Observing Galaxy Mergers at the Epoch of Reionization

Evgenii A. Chaikin¹, Nadezda V. Tyulneva¹, and Alexander A. Kaurov²

¹ Peter the Great St. Petersburg Polytechnic University, Politechnicheskaya 29, 194021, St. Petersburg, Russia
² Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540, USA; kaurov@ias.edu

Received 2017 May 4; revised 2017 November 22; accepted 2017 December 11; published 2018 January 24

Abstract

The galaxies with photometric redshifts observed in a close angular proximity might be either projection coincidences, strongly lensed images of the same galaxy, or separate galaxies that are in a stage of merging. We search for the groups of galaxies in the Hubble Ultra Deep Field (HUDF09) in z ∼ 7 and z ∼ 8 drop-out samples. We find no close pairs among 50 galaxies in the z ∼ 7 sample, while in the z ∼ 8 sample we find that 6 out of 22 galaxies have a companion within ∼1″ (3 pairs). Adopting a numerical simulation and performing forward modeling, we show that even though mergers are unlikely to have such a high fraction, the projection coincidences and the strong lensing are even less likely mechanisms to account for all of these pairs. Alternatively, there is a possibility of the contamination in the drop-out catalog from lower redshifts, which potentially can account for all of the groups. Finally, we make projection on the sensitivity to mergers of the James Webb Space Telescope (JWST), and discuss the possible applications of the high-redshift merging galaxies for decreasing cosmic variance effects on the luminosity function and for improving the accuracy of photometric redshifts in general.

Key words: galaxies: high-redshift – galaxies: photometry

1. Introduction

The next decade will be rich for new space observatories sensitive to the infrared, such as the James Webb Space Telescope (JWST; Gardner et al. 2006) and the Wide Field Infrared Survey Telescope (Spencer et al. 2015), which will allow us to greatly extend the observational horizon well into the epoch of reionization, and the ground-based instruments like the Giant Magellan Telescope (Johns et al. 2012) would provide greater angular resolution in infrared. The main statistics that will be immediately extracted from the deep galaxy surveys is the luminosity function and the average star formation rate (e.g., Bouwens et al. 2015). In this paper, we study another piece of information that can be potentially contained in the existing and upcoming imaging data: the statistics of galaxy mergers at the epoch of cosmic reionization.

The fraction of mergers is well studied at redshifts below ∼4 (i.e., Conselice et al. 2008; López-Sanjuan et al. 2009; Lotz et al. 2011; Man et al. 2016) and was compared with simulations (i.e., recent comparison with the Illustris simulation Snyder et al. 2017). See Rodriguez-Puebla et al. (2017) for a combined review of available data on merger observations. Also, the fraction of clumpy galaxies was studied up to redshifts 8 (see compilation of various studies in Shibuya et al. 2016). The large number of both spectroscopic and photometric galaxies at low redshifts allows one to make magnitude cuts and study major or minor mergers separately. At the redshifts of reionization, on the contrary, there are much less observed galaxies, especially spectroscopically confirmed. Nevertheless, in this paper we attempt to see what is observed at z ∼ 7–8 and how consistent is it with the numerical simulations.

The galaxies at redshifts beyond 6 are expected to be compact and their detailed morphology cannot be presently resolved (e.g., Liu et al. 2017). However, individual blobs (which appear to be gravitationally bound) can be observed. There is already a handful of Lyα emitters (LAEs) observed at high redshifts that exhibit either multiple clumps, merging, or have another galaxy nearby: “Himiko” object has three clumps that may correspond to a triple merger at z ≈ 7 (Ouchi et al. 2009, 2013); “CR7” (COSMOS Redshift 7) at z = 6.6 that exhibits three distinct clumps with significantly different photometric SEDs (Sobral et al. 2015); A1689-zD1 is a potential dusty merger at z ≈ 7.5 (Watson et al. 2015; Knudsen et al. 2017); and the lensed object at z = 6.3 that has two distinct clumps and three images (Ryder et al. 2017).

However, Lyα line is detected only for a fraction of objects. In the absence of spectroscopic redshifts, we have only photometric estimates with much lower precision and accuracy, especially at high redshifts. Therefore, the projection uncertainty fundamentally limits the search of grouped of objects. Nevertheless, we show that there is still some signal that can be extracted. In this paper, we focus on the Hubble Ultra Deep Field 2009 (HUDF09, Beckwith et al. 2006) and adopt the Hubble eXtreme Deep Field legacy data (XDF, Illingworth et al. 2013) and consider drop-out samples of Lyman-break Galaxies at z ∼ 7 and z ∼ 8 from B15. We show that even without spectroscopic redshifts there is a substantial statistical signal in the data (Section 2), and using one group as an example we argue that, indeed, it can be identified as one system (Section 2.1).

The theoretical considerations (Lacey & Cole 1993) also predict a significant fraction of dark-matter halos and, therefore, galaxies to be in some stage of merging. To compare theoretical predictions with observations, we perform forward modeling using numerical simulations of cosmological size boxes with galaxy formation (Section 3.1). We mimic the Hubble Space Telescope’s (HST’s) Wide Field Camera 3 (WFC3), including noise and point-spread functions, and generate mock observations with a depth similar to the XDF (Section 3.2). The mock images are then processed with common pipelines and source catalogs are created
(Section 3.3). In Section 4, we summarize the results of our modeling and compare it with other possible mechanisms that can create observed grouped galaxies. Finally, in Section 5, we discuss the possible application of the grouped galaxies.

2. Observations of Mergers

In this study, we consider the HUDF09 and the XDF project data release, as it is likely to be the deepest imaged region until the JWST will start taking measurements. Also, this field is well imaged by ground-based telescopes and the photometry is conveniently available (Brammer et al. 2012). The Frontier fields (with images of high-redshift galaxies magnified by foreground clusters) effectively provide even deeper imaging; however, for this study we avoid the complications caused by lensing effects.

We adopt the source catalog used to determine the luminosity function at $4 < z < 10$ in B15 which is available online. Another catalog with open access—the HST 3D survey (Skelton et al. 2014)—also covers this field, but it differs in the way that sources are detected, as well as the photometry measurements and final photometric redshifts. We do not use the 3D HST redshift catalog because by construction it assumes the upper limit for the photometric redshifts to be 6, i.e., the objects that are determined as $z > 6$ in the B15 catalog are $z < 6$ in the 3D HST catalog. Also, there is a catalog by Mclure et al. (2011) that covers this region and redshift range. We discuss it in Section 4.4.

The galaxies in B15 catalog have two redshifts. The first one is the photometric redshift evaluated with template fitting code EAZY (Brammer et al. 2008) using the photometric fluxes in different filter bands. The second is the redshift bin based on drop-out technique. We adopted $z \sim 7$ and $z \sim 8$ samples that correspond to the drop-outs from F850LP and F105W HST filters (Bouwens et al. 2011). By its nature, the photometric and drop-out redshifts are much less precise than spectroscopic redshifts and have big uncertainties; therefore, it is not possible to accurately measure the spatial distance between two galaxies with photometric redshifts.

The effect of photometric redshifts on the identifying mergers is studied at lower redshifts (e.g., López-Sanjuan et al. 2015). In case of redshifts beyond 6, the main problem is likely to be not the uncertainty of photometric redshifts (if the probability density function of the redshift is known, one can marginalize over it), but the so-called catastrophic outliers, i.e., when a low-redshift dusty galaxy is misidentified with a high-redshift galaxy (see further discussion in Section 4.4).

The area of the deepest part of the XDF (which is slightly smaller HUDF09) is $\sim 1/9 \times 2/1$ that roughly corresponds to a $4 \times 4 \times 650 \ (h^{-1} \text{Mpc})$ pencil beam at redshifts from 6 to 9. The total number of detected sources in the field at all redshifts is $\sim 7000$ (Illingworth et al. 2013); among them are 50 objects with redshifts $z \sim 7$, and 22 with $z \sim 8$ (B15). As we show in Section 4, the former sample agrees well with the simulations and does not include any groups (defined by 1" threshold), while the latter sample does include groups and therefore it is of greater interest.

Among 22 galaxies at $z \sim 8$, there are three groups (with six galaxies in total), where the distance between galaxies is less than $\sim 1"$ (or $\sim 30 \ h^{-1} \text{kpc}$). This separation criteria is quite arbitrary, and in this particular case it is motivated by the fact that there are no groups with separation $1"-2"$ in this sample (in Section 4.3 and Figure 7 we vary the separation threshold). In Figure 1, these groups in different filters are presented. Also, there are a few cases where a galaxy with $z > 5$ is in a close proximity, which can potentially be a clue for a misidentified redshift.

Even though we use the term “merger” throughout the paper, its definition is quite ambiguous. When two blobs are observed within 1" (projection corresponds to roughly $\sim 30 \ h^{-1} \text{kpc}$), there is no guarantee that they are merging. Alternatively, they can be in a pre-merging state. Therefore, by “mergers” we mean the objects will occasionally merge.

Overall, it is safe to say that even the existing imaging data show the possibility of observing mergers at high redshifts. However, it is hard to quantitatively estimate whether we underestimate or overestimate the total fraction of mergers. The foregrounds definitely decrease the number of observed high-redshift galaxies, and possible catastrophic photometric redshift errors can decrease as well as increase the total number of high-redshift galaxies and the merger fraction.

2.1. Example of a Group

We select one of the groups and study its photometric redshifts in more details to be sure that it can indeed be a merger. The system contains three objects that we mark as A, B, and C. To the best of our knowledge, none of these objects has a spectroscopic redshift. We adopt photometric catalog by 3D HST (Momcheva et al. 2016) and rerun EAZY code to extract the probability density functions. We use the settings and the templates identical to those used in 3D HST except for two things: (1) we increase the hard limit on the maximum redshift from 6 to 12 and (2) we do not apply any prior on the luminosity function. In Figure 2, the cutout of the system, photometry, and the redshift probability density functions, $P(z)$, are presented.

The photometric redshifts for these galaxies in B15 catalog, $z_{\text{Bouwens}}$, and our redshifts of best fit (smallest $\chi^2$), $z_{\text{out}}$, and weighted peak redshift, $z_{\text{peak}}$, are shown in Table 1. The values differ due to different photometry, templates, etc.; however, the object A is identified as a lower redshift galaxy compared with B and C. The width of probability density functions of photometric redshifts (see the bottom left panel in Figure 2) shows that in fact all three galaxies can be at the same redshift. The probability of a simple projection coincidence for two objects, B and C, being nearby is $\sim 5\%$ (see Section 4.1). An alternative interpretation of the group is a strong lens. In this particular case, A is unlikely to be an image of the same object as B and C because it is distinctly visible in F850LP and F775 filters, while B and C are not. Therefore, we explore a possibility of B and C being the images of the same object. Figure 3 shows what point mass at which redshift is necessary to achieve observed separation of $\sim 1"$, and whether such a lens (in case if it is an elliptical galaxy) could be not detected by the HST. The estimation of the sensitivity curve is similar to the one made in Tsai et al. (2015). The conclusion is that it is possible to “hide” the lens.

---

3 http://cdsarc.u-strasbg.fr/viz-bin/Cat?J/ApJ/803/34

4 The corners of the XDF field are taken as ([03:32:45; -27°46′44″495′′], [03:32:48; -27°45′34″68′′], [03:32:32; -27°47′13″72′′], [03:32:39; -27°48′ 25″95′′]) as cited at https://archive.stsci.edu/prepds/xdf/.

5 See Brammer et al. (2008) for details regarding $z_{\text{out}}$ and $z_{\text{peak}}$. 
This system illustrates the uncertainty associated with identifying the nature of such groups. Here, all three possibilities—actual mergers, projection, and strong lens—can take place; however, as we will show in Section 4, the mergers are a slightly more probable interpretation of observing multiple groups.

### 3. Numerical Methods

#### 3.1. Numerical Simulations

To properly model the number of mergers and their separation, we use a numerical simulation with all the necessary physics—hydrodynamics and star formation—included, and where the times of merging are fully modeled. Alternatively, one can perform the same study with a dark-matter-only simulation or even just with a mock-halo catalog and a halo-occupation distribution model. However, here the dynamic timescales that correspond to the merging of galaxies, not only dark-matter halos, might be important. While these timescales can be considered analytically, e.g., Simha & Cole (2017) or Hong et al. (2016), still, such methods are approximate.

We adopt the suite of simulations of the cosmic reionization (Gnedin 2014; Gnedin & Kaurov 2014), in which the spatial resolution of the adaptive grid reaches 100 proper pc. We use a $40 \, h^{-1} \, \text{Mpc}$ box, for which we have only a few snapshots, and a series of ten $10 \, h^{-1} \, \text{Mpc}$ simulations (with the initial conditions drawn randomly with a random DC mode (Gnedin et al. 2011)). Smaller boxes allow us to have better resolution in time, while the larger box is used to confirm our results. The $40 \, h^{-1} \, \text{Mpc}$ box is relatively small compared with some other modern simulations, but it still has a greater statistical power than the volume of the XDF.

#### 3.2. Mimicking the Observations

Because the simulations incorporate star formation, we can calculate the fluxes directly. To properly evaluate the flux for each filter, we first calculate the spectrum for each star particle using Flexible Spectral Population Synthesis code (Conroy et al. 2009) and then apply appropriate redshift and the IGM absorption corrections (we adopt an updated Madau (1995) model by Inoue et al. 2014). The dust extinction of the dim high-redshift galaxies is expected to be low at $z \gg 7$ (e.g.,

![Figure 1. Cutouts of galaxies from $z \sim 8$ (red squares) drop-out sample with separation less than 3″. Only galaxies with $z > 4$ are labeled with cyan circles and corresponding photometric redshift. All panels are 6 arcsec across, and the colormap in all panels spans from $-1\sigma$ to $+2\sigma$. The galaxy in the last cutout with photometric redshift $z = 7.75$ belongs to $z \sim 7$ drop-out sample. Coordinates of groups from up to bottom (03:32:44.7; $-27^\circ46'44''$), (03:32:40; $-27^\circ47'17''$), (03:32:43; $-27^\circ46'27''$).](image-url)
The bright galaxies are affected more; however, it does not suppress their detectability. Therefore, in this study we assume the dust effect to be negligible.

Initially, we produce a projection with a resolution higher than the resolution of the detectors by factor of a few, then we apply the point-spread function corresponding to a given filter and instrument. To mimic the noise of the HST, we manually extract the empty regions (without any objects detected) from the legacy data of the deepest of the deep fields: XDF (Illingworth et al. 2013). Then, we evaluate the noise pattern (the amplitude and smoothness, that correspond to the point-spread function) that we later use for generating noise. Alternatively, we tried to use the noise cutouts and got identical results.

For the JWST, we used the demo version of “JWST Exposure Time Calculator” which allows us to generate noise for a given filter, observation time, background level, and readout pattern. For our estimates, we used 30-hr observations per filter and a “low” background noise. We found that other parameters do not significantly affect our results. A detailed study of different parameters and strategies is beyond the scope of this paper.

Bouwens et al. 2014). The bright galaxies are affected more; however, it does not suppress their detectability. Therefore, in this study we assume the dust effect to be negligible.

Initially, we produce a projection with a resolution higher than the resolution of the detectors by factor of a few, then we apply the point-spread function corresponding to a given filter and instrument.

To mimic the noise of the HST, we manually extract the empty regions (without any objects detected) from the legacy data of the deepest of the deep fields: XDF (Illingworth et al. 2013). Then, we evaluate the noise pattern (the amplitude and smoothness, that correspond to the point-spread function) that we later use for generating noise. Alternatively, we tried to use the noise cutouts and got identical results.

For the JWST, we used the demo version of “JWST Exposure Time Calculator” which allows us to generate noise for a given filter, observation time, background level, and readout pattern. For our estimates, we used 30-hr observations per filter and a “low” background noise. We found that other parameters do not significantly affect our results. A detailed study of different parameters and strategies is beyond the scope of this paper.

Table 1

|       | A        | B        | C        |
|-------|----------|----------|----------|
| $z_{\text{Bowens}}$ | 5.82     | 7.49     | 7.49     |
| $z_{\alpha}$      | 6.17     | 7.41     | 8.02     |
| $z_{\text{peak}}$ | 6.17     | 6.55     | 8.01     |

Figure 2. Top left panel: cutout (3:32:47.9; $-27^\circ 46' 53''$) of a group from the science image combined from F125W, F140W, and F160W filters (Momcheva et al. 2016) with size of 6 arcsec across (colormap spans from $-3\sigma$ to $2\sigma$). This system corresponds to the first row in Figure 1. Right panel: spectral energy distributions for each object. Solid lines are the best fit evaluated by EAZY code (Brammer et al. 2008), yellow points are theoretically predicted flux values and red points are the actual photometric data taken from 3D HST survey (Momcheva et al. 2016). Bottom left panel: corresponding redshift probability distribution $P(z)$ for each of three objects (color coded) and combined probability (solid line).

Figure 3. Maximum separation in arcseconds for two images (solid lines) of an object at $z \approx 7.5$ for a lens at given redshift and with given mass. Detection limit of unresolved lenses for HST WFC3 in F160W filter (dashed line) in the assumption of an elliptical galaxy SED (Polletta et al. 2007) and $M/L_B \sim 5$, following the procedure used in Tsai et al. (2015).

In the final step, we downsample our projections and add them to the noise. We show the projection of the gas density in a 100 kpc box at $z = 7.7$ in Figure 4 and then the
corresponding mock observation by *HST* and *JWST* in Figure 5. For illustrative purposes, the region shown in the figures is selected to be very overdense and corresponds to the proximity of the brightest galaxy in 80 $h^{-1}$ Mpc box at $z = 8$. Because the 80 $h^{-1}$ Mpc box is almost 100 times larger than the effective XDF volume (in the range $7 < z < 10$), such a system is unlikely to be among observed.

3.3. Processing the Observations

To extract the flux from an image, we adopt the widely used SExtractor (Bertin & Arnouts 1996) and an alternative instrument “photutils” (an associated package of “astropy”; Bradley et al. 2016), which provides similar functionality. These software allow one to find an object on a noisy image; determine its size and shape; and then derive its flux. We tried both methods and got similar results when used with similar parameters. We found our results to be more sensitive to the choice of parameters rather than to the software. The results presented below are made with use of photutils.\(^5\)

The objects were detected in the combination of F125W, F140W, and F160W filters for the *HST*, following the common practice (e.g., 3D *HST*; Skelton et al. 2014), and in the F115W, F150W, and F200W combination for the *JWST*. The result of the source detection is shown in the Figure 5. It is immediately apparent that the *JWST* is superior and in 30 hr of observation is capable to improve over the *HST*’s deepest image. The number count mostly increases because of the depth, and the smaller point-spread function accounts only for $\sim 1\%$ of additional detected objects.

Finally, we take into account that drop-out samples $z \sim 7$ and 8 have some distribution in redshift space (see Figure 4 in Bouwens et al. 2011). To do so, we converge this distribution with the redshift dependent results we got from the numerical simulation.

In Figure 6, the total number of observed sources is plotted. The simulations do match the observations as expected because the adopted simulation has been already shown to match the observed luminosity functions of galaxies combined from multiple deep fields (Gnedin & Kaurov 2014). This agreement confirms the legitimacy of the approach.

The foreground galaxies obscure some parts of the field. Therefore, the XDF data that we present here underestimates the actual number of galaxies. Also, notice that due to the photometric redshift uncertainties there should be scatter along $x$-axis, which is not included here.

Even though we tried to follow the source extraction approaches as used in the 3D *HST* survey, we still had some flexibility in the parameters. This flexibility can cause 10%–20% change in the number of the detected sources. We chose parameters to reproduce the observed number counts identically, therefore the match with the data in the Figure 6 is almost perfect.

4. Results

We have shown that there are groups of galaxies in drop-out $z \sim 8$ sample in Bouwens et al. (2015) catalog, while there are none of such groups in $z \sim 7$ sample. We summarize all possible explanations in the subsections below, and the corresponding probabilities are listed in Table 2. Overall, we show that while all three mechanisms are unlikely to individually account for all of three pairs, a combination of

\(^5\) The photutils package provides the functionality of separating objects with multiple blobs into individual, which we did not use in this study.
Figure 6. Number of objects in the XDF field at redshifts above $z$ (black curve) with scatter corresponding to the Poisson noise and cosmic variance (shaded region). For this figure only, we adopted the photometric redshifts derived with template fitting code EAZY from Bouwens et al. (2015), while in the rest of the paper we use drop-out samples from the same paper. The blue curve is the result of the forward modeling mimicking the depth of the XDF, and therefore it matches the data. The red curve is the forecast of the number of objects the JWST will be capable of observing in the same field (assuming 30-hr integration per each of these filters: F115W, F150W, and F200W).

Figure 7. Cumulative fraction of galaxies that have a companion within given angular distance (distance in $h^{-1}$ kpc given for $z = 7.5$) for $z \sim 7$ and $z \sim 8$ drop-out samples along with the results from simulations. The error region (99.7% region) is shown only for the simulation result at $z \sim 8$ for the HST, and it is estimated with Monte Carlo approach. The error is underestimated due to the neglected cosmic variance (see the Appendix). The projection of the JWST measurements as calculated assuming 30-hr integration per each of these filters: F115W, F150W, and F200W.

The lensing effect has been studied in Wyithe et al. (2011), and its effect on the luminosity function in Mason et al. (2015), Fialkov & Loeb (2015), Barone-Nugent et al. (2015). The general conclusion is that it is significant and should be taken into account for the luminosity function derivation. However, the multiple imaging, which is of interest in this paper, is less definite since it partially relies on the faint-end slope of the luminosity function which is not well measured.

In Wyithe et al. (2011), the fraction of observed multiply imaged systems at $z \gtrsim 7$ is estimated to be the order of $\lesssim 1\%$; therefore, we can roughly estimate the probability of three objects to be lensed as $20 \times 1\%^3 = 0.002\%$, i.e., same order of magnitude as the projection coincidences.

### Table 2

| Physical Effect | 1 Pair | 2 Pairs | 3 Pairs |
|-----------------|--------|---------|---------|
| Projection coincidence (4.1) | 5% | 0.1% | 0.001% |
| Strong lensing (Section 4.2) | <20% | <0.2% | <0.002% |
| Mergers (Section 4.3) | 15% | 4% | 0.2% |
| Catalog contamination (Section 4.4) | Unknown |

**Note.** Groups of three objects are neglected, but for each mechanism it is a much lower probability.

them can, and some of the observed pairs are likely to be real mergers.

#### 4.1. Projection Coincidences

To estimate the projection coincidences, we run a Monte Carlo simulation by randomly populating a $2'1 \times 1'9$ field with 22 points and calculating the number of points with a neighbor within 1". The results for zero, two, three, four, five, and six objects with pairs are 95%, 4.8%, 0.012%, 0.1%, 0.0004%, and 0.001%, respectively. Therefore, all three pairs are highly unlikely to be coincidences. Projection coincidences can be reduced with better photometry and completely eliminated with deeper spectroscopic surveys.

#### 4.2. Lensing

In Section 2.1, we showed that a “hidden” lens can produce the pair of images of a $z \sim 8$ galaxy. Here, we discuss the probability of such an event.

#### 4.3. Merger Statistics

In Figure 7 the fractions of objects with a companion within given radius for $z \sim 7$ and $z \sim 8$ are presented. The $z \sim 7$ sample is consistent with the simulations. However, three groups of galaxies in $z \sim 8$ sample cause a noticeable divergence.

The error region for the simulated $z \sim 8$ sample in Figure 7 is calculated using the Monte Carlo approach by generating many mock observations. The discrepancy barely reaches the significance of $3\sigma$. Since some of these pairs can be projection coincidences, we cannot report this as a significant discrepancy. In order to definitively say whether it is an anomaly, one has to improve the statistics. We believe that it is even possible to do with existing data; our preliminary results show that similar groups of galaxies can be found in other deep fields. However, in this study we limited our exploration to the XDF data.
because the inconsistency in the number and the depth of the filters of other deep fields complicates the analysis.

In Park et al. (2017), similar comparison between numerical simulation and observations are made. In contrast to this paper where we use the merger fraction, the authors adopt the angular correlation function. Also, in that study other deep fields are included into consideration in addition to the XDF data, and redshifts are considered starting from 6 and above. The authors find a good agreement with simulation.

In Figure 7, the merger fraction for the JWST is also plotted. As expected, the fraction of observed mergers is higher compared to the HST. The main caveat is likely to be the confusion limit. The total number of sources will greatly increase (almost by factor of 10), including foreground dim galaxies, while the point-spread function of the JWST is not dramatically superior to that of the HST. Therefore, the field will be crowded with galaxies and our criteria for mergers—within certain angular distance and in the same photometric redshift bin—may not be sufficient, and the probabilistic methods would be required.

4.4. Catalog Contamination

Finally, there is a possibility that the observed groups in $z \sim 8$ sample are caused by the contaminations from lower redshifts. The absence of a sufficiently large spectroscopic sample at the redshifts of our interest makes it impossible to perform a similar quantitative analysis. Nevertheless, it is often assumed that the contamination is low, following the arguments presented in Bouwens et al. (2011) (see Section 4.2 therein).

The catalog by Mclure et al. (2011) has a much lower number of high-redshift galaxies (see the discussion and detailed comparison with Bouwens et al. (2011) therein). There are eight galaxies in the region we consider with $z > 7,$ and all of them are present in $z \sim 7$ or eight samples of B15. Only 6 out of 22 objects of $z \sim 8$ have a counterpart in Mclure et al. (2011) catalog; however, there is one of these six galaxies in each of three groups we studied (see Figure 8). Therefore, the tension between the number of groups in B15 catalog and our theoretical predictions might be interpreted as an argument in favor of a more conservative catalog like Mclure et al. (2011).

5. Discussion

Under the assumption that some fraction of the observed groups are real mergers, we discuss how they could be further studied and used. Most of the applications listed below are already applied to the lower redshift data.

As the distribution of mergers is more biased than isolated galaxies (see the Appendix and Furlanetto & Kamionkowski (2005) for an analytical model), the merger fraction can be used as the indicator of the local overdensity and, consequently, it may allow to control the cosmic variance effect on the luminosity function. For instance, this study shows that the XDF field is likely to correspond to an overdense region at $z \sim 8$.

As was mentioned in the introduction, there are a couple of LAE mergers detected at $z > 6$. Because Ly$\alpha$ is one of a few spectroscopic emission lines that can be observed at high redshifts, it is often the only way to measure spectroscopic redshift for reionization galaxies. Also, the studies of Ly$\alpha$ emitters and their morphology at lower redshifts ($z \approx 4$) confirm that the models that suggest that merging and close encounters can activate Ly$\alpha$ emission (Kobayashi et al. 2016). Therefore, mergers can be potential candidates for Ly$\alpha$ follow up observations.

Besides the connection between mergers and LAEs, the statistics of mergers at high redshifts may shed light on other questions; among them is the accordance between dark-matter halos and galaxies. While the luminosity function of galaxies allows us to perform abundance matching with dark-matter halos, the merging statistics contains additional information that allows to perform more complicated analysis and study assembly bias and subhalo abundance matching. Recently, a comprehensive study of Sloan galaxies at lower redshifts, where statistics are orders of magnitudes better, has been done in Guo et al. (2016).

It is known that at lower redshifts, the major mergers enhance star formation, while at intermediate redshifts ($z \sim 1-3$) the effect seems to be less prominent Fensch et al. (2016). Depending on how strong the effect of the enhanced star formation at the epoch of reionization, the stacking methods like Geil et al. (2017) might be able to detect larger ionization bubbles with 21 cm signal around such systems.

The galaxy mergers may also improve photometric redshifts. The galaxies identified within the same merging group allow one to consider them to be at the same redshift, and therefore improve their photometric redshift by increasing precision (combined signal has larger signal-to-noise ratio) and accuracy (avoiding catastrophic errors using prior of the abundance of mergers at different redshifts). The spectroscopic redshifts matched with some members of the groups can further improve the measurements (Ho et al. 2008). A significant amount of work has been already done to incorporate the spatial correlation of galaxies into the determination of photometric redshifts at lower redshifts (e.g., Newman 2008; McQuinn &
White 2013; Ménard et al. 2013), and these techniques can be applied to high-redshift galaxies. Moreover, the uncertain but identical redshifts of galaxies within a merging group will allow one to directly compare their photometric SEDs. Such a comparison will provide information regarding the UV slopes and the dust abundance as function of galaxy mass.

A.A.K. is supported by the Friends of the Institute for Advanced Study. This work is based on observations taken by the 3D HST Treasury Program (GO 12177 and 12328) with the NASA/ESA HST, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

Software: SExtractor (Bertin & Arnouts 1996), photutils (Bradley et al. 2016), yt (Turk et al. 2011), EAZY (Brammer et al. 2008).

Appendix

Rescaling the Simulation

In the main body of this paper, we assumed that the observed volume is a typical region of the universe. However, it might not be the case, for instance, if by chance the XDF region happen to correspond to a few sigma overdensity at redshift \( z \sim 8 \). In this Appendix we explore how this possibility alters our chances to observe the pairs of galaxies.

To mimic overdense or underdense regions, we rescale the simulation by linearly shifting the time. Because we are interested in relatively short time intervals that corresponds to the \( z \sim 8 \) drop-out sample, rescaling in redshift or scale parameter space gives the same result. After rescaling the simulation does not longer match the observed luminosity functions on lower redshifts; however, it is not relevant for this particular numerical experiment.

For each time shift of the simulation, we use Monte Carlo method to generate thousands of mock observations. First, we randomly stack our simulation box to fill the redshift range corresponding to \( z \sim 8 \). Second, we generate pencil beams that match the XDF volume and \( z \sim 8 \) selection function (Bouwens et al. 2011). As a result, we have multiple realizations some of which contain more or less than 22 objects and some number of pairs. We select only those realizations that have exactly 22 objects and record the probability of this event in terms of \( \sigma \)'s of normal distribution. For convenience, we use this \( \sigma \) to parametrize the rescaling of the simulation. With this notation, the original simulation without rescaling corresponds to roughly zero \( \sigma \), because it was tuned to match the observation. In case we delay it, the typical number of objects drops and an observation with 22 galaxies corresponds to a few \( \sigma \).

Also, we keep record of how many pairs (in this case defined by 1’ separation criteria) were present in each realization with 22 objects. In Figure 9, we show the fraction of realizations that have certain number of pairs versus rescaling parameter \( \sigma \). It can be seen, that when we delay our simulation and, therefore, consider an overdense region, we are more likely to see multiple pairs, and observing three pairs is not longer an unlikely event. The numbers presented in Table 2 correspond to the \( \sigma \) equals zero in the Figure 9.

The trends in Figure 9 are the result of the fact that mergers are more biased than “field” galaxies. See Furlanetto & Kamionkowski (2005) for an analytical model that describes this effect.

![Figure 9](https://zenodo.org/record/1039309)
