Discovery of a Protocluster Associated with a Lyα Blob Pair at z = 2.3

Toma Bădescu1, Yujin Yang1,2, Frank Bertoldi1, Ann Zabludoff3, Alexander Karim1, and Benjamin Magnelli1

1 Argelander Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany; toma@astro.uni-bonn.de
2 Korea Astronomy and Space Science Institute, 776 Daedeokdae-ro, Yuseong-gu, Daejeon 34055, Korea
3 Steward Observatory, University of Arizona, 933 North Cherry Avenue, Tucson, AZ 85721, USA

Received 2017 May 2; revised 2017 July 10; accepted 2017 July 20; published 2017 August 23

Abstract

Bright Lyα blobs (LABs)—extended nebulae with sizes of ~100 kpc and Lyα luminosities of ~10^{44} erg s^{-1}—often reside in overdensities of compact Lyα emitters (LAEs) that may be galaxy protoclusters. The number density, variance, and internal kinematics of LABs suggest that they themselves trace group-like halos. Here, we test this hierarchical picture, presenting deep, wide-field Lyα narrowband imaging of a 1° × 0.5° region around a LAB pair at z = 2.3 discovered previously by a blind survey. We find 183 Lyα emitters, including the original LAB pair and three new LABs with Lyα luminosities of (0.9–1.3) × 10^{43} erg s^{-1} and isophotal areas of 16–24 arcsec^2. Using the LAEs as tracers and a new kernel density estimation method, we discover a large-scale overdensity (Boötis J1430+3522) with a surface density contrast of 3.7, a volume density contrast of 2.7, and a projected diameter of ≈20 comoving Mpc. Comparing with cosmological simulations, we conclude that this LAE overdensity will evolve into a present-day Coma-like cluster with log(M/M_☉) ~ 15.1 ± 0.2. In this and three other wide-field LAE surveys re-analyzed here, the extents and peak amplitudes of the largest LAE overdensities are similar, not increasing with survey size, and implying that they were indeed the largest structures then and today evolve into rich clusters. Intriguingly, LABs favor the outskirts of the densest LAE concentrations, i.e., intermediate LAE overdensities of δ_LAE = 1–2. We speculate that these LABs mark infalling protogroups being accreted by the more massive protocluster.

Key words: galaxies: clusters: general – galaxies: formation – galaxies: high-redshift – intergalactic medium – large-scale structure of universe

Supporting material: machine-readable tables

1. Introduction

The study of galaxy clusters plays an important role in understanding cosmological structure formation and the astrophysics of galaxy evolution. Statistics of galaxy cluster size, mass, and redshift distribution provide constraints for cosmological models, while the properties of the galaxies and gas inside clusters give clues about galaxy evolution and the cosmological models, while the properties of the galaxies and size, mass, and redshift distribution provide constraints for astrophysics of galaxy evolution. Statistics of galaxy cluster formation at higher redshifts has been challenging. Since protoclusters lack many of the observational properties of the massive virialized galaxy clusters of today, one of the best ways to find them is to identify galaxy overdensities at high redshift (Overzier 2016). Readily observable populations of galaxies include radio galaxies (Venemans et al. 2002, 2007; Hatch et al. 2011a; Hayashi et al. 2012; Wylezalek et al. 2013; Cooke et al. 2014), submillimeter galaxies (Daddi et al. 2009; Capak et al. 2011; Dannerbauer et al. 2014; Rigby et al. 2014), hydrogen alpha emitters (Hatch et al. 2011b; Hayashi et al. 2012), or Lyman break galaxies, and Lyman alpha emitters (LAEs) (e.g., Taniguchi et al. 2005; Overzier et al. 2006, 2008). LAEs, which are compact galaxies that have strong emission in the Lyα line, are relatively easy to observe over a wide range of redshifts at z ~ 2–6 (e.g., Taniguchi et al. 2005; Gronwall et al. 2007; Nilsson et al. 2009; Guitaia et al. 2010). LAEs are mainly star-forming, low-mass objects, and some may be the progenitors of today’s Milky Way-type galaxies (Gawiser et al. 2007). With wide-field, deep narrowband surveys centered on the Lyα line emission at a given redshift, one can use LAEs to identify galaxy overdensities. Giant Lyα-emitting nebulae, also known as Lyα “blobs” (LABs; Francis et al. 1996; Ivison et al. 1998; Steidel et al. 2000; Matsuda et al. 2004; Palunas et al. 2004), which emit Lyα radiation on large scales (50–100 kpc) and have high Lyα luminosities of 10^{43–44} erg s^{-1}, are also apparent tracers of LAE overdensities (e.g., Matsuda et al. 2004, 2005; Saito et al. 2006; Prescott et al. 2008; Yang et al. 2009, 2010).

What powers the strong extended Lyα emission in blobs is still poorly understood. Possible powering mechanisms include gravitational cooling radiation (Haiman et al. 2000; Fardal et al. 2001; Yang et al. 2006; Dijkstra & Loeb 2009; Faucher-Giguère et al. 2010; Goerdt et al. 2010; Rosdahl & Blaizot 2012), the resonant scattering of Lyα photons produced by star formation (Möller & Warren 1998; Laursen & Sommer-Larsen 2007; Hayes et al. 2011; Steidel et al. 2011; Zheng et al. 2011; Cen & Zheng 2013), and photoionizing radiation from active galactic nuclei (AGNs) (Haiman & Rees 2001; Cantalupo et al. 2005; Geach et al. 2009; Kollmeier et al. 2010; Overzier et al. 2013; Yang et al. 2014a). Another potential source is shock-heating from starburst-driven winds (Taniguchi & Shioya 2000; Mori & Umemura 2006), although recent
studies of the emission of non-resonant lines from eight
Lyα blobs exclude models that require fast galactic winds
by AGNs or supernovae (Yang et al. 2011, 2014a, 2014b; Prescott et al. 2015a).

Regardless of the energy sources of Lyα blobs, the
association of blobs with compact LAE overdensities with
sizes of ∼10–20 Mpc (Matsuda et al. 2004, 2011; Palunas et al.
2004; Yang et al. 2010; Prescott et al. 2012, 2008; Saito et al.
2015), suggests that LABs are good potential markers of large
protoclusters. Furthermore, the number density and variance of
Lyα blobs, as well as the 200–400 km s\(^{-1}\) relative velocities of
their embedded galaxies, suggest that blobs themselves occupy
individual group-like halos of ∼10\(^{13}\) M\(_{\odot}\) (Yang et al. 2010,
2011; Prescott et al. 2012, 2015b). Thus, blobs may be sites of
massive galaxy formation and trace significant components of
the build-up of protoclusters. However, because most previous
LAB studies have been carried out toward known overdense regions
or protoclusters, the observed relationship between Lyα blobs and LAE overdensities may be biased. To probe the
LAB-overdensity connection one should investigate the area
around known Lyα blobs that were identified without prior knowledge of their environments. For example, Prescott et al.
(2008) studied the environment of a Lyα blob that was serendipitously discovered by its strong Spitzer MIPS 24 μm
flux (Dey et al. 2005), finding that this Lyα blob resides in an
overdense region of 20 × 50 Mpc\(^2\).

In this work, we investigate the large-scale environment of a
Lyα blob pair at redshift \(z = 2.3\) that was discovered without
prior knowledge of the environment (Yang et al. 2009). The
paper is organized as follows. In Section 2, we present our
observations and data reduction. In Section 3, we discuss our
selection of Lyα emitters and blobs. In Section 4, we describe
the discovery of an overdensity associated with the Lyα blob pair,
compare its properties with those obtained from three
previous narrowband surveys of other LAE structures, discuss
whether it will evolve into a present-day galaxy cluster, and
show that Lyα blobs are preferentially located in the outskirts
of protoclusters here and in the other surveys. In Section 5, we
summarize the results. Throughout the paper, we adopt the
following cosmological parameters: \(H_0 = 70\) km s\(^{-1}\) Mpc\(^{-1}\),
\(\Omega_M = 0.3\), and \(\Omega_{\Lambda} = 0.7\). All distances presented are in
the comoving scale unless noted otherwise, and all magnitudes are
in the AB system (Oke 1974).

2. Observations and Data Reduction

Yang et al. (2009) conducted a wide-field narrowband survey covering an area of 4.82 deg\(^2\) of the Bootes NDWFS,
targeting Lyα emission at \(z = 2.3\), and obtained an unbiased sample of the largest and brightest Lyα blobs at that redshift.
The redshift was chosen to facilitate future observations of the
extended Lyα gas via the optically thin H\(\alpha\) 6563 Å line, which
is redshifted into a relatively sky-line-free part of the infrared
spectrum. Yang et al. (2009) discovered four Lyα blobs with
luminosities of 1.6–5.3 × 10\(^{43}\) erg s\(^{-1}\) and isophotal areas
28–57 arcsec\(^2\). Two of the four blobs form a pair, with a
separation of only 70″ (550 kpc at \(z = 2.3\)), which makes them
ideal targets for our deeper follow-up Lyα survey to map the
spatial distribution of LAEs and LABs described here.

We obtain narrowband images covering a total area of
\(~1° \times 0.5°\) around the two known Lyα blobs (Yang et al.
2009) using the Mosaic1.1 camera on the Kitt Peak National
Observatory (KPNO) Mayall 4 m telescope. In Figure 1 we
show the areas covered by the National Optical Astronomical
Observatory (NOAO) Deep Wide-field Survey (NDWFS) and the
locations of our two pointings (hereafter Boote1 and Boote2) centered on 14\(^{h}\)31\(^{m}\)42\(\frac{2}{\text{s}}\)2, +35\(^{\circ}\)31\(\frac{19}{\text{s}}\)9 and
14\(^{h}\)28\(^{m}\)54\(\frac{10}{\text{s}}\)8, +35\(^{\circ}\)31\(\frac{19}{\text{s}}\)9. Observations were carried out on
2011 April 29 and 30, with exposure times of 7.3 and 6.0 hr,
respectively. During the two observing nights, the average
seeing was \(\sim 1\arcsec\).

We observe with the custom narrowband filter used in
the discovery of the known Lyα blob pair (Yang et al. 2009).
The filter has a central wavelength of \(\lambda_c = 4030\) Å and a bandwidth
of \(\Delta \lambda_{\text{FWHM}} = 47\) Å, corresponding to the Lyα emission at
\(z = 2.3\) and a line-of-sight depth of \(\sim 46.4\) Mpc (\(\Delta z = 0.0037\). Apart from the narrowband (NB) images, we also
use NDWFS broadband \(B_{\text{W}}, B_r,\) and \(J\)-band images for
continuum estimation.

We reduce the data using the MSCRED package in IRAF
(Tody 1986). We correct the images for cross-talk and bias, then
apply the flat-field correction, using both dome and sky-flats. Bad
pixels and satellite trails are masked, and cosmic rays are
removed using the LA-COSMIC software (van Dokkum 2001).
We flux-calibrate by observing 3–4 spectrophotometric standard
stars per night, with typical uncertainties in flux calibration
of \(~0.02–0.04\) mag. The astrometry of our images is improved with
the mscmatch tasks in IRAF using the USNO-B1.0 (Monet et al.
2003) catalog. After matching the image scales, we stack them
using the mscstack task. The total field of view has dimensions of
69\(^{\prime}\)5 × 35\(^{\prime}\)6 or 112.9 Mpc × 57.9 Mpc, with a
total survey volume of 3.03 \times 10\(^{5}\) Mpc\(^3\).

3. Analysis

3.1. Selection of Lyα Emitters

We run Source Extractor (SExtractor; Bertin & Arnouts
1996) on the NB image and select sources having at
least four adjacent pixels above the 1σ local background rms,
identifying $\sim$45,000 sources. After applying a $3 \times 3$ pixel (0.768 $\times$ 0.768 arcsec) boxcar filter to the NB and $B_W$ images, we extract the NB and $B_W$ magnitudes inside circular 3$''$ apertures centered on the selected sources. From these we determine the Ly$\alpha$ line flux, equivalent width (EW), and underlying continuum flux for each of our objects using the following relations:

$$f_\lambda^{\text{cont}} = \frac{F_{B_W} - F_{NB}}{\Delta \lambda_{B_W} - \Delta \lambda_{NB}},$$

(1)

$$F_{\text{line}} = F_{NB} - \Delta \lambda_{NB} \cdot f_\lambda^{\text{cont}},$$

(2)

$$B_W^{\text{cont}} = -2.5 \log \left( f_\lambda^{\text{cont}} \frac{\lambda_{NB}^2}{c} \right) - 48.6,$$

(3)

where $f_\lambda^{\text{cont}}$ is the continuum flux density, $F_{\text{line}}$ is the Ly$\alpha$ line flux, $F_{NB}$ and $F_{B_W}$ are the fluxes in the NB and $B_W$ bands, respectively, and $\Delta \lambda_{NB}$, $\Delta \lambda_{B_W}$ are the bandwidths of the two filters. $B_W^{\text{cont}}$ is the AB continuum magnitude, without the line contribution, and $\lambda_{NB} = 4030$ Å is the central wavelength of the NB filter.

To identify excess Ly$\alpha$ emission, we calculate the color index ($B_W^{\text{cont}} - NB$) of all our candidate sources. We create the Ly$\alpha$ emitter sample by applying the following selection criteria to the extracted objects:

1. $B_W^{\text{cont}} - NB \geq 1$, corresponding to EW$_{\text{obs}} \geq 67$ Å
2. NB $\leq$ 24.77 (5.5$\sigma$ detection threshold)
3. $B_W^{\text{cont}} - NB \geq 5\sigma_{\text{NB}},$

where $B_W^{\text{cont}}$ is the continuum magnitude of an object, without the Ly$\alpha$ line emission. The 5.5$\sigma$ narrowband detection threshold corresponds to a Ly$\alpha$ luminosity of $1.6 \times 10^{42}$ erg s$^{-1}$, which is $\approx$3 times deeper than the original wide-field survey (Yang et al. 2009).

In Figure 2, we show the NB magnitude versus color index and EW for the Boötes 1 and Boötes 2 fields. The dashed vertical and horizontal lines correspond to our selection criteria in NB magnitude and color, respectively. After applying these cuts, we are left with a sample of 354 objects. The blue solid lines correspond to the cut-imposed requirement that the color index should be larger than 5 times the error in the NB magnitude, which eliminates 77 objects from our sample. Removing objects that are close to bright stars or less than 50 pixels away from the image edges further reduces the size of the sample to 223 objects. Finally, we inspect the sample visually, eliminating obvious false detections, like bright nearby galaxies or image artifacts, producing a final sample of 183 objects. We consider sample contamination from [O II] $\lambda$3727 emission in galaxies at $z \approx 0.08$. The rest-frame EW of [O II] emitters at $z = 0.1-0.2$ is $<50$ Å (Hogg et al. 1998; Ciardullo et al. 2013), below our EW cut. Given that [O II] EWs e-fold with a scale length of 6 Å–14 Å (Ciardullo et al. 2013), we estimate that the probability of finding [O II] interlopers with EW$_{\text{obs}} > 67$ Å is less than 1%. We list the properties of the 183 Ly$\alpha$ emitters in Table 1.

![Figure 2](image-url)

**Figure 2.** Color–magnitude plots of objects in the Boötes1 (left) and Boötes2 (right) fields (black dots). Ly$\alpha$-emitter candidates are marked with red dots. The blue stars represent the two known blobs from Yang et al. (2009). The horizontal dashed line marks the cut in EW$_{\text{obs}} > 67$ Å, while the vertical one represents the cut in NB magnitude at 24.77. The blue curve represents the 5$\sigma$ NB magnitude error cut.

To test how our selection criteria might influence the size and spatial distribution of our LAE sample. We create 81 different Ly$\alpha$-emitter samples by varying the selection criteria around our original values. We vary the color index cuts, from 0.8 to 1.2, in nine steps of 0.05, and the NB magnitude cuts, from 24.69 to 24.85, in nine steps of 0.02 mag. Comparing all the resulting samples to the one we originally adopted for this work, we find that the influence of using these different selection criteria on the large-scale distribution of objects is minimal (see Section 4.1).

### 3.2. Selection of Ly$\alpha$ Blobs

With deeper NB imaging data than those in Yang et al. (2009), we search for Ly$\alpha$ blobs with intermediate luminosities and sizes that our prior shallower survey might have missed. Using Equations (1) and (2), we calculate the Ly$\alpha$ line flux for each pixel. The 1$\sigma$ surface brightness limit of the resulting Ly$\alpha$ line image is $\sim 2.1 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$ arcsec$^{-2}$ per 1 arcsec$^2$ aperture, which makes this survey 1.5–2.2 times deeper than the original wide-field survey that led to our discovery of the LAB pair (Yang et al. 2009). We run SExtractor on the line image, selecting sources with at least 16 adjacent pixels above the 5$\sigma$ surface brightness limit. Then, we cross-match this catalog with our emitter sample above to make sure that the extracted Ly$\alpha$ blob candidates do have a Ly$\alpha$ line excess. We select Ly$\alpha$ blob...
Table 1

| ID | R.A. (J2000) | Decl. (J2000) | log(L_{\text{Ly}\alpha}) | EW (A) |
|----|--------------|---------------|--------------------------|--------|
| 1  | 14:32:36.39  | +35:23:34.7   | 42.41 ± 0.05             | 83     |
| 2  | 14:32:13.86  | +35:14:29.3   | 42.52 ± 0.04             | 113    |
| 3  | 14:30:27.02  | +35:14:32.7   | 42.17 ± 0.08             | 314    |
| 4  | 14:32:08.58  | +35:14:37.6   | 41.63 ± 0.25             | 137    |

(This table is available in its entirety in machine-readable form.)

candidates by requiring that their isophotal area above the surface Ly\alpha brightness threshold of 4.45 \times 10^{-18} \text{erg s}^{-1} \text{cm}^{-2} \text{arcsec}^{-2} is larger than 16 arcsec^2. We initially find seven objects matching this criterion, including the two known blobs. In order to estimate possible sample contamination, we place artificial point sources having \( L(\text{Ly}\alpha) = 10^{41-44} \text{erg s}^{-1} \) in our Ly\alpha images and extract them using the same procedures as for the LABs.

Because the noise and background level of the image can vary across the field, we also test how reliably we can recover extended Ly\alpha emission for the LAB candidates. We cut out 101 \times 101 pixels regions around the candidates from the Ly\alpha line image, centered on the candidates, and place them in 4000 empty sky regions in the Boötes 1 and 2 fields. We then run the source extraction procedure using the same settings used for the real data. The measured size and luminosities of the Ly\alpha blob candidates will vary depending on the position in the field. The variance of the source properties recovered this way gives us the uncertainties on the luminosities and sizes of the candidates introduced by placing the objects in different parts of the field. The recovery fraction is defined as the fraction of times the Ly\alpha blob candidate is recovered with a size above 16 arcsec^2. Out of the seven initial candidates, five candidates—including the already known blob pair—have recovery fractions higher than 90%. We consider these to be our LAB sample. The 90% recovery threshold was chosen because the rest of the recovered blobs have much lower recovery fractions: two blobs with 75% and the rest well below the 50% recovery fraction. In Figure 3, we show the isophotal area of the Ly\alpha blob candidates against their Ly\alpha luminosity, as well as the relations for the simulated point sources. The Ly\alpha blob candidates are located at higher isophotal areas for a given luminosity, clearly separated from the locus of point sources.

In Figure 4, we show all our Ly\alpha blob candidates, including the two known Ly\alpha blobs of Yang et al. (2009), in the NB, Ly\alpha line, \( B_W \), \( R_- \), and \( I_- \) bands, respectively. The shapes of the Ly\alpha blob candidates are irregular and their isophotal areas exceed those of their continuum counterparts. In Table 2, we list the properties of the three new Ly\alpha blobs, including position, luminosity, and size. Their Ly\alpha luminosities lie in the range of \( (0.9-1.4) \times 10^{33} \text{erg s}^{-1} \).

4. Results and Discussion

4.1. Discovery of a LAE-traced Protocluster Associated with LABs

Using our 183 Ly\alpha emitter and blob sample, we investigate the large-scale environment around the known Ly\alpha blob pair (Yang et al. 2009). In Figure 5(a), we show the spatial distribution of our Ly\alpha emitters—which includes the new Ly\alpha blobs—across the 69.3 \times 35.4 field. We mark the locations of our 183 Ly\alpha emitters, and indicate the areas that were excluded from our analysis because of contamination from bright sources such as stars.

To estimate the smooth surface density distribution from the discrete positions of the detected galaxies, one often convolves the position map with a Gaussian kernel of width \( \sigma \). The width of this kernel affects the resulting surface density distribution, yet there is no single way of selecting the smoothing method and size of a smoothing kernel. The kernel size is often chosen to match the mode (Saito et al. 2015) or the median (Matsuda et al. 2011) of the distances between objects in a sample. Matsuda et al. (2005) selected a kernel width that matched the redshift dispersion introduced by the peculiar velocity dispersion of their LAE sample, and Yang et al. (2010) used an adaptive kernel technique to smooth their LAE sample. In this paper, we choose a different approach, one meant to find the kernel size generating the smoothed density field that has the highest probability of representing our LAE sample. This technique is described in detail in the Appendix, and we briefly explain it here.

Assuming our LAEs’ positions are randomly drawn from an unknown underlying density distribution \( f \), we use kernel density estimation (KDE) to find an estimate \( \hat{f} \) for the density distribution function. Our method involves convolving the discrete object map with Gaussian kernels, thus generating smooth density maps. Each map is generated using a different kernel width \( \sigma \). We search for the \( \sigma \) value that maximizes the likelihood to observe our Ly\alpha-emitter sample, given the density distribution estimate \( \hat{f} \). We find this optimum value for the kernel width to be \( \sigma = 2.63 \), which is used for the smoothed image in Figure 5(b).

The Ly\alpha-emitter density map in Figure 5(b) reveals a significant overdensity near the field center (R.A. = 14\textdegree30\textprim35\textsec7, decl. = +35\textdeg22\textprim06\sec2), with a projected radius of \( \sim 10 \text{Mpc} \). This overdense region is in both Boötes1 and 2 image frames, and thus is unlikely to be caused by different observing conditions for the two fields or different sample-selection criteria.

To test if the surface density maps change due to the different selection criteria, we create 81 surface density maps, eachcorresponding to a different cut in color index and NB magnitude as described in Section 3.1. Figure 5(c) shows the mean and variance of the Ly\alpha-emitter surface density of these 81 maps. The average density map shows an overdensity that is very similar in size and position to the one we obtained using our selection criteria. The variance is largest away from the overdense region, indicating that the overall number density and density contrast of the overdense region is not strongly dependent on the LAE selection criteria.

To illustrate the size of the overdense region, we show the radial distribution of Ly\alpha emitters in Figure 6. The Ly\alpha-emitter surface density peaks at \( \Sigma_{\text{overdense}} \sim 0.27 \text{arcmin}^{-2} \Delta z^{-1} \) inside a \( \sim 8 \text{Mpc} (5\prime) \) radius centered on the overdense region, decreasing to \( \Sigma_{\text{field}} = (5.4 \pm 0.9) \times 10^{-2} \text{arcmin}^{-2} \Delta z^{-1} \) at radii larger than 25 Mpc, with an average value of \( \Sigma = (7.4 \pm 0.54) \times 10^{-2} \text{arcmin}^{-2} \Delta z^{-1} \) over the entire survey. The scale of this structure clearly demonstrates that one needs a very wide-field survey over \( \sim 100 \text{Mpc} \) to reliably measure the overdensity relative to the background field region.
All uncertainties for the density measurements so far were calculated assuming only Poissonian noise with a sample variance \( \sigma^2 = N \), where \( N \) is the number of galaxies. Cosmic variance (CV) due to galaxy clustering can exceed sample variance and is dependent on the survey geometry. Although our survey volume is quite large and the CV might not be significant, we also provide the density measurements with uncertainties arising from CV.

We use the CV Calculator (Trenti & Stiavelli 2008) to estimate the CV for our survey volume. Given our survey configuration and a sample completeness of 95%, assuming a halo filling factor of 1, we obtain a relative error due to CV of 25.7%. Adding these errors in quadrature, the resulting relative error is approximately 26.7%.

### 4.2. Comparison with Previous Wide-field LAE Surveys

We compare our LAE number densities with those of other surveys at similar redshifts (Palunas et al. 2004; Prescott et al. 2008; Nilsson et al. 2009; Guaita et al. 2010; Mawatari et al. 2012). Since each survey employs different selection criteria in EW and NB magnitude (\( \text{Ly} \alpha \) luminosity), as well as probes different redshift depths due to different filter widths, we need to correct the reported LAE surface density values in the literature. We scale the LAE surface densities assuming \( \text{Ly} \alpha \) luminosity functions \( f(L) \) and an exponential EW distribution \( \text{exp}(\text{EW} \,/-\, \text{EW}_0) \) with a scale length of \( \text{EW}_0 \). We calculate the following correction factors for each survey:

\[
C_L = \frac{\int_{L_0}^{\infty} \phi(L')dL'}{\int_{L_0}^{\infty} \phi(L)dL'},
\]

\[
C_{\text{EW}} = \frac{\int_{\text{EW}_0}^{\infty} \exp(-\text{EW}'/\text{EW}_0)d\text{EW}'}{\int_{\text{EW}_0}^{\infty} \exp(-\text{EW}'/\text{EW}_0)d\text{EW}'},
\]

\[
C_{\Delta z} = \frac{\Delta z_0}{\Delta z_1},
\]

where \( C_L, C_{\text{EW}}, \) and \( C_{\Delta z} \) are the correction factors for \( \text{Ly} \alpha \) luminosity, EW, and redshift depth, respectively; \( L_0, \text{EW}_0, \) and \( \Delta z_0 \) are the luminosity limits, EW cuts, and redshift depths for different surveys, respectively; and \( L_0, \text{EW}_0, \) and \( \Delta z_0 \) are the values used in our survey. We adopt the results from Gronwall et al. (2007) for the Schechter function, assuming no redshift evolution: \( L^* = 10^{32.66} \text{ erg s}^{-1}, \Phi^* = 1.28 \times 10^{-3} \text{ Mpc}^{-3}, \)

---

**Figure 3.** \( \text{Ly} \alpha \) luminosity vs. the isophotal area of \( \text{Ly} \alpha \) emitters, including \( \text{Ly} \alpha \) blobs, for the Boötes1 field (left) and the Boötes2 field (right). \( \text{Ly} \alpha \) emitters, the three new \( \text{Ly} \alpha \) blob candidates, and the two known blobs (Yang et al. 2009) are shown as black dots, filled orange circles, and blue stars, respectively. The dotted horizontal line marks the selection criteria for our \( \text{Ly} \alpha \) blobs: an isophotal area greater than 16 arcsec\(^2\) above the \( 4.45 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2} \) brightness limit. The gray circles represent simulated point sources in our fields.

**Known \( \text{Ly} \alpha \) Blobs (Yang et al. 2009)**

**New Blobs in the Boötes 1 Field**

**A New Blob in the Boötes 2 Field**

**Figure 4.** Images of the three new \( \text{Ly} \alpha \) blob candidates (bottom three rows) and of the two known blobs from Yang et al. (2009) (top two rows). From left to right: NB, continuum-subtracted \( \text{Ly} \alpha \) image, \( B_w, R_-, \) and \( I_- \) bands, respectively. The contours represent the surface brightness of \( 4.45 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcsec}^{-2} \), The isophotal areas of the new intermediate blobs are \( \approx 16 \text{ arcsec}^2 \), with luminosities of \( (0.9-1.4) \times 10^{43} \text{ erg s}^{-1} \). The distance between the tick marks is 5 arcsec.
\[ \alpha = -1.36, \text{and } w_0 = 74 \text{ Å}. \]

We summarize the results from the previous LAE surveys, the adopted correction factors, and the LAE surface densities corrected to our survey properties in Table 3.

In Figure 7, we show the surface density values from other surveys, with redshifts close to \( z = 2.3 \), and compare their measurements with our peak and average surface densities. Our average surface densities agree with those of Nilsson et al. (2009), Mawatari et al. (2012), and Guita et al. (2010). The LAE density of our overdense region with a radius of 10 Mpc is in agreement with average density values from the two surveys that targeted known dense regions, Palunas et al. (2004) and Prescott et al. (2008), who targeted the J2143–4423 protocluster at \( z = 2.38 \) and the LBD05 protocluster (Dey et al. 2005), respectively.

Given that the LF and EW distribution might evolve between \( z = 2 \) and 3 (Ciardullo et al. 2012), we also test how the correction factors \( C_L \), \( C_{EW} \), and \( C_{\Delta z} \) might be affected by the redshift evolution of the luminosity function. We repeat the previous comparison using the luminosity function of Guita et al. (2010) for \( z = 2 \) with \( L^* = 10^{42.33} \text{ erg s}^{-1} \), \( \Phi^* = 0.64 \times 10^{-3} \text{ Mpc}^{-3} \), and \( \alpha = -1.65 \). The resulting surface density values differ by \( \sim 30\% \) on average and by at most 70\% from the values in Table 3 (Figure 7). Note that we do not show the values for the shallowest Palunas et al. (2004) survey because its sources populate only the bright end of the luminosity function, which introduces large errors when extrapolated to the faint end.

### 4.3. Measurement of Surface and Volume Overdensity

To gauge the significance of the discovered overdense structure, and to compare its properties with the cosmological simulations and other known protoclusters, we estimate the surface and volume overdensity in this section.

The surface density contrast \( \delta_S = (\Sigma_{\text{overdense}} - \bar{\Sigma}) / \bar{\Sigma} \) is 2.7 inside a 8.1 Mpc (5\') radius around the position of peak density. This value increases to \( \delta_S = (\Sigma_{\text{overdense}} - \Sigma_{\text{field}}) / \Sigma_{\text{field}} = 4.1 \) if we compare our overdense region to the field density (\( \Sigma_{\text{field}} \)). Throughout the paper and to be consistent with the definition of density contrast used in the literature, we use the average density of the whole survey (i.e., \( \bar{\Sigma} \)) when calculating overdensities. Calculating contrast densities instead of using the average field (i.e., \( \Sigma_{\text{field}} \)) value would increase the peak overdensity, while the standard definition yields a more conservative result.

Assuming that the overdense region is a sphere with a radius of 10 Mpc, we can estimate the volume density contrast as follows. We find 35 LAEs inside a projected area with a 10 Mpc radius centered on R.A. = 14°30′31″, decl. = +35°25′01″, while only \( \sim 9 \) LAEs are expected given the average volume density over the survey. Thus, we estimate that \( \sim 26 \) more LAEs are located within the assumed spherical overdensity having a volume of \( \sim 4.18 \times 10^3 \text{ Mpc}^3 \). Our survey contains 183 objects in a volume of 3.0 \( \times \) 10^8 Mpc^3. The volume density contrast \( \delta = (\rho_{\text{overdense}} - \bar{\rho}) / \bar{\rho} \) is then \( \sim 10.4 \), where \( \rho_{\text{overdense}} \) is the density inside the spherical region, and \( \bar{\rho} \) is the average density over the whole survey.

Several other surveys also find LAE overdensities at \( z = 2 - 4 \). At \( z = 2.16 \), a protocluster with an overdensity of \( \delta_S \sim 3 \) is associated with the PKS 1138–262 radio galaxy and its extended Ly\( \alpha \) halo (Kurk et al. 2000; Venemans et al. 2007). Targeting a known cluster, J2143–4423, at \( z = 2.38 \), Palunas et al. (2004) found an LAE overdensity of \( \delta_S \approx 2 \). A similar surface overdensity of \( \delta_S \approx 2 \) was seen by Prescott et al. (2008) around a known LAB at \( z = 2.7 \). An overdensity was found by Saito et al. (2015) around the radio galaxy TN J1338-1942 at redshift 3.1, with \( \delta_S = 2.8 \pm 0.5 \). At \( z = 3.78 \), two or three overdensities with similar \( \delta_S \) values were found by Lee et al. (2014) and Dey et al. (2016), with \( \delta_S = 2.5 - 2.8 \). The largest overdensity by far lies in the SSA22 field (Steidel et al. 2000), with a surface density contrast \( \delta_S = 5 \pm 2 \) (Steidel et al. 2000; Matsuda et al. 2004, 2005; Yamada et al. 2012). More recently, Cai et al. (2017b) discovered a massive overdensity at \( z = 2.3 \), having a spectroscopically confirmed volume density contrast of \( \delta \sim 10 \), associated with an extremely large and luminous Ly\( \alpha \) nebula (Cai et al. 2017a). All these surveys probe redshift slices of \( \Delta z \sim 0.03 - 0.16 \). Although it is difficult to directly compare these overdensity contrasts with our own values because of different kernels and field sizes, as well as different \( \Delta z \)’s, the \( \delta_S \) and \( \delta \) of our overdensity are roughly comparable to these protocluster candidates.

### 4.4. Will Boötes J1430+3522 Evolve into a Cluster Today?

Using Ly\( \alpha \) emitters as a density tracer, we discover an overdense region with a projected surface density of \( \delta_S = 2.7 \pm 1.1 \) and a radius of \( \sim 10 \text{ Mpc} \). To address whether this structure could collapse into a virialized galaxy cluster by \( z = 0 \), i.e., whether it is in fact a “protocluster,” we compare our observations with the analysis of structure.
formation from cosmological simulations by Chiang et al. (2013).

Using the Millennium Run (MR; Springel et al. 2005) cosmological simulation, Chiang et al. (2013) identified the mass, extent, and density contrast that galaxy cluster progenitors must have in order to evolve into galaxy clusters at $z = 0$. In their study, a cluster is defined as a virialized dark matter halo with a total mass greater than $10^{14} M_\odot$, at redshift $z = 0$. Based on this definition they tracked the evolution of DM halos and galaxies in $\sim $3000 clusters from early epochs ($z = 7$) to the present day. For this sample, they calculated the correlation between galaxy density contrast of protoclusters at

---

**Figure 5.** (a) Spatial distribution of 183 Ly$\alpha$ emitters and blobs. The filled gray and red circles represent Ly$\alpha$ emitters and new Ly$\alpha$ blob candidates, respectively. The star symbols are the two previously known Ly$\alpha$ blobs (Yang et al. 2009). The radii of the circles are proportional to the logarithm of the Ly$\alpha$ emitters’ luminosities, in the range of $10^{41.4} - 43.4$ erg s$^{-1}$. The field of view is $69\,^\circ\times 35\,^\circ$ ($138.5$ Mpc $\times 57.5$ Mpc). The dotted lines enclose areas that have been excluded from our analysis because of contamination from bright stars or galaxies. (b) Ly$\alpha$ emitter surface density distribution obtained from the KDE method explained in the Appendix. The contour labels show the surface density of Ly$\alpha$ emitters in arcmin$^{-2} \, \text{per} \, \Delta z = 0.037$—the value given by the narrowband filter width. The two known Ly$\alpha$ blobs are marked with stars. The overdense region is clearly visible toward the center of the image. (c) Average of scatter of the Ly$\alpha$-emitter surface density of the 81 surface density maps corresponding to different selection methods. The contour labels represent the average surface density, while the background image represents the scatter around this average map, in percentages. The scatter is largest away from the overdense region, increasing the confidence that the shape and size are not significantly affected by varying the selection criteria.
different redshifts and the mass of its present-day cluster offspring. They also showed how the projected density contrast is affected by the redshift uncertainty $\Delta z$ of a survey, demonstrating that potential protocluster overdensities become observationally indistinguishable from the field, for all except the most massive structures, if the overdensities are measured with $\Delta z > 0.1$. Thus, wide-field narrowband imaging surveys are the key to identifying early stages of cluster formation.

In their analysis, Chiang et al. (2013) measure the density contrast of the structures in the MR data after smoothing it with $(15 \text{ Mpc})^3$ and $(25 \text{ Mpc})^3$ top hat cubic kernels. To match these kernel sizes, we smooth our survey map with $(15 \times 15 \times \Delta z) \text{ Mpc}^3$ and $(25 \times 25 \times \Delta z) \text{ Mpc}^3$ rectangular windows, with a redshift uncertainty of $\Delta z \approx 0.037$. In this configuration, we find $\delta_{z,15} = 3.4$ and $\delta_{z,25} = 2.0$ for Boötes J1430+3522.

According to Chiang et al. (2013), an uncertainty of $\Delta z \approx 0.037$ in the redshift of the Ly$\alpha$ emitters used to trace an overdensity at redshift $z = 2–3$ reduces the apparent surface density contrast by $\sim 50\%$ compared to its original value calculated using $(15 \text{ Mpc})^3$ cubic windows. This is because with increasing redshift uncertainties, more galaxies in the background and the foreground of the overdense regions are included in the analysis and smooth out irregularities in surface density. Correcting for this effect, we obtain $\delta_{15,\text{corrected}} \approx 6.8$. Note that if we assume that the overdensity is confined only within the $(15 \text{ Mpc})^3$ cube, $\delta_{15,\text{corrected}} = 10.6$ would be required to yield the observed $\delta_{z,15} = 3.4$. Therefore, $\delta_{15,\text{corrected}} \approx 6.8$ should be a reasonable value for the density contrast over the $(15 \text{ Mpc})^3$ cubic window.

This density contrast is much higher than $\delta_{15} = 2.88$, the value needed for a $z \sim 2$ structure to evolve into a cluster at $z = 0$ with $>80\%$ probability (Chiang et al. 2013). Here, we have adopted $\delta_{15}$ using galaxies with SFR $> 1 \text{ M}_\odot \text{ yr}^{-1}$, which are analogs to Ly$\alpha$-emitter populations. This $\delta_{15,\text{corrected}}$ is high enough for it to evolve into a present-day cluster with near $100\%$ certainty, even if a wide range of other tracer populations are assumed (see Figure 8 of Chiang et al. 2013). Therefore, we conclude that Boötes J1430+3522 is indeed a “protocluster.”

Finally, using the correlation found by Chiang et al. (2013) between galaxy contrast at a given epoch and present-day cluster mass, we estimate the future mass protocluster to be $\log(M/M_\odot) \sim 15.1 \pm 0.2$ similar to that of Coma cluster.

4.5. Size and Amplitude of Protoclusters

We have discovered a new protocluster traced by LAEs and Ly$\alpha$ blobs. To compare it to other known protoclusters, we compile previous narrowband imaging surveys at $z = 2–3$ that have discovered both Ly$\alpha$ blobs and protoclusters in the same field. These three protoclusters are located in the E-CDFS (Yang et al. 2010), the 53W002 (Mawatari et al. 2012), and SSA22 fields (Matsuda et al. 2011; Yamada et al. 2012) at $z$ = 2.3, 2.4, and 3.1, respectively.4

We reproduce the surface density maps of Ly$\alpha$ emitters and blobs for these three fields and Boötes J1430+3522 field in Figure 8. To make these maps, we use the KDE and cross-validation method presented in this paper with Gaussian kernel widths of $\sigma = 1.43, 1.55, 1.87$, and 2.64 for the 53W002, E-CDFS, SSA22, and Boötes J1430+3522 fields, respectively. These kernel sizes are within 25% of the values originally adopted by each survey: $\sigma = 1.55$ for 53W002 (Mawatari et al. 2012), 1.2–2.2 for E-CDFS (Yang et al. 2010), and 1.5 for SSA22 (Yamada et al. 2012). To test the effect of kernel sizes on our results below, we also produce maps with (1) the values adopted in each reference and (2) a same width (1.5) for all four fields. Our results here do not change with the choice of the kernel size.

Figure 8 shows the contours of surface overdensity $\delta_\Sigma = (\Sigma - \bar{\Sigma})/\bar{\Sigma}$ for each survey. When calculating $\delta_\Sigma$, we estimate $\bar{\Sigma}$ over each survey. For the E-CDFS protocluster (Balestra et al. 2010; Yang et al. 2010,) which almost fills the 30’ × 30’ field, we use the $\bar{\Sigma}$ from our Boötes survey because both surveys used the same narrowband filter and sample-selection methods.

Figure 8 shows that both the peak amplitudes and the sizes of the protoclusters are consistent with each other, despite the wide ranges of survey areas probed in each survey. In particular, three protoclusters in the E-CDFS, Boötes J1430+3522, and SSA22 fields have almost identical peak surface density contrasts of $\delta_\Sigma = 2.8–3.0$. In contrast, the 53W002 protocluster has a smaller size and lower peak amplitude than the others, suggesting it is only moderately rich. The three protoclusters (E-CDFS, Boötes J1430+3522 and SSA22) have 8.5–10 physical Mpc diameters (28–39 comoving Mpc; 17’–21’) if we measure the largest dimension of the $\delta_\Sigma = 1$ contour (dashed). The linear size of the protocluster does not grow bigger than this typical size, even though the survey area increases from E-CDFS (35’), Boötes J1430+3522 (70’4) to the

4 Prescott et al. (2008) and Erb et al. (2011) also found Ly$\alpha$ blobs associated with overdensities traced by LAEs. However, the coordinates of the LAEs in their fields are not available, and it is unknown if the small survey area (220 arcmin$^2$) of Erb et al. (2011) includes the whole overdensity.
SSA22 field (110'). For $\delta_v = 2$ (dotted–dashed) contour, the protoclusters also have similar sizes of 4.6–7.2 physical Mpc (16–24 comoving Mpc) with wider ranges.

Protocluster overdensity profiles from simulations (Chiang et al. 2013) show that even for the most massive protoclusters at redshift $z = 2$–3 (i.e., progenitors of galaxy clusters with a present-day mass greater than $10^{15} M_\odot$), the average diameters of areas with a volume density contrast above $\delta = 1$ and 2 are $\approx 32$ and $\approx 24$ Mpc, respectively. Although it is not straightforward to relate the size measured for a fixed surface density contrast $(\delta_v)$ to that measured for a volume density contrast $(\delta)$, these sizes are in good agreement with the observations discussed above.

The comparable extents of protoclusters at $z = 2$–3, and the fact that their observed sizes do not grow with the extent of the surveyed field, suggest that they are the largest bound structures at that epoch. It is clear from our results that a very wide-field survey over $\sim 1$ degree is required to reliably confirm massive protoclusters at this epoch and to determine their full physical sizes.

### 4.6. Ly$\alpha$ Blobs in Protocluster Outskirts

Visually, all the maps in Figure 8 are striking; the Ly$\alpha$ blobs often lie outside the densest concentration of LAEs. Mawatari et al. (2012) also note that all four of their LABs are located on the edges of high-density regions. To quantify relative, local environments of Ly$\alpha$ emitters and blobs, we measure their local overdensities from the smoothed surface density maps (Figure 9). The distribution of LAEs’ local overdensities is similar to the lognormal distribution that is known to approximate the dark matter distribution (e.g., Coles & Jones 1991; Orsi et al. 2008). The two-sample Kolmogorov–Smirnov test shows that the distributions of the LAE and LAB populations are different at the 3.8$\sigma$ significance level. The distributions differ most at moderate overdensities, $\delta_v = 1$–2, where there is a clear excess of Ly$\alpha$ blobs. While it is not reanalyzed here, the LABd05 blob is also located near the region of $\delta_v \sim 1.3$ (Prescott et al. 2008; see their Figure 3). Likewise, the six Ly$\alpha$ blobs in Erb et al. (2011) appear to lie at the edges of the HS 1700+643 protocluster field. We conclude that Ly$\alpha$ blobs prefer moderately overdense regions of LAEs that are twice or three times denser than the average density of the survey ($\delta_v \approx 0$), perhaps avoiding the densest regions within a protocluster.

Why do Ly$\alpha$ blobs occupy the moderate overdense region or outskirts of protoclusters? One possibility is that Ly$\alpha$ blobs represent protogroups that are accreting into a more massive protocluster from the cluster outskirts. Prescott et al. (2012) found that the LABd05 Ly$\alpha$ blob (Dey et al. 2005) contains numerous compact, small, low-luminosity (<0.1$L_\odot$) galaxies. Similarly, Yang et al. (2011, 2014b) identified several H$\alpha$ or [O III] emitting sources within Ly$\alpha$ blobs with relative line-of-sight velocity differences of $\sim 200$–400 km s$^{-1}$, which are consistent with the velocity dispersions of $\sim 10^{13} M_\odot$ galaxy groups. Furthermore, the number and variance of Ly$\alpha$ blobs is consistent with them occupying $\sim 10^{13} M_\odot$ halos. We speculate that the extended Ly$\alpha$-emitting gas may be the protogroup medium and/or stripped gas originating from galaxy–galaxy interactions within these protogroups.

We test the plausibility of this scenario by checking if the expected number of protogroups in the massive protocluster environment is roughly consistent with that of the Ly$\alpha$ blobs around Boötes J1430+3522. We estimate that our LAE overdensity will evolve into a $\sim 10^{15} M_\odot$ rich cluster today (Section 4.4). In this case, simulations predict that the current protocluster mass is $\sim 10^{14} M_\odot$ and that it accretes $\sim 15 \times 10^{13} M_\odot$.

---

**Table 3**

| Survey                | $z$  | $\Delta z$ | EW Cut | $L$(Ly$\alpha$) | $C_L$ | $C_{EW}$ | $C_L$ | $C_L$ | $\Sigma$ |
|-----------------------|-----|------------|--------|----------------|------|---------|------|------|---------|
| Nilsson et al. (2009) | 2.2 | 0.1061     | 20     | 42.36          | 0.364| 1.000   | 1.565| 0.107|         |
| Guaita et al. (2010)  | 2.1 | 0.0411     | 20     | 41.80          | 0.940| 1.000   | 0.415| 0.075|         |
| Mawatari et al. (2012)| 2.4 | 0.0683     | 25     | 41.99          | 0.566| 1.121   | 0.611| 0.107|         |
| Prescott et al. (2008)| 2.7 | 0.1653     | 40     | 42.18          | 0.234| 1.580   | 0.955| 0.350|         |
| Palunas et al. (2004) | 2.3 | 0.0444     | 36     | 42.78          | 0.870| 1.475   | 8.958| 0.201| 0.074   |

This Work

| $z$  | $\Delta z$ | EW Cut | $L$(Ly$\alpha$) | $C_L$ | $C_{EW}$ | $C_L$ | $C_L$ | $\Sigma$ |
|------|------------|--------|----------------|------|---------|------|------|---------|
| 2.3  | 0.0370     | 20     | 42.19          | ...  | ...     | ...  | ...  | 0.074   |

**Note.** (1) reference for the survey, (2) survey redshift, (3) redshift depths from filter widths, (4)–(5) selection criteria for EWs (Å) and Ly$\alpha$ luminosity ($\log(L/\text{erg s}^{-1})$), (6)–(8) correction factors for the redshift depth, EW, and Ly$\alpha$ luminosity, introduced in Section 4.2, (9) average surface density (arcmin$^{-2}$ $\Delta z^{-1}$) over the entire field corrected for our sample-selection criteria.
halos from $z \sim 2.3$ to 0 (Gao et al. 2004; Giocoli et al. 2008; Jiang & van den Bosch 2016). Thus, the five Ly$\alpha$ blobs that we detect within $\sim$10 Mpc ($\sim$5 virial radii) could plausibly trace some of the group-like halos that build the cluster.

5. Conclusions

We carry out a deep narrowband imaging survey of a $\sim 1^\circ \times 0.5^\circ$ region at $z = 2.3$ around a known bright Ly$\alpha$ blob pair discovered by a blind narrowband survey (Yang et al.
2009). We test whether bright Lyα blobs are indeed a tracer of overdense regions at high redshift.

We find a total of 183 Lyα emitters, including three new intermediate Lyα blobs in our 69.3 × 35.4 field. The average Lyα-emitter surface density in our field is \( \Sigma = (7.4 \pm 1.9) \times 10^{-2} \text{arcmin}^{-2} \Delta z^{-1} \), corresponding to a volume density \( n = (6 \pm 1.5) \times 10^{-3} \text{Mpc}^{-3} \) over the survey volume of 3.03 × 10^5 Mpc^3. The surface density varies from 5.4 × 10^{-2} arcmin^{-2} \Delta z^{-1} in the field region to 0.27 arcmin^{-2} \Delta z^{-1} at the densest part, in good agreement with results from previous surveys that targeted either the densest part, in good agreement with results from previous surveys that targeted either field regions or protoclusters at similar redshifts.

We discover a massive overdensity (Boötes J1430+3522) of Lyα emitters with a surface density contrast of \( \delta_S = 2.7 \pm 1.1 \), a volume density contrast of \( \delta \sim 10.4 \), and a projected diameter of \( \approx 20 \) comoving Mpc. By comparing our measurements with an analysis of the MR cosmological simulation (Chiang et al. 2013), we conclude that this large-scale structure is indeed a protocluster and is likely to evolve into a present-day Coma-like galaxy cluster with log(M/M_\odot) \approx 15.1 \pm 0.2.

In our survey and three others we re-analyze here, the physical extent and peak amplitude of the LAE overdensities are consistent across the surveys. Because these properties do not increase with survey size, it is likely these overdensities are the largest structures at this epoch and will indeed evolve into rich clusters today.

The discovery of a protocluster in the vicinity of the two Lyα blobs, along with the discovery of three new nearby LABs, confirms that bright Lyα blobs are associated with overdense regions of LAEs. Yet, among the four surveys we analyze, LABs tend to avoid the innermost, densest regions of LAEs and are preferentially located in the outskirts at density contrasts of \( \delta_S = 1–2 \). This result, and the likelihood that blobs themselves occupy \( \sim 10^{13} M_\odot \) individual halos (Yang et al. 2010), suggest, that Lyα blobs represent protogroups that will be accreted by the protocluster traced by LAEs. In that case, the extended Lyα-emitting blob gas may be a precursor of the intragroup medium, and ultimately a contributor to the intracluster medium.

We thank the anonymous referee for her or his thorough reading of the manuscript and helpful comments. We thank Lucia Guaita for providing the Lyα-emitter catalogs, and Ashoordin Ashoormaran for his insights on the sub-halo mass function. T.B. and Y.Y. acknowledge support from the BMBF/DLR grant No. 50 OR 1306. Y.Y.’s research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2016R1C1B2007782). Support for B.M. was provided by the DFG priority program 1573 “The physics of the interstellar medium.” T.B., A.K., and F.B. acknowledge support by the Collaborative Research Council 956, sub-project A1, funded by the Deutsche Forschungsgemeinschaft (DFG). A.I.Z. acknowledges support from NSF grant AST-0908280 and NASA grant NNX10AD47G.

Appendix

Estimating LAE Surface Density Using KDE and Cross-validation

To build a continuous Lyα-emitter density map (Figure 5(b)) from the spatial distribution of Lyα emitters (Figure 5(a)), we use the KDE method (Rosenblatt 1956; Parzen 1962) with a cross-validation technique. Assuming that the sky positions of our Lyα-emitter sample \( \{x_1, x_2, \ldots, x_n\} \) are randomly drawn from an underlying unknown surface density distribution \( f(x) \), our goal is to find an estimator \( \hat{f}(x) \) for the true distribution. Using KDE

\[
\hat{f}(x) = \frac{1}{N} \sum_{j=1}^{N} K(x - x_j; \sigma),
\]

where \( K(x; \sigma) \) is a normalized kernel, e.g., in a functional form of uniform, triangular, or Gaussian. The \( \sigma \) is a bandwidth, a free-smoothing parameter that strongly influences the estimate obtained from KDE. Note that \( \sigma \) can be 1D or 2D, as well as different for each datum. In our application, we consider 1D and 2D Gaussian kernels:

\[
K(x; \sigma) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp \left( -\frac{x^2}{2\sigma^2} \right),
\]

\[
K(x; \sigma) = \frac{1}{2\pi\sigma_\gamma} \exp \left( -\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2} \right).
\]

Our goal is to determine the \( \sigma \) that best describes the data itself. KDE is mathematically identical to smoothing a map image with a Gaussian kernel, the approach most often taken in the literature, although the smoothing widths are often chosen rather arbitrarily. We show below that an optimal \( \sigma \) can be determined from the data themselves. For that purpose, we use a leave-one-out cross-validation scheme (e.g., Hogg 2008): let \( \hat{f}_{-i}(x) \) be the KDE of \( f \) that is obtained from our sample, excluding the \( i \)th element. The probability of finding that \( i \)th element at the observed position \( x_i \) is proportional to \( \hat{f}_{-i}(x_i) \):

\[
\hat{f}_{-i}(x_i) = \frac{1}{\sqrt{2\pi \sigma^2}} \exp \left( -\frac{|x_i - x|^2}{2\sigma^2} \right)
\]
We then find the parameters that best predict the observed data by maximizing the likelihood of finding all \( \{ x_i \} \) for a given \( \sigma \):

\[
L(\{x_i\}^N_{i=1}|\sigma) = \prod_{i=1}^{N} \hat{f}_i(x_i).
\]  

(11)

We use a simple grid search to determine the kernel width \( \sigma \). Figure 10 shows the likelihood \( L \) as a function of \( \sigma \) for an \( 1^-5^\prime \) range. The maximum likelihood is obtained for \( \sigma = 2.63 \pm 0.24 \). If we adopt a 2D Gaussian kernel with two smoothing parameters \( (\sigma_x, \sigma_y) \) as in Equation (9), \( \sigma_x = 3.00^{+0.78}_{-0.75} \) and \( \sigma_y = 2.31^{+0.83}_{-0.82} \). The 1D kernel width is within the 68.3% confidence interval of 2D kernel width. We use \( \sigma = 2.63 \) throughout the paper to estimate the underlying density distribution.

**References**

Balestra, I., Mainieri, V., Popesso, P., et al. 2010, A&A, 512, A12

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393

Bleem, L. E., Stalder, B., de Haan, T., et al. 2015, ApJS, 216, 27

Bolton, E. S., & Green, D. 2020, ApJ, 906, L35

Ciardullo, R., Gronwall, C., Hickey, T., et al. 2007, ApJ, 667, 79

Gladders, M. D., & Yee, H. K. C. 2000, ApJ, 537, L5

Haiman, Z., De Breuck, C., Galametz, A., et al. 2011a, MNRAS, 410, 1537

Haiman, Z., & Rees, M. J. 2001, ApJ, 556, 87

Haiman, Z., Spaans, M., & Quataert, E. 2000, ApJL, 537, L5

Hatch, N. A., & De Breuck, C., Galametz, A., et al. 2011a, MNRAS, 410, 1537

Hatch, N. A., Kurk, J. D., Pentericci, L., et al. 2011b, MNRAS, 415, 2993

Hayashi, M., Kodama, T., Tadaki, K.-I., Koyama, Y., & Tanaka, I. 2012, ApJ, 757, 15

Hogg, D. W., 2008, arXiv:0807.4820

Hogg, D. W. 2008, arXiv:0807.4820

Hogg, D. W., Cohen, J. G., Blandford, R., & Pahre, M. A. 1998, ApJ, 504, 622

Ivison, R. J., Smail, I., Le Borgne, J.-F., et al. 1998, MNRAS, 298, 583

Ivison, R. J., Shalloe, R., 2005, MNRAS, 366, 2135

Jiang, F., & van den Bosch, F. C. 2009, ApJ, 700, 1

Kurk, J. D., Röttgering, H. J. A., Pentericci, L., et al. 2006, A&A, 568, L1

Laursen, P., & Sommer-Larsen, J. 2007, ApJL, 657, L69

Lee, K.-S., Dey, A., Hong, S., et al. 2014, ApJ, 796, 126

Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106

Matsuda, Y., Yamada, T., Hayashino, T., et al. 2005, ApJ, 634, L125

Matsuda, Y., Yamada, T., Hayashino, T., et al. 2011, MNRAS, 410, L13

Mawatari, K., Yamada, T., Nakamura, Y., Hayashino, T., & Matsuura, Y. 2012, ApJ, 757, 15

Mullis, C. R., Rosati, P., Lamer, G., et al. 2005, ApJ, 623, L85

Mullis, C. R., Rosati, P., Lamer, G., et al. 2005, ApJ, 623, L85

Mukhovatov, Y., Ostriker, E. C., & Frenk, C. S. 2009, MNRAS, 397, 1150

Oke, J. B. 1974, ApJS, 27, 21

Oke, J. B. 1974, ApJS, 27, 21

Oke, J. B. 1974, ApJS, 27, 21

Oke, J. B. 1974, ApJS, 27, 21

Oke, J. B. 1974, ApJS, 27, 21

Oke, J. B. 1974, ApJS, 27, 21

Oke, J. B. 1974, ApJS, 27, 21
