ALMA REVEALS LARGE OVERDENSITY AND STRONG CLUSTERING OF GALAXIES IN QUASAR ENVIRONMENTS AT $Z \sim 4$

Cristina García-Vergara[1], Matus Rybak[1,2], Jacqueline Hodge[1], Joseph F. Hennawi[3], Roberto Decarli[4], Jorge González-López[5,6], Fabrizio Arrigoni-Battaia[7], Manuel Aravena[5], and Emanuele P. Farina[7]

Draft version September 22, 2021

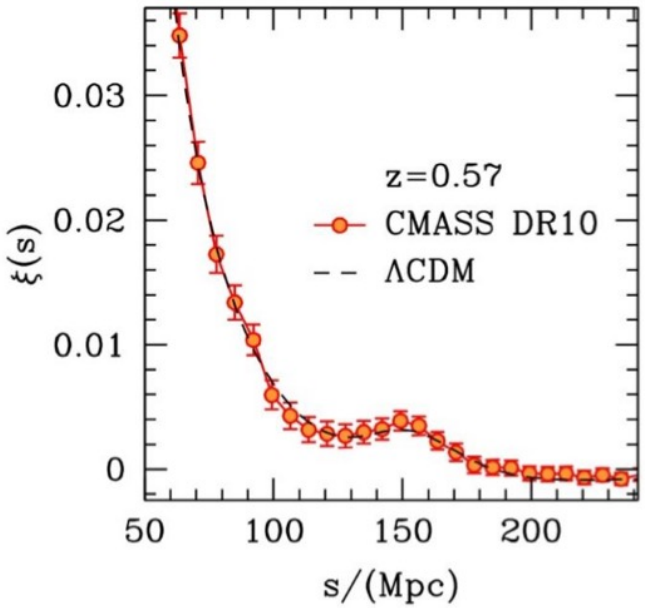
ABSTRACT

We present an Atacama Large Millimeter/submillimeter Array (ALMA) survey of CO(4–3) line emitting galaxies in 17 quasar fields at $z \sim 4$ aimed to perform the first systematic search of dusty galaxies in high−z quasar environments. Our blind search of galaxies around the quasars results in 5 CO emitters with S/N≥ 5.6 within a projected radius of $R \lesssim 1.5\,h^{-1}$ cMpc and a velocity range of $\Delta\,\mathrm{v} = \pm 1000\,\mathrm{km\,s^{-1}}$ around the quasar. In blank fields, we expect to detect only 0.28 CO emitters within the same volume, implying a total overdensity of $17.6^{+11.9}_{-7.6}$ in our fields, and indicating that quasars trace massive structures in the early universe. We quantify this overdensity by measuring the small-scale clustering of CO emitters around quasars, resulting in a cross-correlation length of $r_{0,\mathrm{QG}} = 8.37^{+2.42}_{-2.04}\,h^{-1}$ cMpc, assuming a fixed slope $\gamma = 1.8$. This contradicts the reported mild overdensities (x1.4) of Lyman alpha emitters (LAEs) in the same fields at scales of $R \lesssim 7\,h^{-1}$ cMpc which is well described by a cross-correlation length 3 times lower than that measured for CO emitters. We discuss some possibilities to explain this discrepancy, including low star formation efficiency, and excess of dust in galaxies around quasars. Finally, we constrain, for the first time, the clustering of CO emitters at $z \sim 4$, finding an auto-correlation length of $r_{0,\mathrm{CO}} = 3.14 \pm 1.71\,h^{-1}$ cMpc (with $\gamma = 1.8$). Our work together with the previous study of LAEs around quasars traces simultaneously the clustering properties of both optical and dusty galaxy populations in quasars fields, stressing the importance of multi-wavelength studies, and highlighting important questions about galaxy properties in high−z dense environments.

Keywords: Quasars (1319), Quasar-galaxy pairs (1316), High-redshift galaxies (734), CO line emission (262), Clustering (1908), Large-scale structure of the universe (902), Submillimeter astronomy (1647)

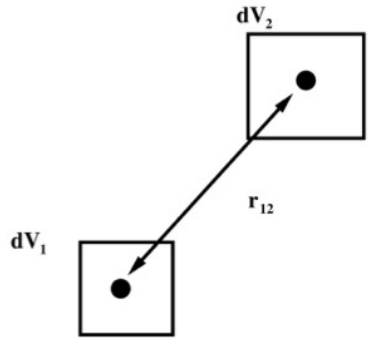
8.1.2 星系成团量度

- 如何定量描述星系空间分布的不均匀性？描述方法包括两点相关函数、功率谱、密度涨落



- 两点相关函数ξ(r)：计算在给定间距，发现成对星系的概率，得出两点相关性与距离的函数关系。
- 数学描述：
  - 假定星系的平均空间密度是 n/Mpc^3；随机选择两个小体积ΔV_1和ΔV_2，
  - 则在ΔV_1中找到一个星系的概率是$n\Delta V_1$；在相距r_{12}的ΔV_2中也找到一个星系的联合概率为：

$$\mathrm{d}P = \bar{n}^2(1+\xi(r_{12}))\,\mathrm{d}V_1\,\mathrm{d}V_2 \qquad \bar{n} \text{ is the mean galaxy density.}$$



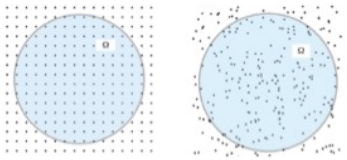
  - ξ(r)：在距给定星系距离r处找到另一星系高出平均的几率
  - 如果在小r处 ξ(r) > 0，则星系成团；如果ξ(r) < 0，则反成团； 较大体积ξ(r) 的平均值应该为0.
  - 在r ≤ 10 h^{-1} Mpc尺度， ξ(r)可近似表达为： $\xi(r) \approx (r/r_0)^{-\gamma}$ $(\gamma > 0)$
  - 若r < r0时，在距离一星系r以内找到一个星系的概率显著大于随机分布，则称r0为相关长度（correlation length）。

Two-Point Correlation Functions

One of the most important statistics used to characterize the spatial distribution of galaxies is the two-point correlation function, defined as the excess number of galaxy pairs of a given separation, r, relative to that expected for a random distribution:

对于不同的间隔r，在观测和随机数分布中，分别计算星系对数目

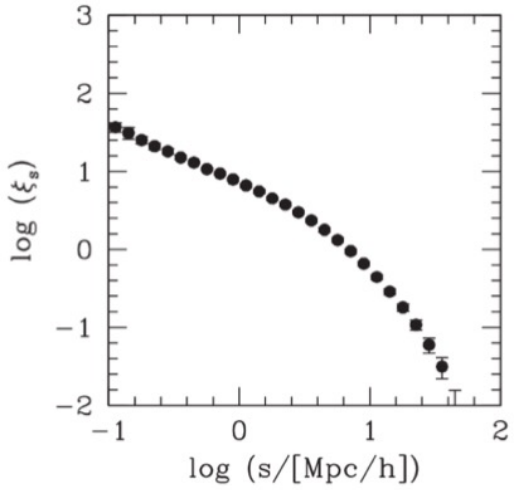
$$\xi(r) = \frac{DD(r)\,\Delta r}{RR(r)\,\Delta r} - 1.$$



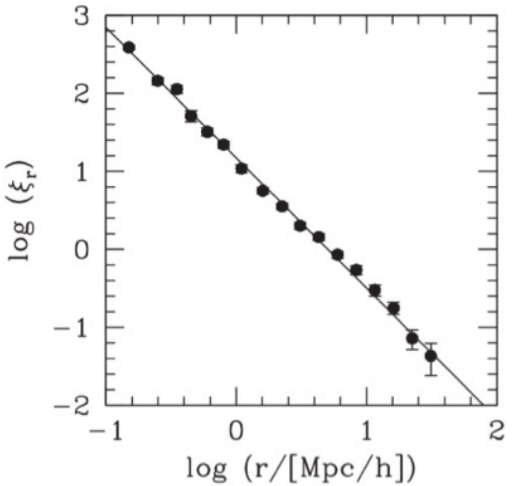
Here $DD(r)\,\Delta r$ is the number of galaxy pairs with separations in the range $r \pm \Delta r/2$, and $RR(r)\,\Delta r$ is the number that would be expected if galaxies were randomly distributed in space. Galaxies are said to be positively correlated on scale r if $\xi(r) > 0$, to be anticorrelated if $\xi(r) < 0$, and to be uncorrelated if $\xi(r) = 0$. Since it is relatively straightforward to measure, the two-point correlation function of galaxies has been estimated from various samples. In many cases, redshifts are used as distances and the corresponding correlation function is called the correlation function in redshift space. Because of peculiar velocities, this redshift-space correlation is different from that in real space. The latter can be estimated from the projected two-point correlation function, in which galaxy pairs are defined by their separations projected onto the plane perpendicular to the line-of-sight so that it is not affected by using redshift as distance (see Chapter 6 for details). Fig. 2.37 shows an example of the redshift-space correlation function and the corresponding real-space correlation function. On scales smaller than about $10\,h^{-1}$Mpc the real-space correlation function can well be described by a power law,[5]

$$\xi(r) = (r/r_0)^{-\gamma}, \tag{2.45}$$

with $\gamma \sim 1.8$ and with a correlation length $r_0 \approx 5\,h^{-1}$Mpc. This shows that galaxies are strongly clustered on scales $\lesssim 5\,h^{-1}$Mpc, and the clustering strength becomes weak on scales much larger than $\sim 10\,h^{-1}$Mpc. The exact values of γ and r_0 are found to depend significantly on the properties of the galaxies. In particular the correlation length, r_0, defined by $\xi(r_0) = 1$, is found to depend on both galaxy luminosity and color in the sense that brighter and redder galaxies are more strongly clustered than their fainter and bluer counterparts (e.g. Norberg et al., 2001,

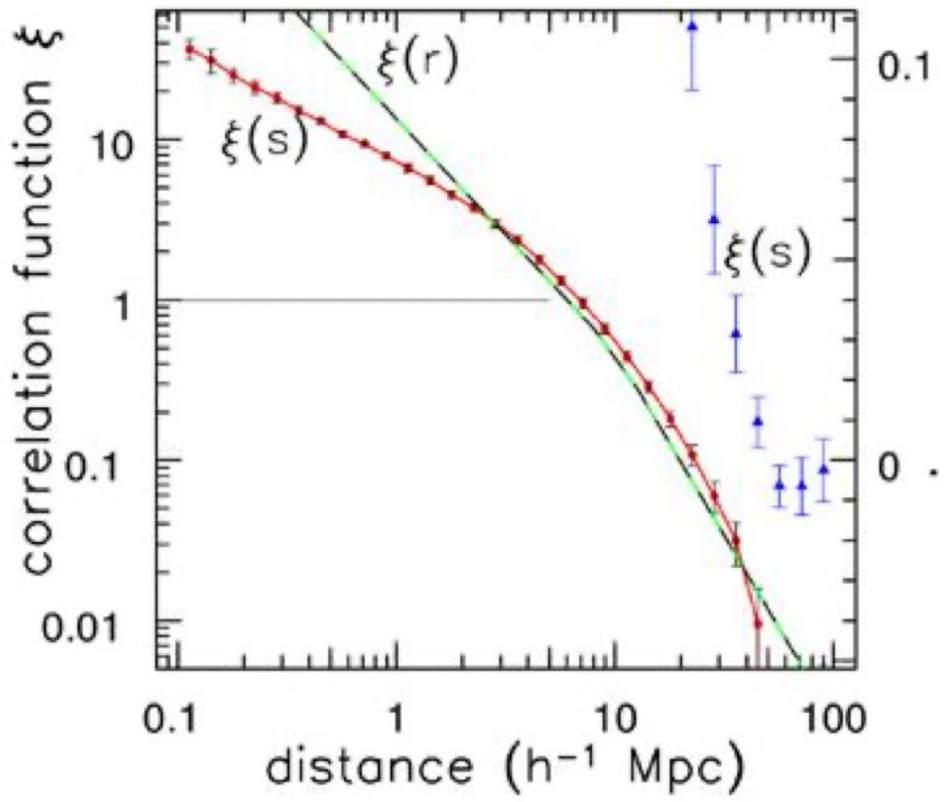


TPCF of galaxies in redshift space.



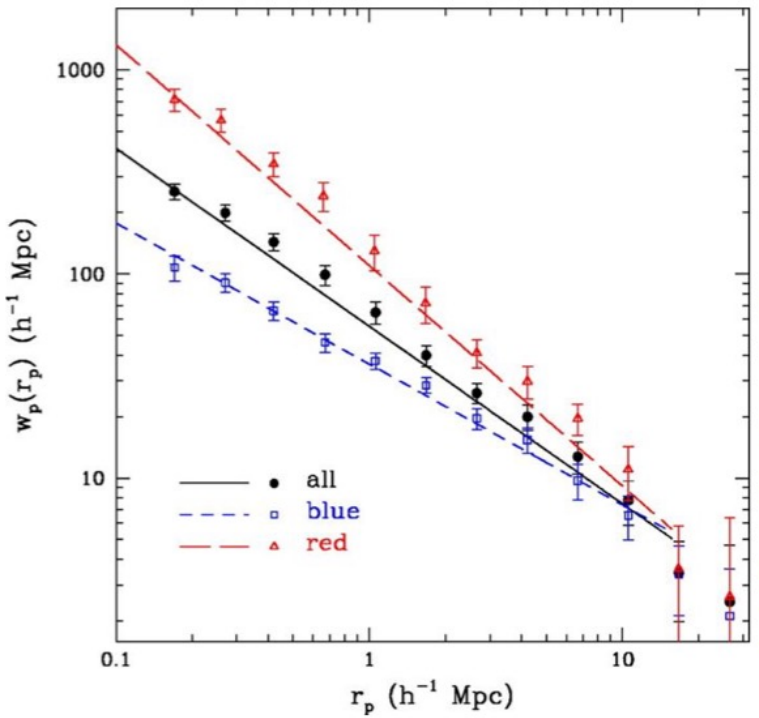
TPCF of galaxies in real space.

2dF巡天中星系的两点相关函数ξ(s)



2dF巡天中星系的两点相关函数。左边的对数标度；右边的线性标度；圆圈为小间隔的相关函数，三角为大间隔的相关函数；ξ(s) 没有改正本动速度效应；ξ(r)为改正了本动速度效应的相关函数.

- 星系相关长度 $r_0 \approx 5\ h^{-1}$ Mpc：椭圆星系的成团性强， $r_0 \approx 6\ h^{-1}$ Mpc；恒星形成星系的r_0较小，成团性弱
- 在r_0附近，斜率$\gamma \approx 1.7$。
- 对于$r \geq 50\ h^{-1}$ Mpc，ξ围绕0附近振荡：在较大尺度上，星系分布相当均匀



The red galaxy has a steeper slope and a higher amplitude at all $r_p < 10h^{-1}$Mpc.

In theoretical studies and simulations, the most massive galaxies in the early universe are placed in the most massive dark matter halos.

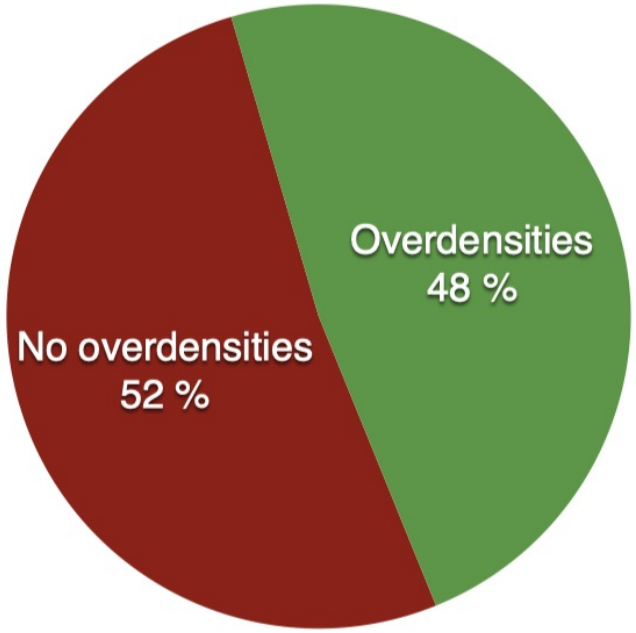
QSOs are expected to be a tracer of most massive dark matter halos because they harbour supermassive black hole

the massive dark matter halo mass can be directly related to the clustering of galaxy population

Therefore, clustering of galaxies around quasars should be observed.

Search for galaxies at optical wavelengths show confusing and contradictory results

~30 z>~4 quasar fields studied



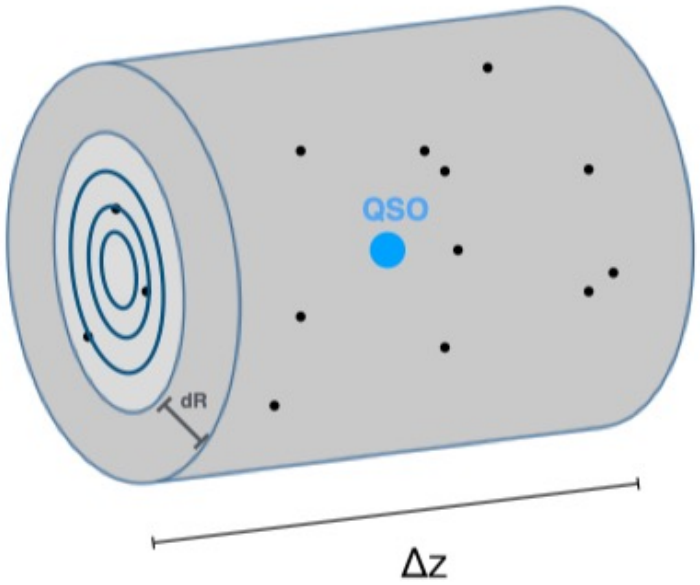
Willott+05, Kim+09, Bañados+13, Husband+13, Simpson+14, Mazzucchelli+17, Kikuta+17, Goto+17, Ota+18.

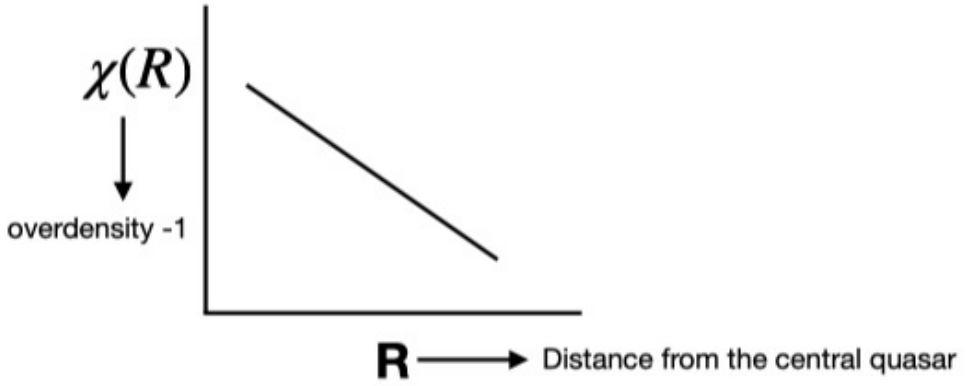
Adams+05, Stiavelli+05, Zheng+06, Kashikawa+07, Kim+09, Utsumi+10, Capak+11, Swinbank+12, Morselli+14, Balmaverde+17, Ota+18.

Some possible reasons:

- Individual quasar fields, so affected by low number statistics and high cosmic variance.
- Incomplete galaxy population traced

i) Expand the quasars sample
ii) multi-wavelength observe to find different galaxy population
iii) Using quasar-galaxy cross-correlation function, which not only quantifies the over/under density of galaxies but also their distribution about the QSO.

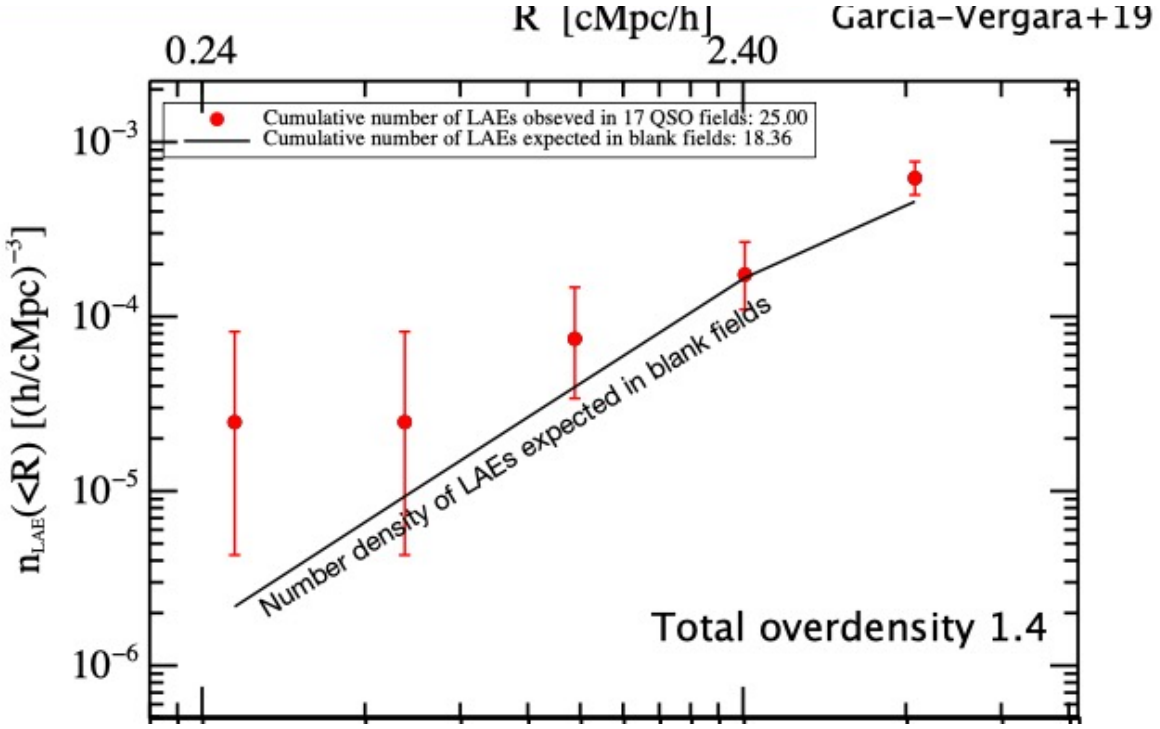


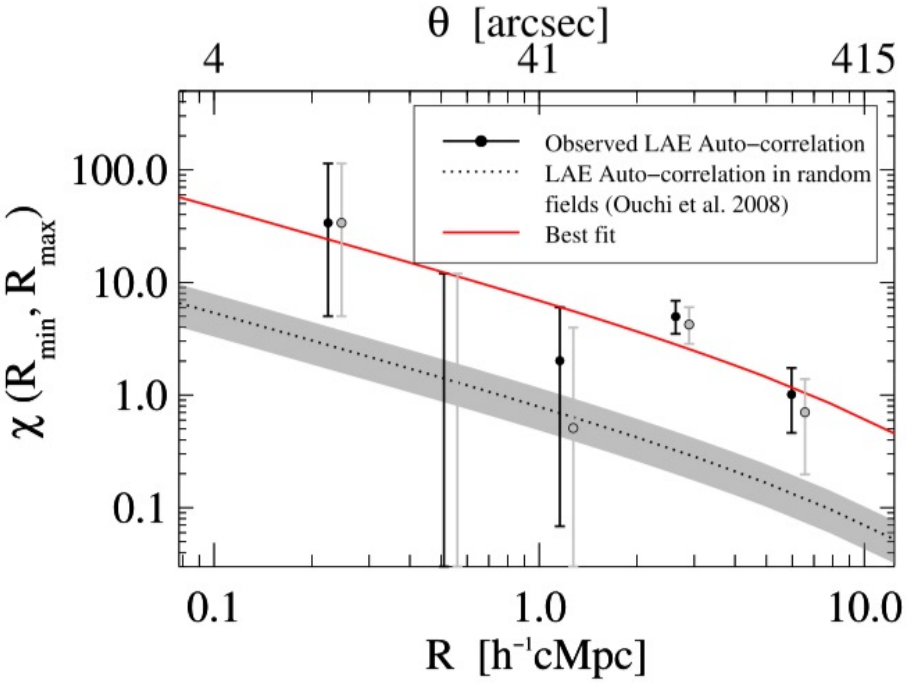


Expectation: Galaxies should be accumulated around quasar implying a large quasar-galaxy cross-correlation function.

García et al. (2017) observe 17 QSO fields at z~4 with VLT/FORS2 (20hrs) to search for Lyman alpha emitters.

They detected 25 LAEs, only 18.4 LAEs are expected in blank fields, therefore they find total overdensity 1.4.





Quasar sample and ALMA observations

the 17 quasars were selected from the Sloan Digital Sky Survey (SDSS) and the Baryon Oscillation Spectroscopic Survey (BOSS) quasar catalog.

$z \sim 3.862 - 3.879$ (corresponding to $\Delta v = 1066$ km/s at z = 3.87)

CO(4–3) rest-frame frequency is 461.04 GHz, corresponding to about 95GHz at z = 3.87.

Observations were taken in ALMA Band 3 (84-116 GHz)

Optical properties of the targeted quasars.

Target ID (SDSS)	RA (J2000)	Dec (J2000)	Redshift	*i* mag
J0040+1706	00:40:17.426	+17:06:19.78	3.873 ± 0.008	18.91
J0042-1020	00:42:19.748	-10:20:09.53	3.865 ± 0.012	18.57
J0047+0423	00:47:30.356	+04:23:04.73	3.864 ± 0.008	19.94
J0119-0342	01:19:59.553	-03:42:16.51	3.873 ± 0.013	20.49
J0149-0552	01:49:06.960	-05:52:18.85	3.866 ± 0.013	19.80
J0202-0650	02:02:53.765	-06:50:44.54	3.876 ± 0.008	20.64
J0240+0357	02:40:33.804	+03:57:01.59	3.872 ± 0.012	20.03
J0850+0629	08:50:13.457	+06:29:46.91	3.875 ± 0.013	20.40
J1026+0329	10:26:32.976	+03:29:50.63	3.878 ± 0.008	19.74
J1044+0950	10:44:27.798	+09:50:47.98	3.862 ± 0.012	20.52
J1138+1303	11:38:05.242	+13:03:32.61	3.868 ± 0.008	19.10
J1205+0143	12:05:39.550	+01:43:56.52	3.867 ± 0.008	19.37
J1211+1224	12:11:46.935	+12:24:19.08	3.862 ± 0.008	19.97
J1224+0746	12:24:20.658	+07:46:56.33	3.867 ± 0.008	19.08
J1258-0130	12:58:42.118	-01:30:22.75	3.862 ± 0.008	19.58
J2250-0846	22:50:52.659	-08:46:00.22	3.869 ± 0.012	19.44
J2350+0025	23:50:32.306	+00:25:17.23	3.876 ± 0.012	20.61

Note. — Quasar positions are determined from optical images (SDSS/BOSS quasar catalog; Pâris et al. 2014), redshifts are based on one or more rest-frame UV emission lines and calibrated to estimate the systemic redshift (see § 2.1), and magnitudes correspond to the *i*-band magnitudes from SDSS.

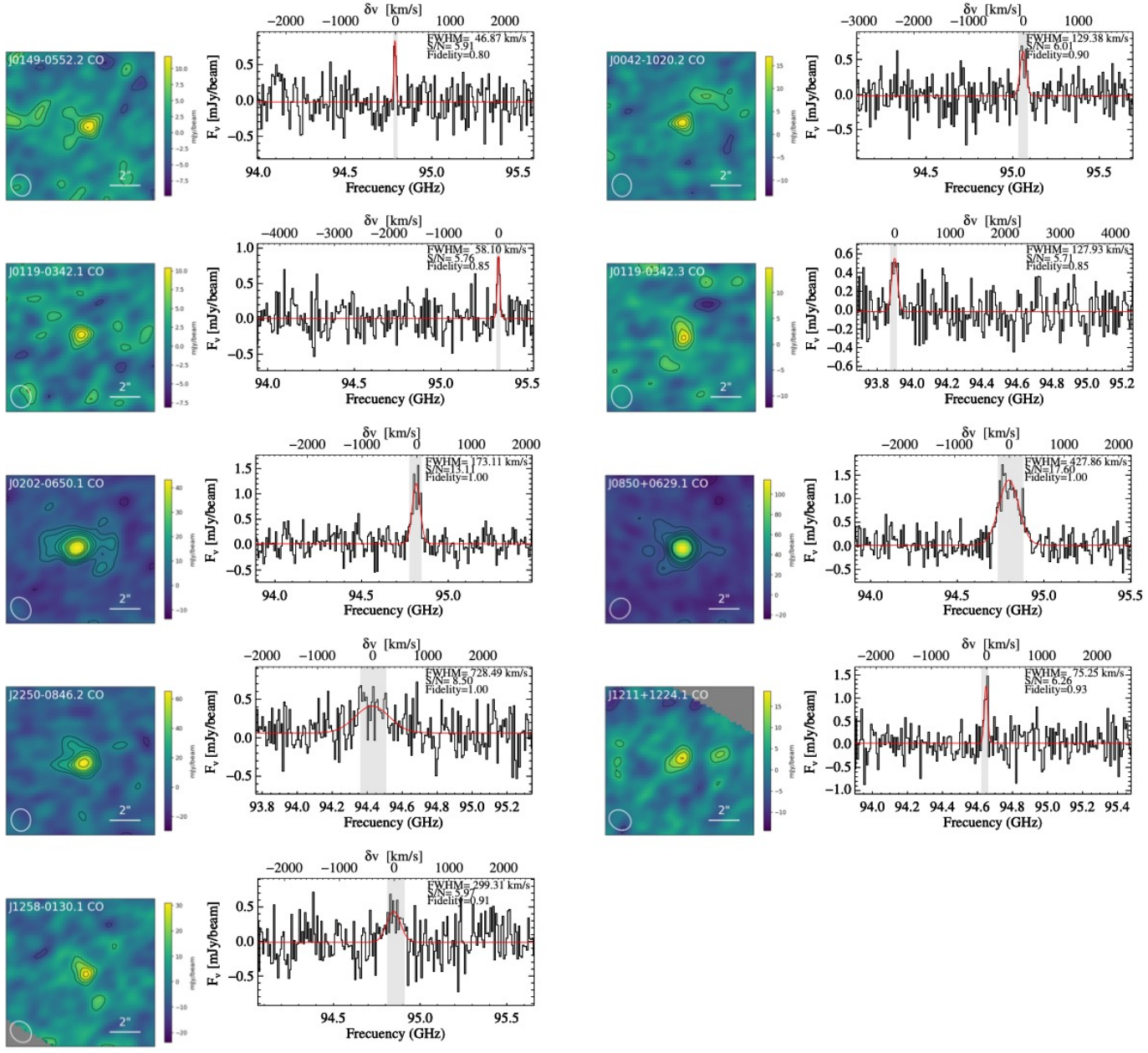
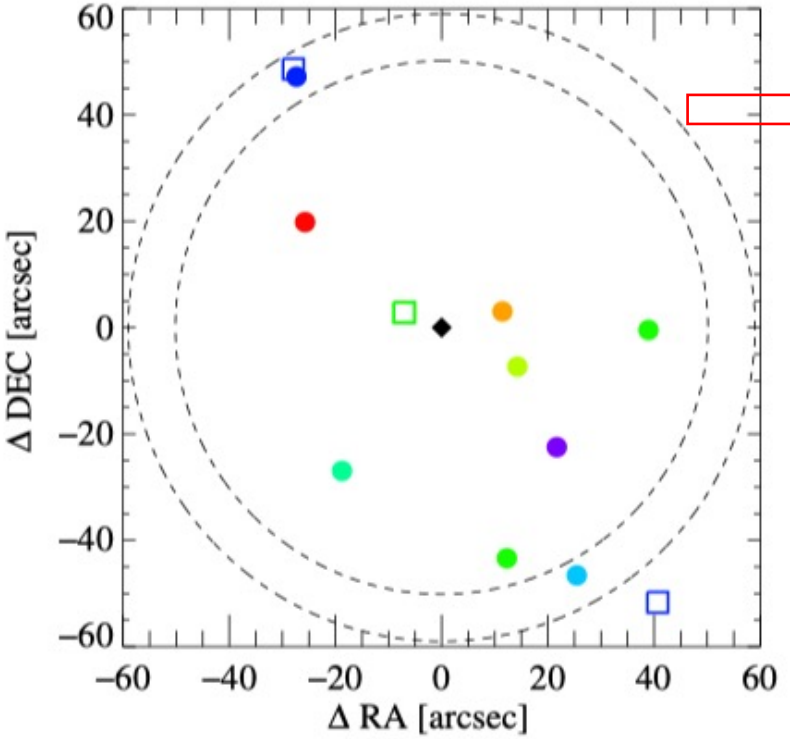


Figure 2. *Left:* Line maps for all the sources detected in our survey with fidelity ≥0.8 (or ≥0.9 if they are located at radius > 50.14″ from the pointing center), integrated over a frequency width given by $2.8\sigma_{\rm line}$ around the center of the line. Contours start at $\pm 2\sigma$ and increase in steps of 1σ. The white ellipse in the lower-left corner shows the FWHM beam size. *Right:* 1D extracted spectra on the brightest pixel of each source. We show the Gaussian plus flat continuum fit as a red curve. The shaded gray area shows the line width as detected by the line search algorithm. In each panel, we report the best-fitted FWHM value, the S/N of the line determined by the line search algorithm, and the fidelity. Line maps and spectra are both extracted from the cleaned cubes, and they are not corrected by the primary beam response.

Considering all the fields together, 9 sources with S/N>5.6 were found.



Sky distribution of the CO(4–3) line emitting galaxies around the central quasar for our 17 fields combined.

如何给出Z？ 通过比较CO(4–3) line emitting galaxies 与quasars的红移来得到

原则上是比较这两种星系CO(4–3)辐射的红移，但是17个quasars中只有10个探测到了高可信度(> 0.8)的CO(4–3)辐射.

对于其它的7个quasars，只能通过光学光谱来确定它们的红移。因此需要考察光学波段和ALMA波段红移之间是否有系统性偏移。

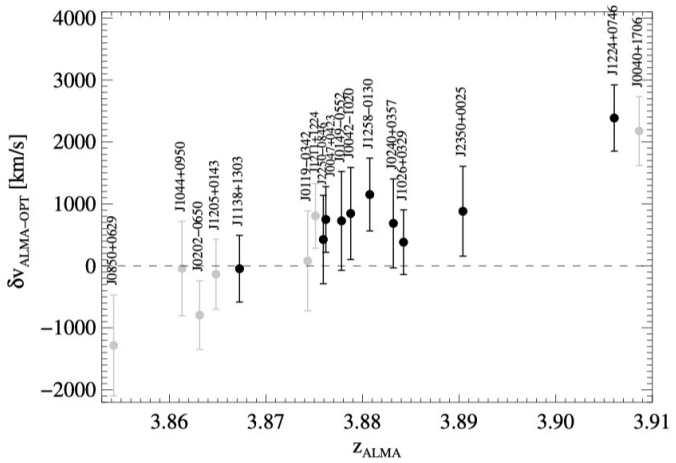


Figure 3. Velocity offset between the quasar redshifts determined from the CO(4–3) emission lines and the rest-frame UV emission lines. We include the secure quasar detections (black) and the low-fidelity detections (gray). Based only on the secure detections, we find a median offset of $|\delta v| = 738 \pm 651$ km s^{-1}.

Table 5
Quasar redshift determined from ALMA observations, and number of emission lines in each field.

Field (1)	z_{opt} (2)	z_{ALMA} (3)	δv [km s^{-1}] (4)	N_{3000} (δ_{CO}) (5)	N_{1000} (δ_{CO}) (6)	δ_{LAE} (7)
J0040+1706	3.873	3.910*	2266*	0 (00.00)	0 (00.00)	2.66
J0042-1020	3.865	3.878	805	1 (17.77)	0 (00.00)	0.00
J0047+0423	3.864	3.877	780	0 (00.00)	0 (00.00)	0.00
J0119-0342	3.873	3.874*	46*	2 (34.90)	0 (00.00)	3.41
J0149-0552	3.866	3.879	781	1 (18.61)	1 (55.87)	0.00
J0202-0650	3.876	3.864*	-741*	1 (17.63)	1 (52.91)	0.00
J0240+0357	3.872	3.883	661	0 (00.00)	0 (00.00)	0.92
J0850+0629	3.875	3.854*	-1310*	1 (19.40)	1 (58.22)	2.41
J1026+0329	3.878	3.884	380	0 (00.00)	0 (00.00)	2.16
J1044+0950	3.862	3.861*	-37*	0 (00.00)	0 (00.00)	0.78
J1138+1303	3.868	3.867	-62	0 (00.00)	0 (00.00)	0.00
J1205+0143	3.867	3.864*	-161*	0 (00.00)	0 (00.00)	1.34
J1211+1224	3.862	3.875*	805*	1 (25.97)	1 (77.95)	3.75
J1224+0746	3.867	3.905	2355	0 (00.00)	0 (00.00)	0.00
J1258-0130	3.862	3.882	1203	1 (27.34)	0 (00.00)	1.08
J2250-0846	3.869	3.876	433	1 (19.61)	1 (58.86)	0.77
J2350+0025	3.876	3.891	920	0 (00.00)	0 (00.00)	0.59
ALL				9 (10.56)	5 (17.60)	1.36

* Indicates the cases where the CO(4–3) line is detected with low (< 0.8) fidelity.

Redshift
3.873 ± 0.008
3.865 ± 0.012
3.864 ± 0.008
3.873 ± 0.013
3.866 ± 0.013
3.876 ± 0.008
3.872 ± 0.012
3.875 ± 0.013
3.878 ± 0.008
3.862 ± 0.012
3.868 ± 0.008
3.867 ± 0.008
3.862 ± 0.008
3.867 ± 0.008
3.862 ± 0.008
3.869 ± 0.012
3.876 ± 0.012

For the secure quasar detections, we find a median velocity offset of $|\delta v| = 738 \pm 651$ km/s. This is consistent with the typical reported uncertainties of the optical-based redshifts

Table 3

Properties of the emission lines detected in all the 17 fields with fidelity ≥0.8 (or ≥0.9 if they are located at radius > 50.14″from the pointing center).

ID	RA (J2000) [deg]	Dec (J2000) [deg]	ν_{line} [GHz]	$FWHM_{line}$ [km s^{-1}]	S/N	Fidelity	$\delta\theta$ [″]	δv [km s^{-1}]	Optical counterpart
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
J0042-1020.2	00:42:18.48	-10:20:36.48	95.06	129	6.01	0.90	33.0	-1767.19	$R > 25.6$, $g > 26.3$
J0119-0342.1	01:20:00.37	-03:42:59.90	95.33	58	5.76	0.85	45.1	-2266.36	$R > 25.7$, $g > 26.4$
J0119-0342.3	01:20:02.15	-03:42:16.96	93.90	128	5.71	0.85	39.0	2270.94	$R > 25.7$, $g > 26.4$
J0149-0552.2*	01:49:08.41	-05:52:41.34	94.79	47	5.91	0.80	31.3	-851.97	$R > 25.7$, $g > 26.5$
J0202-0650.1*	02:02:54.72	-06:50:51.88	94.82	173	13.11	1.00	16.1	-829.85	$R = 24.4$, $g = 25.7$ Lyman break detection with $g - R = 1.3$ located at 1.2″
J0850+0629.1*	08:50:14.22	06:29:49.93	94.80	428	17.60	1.00	11.9	-710.84	$R > 25.7$, $g > 26.3$
J1211+1224.1*	12:11:45.06	12:25:06.25	94.65	75	6.26	0.93	54.9	559.31	$R > 25.3$, $g > 26.3$ LAE located at 1.6″
J1258-0130.1	12:58:43.81	-01:31:09.35	94.85	299	5.97	0.91	53.1	-1220.26	$R > 25.7$, $g > 26.4$
J2250-0846.2*	22:50:50.92	-08:45:40.39	94.43	728	8.50	1.00	32.7	398.63	$R > 25.6$, $g > 26.4$

* included for the clustering measurement.

The authors measure the clustering using $Z\text{max} = -Z\text{min} = 8.19\, h^{-1} cMpc$ (corresponding to $\Delta v = \pm 1000$ km/s at z = 3.87)

This choice results in a total of **five sources** included in the clustering analysis

small velocity range can avoid dilution of the small-scale clustering signal

$$dP = n_G[1 + \xi(r)]dV$$

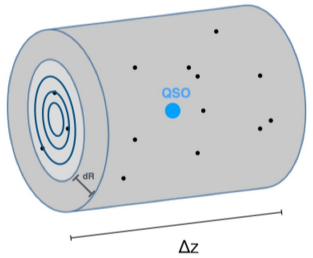
$$\xi(r) = (r/r_{0,\mathrm{QG}})^{-\gamma}$$

$\xi(r)$ is the QSO–galaxy cross-correlation function

$$\chi(R) = \frac{\int_{V_{\mathrm{eff}}} \xi(R,Z)dV}{V_{\mathrm{eff}}}$$

$$r^2 = R^2 + Z^2$$

volume-averaged projected cross-correlation function



V_{eff} is a cylindrical shell centered on quasar

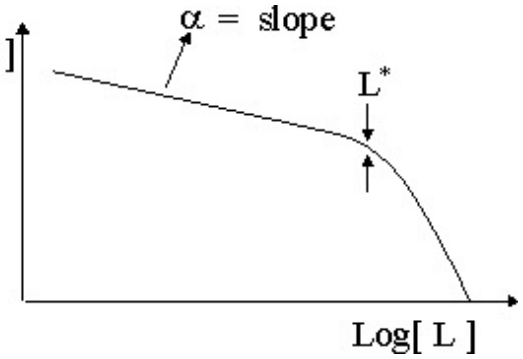
$$\chi(R) = \frac{\langle QG(R)\rangle}{\langle QR(R)\rangle} - 1$$

$\langle QG(R)\rangle$ is the number of quasar-galaxy pairs observed in our survey, within the cylindrical shell volume.

$\langle QR(R)\rangle$ is the number of quasar-galaxy pairs that is expected in random.

$$\langle QR(R)\rangle = \int_{Z\mathrm{min}}^{Z_{\mathrm{max}}} \int_{R\mathrm{min}}^{R_{\mathrm{max}}} n_{CO}(L' \geq L'_{\mathrm{Lim}}) 2\pi R dR dZ$$

$$\phi(L)\mathrm{d}L = \phi^* \left(\frac{L}{L^*}\right)^{\alpha} \exp\left(-\frac{L}{L^*}\right) \frac{\mathrm{d}L}{L^*}.$$



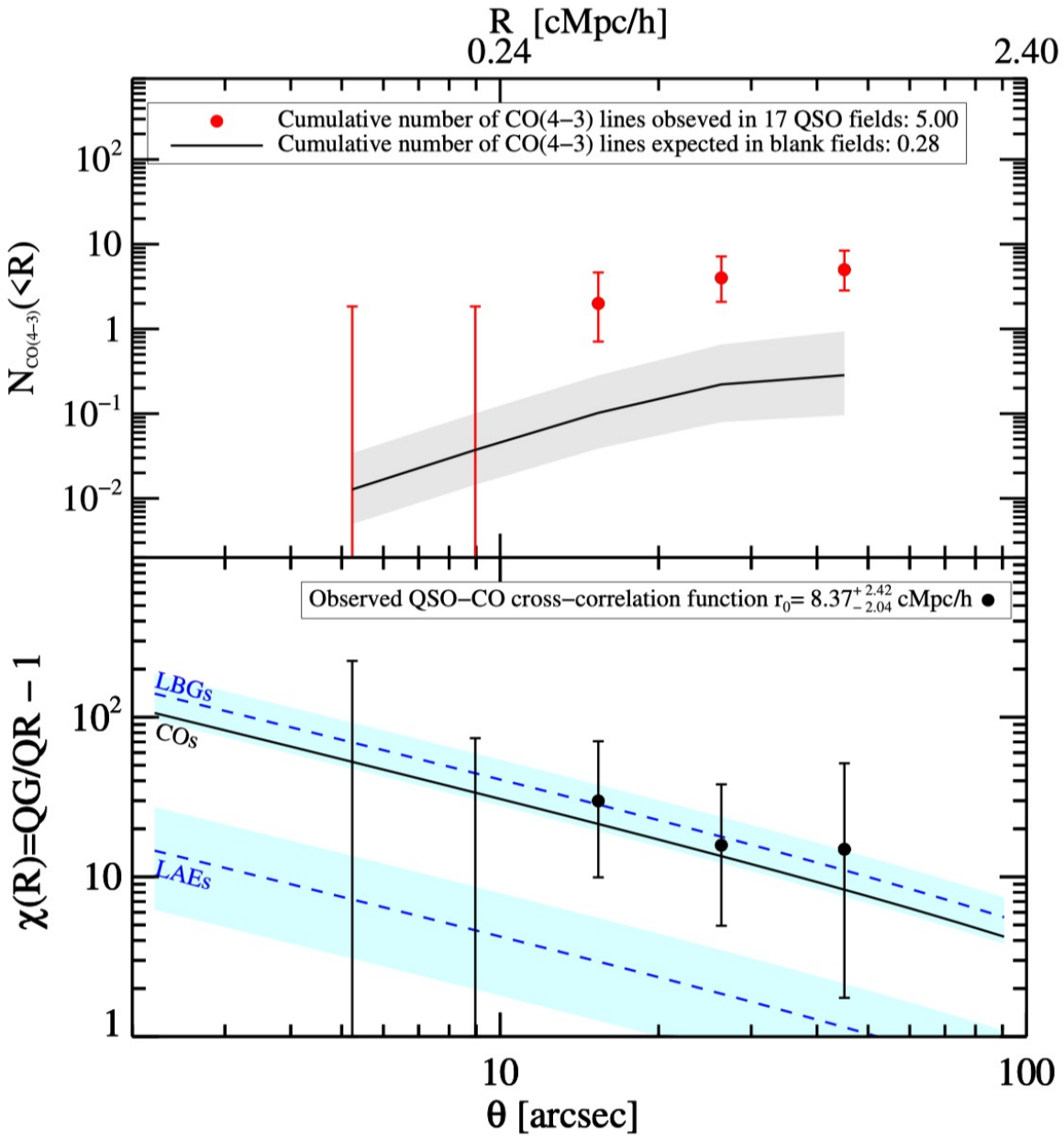
$n_{CO}(L' \geq L'_{Lim})$ is derived from the CO(4–3) luminosity function at z = 3.8 measured by Decarli et al. (2019)

Table 6
Quasar-CO cross correlation function.

R_{min} h^{-1}cMpc	R_{max} h^{-1}cMpc	$\langle QG(R)\rangle$	$\langle QR(R)\rangle$	$\chi(R)$
0.096	0.165	0.000	0.008	$-1.00^{+226.13}_{-0.00}$
0.165	0.282	0.000	0.025	$-1.00^{+74.91}_{-0.00}$
0.282	0.483	2.000	0.065	$29.89^{+40.75}_{-19.96}$
0.483	0.827	2.000	0.119	$15.79^{+22.15}_{-10.85}$
0.827	1.417	1.000	0.063	$14.88^{+36.53}_{-13.14}$

use a Poisson maximum likelihood estimator to determine the real-space cross-correlation parameter $r_{0,QG} = 8.37^{+2.42}_{-2.04}\ h^{-1}Mpc\ \ (\gamma = 1.8)$

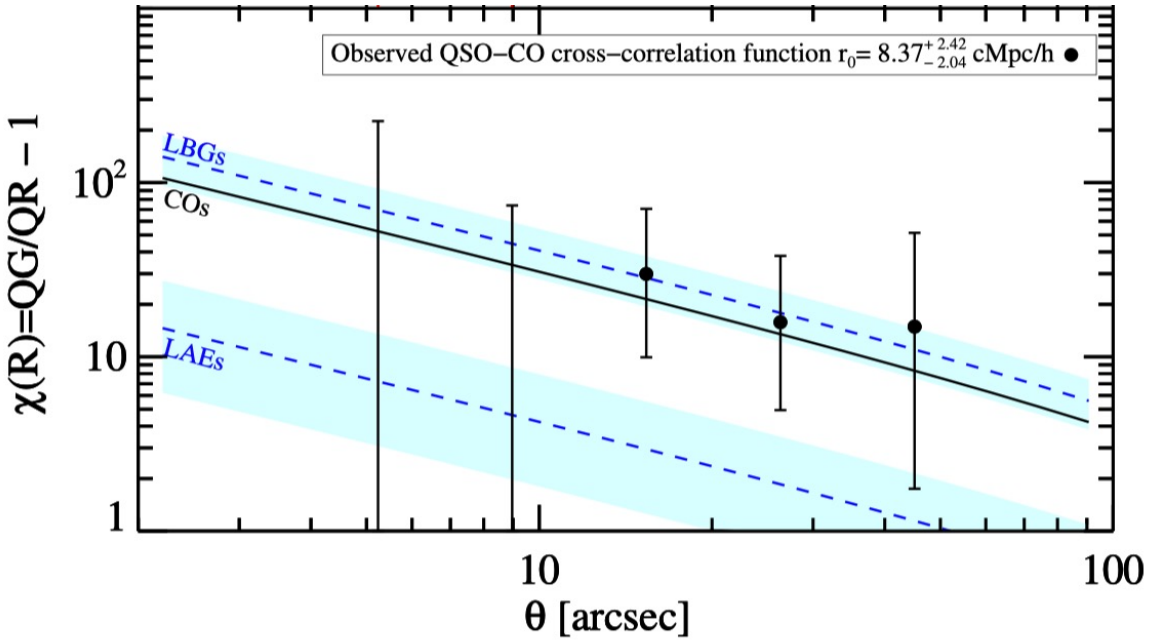
In total, **5 CO(4–3)** lines are detected in quasar fields, only **0.28 CO(4–3)** lines are expected in blank fields, resulting in a total CO(4–3) line **overdensity of** $\mathbf{17.6^{+11.9}_{-7.6}}$ in quasar fields.



$r_{0,Q-CO} = 8.37^{+2.42}_{-2.04}\ h^{-1}Mpc \qquad (\gamma = 1.8)$

$r_{0,Q-LBG} = 9.78^{+1.68}_{-1.86}\ h^{-1}Mpc$ (García et al. 2017)

$r_{0,Q-LAE} = 2.78^{+1.16}_{-1.05}\ h^{-1}Mpc$ (García et al. 2019)



LBG and LAE clustering studies extend up to larger scales ($R \leq 9\ h^{-1}Mpc$) than CO study ($R \leq 1.5\ h^{-1}Mpc$)

The authors assume the small-scale quasar-CO crosscorrelation function can be extrapolated towards larger scales following a single power-law shape.

$r_{0,Q-CO}$ is slightly lower than $r_{0,Q-LBG}$, but is 3 times higher than $r_{0,Q-LAE}$

Auto-correlation of CO emitters at z ~ 4

QSO-galaxy cross-correlation function can be written as $\xi_{QG} = \sqrt{\xi_{QQ}\xi_{GG}}$

assume that they al have the same slope γ = 1.8

$r_{0,QG} = \sqrt{r_{0,QQ}r_{0,GG}}$

$r_{0,QQ} = 22.3 \pm 2.5\ h^{-1}Mpc$ (Shen et al. 2007) → $r_{0,GG} = 3.14 \pm 1.71\ h^{-1}Mpc$

$r_{0,LBG} = 4.1^{+0.2}_{-0.2}\ h^{-1}Mpc$ (Ouchi et al. 2004)

$r_{0,LAE} = 2.74^{+0.58}_{-0.72}\ h^{-1}Mpc$ (Ouchi et al. 2010)

CO emitters would inhabit dark matter halos with similar masses as these hosting LBGs and LAEs

$r_{0,Q-CO} = 8.37^{+2.42}_{-2.04}\ h^{-1}Mpc$

$r_{0,Q-LBG} = 9.78^{+1.68}_{-1.86}\ h^{-1}Mpc$

$r_{0,Q-LAE} = 2.78^{+1.16}_{-1.05}\ h^{-1}Mpc$

The discrepancy may related to physical processes affecting the visibility of the LAEs around quasars.

i) Galaxies around quasars could be more dusty, affecting the visibility of the Lyα line.

ii) A relatively small star formation efficiency in galaxies around quasars could explain the lack of LAEs in these fields.

Summary

- The first quasar sample targeted for clustering studies of both optical and dusty galaxies.
- Strong clustering of CO emitters around quasar is detected (overdensity of $17.6^{+11.9}_{-7.6}$).
- previous studies also find clustering of LBG and LAE around quasar, indicates quasars trace massive structures in the early universe.
- The first constraints on the clustering of auto-correlation of CO emitters at $z \sim 4$.
- The clustering of CO emitters is comparable with the clustering of LBG, but is significantly higher than the LAE clustering around quasars.
- Low SF efficiency in these galaxies or possibly an excess of dust in galaxies around quasars.