

# Primordial environment of super massive black holes: large-scale galaxy overdensities around $z \sim 6$ quasars with LBT

L. Morselli<sup>1</sup>, M. Mignoli<sup>2</sup>, R. Gilli<sup>2</sup>, C. Vignali<sup>3,2</sup>, A. Comastri<sup>2</sup>, E. Sani<sup>4</sup>, N. Cappelluti<sup>2</sup>, G. Zamorani<sup>2</sup>, M. Brusa<sup>3,2</sup>,  
S. Gallozzi<sup>5</sup>, and E. Vanzella<sup>2</sup>

<sup>1</sup> Excellence Cluster Universe, Technische Universität München, Boltzmannstr. 2, 85748, Garching, Germany  
e-mail: laura.morselli@tum.de

<sup>2</sup> INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy

<sup>3</sup> Dipartimento di Fisica e Astronomia, Università degli Studi di Bologna, viale Berti Pichat 6/2, 40127 Bologna, Italy

<sup>4</sup> INAF – Osservatorio Astrofisico di Arcetri, largo E. Fermi 5, 50125 Firenze, Italy

<sup>5</sup> INAF – Osservatorio Astronomico di Roma, via di Frascati 33, 00040 Monteporzio, Italy

Received 20 March 2014 / Accepted 16 June 2014

## ABSTRACT

**Context.** In the current model of structure formation, bright quasars (QSOs) at  $z \sim 6$  are assumed to be hosted by the most massive dark matter halos that collapsed at that time. The large-scale structures in which these halos are embedded may extend up to 10 physical Mpc, and probably can be traced by overdensities of star-forming galaxies. To date, the search for these overdensities has been limited to scales of 1–2 physical Mpc around the QSO and did not produce coherent results.

**Aims.** We aim at studying the environment of  $z \sim 6$  QSOs and verify whether they are associated with large-scale overdensities of galaxies selected at the same redshift as  $i$ -band dropouts.

**Methods.** With the wide-field ( $\sim 23' \times 25'$ ) Large Binocular Camera (LBC) at the Large Binocular Telescope (LBT), we obtained deep  $r$ -,  $i$ - and  $z$ -band imaging of the fields around four high- $z$  QSOs, namely SDSS J1030+0524 ( $z = 6.28$ ), SDSS J1148+5251 ( $z = 6.41$ ), SDSS J1048+4637 ( $z = 6.20$ ), and SDSS J1411+1217 ( $z = 5.95$ ). Our photometric catalogs are based on source detection in the  $z$ -band image ( $5\sigma$ ) and contain from  $\sim 2.3 \times 10^4$  to  $\sim 2.9 \times 10^4$  objects, down to a 50% completeness limit of  $z = 25.0$ – $25.2$  AB mag. We adopted color–color selections within the  $i - z$  vs.  $r - z$  plane to identify samples of  $i$ -band dropouts at the QSO redshift and measured their relative abundance and spatial distribution in the four LBC fields, each covering  $\sim 8 \times 8$  physical Mpc at  $z \sim 6$ . The same selection criteria were then applied to  $z$ -band-selected sources in the  $\sim 1$  deg<sup>2</sup> wide-and-deep Subaru-*XMM Newton* Deep Survey (SXDS) to derive the expected number of dropouts over a blank LBC-sized field ( $\sim 0.14$  deg<sup>2</sup> after removing masked regions).

**Results.** The four observed QSO fields host more candidates than expected in a blank field. By defining objects with  $z_{AB} < 25$  and  $i - z > 1.4$  that are undetected in the  $r$ -band as  $i$ -band dropouts, we found 16, 10, 9, and 12 dropouts in SDSS J1030+0524, SDSS J1148+5251, SDSS J1048+4637, and SDSS J1411+1217, respectively, whereas only 4.3 such objects are expected over a 0.14 deg<sup>2</sup> blank field. This corresponds to overdensity significances of 3.3, 1.9, 1.7, and 2.5 $\sigma$ , respectively, after accounting for cosmic variance and for the contamination by bluer objects in our dropout samples produced by photometric errors. By considering the total number of dropouts in the four LBC fields and comparing it with what is expected in four blank fields of 0.14 deg<sup>2</sup> each, we find that high- $z$  QSOs reside in overdense environments at the 3.7 $\sigma$  level. This is the first direct and unambiguous measurement of the large-scale structures around  $z \sim 6$  QSOs.

**Key words.** quasars: supermassive black holes – large-scale structure of Universe – galaxies: photometry – galaxies: high-redshift

## 1. Introduction

Wide-area optical surveys like the Sloan Digital Sky Survey (SDSS; Fan et al. 2001a) and the Canada France High- $z$  Quasar Survey (Willott et al. 2009, 2010) have discovered a few tens of  $5.7 < z < 6.5$  quasars (Fan et al. 2001a,b,c, 2003) powered by accreting supermassive black holes (SMBHs) with  $M_{\text{BH}} > 10^{8-9} M_{\odot}$  (Kurk et al. 2007; De Rosa et al. 2011). Wide-area infrared surveys, such as the United Kingdom Infrared Telescope (UKIRT) Infrared Deep Sky Survey (UKIDSS; see Lawrence et al. 2007), and the public ESO-VISTA surveys (Arnaboldi et al. 2013) are now moving to  $z > 6.5$  the redshift frontier for quasars (QSOs) discovery. In particular, three QSOs at  $6.5 < z < 7.0$  have just been discovered by the VISTA Kilo-degree Infrared Galaxy survey (VIKING; see Venemans et al. 2013). The most distant QSO known, at  $z = 7.08$ , was discovered two years ago by the UKIDSS Large Area Survey (Mortlock et al. 2011). De Rosa et al. (2013) studied the properties of these QSOs at

$z > 6.5$  and found that they host  $\sim 10^9 M_{\odot}$  BHs. To explain the existence of these massive objects in place only 1 Gyr after the Big Bang, structure formation models assume that SMBHs are hosted in the most massive dark matter halos that collapsed at that time (Volonteri & Rees 2006). A BH seed, born at  $z \sim 20$  as the consequence of the direct collapse of a gas cloud or as end-product of the first generation of stars, may experience major-merging events that trigger gas accretion and can become a  $10^9 M_{\odot}$  SMBH in a relatively short time, shorter than 1 Gyr. If this were true, the fields around high-redshift QSOs are expected to show galaxy overdensities, which probably embed the progenitors of the most massive clusters of the Local Universe, with  $M > 10^{14-15} M_{\odot}$  (Springel et al. 2005). This agrees with the average mass of the  $z = 0$  descendants of the halos that host  $z \sim 6$  QSOs, as found by recent simulations (albeit with significant scatter of  $\sigma = 0.36$  dex; see Angulo et al. 2012).

In the past decade, large efforts have been made to understand the properties of the  $z \sim 6$  QSO fields, which are extremely

rare ( $\sim 1$  per comoving Gpc<sup>3</sup>) and are assumed to be hosted by dark matter halos of  $M_h = 10^{13} M_\odot$  (Fan et al. 2004). Stiavelli et al. (2005, hereafter S05) found an excess of candidate Lyman-break galaxies (LBGs) around the QSO SDSS J1030+0524 at  $z = 6.28$  using ACS observations that cover a field of  $3' \times 3'$  around the central QSO ( $\sim 1 \times 1$  physical Mpc). Willott et al. (2005) studied the field around three  $z \sim 6$  QSOs using the GMOS-North imaging spectrograph, which has a field of view (FoV) of  $5' \times 5'$ , corresponding to  $\sim 1.8 \times 1.8$  physical Mpc at  $z \sim 6$ . They found no evidence for galaxy overdensity in all the three fields. Kim et al. (2009, hereafter K09) observed the fields around five  $z \sim 6$  QSOs with ACS and found two of them to be overdense, two underdense, and a normally dense one. Recently, Bañados et al. (2013) observed the field around the QSO ULAS J0203+0012 at  $z = 5.72$  with the FORS2 spectrograph, which has a FoV of  $6.8' \times 6.8'$  ( $\sim 2.5 \times 2.5$  physical Mpc). They found that the number of Lyman-alpha emitters (LAEs) is consistent with what is expected in blank fields, and tried to interpret the lack of an overdensity with different scenarios, also including the possibility that the strong ionizing radiation from the QSO may prevent star formation in its vicinity, or that  $z \sim 6$  QSOs are not hosted in the most massive dark matter halos. The observational scenario is unclear even at lower redshifts: observations of QSOs at  $2 < z < 5$  revealed galaxy overdensities around some of them (Djorgovski et al. 2003; Kashikawa et al. 2007; Cantalupo et al. 2012; Swinbank et al. 2012; Husband et al. 2013) and average densities around some others (Francis & Bland-Hawthorn 2004). In addition, even the richest environments around them are as rich as other observed structures that do not host a SMBH (Husband et al. 2013).

Recent simulations by Fanidakis et al. (2013) have suggested that halos hosting luminous  $z \sim 6$  QSOs have masses around  $10^{11} - 10^{12} M_\odot$ , that is, at least a factor of 10 smaller than that suggested by Volonteri & Rees (2006). As a consequence, they would evolve in structures that are not the most massive ones in the local Universe, and are probably not associated with galaxy overdensities at  $z \sim 6$ . However, the apparent lack of strong galaxy overdensities around high- $z$  QSOs does not preclude that they are hosted in the peaks of the dark matter distribution. This is shown by the simulation of Overzier et al. (2009), which was based on the Millennium Run (Springel 2005). They showed that galaxy overdensities can extend on scales of 10 physical Mpc or more from the central BHs, that is, on scales larger than those observed using ACS or GMOS. This result sheds new light on the study of the  $z \sim 6$  QSOs environment: to probe the typical density of the fields in which these extreme structures form, wide-area and deep observations are needed. In light of this, Utsumi et al. (2010) carried out a study around the QSO CFHQS J2329-0301 at  $z = 6.43$  using the wide FoV of the Supreme Cam ( $34' \times 27'$ ) at the Subaru Telescope, and found a  $\sim 3\sigma$  overdensity of galaxies around it on scales of  $\sim 5-6$  physical Mpc.

In this work, we study the fields around four of the most distant QSOs known to date using wide-area, deep images in the  $r$ -,  $i$ - and  $z$ - bands, obtained with the Large Binocular Cameras (LBC; Giallongo et al. 2008) at the Large Binocular Telescope (LBT<sup>1</sup>) on Mount Graham in Arizona. LBC is characterized by a unique *etendue*, the product between collecting area and field of

**Table 1.** Celestial coordinates, redshift, BH mass, and  $z_{AB}$  magnitude of the four quasars (from De Rosa et al. 2011).

Field	RA (J2000)	Dec (J2000)	$z$	$M_{BH}$ [ $10^9 M_\odot$ ]	$z_{SDSS}$ [AB mag]
J1030	10:30:27.1	+05:24:55	6.28	1.4	20.0
J1148	11:48:16.1	+52:51:50	6.41	4.9	20.1
J1048	10:48:45.5	+46:37:18	6.20	3.9	19.9
J1411	14:11:11.3	+12:17:37	5.95	1.0	19.6

view, of  $111 \text{ m}^2 \times 0.16 \text{ deg}^2$  (considering both telescopes). This allows studying  $z \sim 6$  overdensities on scales of  $8 \times 8$  physical Mpc around the QSOs ( $\sim 56 \times 56$  comoving Mpc), that is, on areas at least a factor of 20 larger than those explored by previous studies with ACS or GMOS.

The outline of the paper is as follows: in Sect. 2 the observation, data reduction, and the criteria for selecting  $z \sim 6$  galaxies are described. In Sect. 3 the candidates in each field are shown. In Sect. 4 we discuss our results for the four observed fields, while in Sect. 5 we present our conclusions.

We adopt a  $\Lambda$ CDM cosmology, with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$  and  $\Omega_\Lambda = 0.7$ .

## 2. Observations and data analysis

### 2.1. Observations

We investigated the fields around SDSS J1030+0524 ( $z = 6.28$ , Fan et al. 2001b), SDSS J1148+5251 ( $z = 6.41$ , Fan et al. 2003), SDSS J1048+4637 ( $z = 6.20$ , Fan et al. 2003) and SDSS J1411+1217 ( $z = 5.95$ , Fan et al. 2004). The four fields (hereafter J1030, J1148, J1048 and J1411) have been observed with the  $z_{SDSS}$ ,  $i_{SDSS}$  and  $r_{SDSS}$  filters of LBC. Celestial coordinates, redshifts, BH masses (De Rosa et al. 2011) and  $z_{AB}$  magnitudes of the four  $z \sim 6$  QSOs are shown in Table 1. These QSOs have been chosen because they host some of the most massive black holes observed at  $z \sim 6$  ( $1.0-4.9 \times 10^9 M_\odot$ , see Table 1), and for observability reasons.

The FoV of LBC is  $23.6 \times 25.3 \text{ arcmin}^2$ , which corresponds to  $\sim 8.1 \times 8.7$  physical Mpc at  $z \sim 6$ . Each field is observed using three hours of exposure, 1.5 h in  $z_{SDSS}$  and 1.5 h in  $i_{SDSS}$  on the LBC-red channel, and simultaneously 3 h in the  $r_{SDSS}$  filter on the LBC-blue channel. Thus we reach AB sensitivity limits that are efficient in the color selection of  $z \sim 6$  sources (see Sect. 2.3).

Standard data reduction, such as flat fielding, sky subtraction, image alignment, and stacking was performed with the LBC pipeline. In addition, a mask with an exposure map was used to mask low exposure time regions that were the result of the dithering technique. A detailed description of the reduction procedure can be found in Giallongo et al. (2008). Finally, visual inspection was needed to mask spikes caused by saturation, bright star haloes (due to reflections within the cameras), and artifacts.

The seeing FWHM for each band image and each QSO field is shown in Table 2. The seeing values were computed from the brightness profile of reliable point-like objects. Table 2 also shows the effective area of each field, after masking low S/N and contaminated regions, and the total number of detected objects per field.

University; and The Research Corporation, on behalf of The University of Notre Dame, University of Minnesota and University of Virginia.

<sup>1</sup> The LBT is an international collaboration among institutions in the United States, Italy, and Germany. LBT Corporation partners are: The University of Arizona on behalf of the Arizona university system; Istituto Nazionale di Astrofisica, Italy; LBT Beteiligungsgesellschaft, Germany, representing the Max-Planck Society, the Astrophysical Institute Potsdam, and Heidelberg University; The Ohio State

**Table 2.** Seeing (FWHM) in the  $z$ ,  $i$  and  $r$  filters, effective area of the image, number of detected sources, 50% completeness limit in the  $z$ -band image, and number of primary ( $(i - z) - \sigma_{(i-z)} > 1.3$ ) and secondary ( $1.1 < (i - z) - \sigma_{(i-z)} < 1.3$ )  $z \sim 6$  LBG candidates in the four fields.

Field	Filter	Seeing $FWHM$ [arcsec]	Clean area [deg <sup>2</sup> ]	Total objects $z_{TOT}$	Completeness limit	Primary candidates	Secondary candidates
J1030	$z$	$0.64 \pm 0.02$	0.143	29151	25.2	14	10
	$i$	$0.70 \pm 0.02$					
	$r$	$0.74 \pm 0.02$					
J1148	$z$	$0.79 \pm 0.02$	0.144	24447	25.0	8	3 <sup>a</sup>
	$i$	$0.87 \pm 0.02$					
	$r$	$0.93 \pm 0.02$					
J1048	$z$	$0.84 \pm 0.01$	0.143	24087	25.0	6	9
	$i$	$0.87 \pm 0.02$					
	$r$	$1.04 \pm 0.02$					
J1411	$z$	$0.73 \pm 0.01$	0.145	26055	25.0	11	8
	$i$	$0.77 \pm 0.01$					
	$r$	$0.80 \pm 0.01$					

**Notes.** <sup>(a)</sup> In the J1148 field an object that just barely misses our selection criteria for secondary  $z \sim 6$  LBG candidates (i.e. it is undetected in the  $r$ -band, has  $z_{tot} = 23.4$ , but has  $(i - z) - \sigma_{(i-z)} = 1.06$ ), was proven to be a  $z = 5.7$  AGN (Mahabal et al. 2005).

## 2.2. Object detection

Object detection was carried out using the software SExtractor 2.3.2 (Bertin & Arnouts 1996). The  $z$ -band images were used as *master* images on which object detection was performed. Several tests were carried out to obtain the set of input parameters that correspond to the maximum number of real objects while keeping suspected false detections as rare as possible. In particular, we considered an object detected if it consists of 9 connected pixels, each exceeding the local sky background by of a factor of 1.5 (for a total  $4.5\sigma$  detection). We then used SExtractor in dual mode to obtain the photometric measurements on the  $i$  and  $r$  images. We chose an aperture diameter of 7 pixels ( $\sim 1.6$  arcsec) to compute the aperture magnitudes to collect a large fraction of the flux from the object while minimizing the contamination from neighboring sources. The four created catalogs (one per field) typically have  $\sim 2.5 \times 10^4$  detected objects each. The field with the most sources is J1030, which was observed in the  $z$ -band with the best seeing.

The photometric zero points took into account the correction for dust extinction computed from Schlegel et al. (1998). We used the SExtractor MAG\_AUTO magnitudes<sup>2</sup> to estimate the total magnitudes (e.g.,  $z_{TOT}$ ), because of their low uncertainties and small systematic errors at faint magnitudes. Finally, we included an aperture-correction term to optimize the estimate of the aperture magnitude for pointlike objects since we needed to compensate for seeing variations between different fields/filters.

The completeness limit in total magnitude of the images, defined as the value of the magnitude at which the number of detected objects decreases to 50% of the expected value, is 25.2 for the  $z$  image of the J1030 field (the best seeing image, see Table 2), while it is 25.0 for the other three  $z$  images. The  $z$ -band-selected catalogs with associated  $r$ - and  $i$ -band photometry, as well as the reduced  $r$ ,  $i$ ,  $z$ -band images of the four fields are publicly available<sup>3</sup>.

The selection criteria of  $z \sim 6$  source candidates made use of colors computed from aperture magnitudes corrected for seeing variations, therefore we also needed to estimate the completeness limits of the aperture magnitude in the  $i$  and  $r$  bands.

To compute these limits, we adopted a deeper  $3\sigma$  object detection, performed independently in the  $i$  and  $r$  images of the four fields. The  $i$  and  $r$  detected catalogs show very similar photometric properties, therefore, we adopted the  $i$  and  $r$  magnitude aperture limits of 26.6 and 27.2 for the four different fields to achieve uniformity in the selection criteria.

To confirm the reliability of our photometric measurements, we downloaded catalogs of SDSS sources in the same area of the four LBC fields from the SDSS archive and then cross-correlated them with the LBC catalogs, using a correlation radius of 0.8 arcsec (this value was chosen to minimize the number of contaminants). In particular, for point-like sources we compared the SDSS *psfMag* with the total magnitude of LBC catalogs. The mean difference  $\langle \Delta_{mag} \rangle$  between the two magnitudes in the three filters is small<sup>4</sup>, no offset is seen between the four LBC fields, and the *rms* of the  $\Delta_{mag}$  distribution in the three filters is lower than 0.1, confirming the reliability and stability of the photometric measurements of this work. Fig. 1 shows the number counts per deg<sup>2</sup> of sources in  $z_{TOT}$  in the four fields. The number counts in the four fields agree excellently well.

## 2.3. Candidate selection

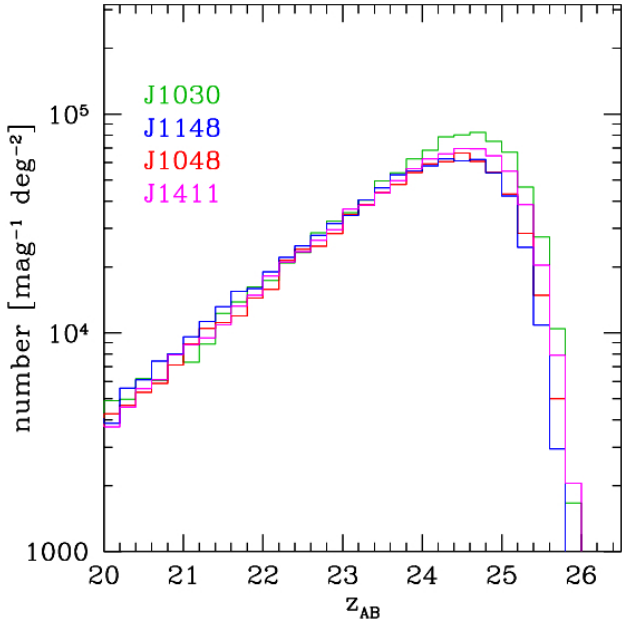
At high redshift, galaxy spectra are characterized by a flux drop blueward of the rest-frame wavelength of the Ly $\alpha$  line ( $\lambda_{Ly\alpha,rest} = 1216 \text{ \AA}$ ) as a consequence of the increasing amount of neutral hydrogen in the intergalactic medium. This drop is the so-called Gunn-Peterson trough (Gunn & Peterson 1965) and is routinely used to identify galaxies at high redshift. At  $z \gtrsim 5.9$ , the observed wavelength of the Ly $\alpha$  line is at  $\gtrsim 8400 \text{ \AA}$ , that is, the line moves in the  $z_{SDSS}$  filter. As a consequence, the flux in the  $i_{SDSS}$  filter is suppressed and the  $z \sim 6$  galaxies can be identified as  $i$ -band dropouts. In particular, Dickinson et al. (2004), Beckwith et al. (2006) and other authors have shown that, by using different spectral templates of mostly star-forming galaxies, objects at  $z > 5.6$  can be efficiently selected among those with  $i - z > 1.3$ .

The  $z \sim 6$  source candidates identified using the  $i - z$  color can be affected by contamination from red objects at lower redshift. In particular, elliptical galaxies are characterized by a large

<sup>2</sup> MAG\_AUTO is computed using an adaptive elliptical aperture around every detected object, following Kron (1980).

<sup>3</sup> <http://www.oabo.inaf.it/~LBTz6>

<sup>4</sup> In particular, the  $\langle \Delta_{mag} \rangle = \text{psfMag}(\text{SDSS}) - \text{MagTot}(\text{LBC})$  in the three filters is:  $-0.10$  in  $z$ ,  $+0.07$  in  $i$  and  $-0.17$  in  $r$ . This implies an  $i - z = 0.17$  color shift between the two photometric systems.



**Fig. 1.** Source number counts in the  $z$ -band filter in the four LBC fields. The green, blue, red, and magenta histograms refer to J1030, J1148, J1048, and J1411, respectively.

Balmer break ( $D_{4000}$ ,  $\lambda_{\text{rest}} = 4000 \text{ \AA}$ ) in their spectra and, at  $z \gtrsim 1.1$ , the  $D_{4000}$  falls in the  $z$  filter, causing the source to be detected as  $i$ -dropout. To reduce the contamination from elliptical galaxies it is possible to use the  $r - z$  color (similarly to the ACS/HST  $V - z$  color criterium adopted by Beckwith et al. 2006 and Bouwens et al. 2006). In fact, the flux of elliptical galaxies blueward of the  $D_{4000}$  does not drop to zero as the expected flux of  $z \sim 6$  galaxies beyond the redshifted Lyman limit. For this reason,  $z \sim 6$  galaxies have redder  $r - z$  color than elliptical galaxies at lower redshift. Moreover, since the profile of elliptical galaxies at  $z \sim 1-2$  is expected to be resolved in our sharp LBC images, we can use the morphology of the selected sources to estimate the contamination from elliptical interlopers in the sample of high-redshift galaxies (see Sect. 4.1).

Cool dwarfs like L and T stars with temperatures in the range 700–2000 K have intrinsic colors similar to those of  $z \sim 6$  galaxies. It is difficult to quantify the degree of contamination by late-type dwarfs in our fields since there is no accurate knowledge of late-type star counts at different sky positions at the faint magnitudes we probe. We limit ourselves to note that our LBC fields are all at high Galactic latitude ( $b > 50$ ), and three of them point far away from the Galactic center ( $240 < l < 140$ ). The only field that points toward the Galactic center is J1411, at  $l = 359$ , which may then suffer from some stellar contamination more than the other fields. We note, however, that the precise knowledge of the stellar contamination is unlikely to bias our results significantly. Indeed, the Subaru *XMM-Newton* Deep Survey that is used in Sect. 4.3 as a reference field to quantify galaxy overdensities, has Galactic coordinates comparable with those of at least three of our four fields, and therefore probably shares a similar level of stellar contamination among  $i$ -band dropouts (see Sect. 4.3).

Colors were estimated using the corrected aperture magnitudes to select the flux always from the same portion of the source. Finally, to obtain a more reliable sample of  $z \sim 6$  candidates, we replaced a sharp color criterion with one that includes the photometric color errors. The criteria implemented in this work to select our primary candidates are then as follows:

- $z_{\text{TOT}} < z_{\text{lim,TOT}}$ ;

- $(i - z) - \sigma_{(i-z)} > 1.3$ ;
- non-detection in the  $r$ -band at  $3\sigma$  ( $r_{\text{AP}} > 27.2$ );

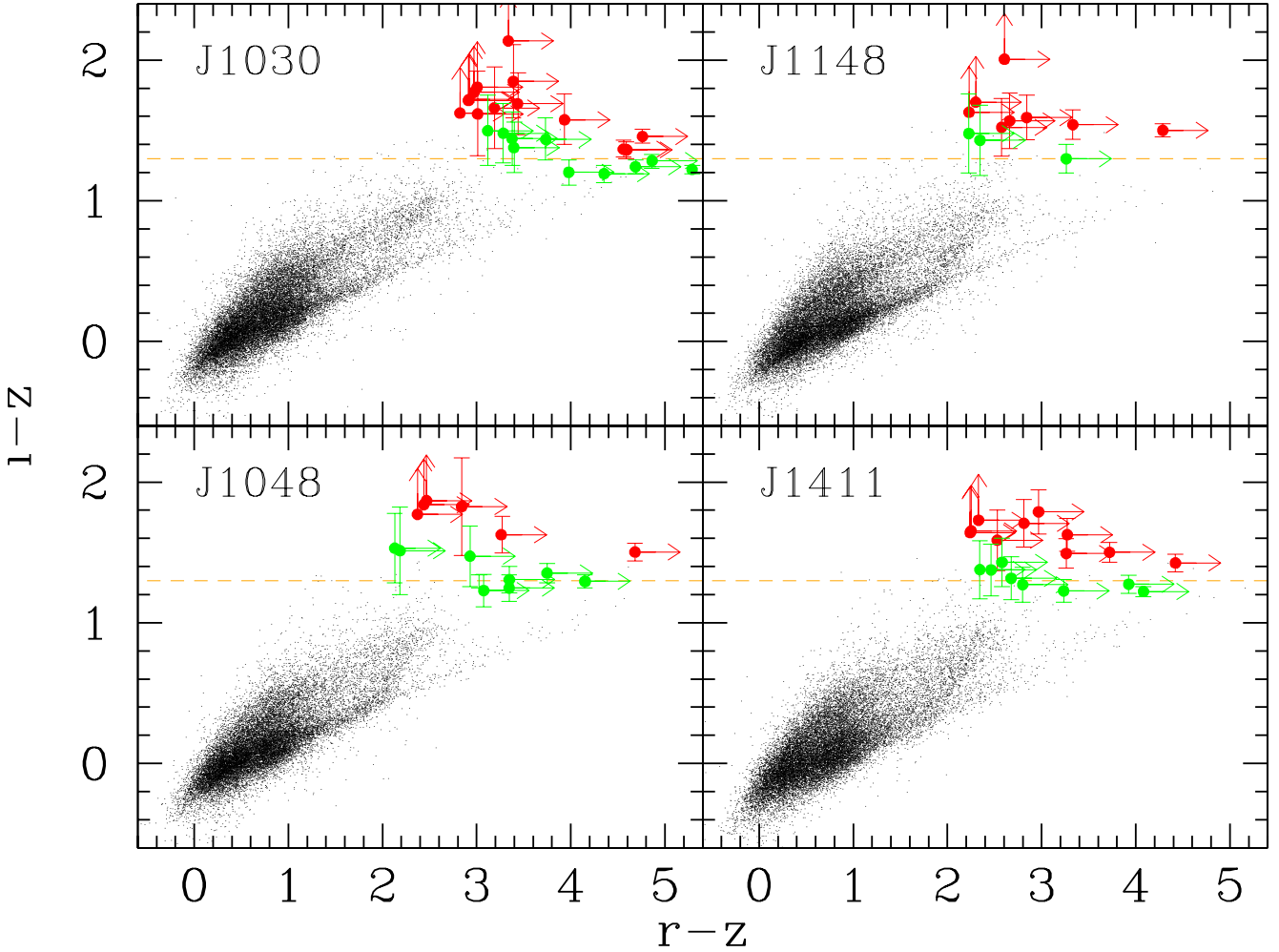
where  $z_{\text{TOT}}$  is the total magnitude in the  $z$  image,  $z_{\text{lim,TOT}}$  is the completeness limit in  $z$ , the  $i - z$  color and its error  $\sigma_{(i-z)}$  are computed with corrected aperture magnitudes, and  $r_{\text{AP}}$  is the corrected aperture magnitude in the  $r$  image.

The spectra of high- $z$  QSOs show a strong Ly $\alpha$  emission line. At  $z \lesssim 5.8$ , the shifted Ly $\alpha$  line falls in the  $i_{\text{SDSS}}$  filter. If the line is strong, it causes an enhancement of the flux in the  $i$  band and, as a consequence, the  $i - z$  color could be smaller than the adopted threshold of 1.3 (see, e.g., Vanzella et al. 2010). For this reason, in addition to a primary selection based on  $(i - z) - \sigma_{(i-z)} > 1.3$ , we also investigated objects with  $1.1 < (i - z) - \sigma_{(i-z)} < 1.3$ . Indeed, it is not uncommon that a significant fraction of objects selected by more relaxed color criteria is made by genuine high- $z$  galaxies (Toshikawa et al. 2012). Candidates with  $1.1 < (i - z) - \sigma_{(i-z)} < 1.3$  form our secondary samples and will constitute the additional targets in our future spectroscopic follow-up campaign. Figure 2 shows the position in the  $i - z$  versus  $r - z$  diagram of the primary and secondary candidates in the four fields.

We note that the adopted selection criteria cover a rather broad redshift range and might in principle select objects that do not belong to the same structure of the QSO. However, the network of LSSs around overdense regions at  $z \sim 6$  can extend up to 10 physical Mpc in radius (Overzier et al. 2009). This corresponds to  $\Delta z \approx 0.2$  at  $z = 6$ , and we therefore preferred to avoid more stringent selection criteria that would span only a narrow redshift window around each of the four QSOs.

### 3. Results

The number of primary and secondary candidates in the four fields are shown in the last two columns of Table 2, while the catalog ID,  $z_{\text{TOT}}$ , morphological class,  $i - z$  and  $r - z$  colors of the primary candidates are reported in Table 3. The morphological class was estimated by analysis of various parameters: objects with  $z_{\text{TOT}} \lesssim 24$  were classified as unresolved (point-like;  $p$ ) if CLASS\_STAR (from SExtractor) is greater than 0.5 and the difference between total and corrected aperture magnitude in the  $z$ -band is lower than 0.1, or as resolved otherwise ( $r$ ). Fainter objects are unclassified ( $u$ ). The spatial distribution of the primary and secondary candidates around the QSOs is shown in Fig. 3, where the  $z \sim 6$  candidate sources found in the previous work of K09 are also indicated. The work of K09 was carried out using ACS observations, and the  $z$ -band magnitudes of their candidates are in the range [25.0:26.5], which is fainter than our completeness limit. We note that  $z_{\text{TOT}} < 25$  corresponds to UV absolute magnitudes of  $M_{1450} < -21.7$  at  $z = 6$ , meaning that we sampled the bright end of the (star-forming) galaxy luminosity function at that redshift ( $\sim 0.5$  mag or more brighter than  $M_{\text{UV}}^*$  at  $z = 6$ , Bouwens et al. 2014). The four fields are characterized by a non-uniform distribution of the candidates around the central QSO. This is an important indication that our primary samples are probably not dominated by star contaminants. In fact, stars are expected to be randomly distributed in the fields, whereas the distributions of  $i$ -band dropouts around early QSOs are expected to be all but symmetric (see e.g. Fig. 19 of Overzier et al. 2009). In J1030, the western region is clearly more populated than the eastern one: 11 out of 14 primary dropouts have RA values lower than the central QSO. Based on the binomial distribution, the probability of observing 11 or more out of 14 objects within half of the field is 2.8% (1.1% when adding secondary



**Fig. 2.**  $i - z$  versus  $r - z$  diagram for the sources in the four QSO fields (small black dots). Primary and secondary  $z \sim 6$  galaxy candidates are marked with red and green dots, respectively. The dashed line represents the adopted threshold of 1.3 in  $(i - z) - \sigma_{i-z}$  used to select primary candidates. A primary candidate in the J1411 field (with  $i - z > 2.82$ ) falls outside the plot. The small black dots above the dashed lines do not match all of our selection criteria (e.g., most of them are detected in the  $r$ -band).

candidates). This peculiarity in J1030 was also found by K09 and S05 over a much smaller area around the QSO. In J1148 the southern region is clearly more populated, with 7 out of 8 primary candidates found at lower declinations than the QSO (corresponding to a probability of 3.5% and 0.26% if the region populated by the 7 southern dropouts is considered to be half or one third of the LBC field, respectively). The distribution is less clear in J1048, while in J1411 the majority of the primary candidates is found in the north western region (the probability of having 7 out of 11 objects within one quadrant is 0.76%), while the secondary sample mostly populates the south eastern area.

### 3.1. J1030

The field around J1030 hosts 14  $z \sim 6$  primary galaxy candidates. Four of them are relatively bright ( $z_{\text{TOT}} < 24.2$ ) and have a reliable morphological classification; only one of these four sources is classified as resolved. The morphological classification of the remaining 10 fainter sources is not reliable, but a visual inspection of the images and a stacking analysis (see Sect. 4.1) confirm that all of them are very compact in the sharp  $z$  image of the field, supporting the idea that they are good high- $z$  candidates. Figure 4 shows three candidates in the J1030 field in the  $z$ -,  $i$ - and  $r$ - band images: an unresolved, point-like candidate

(ID26728), a resolved one (ID3200), and an object with unclassified morphology (ID25971). The secondary sample of this field consists of 10 candidates.

### 3.2. J1148

The primary sample of candidates in the J1148 field consists of 8 sources. The three of them with  $z_{\text{TOT}} < 24$  have a reliable morphological classification; only one source is classified as an unresolved object in our image. This field hosts three secondary candidates. Mahabal et al. (2005) reported the discovery of an optically faint QSO at  $z = 5.70$  at 1.82 arcmin from the QSO SDSS J1148+5251. In our catalog, this QSO is undetected in the  $r$ -band, has  $z_{\text{TOT}} = 23.4$ , and  $i - z = 1.11 \pm 0.05$ , and is then just below the selection criteria adopted for our secondary candidates. This underlines the importance of studying objects with bluer  $i - z$  color (part of our secondary sample) in addition to sources with  $i - z > 1.3$  (our primary sample).

### 3.3. J1048

The field around J1048 is the less populated. It contains only six primary candidates. Three of them have  $z_{\text{TOT}} < 24$ ; two have a

**Table 3.** Object ID,  $z_{\text{TOT}}$  magnitudes, morphological classification (p: unresolved objects, r: resolved objects, u: undefined),  $i - z$  and  $r - z$  colors of the primary candidates in the four fields.

ID	$z_{\text{TOT}}$	Morph	$(i - z)_{\text{AP}}$	$(r - z)_{\text{AP}}$
J1030				
3200	23.20 ± 0.03	r	1.46 ± 0.05	>4.76
18619	23.51 ± 0.03	p	1.36 ± 0.06	>4.59
26728	23.55 ± 0.03	p	1.37 ± 0.06	>4.56
25831	24.16 ± 0.09	p	1.85 ± 0.26	>3.39
12265	24.22 ± 0.11	u	1.58 ± 0.18	>3.93
6024	24.56 ± 0.13	u	1.69 ± 0.22	>3.44
2140	24.75 ± 0.17	u	1.62 ± 0.30	>3.01
21596	24.77 ± 0.11	u	>2.14	>3.34
28941	24.87 ± 0.16	u	>1.62	>2.82
24071	24.88 ± 0.16	u	1.66 ± 0.29	>3.19
25971	24.92 ± 0.17	u	>1.77	>2.97
11963	25.00 ± 0.17	u	>1.71	>2.91
23354	25.07 ± 0.14	u	>1.72	>2.92
21438	25.16 ± 0.12	u	>1.81	>3.01
J1148				
5113	22.99 ± 0.03	p	1.50 ± 0.05	>4.29
3556	23.48 ± 0.07	r	1.59 ± 0.16	>2.84
3605	23.84 ± 0.06	r	1.54 ± 0.11	>3.33
5393	24.04 ± 0.11	u	1.52 ± 0.20	>2.58
3910	24.23 ± 0.12	u	1.57 ± 0.20	>2.66
18368	24.68 ± 0.12	u	>2.00	>2.61
6761	24.85 ± 0.15	u	>1.70	>2.30
1556	24.91 ± 0.16	u	>1.62	>2.23
J1048				
3002	22.65 ± 0.02	p	1.50 ± 0.03	>4.68
9077	23.87 ± 0.06	r	1.63 ± 0.13	>3.26
4811	23.93 ± 0.14	r	1.83 ± 0.35	>2.84
22521	24.43 ± 0.12	u	>1.87	>2.47
2968	24.57 ± 0.15	u	>1.84	>2.44
22931	24.93 ± 0.16	u	>1.77	>2.37
J1411				
25349	22.89 ± 0.02	p	1.42 ± 0.03	>4.42
25638	23.48 ± 0.04	p	1.50 ± 0.07	>3.72
564	23.75 ± 0.05	r	>2.82	>3.42
19681	24.03 ± 0.06	u	1.49 ± 0.10	>3.26
16818	24.04 ± 0.07	u	1.63 ± 0.12	>3.27
17304	24.29 ± 0.07	u	1.79 ± 0.16	>2.97
6136	24.35 ± 0.11	u	1.71 ± 0.17	>2.81
22009	24.49 ± 0.16	u	>1.73	>2.33
24189	24.84 ± 0.09	u	1.59 ± 0.21	>2.53
18095	24.90 ± 0.17	u	>1.65	>2.25
13344	24.99 ± 0.12	u	>1.64	>2.24

**Notes.** Upper limits on  $i - z$  and  $r - z$  are indicated.

resolved but compact morphology, and one is unresolved. The secondary sample consists of nine sources.

### 3.4. J1411

The J1411 field is the second-most populated field. It contains 11  $z \sim 6$  LBG primary candidates, in particular, three bright candidates with  $z_{\text{TOT}} < 24$  (two unresolved and one resolved). The secondary sample of  $z \sim 6$  galaxies around J1411 consists of eight sources.

The  $r, i, z$  cutouts of all the primary and secondary dropout samples in the four fields can be consulted at the project webpage<sup>5</sup>.

<sup>5</sup> <http://www.oabo.inaf.it/~LBTz6>

## 4. Discussion

### 4.1. Morphology and near-IR colors of primary candidates

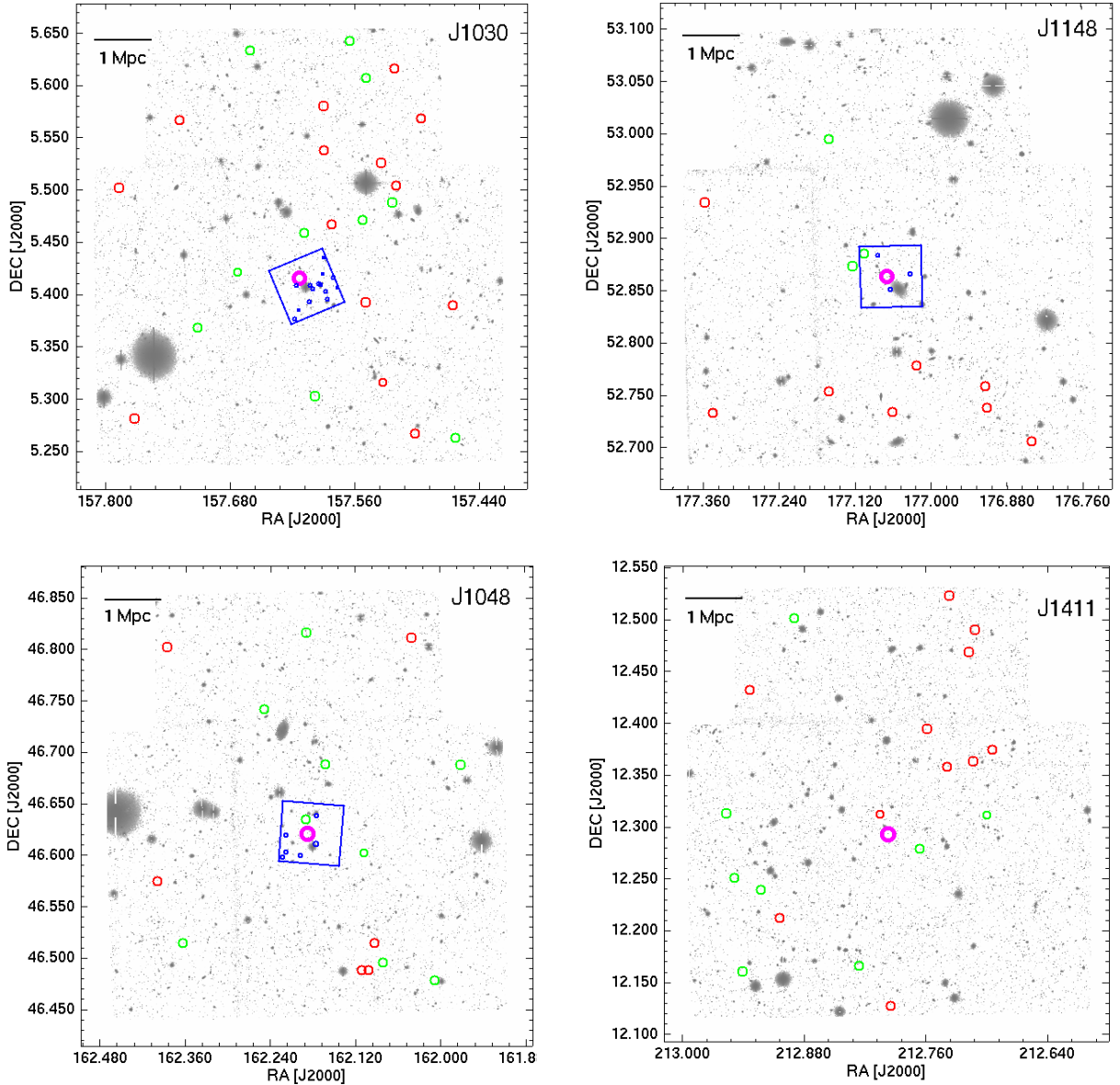
As explained in Sect. 2.3, the  $z \sim 6$  candidates samples obtained using a pure color selection may be affected by contamination from elliptical galaxies at lower redshift. Since the flux of elliptical galaxies blueward of the  $D_{4000}$  does not drop to zero, elliptical galaxies at  $z \sim 1-2$  are characterized by a bluer  $r - z$  color than  $z \sim 6$  sources. In addition, the extended structure of ellipticals at  $z \sim 1-2$  is very likely resolved in the sharp  $z$ -band images used in this work. This is evident from the comparison between the stacking of the 11 primary  $z \sim 6$  candidates in the J1030 field with  $z_{\text{TOT}} = 24-25$ , for which SExtractor cannot give a reliable morphological classification, and that of 12 sources in the same range of magnitude, which, although they are red ( $i - z > 1.3$ ), do not satisfy our stringent selection constraints because of their larger photometric errors or because they are barely detected in the  $r$ -filter. We performed this test in the J1030 field since it features the best seeing and contains most of the  $i$ -band dropouts (see Table 3). Figure 5 shows the comparison between the profile of the two stacks. Clearly, our candidates are unresolved and compact, with a stacked FWHM of the profile of 0.6 arcsec, comparable with the profiles of 55 stars in the field, while  $i$ -band dropouts with bluer  $r - z$  color are resolved, with a FWHM of 0.85 arcsec. The stacked images of the primary candidates are shown in insets of Fig. 5. The  $z$ -band stacking (left panel) displays a high S/N, unresolved source that it is not detected in the  $r$ -band stacked image (right panel) with a limiting magnitude of  $\approx 29$ , estimated from the background rms. The unresolved profile in  $z$ , the non-detection in the  $r$ -band stacked image, and the resulting average color ( $\langle r - z \rangle \gtrsim 4$ ) are all strong clues of the goodness of the adopted color criteria for selecting  $z \sim 6$  sources.

In addition to the morphological analysis, near-IR photometry can be used in principle to test the reliability of LBG candidates. For instance,  $z \sim 6$  LBGs are known to be  $\sim 1.5-2$  mag bluer in  $z - J$  than brown dwarfs with similar  $i - z$  color (Willott et al. 2013). The only deep IR observations that cover a sizeable portion of our fields, however, are those by *Spitzer*/IRAC at  $\lambda \geq 3.6 \mu\text{m}$ . In all fields but J1030, the IRAC coverage is  $\sim 60 \text{ arcmin}^2$  per field, which is  $\sim 12\%$  of the LBC FoV, and on average only two dropouts (mostly secondary) per QSO field have been observed. A large IRAC mosaic is instead available for the J1030 field, covering about half of the LBC FoV, for which we retrieved the mosaic images and downloaded the photometry at 3.6 and 4.5  $\mu\text{m}$  from the *Spitzer* Heritage Archive<sup>6</sup> for the nine dropouts in the mosaics. Although the depth of the IRAC images is sufficient to detect most of the dropouts, color-color diagnostics using optical- to-IRAC photometry are much weaker than those using  $J$ -band data (Willott et al. 2013). Our dropouts indeed appear to have colors that can be produced by both late type brown dwarfs and  $z \sim 6$  LBGs. A detailed SED analysis of our candidates is beyond the scope of this paper and will be presented elsewhere.

### 4.2. Overdensity estimate

We computed the high- $z$  galaxy overdensity in each LBC field as  $\delta = \frac{\rho}{\bar{\rho}} - 1$ , where  $\rho$  is the number of  $i$ -band dropouts observed in each field and  $\bar{\rho}$  is the number of dropouts expected over a blank sky field of area equal to the LBC FoV ( $0.144 \text{ deg}^2$  after correcting for masked regions). To obtain a good estimate

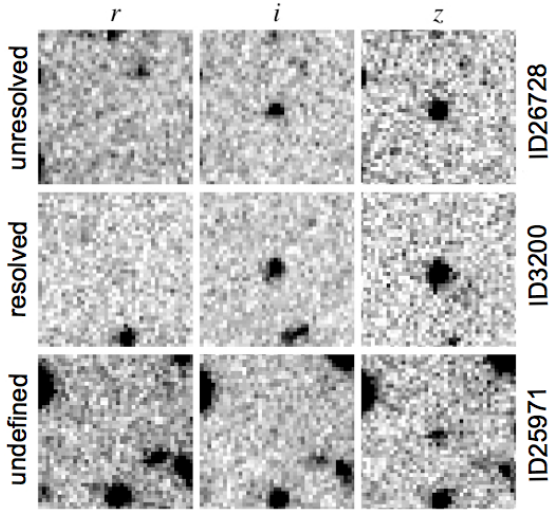
<sup>6</sup> <http://sha.ipac.caltech.edu/applications/Spitzer/SHA>



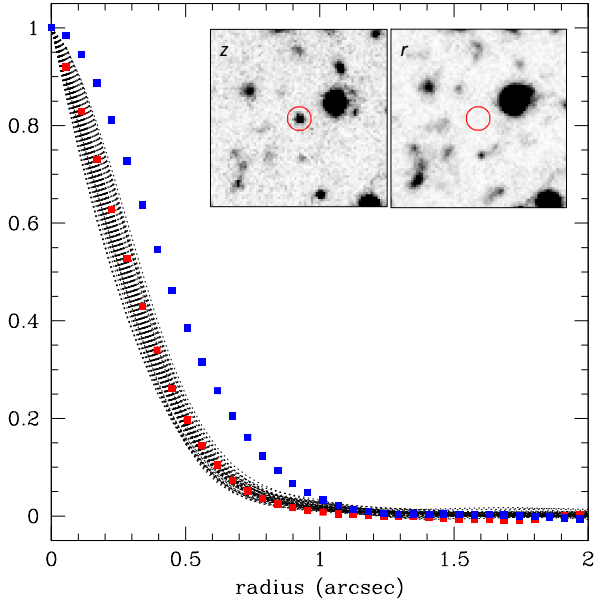
**Fig. 3.** Spatial distribution of  $z \sim 6$  galaxy candidates around the four QSOs (magenta thick circles) overlaid on the LBC  $z$ -band images. Red and green circles indicate our primary and secondary candidates. The small blue dots mark the faint  $i$ -band dropouts ( $25 < z_{850} < 26.5$ ) found in the previous study of K09 carried out with HST/ACS observations. The  $3 \times 3$  arcmin<sup>2</sup> ACS FoV is shown by the blue squares.

of  $\bar{\rho}$ , a blank sky field as wide as possible is needed, which should also feature observations in  $r, i, z$  down to a depth similar to, or greater than, that of our LBC observations. To our knowledge, the best available field that satisfies these criteria is the Subaru/*XMM-Newton* Deep Survey (SXDS; Furusawa et al. 2008), a mosaic of five partially overlapping Suprime-Cam pointings, covering  $\sim 1$  deg<sup>2</sup> in total, down to  $B = 28.4$ ,  $V = 27.8$ ,  $R_c = 27.7$ ,  $i' = 27.7$  and  $z' = 26.6$  ( $5\sigma$  detection limit). The SXDS can be considered a good representation of a typical blank sky field. Its optical source number counts agree well with those of other surveys in similar bands and with similar depth (Furusawa et al. 2008). The AGN content of the SXDS is also similar to that of other wide-and-deep X-ray fields. The X-ray number counts obtained from moderately deep XMM observations (e.g.,  $f_{0.5-2} > 6 \times 10^{-16}$  erg cm<sup>-2</sup> s<sup>-1</sup>) agree well with those from other X-ray surveys and the measured clustering level of X-ray sources (Ueda et al. 2008). In addition, the SXDS is not biased to any known overdense regions at  $z \sim 6$ . Based on spectroscopically confirmed samples of LAEs, Ouchi et al. (2008)

concluded that the  $\sim 1$  deg<sup>2</sup> SXDS has no signature of overdensity or underdensity at  $z \sim 5.7$ , and that the numbers of LAEs in each of its five subfields are consistent with the average field-to-field variations. McLure et al. (2006) identified nine robust LBG candidates at  $z > 5$  in the SXDS, all with  $z_{AB} < 25$ . One of them was spectroscopically identified by Willott et al. (2009) as the faintest known QSO at  $z \sim 6$  (CFHQS J021627-045534, with  $z = 6.01$ ,  $z_{AB} = 24.4$ ,  $M_{1450} = -22.21$ ). We note, however, that it is not clear whether CFHQS J021627-045534 is an example of a bona fide high- $z$  QSO. Its spectrum is indeed very different from that commonly observed for  $z \sim 6$  QSOs, since it features a single prominent Ly $\alpha$  line without any continuum emission. Its spectrum thus resembles those of bright LAEs, and its absolute magnitude is also consistent with those observed at the bright end of the  $z \sim 6$  galaxy luminosity function (Willott et al. 2013; Bouwens et al. 2014). Furthermore, the relatively large FWHM measured by Willott et al. (2009) ( $\sim 1600$  km s<sup>-1</sup>) may be ascribed to gas outflows produced by star formation. At any rate, even assuming that CFHQS J021627-045534 is powered



**Fig. 4.** Postage stamps (9 arcsec on a side) in the  $z$ ,  $i$  and  $r$  filters of three  $z \sim 6$  LBG candidates in the J1030 field. An unresolved object (*top row*), a resolved one (*center row*) and one with unclassified morphology (*bottom row*) are shown.

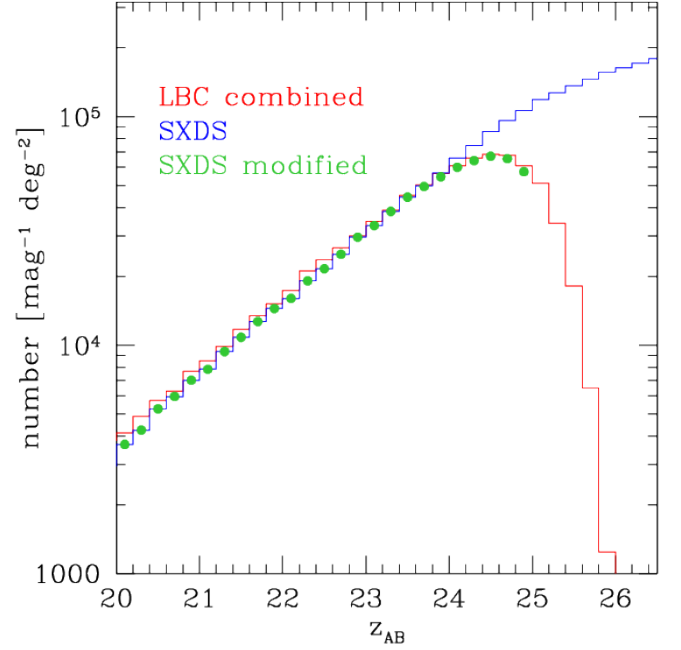


**Fig. 5.** Stacking profiles of the 11  $z \sim 6$  primary candidates in the J1030 field with unclassified morphology (red symbols) and of 12 red ( $i - z > 1.3$ ) objects that do not satisfy our stringent selection criteria (blue symbols). The dotted lines indicate the profiles of 55 stars selected in the same field to represent the mean seeing (FWHM) of the  $z$ -band image. The insets show the  $z$  (left) and  $r$  (right) stacked images of the 11 primary candidates; the size of each panel is  $15 \times 15 \text{ arcsec}^2$ .

by an accreting BH, its mass would be smaller by about two dex than that of the bright QSOs analyzed in this work (assuming it is accreting at the Eddington limit, like most  $z \sim 6$  QSOs do). Therefore, it would most likely trace more typical, less dense environments at  $z \sim 6$ .

Five  $z$ -band selected catalogs, one for each pointing of the Subaru Suprime-Cam, are publicly available for the SXDS<sup>7</sup>. We retrieved and merged them to create a single master catalog, removing multiple measurements of the same object in the overlapping regions. The effective area of the resulting SXDS master

<sup>7</sup> [http://soaps.nao.ac.jp/SXDS/Public/DR1/index\\_dr1.html](http://soaps.nao.ac.jp/SXDS/Public/DR1/index_dr1.html)



**Fig. 6.** Source number counts in the  $z$ -band for the LBC combined catalog (red histogram), the SXDS catalog (blue histogram) and the SXDS modified catalog (green points). The SXDS modified catalog, cut here at  $z_{\text{TOT}} < 25$ , represents the average of the 10 000 simulated SXDS catalogs downgraded to the completeness level of the LBC fields (see text for details).

catalog is  $\sim 1.13 \text{ deg}^2$  (see also Table 3 of Furusawa et al. 2008). To compare our LBC catalogs and the SXDS, a correction factor that takes into account the offset between the different filters was computed by comparing the photometric measurements of SXDS with those of the SDSS in the same area of the sky<sup>8</sup>.

Our catalogs and the SXDS have different completeness limits and consequently different source densities. The LBC catalogs start missing objects at  $z_{\text{TOT}} \approx 23.8$ , whereas the SXDS is still 100% complete. Most LBG candidates are in the range  $24 < z_{\text{TOT}} < 25$ . This means that the overdensity would be underestimated if a correction for this effect were not added. To solve this problem, 10 000 SXDS catalogs were simulated, differentially excluding objects from the SXDS catalog in the range  $24.0 < z_{\text{TOT}} < 25.0$  in such a way that the simulated catalogs reproduce the same trend of the number of counts with  $z$ -magnitude as that of the LBC observations (Fig. 6). It is clear that the average of the SXDS modified catalogs has the same drop of counts of the LBC fields near the completeness limit (cut at  $z_{\text{TOT}} = 25.0$ ). We note the good agreement between the number counts in our LBC combined catalog and those in the SXDS at  $z_{\text{TOT}} < 23.8$ ; a similarly good match between our LBC fields and the SXDS is also found when comparing the source counts in the  $r$  and  $i$  band, as well as the  $i - z$  vs.  $r - z$  color-color diagrams. This makes us confident that the corrections for the different photometric systems were made properly, and therefore the SXDS can be used as a good comparison field to compute galaxy overdensities.

We then verified whether the dropout selection criteria presented in Sect. 2.3 can be safely applied to select  $z \sim 6$  galaxy candidates in the SXDS. The errors on the magnitudes of the SXDS galaxies as reported by Furusawa et al. (2008) appear to be one order of magnitude smaller than what we find for

<sup>8</sup> The average values of  $\text{psfMag}(\text{SDSS}) - \text{MagTot}(\text{SXDS})$  are  $-0.05$ ,  $-0.03$ , and  $0.11$  in  $z, i, r$ , respectively. This leads to  $\text{MagTot}(\text{LBC}) - \text{MagTot}(\text{SXDS}) = 0.05, -0.1, 0.28$  in  $z, i, r$ , respectively.



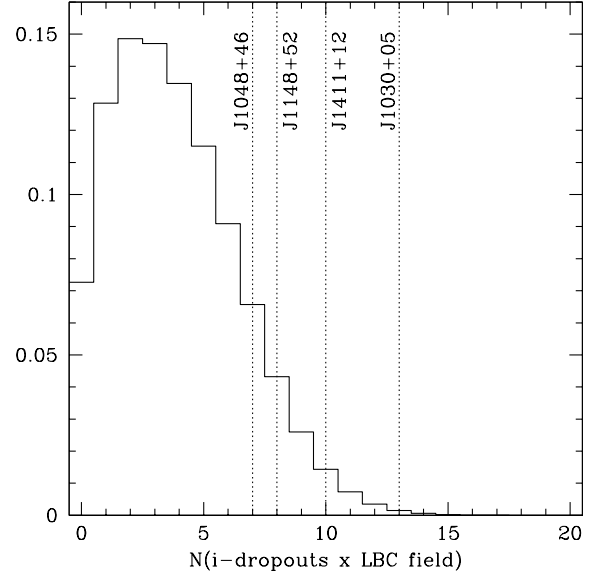
**Table 4.** Summary of measured overdensities.

Field (1)	$\rho$ (2)	$\rho_d$ (3)	$\delta$ (4)	$\sigma_\delta$ (5)
J1030	16	13	2.0	3.3
J1148	10	8	0.9	1.9
J1048	9	7	0.6	1.7
J1411	12	10	1.3	2.5

**Notes.** (1) QSO field. (2) Number of observed dropouts defined as objects with  $z_{\text{TOT}} < 25$ ,  $r_{\text{AP}} > 27.2$ ,  $i - z > 1.4$ . (3) Same as Col. 2, but corrected for the contamination from bluer objects caused by photometric errors (see text for details): these are the numbers adopted to compute the overdensities and their significance, shown in Cols. 4 and 5, respectively. (4) Overdensity  $\delta = \frac{\rho}{\bar{\rho}} - 1$ , where  $\bar{\rho} = 4.3$ , as estimated from the SXDS. (5) Significance of the measured overdensities (expressed in Gaussian-equivalent probabilities; see text for details).

LBC sources of comparable magnitudes. We independently estimated the magnitude errors of SXDS galaxies by considering objects that have been observed twice by Suprime-Cam and for which two independent magnitude estimates are available. We computed the magnitude differences ( $\Delta\text{mag}$ ) for these objects in five magnitude bins: the rms of the  $\Delta\text{mag}$  distributions is significantly larger than the errors quoted in the SXDS catalog at the same magnitudes. Therefore, since we do not have full control of the relative errors between our LBC fields and the SXDS, we preferred to exclude any dropout selection criteria that involve errors on magnitudes. We instead adopted a simple (and more conservative) color cut at  $i - z > 1.4$  to define the  $i$ -band dropout samples to be used for the overdensity estimate. Table 4 shows the number of candidates in the LBC fields that satisfy these criteria. However, because the photometric errors in our LBC catalogs are larger than those in the SXDS, it is likely that our dropout samples suffer from more contamination from objects with  $i - z < 1.4$ , which are more abundant than those with  $i - z > 1.4$ , and thus can artificially enhance the observed number of the latter if photometric errors are significant. We tried to quantify this contamination as follows: we assumed that the  $i - z$  distribution of SXDS objects with  $z_{\text{TOT}} < 25$  and  $r_{\text{AP}} > 27.2$  is the true, underlying  $i - z$  distribution of our LBC dropout candidates (note that this is conservative, since it assumes that the SXDS catalog has negligible photometric errors). Then we randomly extracted an  $i - z$  color from this input distribution, assumed a Gaussian error with  $\sigma = 0.2$  on it (i.e., the average error on the  $i - z$  color measured for our LBC dropout samples; see Table 1), and extracted a new  $i - z$  color from the Gaussian distribution. We repeated this process 10 000 times and built a new  $i - z$  distribution that is indeed broader than the input one because of photometric errors. We compared the input and the “blurred”  $i - z$  distributions and verified that the fraction of objects with  $i - z > 1.4$  has increased by 20% from the input to the blurred one. Therefore, we estimate that 20% of our observed dropouts with  $i - z > 1.4$  are bluer objects that met the selection because of photometric errors, and also quote in Table 4 the number of “statistically decontaminated” dropouts in the four LBC fields.

We then estimated the number of dropouts that is expected in a  $0.144 \text{ deg}^2$  area based on the SXDS catalog. We applied the same selection criteria as above ( $z_{\text{TOT}} < 25$ ,  $r_{\text{AP}} > 27.2$ ,  $i - z > 1.4$ ) to the SXDS, considering only objects in the  $0.95 \text{ deg}^2$  of the SXDS that are free from either strong or weak halos caused by bright stars, and whose photometry is therefore expected to be more reliable. These candidates were visually inspected to remove clear artifacts and spurious sources, exactly



**Fig. 7.** Number of sources measured in 10 000 Poisson realisations based on the completeness-corrected number of dropouts detected in  $23 \times 25 \text{ arcmin}^2$  rectangles randomly placed within the SXDS field. The vertical lines represent the number of decontaminated dropouts with  $i - z > 1.4$  in the four LBC fields (see Table 4 and text for details).

in the same way as we cleaned our LBC catalogs. We finally obtained a clean sample of 41 SXDS dropouts over  $0.95 \text{ deg}^2$ . Based on the 10 000 SXDS modified catalogs, this number has to be reduced by  $\sim 30\%$  to account for the incompleteness of the LBC fields, which translates into 4.3 sources expected per LBC FoV. Based on the number of decontaminated LBC dropouts presented in Table 4, we then measure an overdensity of 2.0, 0.9, 0.6, 1.3 in J1030, J1148, J1048, J1411, respectively. The total overdensity obtained by summing the four fields is 1.2.

To correctly estimate the statistical significance of the overdensity measured in the four LBC fields, the effects of cosmic variance have to be taken into account. Once again we exploited the SXDS to evaluate it, randomly placing area rectangles of  $23 \times 25 \text{ arcmin}^2$  within the SXDS, that is, the LBC FoV, counting the number of dropouts in each rectangle and correcting for incompleteness. An average of four (completeness-corrected) dropouts per rectangle is measured, and no rectangle contains more than eight dropouts. For each rectangle, we then built a Poissonian distribution with the mean equal to the completeness-corrected number of dropouts, and summed all the Poissonian distributions to build the total distribution presented in Fig. 7, which then accounts for both statistical uncertainties and cosmic variance (albeit limited to what can be extracted from the SXDS field). The probability of observing a number of dropouts in this distribution equal to or larger than what is observed in J1030, J1148, J1048, and J1411 is  $1.0 \times 10^{-3}$ ,  $5.4 \times 10^{-2}$ ,  $9.7 \times 10^{-2}$ , and  $1.3 \times 10^{-2}$ , respectively. If the distribution were Gaussian, these probabilities would correspond to  $3.3$ ,  $1.9$ ,  $1.7$ , and  $2.5\sigma$ , respectively. The two fields J1030 and J1411 are the densest, with an overdensity significance of  $> 2\sigma$ . J1148 and J1048 are also denser than average, but with lower significance.

As discussed at the end of Sect. 2.3, the adopted  $i - z$  color criteria cover a broad redshift range that may in principle select objects that are not associated with the QSO LSS. We then performed a test on the J1030 field, where the dropout statistics

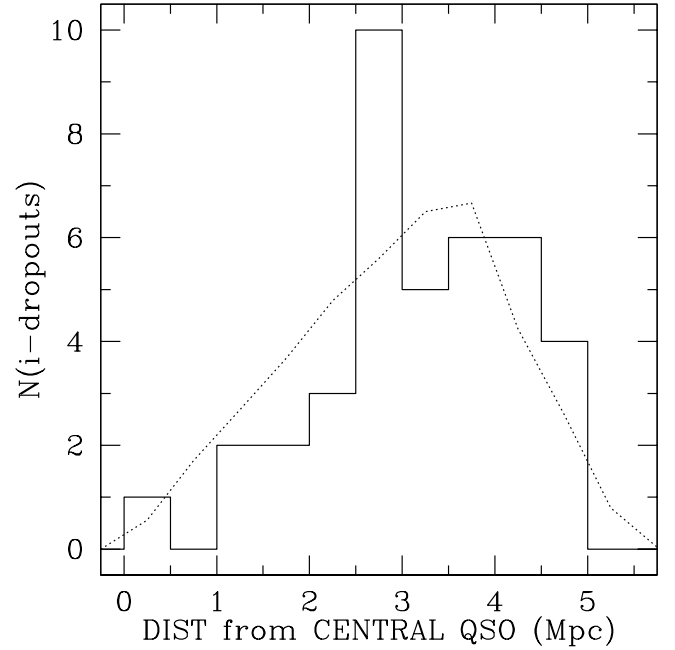
is the highest, by selecting  $i$ -band dropouts with  $i - z > 1.8$ , that are expected to be at  $z > 5.9$  (Beckwith et al. 2006), that is, within  $\sim 20$  physical Mpc from the QSO ( $z = 6.28$ ). By repeating the same analysis as described above, we found that the observed overdensity in J1030 is still significant at the  $\sim 2.5\sigma$  level.

We finally considered the probability of finding a number of dropouts equal to or greater than what is observed in the four LBC fields all together by randomly extracting four of the LBC-like rectangles at a time. We found that  $z \sim 6$  QSOs are surrounded by galaxy overdensities at the  $3.7\sigma$  level.

#### 4.3. Spatial distribution and comparison with previous results

In three of our LBC fields (J1030, J1148, J1048) the density of faint ( $27 < z_{850} < 25$ )  $i$ -band dropouts was previously investigated through observations with HST/ACS on scales a factor of 8 smaller than those sampled in this work (S05, K09). In J1030, both S05 and K09 observed an overdensity with respect to the reference fields (GOODS South and North). The significance of the overdensity obtained in our work, simply based on the dropout counts alone, is higher than measured by K09 and S05 ( $2.8\sigma$ ) on the same basis (S05 and K09 report a higher significance when the dropout color distribution is also taken into account). K09 also found that J1148 and J1048 have a dropout density lower than and similar to that of the field, respectively, whereas we find that, at large scales, both fields are denser than average (albeit with low significance). Although it is difficult to directly compare the dropout densities observed by K09 and those reported in this paper, since they sample different magnitude regimes and have also been selected with slightly different color criteria, it is worth considering some aspects: in J1030, the most overdense field in both K09 and our sample, the distribution of  $i$ -band dropouts is strongly skewed westward of the QSOs in both ACS and LBC data, which then are fully consistent with each other. In J1148, we found that most of the LBC dropouts are distributed south of the QSO, at separations from the QSO larger than 5 arcmin. Therefore, if the LBC dropouts trace the most overdense galaxy region in the J1148 field, this would explain why the ACS observations have missed it. In general, previous results based on narrow-field ( $< 5 \times 5$  arcmin<sup>2</sup>) imaging around high-redshift QSOs have produced contradictory results in terms of galaxy overdensity. Together with the results of Utsumi et al. (2010), who measured a  $\sim 3\sigma$  overdensity of  $i$ -band dropouts around a  $z = 6.43$  QSOs using the wide-field Subaru/Suprime-Cam ( $34 \times 27$  arcmin<sup>2</sup> FoV), our results instead suggest that large-scale galaxy overdensities are common among high- $z$  SMBHs.

In Fig. 8 we compare the distribution of the projected distances from the central QSO (in physical Mpc) of our primary candidates with the distribution of the  $\sim 75\,000$  objects with  $22 < z_{\text{TOT}} < 25$  detected in our LBC fields. The  $i$ -band dropout distribution suggests that there is a deficit of high- $z$  galaxies at  $r < 2.5$  Mpc when compared with the random field galaxy distribution. A K-S test shows that the radial distribution of high- $z$  candidates is different from that of the field galaxies at a 98.5% confidence level ( $2.4\sigma$ ). A similar deficit was also reported by Utsumi et al. (2010) on the same scales. This might be an indication that, on scales of 2–3 physical Mpc around the QSOs, which is comparable with their Stromgren spheres (Kurk et al. 2007; Wythe et al. 2005), galaxy formation is suppressed because of the strong QSO negative feedback (either radiative or even mechanical). This interpretation, among other possibilities, was put forward by Romano-Diaz et al. (2011) to explain why the average number of dropouts per ACS field measured by K09



**Fig. 8.** Distribution of the primary candidate projected distances (in physical Mpc) from the central QSO in all the four fields (solid histogram). The dotted curve represents the distribution of the  $\sim 75\,000$  objects with  $22 < z_{\text{TOT}} < 25$  in the LBC photometric catalogs, normalized to the total number of dropouts (39). The drop in the galaxy density at radii larger than  $\sim 4$  Mpc is caused by the physical limit of the LBC FoV.

(4 per field) was a factor of  $\sim 2$  lower than that predicted by their simulations, which placed early BHs in halos with  $M > 10^{12} M_{\odot}$  surrounded by overdense environments. The other obvious interpretation is that  $z \sim 6$  QSOs are not hosted in the most massive collapsed halos (see, e.g., Fanidakis et al. 2013). However, if our results of large-scale overdensities in all the four QSO fields are confirmed by optical spectroscopy and deeper observations, the negative-feedback interpretation would be favored, and it will have been demonstrated that early SMBHs do indeed live in the densest environments at their time.

## 5. Conclusions

We have studied large-scale galaxy overdensities around four bright QSOs hosting four of the most massive BHs at  $z \sim 6$  known to date (namely SDSS J1030+0524, SDSS J1148+5251, SDSS J1048+4637, and SDSS J1411+1217) using deep images in  $r$ ,  $i$  and  $z$  obtained with the  $23 \times 25$  arcmin<sup>2</sup> Large Binocular Camera (LBC) at the Large Binocular Telescope (LBT). In contrast to most previous studies that were sampling relatively small scales, our observations map a region of  $\sim 8 \times 8$  physical Mpc at  $z \sim 6$ , and can therefore trace galaxy overdensities on scales similar to those expected by simulations of early structure formation. We built  $z$ -band selected catalogs down to  $z_{\text{AB}} = 25$ , performed  $i$ - and  $r$ -band photometry, and applied different color-color selection techniques to identify galaxies at approximately the same redshift as that of the four QSOs. Our main results are as follows:

- The four analyzed fields show a significant number of candidate  $z \sim 6$  galaxies selected as  $i$ -band dropouts. Their spatial distribution in the LBC fields is highly asymmetric, as expected from simulations of galaxy formation in high- $z$

dense environments. This argues against a major contamination from late-type stars, which is expected to be more evenly distributed within the fields. A possible deficit of dropouts within 2.5 physical Mpc from the QSOs is found at a  $2.4\sigma$  level. This suggests that feedback from the QSO may hinder galaxy formation within those scales.

- Most of the  $i$ -band dropouts have  $z_{\text{TOT}} > 24$  and no reliable morphological information can be derived for them. However, the stacking of these faint dropouts in the J1030 field, which features the sharpest and deepest observations, shows that they are unresolved and compact, with a FWHM of the median profile comparable with the seeing of the  $z$ -band image. This excludes severe contamination from lower-redshift early-type galaxies and increases the likelihood that we selected genuine high- $z$  objects.
- The number of  $i$ -band dropouts measured in our fields is always larger than expected in a blank sky field. By considering the Subaru/XMM Deep Survey (SXDS) as a reference blank field, and after accounting for the different LBC vs. SXDS depths and photometric systems, we expect only 4.3  $i$ -band dropouts (defined as objects with  $z_{\text{TOT}} < 25$ ,  $r_{\text{AP}} > 27.2$ ,  $i - z > 1.4$ ) over a  $0.14 \text{ deg}^2$  blank field, whereas 16, 10, 9, and 12 dropouts were observed in the J1030, J1148, J1048, and J1411 field, respectively. When accounting for cosmic variance and for the contamination by bluer objects produced by photometric errors, these numbers correspond to overdensity significancies of 3.3, 1.9, 1.7, and  $2.5\sigma$ , respectively. By considering the dropouts in the four LBC fields combined and comparing them with what is expected in four random blank fields of  $0.14 \text{ deg}^2$  each, we find that high- $z$  QSOs reside in overdense environments at the  $3.7\sigma$  level.
- Three out of four QSO fields were previously observed by the  $3 \times 3 \text{ arcmin}^2$  HST/ACS. The ACS observations revealed an  $i$ -band dropout overdensity only in the J1030 field, whereas the J1148 and J1048 fields were found to be underdense and average dense, respectively. The overdensity we measured in each of the four fields suggests that wide-field observations may be more effective in revealing large-scale structures around early QSOs (albeit the significance is below  $2\sigma$  in two fields). This would also be consistent with the idea that, although early SMBHs reside in dense environments, galaxy formation may be depressed in their immediate vicinity because of their strong radiation field.

Follow-up spectroscopy of the selected  $i$ -band dropouts is obviously needed to confirm whether they mostly consists of genuine  $z \sim 6$  galaxies. In addition, extending our LBC imaging analysis to other QSO fields is highly desirable to obtain a statistically sound sample. We started the process of acquiring optical spectroscopy and near-IR imaging on some of our fields to further refine the selection of  $z \sim 6$  galaxies. Other multiband data, including X-rays, might reveal the presence of AGN within the observed structures at  $z \sim 6$  and would place the first observational constraints on the environment of primordial objects, thus providing a fundamental input to models of early BH and galaxy formation.

*Acknowledgements.* We acknowledge the support from the LBT-Italian Coordination Facility for the execution of observations, data distribution and reduction. We acknowledge support from the Italian Space Agency under the ASI-INAF contract I/009/10/0 and from INAF under the contract PRIN-INAF-2012. M.B. acknowledges support from the FP7 Career Integration Grant “eEASy” (“SMBH evolution through cosmic time: from current surveys to eROSITA-Euclid AGN Synergies”, CIG 321913). The referee is acknowledged for insightful comments that improved the quality of the paper.

## References

- Angulo, R., Springel, V., White, S., et al. 2012, MNRAS, 425, 2722
- Arnaboldi, M., Rejkuba, M., Retzlaff, J., et al. 2013, The Messenger, 154, 18
- Bañados, E., Venemans, B., Walter, F., et al. 2013, ApJ, 773, 178
- Beckwith, S., Stiavelli, M., Koekemoer, A., et al. 2006, AJ, 132, 1729
- Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
- Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., & Franx, M. 2006, ApJ, 653, 53
- Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2014, ApJ, submitted [arXiv:1403.4295]
- Cantalupo, S., Lilly, S. J., & Haehnelt, M. G. 2012, MNRAS, 425, 1992
- De Rosa, G., Decarli, R., Walter, F., et al. 2011, ApJ, 739, 56
- De Rosa, G., Venemans, B. P., Decarli, R., et al. 2013, ApJ, submitted [arXiv:1311.3260]
- Dickinson, M., Stern, D., Giavalisco, M., et al. 2004, ApJ, 600, L99
- Djorgovski, S. G., Stern, D., Mahabal, A. A., & Brunner, R. 2003, ApJ, 596, 67
- Fan, X., Strauss, M. A., Richards, G. T., et al. 2001a, AJ, 121, 31
- Fan, X., Strauss, M. A., Schneider, D. P., et al. 2001b, AJ, 121, 54
- Fan, X., Narayanan, V. K., Lupton, R. H., et al. 2001c, AJ, 122, 2833
- Fan, X., Strauss, M. A., Schneider, D. P., et al. 2003, AJ, 125, 1649
- Fan, X., Hennawi, J. F., Richards, G. T., et al. 2004, AJ, 128, 515
- Fanidakis, N., Macció, A. V., Baugh, C. M., Lacey, C. G., & Frenk, C. S. 2013, MNRAS, 436, 315
- Francis, P. J., & Bland-Hawthorn, J. 2004, MNRAS, 353, 301
- Furusawa, H., Kosugi, G., Akiyama, M., et al. 2008, ApJS, 176, 1
- Giallongo, E., Ragazzoni, R., Grazian, A., et al. 2008, A&A, 482, 349
- Gunn, J., & Peterson, B. 1965, ApJ, 142, 1633
- Husband, K., Bremer, M. N., Stanway, E. R., et al. 2013, MNRAS, 432, 2869
- Kashikawa, N., Kitayama, T., Doi, M., et al. 2007, ApJ, 663, 765
- Kim, S., Stiavelli, M., Trenti, M., et al. 2009, ApJ, 695, 809
- Kron, R. G. 1980, ApJS, 43, 305
- Kurk, J. D., Walter, F., Fan, X., et al. 2007, ApJ, 669, 32
- Lawrence, A., Warren, S. J., Almaini, O., et al. 2007, MNRAS, 379, 1599
- Mahabal, A., Stern, D., Bogosavljević, M., Djorgovsky, S. G., & Thompson, D. 2005, ApJ, 634, L9
- McLure, R. J., Cirasuolo, M., Dunlop, J. S., et al. 2006, MNRAS, 372, 357
- Mortlock, D. J., Warren, S. J., Venemans, B. P., et al. 2011, Nature, 474, 616
- Ouchi, M., Shimasaku, K., Akiyama, M., et al. 2008, ApJS, 176, 301
- Overzier, R. A., Guo, Q., Kauffmann, G., et al. 2009, MNRAS, 394, 5770
- Romano-Diaz, E., Shlosman, I., Trenti, M., & Hoffman, Y. 2011, ApJ, 736, 66
- Schlegel, D. J., Finkbeiner, D., & Davis, M. 1998, ApJ, 500, 525
- Springel, V. 2005, MNRAS, 364, 1105
- Springel, V., Withe, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629
- Stiavelli, M., Djorgovski, S. G., Pavlovsky, C., et al. 2005, ApJ, 622, L1
- Swinbank, J., Baker, J., Barr, J., Hook, I., & Bland-Hawthorn, J. 2012, MNRAS, 422, 2980
- Toshikawa, J., Kashikawa, N., Ota, K., et al. 2012, ApJ, 750, 137
- Ueda, Y., Watson, M. G., Stewart, I. M., et al. 2008, ApJS, 179, 124
- Utsumi, Y., Goto, T., Kashikawa, N., et al. 2010, ApJ, 721, 1680
- Vanzella, E., Grazian, A., Hayes, M., et al. 2010, A&A, 513, A20
- Venemans, B. P., Findlay, J. R., Sutherland, W. J., et al. 2013, ApJ, 779, 24
- Volonteri, M., & Rees, M. J. 2006, ApJ, 650, 669
- Willott, C. J., Percival, W. J., McLure, R. J., et al. 2005, ApJ, 626, 657
- Willott, C. J., Delorme, P., Reylé, C., et al. 2009, AJ, 137, 3541
- Willott, C. J., Delorme, P., Reylé, C., et al. 2010, AJ, 139, 906
- Willott, C. J., McLure, R. J., Hiben, P., et al. 2013, AJ, 145, 4
- Wythe, J. S. B., Loeb, A., & Carilli, C. 2005, ApJ, 628, 575