5-2023

Insights Into Star Formation and AGN Activity in Protocluster Environments From Morphological Studies and SED Fitting

Erik B. Monson
University of Arkansas-Fayetteville

Insights Into Star Formation and AGN Activity in Protocluster Environments From Morphological Studies and SED Fitting

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Physics

by

Erik B. Monson
University of Arkansas
Bachelor of Science in Physics, and Mathematical Sciences, 2017

May 2023
University of Arkansas

This dissertation is approved for recommendation to the Graduate Council.

Bret Lehmer, Ph.D.
Dissertation Director

Daniel Kennefick, Ph.D.
Committee Member

Woodrow Shew, Ph.D.
Committee Member

Abstract

In this dissertation I present work done from 2018-2023 to investigate the growth of galaxies and supermassive black holes (SMBH) in high redshift overdensities (protoclusters) by studying the star-forming galaxy and active galactic nucleus (AGN) population in the SSA22 protocluster.

I examined possible environmental sources of the enhanced star formation and AGN activity in the $z = 3.09$ SSA22 protocluster using *Hubble* WFC3 ~ 1.6 μm observations of the SSA22 field, including new observations centered on eight X-ray selected protocluster AGN. To investigate the role of mergers in the observed AGN and star formation enhancement, quantitative and visual morphological classifications were applied to F160W images of protocluster Lyman break galaxies (LBGs) in the fields of the AGN and $z \sim 3$ field LBGs in SSA22 and the GOODS-N field. No significant differences are found between the morphologies and merger fractions of protocluster and field LBGs, though this conclusion is limited by small number statistics in the protocluster. The UV-to-near-IR spectral energy distributions (SED) of F160W-detected protocluster and field LBGs were fit to characterize their stellar masses and star formation histories (SFH). These fits suggest that the mean protocluster LBG is a factor of ~ 2 times more massive and more attenuated than the mean $z \sim 3$ field LBG. These results suggest that ongoing mergers are no more common among protocluster LBGs than field LBGs, though protocluster LBGs appear to be more massive. The larger mass of the protocluster LBGs may contribute to the enhancement of SMBH mass and accretion rate in the protocluster, in turn driving the protocluster AGN enhancement.

I also analyzed the physical properties of 8 X-ray selected AGN and one candidate proto-quasar system (ADF22A1) in the SSA22 protocluster by fitting their X-ray-to-IR SEDs. SFHs were recovered for 7 of these systems which are well-fit by stellar population and AGN models.

The majority of the protocluster AGN systems, including ADF22A1, are consistent with obscured SMBH growth in star forming galaxies. The SFRs, SMBH accretion rates, and masses estimated from the SED fits for the 9 protocluster AGN systems are consistent with the distributions of SFR, accretion rate, SMBH masses, and stellar masses for a sample of X-ray detected AGN candidates in the Chandra Deep Fields (CDF). The ratio between the sample-averaged SSA22 SMBH mass and CDF SMBH mass is constrained to < 1.41. Furthermore, no statistically significant trends between the AGN or host galaxy properties and their location in the protocluster are found. The similarity of the protocluster and field AGN populations suggests that the observed protocluster AGN fraction enhancement may be driven by the SMBH mass distribution of the entire galaxy population, rather than only the AGN population, consistent with the interpretation of the results of the WFC3 F160W study of LBGs.

The results presented here add to the existing body of knowledge on the growth of galaxies and SMBHs in protoclusters, suggesting that galaxies and their black holes grow by the same mechanisms in protoclusters and field environments, and that the observed AGN enhancement is driven instead by more frequent, secular, triggering of AGN episodes in star forming galaxies.

Acknowledgements

The research presented in this dissertation was supported by STScI grant *HST*-GO-13844.013-A and NASA Astrophysics Data Analysis Program (ADAP) grant 80NSSC20K0444.

Much of the computation described in this work was performed on the Pinnacle cluster at the Arkansas High Performance Computing Center (AHPCC). The AHPCC is funded through multiple National Science Foundation grants and the Arkansas Economic Development Commission.

Thanks also go to my parents for their support, and to anyone who ever let me rant at and/or bounce ideas off of them. It's all part of the process.

Contents

List of Figures

List of Tables

List of Published Papers

Chapters 3 and 4 of this dissertation respectively have appeared or will appear in press as the following:

Monson, E. B., Lehmer, B. D., Doore, K., Eufrasio, R. T., Bonine, B., Alexander, D. M., Harrison, C. M., Kubo, M., Mantha, K. B., Saez, C., Straughn, A., and Umehata, H., 2021, "On the Nature of AGN and Star Formation Enhancement in the z = 3.1 SSA22 Protocluster: The HST WFC3 IR View", ApJ 919, 51, doi: 10.3847/1538-4357/ac0f84

Monson, E. B., Doore, K., Eufrasio, R. T., Lehmer, B. D., Alexander, D. M., Harrison, C. M., Kubo, M., Saez, C., and Umehata, H., 2023, submitted, "Revisiting the Properties of X-ray AGN in the SSA22 Protocluster: Normal SMBH and Host-Galaxy Growth for AGN in a $z = 3.09$ Overdensity", ApJ, in review

For the second article, the publication date of 2023 is anticipated. Monson et al. (2021) is presented here in its published form, while Monson et al. (2023, submitted) is presented as submitted.

Chapter 1

Introduction

1.1 Evolution of Galaxies and Black Holes

One of the fundamental questions in astrophysics is how galaxies evolve throughout the history of the Universe. Our Galaxy and the galaxies in our local neighborhood are the products of billions of years of evolution, driven by complex interactions interactions between both internal and external factors.

The environment around a galaxy is expected to play a key role in its evolution. The most massive galaxies in the local Universe are observed at the centers of galaxy clusters, large, virialized structures containing hundreds of galaxies and $\sim 10^{14}$ $M_{\odot}$. In cosmological simulations, these clusters form as a result of dark-matter mediated mergers of smaller, non-virialized "protoclusters" in the distant Universe. These simulations also predict the rapid growth of galaxies in these distant structures.

In observational studies, stark differences are seen between the galaxies in clusters and protoclusters, mirroring the overall evolution of star formation in the Universe. Cluster galaxies tend to be massive and quenched (i.e., not star-forming), while protocluster galaxies are less massive and more likely to be highly star forming (up to ~ 1000 $M_{\odot}$ yr^{-1}), in line with expectations of rapid galaxy growth in protoclusters.

Every massive galaxy that we know about hosts, at its center, a super-massive black hole (SMBH). These SMBHs accrete matter from their host galaxies, growing symbiotically with the galaxy. When growing rapidly enough, the SMBH can become visible as an active galactic nucleus (AGN). It is thought that all star forming galaxies spend a portion of their evolution as AGN, in the

period where their SMBHs are growing most rapidly. Correlations are also seen between SMBH growth and environment, mirroring the correlations with galaxy growth and environment. In clusters, active galaxies are rarer than in non-cluster (hereafter "field") environments (e.g Martini et al., 2013), whereas a number of studies suggest that high-redshift protoclusters have a larger fraction of galaxies hosting an AGN than field galaxies (e.g., Lehmer et al., 2009; Digby-North et al., 2010; Lehmer et al., 2013; Krishnan et al., 2017; Vito et al., 2020; Tozzi et al., 2022; see also Macuga et al., 2019 for a contrary result)[1]. Figure 1.1 shows a subset of these results plotted as a function of redshift. The implication of this trend is either that SBMHs in protocluster galaxies are growing more rapidly on average than field galaxies or that events that trigger AGN episodes are somehow more common; either explanation could be linked to a number of physical mechanisms, detailed below in Section 1.1.4.

Taken together, the relationships between star formation, AGN activity, and redshift in the protocluster environment suggest overall rapid evolution. Protoclusters, then, are important laboratories for the study of the buildup of stellar mass in the Universe and the assembly of SMBHs, including the evolution of the most massive galaxies in the Universe.

1.1.1 Galaxy Growth

The growth of galaxies is defined (for our purposes) by the formation of stars. The detailed history of stellar mass buildup in a galaxy is encoded in its star formation history (SFH), which describes its star formation rate (SFR) as a function of stellar age. Observationally, we cannot measure the "true," instantaneous SFR and must average over some period of time. The typical

[1]Note that these results are sensitive to how we define AGN (i.e., our X-ray luminosity cutoff for defining an AGN), which galaxies make up the denominator of the AGN fraction (e.g. Hydrogen-α emitters, Lyman-α emitters, Lyman Break galaxies, sub-millimeter galaxies, or a mass-complete sample), and similarly how we construct the field comparison sample.

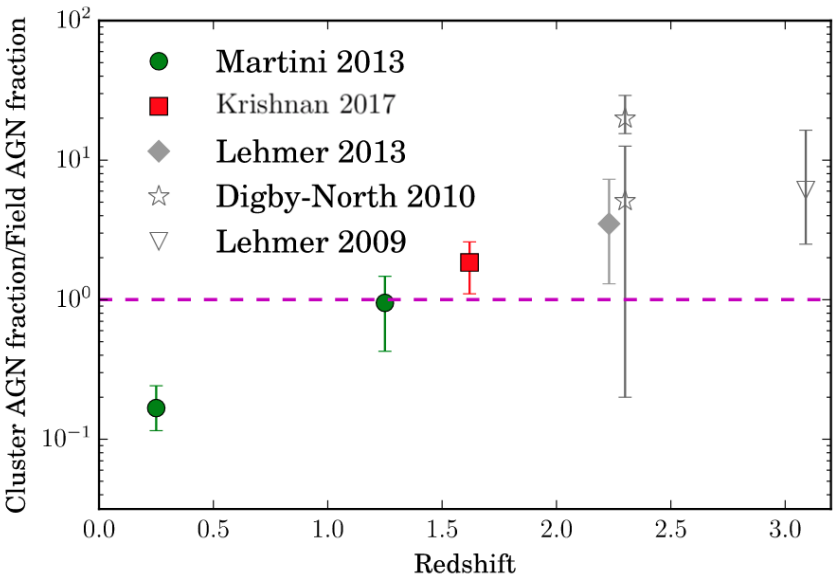

Figure 1.1: The cluster/field X-ray AGN fraction ratio is shown for a selection of clusters and protoclusters as a function of redshift. The high redshift protoclusters are typically seen to have significantly enhanced AGN fractions over the field (note the dashed line at unity), with the inversion from AGN enhancement to suppression occurring for $z \lesssim 1.5$. The high redshift protoclusters are in gray to denote their mass-incomplete sampling of the protocluster population. Filled symbols use a common selection of X-ray AGN with $L_X > 10^{43}$ erg s^{-1}. Adapted from Figure 8 in Krishnan et al. (2017).

tracers of SFR (e.g., intensity of the Hα emission line, far infrared emission, or some combination) are sensitive to the SFR over the past 100 Myr, so it is conventional to report the SFR of a galaxy averaged over the past 100 Myr.

On short timescales, the SFHs of galaxies are typically bursty: following an accretion episode, SFR increases rapidly and then decreases exponentially as fuel for further star formation is exhausted or heated beyond the point where gas clouds can continue to collapse. The source of these accretion episodes may be gas-rich mergers with other galaxies, which are categorized as minor, for accretion of small satellites by a larger galaxy (mass ratio $\mu = M_2/M_1 < 0.25$, where $M_1 > M_2$ by definition) or major, for collisions of galaxies of more equal mass ($\mu \geq 0.25$). From cosmological simulations, major mergers are expected to be a primary growth channel for the most massive galaxies. Mergers produce observable signatures as the two component galaxies interact and coalesce: galaxies observed in close pairs might be in the early stages of a merger, while actively merging or recently merged galaxies may exhibit morphological disturbances such as tidal tails, concentrated off-center star formation, or multiple nuclei.

When averaged over volume, we find that the SFR density (SFRD) of the Universe peaks at a redshift in $z = 1-2$, during the epoch of "cosmic noon," and decreases toward $z = 0$. Higher-mass galaxies are also seen to be less star-forming at present day than lower-mass galaxies, whereas the density of high-mass star-forming galaxies is larger at high redshift. This "downsizing" process is further evidence for the rapid evolution of the most massive galaxies at high redshift.

1.1.2 Co-evolution of Black Holes and Galaxies

The properties of SMBHs are tightly correlated with those of the host galaxy, especially the SMBH mass and the mass or velocity distribution of the host galaxy's central bulge. The very

low scatter in this relationship is taken as evidence that the SMBH and its host galaxy evolve in lockstep throughout the lifetime of the galaxy (see Kormendy & Ho, 2013, for a review). Further evidence for co-evolution comes from the observation of a trend in the SMBH accretion rate density with redshift similar to the trend in SFRD, and observations of a downsizing process in accreting SMBHs comparable to that in star-forming galaxies: actively growing SMBHs at lower redshift are found to be lower mass on average than accreting SMBHs at high redshift.

Co-evolution of star formation and SMBH growth should be expected, given that both processes draw from the same overall reservoir of cold gas in the galaxy. However, the degree of causality in the relationship remains an open question. The two processes operate on different scales: accretion of cold gas by the SMBH requires the gas to be driven down the central few parsecs of the galaxy. The precise mechanism by which gas is driven down to the scale of the SMBH are not currently fully known (see Alexander & Hickox, 2012, for a review), but stellar structures like the bars commonly found in disk galaxies are thought to play a role in the removal of angular momentum from gas, and star formation is also expected to play a direct role, with stellar winds from young, massive stars driving gas further into the potential well of the SMBH.

1.1.3 Active Galactic Nuclei

When enough material infalls to the central ≈ 10 pc of the galaxy and the accretion rate of the SMBH ($\dot{M}$) increases, the SMBH will become visible as an AGN. AGN are complicated systems, as pictured in Figure 1.2, which can produce radiation across the whole of the electromagnetic spectrum. The central engine of the AGN system is a luminous, Keplerian (i.e., orbiting according to classical, non-relativistic mechanics) accretion disk which forms around the SMBH. The disk efficiently converts the gravitational potential energy of the infalling matter into energy: if we

assume the bolometric luminosity L_{bol} of the AGN is

$$L_{bol} = \eta \dot{M} c^2, \tag{1.1}$$

the matter-energy conversion efficiency η can range from $\sim 5-40\%$, depending on the spin of the black hole. The accretion disk can persist down to an inner radius of approximately $3R_g = 6GM/c^2$, approximately 10^{-5} pc for a 10^8 $M_\odot$ SMBH[2] (Shakura & Sunyaev, 1973). The temperature profile of the AGN accretion disk is such that the luminosity density peaks in the UV portion of the electromagnetic spectrum, producing significant ionizing UV radiation with $E > 13.6$ eV. In the immediate vicinity of the accretion disk, photo-ionization of Hydrogen produces a hot electron gas, which interacts with optical and UV photons by reverse Compton scattering, scattering the photons up to X-ray energies.

In the common "unified" picture of the AGN system (see Urry & Padovani, 1995, for a review), a torus of dusty material (i.e., large molecules including silicate and carbonaceous grains) condenses beyond the accretion disk, $\sim 10^5 - 10^6$ $R_g \sim 1-10$ pc. The torus is optically thick to the UV and optical emission from the accretion disk, and it absorbs, scatters, and re-emits the accretion disk emission to near- and mid-IR energies (wavelengths $2-20$ μm). The hard ionizing radiation from the accretion disk and Comptonizing regions also photo-ionizes other atomic species, leading to line emission from recombination. Gas near the central SMBH orbits faster than the gas further out, and as such emission lines generated near the SMBH ($\sim 10^2 - 10^5$ $R_g \sim 10^{-3} - 1$ pc) are Doppler-broadened, while the emission lines generated at and beyond the radius of the torus ($\sim 10^5 - 10^8$ $R_g \sim 1$ pc $- 1$ kpc) in the ionizing cone above the disk are not. The torus is also optically thick to the broad emission lines, leading to the two primary classifications of the unified

[2]For the remainder of this section, size scales in parsecs are given for a 10^8 $M_\odot$ SMBH, corresponding to the scales in Figure 1.2.

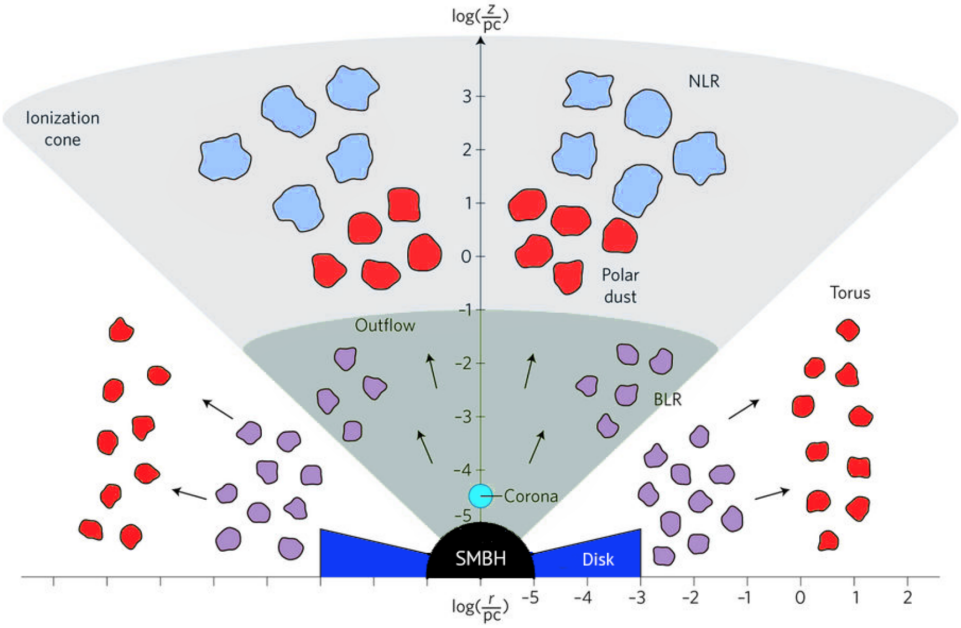

Figure 1.2: A cartoon schematic showing the emissive structures around an SMBH. An accretion disk (dark blue) forms around the central SMBH, surrounded by a toroidal structure of clumpy, dusty material (red). Fast moving clouds of ionized material orbit in the inner region (purple), producing doppler-broadened emission lines which are obscured by the torus on some lines of sight. Slower-moving clouds orbit further from the SMBH (light blue), producing narrow emission lines. A corona of hot electrons (cyan) near the SMBH produces the X-ray spectrum. Jets of charged particles may be launched along the spin axis of the SMBH, producing radio synchrotron emission. The observational signatures of the AGN thus depend on the line of sight through the dusty torus, relative to the axis of the SMBH. The scales are given in log pc; note that for a 10^8 solar mass black hole the gravitational radius is $\approx 10^{-5}$ pc. Figure 2 from Hickox & Alexander (2018), which was originally adapted from Ramos Almeida & Ricci (2017).

model of AGN. AGN where the observer's line of sight to the SMBH is unobscured by the torus have both broad and narrow emission lines in their optical spectrum and are classified as Type 1; AGN where the line of sight is obscured by the torus will have only narrow lines, and are classified as Type 2. In some cases, the threading of magnetic field lines through the accretion disk can cause jets to be launched along the spin axes of the SMBH, which will in turn produce radio synchrotron emission. Observers along the spin axis of the SMBH will then see a "radio-loud" AGN, while observers along other lines of sight may see less-intense or absent radio emission.

AGN are extremely powerful, with quasars reaching luminosities $L_{bol} > 10^{45}$ erg s^{-1} ($\gtrsim 10^{11}$ L$_\odot$), easily capable of outshining their host galaxy. The energy released by the AGN can impact the host galaxy in a variety of ways. Radiative mode feedback can occur when the interstellar medium (ISM) absorbs or scatters radiation from the AGN, thereby increasing the temperature of the ISM; kinetic mode feedback may occur when material launched by the AGN collides with the ISM, producing shocks. The prevalence and precise effects of these feedback mechanisms are a matter of ongoing study; it is as yet unclear under what conditions feedback may be positive (e.g., kinetic feedback triggering collapse of molecular clouds and thereby causing star formation) or negative (e.g., radiative feedback heating the ISM such the collapse of molecular clouds is inhibited, thereby inhibiting star formation, or kinetic feedback removing material from the galaxy). Regardless of the precise mechanism, it is very likely that AGN feedback has a direct effect on the intensity of star formation in the host galaxy.

It is thought that most star-forming galaxies spend a period of their life accreting rapidly enough to be visible as AGN. The fraction of time that such galaxies spend as AGN is the AGN duty cycle. The duty cycle can be estimated by computing the fraction of star forming galaxies in a population that are detected as AGN. However, this is subject to observational biases – Type 2 AGN can be missed by optical surveys, and even X-ray observations can miss the most heavily absorbed AGN (HI column densities $N_H \geq 10^{24}$ cm^{-2}).

1.1.4 The Role of Environment

The differences in the star formation rate densities and AGN fractions of protoclusters and low-redshift clusters suggest that protoclusters are the sites of rapid growth for both galaxies and their SMBHs, and that the growth mechanisms for protocluster galaxies and SMBHs may be different

from the field. In what follows, we will focus our search for differences in the growth mechanisms for SMBHs. There are a variety of possible explanations for how different evolutionary pathways may lead to the observed AGN fraction enhancement, which we may distill to two arguments: mergers or secular processes may lead to more frequent AGN triggering episodes, such that any given galaxy is more likely to become an AGN; or differences in the gas supply for protocluster galaxies may produce a longer AGN duty cycle, such that any given AGN persists longer and thus is more likely to be observed.

In protoclusters, where there is by definition a spatial over-density of galaxies, mergers between protocluster members may be more common than mergers of field galaxies, triggering more frequent bursts of star formation and AGN activity. In classical models of the evolution of luminous AGN (Hopkins et al., 2008), mergers are effective at driving gas down to the SMBH a times ~ 10 Myr after coalescence. However, these gas-rich major mergers may not be the dominant mode of growth in overdense environments (e.g. Romano-Díaz et al., 2014).

At high redshift, a typical galaxy has more cold gas available for star formation and SMBH accretion. Additionally, protocluster environments are often seen to have reservoirs of cold gas in the ISM. If coherent streams of cold gas from these reservoirs are able to penetrate the halo of the galaxy, the continuous replenishment of the galaxy's cold gas reservoir may allow star formation and SMBH accretion to maintained for longer timescales at higher intensities, leading to a longer AGN duty cycle. The greater availability of cold gas may also lead to larger stellar and SMBH masses in the protocluster, which could lead to a greater chance of any given galaxy being detected as an AGN, as accretion rate correlates with SMBH mass.

1.2 Outline of the Dissertation

First, Chapter 2 provides a brief introduction to the modeling procedures used in the remainder of this dissertation.

The following two chapters are articles published in or submitted to The Astrophysical Journal. These articles both approach the problem (stated above) of AGN fraction enhancement in protoclusters, the high-redshift progenitors of local rich galaxy clusters, by investigating possible causes of AGN enhancement in the SSA22 protocluster using 2D model fitting and SED fitting. They are included together in this form due to their common science goal and target.

Chapter 3 covers efforts to constrain the merger fraction of normal star forming galaxies in the SSA22 protocluster using a variety of methods, including 2D model fitting. Increased merger rates have been suggested as one possible cause for an elevated AGN fraction: more frequent mergers between galaxies in the protocluster could result in more frequent and more violent episodes of galaxy growth, in turn leading to larger average SMBH accretion rates and a larger AGN fraction. In this chapter SED fitting is also applied to determine the star formation rates and stellar masses of the star forming galaxies in the SSA22 protocluster which, if elevated over isolated galaxies, could also indicate the rapid galaxy growth needed to produce AGN.

Chapter 4 discusses direct analysis of the X-ray AGN in the SSA22 protocluster by SED fitting, measuring the star formation rates, stellar masses, black hole masses, and black hole accretion rates of the AGN. If any of theses properties are significantly in protocluster AGN than field AGN, it may indicate that the protocluster AGN are produced or grow through a different evolutionary pathway than field AGN.

Finally, Chapter 5 places the results of the previous two chapters in the context of the field, and

gives the outlook for future studies of galaxy and SMBH growth in protocluster environments.

References

Alexander, D. M., & Hickox, R. C. 2012, New A Rev., 56, 93, doi: 10.1016/j.newar.2011.11.003

Digby-North, J. A., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 407, 846, doi: 10.1111/j.1365-2966.2010.16977.x

Hickox, R. C., & Alexander, D. M. 2018, ARA&A, 56, 625, doi: 10.1146/annurev-astro-081817-051803

Hopkins, P. F., Hernquist, L., Cox, T. J., & Kereš, D. 2008, ApJS, 175, 356, doi: 10.1086/524362

Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511, doi: 10.1146/annurev-astro-082708-101811

Krishnan, C., Hatch, N. A., Almaini, O., et al. 2017, MNRAS, 470, 2170, doi: 10.1093/mnras/stx1315

Lehmer, B. D., Alexander, D. M., Geach, J. E., et al. 2009, ApJ, 691, 687, doi: 10.1088/0004-637X/691/1/687

Lehmer, B. D., Lucy, A. B., Alexander, D. M., et al. 2013, ApJ, 765, 87, doi: 10.1088/0004-637X/765/2/87

Macuga, M., Martini, P., Miller, E. D., et al. 2019, ApJ, 874, 54, doi: 10.3847/1538-4357/ab0746

Martini, P., Miller, E. D., Brodwin, M., et al. 2013, ApJ, 768, 1, doi: 10.1088/0004-637X/768/1/1

Ramos Almeida, C., & Ricci, C. 2017, Nature Astronomy, 1, 679, doi: 10.1038/s41550-017-0232-z

Romano-Díaz, E., Shlosman, I., Choi, J.-H., & Sadoun, R. 2014, ApJ, 790, L32, doi: 10.1088/2041-8205/790/2/L32

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Tozzi, P., Pentericci, L., Gilli, R., et al. 2022, A&A, 662, A54, doi: 10.1051/0004-6361/202142333

Urry, C. M., & Padovani, P. 1995, PASP, 107, 803, doi: 10.1086/133630

Vito, F., Brandt, W. N., Lehmer, B. D., et al. 2020, A&A, 642, A149, doi: 10.1051/0004-6361/202038848

Chapter 2

Technical Background

2.1 Forward Modeling in Astrophysics

In astrophysics, we have a hard limit on the number of observable quantities: in most cases, we are limited to measuring only the position, intensity, and wavelength (equivalently, frequency or energy) of an astrophysical signal, which are inevitably degraded or rendered uncertain by our instruments. We cannot directly measure, for example, the mass or star formation rate of a galaxy, or the accretion rate of its SMBH. Rather, we must model the physical conditions that produce the signal, convolve the model with the response functions of our instruments, and compare the result with observations. By doing so we can recover the model (or set of models) most consistent with our observations, and infer the physical conditions that produced the observed signal from the model. This technique, called forward modeling, is ubiquitous in astrophysics and a variety of other fields where direct measurements are infeasible.

2.1.1 Bayesian Analysis

Forward modeling requires a framework in which we can evaluate how well our models reproduce the observed data. Commonly, this is the Bayesian statistical framework. If $\mathbf{y}$[1] are our observations and θ our parameters, then Bayes' theorem states that

$$P(\theta|\mathbf{y}) = \frac{P(\mathbf{y}|\theta)P(\theta)}{P(\mathbf{y})}, \tag{2.1}$$

[1] Boldface text is used to represent vectors.

where $P(\mathbf{y}|\theta)$ is the likelihood (often denoted $\mathscr{L}$) that the model can reproduce the observations, $P(\theta)$ is the prior, representing our assumptions or prior knowledge of θ, $P(\mathbf{y})$ is the evidence, a constant depending on the data, and $P(\theta|\mathbf{y})$ is the posterior probability, combining the likelihood and posterior.

Often it suffices to determine the "maximum likelihood" model, representing the set of model parameters most likely to produce the observations. Under the assumption that our data $\mathbf{y}$ were drawn from a multivariate normal distribution with variances $\sigma\mathbf{y}$, the likelihood of the model is

$$\mathscr{L} = \exp\left(-\chi^2/2\right), \tag{2.2}$$

with

$$\chi^2 = \sum_i \frac{(y_i - \hat{y}_i(\theta))^2}{\sigma y_i^2}, \tag{2.3}$$

where $\hat{y}_i(\theta)$ is the model estimate of the ith observation under parameters θ. This likelihood is appropriate for most astrophysical sources in most wavebands, as we capture large numbers of photons from the source and our uncertainties on the photometry are approximately Gaussian. Note that the maximum likelihood corresponds to the typical minimum-χ^2 "best-fit" model, or, in the case where all of the data are equally weighted ($\sigma y_i = \text{const}$ for all i), the best-fit by least-squares. In cases where it is rare for the detector to capture a photon, it is often more appropriate to use a Poissonian likelihood function (Cash, 1979), or to use an approximation for the uncertainty based on low numbers of counts, such as the Gehrels (1986) approximation.

The maximum likelihood estimates for the model parameters give us the model best suited to the data, but in many cases do not provide the uncertainties on these estimates. Estimates of the uncertainties require us to solve an integral equation on the posterior distribution to construct the

desired confidence intervals. Thus, in cases where we require confidence intervals, we often need to sample the posterior distribution. This is most commonly done with Markov chain Monte Carlo (MCMC) algorithms, a class of Monte Carlo simulation where a Markov chain is constructed such that its stationary distribution approaches the sampling distribution of the posterior. This is accomplished by treating the parameters of the model as random variables, which are sampled from a continuous distribution. The model is computed for each new draw from the distribution and compared to the data using the appropriate likelihood. In the Metropolis Hastings (MH) MCMC algorithm, we define

$$a = \frac{P(\theta_{\mathbf{i+1}}|\mathbf{y})}{P(\theta_{\mathbf{i}}|\mathbf{y})} \tag{2.4}$$

$$= \frac{P(\mathbf{y}|\theta_{\mathbf{i+1}})P(\theta_{\mathbf{i+1}})}{P(\mathbf{y}|\theta_{\mathbf{i}})P(\theta_{\mathbf{i}})}. \tag{2.5}$$

New draws are then accepted with probability $\max\{a, 1\}$. If a draw is accepted it is added to the chain; if it is rejected, we return to the previous accepted draw and add it to the chain. In the MH algorithm the parameters of the proposal distribution are fixed. This may lead to poor sampling in cases where the variance of the proposal distribution is not well matched to the posterior distribution. Alternatives, such as the vanishing adaptive MCMC presented in Andrieu & Thoms (2008), or the affine-invariant MCMC of Goodman & Weare (2010) may be better suited to complicated problems in high-dimensionality parameter spaces. These algorithms efficiently sample the parameter space of the model, but they do not, nor are they designed to, find the maximum likelihood model. Nevertheless it is common to report a "best-fit" model when using an MCMC algorithm for fitting, as the most probable model from the MCMC chain is typically close to the true best fit.

2.1.2 2D Models

One common technique for studying the morphological characteristics of galaxies is to fit their surface brightness distributions (effectively, an image of the galaxy) with a 2D model that gives the spatial distribution of brightness. By convolving these models with the pixel size of the original image and the point spread function (PSF) of the instrument, the models can be compared directly to the images and the underlying surface brightness distribution of the galaxy can be inferred. In this technique we typically construct an uncertainty image, representing, at each pixel, the uncertainty in the value of the flux at that pixel, including both systematic instrumental uncertainties and random measurement uncertainties. We can then use a maximum likelihood method, such as a gradient descent algorithm, to find the best-fitting model. MCMC algorithms are typically not applied when 2D modeling, due to the computational expense of the 2D convolutions involved.

By assuming a functional form for the underlying surface brightness distribution (e.g., a Gaussian, an exponential disk, or the general Sérsic profile), this allows us to measure, for example, the effective sizes and concentrations of galaxies. We can also subtract the convolved models from the original image to search for features in the residual image. This can allow us to uncover signatures associated with mergers: multiple nuclei or other disturbances hidden by the main surface brightness profile, or low-surface brightness features like tidal tails which can be washed out by the outer wings of the main surface brightness profile.

2.1.3 Spectral Energy Distribution Modeling

The underlying physical properties of an astrophysical source are encoded in its spectral energy distribution (SED), the distribution of its emitted power as a function of wavelength. Any light-

measuring ("photometric") observation of the source can be thought of as a convolution of the bandpass and instrumental responses with the underlying SED. As such, modeling the SEDs of galaxies is a powerful tool for inferring their physical properties, such as stellar mass, SFR, and dust mass.

Due to this power, SED fitting is increasingly ubiquitous in the study of galaxy evolution, as deep surveys and large photometric datasets enable us to better model the underlying SEDs of galaxies. As such there are an increasing number of SED codes available: e.g., CIGALE (Boquien et al., 2019; Yang et al., 2020, 2022), PROSPECTOR (Johnson et al., 2021), BAGPIPES (Carnall et al., 2018), MAGPHYS (da Cunha et al., 2008)[2]. Throughout this dissertation I will present SED fits using LIGHTNING (Eufrasio et al., 2017; Doore et al., 2021), an SED-fitting code to which I have contributed AGN and X-ray emission modules (see Chapter 4). The newest version of LIGHTNING is described thoroughly in Doore, Monson, et al. (2023, submitted)[3].

To perform SED fitting, we must have detailed models of each of the components of the galaxy's total SED: the stellar population of the galaxy; the emission from the dust in the ISM, which is heated by radiation from the stellar population; and possibly emission from an AGN, if the galaxy hosts an accreting SMBH. In the below sections I introduce briefly each of these ingredients for SED fitting to provide context for their use in later chapters (for a more exhaustive introduction to the basics of SED fitting, see reviews by Walcher et al., 2011 and Conroy, 2013).

[2]This is necessarily an incomplete list; see e.g. sedfitting.org for a more extensive (yet still incomplete) list.
[3]See also lightning-sed.readthedocs.io

Stellar Population

The future history of a given, isolated[4] star is determined entirely by its initial, zero-age main sequence (ZAMS) mass and chemistry (commonly flattened to the metallicity, Z, the mass fraction of all elements heavier than Helium). To model an entire population of stars, one assumes a set of evolutionary steps, or isochrones, giving the evolutionary state of all stars of a particular age. It is also necessary to assume an initial mass function (IMF), which gives the distribution of ZAMS masses. There are a variety of prescriptions for the IMF (e.g. Salpeter, 1955; Kroupa, 2001; Chabrier, 2003), each with a different shape and normalization. A number of solutions now exist for modeling the evolution of individual stars and populations (e.g. FSPS, Conroy et al., 2009; Conroy & Gunn, 2010; PÉGASE Fioc & Rocca-Volmerange, 1997, 1999).

If we assume that all of the stars in our population were created instantaneously, we can track the evolution of this burst of stars to an arbitrary age. By then integrating these starburst populations over time, we can construct stellar population models that represent the stellar population over a range of ages. In Figure 2.1 we show examples of these instantaneous burst models created with PÉGASE, and in Figure 2.2 we show the integrated models built from these bursts. We can then use the spectra of these populations in our SED model as the components in a linear combination, which allows us to decompose the observed SED into starlight from different ages, giving us a measure of the SFH.

If $\tilde{L}_{\nu,i}$ is the intrinsic luminosity density of the binned stellar population for $t_i \leq t < t_{i+1}$ (i.e., the spectra in Figure 2.2), normalized to a SFR of 1 $M_\odot$ yr^{-1} then we can construct a piecewise approximation of the SFH,

$$\psi(t) = \psi_i,\ t_i \leq t < t_{i+1} \tag{2.6}$$

[4]The evolution of binary stars is more complex, due to factors including mass transfer and tidal forces.

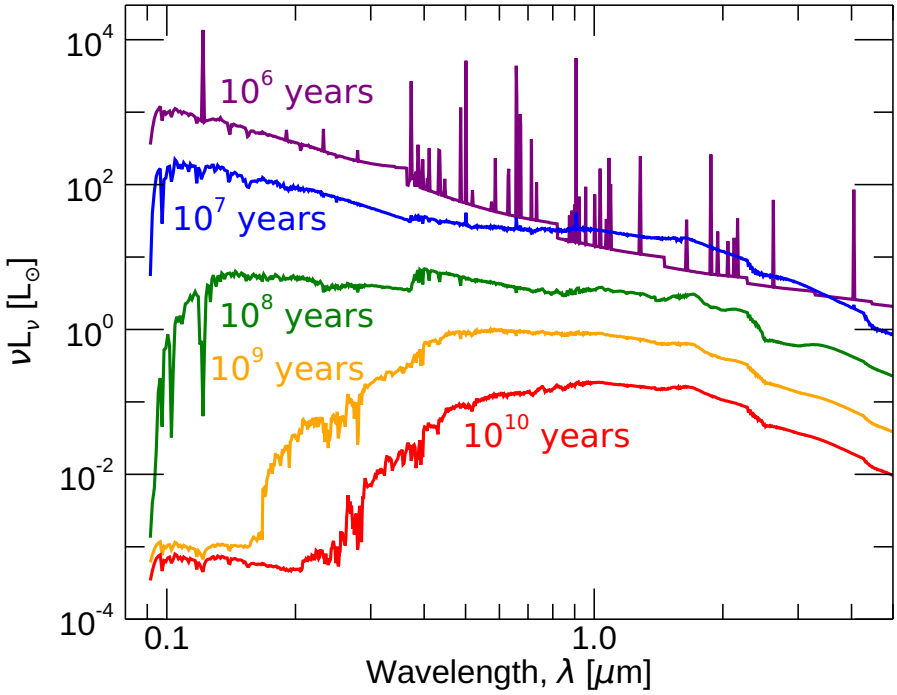

Figure 2.1: Single-age stellar populations with an initial mass of 1 $M_\odot$ created with PÉGASE, at ages from 10^6–10^{10} years. Spectral features associated with cooler, lower mass stars become more prominent as the age of the population increases, as the hotter, high mass stars evolve off the main sequence. In these models, the spectrum of the 1 Myr-old population also contains nebular emission lines, which are generated by photoionization of the gas surrounding the stars by UV radiation from hot, high-mass stars. Figure 1 from Doore, Monson, et al. (2023, submitted).

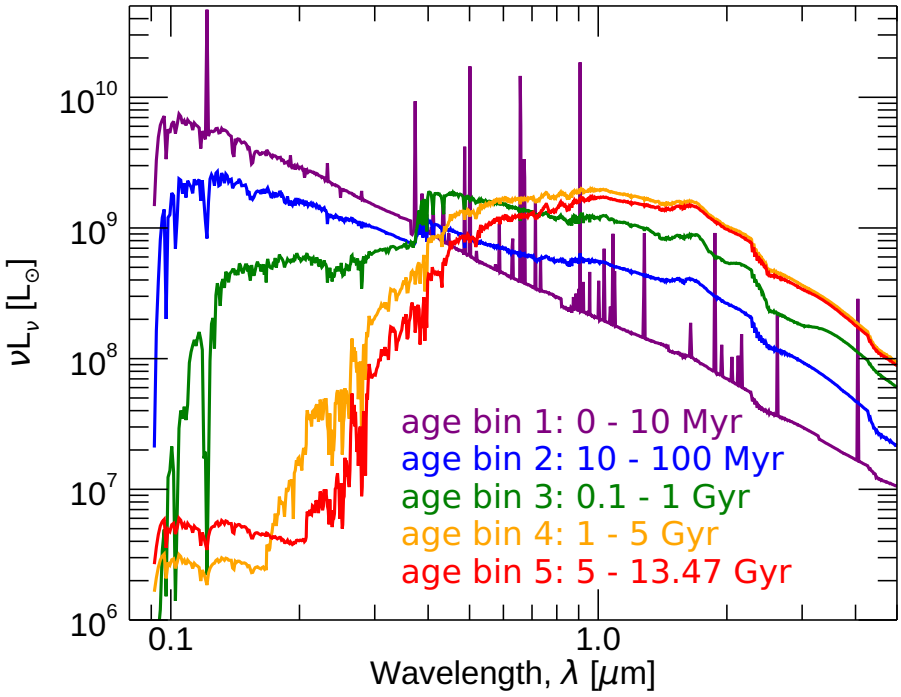

Figure 2.2: Binned stellar populations created by integrating the evolution of the burst from Figure 2.1. As in Figure 2.1, nebular emission lines are visible in the spectrum of the youngest population. Figure 2 from Doore, Monson, et al. (2023, submitted).

by solving the equation

$$\tilde{L}_\nu = \sum_i^N \psi_i \tilde{L}_{\nu,i}, \tag{2.7}$$

where $\tilde{L}_\nu$ is the intrinsic luminosity density of the whole galaxy. Clearly we are not limited to this piecewise-constant approximations of the SFH, and could instead use a functional form of our choosing; piecewise-constant forms are popular due to their combination of flexibility and simplicity. Solution of the equation above is the crux of SED fitting: for a given set of $\tilde{L}_{\nu,i}$ we fold $\tilde{L}_\nu$ through our instrumental responses, compare it to our observations, and search for the set of ψ_i that best models the observations. However, in reality galaxies are not composed entirely and only of stars; there are a number of complicating factors and other components that must be included in the model, detailed below.

Dust Absorption

The stars in galaxies are embedded in the ISM, which can absorb and scatter photons out of the light cone we observe. The effects of absorption and scattering attenuate the light from stars more efficiently at some wavelengths than others. The combined effects of absorption and scattering are often represented as the optical depth τ_λ, such that

$$L_\nu = \tilde{L}_\nu e^{-\tau_\lambda}, \tag{2.8}$$

where L_ν is the observed luminosity density and $\tilde{L}_\nu$ is the intrinsic (i.e. unabsorbed) luminosity density. Since ISM attenuation varies as a function of wavelength, it affects both the intensity and color of the source. It is thus necessary for our SED models to include the effects of ISM attenuation, by modeling $\tau_\lambda(\lambda)$. A variety of models exist to do so, including Calzetti et al. (2000), Noll et al. (2009), and Doore et al. (2021). In the simplest case, these models (e.g. Calzetti et al.,

2000; Noll et al., 2009) assume that the galaxy is a sphereoid, with stars uniformly distributed throughout the ISM, such that the optical depth is the same along any line of sight into the galaxy. More complicated models (e.g. Doore et al., 2021) may assume a disk geometry, where the optical depth varies with the inclination to the line of sight into the galaxy.

Dust Emission

Absorption of starlight by the ISM heats the dust grains in the ISM, which then radiate in the IR. The dominant component of this emission is thermal, but there are wide bands of emission from molecular transitions, including emission by polycyclic aromatic hydrocarbons (PAH). The major assumptions of different dust emission models are the composition and size distribution of the dust grains, and the intensity distribution of the radiation field the dust is exposed to (e.g. Draine & Li, 2007). Implementations of dust emission may also assume energy balance, in which the power attenuated by the ISM dust is identically equal to the re-radiated power (when integrated over all lines of sight). Note that dust models may not always be included in SED fits, especially when only optical-to-near-IR data is available, as the emitted power from dust peaks in the far-IR.

AGN Emission

As described in Chapter 1, AGN can potentially be bright enough to outshine their host galaxy. Measuring the host galaxy properties of AGN hosts by SED fitting thus often requires modeling the emission from the AGN simultaneously. Templates of AGN SEDs are often produced by radiative transfer methods, where an assumed accretion disk spectrum (often a power law; the UV-optical continuum emission from the AGN accretion disk is observed to be relatively flat and featureless) is combined with a physical model of the dusty torus to generate a model spectrum of the entire

accretion disk and torus system. Examples include Fritz et al. (2006) and Stalevski et al. (2016).

It is increasingly common to include the X-ray emission from AGN in SED fitting procedures, since the AGN X-ray luminosity provides a strong constraint on the bolometric AGN luminosity. The AGN X-ray spectrum is typically modeled as a power law, though more complex and physically-motivated models are available (e.g. Kubota & Done, 2018).

References

Andrieu, C., & Thoms, J. 2008, Statistics and Computing, 18, 30, doi: 10.1007/s11222-008-9110-y

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103, doi: 10.1051/0004-6361/201834156

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682, doi: 10.1086/308692

Carnall, A. C., McLure, R. J., Dunlop, J. S., & Davé, R. 2018, MNRAS, 480, 4379, doi: 10.1093/mnras/sty2169

Cash, W. 1979, ApJ, 228, 939, doi: 10.1086/156922

Chabrier, G. 2003, PASP, 115, 763, doi: 10.1086/376392

Conroy, C. 2013, ARA&A, 51, 393, doi: 10.1146/annurev-astro-082812-141017

Conroy, C., & Gunn, J. E. 2010, ApJ, 712, 833, doi: 10.1088/0004-637X/712/2/833

Conroy, C., Gunn, J. E., & White, M. 2009, ApJ, 699, 486, doi: 10.1088/0004-637X/699/1/486

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595, doi: 10.1111/j.1365-2966.2008.13535.x

Doore, K., Eufrasio, R. T., Lehmer, B. D., et al. 2021, ApJ, 923, 26, doi: 10.3847/1538-4357/ac25f3

Doore, K., Monson, E. B., Eufrasio, R. T., et al. 2023, submitted, ApJ, in review

Draine, B. T., & Li, A. 2007, ApJ, 657, 810, doi: 10.1086/511055

Eufrasio, R. T., Lehmer, B. D., Zezas, A., et al. 2017, ApJ, 851, 10, doi: 10.3847/1538-4357/aa9569

Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950

—. 1999, arXiv e-prints, astro. https://arxiv.org/abs/astro-ph/9912179

Fritz, J., Franceschini, A., & Hatziminaoglou, E. 2006, MNRAS, 366, 767, doi: 10.1111/j.1365-2966.2006.09866.x

Gehrels, N. 1986, ApJ, 303, 336, doi: 10.1086/164079

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65, doi: 10.2140/camcos.2010.5.65

Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22, doi: 10.3847/1538-4365/abef67

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Kubota, A., & Done, C. 2018, MNRAS, 480, 1247, doi: 10.1093/mnras/sty1890

Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793, doi: 10.1051/0004-6361/200912497

Salpeter, E. E. 1955, ApJ, 121, 161, doi: 10.1086/145971

Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288, doi: 10.1093/mnras/stw444

Walcher, J., Groves, B., Budavári, T., & Dale, D. 2011, Ap&SS, 331, 1, doi: 10.1007/s10509-010-0458-z

Yang, G., Boquien, M., Buat, V., et al. 2020, MNRAS, 491, 740, doi: 10.1093/mnras/stz3001

Yang, G., Boquien, M., Brandt, W. N., et al. 2022, ApJ, 927, 192, doi: 10.3847/1538-4357/ac4971

Chapter 3

The *HST* WFC3 IR View of Galaxies in the SSA22 Protocluster

Erik B. Monson, Bret D. Lehmer, Keith Doore, Rafael T. Eufrasio, Brett Bonine, David M. Alexander, Chris M. Harrison, Mariko Kubo, Kameswara B. Mantha, Cristian Saez, Amber Straughn, and Hideki Umehata

The following chapter was originally published in The Astrophysical Journal, Volume 919, Issue 1, as article 51, titled "On the Nature of AGN and Star Formation Enhancement in the $z = 3.1$ SSA22 Protocluster: The *HST* WFC3 IR View"

Abstract

We examine possible environmental sources of the enhanced star formation and active galactic nucleus (AGN) activity in the $z = 3.09$ SSA22 protocluster using *Hubble* WFC3 F160W (~ 1.6 μm) observations of the SSA22 field, including new observations centered on eight X-ray selected protocluster AGN. To investigate the role of mergers in the observed AGN and star formation enhancement, we apply both quantitative (Sérsic-fit and Gini-M_{20}) and visual morphological classifications to F160W images of protocluster Lyman break galaxies (LBGs) in the fields of the X-ray AGN and $z \sim 3$ field LBGs in SSA22 and GOODS-N. We find no statistically significant differences between the morphologies and merger fractions of protocluster and field LBGs, though we are limited by small number statistics in the protocluster. We also fit the UV-to-near-IR spectral energy distributions (SED) of F160W-detected protocluster and field LBGs to characterize their stellar masses and star formation histories (SFH). We find that the mean protocluster LBG is by a factor of ~ 2 times more massive and more attenuated than the mean $z \sim 3$ field LBG. We take our results to suggest that ongoing mergers are not more common among protocluster LBGs than field LBGs, though protocluster LBGs appear to be more massive. We speculate that the larger mass of the protocluster LBGs contributes to the enhancement of SMBH mass and accretion rate in the protocluster, which in turn drives the observed protocluster AGN enhancement.

3.1 Introduction

Galaxy clusters, the largest virialized, gravitationally bound structures in the Universe, are currently thought to form by the dark-matter driven mergers of protoclusters, smaller groupings of galaxies in the early universe (i.e. $z \gtrsim 2$; lookback times $\gtrsim$ 10 Gyr). In current cosmological models, the most luminous modern galaxies are assembled in protoclusters by mergers of smaller galaxies, where gas-rich mergers may trigger active galactic nuclei (AGN) and episodes of star formation. These protoclusters can move along dark matter filaments toward dense nodes, where they merge and collapse to form the clusters observed at lower redshifts (e.g., Boylan-Kolchin et al., 2009). Protoclusters have been discovered at redshifts ranging from $z \sim 2$ (e.g. Venemans et al., 2002; Miley et al., 2004; Capak et al., 2011; see Overzier, 2016 for a review) as far as $z \sim 8$ (in the BoRG58 field; Trenti et al., 2012), in the epoch of galactic re-ionization, less than 1 Gyr after the Big Bang. Observations of protoclusters provide direct constraints on galaxy evolution, supermassive black hole (SMBH) growth, and the formation of large-scale structures and their galactic constituents.

The protocluster in the SSA22 survey field (R.A.: $22^{\rm h}17^{\rm m}34.7^{\rm s}$, Dec.: $+0^\circ15'7''$; Cowie et al., 1994) was discovered by Steidel et al. (1998) as a spike in the redshift distribution of Lyman-break galaxies (LBGs) at $z = 3.09$. Further observations of the same region revealed a $\sim$six-fold overdensity of LBGs consistent with a galaxy cluster in the early stages of development (Steidel et al., 2000, 2003). Simulations suggest that the SSA22 protocluster and structures of similar scale and overdensity will evolve toward Coma-like (virialized mass $\sim 10^{14}$ $M_\odot$) clusters at redshift $z = 0$ (Governato et al., 1998). Recent surveys (e.g., Toshikawa et al., 2016, 2018; Higuchi et al., 2019) of candidate protoclusters suggest that progenitors of Coma-like clusters are not uncommon,

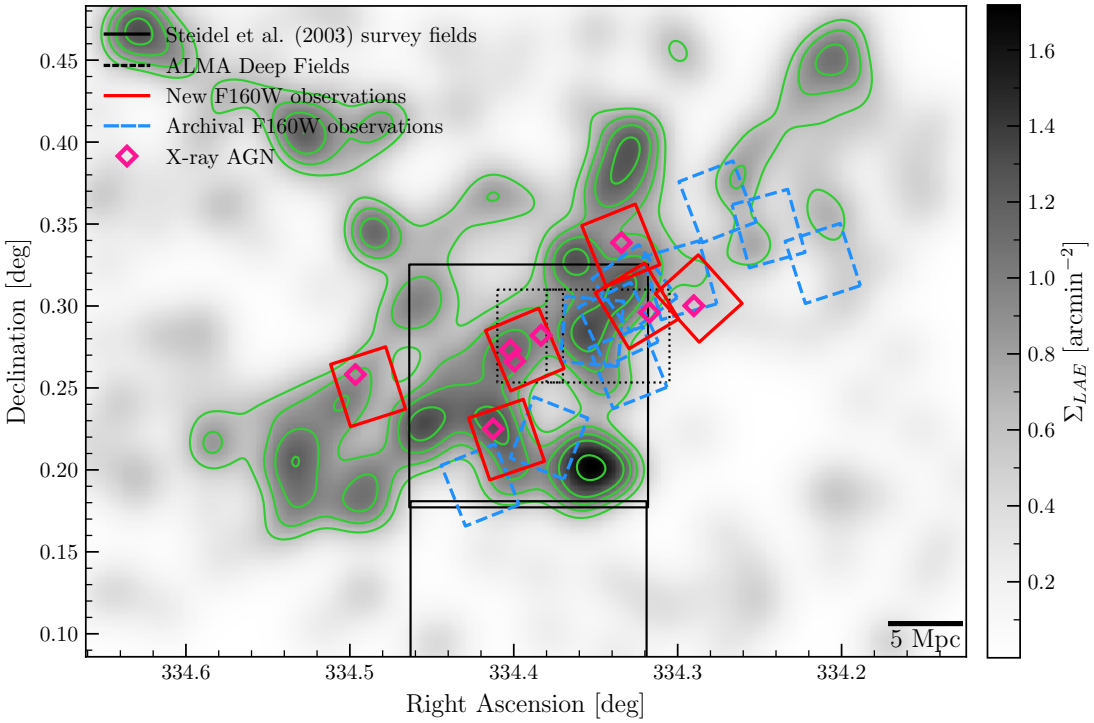

Figure 3.1: Rectangles show the position and orientation of our observations (thick red lines) and archival observations (dashed blue lines) relative to the protocluster, which is shown as a surface density map of $z = 3.1$ LAEs from Hayashino et al. (2004). To highlight the general shape of the protocluster we also show contours at levels $\{0.6, 0.8, 1.0, 1.2, 1.6\}$ arcmin^{-2}. We show the positions of the eight X-ray detected protocluster AGN from Lehmer et al. (2009a) and Alexander et al. (2016) as open magenta diamonds. For reference we also show the SSA22 fields studied in Steidel et al. (2003) (their SSA22a and SSA22b) in solid black lines, the approximate footprints of the ALMA deep fields studied by Umehata et al. (2015, 2017, 2018, 2019) in dotted black lines, and a 5 co-moving Mpc scalebar.

with perhaps 76% of 4σ significant overdensities projected to evolve into clusters with masses $\sim 5 \times 10^{14}$ $M_{\odot}$, though conclusive spectroscopic identifications of protoclusters remain difficult and possibly biased toward younger galaxies due to the reliance on Lyα emission (Toshikawa et al., 2016, 2018).

Follow-up narrowband observations over an area 10 times larger than that studied in Steidel et al. (1998) using SUPRIME-Cam have further identified ∼six-fold overdense bands of Lyα emitting galaxies (LAEs), spectroscopically confirmed as a set of three large-scale filamentary structures (Hayashino et al., 2004; Matsuda et al., 2005). The co-moving scale of the largest of these filaments, shown in Figure 3.1, is on the order of 60 Mpc long and 10 Mpc wide, with a redshift range of 3.088–3.108, making it one of the largest mapped structures at $z \sim 3$ (Matsuda et al., 2005). Large Lyα emitting nebulae (Lyα blobs; LABs) around star forming galaxies have been shown to be associated with these filaments, suggesting that they are the precursors of massive galaxies developing in the regions of greatest overdensity (Matsuda et al., 2005). The filamentary structure of the protocluster has been further established by MUSE spectral-imaging observations mapping the filaments in emission around the protocluster core (Umehata et al., 2019). The scale and detail at which the filamentary structures in SSA22 have been mapped remains relatively unique among high-redshift protoclusters, though this is also changing: Harikane et al. (2019) mapped candidate protoclusters at $z > 6$ on scales of > 100 comoving Mpc, and Daddi et al. (2021) have recently imaged filamentary structures in Lyα emission around an overdensity of galaxies at $z = 2.9$.

The SSA22 region has also been well-studied in millimeter/sub-millimeter bands (Tamura et al., 2009; Umehata et al., 2014, 2015, 2017, 2018, 2019; Alexander et al., 2016). Umehata et al. (2015) identified a concentration of 8 dusty star-forming galaxies (DSFGs) associated with the intersection of the major filamentary structures at the center of the protocluster. *Herschel*

SPIRE measurements suggest that DSFGs in the SSA22 protocluster account for a star formation rate density on the order of 10^3 $M_\odot$ yr^{-1} Mpc^{-3}, a factor of $\sim 10^4$ increase in star formation rate density over the field at redshift $z = 3.09$ (Kato et al., 2016). Additionally, five of the SSA22 DSFGs are associated with X-ray luminous AGN, and two are associated with LABs. Further observations of the protocluster core with higher resolution sub-millimeter instruments have revealed more DSFGs, at least 10 of which are spectroscopically confirmed protocluster members, possibly indicating preferential formation of these galaxies in the densest region of the protocluster (Umehata et al., 2017, 2018). Wide-spectrum (u^*-band to *Spitzer* IRAC 8 μm) photometry and spectral energy distribution (SED) fitting by Kubo et al. (2013) has additionally suggested the presence of very massive galaxies in the densest regions of the protocluster. Overall, these observations support a picture of the SSA22 protocluster environment as one where massive, intensely star-forming galaxies are actively forming.

Chandra X-ray observations of the region have revealed a higher rate of AGN activity compared to field galaxies at $z \sim 3$ (Lehmer et al., 2009a), with $9.5^{+12.7}_{-6.1}$% and $5.1^{+6.8}_{-3.1}$% of protocluster LBGs and LAEs, respectively, hosting AGN with $L_{8-32\ \mathrm{keV}} \gtrsim 3 \times 10^{43}$ ergs s^{-1}. These fractions are elevated by a factor of ≈ 6 compared to non-protocluster galaxies (hereafter, "field" galaxies) at $z \sim 3$, indicating a possible enhancement of SMBH growth in the protocluster. Similar enhancements have been observed in other $z = 2$–3 protoclusters (e.g., HS1700, Digby-North et al., 2010; 2QZCluster, Lehmer et al., 2013; DRC, Vito et al., 2020); there is also evidence that AGN fraction in overdense environments evolves with redshift (Martini et al., 2013), with overdensities at higher redshifts having larger AGN fractions and modern clusters having lower AGN fractions than the field.

The AGN and highly star-forming galaxies in the protocluster are consistently found to be

associated with the larger scale structure of the protocluster, embedded in the intersection of the filaments and giant Lyα nebulae. Of the 8 X-ray detected protocluster AGN from Lehmer et al. (2009a), 4 are associated with LABs, 2 of which are giant LABs larger than 100 kpc in scale (Alexander et al., 2016). The implication, then, is that the enhanced AGN and star formation activity are driven by environmental factors unique to the protocluster. These enhancements may be driven by accretion episodes caused by an elevated merger rate among protocluster members, or by secular gas accretion from shared gas reservoirs (Narayanan et al., 2015) and filamentary structures in the intergalactic medium (Umehata et al., 2019). However, the AGN enhancement may also be driven by the presence of more massive galaxies (and hence SMBH) in the protocluster, as compared to the field at $z \sim 3$; it has been established that galaxies in $2 \lesssim z \lesssim 4$ protoclusters are on average more massive than their field counterparts at the same redshift (e.g., Steidel et al., 2005; Hatch et al., 2011; Cooke et al., 2014).

Motivated by the elevated AGN fraction observed in the protocluster, Hine et al. (2015) used archival *Hubble* ACS F814W optical observations of SSA22 (probing rest-frame UV emission) to visually classify LBGs in the protocluster, finding a marginally enhanced merger fraction among protocluster galaxies ($48 \pm 10\%$) as compared to field galaxies ($30 \pm 6\%$). However these results are limited by the small-number statistics of the protocluster and suffer from ambiguous interpretation due to the patchiness of rest-frame UV observations, which are highly influenced by star formation activity and attenuation. High-resolution near-infrared (i.e., 1–2 μm) observations, which probe rest-frame optical wavelengths at $z = 3.09$, can better trace the stellar mass content of the protocluster galaxies, less influenced by individual bursts of star formation and more sensitive to merger activity.

In this work, we investigate the possible contributions of mergers to the increased AGN and SF

Table 3.1: Summary of *HST* WFC3 F160W survey fields.

ID	R.A (deg)	Dec. (deg)	Pos. Angle (deg)	Exp. Time (s)	Proposal No.	PI
New Observations[a]						
TARGET1	334.29	0.3046	-47.95	5223.50	13844	Lehmer
TARGET2	334.32	0.3001	121.10	5223.50	13844	Lehmer
TARGET3	334.33	0.3371	111.91	5223.50	13844	Lehmer
TARGET4	334.39	0.2734	-67.27	5223.50	13844	Lehmer
TARGET5	334.40	0.2185	-71.17	5223.50	13844	Lehmer
TARGET6	334.49	0.2507	-72.31	5223.50	13844	Lehmer
Archival Observations[b]						
SSA22AC6M4	334.42	0.1908	111.75	2611.75	11735	Mannucci
SSA22AC30	334.33	0.2624	111.77	2611.75	11735	Mannucci
SSA22AM16	334.38	0.2194	68.60	2611.75	11735	Mannucci
SSA22AM38C48	334.33	0.3109	126.09	2611.75	11735	Mannucci
SSA-22A-IR	334.34	0.2888	115.00	2611.75	11636	Siana
SSA-22A-IR2	334.35	0.2846	85.00	2611.75	11636	Siana
SSA22.4.IR	334.21	0.3260	108.00	2611.75	14747	Robertson
SSA22.5.IR	334.24	0.3473	-74.60	2611.75	14747	Robertson
SSA22.6.IR	334.28	0.3636	111.00	2611.75	14747	Robertson
SSA22.7.IR	334.30	0.3157	106.00	2611.75	14747	Robertson

Notes – [a]Shown in red in Figure 3.1. [b]Shown in blue in Figure 3.1.

activity in the protocluster by applying three separate morphological analysis techniques to galaxies detected in *Hubble* WFC3 infrared (F160W, $\lambda_p = 1.537$ μm) observations in SSA22, targeting the environment around the X-ray detected protocluster AGN studied in Alexander et al. (2016). We use parametric model fitting techniques to extract Sérsic model parameters from F160W-detected LBGs, compare the measured morphologies of protocluster LBGs to a sample of field LBGs in GOODS-N, and analyze the residuals after Sérsic model subtraction for indications of merger activity. We additionally use nonparametric morphological measures (the Gini coefficient G, moment of light M_{20}, and concentration C) to compare the morphologies of protocluster and field galaxies and attempt to classify mergers. Lastly, we apply a similar visual analysis as Hine et al. (2015) to our sample of F160W-detected LBGs to compare the rest-frame optical merger fraction for the SSA22 protocluster to the field at $z \sim 3$.

We also investigate the possibility that more massive galaxies are the driver of AGN and SF enhancement by fitting the SEDs of a subset of our F160W-detected SSA22 LBG sample. We compare the distribution of stellar mass and the mean star formation history of protocluster LBGs to a sample of $z \sim 3$ field LBGs in GOODS-N.

The paper is organized as follows: Section 3.2 describes the observations, data reduction, catalog generation, and sample selection; Section 3.3 describes our analysis of the morphological properties of protocluster galaxies; Section 3.4 describes our analysis of the SEDs and physical properties of protocluster galaxies; in Section 4.5 we discuss our results and attempt to connect the morphologies and physical properties of protocluster galaxies to the protocluster environment; lastly, in Section 4.6 we summarize our results and their implications for understanding the galaxy assembly process at $z \sim 3$.

Coordinates in this work are J2000, magnitudes are given in the AB system, and we adopt a

Kroupa (2001) IMF. We adopt a cosmology with $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$, yielding a lookback time of 11.42 Gyr, a 2.04 Gyr universe age, and a physical scale of 7.63 proper kpc arcsec^{-1} at $z = 3.09$. We use proper scales when discussing lengths on the scale of galaxies, and co-moving scales when discussing lengths on the scale of the protocluster itself. Hereafter, we define SSA22 protocluster galaxies as those galaxies in our F160W images with $3.06 \leq z \leq 3.12$. Over this range of redshift, the rest-frame wavelength probed by our WFC3 IR observations ranges from $3740 - 3790$ Å.

3.2 Data Analysis

3.2.1 Data Reduction

We use sixteen *HST* WFC3 F160W images of the SSA22 field for our morphological analyses (*HST* proposals 13844, 11735, 11636, 14747). We summarize the locations and exposure times of these fields in Table 3.1 and show their footprints superimposed on the protocluster structure in Figure 3.1. Six of these fields (*HST* proposal 13844) were new observations obtained to cover the eight protocluster AGN detected in *Chandra* observations by Lehmer et al. (2009a) that were studied by Alexander et al. (2016). These six observations were taken at two-orbit depth for ease of comparison to the CANDELS-Wide fields (Koekemoer et al., 2011; Grogin et al., 2011).

The STScI AstroDrizzle package[1] was used for image recombination and data reduction. We re-bin our images to a scale of 0.065 arcsec pixel^{-1}; at this scale the PSF FWHM of the images is ≈ 2.6 pixels. We adopt an inverse variance weighting scheme when combining the exposures. We use the inverse variance weight maps generated by AstroDrizzle to create a map (a "sigma

[1] http://drizzlepac.stsci.edu

image") of the estimated total standard deviation (in electrons) for each pixel in the sky-subtracted science image as

$$\sigma = \sqrt{p_{>0} + \frac{1}{\text{weight}}},$$

where $p_{>0}$ denotes the greater of the science image pixel value (in electrons) and zero.

We adopt a model PSF based on the median of 33 isolated stars in our two-orbit depth WFC3 F160W images, identified using SExtractor (Bertin & Arnouts, 1996) by making a box selection in the MAG_AUTO–FLUX_RADIUS plane; we select sources with $0.13'' <$ FLUX_RADIUS $< 0.15''$ and $21.1 <$ MAG_AUTO < 22.8 as stars.

3.2.2 Catalog Generation

We use GALAPAGOS-2 (Häussler et al., 2013; Barden et al., 2012) to carry out source detection (using SExtractor), background determination, model fitting (using GALFITM, an updated multi-band version of GALFIT; see Vika et al., 2013), and catalog compilation. We briefly describe the methods adopted for the catalog generation tasks here, since they are relevant to our methodology (for a more thorough description, see Barden et al., 2012).

SExtractor is designed to de-blend sources so that nearby sources are detected and cataloged separately. To prevent over-de-blending of bright sources while still detecting faint sources in deep images, GALAPAGOS-2 runs SExtractor twice, once with a low threshold for detection and once with a high threshold. Detections from the high-threshold run are accepted and cataloged automatically; detections from the low-threshold run are rejected if they fall inside the isophotal ellipse of a high-run source, and accepted otherwise. Science image and sigma image cutouts are then constructed for each source.

SExtractor is known to overestimate the sky level (Häussler et al., 2007), so GALAPAGOS-2 determines a local sky value of its own and holds the sky fixed during the later model-fitting step. The procedure is described in detail in Barden et al. (2012). Briefly, GALAPAGOS-2 identifies the brightest secondary sources in each cutout, enlarges their SExtractor Kron isophotes by a pre-determined factor (default 2.5 times), masks out every pixel inside the resulting ellipses, and takes the median of the remaining pixels as the local background.

To generate our preliminary source catalog, the detection criteria for the "hot" low-threshold run was 15 pixels $\geq 2.5\sigma$ above the background after convolution with a Gaussian kernel with a FWHM of 2 pixels (i.e., the default 5x5 pixel Gaussian kernel supplied with SExtractor). For the "cold," high-threshold run the detection criteria were strengthened to 30 pixels $\geq 3.5\sigma$ above the background; no convolution filter was applied on the cold run. Sources were de-blended using 64 thresholds and a minimum contrast of 0.001. The SExtractor Kron isophotes were enlarged by a factor of 2.5 when combining the hot and cold catalogs.

To estimate the completeness of our preliminary catalog under these SExtractor parameters, we added a total of 3000 simulated stars and 3000 simulated galaxies (generated using GALFIT, with morphological parameters uniformly selected from ranges taken from the van der Wel et al. (2012) fits to $z \sim 3$ GOODS-N galaxies) to our two-orbit images and re-detected them with the same procedures. We find that our preliminary catalog is 90 (50) percent complete to a F160W magnitude of 24.8 (25.7) for galaxies and 25.1 (26.2) for point sources detected in two-orbit images. We repeated the process for the single-orbit images, finding 90 (50) percent completeness limit magnitudes of 24.7 (25.4) for galaxies and 24.8 (26.0) for point sources.

To clean spurious or un-physical SExtractor detections from the catalog, we require detections to have $S/N \geq 5$ in a $1''$ diameter aperture, `FWHM_IMAGE` > 2.6 pix, and 50% `FLUX_RADIUS` $>$

0. Following the above procedures results in a "main catalog" of 7538 F160W-detected sources.

We cross-match detections in our catalog with sources in reference catalogs, provided they are within $1''$ of the position returned by SExtractor. To identify LBGs we use the catalogs produced by Steidel et al. (2003) (hereafter S03) and Micheva et al. (2017) (hereafter M17). The M17 LBG sample is an expanded version of an LBG sample from Iwata et al. (2009), containing U-dropouts with VLT/VIMOS followups that confirm redshifts $z \geq 3.06$; while this sample overlaps in part with the Steidel et al. (2003) LBG sample we keep S03 and M17 LBGs separate in subsequent figures and analysis out of an abundance of caution concerning possible differences in, e.g., color selection criteria. To identify LAEs we use the catalog from Yamada et al. (2012) and the M17 catalog. We identify AGN based on the X-ray point source catalog from Lehmer et al. (2009a) and Lyman-continuum (LyC) emitters based on the M17 catalog. For spectroscopic redshifts, if not available in one of the aforementioned catalogs, we have taken redshifts from the spectroscopic SSA22 surveys by Saez et al. (2015) and Kubo et al. (2015a), and the VLT-VIMOS Deep Survey (VVDS) (Le Fèvre, O. et al., 2013). We find spectroscopic redshifts from the above references for 216 of the sources in our catalog. For the subsequent analysis, we require galaxies to have spectroscopic redshifts in order to conclusively identify them as protocluster members or $z \sim 3$ field galaxies. We find that of these sources, 91 have redshifts $2.9 \leq z \leq 3.3$, with 72 galaxies in the protocluster redshift range ($3.06 \leq z \leq 3.12$), and thus 19 galaxies in the field redshift range. By construction, the M17 LBGs have a spectroscopic redshift; for S03 LBGs we estimate that our requirement of spectroscopic redshifts may exclude as many as four protocluster LBGs from our subsequent analysis. For additional photometry (covering u^*-band to *Spitzer* IRAC 8 μm; see Table 3.6 for the full list of filters), we have used the photometric catalog of Kubo et al. (2013). We also include narrowband magnitudes at 4972 Å (the observed wavelength of the Lyman α line at

Table 3.2: Catalog excerpt for protocluster LBGs with acceptable fits as defined in Section 3.3.1, showing the GALFITM-extracted morphological parameters.

ID	S/N[a]	m_{F160W}[b]	$\delta m_{\mathrm{F160W}}$	r_e[c] (arcsec)	δr_e (arcsec)	n[d]	δn	q[e]	δq	PA[f] (deg)	δPA (deg)	z[g]	z Src.[h]	Σ_{LAE}[i] (arcmin^{-2})	Vis. Class.[j]	S03 Name[k]
J221710.35+001920.8	223.5	23.50	0.07	0.09	0.03	1.15	1.41	0.45	0.22	47.76	17.56	3.103	M17	0.64	–	–
J221704.34+002255.8	188.2	24.80	0.32	0.18	0.11	1.28	3.04	0.72	0.52	-42.55	98.42	3.108	M17	0.77	–	–
J221732.04+001315.6	183.6	22.93	0.06	0.32	0.03	0.65	0.17	0.78	0.06	30.23	14.57	3.065	M17	0.67	C2	–
J221737.92+001344.1	166.5	24.61	0.19	0.19	0.07	0.59	0.95	0.73	0.30	-68.91	49.33	3.094	S03a	1.26	M2	SSA22a-MD14
J221731.69+001657.9	163.4	23.94	0.05	0.28	0.03	0.47	0.19	0.58	0.06	-56.33	7.60	3.088	S03a	0.91	C2	SSA22a-M28
J221720.25+001651.7	157.0	23.32	0.06	0.38	0.03	0.82	0.18	0.73	0.06	-78.20	9.94	3.098	S03a	0.84	C2	SSA22a-C35
J221718.87+001816.2	156.8	24.06	0.06	0.09	0.02	1.78	1.16	0.77	0.20	12.01	32.32	3.089	S03e	1.20	M2	SSA22a-D17
J221718.96+001444.5	135.1	23.78	0.10	0.19	0.03	0.90	0.64	0.53	0.16	21.98	13.89	3.091	M17	0.54	–	–
J221701.38+002031.9	133.6	24.33	0.13	0.32	0.08	0.99	0.65	0.27	0.12	-14.07	9.27	3.073	M17	0.73	–	–
J221720.20+001731.6	123.1	24.61	0.07	0.19	0.02	1.42	0.73	0.22	0.17	-16.46	6.21	3.065	M17	1.07	–	SSA22a-C47
J221718.04+001735.5	120.8	24.43	0.09	0.27	0.05	1.78	1.18	0.30	0.12	31.35	7.55	3.093	M17	0.86	–	–
J221731.51+001631.0	118.2	24.79	0.09	0.26	0.05	0.63	0.38	0.50	0.13	56.67	11.93	3.098	M17	0.88	–	SSA22a-M25
J221740.98+001127.2	108.7	24.28	0.10	0.21	0.04	0.95	0.66	0.32	0.16	69.28	8.14	3.093	M17	0.59	C2	–
J221719.30+001543.8	106.4	23.56	0.07	0.58	0.04	0.51	0.15	0.44	0.04	-75.46	3.67	3.097	S03a	0.64	M3	SSA22a-C30
J221736.90+001712.8	98.1	25.08	0.08	0.29	0.04	1.17	0.51	0.24	0.09	0.22	5.07	3.099	M17	0.95	–	SSA22a-M31
J221721.02+001708.9	79.0	25.13	0.05	0.26	0.00	0.26	0.06	0.41	0.01	68.08	0.15	3.076	S03e	1.03	C2	SSA22a-C39

Notes – [a]Signal-to-noise in a 1″ diameter aperture. [b]Integrated Sérsic model magnitude. [c]Sérsic model effective radius. [d]Sérsic model index. [e]Sérsic model axis ratio b/a. [f]Sérsic model position angle. [g]Redshift from literature; see Section 3.2.2. [h]Redshift source: S03e=Steidel et al. (2003) Lyα emission; S03a=Steidel et al. (2003) absorption-line; M17=Taken from Micheva et al. (2017) catalogs. [i]Local LAE surface density; see Section 3.5.2. [j]Consensus visual classification; see Section 3.3.4. [k]S03 catalog designation.

$z = 3.1$) from Yamada et al. (2012), and we report $\log L_X$ (measured from 2–8 keV, approximately 8–32 rest-frame keV) for the X-ray sources in Lehmer et al. (2009a). We estimate that for the majority of the reference catalogs the number of possible mismatches with our catalog is on the order of a few galaxies. For the larger Kubo et al. (2013) photometric catalog, the number of mismatches could be as large as 200, though this is still $\sim 11\%$ of the overall number of matches we find with the Kubo et al. (2013) catalog. We note that these numbers of false matches are likely overestimated, since the angular separations of the matches are typically much less than an arcsecond, and there is a large number of sources in our main catalog. Excerpts of the main catalog are provided in Table 3.2 and Table 3.3 for protocluster LBGs with acceptable fits from GALFITM (as defined in Section 3.3.1).

In Figure 3.2 we show how we divide our main catalog into sub-samples (based on requirements for spectroscopic redshift, etc.) for the analyses presented in the following sections. Due to the availability of an additional S03 LBG sample in GOODS-N, and concerns about how different LBG color selection criteria might harm any protocluster-field comparisons, we focus the majority of the following analysis on the sample of S03 LBGs in our catalog. Our main catalog contains 26 S03 LBGs with $2.9 \leq z \leq 3.3$, to a maximum $R-$band magnitude of 25.4. These 26 $z \sim 3$ S03 LBGs amount to 13 protocluster LBGs and 13 field LBGs. Our main catalog also contains 13 LBGs in the same redshift range which are unique to the M17 catalog; these have a maximum $R-$band magnitude of 25.5, and amount to 11 protocluster LBGs and 2 field LBGs.

3.2.3 Comparison LBG Sample in GOODS-N

We constructed an additional comparison sample of LBGs in the GOODS-N field based on the the S03 catalog. The S03 GOODS-N catalog contains 40 LBGs with $2.9 \leq z \leq 3.3$, to a maximum

$R-$band magnitude of 25.6. Sérsic parameters for these galaxies were retrieved by cross-matching with the van der Wel et al. (2012) single-Sérsic fitting catalog. We find that a subset of 33 $z \sim 3$ LBGs have acceptable single-Sérsic fits from van der Wel et al. (2012). We use these Sérsic parameters to increase the size of our field galaxy comparison sample in the analysis described in Section 3.3.1.

We use UV-to-mid-IR (U-band to *Spitzer* MIPS 70 μm) photometry from the Barro et al. (2019) catalog in the CANDELS survey areas for our comparison LBG sample in GOODS-N, retrieved from the *Rainbow* database[2]. We searched the catalogs for the closest match within $1''$ to each LBG in the Steidel et al. (2003) GOODS-N sample. We estimate that $\lesssim 10$ galaxies could be mismatched with the CANDELS photometric catalog, though this could amount to a significant fraction of the 40 S03 LBGs in GOODS-N. However, we again expect that this number is overestimated based on the small angular separations between matches. We list the filters used for this photometry in Table 3.6.

3.3 Morphological Analysis

3.3.1 Parametric Morphology Fitting

To analyze the morphologies of our detected galaxies in the SSA22 field, we began by fitting 2D parametric models to the data. The surface brightness profile of an elliptical or spheroidal galaxy without a well-resolved disk is well described by the Sérsic law, a symmetric profile specified by two parameters, the Sérsic index n and effective radius r_e:

$$I(r) = I_e \exp\left[-b_n\left(\left(\frac{r}{r_e}\right)^{\frac{1}{n}} - 1\right)\right], \tag{3.1}$$

[2]http://rainbowx.fis.ucm.es/Rainbow_navigator_public/

where I_e is the surface brightness at $r = r_e$ and b_n satisfies

$$\Gamma(2n) = 2\gamma(2n, b_n); \tag{3.2}$$

Γ and γ are the complete and lower incomplete gamma functions, respectively (Graham & Driver, 2005). In the general case of an elliptical profile, the r in the equation above is a function of the profile's center and elliptical axis ratio $q = b/a$.

Multiple studies have demonstrated that single Sérsic model fitting with GALFIT (Peng et al., 2002) can be used to extract galaxy morphologies from large *HST* datasets (e.g. GEMS, Häussler et al., 2007, and CANDELS, van der Wel et al., 2012). GALFIT has the ability to de-blend nearby sources by simultaneous fitting, allowing accurate photometric measurements in crowded images and the examination of galaxies with close projected companions for evidence of mergers. GALAPAGOS-2 uses GALFITM (Vika et al., 2013) for single Sérsic model fitting. GALFITM is a modified version of GALFIT, which retains all of the same functionality and runs on the same Levenberg-Marquardt algorithm.

The GALAPAGOS-2 fitting procedure distinguishes between the "primary" source in a cutout, i.e. the main source currently being fit, and "secondary" sources, nearby objects bright enough to bias the photometry of the primary. For accurate fitting to the primary source, secondary sources must be fit simultaneously. Sources are sorted and fit in order of decreasing brightness, and every source gets a turn as the primary. If a secondary source is present and was already fit (i.e., if the secondary source is brighter than the current primary), the parameters from that fit are reused and held fixed (provided the secondary is in the same image as the primary source; see Figure 9 in Barden et al., 2012). While secondary sources are fit simultaneously, sources that are present in the cutout but are faint enough compared to the primary are not fit, and instead masked out such

that GALFITM ignores any pixels corresponding to these sources. The fit results presented in Table 3.2 and in what follows represent the primary fit to each source.

Following the generation of the source catalog, GALFITM was used to fit a single Sérsic profile to each F160W detection. We focus on single component fits for individual galaxies, reasoning that for the $z \sim 3$ we are primarily interested in bulge-like and disk-like components were unlikely to be resolved separately. The initial guesses and constraints for our fits come from SExtractor parameters. For a given object, the initial guess for the Sérsic model magnitude is the SExtractor `MAG_BEST`; the initial value for r_e is the 50% `FLUX_RADIUS` raised to the power of 1.4. The initial value for the Sérsic index is 2.5 for all galaxies, and the initial position of the Sérsic model is the position determined by SExtractor.

GALFIT (and by extension, GALFITM) allows parameter value ranges to be limited and coupled. The following bounds on r_e, n, and the Sérsic profile magnitude m were adopted for this work:

$$0.002'' < r_e < 26'';$$

$$0.2 < n < 10;$$

$$|m_{SE} - m| < 5,$$

where m_{SE} is the magnitude reported by SExtractor. In pixel units the constraint on Sérsic radius is $0.3 < r_e < 400$, where the lower limit is hard-coded into GALAPAGOS-2. These constraints are a slight relaxation of the GALAPAGOS-2 defaults, which are themselves selected to do a good job of keeping the fit from wandering into unphysical regions of the parameter space without being overly restrictive. The center of the model is constrained so that it can only move

within 10% of the cutout size from the initial position.

Following van der Wel et al. (2012), we flagged fits where the primary fit is unlikely to represent the galaxy well as unacceptable, based on the following criteria. We flagged fits where any of the final parameter values were equal to one of the bounds listed above and fits where the final Sérsic index was equal to the initial value of 2.5 as unacceptable. The GALFIT algorithm occasionally converges on an arbitrarily small axis ratio for low-S/N objects with small apparent sizes. For this reason, we also flagged fits with axis ratios less than 0.125. Apparently well-converged fits that do not represent the data well also have exceptionally large errors in the resulting total Sérsic magnitude and effective radius, so we also flagged fits where the magnitude error estimated by GALFITM is greater than 5 mag (more typical errors are on the order of 0.08 mag for acceptable fits and 1.4 mag for unacceptable fits), and flagged fits where the recovered effective radius was consistent with 0 within 1σ.

Of our main catalog, 2833 detections (37.6% of detections) have acceptable single Sérsic fits according to the criteria above. Sources with acceptable fits tend to be brighter and have larger effective sizes than sources with bad fits; foreground stars and other point-like sources tend to have bad fits, as do faint, compact galaxies like LAEs. In terms of SExtractor parameters, poorly fit sources tend to have $23.5 \leq$ `MAG_BEST` ≤ 26.1 and (in pixel units) $4.41 \leq$ `FWHM_IMAGE` ≤ 9.84; sources with acceptable fits tend to have $21.9 \leq$ `MAG_BEST` ≤ 24.7 and $5.71 \leq$ `FWHM_IMAGE` ≤ 13.25. The fraction of acceptable fits increases slightly among sources with $2.9 \leq z \leq 3.3$: 40 (44.0%) have acceptable Sérsic fits. In Figure 3.2 we show how the sources with acceptable fits break down between the protocluster and field, and the numbers of LBGs and other categories of galaxies with acceptable fits. Among the sources with acceptable fits, we identify 29 spectroscopically-confirmed protocluster members, 9 of which are LBGs from the S03 catalog

Table 3.3: Catalog excerpt for protocluster LBGs with acceptable fits as defined in Section 3.3.1, showing additional photometry.

ID	u^*	B	V	R	i'	z'	J	K	$3.6\mu m$	$4.5\mu m$	$5.8\mu m$	$8.0\mu m$
J221710.35+001920.8	26.57	25.31	24.31	24.11	23.94	23.91	23.40	22.78	23.07	23.13	22.32	23.05
J221704.34+002255.8	26.77	25.13	24.23	24.08	23.96	23.99	23.04	22.65	23.04	22.95	22.40	22.85
J221732.04+001315.6	27.61	25.66	24.49	24.16	23.85	23.85	23.26	22.34	22.11	22.02	22.12	21.49
J221737.92+001344.1	26.82	26.10	25.21	24.91	24.66	24.61	24.13	23.49	23.25	23.29	–	22.53
J221731.69+001657.9	27.98	26.46	25.42	25.28	24.99	25.16	25.34	23.29	23.25	23.03	22.78	21.98
J221720.25+001651.7	27.57	25.80	24.83	24.56	24.36	24.36	24.16	23.18	23.63	23.69	24.76	23.27
J221718.87+001816.2	26.41	25.50	24.78	24.73	24.62	24.80	24.45	24.20	23.27	23.38	22.84	21.92
J221718.96+001444.5	26.95	25.64	24.69	24.55	24.37	24.33	24.13	23.22	23.15	22.97	23.78	22.86
J221701.38+002031.9	27.50	26.68	25.43	25.08	24.78	24.78	24.58	23.01	23.30	23.01	23.29	22.29
J221720.20+001731.6	–	–	–	–	–	–	–	–	–	–	–	–
J221718.04+001735.5	28.79	26.45	25.64	25.30	25.08	24.91	25.27	23.68	20.42	20.88	22.87	22.42
J221731.51+001631.0	–	–	–	–	–	–	–	–	–	–	–	–
J221740.98+001127.2	26.86	26.01	25.14	24.93	24.75	24.81	23.98	23.78	23.09	23.09	–	–
J221719.30+001543.8	26.80	25.66	24.82	24.68	24.54	24.45	25.43	23.27	22.14	21.55	21.82	21.72
J221736.90+001712.8	–	–	–	–	–	–	–	–	–	–	–	–
J221721.02+001708.9	27.20	26.20	25.43	25.28	25.29	25.39	24.72	23.17	24.22	24.54	–	–

Notes – In each column we show the AB magnitude of the galaxy in the given band, as measured by Kubo et al. (2013). We label the IRAC bands (channel 1–4) by their reference wavelemgth (3.6 μm$-$8.0 μm).

and 7 of which are LBGs exclusive to the M17 catalog. Only 4 of the 8 X-ray detected protocluster AGN from Lehmer et al. (2009a) have acceptable Sérsic fits under the criteria above; the fits to the rest-frame optically bright AGN (e.g., the quasar SSA22a-D12) do not return especially meaningful or well-constrained Sérsic model parameters due to their point source-like profiles.

We present the parameters derived from model fitting in Figure 3.3; to broaden the field galaxy comparison sample, we include single-Sérsic-fit parameters from the 33 galaxies in our GOODS-N comparison sample with good fits from van der Wel et al. (2012) (see Section 3.2.3) in the figure.

To assess whether or not the morphologies of the protocluster and field samples are drawn from the same underlying population, we use 1-D and 2-D two-sample Kolmogorov-Smirnov (hereafter, KS) tests, under the null hypothesis that the results for the protocluster and field samples are drawn from the same distribution. For a fair comparison between LBGs selected by the same color criteria, we initially limited the tests to S03 LBGs. The tests are consistent with the null hypothesis that the S03 protocluster and S03 field LBGs are drawn from the same morphological population. We then performed the same KS tests with the addition of the M17 LBGs, finding again that all the tests on the parametric morphologies are consistent with the null hypothesis. We show the results for both sets of 1-D and 2-D KS tests in Table 3.4.

We note the one-to-two orbit depth of our images is not ideal for parametric model fitting; however, in Section 3.A we use simulated galaxies similar to our LBG sample to investigate how decreasing signal-to-noise affects the reliability of our fits. We find that the fits are generally reliable for low-n galaxies with $S/N \gtrsim 100$. Noting that only two galaxies in Table 3.2 fall below this rough threshold and that their Sérsic model parameters appear well constrained according to the error estimates from GALFITM, we are confident in the reliability of the fits to the LBG samples we use above.

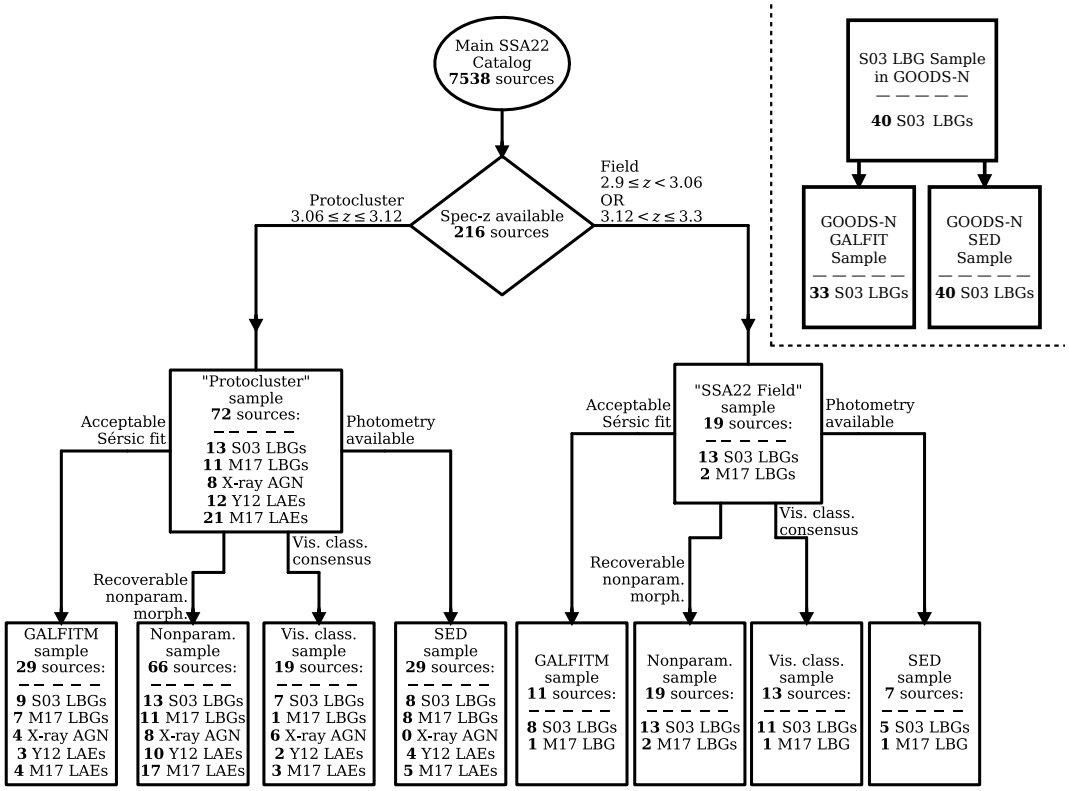

Figure 3.2: We illustrate how each of the protocluster and field subsamples used throughout this paper stem from the main SSA22 catalog. We require galaxies to have spectroscopic redshifts from the literature to distinguish between protocluster and field galaxies. To be included in the analysis of our GALFITM *sample*, we require sources to have acceptable Sérsic fits, as defined in Section 3.3.1. In our analysis of the *non-parametric morphology sample*, we require sources to have a recoverable nonparametric morphology using the procedures described in Section 3.3.3. In our *visual classification sample*, we include only sources for which our classifiers reached a consensus about whether the system was merging or an isolated galaxy, as described in Section 3.3.4. Our *SED sample* includes only galaxies for which we could retrieve photometry from the Kubo et al. (2013) photometric catalog. For each subsample, we show the total number of sources as well as the number in different classes and the reference to the classification in the literature. While we focus on the LBG and AGN populations in subsequent sections we also list the number of LAEs in each sample. We abbreviate the references as: S03 – Steidel et al. (2003); Y12 – Yamada et al. (2012); M17 – Micheva et al. (2017). The X-ray AGN are those studied in Alexander et al. (2016); they are all found in the protocluster, and our SED fitting sample excludes them by construction. Due to the nature of the narrowband LAE selection, there are no field LAEs in our catalog. In the upper right corner we show the breakdown of our *comparison LBG sample in GOODS-N*, made up of 40 GOODS-N LBGs from the S03 catalog, 33 of which have acceptable single-Sérsic fits from van der Wel et al. (2012).

In hierarchical models, galaxies with bulge-dominated early-type morphologies are assembled by past mergers. More evolved galaxies that have experienced a number of past mergers should then have larger Sérsic indices. Since we probe a young population by relying on a Lyman break selected sample, larger-than-expected Sérsic indices could be an indication of more rapid, merger-driven morphological evolution. We note that the majority of our LBGs have more disk-like morphologies with $n < 2.5$, and all of the protocluster LBGs fall below this line. The protocluster LBGs do not tend to have larger n than field galaxies, nor do they tend to be larger, suggesting that the morphologies of LBGs in the protocluster are not evolving faster than their field counterparts at $z \sim 3$. We note that in Section 3.A we find that our Sérsic model fits typically recover n smaller than the "true" value of n due to the broadening effects of the PSF on the Sérsic model. For the 4 protocluster LBGs with recovered Sérsic indices $1.25 < n < 2$, we thus expect that the true Sérsic index could be as much as a factor of 2 larger in the case of $S/N \gtrsim 100$, possibly indicating an underlying bulge dominated morphology smeared out by the effects of the PSF. However, we note that the PSF also affects the field LBGs, and given again that the protocluster and field LBGs cluster together strongly in the space of the Sérsic model parameters we are confident in assessing that the protocluster LBGs are not more morphologically evolved than the field LBGs.

3.3.2 Residual Image Analysis

For the subset of 25 SSA22 LBGs (16 protocluster and 9 field) in our sample with acceptable fits, we show the residuals after model subtraction in Figure 3.4 and Figure 3.5 for protocluster and field galaxies, respectively. We used the `Python`-based tidal feature extraction pipeline[3] from Mantha et al. (2019) to examine our single-Sérsic fit residuals for evidence of potential tidal

[3] https://github.com/AgentM-GEG/residual_feature_extraction

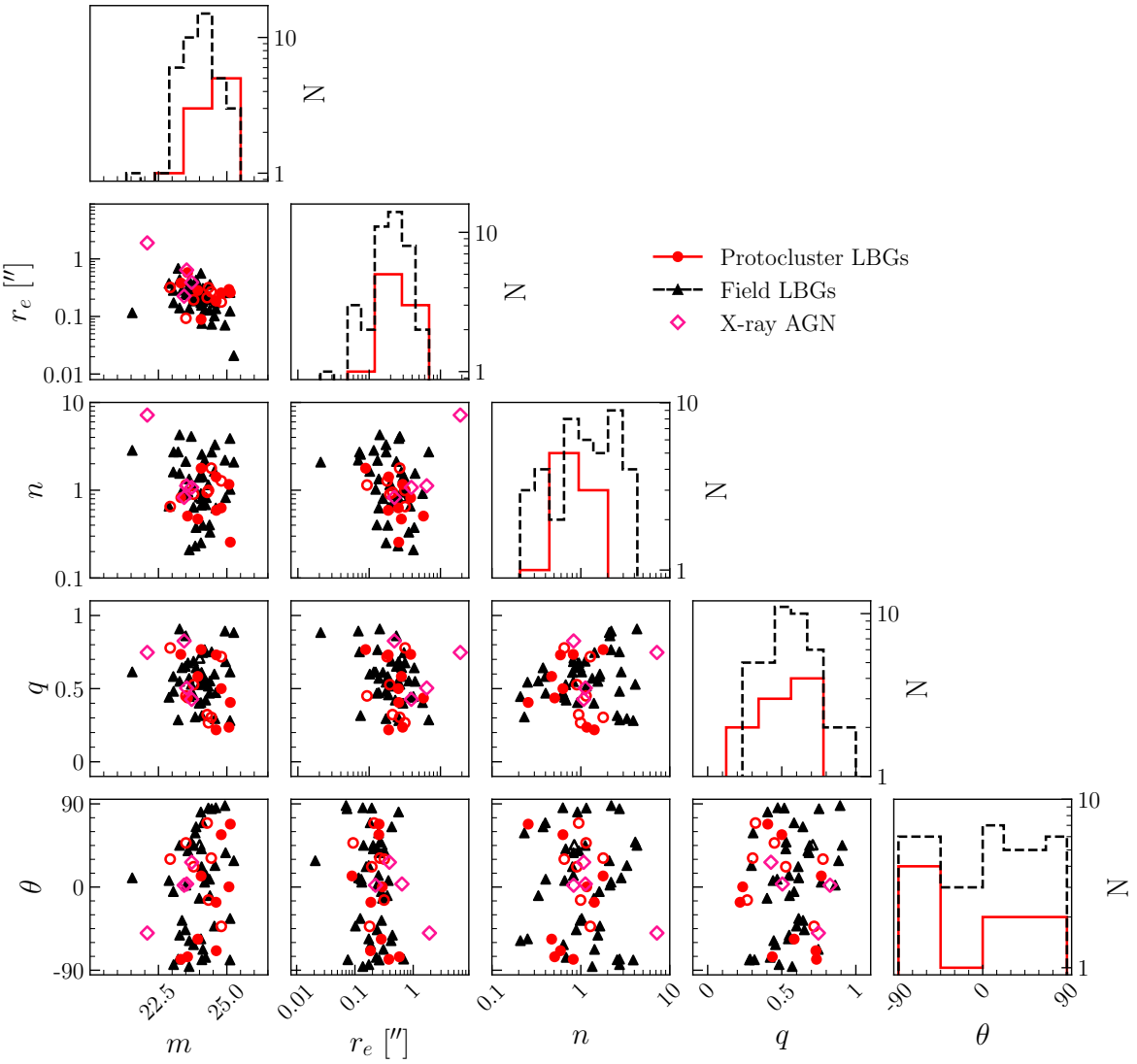

Figure 3.3: On the diagonal, we show the distributions of Sérsic fit parameters – integrated F160W magnitude m, effective radius r_e, Sérsic index n, axis ratio q and position angle θ – for the S03 LBGs with acceptable fits in the final SSA22 catalog and the comparison GOODS-N sample with fits from van der Wel et al. (2012). Off-diagonal, we show correlations between the parameters for S03 LBGs, M17 LBGs, and protocluster AGN with acceptable fits. We show LBGs associated with the protocluster ($3.06 \leq z \leq 3.12$) in red (circles, solid lines), and field LBGs in black (triangles, dashed lines). While the histograms on the diagonal are limited to only S03 LBGs, on the off-diagonal scatter plots we show both S03 LBGs (filled symbols) and M17 LBGs (open symbols). The X-ray detected protocluster AGN are shown on the scatter plots as open magenta diamonds.

Table 3.4: One- and two-dimensional KS test statistics and probabilities for derived properties and correlations of protocluster and field LBGs.

	S03 LBGs[a]				S03 & M17 LBGs[b]			
Protocluster/Field Comparison	$N_{\rm proto}$[c]	$N_{\rm field}$[d]	D_{KS}[e]	p_{KS}[f]	$N_{\rm proto}$	$N_{\rm field}$	D_{KS}	p_{KS}
m	9	41	0.43	0.09	16	42	0.28	0.28
$\log r_e$	9	41	0.30	0.41	16	42	0.26	0.37
$\log n$	9	41	0.36	0.23	16	42	0.31	0.17
q	9	41	0.25	0.66	16	42	0.27	0.31
G	13	13	0.54	0.04	24	15	0.32	0.25
M_{20}	13	13	0.38	0.30	24	15	0.28	0.41
C	13	13	0.38	0.30	24	15	0.23	0.61
$M_\star$	8	45	0.58	0.01	15	46	0.58	0.00
SFR	8	45	0.49	0.05	15	46	0.54	0.00
sSFR	8	45	0.24	0.75	15	46	0.22	0.57
Mass-weighted age	8	45	0.23	0.80	15	46	0.19	0.75
Joint Distribution Tests								
$m - \log r_e$	9	41	0.28	0.59	14	42	0.26	0.42
$m - \log n$	9	41	0.40	0.20	14	42	0.28	0.32
$m - q$	9	41	0.43	0.12	14	42	0.26	0.40
$\log r_e - \log n$	9	41	0.32	0.44	14	42	0.26	0.40
$\log r_e - q$	9	41	0.29	0.55	14	42	0.25	0.45
$\log n - q$	9	41	0.36	0.30	14	42	0.29	0.29
$G - M_{20}$	13	13	0.54	0.05	24	15	0.40	0.10
$C - M_{20}$	13	13	0.46	0.13	24	15	0.32	0.28
$G - C$	13	13	0.54	0.05	24	15	0.40	0.10

Notes – For tests on parametric morphological properties and physical properties we include galaxies from our comparison sample of GOODS-N field LBGs. For the physical properties we show only the results for the SED fits with $Z = 0.655 Z_\odot$; the results of the tests are not significantly different for the fits with $Z = Z_\odot$. We exclude the M17 LBG J221718.04+001735.5 from the KS tests on the SED-fitting derived parameters due to likely contamination of its near-IR photometry by a nearby point source.

[a]I.e., comparing protocluster and field LBGs from the S03 catalog only. [b]I.e., comparing protocluster and field LBGs from both the S03 and M17 catalogs. [c]Number of protocluster LBGs in comparison. [d]Number of field LBGs in comparison. [e]Two sample KS test statistic. [f]Two sample KS test p-value.

features related to recent mergers. Briefly, the Mantha et al. (2019) method identifies flux- and area-wise significant contiguous pixel regions in residual images. We set the pipeline to search for connected regions of pixels $\geq 2\sigma$ above the background after convolution with a boxcar filter 3 pixels wide. Based on the average galaxy size at $z \sim 3$, we searched for residual features within 15 kpc of the main galaxy. To ensure that under-subtracted regions at the center of a galaxy were not erroneously detected as possible tidal features, we also set the pipeline to mask out an ellipse centered on the source position, which has major and minor axes scaled up 2.5 times from the SExtractor detection ellipse. Additionally, since the goal of the pipeline is to extract low surface brightness features associated with mergers, bright features are masked in the image before extraction, in order to exclude bright companion galaxies from being identified as tidal features.

Three galaxies in our LBG sample are deliberately excluded from this analysis: J221732.04+001315.6, a protocluster LBG, which appears to be 2–3 sources blended within 5 kpc, and both J221717.69+001900.3 and J221717.68+001901.0 (SSA22a-M38), a pair of field LBGs which are also blended together within 5 kpc, though they were detected and fit separately by our GALAPAGOS-2 pipeline. We excluded these blended galaxies even though they may be merging systems that are physically associated. In such close associates, the blending makes it difficult to reliably fit and mask the images in a way that prevents the pipeline from extracting under-fit components of the blended system as tidal features.

To ensure the cleanest possible residuals, we re-fit the "original" models from the GALAPAGOS-2 pipeline for five galaxies with small adjustments to improve their positioning. For instance, the fitting cutout for J221718.04+001735.5 contains a bright, unrelated point source which was originally poorly fit with a Sérsic model and adversely affected the positioning of the other models in the image. We fit this cutout again, with the point source properly modeled by a PSF, and allow

the positions of the other models to vary. We also found that secondary galaxies in the cutouts containing J221720.20+001731.6 (SSA22a-C47) and J221719.30+001543.8 (SSA22a-C30) were under-subtracted due to offsets in the positions of their models; we re-fit them with the magnitude and shape parameters fixed, and the position parameters allowed to vary.

In the case of both J221720.25+001651.7 (SSA22a-C35) and J221704.34+002255.8, the main concentration of the galaxy appears offset from the center of the original GALAPAGOS-2 pipeline model, which is fit to both the concentrated component and an apparent fainter, diffuse component that extends asymmetrically to the southeast of the main concentration. To recenter the fit on the dominant, concentrated feature, we computed the centroid of the pixels $\geq 4\sigma$ above the background and re-fit the cutout with the primary model fixed to that position. The best fit model parameters of the primary galaxy do not change significantly in any of these cases, except for the position of the fits to SSA22a-C35 and J221704.34+002255.8, by construction.

We show the features extracted by the tidal feature pipeline in the residual panels of Figure 3.4 and Figure 3.5, along with their surface brightness in mag arcsec^{-2} and the unmasked area in which they were extracted.

The extracted features are all of low surface brightness. The range of 2σ limiting surface brightness in our two-orbit images is 25.0–25.4 mag arcsec^{-2}; in the single-orbit images the range is 24.6–25.1 mag arcsec^{-2}. In terms of surface brightness alone, none of the extracted features are unambiguous; only four galaxies have extracted features with surface brightness brighter than the 2σ limit: J221710.35+001920.8, J221718.87+001816.2 (SSA22a-D17), J221719.30+001543.8 (SSA22a-C30), and J221701.38+002031.9. While the surface brightness of the features in the other galaxies is on the order of the limiting surface brightness, the sizes of the features in most cutouts indicate that they are unlikely to be due to noise alone. In one case, J221731.69+001657.9

(SSA22a-M28), the area of the largest residual feature is small enough to be consistent with noise, and we thus exclude the extracted features in this galaxy cutout from consideration as plausible tidal features; in the remainder of cases we estimate a probability $p < 0.02$ that the largest feature is due to noise, based on simulations of the image background. In classifying residual features as plausible tidal features resulting from merger activity, we focus on three additional criteria: (1) asymmetry with respect to the primary galaxy, (2) extension, and (3) plausible physical association with the primary galaxy. As a rule-of-thumb, we consider features that reach within 5 kpc of the primary galaxy's center to plausibly be physically associated with the galaxy.

While the residual feature in the J221710.35+001920.8 cutout is extended, asymmetric, and plausibly associated with the main galaxy, we note that it is positioned near the expected location of a diffraction spike from the WFC3 PSF. Though the extracted feature is low surface brightness, the galaxy is very concentrated in appearance, and the Sérsic model fit is concentrated and visibly PSF-like in appearance. For these reasons, we do not consider the feature extracted from the J221710.35+001920.8 cutout to be plausibly tidal.

Two features are extracted in the SSA22a-D17 cutout, associated with clumpy features to the north and south of the main galaxy in the original image. The features are asymmetrical in size and shape, and both are within 5 to 10 kpc of the model barycenter. The residual features in the SSA22a-C30 cutout are also associated with clumpy structures which are apparent in the original image, to the southwest of the main galaxy. These clumpy features are also asymmetrical with respect to the main galaxy, and extend between 5 and 10 kpc away from the primary model's barycenter. Similar clumpy residual features have been observed by Mantha et al. (2019) in mock two-orbit F160W observations of merging galaxies in the VELA cosmological simulations (Ceverino et al., 2014; Zolotov et al., 2015). The simulated observations suggest that similar features

may be associated with the late stages of a major merger (i.e. 0.15–0.80 Gyr after the galaxies coalesce), when multiple nuclei are no longer apparent (see Figure 9 in Mantha et al., 2019). The feature extracted in the J221701.38+002031.9 cutout is brighter than the limiting surface brightness in the image, and appears to be asymmetrical and plausibly physically associated with the galaxy. Its shape and offset from the main concentration of the galaxy suggest that it may be tidal in origin, though this galaxy is in a single-orbit image. We are thus less confident in assessing this as a plausible tidal feature.

The apparent association of the small residual feature in the SSA22a-C47 cutout with both galaxies may suggest interaction; residual features bridging the two galaxies seem to be common in mergers (see, e.g., Figures 7 and 9 in Mantha et al., 2019). If these two galaxies are at the same redshift, they might then be a pre-coalescence merging pair, based on the residual feature and the apparent bridge between the galaxies in the original image. However, we have not found a spectroscopic redshift in the literature for the projected companion galaxy, nor do we have independent photometric redshifts for both galaxies, so we are unable to establish whether the apparent companion is physically close to the primary galaxy.

The irregularly shaped galaxy SSA22a-C35 has a diffuse feature offset to the southeast from the main concentration of the galaxy, with a surface brightness of ≈ 25 mag arcsec^{-2}. This feature, based on its shape and plausible physical association with the galaxy, may also be associated with the late, post-coalescence stages of a merger. However, the data we used to fit this galaxy was taken at single-orbit depth, and the quality of the fit is poor. We are thus less confident in assessing this as a plausible tidal feature.

Based on the above, we find 2–5 residual features that may plausibly be tidal in our proto-cluster LBGs. The clumpy features associated with SSA22a-D17 and SSA22a-C30 are the most

plausible, and the diffuse features associated with SSA22a-C35 and J221701.38+002031.9 are the least plausible. If we extrapolate this to a merger fraction based on the number of S03 LBGs we applied the pipeline to, we find a protocluster LBG merger fraction of 0.22–0.44, comparable to the merger fraction we derive by naïve visual classification of S03 LBGs in Section 3.3.4. If we include M17 LBGs we find a protocluster merger fraction of 0.13–0.33.

Of the 7 field LBGs we have applied the pipeline to, J221735.08+001422.7 (SSA22a-MD20), J221726.65+001638.4 (SSA22a-MD37), and J221724.44+001714.4 (SSA22a-C41) have offset features apparent in the original image, though the feature near SSA22a-MD20 is bright enough to be masked by the residual extraction pipeline. For consistency with the above we consider only the features near SSA22a-MD37 and SSA22a-C41. These features are both consistent with the 2σ limiting surface brightness in their respective images. Both features are asymmetric and plausibly physically associated with the galaxy, though the feature in the SSA22a-C41 cutout is smaller and closer to the main concentration of the galaxy. We note the superficial similarities of the extracted features in the SSA22a-MD37 cutout to the clumpy features associated with SSA22a-D17 and SSA22a-C30, and conclude that they may be tidal features. The feature in the SSA22a-C41 cutout is more ambiguous; in the original image it appears that the galaxy is asymmetric, with the eastern side of the galaxy being fainter and more diffuse. The extracted feature is apparently associated with this diffuse region of the main galaxy, similar to SSA22a-C35 above, but smaller in size. We thus find this to be a low-confidence tidal feature, as with SSA22a-C35. Based on the 7 field LBGs we applied the pipeline to, we find a field merger fraction of 0.14–0.28, comparable to the protocluster merger fraction we found above, and the merger fractions derived by visual classification in Section 3.3.4.

We note that conclusive or completely quantitative identification of residual features as being

due to mergers is beyond the scope of this work. The relationship between the observed strength and shape of tidal features and the different stages of a merger is not yet fully explored, and will require comprehensive simulations to establish. Thus, rather than attempting to conclusively identify mergers with this technique, we have classified features only as "plausibly" tidal above, and we treat this method as a supplement to the more established methods of morphological analysis we use in the other sections of this paper.

To roughly estimate the mass of the plausible tidal features above, we used the protocluster and field SED models described in Section 3.4 to calculate mass-to-light ratios in the F160W band. We find $M/L = 0.45$ $M_\odot/L_\odot$ for the protocluster model and $M/L = 0.22$ $M_\odot/L_\odot$ for the field model, both assuming $z = 3.1$. These are in agreement with rest-frame $B-$band mass-to-light ratios calculated from the models using the $B-R$ color relationship in Zibetti et al. (2009, see their Appendix B). Using the F160W mass-to-light ratios we derived, we find that the plausible tidal features associated with the protocluster galaxies have masses on the order of 10^9 $M_\odot$, ranging from $\log_{10} M_\star/M_\odot = 9.10$ to 9.78, suggesting that the largest of the clumps have masses comparable to the Small Magellanic Cloud. Assuming a typical stellar mass of 10^{10} $M_\odot$ for protocluster LBGs (see Figure 3.8) we find feature mass to total stellar mass ratios ranging from 0.13–0.60. These feature mass ratios, along with the compact or clumpy nature of some of the residuals (e.g., SSA22a-C30), suggest that this technique may be sensitive to minor mergers (mass ratio < 0.25), and that some of the features we extract may be infalling satellite galaxies.

3.3.3 Non - Parametric Morphological Analysis

To mitigate the limits imposed by requiring well-converged Sérsic fits to our data, we also pursued non-parametric morphological analyses. We applied this analysis to the $z \sim 3$ SSA22 S03

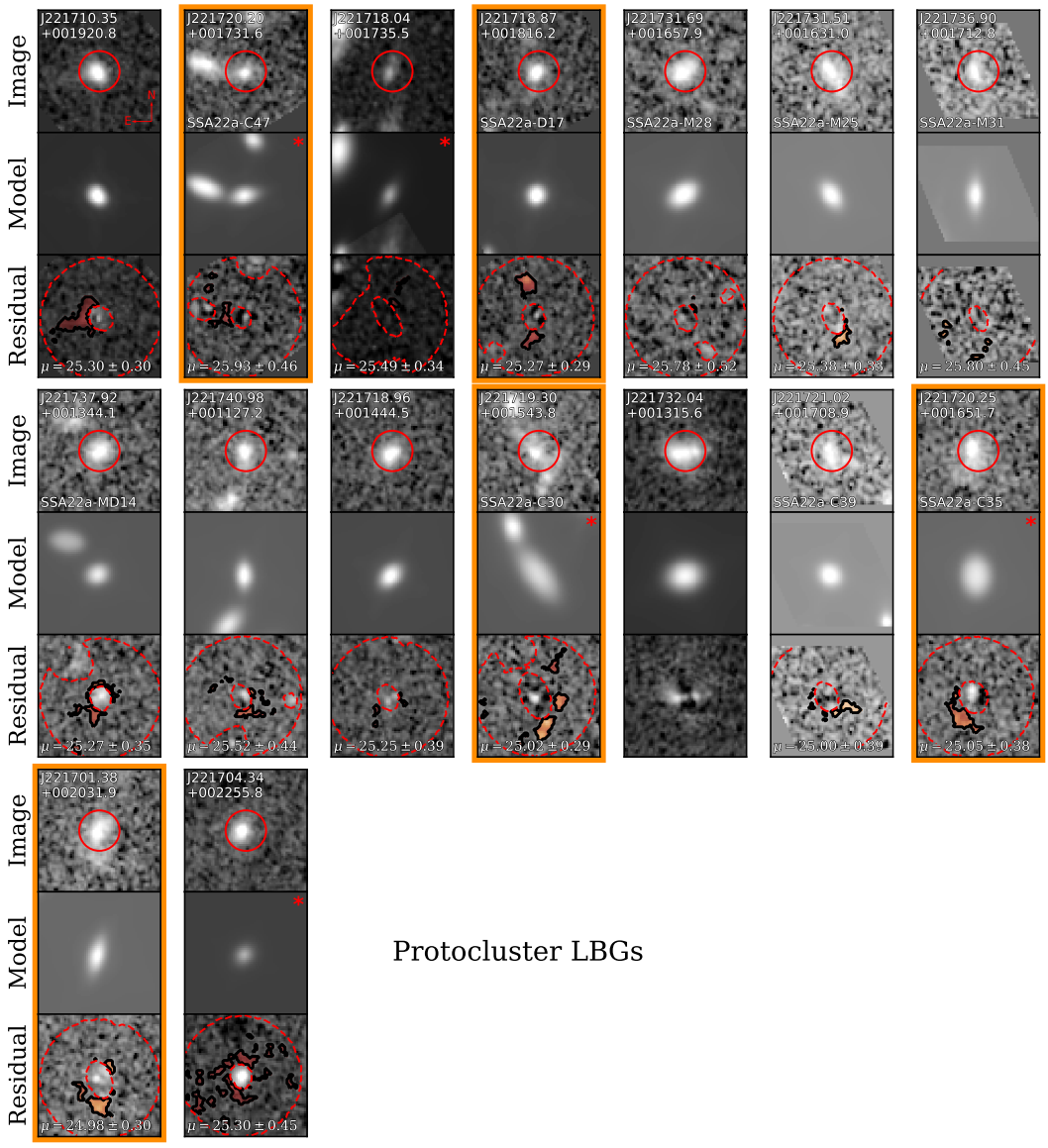

Figure 3.4: SSA22 protocluster LBGs in our sample with acceptable single Sérsic fits. We display the original F160W image, model, and residual at the same normalization and stretch for each galaxy. In four cases we have re-fit the model with adjustments to the position of the models to allow for cleaner tidal feature extraction; we mark these cases with a red star in the model panel. The red circle shows a 10 proper kpc diameter aperture at the position of the primary source. In the top-left corner of each F160W cutout, we print the galaxy's J2000 positional identifier; for LBGs in the S03 catalog we print their catalog designation in the lower-left corner. In the residual panel, we highlight features extracted by the tidal feature pipeline in an orange colormap and print the surface brightness of the extracted features in mag arcsec^{-2}. The regions inside the inner dashed red line and outside the outer dashed red line were excluded from feature extraction. Image blocks for galaxies where we have found a plausible tidal feature are outlined with a thick orange border.

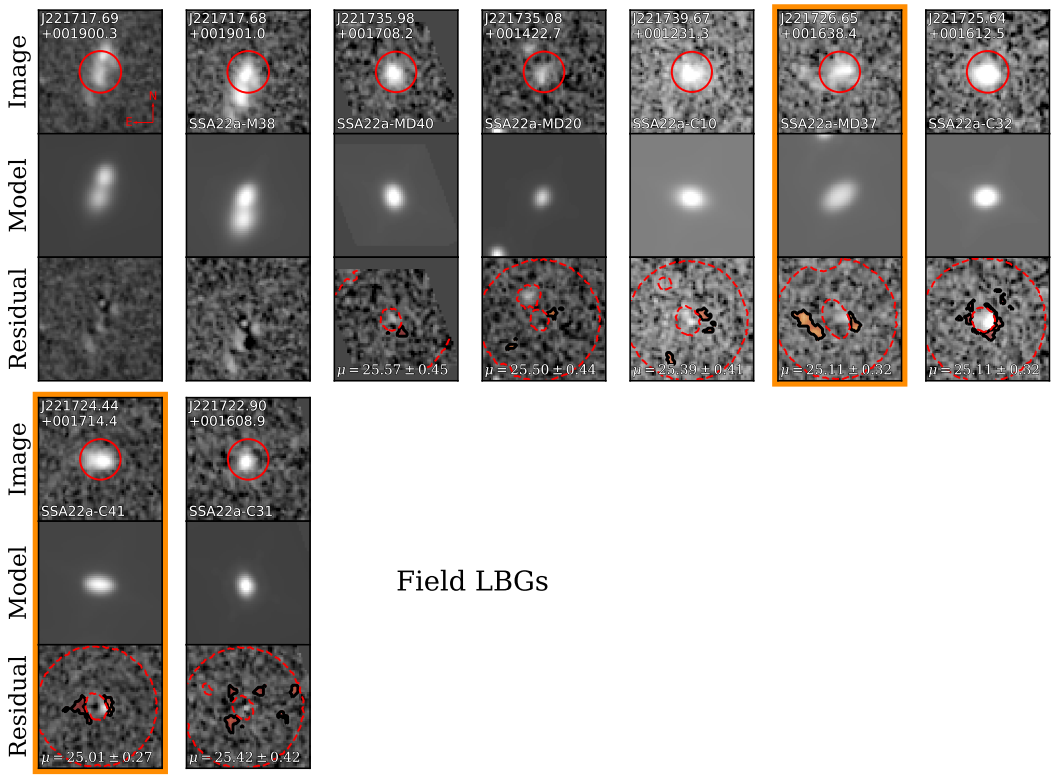

Figure 3.5: SSA22 field LBGs in our sample with acceptable single Sérsic fits. All annotations are the same as Figure 3.4.

and M17 LBGs in our sample, excluding LBGs from the GOODS-N comparison sample. The Gini coefficient G of a galaxy's flux (Abraham et al., 2003) and the second order moment of light statistic M_{20} for the brightest 20% of light from a galaxy (Lotz et al., 2004) can be used in concert to identify merger candidates. We used the definition of G from Lotz et al. (2004):

$$G = \frac{1}{\overline{|f|} n\,(n-1)} \sum_{i}^{n} (2i - n - 1)\,|f_i|, \tag{3.3}$$

where $\{|f_i|\}$ contains the absolute values of the source flux, sorted from smallest to largest, and $\overline{|f|}$ denotes the mean of these values. The Gini coefficient describes the equality of the distribution of light in a galaxy, with values close to 0 indicating an egalitarian distribution of flux and values approaching 1 indicating an unequal distribution.

The total second order moment of light is defined as

$$M_{tot} = \sum_{i}^{n} f_i r_i^2 = \sum_{i}^{n} f_i \left[(x_i - x_c)^2 + (y_i - y_c)^2 \right], \tag{3.4}$$

where r_i is the distance from the pixel containing flux f_i to the center (x_c, y_c) of the galaxy; in this context, the center is defined as the location that minimizes M_{tot}. M_{20} is then computed as

$$M_{20} = \log_{10} \left(\frac{\sum_{j}^{m} f_j r_j^2}{M_{tot}} \right), \tag{3.5}$$

where m is the largest index satisfying

$$\sum_{j}^{m} f_j \leq 0.2 f_{tot} \tag{3.6}$$

when $\{f_j\}$ contains the flux of the pixels sorted from largest to smallest. Values of M_{20} close to 0 indicate excesses of flux further from the galactic centers, which may indicate star forming knots, multiple nuclei, or otherwise disturbed morphologies. Smaller (more negative) values indicate

concentration of light at the center of the galaxy; local quiescent elliptical galaxies have average $M_{20} \sim -2.5$, for example (Lotz et al., 2004).

Locally, the difference between mergers and "normal" galaxies is well-established in the Gini-M_{20} plane; disturbed morphologies create an unequal, off-center distribution of light, so mergers tend to fall above the Gini-M_{20} trend – canonically, the line $G = -0.14M_{20} + 0.33$ (Lotz et al., 2008) – with larger values of M_{20}, while normal galaxies fall below. The Gini coefficient is expected to remain relatively unbiased at $z \gtrsim 2$ given high signal-to-noise and resolution better than 500 proper pc pixel^{-1}; M_{20} may be biased by the flattening of the angular size of features at high redshift, though its large dynamic range may still prove useful in distinguishing between disturbed and undisturbed morphologies (Lotz et al., 2004). The physical resolution of our images is 496 proper pc pixel^{-1} at $z = 3.09$, approaching the limit of what Lotz et al. (2004) recommend. In Section 3.A, we used simulations of galaxies similar to our LBG sample to investigate how G, M_{20}, and C are biased by decreasing signal-to-noise, finding that they are relatively stable over the range of S/N in our sample.

As a non-parametric analog to the Sérsic index, we also calculated the concentration parameter C, defined as

$$C = 5\log\left(\frac{r_{80}}{r_{20}}\right), \tag{3.7}$$

where r_{80} and r_{20} are the radii enclosing 80% and 20% of the total flux, respectively (Conselice et al., 2003). Here, the total flux is defined as the flux contained within 1.5 Petrosian radii. The value of C tends to increase for more concentrated, bulge-dominated morphologies (i.e., with increasing n).

We adopted a similar method as Lotz et al. (2004) to compute G, M_{20}, and C. We defined a new

segmentation map for each detection by first computing its elliptical Petrosian radius r_p. We then smoothed the image of the galaxy with a Gaussian with $\sigma = r_p/5$, measured the mean flux $\mu(r_p)$ of the pixels at r_p, and assigned any pixels with flux greater than $\mu(r_p)$ to the new segmentation map. Finally, we computed G and M_{20} for each detection using all the pixels included in this new segmentation map. It is possible for this process to fail if, for example, the Petrosian radius cannot be computed or the new segmentation map contains multiple disjoint features. In these cases (6 of the 72 galaxies with protocluster redshifts, all of which are LAEs) we do not report G, M_{20}, or C in the catalog. However, we were able to recover all three quantities for all of the protocluster and field LBGs in our F160W images, as well as the 8 protocluster AGN in our images.

We show the derived values of G, M_{20}, and C for protocluster and SSA22 field LBGs in Figure 3.6. Both S03 protocluster and SSA22 field LBGs cluster together with median (G, M_{20}, C) of $(0.37, -1.26, 2.15)$ and $(0.44, -1.52, 2.25)$, respectively. While most of the X-ray AGN fall in the same area of $G - M_{20}$ space as our LBG samples, they tend to be more concentrated than normal galaxies, falling mainly to the upper right of the trend in $G - C$ space, as expected in cases where the AGN dominates the rest-frame optical emission from the galaxy.

In Figure 3.6, we adopt the $G - M_{20}$ classifications from Lotz et al. (2008), which show that galaxies with $G > -0.14M_{20} + 0.33$ are merger-like and galaxies with $G \leq -0.14M_{20} + 0.33$ and $G > 0.14M_{20} + 0.80$ are bulge-dominated (i.e, Hubble classes E, S0, and Sa). The majority of our galaxies occupy the third region defined by these two lines, where galaxies with irregular and disk-dominated morphologies (i.e Hubble classes Sb, Sc, and Ir) fall at low redshift. However, there is significant concern about the use of cuts in the $G - M_{20}$ plane to classify mergers at high redshift. Artificial redshifting of simulated merging systems suggests that the typical $G - M_{20}$ criteria may miss a significant number of mergers at high redshift, and that any apparent trend

toward the merger-like region of the $G-M_{20}$ plane may only be a function of mass (Abruzzo et al., 2018). Snyder et al. (2015) also find that the joint distribution of G and M_{20} is narrow at $z\sim 3$, even for diverse morphological types, due to the $G-M_{20}$ segmentation algorithm excluding the low surface brightness outer regions of the galaxy. As such mergers without clear-cut cases of multiple nuclei may not separate from normal galaxies in the $G-M_{20}$ plane.

It is also apparent from our analysis that the $G-M_{20}$ criteria might miss mergers: the systems in which we see plausible tidal features in Section 3.3.2 do not all fall above the merger-like/disk-like dividing line in Figure 3.6, nor do the galaxies visually classified as mergers in Section 3.3.4. Only two galaxies fall in the merger region: J221740.98+001127.2, a M17 protocluster LBG, and J221726.65+001638.4 (SSA22a-MD37) a S03 field LBG. Neither of these galaxies has obvious multi-nuclear structure; rather, their elevated G and M_{20} values appear to be due to diffuse, asymmetric features that extend away from the main concentration of the galaxy. As we have noted in Section 3.3.2, such features could be tidal structures associated with the late stages of a merger. In Section 3.3.2 we identified clumpy structures associated with SSA22a-MD37, though we found no significant residual features associated with J221740.98+001127.2. Given this inconsistency, and the fact that other galaxies in which we have identified plausible tidal features do not have G and M_{20} values consistent with the Lotz et al. (2008) merger classification, we hesitate to draw any conclusions about the SSA22 protocluster merger fraction from the $G-M_{20}$ classification.

Protocluster LBGs also do not appear to be more morphologically evolved, or bulge-like, by the Lotz et al. (2008) criteria, though given that the merger criterion evidently misses mergers we cannot conclusively apply the bulge criterion here. Additionally, the protocluster and field LBGs cluster along the same locus in the $G-C$ plane, with neither set of galaxies conclusively being more concentrated or bulge-dominated than the other.

Beyond using G and M_{20} to attempt to classify mergers, we also attempted to use them to distinguish between the protocluster and field populations. We show 1-D and 2-D KS test results for the individual and joint distributions of non-parametric morphologies in Table 3.4. We again limited these tests to S03 LBGs, but here we are restricted to the protocluster and the SSA22 field. The KS test results for the S03 LBGs alone suggest that the the protocluster LBGs are flatter than their field counterparts: the test for G admits the rejection of the null hypothesis, and the tests on the $G-M_{20}$ and $G-C$ joint distributions also suggest differences in the distributions of light of protocluster and field galaxies. However, the protocluster and field galaxies do not cleanly separate in any of the projections of G–M_{20}–C space. Additionally, if we include the M17 LBGs, all of the KS tests on the nonparametric morphological measures are consistent with the null hypothesis. Thus we are unable to draw any conclusions about morphological differences between the protocluster and field LBGs.

3.3.4 Visual Classification

For a direct comparison of our rest-frame optical morphologies with Hine et al. (2015), we also pursued a similar visual classification scheme for protocluster candidate LBGs using F160W images. Cutouts of SSA22 LBGs were mixed with cutouts of LBGs from the GOODS-N field and blindly distributed to seven voters, who placed each galaxy in one of six categories, as defined in Hine et al. (2015):

- **C1**: One clearly distinct, compact nucleus
- **C2**: Single nucleus, but less compact or with minor asymmetry

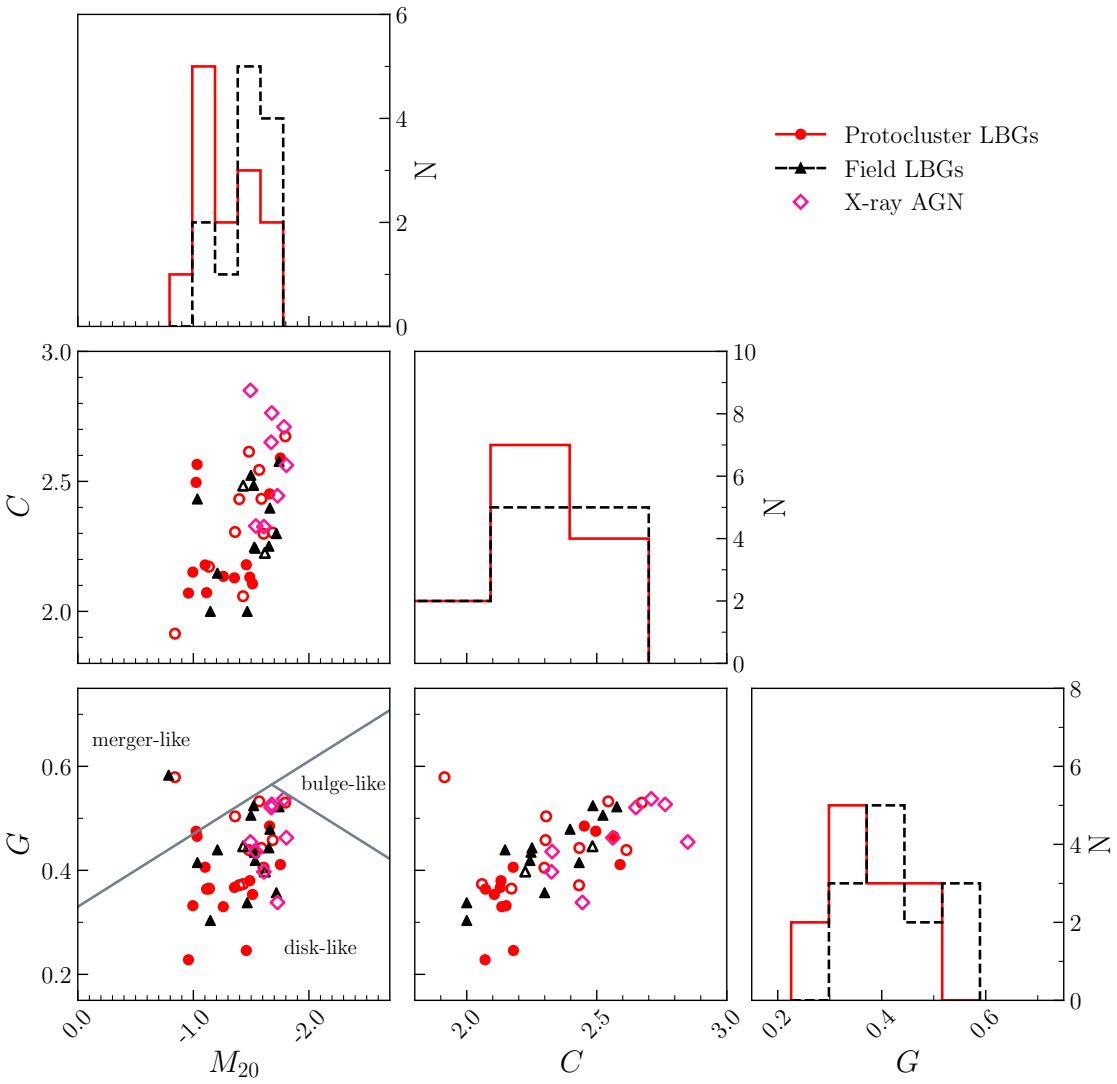

Figure 3.6: We plot the projections in $G - M_{20} - C$ space of the 13 S03 protocluster LBGs (filled circles), 11 M17 (open circles) protocluster LBGs, 13 S03 field LBGs (filled triangles), 2 M17 field LBG (open triangles), and 8 X-ray detected protocluster AGN (open diamonds), where G is the Gini coefficient, M_{20} is the second-order moment of light, and C is the concentration statistic. For reference we also plot the $G - M_{20}$ classification regions of Lotz et al. (2008). While we show both S03 and M17 LBGs in the scatter plots, we restrict the histograms to the S03 LBG sample.

- **M1**: Clear evidence of a second nucleus; all flux falls in a $1''$ diameter aperture centered at the source position
- **M2**: Clear evidence of a second nucleus; some flux falls outside the $1''$ aperture
- **M3**: More than two nuclei or more complex clumpy structure; some flux falls outside the $1''$ aperture
- **M4**: More than two nuclei or more complex clumpy structure; all flux falls inside the $1''$ aperture

Voters were asked to assign a confidence level from 1 (low) to 5 (high) to their classification, which we use to weight the votes. To artificially enforce consensus on our classifications, we summed the confidence scores of the voters for each galaxy and set a confidence threshold of 65%, classifying galaxies with $\geq 65\%$ of the total confidence in categories M1, M2, M3, and M4 as mergers, while galaxies with a $\geq 65\%$ of the total confidence in categories C1 and C2 were classified as isolated. This threshold was chosen to exclude cases where voters were split 4-to-3 between merger and isolated categories. This method excludes 2 protocluster LBGs and 1 X-ray AGN for which the confidence threshold is not met (i.e., the voters were not confident in classifying the galaxy as a merger or isolated). We classify these galaxies as "ambiguous." A plurality of votes for these ambiguous galaxies place them in category C2, which allows for diffuse morphology and asymmetry about a single nucleus. This is consistent with the voting scheme in Hine et al. (2015), where many ambiguous galaxies were classified as C2. We also used the confidence scores to assign each galaxy a final category from the above list by weighting each vote by the voter's assigned confidence. We show the results of this analysis in Table 3.5 and Figure 3.7.

Table 3.5: Number of galaxies in each category and calculated merger fraction for each of our visual classification samples. The categories $M1-4$ and $C1-2$ are defined in Section 3.3.4. The uncertainty on the merger fraction is calculated using Poisson statistics.

	Number in Category								
Sample	M1	M2	M3	M4	C1	C2	Mergers	Isolated	Merger Fraction
Protocluster	0	2	1	0	1	4	3	5	$0.38^{+0.37}_{-0.20}$
X-ray AGN	1	2	0	0	2	1	3	3	$0.50^{+0.49}_{-0.27}$
Total Field	9	6	4	1	17	12	20	29	$0.41^{+0.11}_{-0.09}$
GOODS-N Field	8	6	4	1	11	8	19	19	$0.50^{+0.14}_{-0.11}$
SSA22 Field	1	0	0	0	6	4	1	10	$0.09^{+0.21}_{-0.08}$

The outcome of our visual classification of LBGs supports the results of the parametric and non-parametric morphological analysis, finding no significant difference in the observed fraction of galaxies in mergers between the SSA22 protocluster and the combined $z \sim 3$ field. Specifically, we find merger fractions of $0.38^{+0.37}_{-0.20}$ for protocluster LBGs, $0.50^{+0.49}_{-0.27}$ for the X-ray selected protocluster AGN, and $0.41^{+0.11}_{-0.09}$ for the combined SSA22 and GOODS-N field samples. We note also that if we consider the case where all of the ambiguous cases mentioned above were, in reality, mergers, we would have a protocluster LBG merger fraction of $0.50^{+0.34}_{-0.22}$, which remains consistent with the GOODS-N and combined field merger fractions. However, our reported merger fraction of $0.09^{+0.21}_{-0.08}$ for the SSA22 field is based on one identified merger and is considerably lower than the merger fraction of $0.50^{+0.14}_{-0.11}$ for the GOODS-N field. Additionally, Hine et al. (2015) report a merger fraction of 0.33 ± 0.18 for the SSA22 field based on F814W ACS imaging data. We note that we use a different redshift range to define $z \sim 3$ field galaxies: Hine et al. (2015) used galaxies from $2.5 \leq z \leq 3.5$ in their field samples, while we include only galaxies with $2.9 \leq z \leq 3.3$. However, this does not appear to be a significant driver of the very low merger fraction we observe in the SSA22 field: if we widen our criteria to include galaxies with $2.5 \leq z \leq 3.5$, the field merger fractions remain consistent with the values reported above.

Regardless of the redshift range we adopt, if we assume that the true merger fraction in the SSA22 field is equal to the GOODS-N merger fraction, we find a Poisson probability of $p < 0.03$ that we would observe one or fewer mergers among the SSA22 field LBGs classified here by chance alone. If we compare the SSA22 protocluster and field directly by assuming that the protocluster merger fraction is correct, we find that the elevation seen in Table 3.5 and Figure 3.7 is apparently marginal: there is a Poisson probability $p < 0.09$ of observing a merger fraction less than or equal to the SSA22 field merger fraction, regardless of how we define the field.

Varying WFC3 IR imaging depth across the SSA22 field of view appears to play a role in the classification of galaxies as merging or isolated, and may be the primary driver of the protocluster-over-field elevation of the merger fraction that we observe in SSA22. If we divide the galaxies by depth, we find two-orbit merger fractions $0.67^{+0.89}_{-0.43}$ (2/3) for protocluster LBGs and $0.20^{+0.73}_{-0.06}$ (1/5) for SSA22 field LBGs, and one-orbit merger fractions $0.20^{+0.73}_{-0.06}$ (1/5) and $0.00^{+0.31}_{-0.00}$ (0/6) for the protocluster and SSA22 field LBGs, respectively. Again assuming that the merger fraction calculated from the GOODS-N LBG sample represents the true field galaxy merger fraction at $z \sim 3$, we calculate the Poisson probability of finding a merger fraction less than or equal to the observed SSA22 field merger fraction in the two-orbit (one-orbit) images to be 0.29 (0.05). If we evaluate the significance of the protocluster-over-field elevation after separating the galaxies by depth, we find that the elevation is no longer significant: if we assume the merger fraction from the two-orbit (one-orbit) protocluster galaxies, we find a Poisson probability of 0.16 (0.30) of observing a merger fraction less than or equal to the SSA22 field merger fraction at the same depth.

We take this to mean that our classification misses some mergers at single-orbit depth, consequently underestimating the SSA22 field merger fraction, since the sample of SSA22 field galaxies we classified coincidentally contains a larger proportion of galaxies observed at single-orbit depth. Our merger classifications require that the irregular morphological features induced by a merger be discernible by eye; it is possible that the lower signal-to-noise in the single-orbit images obscures the low surface brightness features associated with the late stages of a merger. Given these issues, the small number of sources in our samples, and the small number of sources observed at two-orbit depth, our constraints on any enhancements on merger fractions in the protocluster compared to the field are weak and not statistically significant at present.

The wavelength range of our observations is also expected to affect how many galaxies are classified as mergers. Hine et al. (2015) used rest-frame UV ACS F814W observations to classify their galaxies. Over their adopted redshift ranges, this probes wavelengths (based on the F814W reference wavelength) $\sim 1960-1990$ Å in the protocluster and $\sim 1790-2310$ Å throughout the range $2.5 \leq z \leq 3.5$. High angular resolution and the ability to trace star formation have made rest-frame UV observations a typical choice for merger classification (e.g. Lotz et al., 2004, 2006), though we note that the patchiness of UV observations due to individual star forming clumps may make otherwise "ordinary" star-forming galaxies look irregular. In our case, our F160W observations probe (again based on the filter reference wavelength) $\sim 3740-3790$ Å in the protocluster and $\sim 3580-3950$ Å throughout the range $2.9 \leq z \leq 3.3$. Over this range of redshift the 4000 Å break moves through the F160W bandpass. By construction and the cosmological constraints of observing at $z \sim 3$, the galaxies in our LBG sample are dominated by young stellar populations and consequently have weak 4000 Å breaks (see, e.g. the model SEDs in Figure 3.9). Since there is significant continuum emission on either side of the break in our galaxies, we do not expect that observing across the 4000 Å break should impact our morphological classifications in any significant way.

3.4 SED Fitting and Physical Property Analysis

To quantify the stellar mass distribution of our sample, we fit the spectral energy distributions (SED) of LBGs in the redshift range $2.9 \leq z \leq 3.3$. To avoid biasing our measurements of stellar masses and star formation rates, we exclude galaxies known from the literature to be hosting AGN and galaxies with rest-frame 8-32 keV luminosities $\geq 10^{43.5}$ erg s^{-1} (Lehmer et al., 2009b). We

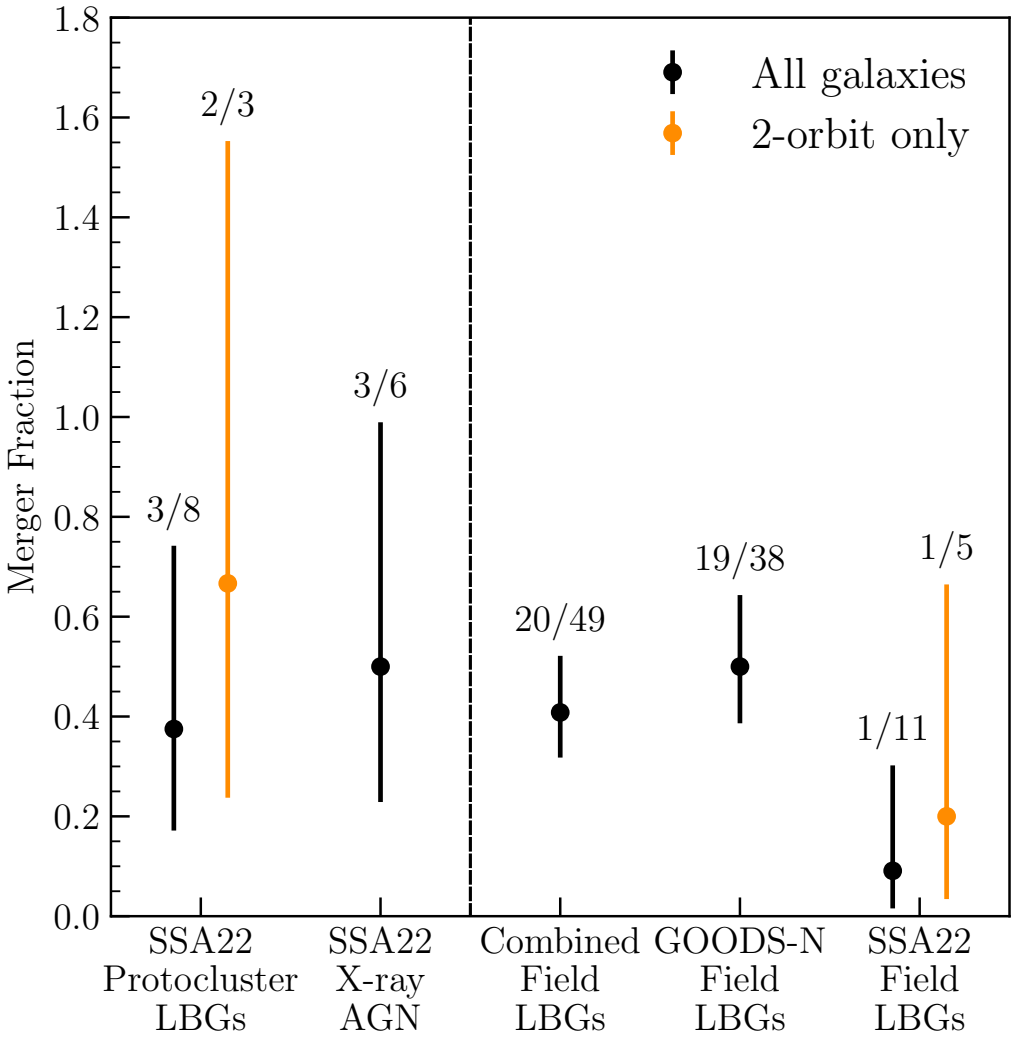

Figure 3.7: We show the fraction of S03 LBGs in classes M1 – 4 for the protocluster, combined field, GOODS-N field, and SSA22 field, along with the fraction of X-ray selected protocluster AGN in merger classes. In orange, we show the merger fractions for the protocluster and SSA22 field if we consider only galaxies in 2-orbit images. The uncertainty on the merger fraction is derived from Poisson statistics. The vertical dashed line separates protocluster and field merger fractions.

require the SSA22 galaxies we fit to have optical-to-NIR photometry available from Kubo et al. (2013), and we extract F160W fluxes from our own images. For consistency with the Kubo et al. (2013) photometry, we deconvolve our F160W images from the PSF described in Section 3.2.1, smooth the result to a Gaussian PSF with a $1''$ FWHM, and extract photometry from a $2''$ diameter circular aperture. We use all of the filters listed in the top half of Table 3.6, where available. Some of our SSA22 galaxies are not detected in the u^* band due to their strong Lyman breaks, and some galaxies do not have photometry available in all 4 IRAC bands. We exclude missing and non-detected bands from the fits.

For the GOODS-N sample, we used the photometric uncertainties derived by (Doore et al., 2021, accepted in ApJ). These uncertainties were recalibrated to include systematic uncertainties beyond single-instrument calibration, including the use of different photometric methods and systems in the observations, uncertainty and variation in Galactic extinction, blending of sources, and systematic effects created by the assumptions of our SED model.

We performed SED fitting using LIGHTNING (Eufrasio et al., 2017), which fits non-parametric star-formation histories (SFH) in discrete, variable or fixed-width stellar age bins. We made use of the most recent update to LIGHTNING, which uses an adaptive Markov Chain Monte Carlo (MCMC) algorithm (Doore et al., 2021). We chose SFH bins of $0-10$ Myr, $10-100$ Myr, 100 Myr $-$ 1 Gyr, and $1-2$ Gyr, where the upper age limit of the final bin is allowed to vary based on the age of the Universe at the redshift of the galaxy being fit. We assumed a Kroupa IMF (Kroupa, 2001) and fit using two metallicities: $Z_\odot$ and $0.655Z_\odot$, corresponding to the average metallicity of the Universe at $z = 3.1$ as given by the best-fit model in Madau & Fragos (2017). To generate the stellar population models, we used PÉGASE (Fioc & Rocca-Volmerange, 1997, 1999), running it once for each metallicity. For intrinsic attenuation, we adopted a Calzetti et al. (2000)

Table 3.6: SED fitting filters for each sample.

Sample	Observatory/Instrument	Filter(s)
SSA22[a]	SUBARU/SUPRIMECAM	u^* B V R i' z'
	SUBARU/MOIRCS	J K_s
	HST/WFC3	F160W
	Spitzer/IRAC	3.6 μm 4.5 μm 5.8 μm 8.0 μm
GOODS-N[b]	KPNO/4m/MOSAIC	U
	LBT/LBC	U
	HST/ACS/WFC	F435W F606W F775W F814W F850LP
	HST/WFC3	F105W F125W F140W F160W
	CFHT/WIRCam	K_s
	SUBARU/MOIRCS	K_s
	Spitzer/IRAC	3.6 μm 4.5 μm 5.8 μm 8.0 μm
	Spitzer/MIPS	24 μm 70 μm

Notes – [a]With the exception of F160W, the SSA22 photometry was measured by Kubo et al. (2013). F160W photometry was extracted from our images using the same procedures as Kubo et al. (2013). [b]See Barro et al. (2019) for description of the procedures used to extract the GOODS-N photometry.

extinction law, modified as in Noll et al. (2009) to include a UV bump at 2175 Å and a parameter δ to control the slope of the attenuation curve. We further modified the attenuation curve by including a birth cloud component, which is applied to the emission from the stars in only the youngest age bin. Our SED model then has a total of 7 free parameters: 4 SFH coefficients and 3 attenuation parameters. For a more thorough description of the stellar emission and attenuation prescriptions available in LIGHTNING, we refer the reader to Eufrasio et al. (2017) and Doore et al. (2021). To account for Galactic reddening, we used the standard Fitzpatrick (1999) curve. The Galactic A_V varies with the position of each galaxy, based on the Galactic dust extinction estimates of Schlafly & Finkbeiner (2011), which we retrieved using the IRSA `DUST` web application[4].

We find the quality of our fits to both SSA22 LBGs and GOODS-N LBGs acceptable based on the distributions of χ^2 for each sample. For the fits with $Z = 0.655Z_{\odot}$ the median and 16th to 84th percentile range of the χ^2_{min} distribution is $5.79^{+3.69}_{-2.98}$ for the SSA22 LBGs we discuss below, with a median of 6 degrees of freedom, and $11.03^{+9.45}_{-4.58}$ for the GOODS-N LBGs, from a median of 10 degrees of freedom (note that the number of degrees of freedom is larger for the GOODS-N galaxies due to the larger number of available bands; see Table 3.6). This corresponds to probabilities $p_{null} = 0.40^{+0.39}_{-0.29}$ for the SSA22 LBGs and $p_{null} = 0.24^{+0.47}_{-0.20}$ for the GOODS-N LBGs; here we define p_{null} as the probability of accepting the hypothesis that the data are generated by the model. The majority of fits are thus not ruled out by a χ^2 test. The quality of the fits does not change significantly for the fits with $Z = Z_{\odot}$.

We show example SED and SFH fit results in Section 3.B. The SFH parameters are sampled from the last 1000 steps of the MCMC chains. We use the sampled SFH to calculate the stellar mass of each galaxy, and compute the recent star formation rate (SFR) over the last 100 Myr as

[4]https://irsa.ipac.caltech.edu/applications/DUST/

the age-bin weighted average of the most recent two bins of the SFH. The mass-weighted age is computed by weighting the average age of the stars in each bin by the mass in the bin.

We show the distributions of stellar mass, SFR, and specific star formation rate (sSFR) in Figure 3.8. The SSA22 protocluster LBGs largely appear to follow the same star-forming main sequence as the field LBGs, though they populate the upper end. The typical star formation rates of our protocluster LBGs, 20–200 $M_\odot$ yr^{-1}, are significantly smaller than the IR-dervied SFRs for the DSFGs in the core region of the protocluster, which typically range from $\sim 10^2$–10^3 $M_\odot$ yr^{-1} (Umehata et al., 2015; Kato et al., 2016); that is, we are not probing the most intensely star-forming population of the protocluster. Three of the eight galaxies hosting X-ray detected AGN, which we excluded from our SED fitting, have ALMA derived SFR $\approx$ 220–410 $M_\odot$ yr^{-1}, and the remaining five, which are not ALMA-detected, have upper limits $<$ 130–210 $M_\odot$ yr^{-1}. These upper limits are consistent with the LBGs in our sample with the largest SFRs.

Two-sample KS tests comparing the SFH-derived properties of S03 protocluster and field LBGs (see Table 3.4) indicate a significant difference between the protocluster and field distributions of stellar mass; in Figure 3.8 the protocluster galaxies cluster at higher masses than field galaxies. One of the M17 LBGs, J221718.04+001735.5, is best fit by an extremely high SFR on the order of 10^3 $M_\odot$ yr^{-1}. Visual inspection of this galaxy (see Figure 3.4) shows that there is a bright unrelated point source nearby, which may be blended with the galaxy in near-IR photometry, producing an IR-heavy, high-attenuation, high-SFR best-fit SED with $p_{null} = 0.99$. We therefore exclude it when we perform KS tests on the combined S03 and M17 LBG samples. The KS tests on the combined LBG samples indicate significant differences between the protocluster and field distributions of stellar mass and SFR, with protocluster galaxies having, on average, larger masses and larger SFR. There does not appear to be a significant difference between the sSFR distributions of the

protocluster and field galaxies (as visible in Figure 3.8) or the mass-weighted age distributions of the protocluster and field galaxies.

We took advantage of our non-parametric SFH fitting technique to investigate the average SFH of SSA22 protocluster LBGs. The SFR ψ_i in each of our four stellar age bins is fit as a free parameter. We constructed a sample average SFH chain for the protocluster and field samples by averaging the ψ_i values across each sample's chains. For this exercise we used the last 1000 values of ψ_i in the MCMC chains, thus yielding sample average SFH chains with 1000 entries. We then sampled the sample average SFH chains to construct the average SFH and model SED of both samples, which we show in Figure 3.9. Regardless of the assumed metallicity, we find that the SFH of the protocluster sample is significantly elevated over the combined field SFH; for the fits with sub-solar metallicity we find that the SFH is more significantly elevated at the earlier times, while for solar metallicity we find that the elevation is more significant for the most recent age bin. For an assumed metallicity of $Z = 0.655Z_{\odot}$ ($Z = Z_{\odot}$) the maximum SFR enhancement for S03 protocluster LBGs is $2.36^{+0.46}_{-0.63}$ ($2.02^{+0.82}_{-0.70}$)[5] in the $10-100$ Gyr ($0-10$ Myr) stellar age bin. We list the protocluster-over-field SFH ratio for both metallicities and each stellar age bin in Table 3.7.

Due to the elevation of the mean protocluster SFH over the mean field SFH, the mass of the mean protocluster LBG as determined from the mean SFH is greater by a factor of 1.99 than the mean field LBG: the mean S03 protocluster LBG has $\log M_*/\mathrm{M}_{\odot} = 10.31^{+0.07}_{-0.07}$ while the mean S03 field LBG has $\log M_*/\mathrm{M}_{\odot} = 10.01^{+0.02}_{-0.02}$. We find that the mean protocluster LBG has a mass-weighted age consistent with the mean field LBG: $\log t_{AGE}/\mathrm{yr} = 8.85^{+0.04}_{-0.04}$ for the protocluster LBGs, and $\log t_{AGE}/\mathrm{yr} = 8.84^{+0.02}_{-0.02}$ for the field LBGs.

[5]SED-fit derived parameters and their uncertainties are reported as the median and 16th to 84th percentile range of the last 1000 steps in the MCMC chain.

Table 3.7: SFR enhancement as a function of time for both assumed metallicities.

	$\mathrm{SFR}/\mathrm{SFR}_{\mathrm{Field}}$	
Epoch	$Z = 0.655 Z_{\odot}$	$Z = Z_{\odot}$
$0 - 10$ Myr	$1.84^{+1.45}_{-0.65}$	$2.02^{+0.82}_{-0.70}$
$10 - 100$ Myr	$2.36^{+0.46}_{-0.63}$	$1.64^{+0.29}_{-0.26}$
$0.1 - 1$ Gyr	$1.57^{+0.59}_{-0.48}$	$1.73^{+0.60}_{-0.47}$
$1 - 2$ Gyr	$2.17^{+1.03}_{-0.81}$	$1.91^{+0.56}_{-0.53}$

Notes – Uncertainties are reported for the 16% to 84% confidence interval.

If we construct sample average attenuation curves by treating the attenuation parameters in the same way as the SFH, we find that the protocluster LBGs are more attenuated than their field counterparts. For the fits with $Z = 0.0655 Z_{\odot}$ the optical depth in the rest-frame $V-$band, τ_V, is $0.49^{+0.06}_{-0.07}$ for the mean S03 protocluster LBG and $0.19^{+0.02}_{-0.02}$ for the mean S03 field LBG.

The larger contribution to the SED from the older stellar population and the increased attenuation together produce a mean protocluster LBG SED slightly redder than the mean field LBG SED. We computed IR colors from the model SEDs in Figure 3.9, finding that protocluster and field LBGs may be distinguished by future observations with JWST bands: $J - \mathrm{F444W} = 1.68^{+0.46}_{-0.43}$ for the protocluster model, while $J - \mathrm{F444W} = 1.31^{+0.17}_{-0.17}$ for the field model. However, both SED models are still dominated by young stars. Given the uncertainties on the model parameters, these colors remain uncertain, and color differences between protocluster and field LBGs remain speculative. Observations of the protocluster with JWST, which, with NIRCam, could reach F444W $S/N \approx 60$ in 1900 s exposures, will vastly improve the constraints of our SED models across the 4000 Å break, allowing to reduce the uncertainties on our models and determine whether the color differences we have extrapolated from the models are real.

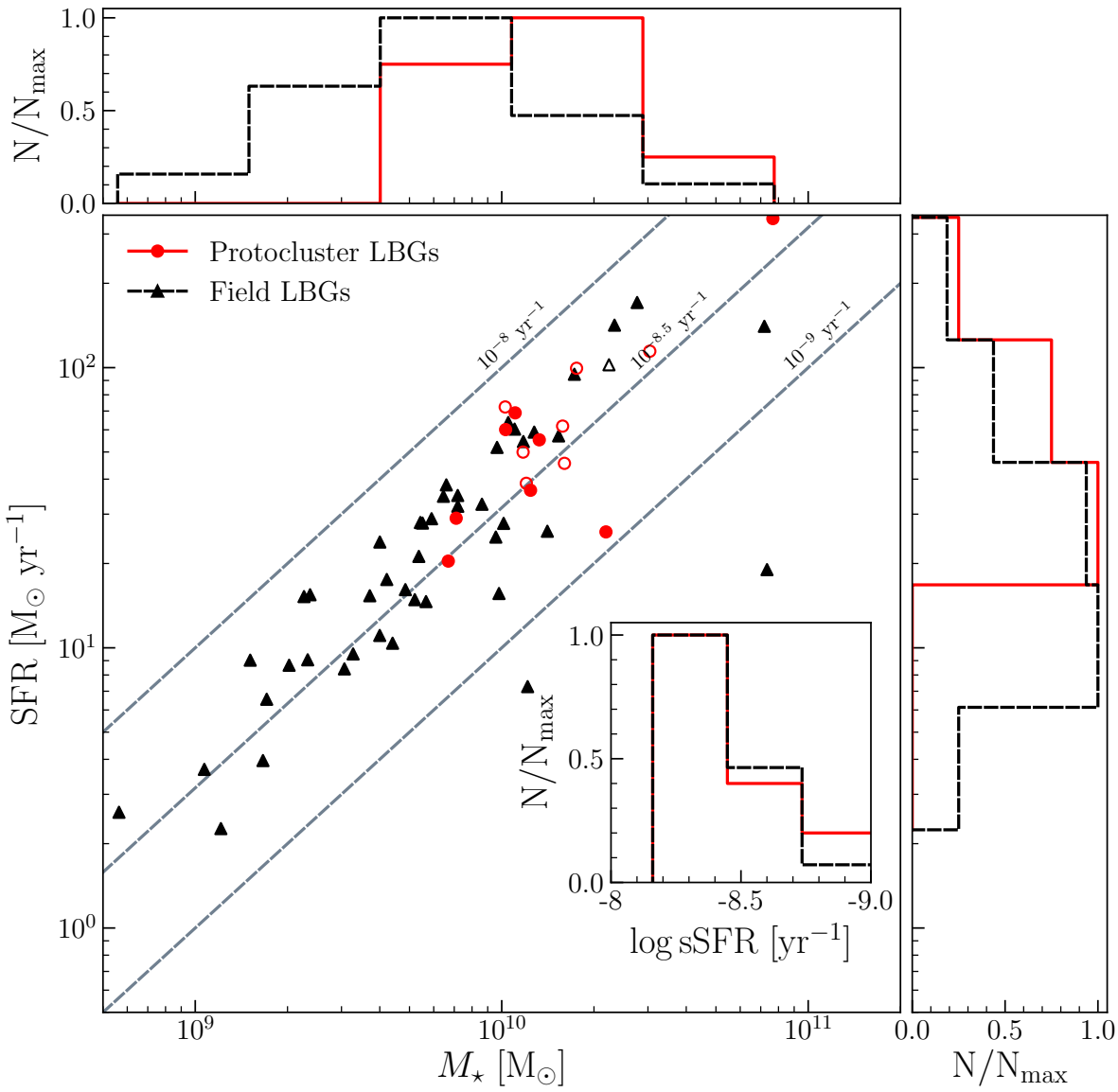

Figure 3.8: Recent SFR versus stellar mass M_* for protocluster and field LBGs from the SED fits with $Z = 0.655Z_\odot$. LBGs from the S03 catalog are shown as filled symbols, and LBGs from the M17 catalog as open symbols. For reference we show dashed lines of constant specific star formation rate (sSFR) in gray, covering 10^{-9} yr^{-1} to 10^{-8} yr^{-1} at increments of 0.5 dex. In the histograms in the margins we show the distributions of SFR and M_*, and in the inset histogram we show the distribution of sSFR. For the histograms we again include only LBGs from the S03 catalog. The protocluster LBGs appear to trend toward larger masses and SFR compared to the field LBGs.

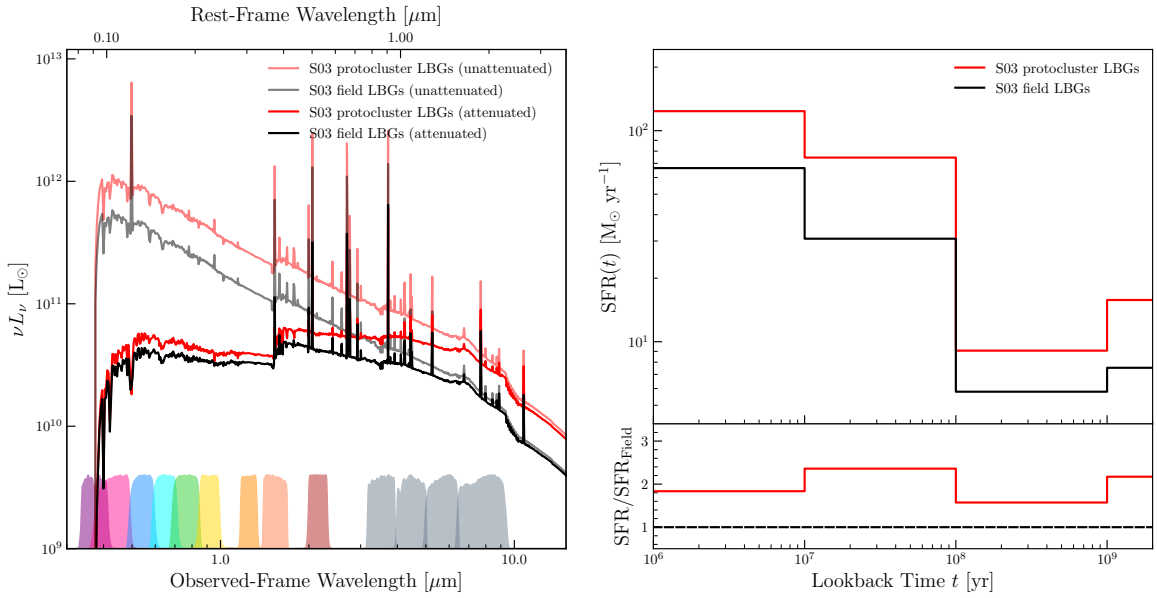

Figure 3.9: *Left panel*: We show the sample average model SED for the 8 S03 protocluster (red) and 45 S03 field (black) LBGs for which we have SED fits, both before (faded) and after (solid) application of the sample average attenuation model. Underneath the SED we show the filter curves for the bands we use for fitting for the SSA22 sample (see Table 3.6), where coverage could be improved throughout the mid-IR to better constrain the attenuation in protocluster galaxies. The protocluster model is more attenuated than the field model. *Right panel*: We show the sample average SFH of the same protocluster (red) and combined field (black) samples. Below, we show the ratio of the protocluster SFH to the field SFH, with a dashed line at unity for reference. The shaded regions show the 16$^{\rm th}$ to 84$^{\rm th}$ percentile range. In both panels we show only the results from fitting with $Z = 0.655 Z_\odot$. For this plot we have set the redshift for both models at $z = 3.1$.

3.5 Discussion

3.5.1 Star Formation and AGN Enhancement in the SSA22 Protocluster

Lehmer et al. (2009a) suggested two plausible explanations for how the SSA22 protocluster environment might lead to the observed AGN fraction excess: SMBH accretion activity may be increased through (1) more frequent accretion episodes - possibly triggered by major mergers - and higher SMBH accretion rates in the dense regions of the protocluster, or (2) an increase in the X-ray luminosity of protocluster SMBHs due to the presence of more massive galaxies (and hence SMBHs) in the protocluster.

We have tested scenario 1 by searching for evidence of major mergers in protocluster LBGs detected in the fields of the X-ray detected protocluster AGN. We note that we have not focused directly on the AGN; at the wavelengths we probe it is difficult to extract morphological information from the AGN, as the AGN contributes significantly to the observed light, resulting in a point-source-like profile superimposed on the host galaxy's light profile (see, e.g., Figure 3.6). We have instead focused on the inactive LBG population, to attempt to discern how mergers contribute to the overall growth of galaxies in the protocluster. If major mergers are a significant factor in the growth of galaxies in the protocluster, we would expect to observe differences in the morphologies of protocluster and field LBGs. Our results from quantitative and visual morphological analyses suggest that this is not the case: our samples of protocluster and field LBGs appear to be drawn from the same morphological population. We find a marginal result that the protocluster LBGs from the S03 catalog appear to have larger values of G (indicating flatter distributions of light) than their field counterparts, but when we take other Lyman-break selected galaxies from the M17 catalog into account, we find that the Sérsic parameters and nonparametric morphologies of pro-

tocluster and field LBGs are consistent with each other. Our results from model-fitting suggest that the majority of protocluster LBGs have Sérsic indices $n < 2.5$, consistent with disk-dominated light profiles, even accounting for the effects of noise and the PSF (see Section 3.A) which may cause n to be underestimated by up to 50%.

While previous work by Hine et al. (2015) found a marginally elevated merger fraction in the protocluster, we find that the merger fraction among X-ray AGN in the protocluster is consistent with both the merger fraction among S03 protocluster and field LBGs, though we are limited by the small number of protocluster LBGs we are able to use. We have attempted to go beyond the typical methods for counting mergers by examining the residuals after subtracting Sérsic models for evidence of mergers, finding a merger fraction in rough agreement with the one derived by visual classification, though this is also limited by small numbers.

The increased merger fractions in Hine et al. (2015) may be due to the influence of star formation on their F814W observations, probing the rest-frame UV. Individual UV-luminous clumps of star formation may be difficult to discern by eye from multiple nuclei: in 6 out of 10 cases, the classifications they assigned to protocluster LBGs often indicated "more than two nuclei or complex clumpy structure" rather than clear-cut cases of a double nucleus. In addition, Hine et al. (2015) found that the rest-frame UV asymmetry in protocluster LBGs, often used as a merger diagnostic, indicates fewer mergers than their visual classification; however, at high redshift, calculation of the asymmetry statistic A suffers from the same limitations on resolution and signal-to-noise per pixel as we have encountered in computing our own nonparametric morphologies. Comparisons to protocluster merger fractions from the literature outside of SSA22 are limited; Lotz et al. (2013) found a merger fraction of $0.57^{+0.13}_{-0.14}$ in a $z = 1.62$ protocluster (XMM-LSS J02182-05102; also called IRC-0218A). By comparison they measure a field merger fraction of 0.11 ± 0.03, indicat-

ing significant enhancement of merger activity in the protocluster. The enhancement of merger activity in an overdense environment is also in line with theory: in studies of Millennium simulation merger trees, Fakhouri & Ma (2009) found that overdensity enhanced merger rate relative to the mean field environment at all redshifts $z \leq 2$. However, their results also suggest that this enhancement grows weaker with increasing redshift, with overdensities at $z \sim 2$ seeing less merger rate enhancement relative to overdensities at lower redshift (see e.g., Figure 5 in Fakhouri & Ma, 2009).

Our discussion thus far has been focused on major mergers (mass ratio ≥ 0.25), which are expected to cause the largest and clearest morphological disturbances and possibly trigger an AGN phase. However, if the overall merger rate is enhanced in overdense environments, minor mergers ($0.10 \leq$ mass ratio < 0.25) should also be more abundant. Our techniques are not particularly well suited for detection of minor mergers, largely due to surface brightness limits making detection of low-mass satellite galaxies difficult. In particular, the algorithm used to create the $G-M_{20}$ segmentation map tends to exclude low surface brightness satellite galaxies, including only the bright core region of the primary galaxy and thus making the $G-M_{20}$ merger diagnostic insensitive to galaxies accreting lower-mass satellites. In cases where low-mass satellites are segmented properly, their presence still may not move the system in the merger region of the $G-M_{20}$ diagram. Lotz et al. (2010) plotted tracks in $G-M_{20}$ space for the course of mergers at a variety of mass ratios and viewing angles, finding that the $G-M_{20}$ diagnostic is not very sensitive to the early stages of low mass ratio mergers. They also find that flyby cases with low mass ratios are not cleanly diagnosed by the $G-M_{20}$ diagnostic and do not trigger long-lasting asymmetries that might be visible by eye. For the visual classifications we performed, the categories we asked the voters to use were designed with major mergers in mind. However, our residual analysis may be more sen-

sitive to minor mergers. Estimating the feature masses and feature mass to total mass ratios of the features we have extracted from the residual images, we found that the masses of some features are consistent with dwarf galaxies, and that the mass ratios approach 0.10 at the low end. Some of the clumpy features we see in the residuals may thus be infalling satellites. While it is possible that low-mass ratio or early stage mergers our techniques are less sensitive to could contribute to the enhancement of star formation in the protocluster, studies vary on whether mergers do (e.g., Zamojski et al., 2011; Kartaltepe et al., 2012) or do not (e.g., Targett et al., 2011) play a significant role in triggering bursts of star formation at $z \sim 2$, suggesting that the influence of mergers on star formation may vary with redshift and among galaxies selected by different methods.

At $z > 1$, the highest luminosity (bolometric luminosity $\gtrsim 10^{46}$ erg s^{-1}) AGN are preferentially found in disturbed systems, believed to be evidence of recent mergers, though AGN luminosity and merger fraction both scale with mass at fixed redshift. McAlpine et al. (2020) find, on the basis of EAGLE simulations, that while high-luminosity AGN are more likely to be found in mergers, major mergers are only an effective driver of AGN fraction enhancement for lower-mass ($< 10^{11}$ M$_\odot$) host galaxies. Five of the eight protocluster X-ray AGN in our F160W images (those without broad optical emission lines) have stellar mass estimates from optical SED fitting by Kubo et al. (2015a), finding stellar masses in the range 0.3–2 $\times 10^{11}$ M$_\odot$, on the edge of where major mergers are expected to be effective triggers for AGN activity from EAGLE simulations.

We note, however, that our reliance on a Lyman-break selected sample excludes some high mass galaxies, which are more likely to be found in mergers. Indeed, Kubo et al. (2017) found some evidence of merger-driven evolution in a group of massive quiescent galaxies at the protocluster core. It may be then that while our results suggest only that the protocluster LBG population (or equivalently, galaxies in an LBG phase of their evolution) are not any more likely to be found in

major mergers than their field counterparts and that the protocluster overall may not be a more merger-rich environment than the field, mergers may still play a role in the evolution of the most massive protocluster galaxies.

While our results suggest that accretion due to major mergers is likely not a major environmental difference between the protocluster and field, and hence not the primary source of the X-ray enhancement in the fields we have studied here, galaxy-scale accretion is not ruled out. Recent simulations (e.g. Romano-Díaz et al., 2014) taken together with IFU observations also suggest that major mergers may not be the dominant mode for the accumulation of stellar mass in the SSA22 protocluster. Rather, smooth accretion of gas flows along filaments of the cosmic web between galaxies may power in-situ star formation, SMBH growth, and AGN activity. These filaments have been imaged in emission in the SSA22 protocluster, and rough calculations estimate that they may contain $\sim 10^{12}\ M_{\odot}$ of gas available for accretion (Umehata et al., 2019). However, it is difficult to establish whether inflows along these filaments exist, or whether they could carry enough gas into a galaxy to power an AGN.

The nodes of the imaged web notably coincide with the massive, star-forming submillimeter galaxies observed in Umehata et al. (2018). Four out of the eight X-ray AGN studied in Alexander et al. (2016) are co-located with large Lyα nebulae. We find that two of the S03 protocluster LBGs in our GALFITM sample, SSA22a-C47 and SSA22a-M28, are associated with Lyα nebulae from the Matsuda et al. (2004) Lyα blob survey: LAB 11 and LAB 12, respectively. The Sérsic model parameters and nonparametric morphologies for these galaxies are comparable to those for the other protocluster and field LBGs in our sample, and SSA22a-M28 is classified as isolated but slightly irregular by our visual classification scheme, while SSA22a-C47 was not classified by our voters. However, we find some evidence of a possible tidal bridge between SSA22a-C47 and its

projected companion when examining its residual after Sérsic model subtraction (see Figure 3.4). With such a small sample we are unable to establish whether the LABs, which appear to be associated with AGN activity, are also associated with major mergers. We attempt to draw more detailed connections between the morphologies and physical properties of our LBG sample and their local, Mpc-scale environment in Section 3.5.2.

We investigated scenario 2 (which already has significant evidence from prior SED fitting and sub-mm studies, e.g. Kubo et al., 2013; Umehata et al., 2018) by fitting the SEDs and non-parametric SFHs of our samples of protocluster and field LBGs. Typical studies of high-redshift AGN focus on hard X-ray detected AGN with large bolometric luminosities, making studies of the physical causes of AGN fraction enhancement difficult to control for mass. The result of Yang et al. (2018) establishes that, at $z > 2$, AGN fraction does not depend on environment when galaxies with similar stellar masses are compared. Galaxies are expected to be more massive in protoclusters at intermediate redshifts (with $2 \lesssim z \lesssim 4$) when compared to the field (e.g., Steidel et al., 2005; Hatch et al., 2011; Cooke et al., 2014). Since SMBH mass should scale with stellar mass (e.g., Ferrarese & Merritt, 2000; Ding et al., 2020) we should thus expect to have enhanced AGN activity in protoclusters with respect to the field at intermediate redshifts (e.g., Lehmer et al., 2009a, 2013; Digby-North et al., 2010; Vito et al., 2020). Using KS tests we found that the protocluster and field samples appear to be drawn from the same distributions of sSFR and mass-weighted stellar age, while the protocluster galaxies have stellar mass and SFR distributions significantly weighted toward higher mass and higher SFR. We have also found that the mean SFH of our sample of protocluster LBGs is elevated by a factor of about 2 over the mean field LBG SFH in the earliest stellar age bins, from 10 Myr to approximately 2 Gyr. This elevated star formation rate in the oldest stellar population leads to a mean protocluster LBG about 2.2 times more massive than the

mean field LBG. Thus, the observed AGN fraction enhancement in the SSA22 protocluster may largely be an effect of the enhanced mass of the typical protocluster galaxy.

3.5.2 Correlations Between Local Environment and Galaxy Properties

To assess the variance in morphology and star formation properties throughout the protocluster, we use a Gaussian kernel density estimation (KDE) to estimate the LAE surface density, Σ_{LAE}. At the location of each $z \approx 3.1$ LAE from (Hayashino et al., 2004) we place a circular 2D Gaussian with a FWHM of $2'$, corresponding to 3.75 co-moving Mpc. The resulting distribution is re-normalized as a surface number density, and sampled at the positions of our protocluster LBG sample. Hayashino et al. (2004) place the threshold of the "high-density region" of the SSA22 protocluster at a LAE surface density of 0.26 arcmin^{-2}. We find that due to the construction of our fields, all of our LBGs are in the high-density region, and they all have local $\Sigma_{LAE} > 0.5$ arcmin^{-2}, which is > 5 times the average density of the control field in Hayashino et al. (2004).

We plot the morphological and physical properties of our samples of protocluster LBGs in Figure 3.10 and Figure 3.11, respectively. To examine whether galaxies move through $G - M_{20}$ space as a function of Σ_{LAE}, we define two additional morphological measures, the merger statistic μ and the bulge statistic β, as the signed perpendicular distance from the lines $G = -0.14M_{20} + 0.33$ and $G = 0.14M_{20} + 0.80$, respectively:

$$\mu = 0.14M_{20} + 0.99G - 0.33 \tag{3.8}$$

$$\beta = -0.14M_{20} + 0.99G - 0.79 \tag{3.9}$$

Galaxies with more merger-like morphologies (in the sense of Lotz et al., 2008; see Figure 3.6) thus have larger values of μ and galaxies with more bulge-like morphologies have larger values of

β.

If we suppose that mergers are more common in the denser regions of the protocluster, we should expect the morphologies of galaxies to have identifiable trends with the projected LAE density. We find no strong correlations between the parametric or nonparametric morphological measurements and Σ_{LAE}. There is an apparent downward trend in G, but a Pearson test shows that it is marginal with $p = 0.09$, and n and C, which also probe the concentration of the galaxies, exhibit no such trend. The only other parameter that shows any marginal correlation with Σ_{LAE} is the Sérsic model magnitude m ($r = 0.50$, $p = 0.05$), growing fainter with increasing density. We find that this is likely due to biases in our images: the two-orbit depth images we use are targeted on the denser regions, so in the less dense regions we preferentially select brighter galaxies, and only brighter galaxies have acceptable Sérsic fits. If we repeat the Pearson test for all S03 and M17 protocluster LBGs regardless of Sérsic fit quality, using the F160W aperture magnitudes we measured for SED fitting, we find that in this larger sample there is no significant correlation between magnitude and Σ_{LAE}. Kubo et al. (2013) and Kubo et al. (2017) found that several massive galaxies in the densest regions of the protocluster have compact sizes and $n > 2.5$ Sérsic profiles similar to local early-type galaxies (ETGs). Our results do not reflect this, though we note that the maximum recovered Sérsic index among our LBGs is < 2 and our F160W imaging and SSA22 LBG samples do not cover the entirety of the "core" region of the protocluster targeted by the ALMA deep fields (ADF). In particular, the AzTEC14 group at the protocluster core (Kubo et al., 2015b), where Kubo et al. (2017) found massive galaxies similar to local ETGs, is not covered by our F160W images. In this region, where the proto-brightest cluster galaxy (BCG) is predicted to be evolving, mergers may be a driving factor behind evolution. However, Kubo et al. (2013) and Kubo et al. (2017) studied only a subset of the massive galaxies in the protocluster and Kubo

et al. (2017) noted a possible deficiency in low-mass galaxies in the AzTEC14 group, implying that a more sensitive mass complete sampling of the protocluster core may be necessary to establish whether mergers are ongoing or in the past.

We find that none of the SED-fit derived physical properties are correlated with Σ_{LAE}. It is well established that there are intensely star forming galaxies (many of which also host AGN) in the densest regions of the protocluster with IR-derived star formation rates on the order of 10^2–10^3 $M_\odot$ yr^{-1} (Umehata et al., 2015; Alexander et al., 2016; Kato et al., 2016, e.g.). By the placement of our fields and the construction of our SED fitting sample, which excludes known AGN, we have excluded these DSFGs. Sub-millimeter observations and previous SED fitting have also shown that there are massive galaxies in the most dense region of the protocluster (e.g. Kubo et al., 2013, 2015b). Our results, which finds no strong trend between the SFR or mass of protocluster LBGs and Σ_{LAE} may not conflict with these established results: it may be that the general star-forming galaxy population is not more massive or more intensely star forming in the denser regions of the protocluster, but rather that some galaxies in the densest regions (i.e., AzTEC14) are exceptionally massive, the possible predecessors of what will become the BCG as the protocluster evolves. In addition to our results that find that the general LBG population of the protocluster does not appear to be rapidly, currently evolving, we thus also find that it appears that LBGs in denser areas of the protocluster are evolving no more rapidly than elsewhere in the protocluster.

3.5.3 JWST Prospects for SSA22

While the primary limits on this study are imposed by the small numbers of protocluster and field galaxies, morphological studies at $z \sim 3$ are also limited by the sensitivity and resolution of

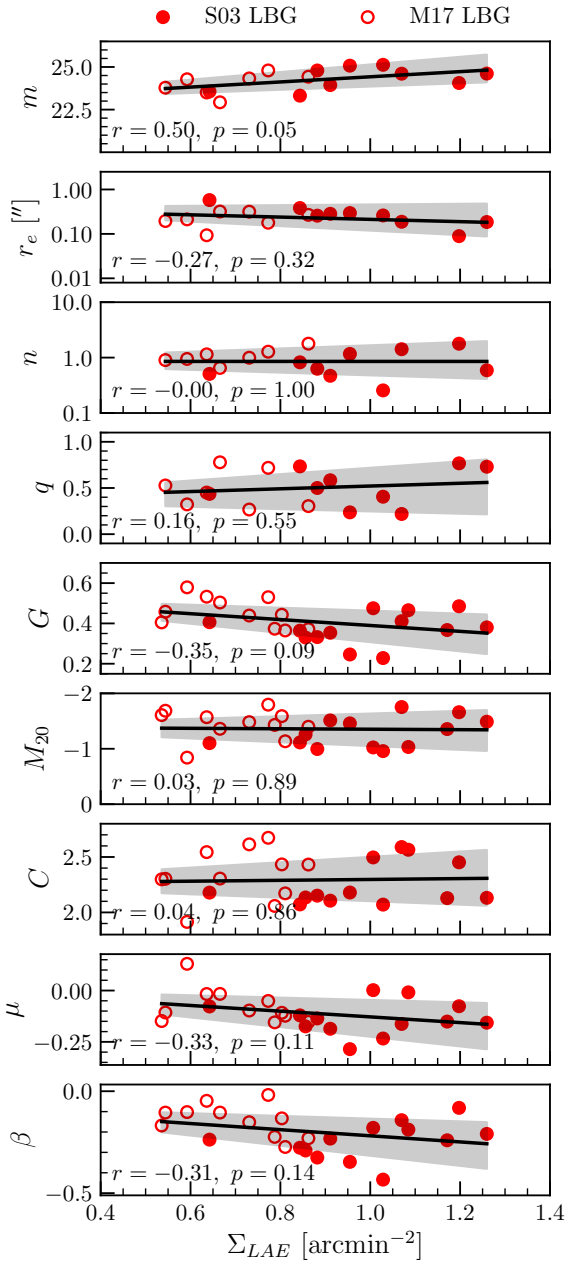

Figure 3.10: We show all of our morphology measurements as a function of LAE surface density for our sample of protocluster LBGs. From the top panel down: Sérsic model F160W magnitude m, Sérsic fit effective radius r_e, Sérsic index n, Sérsic fit axis ratio q, Gini coefficient G, second order moment of light M_{20}, concentration index C, and the $G - M_{20}$ merger and bulge statistics μ and β. We show LBGs from the S03 catalog as filled symbols, and LBGs from the M17 catalog as open symbols. For each property we plot a linear regression to both sets of LBGs with the 16th to 84th percentile interval for the slope (as computed from bootstrap resampling). We print the Pearson test statistic and probability for each measurement *when both S03 and M17 LBGs are included* in the corner of each panel.

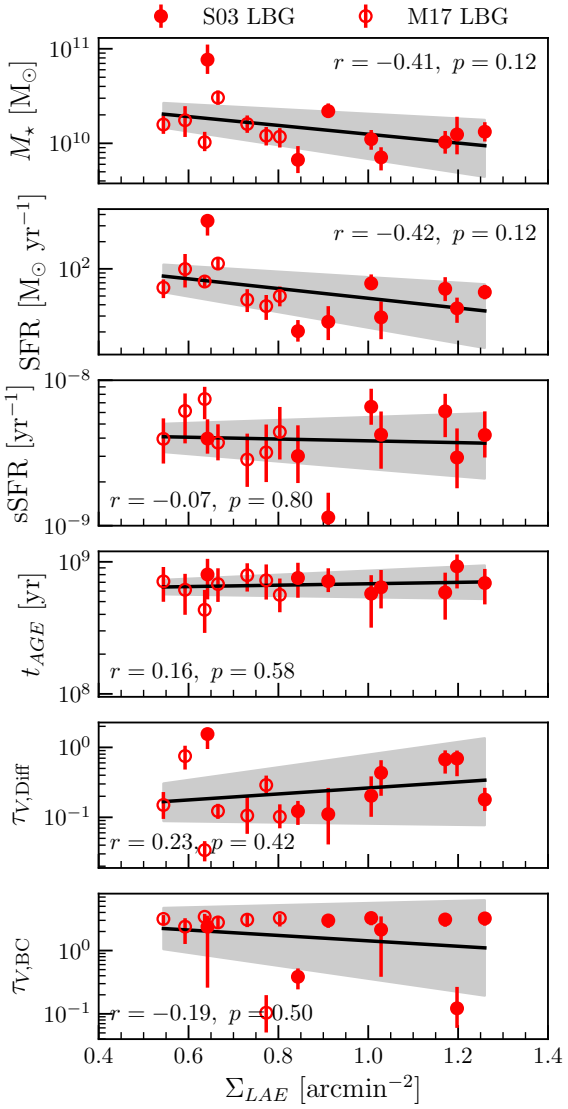

Figure 3.11: We show all of our SFH-derived measurements as a function of LAE surface density for the $Z = 0.655Z_\odot$ SED fits to our sample of protocluster LBGs. From the top panel down: stellar mass $M_\star$, star formation rate over the last 100 Myr, specific star formation rate over the last 100 Myr, mass-weighted age t_{AGE}, diffuse optical depth at the rest-frame $V-$band $\tau_{V,\mathrm{Diff}}$, and birth cloud optical depth at the rest-frame $V-$band $\tau_{V,\mathrm{BC}}$. We show LBGs from the S03 catalog as filled symbols, and LBGs from the M17 catalog as open symbols. For each property we plot a linear regression to both sets of LBGs with the 16th to 84th percentile interval for the slope (as computed from bootstrap resampling). We print the Pearson test statistic and probability for each measurement *when both S03 and M17 LBGs are included* in the corner of each panel.

current near-IR telescopes. Our two-orbit depth F160W images are insufficient to map the full extent of plausible tidal features, and the angular resolution in our images (slightly less than 500 kpc pix^{-1}) is at the limit of what Lotz et al. (2004) recommend for nonparametric morphological measurements. Next-generation space telescopes such as the James Webb Space Telescope (JWST) will present significant improvements in both regards. The search for low-surface brightness tidal features will benefit from improved sensitivity, while quantitative morphological measurements like Sérsic model fitting will also benefit from improved angular resolution throughout the rest-frame optical.

For example, the field of view of our images contains at least one interesting system believed to be a complex merger: the AzTEC1/ADF22.A1 system (Tamura et al., 2010; Umehata et al., 2014, 2015, 2017), shown in Figure 3.12. The primary component of the system is an ALMA 1.1 detection with a coincident *Chandra* detection, with faint companion galaxies that are not visible at wavelengths < 1 μm. The companions, of which there appear to be at least two, are visible in F160W, though they are not included in our catalog due to the S/N cuts we impose, and they are not Lyman-break selected due to their non-detection blueward of 1 μm. The SED of the system appears to be consistent with an AGN buried in a dusty, highly star forming galaxy. The system is thus speculated to be a protoquasar, possibly fueled by the major merger of the F160W-detected galaxies nearby. JWST ETC simulations using the currently observed SED suggest that NIRCam and MIRI observations will produce near- and mid-IR detections of all three components this system at higher resolution than even the currently available *HST* data, allowing us to search for merger features throughout the IR and to treat this system as a case study in whether mergers fuel AGN in the protocluster environment.

While we have been unable to establish whether major mergers are a dominant mode of growth

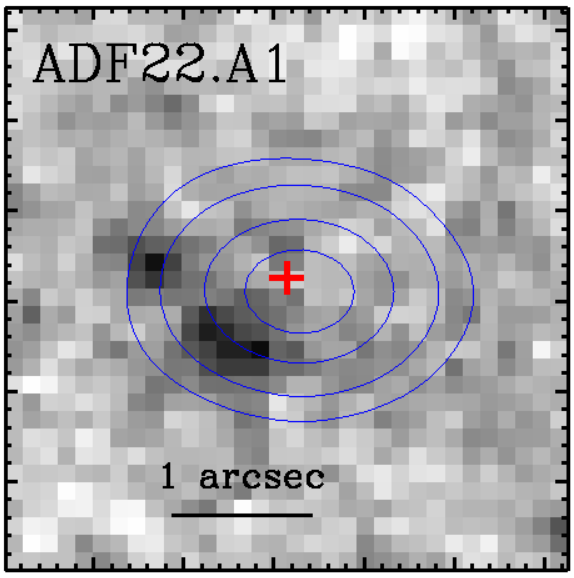

Figure 3.12: The AzTEC1/ADF22.A1 system, as seen by WFC3/F160W. The system consists of at least two closely associated (< 5 kpc) F160W-visible galaxies, with offset Chandra (red cross) and ALMA 1.1 mm (blue contours) detections. The offset source is not detected in ACS bands, and appears very faintly in F160W.

among galaxies throughout the protocluster, we have not examined the very core of the protocluster where the system that will become the BCG may be evolving. The BCGs in Coma-like clusters, which the SSA22 protocluster is expected to become, are thought to form as the result of successive major mergers in the cluster core. A group of massive galaxies has already been identified in the protocluster core by Kubo et al. (2015b), which may be the site where the future BCG is forming. A larger-than-expected portion of these galaxies are already quiescent and bulge-dominated, suggesting significant evolution and possible previous merger activity. JWST observations of this system and morphological analyses of its components could allow the detection of merger signatures, giving us a possible window into the early stages of BCG formation.

3.6 Summary

We have pursued multiple avenues of morphological analysis on protocluster and field galaxies detected in new and archival *HST* WFC3 F160W images of the SSA22 protocluster. We fit single Sérsic models to galaxies detected in our images to extract effective sizes and Sérsic indices, and then examined the residual images after model subtraction for evidence of tidal features. We also calculated the Gini coefficient G, moment of light M_{20}, and concentration statistic C for protocluster and field galaxies detected in our images. For a third point of comparison we used a visual classification scheme modeled on Hine et al. (2015) to examine the observed merger fractions among protocluster and field LBGs.

To supplement our morphological analysis, we fit the UV-to-near-IR SEDs and non-parametric SFHs of a sample of protocluster and field galaxies, in order to measure stellar masses and SFR.

Our main results and conclusions are as follows:

- Using two-sample KS tests to compare the Sérsic fit morphologies of SSA22 protocluster LBGs to a combined sample of field LBGs from SSA22 and GOODS-N, we find no significant differences in the protocluster and field distributions of any of the model parameters, including effective size r_e ($p_{KS} \geq 0.37$) and Sérsic index n ($p_{KS} \geq 0.17$).

- We find evidence of tidal features in the residual images of both SSA22 protocluster and field LBGs after subtracting the best-fit Sérsic model. Based on this analysis we estimate rough merger fractions of 0.13–0.44 for protocluster LBGs and 0.14–0.28 for field LBGs. We estimate that the largest and brightest of the plausible tidal features have masses $\log_{10} M_\star/M_\odot = 9.78$, suggesting that they may be as massive as the Small Magellanic Cloud.

- Using two-sample KS tests comparing the non-parametric morphologies of SSA22 protocluster and field galaxies, we find no significant differences in the protocluster and field distributions of M_{20} ($p_{KS} \geq 0.30$), and C ($p_{KS} \geq 0.30$). We find a marginal difference between the G distributions of protocluster and field S03 LBGs ($p_{KS} = 0.04$). However, this is not supported by the KS tests on any of the other measures of the galaxies' concentration (i.e., n and C), and is not evident when M17 LBGs are included in the KS test ($p_{KS} = 0.25$). We note that only one of the galaxies in which we identify a plausible tidal feature is classified as a merger by the Lotz et al. (2008) cuts in the G–M_{20} plane. We thus hesitate to estimate a merger fraction based on non-parametric morphological analysis.

- By performing visual merger classifications of selected F160W galaxy cutouts for a direct comparison to Hine et al. (2015) we estimate merger fractions $0.38^{+0.37}_{-0.20}$ among S03 protocluster LBGs and $0.41^{+0.11}_{-0.09}$ among S03 field LBGs. We note that visual classifications from our SSA22 images are limited by the number of S03 LBGs in our images, and that the

number of mergers may be undercounted due to limited depth.

- We find that the SED fits to our small sample of protocluster and field LBGs are consistent with elevated star formation in the protocluster's oldest stellar population, with the mean protocluster SFH being significantly elevated over the mean field SFH between lookback times 100 Myr – 2 Gyr. The mean protocluster LBG is also more massive and more attenuated by a factor of 2 compared to the mean field galaxy. The mean protocluster LBG is thus slightly redder in terms of IR color, with $J - \mathrm{F444W} = 1.68^{+0.46}_{-0.43}$, than the mean field LBG, which has $J - \mathrm{F444W} = 1.31^{+0.17}_{-0.17}$. However, young stars are still the dominant contributor to the SEDs of both protocluster and field LBGs.

- Based on our results, we conclude that the observed enhancement in the SSA22 protocluster AGN fraction may be due to the larger average stellar mass (and hence larger average SMBH mass) of galaxies in the protocluster. In addition, the protocluster LBGs we have studied here appear to have formed more stellar mass earlier than their field counterparts.

Our results are limited throughout by the small number statistics of protocluster LBGs; we are only able to identify 24 protocluster LBGs in our F160W images, and requirements on converged Sérsic fits and available photometry mean that in practice we can only use a fraction of them in our analysis. These limits also mean that our analyses are difficult to control for mass; studies of merger fractions in particular are sensitive to mass, as merger fraction increases with stellar mass at fixed redshift. Our focus on Lyman-break selected galaxies in this work may also exclude more massive galaxies with more evolved SED shapes (Wang et al., 2019), which may be involved in ongoing mergers or have morphologies that indicate past mergers. We are hopeful that the increased sensitivity of JWST will allow the construction of true mass-selected samples, which

will allow deeper studies of the connections between stellar mass, AGN fraction, and overdensity.

Appendix

3.A Reliability of Morphological Measurements

We assessed the reliability of our model-fitting and non-parametric analyses with a Monte Carlo technique. We created three sets of GALFITM Sérsic models with indices $n = \{1.0, 1.5, 4.0\}$. The other parameters of the model were allowed to vary randomly over their observed ranges. We convolved these synthetic galaxies with our PSF model, added them to blank sky frames from our images, and re-measured their morphologies with GALFITM and our non-parametric analysis procedure as described in Section 3.3.3. We show the relative differences between the recovered morphologies and morphologies as measured from un-convolved, noise-free models in Figure 3.A.1.

We find that the reliability of the GALFITM-recovered Sérsic radius r_e and axis ratio q are strongly dependent on the signal-to-noise ratio, with GALFITM consistently underestimating the effective radii of low-S/N galaxies by as much as 75% and consistently converging on unrealistically small axis ratios for low-S/N galaxies. Galaxies with bulge-like $n = 4$ profiles are more strongly affected in both cases. There are large errors in the recovered value of the Sérsic index n for all initial values of n and all values of S/N. The recovered value is significantly smaller than the true value in all cases, indicating that the over-representation of galaxies with $n \leq 1.0$ in our results may be due to underestimation of the "true" n for these galaxies. However, we note also that the consistent underestimation of n and r_e is expected from the nature of these simulations; convolution with the PSF spreads out the light profile of the model, making it flatter, and the observed effective radius of the model galaxies is naturally decreased by the addition of noise. Since the effects of PSF blurring and Poisson noise cannot be removed from the images in any practical

situation, the "true" values of n and r_e are not necessarily recoverable, and our measurements serve only as a description of the data.

The non-parametric morphological measures are generally more stable with S/N. M_{20} is the most strongly affected, with low-S/N galaxies measured to have significantly larger M_{20}; that is, the models are observed to be clumpier after convolution with the PSF and the addition of noise.

With the exception of n, r_e, and C (which also depends on measurements of the effective size of the galaxy and is thus biased low even for high S/N), the median relative errors in the morphological measurements are $< 10\%$ within the 2σ range for $S/N \gtrsim 100$ and input Sérsic indices 1.0 and 1.5.

As the majority of galaxies in our catalog have Sérsic indices < 4, we take these results to show that our morphological analyses are reliable for our S03 sample (with one exception, our S03 protocluster LBGs have $S/N \gtrsim 100$, and all of them have $n < 2$). In general, we expect the Sérsic fit results in the catalog to be reliable for the other galaxies in our sample (provided they meet our other criteria for acceptable fits) for $S/N \gtrsim 100$.

3.B SED Fit Results

We show the SED fit results with $Z = 0.655Z_{\odot}$ for our SSA22 LBG samples in Table 3.B.1, Figure 3.B.1, Figure 3.B.2, Figure 3.B.3, and Figure 3.B.4. We also show the SED fit results with $Z = 0.655Z_{\odot}$ for our sample of GOODS-N field LBGs in Table 3.B.2 and example fits in Figure 3.B.5. For the full set of GOODS-N SED plots, please refer to the online version of this article.

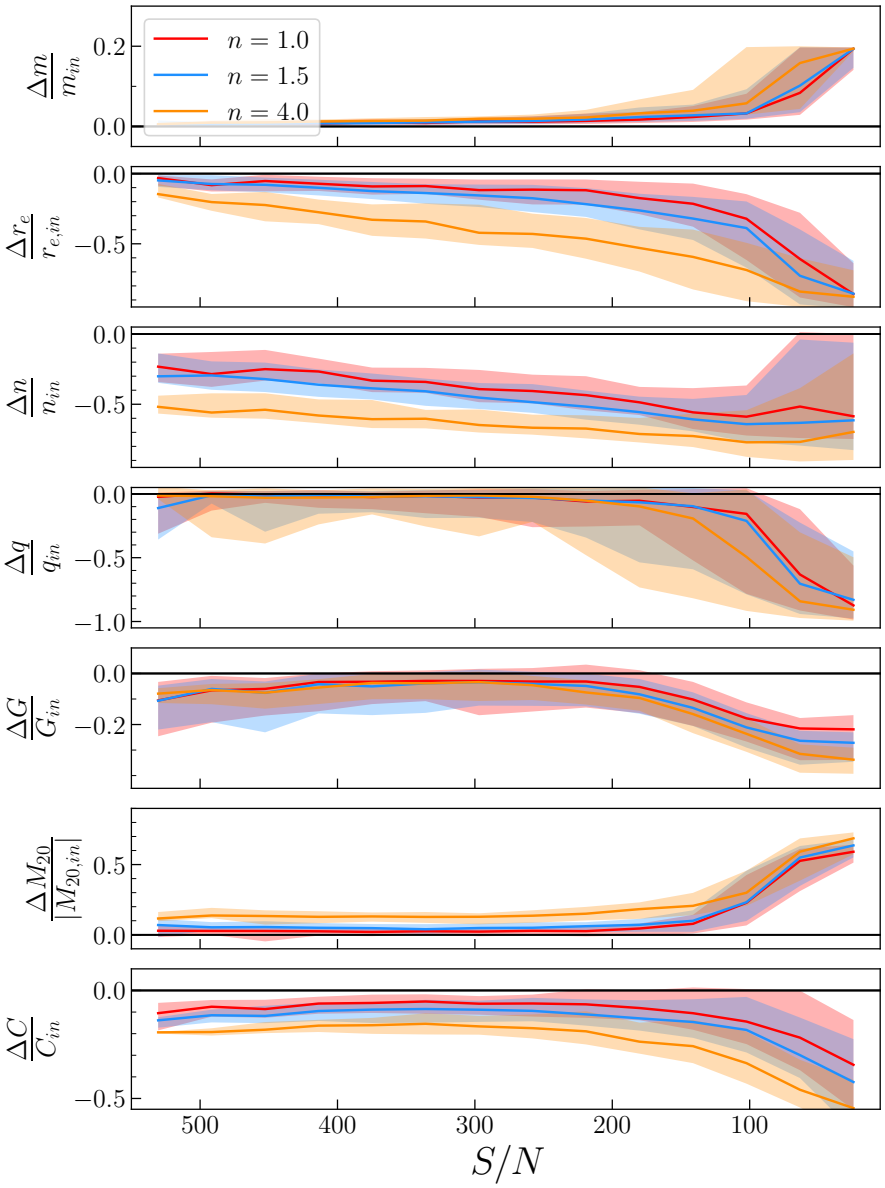

Figure 3.A.1: For each morphological measurement, the median relative difference between the initial values in three sets of noise-free, un-convolved Sérsic models and the extracted values after PSF convolution and the addition of noise is shown as a function of the final signal-to-noise ratio in a $1''$ diameter aperture. The shaded regions show the 16th to 84th percentile interval.

Table 3.B.1: We show the SED-fit derived physical properties for our samples of SSA22 LBGs.

ID	z[a]	χ^2_{min}[b]	$P_{null}(\chi^2_{min})$[c]	M_*[d] (10^9 $M_\odot$)	SFR[e] ($M_\odot$ yr^{-1})	sSFR[f] (10^9 yr^{-1})	Age[g] (10^8 yr)	Ref.[h]
SSA22 Protocluster								
J221721.02+001708.9	3.076	9.56	0.05	7.10	29.01	4.22	6.42	S03
J221718.60+001815.5	3.079	2.95	0.57	10.30	60.10	6.11	5.86	S03
J221727.27+001809.6	3.080	10.99	0.05	11.06	68.97	6.57	5.74	S03
J221731.69+001657.9	3.088	2.73	0.84	21.88	25.92	1.14	7.14	S03
J221718.87+001816.2	3.089	6.32	0.39	12.42	36.50	2.94	9.24	S03
J221737.92+001344.1	3.094	3.54	0.62	13.26	55.25	4.20	6.91	S03
J221719.30+001543.8	3.097	6.51	0.37	76.70	340.37	3.97	8.02	S03
J221720.25+001651.7	3.098	3.86	0.70	6.69	20.40	3.01	7.55	S03
J221732.04+001315.6	3.065	3.98	0.41	30.42	114.46	3.72	6.78	M17
J221701.38+002031.9	3.073	5.27	0.51	16.02	45.55	2.85	7.91	M17
J221718.96+001444.5	3.091	1.95	0.92	15.80	61.85	3.95	7.11	M17
J221718.04+001735.5	3.093	13.50	0.04	197.17	1707.56	9.05	2.42	M17
J221740.98+001127.2	3.093	4.03	0.40	17.55	99.49	6.15	6.14	M17
J221710.35+001920.8	3.103	7.47	0.28	10.27	72.38	7.43	4.33	M17
J221720.55+002046.3	3.103	2.65	0.85	11.74	49.95	4.41	5.62	M17
J221704.34+002255.8	3.108	11.20	0.08	12.03	38.65	3.19	7.24	M17
SSA22 Field								
J221724.44+001714.4	3.018	6.60	0.36	12.76	58.88	4.75	5.65	S03
J221735.98+001708.2	3.018	7.25	0.30	10.14	27.81	2.69	5.94	S03
J221735.30+001723.9	3.019	2.03	0.92	8.62	32.55	3.93	6.25	S03
J221722.90+001608.9	3.019	9.33	0.16	12.15	7.27	0.60	8.39	S03
J221725.64+001612.5	3.290	8.47	0.21	15.33	57.08	3.73	8.01	S03
J221717.69+001900.3	3.288	4.95	0.55	22.39	102.38	4.68	5.23	M17

Notes – [a]Redshift from literature; see Section 3.2.2. [b]Minimum χ^2 of of SED fitting chain. [c]Probability of *accepting* the hypothesis that the data were generated by the best-fit model. [d]Stellar mass produced by median SFH. [e]Star formation rate averaged over the last 100 Myr of the median SFH. [f]Specific star formation rate averaged over the last 100 Myr of the median SFH. [g]Mass-weighted age of the median SFH. [h]Catalog reference: S03=Steidel et al. (2003), M17=Micheva et al. (2017).

Table 3.B.2: We show the SED-fit derived physical properties for our sample of GOODS-N field LBGs from the Steidel et al. (2003) catalog.

ID	z[a]	χ^2_{min}[b]	$P_{null}(\chi^2_{min})$[c]	M_*[d] (10^9 $M_\odot$)	SFR[e] ($M_\odot$ yr^{-1})	sSFR[f] (10^9 yr^{-1})	Age[g] (10^8 yr)
J123627.59+621130.1	2.917	9.52	0.48	9.66	51.92	5.45	6.29
J123712.30+621138.2	2.925	4.06	0.91	2.02	8.67	4.23	7.71
J123706.62+621400.3	2.926	15.81	0.20	17.26	94.71	5.53	5.85
J123703.24+621451.3	2.926	10.06	0.26	2.26	15.24	6.84	5.17
J123650.40+621055.6	2.928	12.93	0.07	0.56	2.59	4.55	6.42
J123644.11+621311.2	2.929	20.97	0.03	27.66	170.82	6.30	5.30
J123617.54+621310.1	2.930	8.78	0.27	2.33	9.06	3.81	7.72
J123707.72+621038.2	2.931	13.56	0.19	5.90	28.90	4.85	6.51
J123647.77+621256.1	2.932	17.39	0.10	11.02	60.34	5.56	5.24
J123717.38+621247.4	2.939	9.71	0.47	6.60	38.16	5.83	6.55
J123709.34+621047.1	2.942	9.42	0.31	1.51	9.04	5.83	5.76
J123647.72+621053.6	2.943	9.48	0.49	14.09	26.07	1.80	7.78
J123639.27+621713.4	2.944	8.62	0.57	10.50	63.71	6.17	5.57
J123642.40+621448.9	2.962	6.46	0.78	2.37	15.47	6.70	4.97
J123646.97+621226.5	2.970	10.73	0.22	1.07	3.68	3.54	7.35
J123651.56+621042.2	2.975	15.80	0.15	23.33	141.81	6.14	4.69
J123637.15+621547.8	2.975	5.67	0.84	5.36	21.20	4.05	5.95
J123635.55+621522.0	2.980	39.49	0.00	3.07	8.41	2.83	7.98
J123622.63+621306.4	2.981	12.44	0.26	73.39	19.02	0.26	11.20
J123645.04+620940.8	2.983	13.30	0.15	4.00	11.07	2.78	9.34
J123647.88+621032.3	2.990	3.79	0.96	5.43	28.01	5.21	6.24
J123626.95+621127.4	2.993	5.99	0.74	5.51	27.89	5.06	6.28
J123640.93+621358.6	3.087	10.38	0.32	9.80	15.63	1.55	9.90
J123650.80+621444.8	3.106	25.27	0.00	1.71	6.55	3.97	7.56
J123648.88+621502.6	3.115	17.99	0.08	9.57	24.85	2.53	7.21
J123619.40+621502.0	3.128	8.92	0.45	4.40	10.39	2.36	8.02
J123658.99+621714.4	3.130	17.36	0.07	4.00	23.86	5.92	4.99
J123649.03+621542.4	3.136	4.24	0.94	6.45	34.75	5.41	5.09
J123721.63+621350.5	3.148	31.89	0.00	3.27	9.51	3.10	8.49
J123645.39+621347.1	3.161	13.75	0.09	1.66	3.96	2.41	7.32
J123651.18+621349.0	3.163	43.97	0.00	4.85	16.13	3.43	7.52
J123634.88+621253.9	3.182	26.63	0.00	1.21	2.26	1.91	8.33
J123653.61+621410.5	3.196	6.45	0.49	5.20	14.86	3.04	6.40
J123641.87+621107.4	3.197	11.33	0.33	3.71	15.34	4.09	6.22
J123702.70+621426.3	3.214	17.47	0.06	71.93	140.47	1.95	8.63
J123641.26+621203.4	3.222	18.91	0.04	5.66	14.62	2.61	6.90
J123645.19+621652.4	3.229	7.35	0.60	11.77	54.49	4.64	6.17
J123703.26+621635.3	3.239	46.47	0.00	7.19	34.98	4.91	5.41
J123637.02+621044.5	3.241	9.79	0.20	4.22	17.54	4.24	6.53
J123706.16+621509.9	3.246	4.84	0.77	7.19	32.01	4.40	6.55

Notes – [a]Redshift from Steidel et al. (2003). [b]Minimum χ^2 of of SED fitting chain. [c]Probability of *accepting* the hypothesis that the data were generated by the best-fit model. [d]Stellar mass produced by median SFH. [e]Star formation rate averaged over the last 100 Myr of the median SFH. [f]Specific star formation rate averaged over the last 100 Myr of the median SFH. [g]Mass-weighted age of the median SFH.

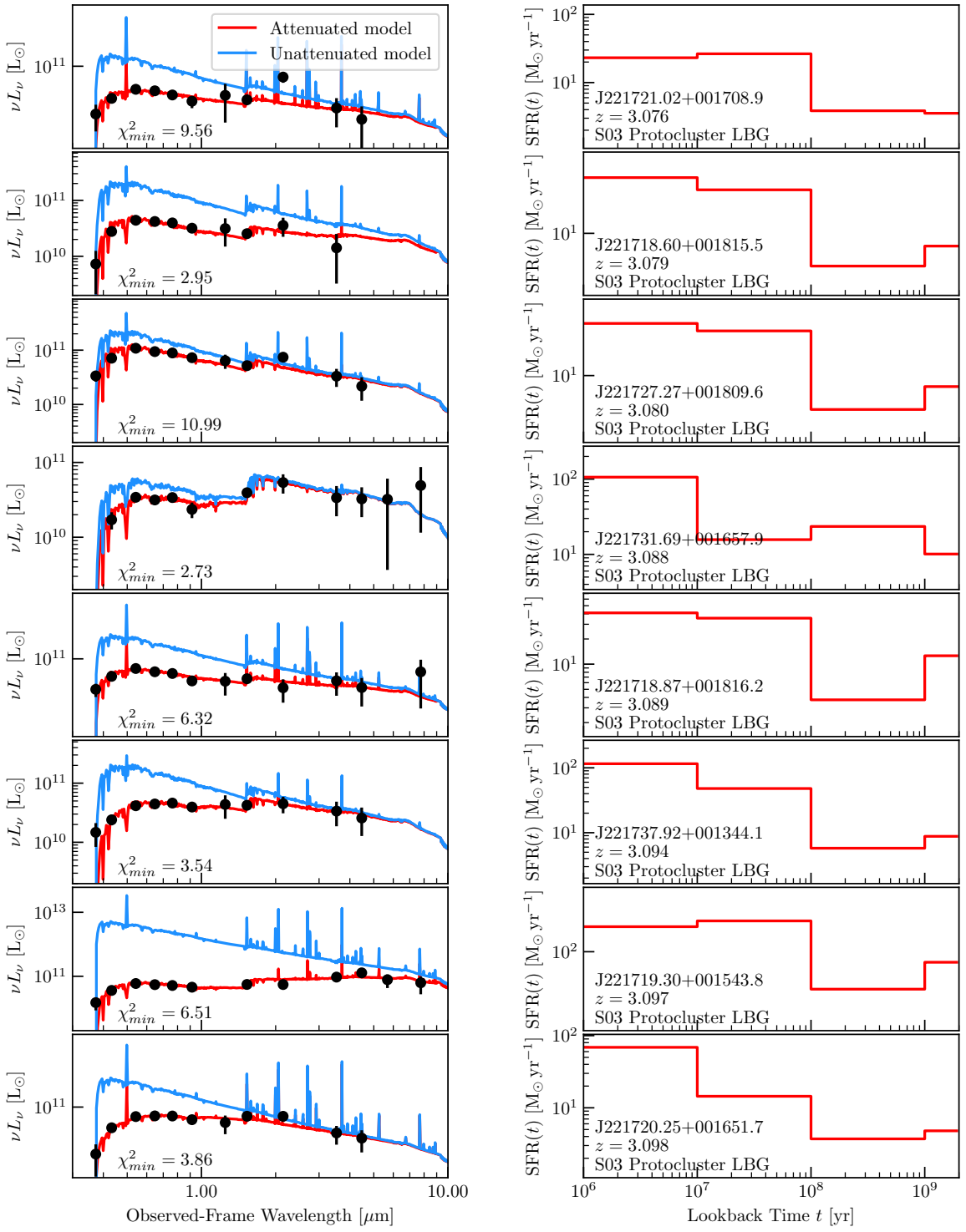

Figure 3.B.1: *Left:* We show the best-fit SED model for each of the 8 galaxies in our sample of Steidel et al. (2003) protocluster LBGs that have SED fits. *Right:* We show the median SFH for the same galaxies. For clarity we have truncated the last bin of the SFH at 2 Gyr. The shaded regions indicate the 16th to 84th percentile interval.

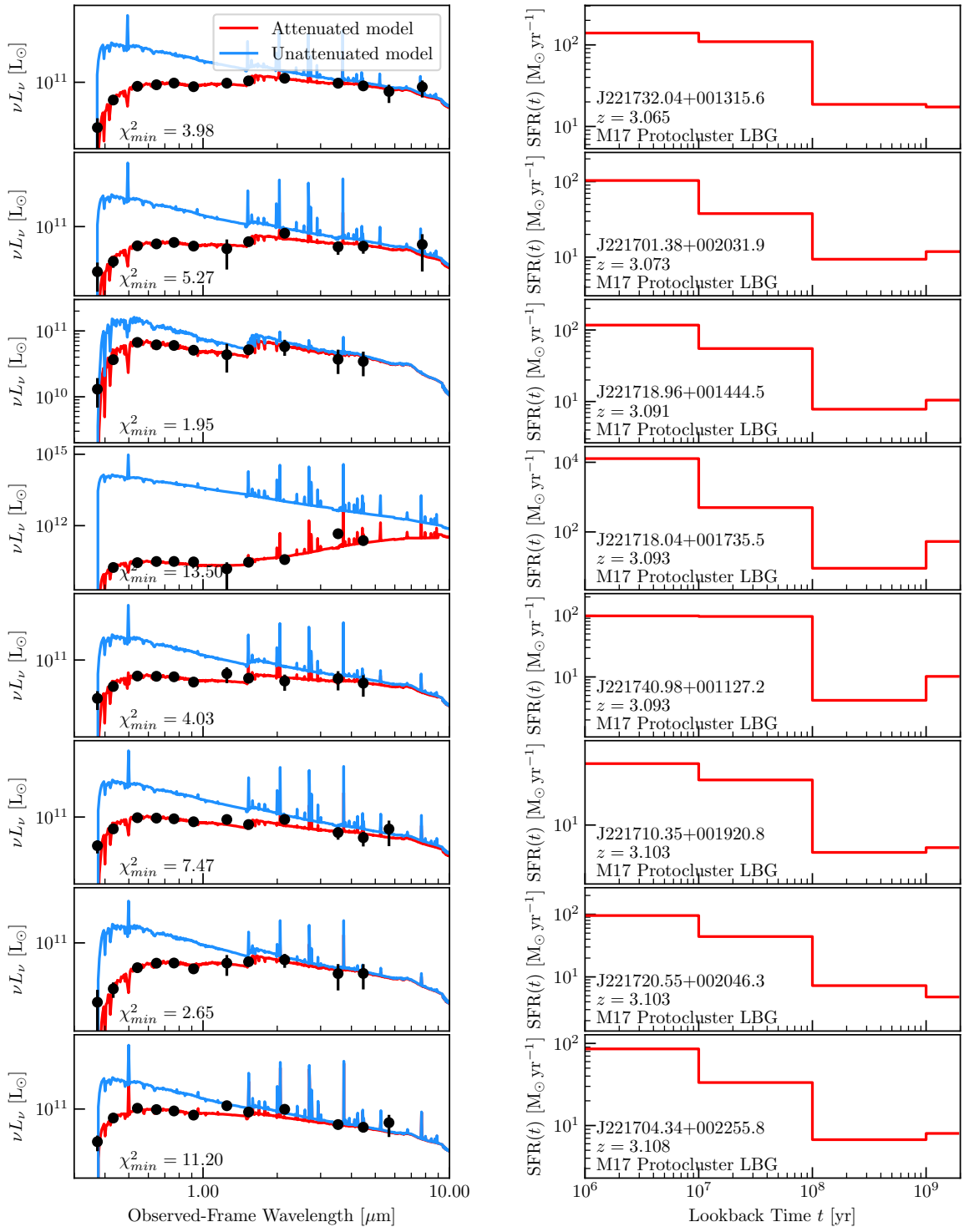

Figure 3.B.2: Same as Figure 3.B.1, showing the 8 M17 protocluster LBGs with SED fits.

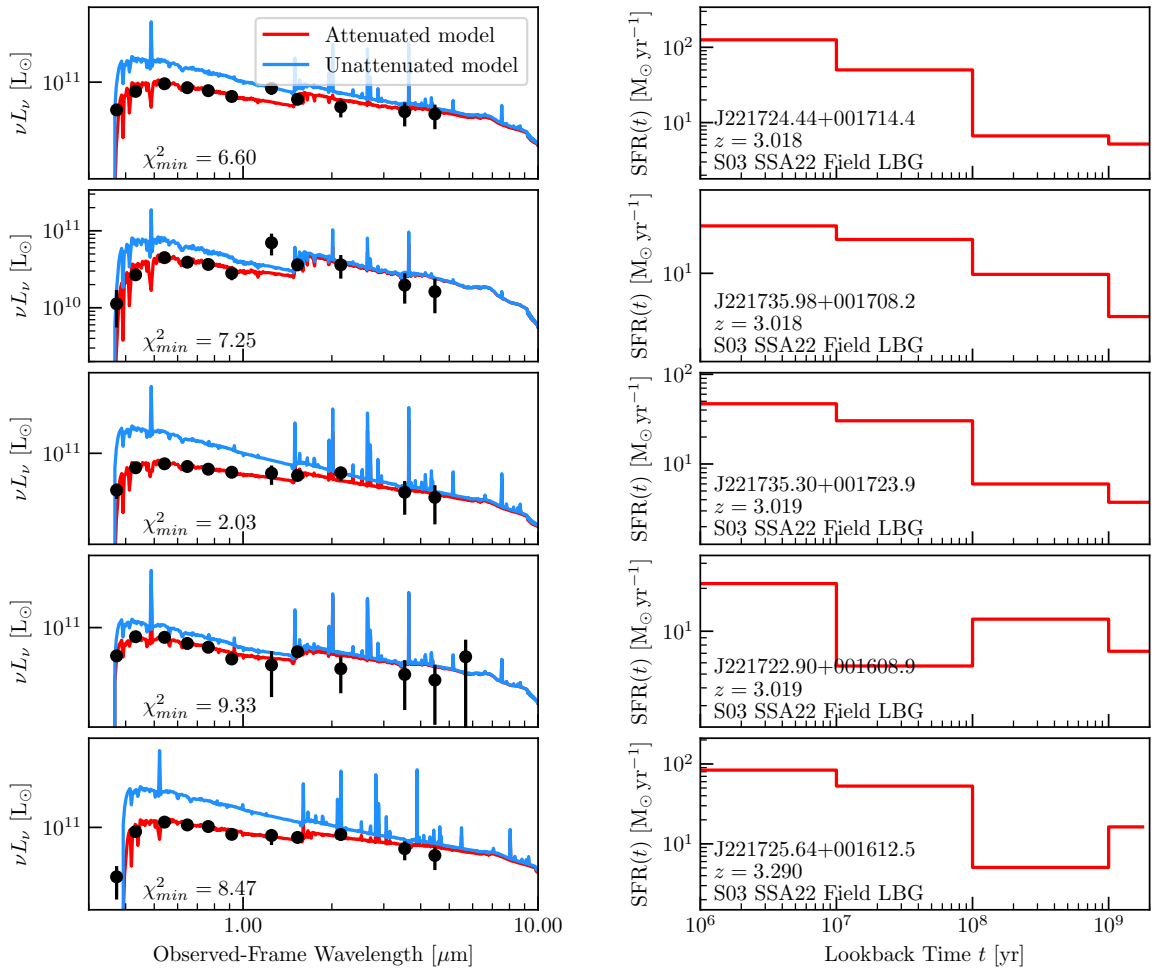

Figure 3.B.3: Same as Figure 3.B.1, showing the 5 S03 SSA22 field LBGs with SED fits.

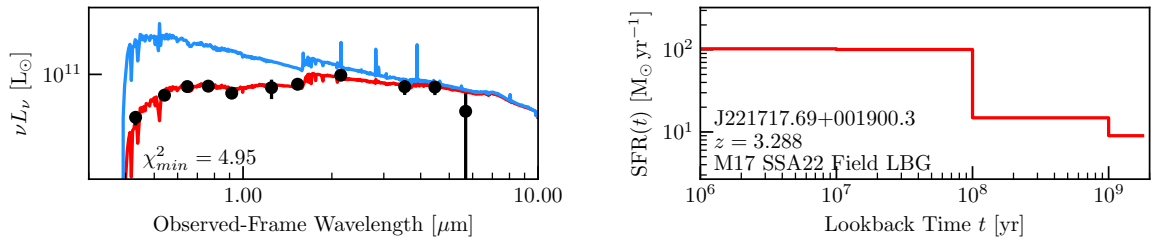

Figure 3.B.4: Same as Figure 3.B.1, showing the single M17 SSA22 field LBG with an SED fit.

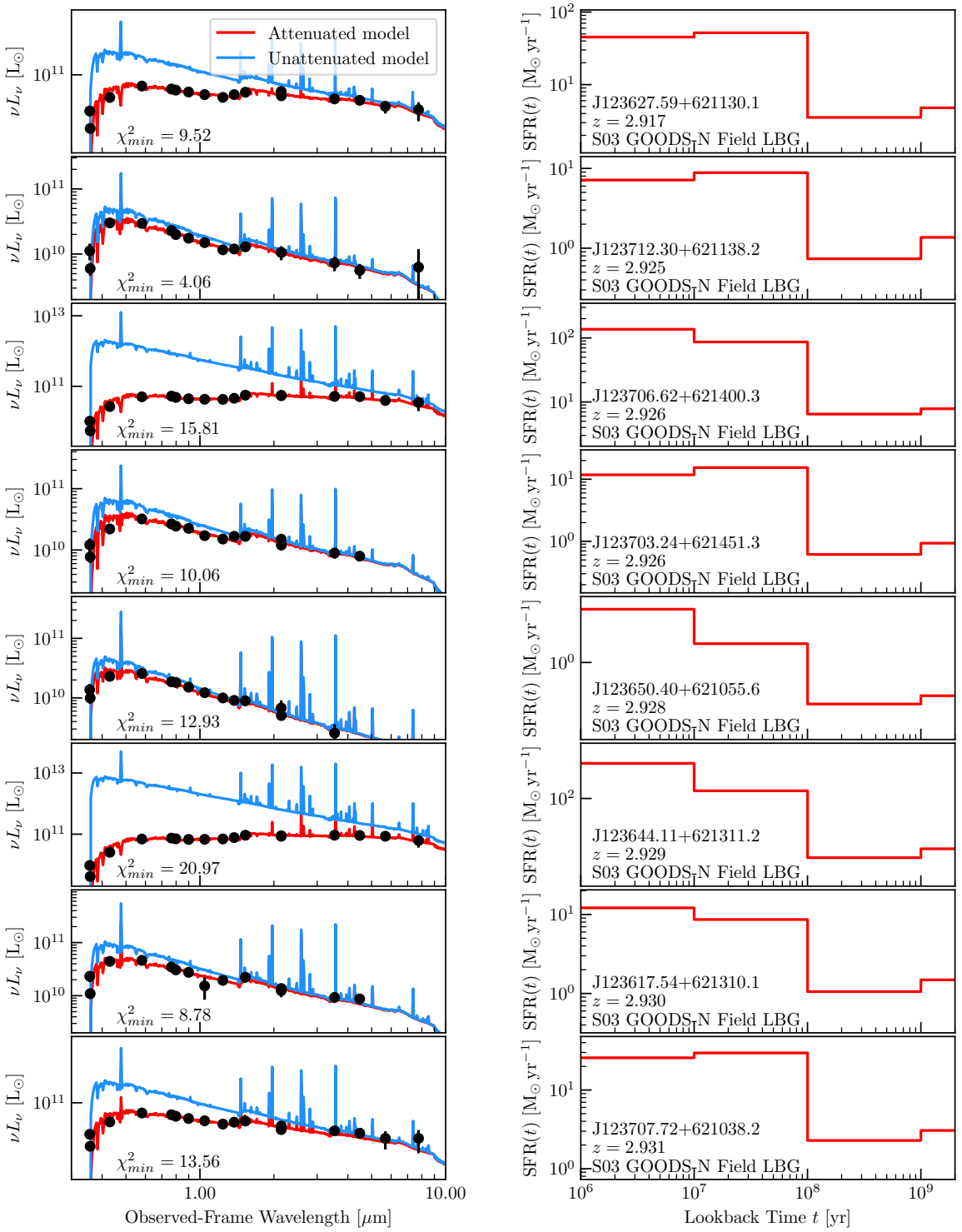

Figure 3.B.5: Same as Figure 3.B.1, showing the first 8 S03 GOODS-N field LBGs listed in Table 3.B.2. We show the SED and SFH fits for the remaining GOODS-N LBGs in the online version of this article.

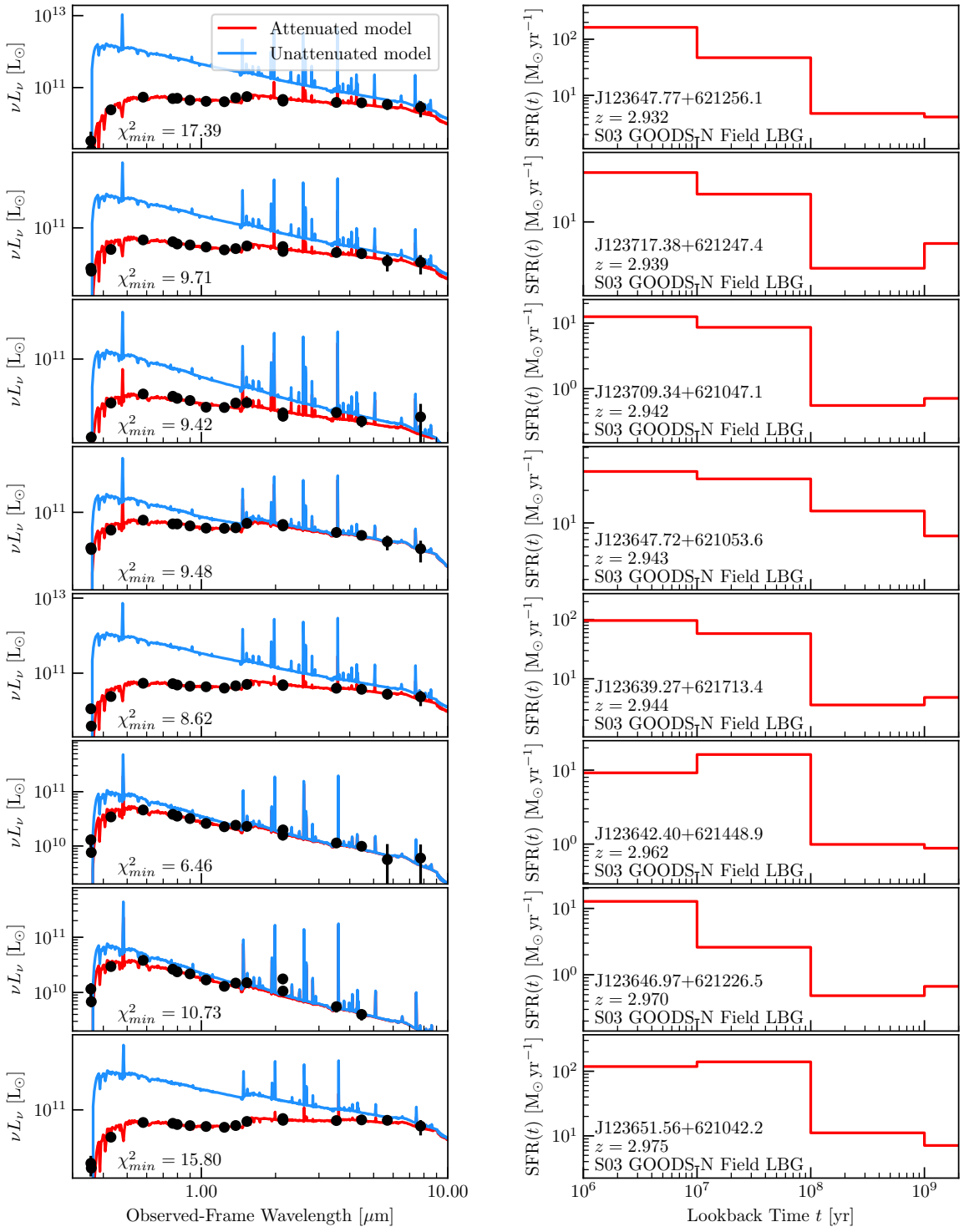

Figure 3.B.5: *Continues.*

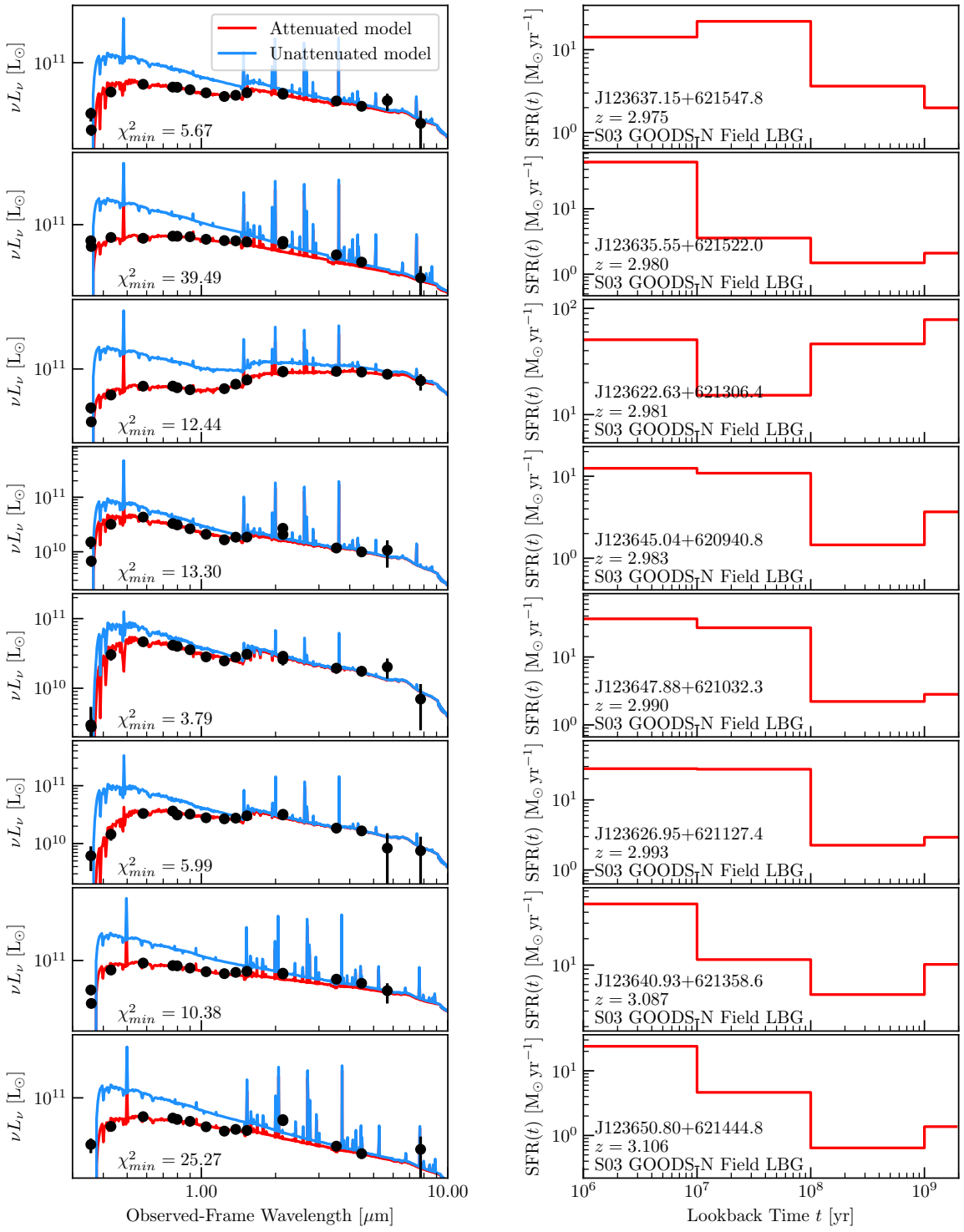

Figure 3.B.5: *Continues.*

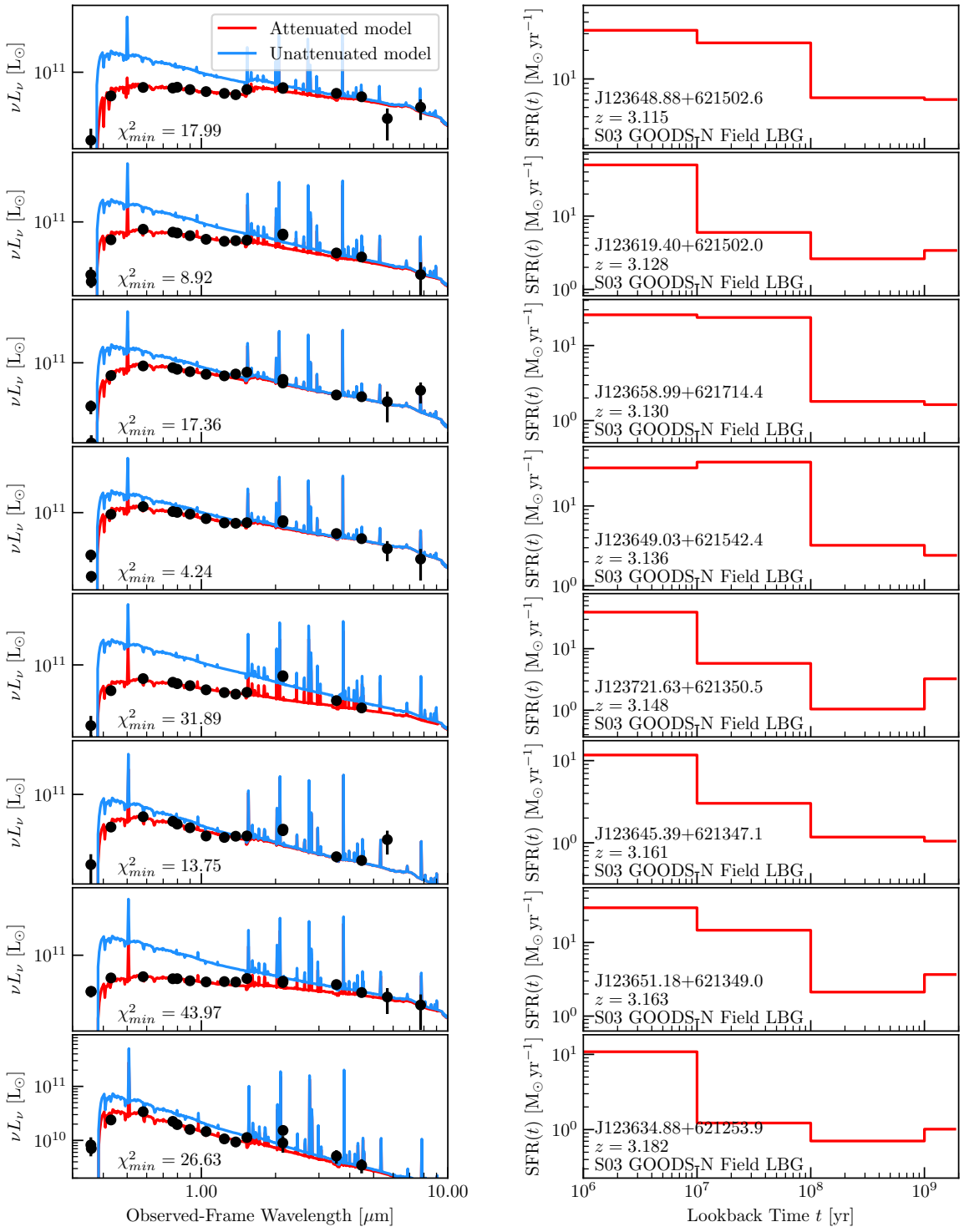

Figure 3.B.5: *Continues.*

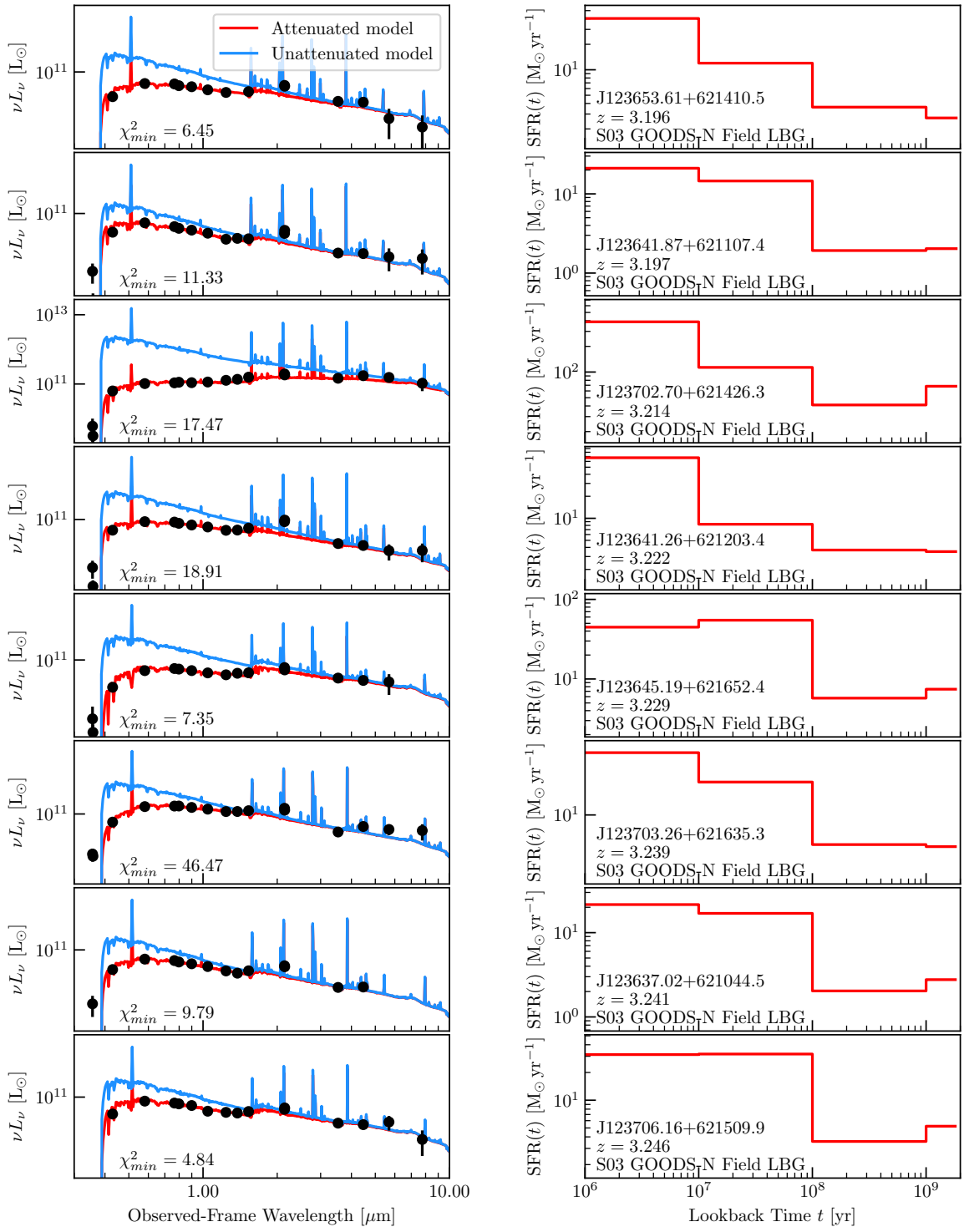

Figure 3.B.5: *Continues.*

References

Abraham, R. G., van den Bergh, S., & Nair, P. 2003, ApJ, 588, 218, doi: 10.1086/373919

Abruzzo, M. W., Narayanan, D., Davé, R., & Thompson, R. 2018, arXiv e-prints, arXiv:1803.02374. https://arxiv.org/abs/1803.02374

Alexander, D. M., Simpson, J. M., Harrison, C. M., et al. 2016, MNRAS, 461, 2944, doi: 10.1093/mnras/stw1509

Barden, M., Häussler, B., Peng, C. Y., McIntosh, D. H., & Guo, Y. 2012, MNRAS, 422, 449, doi: 10.1111/j.1365-2966.2012.20619.x

Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, ApJS, 243, 22, doi: 10.3847/1538-4365/ab23f2

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393, doi: 10.1051/aas:1996164

Boylan-Kolchin, M., Springel, V., White, S. D. M., Jenkins, A., & Lemson, G. 2009, MNRAS, 398, 1150, doi: 10.1111/j.1365-2966.2009.15191.x

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682, doi: 10.1086/308692

Capak, P. L., Riechers, D., Scoville, N. Z., et al. 2011, Nature, 470, 233, doi: 10.1038/nature09681

Ceverino, D., Klypin, A., Klimek, E. S., et al. 2014, MNRAS, 442, 1545, doi: 10.1093/mnras/stu956

Conselice, C. J., Bershady, M. A., Dickinson, M., & Papovich, C. 2003, AJ, 126, 1183, doi: 10.1086/377318

Cooke, E. A., Hatch, N. A., Muldrew, S. I., Rigby, E. E., & Kurk, J. D. 2014, MNRAS, 440, 3262, doi: 10.1093/mnras/stu522

Cowie, L., Gardner, J., Hu, E., et al. 1994, ApJ, 434, 114, doi: 10.1086/174709

Daddi, E., Valentino, F., Rich, R. M., et al. 2021, A&A, 649, A78, doi: 10.1051/0004-6361/202038700

Digby-North, J. A., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 407, 846, doi: 10.1111/j.1365-2966.2010.16977.x

Ding, X., Silverman, J., Treu, T., et al. 2020, ApJ, 888, 37, doi: 10.3847/1538-4357/ab5b90

Doore, K., Eufrasio, R. T., Lehmer, B. D., et al. 2021, arXiv e-prints, arXiv:2109.05039. https://arxiv.org/abs/2109.05039

Eufrasio, R. T., Lehmer, B. D., Zezas, A., et al. 2017, ApJ, 851, 10, doi: 10.3847/1538-4357/

aa9569

Fakhouri, O., & Ma, C.-P. 2009, MNRAS, 394, 1825, doi: 10.1111/j.1365-2966.2009.14480.x

Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9, doi: 10.1086/312838

Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950

—. 1999, arXiv e-prints, astro. https://arxiv.org/abs/astro-ph/9912179

Fitzpatrick, E. L. 1999, PASP, 111, 63, doi: 10.1086/316293

Governato, F., Baugh, C., Frenk, C., et al. 1998, Nature, 392, 359

Graham, A. W., & Driver, S. P. 2005, PASA, 22, 118–127, doi: 10.1071/AS05001

Grogin, N. A., Kocevski, D. D., Faber, S. M., et al. 2011, ApJS, 197, 35, doi: 10.1088/0067-0049/197/2/35

Harikane, Y., Ouchi, M., Ono, Y., et al. 2019, ApJ, 883, 142, doi: 10.3847/1538-4357/ab2cd5

Hatch, N. A., Kurk, J. D., Pentericci, L., et al. 2011, MNRAS, 415, 2993, doi: 10.1111/j.1365-2966.2011.18735.x

Häussler, B., McIntosh, D., Barden, M., et al. 2007, ApJS, 172, 615, doi: 10.1086/518836

Häussler, B., Bamford, S. P., Vika, M., et al. 2013, MNRAS, 430, 330, doi: 10.1093/mnras/sts633

Hayashino, T., Matsuda, Y., Tamura, H., et al. 2004, AJ, 128, 2073, doi: 10.1086/424935

Higuchi, R., Ouchi, M., Ono, Y., et al. 2019, ApJ, 879, 28, doi: 10.3847/1538-4357/ab2192

Hine, N. K., Geach, J. E., Alexander, D. M., et al. 2015, MNRAS, 455, 2363, doi: 10.1093/mnras/stv2448

Iwata, I., Inoue, A. K., Matsuda, Y., et al. 2009, ApJ, 692, 1287, doi: 10.1088/0004-637X/692/2/1287

Kartaltepe, J. S., Dickinson, M., Alexander, D. M., et al. 2012, ApJ, 757, 23, doi: 10.1088/0004-637X/757/1/23

Kato, Y., Matsuda, Y., Smail, I., et al. 2016, MNRAS, 460, 3861, doi: 10.1093/mnras/stw1237

Koekemoer, A. M., Faber, S. M., Ferguson, H. C., et al. 2011, ApJS, 197, 36, doi: 10.1088/0067-0049/197/2/36

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Kubo, M., Yamada, T., Ichikawa, T., et al. 2015a, ApJ, 799, 38, doi: 10.1088/0004-637x/799/1/38

—. 2015b, MNRAS, 455, 3333, doi: 10.1093/mnras/stv2392

Kubo, M., Yamada, T., Ichikawa, T., et al. 2017, MNRAS, 469, 2235, doi: 10.1093/mnras/stx920

Kubo, M., Uchimoto, Y. K., Yamada, T., et al. 2013, ApJ, 778, 170, doi: 10.1088/0004-637x/778/2/170

Le Fèvre, O., Cassata, P., Cucciati, O., et al. 2013, A&A, 559, A14, doi: 10.1051/0004-6361/201322179

Lehmer, B. D., Alexander, D. M., Geach, J. E., et al. 2009a, ApJ, 691, 687, doi: 10.1088/0004-637X/691/1/687

Lehmer, B. D., Alexander, D. M., Chapman, S. C., et al. 2009b, MNRAS, 400, 299, doi: 10.1111/j.1365-2966.2009.15449.x

Lehmer, B. D., Lucy, A. B., Alexander, D. M., et al. 2013, ApJ, 765, 87, doi: 10.1088/0004-637X/765/2/87

Lotz, J. M., Jonsson, P., Cox, T. J., & Primack, J. R. 2010, MNRAS, 404, 575, doi: 10.1111/j.1365-2966.2010.16268.x

Lotz, J. M., Madau, P., Giavalisco, M., Primack, J., & Ferguson, H. C. 2006, ApJ, 636, 592, doi: 10.1086/497950

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163, doi: 10.1086/421849

Lotz, J. M., Davis, M., Faber, S. M., et al. 2008, ApJ, 672, 177, doi: 10.1086/523659

Lotz, J. M., Papovich, C., Faber, S. M., et al. 2013, ApJ, 773, 154, doi: 10.1088/0004-637X/773/2/154

Madau, P., & Fragos, T. 2017, ApJ, 840, 39, doi: 10.3847/1538-4357/aa6af9

Mantha, K. B., McIntosh, D. H., Ciaschi, C. P., et al. 2019, MNRAS, 486, 2643, doi: 10.1093/mnras/stz872

Martini, P., Miller, E. D., Brodwin, M., et al. 2013, ApJ, 768, 1, doi: 10.1088/0004-637X/768/1/1

Matsuda, Y., Yamada, T., Hayashino, T., et al. 2004, AJ, 128, 569, doi: 10.1086/422020

—. 2005, ApJ, 634, L125, doi: 10.1086/499071

McAlpine, S., Harrison, C. M., Rosario, D. J., et al. 2020, MNRAS, 494, 5713, doi: 10.1093/mnras/staa1123

Micheva, G., Iwata, I., Inoue, A. K., et al. 2017, MNRAS, 465, 316, doi: 10.1093/mnras/stw2700

Miley, G. K., Overzier, R. A., Tsvetanov, Z. I., et al. 2004, Nature, 427, 47, doi: 10.1038/nature02125

Narayanan, D., Turk, M., Feldmann, R., et al. 2015, Nature, 525, 496, doi: 10.1038/nature15383

Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793, doi: 10.1051/0004-6361/200912497

Overzier, R. A. 2016, A&A Rev., 24, 14, doi: 10.1007/s00159-016-0100-3

Peng, C. Y., Ho, L. C., Impey, C. W., & Rix, H.-W. 2002, AJ, 124, 266, doi: 10.1086/340952

Romano-Díaz, E., Shlosman, I., Choi, J.-H., & Sadoun, R. 2014, ApJ, 790, L32, doi: 10.1088/2041-8205/790/2/L32

Saez, C., Lehmer, B. D., Bauer, F. E., et al. 2015, MNRAS, 450, 2615, doi: 10.1093/mnras/stv747

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103, doi: 10.1088/0004-637X/737/2/103

Snyder, G. F., Lotz, J., Moody, C., et al. 2015, MNRAS, 451, 4290, doi: 10.1093/mnras/stv1231

Steidel, C. C., Adelberger, K. L., Dickinson, M., et al. 1998, ApJ, 492, 428, doi: 10.1086/305073

Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2005, ApJ, 626, 44, doi: 10.1086/429989

Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2000, ApJ, 532, 170, doi: 10.1086/308568

—. 2003, ApJ, 592, 728, doi: 10.1086/375772

Tamura, Y., Kohno, K., Nakanishi, K., et al. 2009, Nature, 459, 61, doi: 10.1038/nature07947

Tamura, Y., Iono, D., Wilner, D. J., et al. 2010, ApJ, 724, 1270, doi: 10.1088/0004-637X/724/2/1270

Targett, T. A., Dunlop, J. S., McLure, R. J., et al. 2011, MNRAS, 412, 295, doi: 10.1111/j.1365-2966.2010.17905.x

Toshikawa, J., Kashikawa, N., Overzier, R., et al. 2016, ApJ, 826, 114, doi: 10.3847/0004-637X/826/2/114

Toshikawa, J., Uchiyama, H., Kashikawa, N., et al. 2018, PASJ, 70, S12, doi: 10.1093/pasj/psx102

Trenti, M., Bradley, L., Stiavelli, M., et al. 2012, ApJ, 746, 55, doi: 10.1088/0004-637X/746/1/55

Umehata, H., Tamura, Y., Kohno, K., et al. 2014, MNRAS, 440, 3462, doi: 10.1093/mnras/stu447

—. 2015, ApJ, 815, L8, doi: 10.1088/2041-8205/815/1/L8

—. 2017, ApJ, 835, 98, doi: 10.3847/1538-4357/835/1/98

Umehata, H., Hatsukade, B., Smail, I., et al. 2018, PASJ, 70, 65, doi: 10.1093/pasj/psy065

Umehata, H., Fumagalli, M., Smail, I., et al. 2019, Science, 366, 97, doi: 10.1126/science.aaw5949

van der Wel, A., Bell, E. F., Häussler, B., et al. 2012, ApJS, 203, 24, doi: 10.1088/0067-0049/203/2/24

Venemans, B., Kurk, J., Miley, G., et al. 2002, ApJ, 569, L11, doi: 10.1086/340563

Vika, M., Bamford, S. P., Häussler, B., et al. 2013, MNRAS, 435, 623, doi: 10.1093/mnras/stt1320

Vito, F., Brandt, W. N., Lehmer, B. D., et al. 2020, A&A, 642, A149, doi: 10.1051/0004-6361/202038848

Wang, T., Schreiber, C., Elbaz, D., et al. 2019, Nature, 572, 211, doi: 10.1038/s41586-019-1452-4

Yamada, T., Nakamura, Y., Matsuda, Y., et al. 2012, AJ, 143, 79, doi: 10.1088/0004-6256/143/4/79

Yang, G., Brandt, W. N., Darvish, B., et al. 2018, MNRAS, 480, 1022, doi: 10.1093/mnras/sty1910

Zamojski, M., Yan, L., Dasyra, K., et al. 2011, ApJ, 730, 125, doi: 10.1088/0004-637X/730/2/125

Zibetti, S., Charlot, S., & Rix, H.-W. 2009, Monthly Notices of the Royal Astronomical Society, 400, 1181, doi: 10.1111/j.1365-2966.2009.15528.x

Zolotov, A., Dekel, A., Mandelker, N., et al. 2015, MNRAS, 450, 2327, doi: 10.1093/mnras/stv740

Chapter 4

Revisiting the Properties of X-ray AGN in the SSA22 Protocluster: Normal SMBH and Host-Galaxy Growth for AGN in a z = 3.09 Overdensity

Erik B. Monson, Keith Doore, Rafael T. Eufrasio, Bret D. Lehmer, David M. Alexander, Chris M. Harrison, Mariko Kubo, Cristian Saez, and Hideki Umehata

The following chapter has been submitted to The Astrophysical Journal, and is currently in review.

Abstract

We analyze the physical properties of 8 X-ray selected active galactic nuclei (AGN) and one candidate protoquasar system (ADF22A1) in the $z = 3.09$ SSA22 protocluster by fitting their X-ray-to-IR spectral energy distributions (SEDs) using our SED fitting code, LIGHTNING[1]. We recover star formation histories (SFH) for 7 of these systems which are well-fit by composite stellar population plus AGN models. We find indications that 4/9 of the SSA22 AGN systems we study have host galaxies below the main sequence, with $\mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}} \leq -0.4$. The remaining SSA22 systems, including ADF22A1, are consistent with obscured supermassive black hole (SMBH) growth in star forming galaxies. We estimate the SMBH accretion rates and masses, and compare the properties and SFH of the 9 protocluster AGN systems with X-ray detected AGN candidates in the Chandra Deep Fields (CDF), finding that the distributions of SMBH growth rates, star formation rates, SMBH masses, and stellar masses for the protocluster AGN are consistent with field AGN. We constrain the ratio between the sample-averaged SSA22 SMBH mass and CDF SMBH mass to < 1.41. While the AGN are located near the density peaks of the protocluster, we find no statistically significant trends between the AGN or host galaxy properties and their location in the protocluster. We interpret the similarity of the protocluster and field AGN populations together with existing results as suggesting that the observed protocluster AGN fraction enhancement may be driven by the SMBH mass distribution of the entire galaxy population, rather than only the AGN population.

[1] https://www.github.com/rafaeleufrasio/lightning

4.1 Introduction

Galaxies hosting X-ray detected active galactic nuclei (AGN) are observed to be more numerous in the $z \gtrsim 1$ progenitors of galaxy clusters ("protoclusters") than in local galaxy clusters, where the frequency of X-ray AGN (AGN fraction) is suppressed compared to mean-density ("field") environments. The enhancement of AGN fraction has now been observed in clusters over $z \approx 1$–1.5 (Martini et al., 2013) and numerous protoclusters over $z \approx 2$–4 (e.g., Digby-North et al., 2010; Lehmer et al., 2013; Vito et al., 2020; Tozzi et al., 2022), including the $z = 3.1$ SSA22 protocluster, where Lehmer et al. (2009a) (hereafter L09) found a ~ 6-fold AGN fraction enhancement over the field at the same redshift. The SSA22 protocluster is an intersection of filaments ~ 60 comoving Mpc (cMpc) across at its widest extent (see Figure 4.1). It is known to contain an intense overdensity of star forming galaxies (see, e.g., Tamura et al., 2009; Kubo et al., 2013; Umehata et al., 2015) which are found, along with the AGN, to be coincident with intersections of filamentary cold gas reservoirs in the intergalactic medium (IGM) and Lyman-alpha (Lyα) nebulae (Umehata et al., 2019). SSA22, then, represents one of our best laboratories for studying how the protocluster environment may drive enhancements of AGN activity.

Observational constraints and the scatter in the AGN fraction–redshift relationship are such that the source of the apparent AGN fraction enhancement in protocluster environments remains uncertain. Simulations of SSA22-like protocluster environments by Yajima et al. (2022) suggest that the super-massive black holes (SMBH) in the protocluster grow rapidly in galaxies with $M_\star > 10^{10}$ $M_\odot$, potentially reaching masses approaching 10^9 $M_\odot$ by $z = 3$ before quasar feedback stalls the SMBH growth. The observed AGN fraction enhancement over the field could then be due to larger SMBH masses in the protocluster, as we observe a rapid period of SMBH growth with larger

average accretion rates. Alternatively, it could be caused by some modification of the AGN duty cycle in overdense, high-redshift environments, possibly due to more frequent accretion episodes caused by mergers. Monson et al. (2021) argued in favor of the former explanation based on finding larger masses among protocluster Lyman-break galaxies (LBGs) compared to field LBGs, and speculated based on non-parametric SFH fitting that the protocluster galaxy population had undergone an earlier period of growth than the field galaxies. However, this sample excluded AGN due to the difficulty of decomposing the stellar population and AGN emission for SED fitting, and hence excluded many of the most massive SSA22 galaxies. In this work, we return to the X-ray detected AGN in SSA22 and fit their SEDs with an improved version of our SED fitting code, LIGHTNING (Eufrasio et al., 2017; Doore et al., 2021), which now includes models for X-ray to IR AGN emission, allowing us to directly recover SFHs for AGN host galaxies.

The current state-of-the-art SED fitting codes (e.g., CIGALE, Boquien et al., 2019; Yang et al., 2020, 2022, and PROSPECTOR-α/PROSPECTOR, Leja et al., 2018; Johnson et al., 2021) now typically include prescriptions for fitting the continuum emission from an AGN alongside the stellar population of the host galaxy, including, in the case of CIGALE, X-ray emission. At redshifts where mid-IR observation is difficult, X-ray emission is our primary tool for constraining the AGN luminosity, a necessary step in extracting the SFH of an AGN hosting galaxy. In our updates to LIGHTNING, which we describe in detail in Section 4.3, we have added the ability to fit binned X-ray spectra with a physically-motivated X-ray AGN model, which we connect directly to a UV-to-IR AGN model. Our X-ray model allows us to estimate the SMBH masses and accretion rates of the SSA22 AGN, and to connect them to the evolution of the host galaxy.

The paper is organized as follows: Section 4.2 describes the sample selection and data preparation, Section 4.3 introduces our models and fitting procedures, Section 4.4 presents our SED fits

to the SSA22 X-ray AGN, Section 4.5 discusses the implications of our results, and Section 4.6 summarizes our work.

Where necessary we assume a flat ΛCDM cosmology with $H_0 = 70$ km s^{-1} Mpc, $\Omega_{m,0} = 0.3$, and $\Omega_{\Lambda,0} = 0.7$. In this cosmology, the age of the Universe at $z = 3.09$ is 2.04 Gyr. Coordinates are given in the J2000 epoch. Our stellar population synthesis models assume a Kroupa (2001) IMF.

4.2 Samples and Data

4.2.1 SSA22

Our sample contains the 8 X-ray detected AGN from L09 associated with the $z = 3.1$ protocluster that were first identified in a 400 ks Chandra survey of the SSA22 field (ObsIDs 8034, 8035, 8036, and 9717). Two of these AGN (J221736.54+001622.6 and J221739.08+001330.7) were first identified as LBGs and spectroscopically confirmed by Steidel et al. (2003); five (J221709.60+001800.1, J221720.24+002019.3, J221735.84+001559.1, J221739.08+001330.7, and J221759.23+001529.7) were detected as $z \approx 3.09$ Lyman-α emitters (LAEs) by Hayashino et al. (2004) with spectroscopic follow-ups from Matsuda et al. (2005). Of the remaining two systems, J221716.16+001745.8 was spectroscopically confirmed as a protocluster member by Saez et al. (2015), and J221732.00+001655.6 by Kubo et al. (2015). The proto-quasar system ADF22A1, identified as an AzTEC 1.1 mm source in Tamura et al. (2010), is also X-ray detected. We include ADF22A1 in our sample as it possibly represents an early phase of AGN growth in the protocluster, and may (given its projected mass and location in the protocluster) evolve into one of the largest galaxies in the protocluster by $z = 0$. The basic observational properties of these 9 sources are given in Table 4.1. In Figure 4.2, we show F160W cutouts of the X-ray AGN and ADF22A1,

Table 4.1: Observational properties of the SSA22 sample.

ID	Alt. Name(s)[a]	Redshift[b]	$\log L_X$[c] ($\mathrm{erg\ s^{-1}}$)	*Chandra* R.A.[d] (Deg.)	*Chandra* Dec.[d] (Deg.)	ALMA 870μm Flux[e] (mJy)	ALMA R.A.[f] (Deg.)	ALMA Dec.[f] (Deg.)	Association
J221736.54+001622.6	AGN1, ADF22A9	3.084[(1)]	44.17	334.40225	0.2729	1.84 ± 0.21	334.40233	0.2729	–
J221739.08+001330.7	AGN2	3.091[(1)]	43.91	334.41279	0.2252	0.91 ± 0.10[*]	334.41187	0.2260	LAB2
J221709.60+001800.1	AGN3	3.106[(2)]	44.03	334.29025	0.2999	< 1.04	–	–	–
J221720.24+002019.3	AGN4	3.105[(2)]	44.61	334.33442	0.3386	< 1.10	–	–	–
J221735.84+001559.1	AGN5, ADF22A6	3.094[(2)]	44.27	334.39933	0.2664	2.96 ± 0.29	334.39925	0.2664	LAB14
J221759.23+001529.7	AGN6	3.096[(2)]	44.33	334.49675	0.2580	< 1.21 (1.38 ± 0.29)	334.49725	0.2589	LAB3
J221716.16+001745.8	AGN7, ADF22B5	3.098[(3)]	43.98	334.31721	0.2958	< 1.50 (2.25 ± 0.34)	334.31704	0.2964	–
J221732.00+001655.6	AGN8, ADF22A12	3.091[(4)]	43.96	334.38329	0.2821	1.58 ± 0.35	334.38308	0.2822	LAB12
J221732.41+001743.8	ADF22A1	3.092[(5)]	44.33	334.38508	0.2955	–	–	–	AzTEC1

Notes – [a]AGN are numbered as they appear in Lehmer et al. (2009a) and Alexander et al. (2016). [b]Spectroscopic redshift. Sources are as follows: 1 – Steidel et al. (2003); 2 – Matsuda et al. (2005); 3 – Saez et al. (2015); 4 – Kubo et al. (2015); 5 – Umehata et al. (2017). [c]Rest-frame $2-32$ keV luminosity computed from the Lehmer et al. (2009b) $0.5-8.0$ keV flux, assuming the inferred Γ from their catalog. [d]Position of the *Chandra* detection produced by ACISEXTRACT. [e]See Alexander et al. (2016), Tables 1 and 2. Fluxes for offset detections are shown in parentheses; upper limits here are given at 4.5σ. [f]For detected sources we give the position; for non-detected sources with an offset detection, we give the position of the offset detection. [*]This is the ALMA Band 7 flux from Ao et al. (2017); this AGN is non-detected at the flux limit of the Alexander et al. (2016) observations.

along with the positions of the corresponding X-ray detections from Lehmer et al. (2009b).

For this work the X-ray data were re-reduced using CIAO v4.13 and CALDB v4.9.5[2], following the procedures described in Lehmer et al. (2017). Source photometry was extracted using the ACIS EXTRACT (AE) pipeline v2020dec17 (Broos et al., 2010)[3], following the procedures outlined in Lehmer et al. (2019). For our SED fits (see Section 4.3), we grouped the X-ray photometry into fixed-width bins from $0.5-2.0$, $2.0-4.0$, and $4.0-7.0$ keV. We chose these bins as they were well detected for the majority (6/9) of the sample. In cases where the binomial no-source probability for a bin was < 0.05, we combined it with the next highest-energy bin, such that the remaining three galaxies have two bins, from $0.5-4.0$ and $4.0-7.0$ keV.

We use UV–NIR photometry from Kubo et al. (2013) in the $u^*BVR_cizJHK_s$ bands. These photometry were extracted using a $2''$ diameter circular aperture from images smoothed to a common $1''$ FWHM Gaussian PSF. We also use IRAC channel 3.6 μm, 4.5 μm, 5.6 μm, and 8.0 μm photometry from Kubo et al. (2013). The IRAC fluxes were not extracted from images smoothed to a common PSF, but were aperture corrected to the same encircled energy as the ground-based photometry. For ADF22A1, we use an additional MIPS 24 μm measurement calculated by Tamura et al. (2010). We use the *HST* WFC3 F160W fluxes measured in Monson et al. (2021), which were also extracted in a $2''$ diameter circular aperture from images smoothed to a common $1''$ FWHM Gaussian PSF. F160W measurements were available for each of the L09 X-ray AGN, but not ADF22A1, which is non-detected at the S/N threshold of the Monson et al. (2021) catalog.

All of our SSA22 sources of interest have been observed at mm/sub-mm wavelengths. Four of the L09 AGN and ADF22A1 are detected in ALMA Band 6 (~ 1.1 mm) observations from the

[2]https://cxc.cfa.harvard.edu/ciao/download/
[3]http://personal.psu.edu/psb6/TARA/code/

ALMA Deep Field in SSA22 (ADF22) survey (Umehata et al., 2014, 2015, 2017). Alexander et al. (2016) targeted the 8 X-ray AGN from L09 with ALMA Band 7 (~ 870 μm), in which 3 of the AGN are detected and the remaining 5 are given 4.5σ upper limits (see Table 4.1). We use these fluxes and upper limits in our SED fits to those sources. Three of the ALMA Band 7-undetected SSA22 AGN (J221739.08+001330.7, J221759.23+001529.7, and J221716.16+001745.8) have associated, offset ALMA detections from Alexander et al. (2016). In subsequent ALMA Band 7 observations by Ao et al. (2017), a source coincident with the J221739.08+001330.7 X-ray source was detected, indicating that the low-S/N offset source may be spurious. We use the Ao et al. (2017) flux in our fit to J221739.08+001330.7. For the other two AGN with offset detections, we do not account for the offset source in our fits, instead adopting the upper limit associated with the *Chandra* position. However, in the case of J221716.16+001745.8 we find that the offset detection may be closely associated with the AGN (see Section 4.3.4 and Section 4.A).

We identified photometric issues with two of the AGN, J221720.24+002019.3 and J2217529.23+001529.7. J221720.24+002019.3 has J, H, and K_s measurements inconsistent with the rest of the observed SED, producing a very flat spectral slope that we cannot reproduce even with our AGN models. These inconsistencies could be caused by line emission (including possibly broad lines; this source is one of three reported Type 1 AGN in the SSA22 sample), but we also find it likely that these measurements are affected by a nearby projected companion, which is inside the photometric aperture for J221720.24+002019.3. The redshift of the projected companion is unknown. We find also that the ground-based H-band measurement is larger than the F160W measurement from Monson et al. (2021) by a factor of ~ 4. The F160W photometry was measured from an image smoothed to the same FWHM and in the same photometric aperture as the ground-based H-band measurement, but the companion may be blended with the AGN in the original,

lower-resolution ground-based image, accounting for the difference in flux.

J2217529.23+001529.7 has *Spitzer* IRAC 3.6 μm and 4.5 μm measurements that are also inconsistent with the remainder of the observed SED. The IRAC 1 and 2 fluxes are very bright; the IRAC 3.6 μm flux is the largest in our sample. However, the IRAC colors are inconsistent with AGN emission. The decrement we see between the IRAC 3.6 μm and 4.5 μm fluxes is more consistent with stellar emission, and this AGN also has a bright foreground star within $4''$. We suspect that this foreground star may be blended with the AGN in the IRAC data, given the large IRAC PSF.

Out of an abundance of caution, we treat the measurements listed above for J221720.24+002019.3 and J2217529.23+001529.7 as upper limits in our fits.

We correct all of the fluxes for Galactic extinction using the standard Fitzpatrick (1999) Milky Way extinction curve and the Galactic dust extinction map from Schlafly & Finkbeiner (2011), retrieved from the IRSA DUST web application[4]. We add instrumental calibration uncertainties to the measured uncertainties for each band, along with a flat 10% model uncertainty in each band. We summarize the bands used for the SSA22 AGN and the fractional calibration uncertainties adopted in Table 4.2.

4.2.2 Chandra Deep Fields

In order to generate a comparison sample for the SSA22 AGN, we utilize the 2 Ms Chandra Deep Field North (CDF-N) catalog (Xue et al., 2016) and 7 Ms Chandra Deep Field South (CDF-S) catalog (Luo et al., 2017). To select candidate AGN in the $z \sim 3$ field, while accounting for the differing X-ray depth between SSA22 and the *Chandra* deep fields, we converted the rest-frame

[4]https://irsa.ipac.caltech.edu/applications/DUST/

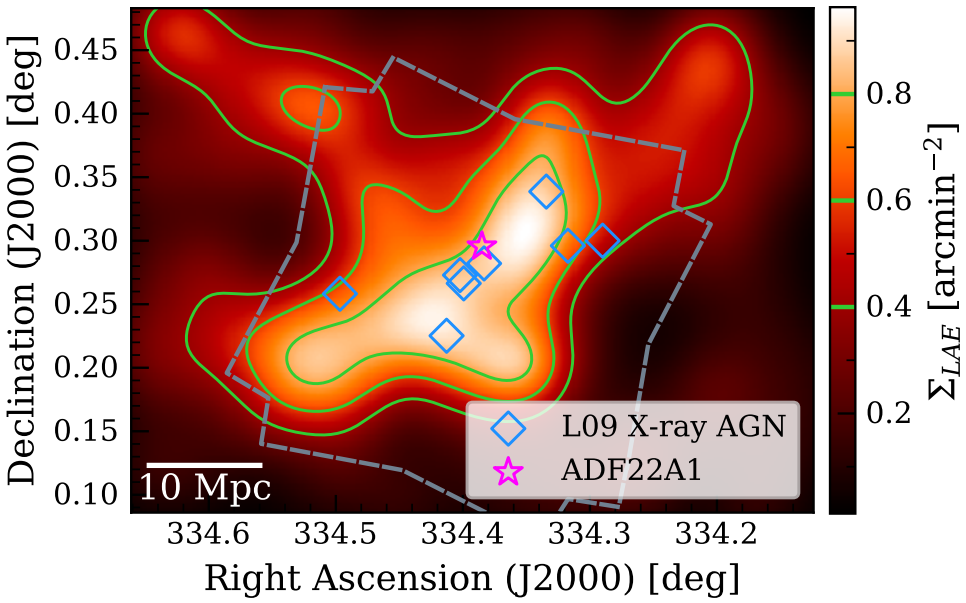

Figure 4.1: We show the projected density map of $z = 3.09$ Lyman-α emitter (LAE) candidates from Hayashino et al. (2004) with the locations of the X-ray AGN shown as blue diamonds and the location of ADF22A1 shown as a magenta star. Density contours are also shown at 0.4, 0.6, and 0.8 LAEs arcmin^{-2}, and the approximate coverage of the *Chandra* field in SSA22 is shown as a dashed gray line. Three of the X-ray AGN are located within ~ 5 cMpc of ADF22A1, which is believed to be near the density peak of the protocluster.

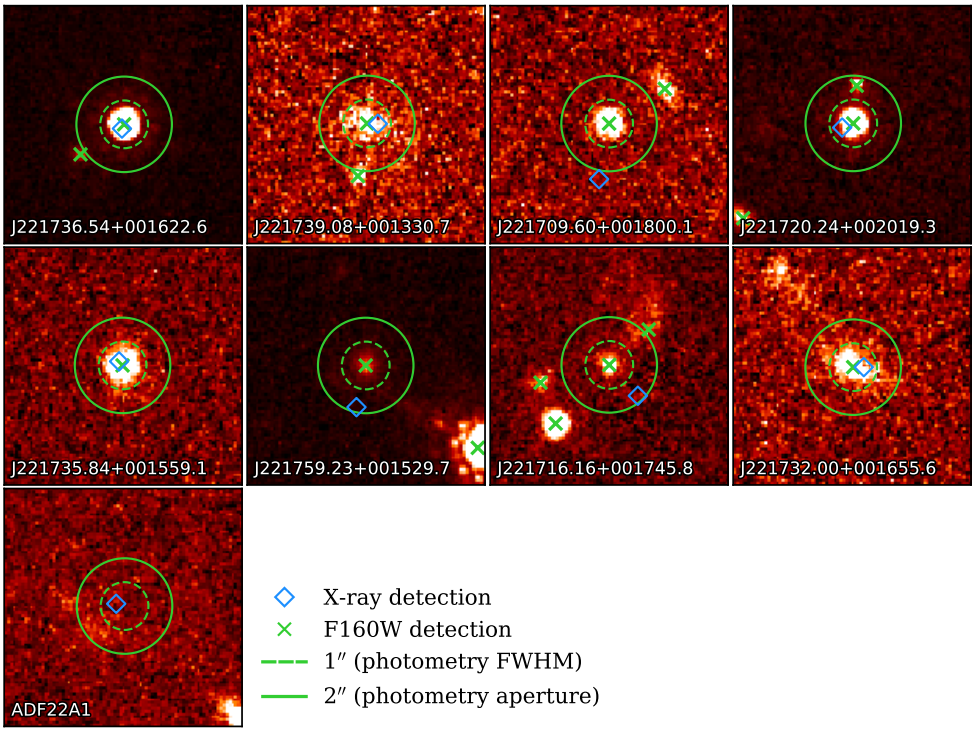

Figure 4.2: We show F160W cutouts of the 8 X-ray detected AGN and ADF22A1. The pixel scale of the images as shown is $0.06''$ pix^{-1}, and the cutouts are $5''$ on each side. The green dashed and solid circles respectively show the smoothed $1''$ FWHM and $2''$ aperture used to measure the $u^*BVR_cizJHK_s$ photometry from Kubo et al. (2013) and F160W photometry from Monson et al. (2021). The *Chandra* detections from L09 are shown as cyan diamonds, and 3σ F160W detections are shown as green crosses. J221720.24+002019.3 has a possible companion within the photometric aperture, which we take as a possible explanation for issues seen with its photometry (see Section 4.2.1). Likewise, we see that J2217529.23+001529.7 has a bright foreground star $< 4''$ from the main galaxy, which may be blended with the main galaxy in the *Spitzer* IRAC photometry.

Table 4.2: Bandpasses and calibration uncertainties used for fits to SSA22 AGN.

Observatory/Instrument	Bandpass	$\sigma_{\rm cal}^{f}$ [a]
SUBARU/SuprimeCam	u^*	0.05
...	B	0.05
...	V	0.05
...	R_c	0.05
...	i'	0.05
...	z'	0.05
SUBARU/MOIRCS	J	0.05
...	H	0.05
...	K_s	0.05
HST/WFC3	F160W	0.02
Spitzer/IRAC	3.6 μm	0.05
...	4.5 μm	0.05
...	5.6 μm	0.05
...	8.0 μm	0.05
Spitzer/MIPS[*]	24 μm	0.05
ALMA	Band 7	0.05
...	Band 6	0.05

Notes – [a]Calibration uncertainty as a fraction of flux: $\sigma_{\rm total} = \sqrt{\sigma^2 + (\sigma_{\rm cal}^{f} f)^2}$. [*]ADF22A1 only.

$0.5-7.0$ keV L_X in the deep field catalogs to a rest-frame $10-30$ keV luminosity, assuming a photon index $\Gamma = 1.8$. We then selected sources with $L_{10-30\,\mathrm{keV}} \geq 2\times10^{43}$ erg s^{-1}, the three-count detection limit from the SSA22 X-ray survey of Lehmer et al. (2009b). We cross matched these candidate AGN with the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS) photometric catalogs in the GOODS-N (Barro et al., 2019) and GOODS-S (Guo et al., 2013) fields to retrieve optical-NIR photometry using a $1''$ match radius. We find 97 candidate AGN with unique photometric matches in GOODS-N, and 108 in GOODS-S.

We also included *Herschel* PACS and SPIRE IR photometry from the CANDELS catalogs (Barro et al., 2019), where available. At $z\sim3$, blending of far-IR detections with *Herschel* PACS and SPIRE is a significant concern. Barro et al. (2019) adopt a prior-based extraction method for their FIR fluxes to mitigate this. To be conservative, we additionally searched for *Spitzer* MIPS 24 μm detections within a 1 FWHM diameter aperture around each *Herschel* detection, in each band (the diameters are respectively $7.0''$, $11.2''$, and $18.0''$ for PACS 100 μm, PACS 160 μm, and SPIRE 250 μm). We then calculated the total 24 μm flux within the 1 FWHM aperture and the fraction of the total 24 μm flux attributed to the primary source most closely matched to the *Herschel* detection. If this fraction is less than 0.50, we suspect that the *Herschel* photometry may be influenced by blending, and exclude it from our fits. This process has the side effect of requiring a MIPS 24 μm detection, which leaves 75 AGN candidates in the CDF-N and 77 in the CDF-S. We removed galaxies from the sample for which all three *Herschel* bands are blended, as this gives us no constraint on the dust emission from the galaxy. This final requirement excludes 1 galaxy from the CDF-N. Our sample thus includes 74 candidate AGN in the CDF-N and 77 in the CDF-S.

In the interest of not biasing our sample toward only high-SFR galaxies, we do not require *Herschel* detections for CDF AGN candidates. For AGN with no *Herschel* far-IR detections (i.e.,

no non-blended detections at observed-frame wavelengths ≥ 100 μm), we calculated PACS upper limits from the PEP/GOODS-*Herschel* DR1 (Magnelli et al., 2013) error maps[5], using the optimum circular apertures and aperture corrections from Pérez-González et al. (2010). We calculated our own upper limits rather than using published limiting fluxes for the fields due primarily to the variation in PACS depth across the GOODS-S field, which causes the published upper limits to be too restrictive for our purposes for galaxies outside the central, ultra-deep region of the field.

To fit the SEDs of the CDF AGN, we require redshifts. Of the CDF-N (CDF-S) AGN, 47/74 (52/77) have reliable spectroscopic redshifts (quality flag ≤ 2 in the CANDELS catalogs). To improve the redshift coverage of our sample we use the photometric redshifts reported in the Guo et al. (2013) and Barro et al. (2019) catalogs where spectroscopic redshifts are unavailable or reported to be unreliable. The Barro et al. (2019) photometric redshifts have three quality tiers, depending on whether they were derived from the broad-band CANDELS photometry (tier 3), CANDELS+SHARDS photometry (tier 2), or CANDELS+SHARDS and grism data (tier 1). Of the 27 CDF-N galaxies in our sample for which we use photometric redshifts, 7 galaxies are at tier 3, 4 at tier 2, and 16 at tier 1. The Guo et al. (2013) photometric redshifts are all derived from broad-band photometry, equivalent to the tier 3 redshifts for CDF-N. We estimate the accuracy of these redshifts by computing $\Delta z/(1+z_{\rm spec}) = (z_{\rm phot} - z_{\rm spec})/(1+z_{\rm spec})$ for the AGN candidates with spectroscopic redshifts. We find that the outlier fraction η of AGN with $\Delta z/(1+z_{\rm spec}) > 0.15$ is $\eta = 12.8\%$ in the CDF-N sample and $\eta = 19.2\%$ in the CDF-S sample, a factor of 3–4 larger than the outlier fractions reported for the full catalogs in Barro et al. (2019) and Guo et al. (2013), respectively. However, the outlier fractions decrease as a function of redshift, with $\eta = 0$ at $z > 2$ for both CDF-N and CDF-S samples. We show the distribution of redshifts for each CDF sample

[5]Retrieved from https://www.mpe.mpg.de/ir/Research/PEP/DR1

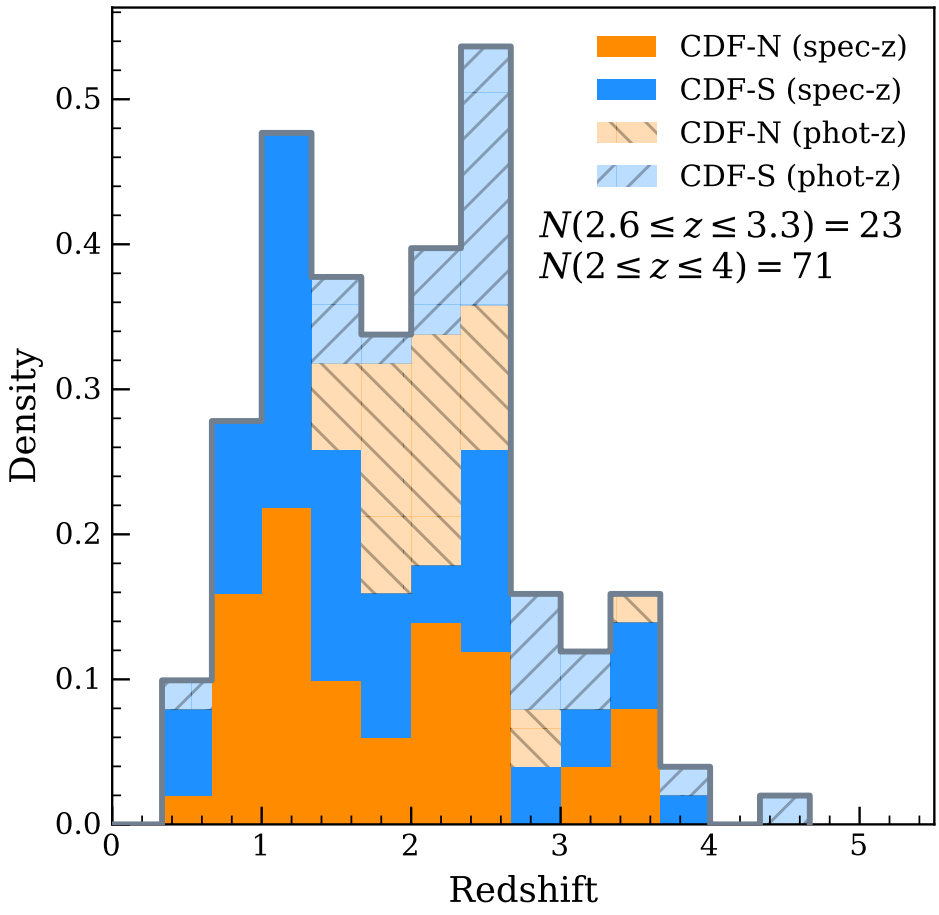

Figure 4.3: We show a stacked histogram of the CDF sample redshifts: solid bars represent spectroscopic redshifts, while desaturated and hatched bars represent photometric redshifts. Including photometric redshifts in our sample improves our coverage at high redshift, crucial to constructing a comparison for the $z = 3.09$ protocluster.

Table 4.3: Bandpasses and calibration uncertainties used for fits to CDF AGN candidates.

Sample	Observatory/Instrument	Bandpass	$\sigma_{\rm cal}^{f}$ [a]	Sample	Observatory/Instrument	Bandpass	$\sigma_{\rm cal}^{f}$
CDF-N	LBT/LBC	U	0.10	CDF-S	VLT/VIMOS	U	0.10
	HST/ACS	F435W	0.02		HST/ACS	F435W	0.02
	...	F606W	0.02		...	F606W	0.02
	...	F775W	0.02		...	F775W	0.02
	...	F814W	0.02		...	F814W	0.02
	...	F850LP	0.02		...	F850LP	0.02
	HST/WFC3 IR	F105W	0.02		HST/WFC3 IR	F098M	0.02
	...	F125W	0.02		...	F105W	0.02
	...	F140W	0.02		...	F125W	0.02
	...	F160W	0.02		...	F160W	0.02
	CFHT/WIRCam	K_s	0.05		VLT/HAWK-I	K_s	0.05
	Spitzer/IRAC	3.6 μm	0.05		*Spitzer*/IRAC	3.6 μm	0.05
	...	4.5 μm	0.05		...	4.5 μm	0.05
	...	5.6 μm	0.05		...	5.6 μm	0.05
	...	8.0 μm	0.05		...	8.0 μm	0.05
	Spitzer/MIPS	24 μm	0.05		*Spitzer*/MIPS	24 μm	0.05
	...	70 μm	0.10		...	70 μm	0.10
	Herschel/PACS	100 μm	0.05		*Herschel*/PACS	100 μm	0.05
	...	160 μm	0.05		...	160 μm	0.05
	Herschel/SPIRE	250 μm	0.15		*Herschel*/SPIRE	250 μm	0.15

Notes – [a]Calibration uncertainty as a fraction of flux: $\sigma_{\rm total} = \sqrt{\sigma^2 + (\sigma_{\rm cal}^{f} f)^2}$.

in Figure 4.3. The full sample of CDF-N AGN ranges from $z = 0.51$–3.65, while the CDF-S AGN range from $z = 0.45$–4.52. Below, we focus on galaxies with $z \geq 2$ when making comparisons to the protocluster; at these high redshifts the spectroscopic redshift fractions are lower: 19/35 for CDF-N and 17/37 for CDF-S. In what follows we distinguish between galaxies with spectroscopic and photometric redshifts where relevant.

Hereafter we refer to these samples as our CDF-N and CDF-S samples. As with SSA22 (Section 4.2.1), we added calibration uncertainties and a 10% model uncertainty to the measured uncertainties for each band. We summarize the calibration uncertainties for each field and bandpass in Table 4.3.

4.3 SED Fitting

4.3.1 Stellar Population and Dust Modeling

The host galaxy stellar populations are modeled using piecewise constant (also called "non-parametric") SFHs, with constant SFR in each of a number of stellar age bins. For the majority of our sample we use 4 SFH bins, spanning 0–10 Myr, 10–100 Myr, 100 Myr–1 Gyr, and 5 Gyr–$t_{\rm age}(z)$, where $t_{\rm age}(z)$ is the age of the Universe at the redshift z of the galaxy. For galaxies with $t_{\rm age}(z) > 5$ Gyr ($z < 1.17$ in our assumed cosmology), we use 5 bins, spanning 0–10 Myr, 10–100 Myr, 100 Myr–1 Gyr, 1–5 Gyr, and 5 Gyr–$t_{\rm age}(z)$. Our single age stellar population models are generated with PÉGASE (Fioc & Rocca-Volmerange, 1997, 1999), assuming solar metallicity, $Z = 0.020$, and a Kroupa (2001) IMF. These stellar populations also include nebular emission associated with star formation (see section 2.4 of Fioc & Rocca-Volmerange, 1997).

Dust attenuation is modeled with the Noll et al. (2009) modified Calzetti et al. (2000) attenuation curve, which includes a UV bump and variable UV attenuation curve slope. For the SSA22 sample we fix the UV slope to match the featureless Calzetti et al. (2000) law, while for the CDF samples we allow the deviation δ in the UV slope to vary, as the presence of IR data at the peak of the warm dust emission and the relatively deep optical data allow us to constrain δ.

Dust emission is modeled using the Draine & Li (2007) templates. We use the "restricted" version of the model set, as recommended by Draine & Li (2007), with the radiation intensity distribution slope fixed at $\alpha = 2$ and the maximum intensity fixed at $U_{\rm max} = 3 \times 10^5$. The minimum intensity, $U_{\rm min}$, cannot be well-constrained by our limited IR data for the SSA22 sample, so we (arbitrarily) fix $U_{\rm min} = 10$, finding that this choice also results in acceptable fits for the CDF AGN. The mass fraction of polycyclic aromatic hydrocarbons (PAH) in the dust mixture is also fixed for

all of our samples at $q_{\rm PAH} = 0.0047$, which is the minimum for the LIGHTNING dust emission implementation, as we do not expect strong PAH emission from the high-redshift galaxies that we are most interested in. We assume energy balance, such that the normalization of the dust emission is set by the UV-optical attenuation of the stellar population and AGN.

4.3.2 UV-IR AGN Modeling

Inspired by the methods adopted in CIGALE (Yang et al., 2020, 2022), we adopt the SKIRTOR UV-IR AGN templates (Stalevski et al., 2016) in our modeling. These models include a broken power law model for the accretion disk emission, which is reprocessed by a two-phase, clumpy torus of dust surrounding the accretion disk. For the accretion disk component, we use the default SKIRTOR broken power law, where

$$\lambda L_\lambda \propto \begin{cases} \lambda^{1.2} & 0.001\ \mu\mathrm{m} \leq \lambda \leq 0.01\ \mu\mathrm{m} \\ \lambda^{0} & 0.01\ \mu\mathrm{m} < \lambda \leq 0.1\ \mu\mathrm{m} \\ \lambda^{-0.5} & 0.1\ \mu\mathrm{m} < \lambda \leq 5\ \mu\mathrm{m} \\ \lambda^{-3} & 5\ \mu\mathrm{m} < \lambda \leq 50\mu\mathrm{m} \end{cases} . \tag{4.1}$$

We use a slice of the model set with the geometry and structure of the torus fixed. We adopt a torus opening angle 40° (corresponding to a covering factor of $\sin 40^\circ \approx 0.64$) based on the findings of Stalevski et al. (2016). The amount of dust in the torus is also fixed: in the SKIRTOR models this is controlled by the parameter $\tau_{9.7}$, the edge-on optical depth of the torus at 9.7 μm (the wavelength of the commonly-seen silicate absorption feature in the spectra of AGN). We adopt $\tau_{9.7} = 7$; we lack the IR data to properly constrain this silicate feature, so we choose the midpoint of the gridded values in the SKIRTOR models. To allow treatment of the viewing angle i as a

continuous parameter (the original models are gridded in 10 degree increments), we interpolate between models in $\cos i$-space.

The UV-optical component of the AGN is attenuated by the same attenuation curve as the stellar population, and we apply energy balance to model heating of the ISM dust by the AGN, such that the power attenuated from the AGN spectrum by ISM dust (integrated over all lines of sight) is added to the luminosity of the dust model.

Our AGN model does not include line emission, which is a strong and ubiquitous feature in the UV-to-IR spectra of AGN. In early tests we found clear discrepancies for some AGN candidates in the CDF between bands probing the Lyman-α line and our models, where the observed flux was significantly larger than the model was capable of producing, while the neighboring bands probing continuum emission were well-fit. Our stellar population models include Lyα, but could not produce the observed fluxes without extreme star formation rates inconsistent with the other observations. We take this to suggest that Lyα emission from the AGN may be contaminating these observations. We note also that X-ray AGN candidates are also often LAEs at redshifts $z \sim 3$, including in SSA22, where five out of eight of the L09 X-ray AGN were originally discovered as LAEs. Since we do not model AGN line emission, for the CDF samples we exclude bands that contain the Lyman-α line (i.e., U-band at $z \sim 2$ and ACS F435W at $z \sim 3$).

Some of the strongest AGN narrow lines, including Hβ, [OIII]λ5007, and [OII]λ3727 fall in the NIR JHK bandpasses at $z = 3.09$. One of the SSA22 galaxies, J221732.00+001655.6, has significantly elevated JHK measurements compared to our best-fit models, and in preliminary fits was unacceptably under-fit. Since there is no evidence of a companion galaxy in the F160W image, we speculate that the elevated photometry is due to contamination of the NIR photometry by AGN narrow line emission, and consequently exclude the J, H, F160W, and K bands from the

fit. We note that if strong narrow lines alone caused this issue we might also expect the $i-$band measurement to be larger than the models could produce, due to contributions from the CIVλ1549 line. We do not see this in our fits, and we cannot conclusively say that narrow lines are the cause. However, we find that the solution we recover without the J, H, F160W, and K bands is consistent with the preliminary, under-fit solution, though with better sampling and larger uncertainties due to the smaller χ^2. Since the SFH solution for this galaxy is apparently robust to the omission of this photometry, we proceed with the second, better-fit solution.

4.3.3 X-ray Modeling

The relationship between the intrinsic optical and X-ray luminosities of AGN are typically quantified empirically with a two-point spectral index between the L_{2500} and $L_{2\ \mathrm{keV}}$, the intrinsic luminosity densities at 2500 Å and 2 keV. There are a number of different calibrations for this relationship, including the $L_{2500} - \alpha_{ox}$ relationship (e.g., Just et al., 2007; as used in CIGALE, Yang et al., 2020, 2022) and the Lusso & Risaliti (2017) $L_{2500} - L_{2\ \mathrm{keV}}$ relationship. The scatter around and variations between these empirical relationships were modeled physically in Kubota & Done (2018) as variations in the properties of the SMBH, including the SBMH mass M_{SMBH} and the Eddington ratio, $\dot{m} = \dot{M}/\dot{M}_{\mathrm{edd}}$. To model the AGN X-ray emission in LIGHTNING, we adopt the `qsosed` models from Kubota & Done (2018), which were found to reproduce the Lusso & Risaliti (2017) $L_{2500} - L_{2\ \mathrm{keV}}$ relationship throughout the $M_{\mathrm{SMBH}} - \dot{m}$ parameter space (see Figure 7 in Kubota & Done, 2018). These models are a subset of the `agnsed` family of models, intended to reproduce the soft X-ray excess seen in some AGN spectra with a three-component model including an outer accretion disk, a warm Comptonizing region covering the disk, and a hot Comptonizing region around the central source, which produces the hard X-ray spectrum. The `qsosed` model

fixes the electron temperatures and geometries of the three regions, such that $M_{\rm SMBH}$ and $\dot{m}$ are the only remaining physical parameters. In LIGHTNING we sample both these parameters, and allow them to set the normalization of the entire X-ray-to-IR AGN model by matching the intrinsic L_{2500} of the SKIRTOR UV-IR AGN model to that of the X-ray model.

The X-ray emission from X-ray binaries within the stellar population of a galaxy can in some cases rival the X-ray emission from an AGN, though in our samples selected by X-ray luminosity they are expected to have small or negligible contributions to the X-ray spectrum. We model the X-ray binary population as a power law with an exponential cutoff at high energies, with photon index $\Gamma = 1.8$ and cutoff energy $E_{\rm cut} = 100$ keV. We determine the X-ray binary luminosities using $L_{\rm X}/M_\star$ relationships based on the stellar-age-dependent parameterizations provided in Gilbertson et al. (2022):

$$\frac{L_{\rm X}^{\rm LMXB}}{M_\star}(\tau) = -1.21(\tau - 9.32)^2 + 29.09 \text{ erg s}^{-1}\text{M}_\odot^{-1} \tag{4.2}$$

and

$$\frac{L_{\rm X}^{\rm HMXB}}{M_\star}(\tau) = -0.24(\tau - 5.23)^2 + 32.54 \text{ erg s}^{-1}\text{M}_\odot^{-1}, \tag{4.3}$$

where τ is the base-10 log of the stellar age in years and $L_{\rm X}$ is measured from rest-frame $2-10$ keV. We thus calculate the HMXB and LMXB luminosity of each stellar age bin of the SFH and sum them to produce the X-ray binary luminosity of a given galaxy.

We model X-ray absorption with the `tbabs` absorption model (Wilms et al., 2000). We use three instances of `tbabs` to model absorption: two `tbabs` in the rest-frame of the galaxy to model the absorption of the AGN and the X-ray binary population, and one in the observed-frame to

account for the absorption by the Milky Way, where the Galactic HI column density along the line of sight (retrieved using the `CIAO colden` procedure) is held constant. Only the HI column density in the nuclear regions is a free parameter in our fits; the HI column density in the galaxy (modifying only the stellar population absorption) is derived from the $V-$band attenuation, A_V (computed in turn from τ_V) assuming that

$$N_H/A_V = 2.24 \times 10^{21}\ \mathrm{cm}^{-2}\ \mathrm{mag}^{-1}, \tag{4.4}$$

chosen as an average of observations of the Milky Way $N_H - A_V$ relationship (Predehl & Schmitt, 1995; Nowak et al., 2012). We note that the results for our samples, which we selected by L_X to include only AGN-dominated sources, are insensitive to the assumed level of absorption for the X-ray binary population.

Our X-ray fitting procedure in LIGHTNING can use either fluxes, which are widely available in high-level X-ray source catalogs but are dependent on an assumed spectral model, or instrumental counts, which are model-independent but require ancillary products to use in fitting. In the former case, the user need only supply fluxes, uncertainties, and bandpasses. Bandpasses are assumed to have uniform sensitivity across their width. In the latter case, we require a user to supply the net counts in each bandpass, the bandpasses, the exposure time, and the auxiliary response function (ARF), which describes the energy-dependent effective area of the detector. The models are converted to photon flux density, multiplied by the ARF, and integrated over the bandpass to get the counts predicted by the model in each bandpass. In this work, we use instrumental counts for our X-ray photometry.

For the CDF-N and CDF-S, we use the ARF nearest the aimpoint of the field and account for its

off-axis variation by using the vignetting-corrected exposure time from the published CDF point source catalogs. In SSA22, we generate an ARF appropriate for each source position, based on our AE runs. Once we have calculated the model counts, we calculate the X-ray model contribution to χ^2 as

$$\chi^2_{\rm Xray} = \sum_{i=1}^{n} \frac{(N_i^{mod} - N_i)^2}{\sigma N_i^2}, \tag{4.5}$$

where N_i is the net counts in bandpass i. For the SSA22 data, we use

$$\sigma N_i^2 = \sigma S_i^2 + (\frac{A_S}{A_B})^2 \sigma B_i^2, \tag{4.6}$$

where S_i and B_i are the source and background counts in bandpass i, with A_S and A_B the respective areas of the source and background regions. We use the upper error of the Gehrels (1986) approximation for the uncertainties:

$$\sigma x = 1 + \sqrt{3/4 + x}. \tag{4.7}$$

For the far deeper CDF data, we assume the source counts dominate over the background counts in both bandpasses, and let

$$\sigma N_i = 1 + \sqrt{3/4 + N_i}. \tag{4.8}$$

4.3.4 Fitting

The parameters outlined above and their allowed ranges for our SED fits are summarized in Table 4.4. We assumed uniform priors on all free parameters. Upper limits are handled in our fitting

procedure by setting the observation in the corresponding band to 0, and setting the uncertainty to the 1σ limiting flux.[6]

To fit the SEDs of our samples, we made use of an implementation of the Goodman & Weare (2010) affine-invariant MCMC algorithm within LIGHTNING. We use an ensemble of 75 MCMC samplers, which we run for 40000 steps. We use the "stretch move" as described in Goodman & Weare (2010). To achieve acceptance fractions $> 20\%$, we change the proposal distribution width parameter a to 1.8 (the original authors recommend using $a = 2$). For our fits, we find that the integrated autocorrelation times of the parameters can be quite long, approaching ~ 1000 steps for some parameters (see, e.g., Goodman & Weare, 2010 and Foreman-Mackey et al., 2013 for discussions of the autocorrelation time in MCMC applications).

To produce the final chains, from which we sample the best-fit parameters and uncertainties, we first discard a number of burn-in steps from the beginning of each chain equal to twice the longest autocorrelation time of any parameter of the chain. We then calculate the autocorrelation time again for the new chain minus burn-in, thin the remaining length of the chains by a factor of half the new longest autocorrelation time, and stack the chains of all the samplers. This results in a final chain consisting of independent samples. We take the last 1000 of these samples for each galaxy as the sampled posterior distribution.

When plotting the SEDs, we show the best fit and the full range of the best 68% of models (in terms of χ^2) drawn from the sampled posterior as the uncertainty in the fit. When quoting individual parameter values and star formation histories, we report the median values from the marginalized posterior distributions, with the 16th to 84th percentile ranges as the uncertainties.

[6]This is only an approximation of the statistically correct handling of upper limits, albeit a widely-used one that produces the expected behavior for upper limits (see, e.g., Figure 4.4). See Appendix A of Sawicki (2012) for a discussion of the model likelihood in the case of nondetections.

We make this distinction to better show the connection between the observational coverage, data uncertainties, and the uncertainties on the fit when plotting the SED.

We use posterior predictive checks (PPC) (Rubin, 1984; Gelman et al., 1996; see also section 5.1 of Chevallard & Charlot, 2016 for an example of PPC used in SED fitting) to assess the suitability of the models used in our fits. Of the SSA22 sample, two galaxies, J221736.54+001622.6 ($p = 0.013$) and J221716.16+001745.8 ($p = 0.000$) have $p-$values that suggest poor fits, though we find that these are likely data-driven rather than model-driven.

While the $p-$value for J221736.54+001622.6 is not extreme (i.e. it is > 0.01), it does suggest under-fitting. Upon inspection, it appears that the under-fitting is due entirely to the u^* measurement. The other galaxies in the SSA22 sample have u^* measurements consistent with either non-detection or a steep fall-off in flux, as expected for $z = 3.1$ galaxies; J221736.54+001622.6 has the largest u^* flux of the SSA22 galaxies (a factor of ~ 10 greater than the next-largest). This suggests that J221736.54+001622.6 is an outlier; given the good agreement between the data and our model for the remainder of the bands, we treat our model as appropriate for this galaxy otherwise.

The fit to J221716.16+001745.8 is significantly ruled out according to the PPC, with $p < 0.001$. Upon inspection, this appears to be driven by the ALMA data: the ALMA Band 6 (1.1 mm) detection is significantly larger than the Band 7 (870 μm) upper limit, suggesting a flat spectral slope that the Rayleigh-Jeans tail of our dust model cannot reproduce. Notably, this source is located close ($< 2''$) to an offset ALMA 870 μm detection not originally associated with this source. However, the offset of the 870 μm source is in the direction of the synthesized beam's major axis, suggesting that it could be a plausible counterpart. We re-fit with the flux of the offset detection in place of the upper limit, finding a much improved fit ($p = 0.164$), and the same SFR and SFH shape as we recover when fitting with the ALMA Band 7 upper limit. See Section 4.A

for details on the alternative SED fit to this AGN.

To establish the degree to which we can constrain the stellar population properties of the SSA22 AGN hosts, we also fit their SEDs with an AGN-only model (i.e., with all SFH coefficients set to 0). Details of these AGN-only SED fits and comparisons to the model with a stellar population can be found in Section 4.B. For all but two of the SSA22 AGN, the PPC rules out the AGN-only model with $p < 0.001$. Our model-comparison tests indicate that the AGN-only model is statistically preferred for both of these AGN, J221720.24+002019.3 and J221759.23+001529.7. Given that the data (under our models) are consistent with zero star formation, in what follows we treat the stellar population properties from our original SED fits as upper limits for these two AGN.

As for the CDF sample, the majority of galaxies are well-fit with $p > 0.05$, though we find 35/151 = 23% (24/151 = 16%) with $p < 0.05$ (0.01). This under-fitting may be model-driven (i.e. the model is not suitable for the data) or data-driven (if, e.g. the data are blended, mismatched, or the uncertainties are overly-constraining). We have taken steps to mitigate blending of the IR data, and used the carefully cross-matched Barro et al. (2019) catalogs to attempt to mitigate mismatches. We also added calibration uncertainties to the data to avoid overly-constrictive uncertainties. However, we note that our choice of model uncertainty (10% in each band) is arbitrary; we may need larger values to appropriately model uncertainties about, e.g. the power law slopes of the AGN accretion disk model, the optical depth of the AGN torus, or the values of any of the other fixed model parameters. To test the effect of model uncertainty on the distribution of $p-$values, we re-fit the CDF samples with a model uncertainty of 12% in each band. This decreased the number of galaxies with $p < 0.01$ to 11/151 (7%). We note also that across the sample, the under-fitting does not appear to be systematically related to any particular component of the model. Given the decrease in under-fitting with a modest and arbitrarily chosen increase in the model uncertainty,

Table 4.4: Parameters and assumptions for SED fits.

Model Component	Parameter	Description	Value/Range[a]
Stellar Population	$\{\psi_i\}_{i=1}^{n}$ [b]	Star Formation History coefficients in solar masses per year	$[0.0, \infty)$
Dust Attenuation	$\tau_{V,\mathrm{Diff}}$	Optical depth of diffuse dust	$[0.0, 10.0]$
	δ	UV attenuation curve power-law slope deviation from Calzetti et al. (2000) law	$[-4.0, 0.3]$[c]
Dust Emission	α	Power law slope of intensity distribution	2.0
	U_{min}	Intensity distribution minimum	10
	U_{max}	Intensity distribution maximum	3×10^5
	q_{PAH}	Mass fraction of PAHs in dust mixture	0.0047
	γ	Mass fraction of dust exposed to intensity distribution	$[0.0, 1.0]$
AGN Emission	$\tau_{9.7\ \mu\mathrm{m}}$	Edge-on optical depth of AGN torus at 9.7 μm	7.0
	$\cos i$	Cosine of AGN torus inclination to the line of sight	$[0.0, 1.0]$
X-ray Emission	$n_H / 10^{20}$ cm^{-2}	Neutral Hydrogen column density along the line of sight	$[10^{-4}, 10^5]$
	M_{SMBH}	SMBH mass parameter in solar masses	$[10^5, 10^{10}]$
	$\log \dot{m}$	$\log_{10}$ of SMBH accretion rate, normalized by the Eddington rate	$[-1.5, 0.3]$

Notes – [a]For free parameters, the allowed ranges are given in brackets. Priors are uniform on all free parameters. [b]Galaxies with $z < 1.17$ have 5 SFH coefficients; galaxies with $z \geq 1.17$ (including the entire SSA22 sample) have 4. [c]δ is only a free parameter for the CDF sample fits; it is fixed at 0 for the SSA22 sample.

and the lack of systematics in the under-fitting, we treat the models as suitable for the CDF sample.

4.4 Results

4.4.1 L09 AGN

In Table 4.5 we summarize the SED-fitting derived physical properties of the SSA22 AGN sample. Best-fit SED models and sampled posterior SFHs for the 8 L09 AGN are shown in Figure 4.4, with ADF22A1 shown in Figure 4.5 (see Section 4.4.2). The stellar masses and star formation rates for the SSA22 and CDF samples are shown in Figure 4.6. We find stellar masses ranging from $10^{10.5} - 10^{11.0}$ $\mathrm{M}_\odot$ and black hole masses ranging from $10^{8.3} - 10^{8.6}$ $\mathrm{M}_\odot$ for the protocluster AGN, though our measurements of the black hole mass may be complicated by the level of circumnuclear obscuration in the sample. Of the galaxies in the SSA22 sample for which we can successfully constrain the star formation history, half (3/6) are classified as sub-main sequence ($-1.3 < \log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -0.4$) or high quiescent ($\log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -1.3$) according to the criteria and redshift-dependent main sequence defined by Aird et al. (2019).

In particular, J221709.60+001800.1 is identified as a plausible quiescent galaxy, with SFR $= 6.80^{+5.20}_{-3.24}$ $M_\odot$ yr^{-1} and $\log \mathrm{SFR}/\mathrm{SFR_{MS}} = -1.38^{+0.26}_{-0.29}$. The galaxies consistent with the main sequence (J221735.84+001559.1, J221716.16+001745.8, and J221732.00+001655.6) are typically ALMA-detected, dusty galaxies, with $\tau_V = 1.4 - 2.3$. Of the galaxies with $\log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -0.4$, only J221736.54+001622.6 is ALMA-detected. Our estimates of the SFR for the ALMA Band 7 detected galaxies are compatible with the SFR estimates from Alexander et al. (2016), which were based on a Kennicutt (1998) scaling relationship corrected to a Chabrier (2003) IMF. The Alexander et al. (2016) estimates are systematically larger by a factor of ~ 2, but this is to be expected when comparing SFR estimates from scaling relations and SPS; our fits to these galaxies typically include significant components in the SFH from stars older than 100 Myr, which also heat dust and contribute to the measured L_{IR}. As Alexander et al. (2016) estimated L_{IR} from the flux at 870 μm, well beyond the peak of the AGN IR emission, their SFR estimates are unlikely to be contaminated by AGN emission.

The majority (6/8) of the fits to the SSA22 AGN sample are consistent with heavily obscured AGN growth ($N_H \geq 5 \times 10^{23}$ cm^{-2}), and J221759.23+001529.7 is moderately obscured ($N_H = 1$–5×10^{23} cm^{-2}) with $N_H = 4.37^{+2.16}_{-1.92} \times 10^{23}$ cm^{-2}. Three of the heavily obscured AGN are Compton-thick (CT) candidates, with $N_H \geq 1 \times 10^{24}$ cm^{-2} (J221709.60+001800.1 is also consistent with Compton-thickness, with $N_H = 9.98^{+45.97}_{-6.17} \times 10^{23}$), which makes the estimates of their black hole masses and particularly accretion rates more uncertain, as our X-ray model does not include line emission or reflection, which make important contributions to the X-ray spectra of CT sources. The level of obscuration we measure from our SED fits is consistent with the (typically hard) spectral indices estimated by Lehmer et al. (2009b): notably, they found $\Gamma < 0.42$ for J221739.08+001330.7 and $\Gamma = 0.16$ for J221732.00+001655.6, both of which we find to be CT

candidates.

Only J221720.24+002019.3 is consistent with being lightly obscured or unobscured, with $N_H = 4.87^{+6.15}_{-3.21} \times 10^{22}$ cm^{-2}. J221720.24+002019.3 also has low UV-optical attenuation ($\tau_V = 0.2 \pm 0.1$) and our proxy for circumnuclear UV-optical obscuration, cos*i*, is consistent with an unobscured view of the SMBH, suggesting that this source is a so-called "blue" quasar. This source, along with J221736.54+001622.6 and J221716.16+001745.8 are presented in the literature as Type 1 AGN with broad lines in their rest-frame optical spectra (J221720.24+002019.3 in Yamada et al., 2012, J221736.54+001622.6 in Steidel et al., 2003, and J221716.16+001745.8 in Saez et al., 2015). We find that AGN continuum emission makes strong contributions to the optical SED fits for J221720.24+002019.3 and J221736.54+001622.6, but not in the case of J221716.16+001745.8. This may suggest additional dust obscuration along the line of sight in J221716.16+001745.8, beyond the circumnuclear region. The broad line emission from unobscured AGN is one of the primary obstacles to extracting detailed star formation histories; we note that while we fit an AGN continuum component to the rest-frame UV-optical emission, we are unable to account for broad lines, and so the SFH we present here are naturally more uncertain. The ALMA Band 7 detections for J221736.54+001622.6 and J221716.16+001745.8 allow us to place good constraints on their recent star formation, though the constraints on the oldest two stellar age bins are weaker, especially in the case of J221736.54+001622.6, where the strength of the AGN continuum emission dominates over the stellar population at rest-frame wavelengths $\gtrsim 0.2$ μm. In the case of J221720.24+002019.3, the overall shape and normalization of the SFH are not well constrained by our original SED fit, though the non-detection in both ALMA bands suggests low star-formation activity. We find a SFR of ≤ 53.83 M$_\odot$ (a factor of 3 lower than the upper limit SFR computed from the ALMA limiting flux in Alexander et al. (2016)), and the upper limit on $\log \mathrm{SFR}/\mathrm{SFR_{MS}}$ is

-0.31, suggesting that this galaxy is at most a factor of ≈ 2 below the location of the star-forming main sequence at $z = 3.09$ and is likely a sub-MS galaxy. The low SFR is also supported by optical spectroscopy: the SDSS spectrum[7] has a clear, strong 4000 Å break, though at this redshift the 4000 Å break is at the edge of the SDSS sensitivity. The low SFR limits we recover for this source, in combination with the high X-ray luminosity and low obscuration (as measured by both N_H and $\cos i$), place it in an interesting position in the "story" of AGN growth in the protocluster, as a possible endpoint of AGN evolution, where star formation has stopped and the AGN has blown away obscuring material.

4.4.2 ADF22A1

ADF22A1 sits at the opposite end of the typical picture of AGN evolution from J221720.24+002019.3. The ADF22A1 proto-quasar system was identified as an AzTEC 1.1 mm source in Tamura et al. (2010). Further study showed that the system consists at least two optical/NIR sources offset from a millimeter and X-ray source, which is not detected in bands up to a few microns. The smoothed FWHM of the SSA22 photometry we use is larger than the projected offset between the galaxies and the millimeter/X-ray source, so we fit the SED of the entire ADF22A1 system. We show the best-fit model and the sampled posterior SFH in Figure 4.5. The fit from our models is consistent with the picture of ADF22A1 harboring a heavily obscured/Compton-thick AGN, with $N_H = 5.19^{+3.24}_{-3.13} \times 10^{24}$ cm^{-2}, $\cos i = 0.33^{+0.21}_{-0.23}$, and $\tau_{V,\mathrm{Diff}} = 2.9^{+0.1}_{-0.1}$. The SFH suggests an initial burst of star formation in the first ~ 1 Gyr after formation, followed by an ongoing, more intense (by a factor of ~ 10) burst of star formation in the past 100 Myr. We measure the average star formation over the past 100 Myr as 524^{+276}_{-258} M$_\odot$ yr^{-1},

[7]Retrieved from http://skyserver.sdss.org/dr17

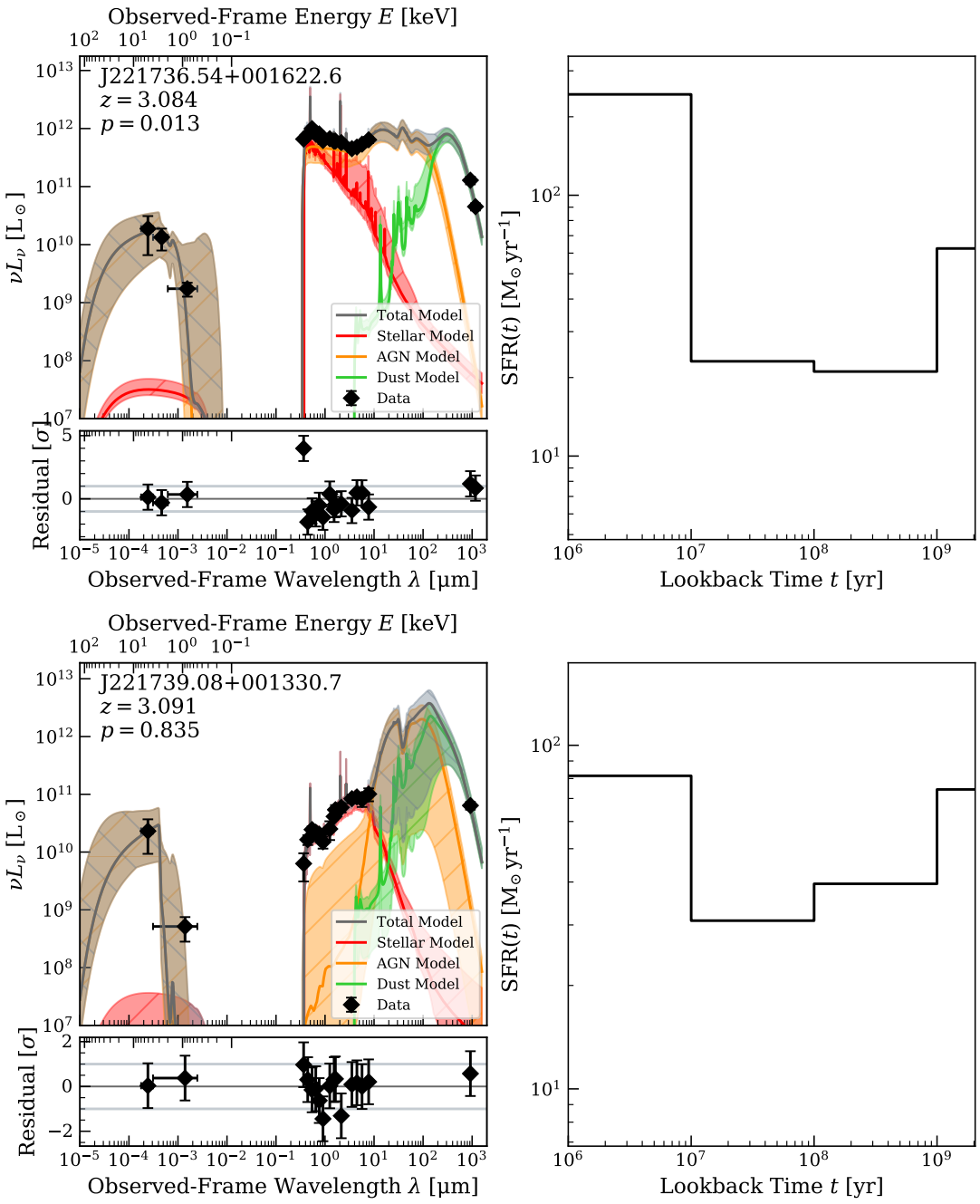

Figure 4.4: *Left panels:* The best-fit SED model for each of the 8 L09 X-ray AGN in the SSA22 is shown with its components. The shaded regions of the SED plot indicate the full range of the best-fitting 68% of models. Upper limits are shown at 3σ. The data $-$ model residuals from the best-fit model are shown below, in units of σ. The $p-$value in the annotation is computed with a posterior predictive check (see Section 4.3.4). *Right panels:* We show the sampled posterior SFH: for galaxies with constrained SFH, the dark line indicates the median of the posterior and the shaded region shows the 16th to 84th percentile range. For J221720.24+002019.3 and J221759.23+001529.7 the dark line indicates the 99th percentile upper limit of the posterior.

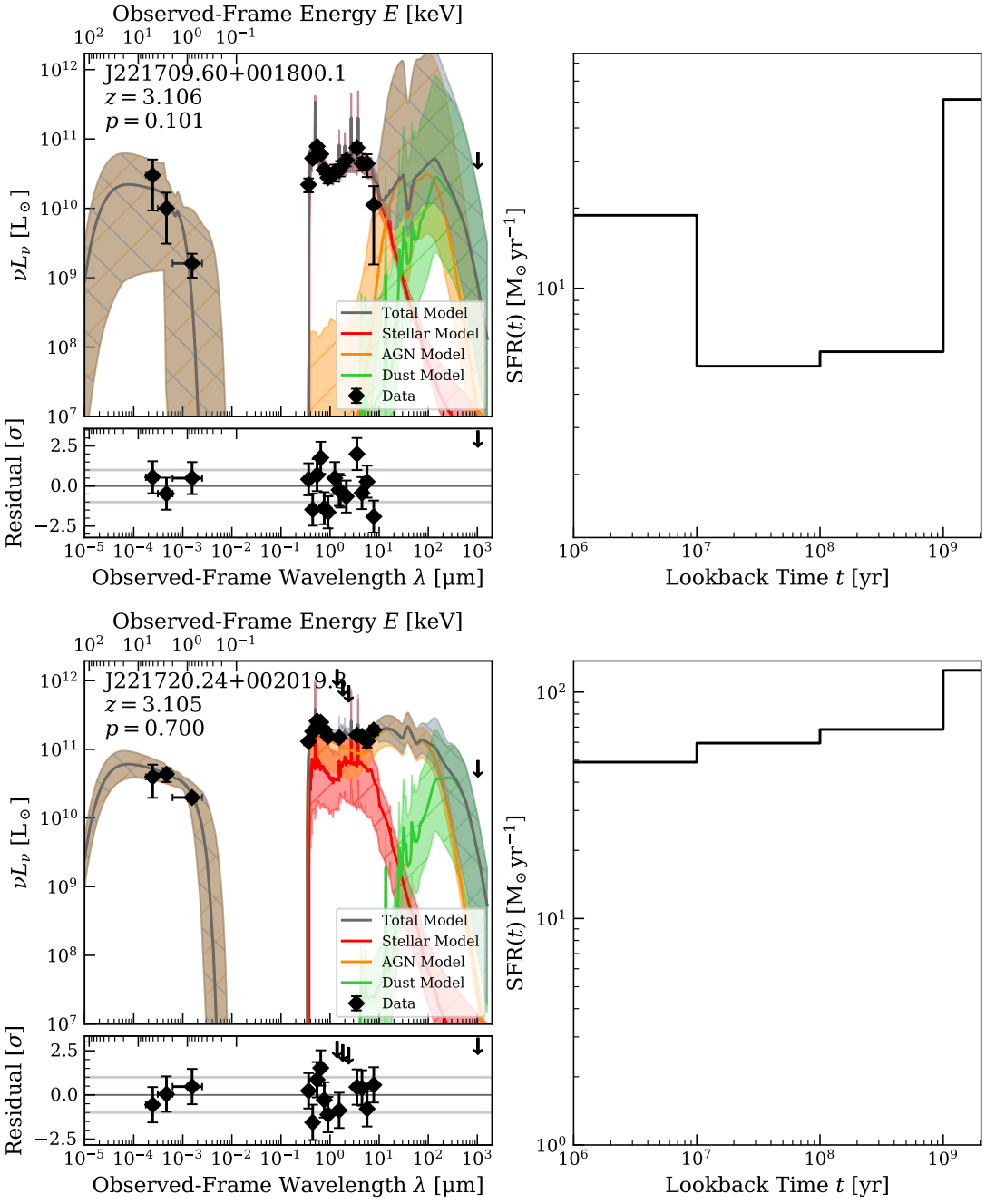

Figure 4.4: *Continues.*

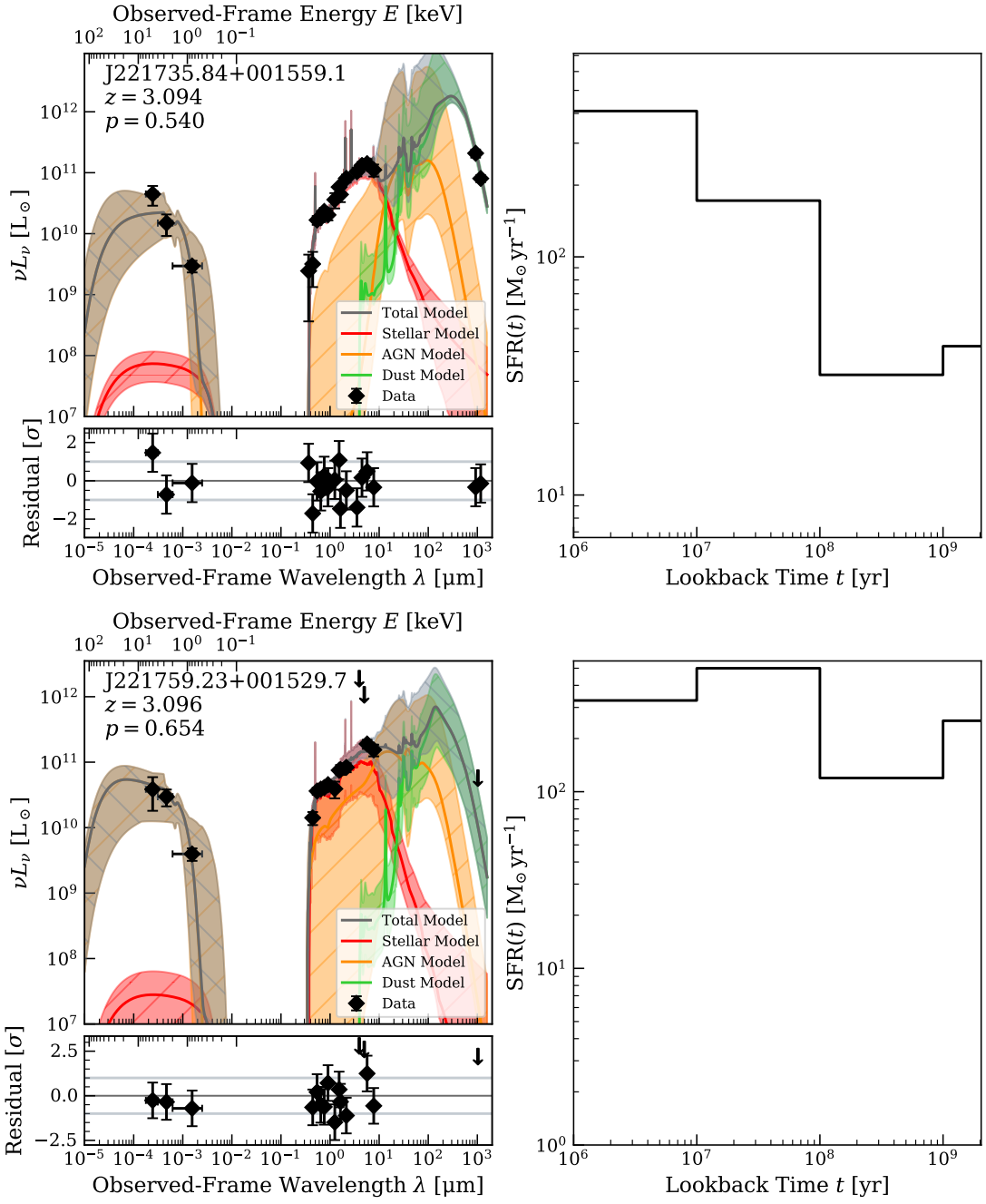

Figure 4.4: *Continues.*

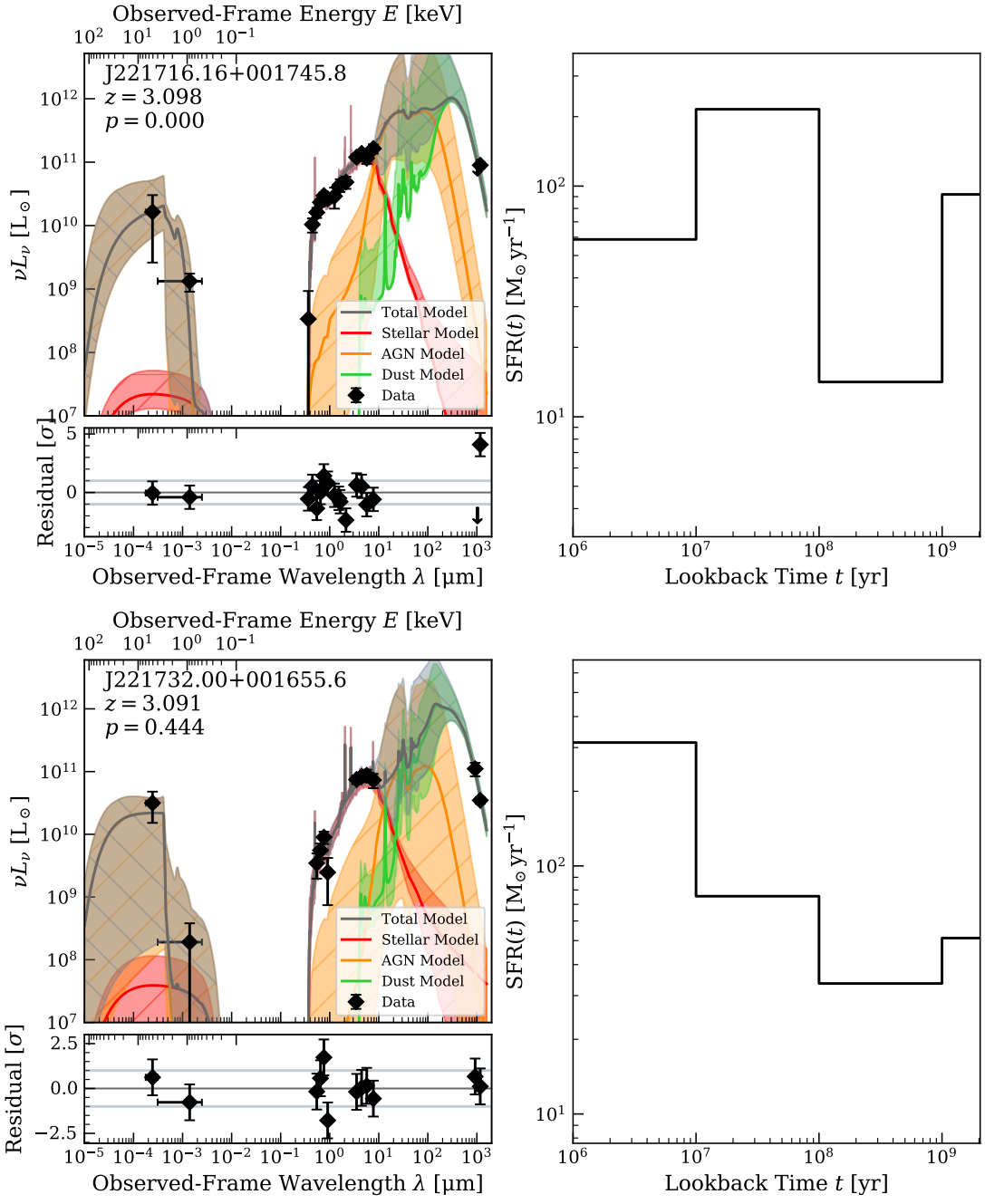

Figure 4.4: *Continues.*

Table 4.5: SED-fit derived properties for the SSA22 sample.

ID	$\log M_\star$[a] ($M_\odot$)	SFR[b] ($M_\odot$ yr^{-1})	logSFR/SFR$_{MS}$[c] (dex)	τ_V[d]	$\cos i$[e]	$\log M_{SMBH}$[f] ($M_\odot$)	$\log \dot{m}$[g]	N_H[h] (10^{22} cm^{-2})
J221736.54+001622.6	$10.8^{+0.2}_{-0.3}$	$45.2^{+31.6}_{-13.8}$	$-0.72^{+0.33}_{-0.25}$	$0.3^{+0.0}_{-0.0}$	$0.80^{+0.10}_{-0.04}$	$8.3^{+0.1}_{-0.1}$	$-0.5^{+0.3}_{-0.2}$	$59.0^{+35.8}_{-23.9}$
J221739.08+001330.7*	$\mathbf{10.9^{+0.1}_{-0.1}}$	$\mathbf{36.8^{+37.7}_{-21.1}}$	$\mathbf{-0.91^{+0.37}_{-0.36}}$	$\mathbf{1.2^{+0.2}_{-0.3}}$	$\mathbf{0.28^{+0.17}_{-0.18}}$	$\mathbf{8.6^{+0.1}_{-0.2}}$	$\mathbf{-0.3^{+0.3}_{-0.4}}$	$\mathbf{572^{+299}_{-268}}$
J221709.60+001800.1	$10.5^{+0.1}_{-0.1}$	$6.80^{+5.20}_{-3.24}$	$-1.38^{+0.26}_{-0.29}$	$0.2^{+0.1}_{-0.1}$	$0.20^{+0.21}_{-0.14}$	$8.2^{+0.2}_{-0.3}$	$-1.0^{+0.5}_{-0.4}$	$99.8^{+460.0}_{-61.7}$
J221720.24+002019.3	$\leq 10.9^{+}$	≤ 53.8	≤ -0.31	$0.2^{+0.1}_{-0.1}$	$0.82^{+0.09}_{-0.06}$	$8.6^{+0.1}_{-0.1}$	$-1.4^{+0.1}_{-0.0}$	$4.87^{+6.15}_{-3.21}$
J221735.84+001559.1	$10.8^{+0.1}_{-0.2}$	198^{+111}_{-90}	$-0.13^{+0.28}_{-0.35}$	$1.9^{+0.1}_{-0.1}$	$0.22^{+0.23}_{-0.16}$	$8.5^{+0.2}_{-0.2}$	$-0.4^{+0.5}_{-0.7}$	$52.6^{+32.2}_{-23.5}$
J221759.23+001529.7	≤ 11.3	≤ 455	≤ 0.34	$1.2^{+0.1}_{-0.1}$	$0.40^{+0.29}_{-0.26}$	$8.5^{+0.1}_{-0.1}$	$-1.2^{+0.4}_{-0.2}$	$43.70^{+21.60}_{-19.10}$
J221716.16+001745.8	$\mathbf{10.9^{+0.1}_{-0.1}}$	$\mathbf{201^{+76}_{-69}}$	$\mathbf{-0.20^{+0.23}_{-0.24}}$	$\mathbf{1.4^{+0.1}_{-0.1}}$	$\mathbf{0.33^{+0.15}_{-0.15}}$	$\mathbf{8.7^{+0.1}_{-0.2}}$	$\mathbf{-0.1^{+0.2}_{-0.3}}$	$\mathbf{425^{+375}_{-238}}$
J221732.00+001655.6	$\mathbf{10.8^{+0.1}_{-0.2}}$	$\mathbf{105^{+79}_{-49}}$	$\mathbf{-0.39^{+0.32}_{-0.36}}$	$\mathbf{2.3^{+0.2}_{-0.3}}$	$\mathbf{0.22^{+0.23}_{-0.15}}$	$\mathbf{8.3^{+0.2}_{-0.4}}$	$\mathbf{-0.5^{+0.5}_{-0.7}}$	$\mathbf{622^{+263}_{-275}}$
ADF22A1	$\mathbf{11.3^{+0.2}_{-0.2}}$	$\mathbf{524^{+276}_{-258}}$	$\mathbf{-0.05^{+0.30}_{-0.40}}$	$\mathbf{2.9^{+0.1}_{-0.1}}$	$\mathbf{0.33^{+0.21}_{-0.23}}$	$\mathbf{8.6^{+0.2}_{-0.2}}$	$\mathbf{-0.8^{+0.3}_{-0.2}}$	$\mathbf{519^{+324}_{-313}}$

Notes – [a]Total stellar mass, integrated over the SFH. [b]Star formation rate, averaged over the last 100 Myr of the SFH. [c]Ratio of SFR to the main sequence SFR of Aird et al. (2019). [d]$V-$band optical depth of diffuse dust. [e]Cosine of the inclination from the AGN axis to the line of sight. [f]Supermassive black hole mass from `qsosed` model. [g]Eddington ratio from `qsosed` model. [h]Neutral Hydrogen column density along the line of sight. *Compton-thick (CT) candidates with $n_H \geq 1 \times 10^{24}$ cm^{-2} are highlighted in bold. +Upper limits for J221720.24+002019.3 and J221759.23+001529.7 are given at the 99th percentile of the posterior.

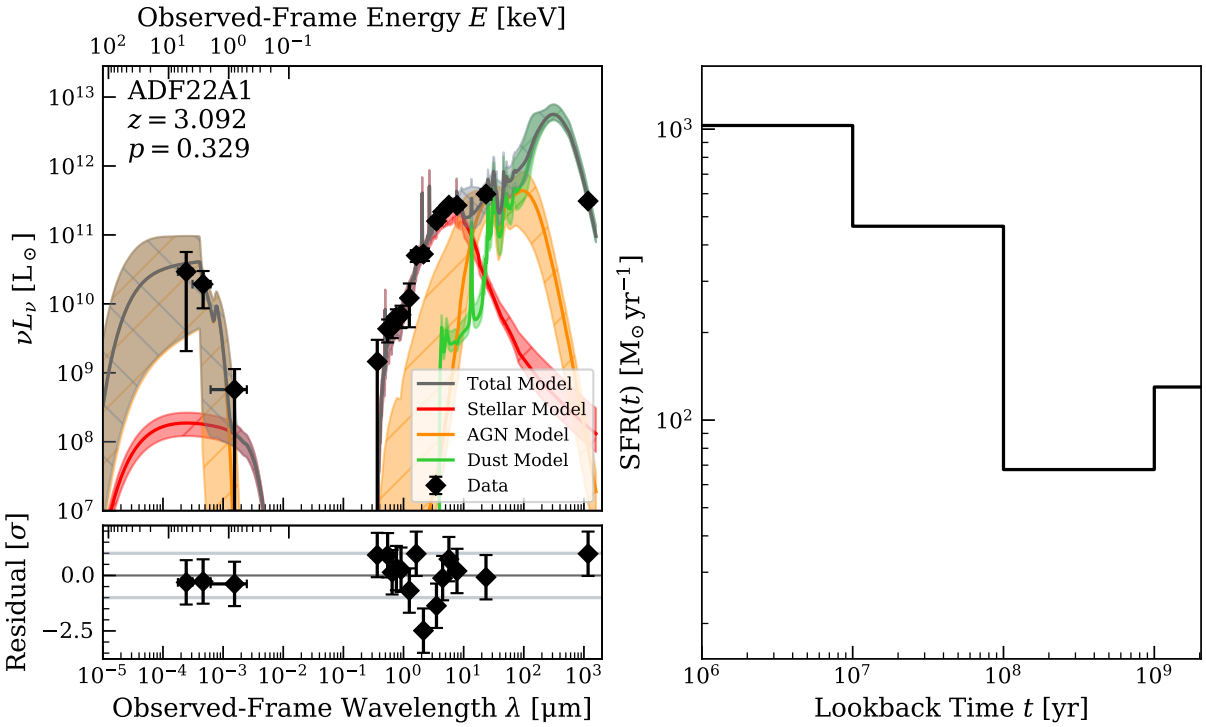

Figure 4.5: *Left panel:* The best fit SED model for ADF22A1 is shown along with its components. The shaded regions of the SED plot indicate the full range of the best-fitting 68% of models. The lower panel shows the data − model residuals from the best-fit model, in units of σ. The $p-$value in the annotation is computed with a posterior predictive check (see Section 4.3.4). *Right panel:* The sampled posterior distribution of the SFH is shown with the dark line indicating the median of the posterior, and the shaded region showing the 16th to 84th percentile range. The SED and SFH of ADF22A1 are consistent with a luminous, heavily obscured AGN, accompanied by intense, obscured star formation. These results are consistent with the identification of ADF22A1 as a sub-millimeter galaxy hosting an obscured "proto-quasar."

a factor of ~ 2 smaller than the estimate from the 1.1 mm flux, but still more than twice the SFR of any of the other X-ray selected AGN in our sample. The mass we recover for the system, $2.00^{+1.16}_{-0.74} \times 10^{11}$ M$_\odot$, is consistent with the high end of the mass estimates from Tamura et al. (2010), $8.2^{+9.0}_{-1.5} \times 10^{10}$ M$_\odot$. The optical attenuation above is also consistent with the value reported in Tamura et al. (2010) ($\tau_V = 3.1^{+0.3}_{-0.2}$), and with the level of $V-$band attenuation typically measured in SMGs. We note again that our model does not include reflection, and as such the uncertainties on our estimate of N_H are large, though we find that ADF22A1 is likely Compton-thick,

as also found by Tamura et al. (2010). As previously mentioned, Compton-thickness renders our estimates of the AGN parameters (particularly $\dot{m}$) more uncertain.

The IR emission from the AGN remains highly uncertain due to the lack of mid-IR photometry; JWST observations with MIRI will not only better constrain the AGN IR emission, but allow the resolution of the system into its separate components, including the central X-ray point source, allowing us to separately characterize each component.

4.5 Discussion

In Section 4.4 we presented the results of our SED fits to the SSA22 AGN, finding that they are typically obscured main sequence galaxies. In the following sections, we compare the measured properties of the SSA22 AGN to the field AGN, in order to establish whether there are measurable differences in the properties of the protocluster AGN-host galaxy systems that could drive the observed AGN enhancement. We also attempt to connect the properties of the AGN-host galaxy systems to their local environment, and finally to place our results in the context of the ongoing question of what drives the AGN enhancement in the protocluster.

4.5.1 Star Formation and Black Hole Growth in the Protocluster AGN

Alexander et al. (2016) previously found, on the basis of IR scaling relations, that the range of SFR of the three IR-detected L09 X-ray AGN is consistent both with AGN in field environments and the overall evolution of the star forming main sequence. On average, we also find that the X-ray detected protocluster AGN hosts are consistent with the star forming main sequence at $z = 3.09$. In Figure 4.7 we show the evolution of the main sequence according to both the Speagle et al. (2014) and Aird et al. (2019) prescriptions, along with the SFR of all the GOODS AGN in our sample,

Table 4.6: 1D KS test $p-$values for SED-fit derived quantities between SSA22 and GOODS.

	$2 \leq z \leq 4$[a]		$2.6 \leq z \leq 3.4$	
Quantity	No CT[b]	CT Included[c]	No CT	CT Included
$M_\star$	$\cdots$	0.113	$\cdots$	0.111
SFR	$\cdots$	0.272	$\cdots$	0.239
$M_{\rm SMBH}$	0.103	0.016	0.280	0.155
$\rm SFR/SFR_{MS}$	0.979	0.610	0.912	0.239
BHAR/SFR	0.505	0.282	0.593	0.445
N_H	0.337	0.197	0.878	0.271

Notes – ADF22A1 is excluded from all KS tests; J221720.24+002019.3 and J221759.23+001529.7 are excluded from the KS tests on $M_\star$, $\rm SFR/SFR_{MS}$, and BHAR/SFR as we cannot effectively constrain their SFH.

[a]Redshift range of field galaxies. [b]$p-$values where galaxies with $N_H \geq 1 \times 10^{24}$ are excluded from the KS test. [c]$p-$values where galaxies with $N_H \geq 1 \times 10^{24}$ are included in the KS test.

up to $z = 3.3$. We also overlay the SFR of the SSA22 AGN and the sample mean SFR. We find that the average SFR of the L09 X-ray AGN in SSA22 is consistent with both the field AGN in the GOODS fields and the main sequence. While our results suggest that $4/8 = 50 \pm 18\%$ of the L09 X-ray AGN are located below the main sequence (as previously observed in the ADF22-QG1 protocluster AGN by Kubo et al., 2021, 2022), we find that sub-main-sequence evolution is just as common in our sample of L_X-selected AGN in the CDF: of the $z > 2$ ($2.4 \leq z \leq 3.4$) AGN, $52/71 = 73 \pm 10\%$ ($17/23 = 74 \pm 18\%$) have $\log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -0.4$. We show KS-tests on the posterior distributions of SFR and $\mathrm{SFR}/\mathrm{SFR_{MS}}$ in Table 4.6, suggesting that the SSA22 AGN are consistent with being drawn from the same underlying population distribution as the high-redshift CDF AGN. We find also that the AGN host galaxies in the protocluster are no more massive on average than AGN hosts in the field with similar X-ray luminosities: excluding the ADF22A1 system and the galaxies for which the AGN-only model is preferred, we find that the ratio of the average stellar mass of the protocluster AGN to the average mass of the GOODS AGN is at most 0.998 (3σ upper limit) over the redshift range $2 \leq z \leq 4$; restricting to $2.6 \leq z \leq 3.4$ we find that the protocluster AGN are on average slightly less massive than the field AGN in our CDF sample, with a mass ratio of at most 0.975.

To further establish the relationship between star formation and the black hole growth in the protocluster, we calculated the dimensionless black hole accretion rate (BHAR) to SFR ratio, $\mathrm{BHAR}/\mathrm{SFR}$, for each of our samples. This ratio traces the evolutionary state of the black hole relative to the host galaxy, and may shed light on the accretion history of the AGN, as the BHAR is shown to lag behind the SFR following major accretion episodes, as gas is driven down to the small scales near the nucleus where it can be accreted by the SMBH. $\mathrm{BHAR}/\mathrm{SFR}$ is found to be independent of redshift when controlled for mass (e.g. Aird et al., 2019). We calculated the BHAR

of our galaxies as BHAR $= \dot{m}\dot{M}_{\rm Edd}$, with

$$\dot{M}_{\rm Edd} = \frac{4\pi G M m_p}{\sigma_T \eta c}, \tag{4.9}$$

and where m_p is the proton mass, σ_T is the Thomson scattering cross section, and η is the radiative efficiency of the BH. We take $\dot{m}$ and M from our SED fits. Following Kubota & Done (2018), we assume a radiative efficiency of $\eta = 0.057$. In Figure 4.8, we show the BHAR/SFR as a function of SFR/SFR$_{\rm MS}$, where SFR$_{\rm MS}$ is computed at the appropriate redshift and stellar mass for each galaxy using the redshift-dependent main sequence from Aird et al. (2019). We find that the SSA22 AGN follow the same trend in BHAR/SFR as a function of SFR/SFR$_{\rm MS}$ as the field galaxies, indicating that, on the timescales we probe, black hole growth in the protocluster is not proceeding out of pace with galaxy growth at any stage of galaxy evolution. A 1D KS test on the distributions of BHAR/SFR indicates that protocluster and field galaxies have consistent distributions of BHAR/SFR regardless of whether we exclude CT candidates ($p = 0.505$) or include them ($p = 0.282$) or whether we limit the test to galaxies within the stellar mass range of the SSA22 AGN, $10.5 \leq \log(M_\star/{\rm M}_\odot) \leq 11.0$ ($p = 0.320$ without CT candidates and $p = 0.180$ with CT candidates). The stochasticity of AGN accretion produces significant scatter in BHAR/SFR; as such it is often useful to calculate the sample average, $\langle$BHAR/SFR$\rangle$. We find $\langle$BHAR/SFR$\rangle = 0.075^{+0.111}_{-0.037}$ for the SSA22 AGN, when we exclude CT candidates. For the CDF samples, $\langle$BHAR/SFR$\rangle$ is consistent with SSA22: $0.075^{+0.037}_{-0.017}$ over the whole range of mass and $z \geq 2$, and $0.058^{+0.031}_{-0.016}$ over $10.5 \leq \log(M_\star/{\rm M}_\odot) \leq 11.0$ and $z \geq 2$, both excluding CT candidates.

Another possible explanation for the observed AGN fraction enhancement in the protocluster is that the black holes in protocluster AGN may simply be more massive than their field counterparts.

In Figure 4.9 we show the black hole mass of our samples plotted against the stellar mass. Our high-redshift galaxy samples are likely not bulge-dominated (Monson et al. (2021) found that none of the SSA22 X-ray AGN are bulge-dominated by G–M_{20} diagnostics, and that the AGN that are not point-like sources in F160W typically have Sérsic indices $n < 2$), and as such we lack estimates of the bulge mass. Thus, we are unable to correlate stellar masses strongly with BH masses. We find, however, that the $z > 2$ CDF AGN and SSA22 AGN occupy the same general region of the $M_{\rm SMBH}$–$M_\star$ plane, and that the $2.6 \leq z \leq 3.4$ CDF AGN track very closely with the SSA22 AGN. This is shown also by the KS test results in Table 4.6; when we limit the field sample to a narrower range of redshifts around $z = 3$, we find that the SMBH mass distributions of the SSA22 and GOODS samples are consistent, with $p > 0.155$. We additionally computed the posterior distributions on the sample-averaged black hole mass (excluding CT candidates), finding that the 3σ upper limit (computed as the 99th percentile of the posterior) on the ratio of the average protocluster SMBH mass (for the 8 L09 AGN) to the average field SMBH mass is 1.41 for field galaxies with $2.6 \leq z \leq 3.4$. While empirical SBMH mass to stellar mass relationships for local bulge-dominated galaxies may not be applicable to our sample, the recent FOREVER22 suite of SSA22-like protocluster simulations (Yajima et al., 2022) allow us a theoretical point of comparison: we find that our estimates of the SMBH masses of our AGN sample are consistent with the SMBH to stellar mass relationship produced by their sample (see Fig. 11 in Yajima et al., 2022). Finally, we caution that our estimates of the black hole mass are (by necessity, for high-z AGN) indirect, based on the theoretical `qsosed` AGN model, and effectively scale with the AGN luminosity. Thus with weak constraints on the AGN luminosity (as in cases where we have few X-ray counts and no mid-IR observations) the estimated black hole masses may be less reliable.

To examine whether the comparison between the SMBH properties of the SSA22 and CDF

AGN is biased by the lack of MIPS 24 μm fluxes (which constrain the NIR AGN torus emission over a wide range of redshift) in the SSA22 SED fits (excluding ADF22A1), we re-fit the field sample without 24 μm measurements. We find that the Eddington ratio parameter $\log\dot{m}$ is the only parameter to show a systematic difference between the two cases: the fits without MIPS have systematically smaller Eddington ratios than those with MIPS. The median offset in $\log\dot{m}$ is less than 0.3 dex for $\log\dot{m} < -1$, increasing to ≈ 1.0 dex as $\log\dot{m}$ approaches 0.0. However, the uncertainties on $\log\dot{m}$ also increase with $\log\dot{m}$, such that the systematic uncertainty is less than the random uncertainties; we find that the difference between the two measurements is consistent with 0.0 within 1σ for the majority of systems. While it is thus likely that we underestimate the Eddington ratio (and consequently BHAR) for some of the SSA22 systems, for which we do not use 24 μm measurements, we expect the systematic offset to be smaller than the random uncertainties on our measurements of $\log\dot{m}$ and the additional uncertainties imposed by the heavy obscuration of the highly accreting SSA22 AGN systems. Our overall results should thus be insensitive to the lack of MIPS 24 μm constraints in our SSA22 SED fits.

4.5.2 The Possible Role of Local Environment

The SSA22 protocluster is well-suited to attempts at linking galaxy properties to the protocluster environment, largely due to the extent to which the cold-gas features of the intergalactic medium (IGM) have been mapped in the protocluster. A number of star-forming galaxies and AGN in the protocluster are associated with Lyman-α blobs (LABs), large scale Lyα emission nebulae (Matsuda et al., 2004; Geach et al., 2009), where the Lyα emission is believed to be powered by heating from the galaxies embedded within (Geach et al., 2009). The LABs are in turn associated with the large-scale Lyα emission filaments imaged with MUSE in Umehata et al. (2019). These reservoirs

of cold gas in the IGM could play a role in maintaining the levels of star formation and black hole accretion in the protocluster galaxies by providing a steady supply of cold gas for the galaxies within.

Four of the X-ray AGN, J221739.08+001330.7, J221735.84+001559.1, J221759.23+001529.7, and J221732.00+001655.6, are associated with LABs. These AGN do not clearly deviate from the other AGN in the SSA22 sample in terms of any of the parameters or quantities derived from our model. However, two, J221739.08+001330.7 and J221732.00+001655.6, have the highest N_H of the 8 L09 X-ray AGN. While Geach et al. (2009) also previously noted that the AGN associated with LABs are heavily obscured, whether or not this indicates a greater cold gas density in LAB-associated AGN would require a larger sample and more detailed X-ray modeling, beyond the scope of this work. J221736.54+001622.6 and ADF22A1 (two of the most intrinsically X-ray luminous sources in our sample) are also notably associated with nodes of the IGM filaments.

While it is believed that the Lyα emission from these nebulae is likely powered by the AGN embedded within them, it remains unclear whether the reservoirs of cold gas in the IGM traced by the filaments and LABs play any role in the fueling of the AGN. However, the LABs and the large-scale filaments are also consistently associated with star formation - the SMGs detected in Umehata et al. (2014) also co-locate with the filaments, and each of the LABs associated with an AGN in our sample contains an ALMA Band 7 detection, either directly associated with the AGN or offset (Alexander et al., 2016), so it remains likely that the IGM gas reservoirs play some role in supplying the galaxies embedded within them with a supply of cold gas.

To attempt to correlate the properties of our AGN sample with their location in the protocluster, we calculated $\Sigma_{\rm LAE}$, the projected surface density of $z \approx 3.09$ LAEs from Hayashino et al. (2004), using the same procedure as in Monson et al. (2021) with a Gaussian kernel density estimate with

a $4'$ (7.5 cMpc) FWHM[8]. We note, however, that the coherent structure of protocluster LAEs also extends by tens of cMpc along the line of sight (Matsuda et al., 2005). The projected surface density of LAEs does not capture the 3D structure of the protocluster, and can only be used as a general gauge of the protocluster density. In Figure 4.10 and Figure 4.11, we show the SED-fitting derived parameters for the SSA22 X-ray AGN and ADF22A1 plotted against the surface density. As previously shown (in, e.g., studies of ADF22: Kubo et al., 2013; Kubo et al., 2022; Umehata et al., 2014, 2015), AGN are preferentially found in the denser regions of the protocluster (ADF22A1 is believed to lie near the density peak of the protocluster), though we find that host galaxy properties and SMBH properties are uncorrelated with $\Sigma_{\rm LAE}$.

We previously attempted to establish the merger states and recent merger histories of the SSA22 X-ray AGN in Monson et al. (2021). In the Gini-M_{20} (calculated based on F160W images) plane, which is commonly used as a merger diagnostic, we found that none of the X-ray AGN have the clumpy, uneven flux distributions associated with recent mergers, and are instead consistent with concentrations of light near the center, as might be expected from mainly star-forming galaxies hosting AGN (see Figure 6 in Monson et al., 2021). We also previously employed blind classifications, with our classifiers voting on whether images contained a merger or an isolated galaxy. Based on the presence of close projected companions, 3 of the X-ray AGN (J221736.54+001622.6, J221739.08+001330.7, J221720.24+002019.3) were classified as possible mergers in this way. However, in these cases the redshifts of the projected companions are unknown and the primary galaxies are not exhibiting the disturbed morphologies associated with an interacting companion galaxy. Our updated estimates of the stellar masses for the SSA22 X-ray AGN place them at

[8]We use a larger FWHM than the $2'$ used in Monson et al. (2021); we found that a KDE with a $2'$ FWHM resulted in overestimates of the LAE density near those AGN which are themselves LAEs.

$M_\star = 10^{10.5}$–$10^{11.3}$ $M_\odot$ (in good agreement with the estimates from Kubo et al., 2015); the EAGLE simulations (McAlpine et al., 2020) suggest that mergers are most effective at triggering SMBH growth at $M_\star < 10^{11.0}$ $M_\odot$, so we again find it unlikely that mergers are a dominant growth mechanism for SMBHs in the protocluster.

4.5.3 What Drives the AGN Fraction Enhancement in the Protocluster?

We have found that the protocluster environment does not strongly impact the physical properties of the X-ray selected protocluster AGN or their host galaxies, as measured by X-ray-to-IR SED fitting. However, these same X-ray AGN account for a $\approx 6-$fold increase in the AGN fraction over the $z \sim 3$ field. In this section we briefly attempt to synthesize these results.

The primary determinant of the number of AGN in a population is the distribution of black hole masses. While we have seen above that the distribution of black hole masses for the X-ray selected AGN in the protocluster and field samples are compatible, the observed AGN fraction enhancement may indicate that the black hole mass distribution of the entire protocluster is shifted to higher masses. As higher SMBH masses correlate with higher accretion rates (and thus luminosities), a top-heavy black hole mass function would make any individual protocluster galaxy more likely to be observed as an AGN, resulting in a larger measured X-ray AGN fraction.

The mechanism by which the protocluster might develop a top-heavy black hole mass function is unclear. SSA22 protocluster galaxies have been shown to be more massive and star forming, on average, than field galaxies, and Monson et al. (2021) found that their SFHs indicated more intense or sustained stellar mass buildup at earlier epochs. Monson et al. (2021) previously found that mergers are no more common among the protocluster population than in the field, and literature indicates that mergers may be effective at driving SMBH growth for only the most massive systems.

However, the constraints on the merger fractions from Monson et al. (2021) are weak, and we do not have information about the historic merger fraction in the protocluster. It remains possible that mergers at $z > 3.1$ contributed to the buildup of stellar and BH masses. Additionally, the IGM in the protocluster, having been imaged in Lyα emission, is known to be denser than in the field, especially in the vicinity of AGN and rapidly growing sub-mm galaxies. Smooth accretion of cold gas from IGM reservoirs thus also remains a compelling explanation for the growth of galaxies and their SMBHs in the protocluster.

4.6 Summary & Conclusions

We have developed a new method for fitting the X-ray to IR SEDs of AGN-hosting galaxies, based on physically-motivated models wherever possible, and implemented it into our SED-fitting code LIGHTNING. We applied our method to the 8 L09 X-ray detected protocluster AGN and the candidate protoquasar ADF22A1 in SSA22, and to a comparison sample of 151 X-ray selected candidate AGN at $0.45 \leq z \leq 4.52$ in the *Chandra* Deep Fields.

The normal (i.e., non-AGN) star-forming population of the protocluster has previously been shown to host an enhancement of SFR and stellar mass over star-forming galaxies in the field at the same redshift (e.g. Kubo et al., 2015; Kato et al., 2016; Monson et al., 2021). However, we find here that the stellar mass and SFR distribution of the X-ray selected protocluster AGN is consistent with the stellar mass and SFR of X-ray selected field AGN, suggesting that AGN hosts in the mass range we probe grow by the same mechanisms in protocluster and field environments. We also find comparable black hole masses and growth rates in the protocluster and field, suggesting normal black hole growth in the the X-ray selected protocluster AGN. We interpret these results together

with the observed AGN fraction enhancement as suggesting that while individual X-ray AGN and the protocluster X-ray AGN population are similar to field X-ray AGN, individual galaxies in the protocluster are more likely to be observed as X-ray AGN. This could be explained by an overall enhancement of the black hole masses of the protocluster galaxy population, supported by the evidence for enhanced star formation and stellar mass in the protocluster.

- We find that SMBH growth in the SSA22 protocluster is largely obscured: 6/8 X-ray selected AGN are heavily obscured, with $N_H \geq 5 \times 10^{23}$ cm^{-2}, and 3/8 are Compton-thick candidates with $N_H \geq 1 \times 10^{24}$ cm^{-2}. The majority of the protocluster AGN, 5/8, also have optical SED fits consistent with optically obscured AGN growth, with $\cos i \lesssim 0.65$.

- At least 3/8 of the SSA22 protocluster AGN have SFHs consistent with sub-MS growth, with $\log \mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}} \leq -0.4$. J221720.24+002019.3, which is better-fit by an AGN-only model, is also likely located below the main sequence, with an upper-limit $\log \mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}} = -0.31$. One of the sub-MS AGN, J221709.60+001800.1 is identified as a probable quiescent galaxy, with $\log \mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}} = -1.38^{+0.26}_{-0.29}$.

- The protocluster AGN hosts are no more massive than the field AGN hosts in our sample: in a KS test comparing the stellar mass distributions of the SSA22 and CDF samples, we find a $p-$value of at least 0.111, suggesting that the distributions are consistent. Likewise we find a KS $p-$value of at least 0.348 for SFR.

- We find that the protocluster AGN hosts are in similar evolutionary states (relative to the main sequence) when compared to the L_X-selected field galaxies: the KS test $p-$value on $\mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}}$ is at least 0.445. Additionally, when controlled for mass, we find that the

sample-averaged BHAR/SFR ratio, measuring the rate of SMBH growth relative to host galaxy growth, is also consistent between protocluster and field samples: $\langle \mathrm{BHAR}/\mathrm{SFR} \rangle = 0.075^{+0.111}_{-0.037}$ for the SSA22 protocluster AGN and $0.058^{+0.031}_{-0.016}$ for the CDF AGN candidates with $10.5 \leq \log(M_\star/\mathrm{M}_\odot) \leq 11.0$ and $z \geq 2$.

- Our estimates of the black hole masses of our samples are also consistent ($p > 0.155$) when we limit the field sample to $2.6 \leq z \leq 3.4$. We constrain the sample-averaged black hole mass of the protocluster AGN to at most 1.41 times that of the the field AGN sample.

- For the ADF22A1 protoquasar system, we find a SFR $= 524^{+276}_{-258}$ $\mathrm{M}_\odot$ and a stellar mass $2.00^{+1.16}_{-0.74} \times 10^{11}$ $\mathrm{M}_\odot$. We estimate the black hole mass at $3.98^{+2.33}_{-1.47} \times 10^8$ $\mathrm{M}_\odot$, though this estimate is rendered more uncertain by the Compton-thickness of the system; we find $N_H = 5.62^{+2.89}_{-2.88} \times 10^{24}$ cm^{-2}. Our new constraints on the properties of this system from panchromatic SED fitting are consistent with the established picture of ADF22A1 as one of the most obscured and highly star-forming systems in the protocluster.

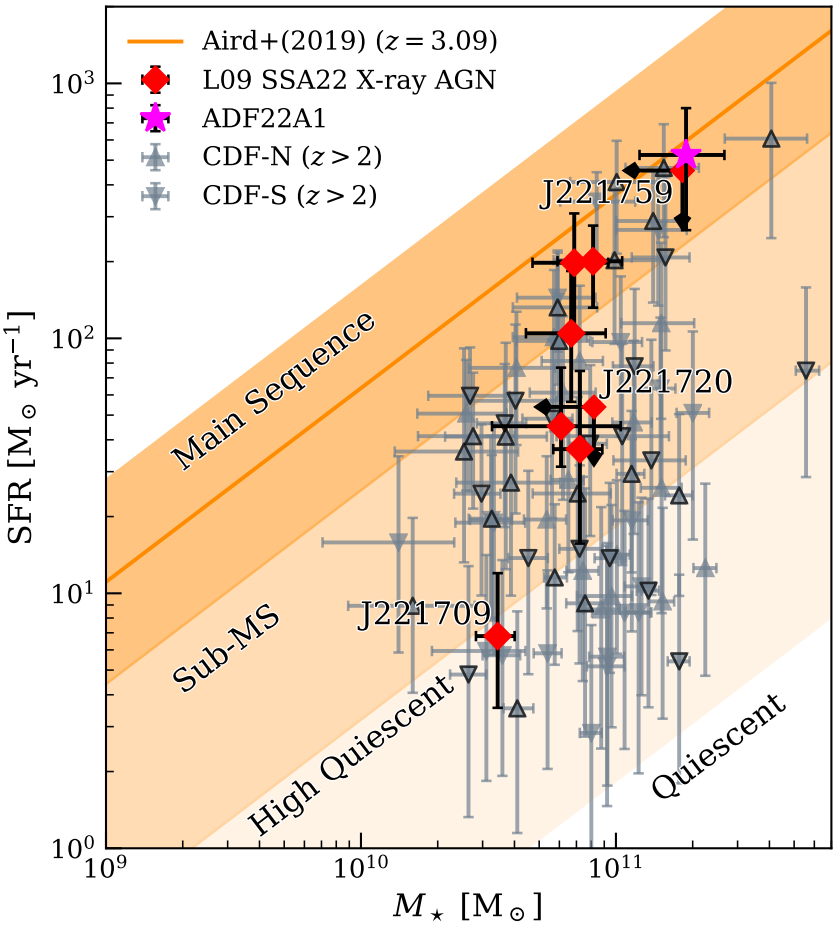

Figure 4.6: We show the SFR averaged over the last 100 Myr plotted against the total stellar mass for the SSA22 X-ray AGN, ADF22A1, and CDF AGN hosts with $z > 2$. The orange line shows the redshift-dependent star-forming main sequence from Aird et al. (2019) at $z = 3.09$, with the shaded regions marking the categories of star formation adopted in that work: main sequence (MS; $-0.4 < \log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq 0.4$), Sub-MS ($-1.3 < \log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -0.4$), high quiescent ($-2.3 < \log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -1.3$), and quiescent ($\log \mathrm{SFR}/\mathrm{SFR_{MS}} \leq -2.3$). We find that 3 of the 6 Lehmer et al. (2009a) X-ray AGN for which we constrain the SFH are categorized as sub-MS or below. The two AGN for which we cannot constrain the SFH are labeled and shown as 99th percentile upper limits. The lowest-SFR AGN in our SSA22 sample is also labeled. The CDF galaxies with spectroscopic redshifts are shown with dark-outlined symbols, and the galaxies with photometric redshifts are shown without dark outlines.

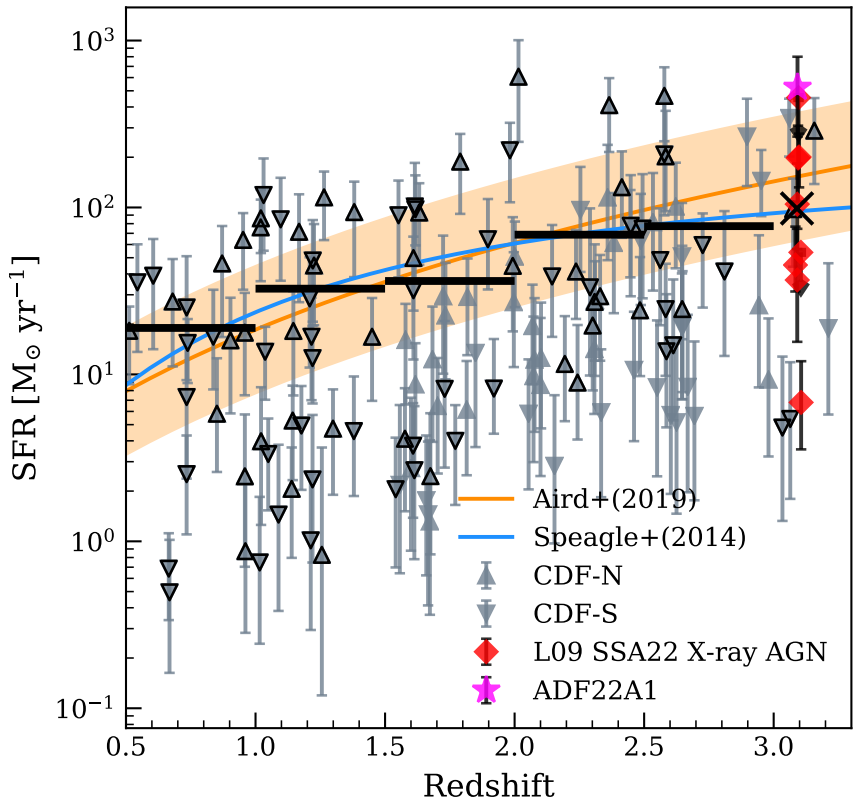

Figure 4.7: We show the SFR of all of our galaxies as a function of redshift, up to $z = 3.3$. CDF galaxies with secure spectroscopic redshifts are shown with dark-outlined symbols. Each galaxy's SFR is averaged over the last 100 Myr of the sampled posterior SFH. The mean SFR of the CDF galaxies is shown by the thick black bars over four ranges of redshift, and the mean SFR of the L09 X-ray AGN is shown as a black cross. The two AGN for which we cannot constrain the SFH are shown as 99th percentile upper limits. In the background we show the evolution of the star forming main sequence according to Aird et al. (2019) and Speagle et al. (2014) for a stellar mass of $10^{10.5}$ M$_\odot$, with the orange shaded region indicating the range of the Aird et al. (2019) main sequence for stellar masses $10^{10} - 10^{11}$ M$_\odot$. The average SFR of the SSA22 protocluster AGN is consistent with both the average SFR of the field AGN and the main sequence at $z = 3.09$.

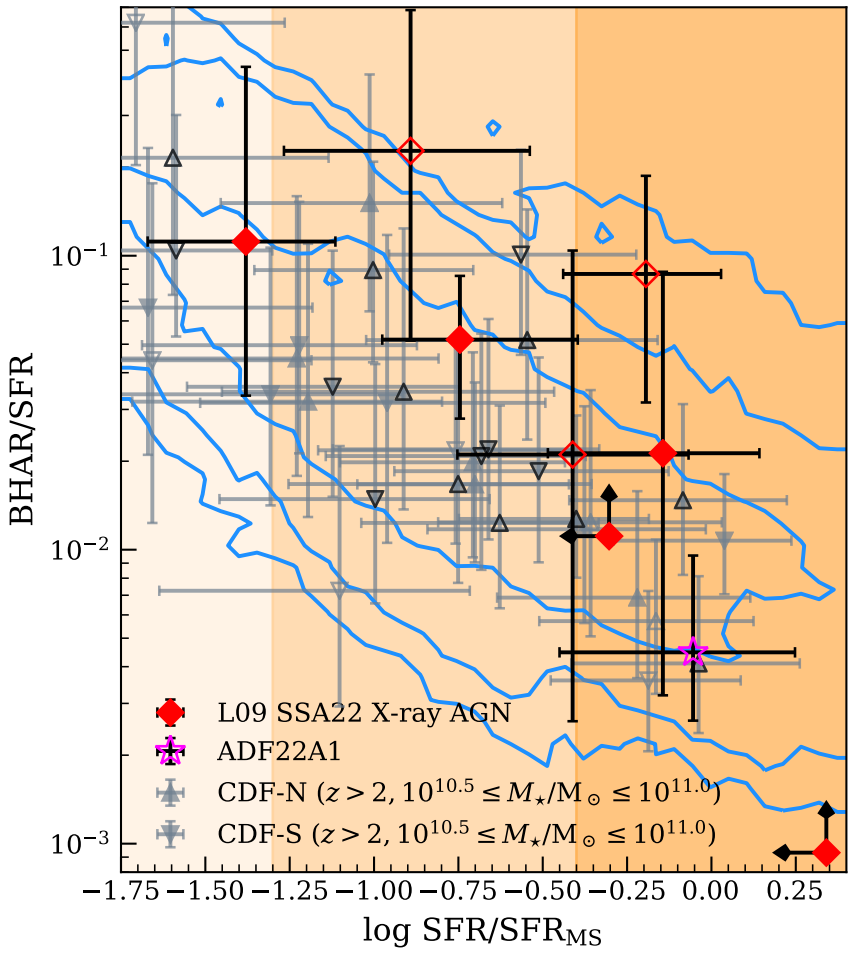

Figure 4.8: We show the BHAR/SFR ratio as a function of SFR/SFR$_{\rm MS}$ for the SSA22 sample and the field samples. The field samples are shown as points for galaxies with $10^{10.5} \leq \log(M_\star/{\rm M}_\odot) \leq 11.0$ (the range of stellar mass for SSA22), while the full set of CDF AGN candidates with $z \geq 2$ and $N_H < 10^{24}$ cm^{-2} are shown as contours, containing 68%, 95%, and 99% of the sample's posterior probability. The vertical shaded regions are the Aird et al. (2019) SFR/SFR$_{\rm MS}$ classifications: from left to right, high quiescent, sub-MS, and MS. CT candidates are shown with unfilled symbols, and field galaxies with spectroscopic redshifts are shown with dark outlined symbols. For the two SSA22 AGN where the AGN-only model is preferred, we show the 99th percentile of SFR/SFR$_{\rm MS}$ as an upper limit and the 1st percentile of BHAR/SFR as a lower limit. The trend in the SSA22 data is well captured by the trend in the field galaxies of the same mass.

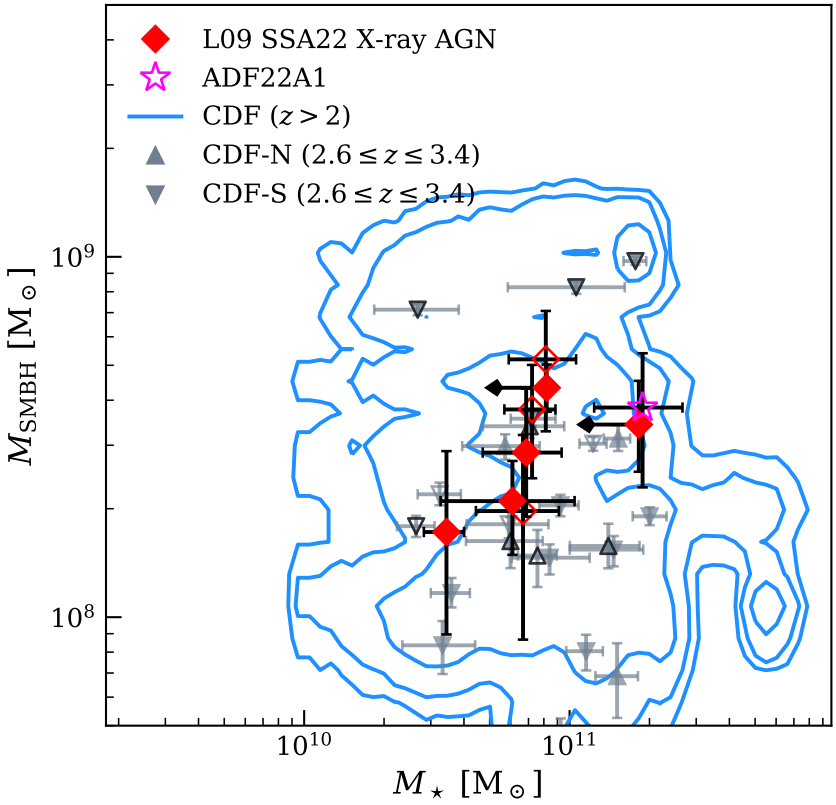

Figure 4.9: We show the SMBH mass of our sample galaxies as a function of stellar mass. As before, unfilled symbols represent CT candidates and CDF symbols with black outlines represent galaxies with spectroscopic redshifts. The $z \geq 2$ CDF galaxies with $N_H < 10^{24}$ cm^{-2} are shown as contours containing 68%, 95%, and 99% of the sample's posterior probability. For the two SSA22 AGN where the AGN-only model is preferred, we show the 99th percentile of the stellar mass as an upper limit. The SSA22 AGN and the CDF AGN with $2.6 \leq z \leq 3.4$ occupy the same region of the $M_{\rm SMBH}$–$M_\star$ plane, and the SSA22 AGN occupy a high-probability region of the $z \geq 2$ posterior, indicating that black holes have not grown out of pace with their hosts in the protocluster.

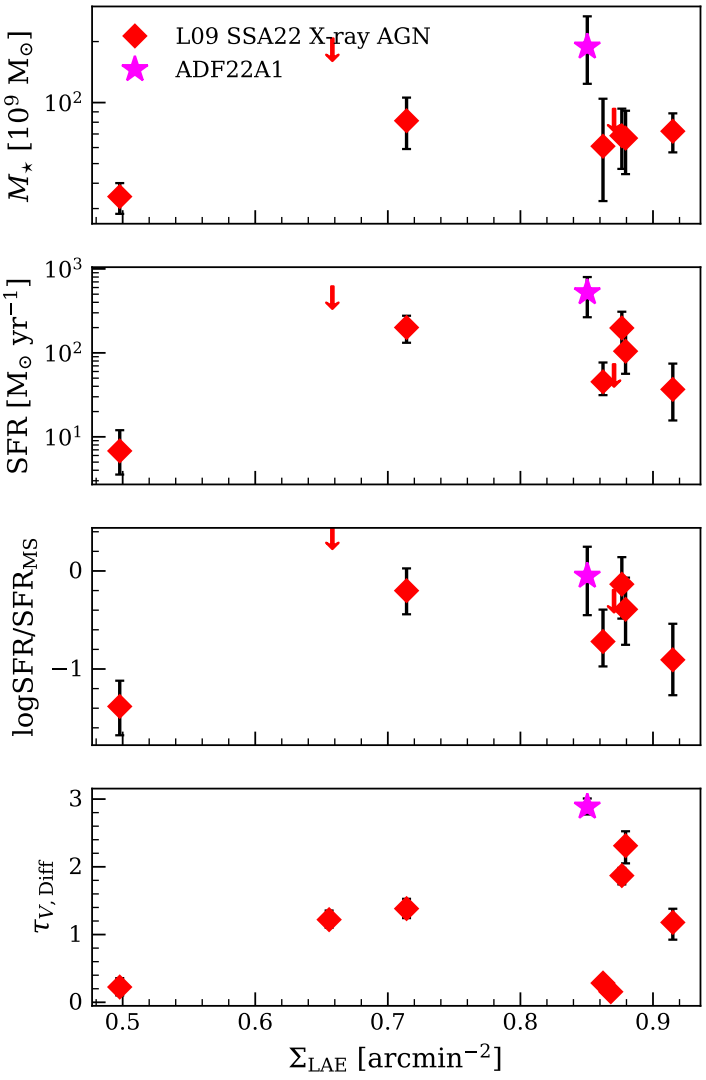

Figure 4.10: We show the SSA22 AGN sample host galaxy properties as a function of the projected surface density of $z \approx 3.09$ LAEs from Hayashino et al. (2004). For the two AGN where the AGN-only model is preferred, we show the 99th percentile upper limit for the SFH-derived properties. There are no significant trends observed with the surface density.

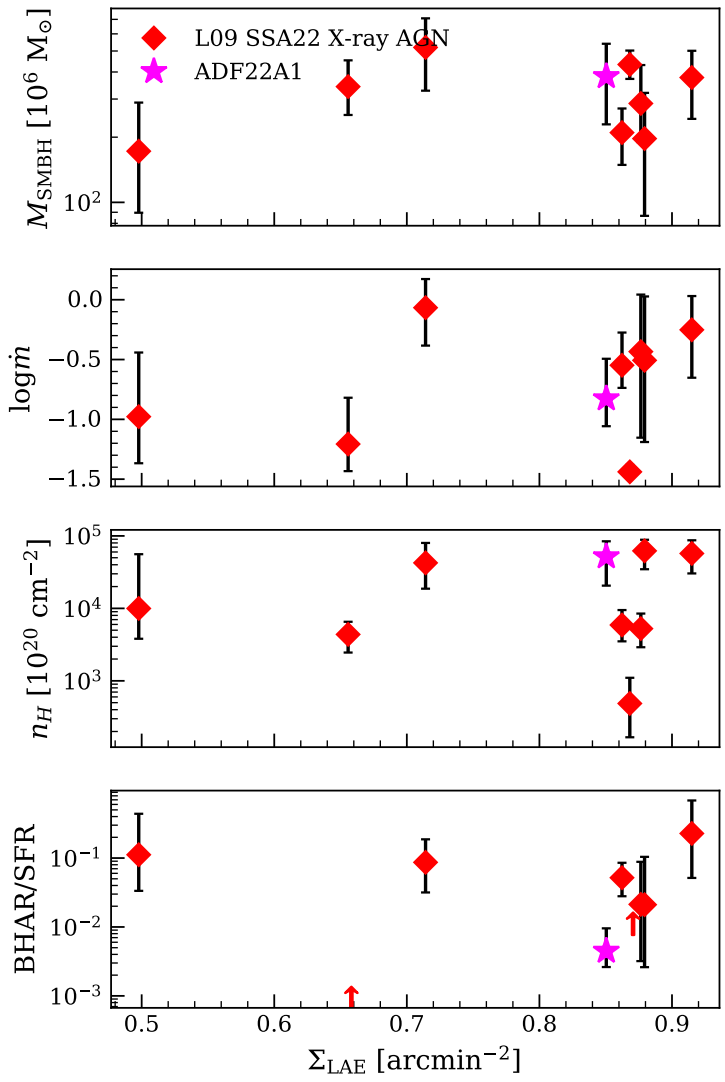

Figure 4.11: We show the SSA22 AGN sample SMBH properties as a function of the projected surface density of $z \approx 3.09$ LAEs from Hayashino et al. (2004). For the two AGN where the AGN-only model is preferred, we show the 1st percentile lower limit on BHAR/SFR. There are no significant trends observed with the surface density.

Appendix

4.A Alternative SED Fit for J221716.16+001745.8

We show the alternative SED fit for J221716.16+001745.8 in Figure 4.A.1 and the derived properties in Table 4.A.1, where we have fit using the ALMA Band 7 offset detection rather than the upper limit from the location of the Chandra detection. Our results for this galaxy are not sensitive to this change, with the exception that we recover a much better fit than when we use the upper limit. In the fit discussed in Section 4.3.4 and Section 4.4, the uncertainty on the IR contribution of the AGN model (compare the orange bands showing the range of the best 68% of models in Figure 4.4 and Figure 4.A.1) is higher due to the larger uncertainty on the IR emission overall. This results in the alternative fit here having a slightly different $\cos i$ and $\dot{m}$, though both are consistent with the original fit and the uncertainty in $\dot{m}$ due to the heavy X-ray obscuration is larger than the shift in $\dot{m}$ introduced by changing the IR data we use for the fit. The shape of the SFH is only slightly different, in the $10^7 - 10^8$ yr bin, resulting in a slightly smaller SFR, though well within the uncertainties.

4.B AGN-Only Model Fits for SSA22 X-ray AGN

To construct the AGN-only model fits, we set all SFH coefficients in the model to identically 0. The priors, assumptions, and free parameters for the fit are otherwise unchanged from Table 4.4.

Table 4.A.1: SED-fit derived properties for J221716.16+001745.8. The column meanings are the same as Table 4.5.

ID	$\log M_\star$ ($M_\odot$)	SFR ($M_\odot$ yr^{-1})	$\log \mathrm{SFR}/\mathrm{SFR}_{\mathrm{MS}}$ (dex)	τ_V	$\cos i$	$\log M_{\mathrm{SMBH}}$ ($M_\odot$)	$\log \dot{m}$	N_H (10^{22} cm^{-2})
J221716.16+001745.8	$\mathbf{10.9^{+0.1}_{-0.2}}$	$\mathbf{184.8^{+71.1}_{-71.0}}$	$\mathbf{-0.20^{+0.26}_{-0.29}}$	$\mathbf{1.4^{+0.2}_{-0.2}}$	$\mathbf{0.30^{+0.15}_{-0.18}}$	$\mathbf{8.8^{+0.1}_{-0.2}}$	$\mathbf{0.0^{+0.2}_{-0.3}}$	$\mathbf{460.35^{+349.95}_{-239.14}}$

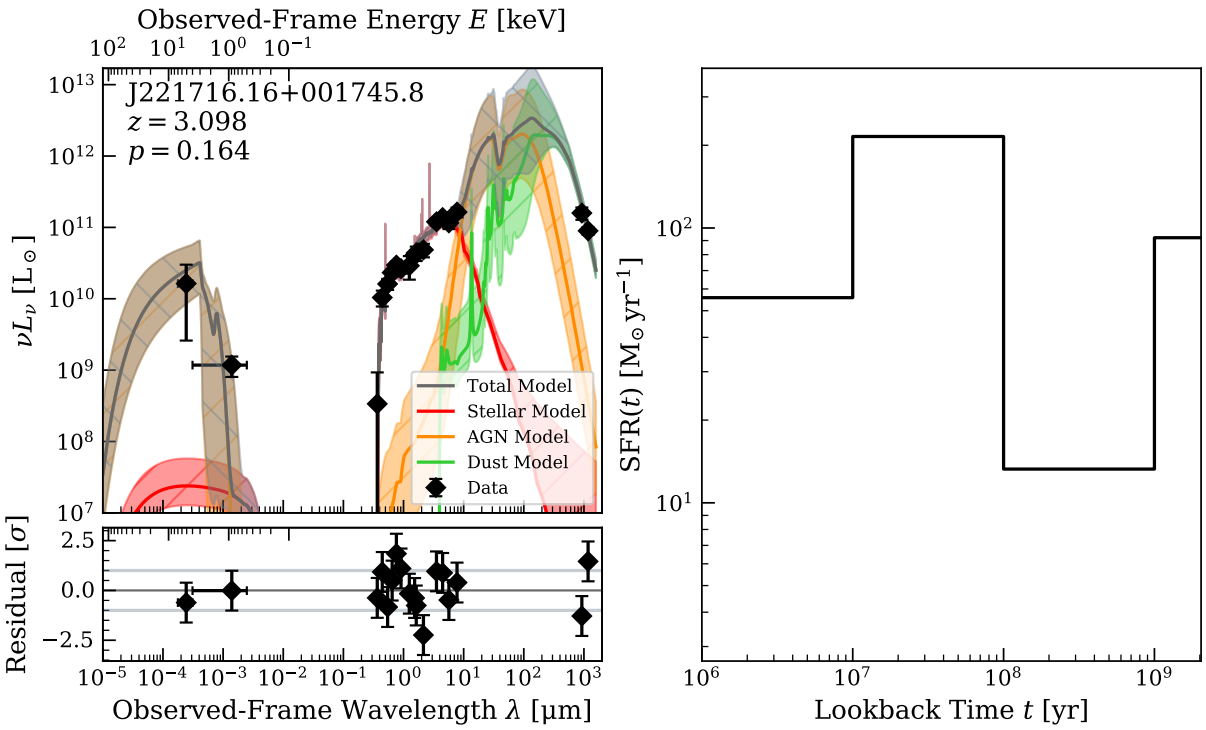

Figure 4.A.1: We show the alternative SED fit for J221716.16+001745.8. The meanings of the symbols and shaded regions are the same as the previous SED plots. We recover the same SFH shape and SFR when we fit this galaxy with the offset ALMA Band 7 detection, rather than the ALMA Band 7 upper limit at the position of the Chandra detection.

Table 4.B.1: AIC values and $F-$test $p-$values comparing models with stellar population emission and AGN-only models.

	AIC		
ID	w/ stellar pop.	w/o stellar pop.	$F-$test $p-$value
J221720.24+002019.3	25.34	19.56	0.5432
J221759.23+001529.7	24.35	17.83	0.6777

We perform the same PPC goodness-of-fit analysis on these fits as described in Section 4.3.4. For all but two AGN, J221720.24+002019.3, and J221759.23+001529.7, the AGN-only model is ruled out with $p < 0.001$. The best-fit AGN-only models for these two AGN are shown in Figure 4.B.1 and Figure 4.B.2.

To compare these fits to our original model, we use the Akaike information criterion (AIC) and $F-$test. AIC values and $F-$test $p-$values for the galaxies where the AGN-only model is not ruled out are shown in Table 4.B.1. As a heuristic, the model with lower AIC should be preferred. The $F-$test $p-$value represents the probability of the null hypothesis that the more complex model (in our case, the model including stellar population emission) does *not* provide a significantly improved fit over the simpler model (AGN emission only). We find that the AGN-only model with no stellar population is preferred in both cases.

The recovered $M_{\rm SMBH}$, $\log \dot{m}$, and N_H, are consistent with the original fits. However, to produce the required amount of optical emission, the AGN-only models necessarily have larger $\cos i$, indicating a less-obscured line-of-sight to the accretion disk, than the models including stellar population emission. While J221720.24+002019.3 is identified in the literature as a spectroscopically confirmed Type 1 AGN, this is not the case for J221759.23+001529.7. The morphology of J221759.23+001529.7 is also extended in F160W, indicating that it is unlikely that AGN continuum emission is dominant across the entire optical-NIR spectrum. Rather, the statistical preference

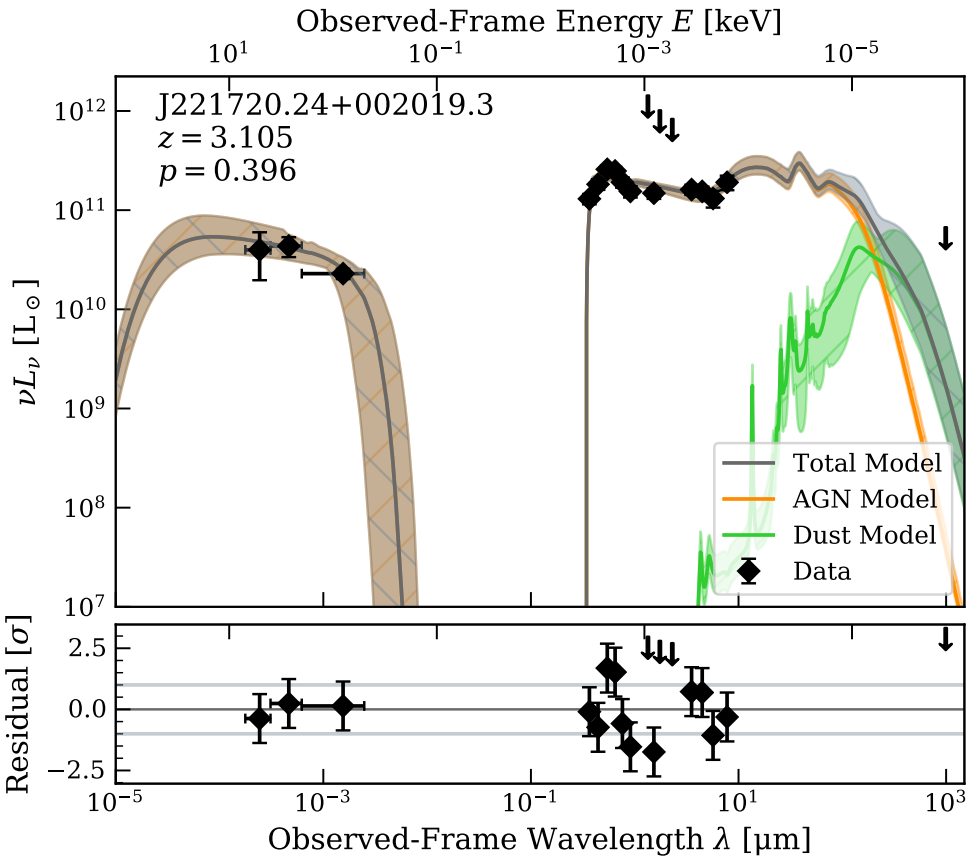

Figure 4.B.1: We show the best-fit AGN-only model for J221720.24+002019.3. All symbols and shaded regions have the same meaning as in the SED plots in Figure 4.4.

for the AGN-only model may be due to limited NIR constraints.

Notably, the AGN for which the AGN-only model is preferred are IR undetected. Our model includes ISM dust heating by the AGN (visible as the green curve in Figure 4.B.1 and Figure 4.B.2), but heating associated with the AGN produces comparatively little sub-mm emission. For the galaxies in our SSA22 sample with sub-mm detections, the sub-mm constraint gives us an effective handle on the level of star formation, ruling out solutions with zero star formation. J221720.24+002019.3 and J221759.23+001529.7 also have limited NIR constraints. For J221720.24+002019.3 we treat the *JHK* measurements as upper limits due to possible contamination, such that the stellar emission across the 4000 Å break is poorly constrained. For

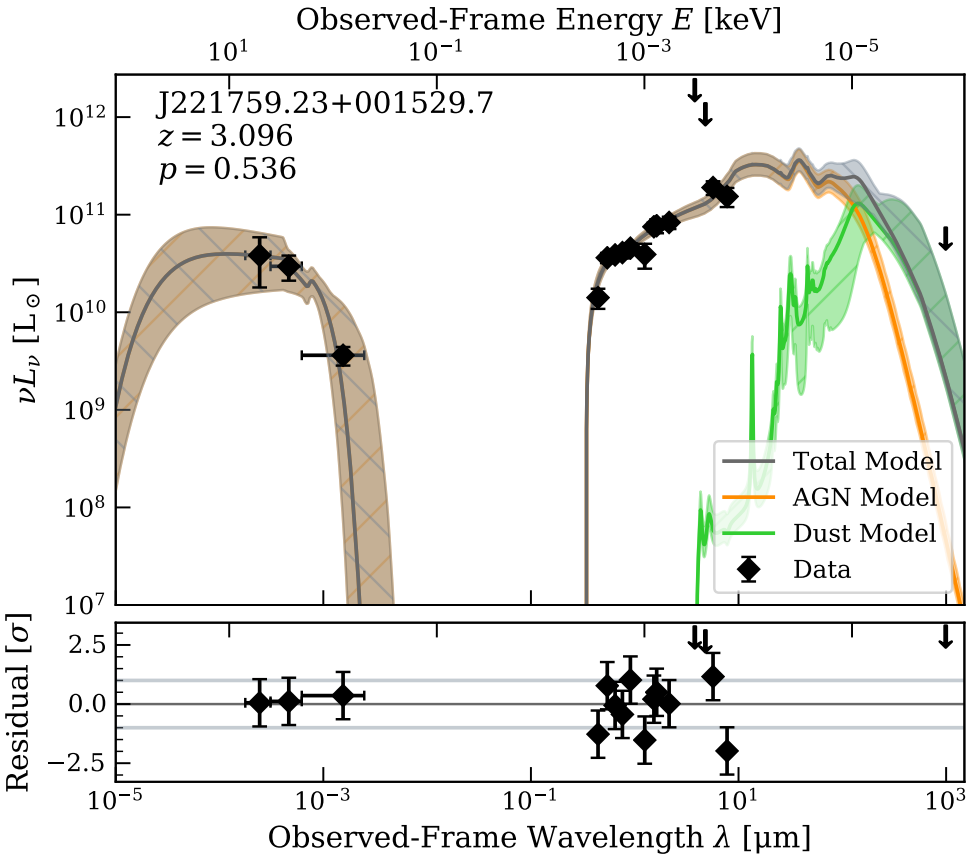

Figure 4.B.2: We show the best-fit AGN-only model for J221759.23+001529.7. All symbols and shaded regions have the same meaning as in the SED plots in Figure 4.4.

J221759.23+001529.7 we treat the *Spitzer* IRAC 3.6 μm and 4.5 μm measurements as upper limits, due to contamination by a foreground star, preventing us from clearly separating the tail of the stellar population emission from the AGN torus emission.

References

Aird, J., Coil, A. L., & Georgakakis, A. 2019, MNRAS, 484, 4360, doi: 10.1093/mnras/stz125

Alexander, D. M., Simpson, J. M., Harrison, C. M., et al. 2016, MNRAS, 461, 2944, doi: 10.1093/mnras/stw1509

Ao, Y., Matsuda, Y., Henkel, C., et al. 2017, ApJ, 850, 178, doi: 10.3847/1538-4357/aa960f

Barro, G., Pérez-González, P. G., Cava, A., et al. 2019, ApJS, 243, 22, doi: 10.3847/1538-4365/ab23f2

Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103, doi: 10.1051/0004-6361/201834156

Broos, P. S., Townsley, L. K., Feigelson, E. D., et al. 2010, ApJ, 714, 1582, doi: 10.1088/0004-637X/714/2/1582

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682, doi: 10.1086/308692

Chabrier, G. 2003, PASP, 115, 763, doi: 10.1086/376392

Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415, doi: 10.1093/mnras/stw1756

Digby-North, J. A., Nandra, K., Laird, E. S., et al. 2010, MNRAS, 407, 846, doi: 10.1111/j.1365-2966.2010.16977.x

Doore, K., Eufrasio, R. T., Lehmer, B. D., et al. 2021, ApJ, 923, 26, doi: 10.3847/1538-4357/ac25f3

Draine, B. T., & Li, A. 2007, ApJ, 657, 810, doi: 10.1086/511055

Eufrasio, R. T., Lehmer, B. D., Zezas, A., et al. 2017, ApJ, 851, 10, doi: 10.3847/1538-4357/aa9569

Fioc, M., & Rocca-Volmerange, B. 1997, A&A, 326, 950

—. 1999, arXiv e-prints, astro. https://arxiv.org/abs/astro-ph/9912179

Fitzpatrick, E. L. 1999, PASP, 111, 63, doi: 10.1086/316293

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306, doi: 10.1086/670067

Geach, J. E., Alexander, D. M., Lehmer, B. D., et al. 2009, ApJ, 700, 1, doi: 10.1088/0004-637X/700/1/1

Gehrels, N. 1986, ApJ, 303, 336, doi: 10.1086/164079

Gelman, A., Meng, X.-L., & Stern, H. 1996, Statistica Sinica, 6, 733

Gilbertson, W., Lehmer, B. D., Doore, K., et al. 2022, ApJ, 926, 28, doi: 10.3847/1538-4357/ac4049

Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65, doi: 10.2140/camcos.2010.5.65

Guo, Y., Ferguson, H. C., Giavalisco, M., et al. 2013, ApJS, 207, 24, doi: 10.1088/0067-0049/207/2/24

Hayashino, T., Matsuda, Y., Tamura, H., et al. 2004, AJ, 128, 2073, doi: 10.1086/424935

Johnson, B. D., Leja, J., Conroy, C., & Speagle, J. S. 2021, ApJS, 254, 22, doi: 10.3847/1538-4365/abef67

Just, D. W., Brandt, W. N., Shemmer, O., et al. 2007, ApJ, 665, 1004, doi: 10.1086/519990

Kato, Y., Matsuda, Y., Smail, I., et al. 2016, MNRAS, 460, 3861, doi: 10.1093/mnras/stw1237

Kennicutt, Robert C., J. 1998, ApJ, 498, 541, doi: 10.1086/305588

Kroupa, P. 2001, MNRAS, 322, 231, doi: 10.1046/j.1365-8711.2001.04022.x

Kubo, M., Yamada, T., Ichikawa, T., et al. 2015, ApJ, 799, 38, doi: 10.1088/0004-637x/799/1/38

Kubo, M., Uchimoto, Y. K., Yamada, T., et al. 2013, ApJ, 778, 170, doi: 10.1088/0004-637x/778/2/170

Kubo, M., Umehata, H., Matsuda, Y., et al. 2021, ApJ, 919, 6, doi: 10.3847/1538-4357/ac0cf8

—. 2022, ApJ, 935, 89, doi: 10.3847/1538-4357/ac7f2d

Kubota, A., & Done, C. 2018, MNRAS, 480, 1247, doi: 10.1093/mnras/sty1890

Lehmer, B. D., Alexander, D. M., Geach, J. E., et al. 2009a, ApJ, 691, 687, doi: 10.1088/0004-637X/691/1/687

Lehmer, B. D., Alexander, D. M., Chapman, S. C., et al. 2009b, MNRAS, 400, 299, doi: 10.1111/j.1365-2966.2009.15449.x

Lehmer, B. D., Lucy, A. B., Alexander, D. M., et al. 2013, ApJ, 765, 87, doi: 10.1088/0004-637X/765/2/87

Lehmer, B. D., Eufrasio, R. T., Markwardt, L., et al. 2017, ApJ, 851, 11, doi: 10.3847/1538-4357/aa9578

Lehmer, B. D., Eufrasio, R. T., Tzanavaris, P., et al. 2019, ApJS, 243, 3, doi: 10.3847/1538-4365/

ab22a8

Leja, J., Johnson, B. D., Conroy, C., & van Dokkum, P. 2018, ApJ, 854, 62, doi: 10.3847/1538-4357/aaa8db

Luo, B., Brandt, W. N., Xue, Y. Q., et al. 2017, ApJS, 228, 2, doi: 10.3847/1538-4365/228/1/2

Lusso, E., & Risaliti, G. 2017, A&A, 602, A79, doi: 10.1051/0004-6361/201630079

Magnelli, B., Popesso, P., Berta, S., et al. 2013, A&A, 553, A132, doi: 10.1051/0004-6361/201321371

Martini, P., Miller, E. D., Brodwin, M., et al. 2013, ApJ, 768, 1, doi: 10.1088/0004-637X/768/1/1

Matsuda, Y., Yamada, T., Hayashino, T., et al. 2004, AJ, 128, 569, doi: 10.1086/422020

—. 2005, ApJ, 634, L125, doi: 10.1086/499071

McAlpine, S., Harrison, C. M., Rosario, D. J., et al. 2020, MNRAS, 494, 5713, doi: 10.1093/mnras/staa1123

Monson, E. B., Lehmer, B. D., Doore, K., et al. 2021, ApJ, 919, 51, doi: 10.3847/1538-4357/ac0f84

Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793, doi: 10.1051/0004-6361/200912497

Nowak, M. A., Neilsen, J., Markoff, S. B., et al. 2012, ApJ, 759, 95, doi: 10.1088/0004-637X/759/2/95

Pérez-González, P. G., Egami, E., Rex, M., et al. 2010, A&A, 518, L15, doi: 10.1051/0004-6361/201014593

Predehl, P., & Schmitt, J. H. M. M. 1995, A&A, 293, 889

Rubin, D. B. 1984, The Annals of Statistics, 12, 1151

Saez, C., Lehmer, B. D., Bauer, F. E., et al. 2015, MNRAS, 450, 2615, doi: 10.1093/mnras/stv747

Sawicki, M. 2012, PASP, 124, 1208, doi: 10.1086/668636

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103, doi: 10.1088/0004-637X/737/2/103

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15, doi: 10.1088/0067-0049/214/2/15

Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288, doi: 10.1093/mnras/stw444

Steidel, C. C., Adelberger, K. L., Shapley, A. E., et al. 2003, ApJ, 592, 728, doi: 10.1086/375772

Tamura, Y., Kohno, K., Nakanishi, K., et al. 2009, Nature, 459, 61, doi: 10.1038/nature07947

Tamura, Y., Iono, D., Wilner, D. J., et al. 2010, ApJ, 724, 1270, doi: 10.1088/0004-637X/724/2/1270

Tozzi, P., Pentericci, L., Gilli, R., et al. 2022, arXiv e-prints, arXiv:2203.02208. https://arxiv.org/abs/2203.02208

Umehata, H., Tamura, Y., Kohno, K., et al. 2014, MNRAS, 440, 3462, doi: 10.1093/mnras/stu447

—. 2015, ApJ, 815, L8, doi: 10.1088/2041-8205/815/1/L8

—. 2017, ApJ, 835, 98, doi: 10.3847/1538-4357/835/1/98

Umehata, H., Fumagalli, M., Smail, I., et al. 2019, Science, 366, 97, doi: 10.1126/science.aaw5949

Vito, F., Brandt, W. N., Lehmer, B. D., et al. 2020, A&A, 642, A149, doi: 10.1051/0004-6361/202038848

Wilms, J., Allen, A., & McCray, R. 2000, ApJ, 542, 914, doi: 10.1086/317016

Xue, Y. Q., Luo, B., Brandt, W. N., et al. 2016, ApJS, 224, 15, doi: 10.3847/0067-0049/224/2/15

Yajima, H., Abe, M., Khochfar, S., et al. 2022, MNRAS, 509, 4037, doi: 10.1093/mnras/stab3092

Yamada, T., Matsuda, Y., Kousai, K., et al. 2012, ApJ, 751, 29, doi: 10.1088/0004-637X/751/1/29

Yang, G., Boquien, M., Buat, V., et al. 2020, MNRAS, 491, 740, doi: 10.1093/mnras/stz3001

Yang, G., Boquien, M., Brandt, W. N., et al. 2022, arXiv e-prints, arXiv:2201.03718. https://arxiv.org/abs/2201.03718

Chapter 5

Conclusion

5.1 The Current Understanding of Galaxy and SMBH Growth in Protoclusters

As it stands, the mechanisms driving the AGN fraction enhancement in high redshift protoclusters (and the apparently inverted relationship that it has with density in the local Universe) remain unclear. In Chapter 3, we have seen that the LBGs in the SSA22 protocluster, representing the "normal" star-forming galaxy population, are on average no more likely to be morphologically disturbed or observed in a merger than LBGs in mean-density environments at the same redshift. However, we have also seen that fitting their SEDs indicates that the protocluster LBGs are more massive on average, and that they may have experienced an earlier epoch of rapid growth at ages older than 100 Myr. In Chapter 4, we saw that both the host galaxies of the SSA22 X-ray AGN and the SMBHs themselves do not appear to be growing any more rapidly than their counterparts in mean-density fields, implying that any difference in the protocluster and field AGN populations is in the mechanism triggering an AGN phase. These results may be compatible in a scenario where the black hole masses of the protocluster galaxies are larger, on average, than the black hole masses of the field galaxies. Then, any given protocluster galaxy has a larger chance of being detected as an AGN, as accretion rate and AGN luminosity scale with black hole mass. However, by selecting X-ray AGN for our analysis, we sample only from the highest SMBH masses, and find that the growth of these high-SMBH-mass objects is broadly similar between the protocluster and field environment.

More direct tests of this explanation are a difficult prospect: we can only measure the mass of a SMBH if it is actively accreting enough to be observed as an AGN. Direct BH mass estimates

are also difficult in general, requiring, detection of broad emission lines in, ideally, multi-epoch spectroscopy. For obscured AGN and AGN at high redshift, this may be prohibitive, hence our reliance on indirect estimates of the SMBH mass in Chapter 4. Direct measurements of the SMBH mass for an inactive or low-accretion-rate SMBH are impossible. However, we may be able to constrain the black hole mass function of protoclusters with deep X-ray imagery (like SSA22) by estimating the accretion rate distribution function of the population. If we assume that each galaxy (above a given stellar mass at which we can be confident that we have detected every galaxy) contains an accreting SMBH, we can marginalize over the set of all possible accretion rates that are consistent with the X-ray data. By doing so for an entire population, we can push below the detection limit of our X-ray data to constrain the low end of the accretion rate distribution (as in Aird et al., 2019). If we were to find that this function is shifted to larger accretion rates than an equivalent constructed for the mean-density field, it could suggest that the AGN enhancement is powered by the SMBH mass enhancement of the entire population.

5.2 Future Prospects for the Study of Galaxy and SMBH Growth in Protoclusters

Measurements of galaxy morphologies and detection of merger features are both limited by the sensitivity and resolution of the telescopes we use. *JWST* makes an unprecedented improvement in both over previous generations of near-IR telescopes, including the near-IR capabilities of *HST*. We have previously seen in Chapter 3 that the existing *HST* imaging in the SSA22 protocluster at 1.6 μm provides only weak constraints on the merger fraction, and that our conclusions from visual analysis are limited by sample size and inconsistent depth. Uniform-depth *JWST* imaging could allow better constraint of the merger fraction among protocluster galaxies, as well as improved

Sérsic fits and nonparametric morphological measurements.

One of the primary limitations of the studies of protocluster galaxies presented above is the reliance on biased selections: the LBG selection in Chapter 3 favors young, massive galaxies, and the X-ray AGN selection in Chapter 4 favors even larger galaxies in the most violent phase of their growth. Our studies of the SSA22 galaxies thus miss a significant fraction of the protocluster population, at masses $\lesssim 10^9$ $M_\odot$. Understanding how these dwarf galaxies are growing can give us a more detailed picture of the evolution of the structure as a whole and the evolution of its most massive galaxies, which are expected to grow by accretion of smaller satellite galaxies. Building a sample that is complete to such low masses has not been possible to date, but will be enabled by *JWST*.

Protoclusters like SSA22 have been observed to have an overdensity of sub-millimeter galaxies: extremely high-SFR galaxies containing significant amounts of dust that obscures their optical signatures. ADF22A1, detected as a 1.1 mm source and believed to be a complex system containing a Compton-thick AGN (Tamura et al., 2010), makes the case that these violently-growing systems may harbor hidden AGN. Compton-thick systems, like ADF22A1 is believed to be, are difficult to detect even with X-ray observations. As such, we may be missing a hidden population of AGN in the protocluster due to our reliance on X-ray selection. Again, *JWST* may improve matters: as obscured AGN radiate strongly in the mid-IR, we can use JWST to select obscured AGN by IR colors (e.g. Kirkpatrick et al., 2017) or by SED fitting, which allows us to leverage existing X-ray-to-IR data (e.g. Zou et al., 2022). SMBHs are expected to grow rapidly in this obscured phase, thus selecting obscured protocluster AGN will give us better constraints on quantities important to understanding protocluster galaxy-SMBH coevolution, like the average BHAR/SFR ratio of the protocluster population.

Protoclusters can be blindly selected by clustering analysis of dropout-selected galaxies (such as LBGs). Surveys like GOLDRUSH have previously generated large samples of LBGs from $z \approx 4-6$ and used clustering techniques to select protocluster candidates (Toshikawa et al., 2018). The deep drilling fields (DDF) of the Legacy Survey of Space and Time (LSST), which will reach up to a magnitude deeper than the GOLDRUSH data in the same filters, will reveal even fainter LBGs (though the repeated observations to reach full depth in the DDFs will be a $\sim$ 10 year project). Another option for relatively unbiased protocluster discovery involves large-angle observations with narrowband filters, which are constructed to target the Lyman-α line in a narrow range of redshift and thus detect Lyman-α emitters, low-mass galaxies with strong Lyman-α emission that are know to give an unbiased tracing of large-scale structure. The SILVERRUSH project has undertaken a large scale narrowband search for protoclusters (Higuchi et al., 2019), as has the ODIN project (survey description paper in prep., but preliminary results are in Ramakrishnan et al., 2023). These surveys have complementary redshift ranges, which should enable the study of how protoclusters and filamentary structures evolve from $z = 7$ to $z = 3$. These surveys are constructed to overlap over large parts of their area with existing X-ray coverage (both *Chandra* and *XMM-Newton*) and deep optical coverage, including future observations in the LSST. Prospects for large samples of consistently- and blindly-selected protocluster candidates are thus good, and may remove some of the current limitations on sample size that hinder current studies, allowing us to better study protoclusters as a population, and to make stronger comparisons between galaxy populations in and out of protoclusters.

References

Aird, J., Coil, A. L., & Georgakakis, A. 2019, MNRAS, 484, 4360, doi: 10.1093/mnras/stz125

Higuchi, R., Ouchi, M., Ono, Y., et al. 2019, ApJ, 879, 28, doi: 10.3847/1538-4357/ab2192

Kirkpatrick, A., Alberts, S., Pope, A., et al. 2017, ApJ, 849, 111, doi: 10.3847/1538-4357/aa911d

Ramakrishnan, V., Moon, B., Hyeok Im, S., et al. 2023, arXiv e-prints, arXiv:2302.07860, doi: 10.48550/arXiv.2302.07860

Tamura, Y., Iono, D., Wilner, D. J., et al. 2010, ApJ, 724, 1270, doi: 10.1088/0004-637X/724/2/1270

Toshikawa, J., Uchiyama, H., Kashikawa, N., et al. 2018, PASJ, 70, S12, doi: 10.1093/pasj/psx102

Zou, F., Brandt, W. N., Chen, C.-T., et al. 2022, ApJS, 262, 15, doi: 10.3847/1538-4365/ac7bdf