

Extensive diffuse Lyman- α emission correlated with cosmic structure

Received: 10 January 2023

Accepted: 21 July 2023

Published online: 28 September 2023



D. Christopher Martin¹✉, Behnam Darvish¹, Zeren Lin¹, Renyue Cen²,
Mateusz Matuszewski¹, Patrick Morrissey³, James D. Neill¹ &
Anna M. Moore⁴

The intergalactic medium represents the dominant reservoir of baryons at high redshift, traces the architecture of the cosmic web dominated by dark matter and fuels ongoing galaxy evolution. The intergalactic medium has been studied using the absorption lines of quasi-stellar objects, including the Lyman- α forest, but these absorption lines are unable to provide the information that emission maps would give. However, because of the low surface brightness and extended, diffuse distribution, direct detection of an emission equivalent to the absorption Lyman- α forest has not been possible with existing instrumentation and observational approaches. Using a purpose-built instrument, with nod-and-shuffle and dual-field subtraction, we have detected an emission Lyman- α forest. The emission forest is highly extended, shows filamentary morphology with filaments connecting galaxies, exhibits statistics like the absorption Lyman- α forest, displays spectra resembling the absorption forest and is correlated with galaxy-traced overdensities consistent with bias like dark matter. We conclude that the emission Lyman- α forest may provide a new tool for tracing a substantial fraction of the cosmic web of baryons and dark matter.

Emission from the intergalactic medium (IGM) was first predicted by ref. 1, with a reference fiducial surface brightness of $SB_0 = 3.7 \times 10^{-20} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ arcsec}^{-2}$ at redshift $z = 3$ (ref. 2). Higher intensities may be expected from Lyman α (Ly α) photon scattering^{3–5}, particularly in overdense regions. Very-high-intensity levels are found in the vicinity of quasi-stellar object (QSO) illumination^{6–11}. Rare, isolated Ly α nebulae with high luminosity and surface brightness have been detected in highly overdense regions, often associated with massive collapsing galaxy groups, submillimetre star-forming galaxies and obscured QSOs^{12–20}, some showing evidence of filamentary extensions and inflow^{20–22}. Ly α halos are ubiquitous around star-forming galaxies at high redshift^{23–25}. A few tentative detections of faint IGM filaments have been presented²⁶. But general Ly α emission is predicted to be low surface brightness, diffuse and highly extended, rendering detection of general Ly α emission from the cosmic web and from the gas responsible for Ly α absorption very challenging.

The Keck Cosmic Web Imager (KCWI) was designed to detect very-low-surface-brightness emission using fast, high-throughput optics, a large image slicer that focuses considerable extended emission-line flux into a single resolution element, transmission down to 3,500 Å to exploit the faint blue sky during dark time, and a deployable nod-and-shuffle (NAS) mask and NAS charge-coupled device (CCD) clocking sequences²⁷. As IGM Ly α emission is expected to be ~0.25% of sky, we used the NAS capability of the KCWI to perform precision sky subtraction²⁸, using a dual-field-subtraction (DFS) strategy. DFS targets two fields (A and B) separated by a transverse distance large enough that the emission redshifts distributions will exhibit low overlap fractions (<5%; Methods). During each exposure, field A is observed for 2 min, followed by field B for the same exposure time. The two spectra are shuffled back and forth beneath the NAS, which covers two-thirds of the CCD. After a total of 20 min for each field, the CCD is read out.

¹Cahill Center for Astrophysics, California Institute of Technology, Pasadena, CA, USA. ²Department of Astrophysical Sciences, Princeton University, Princeton, NJ, USA. ³Jet Propulsion Laboratory, Pasadena, CA, USA. ⁴Research School of Astronomy and Astrophysics, Australian National University, Canberra, Australian Capital Territory, Australia. ✉e-mail: martinc@caltech.edu

Sky background common to both fields, including terrestrial background, airglow and zodiacal light, will subtract out, as long as airglow variations are on longer timescales than the 4 min NAS cadence.

The DFS approach was validated using a cosmological hydrodynamical zoom-in simulation²⁹ and CLOUDY³⁰ to convert each simulation voxel to ionization fraction and Ly α surface brightness. We generated a forward-modelled high-resolution ($\Delta\lambda = 0.1\text{\AA}$) KCWI simulated data cube based on the gas line-of-sight velocity, where λ is observed wavelength, and applying an optically thick line profile³¹ (Methods). Extended Data Fig. 1 shows the simulations that provided evidence that the observation strategy would work. We can calculate the average overlap fraction as a function of transverse field separation. Field separations of >12 arcmin yield average overlap fractions of $<3\%$ with no preselection. The emission discussed below may rise to 5% overlap at low surface brightnesses.

We targeted two nearby regions in the Cosmic Evolution Survey (COSMOS) field³² (Extended Data Fig. 2): field A targeted an overdensity at $z = 2.242$, while field B targeted an overdensity at $z = 2.474$. Fields A and B are separated by 24 arcmin, or 40 Mpc (comoving; cMpc). We use the KCWI large slicer ($20'' \times 33''$, $1.3''$ slice width) and low-resolution grating (BL) centred at $4,060\text{\AA}$. The resulting wavelength coverage (limited by the NAS mask) is $3,835\text{--}4,320\text{\AA}$, with a Ly α redshift range of $2.15 < z < 2.55$ and a comoving depth of 500 cMpc. A $60 \times 60\text{ arcsec}^2$ region of each field is obtained using a mosaic of 2×3 or 3×2 fields depending on the position angle (PA). Roughly half of the exposure is at PA = 0° , the other at PA = 90° (Supplementary Table 1). Exposure to exposure subpixel dithering improves spatial sampling. Each exposure is converted to a data cube using the KCWI pipeline (kderp). The A and B cubes are subtracted to remove sky. The difference cubes are then coadded (with small astrometry corrections) into a final mosaic difference cube. We obtained a total average exposure of 3 h over the entire mosaic in each field in photometric conditions. Field A Ly α emission appears as a positive signal, while field B Ly α emission appears as a negative signal. We estimate that the resulting sky subtraction is accurate to $<0.02\%$ (1σ) of the sky continuum level (Methods). Using the COSMOS source catalogue³³, we point spread function (PSF) subtract all sources brighter than a certain magnitude limit ($m_{\text{lim}} = 24$) that appear in either the A or B regions. We perform sky substitution for sources in the $24 < m_b < 28$ magnitude range (Methods). The cube is then adaptively smoothed^{7,21,34,35}, as described in Methods. Comparison of raw and smoothed images are shown in Extended Data Fig. 3.

As we present in Fig. 1, we have detected extended Ly α emission at multiple redshifts that is correlated with galaxy overdensities. Strong emission is detected at both target overdensity redshifts. We note that the observations cover a cosmic volume of $\sim 2,000\text{ cMpc}^3$, while the overdensities occupy $\sim 50\text{ cMpc}^3$. In fact, extended emission is also detected at multiple redshifts (five for each field in this figure) associated with peaks in the galaxy overdensity versus redshift distributions. We estimate a relative galaxy overdensity (Methods) δ_{gal} in radius $r = 8\text{ h}^{-1}\text{ cMpc}$ spheres, and the average Ly α surface brightness in $\pm 8\text{ h}^{-1}\text{ cMpc}$ -wide slices of the KCWI data cube. The figure shows 5 $\text{\AA}$ slices at five redshifts for each field, each showing both extended Ly α

emission and corresponding to a peak in the galaxy distribution. The emission is often quite extended, filling much of the field of view. Such emission would be difficult to detect using conventional sky subtraction approaches, which require a portion of the field of view to estimate sky. The intensity scale shows both the total surface brightness (SB) and the average surface brightness per observed $\text{\AA}$ (SB $_{\lambda}$). In Fig. 1c, it can be seen that the average surface brightness is well correlated with the galaxy overdensity, with some evidence of saturation for $\delta_{\text{gal}} \gg 10$. The latter could be due to radiative transfer effects, other nonlinearities in the emission process or cosmic variance. For this analysis, we assume that the spectroscopic sample, consisting of 1,100 galaxies in the appropriate redshift range ($2.1 < z < 2.6$) and an average of 0.7 galaxies in $r = 8\text{ h}^{-1}\text{ cMpc}$ spheres, traces the overall galaxy distribution, and that the relative overdensity is a reasonable estimate (other overdensity results in COSMOS are compared in Methods). We discuss redshift sample selection function effects below and in Methods.

We use the following definitions: 1 line unit (LU) ($1\text{ ph cm}^{-2}\text{ s}^{-1}\text{ sr}^{-1}$) = $1.2 \times 10^{-22}\text{ erg cm}^{-2}\text{ s}^{-1}\text{ arcsec}^{-2}$ at $z = 2.35$, and $1\text{ LU } \text{\AA}^{-1}$ is $1\text{ ph cm}^{-2}\text{ s}^{-1}\text{ sr}^{-1}\text{ \AA}^{-1}$ = $1.2 \times 10^{-22}\text{ erg cm}^{-2}\text{ s}^{-1}\text{ \AA}^{-1}\text{ arcsec}^{-2}$. We give total surface brightness (SB) in LU and surface brightness per observed $\text{\AA}$ (SB $_{\lambda}$) in $\text{LU } \text{\AA}^{-1}$. An optically thick Ly α fluorescent source in ref. 2 has a reference surface brightness $\text{SB}_0(z=3) = 3.7 \times 10^{-20}\text{ erg cm}^{-2}\text{ s}^{-1}\text{ arcsec}^{-2}$. As the metagalactic ultraviolet (UV) background is either the same or slightly higher at $z = 2.35$ than at $z = 3$, this maximum surface brightness should be higher by $(1+z)^4$, so the reference surface brightness is $\text{SB}_0(z=2.35) \equiv \text{SB}_0 = 7.4 \times 10^{-20}\text{ erg cm}^{-2}\text{ s}^{-1}\text{ arcsec}^{-2} \approx 640\text{ LU}$. A typical Ly α predicted linewidth is $8\text{\AA}$ (ref. 2), so we define a reference $\text{SB}_{\lambda,0} \equiv \text{SB}_0/8\text{\AA} = 80\text{ LU } \text{\AA}^{-1}$. More recent empirical constraints³⁶ are consistent with at least a factor of two lower, $\text{SB}_{\lambda,0,e} = 40\text{ LU } \text{\AA}^{-1}$.

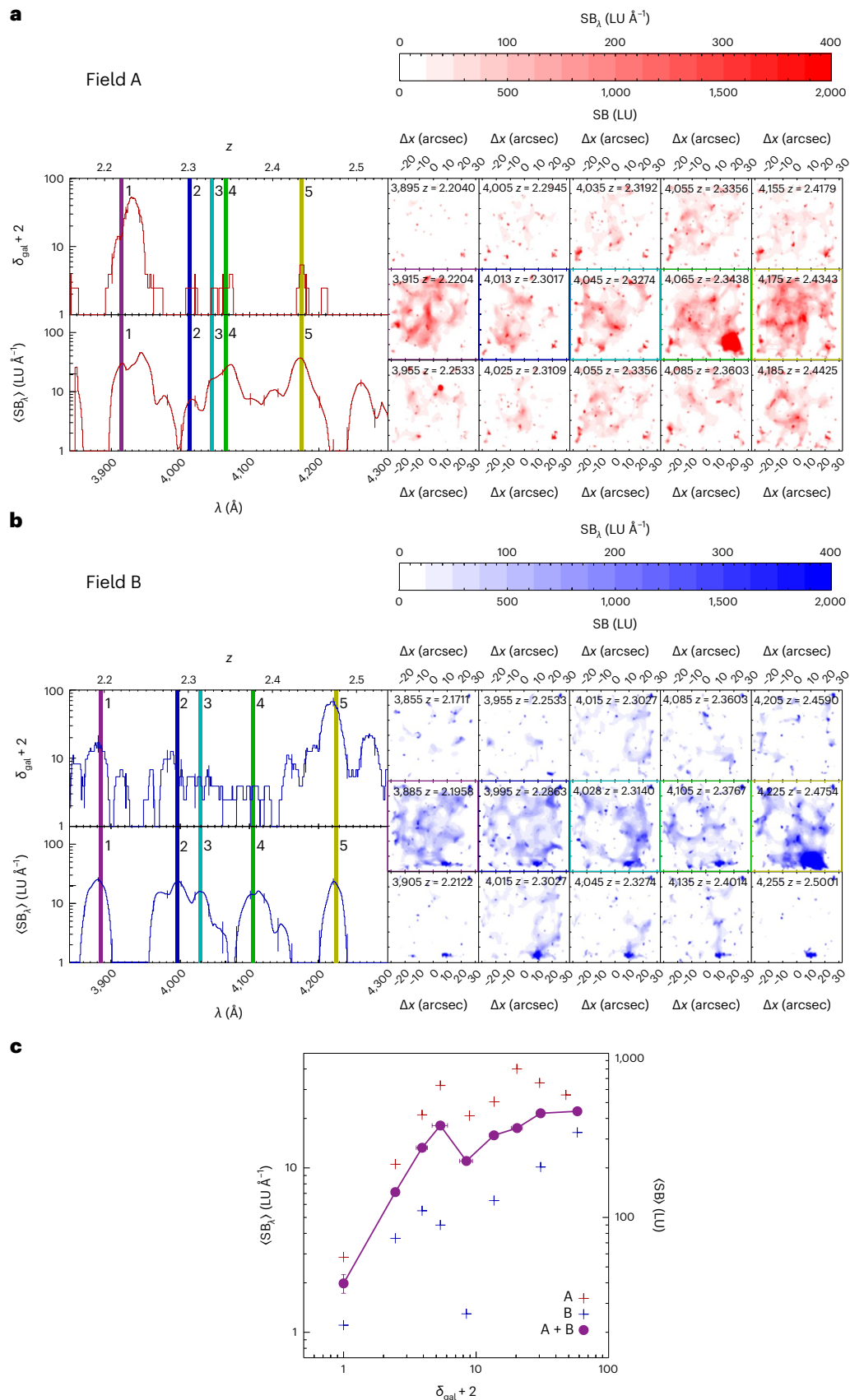
The volume density statistics reveal that the emission is pervasive. We present the distribution of voxel surface brightnesses in Fig. 2a. Each voxel in the smoothed data cube is coded with the smoothing kernel size at which the voxel intensity was detected (Methods). Figure 2a shows that larger smoothing kernels reach lower surface brightnesses. We employed a data cube source detection algorithm to identify discrete sources (Methods). As lower-surface-brightness thresholds and/or smoothing thresholds are reached, extended, low-surface-brightness emission merges distinct higher-surface-brightness sources, thus reducing the source count but increasing the maximum source flux. Emission from the cosmic web detected at very low thresholds would percolate to a small number of sources. As a compromise, recognizing the typical size of brighter emission regions, we choose a maximum smoothing kernel size of full-width at half-maximum (FWHM) $\Delta = 11.3''$. In Fig. 2b, the distribution of source SB and SB $_{\lambda}$ is consistent with the reference values SB_0 and $\text{SB}_{\lambda,0}$. In Fig. 2c, we compare the comoving number density of sources with areas exceeding 3 arcsec^2 , 10 arcsec^2 and 20 arcsec^2 to the predictions of ref. 2, finding agreement. Finally, we show the luminosity function of sources with areas exceeding 3 arcsec^2 and 10 arcsec^2 to that of ref. 37, showing acceptable agreement given the differences in detection methodology. Reference³⁷ also estimated a covering fraction $f \approx 0.2\text{--}1$, which compares well with the covering fraction that can be calculated from Fig. 2a of $f \approx 0.1\text{--}0.3$.

Fig. 1 | The emission Ly α forest at multiple redshifts in two fields. **a**, Field A. Top left: redshift distribution of galaxy overdensities δ_{gal} within radius $8\text{ h}^{-1}\text{ cMpc}$ spheres. As $\delta_{\text{gal}} \geq -1$, we plot $\delta_{\text{gal}} + 2$ on the log scale. Bottom left: redshift distribution of Ly α emission average surface brightness (SB $_{\lambda}$) in $\pm 8\text{ h}^{-1}\text{ cMpc}$ slices. The 1σ root-mean-square error bars (error bars in all figures give standard deviations from the mean based on error propagation through the smoothed data cube, 1σ) are given for both distributions, based on Poisson errors for the number of redshift sample galaxies in each spherical volume ($(\delta_{\text{gal}} + 1) \times 0.7$, as the mean galaxy number per $8\text{ h}^{-1}\text{ cMpc}$ sphere is 0.7), and not including systematic errors. The coloured stripes indicate the data cube slices illustrated in the right panels. Right: each slice exhibiting strong, extended emission is shown in the middle row, with the top and bottom rows showing the nearby low-emission

slices for comparison. The slices are $5\text{\AA}$ wide and the colour bar gives equivalent SB (LU) and SB $_{\lambda}$ ($\text{LU } \text{\AA}^{-1}$). Detected emission must exceed an SNR threshold of 3 (Methods). **b**, The same plots as in **a**, but for field B. **c**, Mean surface brightness versus galaxy overdensity in $r = 8\text{ h}^{-1}\text{ cMpc}$ spheres for field A (red) and field B (blue) and the average (purple line), for which standard root-mean-square error bars are also shown. $\langle \text{SB} \rangle = \langle \text{SB}_{\lambda} \rangle \Delta\lambda$ for a $16\text{ h}^{-1}\text{ cMpc}$ slice. The average surface brightness is well correlated with the galaxy overdensity, with possible saturation for $\delta_{\text{gal}} \gg 10$. The latter could be due to radiative transfer effects or other nonlinearities in the emission process or cosmic variance. Small constant offsets were subtracted from $\langle \text{SB}_{\lambda} \rangle$ to enable log-log plotting of variation with overdensity.

The spectra reinforce the conclusion that the emission is pervasive. In Fig. 3, we show 2.5-Å-wide narrow-band images at five redshifts. We show full spectra in five 5.2×5.2 arcsec² regions

indicated in the image panels. These spectra show many statistically significant emission peaks. Many of these peaks correspond to peaks in the galaxy redshift distribution, as described in the caption.



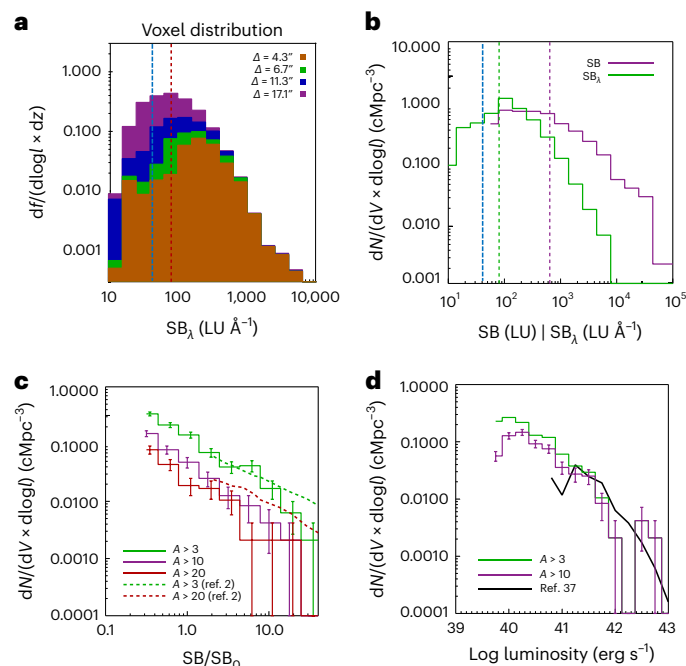


Fig. 2 | Intensity statistics for combined fields A and B. **a**, Differential fraction of voxels with a given intensity in $\text{LU } \text{\AA}^{-1}$ at different smoothing thresholds, given as the FWHM of the Gaussian kernel. Covering fraction per unit redshift is the integral of this, roughly 0.1 for $\Delta = 11.3''$ and 0.25 for $\Delta = 17.1''$. Vertical dotted lines show reference $\text{SB}_{\text{A},0}$ (red) and $\text{SB}_{\text{A},0,e}$ (blue). **b**, Distribution of mean source intensity in LU and in $\text{LU } \text{\AA}^{-1}$. Vertical lines show predicted typical optically thick source intensity SB_0 (purple), $\text{SB}_{0,\text{A}}$ (green)³² and $\text{SB}_{0,\text{A},e}$ (blue)³⁶ in LU for unboosted UV ionizing background at $z = 2.35$, as discussed in text. Thin lines are log-normal fits (Methods). **c**, Source comoving volume density versus intensity relative to SB_0 , for minimum source areas $A > 3, 10, 20 \text{ arcsec}^2$. Dashed lines show predictions from ref. 2. **d**, Source differential luminosity function for two minimum source areas, compared with ref. 37. All plots assume all sources are Ly α and plots **b–d** use a smoothing threshold of $\Delta = 11.3''$. Error bars are 1σ .

We compare the emission spectra to an inverted QSO absorption-line (QAL) spectrum. We have taken a continuum corrected spectrum of a randomly selected QSO J094253–110426 in the same redshift range (f_{QAL}), convolved it with a spectral kernel comparable to the KCWI data, inverted it ($1 - f_{\text{QAL}}$), and multiplied it by 100 to put it on a similar scale as the emission spectrum. The emission and absorption spectra look quite similar. Because of the high density of emission lines and the likeness to the QSO absorption-line distribution, we dub the emission an ‘emission Lyman- α forest’ (ELAF) and refer to the absorption system as the absorption Lyman- α forest (ALAF). We further explore their relationship below. A QSO is present in field B at the overdensity redshift. We performed a cross-correlation analysis between the low-spectral-resolution ALAF present in that spectrum and the field B ELAF. A positive cross-correlation signal is detected (Methods and Extended Data Fig. 4).

The ELAF is filamentary. We have used a subspace-constrained mean shift algorithm^{38,39} to identify and trace filaments in narrow-band images. We show one example in Fig. 3b and more in Extended Data Fig. 5. Galaxies in the field of view with known redshifts are seen to be embedded in and connected by filaments at corresponding redshifts. We also explored the variation in filament properties with SB_λ threshold, and find that the emission becomes less filamentary as the threshold is raised (Extended Data Fig. 6). Median filament transverse widths are $\sim 40|80|120 \text{ pkpc}$ at thresholds $\text{SB}_\lambda > 200|100|30 \text{ LU } \text{\AA}^{-1}$, in the expected range⁴⁰. As we report below, as the surface brightness threshold is lowered the filaments lengthen and a single connected source is revealed to occupy a substantial fraction of the voxels, a sign of percolation.

The observation of galaxies embedded in and connected by filaments (Fig. 3b and Extended Data Figs. 5–7), and the correlation of emission with galaxy-traced overdensities (Fig. 1) suggests that an emission–galaxy cross-correlation signal is measurable. We have calculated the two-dimensional (2D) and one-dimensional (1D) cross-correlation functions between the Ly α emission and the galaxies, ξ_{ge} . We use the COSMOS spectroscopic catalogue⁴¹ for the entire field. See Methods for our approach and Fig. 4 for the results. The emission bias factor b_e can be derived assuming $\xi_{\text{ge}} = b_e b_g \xi_{\text{dm}}$, where b_g is the galaxy bias factor⁴² and ξ_{dm} is the dark matter correlation function, giving $b_e \approx 1$ (based on estimated errors $1^{+0.5}_{-0.4}$), and comparable to dark matter at scales of $r \approx 1 h^{-1} \text{ cMpc}$. Thus, a galaxy–emission cross-correlation signal is detected for $r < 10 h^{-1} \text{ cMpc}$, with slope similar to the galaxy correlation function and an amplitude close to that expected for dark matter. Note that a cross-correlation with emission is far less sensitive to galaxy redshift selection effects than an autocorrelation.

As we discuss in Supplementary Information, the emission is considerably more extended and the surface brightness an order of magnitude lower than that typically observed in Ly α blobs^{10,12–20,22,43} and highly boosted emission near active galactic nuclei^{6–9,11}.

The emission is not likely due to the smoothed summation of numerous faint point sources or Ly α halos (LAHs). We conclude based on two methodologies that LAHs fail by more than an order of magnitude to explain the observed emission distributions (Supplementary Information, Extended Data Fig. 8 and below). Point sources are always detected by the smoothing algorithm as soon as they exceed the signal-to-noise ratio (SNR) threshold. The SNR is always maximum when the smoothing kernel is minimum as a larger kernel incorporates more sky noise. Thus, unless they exceed the confusion limit they cannot explain the observed extended emission.

The MUSE Extremely Deep Field (EDF) exhibited five extended emission sources (two filamentary) with confidence $>95\%$ in overdensities, with surface brightnesses in the same range we are detecting²⁶. The detection significance of the extended regions was limited by an unknown source of noise, as we discuss in Supplementary Information, and depending on how this noise is factored into the comparison the EDF SNR is somewhat lower than or comparable to the KCWI observations (see Supplementary Table 2 and the discussion in Supplementary Information). The authors did not report a pervasive ELAF correlated with galaxies that has statistical properties analogous to the ALAF.

In this preliminary study, we do not discriminate between radiative recombination and photon scattering origins of the emission. Recent modelling^{3–5}, and other radiative transfer simulations^{44,45}, suggests Ly α resonant scattering in overlapping halos, particularly near massive ones, could produce an extended emission plateau reaching $\sim 100\text{--}1,000 \text{ LU}$, while at the same time articulating filaments of the cosmic web. Because the densities are predicted to be low, radiative recombination is unlikely to account for the faintest emission, which, as we discuss below, has statistical properties similar to the absorption Ly α forest.

An additional contribution^{46,47} may be relevant at the lowest surface brightness levels, continuum photon pumping (CPP). Using the continuum metagalactic background at Ly α in the rest frame⁴⁸, we can calculate the rate of CPP as measured at Earth. A perfect H I absorber with a width and equivalent width of $1 \text{ \AA}$ (that is, a perfect rectangular absorption profile) will produce a scattered surface brightness of $\text{SB}_{\text{CPP},0} \approx 166 \text{ LU } \text{\AA}^{-1}$ at $z = 2.35$. While the radiative transfer will be more complex, a first approximation can be obtained from the continuum normalized QSO absorption-line spectrum f_{QAL} with $\text{SB}_{\text{CPP}} \approx \text{SB}_{\text{CPP},0}(1 - f_{\text{QAL}})$. In Fig. 3l, we use a random QSO with high-spectral-resolution data to illustrate the typical appearance that such an emission-line spectrum might take at the KCWI spectral resolution. We used a slightly lower normalization, $\text{SB}_{\text{CPP},0} = 100 \text{ LU } \text{\AA}^{-1}$. There is evidence that the fainter emission features could be explained by this process, and that this could provide emission sensitivity to much of

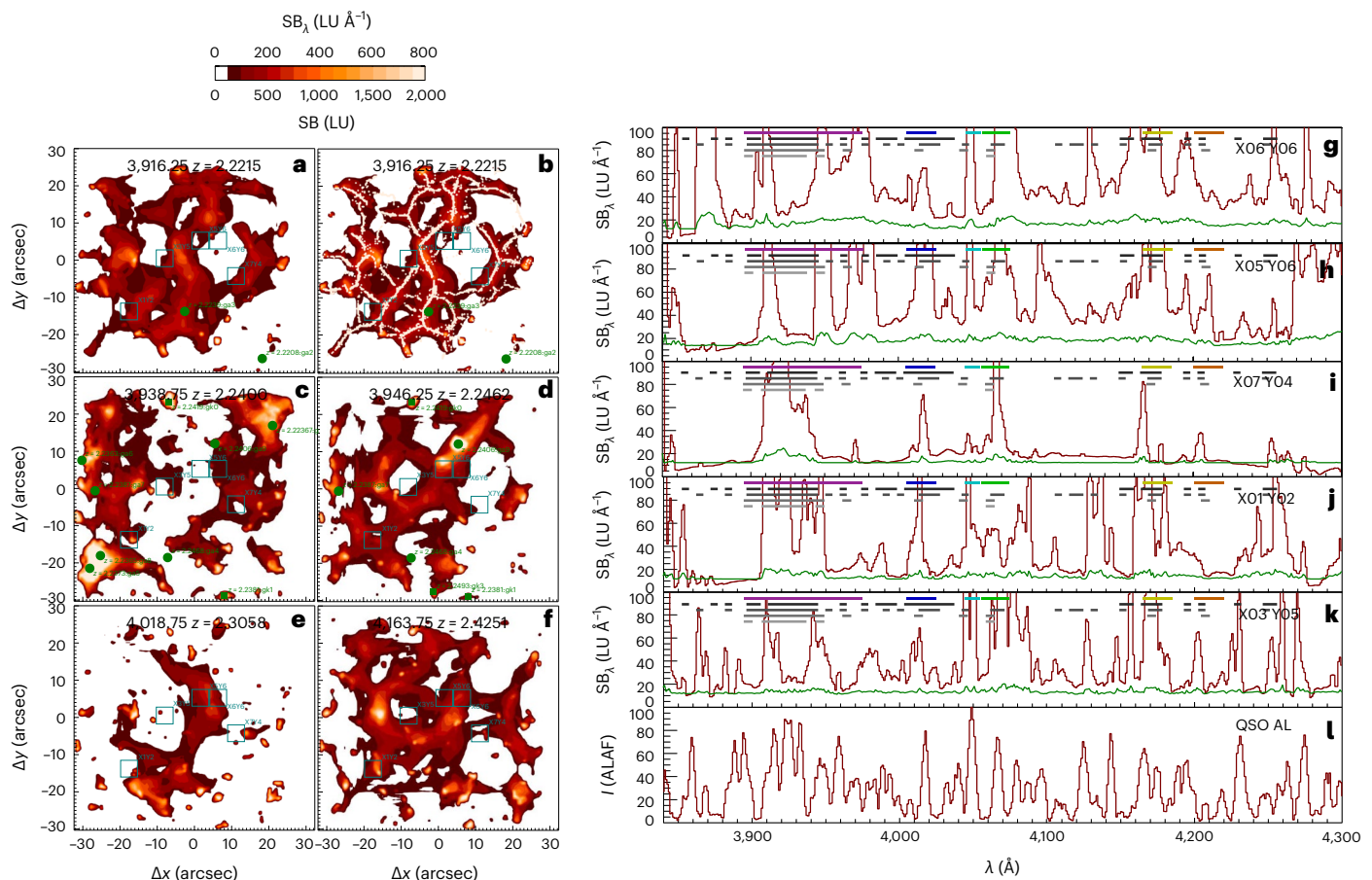


Fig. 3 | The emission Ly- α forest. **a–f**, Illustrative narrow-band images and spectra showing the nature of the observed Ly α emission. Green boxes locate the regions from the spectra in which **g–l** are extracted. Green dots show locations and redshifts of galaxies near the narrow-band image redshift. Codes indicate redshift and galaxy identification. **a–f**, Narrow-band images, width 2.5 Å, with central wavelengths and redshifts given. **a**, The narrow-band image at 3,916.25 Å. **b**, The results of filament finding algorithm in white (see text). On each image, known galaxies with redshifts within 10 Å of the central redshift are shown with green circles. Detected emission must exceed an SNR threshold of 3 (Methods). Spectral extraction regions, each 5.2 arcsec square, are shown in cyan. **c**, The narrow-band image at 3,938.75 Å. **d**, The narrow-band image at 3,946.25 Å. **e**, The narrow-band image at 4,018.75 Å. **f**, The narrow band image at 4,163.65 Å. **g–k**, The corresponding spectra. For each spectral panel, the spectrum is in red and the 1 σ

error is in green (including estimated average systematic errors). Note that errors are correlated (spatially and spectrally) due to adaptive smoothing. Full scale has been chosen to be 100 LU Å⁻¹ to show fainter emission features and make a comparison with the ALAF. The coloured bars correspond to the overdensities shown in Fig. 1. Below these, we show bars that indicate the location of galaxies within (5, 10, 15, 20) cMpc with number counts exceeding (1, 1, 1, 2), proceeding from bottom (grey bar) to top (black bar), respectively (Supplementary Fig. 2). **l**, A QSO absorption-line (AL) spectrum converted to emission by subtraction from one, and multiplied by 100 to make the scales identical to the ELAF spectra. The spectrum is QSO J094253–110426 obtained from the UVES Spectral Quasar Absorption Database (SQUAD)⁶⁶, and it has been smoothed with a resolution kernel comparable to the KCWI spectra.

the ALAF. We note that this process relies on resolving out the sources of continuum emission as discussed in Methods and illustrated in Extended Data Fig. 9. The equivalent of Fig. 3*l*, which is meant to be illustrative, is given for field B and the field B QSO in Extended Data Fig. 10.

Because of the pervasiveness of the ELAF (Fig. 2*a*), its resemblance to the ALAF (Fig. 3*l*), the ELAF/ALAF cross-correlation signal detected in the one field B QSO, the correlation with galaxies and galaxy-traced overdensities with low bias (Figs. 1 and 4), and the possible contribution of CPP, we explore here the hypothesis that the ELAF and ALAF are physically related. We begin by comparing in Fig. 5 the observed ELAF surface brightness distribution to that of the ALAF. To do this, we use the fact that typical ALAF lines are roughly 1 Å wide, which is comparable to our emission voxel depth. The ELAF covering factor/volume distribution is given with the thick purple line. The distribution can also be expressed as a volume filling fraction as in Fig. 2*a*, and the scale for this is given on the left. If we assume that the ELAF and ALAF are physically related, and equate the ELAF and ALAF distributions, we can generate a column density from the ALAF column density-number

count distribution^{40,49}, which is shown on the right axis of the main box. Typical designations for column density ranges are shown with vertical blue/green bars showing Ly α forest (LAF), Lyman limit systems (LLS) and subdamped Ly α absorbers (sDLA). These column densities roughly correspond to overdensities⁵⁰, shown on the rightmost vertical axis. Typical ranges for cosmic web morphological features, knots, filaments and walls are related both to overdensity⁵¹ and volume filling factor⁵².

We now consider emission morphology variations with threshold surface brightness. The red line shows the volume distribution for the single source with maximum voxel count versus the isophote threshold, showing a demonstration of percolation at $SB_{\lambda} \lesssim 150$ LU Å⁻¹. The blue line shows η_{fil} , which we define as the ratio of the filament length to width, again showing a large increase in elongation for $SB_{\lambda} < 800$ LU Å⁻¹. Both of these metrics suggest the onset of percolation at $SB_{\lambda} \lesssim 300$ LU Å⁻¹, as knots become connected by filaments. These indicators of filamentary morphology are well aligned with the expectations from the ALAF overdensity and ELAF volume fraction expectations indicated in the figure, which illustrate that filaments

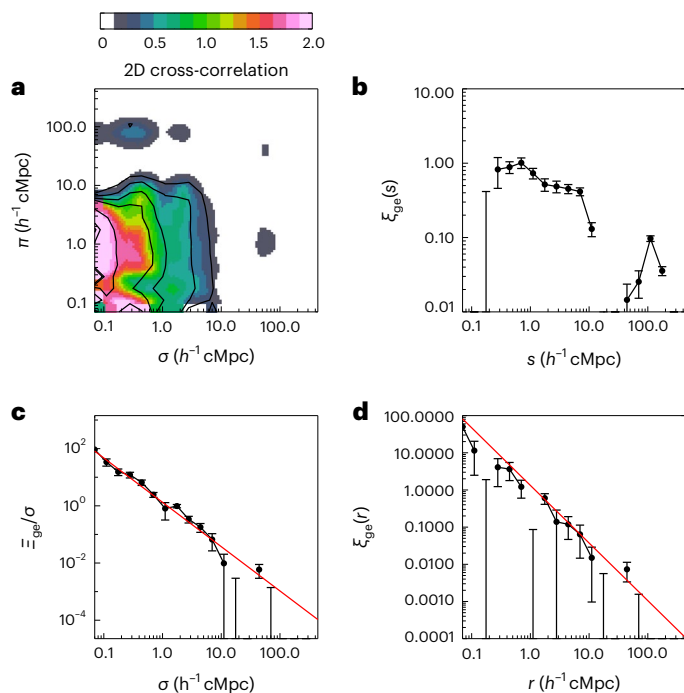


Fig. 4 | Spatial galaxy–emission cross-correlation functions in 1D and 2D for combined A and B fields. **a**, The 2D spatial cross-correlation function $\xi_{\text{ge}}(\sigma, \pi)$ versus comoving (σ -transverse, π -radial) distance. The SNR in the region with detected signal is ~ 3 – 5 (compare Extended Data Fig. 3). Contours are shown at $\xi_{\text{ge}} = 0.25, 0.5, 0.75, 1.0, 2.0, 4.0, 5.0$. We see the usual redshift-space distortions—elongation along the line of sight at small transverse separations and compression along the line of sight at larger separations, indicating inflow⁶⁷. **b**, The 1D redshift correlation function⁵⁹. **c**, The 1D projected redshift distribution $\xi_{\text{ge}}(\sigma)/\sigma$ (ref. 59), compared with a power law with $\gamma = 1.55$ and $r_0 = 1.2 \text{ h}^{-1} \text{ cMpc}$ (ref. 68). **d**, Derived real-space 1D correlation function $\xi_{\text{ge}}(r)$ (ref. 59), compared with a power law with $\gamma = 1.5$ and $r_0 = 0.8 \text{ h}^{-1} \text{ cMpc}$ (Methods). In each case, the corresponding emission–galaxy cross-correlation function can be obtained by multiplying ξ by the average $\langle \text{SB}_{\lambda} \rangle = 20 \text{ LU } \text{\AA}^{-1}$. 1σ error bars shown.

and walls occupy substantial volume fractions and are associated with column densities and inferred overdensities in the surface brightness range $100 \text{ LU } \text{\AA}^{-1} < \text{SB}_{\lambda} < 500 \text{ LU } \text{\AA}^{-1}$.

The volume distribution predicted from the MUSE LAH observations are given by the green line²⁵, falling 1–1.5 orders of magnitude below the observations. The distribution for our conservative simulation of smeared LAHs is given by the grey line, again falling far below the ELAF distribution except at bright surface brightnesses where we expect to have detected LAHs.

Thus we have presented five lines of evidence that the ELAF and ALAF are physically related. These are (1) from comparison of ELAF and ALAF spectra, (2) the large volume/covering fraction of the ELAF analogous to that of the ALAF at lower column densities, (3) the presence of a detectable cross-correlation between the ELAF and ALAF in a field BQSO, (4) the presence of cross-correlation between galaxies and the ELAF, and (5) CPP as a process that could be important in producing a detectable emission signal from the ALAF. The ALAF accounts for the majority of baryons at $z = 2.35$ (Fig. 5), and if related the ELAF may also trace a substantial baryon and dark matter fraction. If so, the ELAF represents an important observational tool for studying the structure and evolution of the cosmic web of baryons and dark matter, and its relationship to and co-evolution with galaxies. Future observations will go deeper, provide similar measurements at higher redshift for comparison, provide higher spectral resolution for further comparison with the ALAF, and increase the field of view to reduce cosmic variance.

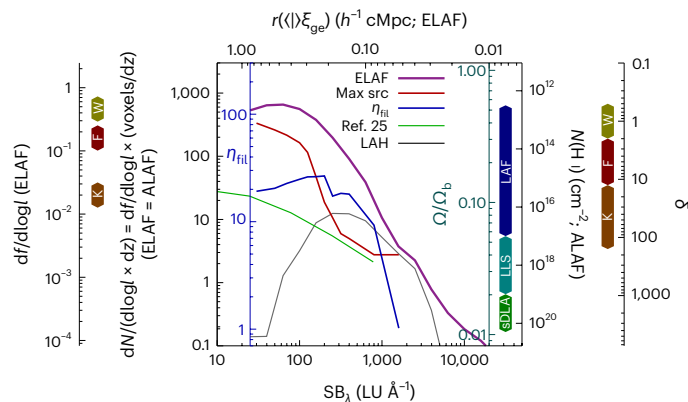


Fig. 5 | Synoptic view of the IGM in ELAF and ALAF. The observed ELAF covering factor/volume distribution (purple line) versus surface brightness $\text{SB}_{\lambda} = I$, also expressed as a volume filling fraction on the leftmost scale. Corresponding ALAF column density ranges shown for LAF, LLS and sDLA. Overdensity corresponding to $N(\text{H I})$ shown on rightmost axis. Cosmic web features knots (K), filaments (F) and walls (W) are shown on right and left axis corresponding to overdensity and volume fraction, respectively. Dark cyan axis on right shows estimated baryon fraction versus ALAF column density⁶⁹. Median volume fraction of maximum source (max src; red line) shows percolation at $\text{SB}_{\lambda} \lesssim 150 \text{ LU } \text{\AA}^{-1}$. Filament parameter (η_{fil} ; blue line, left axis) steeply rises for $\text{SB}_{\lambda} < 800 \text{ LU } \text{\AA}^{-1}$. Comparison with two estimates of LAH contribution shown with green (ref. 25) and grey (LAH) lines, both falling far below the ELAF distribution (Supplementary Information).

Methods

Numerical simulation and forward modelling

We have used a cosmological hydrodynamical simulation²⁹ to generate a forward-modelled simulated KCWI data cube to determine detectability and optimize our observing strategy. We perform cosmological zoom-in simulations of the standard lambda-dominated cold dark matter model with the Adaptive Mesh Refinement (AMR) Eulerian hydro code, Enzo. One is centred on a cluster of total mass of $\sim 3 \times 10^{14} M_{\odot}$ ('C' box) with overdensity parameter $\delta = 2$ and the other centred on a void region ('V' box) ($\delta = 0.5$). The zoom-in region for the C run has a size of $21 \times 24 \times 20 \text{ h}^{-3} \text{ Mpc}^3$ and that for the V run is $31 \times 31 \times 35 \text{ h}^{-3} \text{ Mpc}^3$. The maximum resolution in the zoom-in box is $460 \text{ h}^{-1} \text{ pc}$, which is about two orders of magnitude smaller than the Jeans scale of photoheated IGM at a density 100 times the mean density. The lowest spatial resolution in the zoom-in region initially before refinement is $177 \text{ h}^{-1} \text{ kpc}$, compared with the Jeans scale of $400 \text{ h}^{-1} \text{ kpc}$ of the photoheated IGM at mean density. The C and V boxes are two separate zoom regions with a $120 \text{ h}^{-1} \text{ Mpc}$ periodic box. We used CLOUDY³⁰ to convert each simulation voxel to ionization fraction and Ly α surface brightness (including radiative recombination and collisional excitation), and we generated a high-resolution ($\Delta\lambda = 0.1 \text{\AA}$ (observed frame)) data cube by converting each simulation voxel to a central velocity based on the gas z velocity and applying an optically thick line profile³¹ based on the Ly α optical depth (determined by the H I column density in each cell). Peak separations (line widths) are comparable to peculiar velocities. We account for the impact of a boosted metagalactic UV background to the ionization state and emissivity of each region. Our boost factor refers to the input boost (not effective boost), but the simulations correctly predict the somewhat lower emission enhancement expected². Finally, we used the measured sky spectrum at the W. M. Keck Observatory and the KCWI instrument parameters to generate a simulated observed data cube, assuming 4 h on-source (and 4 h background) exposures.

Extended Data Fig. 1 shows the simulations and the observation strategy. Our survey consists of 6 KCWI large slicer pointings in a roughly $60 \times 60 \text{ arcsec}^2$ source field mosaic using the BL grating

centred at $4,060 \text{ \AA}$ ($z = 2.3$, $\Delta\lambda \approx 500 \text{ \AA}$) (roughly $1.5 \times 1.5 \times 500 \text{ cMpc}^3$). BL is optimal for the typical predicted linewidth as it provides maximal bandpass and its spectral resolution is well matched to the predicted line widths. With the NAS approach, there is also a $60 \times 60 \text{ arcsec}^2$ background field, chosen in this sim to be 12 arcmin away (5 pMpc or 20 cMpc transverse distance). The figure shows emission regions appearing in the source cube as positive (red in the narrow-band images) and those appearing in the background cube as negative (blue in the narrow-band images). Four cases are shown, where b is the ratio of the local ionizing background to the average metagalactic ionizing background: (1) void, boost $b = 1$; (2) cluster, boost $b = 1$; (3) void, boost $b = 10$; (4) cluster, boost $b = 10$. The upper spectrum and images for 4 redshift slices show the simulated intensity (no noise), and the lower spectrum and images show the binned spectrum (full $60 \times 60 \text{ arcsec}^2$ field) and adaptively smoothed images^{7,21}. The spectra are generated by summing the full field of view.

Several things are clear: (1) Ly α emission from the cosmic web is probably detectable even in the case of void, $b = 1$, with more than ~ 10 regions detectable in the source and background region; (2) source and background regions will in general be distinguishable and will not in all but rare cases ‘cancel out’; (3) Ly α in the cluster simulation is easily detectable even with $b = 1$, and we can guarantee selection of an overdensity by choosing a field with an overdensity of galaxies; (4) in the case of regions with a higher boost factor of $b = 10$, emission is detectable (even from small regions) and detailed maps are possible.

Using this simulation, we determined the range of acceptable separations that statistically minimize confusion. In particular, for the simulation a separation of 12 arcmin produces a confusion percentage of $<3\%$ for $SB_{\lambda} > 30 \text{ LU \AA}^{-1}$ and $<5\%$ for $SB_{\lambda} > 10 \text{ LU \AA}^{-1}$. This percentage is the fraction of emission from detectable voxels that is reduced by subtraction of emission at a similar position and redshift. As our final field separation is double this, 24 arcmin, using a correlation function of $\xi_e \approx r^{-\gamma}$ with $\gamma = 1.5$ is the correlation function power law negative index, the expected confusion is $<1\%$ and $<1.5\%$, respectively.

We note that these simulations are not state of the art, as cooling models, the metagalactic UV background, self-shielding analysis and IGM enrichment prescriptions are not current and active galactic nuclei are missing (although implicitly included with varying boost factors). The Ly α predictions are based on CLOUDY, and are limited by the uncertainties in the local temperature, density and radiation field. However, there have not been major changes in assumptions, such as metagalactic UV background, that would affect the results in this application in any notable way. We are not using the simulations to infer detailed physical parameters, rather only to demonstrate that the overlap fraction using DFS is very modest at initio. As the fraction is small, the relative accuracy of this conclusion is not critical, and the measured statistics confirm that the overlap fraction should be small. Even more recent simulations are challenged to predict Ly α emission from the cosmic web because of the unconstrained subgrid physics involved (for example, clumping, multiphase structure, dust and radiative transfer).

Target selection

To increase our chance of Ly α detection from the cosmic web, we perform our observations on two overdense regions in a deep field. For this purpose, we use the COSMOS field⁵³, which has a plethora of ancillary photometric and spectroscopic data over a broad range of wavelengths. As we plan to use the NAS mode and only one-third of the CCD is available in this mode, we need to wisely select the proper diffraction grating, central wavelength and the redshift separation of the two overdensities to be able to detect the potential redshifted Ly α emission. Given this, we use the large slicer and the low-resolution BL grating (central wavelength $\sim 4,060 \text{ \AA}$, resolving power ~ 900), and perform our observations focused on two spectroscopically confirmed overdensities at $z \approx 2.24$ and $z \approx 2.45$ in the COSMOS, named A and B fields, respectively. The $z \approx 2.24$ overdensity is a confirmed protocluster (dubbed CC2.2)

with an estimated total mass of $\sim (1\text{--}2) \times 10^{14} M_{\odot}$, which is expected to collapse to a Coma-type cluster at $z = 0$ (ref. 54). The $z \approx 2.45$ structure is a massive multi-component protocluster (named Hyperion) with an estimated total mass of $\sim 4.8 \times 10^{15} M_{\odot}$ and at least 7 spectroscopically confirmed components in a range of masses $\sim (0.1\text{--}2.7) \times 10^{14} M_{\odot}$ (ref. 55). We design the A (and B) field as an $\sim 60 \times 60 \text{ arcsec}^2$ region that is composed of 6 overlapping KCWI large slicers (Extended Data Fig. 2). We place the field A at coordinates right ascension 150.2074° and declination $+2.0189^{\circ}$, which corresponds to a dense subcomponent of the CC2.2 protocluster. For field B, we place it on the most massive component of the Hyperion protocluster at coordinates right ascension 150.0937° and declination $+2.4049^{\circ}$.

Observations

The observations were conducted over 17 nights from 2019B (2019B is the second observing semester in 2019) to 2021A (the first observing semester in 2021) with the Keck/KCWI integral field spectrograph²⁷ under good observing conditions with the typical seeing of $\sim 1.0 \text{ arcsec}$. We used the NAS mode with the large slicer, low-resolution BL grating and a central wavelength of $4,060 \text{ \AA}$. Observations were done on two fields, A and B, each covering an $\sim 60 \times 60 \text{ arcsec}^2$ region composed of 6 overlapping tiles of the KCWI large slicer at two different position angles of PA = 0° and PA = 90° (Extended Data Fig. 2). Each tile was observed for 2 min, shuffling back and forth between fields A and B for a cycle of 20 min, before moving to the next tile.

We reduced the data using the interactive data language (IDL) version of the KCWI data reduction pipeline *kderp* (<https://github.com/Keck-DataReductionPipelines/KcwiDRP>). For each image, we subtracted the bias and overscan, removed cosmic-rays and corrected for dark current and scattered light. We performed sky subtraction by taking the difference of the two panels in each NAS image. The pipeline converted raw 2D detector images into right ascension–declination and wavelength corrected three-dimensional (3D) data cubes. For each night of observations, we used twilight flats for flat field correction. The pipeline was corrected for differential atmospheric refraction and used standard star observations for flux calibration. We used the image and variance cubes in our analysis, inspecting several intermediate data pipeline products to confirm data quality and processing integrity. Our inspection of the PSF of bright sources in individual frames caused us to discard two images with a compromised PSF due to target reacquisition issues during the NAS exposure. Overall, each tile was observed for a total exposure time of 3.00–3.67 h with a typical depth of 3.33 h, as summarized in Supplementary Table 1.

Sky subtraction using NAS was optimized by allowing a slight difference in normalization of the source and background field, of the order of 1–2%. This was measured using a cross-correlation of each exposure total spectrum difference with the sky spectrum. Using this, we can say that low-order residuals correlated with sky are $<0.07\%$ or $30 \text{ LU \AA}^{-1}$. Such residuals would not in general produce emission-line features. We attribute these differences to either absolute differences in zodiacal light at the two locations or differences in the diffuse galactic light contribution. Using the prominent Ca H/K lines in the sky spectrum from zodiacal light, we can use the lack of detection of these lines in either A or B field to estimate the sky subtraction precision. We find no (3σ) signal associated with these lines $>10 \text{ LU \AA}^{-1}$, which implies given the line contrast of ~ 0.5 means that the 1σ sky subtraction error is $<0.02\%$.

Data analysis

Variance cubes were generated using the calibrated gain curves and recorded photon count per pixel. Pixel variance is reduced when the image is reformatted to generate square 0.291 arcsec pixels, by the stretch factor. Variances are summed (voxel by voxel) during data cube coaddition, and field A and B variances are added when the individual exposures are subtracted. Error propagation is discussed at length in ref. 11. The variance cube is corrected for resampling (effectively by

multiplication by the expansion parameter, which is approximately 4.0) so that the correct variance can later be calculated when the data are smoothed by the adaptive smoothing algorithm. The variance, illustrated in Extended Data Fig. 3d, is dominated by sky background noise, with $\sim 1,200$ electrons per detection pixel ($0.291'' \times 1.35'' \times 1 \text{ \AA}$) in the final mosaic. Read noise of 2.5 electrons root mean square per 20 min exposure, binned 2×2 , adds an additional effective contribution 75 electrons per detection pixel, leading to an increase over sky of a factor of 1.03. The systematic sky subtraction error is $<0.02\%$ of sky (1σ) or 0.2 electrons per pixel in these units. We note here that NAS subtraction increases the first two of these noise sources by $\sqrt{2}$. There are two main additional sources of systematic error. The first is produced by the DFS approach. Redshifts and regions of field A with statistically significant emission will be subtracted from the same redshifts and regions of field B. We estimated an overlap fraction of 5% based on Fig. 2a at $SB_{\lambda} \approx 100 \text{ LU } \text{\AA}^{-1}$, resulting in an average error of $\sim 5 \text{ LU } \text{\AA}^{-1}$ albeit with large deviations in bright overdensities. The second source of error is the source subtraction and sky substitution process. We estimate based on varying the source magnitude parameters that this introduces a typical error of $<5 \text{ LU } \text{\AA}^{-1}$ when averaged over the entire field of view.

Astrometry on each individual exposure was performed using known sources in the fields³³. Note that the A and B coadds are slightly different because of small offsets in the relative A–B NAS positions between each NAS exposure. The A field (A–B) is coaligned to A sources, and the B field (B–A) is coaligned to B sources. These differences amounted to typically 0.5".

After sky subtraction and astrometric correction, the data cubes are coadded, along with variance and exposure cubes. The final flux and variance cubes are exposure weighted, and the variance cube is used to determine the SNR in the adaptive smoothing algorithm. An exposure cube and variance cube slice are shown in Extended Data Fig. 3c,d. Final point source PSFs are ~ 1.4 arcsec (FWHM).

Bright sources in both the A and B fields could impact the derived diffuse emission properties. For field A, bright sources in A produce a positive excursion, in B a negative one. The adaptive smoothing algorithm (see next paragraph) can accommodate positive and negative sources, which are detected and removed from the image and do not produce extended residuals. We run the negative detection loop first to remove sources that would bias the positive detection loop. However, to ensure the minimization of residuals and artefacts, we perform PSF subtraction of all sources brighter than $m_b < 24$. We experimented with two methods to deal with fainter sources, either ignoring them or performing background substitution in the affected voxels. We find that these produce nearly identical results for the diffuse emission, which gives confidence that the presence of modest numbers of faint continuum sources does not add a statistically significant spurious signal or noise component, and we baseline sky background substitution. Background substitution replaces a generous central 3 arcsec region centred on the source with bordering sky. Occasionally more extended sources have larger substitution zones. These ensure that the contribution of any residual halos around these faint galaxies is minimized. We perform this for all galaxies with photometric redshifts in the range $2.1 < z_{\text{phot}} < 3.6$, to minimize the contribution of Ly α from faint galaxies, and to enable the unambiguous detection of CPP (see discussion below). This approach allows the production of an emission cube without the loss of regions with sources so that a reasonable measurement of emission morphology can be performed.

We use the adaptive smoothing algorithm presented in considerable detail in previous work^{7,11,21,34,35,56}. We summarize the algorithm here. We select an SNR threshold SNR_0 . We start with a cube unsmoothed in wavelength. We smooth spatially each wavelength slice with a Gaussian kernel with 5 pixels ($5 \times 0.29'' = 1.46''$). We smooth the variance cube identically. All voxels with $\text{SNR} > \text{SNR}_0$ are added to the smoothed cube and at the end subtracted from the input cube (after the pass). The smoothing length is increased by 2 pixels and the spatial smoothing

loop is repeated. This continues up to 59 pixels ($17.2''$). Then the input cube and variance cubes are smoothed in wavelength by 2 pixels (boxcar). The spatial cycle is repeated. The wavelength kernel is doubled in repeated cycles, followed by the spatial cycle, until the wavelength kernel reaches 128 pixels. Negative emission is detected first and removed, then positive. The B image (B–A) is performed in the same way. As can be seen from Fig. 1, the algorithm produces fair representations of the input diffuse emission morphology. Adaptive smoothing produces covariance between voxels, and interpretation of extended, low-surface-brightness emission regions must be circumspect. The main impact here is that the source detection threshold (discussed below) must account for the reduced number of independent degrees of freedom in covariant voxels.

We show in Extended Data Fig. 3 a $10 \text{ \AA}$ slice of the raw data cube at $3,915 \text{ \AA}$ smoothed to 20 pixels or 5.8 arcsec, the corresponding adaptively smoothed slice, the exposure time for the mosaic and the variance image in the slice. The adaptively smoothed image is a good representation of the raw data, and avoids smoothing compact sources above the noise threshold, which would be added to the extended emission. Extensive simulation tests were performed in refs. 7,21, which showed that extended emission regions become less patchy (for example, as in Extended Data Fig. 1) and low-surface-brightness contours may shift somewhat as the SNR increases.

Smoothing combined with exquisite NAS subtraction allows us to reach the very low minimum sensitivity limits that permit detection of the general emission. For example, in 4 h, the 1σ sensitivity limit in $1 \times 1 \text{ arcsec}^2$ is $540 \text{ LU } \text{\AA}^{-1}$, and $770 \text{ LU } \text{\AA}^{-1}$ with NAS (due to the additional $\sqrt{2}$ noise. When smoothed over $8 \text{ \AA}$ and $10 \times 10 \text{ arcsec}^2$, this is reduced to $27 \text{ LU } \text{\AA}^{-1}$ (1σ), or $\sim 81 \text{ LU } \text{\AA}^{-1}$ (650 LU for an $8\text{-\AA}$ -wide line) at 3σ , our smoothing threshold. This comports well with our voxel luminosity distribution in Fig. 2a, which shows that we are picking up the lowest-surface-brightness voxels for only the largest smoothing kernels. SNR variations due to exposure time differences in the mosaic will produce a modest fixed pattern modulation in the thresholded adaptively smoothed images. The larger smoothing kernels required to obtain the lowest-surface-brightness detections attenuates this effect. For example, with 30 pixel (8.7 arcsec) smoothing, the variation in the SNR is $\sim 12\%$ over most of the field, except in the masked areas. Small artefacts also occur near some edges.

Extensive scattering measurements of the KCWI before and during commissioning²⁷ established that instrument scattering is very low, and quite diffuse, due almost solely to camera ghosts and grating diffuse scatter. These will produce broad wings for bright in-field sources that translate into very diffuse spectral/spatial features that are not spectrally confined. These fields do not have bright sources in or near the field. Even if they did, these low-level scattering wings would not produce spatially diffuse but spectrally coherent line emission.

In Supplementary Fig. 1, we show a subset of the emission peaks with field A and field B to show the impact of DFS on feature overlap and confusion. This figure is analogous to Extended Data Fig. 1.

Source detection and analysis

Sources are detected in the smoothed cubes by selecting all voxels generated at or below a given smoothing threshold (for example, 9 pixels includes 5, 7 and 9), and then running an iterative 3D association algorithm using a 1 voxel nearest-neighbour threshold. This is repeated until there is no change in the source identification of all the above threshold voxels. We repeat this at every smoothing threshold. Sources are characterized by spatial and spectral moments, and principal component analysis is used to determine the principal axes of sources in 3D, giving the major and minor spatial axes and the axis of maximum wavelength shear. We calculate total brightness, surface brightness SB (in LU) and surface brightness SB_{λ} (in $\text{LU } \text{\AA}^{-1}$), and the line profile moments including skew, kurtosis and a parameter sensitive to double peaks we call b-par (bi-peak parameter). We calculate distributions using the cube

comoving volume accounting for the loss of the masked fraction. We assume all line sources are Ly α following ref. 37. As shown in Fig. 2, there is good agreement between the source surface brightness and luminosity predictions from observations³⁷ and simulations².

Redshift distribution

We used the spectroscopic data from the zCOSMOS survey⁴¹. It consists of two samples: zCOSMOS-bright which is an (*I* band (AB) magnitude) $I_{AB} < 22.5$ catalogue targeting $0.1 < z < 1.2$ galaxies in the COSMOS field, and zCOSMOS-deep which is a colour-selected sample aiming at $1.4 < z < 3.0$ sources, typically with (*K* band (AB) magnitude) $K_{AB} < 23.5$. We used the former sample to verify that low-redshift diffuse emission has no contribution to the Ly α signal. For clustering analysis, we used the latter. In addition, we used the spectroscopic data from ref. 54 for the C2.2 protocluster in field A as well.

We verified that low-redshift, diffuse emission does not contribute to the signal. The galaxy redshift distribution shows a single, large peak at $z = 0.121$. If [O II] 3,727 Å contributes to the diffuse signal, we would expect a corresponding peak at 4,182 Å. But field A has a peak at 4,170 Å and field B at 4,220 Å. Thus we find no evidence that low-redshift, diffuse emission is contributing to the Ly α signal.

We estimate the galaxy overdensity δ_{gal} in a preliminary fashion by taking the number of galaxies within $r = 8 h^{-1} \text{ cMpc}$ spheres, divided by the average number of galaxies in the same spheres in our redshift range. The redshift sample includes the entire COSMOS field, but there are modest modulations in the surface density of spectra over the field on 0.2° scales, but no substantial variations over the $8 h^{-1} \text{ cMpc}$ projected radii from the two field centres. When the field B overdensity is excluded, we find that the bin-to-bin galaxy numbers for the whole COSMOS sample is approximately uniform with $\pm 23\%$ variations in $\Delta z = 0.05$ bins, consistent with large-scale structure plus Poisson fluctuations and suggesting no substantial redshift-dependent selection effects in our range. For this preliminary work, we have therefore not corrected for galaxy selection functions or spectroscopic completeness of the sample. As we discuss below the galaxy–emission cross-correlation should be relatively immune to modest selection function and completeness variations.

We note that several of the peaks in the galaxy overdensity and the Ly α emission correspond to overdensities found by ref. 57 (field A, $z = 2.23$, $z = 2.30$; field B, $z = 2.28$, $z = 2.45$). Although calculated in a different fashion, when corrected for our averaging sphere they give similar results. A portion of the multi-component Hyperion supercluster in field B was earlier identified⁵⁸.

A comparison of the emission and galaxy redshift distribution at various distances from the fields is shown in Supplementary Fig. 2.

Filament (and wall) detection

The observed data cubes were processed without masking to assess emission morphology and search for the presence of filaments, as we discussed above. The extended morphology on scales greater than ~ 5 arcsec should be representative of the emission distribution. We use the subspace-constrained mean shift algorithm^{38,39} on pseudo-narrow-band images to determine filament location. Examples of filament maps are shown in Fig. 3b and Extended Data Figs. 5 and 6. The latter figure shows the dependence on surface brightness isophote threshold. The maps are qualitatively similar to those of ref. 39 (Fig. 4) derived for Sloan Digital Sky Survey galaxies, albeit on linear scales 100 times smaller. We define a filament shape parameter η_{fil} as the ratio of filament length to filament width. The median filament parameter is shown versus the isophote level in Fig. 5 (blue line). There is a strong trend towards increasing η_{fil} with decreasing isophote level, an increase of a factor of 20 from $SB_\lambda = 1,600 \text{ LU } \text{\AA}^{-1}$ to $SB_\lambda = 100 \text{ LU } \text{\AA}^{-1}$, where it then flattens.

Detected filaments range in width from 5 arcsec to 15 arcsec (40–120 pkpc). This is the range of filament widths predicted by our numerical simulation, and are consistent with observational

constraints on the spatial coherence of the Ly α forest⁴⁰, which is likely to probe the filaments. The two statistically significant filaments detected by ref. 26 are both ~ 11 arcsec wide, and were detected only after considerable smoothing (low frequency spatial filtering). Examination of recent numerical simulations^{3,52} suggest that minimum filament widths are typically 40 pkpc at $z \approx 2.4$ and there is a range of 40–100 pkpc. In many cases, multiple parallel filaments may be present over a wider width.

We also note that at the lowest isophote level that according to Fig. 5 we are approaching the morphological transition from filaments to walls. Walls will produce very extended 2D structures that are optimally detected by our approach, and may be responsible for a good fraction of the lower-column-density ALAF.

Cross-correlation

We calculate the 1D and 2D galaxy–emission correlation functions as follows. We use the complete zCOSMOS spectroscopic catalogue⁴¹ of all galaxies within $\Delta z = \pm 0.5$ of the extremes of fields A or B, a sample of 2,835 (1,598) galaxies. For field A, and for each galaxy, we calculate the comoving distance to each voxel in the data cube. For 1D, binning in logarithmic bins ($\Delta \log_{10}(s) = 0.2$), for the spectroscopic distance range $s = 0.1 \text{ cMpc}$ to $s = 1,000 \text{ cMpc}$ ($h = 0.7$), we calculate the average emission for the galaxy in each bin, using the field A smoothed cube. We subtract the average voxel surface brightness $\overline{SB}_\lambda$, correcting both for the mask fraction. This gives the emission cross-correlation function $\overline{SB}_\lambda \xi_{ge}(s)$. We then divide by the average voxel surface brightness $\overline{SB}_\lambda$ to get the normalized correlation function $\xi_{ge}(s)$. For two dimensions, we bin in projected distance σ and line-of-sight redshift distance π to obtain $\overline{SB}_\lambda \xi_{ge}(\sigma, \pi)$ and $\xi_{ge}(\sigma, \pi)$, the 2D cross-correlation function, again in logarithmic bins. This is repeated for field B. To facilitate comparison with the literature, we display all results in distance units of $h^{-1} \text{ cMpc}$.

In equation form, the 2D cross-correlation estimator is:

$$\xi_{ge}(\sigma, \pi) = \frac{\sum_{l_{gal}} \sum_{ij} \overline{SB}_\lambda(\sigma_{ij}, \pi_k) \Pi\left(\frac{\log_{10}[\frac{\sigma_{ij}}{l_{gal}} - \frac{\pi_k}{\pi}] - \sigma}{\Delta \log_{10} \sigma}\right) \Pi\left(\frac{\log_{10}[\frac{\pi_k}{l_{gal}} - \frac{\pi_k}{\pi}] - \pi}{\Delta \log_{10} \pi}\right)}{\overline{SB}_\lambda \sum_{l_{gal}} \sum_{ijk} \Pi\left(\frac{\log_{10}[\frac{\sigma_{ij}}{l_{gal}} - \frac{\pi_k}{\pi}] - \sigma}{\Delta \log_{10} \sigma}\right) \Pi\left(\frac{\log_{10}[\frac{\pi_k}{l_{gal}} - \frac{\pi_k}{\pi}] - \pi}{\Delta \log_{10} \pi}\right)} - 1 \quad (1)$$

where $\Pi(x)$ is the rectangle function, ijk is the voxel index, l_{gal} is the galaxy index, σ_{ij} is the projected distance to voxel ijk , π_k is the line-of-sight redshift distance to voxel ijk , $\Delta \log_{10} \sigma$ is the logarithmic interval for σ , $\Delta \log_{10} \pi$ is the logarithmic interval for π (both 0.2), and $\overline{SB}_\lambda$ is the average voxel surface brightness. Note that we are using the simplest form of the correlation function estimator. We calculate this for fields A and B separately, and show the average in Fig. 4.

The 1D cross-correlation function $\xi_{ge}(s)$ is obtained by using $s = \sqrt{\sigma^2 + \pi^2}$. The projected cross-correlation function $\Xi_{ge}(\sigma)/\sigma$ and the real-space cross-correlation function $\xi_{ge}(r)$ are obtained using the methods of ref. 59.

We calculate errors using a bootstrap approach. We separately use a random selection of the catalogue galaxies with replacement in multiple trials, and calculate the galaxy cross-correlation variance in each bin. We use a random selection of voxels in the data cube and calculate the emission cross-correlation variance in each bin. The total variance is the sum of the two.

Fields A and B show some differences in the cross-correlations. In Supplementary Fig. 3, we show the combined 2D cross-correlation, the combined error and the individual functions for A and B. Supplementary Fig. 4 shows the comparison of $\xi_{ge}(s)$, $\Xi_{ge}(\sigma)/\sigma$ and $\xi_{ge}(r)$. Field A shows a higher cross-correlation in the 3–10 $h^{-1} \text{ cMpc}$ range than field B, a difference, which we ascribe to cosmic variance.

As we pointed out earlier, a cross-correlation is much less sensitive to redshift selection effects than an autocorrelation. We checked that

the distribution of galaxies in the observed sample in π , s space was not too distorted from a random uniform distribution in space (ratios typically 0.6–1.5 and smoothly varying). Bootstrap error estimates would show any notable sensitivity to non-uniformity. We also checked by calculating the cross-correlation function with the Landy–Szalay statistic⁶⁰, which gave similar results. We all confirmed that the distribution of galaxy distances used to bin the cross-correlation in the observed redshift sample is similar to that of a random, uniform sample. These tests also provide evidence that the overdensity calculation used to generate Fig. 1 is not strongly affected by the galaxy redshift selection function, particularly as the overdensity is calculated relative to the sample mean.

Comparison of ELAF and ALAF

Motivated by the spectra in Fig. 3, we want to make a quantitative comparison of the distribution of Ly α emission and absorption. In Fig. 5, we show the differential fraction of voxels versus SB_{λ} , $df/d\log I$ where $I = SB_{\lambda}$. We change units to make a comparison to the distribution of the LAF (and stronger) lines. The fraction of voxels with emission f is changed to the number of voxels per unit redshift dN/dz by multiplying by the number of voxels per unit redshift (voxels per $\Delta z = 1,216 \text{ \AA}/(1 \text{ \AA per voxel})$), where N is the number of voxels with detected emission above a given surface brightness threshold. The ALAF distribution ($dN/dz d\log N_{\text{HI}}$) is given in number of absorbers per unit redshift and range of logarithmic $N(\text{H I})$ column density. As a typical absorber has a Doppler parameter of $b \approx 30 \text{ km s}^{-1}$, line FWHMs are roughly $1 \text{ \AA}$ at $z = 2.35$. As the KCWI cube has $1 \text{ \AA}$ voxels, to convert $df/d\log I$ to an approximate equivalent ALAF distribution we simply divide by the redshift width of a single voxel ($\Delta z = 1 \text{ \AA}/1,216 \text{ \AA}$). This is called $dN/(dz \times d\log I) = df/d\log I \times (\text{voxels}/\Delta z)$ and is shown as the left axis of the main box in Fig. 5. The ALAF distribution is well fitted by a power law⁴⁹, and when we equate the ELAF and ALAF distributions we arrive at the right-hand axis of the main box, corresponding to the H I column density $N(\text{H I})$. Typical designations for the column density ranges are $\text{LAF } N(\text{H I}) < 10^{17} \text{ cm}^{-2}$, $\text{LLS } 10^{17} < N(\text{H I}) < 10^{19} \text{ cm}^{-2}$ and sDLA (also called super-LLS) $10^{19} < N(\text{H I}) < 10^{20.3} \text{ cm}^{-2}$. We therefore proceed with the assumption that the ELAF and ALAF distributions are approximately related, and investigate the implications of this.

The ALAF column density can be roughly converted to a matter overdensity δ (ref. 50), shown as the rightmost axis. Next to this axis, we show typical ranges of three morphological features of the cosmic web, knots, filaments and walls, and their correspondence to overdensity from simulations⁵¹. These morphological features can also be related to volume fraction using simulations⁵², and this comparison is given next to the leftmost axis. Based on this, we can identify the majority of the ELAF as coming from filaments, with the brighter emission from knots and the faintest from walls. Further evidence that this is the case is shown by the red and blue lines. The red line shows the median volume/redshift fractions for the single source with the maximum voxel volume. The line shows an abrupt increase in the maximum source volume as SB_{λ} decreases from $300 \text{ LU \AA}^{-1}$ to $100 \text{ LU \AA}^{-1}$. At this brightness, the maximum source occupies a substantial volume fraction. This behaviour is a signature of percolation, as knots become connected by filaments at lower surface brightnesses. This transition occurs exactly at the range predicted by the two other tracers of morphology discussed above. The blue line shows the median filament parameter η_{fil} discussed above versus surface brightness. This parameter also shows a steep increase below $SB_{\lambda} \approx 1,000 \text{ LU \AA}^{-1}$. The flattening of this curve could be due to the finite field of view, sensitivity limits or the appearance of walls, which will exhibit lower filament parameters.

Finally, we relate the surface brightness to the galaxy–emission cross-correlation signal, converting $\langle SB_{\lambda} \rangle_{\text{ge}}$ to a radius (in $h^{-1} \text{ cMpc}$) using the power-law fit discussed in the main text. The typical distance from a galaxy is $>100 h^{-1} \text{ cMpc}$ for the range of intensity we identify as filamentary, a result approximately consistent with simulations.

Continuum photon pumping

Using the continuum metagalactic background at Ly α in the rest frame⁴⁸, we can calculate the rate of CPP as measured at Earth. A perfect H I absorber with a width and equivalent width of $1 \text{ \AA}$ (that is, a perfect rectangular absorption profile) will produce a scattered surface brightness of $SB_{\text{CPP},0} \approx 166 \text{ LU \AA}^{-1}$ at $z = 2.35$. While the radiative transfer will be more complex, a first approximation can be obtained from the continuum normalized QSO absorption-line spectrum f_{QAL} with $SB_{\text{CPP}} \approx SB_{\text{CPP},0}(1 - f_{\text{QAL}})$. In Fig. 3l, we use a random QSO with high-spectral-resolution data to illustrate the typical appearance that such an emission-line spectrum might take at the KCWI spectral resolution. We used a slightly lower normalization, $SB_{\text{CPP},0} = 100 \text{ LU \AA}^{-1}$. There is evidence that the fainter emission features could be explained by this process, and that this could provide emission sensitivity to much of the ALAF. We note that this process relies on resolving out the sources of continuum emission as discussed in Methods and illustrated in Extended Data Fig. 9. The equivalent of Fig. 3l, which is meant to be illustrative is given for field B and the field B QSO in Extended Data Fig. 10.

The CPP process^{46,47} is observable because the sources of continuum emission in the metagalactic UV background are galaxies and QSOs. For the purposes of this discussion we refer to an LAF absorber as a ‘cloud’. We illustrate the CPP concept in Extended Data Fig. 9. In Extended Data Fig. 9a, we show the case of a single illumination source, and note that absorption is observed in the direction of the source, while emission is observed when the source is not in the field of view. Extended Data Fig. 9b illustrates that this situation continues to pertain as long as none of the emission sources enter the field of view of observer B. If there is a large surface density of emission sources, as in Extended Data Fig. 9c, with an average surface brightness comparable to the metagalactic UV background, the absorption from the sources compensates for the emission and no line is observed. In Extended Data Fig. 9d, this situation is extrapolated to the ideal case of a truly diffuse continuum source, for which again, no line is observed at B. Note that once continuum light is resonantly scattered in the LAF, it becomes effectively diffuse, and the radiative transfer situation in local regions of the IGM will become more complex. But this only occurs when the continuum emission has redshifted into the Ly α line region, which occurs near the IGM cloud, and only to the fraction of the continuum absorbed by the LAF. Even if we assume all of the continuum is scattered in the LAF over this redshift range, a fraction 0.75 remains unscattered and resolvable into sources⁶¹. The Ly β forest produces at most an 8% reduction in continuum over the full range.

Another way of thinking about this process is the following. Photons are removed from the continuum of QSOs^{62,63} and galaxies⁶⁴ by H I in the IGM, as observed on Earth. A proper accounting of these photons requires that, as they are not destroyed by any process (such as dust absorption), they must be detectable as emission emanating from the H I regions that produce the absorption, as long as the continuum sources are resolved out.

To ensure that continuum sources are excluded from the KCWI emission measurement, we perform sky substitution for all galaxies whose photometric redshifts fall within the range $2.1 < z < 3.6$ for magnitude range $24 < m_B < 28$. This redshift range is chosen to ensure that galaxies with rest wavelengths from $912 \text{ \AA} < \lambda < 1,220 \text{ \AA}$ are removed from the emission voxels. We use a parameterized UV luminosity function⁶⁵ to derive the fraction of continuum background due to galaxies present at the emission redshift $z \approx 2.35$ versus the limiting magnitude observed on Earth. We assume a flat UV spectrum over the range $912 \text{ \AA} < \lambda < 1,220 \text{ \AA}$. For a limiting magnitude of $m_B < 26|27|28$ we find that the continuum background that is resolved out (and therefore observable as CPP) is $0.47|0.70|0.90$. Note that even if we were to set the limit at $m_B < 26$, the sources would have a low surface density, with a fraction of 0.16 falling in a $10 \times 10 \text{ arcsec}^2$ region. Thus in this case, 84% of voxels would display emission. We set the limit at $m_B < 28$ to be conservative.

We show in Extended Data Fig. 10 the field B QSO absorption spectrum continuum normalized and converted to an emission spectrum as

in Fig. 3l. Vertical red and blue lines indicate features in the absorption spectrum that have counterparts in the emission spectrum within 2 Å.

We measured the cross-correlation between the field B QSO at the overdensity redshift with the field B ELAF. The QSO absorption-line spectrum is the same spectral resolution as the emission data. We excluded wavelengths longer than 4,140 Å to avoid QSO system absorbers, and we continuum corrected the spectrum and inverted it as in Fig. 3l. We estimated errors by bootstrap sampling of the emission cube and the absorption spectrum. A statistically significant cross-correlation is detected in the range $0.3\text{--}8\ h^{-1}\text{cMpc}$, as we show in Extended Data Fig. 4. We show only the 1D cross-correlation (analogous to Fig. 4b), as there is insufficient transverse distance (limited by the size of the data cube) to assess the 2D and 1D radial correlation functions.

Data availability

The observational data used in this paper are available on the Keck Observatory Archive (<https://www2.keck.hawaii.edu/koa/public/koa.php>).

References

- Hogan, C. J. & Weymann, R. J. Lyman-alpha emission from the Lyman-alpha forest. *Mon. Not. R. Astron. Soc.* **225**, 1P–5P (1987).
- Cantalupo, S., Porciani, C., Lilly, S. J. & Miniati, F. Fluorescent Ly α emission from the high-redshift intergalactic medium. *Astrophys. J.* **628**, 61–75 (2005).
- Byrohl, C. et al. The physical origins and dominant emission mechanisms of Lyman alpha haloes: results from the TNG50 simulation in comparison to MUSE observations. *Mon. Not. R. Astron. Soc.* **506**, 5129–5152 (2021).
- Elias, L. M. et al. Detecting the cosmic web: Ly α emission from simulated filaments at $z=3$. *Mon. Not. R. Astron. Soc.* **494**, 5439–5448 (2020).
- Witstok, J., Puchwein, E., Kulkarni, G., Smit, R. & Haehnelt, M. G. Prospects for observing the low-density cosmic web in Lyman- α emission. *Astron. Astrophys.* **650**, A98 (2021).
- Cantalupo, S., Arrigoni-Battaia, F., Prochaska, J. X., Hennawi, J. F. & Madau, P. A cosmic web filament revealed in Lyman- α emission around a luminous high-redshift quasar. *Nature* **506**, 63–66 (2014).
- Martin, D. C., Chang, D. & Matuszewski, M. et al. Intergalactic medium emission observations with the Cosmic Web Imager. I. The circum-QSO medium of QSO 1549+19, and evidence for a filamentary gas inflow. *Astrophys. J.* **786**, 106 (2014).
- Borisova, E., Cantalupo, S. & Lilly, S. J. et al. Ubiquitous giant Ly α nebulae around the brightest quasars at $z\sim 3.5$ revealed with MUSE. *Astrophys. J.* **831**, 39 (2016).
- Umehata, H. et al. Gas filaments of the cosmic web located around active galaxies in a protocluster. *Science* **366**, 97–100 (2019).
- Arrigoni Battaia, F. et al. Discovery of intergalactic bridges connecting two faint $z\sim 3$ quasars. *Astron. Astrophys.* **631**, A18 (2019).
- O’Sullivan, D. B. et al. The FLASHES Survey. I. Integral field spectroscopy of the CGM around 48 $z=2.3\text{--}3.1$ QSOs. *Astrophys. J.* **894**, 3 (2020).
- Steidel, C. C., Adelberger, K. L. & Shapley, A. E. et al. Ly α imaging of a proto-cluster region at $z>3.09$. *Astrophys. J.* **532**, 170–182 (2000).
- Matsuda, Y. et al. Large-scale filamentary structure around the protocluster at redshift $z=3.1$. *Astrophys. J. Lett.* **634**, L125–L128 (2005).
- Matsuda, Y. et al. The Subaru Ly α blob survey: a sample of 100-kpc Ly α blobs at $z=3$. *Mon. Not. R. Astron. Soc.* **410**, L13–L17 (2011).
- Yang, Y., Zabludoff, A., Tremonti, C., Eisenstein, D. & Davé, R. Extended Ly α nebulae at $z\sim 2.3$: an extremely rare and strongly clustered population? *Astrophys. J.* **693**, 1579–1587 (2009).
- Yang, Y., Zabludoff, A., Eisenstein, D. & Davé, R. Strong field-to-field variation of Ly α nebulae populations at $z\sim 2.3$. *Astrophys. J.* **719**, 1654–1671 (2010).
- Yang, Y., Zabludoff, A., Jahnke, K. & Davé, R. The properties of Ly α nebulae: gas kinematics from nonresonant lines. *Astrophys. J.* **793**, 114 (2014).
- Cai, Z. et al. Discovery of an enormous Ly α nebula in a massive galaxy overdensity at $z=2.3$. *Astrophys. J.* **837**, 71 (2017).
- Herenz, E. C., Hayes, M. & Scarlata, C. Deciphering the Lyman α blob 1 with deep MUSE observations. *Astron. Astrophys.* **642**, A55 (2020).
- Daddi, E. et al. Three Lyman- α -emitting filaments converging to a massive galaxy group at $z=2.91$: discussing the case for cold gas infall. *Astron. Astrophys.* **649**, A78 (2021).
- Martin, D. C., Chang, D. & Matuszewski, M. et al. Intergalactic medium emission observations with the Cosmic Web Imager. II. Discovery of extended, kinematically linked emission around SSA22 Ly α blob 2. *Astrophys. J.* **786**, 107 (2014).
- Ao, Y. et al. Infalling gas in a Lyman- α blob. *Nat. Astron.* **4**, 670–674 (2020).
- Steidel, C. C., Bogosavljević, M. & Shapley, A. E. et al. Diffuse Ly α emitting halos: a generic property of high-redshift star-forming galaxies. *Astrophys. J.* **736**, 160 (2011).
- Wisotzki, L. et al. Extended Lyman α haloes around individual high-redshift galaxies revealed by MUSE. *Astron. Astrophys.* **587**, A98 (2016).
- Wisotzki, L. et al. Nearly all the sky is covered by Lyman- α emission around high-redshift galaxies. *Nature* **562**, 229–232 (2018).
- Bacon, R. et al. The MUSE Extremely Deep Field: the cosmic web in emission at high redshift. *Astron. Astrophys.* **647**, A107 (2021).
- Morrissey, P. et al. The Keck Cosmic Web Imager Integral Field Spectrograph. *Astrophys. J.* **864**, 93 (2018).
- Sembach, K. R. & Tonry, J. L. Accurate sky subtraction of long-slit spectra: velocity dispersions at $\Sigma_v=24.0\text{ mag/arcsec}^2$. *Astron. J.* **112**, 797 (1996).
- Cen, R. The nature of damped Ly α systems and their hosts in the standard cold dark matter universe. *Astrophys. J.* **748**, 121 (2012).
- Ferland, G. J. et al. CLOUDY 90: numerical simulation of plasmas and their spectra. *Publ. Astron. Soc. Pac.* **110**, 761–778 (1998).
- Neufeld, D. A. The transfer of resonance-line radiation in static astrophysical media. *Astrophys. J.* **350**, 216–241 (1990).
- Scoville, N. et al. Evolution of galaxies and their environments at $z=0.1\text{--}3$ in COSMOS. *Astrophys. J. Suppl. Ser.* **206**, 3 (2013).
- Laigle, C. et al. The COSMOS2015 catalog: exploring the $1<z<6$ Universe with half a million galaxies. *Astrophys. J. Suppl. Ser.* **224**, 24 (2016).
- Martin, D. C. et al. A newly forming cold flow protogalactic disk, a signature of cold accretion from the cosmic web. *Astrophys. J. Lett.* **824**, L5 (2016).
- Martin, D. C. et al. Multi-filament gas inflows fuelling young star-forming galaxies. *Nat. Astron.* **3**, 822–831 (2019).
- Gallego, S. G. et al. Stacking the cosmic web in fluorescent Ly α emission with MUSE. *Mon. Not. R. Astron. Soc.* **475**, 3854–3869 (2018).
- Rauch, M. et al. A population of faint extended line emitters and the host galaxies of optically thick QSO absorption systems. *Astrophys. J.* **681**, 856–880 (2008).
- Chen, Y.-C., Ho, S., Freeman, P. E., Genovese, C. R. & Wasserman, L. Cosmic web reconstruction through density ridges: method and algorithm. *Mon. Not. R. Astron. Soc.* **454**, 1140–1156 (2015).
- Chen, Y.-C. et al. Cosmic web reconstruction through density ridges: catalogue. *Mon. Not. R. Astron. Soc.* **461**, 3896–3909 (2016).
- Meiksin, A. A. The physics of the intergalactic medium. *Rev. Mod. Phys.* **81**, 1405–1469 (2009).
- Lilly, S. J. et al. zCOSMOS: a large VLT/VIMOS redshift survey covering $0<z<3$ in the COSMOS field. *Astrophys. J. Suppl. Ser.* **172**, 70–85 (2007).

42. Bullock, J. S. et al. Profiles of dark haloes: evolution, scatter and environment. *Mon. Not. R. Astron. Soc.* **321**, 559–575 (2001).
43. Lusso, E. et al. The MUSE Ultra Deep Field (MUDF)—I. Discovery of a group of Ly α nebulae associated with a bright $z \approx 3.23$ quasar pair. *Mon. Not. R. Astron. Soc.* **485**, L62–L67 (2019).
44. Smith, A. et al. The physics of Lyman α escape from high-redshift galaxies. *Mon. Not. R. Astron. Soc.* **484**, 39–59 (2019).
45. Mitchell, P. D. et al. Tracing the simulated high-redshift circumgalactic medium with Lyman α emission. *Mon. Not. R. Astron. Soc.* **501**, 5757–5775 (2021).
46. Silva, M. B., Santos, M. G., Gong, Y., Cooray, A. & Bock, J. Intensity mapping of Ly α emission during the epoch of reionization. *Astrophys. J.* **763**, 132 (2013).
47. Pullen, A. R., Doré, O. & Bock, J. Intensity mapping across cosmic times with the Ly α line. *Astrophys. J.* **786**, 111 (2014).
48. Haardt, F. & Madau, P. Radiative transfer in a clumpy universe. IV. New synthesis models of the cosmic UV/X-ray background. *Astrophys. J.* **746**, 125 (2012).
49. Tytler, D. The distribution of QSO absorption system column densities: evidence for a single population. *Astrophys. J.* **321**, 49 (1987).
50. Zhang, Y., Meiksin, A., Anninos, P. & Norman, M. L. Physical properties of the Ly α forest in a cold dark matter cosmology. *Astrophys. J.* **495**, 63–79 (1998).
51. Haider, M. et al. Large-scale mass distribution in the Illustris simulation. *Mon. Not. R. Astron. Soc.* **457**, 3024–3035 (2016).
52. Martizzi, D. et al. Baryons in the cosmic web of IllustrisTNG—I: gas in knots, filaments, sheets, and voids. *Mon. Not. R. Astron. Soc.* **486**, 3766–3787 (2019).
53. Scoville, N. et al. The Cosmic Evolution Survey (COSMOS): overview. *Astrophys. J. Suppl. Ser.* **172**, 1–8 (2007).
54. Darvish, B. et al. Spectroscopic confirmation of a coma cluster progenitor at $z \approx 2.2$. *Astrophys. J.* **892**, 8 (2020).
55. Cucciati, O. et al. The progeny of a cosmic titan: a massive multi-component proto-supercluster in formation at $z = 2.45$ in VUDS. *Astron. Astrophys.* **619**, A49 (2018).
56. Martin, D. C. et al. A giant protogalactic disk linked to the cosmic web. *Nature* **524**, 192–195 (2015).
57. Diener, C. et al. Proto-groups at $1.8 < z < 3$ in the zCOSMOS-deep sample. *Astrophys. J.* **765**, 109 (2013).
58. Diener, C. et al. A protocluster at $z = 2.45$. *Astrophys. J.* **802**, 31 (2015).
59. Hawkins, E. et al. The 2dF Galaxy Redshift Survey: correlation functions, peculiar velocities and the matter density of the Universe. *Mon. Not. R. Astron. Soc.* **346**, 78–96 (2003).
60. Landy, S. D. & Szalay, A. S. Bias and variance of angular correlation functions. *Astrophys. J.* **412**, 64 (1993).
61. Fan, X. et al. Constraining the evolution of the ionizing background and the epoch of reionization with $z \approx 6$ quasars. II. A sample of 19 quasars. *Astron. J.* **132**, 117–136 (2006).
62. Gunn, J. E. & Peterson, B. A. On the density of neutral hydrogen in intergalactic space. *Astrophys. J.* **142**, 1633–1636 (1965).
63. Bahcall, J. N. & Salpeter, E. E. On the interaction of radiation from distant sources with the intervening medium. *Astrophys. J.* **142**, 1677–1680 (1965).
64. Lee, K.-G. et al. Ly α forest tomography from background galaxies: the first megaparsec-resolution large-scale structure map at $z > 2$. *Astrophys. J. Lett.* **795**, L12 (2014).
65. Parsa, S., Dunlop, J. S., McLure, R. J. & Mortlock, A. The galaxy UV luminosity function at $z \approx 2\text{--}4$; new results on faint-end slope and the evolution of luminosity density. *Mon. Not. R. Astron. Soc.* **456**, 3194–3211 (2016).
66. Murphy, M. T., Kacprzak, G. G., Savorgnan, G. A. D. & Carswell, R. F. The UVES Spectral Quasar Absorption Database (SQUAD) Data Release 1: the first 10 million seconds. *Mon. Not. R. Astron. Soc.* **482**, 3458–3479 (2018).
67. Peebles, P. J. E. The mean mass density estimated from the Kirshner, Oemler, Schechter galaxy redshift sample. *Astron. J.* **84**, 730–734 (1979).
68. Adelberger, K. L. et al. The spatial clustering of star-forming galaxies at redshifts $1.4 \lesssim z \lesssim 3.5$. *Astrophys. J.* **619**, 697–713 (2005).
69. Prochaska, J. X. & Tumlinson, J. Baryons: what, when and where? In *Astrophysics in the Next Decade*, *Proc. Astrophysics and Space Science* Vol. 10, 419 (2009).

Acknowledgements

Resources supporting this work were provided by NSF AAG Grant 1716907 and the California Institute of Technology. The data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The observatory was made possible by the generous financial support of the W. M. Keck Foundation. We recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the Indigenous Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

Author contributions

D.C.M. is the principal investigator of the KCWI, conceived of and designed the observing strategy, performed 75% of the data analysis and wrote the main paper. B.D. provided the targeting strategy, and contributed to the observations, paper writing and editing. Z.L. led most of the observations and contributed to the data analysis and paper. M.M. was instrument scientist on the KCWI and provided editing support to the paper. R.C. provided the numerical modelling that supported the observation planning. J.D.N. provided the data reduction pipeline for KCWI. P.M. was the technical lead on the KCWI. A.M.M. was the project manager for the KCWI.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41550-023-02054-1>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41550-023-02054-1>.

Correspondence and requests for materials should be addressed to D. Christopher Martin.

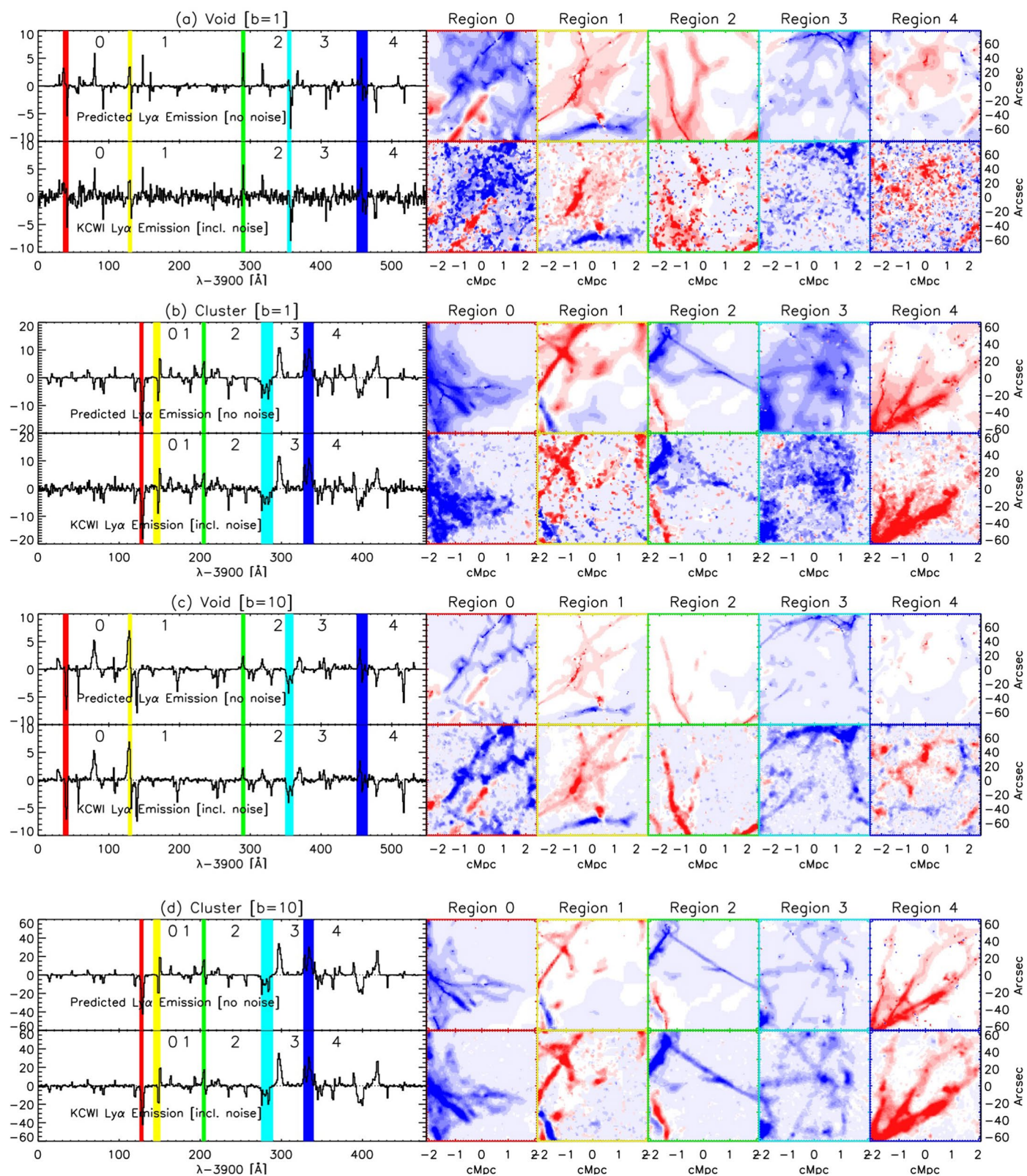
Peer review information *Nature Astronomy* thanks the anonymous reviewers for their contribution to the peer review of this work.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

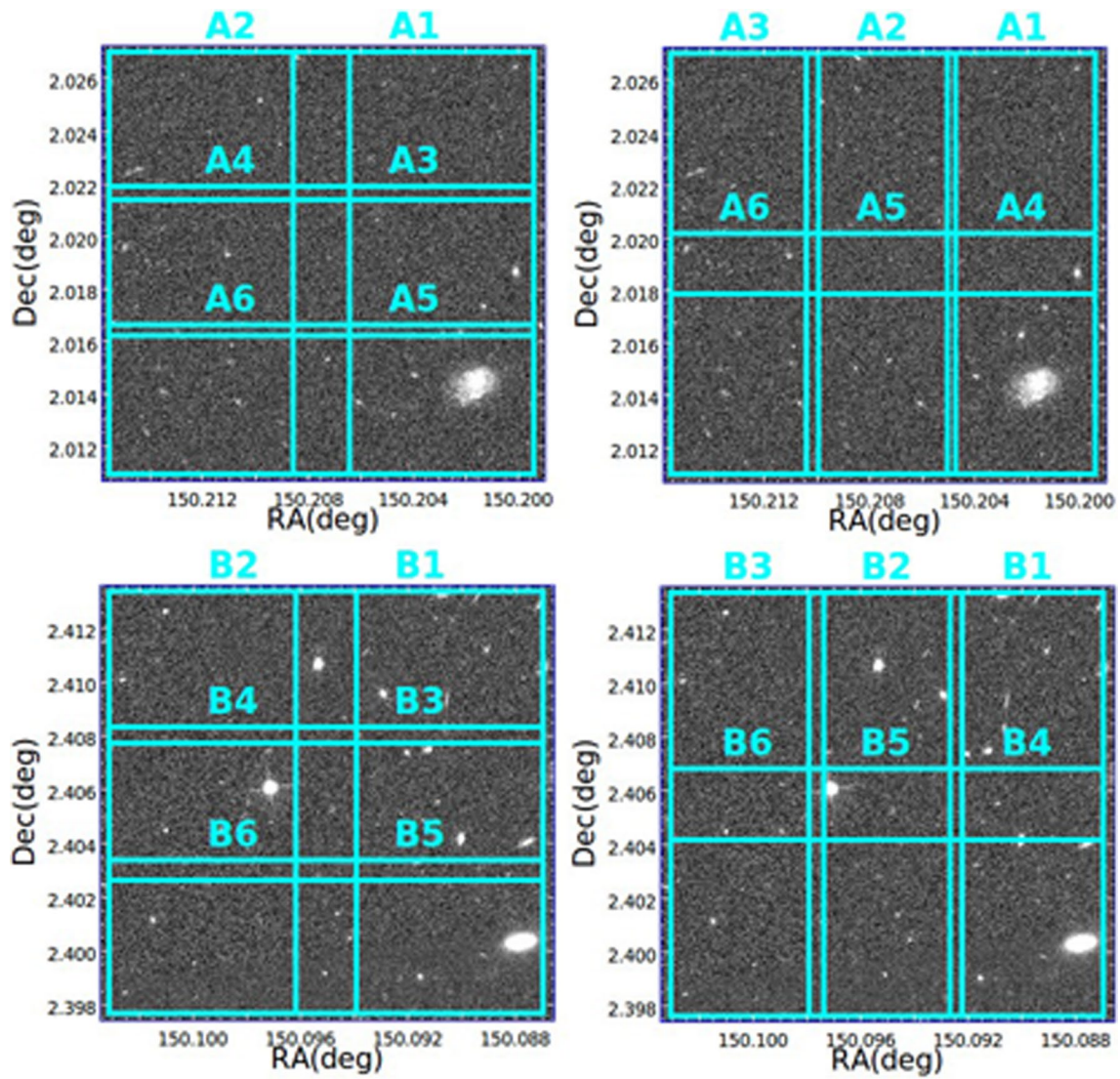
Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023

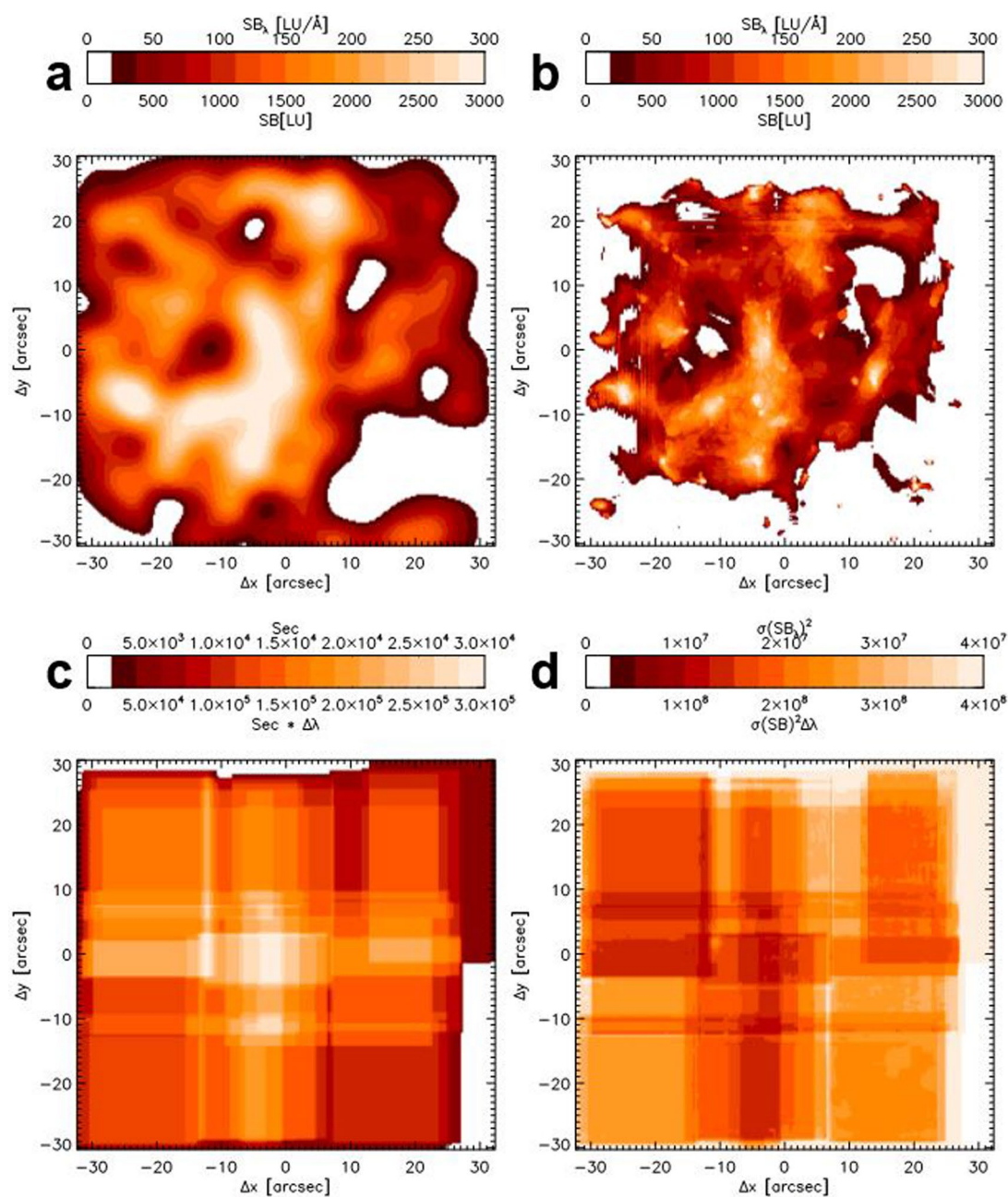


Extended Data Fig. 1 | Forwarded modeled numerical simulation of Ly α emission from the general cosmic web at $z \sim 3$ predicted that the KCWI observation would detect emission. Upper spectrum is sum of enlarged 2×2 arcmin 2 difference field (source-background) model emission, middle spectrum is observed spectrum with KCWI, parameters discussed in the text, with 1 LU

$= 1.2 \times 10^{-22} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ arcsec}^{-2}$. Four narrow-band images are shown for the colored wavelength intervals, upper shows with no noise, lower is adaptively smoothed image^{1,2} based on simulated observed KCWI data cube, with full-scale comparable to spectrum. Ordinate is average surface brightness ($\langle SB_i \rangle$) over displayed image.



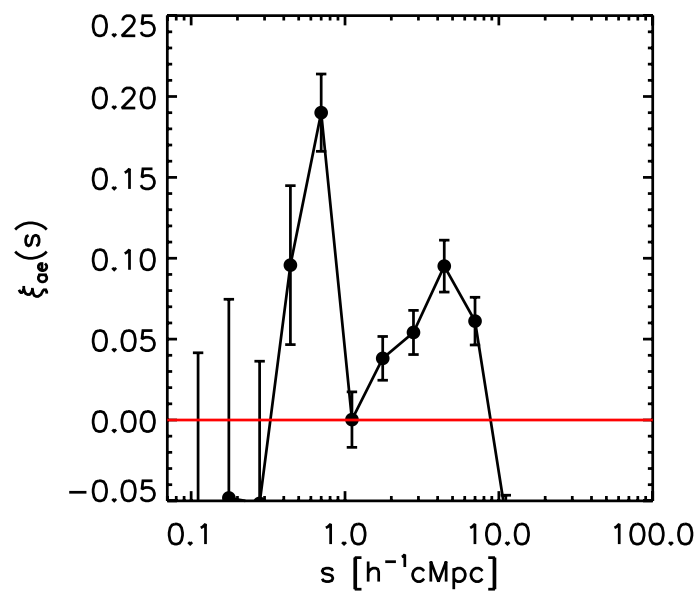
Extended Data Fig. 2 | Tiling overlay of Fields A (top) and B (bottom) at position angles PA=0° (left) and PA=90° (right). An array of 6 overlapping KCWI Large slicers, covering an area of $\sim 60 \times 60 \text{ arcsec}^2$. The image cutouts are from the COSMOS HST/F814w data.



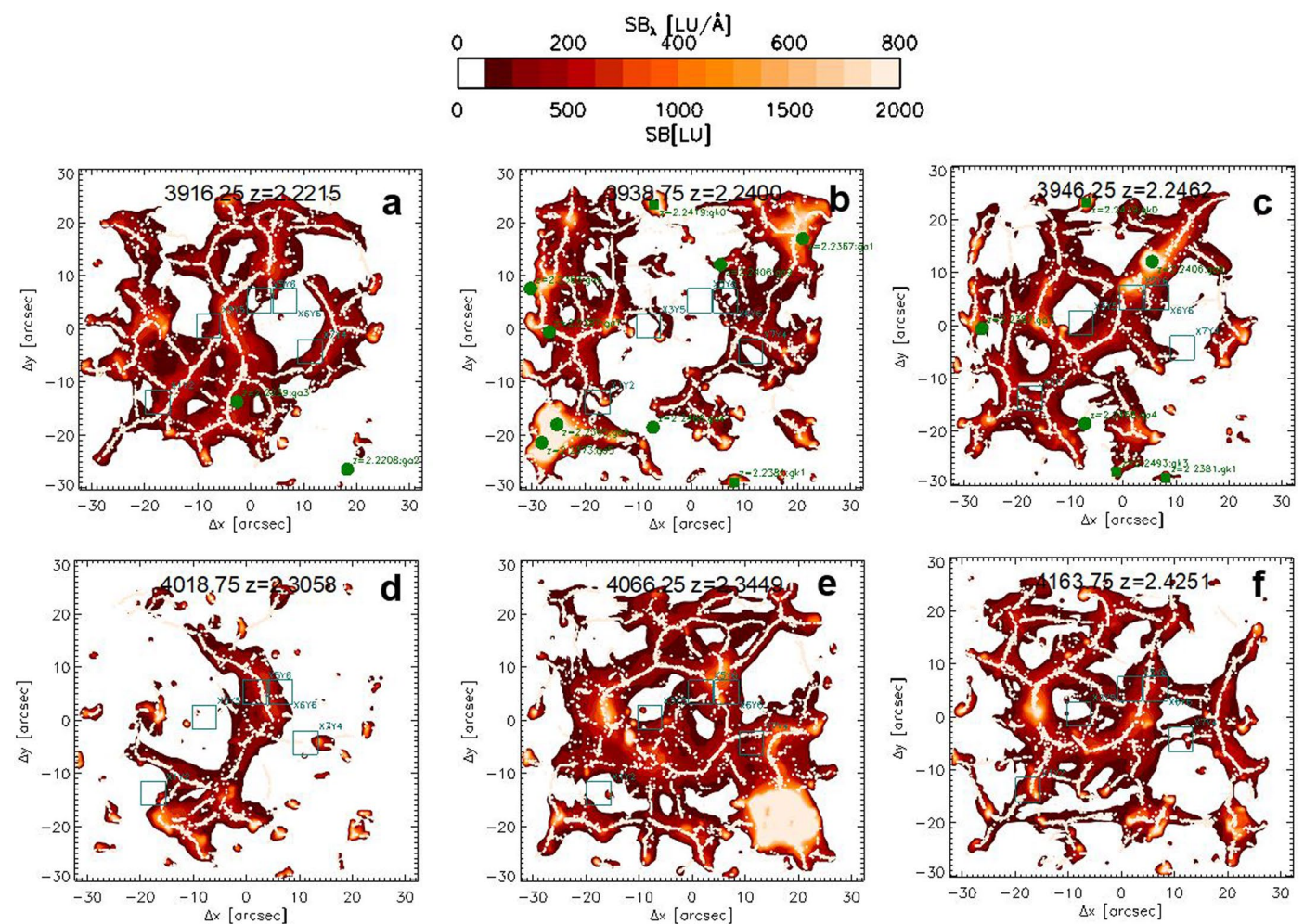
Extended Data Fig. 3 | Raw, smoothed, exposure, and variance images.

a. Raw image slice 3910–3920 Å, smoothed to 20 pixels or 5.8 arcsec. b. Adaptively smoothed image slice. Note that adaptively smoothing reduces the extended

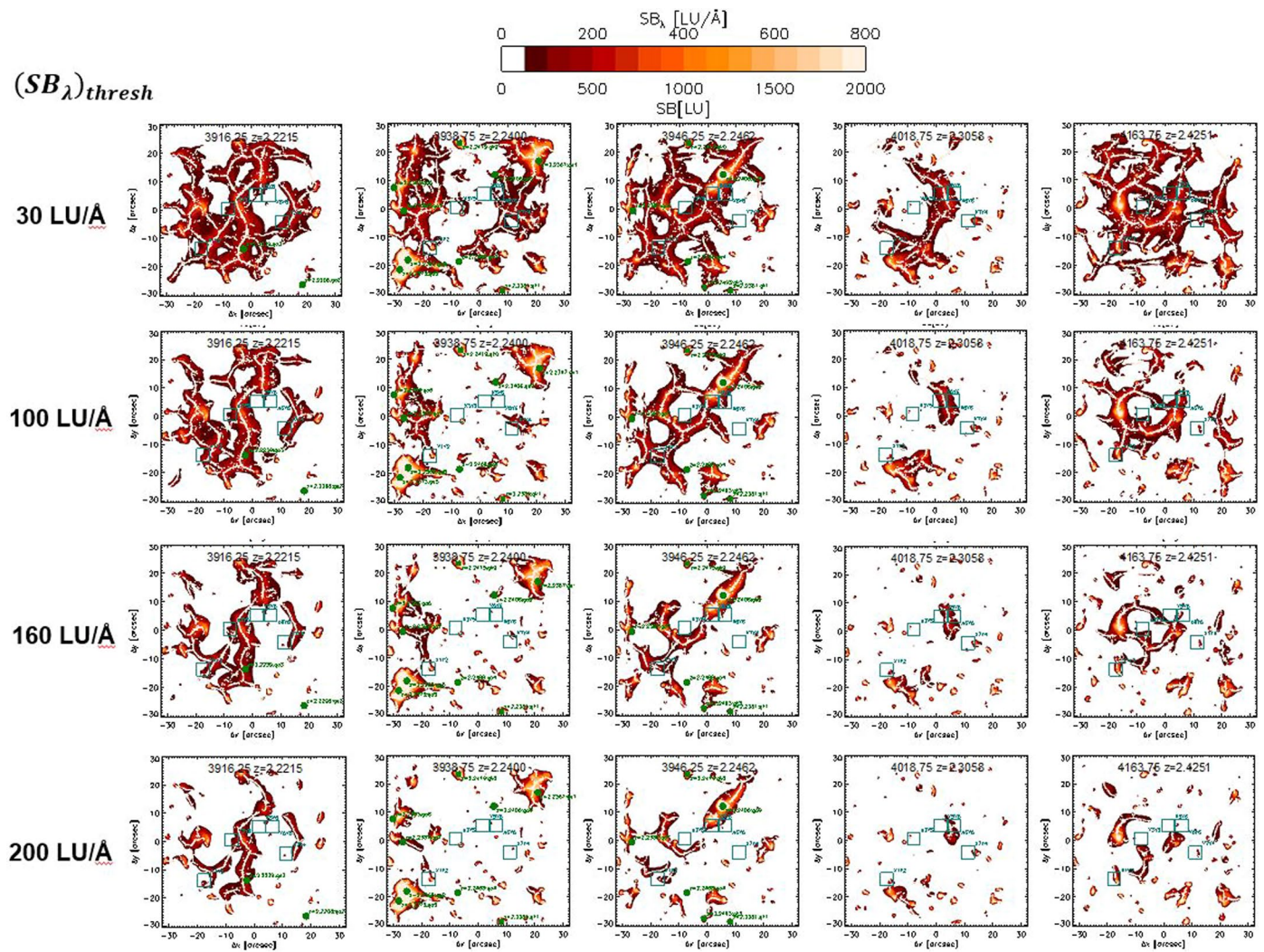
emission levels because bright compact regions above the noise threshold are removed. c. Exposure time (top scale gives seconds, bottom seconds times $\Delta\lambda$). d. Variance cube slice.



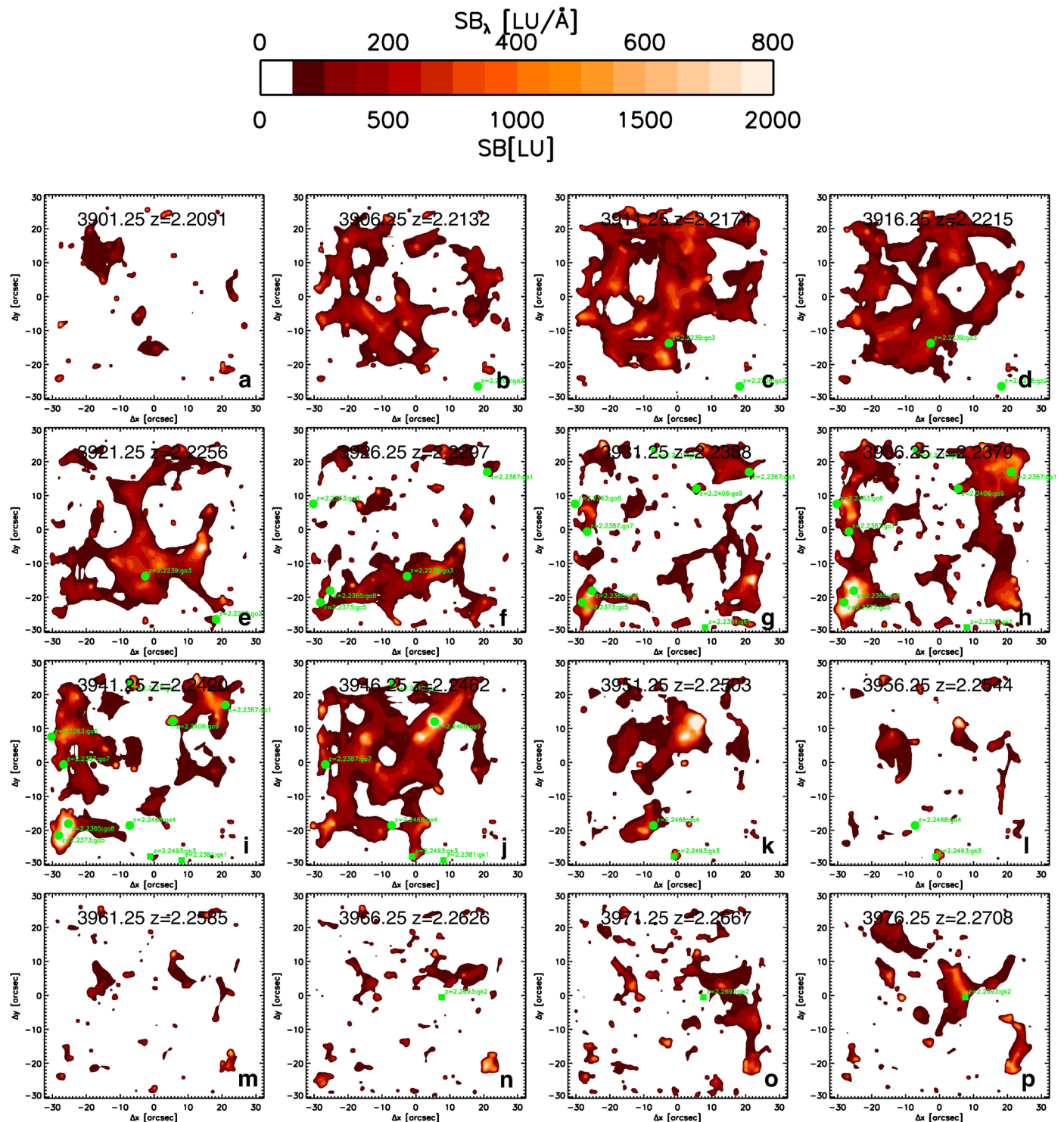
Extended Data Fig. 4 | Cross correlation between ELAF and ALAF. Estimated 1D cross-correlation function between Field B ELAF and the ALAF present in the field QSO. Errors estimated by bootstrapping both the absorption and emission line data.



Extended Data Fig. 5 | Filament plots. Six 3Å narrow-band images showing filament detection locations. Note that most galaxies are connected by filaments.

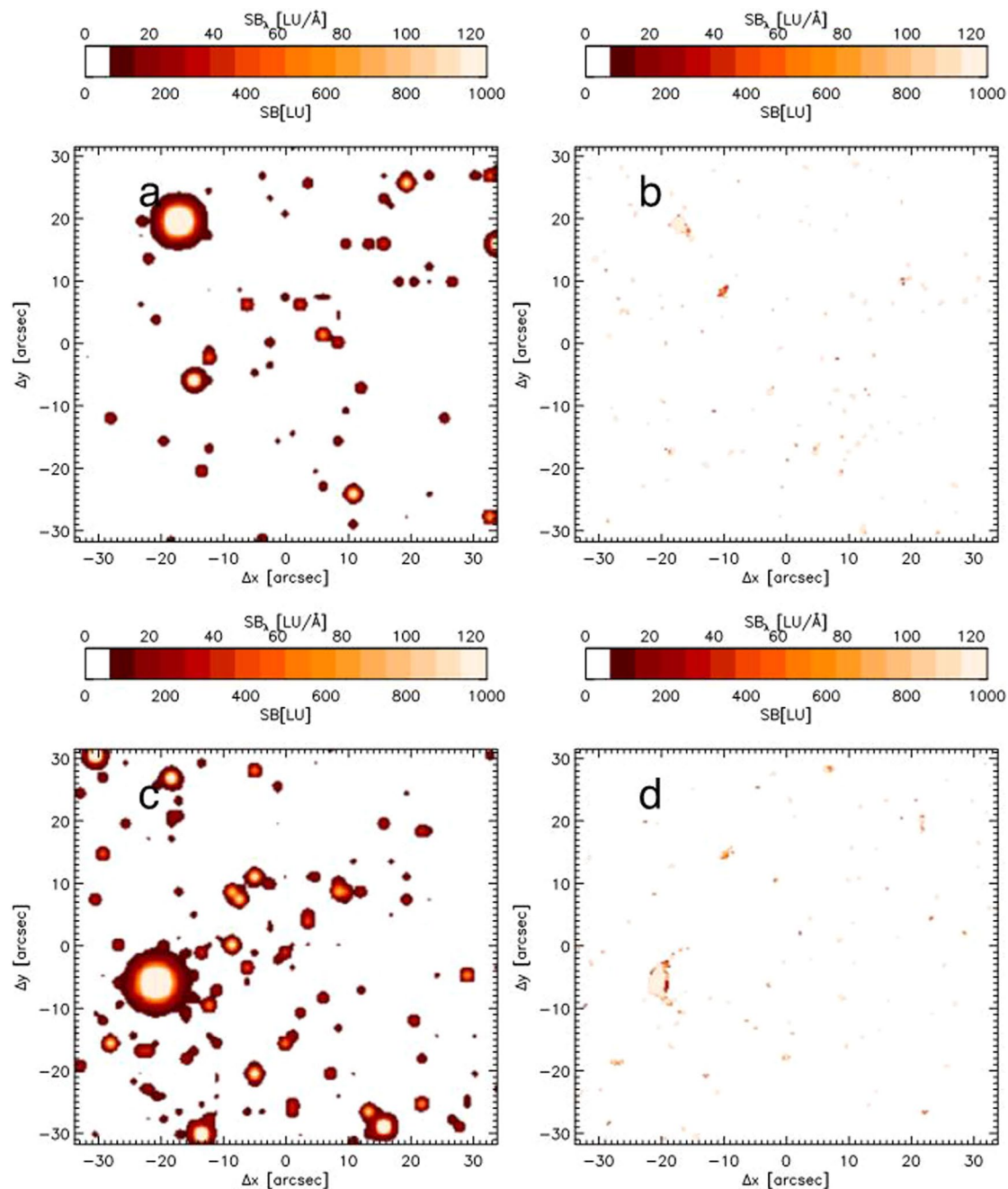


Extended Data Fig. 6 | Filament plots vs. threshold. Five 3Å narrow-band images showing filament detection locations vs. surface brightness threshold. Filament lengths and interconnectivity decrease with increasing surface brightness threshold.



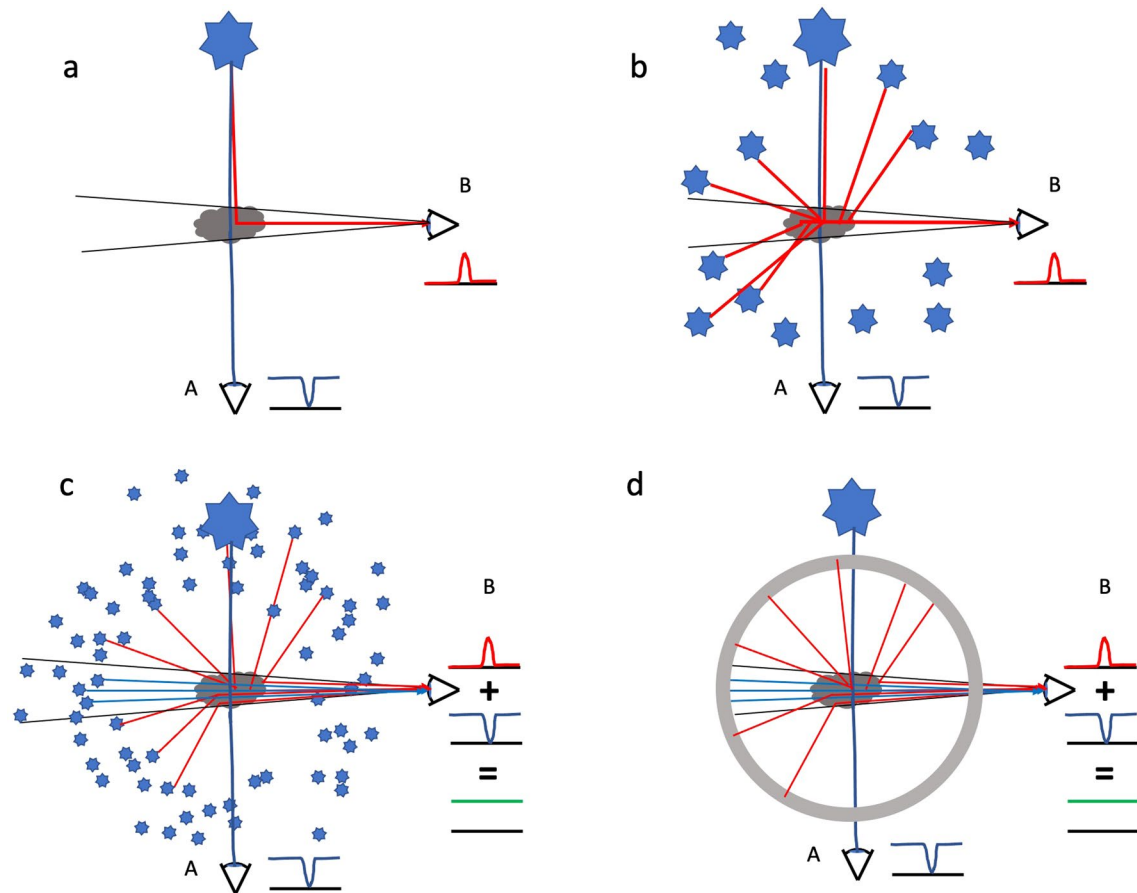
Extended Data Fig. 7 | Field A overdensity at $z=2.23$. Mosaic of 16 3Å slices covering the $z - 2.23$ overdensity. Each central wavelength and corresponding Ly α redshift is shown, and galaxies are displayed for which Ly α falls within 10Å

of the slice. Most of the galaxies are embedded in emission at a corresponding redshift, and the majority have emission spanning multiple galaxies at a given redshift.



Extended Data Fig. 8 | Narrow-band slices from LAH Model. LAHs were placed randomly using simulation following density structures in the cosmic web. Two 8 Å slices shown from top to bottom. On left, the input surface brightness

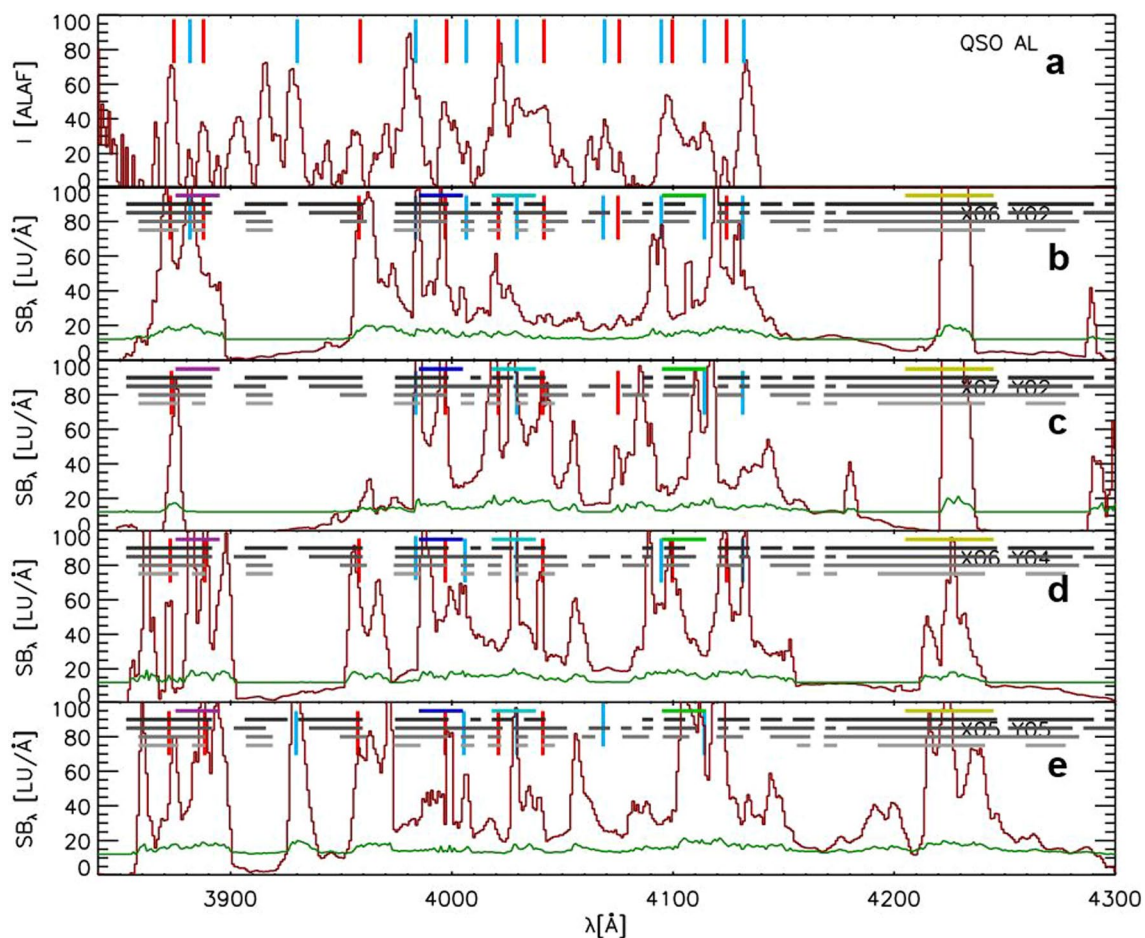
variance cube. Bright cores of the brighter LAHs are detected, but only part of the halo of the brightest LAH. There is no signal from smoothed out faint halos, even though they are distributed in overdense filamentary structures. The derived voxel volume distribution is given as the grey line (LAH) in Fig. 5.



Extended Data Fig. 9 | Illustration of Continuum Photon Pumping Process.

a. With a single illumination source, IGM cloud absorbs in Ly α for observer A, and emits due to resonance scattering in the Ly α into the field of view of observer B. b. With a distribution of line and continuum illumination sources, observer A again sees absorption, which observer B, with illuminating sources resolved and therefore no illuminating sources in the field of view, sees an emission line. c. With a large number of faint illuminating sources, sufficiently numerous that

they occupy the B field of view with the average metagalactic surface brightness, the absorption from the cloud exactly compensates for the scattered emission, and no line is observed. d. With a diffuse, uniform source of line and continuum emission the emission and absorption exactly cancel. In this paper, the scenario in panel b. pertains since the majority of the illuminating sources are resolved out and excluded from the emission measurement.



Extended Data Fig. 10 | QSO absorption line spectrum converted to emission line spectrum. a. Field B QSO absorption line spectrum converted to emission line spectrum following Fig. 3. Panels b-e show emission from individual regions

near QSO, with vertical lines indicating features seen in the QSO absorption line spectrum which have counterparts (within $2\text{\AA}$) in the emission spectra. Line colors alternate for clarity.