

The Stellar Mass and Age Distributions of Star-Forming Clumps at $0.5 < z < 5$ in JWST CANUCS: Implications for Clump Formation and Destruction

VISAL SOK ¹, ADAM MUZZIN ¹, VIVIAN YUN YAN TAN ¹, YOSHIHISA ASADA ^{2,3}, MARUŠA BRADAČ ^{4,5},
VINCE ESTRADA-CARPENTER ^{6,7,8}, KARTHEIK G. IYER ^{9,10}, NICHOLAS S. MARTIS ⁴, GAËL NOIROT,¹¹
GHASSAN T. E. SARROUH ¹, MARCIN SAWICKI ⁸, CHRIS J. WILLOTT ¹², SUNNA WITHERS ¹,
SAMANTHA C. BEREK ^{13,2,14} AND KATHERINE MYERS ¹

¹Department of Physics and Astronomy, York University, 4700 Keele Street, Toronto, ON, M3J 1P3, Canada

²Dunlap Institute for Astronomy and Astrophysics, 50 St. George Street, Toronto, Ontario, M5S 3H4, Canada

³Waseda Research Institute for Science and Engineering, Faculty of Science and Engineering, Waseda University, 3-4-1 Okubo, Shinjuku, Tokyo 169-8555, Japan

⁴Faculty of Mathematics and Physics, Jadranska ulica 19, SI-1000 Ljubljana, Slovenia

⁵Department of Physics and Astronomy, University of California Davis, 1 Shields Avenue, Davis, CA 95616, USA

⁶School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287, USA

⁷Beus Center for Cosmic Foundations, Arizona State University, Tempe, AZ 85287, USA

⁸Department of Astronomy and Physics and Institute for Computational Astrophysics, Saint Mary's University, 923 Robie Street, Halifax, Nova Scotia B3H 3C3, Canada

⁹Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA

¹⁰Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

¹¹Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, Maryland 21218, USA

¹²National Research Council of Canada, Herzberg Astronomy & Astrophysics Research Centre, 5071 West Saanich Road, Victoria, BC, V9E 2E7, Canada

¹³David A. Dunlap Department of Astronomy and Astrophysics, University of Toronto, 50 St. George Street, Toronto, Ontario, M5S 3H4, Canada

¹⁴Data Sciences Institute, University of Toronto, 17th Floor, Ontario Power Building, 700 University Ave, Toronto, ON M5G 1Z5, Canada

ABSTRACT

We investigate the resolved properties of star-forming clumps and their host galaxies at $0.5 < z < 5$ in the JWST CANUCS fields. We find that the fraction of clumpy galaxies peaks near $z \sim 2$ for galaxies with masses of $\log(M_{g,*}/M_{\odot}) \geq 10$. Galaxies with masses of $8.5 \leq \log(M_{g,*}/M_{\odot}) < 10$ show lower clumpy fractions with little redshift evolution. We identify and measure individual clump masses, finding that the aggregated clump stellar mass function (cSMF) follows a power-law slope of $\alpha = -2$ across all redshift bins, broadly consistent with *in-situ* clump formation. However, when split by galaxy mass, the cSMF is found to be flatter ($\alpha \sim -1.6$) for massive galaxies and steeper ($\alpha \sim -2.3$) for lower mass galaxies, with little redshift evolution. We further explore how clump formation and disruption mechanisms shape the cSMF. In particular, we find that the cSMF slope is flatter with increasing inferred gas fraction in younger clump populations (< 300 Myr), suggesting that gas availability favors the formation of more massive clumps. Alternatively, many high-redshift galaxies in our sample have disturbed morphologies and simulations show that clumps of *ex-situ* origins are comparably more massive to *in-situ* clumps, which can flatten the cSMF. We also examine clump evolution, finding the cSMF slope become flatter as clumps evolve and age. We interpret this as an indication of the long-term survivability of massive clumps, where feedback mechanisms preferentially disrupt low-mass clumps. Overall, the galaxy-mass dependent cSMF and age distribution point to a complex history for clumps, involving different and competing mechanisms for their formation and destruction.

Keywords: High-redshift galaxies (734), Star formation (1569), Star forming regions (1565)

1. INTRODUCTION

Star-forming galaxies at $z > 1$ commonly exhibit irregular morphologies and often host multiple kiloparsec-scale, rest-frame UV bright star-forming clumps (e.g., D. M. Elmegreen et al. 2005; M. Puech 2010; N. M. Förster Schreiber et al. 2011a,b; E. Wisnioski et al. 2011; S. Wuyts et al. 2012, 2013; V. Y. Y. Tan et al. 2024). Particularly, the fraction of clumpy galaxies is found to increase with redshift (e.g., K. L. Murata et al. 2014; Y. Guo et al. 2015; M. Huertas-Company et al. 2020; V. Sok et al. 2022; W. Mercier et al. 2025), with a peak around $z \sim 2$, coincident with the peak of the cosmic star formation density (T. Shibuya et al. 2016; Z. Sattari et al. 2023; A. de la Vega et al. 2025). With the recent launch of JWST, there is also a growing number of spectacularly gravitationally lensed sources observed after cosmic noon ($z \gtrsim 5$) that host compact stellar structures (e.g., A. Adamo et al. 2024; L. Mowla et al. 2024; S. Fujimoto et al. 2025; E. Vanzella et al. 2025). While many of these objects are likely to be globular progenitors, it naturally raises important questions about how clumpy star formation occurs, what formation pathways dominate, and how these different mechanisms may evolve across cosmic time.

The commonly suggested mechanisms for clump formation at $z \sim 2$ include both *in-situ* and *ex-situ* processes. Clumps of *in-situ* origins are typically linked to violent disk instability (VDI) as the result of sustained cosmological gas accretion, with gravitational fragmentation governed by the Toomre Q instability parameter (e.g., A. Toomre 1964; O. Agertz et al. 2009; A. Dekel et al. 2009a,b; D. Ceverino et al. 2010; M. Cacciato et al. 2012). This is motivated by observations of very gas-rich galaxies at higher redshifts (e.g., E. Daddi et al. 2010; L. J. Tacconi et al. 2010; T.-M. Wang et al. 2022), with well-ordered rotation (e.g., N. M. Förster Schreiber et al. 2018; E. Wisnioski et al. 2019). The fragmentation of clumps in gas-rich disks is also well-studied in simulations of galaxies in both isolation and cosmological context (e.g., D. Ceverino et al. 2016; J. Fensch & F. Bournaud 2021; F. Renaud et al. 2024). Alternatively, clumps of *ex-situ* origins are attributed to gas compression and fragmentation triggered by galaxy interactions (e.g., M. Puech 2010; A. Calabrò et al. 2019; V. Estrada-Carpenter et al. 2024). These clumps could be remnants of accreted satellites as shown in simulations (N. Mandelker et al. 2017) and suggested in observations (B. Ribeiro et al. 2017; A. Zanella et al. 2019). Further supports for *ex-situ* clumps are also found in

V. Estrada-Carpenter et al. (2025), where some clumps are found to have lower metallicity offsets compared to their host gas disks. These metallicity offsets are not expected if clumps are formed through disk instabilities, suggesting that these are perhaps satellites or even accreting clumps of near-pristine gas. Such metallicity offsets are also reported in V. Sok et al. (2025), where clumpy galaxies are found to have lower metallicities compared to nonclumpy galaxies.

Clump formation due to VDI has been previously argued to be more efficient near cosmic noon due to both the declining of cold gas accretion toward lower redshifts (e.g., E. Daddi et al. 2022; S. Waterval et al. 2025) and the generally larger disks toward lower redshifts. Recent studies of galaxy morphologies at higher- z , however, suggest that VDI-driven clump formation may still be relevant at earlier epochs, given the relatively high fraction of disk-like morphologies observed out to $z \sim 7$ (e.g., L. Ferreira et al. 2023; J. S. Kartaltepe et al. 2023). At $z \gtrsim 5$, both VDI and mergers are linked to lensed galaxies that exhibit clumpy morphologies as observed by JWST/NIRSpec IFU observations (e.g., S. Fujimoto et al. 2025; T. Morishita et al. 2025). Simulations also offer contrasting pictures. Y. Nakazato et al. (2024) found that clump formation due to VDI did not occur until $z \sim 6$ in their FirstLight simulation, with their clumps found at $5.5 < z < 9$ are associated with mergers. In contrast, hydrodynamical simulations in L. Mayer et al. (2025) showed clump formation due to fragmentation at $z > 7$ is possible.

Larger, statistical samples of clumps in galaxies are needed to better constrain clump formation mechanisms. Particularly, the clump stellar mass distribution also provides insights into these processes. Detecting clumps and measuring their stellar masses accurately depend both on the spatial resolution and sensitivity of the observations. High- z star-forming clumps are generally observed to be larger and more massive, often estimated to be 10^6 - 10^9 solar masses (e.g., B. S. Kalita et al. 2025a,b), compared to their star-forming region counterparts in the local universe. While observations of gravitationally lensed galaxies have shown that clump masses are overestimated in those measured in unlensed galaxies (e.g., R. C. Livermore et al. 2015; A. Cava et al. 2018; U. Meštrić et al. 2022; M. Messa et al. 2024; A. Claeysens et al. 2025), the effect of degraded resolution is found to be negligible compared to the impact of observational sensitivity, which can overestimate masses by an order of magnitude (M. Dessauges-Zavadsky et al. 2017; V.

Tamburello et al. 2017). M. Dessauges-Zavadsky & A. Adamo (2018) also showed that the slope of the clump stellar mass function is unaffected when limiting their sample to just blended clumps. Any inferred physics using the measured clump properties must therefore account for these sensitivity limits to ensure robust conclusions.

Initial analyses on the redshift evolution of the clump stellar mass distribution at $z < 5$ include the work of M. Dessauges-Zavadsky & A. Adamo (2018) and A. Claeysens et al. (2025), where both studies found that clump formation is broadly consistent with being *in-situ*. In this paper, we expand on these analyses by utilizing the deep, multi-wavelength imaging of the Canadian NIRISS Unbiased Cluster Survey (CANUCS) to constrain the clump properties of a larger sample of ~ 5000 star-forming galaxies (SFGs) at $0.5 < z < 5$. In particular, we analyze the redshift evolution of the clump stellar mass function in different galaxy mass bins, with clump completeness considerations to mitigate sensitivity biases. We further investigate the relationship between the mass and age of clumps and their host galaxy properties within the context of the mechanisms for their formation and destruction. In order to obtain robust estimates of the stellar mass, age and star formation rate of clumps, we use images that are PSF-matched to the F444W filter. This gives an angular resolution of $\sim 0.16''$, corresponding to an effective spatial resolution of $\sim 1 - 1.3$ kpc for over our redshift range. In §2, we briefly describe the CANUCS dataset, while §3 covers our galaxy selection criteria for the analyses. In §4, we describe our methodology, which includes spatially fitting the SED of galaxies, extracting clumps, and estimating our clump detection limits. We discuss the redshift evolution of the fraction of clumpy galaxies in §5, as well as the clump properties and the clump stellar mass distribution in §6.

Throughout the paper, all magnitudes are expressed in the AB system (J. B. Oke & J. E. Gunn 1983). We adopt a Λ CDM cosmological model of the universe with $\Omega_\lambda = 0.7$, $\Omega_M = 0.3$, and a Hubble constant of $H_0 = 70$ km/s/Mpc.

2. DATA

This paper makes use of deep and extensive photometric datasets from the Canadian NIRISS Unbiased Cluster Survey (CANUCS, C. J. Willott et al. 2022). A detailed overview of the first data release is presented by G. T. E. Sarrouh et al. (2025). Briefly, the full datasets for CANUCS include HST, JWST/NIRCam and JWST/NIRISS photometry, along with spectroscopy from JWST/NIRISS and JWST/NIRSpec of

five strong lensing clusters; Abell 370 ($z = 0.375$, hereafter A370, J. M. Lotz et al. 2017), MACS J0416.1-2403 ($z = 0.395$, hereafter MACS0416, J. M. Lotz et al. 2017), MACS J0417.5-1154 ($z = 0.443$, hereafter MACS0417, M. Postman et al. 2012), MACS J1149.5+2223 ($z = 0.543$, hereafter MACS1149, J. M. Lotz et al. 2017), and MACS J1423.8+2404 ($z = 0.545$, hereafter MACS1423, B. Salmon et al. 2020). The primary cluster fields consist of imaging from NIRISS, including filters F090WN, F115WN, F150WN, and F200WN (where the suffix “N” denoting the NIRISS filter designation), with NIRCam filters F090W, F115W, F150W, F200W, F277W, F356W, F410M, and F444W. Also included are supporting HST imaging, with filters F435W, F606W, F814W, F105W, F125W, F140W, and F160W. The typical 3σ magnitude limits are $rm_{AB} \approx 29.5$ for NIRISS, $m_{AB} \approx 29.5-30.0$ for NIRCam, and $m_{AB} \approx 28.0-30.0$ for the HST/ACS and HST/WFC3 filters.

In addition to the primary cluster fields, there are a total of ten flanking fields, comprising of a NIRCam flanking (NCF) field and a NIRISS flanking (NSF) field that are observed by the NIRCam and NIRISS module on either side of each cluster field, respectively. In this work, we utilize the five NCF fields. Three of these five NCF fields (A370, MACS0416, MACS1149) include almost the full set of NIRCam imaging, while the other two NCF fields contain nine medium bands and four wide-band filters. In brief, the available filters for the NCF fields include wide-band filters F070W, F090W, F115W, F150W, F200W, F277W, F356W, and F444W, and medium-band filters F140M, F162M, F182M, F210M, F250M, F300M, F335M, F360M, F410M, F430M, F460M, and F480M. Similarly, the 3σ magnitude limits for the medium- and wide-band filters in the NCF fields are $m_{AB} \approx 29.5-30.5$ for NIRCam, and $m_{AB} \approx 27.5-30.0$ for the HST/ACS and HST/WFC3 filters. We refer the reader to G. T. E. Sarrouh et al. (2025) for the detailed list of available filters and limiting magnitude for each cluster and flanking field.

Details of the PSF-matching, object detection, photometry and catalog creation are described in G. T. E. Sarrouh et al. (2025). Briefly, source detection is done based on a combined chi-mean detection image, with aperture photometry done using the PHOTUTILS package (L. Bradley et al. 2024). Each cluster field is also affected by varying level of gravitational lensing, with one NIRCam module centered on the cluster and the other generally sampling a lower-magnification region. The lensing models for A370 are provided in R. Gledhill et al. (2024), for MACS0416 in G. Rihtaršič et al. (2025), MACS1149 in G. Rihtaršič et al. (in prep), and

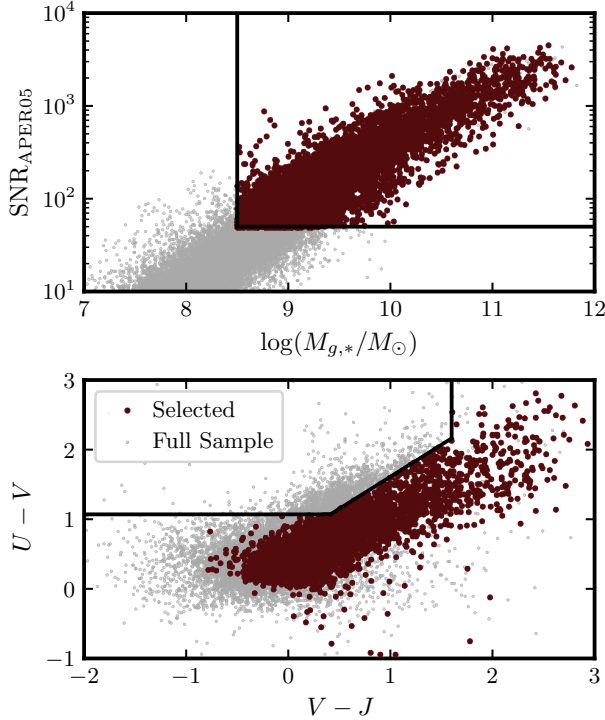


Figure 1. We selected galaxies based on their integrated SNR within a $0.5''$ aperture and stellar mass. Additionally, we used the UVJ color selection to select star-forming galaxies (see text for details). Only $\sim 20\%$ of all sources with `USE_PHOT = True` and `FLAG_POINTSRC = False` satisfies our selection criteria. It should also be noted that not all sources falling within each selection region necessarily satisfy all the selection criteria in this figure.

MACS0417 and MACS1423 in G. Desprez et al. (in prep). The lensing models are available on the CANUCS website¹⁵.

3. SAMPLE SELECTION

We make use of imaging and photometry products from Data Release 1 (G. T. E. Sarrouh et al. 2025) to select the galaxy sample for this work. Objects with unreliable photometry, `USE_PHOT = False`, are excluded. These include sources that are bright cluster galaxies, sources contaminated by the bright cluster galaxy subtraction, and those with Kron aperture overlapping masked and bad pixels. In addition, point-like objects are also removed to avoid stellar contamination, using the flag `FLAG_POINTSRC`.

Figure 1 shows the main selection criteria of the sample. The selection criteria include cuts into stellar masses, redshift, integrate signal-to-noise ratio (SNR_{int} , measured within an aperture diameter of $0.5''$ based on

the F444W image), and rest-frame colors. The mass completeness of the CANUCS sample is determined to be 95% complete down to stellar masses of $10^{8.2} M_{\odot}$ at $z \sim 5$ (Sarrouh et al. in prep.). We adopt a conservative cut at $10^{8.5} M_{\odot}$, excluding lower-mass galaxies to enable a fair comparison of the clump stellar mass function across redshift. We exclude galaxies that have a $\text{SNR}_{\text{APER05}}$ (SNR within an aperture of $0.5''$) less than 50 to ensure that each galaxy can be spatially binned using Voronoi tessellation, as described in §4.1.1, for resolved SED fitting. A lower redshift cut of $z = 0.5$ ($z = 0.65$ for MACS1149 and MACS1423 as these clusters are at a higher redshift) is also applied to exclude galaxies associated with the clusters, as our primary focus is on clump formation irrespective of environment. Due to the constraints on stellar masses and $\text{SNR}_{\text{APER05}}$, galaxies at $z > 5$ are also excluded as the number of galaxies that satisfy these criteria becomes limited.

Since clumps are regions of intense star formation, we select star-forming galaxies based on their rest-frame UVJ colors. The UVJ selection for SFGs at all redshift bins follows the prescription from J. Antwi-Danso et al. (2023), and is defined by the following color cuts,

$$\begin{aligned} (U - V) &< 1.07 \\ (V - J) &> 1.6 \\ (U - V) &< (V - J) \times 0.92 + 0.69 \end{aligned} \quad (1)$$

These lines are calibrated for the CANUCS parent galaxy sample between $0.5 < z < 6$, where the slope and zero-point of the diagonal line is iteratively updated until a clear bimodality is determined between the red and blue sequence of galaxies (see V. Y. Y. Tan et al. 2024). Using a broad redshift interval here ensures that the calibration accounts for any redshift evolution in the zero-point. While the UVJ color selection is known to lead to contamination of SFGs in the quiescent sample (e.g., C. Schreiber et al. 2018), it is still an efficient method to select different galaxy populations. We note other color selections can yield a more complete selection (e.g., G. Noirot et al. 2022; J. Antwi-Danso et al. 2023). However, since our analyses combine clump populations across many host galaxies, the results are not sensitive to the specific choice of color-color selection.

We also exclude galaxies with magnification factors of $\mu > 2$ as the majority of the galaxies in our sample is unlensed. This threshold removes ~ 600 galaxies that are potentially subjected to significant differential magnification, but ensures that the analyses are performed on images with similar spatial resolution. We apply a correction to the derived properties for magnification by accounting for a factor of μ where applicable (see §4.2.1).

¹⁵ <https://niriss.github.io/data.html>

Lastly, we exclude highly edge-on galaxies with axial ratio $q < 0.2$, as projection effects and increased dust attenuation in these systems can obscure clumps, leading to biases in clump detection and unreliable estimates of their physical properties. The axial ratio is determined by fitting a Sérsic profile to each galaxy in the F444W filter with GALFIT (C. Y. Peng et al. 2002). More details for the morphological parameter fitting is presented in Merchant et al. in prep. Finally, the number of galaxies in our sample given these criteria is approximately 4500.

4. METHODOLOGY

This section describes the methodology for constructing resolved stellar mass density and rest-frame UV ($\lambda_{\text{rest}} \approx 3600 \text{ \AA}$) surface brightness maps. Pixels are spatially binned to obtain the spectral energy distribution (SED). These SEDs are fitted to produce resolved maps of the stellar mass density and U_{rest} surface brightness maps. We adopt the recalibrated U filter from J. Maíz Apellániz (2006), which is based on the Johnson system, to retrieve U_{rest} . We also discuss how clumps are identified within these U_{rest} brightness maps. Galaxies are first classified as clumpy and nonclumpy based on their U_{rest} brightness normalized profile, and individual clumps are identified in those galaxies which are classified as clumpy. Lastly, this section outlines how clump masses are retrieved from the two-dimensional stellar mass density maps and discuss our clump detection limits. Again, the following analyses are done using images that are PSF-matched to the F444W image, with pixel resolution of 40 mas.

4.1. Spatially-resolved SED fitting

4.1.1. Voronoi Binning

Spatially-resolved SED fitting provides better estimates of stellar masses compared to unresolved fitting (e.g., R. Sorba & M. Sawicki 2015, 2018). However, rather than fitting the SED on a pixel-by-pixel basis, as individual pixels generally have low SNR, we first bin pixel together to acquire a minimum SNR.

In this paper, the binning step is done based on Voronoi tessellation using VORBIN (M. Cappellari & Y. Copin 2003), where pixels are binned until a minimum S/N threshold is reached in the F444W band. We adopt F444W for the binning as it provides the most reliable tracer of the underlying stellar mass. The minimum SNR per Voronoi bin is defined by the integrated SNR of the galaxies in the F444W image (i.e., $\text{SNR}_{\text{APER05}}$ as measured within a aperture diameter of $0.5''$). We set the minimum SNR to be 5 and 10 for galaxies with $\text{SNR}_{\text{APER05}} < 500$ and $\text{SNR}_{\text{APER05}} \geq 500$, respectively.

The choice for varying the minimum SNR on a galaxy to galaxy basis is because our sample consists of galaxies at a wide range of stellar masses and redshifts. A fixed threshold can otherwise yield fine binning in high-SNR galaxies and unnecessarily coarser binning in low-SNR galaxies.

4.1.2. Inferring physical properties

We determine the flux for each bin by summing the pixel values within that bin. Similar to Y. Han & Z. Han (2019), the flux uncertainty for each bin is estimated as the quadrature sum of the statistical uncertainty and the systematic uncertainty. Here, the statistical uncertainty is the aperture flux uncertainty, while the systematic component accounts for potential biases during data reduction and calibration. Since each bin consists of multiple pixels, we assume that the flux uncertainty for a given bin is similar to that within a circular aperture encompassing the same area (see G. T. E. Sarrouh et al. 2025 for estimations of the aperture flux uncertainty). We assume that the systematic uncertainty is just a multiplicative factor of the flux. A multiplicative value of 0.05 is used to account for systematic uncertainties in this analyses, similar to that proposed in Abdurro'uf et al. (2021).

We derive spatially resolved stellar mass and SED-based star formation rate surface densities for each galaxy by fitting the SED of every Voronoi bin with DENSE BASIS (K. G. Iyer et al. 2019). DENSE BASIS employs an amortized brute-force Bayesian approach to determine the star formation history non-parametrically, which has been shown help mitigate biases against outshining from younger stellar populations and lead to more robust measurements of the SFH (e.g., K. Iyer & E. Gawiser 2017; J. Leja et al. 2019; S. Lower et al. 2020). We adopt a Chabrier initial mass function (G. Chabrier 2003) and Calzetti attenuation law (D. Calzetti et al. 2000). In later discussions, we also determine the age of clumps based on background subtracted aperture photometry. The clump age is defined to be the lookback time at which 75% of its total stellar mass formed (denoted as t_{75}). We adopt t_{75} to account for more recent star formation. To determine t_{75} , we sample the star formation history (SFH) posteriors, integrate each SFH to obtain the cumulative mass formed, and calculate the lookback time at which the cumulative mass fraction reaches 0.75. The adopted t_{75} for each clump is taken as the median of the its posterior distribution. Additionally, we also spatially reconstruct the rest-frame UV surface brightness maps with EAZY-PY (G. B. Brammer et al. 2008), using the recalibrated U filter from J. Maíz

Apellániz (2006). To this end, we fix the redshift of each SED to the redshift of its parent galaxy.

Binning pixels together also effectively degrades the spatial resolution of our maps, especially at the outer region of galaxies where the SNR is lowest. This results in Voronoi bins that are generally larger at the outskirts. Large bins can also smear out clumps in the rest-frame UV brightness maps as binning is done based on the F444W image. To recover spatial resolution in the stellar mass density and U_{rest} brightness maps, we “dezonify” the Voronoi bins by redistributing the binned values according to the flux in the JWST images, similar to the dezonification procedure applied in previous works (e.g., R. Cid Fernandes et al. 2013). Specifically, for each pixel i within a given Voronoi bin, the value is rescaled as,

$$Q'_i = Q_{\text{bin}} \frac{F_i}{F_{\text{bin}}}, \quad (2)$$

where Q_{bin} is the SED-derived quantity for that bin (i.e., either mass or U_{rest} luminosity), F_i is the flux at pixel i (i.e., either using F444W or another image), and F_{bin} is the sum across all pixels in the bin. Explicitly, the F444W image is used to scale the stellar mass maps as F444W closely traces masses, while we use whichever image probes the rest-frame UV at a given redshift to scale the rest-frame UV surface brightness maps at that redshift. We remind the reader that all the images used in this paper are convolved to match the F444W PSF. This rescaling therefore preserves the total measured quantity for each bin while reintroducing spatial resolution at an angular resolution of $\sim 0.16''$.

4.2. Identifying clumps in the rest-frame UV maps

We identify clumps in galaxies through two main steps. First, we select clumpy galaxy candidates using the normalized light profile method of S. Wuyts et al. (2012). This method identifies broad regions of enhanced rest-frame UV emission, which include both extended and more localized clumpy star formation regions. Second, we apply a clump finding algorithm to find localized clumps in these clumpy candidates.

The normalized light profile is constructed based on the rest-frame UV surface brightness maps (see S. Wuyts et al. 2012; V. Sok et al. 2022, 2025). We refer the reader to these papers for more details, however we briefly describe it here. Galaxies with smooth disk will have a normalized light profile that is brighter near its galactic center and tapers off with radius. As star-forming clumps are bright UV regions, superimposed on an otherwise smooth disk, clumpy regions will therefore appear as bumps in this one-dimensional normalized light profile. In order to make the normalized light profile for a

galaxy, we first define two structural parameters of the galaxy; the axial ratio and the position angle. These elliptical parameters are inferred from the segmentation map of the galaxy. The apertures are then used to construct a curve of growth, with apertures centered on the mass-weighted center. We then define the half-light radius R_e as the radius that contains half of the total light, and the effective surface brightness as the mean surface brightness within R_e . The normalized light profile is obtained by normalizing the distance of each pixel to the mass-weighted center of the galaxy by R_e , and the rest-frame UV surface brightness of each pixel by the effective surface brightness. Within this normalized parameter space, S. Wuyts et al. (2012) defined a region to identify clumpy pixels. Here, we classify a galaxy as a clumpy candidate if more than 5% of its total rest-frame UV light arises from this clumpy region. The choice of 5% is generally consistent with thresholds used in previous work to classify clumpy galaxies (e.g., Y. Guo et al. 2018; V. Sok et al. 2022; Z. Sattari et al. 2023).

The normalized light profiles only determine broad regions of enhanced rest-frame UV brightness, and will identify extended regions of star formation as clumpy. We identify star-forming clumps in these galaxies by using a clump finding algorithm similar to what is presented in B. S. Kalita et al. (2025a). We first subtract the large-scale structures from the rest-frame UV images to enhance small-scale structures such as clumps. This is achieved by decomposing the rest-frame UV images into six spatial scales using a wavelet-based approach, and reconstructing a smooth model of each imaging by setting scale one and two to zero. Specifically, we adopt the Isotropic Undecimated Wavelet Transform, or more commonly known as starlet transform, as it is well-suited for astronomical data analysis (J.-L. Starck et al. 2007). The starlet transform has been implemented in the python package SCARLET (P. Melchior et al. 2018).

Figure 2 illustrates the steps to construct the high-contrast residual image. It should be noted that the third wavelet scale corresponds to an angular resolution of $\sim 0.2''$. Subtracting wavelets of this scale and higher effectively removes any diffuse background emission and only leaves structures smaller than this scale. The resolution at scale 3 is determined by injecting a Gaussian profile with FWHM = $0.16''$ into the image and analyzing its FWHM across wavelet scales. We apply a peak finding algorithm to find clump candidates using a threshold of 5σ and minimum separation of three pixels. To mitigate spurious detection, we further require that each detected clump has a flux above 3σ within a $0.2''$ -diameter aperture in the rest-frame UV.

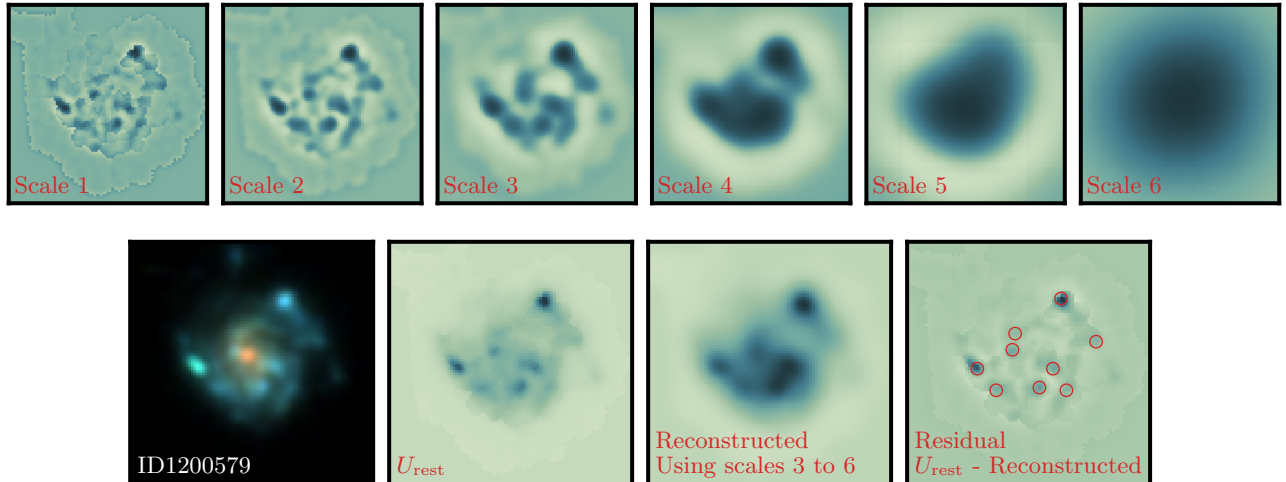


Figure 2. An illustration showing how clumps are identified with starlet wavelet transform. The rest-frame U image is first decomposed into multiple spatial scales. The top row shows the resulting wavelet scales. A smooth model of the galaxy is then reconstructed by combining only the large-scale components (scales 3 and higher), while setting scales 1 and 2 to zero. The bottom row shows the composite color image, the rest-frame U image, the reconstructed smooth image, and finally the high-contrast residual image obtained by subtracting the smooth model from the original. Clumps are identified using a peak finding algorithm on the high-contrast image (see text) and are marked by red circles with a diameter of $0.2''$. The cutouts shown here are $3.6'' \times 3.6''$.

We note that the clump finding algorithm does not discriminate between star-forming clumps and bulge-like components. To remove contamination of bulge-like structures in the subsequent analyses, we assume that the center of mass determined from the surface mass density map traces the galaxy’s gravitational potential. Any clumps located within 0.24 arcsec of the center of mass are therefore considered as part of the bulge structures and are removed from the analyses (see Figure 3 for examples of these bulge-like components). Finally, we define clumpy galaxies as those with detected clumps that contribute to more than 5% of the total UV luminosity. Only clumps detected in clumpy galaxies are considered for the remainder of the paper.

4.2.1. Galaxy subtraction

We measure clump stellar masses directly from the stellar mass surface density maps using fixed aperture photometry. Since clumps are generally superimposed within a galactic disk, it is essential to account for this underlying background to avoid overestimating their physical properties.

We assess the robustness of different background subtraction techniques in Appendix A. In particular, we test three different methods, including (1) reconstructing the smooth underlying disk by decomposing the mass map into different wavelet scales and removing the contribution of the two lowest scales (similar to §4.2 when creating a smooth image of the galaxies), (2) modeling the disk by fitting a Sérsic profile, and (3) estimat-

ing the local background via an annular aperture. In general, we find that all three methods reliably recover masses of artificially injected clumps. In the end, we adopt the annular subtraction method as it relies on simpler assumptions of the underlying disk, compared to parametric modeling approaches such as a Sérsic profile. Other studies such as Z. Sattari et al. (2023) have also used the annular background subtraction to estimate physical properties of clumps. Overall, performing background subtraction generally leads to correction of ~ 0.2 dex per clump.

As the clumps in this study are typically unresolved and will have a FWHM of $\sim 0.16''$, we use an aperture of radius $0.2''$ to extract clump masses. This aperture size encloses the majority of the clump’s mass (and similarly, light) while minimizing contamination from the underlying disk. Local background estimation is done by calculating the mean stellar mass density between two annuli at 3.5 and 5 pixels around the clumps. We also mask other clumps in the galaxy image while measuring the local stellar mass density to avoid contamination. The local stellar mass density is subtracted from the aperture measurement to isolate the stellar mass associated with each clump. We account for magnification, if any, by dividing the retrieved clump mass by μ .

We note that clump masses can be alternatively retrieved by fitting the SED to photometry measured directly from the multi-band images. To compare this approach to extracting the clump masses from the two-dimensional mass maps, we perform similar aperture

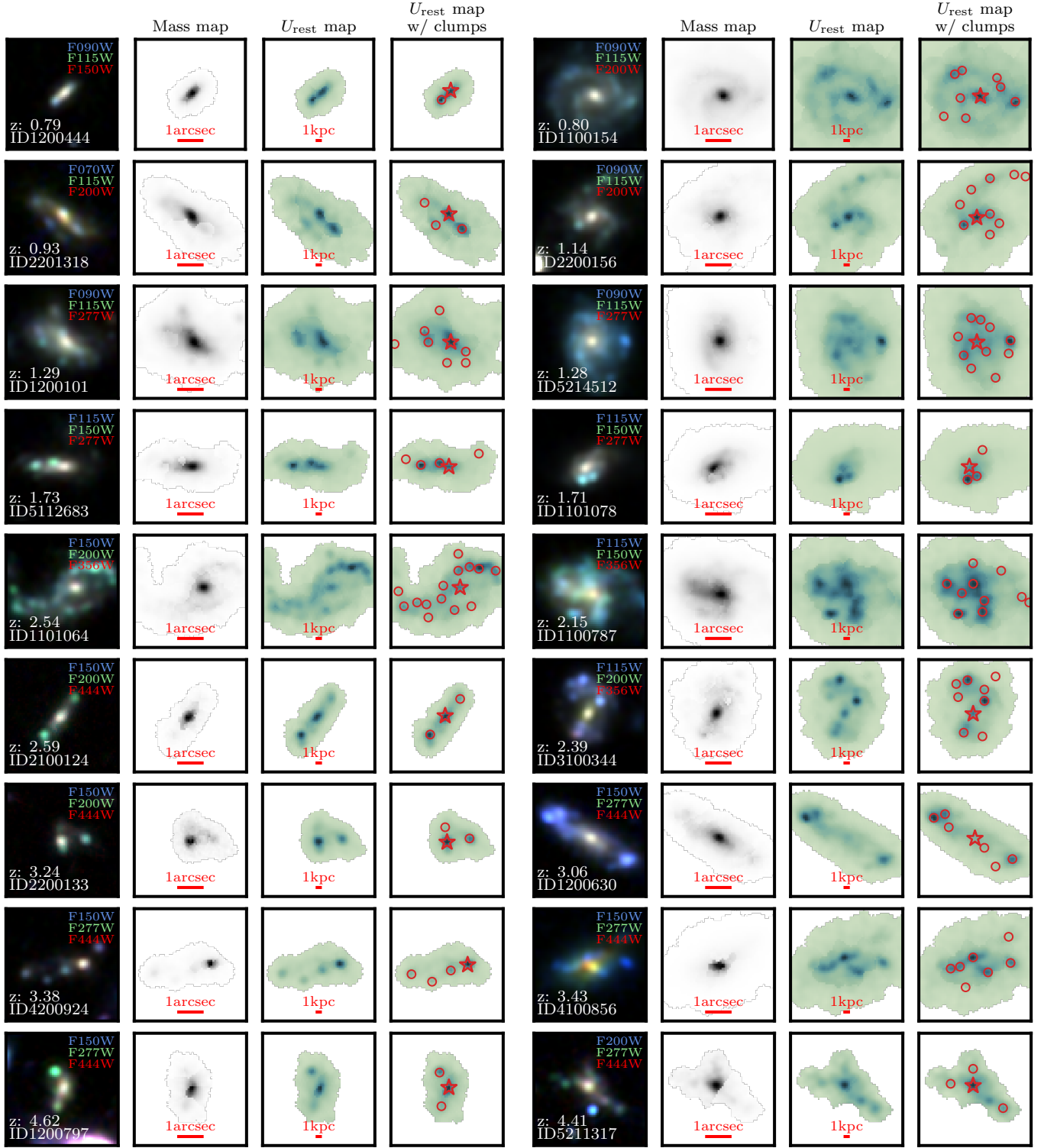


Figure 3. Examples of clumpy galaxies in the CANUCS fields. These galaxies are selected based on their clumpy morphology to illustrate our clump detection algorithm, and are ordered by redshift, with less massive galaxies ($\log(M_{g,*}/M_{\odot}) \lesssim 10$) shown on the left and more massive galaxies ($\log(M_{g,*}/M_{\odot}) \gtrsim 10$) on the right. In each panel, we show the composite RGB image, the stellar mass maps, the rest-frame UV map, and the same UV map with detected clumps (shown as the circle markers) and bulges (shown as the star markers).

photometry on images using the same background sub-

traction procedure. Here, magnification is accounted by

dividing each flux by the corresponding μ . We find that clump masses from the two approaches are consistent, with a scatter of ~ 0.3 dex.

4.2.2. Clump completeness

Understanding our clump detection thresholds is a crucial step towards interpreting our results, including the clump stellar mass function. We determine the clump detection limit by injecting clumps into galaxy images across all available filters and applying our spatially resolved SED-fitting and clump-finding algorithm to quantify our clump recovery.

In Figure 4, we first look at the distribution of all clumps detected as a function of their luminosity compared to the host galaxy’s luminosity (i.e., the fractional luminosity function). We find that we are able to detect fainter clumps compared to the host galaxy’s luminosity at lower redshifts. However, at higher redshifts, the distribution is shifted to higher $L_{\text{clump}}/L_{\text{gal}}$ values as clumps become harder to detect. To determine the completeness of our detection, we simulate clumpy galaxies by injecting clumps into them. Rather than inject clumps from an assumed set of parameters, we instead sample directly from a parent sample of detected clumps and inject them into different galaxies. This approach ensures that the simulated clumps reflect the intrinsic range of observable clump properties. Since clumps are detected based on their rest-frame UV brightness, we explore mass detection limits by randomly scaling clump fluxes so that their final mass span between $7.4 < \log(M_c/M_\odot) < 8.6$.

The parent sample of clumps used for simulating clumpy galaxies consists of clumps detected between $0.5 < z < 4$. We limit this sample to clumps with masses at $7.4 < \log(M_c/M_\odot) < 8.6$ to avoid selecting either clumps that are well below the detection limit, or brighter clumps that could be satellites. To simulate clumpy galaxies, we randomly select galaxies from our galaxy sample. For each simulated galaxy, we inject “donor” clumps from the parent clump sample whose redshifts are within $\Delta z = 1$ of the simulated galaxy’s redshift. The redshift constraint here ensures that these clumps are representative of those observed at similar epoch. Since these donor clumps are generally at a different redshift compared to the target galaxy, we shift their spectral energy distributions accordingly to the new redshift. The SED of clumps is obtained from DENSE BASIS, using a $0.2''$ -diameter circular aperture centered on the clump, with background subtraction applied in each filter (see Appendix A). The flux of the redshifted SED through each JWST filter is measured with DENSE BASIS, which account for the filter response.

These are then injected into the corresponding images of the simulated galaxy. Each clump is modeled as a 2D elliptical Gaussian with semi-minor axis of FWHM of $0.16''$, and is scaled so that the total flux is within an aperture of $0.2''$. We randomize the ellipticity between $0.1 < e < 0.2$ and the position angle. For each simulated galaxy, up to four donor clumps are injected into them. The uncertainties of the clump mass completeness are determined by performing twenty independent realizations of this injection-recovery procedure. In each realization, there are approximately 200 simulated galaxies with ~ 800 injected clumps.

The top panels of Figure 5 show the completeness as a function of the fractional rest-frame UV luminosity in clumps per redshift bin, separated by galaxy masses. In general, we find that the completeness is lower for less-massive galaxies at fixed fractional clumpy luminosity. This arises because lower mass galaxies will have lower total luminosities, so clumps corresponding to a given fraction of that luminosity are intrinsically fainter and therefore more difficult to detect. Similarly, the bottom panels of Figure 5 show the completeness as a function of clump mass. We model the completeness as a sigmoid function, described as,

$$C(M_{c,*}) = \frac{100}{1 + e^{-k(M_{c,*} - M_o)}}, \quad (3)$$

where k and M_o are parameters. At clump masses of $\log(M_{c,*}/M_\odot) \sim 8.0$, the clump mass completeness is at 80%. Given the uncertainties, we find that the clump mass completeness is similar for both low- and high-mass galaxies. However, we note that the mass completeness seems lower for massive galaxies at lower redshifts. This could be driven due to the contrast of the image, where the light from a clump is weaker compared to more massive galaxies.

5. THE FRACTION OF CLUMPY GALAXIES

In this section, we examine the abundance of clumpy galaxies between $0.5 < z < 5$. Clumpy galaxies are defined as those with clumps that contribute to more than 5% of their total UV luminosity. We denote the f_{clumpy} as the fraction of star-forming galaxies that are clumpy. Figure 6 shows the evolution of the fraction of clumpy galaxies at different redshift bins, and the error bars represent the 1σ uncertainties assuming binomial statistics. We find differences in the redshift evolution based on galaxy masses. Specifically, massive galaxies with $\log(M_{g,*}/M_\odot) \geq 10$ exhibit a stronger evolution with redshift relative to less massive galaxies with $8.5 < \log(M_{g,*}/M_\odot) < 10$. The clumpy fraction for massive galaxies peaks at around 60% near $z \sim 2$ and remains

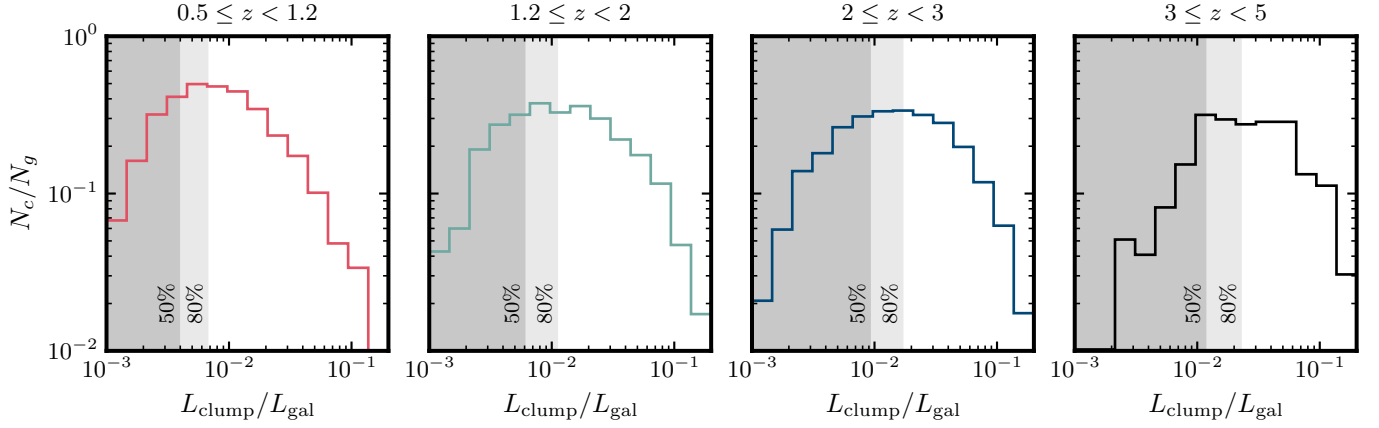


Figure 4. The fractional luminosity function of clumps as a function of redshift. In general, we find that the distribution is shifted toward higher $L_{\text{clump}}/L_{\text{gal}}$ values at higher redshifts as fainter clumps become harder to detect. The gray regions denote the 50% and 80% completeness limit determined by simulating clumpy galaxies and determining how many clumps are retrieved back.

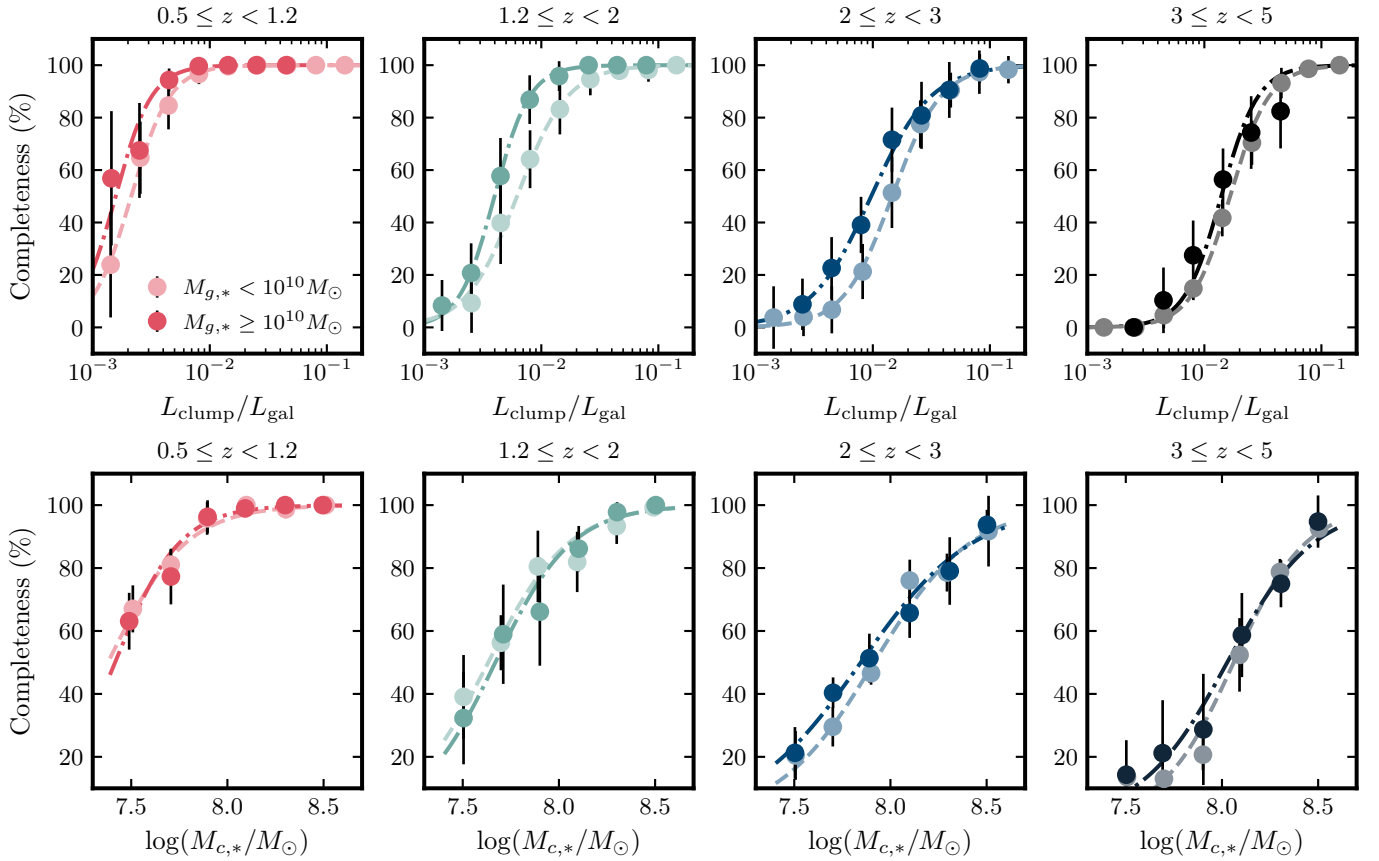


Figure 5. The completeness limit as a function of host galaxy mass and redshift bins. In the top panels, we show the completeness as a function of the fractional luminosity of clumps. The bottom panels show the completeness as a function of clump stellar mass. We determine the completeness by injecting clumps into galaxy images and rerunning our SED fitting and clump detection pipelines (see text for details). In general, we find that the clump completeness is similar for both galaxy mass bins.

at around 40% at $z \sim 4$. For less massive galaxies, the clumpy fraction is observed to be at 30% for all redshift bins.

The trend in the redshift evolution of the clumpy fraction can be driven due to incompleteness, where clumps become fainter and harder to detect at higher redshift.

To estimate a correction to the clumpy fraction, we determine the incompleteness of our classification based on a Monte Carlo approach. For each redshift bin, we integrate both the observed and corrected fractional luminosity function from $\log(L_{\text{clump}}/L_{\text{gal}}) > -2$ (see §4.2.2) to determine the expected number of clumps per galaxy. In each realization, we then sample that number of clumps directly from the fractional luminosity functions to determine if the sampled clumps add up to more than 5%. This provides an estimate for the classification completeness X of our sample based on the corrected fractional luminosity functions. The corrected f_{clumpy} is then calculated following a similar equation in Y. Guo et al. (2015),

$$f_{\text{clumpy}}^{\text{corr}} = f_{\text{clumpy}}^{\text{old}} + \left(\frac{1}{X} - 1\right) f_{\text{clumpy}}^{\text{old}}. \quad (4)$$

The second term on the right hand side accounts for the fraction of galaxy that may have been clumpy but classified as non-clumpy due to non-detections of clumps. In general, we find that the classification completeness is around 80-95% depending on the redshift bins. These values are likely lower if clumps of lower fractional luminosity are included, however we limit the correction to clumps with $\log(L_{\text{clump}}/L_{\text{gal}}) > -2$, where the completeness is approximately at 50% and the correction is more reliable.

We also compare our results to other studies. While the differences in the normalization of the clumpy fraction can be affected by choices in sample selection and clump detection thresholding, we generally find that the overall redshift evolution of the clumpy fraction is consistent with previous studies for galaxies of similar mass bins. This is shown as the gray markers in Figure 6. In particular, studies of lower-redshift clumpy galaxies suggest a steep decline in the clumpy fraction at $z < 1$ (K. L. Murata et al. 2014; R. Hinojosa-Goni et al. 2016; D. Adams et al. 2022). T. Shibuya et al. (2016) further noted that the clumpy fraction peaks near cosmic noon and declines at earlier times, although the true clumpy fraction may drop more slowly at $z > 2$ due to incompleteness. They suggested that the clumpy fraction broadly traces the evolution of the cosmic star formation rate density (SFRD, P. Madau & M. Dickinson 2014). This redshift evolution is also observed by A. de la Vega et al. (2025). Their higher clumpy fraction arises from using a lower threshold at 2% compared to the 5% of this study.

We reproduce the redshift evolution by fitting our uncorrected clumpy fraction of galaxies with $\log(M_{g,*}/M_{\odot}) \geq 10$ using the empirical prescription

from P. Madau & M. Dickinson (2014),

$$f_{\text{clumpy}} = a \frac{(1+z)^b}{1 + [(1+z)/c]^d}, \quad (5)$$

where we fix the parameters b , c and d to the values reported in F. Walter et al. (2020), and only refit for a to renormalize the best-fit curve. This is plotted as the dashed gray line in Figure 6. Similarly, the renormalized molecular gas density is shown as a dotted line. The observed clumpy fraction broadly follows the redshift evolution of both the molecular and SFR densities.

Across all redshifts, we find that massive galaxies are more likely to be clumpy than low mass galaxies. While previous studies (e.g., V. Sok et al. 2022; Z. Sattari et al. 2023) have reported a negative correlation between galaxy mass and clumpy fraction, it should be noted that our sample contains galaxies with stellar masses approximately an order of magnitude lower than those previous works. For example, the clumpy fraction as observed in both K. L. Murata et al. (2014) and M. Huertas-Company et al. (2020) is seemingly peaked at $\log(M_*/M_{\odot}) \sim 10$ before declining at lower masses. In addition to stellar masses, the clumpy fraction is also dependent on additional parameter such as star formation rate (e.g., K. L. Murata et al. 2014; V. Sok et al. 2022).

6. CLUMP PROPERTIES

In this section, we examine the relationship between clump properties and the global properties of their host galaxies, with a particular focus on stellar mass. We further analyze the clump mass distribution and how it evolves with redshift. As noted in §4.2, these analyses are restricted to clumps and do not include bulge-like components detected in galaxies.

6.1. Clumps/Host Galaxies Relations

6.1.1. Number of clumps per galaxy

Figure 7 shows the number of clumps found in galaxies as a function of the host galaxy mass and star formation rate. The marginal probability distributions are also shown the top and the right panels. We find that the number of clumps is generally higher for more massive galaxies at fixed redshift, similar to what was found in E. Soto et al. (2017). When looking at star formation rates, we find that galaxies with higher star formation rate do not seem to preferentially host more clumps compared to those with lower SFRs. This is seen more clearly in the lower-right panel, which shows the marginal probability distribution of SFR, color-coded by the number of clumps per galaxy. In general, the distributions peak at similar SFR values regardless of clump

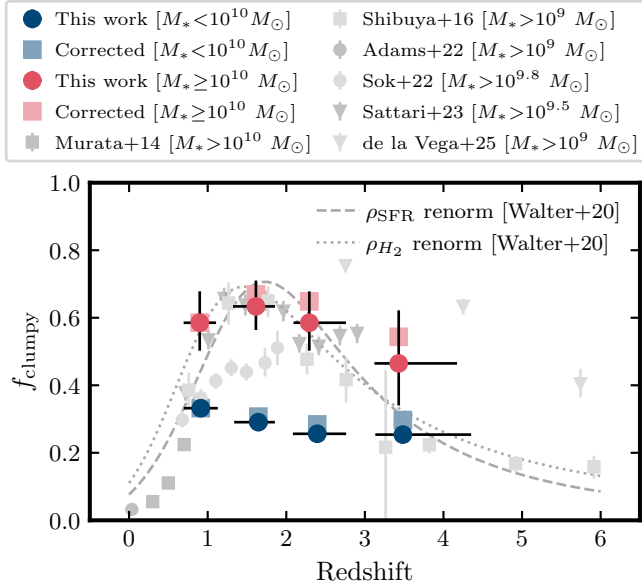


Figure 6. The fraction of clumpy star-forming galaxies for two different galaxy mass bins, where the red markers denote galaxies with $\log(M_{g,*}/M_{\odot}) > 10$ and the blue markers denotes galaxies with $8.5 < \log(M_{g,*}/M_{\odot}) < 10$. We further correct the clumpy fraction based on a Monte Carlo approach where we sample clumps from both the corrected and observed fractional luminosity function to determine classification incompleteness (see text for details). We also compile and show previous clumpy fraction measurements, shown as the lighter gray markers. The dashed and dotted line show the cosmic SFR and molecular gas density from [F. Walter et al. \(2020\)](#), renormalized to fit our measured clumpy fraction of galaxies with $\log(M_{g,*}/M_{\odot}) > 10$.

count. Above $z > 3$, the number of detected clumps decreases. However, this is likely driven by detection limits, where clumps of a fixed fractional luminosity of their host galaxies are less complete.

6.1.2. Clump masses

Figure 8 shows the distribution between clump mass and the host galaxy mass for all clumps in our sample as the contours. We also examine the relationship between the mass of the most massive clump in each galaxy and its host galaxy mass. This is shown as the blue markers, which represent the median mass of these most massive clumps in host galaxy mass bins of 10^8 , 10^9 , 10^{10} , 10^{11} , and $10^{12} M_{\odot}$. Similar trends have been reported in previous studies; for example, [A. Claeysens et al. \(2025\)](#) also found a strong positive correlation between the maximum clump mass and the host galaxy mass (reproduced in Figure 8). We further examine the correlation between the maximum clump mass and host galaxy mass by computing the Pearson correlation coefficient on the binned data points. The Pearson test

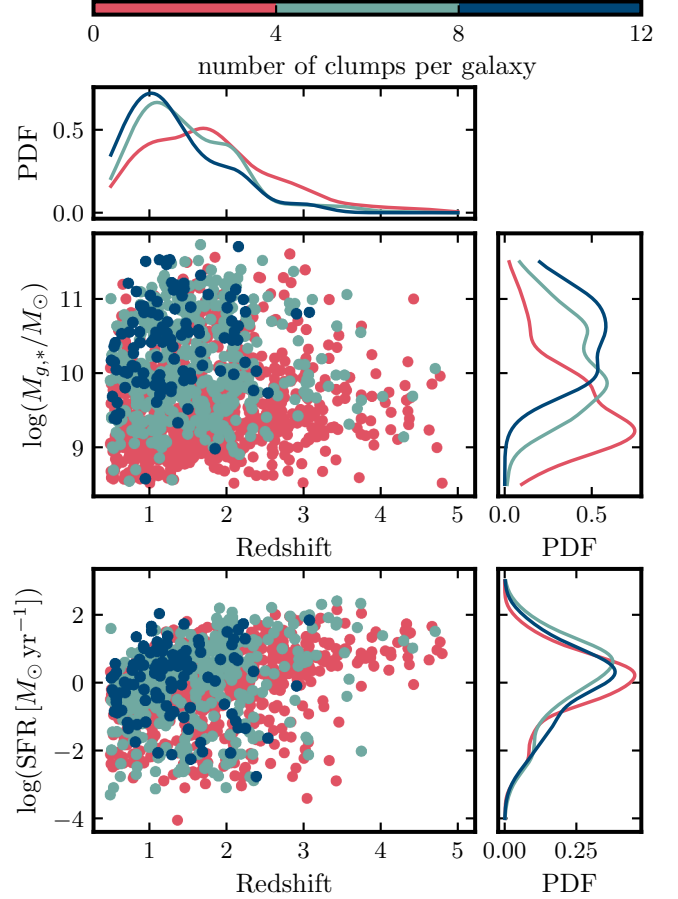


Figure 7. Number of clumps in clumpy galaxies as a function of their host galaxy mass and SFR. The marginal probability distributions, color-coded by the number of clumps, are shown in the top and right panels. Massive galaxies tend to host more clumps relative to less massive galaxies, however we find no preference with the total SFR.

measures the linear correlation between two variables, with coefficients ranging from -1 to 1. A value of one implies positive correlation. We find a strong and significant correlation, with a Pearson coefficient of 0.73 and p-value less than 0.001. The p-value here indicates that the probability of obtaining such a correlation by chance, if no correlation existed, is less than 0.1%.

In general, the correlation between clump mass and host galaxy mass, as shown by the contours of Figure 8, suggests a fundamental link between clump formation and the broader physical context of their host galaxy. In particular, while massive galaxies tend to host the most massive clumps, the contours also suggest that low-mass clumps are less common in massive galaxies. This becomes clearer in §6.2 when we examine the clump mass distribution as a function of galaxy mass and find that galaxies with $\log(M_{g,*}/M_{\odot}) > 10$ appear to have a flatter slope relative to less massive galaxies. This

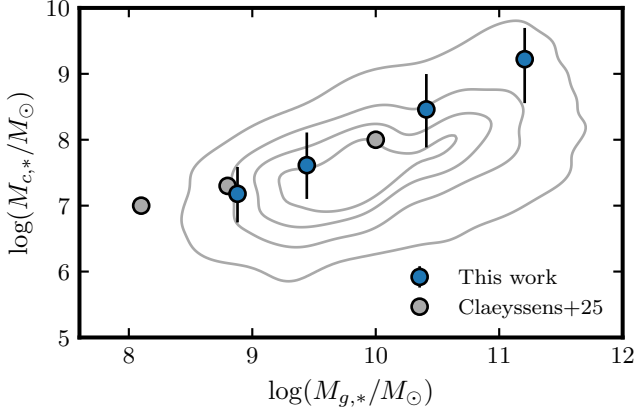


Figure 8. The relationship between clump masses and their host galaxy masses. The contours show the full distribution of the clump sample. To further investigate the relationship between clump mass and host galaxy mass, we select the most massive clump in each galaxy, and compute the median mass of these clumps within bins of host galaxy stellar mass. These median values are shown as blue markers. Similar correlation between the most massive clump mass and host galaxy mass is also observed by [A. Claeysens et al. \(2025\)](#), who performed analyses on a compilation of clumps found in lensed sources between $0.6 < z < 5$.

suggests that massive galaxies have a higher fraction of massive clumps relative to lower mass clumps. We do not expect the absence of low-mass clumps in massive galaxies to be primarily driven by detection biases as we note that our completeness limit is generally similar for all redshift bins at the stellar mass range used for our analysis (see Figure 5). The next section shows that the clump mass distribution across galaxy masses are still different even with a completeness correction.

6.2. The Clump Stellar Mass Function

The clump stellar mass function (cSMF) contains crucial insights on the formation mechanism of clumps. Observations in the local Universe indicate that stellar mass distributions of star clusters often follow a power-law form with a slope of $\alpha = -2$ (e.g., [M. R. Krumholz et al. 2019](#); [A. Adamo et al. 2020](#)), which is expected for structures formed via hierarchical, scale-free fragmentation (e.g., [P. F. Hopkins 2013](#); [D. Guszejnov et al. 2018](#)).

Due to the limited number of clumps identified in individual galaxies, a common practice is to combine clumps across multiple galaxies to analyze their aggregated clump stellar mass function. In the following sections, we investigate the redshift evolution of the aggregated cSMF, as well as the cSMFs for two different galaxy mass bins. In order to measure the clump stellar mass function as a function of redshift, we define the fiducial redshift bin edges at 0.5, 1.2, 2, 3 and 5. We

also account for incompleteness in the following analyses. Our sample reaches an 80% completeness limit for $\log(M_{c,*}/M_{\odot}) \sim 7.7 - 8.3$ between our lowest and highest bins.

6.2.1. Fitting the cSMF

The clump mass function is modeled as a power-law,

$$\frac{dN}{dM_{c,*}} \propto M_{c,*}^{\alpha}. \quad (6)$$

The distribution is first constructed by using uniform bin width of 0.3 dex spanning $7.5 < \log(M_{c,*}/M_{\odot}) < 9.6$. Within each bin, we count the number of clumps and measure the mean clump mass. We apply the completeness correction and restrict the fit to the bins where the average clump mass lies above the 50% completeness limit. The uncertainties of each bin are determined by bootstrap resampling. Specifically, we generate 1000 realizations of the clump sample by randomly sampling the clump masses with replacement. We then compute the binned cSMF for each realization and adopt the dispersion of the resulting distributions as the 1σ uncertainty.

We also fit the clump stellar mass function using a Bayesian approach. We explore the posterior distribution of α using a Markov Chain Monte Carlo (MCMC) method implemented in EMCEE ([D. Foreman-Mackey et al. 2013](#)). The log-likelihood function follows a similar form to the one outlined in [E. Giunchi et al. \(2025\)](#),

$$\ln \mathcal{L}(\alpha) = \sum_{i=1}^{N_{cl}} \ln \left[\frac{\int \text{MF}(x; \alpha) C(x) G_i(y_i | x) dx}{\int \text{MF}(x; \alpha) C(x) dx} \right], \quad (7)$$

where N_{cl} is the number of clumps above the 50% mass completeness, $C(x)$ is the completeness (see §4.2.2), $x = \log(M_{c,*})$, $G_i(y_i | x)$ is the measurement uncertainties of the i -th clump mass described as a Gaussian, and MF is the clump mass function. Similarly, we limit the analyses to clump masses that are above our 50% completeness limit. The values for both methods are provided in Table 1. Both methods yield similar results that are within uncertainties. All quoted cSMF slopes in the paper are based on the binned method.

6.2.2. cSMF for different galaxy mass bins

In this section, we divide the clump mass function by two host galaxy masses, using the fiducial mass bin of $8.5 < \log(M_{g,*}/M_{\odot}) < 10$ and $\log(M_{g,*}/M_{\odot}) \geq 10$ to represent the low- and high-galaxy mass bins, respectively. Figure 9 presents the resulting cSMFs across different redshift and galaxy mass bins. The histograms

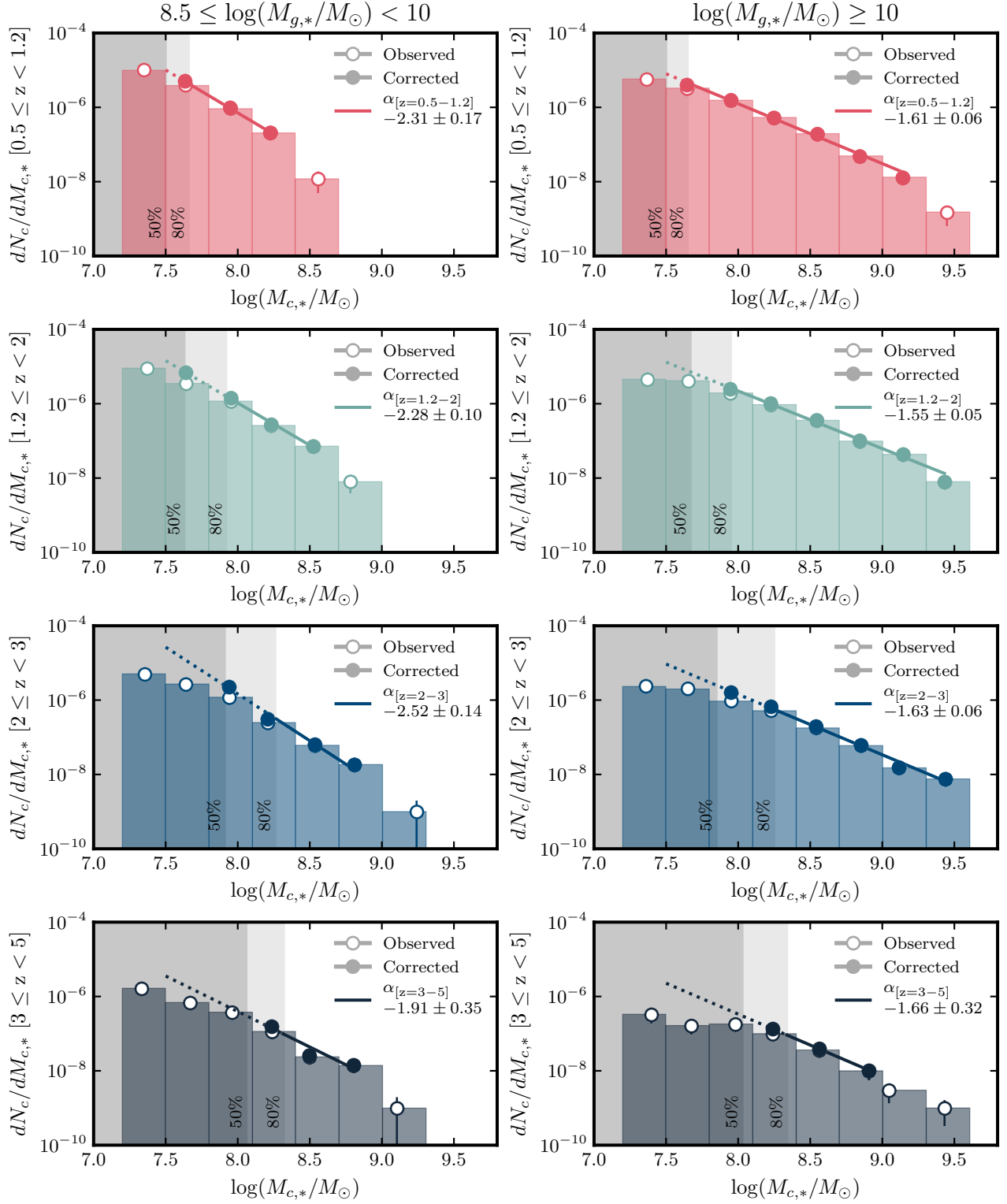


Figure 9. The clump stellar mass function (cSMF) in relation to redshifts and galaxy masses. Open markers denote the observed cSMF, while solid markers show the corrected cSMF, which has a correction factor based on our clump mass completeness limit. The darker and lighter gray regions denote where we are mass complete at 50% and 80%, respectively. The slope of the mass function is only fitted using clump mass bins above our 50% completeness limit and bins that contain more than five clumps.

show the observed (uncorrected) cSMFs, while the solid markers denote the mass-corrected distributions.

We find that the power-law slope of the cSMF is generally flatter for more massive galaxies compared to less massive galaxies, with typical values of $\alpha_{[\text{high-mass}]} \sim -1.6$ compared to $\alpha_{[\text{low-mass}]} \sim -2.3$. This result is consistent with §6.1.2, where we find that massive galaxies tends to host more massive clumps, likely contributing to the flatter cSMF slopes. We also note that the steeper slopes found in less massive galaxies could be partly driven by the physical limit that clump mass cannot be more massive than their host galaxy mass. The cSMF slopes also show little to no evolution with redshift for both galaxy bins, given the large uncertainty associated with each measurement.

We also fit the clump stellar mass functions using only clump mass bins above the 80% completeness limit to investigate whether the observed differences in slope are driven by the incompleteness correction. This is summarized in Table 1. Some redshift bins, particularly for the low galaxy mass bin, have a limited number of clumps, so in these cases the slope is constrained by only two data points. Noting this caveat, we generally find that the results are broadly consistent to the mass completeness correction to the 50% threshold. That is, we find that massive galaxies tend to have flatter slope compared to lower mass systems, with little to no redshift evolution.

6.2.3. Combined cSMF

Figure 10 shows the combined cSMF for the different redshift bins. Again, the sloped line shows the fit to the cSMF, where the solid portion of the line highlights the range of masses where we are mass-complete. In general, we find power-law slopes that are slighter lower than $\alpha = -2$, but is consistent to what are reported by previous studies. Specifically, the typical reported slope ranges between -1.5 to -2 in other studies at redshift of 2 (e.g., M. Dessauges-Zavadsky & A. Adamo 2018; B. S. Kalita et al. 2025a,b). L. Ambachew et al. (2022) found a flatter slope of -1.4 for clumps observed in the DYNAMO sample. This sample consists of galaxies with rotating, marginally stable disk at $z \sim 0.1$, which are proposed to be analogs of high- z clumpy galaxies (e.g., D. B. Fisher et al. 2017). Using the VELA simulations, M. Huertas-Company et al. (2020) reported a shallower slope of $\alpha = -1.5$. However, they found an even shallower slope of $\lesssim -1$ for clumps in galaxies with $\log(M_{g,*}/M_{\odot}) > 9$ at $1 < z < 3$ in CANDELS.

In general, we find no redshift evolution in the slope. This differs from our earlier results where we find a slight redshift evolution that gets flatter with redshift

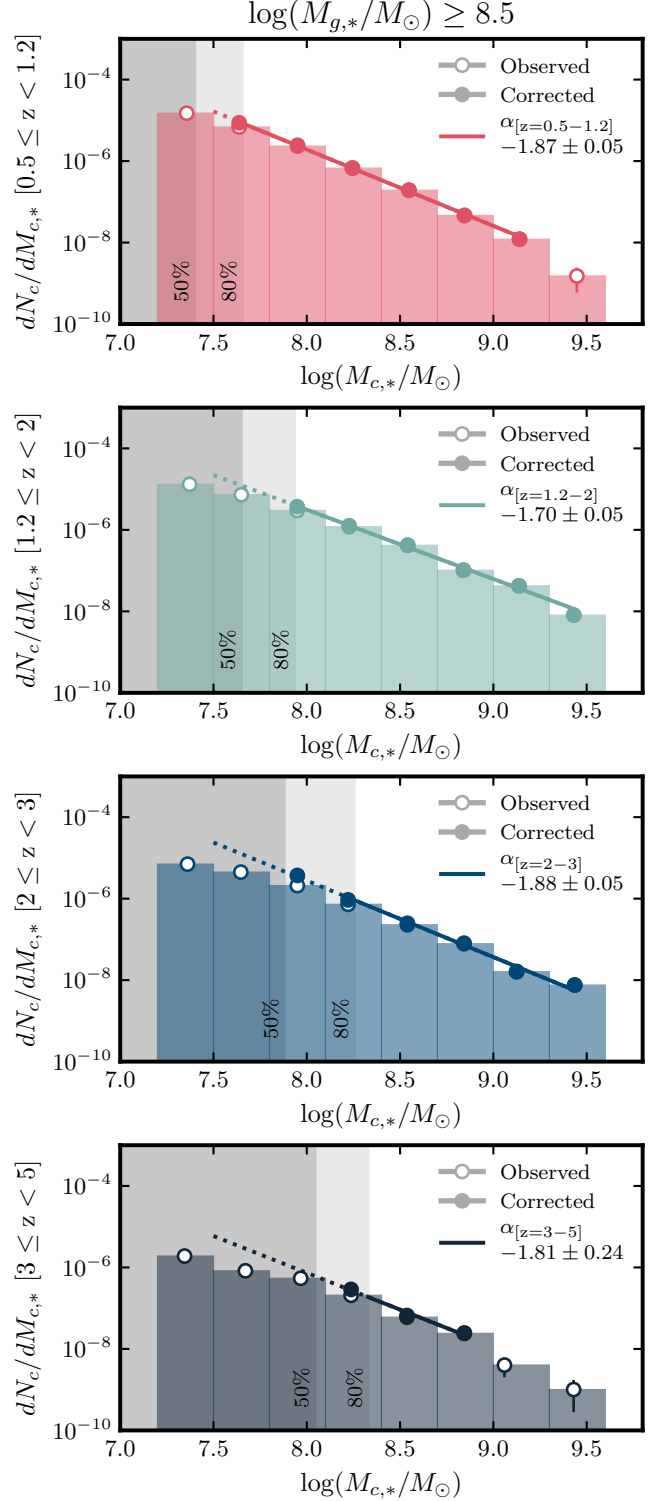


Figure 10. Similar to Fig. 9 but for the combined cSMF. Before combining, a correction factor is applied to each cSMF to account for any completeness differences between the galaxy mass bins. The darker and lighter gray region denote where the mass completeness is 50% and 80% complete, respective.

Table 1. The slope of the clump stellar mass function as a function of redshift and galaxy mass. We include fitted values using both a binned and Bayesian approach. We account for incompleteness down to the 50% limit. In the binned method, uncorrected slopes are measured using only data points above the 80% limit, with no correction applied.

Galaxy Mass Bin	Method	$0.5 < z < 1.2$	$1.2 < z < 2$	$2 < z < 3$	$3 < z < 5$
$8.5 \leq \log(M_{g,*}/M_{\odot}) < 10$	Binned + Uncorrected	-2.47 ± 0.37	-2.18 ± 0.19	-1.91 ± 0.65	-0.77 ± 0.77
	Binned + Corrected	-2.36 ± 0.17	-2.27 ± 0.10	-2.52 ± 0.14	-1.89 ± 0.33
	Bayesian	-2.47 ± 0.10	-2.22 ± 0.12	-2.52 ± 0.19	-2.36 ± 0.46
$\log(M_{g,*}/M_{\odot}) > 10$	Binned + Uncorrected	-1.66 ± 0.08	-1.58 ± 0.06	-1.57 ± 0.13	-1.62 ± 0.67
	Binned + Corrected	-1.61 ± 0.06	-1.56 ± 0.05	-1.61 ± 0.06	-1.66 ± 0.33
	Bayesian	-1.59 ± 0.05	-1.56 ± 0.04	-1.60 ± 0.07	-1.71 ± 0.20
Combined	Binned + Uncorrected	-1.93 ± 0.12	-1.69 ± 0.06	-1.67 ± 0.08	-1.57 ± 0.24
	Binned + Corrected	-1.87 ± 0.05	-1.70 ± 0.05	-1.88 ± 0.05	-1.81 ± 0.24
	Bayesian	-1.91 ± 0.05	-1.73 ± 0.04	-1.87 ± 0.06	-2.04 ± 0.16

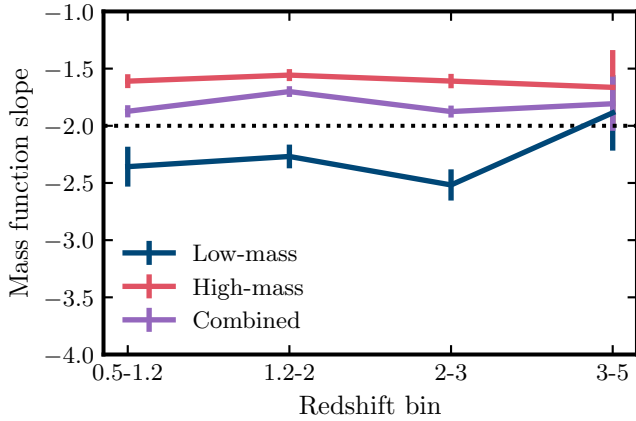


Figure 11. The redshift evolution of the cSMF based on different galaxy mass bins.

for galaxies of lower mass. To assess the impact of completeness corrections on the result, we also fit the slope using clump mass bins that are above the 80% mass completeness limit (e.g., $\log(M_{g,*}/M_{\odot}) \gtrsim 8$). Without any correction, we find similar slope values. There is a more noticeable redshift evolution in the uncorrected cSMF slopes, where it flattens strongly toward higher redshifts. This is likely driven by detection limits, where less massive clumps are becoming increasingly difficult to detect at higher z .

A summary of the slopes is provided in Table 1, and Figure 11 shows the redshift evolution of the slopes for the different galaxy mass bins. When aggregating the entire clump sample, we find that the slope is fairly consistent throughout redshift with a slope of $\alpha \sim -2$. Massive galaxies tend to have flatter slopes compared to less massive galaxies, with $\alpha_{[\text{high-mass}]} \sim -1.6$ compared to $\alpha_{[\text{low-mass}]} \sim -2.3$. This could reflect a mixing

of galaxy population. For example, while differing slope evolution in low- and high-mass galaxies are observed, combining the relative abundances of clumps in different populations can average out these underlying trends. In particular, the cSMF slopes of the combined sample are closer to the high galaxy mass bin, where there are more detected clumps.

7. DISCUSSION

In this work, we identify clumps in star-forming galaxies in CANUCS between $0.5 < z < 5$. We measure the slope of the cSMF, limiting our analyses to clump masses above the 50% completeness limit and applying corrections for incompleteness. When aggregating the clump sample for all galaxies, our results show little evolution in the slope of the clump mass function between $0.5 < z < 5$. Our finding is consistent with what is reported by M. Dessauges-Zavadsky & A. Adamo (2018); A. Claeysens et al. (2025). In the latter case, A. Claeysens et al. (2025) compiled one of the largest sample of ~ 1700 lensed clumps and found that their cumulative mass functions are largely consistent with a power-law slope of $\alpha = -2$ for $0.7 < z < 5.5$. As discussed in M. Dessauges-Zavadsky & A. Adamo (2018), these results are consistent with a scenario where clumps form predominantly *in-situ*, with their mass distribution arising from turbulence-induced gravitational fragmentation. The comparable cSMF slopes across redshifts suggest similar formation conditions.

However, our findings suggest differences in the cSMF for the two different galaxy mass bins, where galaxies with masses of $\log(M_{g,*}/M_{\odot}) > 10$ have slopes of $\alpha_{[\text{high-mass}]} \sim -1.6$ while galaxies with masses of $8.5 < \log(M_{g,*}/M_{\odot}) < 10$ have steeper slopes of

$\alpha_{[\text{low-mass}]} \sim -2.3$. We find that there is little redshift evolution in the slopes for both mass bins. Similarly, [M. Huertas-Company et al. \(2020\)](#) found flatter slope in massive galaxies in the CANDELS sample, but attributed this to incompleteness with low-mass clumps being less complete in massive galaxies. However, we find that this difference cannot be fully explained even with a completeness correction in our analyses.

Taken together, the observed trends of the cSMF likely reflect an interplay between the conditions for clump formation and the mechanisms that destroy them over cosmic time. A flatter slope for a cSMF could reflect either an intrinsically more top-heavy distribution of clumps at formation, or a universal initial cSMF with preferential destruction of low-mass clumps over time, or a combination of both. We explore some of these possibilities in the following discussion. First, we consider potential mechanisms for clump formation and how it may flatten the cSMF slopes, focusing on the roles of gas and mergers. We then examine how clump destruction can also flatten the cSMFs by preferentially disrupting clumps of lower masses.

7.1. Clump formation origins

7.1.1. Are the slopes of the cSMF driven by gas fraction?

The prevalent theory for clump formation is through violent disk instability, regulated by the Toomre Q parameter, which depends on the dynamical state of the disk and its gas content (e.g., [A. Dekel et al. 2009a](#); [D. Ceverino et al. 2010](#)). Consequently, gas fraction plays a role in setting the conditions for instabilities and the resulting clump population. For example, several simulations suggest that higher gas fractions flatten the cSMF slope by enabling the formation of more massive clumps (e.g., [J. Fensch & F. Bournaud 2021](#); [F. Renaud et al. 2024](#)).

We further explore this relation using our observed slopes. We first limit the discussion to galaxies with stellar masses above 10^{10} solar masses due to the larger number of detected clumps in this mass bin. While we do not have a direct measurement of molecular gas, we instead estimate the gas fraction based on the mean stellar masses and SFRs for each of our redshift bin, following the relation described in [T.-M. Wang et al. \(2022\)](#). Figure 12 therefore shows the relationship between increasing gas fraction and the cSMF slope, as well as values reported in [F. Renaud et al. \(2024\)](#). At fixed masses, the galaxies in our redshift bins generally have a higher gas fraction with increasing redshifts. Contrast to what is reported by [F. Renaud et al. \(2024\)](#), we find little correlation between gas fraction and the cSMF slope (open markers).

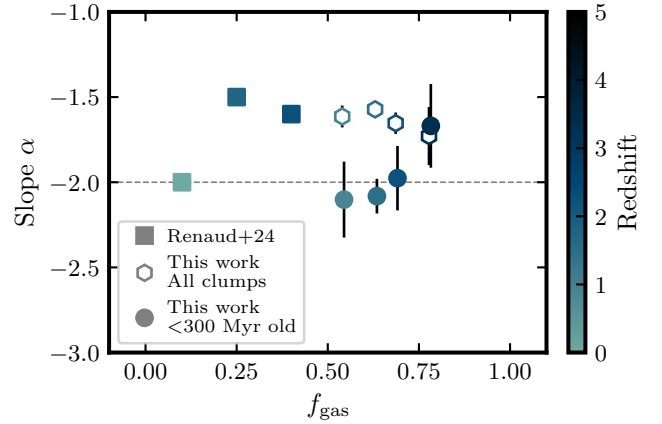


Figure 12. The relationship between the cSMF slope and gas fraction. The data from [F. Renaud et al. \(2024\)](#) represents simulations of isolated galaxies at varying f_{gas} , and are analogous of the Milky-Way at different redshifts, with mass of $\sim 10^{10.7}$ solar masses. In comparison, we plot the slope of cSMF for galaxies between $\log(M_{g,*}/M_{\odot}) \geq 10$ for the different redshift bins and their corresponding f_{gas} , derived using the relation from [T.-M. Wang et al. \(2022\)](#). The open circle markers denote the cSMF slopes derived using all clumps, while the filled circle markers only account for clumps that are < 300 Myr old. The filled circle markers have been shifted by 0.005 in f_{gas} for visual clarity.

However, [J. Fensch & F. Bournaud \(2021\)](#) noted the cSMF slopes tend to flatten from an initial cSMF as clumps evolve and gain mass in their simulation. It could therefore be that initial cSMF is dependent on gas fraction, where more massive clumps are formed in gas-rich disks at beginning, but population mixing as clumps evolve will average out the cSMF slope. To remove the effects of different clump population, we further focus on clumps with ages less than 300 Myr old. We remind the reader that the age is defined to be the lookback time at which 75% of the total stellar mass had formed, derived from SED fitting with a non-parametric SFH and typically more than ten photometric bands. The age criterion therefore provides a closer probe of the initial clump population while retaining a large clump sample for the analyses. The filled circle markers in Figure 12 show the cSMF slope for clumps with ages less than 300 Myr old. We find that the slope of the cSMF of younger clumps appears to correlate with gas fraction, albeit with larger uncertainties.

Contradictory, the clump stellar mass function slopes for low-mass galaxies are steeper than those in high-mass systems, despite potentially having higher molecular gas fractions (i.e., if the relation from [T.-M. Wang et al. \(2022\)](#) extends to lower masses). However, this paradox can be resolved by two effects that regulate clump formation. Firstly, massive clumps are preferentially formed

in massive disks that have larger total gas reservoirs. From Toomre analyses, the fragmentation scale in a differentially rotating disk is given by $\lambda_c \sim G\Sigma/\Omega^2$ (e.g., J. Binney & S. Tremaine 2008, Chapter 6.2.3), where G is the gravitational constant, Σ is the gas density, and Ω is the angular circular velocity. A. Dekel et al. (2009a) showed that this sets the maximum clump scale, with $\lambda_c \equiv R_c \propto \delta R_d$, and the mass of clumps as $M_c \propto \delta^2 M_d$. Here, δ is the disk-to-total mass fraction, R_d and M_d are the radius and mass of the gas disk, respectively. This scaling implies that more massive disks preferentially host more massive clumps. Secondly, the number of clump formed is related to the dynamical states of the disk, with more clumps typically produced in rotationally supported systems (e.g., M. E. Orr & D. Rennehan 2024). Less massive galaxies, which are generally more dispersion-dominated (e.g., M. A. C. de los Reyes et al. 2023), will generally form both fewer and lower mass clumps.

It is important to note a few caveats in this discussion. The simulated galaxies in F. Renaud et al. (2024) represent Milk-Way analogues, spanning from local universe to $z \sim 2.5$, and are isolated systems without any evolutionary, cosmological contexts. Their gas fraction f_{gas} also span between 0.1 – 0.4. Our galaxies span between $0.5 < z < 5$, and are estimated to be more gas-rich based on the relation from T.-M. Wang et al. (2022), with $0.5 < f_{\text{gas}} < 0.7$. While our galaxies are different compared to those in F. Renaud et al. (2024), we emphasize that this comparison only aims to investigate how the slope of the cSMF changes in different gas environment, but for galaxies of comparable masses. While the result suggests a trend between increasing gas fractions and flatter cSMF slope for massive galaxies, this correlation cannot account for the differences in the slope across galaxy mass bins. This points to other factors that affect the shape of the clump mass distribution.

7.1.2. Can mergers affect the cSMF?

In this section, we now consider how mergers, specifically clumps that are of *ex-situ* origins, affect the slope of the cSMF. Clumps of *ex-situ* origins, possibly accreting satellites, generally have higher masses relative to those formed *in-situ* (e.g., N. Mandelker et al. 2014, 2017), possibly flattening the slope of the mass function. Observational studies, including B. Ribeiro et al. (2017); A. Zanella et al. (2019), have also suggested that some clumps are remnants of accreting satellites, due to their extended sizes and larger masses relative to other clump populations.

The major merger pair fraction is observed to increase from $z = 0$ to $z = 3$, but remains flat approximately af-

ter $z \sim 3$ (e.g., Q. Duan et al. 2025; D. Puskás et al. 2025). If *ex-situ* clumps are typically larger and more massive compared to *in-situ* clumps, this suggests that the cSMF slopes may flatten with redshift when mergers are more frequent. A visual inspection of Figure 3 shows a lack of disk galaxy at higher redshifts, where, instead, clumpy galaxies resemble galaxy pairs or satellites. This hints at the possibility a part of the clump populations is of *ex-situ* origins. Indeed, a recent resolved spectroscopic observation with JWST/NIRISS indicates metal-poor gas accretion in cosmic noon galaxies, in which a non-negligible fraction of clumps is observed having lower gas-phase metallicity compared to the host disks (V. Estrada-Carpenter et al. 2025). Simulations have shown that proto-clumps can form along gas filaments accreting onto the host galaxies (F. van Donkelaar et al. 2023). In these scenarios, such metallicity offsets in clumps point to other formation channels such as accreting gas clumps.

In addition to gas fraction, clumps of *ex-situ* origins may also contribute to the clump mass distribution by flattening the distribution if they represent a higher proportion of the clump population relative to those formed *in-situ*. Identifying satellites based on imaging alone is challenging, and future works involving large IFU surveys are needed to be able to distinguish *in-situ* and *ex-situ* clumps.

7.2. Clump destruction

In the previous discussions, we investigate how the slopes of the cSMF are affected by formation mechanisms. Here, we consider whether the cSMF is shaped by processes that destroy clumps. For example, massive galaxies tend to exhibit a flatter clump mass function slope compared to low-mass galaxies, which could result from either the creation of more massive clumps, or from the preferential destruction of low-mass clumps. Feedback processes, such as supernovae and radiation pressure, limit clump survivability, particularly at lower clump masses (e.g., C. E. Moody et al. 2014; N. Mandelker et al. 2017; O. Ginzburg et al. 2021; A. Dekel et al. 2022). Their survivability is also further dependent on clump and disk kinematics, along with bulge and disk mass fraction (A. Dekel et al. 2023).

Here, we test whether lower mass clumps are preferentially disrupted by comparing clump mass and age, using clump ages to define two samples of young and old clumps. The top of Figure 13 shows the distribution of clump ages in low- and high-mass galaxies. The bottom panel shows the relationship between clump mass, their age, and the host galaxy mass. As suggested in the top panel, in addition to the formation of new young clumps,

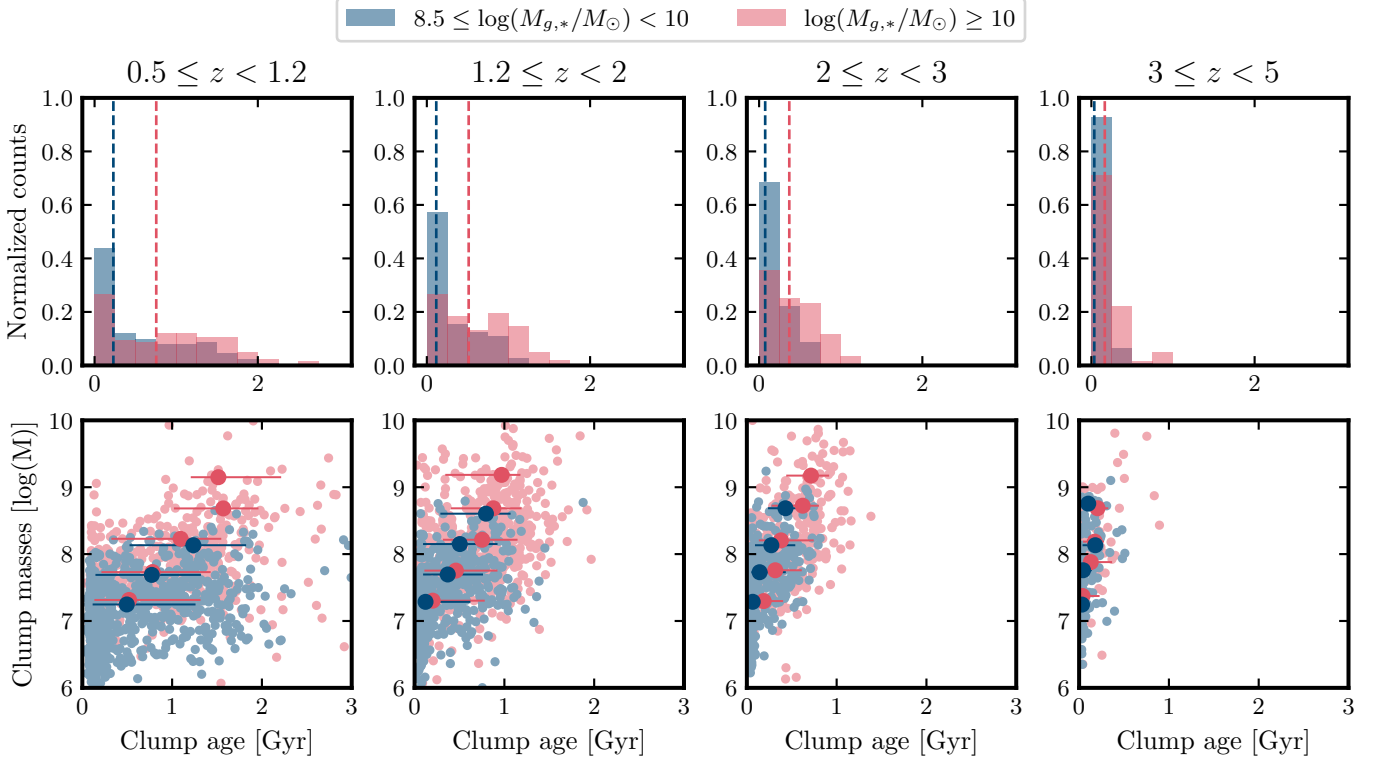


Figure 13. The top panel shows the normalized distribution of clump ages, while the bottom panels show the distribution of clump ages with respect to their masses. The error bars denote the 16th, 50th and 84th percentile. We find that the age of clumps is correlated with their mass.

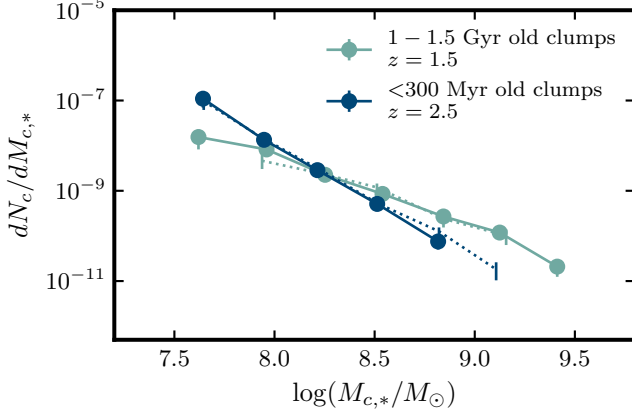


Figure 14. The evolution of the normalized clump mass function for young clumps (<300 Myr old) over ~ 1.5 Gyr of cosmic time. As clumps evolve, the mass function shows a relative absence of lower mass clumps and the abundance of massive clumps, consistent with a scenario for the preferential destruction of lower mass clumps. The arrow denotes the average mass that would be gained from star formation assuming that the clump's star formation rate remains constant over this period. This result is independent of the choice for the age selection (e.g., shifting the selection window by 200 Myr does not affect the overall flattening of the clump mass function, which is shown as the dotted line).

there is also a population of clumps that appear to be aging, as indicated by the shift of the age distribution toward older values at lower redshifts.

If clump disruption is not mass-preferential, the shape of the cSMF should remain similar between young and old populations. This scenario is tested by selecting a population of young clumps (< 300 Myr) in our high-mass galaxy bin at $z \sim 2.5$ and compare its cSMF to a population of older clumps (> 1 Gyr and < 1.5 Gyr) at $z \sim 1.5$ at the same galaxy mass bin. Again, we focus on the high galaxy mass bin due to the larger clump sample size. The choice of redshift intervals here is motivated by the fact that there is approximately 1.5 Gyr of cosmic time between $z \sim 2.5$ to $z \sim 1.5$, roughly tracing the age evolution of the clump populations under the assumption that the young clumps at $z \sim 2.5$ will evolve to become those found at $z \sim 1.5$. Figure 14 shows the cSMF of the young and old clump populations, normalized by the number of clumps in each sample. We find that the mass function slope is steeper for younger clumps, pointing to a higher abundance of low-mass clumps. However, the cSMF becomes flatter for as the clump population ages. This trend supports the scenario in which lower-mass clumps are preferentially disrupted, while massive clumps are more likely to sur-

vive and age as their host galaxies evolve. This finding suggests that the clump mass function is not only affected by formation mechanisms, but also by mass-dependent disruption processes.

8. CONCLUSION

In this work, we use imaging from the CANUCS fields to measure the properties of clumpy star-forming galaxies and their clumps at $0.5 < z < 5$. Particularly, we measure the evolution of the fraction of clumpy SFGs with cosmic time, and investigate various relationships between clump properties and their host galaxies. We also examine the evolution of the clump mass distribution as a function of galaxy masses and redshifts. We summarize our main finding below:

- The fraction of clumpy galaxies (f_{clumpy}) evolves differently for different galaxy mass bins. For galaxies with $\log(M_{g,*}/M_{\odot}) \geq 10$, the fraction of clumpy SFGs peaks near cosmic noon, and decreases at earlier and later times. We find little to no redshift evolution in the fraction of clumpy SFGs of $8.5 \leq \log(M_{g,*}/M_{\odot}) < 10$.
- Massive galaxies tend to host more clumps compared to less massive galaxies. However, we do not find the number of clumps per galaxies to be dependent on the integrated star formation rate of the host galaxies.
- We find that the mass of clumps appears to correlate with the masses of the host galaxies, implying that massive clumps are generally found in galaxies with massive disks.
- When aggregating all clumps, we find little redshift evolution in the clump stellar mass function slope.
- When separating the clump stellar mass function based on galaxy mass bin, we find that galaxies with $\log(M_{g,*}/M_{\odot}) \geq 10$ tend to have flatter slopes ($\alpha_{[\text{high-mass}]} \sim -1.6$) relative to less massive galaxies ($\alpha_{[\text{low-mass}]} \sim -2.3$). This suggests either an abundance of massive clumps or an absence of less massive clumps in massive galaxies.
- We also explore the relationship between the cSMF slope and inferred gas fraction. We find that when considering all clumps, the cSMF slope of galaxies with $\log(M_{g,*}/M_{\odot}) \geq 10$ has is uncorrelated to gas fraction. However, when considering younger clumps with age less than 300 Myr old, the slope of the cSMF is flatter with increasing

gas fraction. This suggests that higher gas fraction leads to more massive clumps at the time of formation.

- By comparing the slope of a young clump population (< 300 Myr old) at $z \sim 2.5$ to that of an older clump population with ages between 1 to 1.5 Gyr old at $z \sim 1.5$, we find that the slope is steeper for the younger clump population. Assuming that the young clump population evolves to become the older clump population, the flattening of the cSMF slope suggests that lower mass clumps are preferentially destroyed, while more massive clumps are likely to survive disruptive feedback processes.

Overall, our results show that the aggregated clump mass function remains relatively constant with redshift. However, when considering the cSMF by host galaxy mass bin, we find that it is sensitive to galaxy masses. This likely reflects an interplay between the conditions for clump formation and feedback regulation that disrupt them. Our results are consistent with the interpretation that clump formation in massive galaxies ($\log(M_{g,*}/M_{\odot}) \geq 10$) happens as a result of instabilities and fragmentation within gas-rich disks. Higher gas fraction can lead to the formation of more massive clumps, which flattens the slope of the clump mass function. However, alternative formation mechanisms, such as accreting satellites, can also be responsible for creating flatter slopes as clumps of *ex-situ* origins are generally considered to be more massive than formed *in-situ*. Together with feedback mechanisms that preferentially disrupt lower-mass clumps, these processes shape the clump mass distribution across different galaxy populations. Additional observational studies will be critical to better constrain the origin of clumps. Particularly, resolved IFU spectroscopy can disentangle the relative contributions of clumps formed via *in-situ* fragmentation and *ex-situ* accretion by exploring the resolved kinematics of clumps with respect to the host disks. Such studies will provide essential insights into the origin of clumps and their role in galaxy evolution.

ACKNOWLEDGMENTS

This research was enabled by grant 18JWST-GTO1 from the Canadian Space Agency, and Discovery Grant funding from the Natural Sciences and Engineering Research Council (NSERC) of Canada.

MB acknowledges support from the ERC Grant FIRSTLIGHT, Slovenian national research agency ARIS through grants N1-0238 and P1-0188, and the pro-

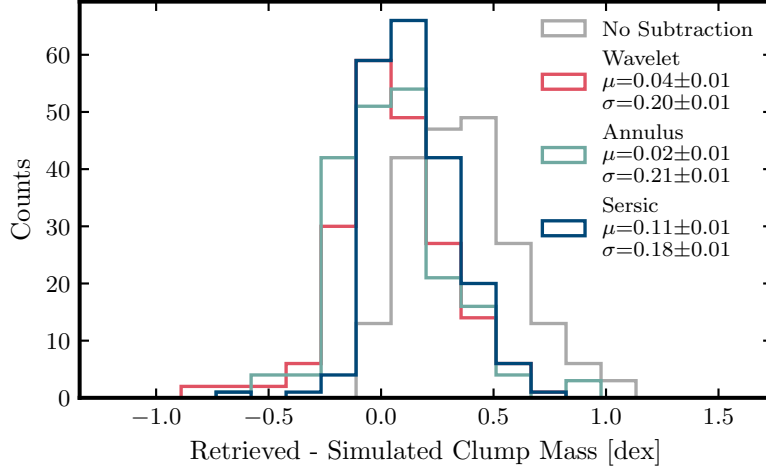


Figure 15. A comparison of different background subtraction methods for retrieving clump masses. In particular, we tested three different methods to estimate the underlying stellar disk; by fitting the disk as a 2D Sérsic profile, using wavelet transformation to construct a smooth disk (see text for details), and by performing an annular subtraction to estimate local background. We find that all three methods can recover clump masses, but the annular subtraction method generally results with more consistent clump masses.

gram HST-GO-16667, provided through a grant from the STScI under NASA contract NAS5-26555.

This research used the Canadian Advanced Network For Astronomy Research (CANFAR) operated in part-

nership with the Canadian Astronomy Data Centre and The Digital Research Alliance of Canada, with support from the National Research Council of Canada, the Canadian Space Agency, CANARIE, and the Canadian Foundation for Innovation.

APPENDIX

A. MASS RETRIEVAL WITH DIFFERENT BACKGROUND SUBTRACTION

We test different disk subtraction methods to determine the most reliable approach to extract clump masses. Specifically, we estimate the underlying stellar disk contribution using three techniques: (1) fitting a 2D Sérsic profile, (2) modeling a smooth disk with wavelet transformation and by setting lower scales to zero, and (3) performing local background subtraction by estimating the mean background within an annular aperture around each clump.

To model the disk using a Sérsic profile, we follow the methodology described in [V. Y. Y. Tan et al. \(2022\)](#). In brief, we normalize the stellar mass maps to the F444W images, and inject Gaussian noise sampled from the pixel-to-pixel noise of the F444W image. We then use STATMORPH ([V. Rodriguez-Gomez et al. 2019](#)) to determine the best Sérsic model, which is then scaled back to the original stellar mass range. For wavelet transformation, similar to what is described in §4.2, we decompose the mass maps into 6 different wavelet scales using SCARLET and estimate the underlying disk by setting the first two scales to zero. This effectively remove structures with scales less than $0.2''$. These estimates of the stellar disk are then subtracted from the stellar mass maps to produce background-subtracted mass maps. Finally, the local background using the annular method involves measuring the mean background within two annuli with radii of 3.5 and 4.5 pixels (corresponding to 0.14 and 0.18 arcsec, respectively). This estimated background is then subtracted from the aperture measurement.

To test the robustness of each method, we inject simulated clump of mass between $7.0 < \log(M_*/M_\odot) < 7.8$ into the stellar mass map of nonclumpy galaxies. Nonclumpy galaxies have smoother mass distributions, and therefore present baseline for these tests. We add two clumps to each galaxy at random location, with FWHM of $0.2''$. We randomize the position angle of each clumps, with ellipticity ranging between $0.1 < e < 0.2$. The mass of each injected clump is recovered by measuring the mass within a aperture of radius $0.18''$. It should be noted that the underlying stellar disk can be overestimated for some clumps, leading to over-subtraction of the clump mass. This occurs more frequently for both the Sérsic and wavelet methods.

Figure 15 shows the distribution of the retrieved clump masses relative to their clump mass. In general, all three methods are able to account for the local stellar disk, resulting in typical mass correction of 0.4 dex. While each method has its strengths and weaknesses, we adopt the annular subtraction method for this work. This is because using wavelets tends to overestimate the disk contribution, while the Sérsic model does not always capture irregular features of the disk.

REFERENCES

- Abdurro'uf, Lin, Y.-T., Wu, P.-F., & Akiyama, M. 2021, *ApJS*, 254, 15, doi: [10.3847/1538-4365/abebe2](https://doi.org/10.3847/1538-4365/abebe2)
- Adamo, A., Hollyhead, K., Messa, M., et al. 2020, *MNRAS*, 499, 3267, doi: [10.1093/mnras/staa2380](https://doi.org/10.1093/mnras/staa2380)
- Adamo, A., Bradley, L. D., Vanzella, E., et al. 2024, *Nature*, 632, 513, doi: [10.1038/s41586-024-07703-7](https://doi.org/10.1038/s41586-024-07703-7)
- Adams, D., Mehta, V., Dickinson, H., et al. 2022, *ApJ*, 931, 16, doi: [10.3847/1538-4357/ac6512](https://doi.org/10.3847/1538-4357/ac6512)
- Agertz, O., Teyssier, R., & Moore, B. 2009, *MNRAS*, 397, L64, doi: [10.1111/j.1745-3933.2009.00685.x](https://doi.org/10.1111/j.1745-3933.2009.00685.x)
- Ambachew, L., Fisher, D. B., Glazebrook, K., et al. 2022, *MNRAS*, 512, 3079, doi: [10.1093/mnras/stac622](https://doi.org/10.1093/mnras/stac622)
- Antwi-Danso, J., Papovich, C., Leja, J., et al. 2023, *ApJ*, 943, 166, doi: [10.3847/1538-4357/aca294](https://doi.org/10.3847/1538-4357/aca294)
- Binney, J., & Tremaine, S. 2008, *Galactic Dynamics: Second Edition*
- Bradley, L., Sipőcz, B., Robitaille, T., et al. 2024, 2.0.2 Zenodo, doi: [10.5281/zenodo.13989456](https://doi.org/10.5281/zenodo.13989456)
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503, doi: [10.1086/591786](https://doi.org/10.1086/591786)
- Cacciato, M., Dekel, A., & Genel, S. 2012, *MNRAS*, 421, 818, doi: [10.1111/j.1365-2966.2011.20359.x](https://doi.org/10.1111/j.1365-2966.2011.20359.x)
- Calabrò, A., Daddi, E., Fensch, J., et al. 2019, *A&A*, 632, A98, doi: [10.1051/0004-6361/201935778](https://doi.org/10.1051/0004-6361/201935778)
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682, doi: [10.1086/308692](https://doi.org/10.1086/308692)
- Cappellari, M., & Copin, Y. 2003, *MNRAS*, 342, 345, doi: [10.1046/j.1365-8711.2003.06541.x](https://doi.org/10.1046/j.1365-8711.2003.06541.x)
- Cava, A., Schaerer, D., Richard, J., et al. 2018, *Nature Astronomy*, 2, 76, doi: [10.1038/s41550-017-0295-x](https://doi.org/10.1038/s41550-017-0295-x)
- Ceverino, D., Dekel, A., & Bournaud, F. 2010, *MNRAS*, 404, 2151, doi: [10.1111/j.1365-2966.2010.16433.x](https://doi.org/10.1111/j.1365-2966.2010.16433.x)
- Ceverino, D., Sánchez Almeida, J., Muñoz Tuñón, C., et al. 2016, *MNRAS*, 457, 2605, doi: [10.1093/mnras/stw064](https://doi.org/10.1093/mnras/stw064)
- Chabrier, G. 2003, *PASP*, 115, 763, doi: [10.1086/376392](https://doi.org/10.1086/376392)
- Cid Fernandes, R., Pérez, E., García Benito, R., et al. 2013, *A&A*, 557, A86, doi: [10.1051/0004-6361/201220616](https://doi.org/10.1051/0004-6361/201220616)
- Claeyssens, A., Adamo, A., Messa, M., et al. 2025, *MNRAS*, 537, 2535, doi: [10.1093/mnras/staf058](https://doi.org/10.1093/mnras/staf058)
- Daddi, E., Bournaud, F., Walter, F., et al. 2010, *ApJ*, 713, 686, doi: [10.1088/0004-637X/713/1/686](https://doi.org/10.1088/0004-637X/713/1/686)
- Daddi, E., Delvecchio, I., Dimauro, P., et al. 2022, *A&A*, 661, L7, doi: [10.1051/0004-6361/202243574](https://doi.org/10.1051/0004-6361/202243574)
- de la Vega, A., Mobasher, B., Manesh, F., et al. 2025, *arXiv e-prints*, arXiv:2508.14972, <https://arxiv.org/abs/2508.14972>
- de los Reyes, M. A. C., Kirby, E. N., Zhuang, Z., et al. 2023, *ApJ*, 951, 52, doi: [10.3847/1538-4357/acd189](https://doi.org/10.3847/1538-4357/acd189)
- Dekel, A., Mandelker, N., Bournaud, F., et al. 2022, *MNRAS*, 511, 316, doi: [10.1093/mnras/stab3810](https://doi.org/10.1093/mnras/stab3810)
- Dekel, A., Sari, R., & Ceverino, D. 2009a, *ApJ*, 703, 785, doi: [10.1088/0004-637X/703/1/785](https://doi.org/10.1088/0004-637X/703/1/785)
- Dekel, A., Tziperman, O., Sarkar, K. C., et al. 2023, *MNRAS*, 521, 4299, doi: [10.1093/mnras/stad855](https://doi.org/10.1093/mnras/stad855)
- Dekel, A., Birnboim, Y., Engel, G., et al. 2009b, *Nature*, 457, 451, doi: [10.1038/nature07648](https://doi.org/10.1038/nature07648)
- Dessauges-Zavadsky, M., & Adamo, A. 2018, *MNRAS*, 479, L118, doi: [10.1093/mnrasl/sly112](https://doi.org/10.1093/mnrasl/sly112)
- Dessauges-Zavadsky, M., Schaerer, D., Cava, A., Mayer, L., & Tamburello, V. 2017, *ApJL*, 836, L22, doi: [10.3847/2041-8213/aa5d52](https://doi.org/10.3847/2041-8213/aa5d52)
- Duan, Q., Conselice, C. J., Li, Q., et al. 2025, *MNRAS*, 540, 774, doi: [10.1093/mnras/staf638](https://doi.org/10.1093/mnras/staf638)
- Elmegreen, D. M., Elmegreen, B. G., Rubin, D. S., & Schaffer, M. A. 2005, *ApJ*, 631, 85, doi: [10.1086/432502](https://doi.org/10.1086/432502)
- Estrada-Carpenter, V., Sawicki, M., Brammer, G., et al. 2024, *MNRAS*, 532, 577, doi: [10.1093/mnras/stae1368](https://doi.org/10.1093/mnras/stae1368)
- Estrada-Carpenter, V., Sawicki, M., Abraham, R., et al. 2025, *arXiv e-prints*, arXiv:2508.00985, doi: [10.48550/arXiv.2508.00985](https://doi.org/10.48550/arXiv.2508.00985)
- Fensch, J., & Bournaud, F. 2021, *MNRAS*, 505, 3579, doi: [10.1093/mnras/stab1489](https://doi.org/10.1093/mnras/stab1489)
- Ferreira, L., Conselice, C. J., Sazonova, E., et al. 2023, *ApJ*, 955, 94, doi: [10.3847/1538-4357/acec76](https://doi.org/10.3847/1538-4357/acec76)
- Fisher, D. B., Glazebrook, K., Abraham, R. G., et al. 2017, *ApJL*, 839, L5, doi: [10.3847/2041-8213/aa6478](https://doi.org/10.3847/2041-8213/aa6478)
- Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *PASP*, 125, 306, doi: [10.1086/670067](https://doi.org/10.1086/670067)
- Förster Schreiber, N. M., Shapley, A. E., Erb, D. K., et al. 2011a, *ApJ*, 731, 65, doi: [10.1088/0004-637X/731/1/65](https://doi.org/10.1088/0004-637X/731/1/65)
- Förster Schreiber, N. M., Shapley, A. E., Genzel, R., et al. 2011b, *ApJ*, 739, 45, doi: [10.1088/0004-637X/739/1/45](https://doi.org/10.1088/0004-637X/739/1/45)

- Förster Schreiber, N. M., Renzini, A., Mancini, C., et al. 2018, *ApJS*, 238, 21, doi: [10.3847/1538-4365/aadd49](https://doi.org/10.3847/1538-4365/aadd49)
- Fujimoto, S., Ouchi, M., Kohno, K., et al. 2025, *Nature Astronomy*, doi: [10.1038/s41550-025-02592-w](https://doi.org/10.1038/s41550-025-02592-w)
- Ginzburg, O., Huertas-Company, M., Dekel, A., et al. 2021, *MNRAS*, 501, 730, doi: [10.1093/mnras/staa3778](https://doi.org/10.1093/mnras/staa3778)
- Giunchi, E., Scarlata, C., Werle, A., et al. 2025, *A&A*, 696, A228, doi: [10.1051/0004-6361/202452983](https://doi.org/10.1051/0004-6361/202452983)
- Gledhill, R., Strait, V., Desprez, G., et al. 2024, *ApJ*, 973, 77, doi: [10.3847/1538-4357/ad684a](https://doi.org/10.3847/1538-4357/ad684a)
- Guo, Y., Ferguson, H. C., Bell, E. F., et al. 2015, *ApJ*, 800, 39, doi: [10.1088/0004-637X/800/1/39](https://doi.org/10.1088/0004-637X/800/1/39)
- Guo, Y., Rafelski, M., Bell, E. F., et al. 2018, *ApJ*, 853, 108, doi: [10.3847/1538-4357/aaa018](https://doi.org/10.3847/1538-4357/aaa018)
- Guszejnov, D., Hopkins, P. F., & Grudić, M. Y. 2018, *MNRAS*, 477, 5139, doi: [10.1093/mnras/sty920](https://doi.org/10.1093/mnras/sty920)
- Han, Y., & Han, Z. 2019, *ApJS*, 240, 3, doi: [10.3847/1538-4365/aaeffa](https://doi.org/10.3847/1538-4365/aaeffa)
- Hinojosa-Goni, R., Muñoz-Tuñón, C., & Méndez-Abreu, J. 2016, *A&A*, 592, A122, doi: [10.1051/0004-6361/201527066](https://doi.org/10.1051/0004-6361/201527066)
- Hopkins, P. F. 2013, *MNRAS*, 430, 1653, doi: [10.1093/mnras/sts704](https://doi.org/10.1093/mnras/sts704)
- Huertas-Company, M., Guo, Y., Ginzburg, O., et al. 2020, *MNRAS*, 499, 814, doi: [10.1093/mnras/staa2777](https://doi.org/10.1093/mnras/staa2777)
- Iyer, K., & Gawiser, E. 2017, *ApJ*, 838, 127, doi: [10.3847/1538-4357/aa63f0](https://doi.org/10.3847/1538-4357/aa63f0)
- Iyer, K. G., Gawiser, E., Faber, S. M., et al. 2019, *ApJ*, 879, 116, doi: [10.3847/1538-4357/ab2052](https://doi.org/10.3847/1538-4357/ab2052)
- Kalita, B. S., Silverman, J. D., Daddi, E., et al. 2025a, *MNRAS*, 537, 402, doi: [10.1093/mnras/staf031](https://doi.org/10.1093/mnras/staf031)
- Kalita, B. S., Suzuki, T. L., Kashino, D., et al. 2025b, *MNRAS*, 536, 3090, doi: [10.1093/mnras/stae2781](https://doi.org/10.1093/mnras/stae2781)
- Kartalpe, J. S., Rose, C., Vanderhoof, B. N., et al. 2023, *ApJL*, 946, L15, doi: [10.3847/2041-8213/acad01](https://doi.org/10.3847/2041-8213/acad01)
- Krumholz, M. R., McKee, C. F., & Bland-Hawthorn, J. 2019, *ARA&A*, 57, 227, doi: [10.1146/annurev-astro-091918-104430](https://doi.org/10.1146/annurev-astro-091918-104430)
- Leja, J., Carnall, A. C., Johnson, B. D., Conroy, C., & Speagle, J. S. 2019, *ApJ*, 876, 3, doi: [10.3847/1538-4357/ab133c](https://doi.org/10.3847/1538-4357/ab133c)
- Livermore, R. C., Jones, T. A., Richard, J., et al. 2015, *MNRAS*, 450, 1812, doi: [10.1093/mnras/stv686](https://doi.org/10.1093/mnras/stv686)
- Lotz, J. M., Koekemoer, A., Coe, D., et al. 2017, *ApJ*, 837, 97, doi: [10.3847/1538-4357/837/1/97](https://doi.org/10.3847/1538-4357/837/1/97)
- Lower, S., Narayanan, D., Leja, J., et al. 2020, *ApJ*, 904, 33, doi: [10.3847/1538-4357/abbfa7](https://doi.org/10.3847/1538-4357/abbfa7)
- Madau, P., & Dickinson, M. 2014, *ARA&A*, 52, 415, doi: [10.1146/annurev-astro-081811-125615](https://doi.org/10.1146/annurev-astro-081811-125615)
- Maíz Apellániz, J. 2006, *AJ*, 131, 1184, doi: [10.1086/499158](https://doi.org/10.1086/499158)
- Mandelker, N., Dekel, A., Ceverino, D., et al. 2017, *MNRAS*, 464, 635, doi: [10.1093/mnras/stw2358](https://doi.org/10.1093/mnras/stw2358)
- Mandelker, N., Dekel, A., Ceverino, D., et al. 2014, *MNRAS*, 443, 3675, doi: [10.1093/mnras/stu1340](https://doi.org/10.1093/mnras/stu1340)
- Mayer, L., van Dokelaar, F., Messa, M., Capelo, P. R., & Adamo, A. 2025, *ApJL*, 981, L28, doi: [10.3847/2041-8213/adadfe](https://doi.org/10.3847/2041-8213/adadfe)
- Melchior, P., Moolekamp, F., Jerdee, M., et al. 2018, *Astronomy and Computing*, 24, 129, doi: [10.1016/j.ascom.2018.07.001](https://doi.org/10.1016/j.ascom.2018.07.001)
- Mercier, W., Kalita, B. S., Shuntov, M., et al. 2025, *arXiv e-prints*, arXiv:2506.13881, doi: [10.48550/arXiv.2506.13881](https://doi.org/10.48550/arXiv.2506.13881)
- Messa, M., Dessauges-Zavadsky, M., Adamo, A., Richard, J., & Claeysens, A. 2024, *MNRAS*, 529, 2162, doi: [10.1093/mnras/stae565](https://doi.org/10.1093/mnras/stae565)
- Meštrić, U., Vanzella, E., Zanella, A., et al. 2022, *MNRAS*, 516, 3532, doi: [10.1093/mnras/stac2309](https://doi.org/10.1093/mnras/stac2309)
- Moody, C. E., Guo, Y., Mandelker, N., et al. 2014, *MNRAS*, 444, 1389, doi: [10.1093/mnras/stu1534](https://doi.org/10.1093/mnras/stu1534)
- Morishita, T., Stiavelli, M., Vanzella, E., et al. 2025, *ApJ*, 985, 83, doi: [10.3847/1538-4357/adc4c3](https://doi.org/10.3847/1538-4357/adc4c3)
- Mowla, L., Iyer, K., Asada, Y., et al. 2024, *Nature*, 636, 332, doi: [10.1038/s41586-024-08293-0](https://doi.org/10.1038/s41586-024-08293-0)
- Murata, K. L., Kajisawa, M., Taniguchi, Y., et al. 2014, *ApJ*, 786, 15, doi: [10.1088/0004-637X/786/1/15](https://doi.org/10.1088/0004-637X/786/1/15)
- Nakazato, Y., Ceverino, D., & Yoshida, N. 2024, *ApJ*, 975, 238, doi: [10.3847/1538-4357/ad7d0b](https://doi.org/10.3847/1538-4357/ad7d0b)
- Noiro, G., Sawicki, M., Abraham, R., et al. 2022, *MNRAS*, 512, 3566, doi: [10.1093/mnras/stac668](https://doi.org/10.1093/mnras/stac668)
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, 266, 713, doi: [10.1086/160817](https://doi.org/10.1086/160817)
- Orr, M. E., & Rennehan, D. 2024, *arXiv e-prints*, arXiv:2410.23337, doi: [10.48550/arXiv.2410.23337](https://doi.org/10.48550/arXiv.2410.23337)
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, *AJ*, 124, 266, doi: [10.1086/340952](https://doi.org/10.1086/340952)
- Postman, M., Coe, D., Benítez, N., et al. 2012, *ApJS*, 199, 25, doi: [10.1088/0067-0049/199/2/25](https://doi.org/10.1088/0067-0049/199/2/25)
- Puech, M. 2010, *MNRAS*, 406, 535, doi: [10.1111/j.1365-2966.2010.16689.x](https://doi.org/10.1111/j.1365-2966.2010.16689.x)
- Puskás, D., Tacchella, S., Simmonds, C., et al. 2025, *MNRAS*, 540, 2146, doi: [10.1093/mnras/staf813](https://doi.org/10.1093/mnras/staf813)
- Renaud, F., Agertz, O., & Romeo, A. B. 2024, *A&A*, 687, A91, doi: [10.1051/0004-6361/202449721](https://doi.org/10.1051/0004-6361/202449721)
- Ribeiro, B., Le Fèvre, O., Cassata, P., et al. 2017, *A&A*, 608, A16, doi: [10.1051/0004-6361/201630057](https://doi.org/10.1051/0004-6361/201630057)
- Rihtaršič, G., Bradač, M., Desprez, G., et al. 2025, *A&A*, 696, A15, doi: [10.1051/0004-6361/202451117](https://doi.org/10.1051/0004-6361/202451117)
- Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, *MNRAS*, 483, 4140, doi: [10.1093/mnras/sty3345](https://doi.org/10.1093/mnras/sty3345)

- Salmon, B., Coe, D., Bradley, L., et al. 2020, *ApJ*, 889, 189, doi: [10.3847/1538-4357/ab5a8b](https://doi.org/10.3847/1538-4357/ab5a8b)
- Sarrouh, G. T. E., Asada, Y., Martis, N. S., et al. 2025, arXiv e-prints, arXiv:2506.21685. <https://arxiv.org/abs/2506.21685>
- Sattari, Z., Mobasher, B., Chartab, N., et al. 2023, *ApJ*, 951, 147, doi: [10.3847/1538-4357/acd5d6](https://doi.org/10.3847/1538-4357/acd5d6)
- Schreiber, C., Glazebrook, K., Nanayakkara, T., et al. 2018, *A&A*, 618, A85, doi: [10.1051/0004-6361/201833070](https://doi.org/10.1051/0004-6361/201833070)
- Shibuya, T., Ouchi, M., Kubo, M., & Harikane, Y. 2016, *ApJ*, 821, 72, doi: [10.3847/0004-637X/821/2/72](https://doi.org/10.3847/0004-637X/821/2/72)
- Sok, V., Muzzin, A., Jablonka, P., et al. 2022, *ApJ*, 924, 7, doi: [10.3847/1538-4357/ac2f40](https://doi.org/10.3847/1538-4357/ac2f40)
- Sok, V., Muzzin, A., Jablonka, P., et al. 2025, *ApJ*, 979, 14, doi: [10.3847/1538-4357/ad96bb](https://doi.org/10.3847/1538-4357/ad96bb)
- Sorba, R., & Sawicki, M. 2015, *MNRAS*, 452, 235, doi: [10.1093/mnras/stv1235](https://doi.org/10.1093/mnras/stv1235)
- Sorba, R., & Sawicki, M. 2018, *MNRAS*, 476, 1532, doi: [10.1093/mnras/sty186](https://doi.org/10.1093/mnras/sty186)
- Soto, E., de Mello, D. F., Rafelski, M., et al. 2017, *ApJ*, 837, 6, doi: [10.3847/1538-4357/aa5da3](https://doi.org/10.3847/1538-4357/aa5da3)
- Starck, J.-L., Fadili, J., & Murtagh, F. 2007, *IEEE Transactions on Image Processing*, 16, 297, doi: [10.1109/TIP.2006.887733](https://doi.org/10.1109/TIP.2006.887733)
- Tacconi, L. J., Genzel, R., Neri, R., et al. 2010, *Nature*, 463, 781, doi: [10.1038/nature08773](https://doi.org/10.1038/nature08773)
- Tamburello, V., Rahmati, A., Mayer, L., et al. 2017, *MNRAS*, 468, 4792, doi: [10.1093/mnras/stx784](https://doi.org/10.1093/mnras/stx784)
- Tan, V. Y. Y., Muzzin, A., Marsan, Z. C., et al. 2022, *ApJ*, 933, 30, doi: [10.3847/1538-4357/ac7051](https://doi.org/10.3847/1538-4357/ac7051)
- Tan, V. Y. Y., Muzzin, A., Sarrouh, G. T. E., et al. 2024, arXiv e-prints, arXiv:2412.07829, doi: [10.48550/arXiv.2412.07829](https://doi.org/10.48550/arXiv.2412.07829)
- Toomre, A. 1964, *ApJ*, 139, 1217, doi: [10.1086/147861](https://doi.org/10.1086/147861)
- van Dokkelaar, F., Mayer, L., Capelo, P. R., et al. 2023, *MNRAS*, 522, 1726, doi: [10.1093/mnras/stad946](https://doi.org/10.1093/mnras/stad946)
- Vanzella, E., Messa, M., Adamo, A., et al. 2025, arXiv e-prints, arXiv:2507.18699, doi: [10.48550/arXiv.2507.18699](https://doi.org/10.48550/arXiv.2507.18699)
- Walter, F., Carilli, C., Neeleman, M., et al. 2020, *ApJ*, 902, 111, doi: [10.3847/1538-4357/abb82e](https://doi.org/10.3847/1538-4357/abb82e)
- Wang, T.-M., Magnelli, B., Schinnerer, E., et al. 2022, *A&A*, 660, A142, doi: [10.1051/0004-6361/202142299](https://doi.org/10.1051/0004-6361/202142299)
- Waterval, S., Cannarozzo, C., & Macciò, A. V. 2025, *MNRAS*, 537, 2726, doi: [10.1093/mnras/staf198](https://doi.org/10.1093/mnras/staf198)
- Willott, C. J., Doyon, R., Albert, L., et al. 2022, *PASP*, 134, 025002, doi: [10.1088/1538-3873/ac5158](https://doi.org/10.1088/1538-3873/ac5158)
- Wisnioski, E., Glazebrook, K., Blake, C., et al. 2011, *MNRAS*, 417, 2601, doi: [10.1111/j.1365-2966.2011.19429.x](https://doi.org/10.1111/j.1365-2966.2011.19429.x)
- Wisnioski, E., Förster Schreiber, N. M., Fossati, M., et al. 2019, *ApJ*, 886, 124, doi: [10.3847/1538-4357/ab4db8](https://doi.org/10.3847/1538-4357/ab4db8)
- Wuyts, S., Förster Schreiber, N. M., Genzel, R., et al. 2012, *ApJ*, 753, 114, doi: [10.1088/0004-637X/753/2/114](https://doi.org/10.1088/0004-637X/753/2/114)
- Wuyts, S., Förster Schreiber, N. M., Nelson, E. J., et al. 2013, *ApJ*, 779, 135, doi: [10.1088/0004-637X/779/2/135](https://doi.org/10.1088/0004-637X/779/2/135)
- Zanella, A., Le Floc'h, E., Harrison, C. M., et al. 2019, *MNRAS*, 489, 2792, doi: [10.1093/mnras/stz2099](https://doi.org/10.1093/mnras/stz2099)