THE UBIQUITY OF COEVAL STARBURSTS IN MASSIVE GALAXY CLUSTER PROGENITORS

Caitlin M. Casey

Department of Astronomy, The University of Texas at Austin, 2515 Speedway Boulevard Stop C1400, Austin, TX 78712, USA; cmcasey@astro.as.utexas.edu

Received 2015 December 11; accepted 2016 March 10; published 2016 June 8

ABSTRACT

The universe’s largest galaxy clusters likely built the majority of their massive $>10^{11}$ $M_\odot$ galaxies in simultaneous, short-lived bursts of activity well before virialization. This conclusion is reached based on emerging data sets for $z > 2$ proto-clusters and the characteristics of their member galaxies, in particular, rare starbursts and ultraluminous active galactic nuclei (AGN). The most challenging observational hurdle in identifying such structures is their very large volumes, $\sim 10^4$ comoving Mpc$^3$ at $z > 2$, subtending areas of approximately half a degree on the sky. Thus, the contrast afforded by an overabundance of very rare galaxies in comparison to the background can more easily distinguish overdense structures from the surrounding, normal density field. Five $2 \lesssim z \lesssim 3$ proto-clusters from the literature are discussed in detail and are found to contain up to 12 dusty starbursts or luminous AGN galaxies each, a phenomenon that is unlikely to occur by chance even in overdense environments. These are contrasted with three higher-redshift ($4 \lesssim z \lesssim 5.5$) dusty star-forming galaxy (DSFG) groups, whose evolutionary fate is less clear. Measurements of DSFGs’ gas depletion times suggest that they are indeed short-lived on $\sim 100$ Myr timescales, and accordingly the probability of finding a structure containing more than 8 such systems is $\sim 0.2\%$, unless their “triggering” is correlated on very large spatial scales, $\sim 10$ Mpc across. The volume density of DSFG-rich proto-clusters is found to be comparable to all of the $>10^{15}$ $M_\odot$ galaxy clusters in the nearby universe, which is a factor of five larger than expected in some simulations. Some tension still exists between measurements of the volume density of DSFG-rich proto-clusters and the expectation that they are generated via short-lived episodes, as the latter suggests that only a fraction ($< \frac{1}{2}$) of all proto-clusters should be rich with DSFGs. However, improved observations of proto-clusters over large regions of sky will certainly shed more light on the assembly of galaxy clusters, and whether or not they build their galaxies through episodic bursts as suggested here.

Key words: galaxies: clusters: general – galaxies: star formation – large-scale structure of universe – submillimeter: galaxies

1. INTRODUCTION

The environmental dependence of galaxies’ evolution is observationally elusive. Locally, it is clear that galaxies residing in the most massive environments exhibit characteristics that are markedly different from their counterparts in the field: they are more massive (e.g., Collins et al. 2009; van der Burg et al. 2013), they are forming relatively few stars (Balogh et al. 1998; Lewis et al. 2002), they are preferentially red (Wake et al. 2005), and they lack spiral structure (Skibba et al. 2009). At their cores, hot inter-cluster gas—containing $\sim 90\%$ of the cluster’s baryonic matter—renders these massive systems easy to detect via their emission of Bremsstrahlung radiation in the X-ray (see review of Kravtsov & Borgani 2012). These threads of observational evidence, combined with knowledge of density fluctuations in the early universe imprinted on the cosmic microwave background (Sheth et al. 2001), have formed the backbone of our understanding of hierarchical growth in galaxy formation (Springel 2005; Vogelsberger et al. 2014). Higher-density environments saw accelerated evolution by forming most of their galaxies early and coalescing at earlier times. What does this imply for observations of overdense environments at high redshift?

In agreement with hierarchical expectations, some works have observed a reversal of the star formation-density relation at $z \sim 1$ (Elbaz et al. 2007; Cooper et al. 2008), whereby galaxies in overdense environments at high redshift are more likely to be star-forming than field galaxies or similar-mass galaxies in overdensities in the local universe. However, several other works do not see this reversal (Patel et al. 2009; Bolzonella et al. 2010; Cucciati et al. 2010; Scoville et al. 2013), which has led to some uncertainty concerning the processes driving the evolution of clusters at early times.

Observations of clusters in the early universe ($z > 1$) themselves also have considerable potential as tools for testing galaxy formation theory in a cosmological context and placing independent constraints on fundamental cosmological parameters. For example, discovering a single cluster of sufficient mass at $z \geq 2$ ($M_{\text{halo}} \sim 5 \times 10^{14} M_\odot$) can place significant constraints on current cosmological models (e.g., Harrison & Coles 2012), just as the discovery of a population of early massive galaxies may already challenge that paradigm (Steinhardt et al. 2015). Hence, several observational efforts to identify high-redshift overdensities have been pursued over the past few decades (Subramanian & Swarup 1992; Steidel et al. 1998, 2005; Stevens et al. 2003; Miley et al. 2004; Doherty et al. 2010; Noble et al. 2013; Clements et al. 2014; Rigby et al. 2014; Planck Collaboration et al. 2015).

Unfortunately, detecting galaxy clusters at $z > 2$ has proved especially challenging. While X-ray searches are efficient at selecting massive clusters through the emission of hot gas at $z \lesssim 1.5$ (e.g., Rosati et al. 2002), the rapid surface brightness dimming of X-ray emission makes it an inefficient observable at high redshift. Other techniques for identifying cluster environments are similarly limited to $z \lesssim 2$, such as optical searches for the galaxy red sequence, which demonstrates the presence of an evolved galaxy population (Gladders &
2. DSFG/AGN-RICH PROTO-CLUSTERS

Here, I present the existing observational characteristics for overdense structures at $z \geq 2$ with robust spectroscopic redshifts and an overabundance of DSFGs or luminous AGN. The importance of an overabundance of DSFGs or luminous AGN is key: these types of galaxies are $\geq 100$ times more rare than most “normal” $L_\star$ galaxies across all epochs. Their rarity is what makes them useful for studying high-redshift overdensities, not only because they represent a potentially critical evolutionary stage for early, massive galaxy formation (Toft et al. 2014), but also because a group of them in close proximity is exceedingly rare and can easily identify an overdense structure which is too large to be identified through more common galaxy populations. Furthermore, as will be discussed in later sections, they can place unique constraints on the assembly histories of proto-clusters.

Their star formation rates, dark matter halo masses, structure volumes, and respective population overdensities are estimated below and are discussed in the context of each proto-cluster’s observations. The star formation rates are computed with careful treatment of individual dust-obscured starbursts, which will dominate the calculations of SFRs, as well as a rough constraint on the contribution from other optically selected members such as LBGs. Dark matter halo masses are estimated using abundance matching techniques (Behroozi et al. 2013), requiring estimates of each individual member’s stellar mass, unless stated otherwise. Due to shear numbers, the halo mass is dominated by optically identified member galaxies. Third, any available information on the physical extent of the structure is summarized, e.g., its occupied volume, although the uncertainty of such an estimate should be emphasized. Due to spectroscopic incompleteness, all of these estimates may be viewed as lower limits in physical terms, but can be regarded as representative of existing observable constraints.

Galaxy overdensities are quantified with the measurement of $\delta_{\text{gal}} = (N_{\text{gal}} - N_{\exp})/N_{\exp}$, where $N_{\text{gal}}$ is the observed number of member galaxies and $N_{\exp}$ is the expected number of galaxies in the same volume of blank field or the expected cosmic density. The expected number of galaxies is determined using known luminosity functions for “normal” galaxies like LBGs (e.g., Reddy & Steidel 2009; van der Burg et al. 2010), X-ray AGN (Silverman et al. 2008), and DSFGs (Casey et al. 2014). Different survey depths of different fields are taken into account when determining how prevalent a given population may be. The observational characteristics of all proto-clusters, as discussed in this section, are summarized in Table 1.

2.1. GOODS-N Structure at $z = 1.99$

Blain et al. (2004) and Chapman et al. (2009) identified a particularly DSFG-rich proto-cluster at $z = 1.99$ in the Hubble Deep Field North. The structure contains at least 24 optically selected, spectroscopically confirmed members in addition to 11 rare types of galaxies spanning the entire GOODS-N field of view (4 submillimeter galaxies, 10 radio galaxies, and 6 X-ray galaxies, all of which have substantial overlap). While Chapman et al. (2009) present potentially as many as nine DSFGs, a few of those are spurious spikes in the original SCUBA maps and others are only bright radio galaxies (A. Barger 2016, private communication).

Using deep data in the HDF and more recent collections of deep submillimeter data from Herschel (Oliver et al. 2012) and...
Table 1

| Name                        | $z$  | $N_{gal}$ | $\delta_{gal}$ | $N_{star}$ | $\delta_{star}$ | References | Total Stellar Mass ($M_\odot$) | Halo Mass at $z$ ($M_\odot$) | Volume ($cMpc^3$) | Total SFR ($M_\odot$ yr$^{-1}$) |
|-----------------------------|------|-----------|----------------|------------|----------------|------------|--------------------------------|-------------------------|----------------|--------------------------------|
| **Genuine DSFG-rich Proto-clusters:** |      |           |                |            |                |            |                                |                         |                |                               |
| GOODS-N proto-cluster       | 1.99 | 34        | 2.5            | 11         | 10             | 3, 5       | $6.5 \times 10^{11}$            | $(6 \pm 3) \times 10^{13}$ | 9000           | 2600 $\pm$ 300                 |
| COSMOS $z = 2.10$ proto-cluster | 2.10 | $\sim$100 | 8              | 10         | 13             | 13, 19     | $1.9 \times 10^{12}$            | $(1.7 \pm 1.2) \times 10^{14}$ | 15000          | 5300 $\pm$ 600                  |
| MRC 1138–256 proto-cluster  | 2.16 | $\sim$80  | 12             | 5          | 12             | 2, 10, 14  | $\sim 1 \times 10^{12}$         | $\sim 1 \times 10^{14}$      | 8000           | 2200 $\pm$ 500                  |
| COSMOS $z = 2.47$ proto-cluster | 2.47 | 57        | 3.3            | 12         | 10             | 15, 16, 17 | $1.0 \times 10^{12}$            | $(8 \pm 3) \times 10^{13}$   | 15000          | 4500 $\pm$ 500                  |
| SSA22 proto-cluster         | 3.09 | $\sim$280 | 39             | 12         | 10             | 1, 4, 7, 8, 18 | $\sim 2 \times 10^{10}$         | $(8 \pm 4) \times 10^{13}$   | 21000          | 5700 $\pm$ 800                  |
| **Other identified DSFG-rich Overdensities:** |      |           |                |            |                |            |                                |                         |                |                               |
| GN20 overdensity            | 4.05 | 8         | $\ldots$       | 3          | $>100$         | 6, 12      | $2.8 \times 10^{11}$            | $(2 \pm 0.4) \times 10^{12}$ | $^a$            | 1500 $\pm$ 800                  |
| HDF 850.1 overdensity       | 5.18 | 13        | 3.6            | 2          | 6              | 11         | $\ldots$                       | $\ldots$                  | 20000          | 850 $\pm$ 300                   |
| AzTEC-3 overdensity         | 5.30 | 11        | 30             | 2          | 80             | 9          | $\sim 2 \times 10^{10}$         | $\sim 4 \times 10^{11}$      | $\gtrsim 500$   | 1600 $\pm$ 500                  |

Note. $^a$ The GN20 structure is notably small as an association of three galaxies (or a total of eight, including candidates); the estimation of its occupied volume is thus quite uncertain.

References. References are 1—Steidel et al. (1998), 2—Kurk et al. (2000), 3—Blain et al. (2004), 4—Hayashino et al. (2004), 5—Chapman et al. (2009), 6—Daddi et al. (2009), 7—Tamura et al. (2009), 8—Lehmer et al. (2009), 9—Capak et al. (2011), 10—Kuiper et al. (2011), 11—Walter et al. (2012), 12—Hodge et al. (2013a), 13—Yuan et al. (2014), 14—Dannerbauer et al. (2014), 15—Casey et al. (2015), 16—Diener et al. (2015), 17—Chiang et al. (2015), 18—Umehata et al. (2015), and 19—C.-L. Hung et al. (2016, in preparation).

Table 2

| Name                        | $z$  | Position | Solid Angle (deg$^2$) | Galaxy Density (deg$^{-2}$) |
|-----------------------------|------|----------|----------------------|-----------------------------|
| GOODS-N proto-cluster       | 1.99 | 12:36:30+62:13:00 | 0.17 $\times$ 0.17     | 1200                         |
| COSMOS $z = 2.10$ proto-cluster | 2.10 | 10:00:23+02:15:07 | 0.34 $\times$ 0.13     | 2300                         |
| MRC 1138–256 proto-cluster  | 2.16 | 11:40:48–26:28:00 | 0.20 $\times$ 0.10     | 4000                         |
| COSMOS $z = 2.47$ proto-cluster | 2.47 | 10:00:31+02:22:22 | 0.33 $\times$ 0.42     | 400                          |
| SSA22 proto-cluster         | 3.09 | 22:17:34+00:15:01 | 0.33 $\times$ 0.50     | 1700                         |
| GN20 overdensity            | 4.05 | 12:37:11+62:22:05 | 0.01 $\times$ 0.01     | 600                          |
| HDF 850.1 overdensity       | 5.18 | 12:36:52+62:12:26 | 0.10 $\times$ 0.13     | 1000                         |
| AzTEC-3 overdensity         | 5.30 | 10:00:20+02:35:20 | 0.003 $\times$ 0.003   | $1.2 \times 10^6$            |

Notes. Positions and observed “sizes” of high-$z$ DSFG-rich proto-clusters. Note the large variation in proto-cluster solid angle and confirmed galaxy density (i.e., number of galaxies belonging to the proto-cluster in its solid angle). It is likely that these sizes and perceived galaxy densities are limited by observational selection effects and are not representative of the structures’ true physical characteristics.

Scuba-2 (Chen et al. 2013b), I re-derive far-infrared spectral energy distributions (SEDs) for this GOODS-N structure’s DSFGs using a simple modified blackbody and power-law prescription (Casey 2012). The modified blackbody dominates the SED fit at rest-frame wavelengths $\gtrsim 40$ μm, and the mid-infrared power law dominates from $5 \lesssim \lambda \lesssim 40$ μm. Note that the calculation of the star formation rates is largely insensitive to the far-infrared SED fitting technique used, as differences in methods are typically much less than measurement uncertainty (see Section 4.2 of Casey et al. 2014). The far-infrared photometry is provided in Table 3. Stellar masses and star formation rates for non-DSFG members are determined via detailed optical and near-infrared SED fitting with MAGPHYS (da Cunha et al. 2008) to rest-frame UV data from Spitzer IRAC available in GOODS-N (Capak et al. 2004). The median stellar mass for non-DSFG members is $6 \times 10^{10} M_\odot$ and the median SFR is $20 M_\odot$ yr$^{-1}$. The total stellar mass for identified cluster members is $6.5 \times 10^{11} M_\odot$ and the total star formation rate is $2600 \pm 300 M_\odot$ yr$^{-1}$. The aggregate star formation rate is dominated (88%) by the DSFGs, as is the stellar mass total (70%).

The stellar masses of the GOODS-N proto-cluster members can be checked by extrapolating Spitzer IRAC photometry to rest-frame 1.6 μm (Hainline et al. 2009), as done in Chapman et al. (2009). Although the star formation histories of DSFGs are quite uncertain, which compounds in the assumed mass-to-light ratio, I apply a $L_{H\alpha}/M_* = 7.9^{+0.8}_{-0.6}$ $L_\odot$/mag (Hainline et al. 2011) for DSFGs and LBGs alike and derive a total integrated stellar mass of $1.3 \times 10^{12} M_\odot$, which is within a factor of two of the SED estimate. Treating each galaxy as its own halo (which is appropriate given the spatial distribution of such structures), a total dark matter halo mass is inferred for the proto-cluster of $(6 \pm 3) \times 10^{13} M_\odot$ at $z = 1.99$. Assuming an exponential growth in agreement with large-box simulations (Wechsler et al. 2002), this proto-cluster would grow to a mass of $(9 \pm 5) \times 10^{14}$ at $z = 0$.

The comoving volume is calculated within a $10^7 \times 10^9$ area and approximate redshift bounds of $1.982 < z < 2.010$ as $9000 cMpc^3$. Most of this is along the line of sight, as the spatial coverage for deep spectra does not extend significantly beyond the deep GOODS-N Hubble Space Telescope coverage. It would be surprising if this structure is not extended.
Table 3

FIR Photometric Characteristics of DSFGs in the HDF $z = 1.99$ Structure

| Name          | $\bar{z}$ | Alt Name* | $S_{250}$ (mJy) | $S_{150}$ (mJy) | $S_{100}$ (mJy) | $S_{850}$ (mJy) | $S_{14}$ (μJy) | $L_{IR}$ ($L_\odot$) | SFR ($M_\odot$ yr$^{-1}$) | References |
|---------------|-----------|-----------|-----------------|-----------------|-----------------|-----------------|--------------|-----------------------|--------------------------|------------|
| DSFG J123600.13+621047.2 | 1.994 | SMG-93 | ... | 12.9 ± 4.9 | 13.1 ± 4.5 | 7.9 ± 2.4b | 128.5 ± 8.1 | (1.1$_{0.6}^{+1.1}$)$_{12}^{12}$ | 100$_{90}^{170}$ | 1 |
| DSFG J123618.32+621550.5 | 1.994 | SMG132 | 22.9 ± 4.5 | 30.0 ± 5.3 | 24.8 ± 5.4 | 7.3 ± 1.1 | 172.0 ± 8.4 | (3.5$_{0.9}^{+1.5}$)$_{12}^{12}$ | 330$_{150}^{210}$ | 1, 4 |
| DSFG J123621.25+621708.3 | 1.988 | 140e+w | 25.1 ± 4.5 | 19.8 ± 4.9 | 7.5 ± 4.8 | 7.8 ± 1.9 | 169.4 ± 8.8 | (3.3$_{1.2}^{+3.3}$)$_{12}^{12}$ | 310$_{150}^{310}$ | 1 |
| DSFG J123635.57+621424.0 | 2.001 | SMG172 | 21.1 ± 4.5 | 11.0 ± 5.0 | ... | 5.5 ± 1.4 | 77.0 ± 7.8 | (2.8$_{1.0}^{+3.8}$)$_{12}^{12}$ | 260$_{150}^{310}$ | 1 |
| DSFG J123711.99+621325.6 | 1.992 | SMG255 | 14.9 ± 4.5 | 13.1 ± 5.0 | ... | 4.2 ± 1.4 | 50.2 ± 8.1 | (2.0$_{1.0}^{+2.0}$)$_{12}^{12}$ | 190$_{150}^{210}$ | 1, 2, 4 |
| DSFG J123711.32+621330.9 | 1.993 | SFRG254 | 38.0 ± 4.5 | 34.7 ± 5.0 | 25.0 ± 5.0 | ... | 79.6 ± 17.2 | (5.7$_{2.3}^{+3.3}$)$_{12}^{12}$ | 540$_{400}^{650}$ | 1, 2, 4 |
| RAD J123632.53+620759.8 | 1.993 | SMG169 | ... | ... | ... | 5.5 ± 1.3b | 80.4 ± 8.6 | (3.9$_{0.6}^{+1.1}$)$_{12}^{12}$ | 40$_{30}^{50}$ | 1 |
| RAD J123617.54+621540.7 | 1.993 | d130 | ... | ... | ... | ... | 200.0 ± 12.8 | ... | <12 | 1, 3 |
| RAD J123640.73+621011.0 | 1.977 | SFRG179 | ... | ... | ... | ... | 72.5 ± 8.3 | ... | <50 | 1 |

Notes. References noted in the last column are 1—Chapman et al. (2009), 2—Casey et al. (2009a), 3—Casey et al. (2009b), 4—Bothwell et al. (2010).

* Alt Name is the alternate name used for this source throughout the literature, and as stated in Chapman et al. (2009).

b The original 850 μm flux densities as measured by SCUBA for SMG-93 and SMG169 are inconsistent with more recent 850 μm follow-up with SCUBA-2 (Chen et al. 2013b).

c Source SMG140e+w is a double radio source within a single SCUBA beam; the second radio source has flux density 63.4 ± 10.6 μJy.

d Source SMG130 was originally thought to be a submillimeter-faint star-forming radio galaxy (SFRG/OFRG) in Chapman et al. (2004b) but was later revealed through high-resolution radio imaging to be a low-luminosity AGN in an evolved galaxy (Casey et al. 2009b).
spatially beyond the limited field of view of the GOODS-N pencil-beam survey.

2.2. COSMOS Structure at \( z = 2.10 \)

Yuan et al. (2014) identify a Virgo-like progenitor in the COSMOS field at \( z = 2.095 \) with 57 spectroscopic members with a cluster velocity dispersion measured to be \( \sigma = 552 \pm 52 \, \text{km s}^{-1} \). The proto-cluster was revealed through spectroscopic follow-up of a zFOURGE candidate cluster at \( z = 2.2 \) identified with photometric techniques by Spitzer et al. (2012), and they predict a halo mass at \( z \sim 0 \) of \( 10^{14.4 \pm 0.3} M_\odot \).

From our own Keck MOSFIRE programs to follow-up on SCUBA-2 and Herschel-selected DSFGs in the COSMOS field, there are seven spectroscopically identified DSFGs coincident with this structure, four of which are published in Casey et al. (2012b). The details of this proto-cluster, as well as its remaining DSFGs and AGN, of which there are 10 total, will be discussed in more detail in C.-L. Hung et al. (2016, in preparation). The DSFGs reach well beyond the original bounds of the structure identified in Yuan et al. (2014), and an LBG overdensity exists across \( \sim 30' \) scales from \( z\text{COSMOS} \) samples (Lilly et al. 2009). The DSFG overdensity, centered at \( z = 2.10 \), is measured to be \( \delta_{\text{DSFG}} = 13 \), with a corresponding LBG overdensity (measured from \( z\text{COSMOS} \)) of \( \delta_{\text{LBG}} = 8 \).

The extensive 30+ bands of imaging in the COSMOS field are used to infer stellar masses and star formation rates from SEDs with MAGPHYS, all of the details of which will be given in Hung et al. The aggregate stellar mass for these sources is \( 1.9 \times 10^{12} M_\odot \) and the star formation rate is \( 5300 \pm 600 \, M_\odot \, \text{yr}^{-1} \). A lower limit on the volume for this structure is placed at \( 15,000 \, \text{cMpc}^3 \), using a sky area coverage of \( 8' \times 20' \). While one of the DSFGs lies significantly outside of this area, and could easily justify a doubling of the volume, spectroscopic incompleteness in that patch of sky significantly limits our ability to assess the structure’s extent.

2.3. MRC1138–256, or the “Spiderweb Galaxy” Structure at \( z = 2.16 \)

This structure was originally characterized in Kurk et al. (2000) and has a number of candidate LAEs as well as HAEs. The most notable member is the “Spiderweb Galaxy” described by Kuiper et al. (2011): a radio-loud starburst with a luminous AGN and a giant Ly\( \alpha \) halo. Dannerbauer et al. (2014) present submillimeter data of the area and point to a number of identified DSFGs that could reside within the structure. From their work, five DSFGs have secure spectroscopic confirmation within a much more spatially compact region.

The stellar masses of these DSFGs are estimated in Dannerbauer et al. (2014), averaging around \( 10^{11} M_\odot \). Lacking stellar mass estimates on the other spectroscopically identified proto-cluster members, the aggregate stellar mass can be estimated roughly at \( \sim 1 \times 10^{12} M_\odot \) and an inferred halo mass of \( 1 \times 10^{14} M_\odot \) if abundance matching is used to separately scale to halo mass. This is perhaps less appropriate in this structure than in the others given the compact spatial arrangement. It is possible that the mass surrounding the identified DSFGs in this sub-halo has virialized. Further observations will be crucial for interpreting the size and mass of this structure (J. D. Kurk et al. 2016, in preparation), and how it compares to the other high-\( z \) structures in the literature.

Without detailed SED information on each proto-cluster member, it is not possible to directly derive a total star formation rate to the system. However, given the far-infrared photometry provided in Dannerbauer et al. (2014), the SFR estimates are re-derived in a self-consistent way and arrive at \( 2200 \pm 500 \, M_\odot \, \text{yr}^{-1} \) as the total for the structure. Note that this may be overestimated due to the lack of correction for confusion boosting on the far-infrared photometry, but also could be an underestimate due to the lack of inclusion of all of the proto-cluster members.

The volume estimate of \( 3000 \, \text{cMpc}^3 \) for the structure surrounding MRC1138–256 uses a sky area of roughly \( 6' \times 9' \) with a redshift interval of \( 2.154 < z < 2.171 \). Similar to the GOODS-N structure, MRC1138–256 is limited by a narrow field of view for multiwavelength follow-up, and so all of the estimated parameters should be regarded as lower limits which are perhaps only representative of a smaller sub-halo in a larger overdensity.

2.4. COSMOS Structure at \( z = 2.47 \)

Casey et al. (2015) describe an extended structure in the COSMOS field at \( z = 2.47 \) which contains seven spectroscopically confirmed DSFGs and five additional AGN. The large-field coverage of COSMOS is uniquely useful for the identification of this overdensity, as the LBG excess is only moderate on smaller scales (<1'). Intriguingly, a few other works identify a neighboring overdensity of LAEs (Chiang et al. 2015; Diener et al. 2015) at \( z = 2.44–2.45 \). While this LAE-rich structure is offset both spatially and in redshift, by ~50 cMpc, it could be associated as part of a colossally large overdensity. Lee et al. (2016) detect this \( z = 2.44–2.45 \) structure using the absorption of neutral hydrogen in the IGM, although existing data is limited to the coincident spatial region and does not cover the \( z \sim 2.47 \) DSFG-rich structure. More work is currently being carried out to determine the possible filamentary connection between the two structures and whether or not this also relates to a possible overdensity of DSFGs detected at \( z = 2.51–2.55 \) in the same field. Note that the number of galaxies in this structure has increased since its initial publication in Casey et al. (2015); the public release of results from the VIMOS Ultra Deep Survey (VUDS) in 2016 February revealed an additional 15 previously unidentified, spectroscopically confirmed proto-cluster members.

The detailed calculation of this structure’s net star formation rate of \( 4500 \pm 500 \, M_\odot \, \text{yr}^{-1} \), total stellar mass of \( 1.0 \times 10^{12} \), halo mass of \( (8 \pm 3) \times 10^{13} \), and volume of 15,000 cMpc\(^3\) is given in Casey et al. (2015) and is calculated in a manner which is fully consistent with the other structures described in this paper.

2.5. SSA22 \( z = 3.09 \) Structure

The SSA22 structure was originally revealed in Steidel et al. (1998) as one of the first high-\( z \) proto-clusters ever detected in LBGs, and as such is probably one of the best-studied proto-clusters in the literature. Narrow-band Ly\( \alpha \) follow-up has revealed an extended excess of \( z = 3.1 \) LAEs extending as far as 60 Mpc comoving (Hayashino et al. 2004; Matsuda et al. 2005; Yamada et al. 2012). The full extent of the structure is shown in Hayashino et al. (2004) in LAEs as reaching over three distinct filaments of about \( 20' \times 3' \), \( 10' \times 4' \), and \( 8' \times 8' \).
across; the implied volume in the redshift range $3.07 < z < 3.11$ is $\approx 21000 \text{cMpc}^3$.

The structure is also home to an excess of DSFGs (Tamura et al. 2009). Three DSFGs were spectroscopically confirmed as proto-cluster members in Chapman et al. (2005), three more were identified as Ly$\alpha$ emitters with submillimeter detections in Geach et al. (2005), and most recently some ALMA-detected submillimeter sources have spectroscopic confirmations from the node of the proto-cluster (Kubo et al. 2015; Umehata et al. 2015). Four of these sources are significantly fainter than the other eight, and so are excluded from the DSFG-overdensity calculation, although they are still considered for their bulk contributions to the SFR. The FIR characteristics of these 12 DSFGs are given in Table 4. Extrapolating from the 850 $\mu$m and 1.1 mm flux density and a 35 K modified blackbody template, the SFRs measured for SSA22 DSFGs range from 120 to 1400 $M_\odot$ yr$^{-1}$ and total 5670 $M_\odot$ yr$^{-1}$. In addition, there are 12 X-ray luminous AGN present in the proto-cluster (Lehmer et al. 2009), 4 of which overlap with the DSFGs, bringing the total rare galaxy count to 13. Lehmer et al. (2009) also find evidence that the LBGs in SSA22 are slightly more massive (by factors of 1.2–1.8) than LBGs in the field, and Hine et al. (2016) show evidence for enhanced merger rates in the proto-cluster member galaxies.

It is difficult to precisely identify how many spectroscopically confirmed proto-cluster members sit in the SSA22 proto-cluster. The original spectroscopic sample only has 16 members, while the narrow-band follow-up imaging around Ly$\alpha$ has 283 confident candidates extending approximately half a degree across the sky. In addition, there have been several additional spectroscopic campaigns in the field, confirming a handful of interesting sources. No stellar mass estimates are given for this structure, although Steidel et al. (1998) do provide an estimate of the total halo mass of $(8 \pm 4) \times 10^{13} M_\odot$ computed using the implied bias from the LBG overdensity.

### Table 4

| Name            | $z$ | $S_{850}$ (mJy) | $S_{1.1\text{mm}}$ (mJy) | $L_{\text{IR}}$ ($L_\odot$) | SFR ($M_\odot$ yr$^{-1}$) | X-ray AGN | References |
|-----------------|-----|----------------|--------------------------|-----------------------------|---------------------------|-----------|------------|
| DSFG J221732.41+001743.8 | 3.092 | ...          | 6.4 $\pm$ 0.2           | $(1.1^{+0.9}_{-0.3}) \times 10^{13}$ | 1000$^{2000}_{-500}$ Y | 3         |
| DSFG J221735.15+001537.3 | 3.096/8 | 6.3 $\pm$ 1.3 | 2.3 $\pm$ 0.1           | $(4.5^{+1.7}_{-2.0}) \times 10^{12}$ | 420$^{700}_{-150}$ N | 1, 3      |
| DSFG J221735.83+001559.0 | 3.089 | 4.9 $\pm$ 1.3 | 1.8 $\pm$ 0.1           | $(3.5^{+2.9}_{-1.6}) \times 10^{12}$ | 330$^{500}_{-170}$ Y | 1, 3      |
| DSFG J221732.01+001655.4 | 3.091 | 3.2 $\pm$ 1.6 | 0.7 $\pm$ 0.1           | $(1.8^{+0.1}_{-0.4}) \times 10^{12}$ | 160$^{300}_{-40}$ Y | 2, 3      |
| DSFG J221725.97+001238.9 | 3.102 | 17.4 $\pm$ 2.9 | ...                     | $(1.4^{+1.2}_{-0.6}) \times 10^{13}$ | 1400$^{2000}_{-600}$ N | 1, 2      |
| LAB J221711.67+001644.9 | 3.06–3.13 | 5.2 $\pm$ 1.4 | ...                     | $(4.3^{+3.0}_{-1.9}) \times 10^{12}$ | 400$^{700}_{-350}$ N | 2         |
| LAB J221802.27+002556.9 | 3.06–3.13 | 6.1 $\pm$ 1.4 | ...                     | $(5.1^{+2.3}_{-1.7}) \times 10^{12}$ | 480$^{700}_{-390}$ N | 2         |
| LAB J221728.90+000751.0 | 3.06–3.13 | 11.0 $\pm$ 1.5 | ...                     | $(9.1^{+1.1}_{-1.0}) \times 10^{12}$ | 860$^{1400}_{-290}$ N | 2         |
| † J221737.11+001712.4 | 3.090 | ...          | 1.1 $\pm$ 0.1           | $(1.8^{+1.5}_{-0.8}) \times 10^{12}$ | 170$^{300}_{-140}$ N | 3         |
| † J221736.54+001622.7 | 3.095 | ...          | 1.0 $\pm$ 0.1           | $(1.7^{+1.4}_{-0.7}) \times 10^{12}$ | 160$^{310}_{-130}$ Y | 3         |
| † J221737.05+001822.4 | 3.086 | ...          | 1.1 $\pm$ 0.1           | $(1.8^{+1.5}_{-0.8}) \times 10^{12}$ | 170$^{310}_{-140}$ N | 3         |
| † J221736.81+001818.1 | 3.085 | ...          | 0.8 $\pm$ 0.2           | $(1.3^{+1.1}_{-0.6}) \times 10^{12}$ | 120$^{240}_{-80}$ N | 3         |

Note. Sources with † preceding the name are not included in the calculation of SSA22’s rare object overdensity, $\delta_{\text{m}}$, as they have much lower luminosities than the other DSFGs in the sample, detected over much larger areas. References are 1—Chapman et al. (2005), 2—Geach et al. (2005), 3—Umehata et al. (2015). 850 $\mu$m flux densities are taken from Chapman et al. (2005) and Geach et al. (2005), while 1.1 mm flux densities from ALMA are given in Umehata et al. (2015). Note that 850 $\mu$m coverage extends over a much larger area than the “ALMA Deep Field” of the SSA22 proto-cluster node but with shallower depth. The redshifts of the three LAEs are not precisely known as they were identified through narrow-band imaging and not direct spectroscopic observations. The X-ray AGN column indicates whether or not the DSFG is matched to an X-ray source in Lehmer et al. (2009). Total infrared luminosities and star formation rates are derived by assuming a 35 K dust modified blackbody plus mid-infrared power law.

2.6. The GN20 Overdensity at $z = 4.05$

One of the brightest submillimeter galaxies from the original SCUBA surveys, GN20, eluded redshift identification for many years until Daddi et al. (2009) confirmed it at $z = 4.055$ through a serendipitous CO detection. Follow-up work revealed two accompanying galaxies, which are themselves submillimeter emitters, at the same redshift. This GN20 system is discussed in detail in Hodge et al. (2013a). This overdensity is significantly different than the structures discussed so far, with far fewer proto-cluster members identified through spectroscopy. This may indicate that it is intrinsically less massive than the other structures, or that spectroscopic incompleteness is quite severe. Because the structure sits in the well-studied GOODS-N structure at $z = 1.99$) spectroscopic incompleteness is less likely, particularly at a redshift where detecting Ly$\alpha$ emitters would be fairly straightforward with ground-based, optical multi-object spectrographs (Cowie et al. 2004; Wirth et al. 2004).

Stellar mass estimates for this group are given in Daddi et al. (2009) and Hodge et al. (2013a) for the three DSFGs: GN20, GN20.2a, and GN20.2b. The sum of their stellar masses is $\sim 3 \times 10^{11} M_\odot$, and the total star formation rate is 1500 $\pm 800 M_\odot$ yr$^{-1}$. Hodge et al. reveal six tentative CO(2-1) detections surrounding the GN20 complex, and the 50 cMpc$^3$ volume for the structure is thus estimated within a $4^\prime \times 3^\prime$ area and a redshift interval of $\Delta z = 0.0014$ at $z = 4.055$. Similar to the lack of large numbers of spectroscopic confirmations, the estimated volume is quite a bit smaller than the other structures presented here, which may be due to the fact that we are looking at a sub-halo in a larger structure or, more likely, a group which is intrinsically less massive than the five structures presented so far that sit at lower redshift.
2.7. The HDF 850.1 Overdensity at \( z = 5.18 \)

Walter et al. (2012) describe the massive starbursting submillimeter galaxy HDF 850.1 and the structure surrounding it at \( z \approx 5.2 \). Like GN20, HDF 850.1 eluded redshift confirmation for over a decade and was only confirmed via the detection of molecular gas. While it is the only DSFG in this \( z \approx 5.2 \) overdensity, there is an accompanying QSO and 11 other spectroscopically confirmed galaxies at the same redshift. This overdensity extends across a large filamentary area of \( 10' \times 30' \). Its total star formation rate is estimated at \( 850 \pm 300 M_{\odot} \) yr\(^{-1}\) simply using the single submillimeter source due to the lack of adequate photometric constraints on the other proto-cluster members. Similarly, given the high redshift of this structure, stellar masses are unconstrained due to a lack of atmospheric transmission around rest-frame 1.6 \( \mu \)m. However, note that there is a dynamical mass constraint on the galaxy HDF 850.1 of \( (1.3 \pm 0.4) \times 10^{11} M_{\odot} \), which can be used as a lower limit to the halo mass of the system at \( z \approx 5.2 \). The volume estimate of 20,000 \( c \)Mpc\(^3\) is derived assuming the above solid angle and a redshift range of 5.183 < \( z < 5.213 \).

2.8. The AzTEC-3 Overdensity at \( z = 5.30 \)

Capak et al. (2011) report the discovery of an overdensity surrounding the interesting luminous DSFG named AzTEC-3 in the COSMOS field. Within a 1' diameter region, there appear to be 12 proto-cluster members at \( z \approx 5.3 \), including the single DSFG AzTEC-3 and one X-ray detected quasar at a distance of 13 Mpc from the starburst. Similar to the HDF 850.1 overdensity, estimating stellar masses for these sources is quite challenging, although Capak et al. (2011) offer this computation directly, totaling >2 \( \times 10^{10} M_{\odot} \). They extrapolate this to a halo mass using abundance matching techniques and estimate a lower limit of >4 \( \times 10^{11} M_{\odot} \). The total SFR estimate is again taken for the sole DSFG member at 1600 \( \pm 500 M_{\odot} \) yr\(^{-1}\). The volume of the structure is estimated within a 0.5' radius and a \( \Delta z = 0.03 \) interval, arriving at a lower limit of >500 cMpc\(^3\). As is the case with the other high-redshift overdensities, it is important to stress that the AzTEC-3 system could be the progenitor of a less massive overdensity.

2.9. Candidate DSFG-rich Proto-clusters

It is important to emphasize again that a number of candidate high-\( z \) DSFG-rich proto-clusters have recently been found thanks to wide-area surveys, like those from Planck and Herschel, but are awaiting spectroscopic confirmation (Clements et al. 2014; Planck Collaboration et al. 2015; Flores-Cacho et al. 2016). It is similarly important to stress that not all other spectroscopically identified \( z > 2 \) proto-clusters have the sensitive submillimeter data sets needed to detect potential DSFG member galaxies (e.g., Lee et al. 2014).

3. FROM DSFG-RICH PROTO-CLUSTERS TO \( z \sim 0 \) CLUSTERS

Such physically large, extended structures—like those observationally identified in Section 2—are not certain to collapse into massive galaxy clusters. How can we adequately determine whether or not these structures will collapse by \( z \sim 0 \)? And does their number density agree with what is known about galaxy clusters at \( z \sim 0 \)?

3.1. Will they Collapse?

Two schools of thought have been used to address this question. The first draws on the Press–Schechter formalism (Press & Schechter 1974) for spherical collapse within large-scale structures (Mo & White 1996), whereby a certain mass overdensity, \( \delta_{\text{mass}} \), is required to exceed a specific critical value \( \delta_c \) to collapse by \( z \sim 0 \) (Peacock 1999). Because the mass overdensity is not directly observable, linear galaxy bias is assumed whereby \( 1 + b \delta_{\text{mass}} = C(1 + \delta_{\text{gal}}) \), \( \delta_{\text{gal}} \) is the observed galaxy overdensity, \( b \) is the bias associated with that galaxy population (i.e., how well they trace the dark matter halo mass), and \( C \) is a redshift distortion factor accounting for unknown peculiar velocities.

For example, the analysis of the GOODS-N \( z = 1.99 \) structure in Chapman et al. (2009) finds an SMG overdensity of \( \delta = 10 \), which is sufficient to cause collapse, however, the underlying LBG population overdensity, \( \delta_{\text{LBG}} = 2.5 \), is not significant enough to cause collapse. These two assessments of the structure are seemingly contradictory, but the authors address this contradiction by suggesting that either the bias of the submillimeter galaxy population is sufficiently different than for LBGs or there could be a large population of massive galaxies that have not been detected surrounding the structure. Given the depth of multwavelength imaging in GOODS-N, the latter is unlikely. Thus, Chapman et al. determined that the bias for SMGs (or DSFGs) and LBGs was sufficiently different, and so even a large overdensity of SMGs may not probe massive clusters in formation.

This conclusion is further supported by Miller et al. (2015), who use large-volume, semi-analytic simulations from Klypin et al. (2011) to argue that SMGs are “poor tracers” of the most massive structures at \( z \sim 2 \). They use an SMG prescription largely driven by disk instabilities and a mildly top-heavy IMF, observing very few massive structures containing more than \( 1-3 \) SMGs. The structures observed with >5 SMGs are indeed among the most massive, but are exceedingly rare in the simulation, indeed, much more so than the observations in Section 2 suggest. This discrepancy between their predicted number of DSFG-rich proto-clusters and our observations are shown as green and blue points on the cluster mass function plot in Figure 1, and is discussed in more detail in the next subsection.

Note that other simulation groups (Granato et al. 2015; Lacey et al. 2015) have been working to understand the turn-on of luminous DSFGs in large-box simulations where the collapse of the most massive structures can be observed. However, as highlighted in those works, it is still incredibly challenging to carry out proper radiative transfer in such large environments, especially on \( \sim 20 \) cMpc scales before clusters have collapsed.

The second school of thought draws on recent cosmological simulations of hierarchical growth, which produce somewhat different predictions than those relying on analytic descriptions of structure formation theory. For example, Chiang et al. (2013) present a clear argument as to why spherical collapse models and the assumed linear regime for overdensities may introduce systematic errors in mass measurements for non-virialized proto-clusters. These direct predictions from simulations suggest the following: (a) the median observed galaxy overdensity, \( \delta_{\text{gal}} \), rarely, if ever, exceeds 10 (this agrees with the predictions given in Miller et al. 2015), (b) \( \delta_{\text{gal}} \) at these epochs also depends strongly on the observational characteristics being
selected for, for example, SFR or stellar mass, and sensibly vary between DSFG populations (very high SFR-selected samples) and LBG populations (a combination of SFR and mass-selected, at much deeper detection thresholds), (c) the progenitors of massive galaxy clusters at \(z > 2\) occupy very large Lagrange volumes, \(\geq 10,000\) cMpc\(^3\) (see also Oñorbe et al. 2014), and (d) \(\delta_{\text{gal}}\) will vary for structures of the same mass depending on the “window size” of observations, or the presumed volume, given intrinsic variations in the underlying density along filaments.

For example, a close inspection of Figure 8 in Chiang et al. (2013)—a plot of the cumulative fraction of proto-clusters with observed galaxy overdensities \(\delta_{\text{gal}}\) at \(z = 2, 3, 4,\) and 5—provides a backdrop against which to interpret the likelihood of proto-cluster collapse. Among the five rich 1.99 < \(z\) < 3.09 proto-clusters described in Section 2, all of the structures are expected to collapse by \(z \sim 0\). The structure with the least remarkable LBG overdensity at \(\delta_{\text{gal}} = 2.5\), the GOODS-N \(z = 1.99\) structure, is still among the top 30% of collapsing structures. The remainder are in the top 5%–10%.

It should be clarified, however, that the three highest-redshift overdensities discussed in Section 2 and summarized in Table 1 have a less clear fate. With far fewer numbers of galaxies (in both rare sub-types and total number), Poisson noise dominates the calculation of the overdensities, causing a wide margin of error on the structures’ predicted state at \(z \sim 0\). These are the types of structures which may either be prone to mass overestimation, due to the effects discussed in Miller et al. (2015), or suffer from incomplete spectroscopic descriptions, although the latter interpretation may be limited by constraints set by the volume density of DSFG-rich proto-clusters as a whole.

3.2. How Common are they?

While the argument for the eventual collapse of DSFG-rich proto-clusters into the most massive \(z \sim 0\) clusters has been made in Section 3, it is not immediately obvious that this evolutionary picture is feasible or likely, given the relatively small number of high-mass clusters at \(z \sim 0\). In Figure 1, the cluster mass function at \(z \lesssim 0.2\) is shown from the Sloan Digital Sky Survey (Bahcall & Cen 1993; Bahcall et al. 2003). This tells us that there is one \(\geq 10^{15} M_\odot\) cluster per every 1–2 million Mpc\(^3\), or per ~120 \(\times\) \(\times\) 120 Mpc\(^3\) comoving box.

We can also work out a rough estimate of the volume density of DSFG-rich proto-clusters for comparison. A significant discrepancy between the volume density of \(z \sim 0\) massive clusters and \(z \gtrsim 2\) DSFG-rich proto-clusters is a sign that the two populations are not likely related. This was claimed to be the case in Blain et al. (2004), after analyzing the overdensity associated with the GOODS-N \(z = 1.99\) structure, and a few other potential SMG-rich overdensities perceived in the first few square degrees of deep submillimeter imaging. Blain et al. determined that DSFG-rich structures were unlikely to be the progenitors of massive clusters in formation because they are ~10 times more common at \(z \gtrsim 2\) than their \(z \sim 0\) descendants, which was reflective of the best data on hand at the time. Here, this estimate is reassessed using improved data sets.

To estimate the volume density of DSFG-rich proto-clusters at \(z \gtrsim 2\), understanding survey area and selection bias is critical. The survey area in this case is set by the solid angle of sky covered to sufficient depth to recover DSFG-rich structures at high redshift. This requires both spectroscopically complete samples and confusion-limited submillimeter blank-field maps. Both are extremely limited by current observational resources. The former is limited by the need for several tens of nights on 8–10 m class optical/near-infrared telescopes for multi-object spectroscopy of faint \(i \sim 22–26\) sources (of which only a few fields have truly complete coverage, e.g., GOODS-N, the central portion of COSMOS, ECDFS), and the latter is limited by the historically slow mapping speeds of single-dish bolometer array instruments like SCUBA (also LABOCA, MAMBO, AzTEC, and now SCUBA-2). The intersection of these two data sets is therefore limited to the following:

1. about 0.4 deg\(^2\) in GOODS-N, a field with confusion-limited 850 \(\mu\)m data (Barger et al. 1998; Chen et al. 2013b) and extensive spectroscopic completeness (Cowie et al. 2004; Wirth et al. 2004; Reddy et al. 2006; Barger et al. 2008);
2. the central 1 deg\(^2\) of the COSMOS field, which has published confusion-limited 850 \(\mu\)m data covering 0.2 deg\(^2\) (more yet in J. Geach et al. 2016, in preparation) and 1 deg\(^2\) of deep spectroscopic data from the zCOSMO\(\bar{S}\) team (Lilly et al. 2009);
3. about 0.5 deg\(^2\) in ECDFS with confusion-limited submillimeter data from LABOCA and ALMA (Weiß et al. 2009; Hodge et al. 2013b) and spectroscopic follow-up from Popesso et al. (2009), Balestra et al. (2010), and Le Fèvre et al. (2005);
4. a 0.5 deg\(^2\) portion of the Lockman Hole (SHADES) field for which a significant number of DSFGs have been

\[\text{Either they are not likely related or, if they are, then DSFG-rich proto-clusters are probably much more rare than most "normal" proto-clusters.}\]
spectroscopically confirmed (Chapman et al. 2005; Lindner et al. 2011; Casey et al. 2012a); and 5. other deep submillimeter fields, which include the backgrounds of low-redshift Abell clusters (e.g., Smail et al. 1997; Chen et al. 2013a) and the SSA13 and SSA22 fields, and cumulatively add up to about ~0.5 deg$^2$.

This collection of deep surveys adds up to a total effective survey solid angle of ~3 deg$^2$, with an uncertainty of about ~0.5 deg$^2$ to account for variable levels of spectroscopic completeness and submillimeter data quality and depth. While it should be noted that Herschel coverage also spans all of these legacy fields and much larger areas, the intersection with spectroscopic samples is the main limiting factor in making use of it for this analysis. In addition, Herschel is most efficient at identifying DSFGs at $z < 2$ (Casey et al. 2012a), which is a characteristic of its shorter-wavelength selection compared to ground-based submillimeter data sets. Color selection with the Herschel bands seems like an efficient method for recovering a higher-redshift sample (e.g., Dowell et al. 2014; Asboth et al. 2016), although the depth and completeness of these “500 μm peakers” is less well characterized. It is important to emphasize that this estimate is rough, as the complexity of these data sets is incredibly difficult to quantify in a simple analysis.

The corresponding solid angle to this 3 deg$^2$ is converted to a cosmological comoving volume within the redshift interval of interest, which is approximated as 1.9 < $z < 4.5$, with the lower limit defined by the limit of known virialized clusters and the upper limit constrained by the low completeness in most of the large spectroscopic surveys summarized above. Allowing for some additional uncertainty in the redshift interval, the total volume accessible is $(9 \pm 3) \times 10^7$ cMpc$^3$. By chance, this is approximately the same volume probed by deep SDSS cluster surveys, ~400 deg$^2$ out to $z \sim 0.1$–0.2 (Bahcall et al. 2003), from which the nearby cluster mass function is measured.

Although there are clearly these five bona-fide, DSFG-rich proto-clusters identified in the literature, one is substantially impacted by a possible selection bias associated with the proto-cluster. Much of the deep data associated with MRC1138−256 at $z = 2.16$ was obtained with the explicit knowledge of the proto-clusters’ presence, and so it cannot be included in the calculation estimating their volume density. Thus, four DSFG-rich proto-clusters are left for the volume density calculation: GOODS-N at $z = 1.99$, COSMOS at $z = 2.10$, COSMOS at $z = 2.47$, and SSA22 at $z = 3.09$. A Poisson uncertainty is assumed for the number of DSFG-rich proto-clusters. The implied volume density is then $\sim 5 \times 10^{-3}$ cMpc$^{-3}$ for DSFG-rich proto-clusters. This is depicted by the blue point in Figure 1 and is in rough agreement with the observed $z \sim 0$ cluster mass function.

There is one remaining concern with this calculation. If this estimate is consistent with the $z \sim 0$ local cluster mass function, then it may imply that every $z \gtrsim 2$ proto-cluster should be DSFG-rich. This is not obviously the case. Before addressing this issue further, one must first consider the timescales of DSFGs and their implications on the clusters’ assembly histories.

### 3.3. Star Formation in DSFG-rich Proto-clusters

Placing DSFG-rich proto-clusters in context requires a more detailed look at their observable star formation characteristics in comparison to the field (i.e., normal density regions), and lower-redshift virialized clusters. Figure 2 shows the cosmic star formation rate density from $0 < z < 4$, as compiled by Hopkins & Beacom (2006) for the field, against similar measures for overdense environments.

DSFG-rich proto-clusters at $2 < z < 3$ have only slightly elevated $\rho_{\text{SFR}}$ compared to the field, thanks primarily to the large volumes they occupy prior to virialization. On the other hand, virialized clusters at $z < 1$ have substantially higher $\rho_{\text{SFR}}$, peaking around $0.5 < z < 1.0$, while potentially experiencing suppressed star formation at lower redshifts brought on by different environmental mechanisms. Note that the comparison between virialized clusters and the field uses the comoving volume, as opposed to proper volume, for fair comparison with structures that have not yet collapsed and decoupled from the Hubble flow. All of the values for SFR are converted to a Chabrier IMF (Chabrier 2003). The gray band marks the evolution of a hypothetical cluster that sustains an aggregate SFR of $3000 M_\odot$ yr$^{-1}$ from $z = 4$ to $z = 0$ while undergoing collapse as predicted from large N-body simulations (Onorbe et al. 2014). This highlights, through one variable, how galaxies in proto-clusters more closely emulate galaxies in the field than those in $z \sim 1$ clusters that have collapsed.

Figure 3 takes a closer look at the breakdown of the star formation rate function or the luminosity function within a DSFG-rich proto-cluster in comparison to the field. For context, the Lyman-break Galaxy luminosity function of Reddy & Steidel (2009) is converted to a SFR function using the UV-scaling in Kennicutt (1998) and by applying a factor of five correction for extinction (i.e., most LBGs are 80% obscured; Reddy et al. 2012). The highest-redshift luminosity function from the infrared (Gruppioni et al. 2013) is converted to a SFR also using the Kennicutt prescription, adjusted for a Chabrier IMF. Against these field measurements, the SFR function of DSFG-rich proto-clusters is shown: all of the known members of the COSMOS $z = 2.47$ proto-cluster (Casey et al. 2015) in red, and the DSFG member galaxies of all five $1.99 < z < 3.09$ structures as black stars. The key distinguishing characteristic of DSFG-rich proto-clusters is the flattening of the luminosity function toward high SFRs. While there may be an excess of LBGs observed in high-$z$ proto-clusters, the excess is not as great as the factor of $\gtrsim 10$ excess toward the highest SFRs.

### 4. SIMULTANEOUS TRIGGERING, OR NOT?

Here, the likelihood of several rare types of galaxies being observed simultaneously within a large structure is explored. If you work based on the premise that both populations of DSFGs and AGN are short-lived on 100 Myr timescales, then one can ask what the probability is of observing $N$ of them simultaneously in one structure (where $N \gtrsim 5$). If the probability is low, and yet the prevalence of such DSFG-rich structures is high, then one may think that this is evidence that clusters themselves assemble in rapid bursts, even when extended over very large volumes $\gtrsim 10,000$ cMpc$^3$ (as suggested in Casey et al. 2015).

Care should be taken when correcting for the dynamical time of each DSFG at different redshifts, as discussed in Simpson et al. (2014). At higher redshifts, a fixed $dz$ element probes shorter and shorter timescales, such that the probability of observing all of the DSFGs that have been triggered during that time element $dz$ increases from low fractions at low-$z$ to
Figure 2. Star formation rate density ($\rho_{\text{SFR}}$, in $M_\odot$ yr$^{-1}$ Mpc$^{-3}$) of proto-clusters and clusters compared to galaxies in the field from Hopkins & Beacom (2006; black points). The five DSFG-rich proto-clusters from this work are shown as blue stars, seven individual low-redshift clusters from Geach et al. (2006) are shown as green circles, and the redshift-averaged results from 42 clusters at $0.3 < z < 1.0$ in Webb et al. (2013) are shown in purple. The virialized clusters have an ~2 Mpc proper radius, and the associated volume is converted into comoving units for fair comparison with the field and proto-clusters. The gray stripe represents the track of a hypothetical and idealized proto-cluster which sustains a constant SFR as black stars. SFR function for all blue regions and dashed lines is shown for all known members. No correction has been made for incompleteness (hashed gray regions and dashed lines), which dominates at SFRs $\lesssim 20 M_\odot$ yr$^{-1}$ (for UV-selected samples) and at SFRs $\approx 100$–200 $M_\odot$ yr$^{-1}$ (for DSFGs). The net SFR function for all five DSFG-rich proto-clusters at $1.99 < z < 3.09$ is shown as black stars.

Figure 3. Star formation rate function of DSFG-rich proto-clusters compared to the field. The luminosity functions of Lyman-break galaxies (blue line; Reddy & Steidel 2009) and IR-selected galaxies (orange line; Gruppioni et al. 2013) in the field are shown for context; the black line is the sum of the two. The SFR function of the COSMOS $z = 2.47$ proto-cluster (red points) is shown for all known members. No correction has been made for incompleteness (hashed gray regions and dashed lines), which dominates at SFRs $\lesssim 20 M_\odot$ yr$^{-1}$ (for UV-selected samples) and at SFRs $\approx 100$–200 $M_\odot$ yr$^{-1}$ (for DSFGs). The net SFR function for all five DSFG-rich proto-clusters at $1.99 < z < 3.09$ is shown as black stars.

100% at high-$z$. While large redshift bins with widths $\Delta z = 0.1$–0.2 will probe all such episodes, it is important to note that the redshift range probed by a single coherent structure, $dz \approx 0.02$, only corresponds to a crossing time of $\approx 20$ Myr, which is shorter than the expected duration of the burst phase. If this itself were to exceed the estimated lifetimes of our rare galaxies, then that could provide an easy explanation as to why we observe structures that are quite rich in DSFGs and luminous AGN. However, that is not the case.

Another possible explanation for the plethora of rare galaxies is that we actually expect nearly all $z \sim 0$ galaxy cluster members to have gone through such a rare phase at some time in its past, probably around $z \sim 2$–3. But in investigating this further, there is a problem. The most massive galaxy clusters at $z \sim 1$ only have $40 \pm 10$ galaxies above a stellar mass of $10^{11} M_\odot$ (van der Burg et al. 2013). If one presumes that all of these galaxies have gone through a DSFG phase at some point during their massive buildup (as most of them are quiescent by $z = 0.5$–1), then by working backwards, the likelihood of observing $N$ of them in a DSFG or luminous AGN phase simultaneously can be worked out. Here, the time $T$ it takes for the structure to collapse from its primordial fluctuations is relatively unknown, but is loosely constrained by the redshift interval $1 < z < 6$ ($\approx 5$ Gyr), or $2 < z < 5$ ($\approx 2$ Gyr).

Figure 4 shows the probability of simultaneously observing $\geq N$ DSFGs/AGN within one structure forming over the course of 2 Gyr. Assuming a 2 Gyr timescale renders the probability calculations in Figure 4 conservative, since allowing for longer
fall-in times makes the probabilities of observing multiple DSFGs simultaneously lower. Four different rare galaxy timescales are assumed (where “rare” can refer to either the DSFGs or the short-lived, luminous AGN in this case): 50 Myr (in line with what is observed in local ULIRGs; Solomon & Sage 1988), 100 Myr (the typical DSFG timescale and upper limit to QSO lifetimes; Martini 2004; Greve et al. 2005), 150 Myr (a depletion time typical of some longer-lived DSFGs at high redshift; Swinbank et al. 2014), and 500 Myr (a DSFG timescale that would rely on some sustained gas fueling, which some assert is likely the case at the massive end of the galaxy “main sequence”; Elbaz et al. 2011). This figure illustrates that the assumed timescale for DSFGs and luminous AGN is rather important to our understanding of cluster assembly. Over a 2 Gyr build time, if DSFGs/AGN are short-lived, then the probability of observing >5 such sources in one proto-cluster structure is <0.5% (50 Myr), 6.1% (100 Myr), 19% (150 Myr), and 77% (500 Myr). However, structures like the COSMOS group contain 12 rare sources each.

With a short-lived phase, this is virtually impossible through uncorrelated triggering (<1 × 10^{-4}%), and yet is still unlikely for long-duration events (<25%). If such phenomena are short-lived, then they most certainly are triggered simultaneously in an event that stretches across very large volumes. One can imagine this triggering is brought on by the rapid collapse of filamentary structure that extends across several tens of Mpc.

On the other hand, the test above seems to suggest that longer lifetimes are far more likely (by over a factor of 10) for DSFGs and luminous AGN in proto-clusters. Recent simulation work (Narayanan et al. 2015) suggests that even somewhat isolated DSFGs could sustain sufficiently high star formation rates (≥500 M⊙ yr^{-1}) for 0.75 Gyr. Physically, this sounds plausible, particularly in dense environments where high star formation rates may be sustained over longer periods of time if the galaxies are continually fed fresh supplies of gas from the surrounding, rich medium. In the next few subsections, I explore observations which support both rapid collapse and heightened gas supply scenarios.

### 4.1. Molecular Gas Depletion Time

Determining the correct interpretation of the assembly history of galaxy clusters requires direct constraints of the molecular gas potential wells in proto-cluster DSFGs. This provides us with critical information concerning the galaxies’ current gas supply, and over what time period such high star formation rates would be continuously sustainable. To reiterate, this is a particularly useful measurement in DSFGs due to their rarity, as demonstrated in the previous section.

Table 5 summarizes existing CO observations of proto-cluster DSFGs from the literature. Though limited in number and heterogeneous in transition and depth, these data can allow us to begin to discern the plausibility of short-lived versus long-lived interpretations. However, as with most previous work on high-z CO observations, it is very important to recognize that the conversion from observed CO line strength to H₂ gas mass is highly uncertain. It first requires a conversion from a high-J CO transition to the ground state CO(1-0), which requires knowledge of the galaxy’s mean CO excitation ladder or kinetic gas temperature. Second, the conversion from CO(1-0) to M_H2, known as α_CO or ε_CO, can also range by factors of 5-100 depending on the gas conditions in the ISM. For example, the Milky Way has a gas conversion rate of ε_CO = 4.5 M⊙ (K km s^{-1} pc^{2})^{-1} (Bloemen et al. 1986; Solomon et al. 1987), while typical local ULIRGs have ε_CO = 0.8 M⊙ (K km s^{-1} pc^{2})^{-1} (Downes & Solomon 1998). The uncertainties in these two conversions alone can account for a factor of ≥10 in the predicted gas mass, which could dramatically affect the interpretation of the depletion timescale, τ_depl = M_H2/ SFR.

For those proto-cluster DSFGs without CO(1-0) observations, a CO gas excitation ladder, and associated uncertainties, are assumed as given in Bothwell et al. (2013), with the median excitation seen in all of the observed DSFGs to-date. Their Figure 3 shows this median DSFG spectral line-energy distribution (SLED). Each high-J CO line luminosity in Table 5 is thus converted to an estimated CO(1-0) line luminosity via L_{CO(1-0)}/L_{CO(J-1)} = (S_{CO(1-0)}/S_{CO(J-1)})(1/J)^2. The uncertainty in the CO SLED is reflected in the resulting uncertainty of the CO(1-0) line luminosity. The conversion from L_{CO(1-0)} to M_H2 assumes α_CO = 1.0 M⊙ H₂ (K km s^{-1} pc^{2})^{-1}, which is the same value adopted in Bothwell et al. (2013) and is justified generally through some limited dynamical mass constraints. The resulting gas masses M_H2 are given in Table 5, with some proto-cluster DSFGs containing multiple components. In the case where multiple high-J CO transitions are observed for a single galaxy, a molecular gas mass is derived for each independently, and then averaged. Depletion times are then calculated by taking the total molecular gas mass estimated to be present in the system and dividing by the current star formation rate, as calculated in Section 2. The probability distribution in depletion times is shown in Figure 5. Although quite sparse, the majority of...
DSFGs in Proto-clusters with CO Measurements

| DSFG Name          | $z$ | Transition | $L^{\text{CO}}$ (K km s$^{-1}$ pc$^2$) | $M_{\text{H}_2}$ ($M_\odot$) | SFR ($M_\odot$ yr$^{-1}$) | $\tau_{\text{dep}}$ (Myr) | References                  |
|--------------------|-----|------------|--------------------------------------|----------------------------|---------------------------|----------------------------|----------------------------|
| DSGF J123618 +621550 | 1.996 | CO(4-3) | ($9.4 \pm 1.4$) $\times 10^{10}$ | ($2.6 \pm 0.6$) $\times 10^{11}$ | $330^{+110}_{-80}$ | 1300 $\pm$ 400 | Bothwell et al. (2010) |
| DSGF J114046 +622913 | 2.001 | CO(4-3) | ($6.5 \pm 0.9$) $\times 10^{10}$ | ($1.8 \pm 0.4$) $\times 10^{11}$ | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Bothwell et al. (2010) |
| DSGF J114046 +621331 | 1.998 | CO(4-3) | ($1.3 \pm 0.2$) $\times 10^{10}$ | ($3.6 \pm 0.8$) $\times 10^{10}$ | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Casey et al. (2011) |
| DSGF J123712 +621322 | 1.996 | CO(4-3) | ($7.8 \pm 1.1$) $\times 10^{10}$ | ($2.2 \pm 0.5$) $\times 10^{10}$ | $540^{+120}_{-20}$ | 110 $\pm$ 60 | Casey et al. (2011) |
| DSGF J123632 +620800 | 1.996 | CO(3-2) | ($6.8 \pm 1.5$) $\times 10^{9}$ | ($1.9 \pm 0.5$) $\times 10^{10}$ | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Casey et al. (2011) |
| DSGF J114408 -262908 | 2.163 | CO(1-0) | ($6.5 \pm 0.6$) $\times 10^{10}$ | ($6.5 \pm 0.6$) $\times 10^{10}$ | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Casey et al. (2011) |
| DSGF J114408 -262913 | 2.150 | CO(1-0) | ($6.9 \pm 2.3$) $\times 10^{9}$ | ($6.9 \pm 2.3$) $\times 10^{9}$ | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Bothwell et al. (2013) |
| DSGF J221735 +001537 | 3.096 | CO(3-2) | ($3.8 \pm 1.0$) $\times 10^{10}$ | ($8.5 \pm 2.5$) $\times 10^{10}$ | $540^{+120}_{-20}$ | 110 $\pm$ 60 | Casey et al. (2011) |
| DSGF J221726 +001239 | 3.102 | CO(4-3) | ($6.7 \pm 2.1$) $\times 10^{10}$ | ($1.9 \pm 0.7$) $\times 10^{11}$ | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Casey et al. (2011) |
| DSGF J221732 +001744 | 3.092 | CO(3-2) | … | … | $330^{+400}_{-120}$ | 1300 $\pm$ 400 | Casey et al. (2011) |

Notes.

* Gas masses estimated assuming a fixed $\alpha_{\text{CO}}$ gas conversion factor of $\alpha = 1.0$ (as in Bothwell et al. 2013).
* Tentative detection of CO.
* SFR for the Spiderweb galaxy is calculated from the starburst component of the FIR SED as presented in Seymour et al. (2012).
* SFR for HAE source at $z = 2.147$ is taken from H$_\alpha$ measurements from Kurk et al. (2004) as $23 \pm 1$; however, we refit the SFR given the FIR photometry measured in Dannerbauer et al. (2014, their Table 4).
* SFR for the SSA22 galaxy calculated using an SED with temperature 35K from Chapman et al. (2005), also accounting for a deboosting factor $\sim 1.5$, which is consistent with more recent submillimeter data sets.
* SFR calculated as in Umehata et al. (2015).
* SMG-93, also known as SMM J123600+621047, is mistakenly labeled as SMM J123600+620253 in Bothwell et al. (2013), but all of the physical characteristics listed in Bothwell et al. are indeed for SMG-93.

sources ($5/7 \approx 71\%$) are estimated to be short-lived with $\tau_{\text{dep}} \lesssim 150$ Myr.

4.2. Evidence Supporting Rapid Bursts in Proto-clusters

The discussion presented on the measured molecular gas depletion time of DSFGs in proto-clusters heavily favors a rapid collapse model, whereby the massive galaxies in clusters are built in short-lived, extreme episodes that permeate the entire volume of the not-yet-virialized proto-cluster. The measured gas depletion times for proto-cluster DSFGs (as presented in Table 5 and Figure 5) are the most crucial constraint on this argument, but it is significantly strengthened by inferred constraints on the lifetimes of AGN with comparable luminosities to unobscured quasars (Marconi et al. 2004). The strong evidence for short lifetimes, combined with the low probability of observing $N \geq 5$ of these rare galaxies in one structure, argue for correlated, simultaneous triggering. If correct, then the result is rather extraordinary, as it represents the only type of direct observation of a temporal “event” on cosmological scales, spanning a volume of $\sim 10^3$ cMpc$^3$. In the next subsection, I briefly explore evidence which supports the contrary conclusion.

4.3. Evidence in Favor of Gradually Built Proto-Clusters

Although analysis of literature DSFGs in proto-clusters suggests that they are mostly short-lived, the impact of our
high-J CO to gas mass assumptions should be revisited. If our assumptions were to be revised in favor of more “Milky Way”-type gas excitation and higher intrinsic value of $\alpha_{\text{CO}}$, then the CO(1-0) line luminosities would be a factor of $\sim$3 higher and the gas masses would be a factor of $\sim$20 higher. The median depletion time of 110 Myr would instead be 2.2 Gyr, which is much more in line with the predicted long lifetimes of DSFGs in some cosmological simulations (Narayanan et al. 2015). Reducing the intrinsic uncertainty in this measurement requires CO(1-0) measurements of a larger sample of proto-cluster DSFGs with additional resolved dynamical mass constraints to hone in on the correctly applicable $\alpha_{\text{CO}}$. Some of these observations are currently underway at the Jansky Very Large Array. However, it should be noted that there is a known upper limit to how long DSFGs can sustain high SFRs, given by the stellar mass of the Milky Way. It should be noted that there is a known upper limit to how long DSFGs can sustain high SFRs, given by the stellar mass of the Milky Way.

Another possible caveat to our rapid collapse argument is the possibly heightened replenishment of gas reservoirs from the IGM. It has recently become clear that galaxies recycle gas through ejective feedback and outflows, as well as the eventual reaccretion of material on $\sim$Gyr timescales (Christensen et al. 2015); however, it is unclear how dense environments at the intersections of filaments in the IGM might shorten the gas recycling timescale and potentially heighten the inflow of pristine material. If molecular gas is fed into galaxies more efficiently in proto-clusters than in the field, particularly on $\sim$100 Myr timescales, then the depletion time measurement might not be an accurate reflection of the lifetime of high-SFR systems. However, such a dramatically fast ($\lesssim$100 Myr) replenishment of a $\sim 10^{10} M_\odot$ gas reservoir is unlikely, again due to the upper limit placed on the high-SFR timescale from observed stellar mass functions.

Finally, as mentioned at the end of Section 3.2, the frequency of DSFG-rich proto-clusters among the population of all proto-clusters raises a potential concern. If the timescale of the DSFG-rich phase is short-lived and unique, then one may only expect a small subset of observed $z > 2$ proto-clusters to have such DSFG excesses. To gauge the plausibility of this argument, we should consider how many galaxies we expect to go through such a phase over the course of a cluster’s lifetime. In Section 4, this was approximated as $40 \pm 10$ massive $> 10^{11} M_\odot$ galaxies. If there are 5–10 rare galaxies per proto-cluster, then we may expect such structures to go through 4–8 “episodes” of heightened activity before virialization at $z < 2$. If these episodes are assumed to all occur in the range $2 < z < 5$ ($\approx$ 2 Gyr), then one would expect $\sim$20%–40% of all proto-clusters of that epoch to be DSFG-rich assuming a 100 Myr “burst” lifetime. With a 150 Myr lifetime, the fraction shifts to $\sim$30%–60%, and at 50 Myr only $\sim$10%–20%. Although these fractions are certainly non-negligible, it is clear that it would be nearly impossible for all $z > 2$ proto-clusters to be DSFG-rich if they are short-lived, and therefore our comparison to the measured cluster mass function at $z \sim 0$ might disfavor short timescales. It is certainly clear that refining measurements of the volume density of high-$z$ proto-clusters is needed before ruling out different histories of their assembly.

### 5. Predictions

#### 5.1. Future Observations

The most important observational characteristic of massive galaxy clusters is the large area they subtend on the sky, approximately half a degree across. While some recent works have recognized the importance of this (e.g., Muldrew et al. 2015), the observational community that works on proto-cluster science has largely overlooked the shear scale of early, overdense structures. It is critical to address this issue if we desire to move beyond simple proto-cluster discoveries and learn about the collapse of large-scale structures from an observational perspective.

The next generation of wide field (and sufficiently deep) surveys—on the order of tens of square degrees—will be of great importance for identifying statistically large samples ($\sim$100) of proto-clusters, both with and without rare galaxies. The most efficient means of confirming high-redshift overdensities like these will be through direct far-infrared/millimeter molecular line detection, which may only be efficient on large scales with the next generation of submillimeter, single-dish, multi-pixel spectrometers. The relative fraction of such structures with rare galaxies will, in turn, allow for a more detailed look at all of the clusters’ temporal evolution.

On slightly smaller angular scales, recent work from Clements et al. (2014), Planck Collaboration et al. (2015), and Flores-Cacho et al. (2016) search out proto-clusters that are rich in dusty star formation by leveraging the poor spatial resolution of the Planck satellite, which covers the entire sky. It is hoped that following-up Planck’s $\sim$5° point sources with the higher-resolution Herschel Space Observatory will be an efficient method of identifying early clusters in formation. While none have yet been spectroscopically confirmed, over 200 candidate
high-redshift clusters have been identified with an excess of dusty starbursts peaking at \( \gtrsim 350 \mu \text{m} \). This technique is certainly promising, although it will be quite incomplete for the type of structure discussed in this paper, as many dusty starbursts would need to fall in one Planck beam, which is much smaller than the previously discussed half-degree scale.

In terms of characterizing known structures more fully, narrow-band imaging should provide the most complete mapping of filamentary structures on the largest scales. This is the case in the SSA22 \( z = 3.09 \) structure, as well as some structures not observed in the submillimeter (e.g., the Boötes \( z = 3.78 \) structure; Lee et al. 2014), but has not been pursued over sufficiently large angular scales for most proto-clusters. Similarly, wide-field IFU spectroscopic follow-up from facilities like the VIRUS instrument on the Hobby–Eberly Telescope will be quite valuable.

It is clear that understanding galaxies’ gas supply is an essential element in discerning the assembly history of proto-clusters, and in the age of ALMA and the Jansky VLA, is not limited to the most luminous, rare galaxies. Scaling of long-wavelength dust continuum to an ISM mass has shown to be a useful proxy (Scoville et al. 2014, 2015) to galaxies’ star-forming molecular gas masses. Thus, fairly inexpensive observational campaigns to constrain the gas content of proto-clusters’ normal galaxy members might provide more important clues as to how environment influences galaxies’ evolution.

5.2. Simulations

Simulations of large-scale structure collapse on cosmological scales play a crucial role in our current picture of galaxy cluster formation, linking the huge gap between observations of nearby virialized clusters and the imprint of density perturbations on the cosmic microwave background. Large-box \( > 100 \) Mpc simulations are certainly needed to analyze \( > 10^{15} \) M\( \odot \) halos. Their enormous volumes limit the incorporation of baryonic physics and force the implementation of ultraluminous starbursts, or luminous AGN, to be somewhat crude. However, there are some basic measurements which could be extracted from the current generation of simulations that would shed ample light on the proposed assembly history of massive galaxy clusters. In dark-matter only simulations, the most direct probe of cluster assembly is the merging of dark matter halos with time. These merger trees, mapped with spatial distribution, could directly trace whether or not the growth of halos, and thus the galaxies living in them, is episodic or steady.

Beyond the measurement of stochasticity in assembly, simulations will be needed to more accurately constrain dark matter halo masses from observations. While it is clear that linear bias assumptions break down under certain contexts, it is not entirely appropriate to use normal abundance matching techniques which are more ideally suited for isolated halos. An in-depth look at halo mass distributions in proto-clusters prior to virialization might provide crucial insight that bridges our theoretical understanding to observational constraints.

6. CONCLUSIONS

This paper has employed literature data sets to demonstrate that several high-redshift proto-cluster environments are rich with rare galaxies; both dusty star-forming galaxies and ultraluminous AGN. These proto-clusters subtend \( 10' \) to a half degree in the sky because they have not yet relaxed into virialized galaxy clusters. By virtue of their large occupied volumes at \( z \gtrsim 2 \) (factors of a few hundred larger than at \( z \sim 0 \)), it is very difficult to detect their significance via an overdensity of “normal” galaxies on appropriately large scales, which are only slightly more dense than the field. Instead, an unexpected excess of rare galaxies (\( \gtrsim 5 \) per \( \sim 10^3 \) cMpc\(^3 \)) can demonstrate a more compelling argument for a large-scale proto-cluster in formation.

Five bona-fide DSFG-rich proto-clusters have been identified to date within \( 1.99 < z < 3.09 \). Estimates of their volume density—constrained by deep spectroscopic and submm data sets—are \( \sim 5 \times 10^{-8} \) cMpc\(^3 \), which is similar to the density of observed \( >10^{12} \) M\( \odot \) clusters at \( z < 0.2 \). Some simulations expect the volume density of DSFG-rich structures to be a factor of \( \sim 5 \) less than observed.

The rarity of DSFGs and luminous AGN relates to their intrinsically short duty cycle. If this population is predominantly short-lived, then it can be used as a constraint on the assembly history of galaxy clusters in the time before virialization. For example, the probability of observing 10 or more 100 Myr-duration rare galaxies in one structure is \(<0.01\%\). This suggests that the phenomenon is exceedingly rare, and yet there are several multiple DSFG-rich proto-clusters in only a few square degrees of data. The existence of these structures provides direct observational evidence that proto-clusters assemble in short-lived, stochastic bursts that likely correspond to the collapse of large-scale filaments on 10 Mpc scales. In this sense, such episodes represent “events” observed on the largest scales observed since the imprint of recombination from the CMB.

An alternate view may be that the gas potential wells of DSFGs in proto-clusters are much deeper, fueled by an excess of gas in the surrounding IGM. This point of view would argue for more long-lived DSFGs. If this is the case, then it is more likely that DSFGs in proto-clusters are triggered at somewhat arbitrary times determined only by their local \( < 1 \) Mpc surrounds. As a result, it is also likely that nearly every observed proto-cluster is DSFG-rich. The evidence that supports this claim is our estimate of the volume density to DSFG-rich proto-clusters and its agreement with the cluster mass function. If such DSFGs are short-lived, then at most \( \sim\)half of high-\( z \) proto-clusters should exhibit an enhanced DSFG-rich phase.

While different pieces of evidence support both possible explanations—short-lived, bursting proto-clusters or gas-enhanced proto-clusters—measurements of gas depletion times for DSFGs sitting in these structures suggests that they are indeed short-lived. Therefore, the former evolutionary scenario is favored, where DSFG-rich structures represent a short-lived phase of rapid growth across incredibly large filamentary structures in the IGM. More observations of such structures are needed to constrain the overall population of high-z overdensities, the diversity of their star formation histories, and to characterize the galaxies within such structures to learn how galaxy growth is governed by environment.

C.M.C. thanks the anonymous referee for helpful comments and the University of Texas at Austin, College of Natural Science for support. C.M.C. also thanks many colleagues for fruitful conversations which led up to this paper, including Chao-Ling Hung, Yi-Kuan Chiang, Asantha Cooray, Joel Primack, Amy Barger, Len Cowie, Cedric Lacey, William Cowley, James Bullock, Shea Garrison-Kimmel, Scott Chapman, Chuck Steidel, and Nick Scoville. C.M.C. would also like to acknowledge financial support and fruitful discussions from
