80%) of all MUSE-Wide objects have a photometric counterpart in Skelton et al. (2014), catalogue (the Guo et al. 2013 catalogue) within 0.5''. Visual inspection of F814W thumbnails with catalogue positions overlaid (see examples in Fig. 10) reveals that all photometric counterparts with  $\Delta\rho \leq 0.5''$  are certain associations. Based on this visual screening we included 34 and 36 additional counterparts with angular separations  $\geq 0.5''$  from the Skelton et al. (2014) and Guo et al. (2013) catalogue, respectively. For these sources either the broad-band photometric centroid appears to be offset from the emission-line centroid, or in the case of some low-redshift sources the photometric catalogues tend to separate one source into multiple detections (see, e.g., ID 101200036 in bottom row of Fig. 10).



**Fig. 10.** Six example MUSE-Wide emission line sources CANDELS HST F814W  $6'' \times 6''$  cutouts. The red circle of  $1''$  diameter is centred on our catalogued position, the blue crosses with labels indicate the positions and IDs given in the Skelton et al. (2014) 3D-HST photometric catalogue. The *top row* shows LAEs, one with a catalogued photometric counterpart, one with a photometric counterpart and no catalogue entry and one with no counterpart in the image. The *bottom row* shows three typical  $z < 1.5$  counterparts.

#### 4.2. Emission line table

For each of the 1652 detected emission lines listed in columns LEAD\_LINE and OTHER\_LINES of the object catalogue LSDCat outputs a set of measurements. We supplement the object catalogue by providing a table that contains these measurements as well as additional information for all emission lines. This table is available only at the CDS as a FITS table and contains the following columns, which are also briefly summarised in Table 4:

- UNIQUE\_ID contains the unique identifier for each detected galaxy in our survey (cf. Sect. 4.1). Therefore, this column establishes the link between the object table and the emission line table.
- OBJECT\_ID is a unique integer identifier for each identified object within a MUSE-Wide pointing. This identifier separates spatially overlapping objects, i.e. two galaxies at different redshifts but at the same position on the sky. This integer number, padded with zero digits on the left, comprise the digits CCC of UNIQUE\_ID. While in the present catalogue never more than 100 objects are within a pointing, we reserve the third digit for future MUSE-Wide catalogue publications to ensure a unified system across different MUSE-Wide catalogues.
- RID is the running index from our initial LSDCat catalogues after thresholding (Sect. 3.1.4) and thereby uniquely indexes each detected emission line in a pointing. This integer number, padded with zeros on the left, comprise the digits DDD of UNIQUE\_ID. Since we removed 951 of the detections in the classification and cleaning process of the initial catalogue (Sect. 3.2), the RID column is not a running integer index per pointing in the final catalogue.

- IDENT contains the identification of each catalogued emission line established in the classification process (Sect. 3.2). We encode line identifications in a string of four alphanumeric characters. A legend of all those emission lines identifiers, their common designations and their wavelengths is given in Table 5. In this table we also list the total number of detections of a particular emission line in our catalogue.
- COMMENT: during the consolidation of our classifications with QtClassify we added comments when necessary. These are stored in this column.
- BORDER\_FLAG contains a logical flag, indicating whether the 3 Kron-radii extraction aperture for the emission line flux measurement overlaps with the FoV borders of the particular pointing in which the flux was extracted.

The other columns in the emission line table stem directly from LSDCat:

- SN contains the detection significance of a particular emission line.
- RA\_SN, DEC\_SN, and LAMBDA\_SN contain the 3D S/N weighted right ascension, declination, and wavelength coordinate.
- RA\_PEAK\_SN, DEC\_PEAK\_SN, and LAMBDA\_PEAK\_SN contain the position of the emission line peak in the S/N cubes.
- RA\_1MOM, and DEC\_1MOM contain the first central image moment coordinate determined in a synthesised narrow band image of the emission lines (cf. Sect. 3.2).
- R\_KRON contains the Kron-radius of a line detection determined in the same synthesised narrow-band (cf. Sect. 3.2).

- LAMBDA\_NB\_MIN and LAMBDA\_NB\_MAX contain the maximum and minimum wavelength of the synthesised narrow band.
- F\_KRON, F\_2KRON, F\_3KRON, and F\_3KRON contain the integrated flux over the emission line in units of  $10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2}$  in the synthesised narrow-band within a circular aperture of radius  $R_{\text{Kron}}$ ,  $2 \times R_{\text{Kron}}$ ,  $3 \times R_{\text{Kron}}$ , and  $4 \times R_{\text{Kron}}$  respectively (see also Sect. 3.2).
- F\_KRON\_ERR, F\_2KRON\_ERR, F\_3KRON\_ERR, and F\_3KRON\_ERR contain the propagated errors on the flux measurements.

For algorithmic details on the above listed measurements we refer the reader to the LSDCat paper (HW17).

#### 4.3. 1D spectra

We provide a 1D spectrum for each of our objects. These are created by utilising a 2D weighted extraction in each spectral layer. As weights we adopt a normalised 2D Gaussian PSF, where we use the same parameterisation that we used for the matched-filtering explained in Sect. 3.1.3. For compact, unresolved sources, the spectral pixels of the so extracted 1D spectra give a nearly unbiased and S/N-optimised estimate of the total flux within each spectral layer. However, for extended sources (or for extended emission line profiles in compact sources), these spectra underestimate the total flux. These spectra should therefore only be used to get a first impression of the spectral characteristics of a source in our catalogue, while measurements should be performed on the source datacubes described in Sect. 4.4 below.

We distribute the 1D spectra as FITS binary tables with file names `spectrum_ABBCCDDDD.fits`, where ABBCCDDDD is the unique ID introduced in Sect. 4.1. The tables contain four columns: AIR\_WAVE and AIR\_VAC contain the air- and vacuum wavelength, respectively, in Ångstrom (see note below Table 5 for details on the applied conversion). FLUX and FLUX\_ERROR contain the flux and the propagated error in the extraction in units of  $10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1}$ .

#### 4.4. Source datacubes

For each emission line galaxy we also provide three-dimensional source datacubes. These source datacubes are created from the per-pointing datacubes whose reduction was described in Sect. 2. Each source datacube is centred on the coordinate given in column RA and DEC of the object table (Sect. 4.1). The spatial dimensions of the extracted cubes are set to encompass 4.5 times the maximum Kron-radius from the set of an object's detected emission lines. We do not remove or mask neighbouring objects that fall into the extracted FoV of the source datacubes. Moreover, if objects are close to the border of the FoV, the source datacubes contain voxels set to nan (not a number) in regions where no data were available.

We distribute these source datacubes as FITS files named `objcube_ABBCCDDDD.fits`, where ABBCCDDDD is the unique ID introduced in Sect. 4.1. These FITS files contain three HDUs. The first HDU stores the actual minicube in a 3D flux array in units of  $10^{-20} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ Å}^{-1}$ , while the second HDU stores the empirical noise estimate  $\sigma_{\text{emp}}$  (Sect. 3.1.1) as an 1D array, in the same units, for the pointing where the mini-cube was extracted.

The third HDU of the mini-cube FITS file contains a 2D exposure map of the same spatial dimensions as the mini-cube itself. Each pixel of this map shows the average number of

exposures that went into the corresponding spaxel. Since each MUSE-Wide cube consists of four exposures, the exposure map values range from zero to four. Values less than four are encountered either when single exposures were affected by cosmic rays on the detector and thus was masked out in the `muse_scipost` resampling process (Sect. 2), or near the FoV borders where several spaxels contain less than four exposures because of the drizzling pattern. Outside the FoV the values of the exposure map are zero.

### 5. Characteristics of the sample

While an extensive scientific exploration of the presented sample is beyond the scope of this paper, we provide in this section some basic characteristics of our emission line selected galaxies from the first 24 MUSE-Wide pointings.

#### 5.1. Redshift distribution

In Fig. 11 we show a histogram of the redshift-distribution obtained from the emission line fits described in Sect. 3.3. We have 595 galaxies at  $z < 2$  detected by their rest-frame optical emission lines and 238  $z > 2.95$  galaxies, of which 237 where detected by strong Ly $\alpha$  emission and a single object where the CIV had a higher S/N than Ly $\alpha$ . At the depth of MUSE-Wide the absence of strong nebular lines between [O II] and Ly $\alpha$  results in a “redshift desert” in the interval  $1.5 \lesssim z \lesssim 3$ .

#### 5.2. Redshift comparison with existing photometric and spectroscopic catalogues

We now compare our 831 emission line selected galaxies to existing photometric and spectroscopic redshift measurements from the literature within our survey area.

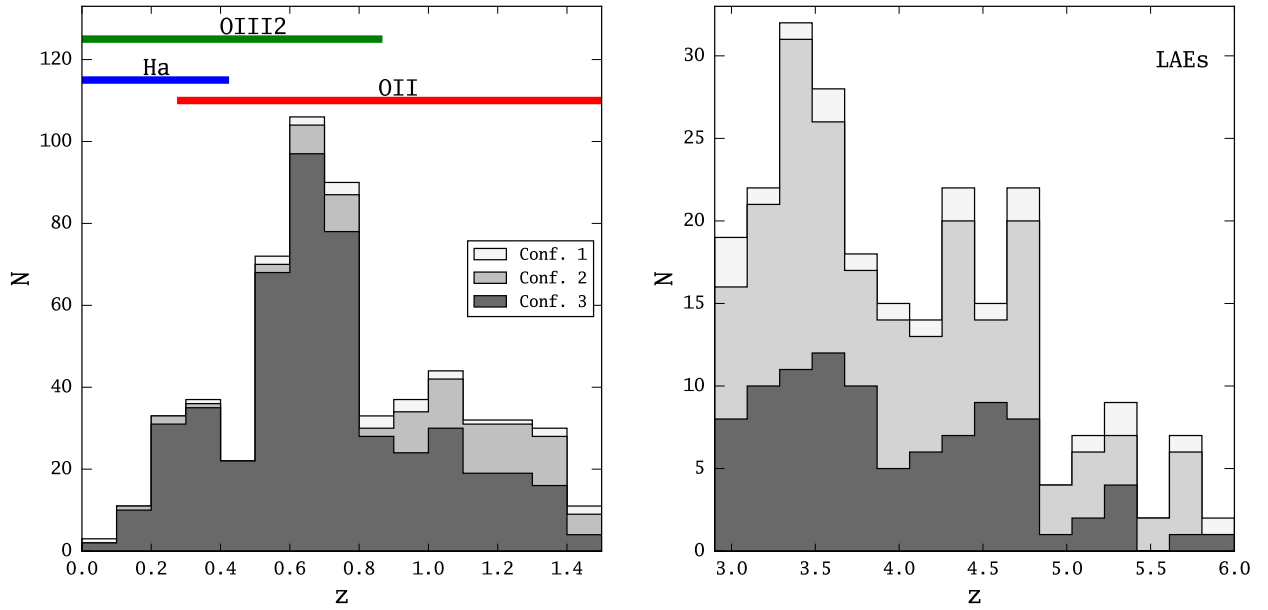
Momcheva et al. (2016) presented a redshift catalogue of  $\sim 10^5$  sources in all CANDELS fields containing grism redshifts from the 3D-HST survey (Brammer et al. 2012; Momcheva et al. 2016), photometric redshifts from 3D-HST/CANDELS photometry (Skelton et al. 2014), as well as a compilation of ground-based spectroscopic redshifts also contained in Skelton et al. (2014). For the CDF-S region under scrutiny here the ground-based redshifts in this compilation are taken from the compilation by Wuyts et al. (2008; see their Table 3 for individual spectroscopic campaigns).

To be conservative, we here include only objects that have a photometric counterpart within a  $0.5''$  radius of the MUSE-Wide position, i.e. we discard all manually assigned cross-matches from Sect. 4.1. 141 of the remaining 669 secure associations have a ground based spectroscopic redshift (139 at  $z \lesssim 2$  and 2 at  $z \gtrsim 3$ ), 184 have a grism redshift (182 at  $z \lesssim 2$  and 2 at  $z \gtrsim 3$ ), and 344 have a photometric redshift (241 at  $z \lesssim 2$  and 103 at  $z \gtrsim 3$ ).

Using the search radius of  $0.5''$  we also cross-matched our objects with the first data release of the VIMOS Ultra Deep Survey (Le Fèvre et al. 2015; Tasca et al. 2017). We find matches to 10 of our objects in their spectroscopic redshift catalogue: seven at  $z < 1.5$ , and three at  $z \gtrsim 3$  galaxies.

In Fig. 12 we compare the literature redshifts for the 679 objects that could be cross-matched with the 3D-HST or VUDS catalogues. Following the literature (e.g. Skelton et al. 2014), we define a catastrophic redshift mismatch between two catalogues as one with  $|z_{\text{MW}} - z_{\text{Lit}}|/(1 + z_{\text{MW}}) > 10\%$ , where  $z_{\text{MW}}$  is the MUSE-Wide redshift and  $z_{\text{Lit}}$  is the corresponding literature





**Fig. 11.** Histograms showing the redshift distribution of the 831 emission line selected galaxies from the first year of MUSE-Wide observations. Redshifts with confidence values 3, 2, and 1 are shown in dark grey, grey, and light grey, respectively. *Left panel:* redshift distribution for the rest-frame optical emission line selected galaxies. Horizontal bars indicate the redshift range of the MUSE wavelength coverage for the strongest emission lines of star-forming galaxies. The bin size is  $\Delta z = 0.1$ . *Right panel:* redshift distribution for the 238 high- $z$  galaxies. The bin size is  $\Delta z = 0.2$ .

**Table 6.** Catastrophic redshift mismatches (defined as  $|z_{\text{MW}} - z_{\text{Lit}}|/(1 + z_{\text{MW}}) > 0.1$ ) between MUSE-Wide and literature redshifts for objects within a  $0.5''$  search radius around a MUSE-Wide position.

| Redshift source                     | $z \lesssim 2$ | $z \gtrsim 3$ |
|-------------------------------------|----------------|---------------|
| Spectroscopy (Wuyts et al. 2008)    | 7/139 (5%)     | 0/2 (0%)      |
| Spectroscopy (Tasca et al. 2017)    | 0/7 (0%)       | 0/3 (0%)      |
| 3D-HST grism (Momcheva et al. 2016) | 6/182 (3%)     | 1/2 (50%)     |
| Photometric (Skelton et al. 2014)   | 19/241 (8%)    | 31/103 (30%)  |
| Total                               | 32/569 (5%)    | 32/110 (29%)  |

**Notes.** In the form  $x/y$ , where  $x$  is the number of redshift mismatches and  $y$  the total number of secure counterparts with literature redshifts. In brackets this fraction is expressed as percent.

redshift. Based on this definition 64 of the 3D-HST redshifts (i.e., 9.6% of secure cross-matches within a search radius  $0.5''$ ) are in disagreement. On the other hand the 10 VUDS redshifts are in excellent agreement with the MUSE-Wide redshift, all characterised by  $|z_{\text{MW}} - z_{\text{VUDS}}|/(1 + z_{\text{MW}}) < 10^{-3}$ .

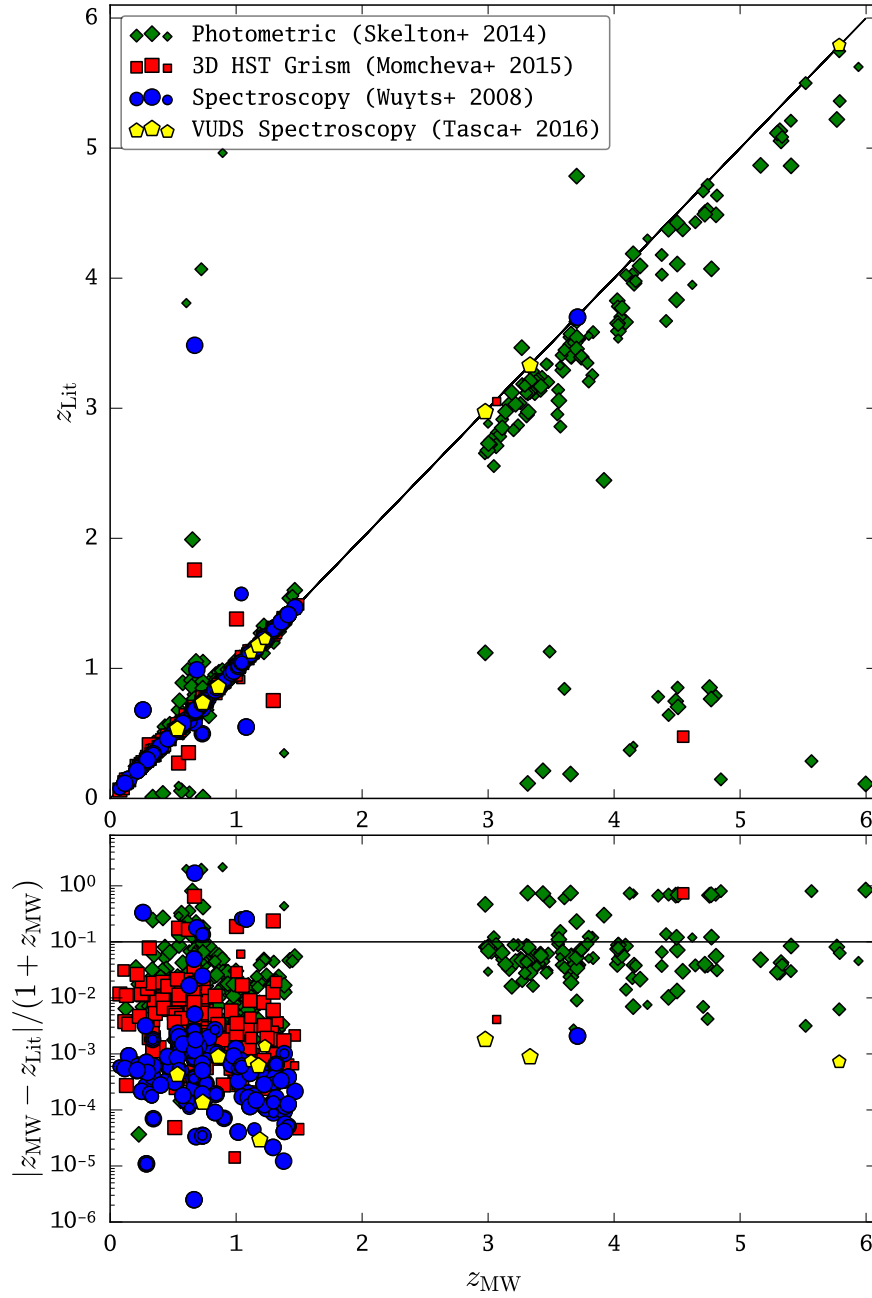
In Table 6 we summarise the statistics on the redshift mismatches and indicate whether the mismatch is a low- or high- $z$  source. We find the highest rate of catastrophic mismatches (30%) between MUSE-Wide and literature redshifts amongst photometrically determined redshifts at  $z \gtrsim 3$ . For the 3D-HST grism redshifts the agreement is mostly excellent, except for a few cases at low-redshift and one of the only two grism redshift at high- $z$ . Notably, all low- $z$  mismatches have the highest confidence value in our catalogue, meaning that the MUSE-Wide data leaves no doubt on the obtained redshift. The high- $z$  grism redshift mismatch could be reconciled if we would classify this single emission line detection as [O II] instead of Ly $\alpha$ . Based on our visual inspection of the line profile, however, we flagged this source with confidence = 2, meaning that we have at most minor doubts on our line classification. Finally, only 5 from 140 spectroscopic redshifts from Wuyts et al. (2008) are in disagreement with our catalogue. Four of those were marked with the highest confidence flag in our classification and only in one case we did

have minor doubts. Notably, one single line detection classified as Ly $\alpha$  in the Wuyts et al. (2008) compilation clearly shows a characteristic O II profile that likely was not resolved in the previous spectrum.

We notice that for high- $z$  LAEs in our sample a systematic offset between photometric and spectroscopic redshifts with  $|z_{\text{MW}} - z_{\text{3DHST}}|/(1 + z_{\text{MW}}) \sim 10^{-1.5}$  exists. 23% of the LAEs have  $\Delta z = z_{\text{MW}} - z_{\text{Lit}} > +0.1$  and the median  $\Delta z$  is  $+0.26$ , with the upper and lower quartiles being  $+0.17$  and  $+0.44$ , respectively. This systematic offset is much larger than the known systematic offset of the Ly $\alpha$  line to the systemic redshift discussed in Sect. 3.3.1. Recently, Oyarzún et al. (2016) reported very similar offsets between photometric and spectroscopic redshifts of LAEs. They found that the magnitude of the offset correlates with the Ly $\alpha$  equivalent width. Hence, a possible source of the discrepancy is that Ly $\alpha$  line flux is missing from the spectral energy distribution templates which are used in the photometric redshift fitting.

### 5.3. Emission line fluxes and continuum magnitudes

To characterise our sample we compare in Fig. 13 the measured line fluxes in  $3 \times R_{\text{Kron}}$  apertures with the continuum



**Fig. 12.** Redshift comparison between our MUSE-Wide emission line selected galaxy sample in the CDFS and literature redshifts (see text for cross-matching details). Literature redshifts are from the catalogues from Momcheva et al. (2016) and Tasca et al. (2017), where the former includes a compilation of ground based spectroscopic redshifts from Wuyts et al. (2008) and photometric-redshifts from Skelton et al. (2014). See text for further details on the cross-matching procedure. *Top panel:* literature redshifts ( $z_{\text{Lit}}$ ) versus MUSE-Wide redshifts ( $z_{\text{MW}}$ ). Symbol size encodes the confidence on the MUSE-Wide emission line classification (Sect. 3.2), with the largest symbols representing “Confidence 3” sources, the medium sized symbols representing “Confidence 2” sources, and the smallest symbols representing “Confidence 1”. The different symbols encode the source of the catalogue redshifts according to the legend. *Bottom panel:* absolute relative difference between literature and MUSE-Wide redshifts. The horizontal line denotes the boundary  $|z_{\text{MW}} - z_{\text{Lit}}|/(1 + z_{\text{MW}}) = 0.1$  where we distinguish between redshift equality and redshift mismatches.

magnitudes in the  $JH_{\text{IR}}$  magnitude from the 3D-HST catalogues. As described in Sect. 3.6 of Momcheva et al. (2016, see also their Table 5) the  $JH_{\text{IR}}$  magnitudes have been measured with SExtractor’s MAG\_AUTO in a coadded image created from the HST WFC3  $J_{125}$ ,  $JH_{140}$ , and  $H_{160}$  images.

Redshifts from the 3D-HST grism survey are provided for objects with  $JH \leq 24$ . Similarly, previous ground based spectroscopic deep-field follow-up campaigns target objects to a limiting photometric depth. For example, the VIMOS VLT Deep

sample contains objects with  $i_{\text{AB}}$ -magnitudes brighter than 24.75 (Le Fèvre et al. 2013). However, as can be seen from Fig. 13, a significant fraction of galaxies in MUSE-Wide is characterised by very faint continuum emission. This is especially so for the Ly $\alpha$  emitting high- $z$  galaxy population in our catalogue and thus explains the previous lack of spectroscopic redshifts for those galaxies.

The  $JH_{\text{IR}}$  distribution of the LAE population peaks at  $JH_{\text{IR}} \sim 27$  for objects that have a photometric counterpart in

**Table 7.** Crossmatch between MUSE emission line galaxies and X-ray sources flagged as AGN from the *Chandra* 7Ms source catalogue (Luo et al. 2017).

| ID<br>MUSE Wide | $z$<br>MUSE | ID<br><i>Chandra</i> 7Ms | Separation<br>( $''$ ) | X-ray flux<br>(erg/s/cm $^2$ ) | $z$<br><i>Chandra</i> |
|-----------------|-------------|--------------------------|------------------------|--------------------------------|-----------------------|
| 102007068       | 0.338       | 304                      | 0.20                   | 2.982e-16                      | 0.340                 |
| 102008071       | 0.338       | 312                      | 0.86                   | 6.491e-17                      | 0.336                 |
| 102031144       | 0.665       | 290                      | 0.20                   | 3.498e-16                      | 0.664                 |
| 102037154       | 1.412       | 287                      | 0.51                   | 5.925e-17                      | 1.413                 |
| 103022086       | 0.670       | 322                      | 0.24                   | 1.510e-16                      | 0.671                 |
| 104014050       | 3.662       | 337                      | 0.17                   | 1.269e-15                      | 3.660                 |
| 106036089       | 0.904       | 344                      | 0.82                   | 6.638e-17                      | 0.956 <sup>a</sup>    |
| 106048103       | 0.665       | 340                      | 0.15                   | 2.406e-15                      | 0.666                 |
| 108025145       | 0.736       | 407                      | 0.49                   | 7.387e-17                      | 0.736                 |
| 109030090       | 1.044       | 447                      | 0.17                   | 5.737e-16                      | 1.043                 |
| 111004005       | 0.604       | 367                      | 0.28                   | 4.594e-15                      | 0.604                 |
| 113001007       | 0.232       | 508                      | 0.94                   | 2.411e-17                      | 0.220 <sup>a</sup>    |
| 113010038       | 0.577       | 460                      | 0.35                   | 2.778e-16                      | 0.577                 |
| 114024110       | 1.035       | 443                      | 0.26                   | 8.709e-16                      | 1.036                 |
| 114028115       | 1.098       | 509                      | 0.30                   | 1.525e-16                      | 1.097                 |
| 115003085       | 3.710       | 551                      | 0.11                   | 2.158e-15                      | 3.700                 |
| 116003060       | 1.364       | 634                      | 0.14                   | 3.899e-17                      | 1.363 <sup>b</sup>    |
| 117034085       | 0.228       | 693                      | 1.13                   | 6.435e-17                      | 2.302                 |
| 119034073       | 1.015       | 814                      | 0.15                   | 4.348e-15                      | 1.016                 |
| 120023032       | 1.118       | 861                      | 0.25                   | 1.744e-16                      | 1.120                 |
| 123005089       | 0.544       | 640                      | 0.51                   | 4.219e-17                      | 0.552 <sup>b</sup>    |
| 123051191       | 4.510       | 625                      | 1.33                   | 2.102e-17                      | 2.616 <sup>b</sup>    |

**Notes.** All redshifts quoted are spectroscopic from extensive identification campaigns except for objects marked with <sup>a</sup> and <sup>b</sup> which are photometric redshifts from Skelton et al. (2014) or Hsu et al. (2014), respectively.

the CANDELS 3D-HST catalogue. However, 87 MUSE-Wide LAEs have no cross-matches to the deep 3D-HST catalogue within  $0.5''$ . We place those objects arbitrarily at  $JH_{\text{IR}} = 29$  in Fig. 13. We caution that most of these objects in fact show prominent photometric counterparts but were missed in the 3D-HST NIR-based detection. Therefore we expect most of the upper limits in the bottom right panel of Fig. 13 moving towards brighter magnitudes. A detailed evaluation of the continuum-faint population of MUSE-Wide LAEs is beyond the scope of this catalogue release and will be subject of a future study.

#### 5.4. Active galactic nuclei (AGNs)

Traditionally blind spectroscopic surveys hunting for emission lines galaxies unveiled numerous active galaxies (e.g. Zamorano et al. 1994), so naturally we want to check that in our sample as well. At low redshift, where  $H\alpha$  and nearby lines fall within the wavelength range covered by MUSE, we can employ the BPT diagram (Baldwin et al. 1981) to distinguish AGN from star forming galaxies. However, for larger redshifts, we need to rely on ancillary spectroscopic data from the near-IR to use the BPT diagram. This is beyond the scope of this paper.

We therefore opt to crossmatch our emission line sources with the main catalogue of the 7Ms *Chandra* Deep Field South (Luo et al. 2017). A match was deemed successful if our emission line position was within 3 times the X-ray positional accuracy (SIGMAX or Col. 6 in Table 4 of the Luo et al. 2017 catalogue). In addition to an X-ray match, we required that the object type column (Col. 70) of the X-ray catalogue be “AGN” as to not include any X-rays from star formation regions. We find 22 candidate AGNs in our emission line catalogue which we present in Table 7.

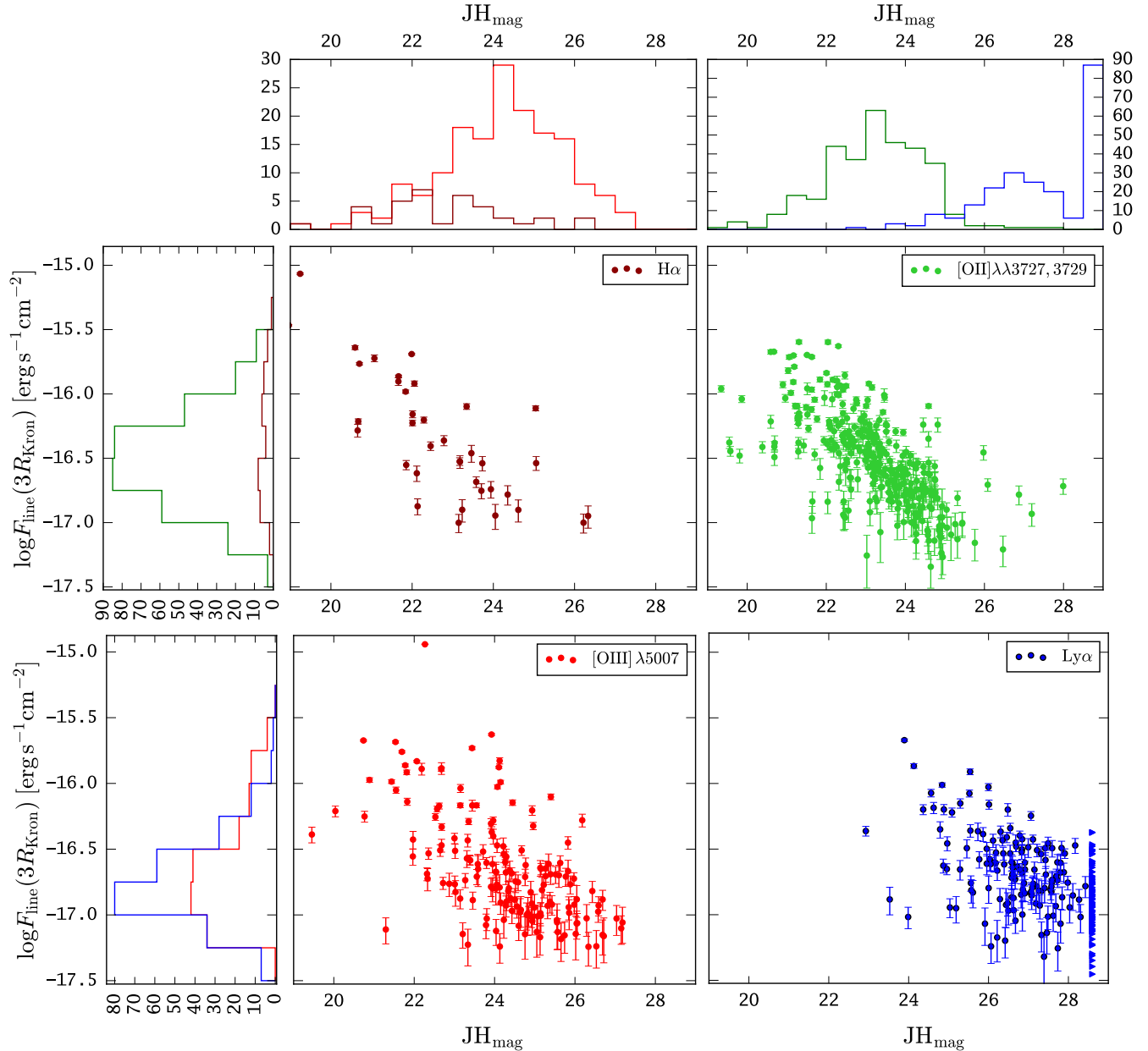
Most of the matched sources already have been spectroscopically identified in previous campaigns to identify optical counterparts to X-ray sources (e.g. Szokoly et al. 2004, see Luo et al. 2017; for an extensive list of 26 identification references). Nevertheless, we can assign spectroscopic redshifts to 5 objects which previously only had photometric redshifts determined. All, but two of the redshifts are in excellent with the MUSE determined redshift. These are the ones with the largest distance between the X-ray and MUSE position, so most likely a different X-ray counterpart was assigned. Among the 17 AGN are two well known type 2 high redshift quasars (MUSE-WIDE IDs: 104014050, 115003085), both of which have the highest  $\text{Ly}\alpha$  flux in our survey (Norman et al. 2002; Mainieri et al. 2005).

## 6. Conclusions and outlook

We present the first results from the ongoing MUSE-Wide survey. Using LSDCat, a novel 3D source detection algorithm based on matched filtering, we constructed a catalogue of 831 galaxies with altogether 1656 detected emission lines, all located in a footprint of  $22.2 \text{ arcmin}^2$  within the CANDELS-Deep/CDFS region. More than half of these galaxies did not have spectroscopic or HST grism redshifts until now.

Because of the emission line selection, the properties of these galaxies are quite different from those of photometrically pre-selected samples: The redshift distribution features two disjoint domains, a low to intermediate redshift category with  $z < 1.5$  where rest-frame optical nebular lines are detected, and a high-redshift ( $z > 2.9$ ) category of strong  $\text{Ly}\alpha$  emitters; in between these domains, MUSE-Wide suffers from the well-known “redshift desert” effect caused by the dearth of strong emission lines in the UV. Nevertheless, compared to the already extensive





**Fig. 13.** Line fluxes of the strongest emission lines vs. continuum  $JH_{\text{IR}}$  magnitudes from 3D-HST (Momcheva et al. 2016) for 669 of 831 MUSE-Wide galaxies that have a 3D-HST catalogue counterpart within  $0.5''$ .  $H\alpha$ ,  $[O II]$ ,  $[O III]$ , and  $Ly\alpha$  lines are shown in brown, green, red, and blue, respectively. With the same colours we also show histograms of the scatter-clouds.  $Ly\alpha$  objects not having a counterpart within  $r = 0.5''$  in the 3D-HST catalogue are shown as blue triangles and have been put into the  $JH_{\text{IR}} = 29$  mag bin. We point out, however, that most of the objects show photometric counterparts that are not present in the catalogue.

spectroscopic coverage in these fields, MUSE-Wide adds significantly to both the low/intermediate- and the high-redshift domain by achieving a spatial target sampling rate of essentially 100%, irrespective of the distribution and shapes of sources in the sky.

While our adopted selection criteria reveal objects that are similar to those found in narrowband imaging, there are also some notable differences. Rather than by covering a large solid angle in the sky, at predefined small redshift windows, such as obtained by instruments like Hyper Suprime-Cam (Miyazaki et al. 2012), MUSE-Wide gains its survey volume through its high redshift path length. Since all our survey fields are fully included within the footprints of very deep multi-band HST imaging and other multi-wavelength deep field efforts, the

amount of available information per object is maximised. An obvious further advantage over narrowband-selected samples is the fact that all MUSE-Wide sources are already spectroscopically confirmed.

In order to promote the legacy aspect of our survey we provide not only the spectra, catalogue data, and flux measurements, but also datacube cutouts centred on each of the 831 objects. Several of the low and intermediate-redshift objects in our sample are clearly spatially extended. With HST prior morphological information and using our datacube cutouts it will be possible, e.g., to perform spatially resolved kinematic analyses (such as in Contini et al. 2016) or investigate gas-phase metallicity gradients, thus conduct studies that normally come at considerable extra observational costs.

Of particular scientific interest is the high-redshift part of this sample, because of the unique combination of MUSE IFU with deep HST data. Remarkably, only 3 out of the 238 high- $z$  emission line objects in the surveyed 22.2 arcmin<sup>2</sup> were previously known as spectroscopically confirmed Ly $\alpha$  emitters. It is now established that Ly $\alpha$  in high- $z$  galaxies essentially always has a significant circumgalactic component that extends much beyond the stellar continuum emission region (Wisotzki et al. 2016). MUSE now facilitates the inclusion of this extended nature of the Ly $\alpha$  line into demographic studies, especially important for quantifying the completeness of LAE selection and the impact on the Ly $\alpha$  luminosity function (Herenz et al., in prep.). We are also investigating several other aspects of LAE demographics which will be presented in a suite of forthcoming papers.

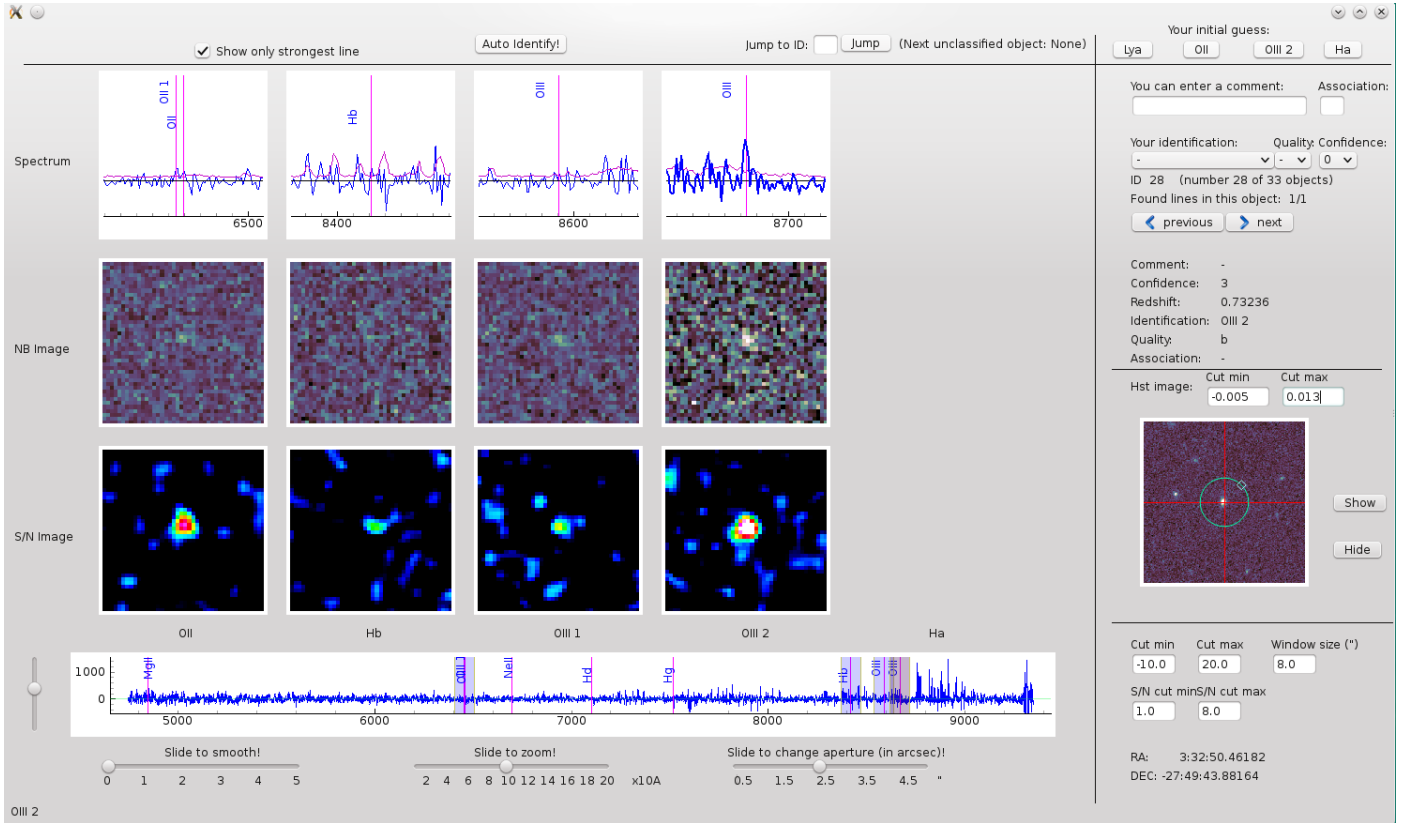
The present sample is by construction restricted to galaxies with significantly detected emission lines. As a next step we are working on an extended redshift survey that includes also continuum-selected objects with only weak or without emission lines. This extension will further enhance the legacy value of MUSE-Wide, especially for relatively low redshifts. Furthermore, results from the deep and ultra-deep MUSE surveys (Bacon et al. 2015, and in prep.) will provide constraints on the faintest parts of the galaxy population at all redshifts accessible to MUSE. Combining such multi-tier approaches has always been a winning strategy for astronomical surveys, and MUSE makes no exception here.

At the time of writing, observations for MUSE-Wide are proceeding well and the goal of observing 100 MUSE pointings was reached February 2017. Besides serving our own scientific interests, this survey will remain a very valuable community resource, and we have committed ourselves to release the data to the public as soon as possible, after careful quality control and postprocessing where needed. The first comprehensive data release (DR1, Urrutia et al., in prep.) is currently anticipated for 2017, covering 44 MUSE fields in the CDFS region. Further data releases will follow.

**Acknowledgements.** We thank the support staff at ESOs VLT for help with the visitor mode observations. This research made extensive use of the *astropy* package (Astropy Collaboration et al. 2013). All plots in this paper (except Fig. A.1) were created using *matplotlib* (Hunter 2007). E.C.H., J.C., J.K., R.S., T.U., and L.W. acknowledge funding by the Competitive Fund of the Leibniz Association through grants SAW-2013-AIP-4 and SAW-2015-AIP-2. R.B. acknowledges support from the ERC advanced grant 339659-MUSICOS.

## References

- Adams, J. J., Blanc, G. A., Hill, G. J., et al. 2011, *ApJS*, **192**, 5
- Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, *A&A*, **558**, A33
- Bacon, R., Bauer, S., Brau-Nogu  , S., et al. 2009, in *Science with the VLT in the ELT Era*, ed. A. Moorwood, 331
- Bacon, R., Vernet, J., Borisova, E., et al. 2014, *The Messenger*, **157**, 13
- Bacon, R., Brinchmann, J., Richard, J., et al. 2015, *A&A*, **575**, A75
- Bacon, R., Piqueras, L., Conseil, S., Richard, J., & Shepherd, M. 2016, *Astrophysics Source Code Library* [record asc1:1611.003]
- Bacon, R., Conseil, S., Mary, D., et al. 2017, *A&A*, in press, DOI: 10.1051/0004-6361/201730833
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, *PASP*, **93**, 5
- Balestra, I., Mainieri, V., Popesso, P., et al. 2010, *A&A*, **512**, A12
- Bertin, E. 2001, in *Mining the Sky*, eds. A. J. Banday, S. Zaroubi, & M. Bartelmann, 353
- Bertin, E., & Arnouts, S. 1996, *A&AS*, **117**, 393
- Borisova, E., Cantalupo, S., Lilly, S. J., et al. 2016, *ApJ*, **831**, 39
- Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, *ApJS*, **200**, 13
- Caillier, P., Accardo, M., Adjali, L., et al. 2014, in *Modeling, Systems Engineering, and Project Management for Astronomy VI*, *Proc. SPIE*, **9150**, 91500D
- Chonis, T. S., Blanc, G. A., Hill, G. J., et al. 2013, *ApJ*, **775**, 99
- Conseil, S., Bacon, R., Piqueras, L., & Shepherd, M. 2016, *ArXiv e-prints* [arXiv:1612.05308]
- Contini, T., Epinat, B., Bouch  , N., et al. 2016, *A&A*, **591**, A49
- Das, P. K. 1991, *Optical Signal Processing* (Springer Science + Business Media)
- Dawson, S., Rhoads, J. E., Malhotra, S., et al. 2007, *ApJ*, **671**, 1227
- Erb, D. K., Steidel, C. C., Trainor, R. F., et al. 2014, *ApJ*, **795**, 33
- Erben, T., Schirmer, M., Dietrich, J. P., et al. 2005, *Astron. Nachr.*, **326**, 432
- Filippenko, A. V. 1982, *PASP*, **94**, 715
- Garilli, B., Fumana, M., Franzetti, P., et al. 2010, *PASP*, **122**, 827
- Gialalisco, M., Ferguson, H. C., Koekemoer, A. M., et al. 2004, *ApJ*, **600**, L93
- Graham, A. W., & Driver, S. P. 2005, *PASA*, **22**, 118
- Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, *ApJS*, **197**, 35
- Guo, Y., Ferguson, H. C., Gialalisco, M., et al. 2013, *ApJS*, **207**, 24
- Hashimoto, T., Verhamme, A., Ouchi, M., et al. 2015, *ApJ*, **812**, 157
- Herenz, E. C., & Wisotzki, L. 2016, *Astrophysics Source Code Library* [record asc1:1612.002]
- Herenz, E. C., & Wisotzki, L. 2017, *A&A*, **602**, A111
- Hickson, P. 2014, *A&ARv*, **22**, 76
- Hildebrandt, H., Erben, T., Dietrich, J. P., et al. 2006, *A&A*, **452**, 1121
- Hsu, L.-T., Salvato, M., Nandra, K., et al. 2014, *ApJ*, **796**, 60
- Hunter, J. D. 2007, *Comput. Sci. Eng.*, **9**, 90
- Ivezi  , Z., Connelly, A. J., VanderPlas, J. T., & Gray, A. 2014, *Statistics, Data Mining, and Machine Learning in Astronomy* (Princeton Univ. Press)
- Kamann, S. 2013, Ph.D. Thesis, Universit  t Potsdam
- Kamann, S., Wisotzki, L., & Roth, M. M. 2013, *A&A*, **549**, A71
- Kennicutt, Jr., R. C. 1992, *ApJS*, **79**, 255
- Kerutt, J. 2017, *Astrophysics Source Code Library* [record asc1:1703.011]
- Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, *ApJS*, **197**, 36
- Kron, R. G. 1980, *ApJS*, **43**, 305
- Le F  vre, O., Cassata, P., Cucciati, O., et al. 2013, *A&A*, **559**, A14
- Le F  vre, O., Tasca, L. A. M., Cassata, P., et al. 2015, *A&A*, **576**, A79
- Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, *ApJS*, **228**, 2
- Mainieri, V., Rigopoulou, D., Lehmann, K., et al. 2005, *MNRAS*, **356**, 1571
- Mallery, R. P., Mobasher, B., Capak, P., et al. 2012, *ApJ*, **760**, 128
- Masters, D., & Capak, P. 2011, *PASP*, **123**, 638
- McLinden, E. M., Finkelstein, S. L., Rhoads, J. E., et al. 2011, *ApJ*, **730**, 136
- Miyazaki, S., Komiyama, Y., Nakaya, H., et al. 2012, in *Ground-based and Airborne Instrumentation for Astronomy IV*, *Proc. SPIE*, **8446**, 84460Z
- Momcheva, I. G., Brammer, G. B., van Dokkum, P. G., et al. 2016, *ApJS*, **225**, 27
- Noll, S., Mehlert, D., Appenzeller, I., et al. 2004, *A&A*, **418**, 885
- Norman, C., Hasinger, G., Giacconi, R., et al. 2002, *ApJ*, **571**, 218
- Oyarz  n, G. A., Blanc, G. A., Gonz  lez, V., et al. 2016, *ApJ*, **821**, L14
- Pence, W. D., Chiappetti, L., Page, C. G., Shaw, R. A., & Stobie, E. 2010, *A&A*, **524**, A42
- Popesso, P., Dickinson, M., Nonino, M., et al. 2009, *A&A*, **494**, 443
- Rakic, O., Schaye, J., Steidel, C. C., & Rudie, G. C. 2011, *MNRAS*, **414**, 3265
- Ryabchikova, T., Piskunov, N., Kurucz, R. L., et al. 2015, *Phys. Scr.*, **90**, 054005
- Schwartz, M., & Shaw, L. 1975, *Linear processing: discrete spectral analysis, detection, and estimation* (Tokyo: McGraw-Hill Kogakusha, Ltd.)
- Shibuya, T., Ouchi, M., Nakajima, K., et al. 2014, *ApJ*, **788**, 74
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, **214**, 24
- Song, M., Finkelstein, S. L., Gebhardt, K., et al. 2014, *ApJ*, **791**, 3
- Soto, K. T., Lilly, S. J., Bacon, R., Richard, J., & Conseil, S. 2016, *MNRAS*, **458**, 3210
- Streicher, O., Weilbacher, P. M., Bacon, R., & Jarno, A. 2011, in *Astronomical Data Analysis Software and Systems XX*, eds. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, *ASP Conf. Ser.*, **442**, 257
- Szokoly, G. P., Bergeron, J., Hasinger, G., et al. 2004, *ApJS*, **155**, 271
- Tasca, L. A. M., Le F  vre, O., Ribeiro, B., et al. 2017, *A&A*, **600**, A110
- Trainor, R. F., Steidel, C. C., Strom, A. L., & Rudie, G. C. 2015, *ApJ*, **809**, 89
- van Breukelen, C., Jarvis, M. J., & Venemans, B. P. 2005, *MNRAS*, **359**, 895
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2005, *A&A*, **434**, 53
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2006, *A&A*, **454**, 423
- Vanzella, E., Cristiani, S., Dickinson, M., et al. 2008, *A&A*, **478**, 83
- Vio, R., & Andreani, P. 2016, *A&A*, **589**, A20
- Weilbacher, P. M., Streicher, O., Urrutia, T., et al. 2014, in *Astronomical Data Analysis Software and Systems XXIII*, eds. N. Manset, & P. Forshay, *ASP Conf. Ser.*, **485**, 451
- Wisotzki, L., Bacon, R., Blaizot, J., et al. 2016, *A&A*, **587**, A98
- Wuyts, S., Labb  , I., F  rster Schreiber, N. M., et al. 2008, *ApJ*, **682**, 985
- Yamada, T., Matsuda, Y., Kousai, K., et al. 2012, *ApJ*, **751**, 29
- Zackay, B., & Ofek, E. O. 2017, *ApJ*, **836**, 187
- Zamorano, J., Rego, M., Gallego, J. G., et al. 1994, *ApJS*, **95**, 387



**Fig. A.1.** Screenshot of QtClassify, the software for classification of emission line sources detected in wide-field IFS datacubes (Kerutt 2017). All graphical panels are interactive. The *top row panels* display sections of a spectrum extracted from the datacube at the position of the detected emission line under scrutiny. The aperture radius for extraction, the limits of the  $y$ -axis, as well as the degree of smoothing of the spectrum can be changed interactively using the sliders at the bottom. The positions of displayed sections are adjusted to show regions around possible veto lines under the users guess for the nature of the detected emission line. *Below the spectra* pseudo-narrow band images at the veto line positions are displayed. Here the cut-levels as well as the width of the narrow-band window can be altered interactively at the bottom right section of the GUI. The *bottom row of panels* displays the corresponding layers from LSDCat’s S/N-cube. Both the spectral position of the *centre-row panels* and the *bottom-row panels* can be changed simultaneously with a slider or by moving the pointer in the top-row panels. On the bottom of the window an overview over the whole extracted spectrum is displayed, and at the right side of the window a cut-out from 2D imaging data can be shown. The example shown here demonstrates the classification of a rather weak [O III] emission line. For this object other lines (O II, H $\beta$ , and [O III]  $\lambda$ 4959) are below the detection threshold, but can still be visually identified in the different panels of the window.

## Appendix A: QtClassify

QtClassify is a graphical user interface (GUI) that helps to classify emission line objects detected in wide-field IFS data. Its development was loosely inspired by the SpecPRO GUI used in classical imaging and spectroscopic surveys (Masters & Capak 2011). QtClassify is publicly available and can be downloaded via the Astrophysics Source Code Library (Kerutt 2017)<sup>7</sup>.

In Fig. A.1 we display a screenshot of QtClassify. As input it requires a – preferentially continuum-subtracted – datacube, a catalogue tabulating the positions of the emission lines, and a S/N datacube. Optionally, broadband imaging data can also be loaded into the software. For creating the catalogue and S/N cube QtClassify is currently geared towards the outputs of LSDCat (Herenz & Wisotzki 2017). However, in principle it can be used with any IFS datacube and any catalogue that contains one entry for each emission line with all entries grouped into objects by position.

The main idea of QtClassify is to use the full 3D information in the datacubes to classify emission lines by simultaneously displaying regions in the cube where other lines would be expected given an initial guess for the identification of the

strongest detected line. Strongest here means the line with the highest signal-to-noise ratio (determined e.g. by the matched-filtering procedure in LSDCat, but it is also possible to use line flux instead). The default guesses are Ly $\alpha$   $\lambda$ 1216, [O II]  $\lambda$ 3726, 3729, [O III]  $\lambda$ 5007 and H $\alpha$   $\lambda$ 6563, but they be changed by customising the input list of options.

To help classify the lines, there are several rows with information (see Fig. A.1). When clicking on one of the possible guesses, the columns change to the positions where you would expect other lines. The first row (“Spectrum”) displays the part of the automatically extracted spectrum with an indication where the line is expected. The second row (“NB Image”) shows a monochromatic layer of the actual datacube and the third row (“S/N Image”) shows a monochromatic layer of the S/N cube. When interacting with the programme, the user can move the mouse through the spectral parts in the first row, which will automatically also move through the corresponding wavelength layers in the second and third rows generating a “movie-like” sequence of monochromatic slices in the datacube. The bottom panel shows the full spectrum where the zoomed-in parts of the first row are indicated by shaded areas. On the right of the window, a broadband image (in the example displayed in Fig. A.1 from the HST CANDELS data) can be displayed if needed. The

<sup>7</sup> <http://ascl.net/1703.011>



cuts for the broadband image as well as for the monochromatic images of the datacube and the S/N cube can be interactively adjusted. Several sliders at the bottom of the window allow interactive smoothing and zooming of the spectrum, as well as changing the size of the extraction aperture for the displayed spectrum.

In order to make it less tedious to go through all the detected emission lines, there is a button at the top that automatically identifies all lines where a unique redshift solution can be automatically anchored. For this, QtClassify looks at all objects with multiple emission lines and internally tries to determine a redshift using a list of typical emission lines (similar to the peak correlation algorithm by [Garilli et al. 2010](#)). This list can be adjusted and extended. This automatic procedure fails, of course, when only a single, maybe even spurious, emission line is detected, or when a superposition of objects at different redshifts occurs at the same position on the sky. Those objects have to be manually classified by the QtClassify user.

The interactive classification is done via a drop-down menu that presents a list of emission lines to the user, including various classes of spurious detections (e.g. cosmics, noise-peaks, telluric-, or continuum-subtraction residuals). In addition the classifier can enter confidence value and quality flag (cf. Sect. 3.2). Cases of overlaps, where one entry in the catalogue has emission lines from multiple objects, are not covered by QtClassify and have to be sorted out separately in a later stage of the catalogue construction. Here the comment function enables one to take notes for this process. Multiple entries at different positions on the sky that belong to the same object can be marked as associations. It is possible to stop classifying at any point and resume later with the output catalogue of QtClassify as input, since it automatically saves the classifications in columns that are appended to the input catalogue.