

# The VIMOS Public Extragalactic Redshift Survey (VIPERS)

## Full spectroscopic data and auxiliary information release (PDR-2)<sup>★</sup>

M. Scodeggio<sup>1</sup>, L. Guzzo<sup>2,3</sup>, B. Garilli<sup>1</sup>, B. R. Granett<sup>2,3</sup>, M. Bolzonella<sup>4</sup>, S. de la Torre<sup>5</sup>, U. Abbas<sup>6</sup>, C. Adami<sup>5</sup>, S. Arnouts<sup>5</sup>, D. Bottini<sup>1</sup>, A. Cappi<sup>4,7</sup>, J. Coupon<sup>8</sup>, O. Cucciati<sup>4,9</sup>, I. Davidzon<sup>4,5</sup>, P. Franzetti<sup>1</sup>, A. Fritz<sup>1</sup>, A. Iovino<sup>2</sup>, J. Krywult<sup>10</sup>, V. Le Brun<sup>5</sup>, O. Le Fèvre<sup>5</sup>, D. Maccagni<sup>1</sup>, K. Małek<sup>11</sup>, A. Marchetti<sup>1</sup>, F. Marulli<sup>4,9,12</sup>, M. Polletta<sup>1,13,14</sup>, A. Pollo<sup>11,15</sup>, L. A. M. Tasca<sup>5</sup>, R. Tojeiro<sup>16</sup>, D. Vergani<sup>17</sup>, A. Zanicelli<sup>18</sup>, J. Bel<sup>19</sup>, E. Branchini<sup>20,21,22</sup>, G. De Lucia<sup>23</sup>, O. Ilbert<sup>5</sup>, H. J. McCracken<sup>24</sup>, T. Moutard<sup>5,25</sup>, J. A. Peacock<sup>26</sup>, G. Zamorani<sup>4</sup>, A. Burden<sup>27</sup>, M. Fumana<sup>1</sup>, E. Jullo<sup>5</sup>, C. Marinoni<sup>19,28</sup>, Y. Mellier<sup>24</sup>, L. Moscardini<sup>9,12,24</sup>, and W. J. Percival<sup>27</sup>

(Affiliations can be found after the references)

Received 22 November 2016 / Accepted 28 April 2017

### ABSTRACT

We present the full public data release (PDR-2) of the VIMOS Public Extragalactic Redshift Survey (VIPERS), performed at the ESO VLT. We release redshifts, spectra, CFHTLS magnitudes and ancillary information (as masks and weights) for a complete sample of 86 775 galaxies (plus 4732 other objects, including stars and serendipitous galaxies); we also include their full photometrically-selected parent catalogue. The sample is magnitude limited to  $i_{AB} \leq 22.5$ , with an additional colour-colour pre-selection devised as to exclude galaxies at  $z < 0.5$ . This practically doubles the effective sampling of the VIMOS spectrograph over the range  $0.5 < z < 1.2$  (reaching 47% on average), yielding a final median local galaxy density close to  $5 \times 10^{-3} h^3 \text{Mpc}^{-3}$ . The total area spanned by the final data set is  $\approx 23.5 \text{ deg}^2$ , corresponding to 288 VIMOS fields with marginal overlaps, split over two regions within the CFHTLS-Wide W1 and W4 equatorial fields (at RA  $\approx 2$  and  $\approx 22$  h, respectively). Spectra were observed at a resolution  $R = 220$ , covering a wavelength range 5500–9500 Å. Data reduction and redshift measurements were performed through a fully automated pipeline; all redshift determinations were then visually validated and assigned a quality flag. Measurements with a quality flag  $\geq 2$  are shown to have a confidence level of 96% or larger and make up 88% of all measured galaxy redshifts (76 552 out of 86 775), constituting the VIPERS prime catalogue for statistical investigations. For this sample the rms redshift error, estimated using repeated measurements of about 3000 galaxies, is found to be  $\sigma_z = 0.00054(1+z)$ . All data are available at <http://vipers.inaf.it> and on the ESO Archive.

**Key words.** cosmology: observations – large-scale structure of Universe – galaxies: distances and redshifts – galaxies: statistics – surveys

## 1. Introduction

Large photometric and spectroscopic galaxy surveys have played a key role in building our current understanding of the Universe. At  $z \leq 0.2$ , the 2dFGRS (Colless et al. 2003) and SDSS (York et al. 2000; Abazajian et al. 2009) redshift surveys have assembled samples of over a million objects, precisely characterising large-scale structure and galaxy properties in the nearby Universe on scales ranging from 0.1 to  $100 h^{-1} \text{Mpc}$ . The SDSS has then extended its reach, first by using luminous red galaxies (LRG) to push to  $z \approx 0.35$  (SDSS-II; Eisenstein et al. 2011; Ahn et al. 2012), and more recently (to  $z \approx 0.5$ ) by using a more heterogeneous set of colour-selected objects to trace large

volumes of the Universe in a highly effective way, notwithstanding a rather dilute sampling of the total galaxy population (SDSS-III BOSS; Alam et al. 2015). [A more complete account of the development of galaxy redshift surveys over the past two decades was given in Guzzo et al. 2014].

The VIMOS Public Extragalactic Redshift Survey (VIPERS) adopted the original broad approach of SDSS-I, transposed to the redshift range  $0.5 < z < 1.2$ , essentially extending to a much larger volume of the Universe the exploration initiated with smaller-area VIMOS precursors, i.e. VVDS (Le Fèvre et al. 2013a; Garilli et al. 2008) and zCOSMOS (Lilly et al. 2009). In practice, VIPERS was conceived to obtain a large-volume, dense sample of the general galaxy population, characterised by a simple, broad selection function, complete to a given flux limit within a well-defined redshift range and complemented by extended photometric information.

This paper accompanies the Public Data Release 2 (PDR-2) of the complete VIPERS data set and is organised as follows: in Sect. 2.1 we summarise the survey design and scope, which we discussed in detail in the papers by Guzzo et al. (2014) and Garilli et al. (2014), which accompanied the first data release (PDR-1); in Sect. 3 we present the final survey mask and completeness estimates, while redshift measurements are summarised in Sect. 4, and the overall properties of the PDR-2 sample are presented in Sect. 5.

<sup>★</sup> Based on observations collected at the European Southern Observatory, Cerro Paranal, Chile, using the Very Large Telescope under programmes 182.A-0886 and partly 070.A-9007. Also based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT), which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS. The VIPERS web site is <http://www.vipers.inaf.it/>

## 2. Summary of survey design and execution

### 2.1. Survey design

VIPERS was designed to sample, at a median redshift  $z \approx 0.7$ , a volume comparable to the one covered by redshift surveys mapping the local Universe (2dFGRS and SDSS), with a similarly high sampling density of the galaxy population. To achieve useful spectral quality in a limited exposure time using the VIMOS spectrograph (Le Fèvre et al. 2003b), a relatively bright limit of  $i_{AB} \leq 22.5$  was adopted, and this generated two main issues for efficient sample selection. At this depth, many galaxies will lie below the redshift range of interest, i.e.  $0.5 < z < 1.2$  (see for example Le Fèvre et al. 2005, 2013a; Lilly et al. 2009); also, it was known from previous similar studies such as the VVDS-Wide (Garilli et al. 2008) that such a purely magnitude-limited sample would suffer from approximately 30% stellar contamination. Here we give a brief summary of the steps taken to overcome these difficulties (see Guzzo et al. 2014; and Garilli et al. 2014; for fuller details).

The VIPERS target selection was derived from the ‘T0005’ release of the CFHTLS Wide photometric survey, completed and improved using the subsequent T0006 release. A preliminary multi-band catalogue, including all objects with extinction-corrected apparent magnitude  $i_{AB} \leq 22.5$ , was built starting from the individual CFHTLS 1-deg<sup>2</sup> tiles. Particular care was taken to verify the homogeneity of these original single-tile catalogues: by analyzing the colour-colour stellar locus within each such catalogue, we were able to identify significant tile-to-tile offsets in the photometric zero-points for different photometric bands. To ensure that the final VIPERS parent photometric catalogue was as spatially homogeneous as possible, a tile-to-tile offset correction to the observed colours was estimated and applied. As discussed extensively by Guzzo et al. (2014), this offset was obtained by comparing the position in colour space of the (well-defined) stellar locus with that of a reference tile (the one overlapping the VVDS F02 survey field, Le Fèvre et al. 2013a).

The homogenisation of galaxy colours over the full area was particularly crucial for the subsequent removal of low-redshift galaxies (nominally  $z < 0.5$ ), which was implemented via a colour-colour selection in the  $(r-i)$  vs.  $(u-g)$  plane: only galaxies whose colours obey the following relation:

$$(r-i) > 0.5 \times (u-g) \quad \text{OR} \quad (r-i) > 0.7 \quad (1)$$

are included in the VIPERS parent sample. This relatively simple criterion, which was preferred to a more complex one based on photometric redshifts because it could be easily implemented for comparisons using other data sets, either real or simulated ones, was tuned using the VVDS complete redshift data (details can be found in Guzzo et al. 2014). The effectiveness of the criterion is due to the fact that galaxy spectral type, as measured by the strength of the D4000 spectral break, strongly correlates with the  $(u-g)$  colour, whereas the  $(r-i)$  one depends strongly on galaxy redshift. Therefore when moving from blue to red  $(u-g)$  colour along the selection boundary one finds galaxies of all spectral types, from the star-forming (low D4000) to the passively evolving (high D4000) ones. On the contrary, when moving from blue to red  $(r-i)$  colour away from the selection boundary, one finds galaxies of progressively larger redshift. However, the natural spread in galaxy properties is somewhat limiting the effectiveness of the colour criterion: galaxies in any redshift interval  $\Delta(z) = 0.1$  occupy a region approximately 0.2 mag wide in  $(r-i)$  colour, significantly broader than the intrinsic uncertainty with which this colour is measured in the

photometric catalogue (approximately 0.05 mag). As a result, the transition region from full rejection of low redshift galaxies to full inclusion of high redshift ones spans the redshift range  $0.4 < z < 0.6$  (see Sect. 3.3).

Finally, stellar objects were removed using a combination of two methods: for objects brighter than  $i_{AB} = 21.0$ , stars were identified on the basis of their half-flux radius, as measured on the  $i$ -band CFHTLS images; for fainter objects, a combination of image size and SED fitting of the 5-band CFHTLS photometry was used (see Appendix A of Guzzo et al. 2014; and Sect. 2.1 of Garilli et al. 2014).

Overall, some 21% of the objects in the total photometric catalogue have been removed because they were classified as stars, 32% were removed because they were classified as low redshift galaxies, and the remaining 47% became the VIPERS parent photometric sample, which was then supplemented with a small additional sample of AGN candidates, chosen from objects that were initially classified as stars on the basis of a colour-colour criterion (see Sect. 2.2 of Garilli et al. 2014). This sample contributes on average 2–5 objects per VIMOS quadrant (against about 90 galaxy targets) with negligible impact on the galaxy selection function. In the PDR-2 catalogue these additional objects can be easily identified and separated from the main galaxy sample through an appropriate keyword, as described in Sect. 5.

### 2.2. Improvements in CFHTLS photometry during the construction of VIPERS

The tile-to-tile colour shifts in the T0005 data discussed in the previous section were a clear evidence that the initial global photometric calibration could be significantly improved. Such a step forward was provided by the CFHTLS T0007 revision (Hudelot et al. 2012). For VIPERS, the most important feature of T0007 compared to previous releases is that each tile in the CFHTLS was rescaled to an absolute calibration provided by a new dedicated survey of calibrators carried out at the CFHT. In addition, in order to ensure that seeing variations between tiles and filters were correctly accounted for, aperture fluxes were rescaled to allow for the seeing of each individual tile. These aperture fluxes have then become the basis of the new photometric catalogue for VIPERS, since it has been shown that Kron (mag auto) flux estimates provide less accurate colour estimates, which lead, among other things, to worse photo- $z$ 's (Moutard et al. 2016b; Hildebrandt et al. 2012). A new photometric catalogue was therefore created, based on Terapix T0007 isophotal aperture magnitudes, with the addition of UV photometry from GALEX (Martin et al. 2005), and NIR  $K_s$ -band photometry from WIRCam (Puget et al. 2004) or from VISTA (Emerson et al. 2004), obtained as part of the VIPERS Multi-Lambda Survey (VIPERS-MLS; Moutard et al. 2016b) or the VISTA Deep Extragalactic Observations (VIDEO; Jarvis et al. 2013), respectively. The isophotal magnitudes were then corrected to pseudo-total ones using an aperture correction for each individual object, obtained as the average of the aperture corrections obtained for the  $g$ ,  $r$ ,  $i$ , and  $K_s$  bands. Details can be found in Moutard et al. (2016b).

This new photometric catalogue became available after VIPERS was well under way, and it was therefore decided not to replace the original parent VIPERS catalogue, because such a substitution would have turned the original well-defined magnitude limit into a somewhat fuzzy limit, resulting from the approximately 0.05 mag scatter in the magnitude comparison between the T0005 and T0007 photometry (at the 22.5 mag limit of the catalogue). A choice was therefore made to match the two

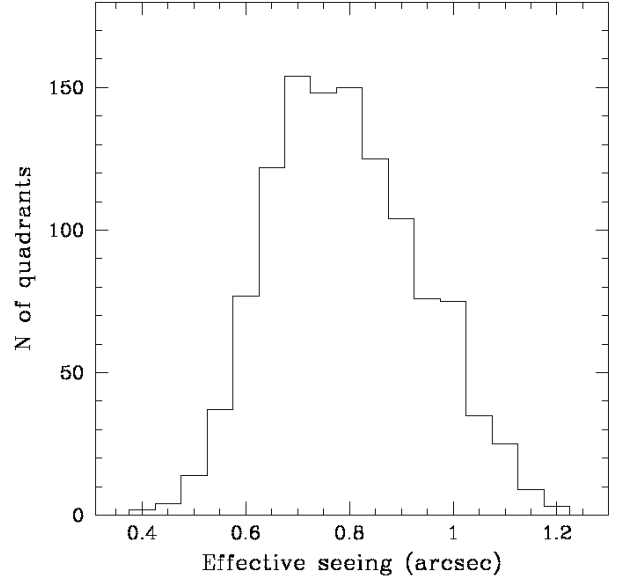
catalogues, and to provide T0007 photometry for all the objects in the original T0005-based catalogue. The matching was carried out using a match circle of 0.6 arcsec, which ensures a 97% matching success rate, with an almost null incidence of multiple object matches. Although the original VIPERS catalogue is limited to  $i_{AB} \leq 22.5$ , the matching was carried out limiting both the old and the new catalogue to  $i_{AB} \leq 23.0$ , to avoid the scatter at the catalogue limit to affect the resulting match. For the small fraction of objects that could not be satisfactorily matched (mostly due to small differences in the source deblending procedure), we computed pseudo-T0007 magnitudes: we estimated the median offset between the T0005 and T0007 catalogues for each photometric band and each CFHTLS parent tile using objects with  $i_{AB} \leq 21.0$ , and added this offset to the original T0005 magnitudes (the magnitude uncertainty was kept equal to the T0005 value). Areas in the photometric catalogue with poor quality, corrupted source extraction, or bright stars are described by a binary mask, as discussed in Sect. 3.

The small magnitude differences between the two catalogues are not expected to have any significant impact on the VIPERS parent sample selection. As already discussed in Sect. 2.1, galaxies of any spectral type are present close to the selection line in the colour-colour plane, and the colour span for galaxies in a small redshift range of  $\Delta(z) = 0.1$  is of approximately 0.2 mag in  $(r - i)$  colour, significantly broader than the intrinsic uncertainty with which this colour is measured in the photometric catalogue. As a consequence, we observe a smooth transition from full rejection to complete inclusion in the VIPERS parent sample across all galaxy types.

### 2.3. Spectroscopic observations

All VIPERS observations were carried out using VIMOS (Visible Multi-Object Spectrograph), on “Melipal”, Unit 3 of the ESO Very Large Telescope (VLT) – see Le Fèvre et al. (2003b). VIMOS is a 4-channel imaging spectrograph; each channel (a ‘quadrant’) covers  $\sim 7 \times 8$  arcmin<sup>2</sup> for a total field of view (a ‘pointing’) of  $\sim 224$  arcmin<sup>2</sup>. Each channel is a complete spectrograph with the possibility of inserting  $30 \times 30$  cm<sup>2</sup> slit masks at the entrance focal plane, as well as broad-band filters or grisms. The precise sizes of the quadrants are in principle all slightly different from each other: the four channels of VIMOS all differ slightly, and they also changed with time during the survey development, when the VIMOS CCDs were refurbished (see below). There is also variation from one pointing to another, e.g. due to vignetting by the guide star probe. All these pieces of information are quantified accurately by the mask files associated with the PDR-2 release, which we discuss in the following section.

The pixel scale on the CCD detectors is 0.205 arcsec/pixel, providing excellent sampling of the Paranal mean image quality and Nyquist sampling for a slit 0.5 arcsec in width. For VIPERS, we used a slit width of 1 arcsec, together with the “low-resolution red” (LR-Red) grism, resulting in a spectral resolution  $R \approx 220$  at the centre of the wavelength range covered by this grism (i.e.  $\sim 5500$ – $9500$  Å). In summer 2010, VIMOS was upgraded with new red-sensitive CCDs in each of the four channels, as well as with a new active flexure compensation system. The reliability of the mask exchange system was also improved (Hammersley et al. 2010). The original thinned E2V detectors were replaced by twice-thicker E2V devices, considerably lowering the fringing and increasing the global instrument efficiency by up to a factor 2.5 (one magnitude) in the redder part of the



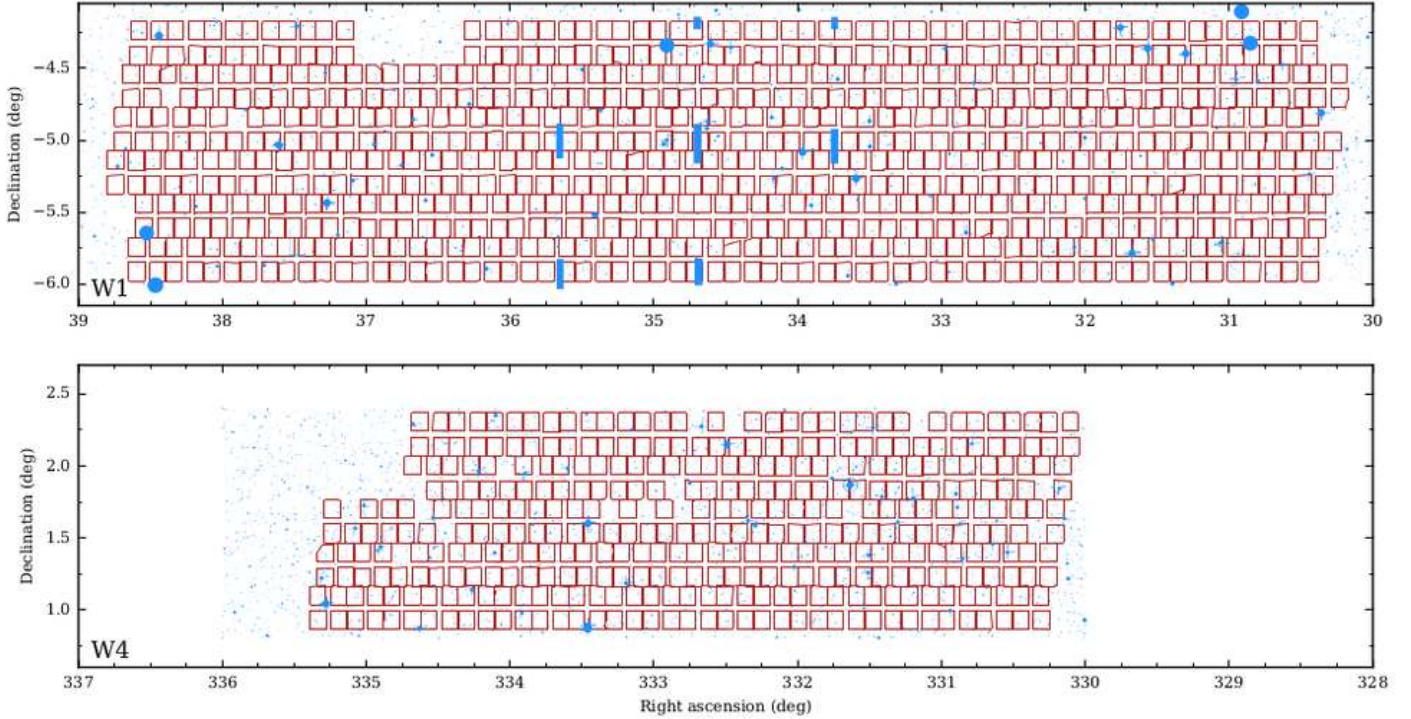
**Fig. 1.** The effective seeing distribution for the VIPERS observations. The seeing value is obtained from the measurement of the FWHM size of the spectral traces for bright objects in the spectroscopic exposures.

wavelength range. This upgrade significantly improved the average quality of VIPERS spectra, resulting in a significantly higher redshift measurement success rate.

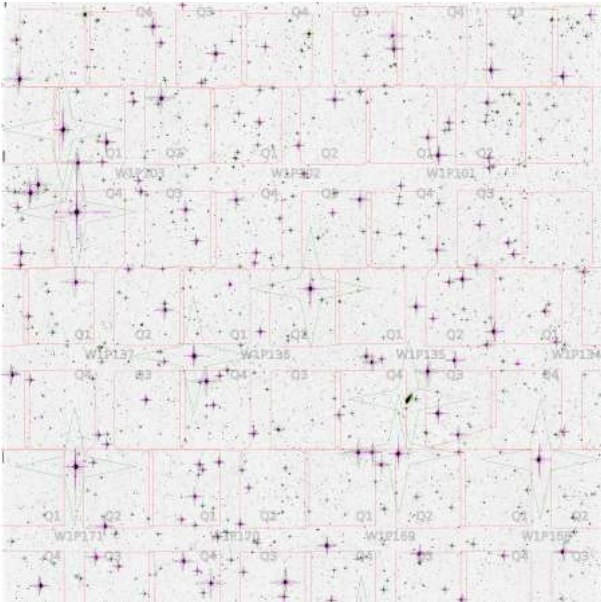
The complete VIPERS survey consists of 288 VIMOS pointings, 192 over the W1 area, and 96 over the W4 area of the CFHTLS, overlapping a total sky area of about 23.5 square degrees. Due to the specific footprint of VIMOS, failed quadrants and masked regions, this corresponds to an effectively covered area of 16.3 square degrees. The number of slits in the spectroscopic masks ranged from 60 to 121 per VIMOS quadrant, with a median value of 87, for a total of 96 929 slits over the whole survey. For nine pointings, observations were repeated because the original observation was carried out under sub-optimal seeing and/or atmospheric conditions, while four pointings already observed before the VIMOS 2010 upgrade were re-observed as part of the related re-commissioning. Overall, observations were carried out starting in November 2008, and were completed by December 2014. For 23 pointings (6 in the W1 area, and 17 in the W4 one) some mask insertion problem prevented the acquisition of useful spectroscopic data in one of the four VIMOS quadrants, leaving some small “holes” in the survey sky coverage (see Fig. 2). These are termed “failed” quadrants. Airmass during the observations ranged from 1.06 to 1.44, with a median value of 1.14, while the effective seeing (measured directly from the observed size of the reference objects used to align the VIMOS masks) ranged from 0.41 to 1.21 arcsec, with a median value of 0.78 arcsec. Figure 1 shows the distribution of these effective seeing values.

### 3. Sky coverage, angular selection functions and completeness

The VIPERS angular selection function is the result of the combination of several different angular completeness functions. Two of these are binary masks (i.e. describing areas that are fully used or fully lost). The first mask (that we call the photometric mask) is related to defects in the parent photometric sample (mostly areas masked by bright stars) and the other (that we call the spectroscopic mask) to the specific footprint of



**Fig. 2.** Layout on the sky of all pointings that contribute to the PDR-2 final release, for the two fields W1 and W4, superimposed on the photometric survey mask. The contours of each of the four quadrants that comprise all VIMOS pointings are shown in red. The blue (grey) areas in the background correspond to areas where the parent photometry is corrupted or observations are not possible due to the presence of foreground objects, particularly bright stars and their diffraction spikes. Note that in this plot failed quadrants and other features introduced by the “spectroscopic mask” are not indicated (see Sect. 3.2 and Figs. 3 and 5).



**Fig. 3.** A  $1 \text{ deg}^2$  detail of the masks developed for VIPERS. The revised photometric mask built for VIPERS corresponds to the magenta circles and cross patterns; for comparison, the original, more conservative mask distributed by Terapix is shown in green. The quadrants that make up the VIPERS pointings are plotted in red. In the background is the CFHTLS T0006  $\chi^2$  image of the field 020631–050800 produced by Terapix. Note the significant gain in usable sky obtained with the new VIPERS-specific mask.

VIMOS and how the different pointings are tailored together to mosaic the VIPERS area. The other completeness functions are

provided instead on a per-galaxy basis: 1) within each of the four VIMOS quadrants on average only 47% of the available targets satisfying the selection criteria are actually placed behind a slit and observed, defining what we call the target sampling rate; 2) since the set of available targets is defined based on the observed colour, as discussed in Sect. 2.1, a colour sampling rate is needed to keep this selection effect into consideration; 3) varying observing conditions and technical issues determine a variation from quadrant to quadrant of the actual number of redshifts measured with respect to the number of targeted galaxies, while our capability to measure the redshift depends on intrinsic galaxy parameters, as we shall discuss in Sect. 3.5 when introducing the spectroscopic success rate.

Detailed knowledge of all these contributions is a crucial ingredient for computing any quantitative statistics of the galaxy distribution, as e.g. its first and second moments (i.e. luminosity/stellar mass functions and two-point correlation functions, respectively).

### 3.1. CFHTLS-VIPERS photometric mask

The photometric quality across the CFHTLS images is tracked with a set of masks that account for imaging artefacts and non-uniform coverage. We use the masks to exclude regions from the survey area with corrupted source extraction or degraded photometric quality. The masks consist primarily of patches around bright stars ( $B_{\text{Vega}} < 17.5$ ) owing to the broad diffraction pattern and internal reflections in the telescope optics. At the core of a saturated stellar halo there are no reliable detections, leaving a hole in the source catalogue, while in the halo and diffraction spikes spurious sources may appear in the catalogue due to false detections. We also add to the mask extended extragalactic

sources that may be fragmented into multiple detections or that may obscure potential VIPERS sources. The details of the revised photometric mask construction were given in Guzzo et al. (2014) and a visual rendition of the two W1 and W4 masks is given in Fig. 2, while Fig. 3 provides a zoom into a smaller area, in particular showing the details of the custom-developed VIPERS photometric mask, compared to the original CFHTLS mask.

It is important to stress here that a small fraction of spectroscopically observed galaxies actually fall within regions forbidden by the photometric mask. These are objects for which, typically, one of the photometric bands had too large an error to be acceptable, but were nevertheless observed as fillers. As such, in any computation of spatial statistics the photometric and spectroscopic masks must be applied not only to any auxiliary random sample (as typically needed for two-point clustering measurements), but also re-applied to the observed spectroscopic catalogue itself. This is required in order to trim these little “leakages” within a few specific areas.

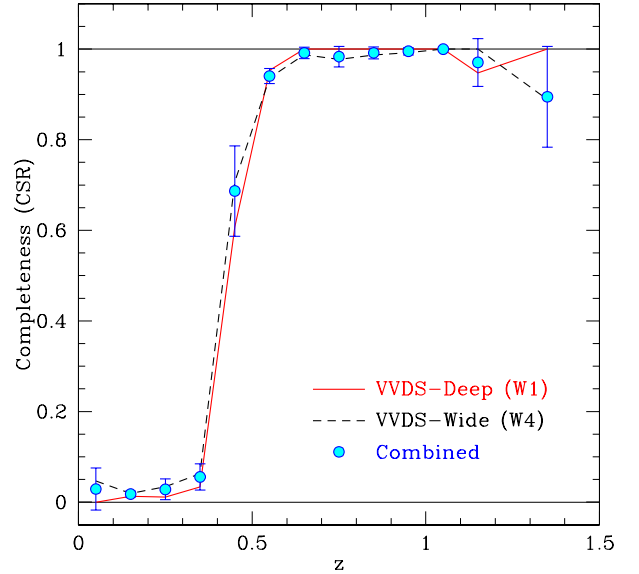
### 3.2. VIPERS spectroscopic masks

The general layout of VIMOS is well known, but the precise geometry of each quadrant has to be specified carefully for each observation, in order to perform precise clustering measurements with the VIPERS data. For example, although it rarely happens, a quadrant may be partly vignetted by the VLT guide probe arm; in addition, the size and geometry of each quadrant changed slightly between the pre- and post-refurbishment data (i.e. from mid-2010 on), due to the dismantling of the instrument and the technical features of the new CCDs. We therefore had to build our own extra mask for the spectroscopic data, accounting for all these aspects at any given point on the sky covered by the survey.

The masks for the W1 and W4 data were constructed from the pre-imaging observations by running an image analysis routine that identifies “good” regions within those images. First, a polygon is defined that traces the edges of the image. The mean and variance of the pixels are computed in small patches at the vertices of the polygon, and these measurements are compared to the statistics at the centre of the image. The vertices of the polygon are then iteratively moved inward toward the centre until the statistics along the boundary are within an acceptable range of those measured at the centre. The boundary that results from this algorithm is used as the basis for the field geometry. The polygon is next simplified to reduce the vertex count: short segments that are nearly co-linear are replaced by long segments. The World Coordinate System information in the fits header is used to convert from pixel coordinates to sky coordinates. Each mask was then examined by eye. Features due to stars at the edge of an image were removed, wiggly segments were straightened and artefacts due to moon reflections were corrected. The red lines in Fig. 3 show the detailed borders of the VIMOS quadrants, describing the spectroscopic mask.

### 3.3. Colour sampling rate

The completeness of the colour-colour pre-selection applied to ideally isolate  $z > 0.5$  galaxies using the CFHTLS corrected photometry, has been quantified and discussed in Guzzo et al. (2014). Using the data from the VVDS survey, the colour sampling rate (CSR) was estimated as a function of redshift. This is shown in Fig. 4 (originally from Guzzo et al. 2014), which we



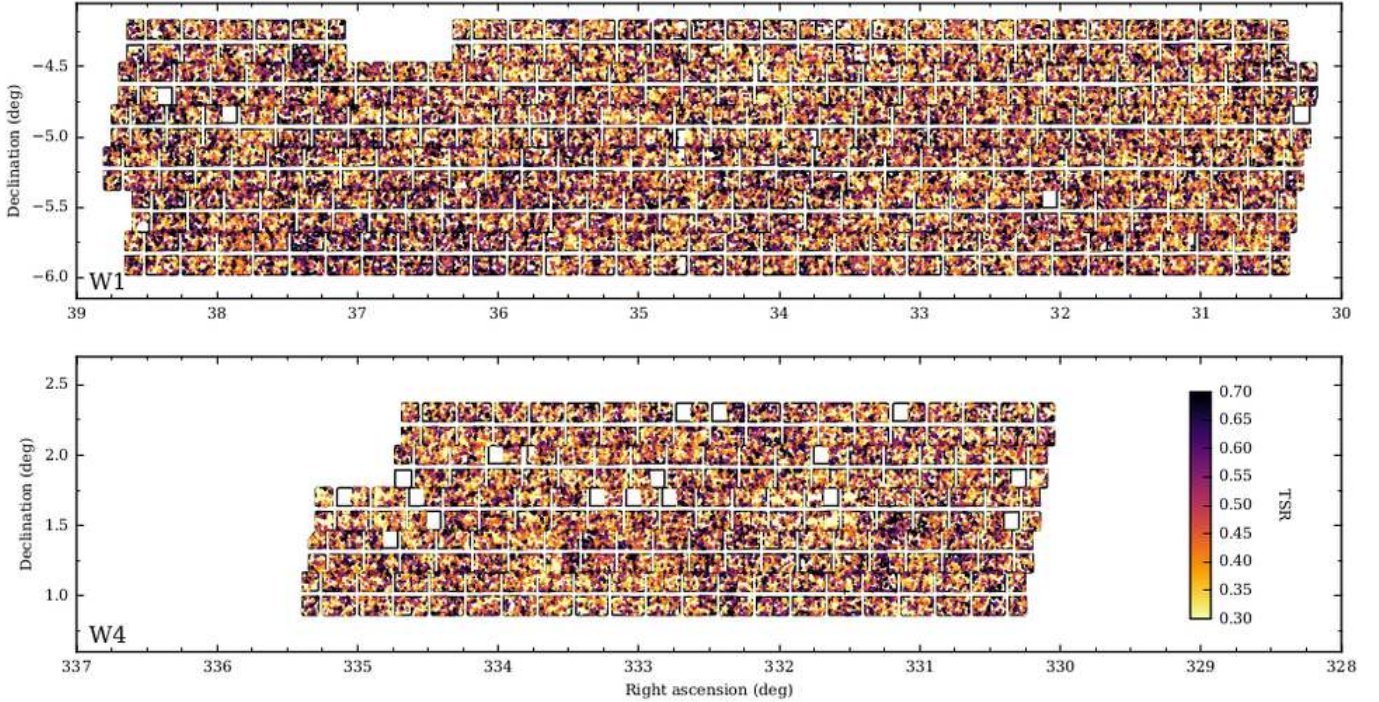
**Fig. 4.** Estimate of the colour sampling rate (CSR) of VIPERS. The plot shows the fraction of galaxies that are selected by the VIPERS colour-colour criteria as a function of redshift, when applied to a sample of galaxies from the VVDS-Deep and VVDS-Wide surveys (Le Fèvre et al. 2013a). This is a highly significant test, given that the original colour-colour boundaries to select VIPERS targets were calibrated on the same VVDS data. Both W1 and W4 fields provide consistent selection functions, yielding a colour selection function that is essentially unity above  $z = 0.6$  and can be consistently modelled in the transition region  $0.4 < z < 0.6$ .

reproduce here for completeness: from this figure it is quite clear how the VIPERS catalogue is virtually 100% complete above  $z = 0.6$ , when compared to a corresponding purely magnitude-limited sample.

### 3.4. Target sampling rate

A Multi-Object Spectrograph (MOS) survey inevitably has to deal with the limitation of MOS slits creating a shadow effect in the targeting of potential sources that is strongly density-dependent. In practice the high-density peaks of the projected galaxy density field are under-sampled with respect to the low-density regions, because the MOS slit length imposes a minimum angular pair separation in the spectroscopic target selection. In VIPERS, this was performed using the SPOC algorithm (Bottini et al. 2005, within the VMPS software distributed by ESO), which maximises the number of slits observed in each quadrant. As a result, (a) very close pairs below a certain scale are practically unobservable; (b) the angular distribution of slits is more uniform than the underlying galaxy distribution. In VIPERS, the first effect suppresses angular clustering below a scale of 5 arcsec, producing a scale-dependent damping of the observed clustering below  $\approx 1 h^{-1}$  Mpc; the second is instead responsible for a nearly scale-independent reduction of the two-point correlation function amplitude above this scale. These effects and their correction are discussed in detail in the parallel paper by Pezzotta et al. (2017). The method builds upon the original approach of de la Torre et al. (2013, by up-weighting galaxies on the basis of the target sampling rate (TSR), namely

$$w_i = \frac{1}{\text{TSR}_i}. \quad (2)$$



**Fig. 5.** Angular distribution over the survey areas of the target sampling rate (TSR, colour scale), estimated locally for each galaxy as described in the text. In this plot, each value has been smoothed on a scale of 2 arcmin to enhance the (inverse) relationship of the TSR with the projected large-scale structure in the galaxy distribution.

In [de la Torre et al. \(2013\)](#), however, the TSR was evaluated only on a quadrant to quadrant basis so that all the targeted galaxies falling in the same quadrant are up-weighted by the same factor. This procedure does not recover the total missing power, since it does not account for the above effects on sub-quadrant scales. The new corrective approach is similar, but uses a local TSR that accounts much more effectively for the angular inhomogeneity of the selection function. This is estimated for each galaxy as the ratio of the local surface densities of target and parent galaxies (i.e. before and after applying the target selection), properly estimated and then averaged within an aperture of appropriate shape and size. If we call these quantities  $\delta_i^s$  and  $\delta_i^p$ , then the  $\text{TSR}_i$  is defined as

$$\text{TSR}_i = \frac{\delta_i^s}{\delta_i^p}. \quad (3)$$

A continuous  $\delta$  field is obtained, starting from the discrete surface distribution, by first using a classical Delaunay tessellation to get the density at the position of each galaxy, and then linearly interpolating. This is finally integrated around the position of each observed galaxy within a rectangular aperture with size  $60 \times 100 \text{ arcsec}^2$  to obtain the local values of  $\delta_i^s$  and  $\delta_i^p$ . It can be shown ([Pezzotta et al. 2017](#)) that a rectangular aperture more efficiently accounts for the angular anisotropy in the distribution of targets within a quadrant introduced by the shadowing effect of the MOS slits.

The resulting distribution of the TSR values over the survey regions is shown in [Fig. 5](#). Thanks to the adopted strategy (i.e. having discarded through the colour selection almost half of the magnitude-limited sample lying at  $z < 0.5$ ), the average TSR of VIPERS is  $\approx 47\%$ , a high value that represents one of the specific important features of VIPERS. For a comparison, with the VVDS-Wide sample, selected to the same magnitude limit, but without colour pre-selection and star rejection, the sampling rate

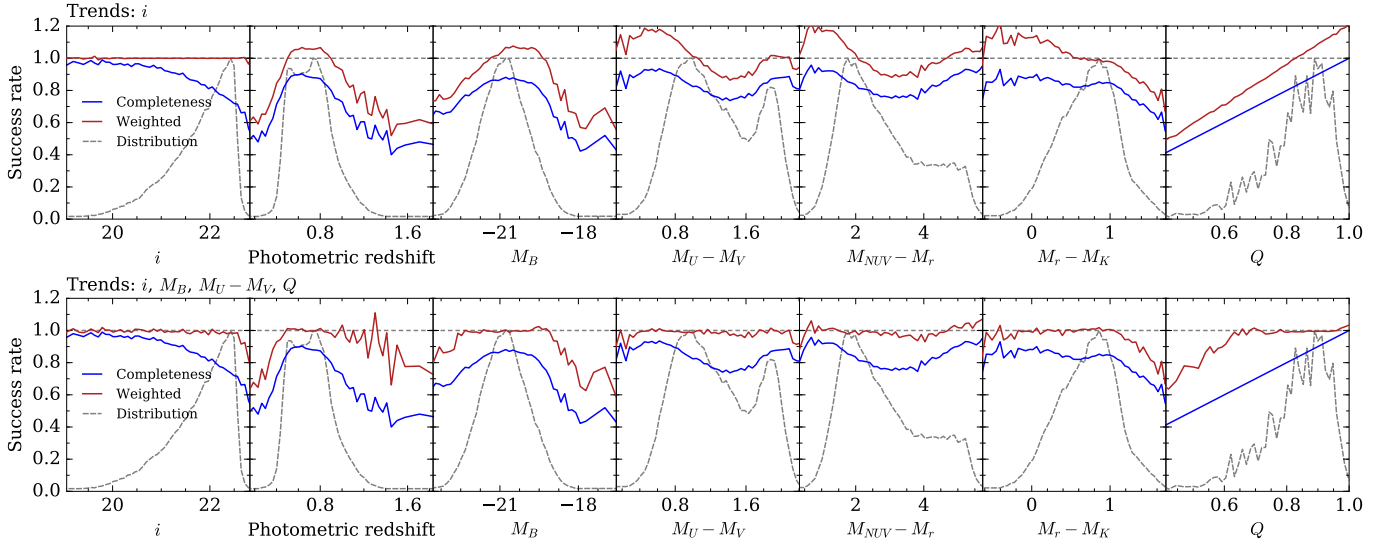
was about 23% ([Garilli et al. 2008](#)), i.e. half of what we have achieved here. We remark how the TSR essentially mirrors the intrinsic fluctuations in the number density of galaxies as a function of position on the sky, and how single quadrants sometimes have a strong internal inhomogeneity in the sampling of galaxies.

### 3.5. Spectroscopic success rate

We quantify the VIPERS redshift measurement success via the spectroscopic success rate (SSR), which is defined as the ratio between the number of objects for which we have successfully measured a redshift  $N_{\text{success}}$  and the number of objects targeted by the spectroscopic observations  $N_{\text{target}}$ . We define the success of a redshift measurement on the basis of the redshift quality flag discussed in detail in [Sect. 4.1](#); generally we adopt the 95% measurement confidence threshold (see [Sect. 4.2](#)) and accept flags 2, 3, 4 and 9 as markers of a successful measurement. Targeted objects for which the spectral extraction completely failed (often these are spurious objects in the photometric catalogue) are not counted among the targets. Additionally, we correct for the stellar contamination by subtracting the number of spectroscopically confirmed stars  $N_{\text{star}}$  from the numerator and denominator. The SSR estimator is therefore:

$$\text{SSR} = \frac{N_{\text{success}} - N_{\text{star}}}{N_{\text{target}} - N_{\text{star}}}. \quad (4)$$

The ability to measure a redshift with confidence depends on a number of factors, starting from the observing conditions for a given survey pointing, and the apparent flux of a given galaxy, but also including intrinsic galaxy properties, such as its spectral type and redshift. The top part of [Fig. 6](#) shows that, if we use just the galaxy apparent  $i$ -band magnitude to parametrize the SSR, we cannot successfully reproduce the SSR dependence on



**Fig. 6.** Spectroscopic success rate (SSR) as a function of different observed photometric properties (blue solid curve), compared to the result of applying the weight to correct incompleteness (red solid curve). The two rows of plots show how the completeness correction changes when one only considers a simple SSR dependence on the selection  $i$ -band magnitude (*top*), or rather includes the more subtle dependencies on observed colours and, in particular, quality of the specific VIMOS quadrant (“ $Q$ ” parameter, quantified via the mean SSR for all galaxies in that quadrant). The differential distribution of each parameter is also plotted in each panel (dashed curve).

other parameters, such as the galaxy rest-frame colour. When instead we use multiple parameters, including the apparent magnitude, the rest-frame colour, the galaxy  $B$ -band luminosity, and the overall quality of the specific VIMOS quadrant (quantified via the mean SSR for all galaxies in that quadrant), we obtain a significant improvement on the SSR capability of describing the VIPERS sample, as shown in the bottom part of Fig. 6. In both rows of plots in that figure we show the SSR as a function of different observed photometric properties (blue solid curve), compared to the result of applying the SSR weight to correct for incompleteness (red solid curve) the total VIPERS sample. Rest-frame properties (galaxy luminosity and colour) in this case are computed on the basis of the galaxy photometric redshift, in order to enable their computation also for galaxies without a spectroscopic redshift measurement (that enter into the denominator of Eq. (3)). Photometric redshift estimates are taken from Moutard et al. (2016b), and have a typical accuracy of  $\sigma_z \leq 0.04$ , with a fraction of catastrophic failures smaller than 2%.

The SSR is computed adaptively using a nearest-neighbour algorithm. Depending on which parameters we want to use to parametrize the SSR, we build an  $N$ -dimensional dataset (with  $N = 1$  when we use only the apparent magnitude, and  $N = 4$  when we add also the rest-frame colour, the luminosity, and the quadrant quality); then, for each object in this  $N$ -dimensional space, we determine the distance  $R_K$  to its  $K$ th nearest neighbour (we use  $K = 100$ ). We then count the number of sources in the successfully measured sample that are contained within the radius  $R_K$ :  $N_{\text{success}}(\leq R_K)$ . The SSR at the specified point is given by the fraction  $SSR = N_{\text{success}}(\leq R_K)/K$ . Distances in this  $N$ -dimensional space are computed using the rank distance measure (i.e. the ranks of each parameter in the sample are used when computing separations).

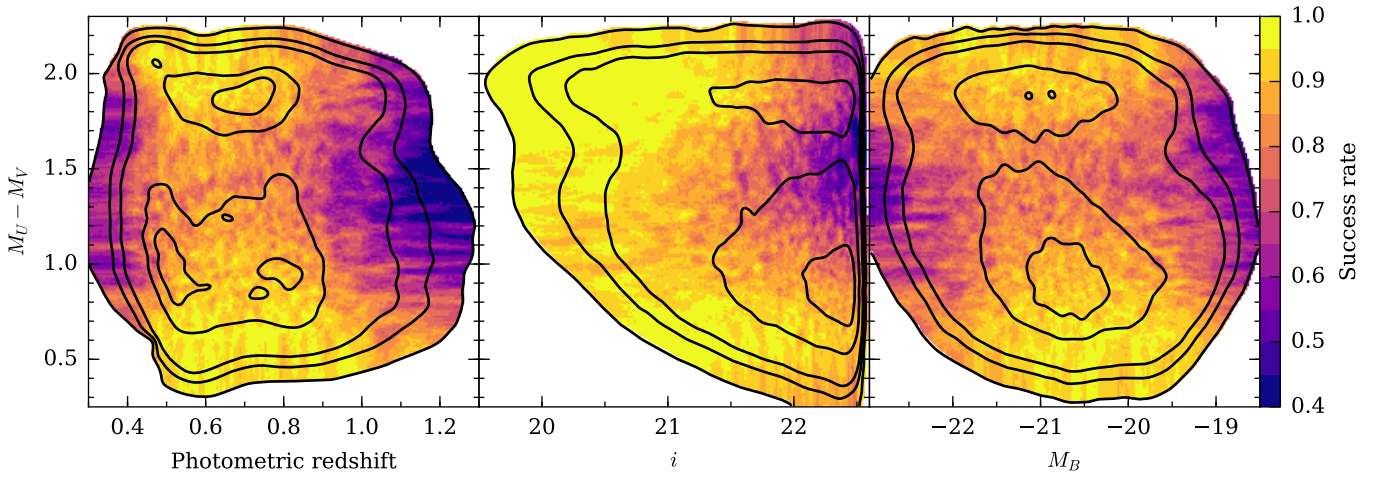
Figure 7 shows the bivariate distribution of SSR values as a function of rest-frame colour and redshift, apparent  $i$ -band magnitude, and rest-frame  $B$ -band luminosity, respectively. The mean SSR is about 83%, but it is clear from this figure that complex SSR variations exist as a function of galaxy properties. An obvious apparent magnitude trend is clearly visible in the

figure’s middle panel, but all panels quite clearly show how the lowest SSR values are characteristic of galaxies with intermediate rest-frame colour. These are objects whose spectra contain neither strong emission lines (as would be seen in the bluest part of the sample) nor a strong 4000 Å break (as would be seen in the reddest part of the sample). This general feature was already observed for the zCOMOS bright survey (Lilly et al. 2009, see their Fig. 2).

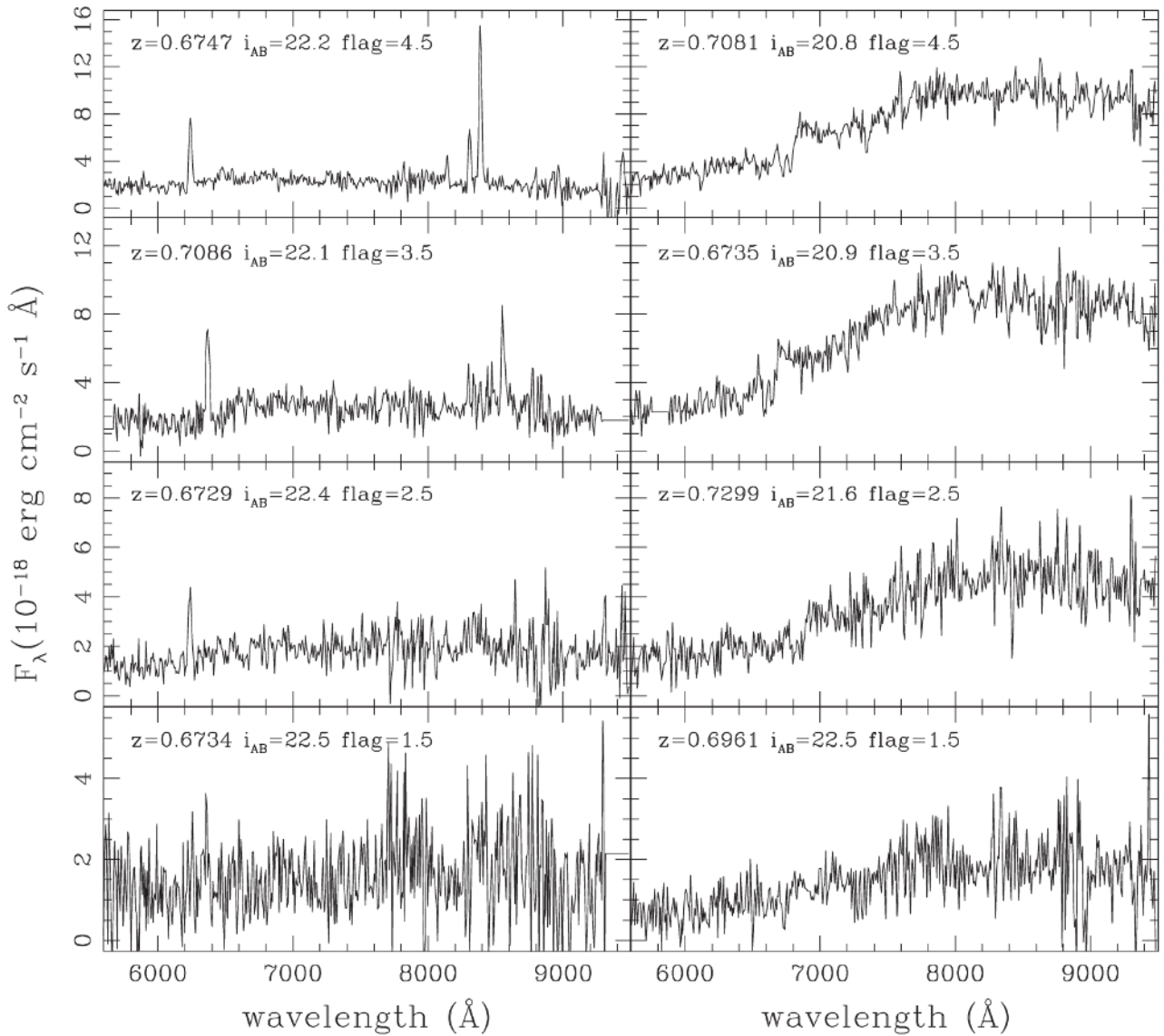
#### 4. Redshift measurements, confidence flags and statistical uncertainties

Figure 8 shows a representative set of spectra of different quality, as available at the end of the automatic data reduction pipeline of VIPERS. This, together with the procedure for redshift validation, have been extensively described in Garilli et al. (2010), Garilli et al. (2012), Guzzo et al. (2014) and Garilli et al. (2014). Here we briefly summarize the last part of this process, i.e. how redshifts are measured and their quality evaluated.

As the final step of the VIPERS automated data-reduction pipeline developed at INAF–IASF Milano (Garilli et al. 2014), redshifts were estimated using the EZ code (Garilli et al. 2010). First, a set of potential emission lines is identified in the 1-dimensional spectrum, and a list of potential redshift estimates is associated to each such line, based on a pre-defined list of typical lines observed in galaxy spectra. If two or more strong emission lines (i.e. lines detected with a signal-to-noise ratio  $S/N > 5$ ) give a matching redshift value, this value is kept as the final measurement. Otherwise full cross-correlations between the 1-dimensional spectrum and a set of galaxy spectra templates are carried out, and the 5 strongest correlation peaks are kept as potential redshift estimates. In this case the final redshift measurement is associated to the galaxy template and redshift estimate combination that provides the best overall fit to the observed 1-dimensional spectrum. Finally, an estimate of the redshift measurement reliability is obtained by repeating the cross-correlation of the observed spectrum with the best fitting template, limited to a pre-defined set of wavelength ranges



**Fig. 7.** 2D plots of the SSR in the photometric parameter space. We show the dependence on the rest-frame  $U - V$  colour jointly with three photometric parameters: photometric redshift (*left*), apparent  $i$ -band magnitude (*center*), and rest-frame  $B$ -band magnitude (*right*). The solid contours contain 10, 50, 90 and 99% of the sample.



**Fig. 8.** Examples of VIPERS spectra: one late-type and one early-type galaxy spectrum is shown for the different redshift measurement quality flags. The measured redshift for all objects is close to  $z = 0.7$ , the peak of the VIPERS redshift distribution.



where, depending on the template and the redshift measurement, we can expect to observe the strongest spectral features (Balmer-series Hydrogen lines and Oxygen lines for late-type templates, D4000, Ca H+K, G-band for early-type templates).

These measurements are then reviewed by two team members independently, using EZ in interactive mode through a user-friendly dedicated interface. If the automated measurement is obviously correct, nothing is modified, and the measurement is stored in the final VIPERS database. Otherwise, the full measurement procedure is re-run by the reviewer, until a satisfactory redshift estimate is achieved. The results of the two reviews are eventually matched, and in case of disagreement the full set of spectroscopic information available (including the one and two-dimensional spectra, the sky background spectrum, and all the cross-correlation results) is visually examined by the two reviewers together, until agreement is reached on a most likely redshift estimate and its associated quality flag.

All redshift measurements presented in the PDR-2 catalogue are as observed and have not been corrected to a heliocentric or Local-Group reference frame. Information to perform these corrections is nevertheless contained in the FITS header of the spectra.

#### 4.1. Redshift quality flags

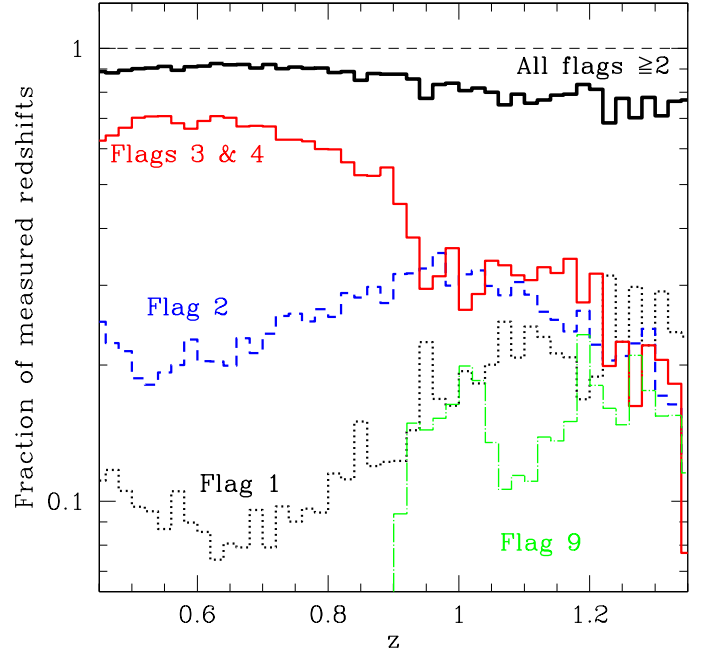
The quality flag system adopted by the VIPERS survey has been inspired by and is in fact very close to those of other precursor surveys (e.g. Le Fèvre et al. 2005; Lilly et al. 2009). The meaning of the various flags has been described in detail in Garilli et al. (2014) and Guzzo et al. (2014); here we repeat the meaning of the flags for those objects released with PDR-2:

- Flags 4.X and 3.X: highly secure redshift, with confidence >99%.
- Flag 2.X: still fairly secure, >95% confidence level.
- Flag 1.X: tentative redshift measurement, with ~50% chance to be wrong.
- Flag 9.X: redshift based a single emission feature, usually [OII]3727 Å. With the PDR-1 data we showed that the confidence level of this class is ~90%.

After the human validation procedure has produced the integer part of the redshift quality flag, a decimal fraction is added to it, with possible values 0.2, 0.4, 0.5, to indicate respectively no, marginal or good agreement of the spectroscopic measurement with the object photometric redshift (see Guzzo et al. 2014, for the specific criteria defining this agreement). If no photometric redshift exists for that object, the decimal part is set to 0.1.

A “1” in front of the above flags indicates a broad lines AGN spectrum, while a “2” indicates a second object serendipitously observed within the slit of a VIPERS target.

In all VIPERS papers, objects with a redshift flag between 2.X and 9.X are referred to as *reliable (or secure) redshifts* and are the only ones normally used in the science analyses. In Garilli et al. (2014) we discussed in detail the reliability of flag 9.X objects. In Fig. 9 the fraction of redshift measurements with a given quality flag is shown as a function of redshift, limited to the main redshift range covered by the VIPERS sample. Notice how the highest quality subset of redshift measurements is always the largest one in the survey, up to  $z \simeq 1.2$  and also how, very importantly, the fraction of measured redshifts with flag 2.X or larger (i.e. the reliable sample to be used for science investigations) is essentially constant out to at least  $z = 1.2$ .



**Fig. 9.** Quality of the VIPERS redshift measurements for different redshifts. Specifically, the plot shows how the fraction of measurements for different quality flags changes as a function of redshift. Note how the “reliable sample”, i.e. that with quality flag  $\geq 2$ , to be used for statistical analyses, shows a stable measured fraction out to the limit of the survey.

#### 4.2. Updated estimate of redshift reliability

As was done for PDR-1, we estimate redshift errors by comparing independent redshift measurements that are available for a subset of galaxies. Some VIPERS targets were observed more than once within the survey, or are in common with other surveys. This also gives us a way to quantify the confidence level of our quality flags.

At the end of the survey, the total number of targets with repeated observations is 3556, compared to 1941 that were available at the time of PDR-1. For 3114 of these, two redshift measurements are available, including any value for the quality flag (see Table 1).

Considering the distribution of the differences between the two redshift measurements  $\Delta z$ , we define matching pairs as those that satisfy the condition  $|\Delta z| < 0.005$ . This threshold has been set on the basis of the first visible gap in the  $\Delta z$  distribution. This identifies 2626 matching pairs (i.e. a matching fraction of 84%). This sample still includes some redshifts with quality flag 1, i.e. redshifts with confidence level ~50–60% (see Guzzo et al. 2014), which are in general not reliable for statistical analyses and have been excluded from all VIPERS investigations so far. Restricting consideration to pairs with both flag 2.X or above (i.e. the reliable measurements), the matching fraction rises to 92.3% (2275 out of 2466). Even further, if we consider only flags either 3.X or 4.X, i.e. the highest quality spectra, the matching reaches 99.1% (1238 out of 1249). Using Eq. (7) in Garilli et al. (2014), we can employ these figures to estimate the average confidence level of single measurements in the reliable redshift sample

$$C_{\text{flag} \geq 2} = \sqrt{0.923} = 96.1\%, \quad (5)$$

and the one of the high-quality redshifts

$$C_{\text{flag} 3,4} = \sqrt{0.991} = 99.54\%, \quad (6)$$

**Table 1.** Statistics of double redshift measurements, both internal to the VIPERS sample and against external data.

Sample	$N_{\text{gal}}$	$N_{\text{agree}}$	Matching (%)	$\sigma_2(\Delta z/(1+z))$	$\sigma_z = \sigma_2/\sqrt{2}$	$\sigma_v = c\sigma_z$ (km s <sup>-1</sup> )	Mean $\Delta v$ (km s <sup>-1</sup> )
Internal Comparison							
All measurements	3114	2626	84.3%	0.00079	0.00056	166	–
Both flags $\geq 2$	2466	2275	92.3%	0.00077	0.00054	163	–
Both flags 3,4	1249	1238	99.1%	0.00075	0.00053	159	–
Both flags $\geq 2$ ; faint	–	1099	–	0.00077	0.00055	164	–
Both flags $\geq 2$ ; bright	–	336	–	0.00071	0.00050	150	–
Both flags $\geq 2$ ; weak EL	–	1153	–	0.00079	0.00056	167	–
Both flags $\geq 2$ ; strong EL	–	508	–	0.00066	0.00046	139	–
External Comparison with VVDS							
All measurements	737	629	85.3%	0.00093	0.00066	198	26
Both flags 3, 4	358	350	97.8%	0.00083	0.00059	177	40
External Comparison with BOSS							
All measurements	747	736	98.5%	0.00064	0.00045*	136*	–108
VIPERS flags 3, 4	690	684	99.1%	0.00061	0.00043*	130*	–113

**Notes.** This is probably an underestimate of the uncertainty on the VIPERS side, because the assumption of uncertainty equipartition does not fully apply in the BOSS comparison.

which agrees very well with the value obtained in Guzzo et al. (2014) and Garilli et al. (2014) using the PDR-1 data.

Comparable results are obtained when comparing the VIPERS measurements to external data. The VIPERS sky areas have a non negligible overlap with the VVDS Wide F22 sample Garilli et al. (2008, 844 galaxies) and the BOSS sample (Dawson et al. 2013; 751 galaxies). The results of these comparisons are shown in Table 1. Note the significantly higher matching fraction in the case of BOSS. This is easily understood when considering that BOSS CMASS galaxies have a magnitude limit brighter than  $i \simeq 20$  and so all matches must correspond to bright VIPERS galaxies. Not surprisingly, then, the matching fraction in this case is comparable to that of the highest quality VIPERS class.

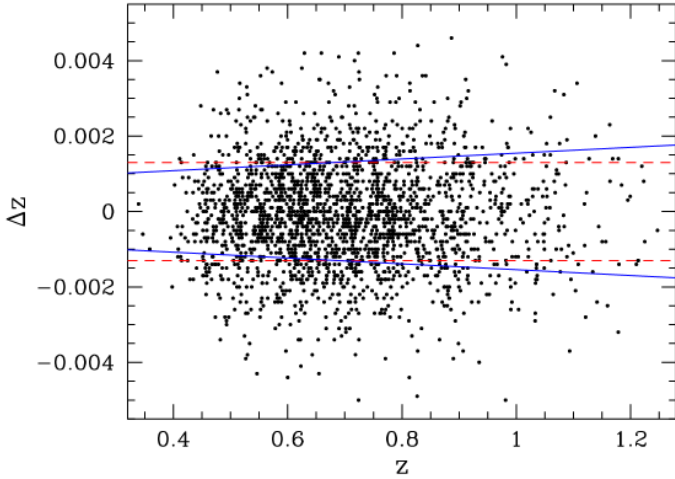
#### 4.3. Updated estimate of redshift errors

An accurate knowledge of the typical redshift measurement uncertainty is clearly important for the scientific analysis of the VIPERS sample, in particular when modelling the observed shape of the power spectrum or the effect of Redshift Space Distortions (see the parallel papers by Rota et al. 2017; de la Torre et al. 2017; Pezzotta et al. 2017; Wilson et al., in prep.). It is also important to probe the possible redshift dependence of this uncertainty. We have used the repeated redshift measurements discussed in the previous section to update the estimate originally presented in Guzzo et al. (2014) and Garilli et al. (2014). To obtain a robust estimate of the measurement error we used the matching pairs of measurements (as defined in the previous section), computed the median absolute deviation (MAD) of the  $\Delta z$  values, and scaled it to the standard deviation equivalent, which for a Gaussian distribution is given by  $\sigma = 1.4826 \times \text{MAD}$ . The resulting scatter, when we consider only reliable redshift measurements (quality flag 2.X and above, 2275 pairs), is  $\sigma_{\Delta z} = 0.0013$ . A very similar result is obtained by fitting a Gaussian to the distribution of  $\Delta z$  values.

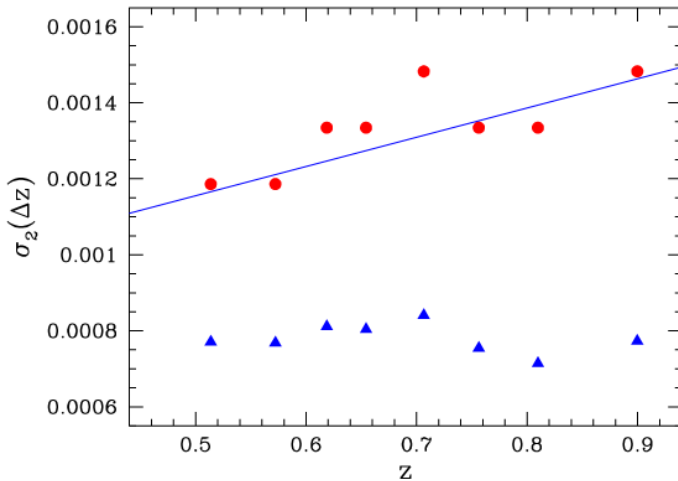
With the current large set of duplicate measurements, we can also re-consider the overall approach used to characterise

the redshift measurement uncertainty. For PDR-1 we adopted the common assumption that redshift uncertainties scale with the redshift itself as  $(1+z)$ . This assumption would apply in the simple case of a spectrograph that yields spectra with a resolution and sensitivity that are constant and independent of wavelength, but neither of these criteria are satisfied by VIMOS spectra. The spectrum signal-to-noise ratio is influenced by the observing conditions, but is nevertheless mostly driven by the galaxy apparent magnitude. This in turn depends (albeit with a significant scatter) on the galaxy redshift, and therefore must induce some increase of the redshift measurement error with redshift. Also, when redshifts are measured through cross-correlation with templates, as in the case of VIPERS, some scaling of the uncertainty with redshift could be expected because of the  $(1+z)$  shrinking of the available rest-frame wavelength range (although when the redshift measurement is dominated by a few key features in the spectrum, which remain observable over most of the survey redshift range – e.g. the [OII] $\lambda 3727$  line and 4000 Å break region – this effect should be negligible). In short, a redshift dependence of the measurement error can be expected, but not necessarily with a linear dependence on  $(1+z)$ .

The observed distribution of  $\Delta z$  values as a function of galaxy redshift is shown in Fig. 10, limited to the matching pairs of measurements. Using these data points, we have determined the scatter in both the  $\Delta z$  and the  $\Delta z/(1+z)$  values within separate redshift bins, still using the MAD estimator as for the whole sample. The results of this estimate are shown in Fig. 11, where it is clearly seen that the observed scatter of the  $\Delta z/(1+z)$  differences is substantially constant, whereas the scatter of the  $\Delta z$  differences increases with the median redshift of the bin, with a scaling that is roughly proportional to  $(1+z)$ . We therefore conclude that this simple scaling does after all provide an adequate description of the effective uncertainty of the VIPERS redshift measurements. We do not extend this exercise to redshifts above 1.1, because of the limited number of repeated measurements available at those redshifts. The final estimate we obtain for the single redshift measurement uncertainty is therefore  $\sigma_z = 0.00054 \times (1+z)$ , which we can compare with the figure of  $\sigma_z = 0.00047 \times (1+z)$  given in the PDR-1 paper (on the basis



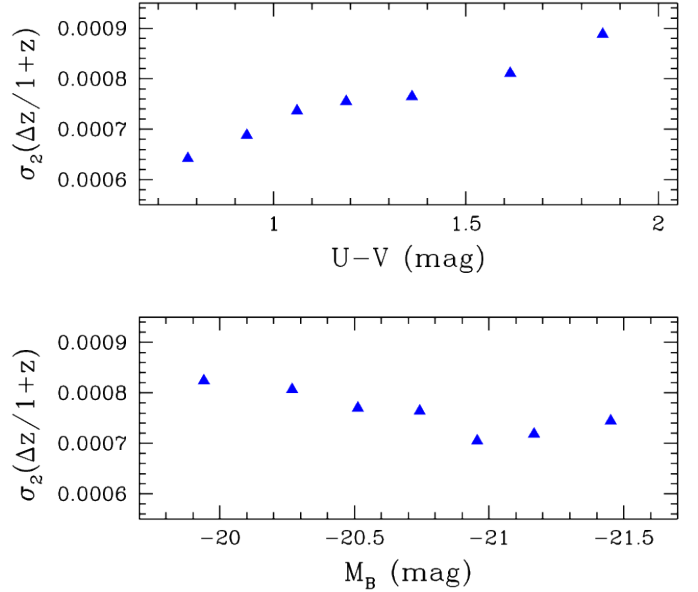
**Fig. 10.** Redshift measurement differences between two independent observations of the same object, obtained from a set of 2466 VIPERS galaxies with quality flag  $\geq 2$ , plotted as a function of the object redshift. Non-matching measurements (defined as being discrepant by more than  $|\Delta z| = 0.005$ ) have been excluded from the plot. The dashed red and continuous blue lines show the 1-sigma scatter estimates obtained assuming the scatter to be independent from  $z$ , or scaling as  $(1+z)$ , respectively.



**Fig. 11.** Scatter in the  $\Delta z$  values (red circles) and in the  $\Delta z/(1+z)$  values (blue triangles) for objects in a narrow redshift bin, as a function of the median redshift for the bin. The bins are built to contain a fixed number of objects (500), and they are not statistically independent since they partly overlap each other. The blue line shows the mean scatter derived from the blue points, with the  $(1+z)$  scaling applied to it.

of 1235 measurement pairs). Notice that this is the estimate for the single measurement uncertainty, which we derive from the scatter measurement listed in Table 1 by assuming an equal contribution from the two measurements to the overall uncertainty, and therefore the single measurement uncertainty is obtained as the scatter value divided by a factor of  $\sqrt{2}$ .

Using the sample of 2275 reliable pairs of measurements (flags 2, 3, 4, 9), it is interesting to further explore how this scatter varies within the sample, as a function of galaxy apparent magnitude, luminosity, and rest-frame colour, or when strong emission lines are present in the spectrum. We isolate bright and faint subsets (galaxies with  $i_{AB} \leq 20.75$  and  $i_{AB} \geq 21.75$ , respectively), and subsets with strong or weak emission-line spectra (galaxies with  $\text{flux}(\text{[OII]}) \geq 1.5 \times 10^{-16}$  and between  $2.0 \times 10^{-17}$  and  $1.5 \times 10^{-16}$   $\text{erg cm}^{-2} \text{s}^{-1} \text{Å}^{-1}$ , respectively).



**Fig. 12.** Scatter in the  $\Delta z/(1+z)$  values for objects in a narrow rest-frame  $U-V$  color bin, as a function of the median color for the bin (*top panel*) and for objects in a narrow  $B$ -band absolute magnitude bin, as a function of the median magnitude for the bin (*bottom panel*). In both cases the bins are built to contain a fixed number of objects (500), and they are not statistically independent since they partly overlap each other.

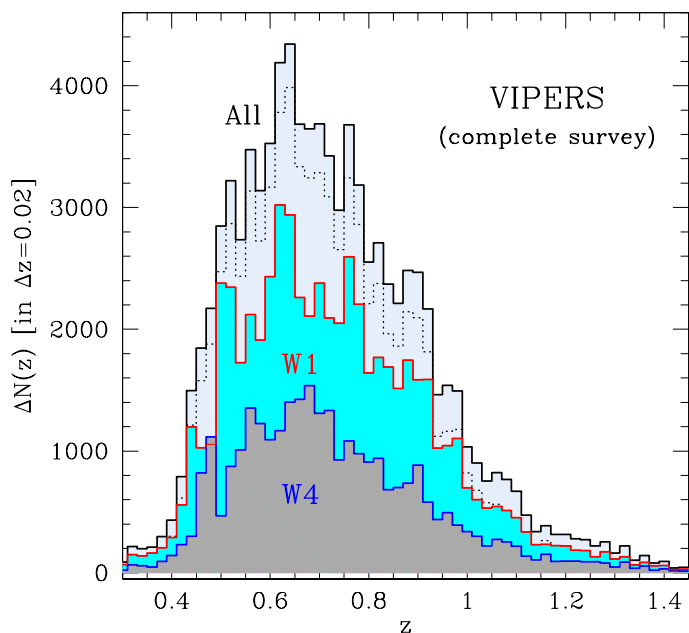
As expected, significant differences are observed between these subsets. Table 1 shows that galaxies which are brighter or have strong emission lines in their spectra provide on average significantly more accurate redshift measurements, respectively by about 10 and 20%. The same result is obtained when we examine the scatter as a function of galaxy rest-frame colour, with blue star-forming galaxies with strong emission lines in their spectra having more accurate redshift measurements than red passive ones (see the top part of Fig. 12). On the contrary, galaxy luminosity has limited influence on the redshift measurement accuracy, as can be seen from the bottom part of Fig. 12, with less than a 10% variation in accuracy between high and low luminosity galaxies.

As in the case of the confidence level estimates, the comparison with independent surveys further confirms the accuracy of the VIPERS redshift measurements. The scatter between VIPERS and the VVDS is only marginally larger than that observed in the internal comparison. The scatter against BOSS is instead significantly smaller, because the higher spectral resolution of the BOSS spectra and the brighter magnitude limit of the BOSS sample both contribute to a higher redshift measurement accuracy for the BOSS galaxies, and therefore in this case we do not have equal contributions to the observed scatter, with the result that this is somewhat reduced in its observed value.

## 5. The PDR-2 dataset

### 5.1. The PDR-2 sample

In total we have obtained spectra for 97 414 objects (see summary in Table 2): 94 335 main survey targets (i.e. the objects selected on the basis of the colour-colour selection criterion discussed in Sect. 2.1), 1478 serendipitous targets in the slits, and 1601 AGN candidates, originally selected on the basis of their colour. The total number of measured redshifts is 91 507 (quality



**Fig. 13.** Galaxy redshift distribution in the final VIPERS PDR-2 catalogue (black solid line), and separately within the W1 and W4 fields (red and blue solid lines, respectively). These include all measured redshifts, with flag 1 or larger. The dotted line shows the result of plotting only the flag  $\geq 2$  galaxies, i.e. those that can be used reliably in statistical analyses.

**Table 2.** VIPERS PDR-2 spectroscopic sample.

Sample	Number
Spectroscopically observed	97 414
— Main survey targets	94 335
— Serendipitous targets	1478
— AGN candidates (not part of main survey)	1601
Measured redshifts	Number
All measured	91 507
Main survey, all targets	89 022
— galaxies	86 775
— stars	2247
Flag $\geq 2$ main survey, all targets	78 586
Flag $\geq 2$ main survey, galaxies	76 552

flags 1 and above); of these, 86 775 make up the main galaxy survey, with the remaining ones belonging to stars and colour selected AGN. The galaxy sample with reliable redshift measurements (quality flag from 2.X to 9.X), contains 76 552 objects, with a median redshift of 0.69, and with 90% of the objects located within the redshift range (0.43, 1.04).

Figure 13 shows the redshift distribution of the final data set, providing also the two separate distributions for the two survey areas W1 and W4. This gives a visual impression of the level of sample variance still present in the redshift distribution averaged over areas of this size. The two panels of Fig. 14 show instead the distribution of  $B$ -band luminosity and of stellar mass for the reliable redshift sample, to give a quantitative idea of the distribution of these two important galaxy properties within the VIPERS sample. Luminosities and stellar masses have been estimated through SED fitting of the available photometry (from UV to  $K$ ) as described in Moutard et al. (2016a). It is clear from

these figures that the colour-colour selection used to select the VIPERS parent sample (see Sect. 2.1) has been highly effective in selecting a well defined galaxy sample at  $z > 0.5$ , which includes a sizeable set of very massive and bright galaxies.

Some detail about the large-scale distribution of the VIPERS galaxies is instead provided by the cone diagrams of Fig. 15, which show quite clearly the abundance of structures sampled by VIPERS, and the amount of segregation of the overall galaxy population as a function of the local galaxy density.

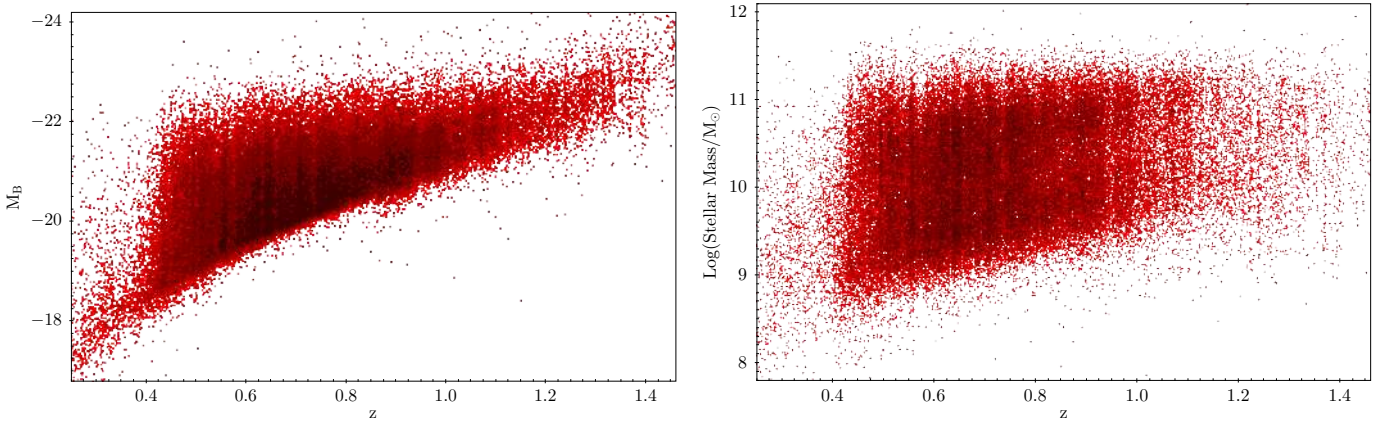
## 5.2. Access to publicly released data

The PDR-2 data release includes:

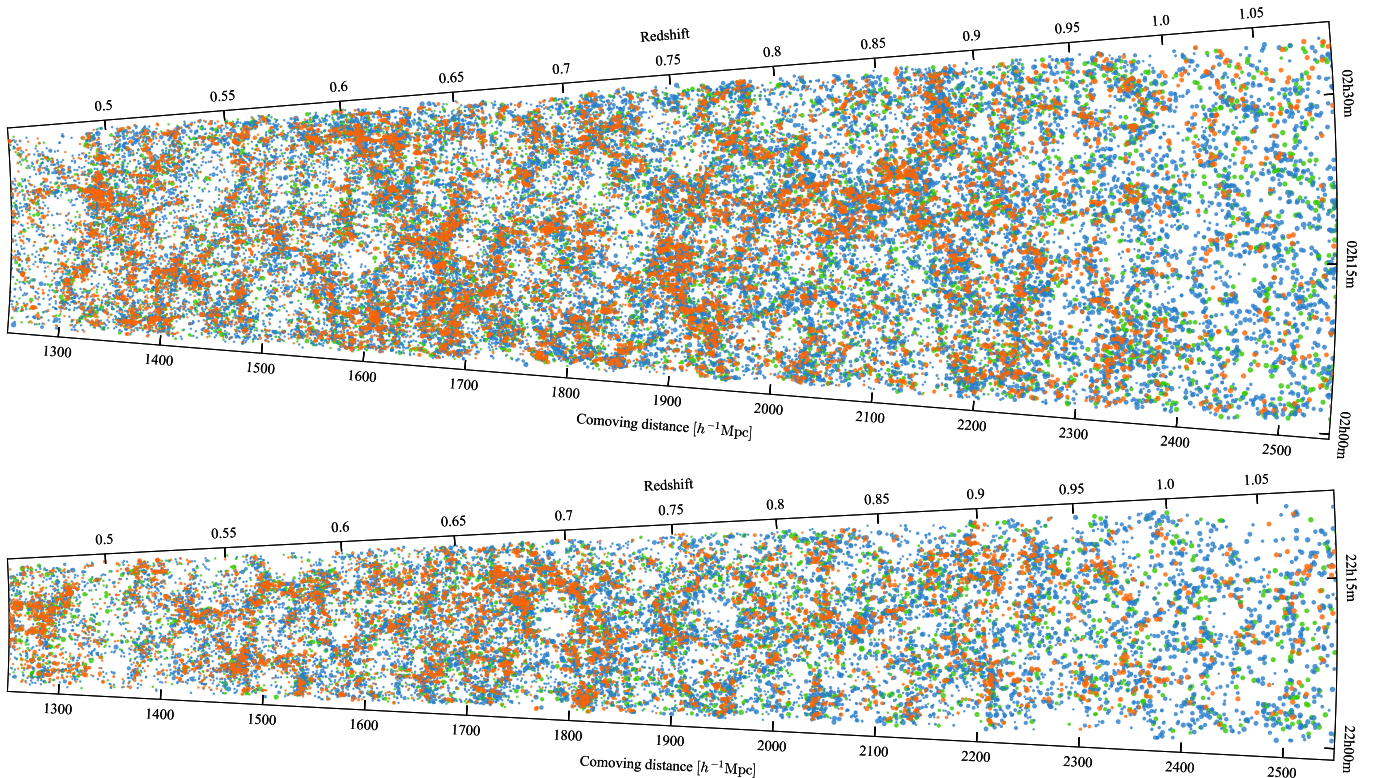
- The total reference photometric catalogue, including both the VIPERS parent photometric sample, the objects excluded from this sample on the basis of the colour-colour criterion described in Sect. 2.1, and the objects classified as stars; the parameter classFlag is used in the catalogue to identify these categories (objects in the VIPERS parent photometric sample have classFlag = 1). The catalogue contains both the CFHTLS release T0005 photometry used for the VIPERS sample selection (see Sect. 2.1), and the VIPERS Multi-Lambda photometry (CFHTLS release T0007 supplemented with UV and  $K$ -band data), described in Sect. 2.2.
- The photometric and spectroscopic masks described in Sects. 3.1 and 3.2, in the form of DS9 region files.
- The redshift catalogue with the parameters described in Sect. 4, and which includes the estimates of the target sampling rate and spectroscopic success rate described in Sects. 3.4 and 3.5. This catalogue includes all the available redshift measurements (quality flag  $\geq 1$ ), and not only those in the reliable redshift sample. Also in this catalogue the parameter classFlag can be used to identify the objects that belong to the main survey sample (objects in this sample have classFlag = 1).
- The wavelength- and flux-calibrated one-dimensional spectra for the objects included in the redshift catalog; these are provided as FITS tables containing the following columns: 1) wavelength (in Angstrom); 2) cleaned-spectrum flux (in  $\text{erg cm}^{-2} \text{s}^{-1} \text{A}^{-1}$ ); the spectral cleaning procedure, based on a PCA reconstruction of the observed spectra to remove the strongest residuals from the sky subtraction, is described in detail in Marchetti et al. (2017); fluxes are also corrected for atmospheric absorption effects; 3) flux uncertainty estimate (same units as the flux); 4) the sky intensity estimate for the slit (in counts); 5) original spectrum flux (same units as the flux), with only the correction for atmospheric absorption applied; 6) the cleaned-spectrum mask (digital mask to identify the pixels edited by the cleaning procedure).
- Optionally, also the two-dimensional spectra are available; these are wavelength calibrated, but without the flux calibration nor the correction for atmospheric absorption applied.

The complete data release of VIPERS is available at <http://vipers.inaf.it>, while the spectroscopic part is also available via the ESO archive at <http://eso.org/rm/publicAccess#/dataReleases>.

*Acknowledgements.* We acknowledge the crucial contribution of the ESO staff for the management of service observations. In particular, we are deeply grateful to M. Hilker for his constant help and support of this program. Italian participation to VIPERS has been funded by INAF through PRIN 2008, 2010, and 2014 programs. L.G., J.B. and B.R.G. acknowledge support from the European Research Council through grant No. 291521. O.L.F. acknowledges support from the European Research Council through grant No. 268107. J.A.P. acknowledges support of the European Research Council through grant No. 67093.



**Fig. 14.** Distribution of  $B$ -band luminosities (*left*) and stellar masses (*right*) computed for all objects with reliable redshift ( $\text{flag} > 2.0$ ) in the PDR-2 catalogue. These are estimated through SED fitting of all available photometry (from UV to  $K$ ) as described in Moutard et al. (2016a). Note that an explicit value of  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$  for the Hubble constant is used here.



**Fig. 15.** Distribution of galaxies at  $0.45 < z < 1.1$  as shown by the final VIPERS PDR-2 catalogue in the W1 and W4 fields (top and bottom cone diagram, respectively). Galaxy positions are projected along declination (1.5 degrees for W4, about 1.8 degrees for W1). Each galaxy is represented by a filled circle of size proportional to its  $B$ -band luminosity and coloured according to its rest-frame  $U - V$  colour.

M.P. acknowledges the financial support from Labex OCEVU (ANR-11-LABX-0060). W.J.P. and R.T. acknowledge financial support from the European Research Council through grant No. 202686. A.P., K.M., and J.K. have been supported by the National Science Centre (grants UMO-2012/07/B/ST9/04425 and UMO-2013/09/D/ST9/04030). W.J.P. is also grateful for support from the UK Science and Technology Facilities Council through the grant ST/I001204/1. E.B., F.M. and L.M. acknowledge the support from grants ASI-INAF I/023/12/0 and PRIN MIUR 2010-2011. L.M. also acknowledges financial support from PRIN INAF 2012. Y.M. acknowledges support from CNRS/INSU (Institut National des Sciences de l'Univers). T.M. and S.A. acknowledge financial support from the ANR Spin(e) through the french grant ANR-13-BS05-0005. C.M. is grateful for support from specific project funding of the *Institut Universitaire de France*. S.D.L.T. and C.M. acknowledge the support of the OCEVU Labex (ANR-11-LABX-0060) and the A\*MIDEX project (ANR-11-IDEX-0001-02) funded by the ‘‘Investissements d’Avenir’’ French government program managed

by the ANR and the Programme National Galaxies et Cosmologie (PNCG). Research conducted within the scope of the HECOLS International Associated Laboratory, supported in part by the Polish NCN grant DEC-2013/08/M/ST9/00664.

## References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, **182**, 543
- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, *ApJS*, **203**, 21
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, **219**, 12
- Bottini, D., Garilli, B., Maccagni, D., et al. 2005, *PASP*, **117**, 996
- Colless, M., Peterson, B. A., Jackson, C., et al. 2003, ArXiv e-print, [astro-ph/0306581]

- Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al. 2013, *AJ*, **145**, 10
- de la Torre, S., Guzzo, L., Peacock, J. A., et al. 2013, *A&A*, **557**, A54
- de la Torre, S., Jullo, E., Giocoli, C., et al. 2017, *A&A*, **608**, A44
- Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *AJ*, **142**, 72
- Emerson, J. P., Sutherland, W. J., McPherson, A. M., et al. 2004, *The Messenger*, **117**, 27
- Garilli, B., Le Fèvre, O., Guzzo, L., et al. 2008, *A&A*, **486**, 683
- Garilli, B., Fumana, M., Franzetti, P., et al. 2010, *PASP*, **122**, 827
- Garilli, B., Paioro, L., Scodreggio, M., et al. 2012, *PASP*, **124**, 1232
- Garilli, B., Guzzo, L., Scodreggio, M., et al. 2014, *A&A*, **562**, A23
- Guzzo, L., Scodreggio, M., Garilli, B., et al. 2014, *A&A*, **566**, A108
- Hammersley, P., Christensen, L., Dekker, H., et al. 2010, *The Messenger*, **142**, 8
- Hildebrandt, H., Erben, T., Kuijken, K., et al. 2012, *MNRAS*, **421**, 2355
- Hudelot, P., Cuillandre, J.-C., Withington, K., et al. 2012, *VizieR Online Data Catalog*, 2317
- Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al. 2013, *MNRAS*, **428**, 1281
- Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, *A&A*, **439**, 845
- Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013a, *A&A*, **559**, A14
- Le Fèvre, O., Saisse, M., Mancini, D., et al. 2003b, in *Proc. SPIE*, eds. M. Iye, & A. F. M. Moorwood, 4841, 1670
- Lilly, S. J., Le Brun, V., Maier, C., et al. 2009, *ApJS*, **184**, 218
- Marchetti, A., Garilli, B., Granett, B. R., et al. 2017, *A&A*, **600**, A54
- Martin, D. C., Fanson, J., Schiminovich, D., et al. 2005, *ApJ*, **619**, L1
- Moutard, T., Arnouts, S., Ilbert, O., et al. 2016a, *A&A*, **590**, A103
- Moutard, T., Arnouts, S., Ilbert, O., et al. 2016b, *A&A*, **590**, A102
- Pezzotta, A., de la Torre, S., Bel, J., et al. 2017, *A&A*, **604**, A33
- Puget, P., Stadler, E., Doyon, R., et al. 2004, in *Ground-based Instrumentation for Astronomy*, eds. A. F. M. Moorwood, & M. Iye, *Proc. SPIE*, **5492**, 978
- Rota, S., Granett, B. R., Bel, J., et al. 2017, *A&A*, **601**, A144
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, **120**, 1579
- <sup>7</sup> Laboratoire Lagrange, UMR7293, Université de Nice Sophia Antipolis, CNRS, Observatoire de la Côte d'Azur, 06300 Nice, France
- <sup>8</sup> Department of Astronomy, University of Geneva, ch. d'Ecogia 16, 1290 Versoix, Switzerland
- <sup>9</sup> Dipartimento di Fisica e Astronomia – Alma Mater Studiorum Università di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
- <sup>10</sup> Institute of Physics, Jan Kochanowski University, ul. Swietokrzyska 15, 25-406 Kielce, Poland
- <sup>11</sup> National Centre for Nuclear Research, ul. Hoza 69, 00-681 Warszawa, Poland
- <sup>12</sup> INFN, Sezione di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy
- <sup>13</sup> IRAP (Institut de Recherche en Astrophysique et Planétologie), Université de Toulouse, CNRS, UPS, Toulouse, France
- <sup>14</sup> Aix-Marseille Université – Pharo – 58 bd Charles Livon Jardin du Pharo – 13007 Marseille, France
- <sup>15</sup> Astronomical Observatory of the Jagiellonian University, Orla 171, 30-001 Cracow, Poland
- <sup>16</sup> School of Physics and Astronomy, University of St Andrews, St Andrews KY16 9SS, UK
- <sup>17</sup> INAF – Istituto di Astrofisica Spaziale e Fisica Cosmica Bologna, via Gobetti 101, 40129 Bologna, Italy
- <sup>18</sup> INAF – Istituto di Radioastronomia, via Gobetti 101, 40129 Bologna, Italy
- <sup>19</sup> Aix Marseille Univ., Univ. Toulon, CNRS, CPT, Marseille, France
- <sup>20</sup> Dipartimento di Matematica e Fisica, Università degli Studi Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy
- <sup>21</sup> INFN, Sezione di Roma Tre, via della Vasca Navale 84, 00146 Roma, Italy
- <sup>22</sup> INAF – Osservatorio Astronomico di Roma, via Frascati 33, 00040 Monte Porzio Catone (RM), Italy
- <sup>23</sup> INAF – Osservatorio Astronomico di Trieste, via G. B. Tiepolo 11, 34143 Trieste, Italy
- <sup>24</sup> Institute d'Astrophysique de Paris, UMR 7095 CNRS, Université Pierre et Marie Curie, 98 bis Boulevard Arago, 75014 Paris, France
- <sup>25</sup> Department of Astronomy & Physics, Saint Mary's University, 923 Robie Street, Halifax, Nova Scotia, B3H 3C3, Canada
- <sup>26</sup> Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK
- <sup>27</sup> Institute of Cosmology and Gravitation, Dennis Sciama Building, University of Portsmouth, Burnaby Road, Portsmouth, PO13FX, UK
- <sup>28</sup> Institut Universitaire de France, 75231 Paris, France
- 
- <sup>1</sup> INAF – Istituto di Astrofisica Spaziale e Fisica Cosmica Milano, via Bassini 15, 20133 Milano, Italy  
e-mail: marcos@lambrate.inaf.it
- <sup>2</sup> INAF–Osservatorio Astronomico di Brera, via Brera 28, 20122 Milano – via E. Bianchi 46, 23807 Merate, Italy
- <sup>3</sup> Università degli Studi di Milano, via G. Celoria 16, 20133 Milano, Italy
- <sup>4</sup> INAF–Osservatorio Astronomico di Bologna, via Gobetti 93/3, 40129 Bologna, Italy
- <sup>5</sup> Aix Marseille Univ., CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France
- <sup>6</sup> INAF–Osservatorio Astrofisico di Torino, 10025 Pino Torinese, Italy