# The Rise and Fall of Star-Forming Clumps: Formation and Evolution of the Building Blocks of Galaxies across Cosmic Time

Visal Sok

A Dissertation submitted to the Faculty of Graduate Studies in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

Department of Physics & Astronomy
York University
Toronto, Ontario

December 2025

# Abstract

It is well established that many high-redshift star-forming galaxies host massive clumpy regions of intense star formation. However, the underlying physical processes that drive clump formation remain unclear, mainly because there are few observations capable of directly linking clump formation to specific processes.

In this dissertation, we address clump formation and evolution through three complementary studies. First, we investigate whether clump formation is linked to an inflow of metal-poor gas. Such an accretion event is expected to dilute the metal content of galaxies. This chapter, therefore, aims to provide the first observational constraint by measuring the gas-phase metallicity of a small sample ($<$30) of clumpy and non-clumpy galaxies using GNIRS on the Gemini-North telescope. We find that clump galaxies show, on average, lower metallicities than the mass–metallicity relation, consistent with a picture where clump formation is occurring due to an inflow of gas.

The second chapter further investigates clump formation and evolution by using resolved imaging from the *James Webb Space Telescope*(JWST) of the CANUCS extragalactic fields. This work enables us to study clump populations and compare their mass distributions with theoretical predictions. The work shows that the slope of the aggregated clump mass function is consistent with clumps being formed *in-situ*. However, more/less massive galaxies will have flatter/steeper clump mass function slope. We also investigate the evolution of clump populations, finding that the clump mass function slope becomes flatter as clumps age. This result may reflect a combination of different evolutionary processes, such as destructive feedback mechanisms and clump–clump mergers.

Lastly, we combine spectroscopic and photometric data to investigate clump formation further. Specifically, this chapter builds on the first study to probe the gas-phase metallicity of $\sim$ 300 star-forming galaxies. A fraction of these galaxies also have JWST imaging, enabling us to connect clump properties to the processes driving clump formation. Specifically, we find that clumpy galaxies with lower gas metallicities tend to host younger, more actively star-forming clumps. Together, these results provide the first cohesive observational picture connecting gas accretion to clump formation.

Dedicated to mom and dad.

# Acknowledgements

I would like to express my deepest gratitude to my supervisor, Dr. Adam Muzzin, for his invaluable guidance, support, and patience throughout my M.Sc. and Ph.D. Without you reaching out to me when I first applied to York, I think I would also have continue a different (and probably less fun) path in graduate studies. Also, thank you for sending me to Australia – it worked out in the end, and now it's great Ph.D. lore!

To Dr. Michael De Robertis, Dr. Patrick Hall, and Dr. Matthew Johnson, while Q&A during during my research evaluations were always scary, thank you for your insightful perspectives. I am equally grateful to Dr. John Moores for serving as my internal committee member and to Dr. Angela Adamo for agreeing to act as my external examiner. Reading a hundred pages of mostly unrelated science to your field may not be the most thrilling task you have, but I hope that you find at least one or two interesting things from my work.

Additionally, I would like to extend my gratitude to all my collaborators for their invaluable insights and encouragement. And to those collaborators in which we have survived through the thick and thin of proposal seasons. To Dr. Gillian Wilson, thank you for give me (and other York students) the opportunity to operate a multi-million-dollar telescope in Hawai'i. To Dr. Ben Forrest, for your help and supports in the numerous nights of operating the MOSFIRE spectrograph and for just making sure we didn't burn the telescope to the ground. Both of your support have directly led to the data presented in Chapter 4.

Speaking of which, some of the data presented in this work is only possible due to observations made from both the Gemini-North and Keck I telescopes, which are located within the Maunakea Science Reserve and adjacent to the summit of Maunakea. I would like to express my gratitude for the privilege of observing the Universe from a place unique in both its astronomical quality and its cultural significance.

On a personal note, I am deeply thankful to my family for their unwavering love and support. To my friends, for keeping me grounded and reminding me of the world beyond research. And to coffee, for being a loyal companion and supplying the energy needed to tackle the endless stream of my problems.

Finally, for those who have stumbled upon this work, I hope to take you on a journey from the very first images of the *Hubble Space Telescope* to the breathtaking views from the *James Webb Space Telescope*. Along the way, I hope you share in the excitement I have found in unravelling the mysteries of the Universe. The process has not been without frustration (and a few hairs lost), but all in an effort to contribute, however modestly, to this exciting field and to share that journey with you.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

> "To see the world, things dangerous to come to, to see behind walls, draw closer, to find each other, and to feel. That is the purpose of life."
> — Sean O'Connell, The Secret Life of Walter Mitty

The Universe, in all its grandeur and splendour, has captivated the minds of so many. Given its immense scale, it is not surprising that astronomy encompasses such a wide range of research topics. Of particular interest (and totally not biased from my point of view, by the way) is the evolution of galaxies over cosmic time. The quest to understand galaxy evolution and star formation is intimately tied to understanding the history of our own Milky Way and the formation of the Sun, which has supported life on Earth for billions of years.

Studying galaxy evolution is challenging because we observe galaxies at a single moment in their history. Yet, due to cosmology and the finite speed of light, observing distant galaxies allows us to see across the history of the Universe. Deep field surveys that probe galaxies at different redshifts, therefore, enable us to trace different populations of galaxies, from the young galaxies in the early Universe to the grand-design spirals of the nearby Universe. By connecting galaxy populations across epochs using techniques such as abundance matching, astronomers can reconstruct the evolutionary paths of galaxies. Intriguingly, as we look further back into the high-redshift Universe, we find that the Universe had a clumpier past, in which galaxies often host massive, localized sites of star formation. These star-forming sites can reach sizes of several hundred parsecs and have masses of approximately $10^6 - 10^8\,M_\odot$ (e.g., Dessauges-Zavadsky and Adamo 2018; Claeyssens et al. 2025), making them significantly larger and more massive than local analogues, which are typically parsec-scale systems with lower masses. High-redshift star-forming regions are therefore commonly dubbed as star-forming "clumps". Their host galaxies are subsequently referred to as "clumpy" galaxies.

To set the stage for my work about the very nature of star-forming clumps, I will first provide a general overview of galaxy formation and evolution in Section 1.1. This discussion naturally leads to the early discoveries of galaxies with irregular morphologies and prominent star-forming regions. The central question of "how did these structures form and evolve" remains at the heart of ongoing research. Section 1.2 summarizes our current understanding of clump formation and evolution, while Section 1.3 highlights open questions that have motivated much of my own research over the past seven years. Section 1.4 introduces the data and tools I used to address these questions. Finally, Section 1.5 provides an overview of the body of my dissertation.

## 1.1 Galaxy Assembly

An important goal in extragalactic astronomy is to understand how galaxies assemble their stellar mass over cosmic time. Progress in this area has been driven by multiwavelength surveys with facilities such as the *Hubble Space Telescope* and, more recently, the *James Webb Space Telescope*, which allow us to trace the resolved galaxy populations across redshift and infer their evolutionary pathways.

Within the current paradigm of galaxy assembly, galaxies grow hierarchically by clustering within dark-matter halos, which are overdense regions seeded by primordial density fluctuations after the Big Bang. Low-mass halos form first, then merge hierarchically to form larger structures. On top of this hierarchical framework, a variety of physical processes act on different scales to shape the morphologies, colours, and star formation properties of galaxies. These processes include the accretion and cooling of gas, its collapse into dense sites of star formation, and feedback from stellar winds, radiation, and supernovae that regulate further star formation and drive outflows.

Diving into the detailed processes that govern mass assembly within galaxies would require several review papers. However, the broader picture is imprinted in the cosmic star formation rate density, which describes the rate of star formation per unit volume in the Universe. Over the past few decades, a coherent picture of the cosmic star formation history has emerged, showing that the Universe reached a peak in global star formation activity at $z \sim 2$. (see Figure 1.1). Approximately 25% of the total stellar mass density (SMD) that we observe today was formed at $z > 2$, while a comparable fraction is formed at $z \lesssim 0.7$ (Madau and Dickinson, 2014). This highlights that the bulk of galaxy mass assembly occurred during the critical epoch of $1 < z < 3$, or around cosmic noon.

The key to painting a clearer picture of galaxy assembly is understanding the underlying physical processes that drive star formation. With this goal in mind, my research has focused on the dominant modes of star formation during this peak epoch of mass assembly. Since the bulk of stellar mass was assembled at cosmic noon, it is natural to question what galaxies look like during this period. In the next Section, we will focus on morphological studies of galaxies, which further provide direct insight into how stars formed within them.

### 1.1.1 A Clumpy Mass Assembly

Early efforts to classify galaxies began with the Hubble Sequence, which broadly divided galaxies into early- and late-type galaxies. Although the naming convention suggests an evolutionary pathway, this is solely a morphological classification nomenclature. Late-type galaxies refer to spiral galaxies, while early-type galaxies are elliptical galaxies. The morphology of spirals and ellipticals is relatively well-defined in the sense that we can clearly distinguish when a galaxy is spiral or elliptical.

With the advent of the *Hubble Space Telescope*, deep imaging revealed that galaxies at $z > 1$ are not only smaller in masses, but are morphologically clumpy (e.g., see Fig. 1.2; Cowie et al. 1995;

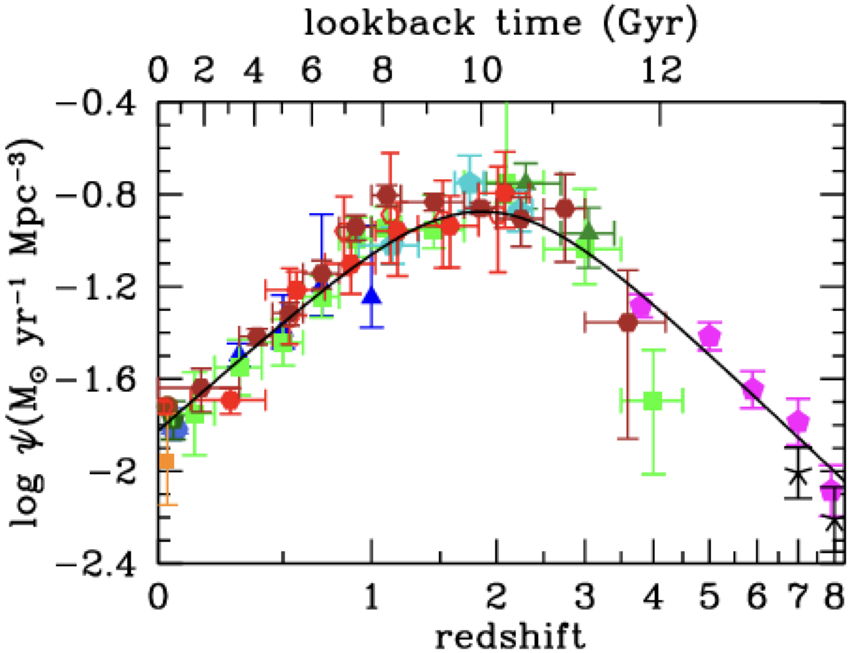


Figure 1.1: Taken from Madau and Dickinson (2014). The history of cosmic star formation in the Universe using data from far-ultraviolet (top right panel), infrared (bottom right panel), and the combined far-ultraviolet and infrared (left panel) rest-frame measurements. The Universe was more actively forming stars at $z \sim 2$ as compared to present-day.

Abraham et al. 1996) and do not fit the classic Hubble sequence. In this context, “clumpy" generally refers to high-redshift galaxies hosting bright star-forming regions. These regions typically span a few hundred parsec to sub-kiloparsec scale, and are therefore substantially larger compared to local star-forming regions such as giant molecular clouds.

A more formal definition for clumpy galaxies is based on the fractional contribution of these regions to the galaxy’s rest-frame UV emission. Galaxies are classified as clumpy if their clumps contribute more than 8% of the total rest-frame UV luminosity. This threshold was proposed by Guo et al. (2015) by comparing the UV luminosity functions of clumps in high-redshift galaxies to those of local star-forming regions, which do not extend beyond the 8% level. However, note that this criterion is not universal and varies across studies. With data from the HST, we now understand that clumpy galaxies are abundant at $z > 1$, with $\gtrsim$40% of star-forming galaxies hosting clumps (Wuyts et al., 2012; Murata et al., 2014; Guo et al., 2015; Sok et al., 2022). We also know that the fraction of clumpy galaxies is lower at $z > 3$ (e.g., Shibuya et al., 2016; Sattari et al., 2023).

The launch of the *James Webb Space Telescope*, with its significantly improved sensitivity as compared to the *Hubble Space Telescope*, combined with the magnification power of gravitational lensing, has enabled us to resolve galaxies at $z > 8$ (e.g., Mowla et al. 2024; Adamo et al. 2024; Vanzella et al. 2025). Strikingly, even these early galaxies often contain clumpy stellar structures. In some cases, individual clumps account for at least 30% of the stellar mass of their host galaxies, suggesting that clumpy star formation played a central role in the earliest phases of galaxy assembly.

When we look at the number of star-forming galaxies that host clumps over cosmic time, we find

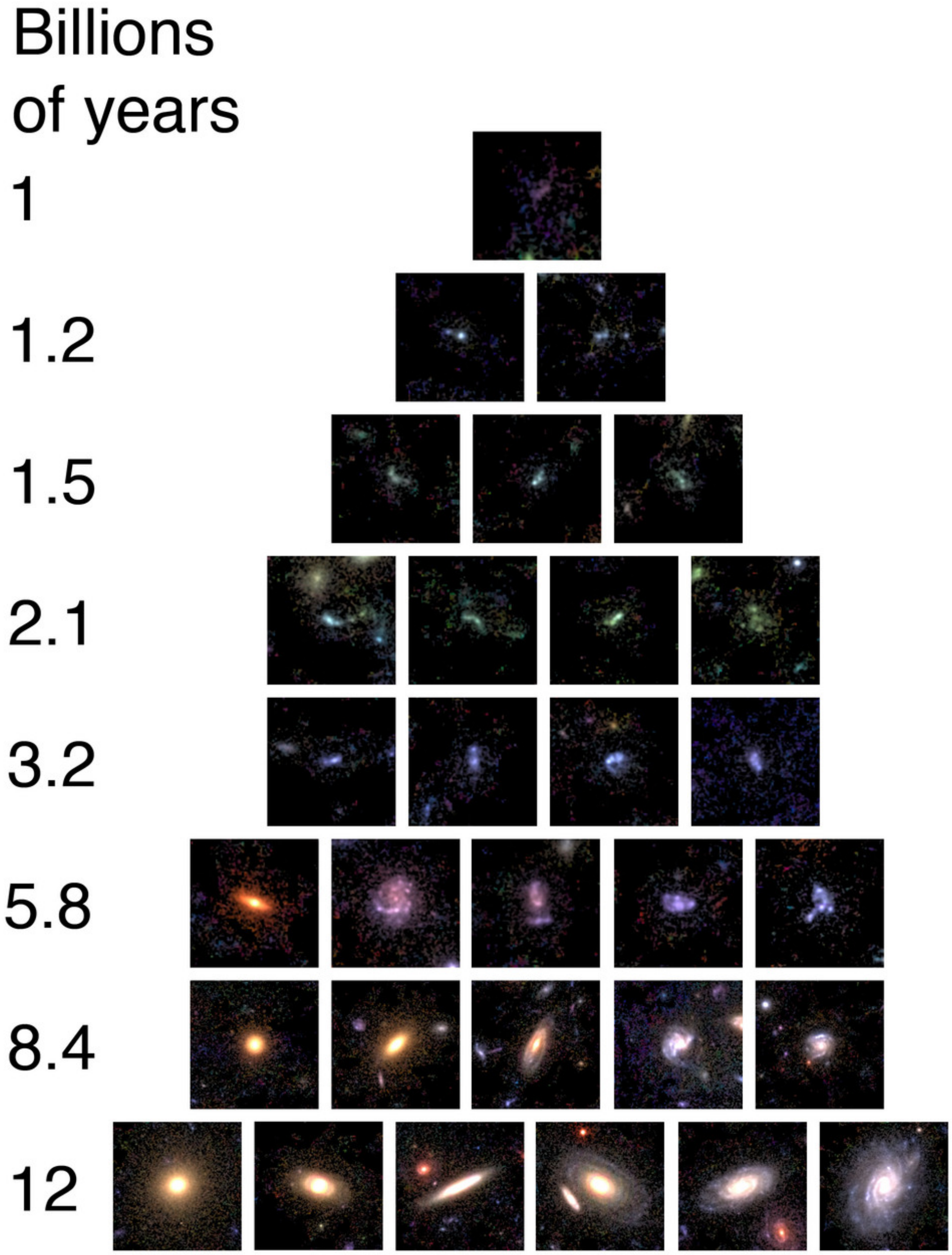


Figure 1.2: The morphology of galaxies in relation to the age of the Universe, taken from the COSMOS webpage (`https://cosmos.astro.caltech.edu/`). Present-day galaxies (i.e., when the Universe is closer to 14 billion years old) are generally well-described as either spiral or elliptical. However, at earlier times, galaxies are more irregular, and often host star-forming clumps on kiloparsec-scale. In addition to the morphological evolution, galaxies are becoming more massive toward the present day.

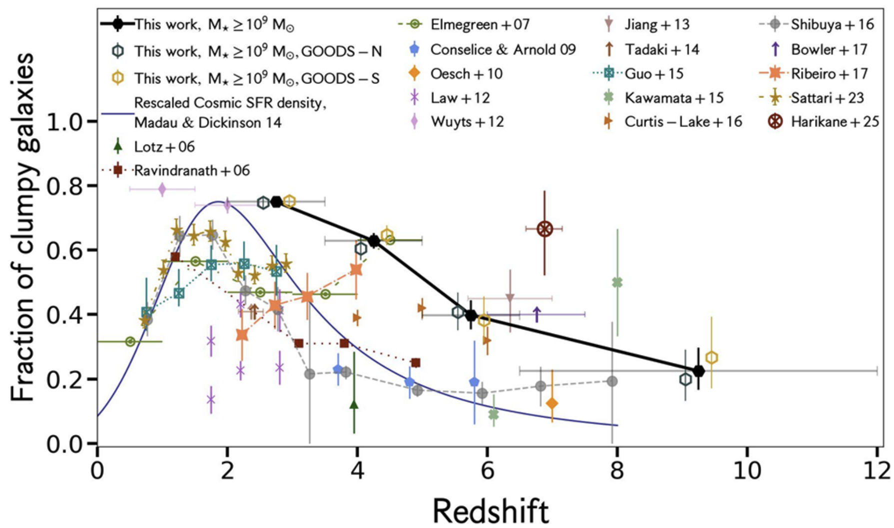


Figure 1.3: Taken from de la Vega et al. (2025). A compilation of the measured fraction of clumpy galaxies. While these measurements are not corrected for incompleteness, namely due to the difficulty in detecting clumps at higher redshifts, the figure still suggests that the fraction of clumpy galaxies peak at $z \sim 2$, and follows a similar redshift evolution to that of the cosmic star formation rate density in Figure 1.1.

that it broadly traces cosmic star formation rate density (e.g., see Figure 1.3, Shibuya et al., 2016; de la Vega et al., 2025). Together, these observations strongly suggest that clumpy regions of star formation are closely tied to the assembly of stellar mass in galaxies. However, observational studies are still inconclusive about the connections between star-forming structures in nearby galaxies and those in clumpy disks at higher $z$. In particular, it is not clear whether the processes driving star-forming regions are universal across cosmic time, or whether clumpy star formation found at high redshift is uniquely the result of the extreme physical conditions of gas-rich galaxies typical at earlier times. If we aim to understand galaxy formation in detail, it is therefore essential to uncover the physical mechanisms that drive clump formation and evolution.

## 1.2 A Crash Course in Clump Formation and Evolution

### 1.2.1 Clump Formation

Despite the ubiquity of star-forming clumps at high redshift, the origin and evolution of clumps is still an on-going topic of research. The formation of clumps at cosmic noon are generally linked to two different pathways. The first is linked to the accretion of gas from cosmological streams, and the second is linked to mergers and accreting satellites.

#### Gas Accretion

The leading framework for clump formation suggests that clumps are formed within gas-enriched disk that are undergoing violent disk instability (VDI, e.g., Agertz et al., 2009; Dekel et al., 2009a,b). This is motivated by observations of very gas-rich galaxies at higher redshifts (e.g., Daddi et al., 2010; Tacconi et al., 2010; Wang et al., 2022). The well-ordered rotations of high-$z$ disks (e.g., Förster

Schreiber et al. 2018b; Wisnioski et al. 2019) also favour VDI-driven clumps over clump formation caused by ongoing merger (which is discussed in the next section). The nature of these gas-rich disks are suggested to be fuelled by gas accretion, through collimated, cosmological filaments of cold and dense gas streams (e.g., Kereš et al., 2005, 2009; van de Voort and Schaye, 2012; Waterval et al., 2025).

In the VDI-driven clump formation scenario, the instability of the disk is often described by the Toomre $Q$ parameter (Toomre, 1964),

$$Q = \frac{\sigma\kappa}{\pi G \Sigma}. \tag{1.1}$$

Here, $\sigma$ and $\Sigma$ denote the gas velocity dispersion and gas density. $\kappa$ is the epicyclic frequency and is related to the angular velocity $\Omega$ of the disk. The disk is considered marginally unstable, with $Q \sim 1$. In this simple model of disk instability, we find that increasing the gas density parameter leads to $Q$ becoming less than 1 (i.e., an unstable disk). As clumps form, their interactions within the disk can drive turbulence again, pushing $Q$ back to unity and stabilizing the disk (Dekel et al., 2009b).

The fragmentation of clumps in gas-rich disks is also well-studied in simulations of galaxies in both isolation and cosmological context (e.g., Dekel et al., 2009a; Ceverino et al., 2016; Fensch and Bournaud, 2021; Renaud et al., 2024). In general, metal-poor, accreting gas is responsible for the gas-rich nature of higher-redshift galaxies (van de Voort and Schaye, 2012). Ceverino et al. (2016) further investigated gas accretion events in the VELA simulations and found that star-forming clumps generally have a lower gas-phase metallicity as compared to the disks. These star-forming clumps are further found to have a lower metallicity offset of 0.3 dex as compared to the disk, consistent with observations (e.g., Sánchez Almeida et al., 2015; Estrada-Carpenter et al., 2025). The timescale for gas dilution is a few disk dynamical times ($\sim$ 50 Myr) (Ceverino et al., 2016) as stellar evolution enriches and disperses the gas. Simulations that look at the effects of varying gas fraction on star-forming clumps also suggest that there is a relationship between gas conditions at the time of formation and the resulting clump populations. Massive clumps are generally found in simulated galaxies with higher gas fractions ($\sim 40 - 50\%$), resulting in a flatter clump stellar mass function (Fensch and Bournaud, 2021; Renaud et al., 2024). Based on previous studies such as Ceverino et al. (2016); Estrada-Carpenter et al. (2025), measuring the gas-phase metallicity of star-forming clumps will provide additional observational constraints for such formation.

#### Merger Events

Alternatively, clumps can also arise either from gas fragmentation triggered during galaxy–galaxy interactions (e.g. Bournaud et al., 2011; Renaud et al., 2014) or from the direct accretion of star-forming satellite galaxies (e.g. Zanella et al., 2019). The efficiency of clump formation during mergers as observed in simulations depends on the initial conditions, including the implementation of feedback and turbulence, and the gas fraction of colliding galaxies (e.g. Di Matteo et al., 2008; Fensch et al., 2017).

Observationally, Puech (2010) studied a sample of 11 clumpy galaxies at $z \sim 0.6$ and suggested that the complex kinematics of these clumpy galaxies can only be attributed to mergers. Other observational supports for merger-driven clump formation include the comparison between the red-shift evolution of the fraction of clumpy galaxies to the fraction of mergers. It is argued that both

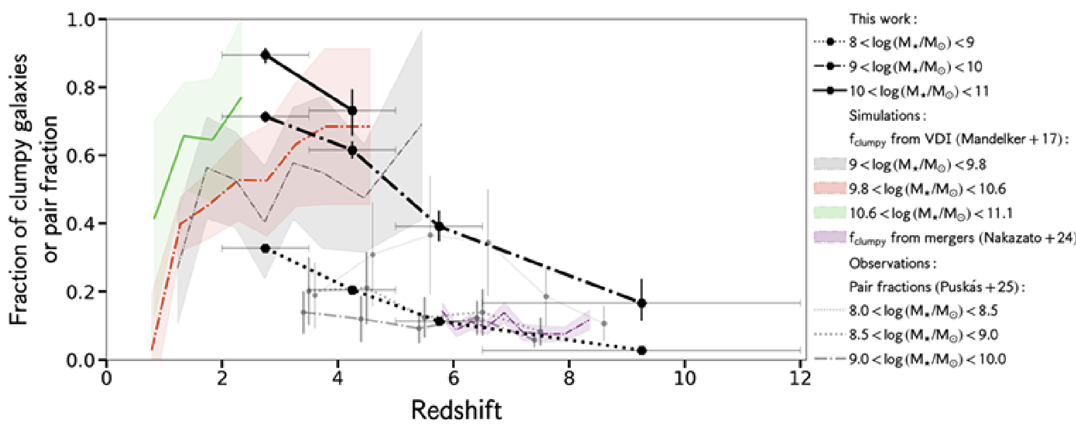


Figure 1.4: Taken from de la Vega et al. (2025). A compilation of the measured fraction of clumpy galaxies over the past decade as compared to the pair fractions from Puskás et al. (2025).

fractions should have a similar redshift evolution if clumps mainly formed due to mergers. This was originally investigated in Guo et al. (2015), where it was suggested that minor mergers likely play a role in clump formation at $z < 2$ for less massive galaxies, while major mergers becomes a more prominent mechanism at higher redshifts. Similarly, a recent study by de la Vega et al. (2025) measured the fraction of clumpy galaxies out to $z \sim 10$ and compared their measurements to studies that measured the pair fractions of galaxies (e.g., Duan et al., 2025; Dalmasso et al., 2024; Puskás et al., 2025) (see Figure 1.4). They suggested that major mergers are the dominant mechanism for clumpy galaxy formation for lower mass galaxies ($M_* < 10^9$ solar masses) for $z > 5$. With currently available data, the observational challenges of separating mergers from non-mergers (and other complex morphologies) leads to difficulty determining the relative abundances of merger-formed clumps. Resolved spectroscopic observations are generally needed to identify the kinematic signatures of mergers.

### 1.2.2 Clump Evolution

The role of star-forming clumps in galaxy evolution remains uncertain, as their fate is still poorly constrained. Their fate depends on how long they survive within the host galaxies. Long-lived clumps could evolve within their host, while clumps that are shorter-lived are suggested to dissipate within a disk dynamical time. In this Section, I discuss the various supports and implications for these two scenarios for clump evolution.

Simulations suggest that long-lived clumps can migrate toward the centres of their hosts and contribute to the growth of bulges through dynamical friction and interactions with other clumps (e.g., Noguchi 1999; Inoue and Saitoh 2012; Mandelker et al. 2014, 2017). During this migration, the stellar populations of the clump evolve as the clumps experience gas accretion, feedback-driven outflows, and disruption by tidal shear in the surrounding gas rich disk (Bournaud et al., 2007, 2014). The longevity of clumps is arguably tied to feedback in galaxies. In particular, stronger feedback is suggested to result in shorter-lived clumps, as the strong gas outflow halts star formation. Simulations that have implemented enhanced feedback observed to produce very short-lived clumps, with timescales of approximately the disc dynamical time of $t_d = r_d/v_d \sim 10 - 50$ Myr at $z \sim 2$

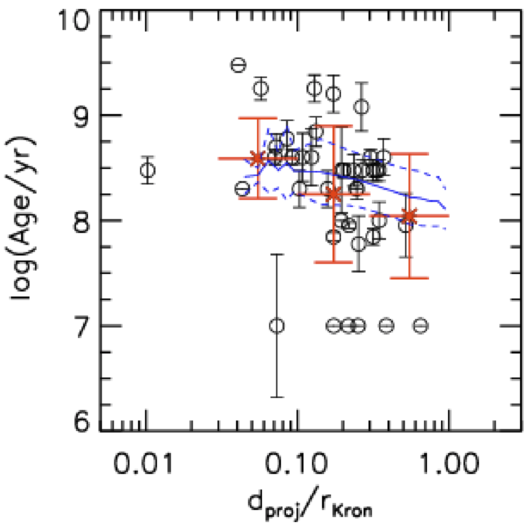


Figure 1.5: Taken from Guo et al. (2012). The red stars show the mean clump age in each distance bin, while the solid line shows the mean age of the interclump regions for all their galaxies. Guo et al. (2012) found that clumps are generally younger than their surrounding disks and exhibit a stronger radial age gradient too, consistent with expectations from clump migration.

(e.g., Hopkins et al., 2012; Oklopčić et al., 2017). Similarly, further investigation into the effects of feedback found that massive clumps (e.g., $\log(M_*/M_\odot) > 8.5$) are generally less affected by feedback (e.g., Bournaud et al., 2014; Moody et al., 2014; Mandelker et al., 2017; Dekel et al., 2022). These clumps can survive longer, lasting $\sim 300$ Myr or more. The survivability of clumps is likely further complicated by clump and disk kinematics (e.g., Dekel et al. 2023). Simulations of isolated galaxies also emphasized the dependence of clump survivability on the disk gas fraction (Fensch and Bournaud, 2021; Renaud et al., 2024), with short-lived clumps found in lower-gas-fraction host galaxies. We anticipate clump populations to be a composition of short- and longer-lived clumps[1].

Observational studies that seek to constrain the evolution of clumps primarily rely on estimating the clump ages and their galactocentric distances. In particular, a radially decreasing age gradient is expected, as clumps are thought to form in the outskirts and migrate inward as they evolve Guo et al. (2012, 2015); Zanella et al. (2019) (see Figure 1.5 as an example). While a correlation has been predicted, it should be noted that individual estimates of the clump masses, SFRs, and ages can generally have significant uncertainties (Huertas-Company et al., 2020; Ginzburg et al., 2022). It is therefore essential to perform these analyses on a large sample of clumps. Correlations between clump physical properties and galactocentric distance can be affected by the background-subtraction method. Clumps are situated on a gas disk, and the underlying light from the disk will affect the light of the clumps. In general, galaxies are redder and older, with more dust near the center, which suggests that improper background subtraction can lead to artificial aging of clumps near the center. Moreover, interpreting clump ages depends critically on their star formation histories. It remains unclear whether clumps form stars continuously over extended periods or through shorter, intense bursts of star formation. Distinguishing between these scenarios is crucial for determining whether clumps are long-lived structures or transient features.

[1]A recurring theme here is that not everything is as black and white as we make it out to be.

Overall, a picture is emerging that clump formation is tied to the availability of gas, either supplied through cosmic gas accretion or through galaxy interactions. Clumps can be disrupted or survive long enough to shape the evolution of their host galaxies, depending on their mass and environment. Untangling these evolutionary pathways requires estimates of properties including stellar and gas masses and ages for larger samples of clumps across cosmic time, which motivates some of the work presented in this dissertation.

## 1.3 Aims and Approach

### 1.3.1 What Still Bothers Us at Night?

The previous sections outline the importance of star-forming clumps in galaxy assembly and broadly discuss the diverse formation channels and evolutionary scenarios proposed for clumps. We broadly know that clumpy star formation is a common feature in most cosmic noon star-forming galaxies (e.g., Shibuya et al. 2016; de la Vega et al. 2025). Recent observations with JWST of clumpy stellar structures in galaxies at $z > 5$ (e.g., Mowla et al., 2024; Fujimoto et al., 2025) further suggest that galaxy formation in the early universe may also have been fundamentally clumpy in nature.

However, the underlying processes that determine clump formation across cosmic time remain elusive. In particular, it remains unclear which processes are dominant and whether they vary with redshift. While previous studies have laid important groundwork for understanding clump formation and evolution, but key aspects of how clumps originate and how they relate to the overall assembly of galaxies remain elusive. Part of the problem lies in the lack of observational studies linking the resolved properties of clumps to the global properties of the host galaxies. Establishing these links will allow for direct testing of clump formation in the environments and will help determine how those environments influence clump evolution and survivability. Additionally, there are limited observations before JWST that can enable astronomers to probe different clump populations across redshift.

### 1.3.2 How Can We Chase the Monster Away?

This dissertation directly addresses these challenges by leveraging the latest multi-wavelength data, from imaging survey of the JWST CANUCS fields (Willott et al., 2022) to near-infrared spectroscopic data, such as the LEGA-C (van der Wel et al., 2021) and MOSDEF (Kriek et al., 2015) survey. This endeavour will ultimately tie different clump populations to the environment that they form in. If processes such as metal-poor gas accretion are responsible for global instabilities of the gas disk, then a missing piece of the puzzle for understanding clump formation is the kinematics and physical conditions of the gas disk. For example, cosmic gas accretion is expected to be near-pristine (van de Voort and Schaye, 2012), which can then dilute the galaxies' gas-phase metallicity (e.g., Ceverino et al. 2016; Estrada-Carpenter et al. 2025). Performing a large systematic study of clumpy and nonclumpy galaxies by probing their gas-phase metallicity will offer one of the first observational constraints that directly link clump formation to a physical process such as gas inflow.

Other observational constraints come from cataloguing the diverse properties of clumps, including their stellar masses, ages, and SFRs. In particular, the global gas conditions leave imprints on the distribution of clump mass that are formed (e.g., Fensch and Bournaud 2021; Renaud et al. 2024).

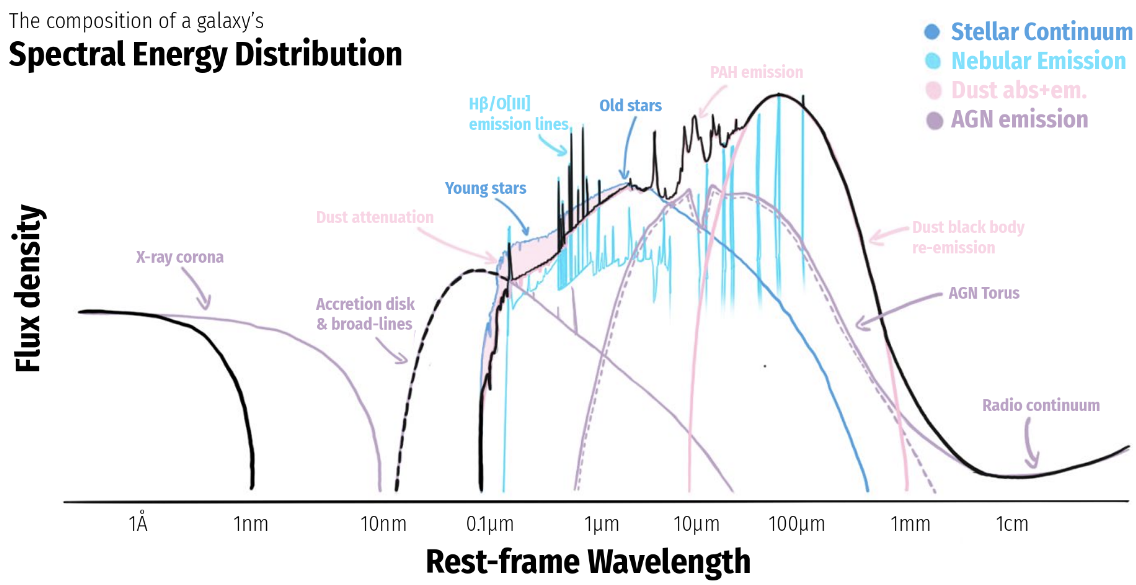


Figure 1.6: Taken from Iyer et al. (2025). A schematic showing the spectral energy distribution of a galaxy, and the different sources of emission at different wavelength.

By probing the clump mass distribution, we gain insight into the environment that sets the stage for clump formation and link the rise and fall of clumps to an evolving cosmic gas density. Before *James Webb Space Telescope*, the first constraint of the clump stellar mass function came from the work of Dessauges-Zavadsky and Adamo (2018), who studied clumps from $1 < z < 3$. With the number of increased surveyed extragalactic fields by *James Webb Space Telescope*, we are now at a unique time to map the clump stellar mass function as a function of redshift.

## 1.4 Field Guide to Being a Clump Expert

The majority of my work involves using various tools and techniques to infer clump properties. To provide you with the necessary knowledge to understand how these properties are obtained, this Section provides a general overview of the different tools. In doing so, I hope to also provide you with the tools to solve the problems of the Universe.

A recurring theme in the later chapters of this dissertation is the extraction of stellar masses from clumps and disks, as well as the inference of gas-phase metallicity from strong emission lines observed in galaxy spectra. The following sections, therefore, describe how we can infer these estimates.

### 1.4.1 How to Infer Mass From the Spectral Energy Distribution

Galaxies consist of different components, including gas, dust, stars and some dark matter that cannot be observed. The spectral energy distribution (SED) of a galaxy is therefore composed of the light emitted from stars, ionized gas, dust and, if presented, active galactic nuclei (e.g., see Figure 1.6). Information about the current state and past star formation activity is also encoded in the SED. In the simplest terms, inferring a galaxy's physical properties means modelling its SED by summing the emissions from its components.

The core part of a galaxy's SED consists of the combined light from all its stars, also known as the stellar continuum. The detailed shape of the stellar continuum depends primarily on the ages, metallicities and relative abundances of the stellar population of the galaxy. A crucial step in creating a composite stellar continuum is the Initial Mass Function (IMF), which describes the relative abundance of stars of different masses at birth. Common choices for the IMF include the Salpeter (Salpeter, 1955), Chabrier (Chabrier, 2003), or Kroupa (Kroupa, 2001) IMFs.

Early versions of SED modelling also often assumed that all of the star formation occurred in a single burst. This is referred to as a Simple Stellar Population (SSP), where all the stars of a given IMF are assumed to form at a single lookback time and with the same metallicity. This motivates the need for more complicated models involving "composite stellar populations" arise later with more complex star formation histories (SFHs). The SFH of a galaxy can be visualized as the number of stars formed over time. Given an SFH, Stellar Population Synthesis (SPS) codes such as FSPS (Conroy and Gunn, 2010) enable the creation of different composite stellar continua by weighting and summing different SSPs according to the galaxy's star formation history.

Assumptions of the star formation history can lead to systematically different derived ages and masses, particularly for star-forming clumps. For example, as noted in Messa et al. (2022), assuming a single stellar population or a constant star formation history generally gives lower stellar masses than models that allow star formation to extend over longer timescales. This therefore motivates the uses for fitting for non-parametric SFH, which provide the flexibility needed to capture the diversity of SFHs expected in galaxies and their clumps.

It should be clear by now that inferring the mass distribution and star formation history of a galaxy is (and may I add, *very complicated*) inverse problem given the observed SED. However, the explanation above is a gross oversimplification of SED modelling. Other factors, including chemical enrichment of the interstellar medium and dust attenuation, also affect the shape of the SED and must therefore be considered. For example, dust can attenuate the rest-frame ultraviolet light, with its re-emission in the longer near-infrared. Nebular lines, arising from star-formation regions, also affect the photometry by boosting certain bands based on their wavelengths (e.g., Withers et al. 2023).

There are numerous SED fitting codes available. In the chapters ahead, I adopt Dense Basis (Iyer and Gawiser, 2017; Iyer et al., 2019) to infer physical properties of galaxies. Dense Basis employs the core concepts outlined above by constructing an atlas consisted of SEDs corresponding to a wide set of SFHs. However, instead of assuming a parametric SFH (e.g., an exponentially declining SFH), Dense Basis infers the SFH nonparametrically from the data. In this framework, the SFH is described by the lookback times at which a galaxy (or clump) assembled fixed quantiles of its total stellar mass, with Gaussian Process interpolation used to generate a smooth SFH between these points. However, the concept of a single "age" is no longer uniquely defined unlike using a single burst SFH model, as star formation proceeds continuously over time. This dissertation therefore adopt a mass-based definition of age, which I will later describe in detail in Chapters 3 and 4. Dense Basis uses the observed photometry of each galaxy to compare to the atlas of precomputed SEDs within a Bayesian framework, providing posterior distributions for stellar population parameters and the reconstructed SFH.

### 1.4.2 How to Infer Gas-Phase Metallicity from Strong Line Diagnostics

A key part of understanding galaxy formation lies in tracing the chemical history of galaxies, which is probed through measurements of their gas-phase metallicity. In this work, the gas-phase metallicity (or sometimes just metallicity) refers to the oxygen abundance. Typically, this is inferred from nebular emission lines spanning wavelengths from the ultraviolet to the infrared.

One method to determine the gas-phase metallicity is the direct method. As noted by Kewley et al. (2019), this is a misleading term because several assumptions are still made to derive metallicity here. In the direct method, the electron density ($n_e$) and temperature ($T_e$) are measured using density-sensitive doublets and temperature-sensitive auroral-to-nebular line ratios such as [OII]$\lambda$4363/$\lambda$5007. These quantities determine the emissivities of collisionally excited metal lines, which are proportional to the abundances of the corresponding ionic species. Measuring the fluxes of these ionic lines relative to hydrogen recombination lines therefore allows the oxygen-to-hydrogen abundance (O/H), and hence the gas-phase metallicity, to be inferred.

However, auroral lines are typically faint and are especially difficult to detect in galaxies with high metallicities. An alternative method is the use of strong line calibrations. The strong line methods rely on the flux ratios of strong lines, which do not directly measure metallicity but are metallicity-sensitive. Therefore, strong line ratios can be calibrated to estimate gas-phase metallicity in three ways: (1) empirically, using metallicities derived from the direct method, (2) theoretically, using photoionization models, or (3) semi-empirically, by combining direct-method metallicities with model-based calibrations (e.g., see Kewley et al. 2019 for reviews). Strong line methods are more widely used because the relevant lines are brighter and more readily observed.

In the upcoming chapters, I employ several robust strong-line metallicity diagnostics to infer the gas-phase metallicity of star-forming galaxies near cosmic noon. In particular, these diagnostics are the $R23$ and the $N2$ indices. The $R23$ index is defined as,

$$R23 = \frac{[\mathrm{O\,II}]\lambda, \lambda 3727, 9 + [\mathrm{O\,III}]\lambda, \lambda 4959, 5007}{H\beta}, \tag{1.2}$$

First proposed in Pagel et al. (1979), the $R23$ line index is sensitive both to metallicity and to the ionization state of the gas, or "ionization parameter". However, its dependency on the ionization parameter can be corrected by including the $O32$ index (e.g., Kobulnicky and Kewley 2004). Another commonly used line diagnostic is the $N2$ line index, defined as,

$$N2 = \frac{[\mathrm{N\,II}]\lambda 6583}{H\alpha}. \tag{1.3}$$

Here, $N2$ is empirically calibrated to the gas-phase metallicity (Pettini and Pagel, 2004). Although it is affected by the ionization parameter, inferring gas-phase metallicity from a large sample still reveals underlying trends. The $N2$ line diagnostic is also very practical for high-redshift observations, as it is less sensitive to dust attenuation and instrumental effects. For this dissertation, the emission lines for the two strong line calibrations are either obtained from existing spectroscopic surveys or from new spectroscopic observations.

It should be further noted that these strong-line diagnostics generally have an intrinsic scatter, where a specific value of $R23$ or $N2$ correspond to a metallicity with scatter of approximately 0.1-0.2 dex. In the later chapters where I will be comparing the metallicity of clumpy and nonclumpy galaxies, such larger scatter can affect the robustness of the metallicity differences. However, this scatter can be mitigated and the robustness thus improves by using a larger sample of galaxies similar to that presented in Chapter 4.

## 1.5 Dissertation Overview

With the stage now set, we turn to the real science. The remainder of this dissertation focuses on my work over the course of my studies and the resulting insights into the formation and evolution of star-forming clumps, their rise and fall, and the imprints they leave on galaxy assembly. Chapter 2 presents one of the first pilot studies examining the gas-phase metallicities of a population of star-forming galaxies. These metallicities are inferred from the ratio of [N II]$\lambda$6583 and $H\alpha$, measured using near-infrared spectroscopic observations from GNIRS on Gemini-North. The results reveal a metallicity difference between clumpy and nonclumpy galaxies, with clumpy galaxies having lower metallicity than nonclumpy galaxies. The discussion investigates the possible physical mechanisms responsible for this metallicity dilution. Chapter 3 shifts gears to explore the diverse populations of clumps, focusing on their stellar masses and ages. The chapter characterizes clump mass distributions across cosmic time. It examines how the slope of the clump stellar mass function relates to clump ages and to the global properties of the host galaxies. The discussion highlights how different clump formation mechanisms can be inferred from the slope and how the environment may influence clump evolution. Finally, Chapter 4 expands on the previous two chapters by connecting resolved properties of clumps to the global gas conditions of the hosts. In particular, the study provides gas-phase metallicity measurements for a larger sample of star-forming galaxies. The study explores how observed properties of clumps relate to the metal content of their host galaxies, specifically whether galaxies are metal-enriched or diluted. The discussion then considers how clump properties relate to larger-scale processes, such as gas accretion. Finally, I summarize the findings from these studies in Chapter 5, where I discuss the exciting new insights this work has yielded. I conclude by discussing the open questions that remain and the exciting directions that can be pursued with the next generation of telescopes. Without further ado, I present the works that form the core of my Ph.D. research.

# Chapter 2

# Evidence for metal-poor gas accretion in clumpy star-forming galaxies

"Well, this little maneuver's gonna cost us 51 years!"
— Cooper, Interstellar

*A version of the following chapter was originally published in the January 2025 issue of the Astrophysical Journal, Volume 979, Issue 1 as Sok et al. (2025a). The co-authors were Adam Muzzin, Pascale Jablonka, Vivian Yun Yan Tan, Z. Cemile Marsan, Danilo Marchesini, Gillian Wilson and Leo Y. Alcorn. In this chapter, we explore the gas-phase metallicity inferred from [NII]$\lambda$6584/H$\alpha$ for a sample of clumpy and non-clumpy galaxies. In particular, we test whether clumpy galaxies have lower gas-phase metallicity compared to nonclumpy galaxies. If so, this would suggest that clumpy galaxies are experiencing gas inflow. This work is one of the first observational studies that provide insights and constraints on the formation of clumps in star-forming galaxies.*

## Abstract

Despite the ubiquity of clumpy star-forming galaxies at high-redshift, the origin of clumps are still largely unconstrained due to the limited observations that can validate the mechanisms for clump formation. We postulate that if clumps form due to the accretion of metal-poor gas that leads to violent disk instability, clumpy galaxies should have lower gas-phase metallicities compared to non-clumpy galaxies. In this work, we obtain the near-infrared spectrum for 42 clumpy and non-clumpy star-forming galaxies of similar masses, SFRs, and colors at $z \approx 0.7$ using the Gemini Near-Infrared Spectrograph (GNIRS) and infer their gas-phase metallicity from the [NII]$\lambda$6584 and H$\alpha$ line ratio. We find that clumpy galaxies have lower metallicities compared to non-clumpy galaxies, with an offset in the weighted average metallicity of $0.07 \pm 0.02$ dex. We also find an offset of $0.06 \pm 0.02$ dex between clumpy and non-clumpy galaxies in a comparable sample of 23 star-forming galaxies at $z \approx 1.5$ using existing data from the FMOS-COSMOS survey. Similarly, lower [NII]$\lambda$6584/H$\alpha$ ratio are typically found in galaxies that have more of their $UV_{rest}$ luminosity originating from clumps,

suggesting that “clumpier" galaxies are more metal poor. We also derive the intrinsic velocity dispersion and line-of-sight rotational velocity for galaxies from the GNIRS sample. The majority of galaxies have $\sigma_0/v_c \approx 0.2$, with no significant difference between clumpy and non-clumpy galaxies. Our result indicates that clump formation may be related to the inflow of metal-poor gas; however, the process that forms them does not necessarily require significant, long-term kinematic instability in the disk.

## 2.1 Introduction

An important step to understanding the formation and evolution of galaxies is to trace where star formation is occurring within galaxies through cosmic time. Deep field observations reveal an abundance of high-redshift star-forming galaxies (SFGs) that host prominent, clumpy star formation (Wuyts et al., 2012; Murata et al., 2014; Guo et al., 2015; Förster Schreiber et al., 2018a; Sok et al., 2022; Sattari et al., 2023). Interestingly, the fraction of clumpy galaxies evolves with the cosmic star formation rate density, where it peaks at cosmic noon before declining toward higher redshift (Shibuya et al., 2016); however recent results from gravitationally-lensed galaxies observed with JWST suggest clumpy star formation may dominate the mass assembly of galaxies at $z > 4$ (e.g., Vanzella et al. 2022; Adamo et al. 2024; Mowla et al. 2024). Star-forming clumps are observed as UV-bright, kiloparsec-scale structures, but are also detected in H$\alpha$ emission. The specific star formation rate (sSFR) of clumps is elevated compared to the surrounding regions (Wuyts et al. 2012, 2013). A gradient in the sSFR of clumps is also reported by Guo et al. (2018), where clumps that are further away from the galactic centers typically have higher sSFR. They found a similar negative gradient in stellar masses and ages, possibly suggesting an inward migration of clumps. In general, clumps have an estimated stellar masses of $10^7 - 10^9$ M$_\odot$ (e.g., Soto et al. 2017; Zanella et al. 2019; Huertas-Company et al. 2020), with lower mass estimates report in lensed galaxies (Cava et al., 2018). Nonetheless, the ubiquity and unique mode of clumpy star formation during a period known for galaxy mass assembly suggest that clumps play an integral role in galaxy formation.

The two broadly discussed pathways for the formation of star-forming clumps are *in-situ* formation through violent disk instability and gravitational fragmentation within the host galaxy's gas disk, and *ex-situ* origins via the accretion of external satellites. In the *in-situ* channel, clumps arise from the fragmentation of gas-rich, turbulent disks, a process commonly described by the Toomre stability parameter $Q$ (e.g., Toomre 1964; Agertz et al. 2009; Dekel et al. 2009b; Ceverino et al. 2010). At high redshift, such gas-rich disks are maintained in a marginally unstable state by continuous gas accretion from the cosmic web (e.g., Dekel et al. 2009a).

Observational support for *in-situ* fragmentation due to gas accretion comes from studies of the co-evolution of the clumpy fraction, merger rates, and disk instability indicators, which show that mergers alone cannot account for the prevalence of clumps at $z \sim 2$ (e.g., Guo et al. 2015; Adams et al. 2022; Sattari et al. 2023). Moreover, the elevated star formation activity in clumpy galaxies relative to nonclumpy systems is difficult to explain solely through merger-driven triggering and instead points to sustained internal gas inflows and disk instabilities (Martin et al. 2023). The observed stellar mass function of clumps is likewise consistent with an *in-situ* origin from disk fragmentation (Dessauges-Zavadsky and Adamo 2018; Huertas-Company et al. 2020), whereas clumps of *ex-situ* origin (i.e., accreted satellites) are expected to produce a flatter mass function (e.g., Mandelker et al.

2017).

Under the gas accretion leading to VDI framework, we postulate that if clumps form within a gas-rich disk that is replenished by (relatively) pristine (or metal-poor) gas from cosmological accretion, it is expected that the accreting gas would dilute the metal contents within the disk, leading to lower metallicities in clumpy galaxies compared to non-clumpy galaxies. Indeed, such a scenario has been observed in a simulation by Ceverino et al. (2016), where clumps are found to have lower metallicities by 0.3 dex compared to the surrounding regions. Therefore, one way for testing *in-situ* clumpy formation is to compare the metallicity of the gas disk for clumpy and non-clumpy galaxies (i.e., galaxies that do not host star-forming clumps).

In this work, we use the Gemini-North Near-Infrared Spectrograph (GNIRS) on the Gemini-North observatory to indirectly measure the gas-phase metallicity for a sample of clumpy and non-clumpy galaxies at $z \approx 0.7$ and test the validity of clump formation under violent disk instability. We also make use of another large near-infrared spectroscopic survey (FMOS-COSMOS; Silverman et al. 2015; Kashino et al. 2019), which will serve as an independent test to see if the metallicity offset is observed in a different sample of SFGs at $z \approx 1.5$. The metallicity is inferred from the [NII]$\lambda$6584 and H$\alpha$ line ratio, following the calibration from Pettini and Pagel (2004). This line ratio is commonly used as both lines are easily accessible from the same filter, and their proximity mitigates the effect of dust in the inferred metallicity. The structure of the chapter is as follows. First, we discuss our galaxy sample, as well as define the nomenclature for clumpy and non-clumpy galaxies in Section 2.2. Overviews of the GNIRS observations and data reduction are then reported in Section 2.3 and 2.4, respectively. In Section 2.5, we present our metallicity measurements and compare them to other studies. Given the high signal-to-noise ratio of some of our spectra, we also investigate the relation between disk kinematics and morphologies (i.e., clumpy vs non-clumpy) in Section 2.6. Finally, we discuss the implications of our results in the context of clump formation due to violent disk instability in Section 2.7.

Throughout the chapter, we adopt a $\Lambda$CDM cosmological model of the universe with $\Omega_\lambda = 0.7$, $\Omega_{\rm M} = 0.3$, and a Hubble constant of $H_0 = 70$ km/s/Mpc. All physical parameters are determined by assuming a Chabrier initial mass function (Chabrier, 2003), and our magnitudes are reported in AB magnitudes.

## 2.2 Galaxy Sample

### 2.2.1 GNIRS Sample

Our targets are selected using the publicly available photometry of the COSMOS/UltraVISTA catalog (Muzzin et al., 2013). A number of criteria are employed to determine the candidates. We first require that the galaxies are at a redshift, either photometric or spectroscopic where available, where both H$\alpha$ and [NII]$\lambda$6584 lie within the wavelength coverage of the $X$ and $J$ filter of GNIRS, which avoid the highly contaminated atmospheric skylines in the $H$-band. The choice in the two GNIRS filters correspond to the detection of H$\alpha$ and [NII]$\lambda$6584 between the redshift range of $0.6 < z < 0.9$.

Using the stellar masses and SFRs from the UltraVISTA catalog, we then select galaxies of similar masses and SFRs to ensure that any observed metallicity offset is not driven by the fundamental metallicity relation. The masses were derived by fitting the spectral energy distribution of 30 photometric filters, extending from $0.15-24\ \mu$m, and the SFRs were determined from the rest-frame UV and infrared luminosity. We select primarily galaxies with $\log(M_*/M_\odot) \approx 10.5$. This corresponds to star-forming galaxies with a high SFR ($\gtrsim$10 $M_\odot yr^{-1}$), and therefore increases the likelihood of a strong detection for both H$\alpha$ and [NII]$\lambda$6584. Lastly, we select galaxies of similar colors (i.e., $U-V$ and $V-J$) to mitigate the effect of dust in our analyses as galaxies with higher dust attenuation generally have higher metallicities.

Figure 2.1 summarizes our GNIRS galaxy sample. We select 42 galaxies; 23 are clumpy galaxies and 19 are non-clumpy galaxies (the classification of clumpy and non-clumpy is presented in Section 2.2.4). The solid markers denote galaxies that have a detection of SNR $>$ 3 in both H$\alpha$ and [NII]$\lambda$6584. Throughout the analysis we use only these S/N $>$ 3 galaxies, which reduces our sample to 32 galaxies. The average mass for the clumpy galaxies is $\log(M_*/M_\odot) = 10.48^{+0.15}_{-0.02}$, with an average SFR of $\log(\mathrm{SFR}\ [M_\odot\ yr^{-1}]) = 1.23^{+0.19}_{-0.14}$. The average mass for non-clumpy galaxies is $\log(M_*/M_\odot) = 10.54^{+0.04}_{-0.06}$, and their average SFR is $\log(\mathrm{SFR}\ [M_\odot\ yr^{-1}]) = 1.30^{+0.12}_{-0.28}$. The average specific SFR for both clumpy and non-clumpy galaxies is approximately $\log(\mathrm{sSFR}\ [yr^{-1}]) \approx -9.2$. Similarly, clumpy galaxies have an average $U-J = 1.21^{+0.11}_{-0.07}$ and $V-J = 0.95^{+0.11}_{-0.08}$, and for non-clumpy galaxies, $U-J = 1.32^{+0.08}_{-0.11}$ and $V-J = 1.05^{+0.06}_{-0.08}$. Here, the lower and upper errors are derived from the 16-th and 84-th percentiles.

### 2.2.2 FMOS-COSMOS Sample

The FMOS-COSMOS near-infrared spectroscopic survey (Silverman et al., 2015; Kashino et al., 2019) utilized the Fiber-Multi Object Spectrograph on the Subaru Telescope, and comprised of an observing mode that primarily targets star-forming galaxies at $z \approx 1.5$. In particular, the survey provides fluxes for both H$\alpha$ and [NII]$\lambda$6584. The primary goal of this sample is to test whether a metallicity offset is also detected for a completely different sample of galaxies that have been observed with a different observing strategy, but are selected similarly in masses, SFRs, and colors as the GNIRS sample.

We cross-match galaxies from the FMOS catalog to the UltraVISTA catalog, keeping only galaxies that have a measured RA and Dec that are within 0.2 arcsec of both catalogs. This angular scale corresponds to less than two imaging pixels of the UltraVISTA/COSMOS mosaics (which has a pixel scale of 0.15 arcsec). This is also well below the size of the images' point spread function, and therefore ensures that there is no contamination from close-by neighbours. We select FMOS galaxies that have H$\alpha$ and [NII]$\lambda$6584 fluxes detected to at least 3 sigma. Similar to the GNIRS sample, we also select galaxies within a mass range between $10.4 < \log(M_*/M_\odot) < 10.6$, and limit the analyses to galaxies of similar colors, i.e., $0.9 < U-V < 1.2$ and $0.7 < V-J < 1.1$. In total, there are 23 FMOS galaxies at $z \approx 1.5$ that meet our selection criteria. We reiterate that the goal for this sample is mainly to determine whether a metallicity offset is similarly observed between clumpy and non-clumpy galaxies when using a different sample of galaxies. In the following sections, we discuss how galaxies are classified as clumpy and non-clumpy from their ground-based images.

### 2.2.3 Resolving clumps with finite-resolution deconvolution

While clumpy galaxies are ubiquitous at high-redshift, identifying galaxies as clumpy is difficult due to the fact that high-resolution observations are required to resolve the kiloparsec-scale structures.

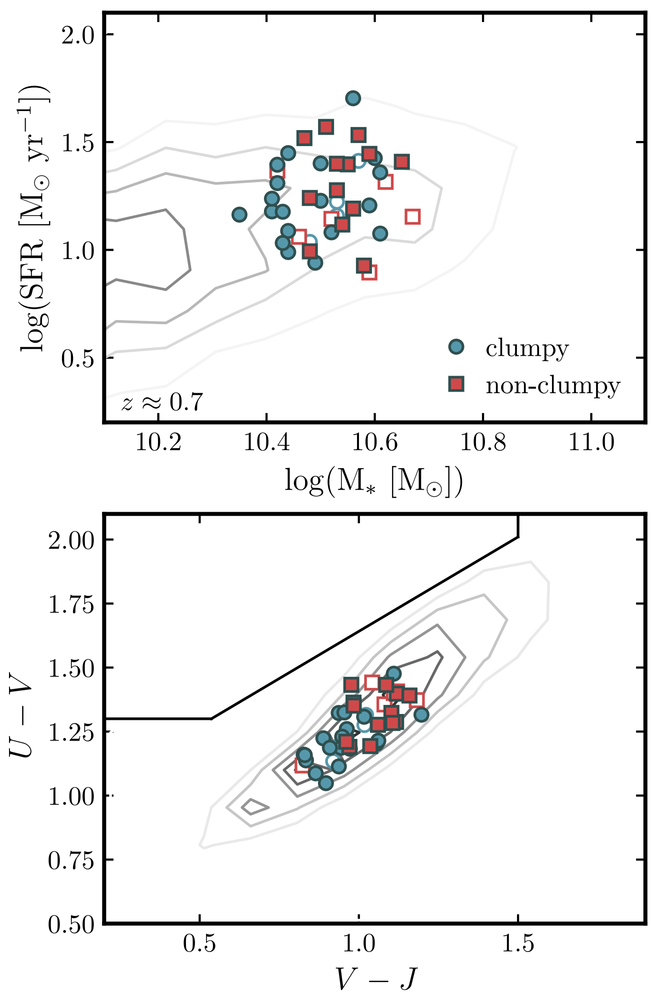


Figure 2.1: Top: The contour lines show the star-forming main sequence for the parent sample of galaxies at $z \approx 0.7$ in the COSMOS field. The blue circle and red square markers denote clumpy and non-clumpy galaxies that were observed with GNIRS, respectively. The solid markers denote galaxies with detection in both H$\alpha$ and [NII]$\lambda$6584 (SNR$>$ 3). Bottom: Similar to the top panel, but is now showing the $UVJ$ diagram. In general, the selection process is limited to galaxies of similar masses, SFRs and colors. This ensures that we can directly test for a relation between clumpy morphologies and metallicities.

Furthermore, multi-wavelength images are typically needed to account for the morphological K-correction effect, in which the morphology of galaxies may change depending on which wavelength they are observed in. This becomes particularly important when the galaxies are observed over a wide range of redshifts.

A unique solution to obtaining high-resolution and multi-wavelength imaging for a large volume of the sky is to use image deconvolution. This had been previously done in Sok et al. (2022) using finite-resolution deconvolution (`FIREDEC`; Cantale et al. 2016a), where we deconvolved the multi-band imaging of approximately 20,000 SFGs in the COSMOS field at $0.5 < z < 2$, and created resolved stellar mass and surface brightness maps for those galaxies. That work showed that `FIREDEC` recovers clumps in galaxies from ground-based imaging. Comparing the deconvolved images to HST observations confirmed that these clumpy features are intrinsic to the galaxies and not artifacts of deconvolution.

A technical overview of image deconvolution and `FIREDEC` is provided in Cantale et al. (2016a), with practical applications discussed in Cantale et al. (2016b) and Sok et al. (2022). In summary, this deconvolution algorithm aims to to deconvolve images to a finite resolution, ensuring that the resulting image retains its own PSF (i.e., as opposed to completely removing the PSF to achieve an infinite angular resolution, which would violate the Nyquist's sampling theorem). In practice, a target resolution is defined to be a 2D Gaussian function. The deconvolution kernel is then constructed through a combination of analytical and numerical fits, such that the convolution between the deconvolution kernel and the Gaussian functions gives back the PSF of the image.

In practice, the ground-based images have a pixel scale of $0.15''$, with typical seeing greater than $0.65''$. In order to deconvolve images in different filters of the COSMOS field, we constructed deconvolution kernels for each filter. This is required as the PSF can vary from filter to filter, which arises due to the different observing conditions (see Sok et al. 2022 for the PSF variation for each filter of the COSMOS field). Here, we deconvolved to a seeing with angular resolution of 0.3" which probes a physical scale of approximately 2.3 kpc for galaxies at $z \approx 0.7$. While this angular resolution only partially resolves star-forming clumps, it is sufficient for identifying whether a galaxy is clumpy or non-clumpy. A comparison between the ground-based imaging and the deconvolved images for a few galaxies in this study is shown in Figure 2.2. These galaxies are chosen as examples because they have multi-band HST imaging, which enables us to directly compare the deconvolved images to HST images. Galaxies that are classified as clumpy are denoted by the asterisk in the ID name in the right panel. In general, deconvolution reveals multiple clumps that are not detectable with ground-based imaging. These clumps are also detected by HST, suggesting that the resolved clumps are intrinsic to the galaxies, and not deconvolution artifacts.

### 2.2.4 Clumpy/Non-clumpy SFGs Classification

A number of methods have been used to classify whether a galaxy is clumpy or non-clumpy. Clumps can be identified by creating a contrast image, in which an image is passed through a high-pass filter (e.g., Zanella et al. 2019; Claeyssens et al. 2023). The flux and size of the clumps can be obtained by fitting each clump with a Gaussian model. Recent studies have also employed a convolutional neural network to identify the location of clumps from an image (Huertas-Company et al., 2020).

In this work, we follow the method described in Wuyts et al. (2012), which utilizes the normalized rest-frame $U$-band light profile to identify clumpy regions in a galaxy. The rationale of

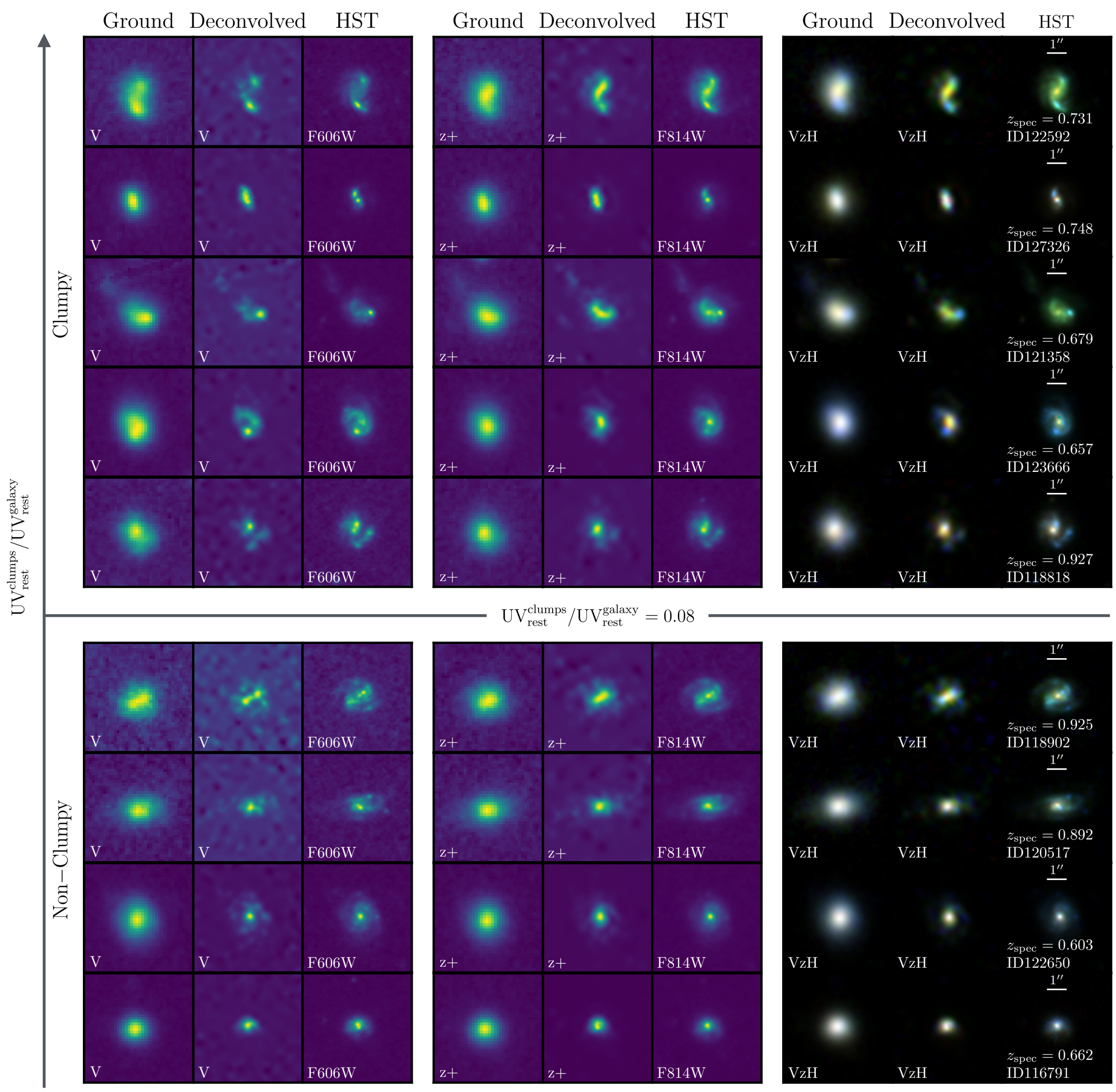


Figure 2.2: Comparison of deconvolved images to the corresponding HST images. From the left to right panel, we compare the $V$-band imaging to *F606W*, $z$+-band imaging to *F814W*, and the composite $VzH$ image to a composite image constructed from *F606W*+*F814W*+*F160W*. Each stamp is $7.8'' \times 7.8''$. For each panel, the left to right column shows the ground-based image, the deconvolved image, and the corresponding HST image. Only 9 out of the 42 galaxies from our sample have HST imaging in the 3 HST filters. The images are deconvolved to a resolution of 0.3", corresponding to a physical scale of ~2.3 kpc at $z \approx 0.7$. We arrange the figure such that the galaxies are shown in ascending order based on their clumpiness, as defined by the ratio of rest-frame UV light emitted from clumps relative to the host galaxy. Clumpy galaxies are those with clumps that have at least 8% of the total rest-frame UV luminosity.

a normalized profile relies on the fact that regions with relatively high surface brightness, such as UV-bright clumps, will appear as a bump in an otherwise smooth light profile of a galaxy. We will briefly describe the steps to making a normalized light profile, however we refer the reader to Wuyts et al. (2012) for a more detailed process. We first obtain $U_{\rm rest}$ by spatially fitting the multi-band deconvolved images with `EAZY` (Brammer et al., 2008). Instead of fitting the spectral energy distribution (SED) over individual pixels, which can introduce biases for pixels with low signal-to-noise ratio (SNR), we instead bin pixels using a Voronoi binning technique (`vorbin`; Cappellari and Copin 2003) so that each Voronoi bin has a minimum SNR of 5 in the $Ks$-band. A similar binning technique has been employed to construct stellar mass maps from spatially SED fits (Wuyts et al., 2012; Sok et al., 2022; Tan et al., 2022).

Once a map of the $U_{\rm rest}$ surface brightness is created, the normalized light profile for a galaxy is constructed. We characterize the galaxies by two elliptical parameters; the ellipticity and position angle. This is done by fitting an elliptical function to a segmentation map of the galaxy. These parameters are then used to define elliptical apertures centered on the galaxy. We then define the effective radius of the galaxy as the radius at which half of the light from the galaxy is within that radius. This is measured from a curve of growth based on the elliptical apertures. The effective surface brightness is then measured as the mean surface brightness within the effective radius. The normalized profile is constructed by normalizing the $U_{\rm rest}$ surface brightness at each pixel by the effective surface brightness, and the galactocentric distance of each pixel by the effective radius. In essence, the normalized profile illustrates how the normalized surface brightness varies with distance from the center of the galaxy. Figure 2.3 shows two examples of the normalized light profile for a clumpy and non-clumpy galaxy.

Clumps are situated in the same parameter space within the normalized profile. The division lines of the inner, outer, and clumpy regions, shown in Figure 2.3, are defined in Eqn. 6-8 of Wuyts et al. (2012), based on their visual inspections of the stacked profiles of their sample. Galaxies are defined as clumpy if they have pixels that originate from the clumpy regime in the normalized profile, and that contribute $\geq$8% of the total $U_{\rm rest}$ luminosity of the host galaxy. This criterion is commonly used to identify star-forming clumps in high-redshift galaxies in other studies, and therefore represents a consistent way of classifying clumpy galaxies.

## 2.3 GNIRS Observations

A total of 42 galaxies in the COSMOS field were observed in seeing limited conditions with GNIRS on the Gemini-North observatory. These observations took places between 2019 and 2022 as part of several GNIRS observation programs, with program IDs: GN-2019B-Q-131, GN-2019B-Q-229, GN-2020A-Q-307, GN-2020B-Q-215, GN-2021B-Q-212, GN-2022A-Q-213. The observations were done through the long-slit spectroscopy mode in the $X$ or $J$ filter and the 111 line per mm grating. We use the short camera, corresponding to a pixel resolution of 0.15 arcsec per pixel, with a slit opening of 1 arcsec. The observational setup therefore corresponds to a wavelength coverage between 1.03 - 1.37 microns, at a spectral resolution of $R \approx 2000$ ($\sigma_{\rm inst} \sim 50$ km/s). This spectral resolution is sufficient for differentiating the H$\alpha$ and [NII]$\lambda$6584 emission lines. The on-source integration time for each galaxy is set either to 40, or 60 minutes. The shorter integration time is used when the observations took place with good seeing (i.e., Band 1) condition, while the longer 60 min of integration is needed

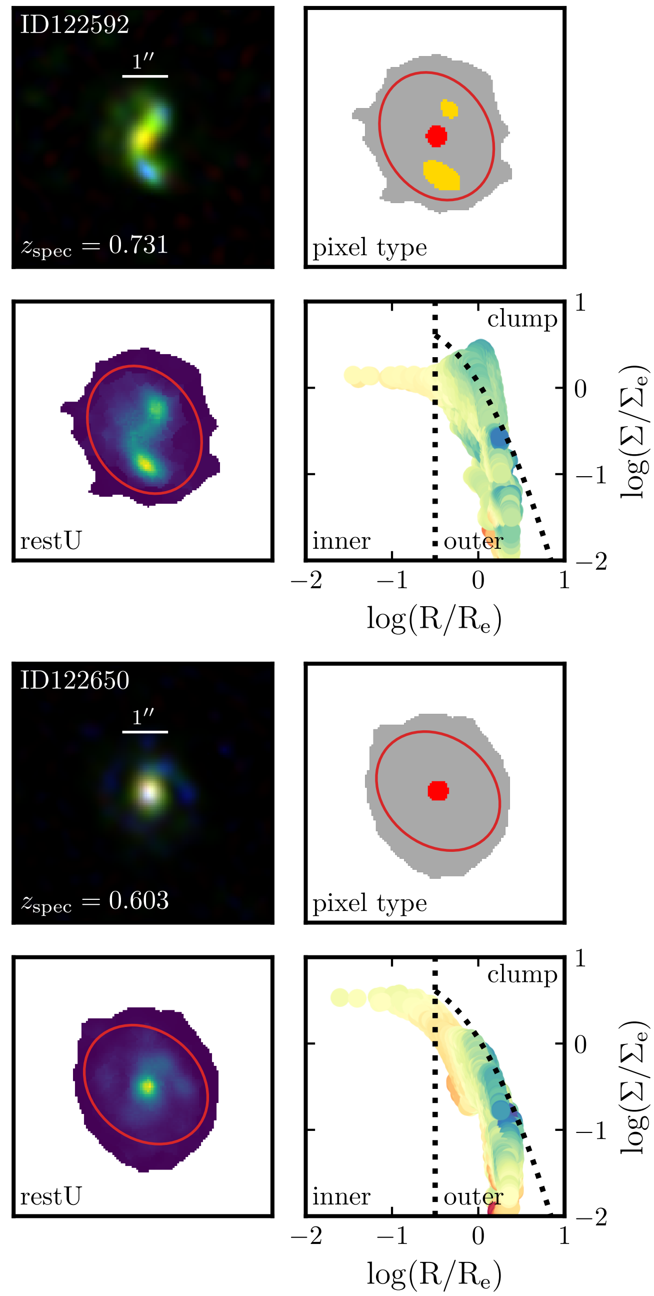


Figure 2.3: Two examples showing how the normalized light profile is used to identify clumps. The top figure shows the normalized profile for a clumpy galaxy while the bottom figure shows the profile for a non-clumpy galaxy. In each figure, the top left panel shows the composite $VzH$ image ($6'' \times 6''$), the bottom left shows the surface brightness, the bottom right shows the constructed normalized $U_{\rm rest}$ profile, and top left shows the pixel type map based on the division lines in the normalized light profile. The clump regime is defined as the pixel residing above the curved dotted line. As shown in the top figure in the pixel type map, clumps are identified easily using the normalized profile. The color-coding in the normalized profile respond to the $(U - V)_{\rm rest}$. Pixels in the clumpy regime are UV-bright.

when observing in the degraded seeing of Band 3 to account for greater light loss through the 1" wide slit opening.

A blind offset from a reference bright star (or a reference object when there are no nearby star) is needed in order to center the slit onto our faint galaxies. The position of the slit is fixed along the major axis of the galaxy to ensure that most of the light coming from the galaxy is captured by the slit. We observe each galaxy using the ABB′A′ pattern, which corresponds to the dithering nod of -1", 6", -6", 1" along the slit direction. This pattern is used for sky subtraction, but helps mitigate pixel persistence when observing bright objects such as a standard star. At each nodding position, we have 300 seconds of integration. The standard star is observed either before, or after the exposures on the galaxy using the same observational setup. The spectroscopic observations of a standard star is needed in order to apply flux calibration for telluric absorption features in the near-infrared.

## 2.4 Near-Infrared Spectra

The reduction of the raw data is done using `Gemini IRAF` (Gemini Observatory and AURA, 2016). A number of raw exposure frames for the science observations were affected by striping patterns; however these patterns were removed with `cleanir`[1]. Telluric and science observations are then flat-fielded, sky-subtracted and wavelength-calibrated. The master flat is obtained by stacking flat frames that do not have anomalous mean, or standard deviation. The wavelength solution is computed by running `nswavelength` on the included Argon lamp observations. We extract the one-dimensional spectrum for the standard star using `nextract`. In the following sections, we describe the steps to extract the one-dimensional spectra from the reduced two-dimensional science spectra.

### 2.4.1 Spectra and velocity curves

One simple method for extracting a one-dimensional spectrum from the two-dimensional spectrum is the boxcar method, in which the two-dimensional spectrum is summed along the slit direction. However, this can lower the SNR of the emission lines if the galaxy's rotation is not accounted for. In this section, we describe the steps to extract the one-dimension spectrum by modeling the kinematics of the galaxy.

The kinematics modeling is performed using similar steps as outlined in Weiner et al. (2006). We first obtain the velocity and dispersion profile of a galaxy by fitting a Gaussian function to its emission lines along the slit. To this end, we bin the two-dimensional spectrum along the spatial direction by 2 pixels to increase the SNR of the emission lines. Here, H$\alpha$ is primarily used when fitting the Gaussian function as it has a higher SNR compared to [NII]$\lambda$6584. However, [NII]$\lambda$6584 is used when H$\alpha$ lies next to a strong skyline as the centroiding of the Gaussian function may be affected by the skyline. The fitted parameters are determined through a least-squares Levenberg-Marquardt minimization algorithm with respect to the observed data using `lmfit` (Newville et al., 2016). The retrieved error represents an uncertainty of one standard deviation around the fitted values. A few examples of the measured velocity profile are plotted on the right panels of Figure 2.4.

[1]https://github.com/andrewwstephens/cleanir

The kinematics model consists of two components; a rotation curve and the dispersion term. We assume that the intrinsic velocity dispersion $\sigma_0$ is constant with radius for each galaxy. The inverse tangent function is adopted for the rotation of the galaxies,

$$v(r) = v_0 + \frac{2}{\pi} v_a \arctan\left(\frac{r - r_0}{r_t}\right), \tag{2.1}$$

where $v_a$ is asymptotic rotation velocity, $r_o$ is the spatial center of the galaxy, and $r_t$ is the knee radius to the flat part of the curve. When modeling the emission line, we vary values for $v_0$, $r_0$, $v_a$ and $\sigma_0$, while keeping $r_t$ constant at a fiducial value of $0.2''$ due to the limited seeing. The modeled emission line is then spatially convolved with a Gaussian seeing kernel. We determine the appropriate FWHM from the acquisition images of the bright offset star, or from the acquisition images of the standard star (which are observed before or after the science integration). The modeled emission line is also convolved by the line spread function. We use `emcee` (Foreman-Mackey et al., 2013) to explore the parameter space, assuming a flat prior for each parameter. The best-fit parameters and their errors are taken to be the median and standard deviation, respectively. We note that $\sigma_0$ is not resolved for a few galaxies due to our relatively coarse spectral resolution of R $\sim$ 2000.

The rotation models are then used to align the emission lines along the slit. The one-dimensional, velocity-corrected spectrum are then optimally extracted, using the extraction algorithm from Horne (1986). The weighting profile is determined from the spatial profile of the emission lines. In cases where the emission lines have low SNR, we perform a standard boxcar extraction over some aperture width (i.e., $2.1''$ to $2.7''$ depending on the size of the galaxy). Finally, we do not apply an aperture correction for possible slit losses. While the size of galaxies varies and can therefore affect the measured [NII]$\lambda$6584/H$\alpha$ ratio, we mitigate this by placing the slit along the major axis that contains clumps. A slit correction would also assume some generalization of the metallicity gradient, which can be diverse (e.g., Gillman et al. 2021). However, larger galaxies typically have a negative metallicity gradient (Carton et al., 2018). Since clumpy galaxies are generally larger compared to non-clumpy galaxies (e.g., Martin et al. 2023), and assuming a negative metallicity gradient for clumpy galaxies, we expect the measured metallicity for clumpy galaxies to be overestimated compared to non-clumpy galaxies when using a slit that is centered on the central region of the galaxies. Given these reasons, we choose to not apply an uncertain slit-loss correction and present the data as-is. Finally, we note that the size difference between non-clumpy and clumpy galaxies is observed for the GNIRS sample (e.g., see Fig. 2.10), where we find the average effective radius of clumpy and non-clumpy galaxies to be $5.1 \pm 0.3$ and $3.4 \pm 0.3$ kpc (corresponding to $0.7''$ and $0.57''$), respectively. Figure 2.4 summarizes this section, showing examples of the observed rotation in the two-dimensional spectrum, the derived velocity curve, and the extracted spectrum after correcting for the velocity curve.

### 2.4.2 Telluric corrections

Molecules in the atmosphere can absorb and subsequent suppress certain spectral features in the near- and mid-infrared. These telluric absorption features may therefore affect the ratio between [NII]$\lambda$6584 and H$\alpha$ if the two lines fall within regions of strong absorption. Correcting for the absorption features is done by observing the spectrum of a standard star. Each of our observing blocks included a near-infrared spectroscopic observation of a standard star. The standard stars are

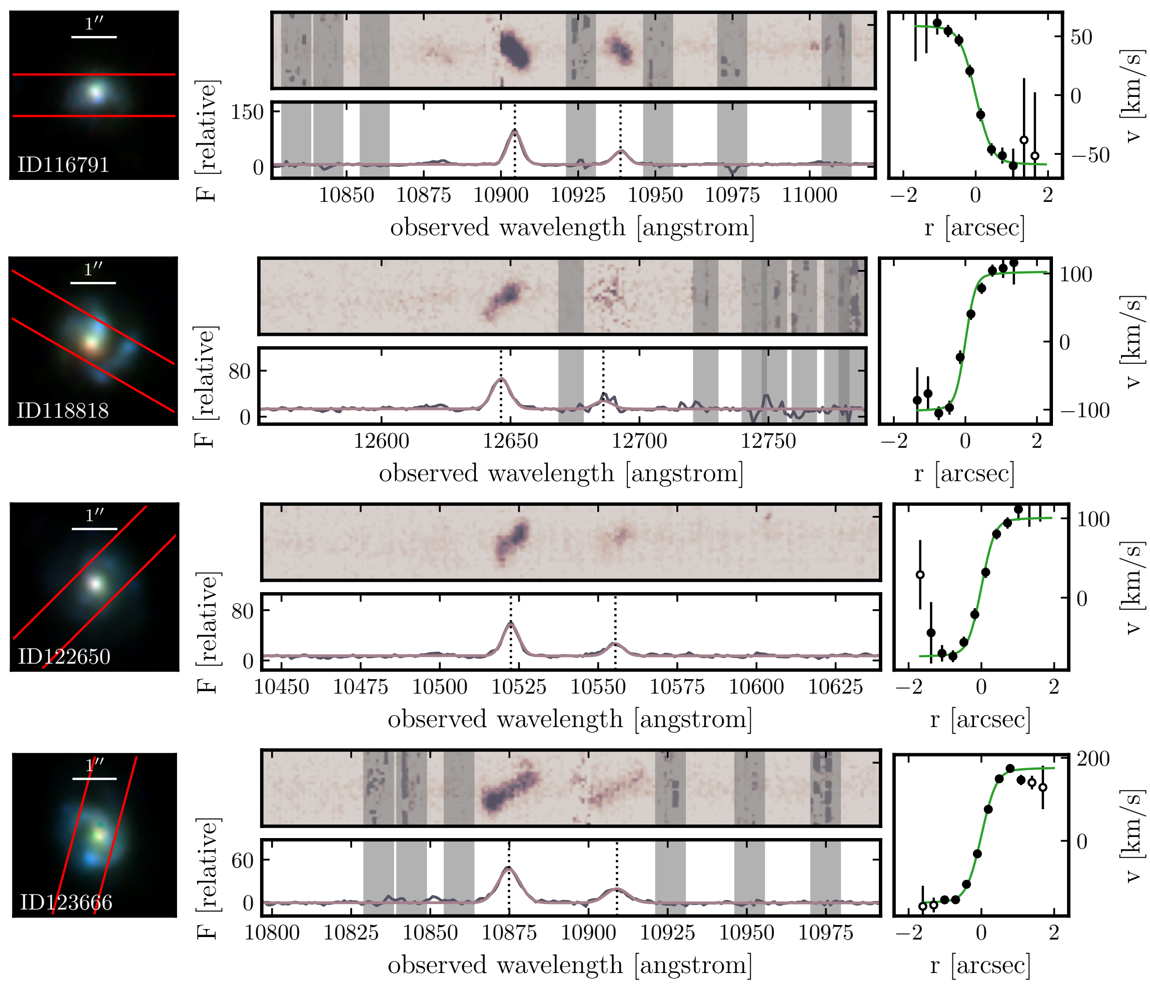


Figure 2.4: Examples of spectra along with the rotation curve of the galaxies. The left panel shows the composite image of the galaxies. The red line denotes the slit, which is purposely chosen to lie along the axis containing clumps. The top middle panel shows the two-dimensional spectrum where the y-axis is the slit (or spatial) direction and the x-axis is the dispersion direction. The bottom middle panel shows the extracted 1D spectrum after correcting for the rotation curve. The grey highlights are regions that are contaminated by skylines. The right panel shows the rotation curve of the galaxies, where the green line is the fitted rotation curve, and the black markers are the measured centroids of the emission line at each spatial bin along the two-dimensional spectrum.

selected to be early-type stars at a similar airmass, and are selected primarily because their spectra are generally devoid of intrinsic stellar absorption features in the near-infrared (between $1.05 - 1.3$ microns), with the exception of hydrogen recombination lines. However, these hydrogen lines can be fitted and removed. The steps to apply the telluric correction to our observations are listed below.

1. Hydrogen absorption lines in the standard star telluric observations are first masked out. The continuum of the masked spectrum is modeled as the sum of Chebyshev polynomials using `specutils`.

2. The reference spectrum for each standard star is obtained from the stellar spectral library of Pickles (1998). These reference spectra are then flux-calibrated so that their $V$-band magni-

tude is equal to the observed magnitude of the respective star. We mask the hydrogen lines in the flux-calibrated reference spectra and model their continuum using a similar step as above.

3. The correction factor at each wavelength is obtained by dividing the flux-calibrated spectrum by the observed telluric spectrum.

In general, we find that the correction factor for the ratio of [NII]$\lambda$6584 and H$\alpha$ ranges between 0.8 and 1.5. On average, the correction factor for the clumpy and non-clumpy sample is 1.05 and 1.09, respectively.

## 2.5 Comparison between Clumpy & Non-Clumpy Galaxies

### 2.5.1 [NII]$\lambda$6584/H$\alpha$

Figure 2.4 shows four examples of the reduced one-dimensional, telluric-corrected spectrum. The flux for each emission line is obtained by fitting a double-Gaussian function using `lmfit`. To this end, we fix the relative wavelength offset between the two emission lines, and assume that the line width for both lines are the same. Therefore, our fitting function consists of five parameters; the peak intensity ($a$) for both [NII]$\lambda$6584 and H$\alpha$, the line width ($\sigma$) for the emission lines, the wavelength center for H$\alpha$, and a constant continuum/background noise. The integrated flux for both emission lines is calculated as $a\sigma\sqrt{2\pi}$, using the respective line intensity. Similarly, the flux error is obtained by propagating the uncertainties of the fitted parameters $a$ and $\sigma$.

Out of the 42 galaxies, 2 galaxies are not detected in either H$\alpha$ and [NII]$\lambda$6584. Only 32 out of the remaining 40 galaxies have H$\alpha$ and [NII]$\lambda$6584 lines that are not significantly affected by atmospheric skylines, and have an SNR $> 3$. Table 2.1 lists their measured line ratios and their associated error. We find the weighted average of the [NII]$\lambda$6584/H$\alpha$ ratio for the entire sample to be $0.36 \pm 0.01$. When separated into clumpy and non-clumpy sample, we find that clumpy galaxies have a weighted flux ratio average of $0.31 \pm 0.02$, while non-clumpy galaxies have a flux ratio of $0.42 \pm 0.02$. The difference in the [NII]$\lambda$6584 and H$\alpha$ flux ratio is $0.11 \pm 0.02$. Taken at face value, this suggests a clear metallicity difference between the two samples; however, in the next section we examine this difference more closely via the mass-metallicity relation and the fundamental metallicity relation.

### 2.5.2 The mass-metallicity relation

The gas-phase metallicity of each galaxy is inferred by converting the [NII]$\lambda$6584/H$\alpha$ line ratio to an oxygen abundance, using the N2 line ratio calibration from Pettini and Pagel (2004), which has the following form,

$$12 + \log(\mathrm{O/H}) = 8.9 + 0.57 \times \log([\mathrm{NII}]\lambda 6584/\mathrm{H}\alpha) \quad (2.2)$$

This calibration is valid for $-2.5 < \log([\mathrm{NII}]/\mathrm{H}\alpha) < -0.3$. The error associated with the slope and intercept are 0.03 and 0.04, respectively. The [NII]$\lambda$6584/H$\alpha$ line ratio and inferred metallicities are listed in Table 2.1.

The left side of Figure 2.5 shows the mass-metallicity relation for 32 SFGs with a detected SNR of at least 3 for both H$\alpha$ or [NII]$\lambda$6584, color-coded as clumpy vs non-clumpy. The weighted averages

Table 2.1: Derived values for the GNIRS observations. We list only galaxies with detected H$\alpha$ and [N II] lines. The [N II]/H$\alpha$ line ratio is used to obtain the gas-phase metallicity using the calibration from Pettini and Pagel (2004).

| COSMOS ID | RA [deg] | Dec [deg] | $z_{\rm spec}$ | $\log(M_*/M_\odot)$ | $\log({\rm SFR}_{\rm UV+IR})$ | ${\rm UV}_{\rm clumpy}$ | [N II]/H$\alpha$ | $12+\log({\rm O/H})$ | $v_c$ [km s$^{-1}$] | $\sigma_0$ [km s$^{-1}$] |
|---|---|---|---|---|---|---|---|---|---|---|
| 2289 | 150.48253 | 1.7125193 | 0.712 | 10.44 | 0.99 | True | 0.26 ± 0.02 | 8.56 ± 0.02 | 172.3 ± 14.8 | 13.7 ± 6.1 |
| 40738 | 150.24486 | 1.9678429 | 0.936 | 10.44 | 1.45 | True | 0.21 ± 0.05 | 8.51 ± 0.06 | – | 35.8 ± 23.7 |
| 62945 | 149.72485 | 1.8760713 | 0.694 | 10.44 | 1.09 | True | 0.28 ± 0.07 | 8.58 ± 0.06 | – | – |
| 64167 | 149.66539 | 1.8908899 | 0.749 | 10.48 | 0.99 | False | 0.28 ± 0.06 | 8.58 ± 0.05 | – | – |
| 73816 | 150.50740 | 2.0096676 | 0.702 | 10.59 | 1.45 | False | 0.38 ± 0.04 | 8.66 ± 0.02 | 212.5 ± 22.2 | 4.1 ± 7.5 |
| 85093 | 150.67809 | 2.1846418 | 0.795 | 10.52 | 1.08 | True | 0.44 ± 0.14 | 8.70 ± 0.08 | – | – |
| 91389 | 150.72302 | 2.2811966 | 0.957 | 10.41 | 1.18 | True | 0.25 ± 0.07 | 8.55 ± 0.07 | – | 26.3 ± 17.1 |
| 91840 | 150.68439 | 2.2863963 | 0.607 | 10.61 | 1.08 | True | 0.49 ± 0.12 | 8.72 ± 0.06 | – | 21.5 ± 13.9 |
| 94152 | 150.45470 | 2.3239200 | 0.662 | 10.58 | 0.93 | False | 0.37 ± 0.08 | 8.65 ± 0.06 | 186.7 ± 46.3 | 37.3 ± 20.4 |
| 111826 | 150.16794 | 2.0991180 | 0.961 | 10.65 | 1.41 | False | 0.40 ± 0.08 | 8.67 ± 0.05 | – | 23.3 ± 15.2 |
| 116791 | 150.07150 | 2.1623764 | 0.662 | 10.57 | 1.53 | False | 0.42 ± 0.01 | 8.68 ± 0.01 | 108.2 ± 13.7 | 23.3 ± 7.0 |
| 118818 | 150.15173 | 2.1886752 | 0.927 | 10.56 | 1.70 | True | 0.27 ± 0.04 | 8.58 ± 0.04 | 151.5 ± 17.8 | 54.6 ± 7.5 |
| 118902 | 150.15822 | 2.1901271 | 0.925 | 10.51 | 1.57 | False | 0.34 ± 0.05 | 8.63 ± 0.04 | 179.7 ± 37.6 | 17.4 ± 12.4 |
| 120517 | 150.18661 | 2.2085881 | 0.892 | 10.47 | 1.52 | False | 0.41 ± 0.05 | 8.68 ± 0.03 | 201.8 ± 30.1 | 79.4 ± 13.1 |
| 121358 | 150.17503 | 2.2170861 | 0.679 | 10.42 | 1.31 | True | 0.27 ± 0.02 | 8.57 ± 0.01 | 173.3 ± 28.7 | 35.7 ± 6.9 |
| 122592 | 150.14954 | 2.2334945 | 0.731 | 10.35 | 1.16 | True | 0.27 ± 0.03 | 8.58 ± 0.03 | 68.2 ± 15.0 | 21.3 ± 7.7 |
| 122650 | 150.07106 | 2.2330399 | 0.603 | 10.53 | 1.28 | False | 0.39 ± 0.02 | 8.67 ± 0.01 | 167.4 ± 20.6 | 14.8 ± 6.7 |
| 123666 | 150.21819 | 2.2458375 | 0.657 | 10.61 | 1.36 | True | 0.42 ± 0.02 | 8.69 ± 0.01 | 255.1 ± 16.3 | 40.0 ± 10.6 |
| 125867 | 150.36453 | 2.2735415 | 0.749 | 10.41 | 1.24 | True | 0.22 ± 0.05 | 8.53 ± 0.05 | 127.1 ± 36.3 | 32.1 ± 12.4 |
| 127326 | 150.17822 | 2.2921340 | 0.748 | 10.42 | 1.40 | True | 0.34 ± 0.09 | 8.63 ± 0.06 | – | 49.9 ± 20.1 |
| 151806 | 149.88297 | 2.1179976 | 0.677 | 10.49 | 0.94 | True | 0.30 ± 0.04 | 8.61 ± 0.03 | 149.4 ± 15.1 | 28.4 ± 8.6 |
| 153139 | 149.73523 | 2.1352146 | 0.698 | 10.43 | 1.18 | True | 0.41 ± 0.04 | 8.68 ± 0.02 | 171.7 ± 17.1 | 33.4 ± 11.0 |
| 159971 | 149.61758 | 2.2155590 | 0.675 | 10.54 | 1.12 | False | 0.49 ± 0.06 | 8.72 ± 0.03 | 190.9 ± 19.4 | 16.2 ± 11.8 |
| 161993 | 149.88490 | 2.2384937 | 0.703 | 10.55 | 1.40 | False | 0.66 ± 0.05 | 8.80 ± 0.02 | – | – |
| 162807 | 149.75438 | 2.2461481 | 0.730 | 10.59 | 1.21 | True | 0.40 ± 0.14 | 8.67 ± 0.08 | – | – |
| 185948 | 150.51099 | 2.5288169 | 0.852 | 10.56 | 1.19 | False | 0.38 ± 0.15 | 8.66 ± 0.09 | – | – |
| 191595 | 150.53107 | 2.6109066 | 0.815 | 10.53 | 1.40 | False | 0.51 ± 0.09 | 8.73 ± 0.04 | – | – |
| 192043 | 150.65984 | 2.6178505 | 0.891 | 10.50 | 1.23 | True | 0.31 ± 0.07 | 8.61 ± 0.05 | – | 35.6 ± 23.1 |
| 207078 | 149.98328 | 2.4797268 | 0.704 | 10.50 | 1.40 | True | 0.30 ± 0.03 | 8.60 ± 0.02 | 75.9 ± 26.2 | 20.0 ± 9.2 |
| 207362 | 150.17223 | 2.4845951 | 0.800 | 10.48 | 1.24 | False | 0.53 ± 0.09 | 8.74 ± 0.04 | – | – |
| 239578 | 149.84000 | 2.5136211 | 0.679 | 10.60 | 1.43 | False | 0.48 ± 0.05 | 8.72 ± 0.03 | 220.0 ± 26.6 | 15.6 ± 11.2 |
| 239638 | 149.72881 | 2.5153310 | 0.696 | 10.43 | 1.03 | True | 0.30 ± 0.07 | 8.60 ± 0.06 | 145.6 ± 23.1 | 38.8 ± 14.7 |

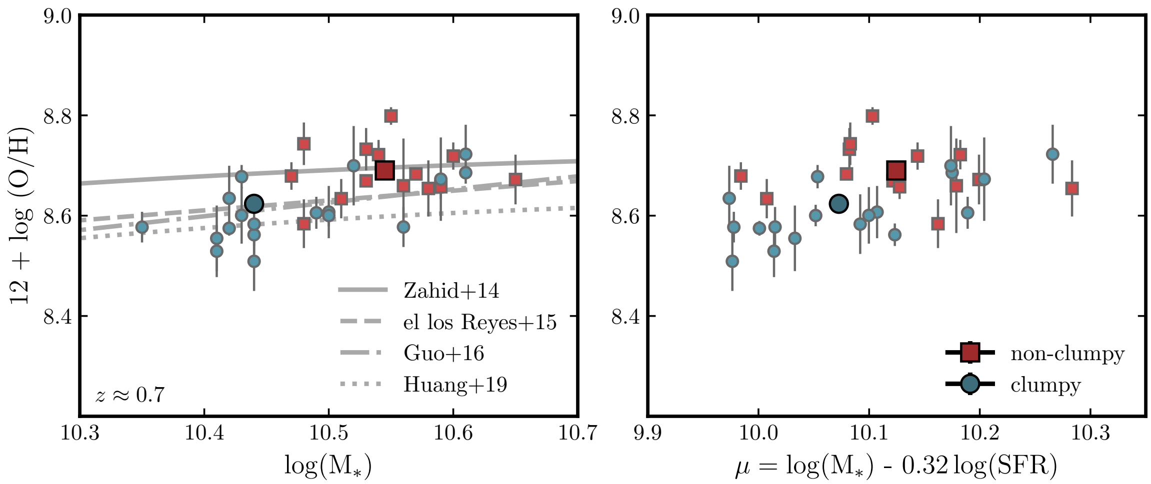


Figure 2.5: Left: the metallicities of clumpy and non-clumpy galaxies in relation to stellar masses. The dark, larger markers show the weighted average. The different lines show the mass-metallicity relation for galaxies at $z \approx 0.8$ from Zahid et al. (2014); de los Reyes et al. (2015); Guo et al. (2016), and Huang et al. (2019). The error is calculated as the weighted standard error. Right: The metallicities in relation to the projected stellar masses and star formation rate.

are calculated using the inverse variance weighting, and are denoted as the dark, large markers, and their errors are calculated as the weighted standard error. We also plot the mass-metallicity relation from other studies for galaxies at a similar redshift range. However, since these studies used a different metallicity calibration, which can lead to discrepancies in metallicities (e.g., Kewley and Ellison 2008), we also apply a polynomial correction to each mass-metallicity relation to convert them to the N2 calibration. The polynomial correction are taken from Teimoorinia et al. (2021).

The grey curves are the corrected mass-metallicity relation from Zahid et al. (2014), de los Reyes et al. (2015), Guo et al. (2016), and Huang et al. (2019), respectively. In general, we find that our values are consistent with other studies. Massive galaxies typically have higher metallicities, while lower-mass systems having lower metallicities. When split into a clumpy and non-clumpy sample, we find that the sample of clumpy SFGs have a lower median metallicity, with an offset of $0.07 \pm 0.02$ dex compared to their non-clumpy counterparts.

While the mass-metallicity relation of Figure 2.5 suggests a metallicity offset between clumpy galaxies and non-clumpy galaxies, we acknowledge that there is a small (but non-negligible) mass offset between the two samples. Our clumpy sample has a lower mean stellar mass compared to the non-clumpy galaxy sample. Although the original sample of targets were chosen so that both the stellar masses and star formation rates are similar, some galaxies were affected by skylines. This resulted in two samples of galaxies with a mean mass offset of 0.1 dex. However, we do not expect the mass difference to affect the observed metallicity offset by much since the mass-metallicity relations suggest only a slight change in metallicities with our given mass offset. For example, given the mass offset of 0.1 dex, the metallicity is expected to change by 0.01-0.02 based on the different mass-metallicity relations.

We further quantify this by fitting a mass-metallicity relation to our data. Using the analytical functions from Zahid et al. (2014), de los Reyes et al. (2015), Guo et al. (2016), and Huang et al. (2019), we find the best fitting mass-metallicity relation for our observed metallicities, by fixing

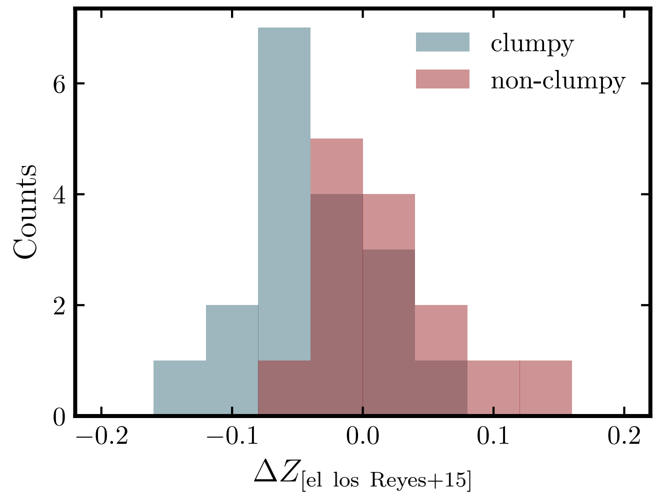


Figure 2.6: The histogram of the differences between the observed metallicity and the re-normalized mass-metallicity relation from de los Reyes et al. (2015) for clumpy and non-clumpy galaxies. Using a KS test, we find the two distributions are statistically different, with a p-value of 0.004.

the slope and only fitting the y-intercept. We opt to do this instead of fitting for both the slope and normalization factor due to our limited sample size and mass range. We find that the mass-metallicity relation from de los Reyes et al. (2015) gives that smallest scatter in the residuals. Figure 2.6 shows the histogram of the differences between the metallicity of clumpy and non-clumpy galaxies to the fitted mass-metallicity relation. A two-sample Kolmogorov-Smirnov (KS) test is performed on the two distributions to determine whether the distributions are statistically similar ($p > 0.05$), or different ($p \leq 0.05$). We obtain a p-value of 0.004, which rejects the null hypothesis of the same distribution.

Finally, we note that the slope of the N2 relation from Pettini and Pagel (2004) is shallower (particularly at higher N2 regime - log(N2) $\sim$ 0.5) relative to other N2 calibrations (e.g., Marino et al. 2013). This could lead to a more defined metallicity offset. To test whether the observed metallicity offset is affected by the choice of line calibration, we perform a similar KS test as above, but using metallicities derived from the steeper line calibration of Marino et al. (2013). Here, We obtain a p-value of 0.03. While the result is less significant than before, the null hypothesis that clumpy galaxies have the metallicities compared to non-clumpy galaxies can still be rejected.

### 2.5.3 The fundamental metallicity relation

It is known that the gas-phase metallicity has a secondary dependence on the star formation rate of the galaxies. The metallicity is found to decrease with increasing SFR at a given stellar mass. The combined stellar mass, SFR and metallicity is suggested to form a fundamental plane, and is known as the fundamental metallicity relation (FMR; Mannucci et al. 2010). We use the following definition from Mannucci et al. (2010) to project the three-dimensional parameter space into two dimensions,

$$\mu = \log(M_*) - 0.32\log(\text{SFR}), \tag{2.3}$$

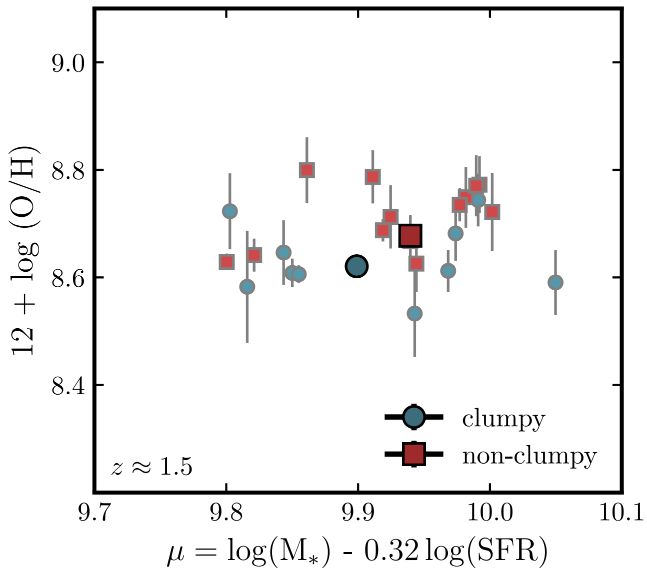


Figure 2.7: The fundamental metallicity relation of clumpy and non-clumpy galaxies in the FMOS-COSMOS sample. The darker markers shows the weighted average, and the error is taken as the weighted standard error. In general, the metallicity of clumpy and non-clumpy galaxies in the FMOS-COSMOS sample are similar to the GNIRS sample. The metallicity offset is measured to be $0.06 \pm 0.02$ dex.

This projection of the FMR is chosen as it minimizes the scatter of the gas metallicity.

The right side of Figure 2.5 shows the projected FMR for our galaxies. As shown in the figure, the distribution of $\mu$ for the clumpy and non-clumpy sample is slightly offset, with a mean difference of $\Delta\mu = 0.04$. When comparing the two distributions of $\mu$, a $p$-value of 0.21 is obtained, suggesting that the distribution of $\mu$ (i.e., the stellar masses and SFRs) for clumpy and non-clumpy galaxies are drawn from the same parent sample. However, if we perform a KS test on the metallicity distributions for clumpy and non-clumpy galaxies, we obtain a $p$-value of 0.004. These results suggest that lower metallicities in clumpy galaxies are not driven by the mass of the galaxies and/or their SFR, and are therefore intrinsic to the clumpy population.

### 2.5.4 Metallicity Difference in FMOS-COSMOS

In this section, we investigate whether there is a metallicity difference between clumpy and non-clumpy galaxies at a higher redshift. We use the [NII]$\lambda$6584 and H$\alpha$ observations from the FMOS-COSMOS survey, which targets galaxies at $z \approx 1.5$.

The selection criteria was described in Section 2.2.2. In essence, we select star-forming galaxies of similar masses, SFRs and colors, with $\log(\text{sSFR [yr}^{-1}]) = -8.70$ and -8.67 for clumpy and non-clumpy galaxies, respectively. Furthermore, only galaxies with a robust detection in H$\alpha$ and [NII]$\lambda$6584 are selected (SNR $> 3$). As these galaxies are located in the COSMOS field, we use the deconvolved dataset from Sok et al. (2022) to identify clumpy and non-clumpy galaxies (see Section 2.2.4). This results in a final sample size of 23 galaxies.

Figure 2.7 shows the relation between metallicities and the projected stellar masses and SFRs for clumpy and non-clumpy galaxies from the FMOS sample. In general, we find that the metallicities of the FMOS sample are comparable to our values. This is also in agreement with other studies of the FMR, which indicated that there is no observed evolution in the FMR with redshift (Sanders et al., 2021). We also find that clumpy galaxies have a lower median metallicity compared to non-clumpy galaxies, with an offset of $0.06 \pm 0.02$ dex. A similar two-sample KS test was performed on the metallicity distribution for clumpy and non-clumpy galaxies, and we obtain a p-value of 0.02. Given that the metallicity offset is observed at two different redshifts for a selected sample of galaxies with similar stellar masses and SFRs, these results suggest that clumpy galaxies intrinsically have lower metallicities compared to non-clumpy galaxies, and this differences cannot be explained by differences in their stellar mass, or SFR alone.

## 2.6 Kinematics of Clumpy Galaxies

Having established that clumpy and non-clumpy galaxies have different metallicities, we now explore if additional signatures between the two samples are observed. Kinematics are interesting as we might expect to see signs of increased disk instability in clumpy galaxies if clumps are formed via VDI. Figure 2.8 shows the relation between the gas fraction $f_{\rm gas}$ and $\sigma_0/v_c$. Here, $v_c$ is the asymptotic rotational velocity, corrected for the inclination of the galaxy (i.e., $v_c = v_a/\sin i$). The inclination $i$ is estimated as,

$$\cos i = \sqrt{\frac{q^2 - q_0^2}{1 - q_0^2}}. \tag{2.4}$$

$q$ is the axial ratio of the galaxy, and $q_0$ is the intrinsic axial ratio for an edge-on galaxy, taken to be $q_0 = 0.2$ (Ryden, 2006). The gas fraction is inferred using the relation between gas and SFR density (e.g., the Kennicutt-Schmidt relation; Kennicutt 1998). The SFR density is derived by assuming that star formation is contained within the effective radius. We obtain the effective radius and $q$ by fitting a Sérsic profile to the *F814W* images using `statmorph` (Rodriguez-Gomez et al., 2019). We further limit the analyses to "kinematically-resolved" galaxies, where $v_c/\delta v_c$ is greater than 1 ($\delta v_c$ is the measurement uncertainty). We also note that some of the derived values $\sigma_0$ are unreliable due to our relatively coarse spectral resolution. In cases where the observed velocity dispersion is within one sigma uncertainty of the instrumental dispersion, we plot the observed dispersion (i.e., $\sigma_{\rm obs} = \sqrt{\sigma_0^2 + \sigma_{\rm inst}^2}$) as the upper limit for the intrinsic dispersion. These are denoted as open markers in the figure.

The dotted horizontal line in Figure 2.8 denotes the cutoff between dispersion- and rotationally-dominated disks, with the former having $\sigma_0/v_c > 1.25$ (e.g., Förster Schreiber et al. 2009). We find that both clumpy and non-clumpy galaxies have a range of kinematics, but can be viewed as rotationally-dominated systems. Given the upper limits in the dispersion, it is also possible for the non-clumpy sample to have lower dispersion compared to clumpy galaxies.

The sloped line in Figure 2.8 is derived from the Toomre's $Q$ parameter (Toomre, 1964). Instability within a gas disk is predicted when $Q$ falls before unity. $Q$ is typically expressed as,

$$Q = \frac{\sigma\kappa}{\pi G \Sigma_{\rm gas}}, \tag{2.5}$$

where $\sigma$ is the local dispersion, $\kappa$ is the epicyclic frequency, and $\Sigma_{\rm gas}$ is the gas density. $\kappa$ is related

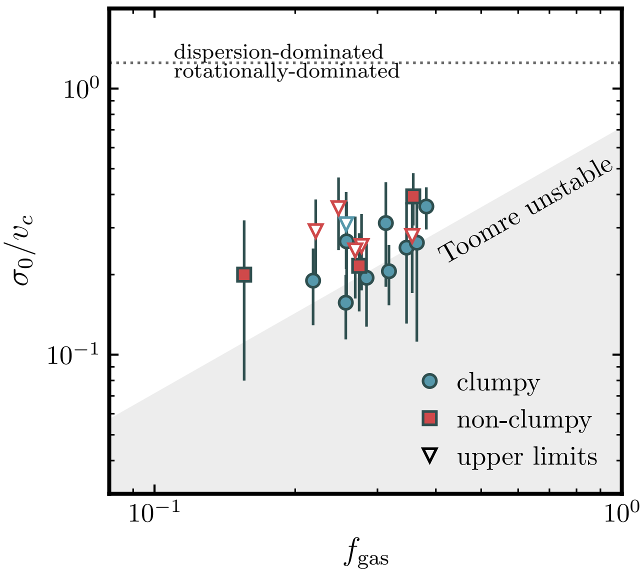


Figure 2.8: The relation between $\sigma_0/v_c$ and $f_{\rm gas}$ for clumpy and non-clumpy galaxies. $f_{\rm gas}$ is derived from the integrated SFR, using the Kennicutt-Schmidt relation. The darker color markers show the mean value, and the error is calculated as the standard error of the mean. The dotted line shows the typical cutoff between rotationally- and dispersion-dominated systems Förster Schreiber et al. (2009). The shaded region shows the parameter space in which the Toomre Q value is less than 1 (i.e., unstable disks). This relation is defined by Eqn. 2.7 (see text for the derivation). Based on the integrated values, clumpy and non-clumpy galaxies have stable disks, with no significant differences between their kinematics.

to the angular velocity $\Omega$ such that $\kappa = a\Omega$, and $a$ ranges from 1 for a Keplerian orbit to $\sqrt{3}$ for a uniform disk (Dekel et al., 2009b). As shown by Genzel et al. (2011), the Toomre parameter can then be reduced to,

$$\begin{aligned} Q &= \frac{\sigma\kappa}{\pi G\Sigma_{\rm gas}} = \frac{\sigma(av_c/R_{\rm disk})}{\pi G\Sigma_{\rm gas}} \\ &= a\Big(\frac{\sigma}{v_c}\Big)\Big(\frac{v_c^2 R_{\rm disk}/G}{\pi R_{\rm disk}^2\Sigma_{\rm gas}}\Big) = a\Big(\frac{\sigma}{v_c}\Big)\Big(\frac{M_{\rm tot}}{M_{\rm gas}}\Big) \\ Q &= af_{\rm gas}^{-1}\frac{\sigma}{v_c}, \end{aligned} \tag{2.6}$$

and $f_{\rm gas}$ is the gas fraction. Assuming a flat rotation with $a = \sqrt{2}$, the following equation is obtained

$$\frac{\sigma}{v_c} = Q\frac{f_{\rm gas}}{\sqrt{2}}. \tag{2.7}$$

Setting $Q = 1$ for the instability criterion gives the sloped relation in Figure 2.8. Lower values for $Q$ indicate unstable disks (i.e., the shaded region). Taking Figure 2.8 at face value, it appears that non-clumpy galaxies might be more kinematically stable than clumpy galaxies, however, given that many non-clumpy galaxies have only upper limits in the dispersion we do not find a strong preference for clumpy galaxies to have lower Toomre values compared to non-clumpy galaxies.

It is not surprising that most galaxies are found to be marginally stable with $Q \sim 1$, if we consider the result within the context of a self-regulating disk. If the gas is accreted smoothly such as there is no injection of turbulence, $Q < 1$ is more readily achieved in the outer region of the disk given the increase in gas density and flatter rotation curve. This naturally leads to the fragmentation and formation of clumps. Several processes can then regulate $Q$ by driving up turbulence. Clump-clump interaction can lead to the transport of mass centrally, and the gain in gravitational energy during this process can increase gas turbulence. Similarly, clumpy accretion of gas is also suggested to drive up turbulence.

It is important to address several caveats in this analysis. We note that the use of $Q$ here only accounts for the gaseous component of the disk. An effective $Q$ that takes into account both the gaseous and stellar component may better predict the stability of the disk. For example, Puech (2010) found that while clumpy galaxies at $z \sim 0.6$ have $Q_{\rm gas}$ that is stable, their $Q_{\rm eff}$ is below unity. Likewise, Eqn. 2.5 describes the local instability within a gas disk. Since we are assuming a gas fraction that is derived from the integrated SFR, using the Kennicutt-Schmidt relation to estimate the gas fraction may be incorrect as clumpy galaxies have localized star formation. Similarly, the kinematics are measured based on the ionized gas, which can be susceptible to processes such as outflows. It may be insightful to do similar analyses in the future at higher resolution with molecular gas observations, as such observations can provide better constraints.

## 2.7 Discussion

### 2.7.1 Gas inflow driving the MZR offset

The simplest explanation for the metallicity offset between galaxies of similar masses and SFRs is with the accretion of metal-poor gas; either induced by mergers, or cosmological gas accretion. Indeed, simulations have shown that the scatter in the mass-metallicity relation can be reproduced as the variations in SFR and metallicity (Torrey et al., 2018). The fluctuation of the gas supply can play a role in these variations, where lower gas inflow leads to the increase of metallicity (De Lucia et al., 2020). Similarly, van Loon et al. (2021) found that gas inflow influences the scatter of the mass-metallicity the most compared to other parameters such as outflows and star formation rate in the `EAGLE` simulation.

Observationally, lower metallicities in high-redshift ($z \approx 3$) galaxies can also be explained through gas accretion. By using a chemical evolution model from Erb (2008), Cresci et al. (2010) found that the relation between their observed metallicities and the inferred gas fraction is better explained as the accretion of in-falling metal-poor gas as opposed to outflows due to AGN, or stellar winds. In this model, both the accretion and outflow rate is assumed to be a constant inflow and outflow fraction compared to the SFR. In the former case, this can be viewed as a continuous process or average of discrete accretion event. Similarly, using the MUSE data from the MEGAFLOW survey, Langan et al. (2023) found that galaxies with inflows are typically found at the lower end of the mass-metallicity relation.

### 2.7.2 Clump formation due to gas accretion

Chemical inhomogeneities in the gas disk have been observed both locally and at high-redshift, with localized, kilo-parsec starbursts reportedly having lower metallicities compared to the surrounding

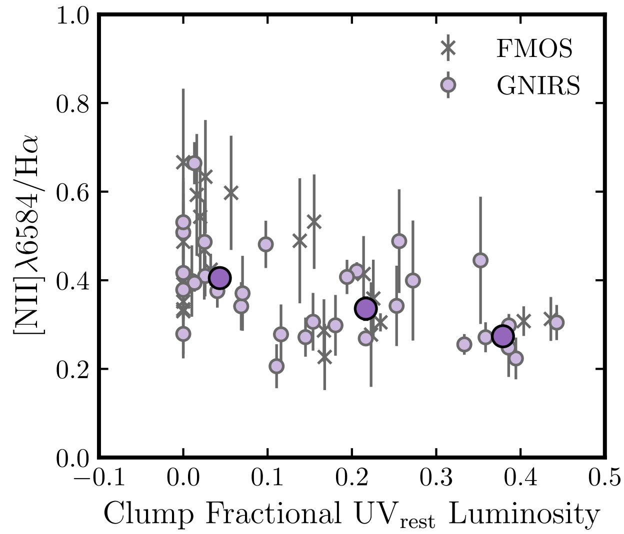


Figure 2.9: The relation between the [NII]$\lambda$6584/H$\alpha$ line ratio and the fractional $UV_{rest}$ contribution for the GNIRS and FMOS samples. The fractional $UV_{rest}$ contribution is defined as the ratio of the total rest-frame UV luminosity of all clumps in a galaxy to the total rest-frame UV luminosity of the galaxy. The darker markers show the mean the mean fractional $UV_{rest}$ contribution and the weighted mean N2 line ratio for the GNIR sample, using discrete bins along the fractional $UV_{rest}$ contribution at between 0-0.15, 0.15-0.3, and 0.3-0.5. Each bin, starting from the lowest bin, has 14, 6, and 6 galaxies. If the fractional $UV_{rest}$ contribution is taken as the size/number of clumps in galaxies, the result suggests that clumpier galaxies are more likely to have lower metallicities.

regions (e.g., Sánchez Almeida et al. 2015). In the context of cosmological gas accretion, lower metallicities in such clumpy galaxies can be explained through the accretion of metal-poor gas. Such gas accretion can either triggers a starburst as the gas compresses through its approach to the disk, or results in a build up of the gas disk, which may subsequently form star-forming clumps due to internal instabilities. Using a zoom-in AMR cosmological simulation, Ceverino et al. (2016) found that the accretion of gas leads to the formation of star-forming clumps in galaxies with $M_* < 10^9\ M_\odot$. In particular, in all the cases where clumps have a metallicity drop compared to the surrounding interstellar medium, the gas mass of the host galaxies steeply increases, and can double prior to clump formation. Most of their observed clumps have a metallicity drop of around 0.3 dex, and are typically dispersed within a few dynamical times. While this offset is larger than what we observe, similar metallicity depressions of $\sim 0.1$ dex have been reported for samples of clumpy star-forming galaxies (e.g., Estrada-Carpenter et al. 2025). Furthermore, as noted in their work, the amplitude of the metallicity offset depends on both the recency and the amount of accreted gas.

### 2.7.3 Clump formation due to mergers

Wet mergers have also been shown to lead to a dilution of metal contents in galaxies, with major mergers resulting in a larger metallicity drop in comparison to minor mergers (Bustamante et al., 2018). The study found that the metallicity dilution can range from approximately 0.05, 0.1, 0.15 dex for minor (mass ratio $< 1/10$), intermediate ($1/10 <$ mass ratio $< 1/3$), and major (mass ratio $> 1/3$) mergers, respectively. Our observed metallicity offset of $0.06 - 0.07$ dex for both the GNIRS and FMOS sample suggests that both minor and intermediate mergers could be the main driver for clump formation. This is broadly consistent with other studies that suggested that clump formation could be formed to minor mergers at low-redshift. In particular, the major merger rate is relatively low at $z \sim 1$ and is therefore not expected to be a major contributor of clump formation, while minor mergers are still relatively common at $z \sim 1$ (e.g., Guo et al. 2015; Shibuya et al. 2016). It should also be noted that the galaxy major merger rate is expected to be higher toward high-$z$ (Romano et al., 2021; Shibuya et al., 2022; Duan et al., 2025). Clumps observed at $z > 4$ (e.g., Shibuya et al. 2016; Ribeiro et al. 2017; Hainline et al. 2024) could therefore be formed due to major mergers. In fact, Nakazato et al. (2024) found that most of their detected clumps at the redshift range of 5.5 and 9.5 were induced by mergers using the FirstLight simulation. We note that most of our galaxies are rotationally-dominated, which may imply that the observed metallicity offset is not driven by major mergers. However, such kinematics signatures may not be a good indicator for major mergers as the galaxies may also be in their late merger stage.

### 2.7.4 What are the properties of the host SFGs?

Within the picture of gas inflow, the accreting gas would lead to the dilution of metallicity, and a burst of star formation. This relation between UV-bright, star-forming clumps and lower metallicities is also observed in our results. Figure 2.9 show the relation between the mean [NII]$\lambda$6584/H$\alpha$ line ratio as a function of the fractional rest-frame UV contribution from all clumps in a galaxy compared to the integrated value of its host galaxy for both the GNIRS and FMOS samples. The fractional UV contribution can be taken as the size and/or number of the clumps, where higher values can indicate that a galaxy that hosts more and/or massive clumps. We find a slight correlation between the [NII]$\lambda$6584/H$\alpha$ line ratio and fractional rest-frame UV contribution. A Spearman correlation test yields a coefficient of -0.47 with a p-value of approximately 0.006. The result indicates that galaxies with a higher fractional UV contribution from clumps tend to have lower metallicities.

In addition, if gas accretion helps build up the disk of the galaxies, leading to violent disk instability, we would expect the galaxies to have a disk-like morphology. The relation between clumpy and disk-like morphologies is reported by Shibuya et al. (2016), who observed that the clumpy fraction is higher for galaxies with lower Sérsic $n$ index (i.e., $n \sim 1$ indicates a disk-like component). Martin et al. (2023) reported no significant differences between the Sérsic index between clumpy and non-clumpy galaxies. We perform a similar analysis on our sample. Figure 2.10 shows the relation between the Sérsic index and the effective radius to the fractional UV contribution. Our analysis indicates that most of our galaxies exhibit disk-like structures. Additionally, we identify a slight correlation between the fractional $UV_{\rm rest}$ contribution of clumps and the Sérsic index of the galaxies. Utilizing the Spearman correlation coefficient test, we obtain a coefficient of -0.44, with a corresponding p-value of 0.01. On the other hand, the fractional $\rm UV_{rest}$ contribution of clumps is correlated with the effective radius of the galaxies, with a Spearman coefficient of 0.78 (p-value $< 0.001$), suggesting that clumpier galaxies are larger in size. This can be explained within the framework of violent disk instability (e.g., see Dekel et al. 2009b). Firstly, gas accretion is anticipated to accumulate in the outer regions of the disk, leading to increased gas density in these areas. With smooth accretion, gas turbulence would remain unaffected. Finally, assuming a flat rotation curve, the angular velocity is expected to fall off as $\Omega \sim r^{-1}$. Together, these factors naturally lead to a decreasing Toomre parameter, $Q \sim \sigma\Omega/\Sigma$. We note that the relation between the size and fractional

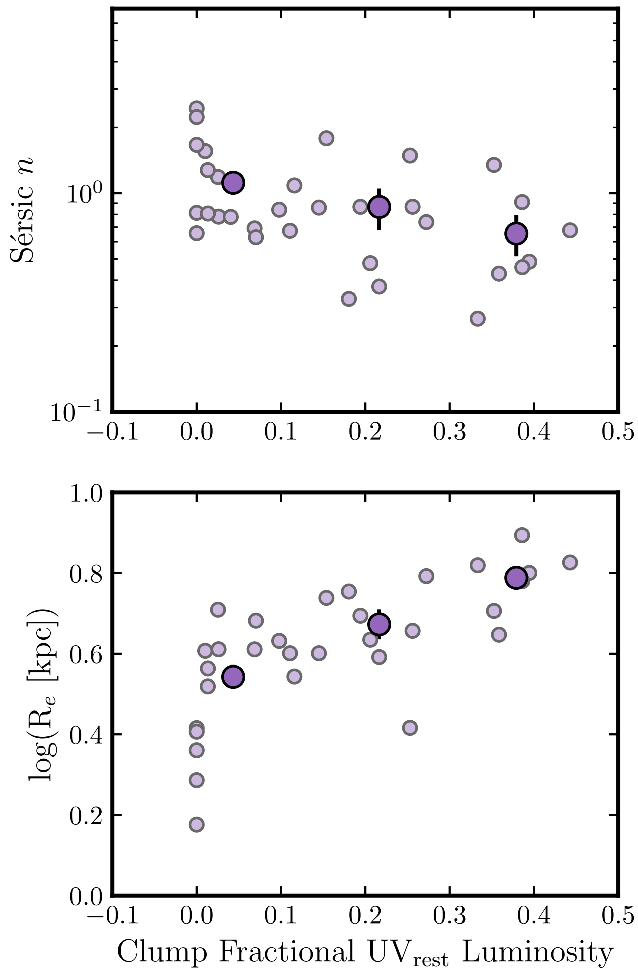


Figure 2.10: Similar to Fig. 2.9, but is now showing the comparison between the Sérsic index (top panel) and $R_e$ (bottom panel) to the fractional $UV_{\rm rest}$ contribution of clumps. The darker markers now denote the mean, and the error here is taken as the standard error. The galaxies in our sample are consistent with being a disky galaxies. Clumpier galaxies are generally bigger in size. However, we note that the relation between the size and clumpiness could be an observational artifact where it is easier to identify clumps that are further extended from the central regions.

$UV_{rest}$ contribution could also be an observational artifact where it is easier to identify clumps that are further extended from the central regions of the galaxies.

## 2.8 Chapter 2 Conclusion

Star-forming clumps in high-redshift SFGs are suggested to form as the result of violent disk instabilities within the gas-rich disk. Such disks are believed to be sustained by the continuous accretion of metal-poor gas from cosmological stream. If such gas accretion event leads to clump formation due to VDI, we expect to find lower metallicities in clumpy galaxies.

In this work, we obtain the [NII]$\lambda$6584/H$\alpha$ line ratio for a sample of 32 SFGs at $z \approx 0.7$ and 23 SFGs at $z \approx 1.5$ in the COSMOS field in order to validate the above scenario by testing whether clumpy galaxies have lower metallicities compared to non-clumpy galaxies. The sample of galaxies consists of both clumpy and non-clumpy galaxies that are selected to have similar stellar masses and star formation rates. The classification of clumpy and non-clumpy is made based on whether the galaxy hosts clumps that account for more than 8% of the total rest-frame UV luminosity of the galaxy. We infer the gas-phase metallicity based on the line ratio of [NII]$\lambda$6584 and H$\alpha$. The main takeaways of the chapter are as follows.

- We measure the mass-metallicity relation and find that it is in agreement with other studies of the relation at similar redshift ranges. Our weighted metallicity average is $8.67 \pm 0.01$ dex.
- Clumpy galaxies have lower metallicities compared to non-clumpy galaxies, with a difference in the metallicity of $0.07 \pm 0.02$ dex.
- Furthermore, we find a similar metallicity offset at a higher redshift by performing similar analyses using data from the FMOS-COSMOS survey. FMOS-COSMOS contains fluxes for both H$\alpha$ and [NII]$\lambda$6584 for galaxies in the COSMOS field at $z \approx 1.5$. We apply the same morphological classification to these galaxies, and observe a metallicity offset of $0.06 \pm 0.02$ dex.
- Both clumpy and non-clumpy galaxies exhibit a wide range of disk kinematics, with $\sigma_0/v_c$ ranging between 0.02 and 0.4. This indicates that the majority of our SFGs are rotationally-dominated.
- We do not find a significant difference in the kinematics of clumpy and non-clumpy galaxies.
- The [NII]$\lambda$6584/H$\alpha$ ratio is negatively correlated with the fractional $UV_{rest}$ luminosity contribution from clumps. Interpreting the fractional $UV_{rest}$ luminosity as the number and/or size of clumps, this result suggests that “clumpier" galaxies tend to have lower metallicities.

Our results indicate a metallicity offset between clumpy and non-clumpy galaxies, as well as a clear relation between metallicity and the number of clumps. The lower metallicities observed in clumpy galaxies suggest that clump formation is driven by the inflow of metal-poor gas, and that the processes that form them may not require significant kinematics instability in the gas disk. However, given the observational uncertainties, performing similar analyses with direct molecular gas observations would be valuable to better constraints for these parameters.

## Acknowledgements

The authors would like to thank the anomynous referee for the constructive feedback. GW gratefully acknowledges support from the National Science Foundation through grant AST-2205189 and from HST program number GO-16300. This chapter is based on observations obtained at the international Gemini Observatory, a program of NSF's NOIRLab, which is managed by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation on behalf of the Gemini Observatory partnership: the National Science Foundation (United States), National Research Council (Canada), Agencia Nacional de Investigación y Desarrollo (Chile), Ministerio de Ciencia, Tecnología e Innovación (Argentina), Ministério da Ciência, Tecnologia, Inovações e Comunicações (Brazil), and Korea Astronomy and Space Science Institute (Republic of Korea).

# Chapter 3

# The Stellar Mass and Age Distributions of Star-Forming Clumps

"Legolas! Two already!"
"I'm on seventeen!"
"Huh? I'll have no pointy-ear outscoring me!"
— Gimli and Legolas, The Lord of the Rings: The Two Towers

---

*A version of the following chapter was initially submitted to the Astrophysical Journal. The co-authors were Adam Muzzin, Vivian Yun Yan Tan, Yoshihisa Asada, Maruša Bradač, Vicente Estrada-Carpenter, Kartheik Iyer, Nicholas S. Martis, Gaël Noirot, Ghassan T. E. Sarrouh, Marcin Sawicki, Chris J. Willott, Sunna Withers, Samantha C. Berek, and Katherine Myers. In this chapter, we investigate the stellar properties of clumps. Specifically, we look at the clump stellar mass distribution as a function of host galaxy mass. We also explore the relationship between the slope of the clump stellar mass function and clump age to provide insights into clump evolution.*

## 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 $\log(M_{g,*}/M_{\odot}) \geq 10$, while 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 masses, 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 in both cases. We explore how different clump formation mechanisms and evolutionary processes affect the shape of the clump mass function. In particular, we find that the cSMF slope is flatter with increasing gas fractions

in younger clump populations ($< 300$ Myr old), suggesting that higher gas availability leads to more massive clumps forming at the time of formation. Alternatively, many high-redshift galaxies in the sample have disturbed morphologies and simulations show that clumps of *ex-situ* origins can flatten the cSMF slope. We also investigate the evolution of clump populations and find that the slope of the cSMF becomes flatter with increasing clump age. Such an evolution may reflect the combined effects of multiple physical processes, including feedback and clump–clump interactions, which remove lower 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.

## 3.1 Introduction

Over the past decades, observations have revealed that 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., Elmegreen et al., 2005; Puech, 2010; Förster Schreiber et al., 2011a,c; Wisnioski et al., 2011; Wuyts et al., 2012, 2013). Particularly, the fraction of clumpy galaxies is found to increase with redshifts (e.g., Murata et al., 2014; Guo et al., 2015; Huertas-Company et al., 2020; Sok et al., 2022; Mercier et al., 2025) but peak at around $z \sim 2$, following the peak of the cosmic star formation density (Shibuya et al., 2016; Sattari et al., 2023; de la Vega et al., 2025). With the recent launch of JWST, there is also a growing number of spectacularly gravitationally lensed sources observed at $z \gtrsim 5$ that host compact stellar structures (e.g., Adamo et al., 2024; Mowla et al., 2024; Fujimoto et al., 2025; Vanzella et al., 2025). While many of these objects are likely to be globular cluster 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 discussed pathways for clump formation at $z \sim 2$ can be broadly divided into *in-situ* and *ex-situ* origins. *In-situ* clumps are often linked to violent disk instability, driven by high gas fractions and turbulence. This process is commonly described in terms of the Toomre instability parameter, $Q$, which governs the susceptibility of a rotating gas disk to fragmentation (e.g., Toomre 1964; Agertz et al. 2009; Dekel et al. 2009a,b; Ceverino et al. 2010; Cacciato et al. 2012). Such conditions are observed in gas-rich, rotationally supported galaxies at high redshift (e.g., Daddi et al. 2010; Tacconi et al. 2010; Wang et al. 2022). The gas-rich nature of these disks is explained as a result of the accretion of gas from cosmological streams (e.g., Kereš et al. 2009; Dekel et al. 2009b). We note there that galaxy interactions and mergers can also cause *in-situ* fragmentation and formation of massive clumps (e.g., Puech 2010; Calabrò et al. 2019; Estrada-Carpenter et al. 2024).

In contrast, *ex-situ* clumps originate as external bound systems and are accreted onto the galaxies. Such clumps are seen in simulations as accreting satellites (e.g., Mandelker et al. 2017) and are inferred observationally (e.g., Ribeiro et al. 2017; Zanella et al. 2019). The key diagnostic for such clumps can include their size, mass, and metallicity. For example, a metallicity offset between clumps and the surrounding disk is expected if clumps are not formed *in-situ*ly. In fact, the low-metallicity clumps reported by Estrada-Carpenter et al. (2025) can be naturally interpreted as candidates for *ex-situ* accretion rather than *in-situ* disk fragmentation. Similar metallicity offsets are also reported in Sok et al. (2025a), where clumpy galaxies are found to have lower metallicities compared to nonclumpy galaxies.

While clump formation has been previously argued to be more efficient near cosmic noon, when high gas accretion rates set the conditions for VDI. it remains unclear what the dominant driver for clump formation is at earlier epochs. Recent morphological studies show that a large fraction of galaxies remain disk-like out to $z \sim 7$, implying that disk fragmentation may still operate efficiently even in the early Universe (e.g., Ferreira et al., 2023; Kartaltepe et al., 2023). In contrast, some clumpy galaxies may instead be ongoing mergers, as indicated by metallicity offsets (e.g., Morishita et al., 2025). Simulations also offer contrasting pictures. Nakazato et al. (2024) found that in the FirstLight simulation, clumps at $5.5 < z < 9$ predominantly arise in merger-driven, with classical disk-driven violent disk instability only becoming important once rotationally supported disks emerge at $z \sim 5$–6. In contrast, hydrodynamical simulations by Mayer et al. (2025) showed that clump formation via gravitational fragmentation of gas-rich disks is already possible at $z > 7$.

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 clumps of *in-situ* and *ex-situ* origins. Detecting clumps and measuring their stellar masses accurately depend both on the spatial resolution and sensitivity of the observations. High-z star-forming clumps in field galaxies are generally observed to be larger and more massive, often estimated to be $10^6$-$10^9$ solar masses (e.g., Kalita et al., 2025a,b), compared to their star-forming region counterparts in the local universe. However, these masses can be overestimated by an order of magnitude due to limited resolution and sensitivity of the observations (e.g., Livermore et al., 2015; Cava et al., 2018; Meštrić et al., 2022; Messa et al., 2024; Claeyssens et al., 2025), with sensitivity having a greater impact on derived clump masses (Dessauges-Zavadsky et al., 2017; Tamburello et al., 2017). Any inferred physics using the measured clump properties must therefore account for these limitation to ensure robust conclusions.

Initial analyses on the redshift evolution of the clump stellar mass distribution at $z < 5$ include the work of Dessauges-Zavadsky and Adamo (2018) and Claeyssens et al. (2025), where both studies found that clump formation is broadly consistent with being *in-situ*. In this chapter, 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 $\sim 5000$ 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 §3.2, we briefly describe the CANUCS dataset, while §3.3 covers our galaxy selection criteria for the analyses. In §3.4, 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 §3.5, as well as the clump properties and the clump stellar mass distribution in §3.6.

Throughout the chapter, all magnitudes are expressed in the AB system (Oke and Gunn, 1983). We adopt a ΛCDM cosmological model of the universe with $\Omega_\lambda = 0.7$, $\Omega_{\rm M} = 0.3$, and a Hubble constant of $H_0 = 70$ km/s/Mpc.

## 3.2 Data

This chapter makes use of deep and extensive photometric datasets from the Canadian NIRISS Unbiased Cluster Survey (CANUCS, Willott et al. 2022). A detailed overview of the first data release is presented by 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, Lotz et al. 2017), MACS J0416.1-2403 ($z = 0.395$, hereafter MACS0416, Lotz et al. 2017), MACS J0417.5-1154 ($z = 0.443$, hereafter MACS0417, Postman et al. 2012), MACS J1149.5+2223 ($z = 0.543$, hereafter MACS1149, Lotz et al. 2017), and MACS J1423.8+2404 ($z = 0.545$, hereafter MACS1423, 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 $3\sigma$ magnitude limits are $rm_{\rm AB} \approx 29.5$ for NIRISS, $m_{\rm AB} \approx 29.5-30.0$ for NIRCam, and $m_{\rm AB} \approx 28.0-30.0$ 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 $3\sigma$ magnitude limits for the medium- and wide-band filters in the NCF fields are $m_{\rm AB} \approx 29.5-30.5$ for NIRCam, and $m_{\rm AB} \approx 27.5-30.0$ for the HST/ACS and HST/WFC3 filters. We refer the reader to 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 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 (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 Gledhill et al. (2024), for MACS0416 in Rihtaršič et al. (2025), MACS1149 in G. Rhitaršič et al. (in prep), and MACS0417 and MACS1423 in G. Desprez et al. (in prep). The lensing models are available on the CANUCS website[1].

[1] `https://niriss.github.io/data.html`

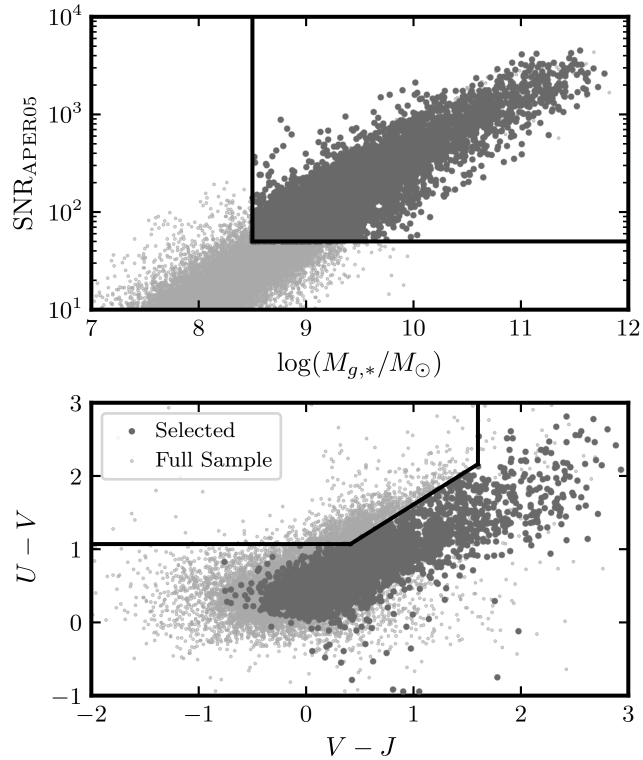


Figure 3.1: We selected galaxies based on their integrated SNR within a 0.5″ aperture and stellar mass. Additionally, we used the $UVJ$ color selection to select star-forming galaxies (see text for details). Note that not all sources falling within each selection region necessarily satisfy all the selection criteria.

## 3.3 Sample Selection

We make use of the imaging and photometric products from Data Release 1 (Sarrouh et al., 2025) to select the galaxy sample for this work. The data release also included photometric redshift (and spectroscopic redshift where available) and derived galaxy property. 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 3.1 shows the main selection criteria of the sample. The selection criteria include cuts into stellar masses, redshift, integrated signal-to-noise ratio ($\mathrm{SNR}_{\mathrm{int}}$, 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 $10^{8.2}$ $\mathrm{M}_{\odot}$ at $z \sim 5$ (Sarrouh et al. in prep.). We adopt a conservative cut at $10^{8.5}$ $\mathrm{M}_{\odot}$, excluding lower-mass galaxies to enable a fair comparison of the clump stellar mass function across redshift. We exclude galaxies that have a $\mathrm{SNR}_{\mathrm{APER05}}$ (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 §3.4.1, 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 $\mathrm{SNR_{APER05}}$, 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 Antwi-Danso et al. (2023), and is defined by the following color cuts,

$$\begin{aligned} &(U-V) < 1.07 \\ &(V-J) > 1.6 \\ &(U-V) < (V-J) \times 0.92 + 0.69 \end{aligned} \tag{3.1}$$

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. 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., 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., Noirot et al. 2022; 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 $\mu$ where applicable (see §3.4.2). 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 (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 4300.

## 3.4 Methodology

This section describes the methodology for constructing resolved stellar mass density and rest-frame UV ($\lambda_{\mathrm{rest}} \approx 3600\,\mathrm{\AA}$) 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 $U_{\mathrm{rest}}$ surface brightness maps. We adopt the recalibrated $U$ filter from Maíz Apellániz (2006), which is based on the Johnson system, to retrieve $U_{\mathrm{rest}}$. We also discuss how clumps are identified within these $U_{\mathrm{rest}}$ brightness maps. Galaxies are first classified as clumpy and nonclumpy based on their $U_{\mathrm{rest}}$ 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.

### 3.4.1 Spatially-resolved SED fitting

**Voronoi Binning**

Spatially-resolved SED fitting provides better estimates of stellar masses compared to unresolved fitting (e.g., Sorba and 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. There exist several binning techniques, including VORBIN, a Voronoi tessellation technique (Cappellari and Copin, 2003), and PIXEDFIT (Abdurro'uf et al., 2021). The former method allows for compact bins of similar SNR and preserve some spatial information, while PIXEDFIT allows for SNR constraints in multiple different filters and performs binning based on similarity in the SEDs.

In this chapter, the binning step is done based on Voronoi tessellation using VORBIN. The code consists of two main steps; a bin accretion step and a regularization step to generate the Voronoi bins. We modify the bin accretion step to also take into account the shape of the SED, making this approach more similar to that of Abdurro'uf et al. (2021). In practice, an unbinned pixel is only accreted to a bin if its SED shape is similar to the mean SED shape of the pixels already in the bin. This is quantified by the cosine similarity score of the SEDs. The cosine similarity score measures the similarity between two vectors by computing the cosine of the angle between them in vector space. Specifically, we calculate the cosine similarity score between the normalized SED of the unbinned pixel and that of the average SED in the current bin. The cosine similarity score ranges from -1 to 1, where a value closer to one means the vectors are very similar. A minimum cosine similarity score of 0.95 is required for bin accretion.

We bin pixels together until we acquire a minimum SNR for each bin based on the F444W photometry. The minimum SNR per Voronoi bin is defined by the integrated SNR of the galaxies in the F444W image (i.e., $\mathrm{SNR}_{\mathrm{APER05}}$ as measured within a aperture diameter of $0.5''$). We set the minimum SNR to be 5 and 10 for galaxies with $\mathrm{SNR}_{\mathrm{APER05}} < 500$ and $\mathrm{SNR}_{\mathrm{APER05}} \geq 500$, 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.

**Inferring physical properties**

We determine the flux for each bin by summing the pixel values within that bin. Similar to Han and 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 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 (Iyer et al., 2019). Dense Basis employs a Bayesian nested sampler 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., Iyer and Gawiser 2017; Leja et al. 2019). We adopt a Chabrier initial mass function (Chabrier, 2003) and Calzetti attenuation law (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 the redshift of the galaxy hosting the clump where 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 (Brammer et al., 2008), using the recalibrated $U$ filter from Maíz Apellániz (2006). To this end, we fix the redshift of each SED to the redshift of its parent galaxy.

Voronoi binning improves the signal-to-noise ratio of our SED fits, but it also degrades the spatial resolution of the resulting maps, particularly in the low S/N outer regions of galaxies where bins become large. This can smear out small-scale structure such as clumps in the rest-frame UV brightness maps. To recover high spatial resolution, we redistribute each SED-derived quantity within a bin using the JWST images. For each pixel $i$ in a bin, we assign

$$Q'_i = Q_{\rm bin} \frac{F_i}{F_{\rm bin}}, \tag{3.2}$$

where $Q_{\rm bin}$ is the SED-derived quantity for the bin (e.g., stellar mass or rest-frame UV luminosity), $F_i$ is the flux in pixel $i$, and $F_{\rm bin}$ is the total flux in the bin. We use the F444W image to redistribute stellar mass, since it closely traces stellar mass, and the appropriate rest-frame UV image to redistribute UV luminosity. Because all images are PSF-matched to F444W, this procedure preserves the total quantity in each bin while restoring spatial information at the native $\sim 0.16''$ resolution.

### 3.4.2 Identifying clumps in the rest-frame UV maps

We identify clumps in galaxies based on the derived rest-frame UV images. This approach ensures all galaxies are analyzed at a common rest-frame wavelength, and avoid redshift- and band-dependent selection effects that would arise from using real observed images. To identify clumps, we first select clumpy galaxy candidates using the normalized light profile method of Wuyts et al. (2012). This method only identifies broad regions of enhanced rest-frame UV emission. We then apply a clump finding algorithm to find localized clumps in these clumpy candidates. We note that the selection based on the rest-frame UV image therefore primarily targets UV-bright, actively star-forming clumps and biases against clumps with low star formation rates or high attenuation.

The normalized light profile is constructed based on the rest-frame UV surface brightness maps (see Wuyts et al., 2012; Sok et al., 2022, 2025a). We refer the reader to these papers for more details, however we briefly describe it here. Galaxies with a smooth disk will have a normalized light profile that is brighter near its galactic center and tapers off with radius. As star-forming clumps

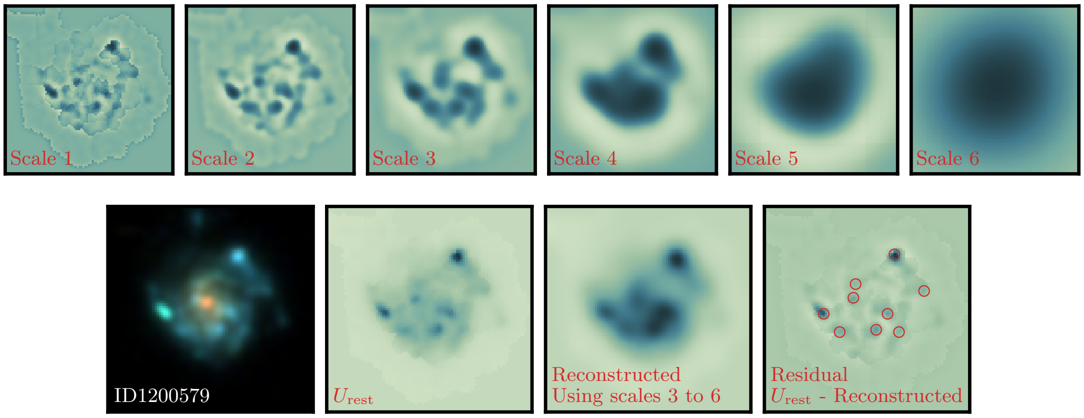


Figure 3.2: An illustration showing how clumps are identified with starlet wavelet transform. The rest-frame $U$ image is first decomposed into multiple spatial scales. The top row shows the resulting wavelet scales. A smooth model of the galaxy is then reconstructed by combining only the large-scale 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 $3.6'' \times 3.6''$.

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, 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., Guo et al., 2018; Sok et al., 2022; 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 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 (Starck et al., 2007). The starlet transform has been implemented in the python package SCARLET (Melchior et al., 2018).

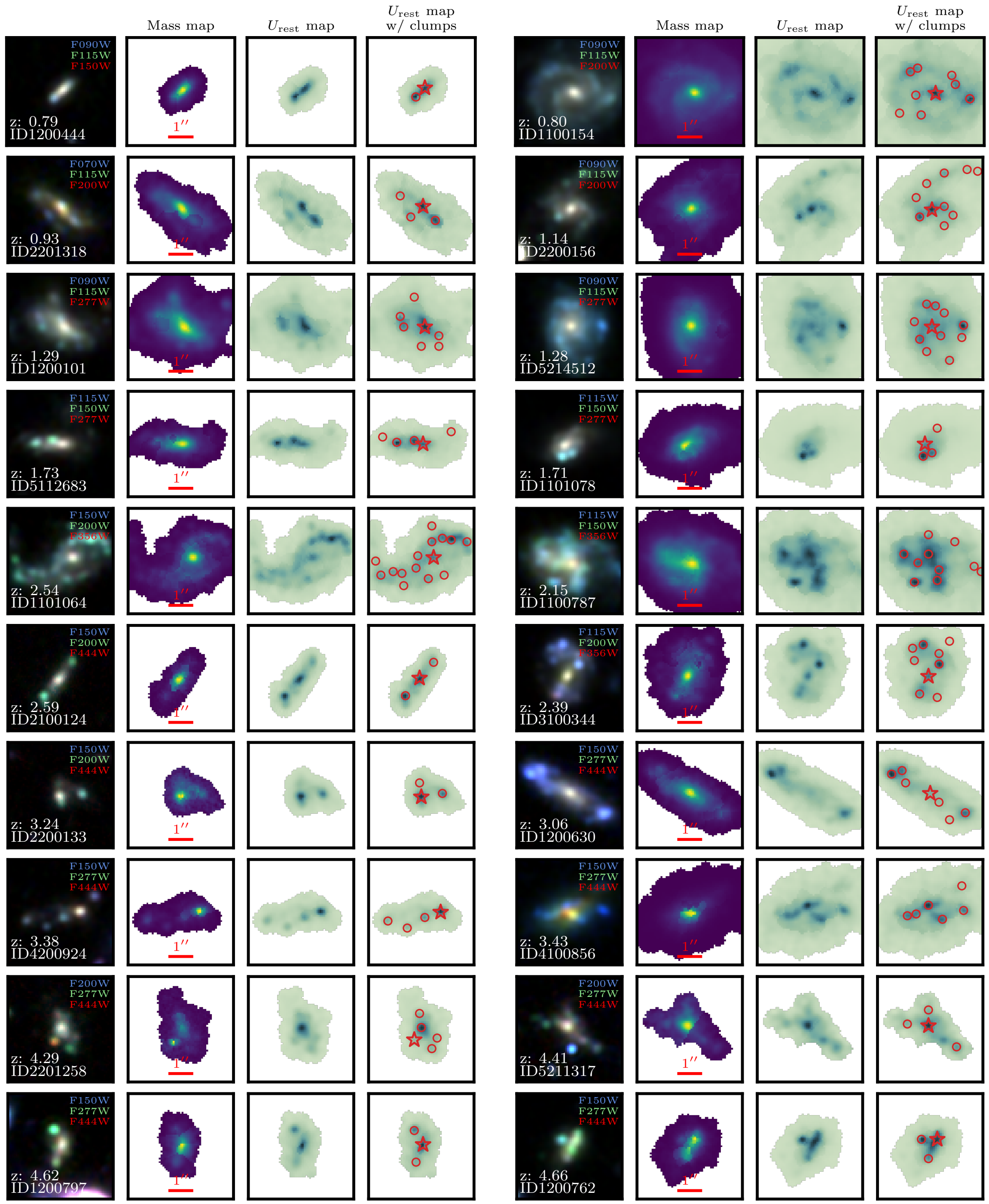


Figure 3.3: Examples of clumpy galaxies in the CANUCS fields. These galaxies are selected based on their clumpy morphology to illustrate our clump detection algorithm, and are ordered by redshift, with less massive galaxies ($\log(M_{g,*}/M_{\odot}) \lesssim 10$) shown on the left and more massive galaxies ($\log(M_{g,*}/M_{\odot}) \gtrsim 10$) on the right. In each panel, we show the composite RGB image, the stellar mass maps, the rest-frame UV map, and the same map superimposed with detected clumps (shown as the circle markers) and bulges (shown as the star markers).

Figure 3.2 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 $5\sigma$ and minimum separation of three pixels. To mitigate spurious detection, we further require that each detected clump has a flux above $3\sigma$ within a $0.2''$-diameter aperture in the rest-frame UV.

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 3.3 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. Here, we use a lower threshold as compared to the 8% in Sok et al. (2022) to select fainter clumps. Only clumps detected in clumpy galaxies are considered for the remainder of the chapter. This gives a total of $\sim$1400 clumpy galaxies with 3800 clumps.

#### 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 3.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 §3.4.2 when creating a smooth image of the galaxies), (2) modeling the disk by fitting a Sérsic profile, and (3) estimating 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 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 $\sim$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 $\mu$.

We note that clump masses can be alternatively retrieved by fitting the SED to photometry

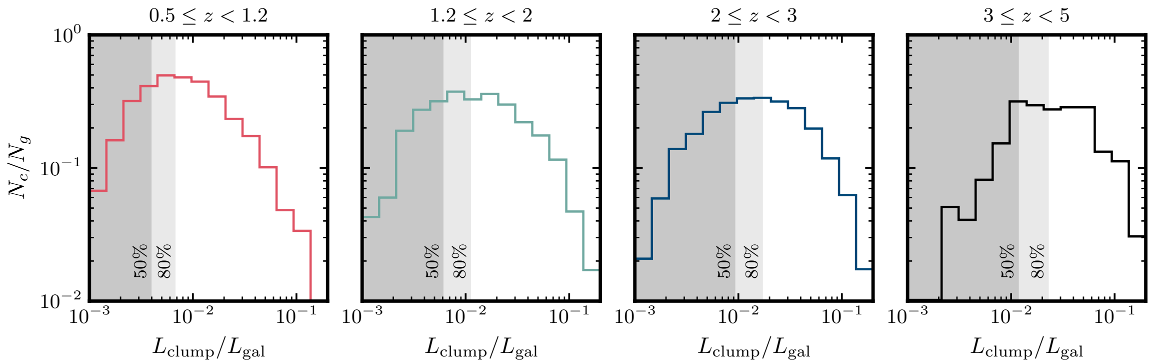


Figure 3.4: The fractional luminosity function of clumps as a function of redshift. In general, we find that the distribution is shifted toward higher $L_{\rm clump}/L_{\rm gal}$ values at higher redshifts as fainter clumps become harder to detect. The gray regions denote the 50% and 80% completeness limit determined by simulating clumpy galaxies and determining how many clumps are retrieved back.

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 photometry on images using the same background subtraction procedure. Here, magnification is accounted by dividing each flux by the corresponding $\mu$. We find that clump masses from the two approaches are consistent, with a scatter of ∼0.3 dex.

**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 3.4, we examine 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 $L_{\rm clump}/L_{\rm gal}$ 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 both clumpy and nonclumpy 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

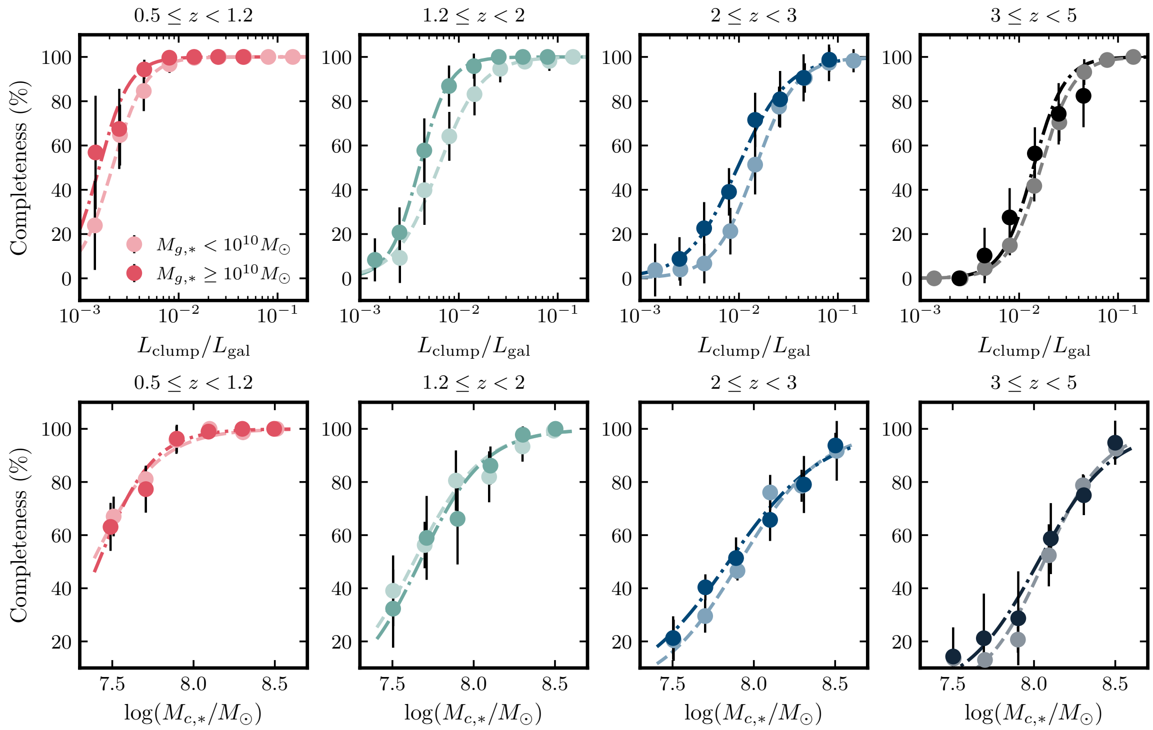


Figure 3.5: The completeness limit as a function of host galaxy mass and redshift bins. In the top panels, we show the completeness as a function of the fractional luminosity of clumps. The bottom panels show the completeness as a function of clump stellar mass. We determine the completeness by injecting 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.

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. 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 $\sim$800 injected clumps.

The top panels of Figure 3.5 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 3.5 show the completeness as a function of clump mass. We model the completeness as a sigmoid function, described as,

$$C(M_{c,*}) = \frac{100}{1 + e^{-k(M_{c,*} - M_o)}}, \tag{3.3}$$

where $k$ and $M_o$ are model 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.

## 3.5 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 $f_{\rm clumpy}$ as the fraction of star-forming galaxies that are clumpy. Figure 3.6 shows the evolution of the fraction of clumpy galaxies at different redshift bins, and the error bars represent the $1\sigma$ 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 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 $\log(L_{\rm clump}/L_{\rm gal}) > -2$ (see §3.4.2) 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 $f_{\rm clumpy}$ is then calculated following a similar equation in Guo et al. (2015),

$$f^{\rm corr}_{\rm clumpy} = f^{\rm old}_{\rm clumpy} + \big(\frac{1}{X} - 1\big) f^{\rm old}_{\rm clumpy}. \tag{3.4}$$

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 $\log(L_{\rm clump}/L_{\rm gal}) > -2$, 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 3.6. Shibuya et al. (2016) further noted that the clumpy fraction peaks near cosmic noon and declines at both earlier and later times. They suggested that the clumpy fraction broadly traces the evolution of the cosmic star formation rate density (SFRD, Madau and Dickinson 2014). This redshift evolution is also observed by 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 trend from these studies by fitting our uncorrected clumpy fraction of galaxies with $\log(M_{g,*}/M_\odot) \geq 10$ using the empirical prescription from Madau and Dickinson (2014),

$$f_{\text{clumpy}} = a\frac{(1+z)^b}{1+[(1+z)/c]^d}, \tag{3.5}$$

where we fix the parameters $b$, $c$ and $d$ to the values reported in Madau and Dickinson (2014), and only refit for $a$ to renormalize the best-fit curve. This is plotted as the dashed gray line in Figure 3.6. While the observed clumpy fraction seem to broadly follow the SFRD redshift evolution, the true clumpy fraction drops more slowly at $z > 2$ due to incompleteness. Further, studies of lower-redshift clumpy galaxies suggest a steep decline in the clumpy fraction at $z < 1$ (Murata et al., 2014; Adams et al., 2022).

Across all redshifts, we find that massive galaxies are more likely to be clumpy than low mass galaxies. While previous studies (e.g., Sok et al. 2022; 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 Murata et al. (2014) and 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., Murata et al. 2014; Sok et al. 2022).

## 3.6 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 §3.4.2, these analyses are restricted to clumps and do not include bulge-like components detected in galaxies.

### 3.6.1 Clump-Host Relations

**Number of clumps per galaxy**

Figure 3.7 shows the number of clumps found in galaxies as a function of the host galaxy mass and star formation rate. The histograms on the top and the right show the normalized distributions of galaxies with different clump counts. We find that the number of clumps is generally higher for more massive galaxies at fixed redshift, similar to what was found in 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 can be seen more clearly in the histogram on the lower-right panel, where it shows the SFR distributions, normalized by the number of galaxies and color-coded by the number of clumps per galaxy. In general, the distributions seem to peak at the same star formation rate for all clump counts. Above $z > 3$,

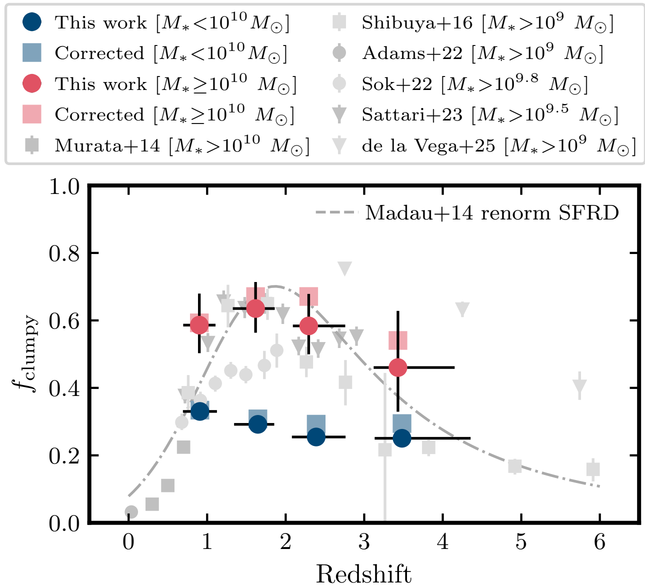


Figure 3.6: The top panel shows the fraction of clumpy star-forming galaxies with $\log(M_{g,*}/M_\odot) >$ 8.5. The bottom panel shows the fraction of clumpy star-forming galaxies in two different galaxy mass bins, where the red markers denote galaxies with $\log(M_{g,*}/M_\odot) > 10$ and the blue 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 line shows a fit to our measured clumpy fraction for galaxies with $\log(M_{g,*}/M_\odot) > 10$, using the functional form of the star-formation rate density equation from Madau and Dickinson (2014).

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.

**Clump masses**

Figure 3.8 shows the distribution of 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 $10^{12}$ $M_\odot$. The blue marker do not trace the peak of the contour distributions since they represent only the most massive clump in each galaxy, whereas the contours include the full clump population. Similar trends have been reported in previous studies; for example, Claeyssens et al. (2025) also found a strong positive correlation between the maximum clump mass and the host galaxy mass (reproduced in Figure 3.8). 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 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 corelation, 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

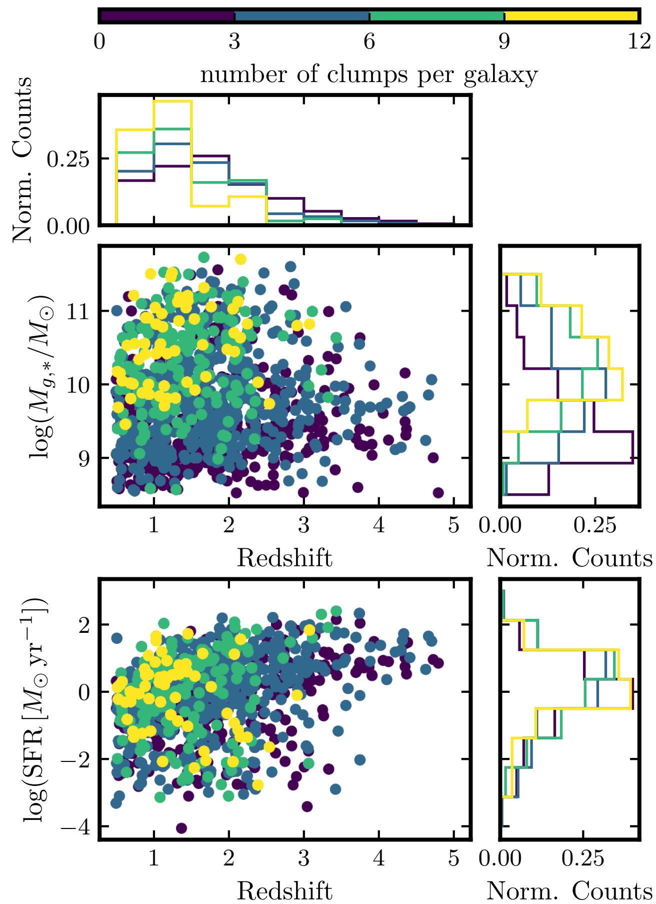


Figure 3.7: Number of clumps in clumpy galaxies as a function of their host galaxy mass and SFR. The histograms show the distributions, which are normalized by the number of galaxies and color-coded by the number of clumps they have. Massive galaxies tend to host more clumps relative to less massive galaxies, however we find no preference with the total SFR.

of Figure 3.8, 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 are less common in massive galaxies. This becomes clearer in §3.6.2 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 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 3.5). The next section shows that the clump mass distribution across galaxy masses are still different even with a completeness correction.

### 3.6.2 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., Krumholz et al., 2019; Adamo et al., 2020),

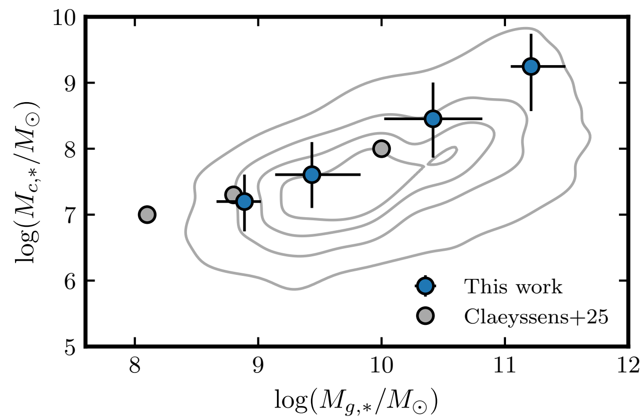


Figure 3.8: The relationship between clump masses and their host galaxy masses. The contours show the full distribution of the clump sample. To further investigate the relationship between clump mass and host galaxy mass, we select the most massive clump in each galaxy, and compute the median mass of 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 Claeyssens et al. (2025), who performed analyses on a compilation of clumps found in lensed sources between $0.6 < z < 5$.

which is expected for structures formed via hierarchical, scale-free fragmentation (e.g., Hopkins 2013; 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, we define the fiducial redshift bin between 0.5, 1.2, 2, 3 and 5 to trace the evolution of the cSMF across cosmic time. The clump mass distribution is 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 note that at lower clump masses, we may not detected all clumps, which is shown in § 3.4.2. This also varies as a function of redshift, where clumps are harder to detect. To account for incompleteness, we correct the number of detected clumps per clump mass bin by multiplying by the reciprocal of the completeness fraction, as described in § 3.4.2. Our sample reaches an 80% completeness limit for $\log(M_{c,*}/M_\odot) \sim 7.7 - 8.3$ between our lowest and highest bins. We therefore apply an incompleteness correction to extend the analyses to lower masses. This correction is only applied to clump masses that are above our 50% completeness limit, as opposed to lower masses where the incompleteness correction is more uncertain.

#### 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 3.9 presents the resulting cSMFs across different redshift and galaxy mass bins. The histograms show the observed (uncorrected) cSMFs, while the solid markers denote the mass-corrected distributions. A power-law relation is fitted to the corrected cSMF using the

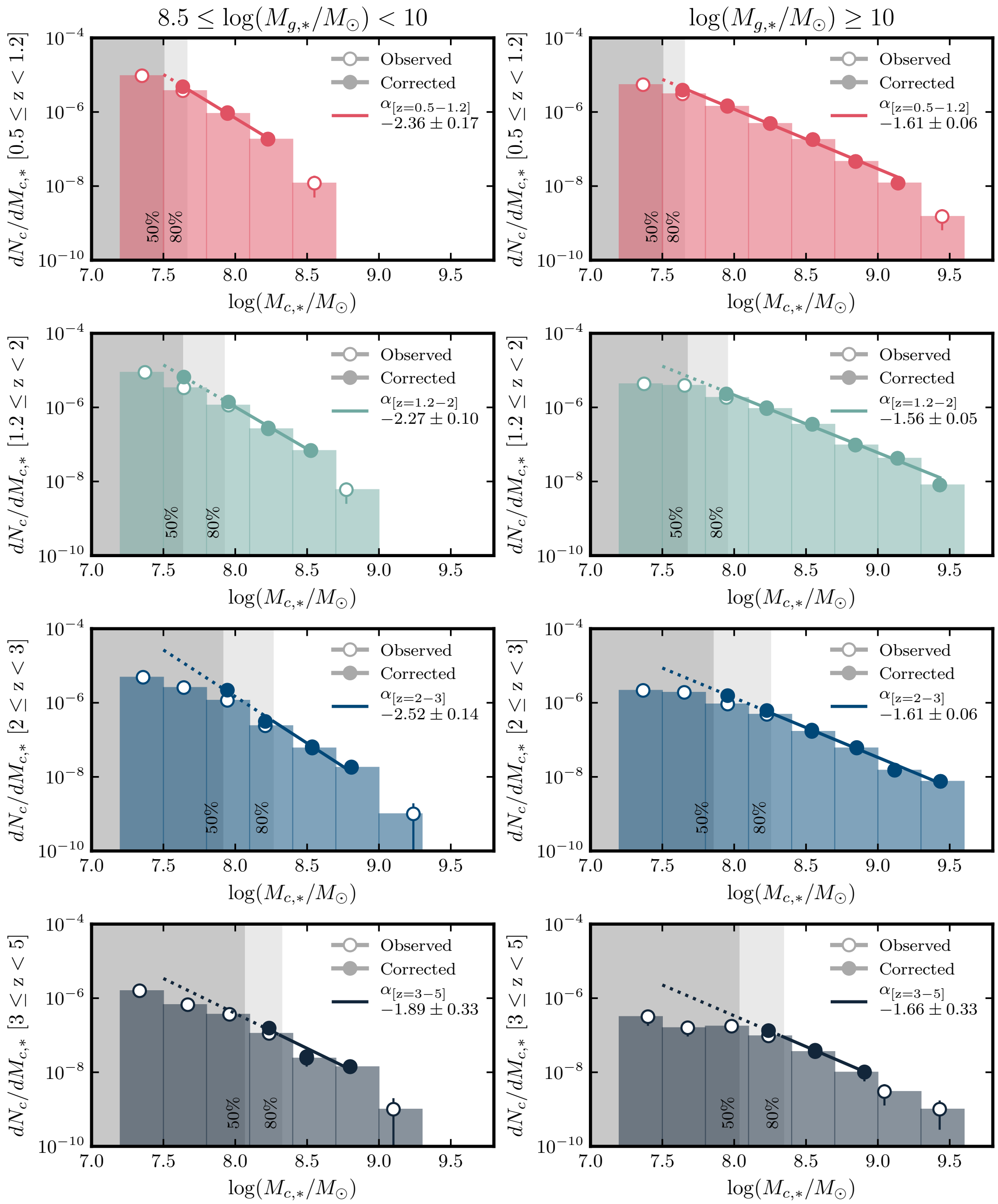


Figure 3.9: The clump stellar mass function (cSMF) in relation to redshifts and galaxy masses. Open markers denote the observed cSMF, while solid markers show the corrected cSMF, which has a correction factor based on our clump mass completeness limit. The darker and lighter gray regions denote where 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.

following equation,

$$\log(dN_c/dM_{c,*}) = \alpha \log(M_{c,*}) - m, \quad (3.6)$$

where $\alpha$ is the slope and $m$ is the normalization constant. The best-fit power law is plotted in the figure as a dashed line, with the solid segment indicating the mass range above the 80% completeness limit. Only the clump mass bins containing more than five clumps and that are above the 50% completeness limit are included in fitting the stellar mass function to ensure robustness.

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 $\alpha_{[\mathrm{high-mass}]} \sim -1.6$ compared to $\alpha_{[\mathrm{low-mass}]} \sim -2.3$. This result is consistent with §3.6.1, 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 3.1. 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.

**Combined cSMF**

Figure 3.10 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., Dessauges-Zavadsky and Adamo 2018; Kalita et al. 2025a,b). 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., Fisher et al. 2017). Using the VELA simulations, Huertas-Company et al. (2020) reported a shallower slope of $\alpha = -1.5$. However, they found a flatter slope of $\lesssim -1$ for clumps observed 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 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 3.1, and Figure 3.11 shows the redshift evolution of the slopes for the different galaxy mass bins. When aggregating the entire clump sample, we

Table 3.1: The slope of the clump stellar mass function as a function of redshift and galaxy mass, with and without a clump mass incompleteness correction. Corrected slopes account for incompleteness down to the 50% limit. Uncorrected slopes are measured using only data points above the 80% limit, with no correction applied.

| Galaxy Mass Bin | Correction | $0.5 < z < 1.2$ | $1.2 < z < 2$ | $2 < z < 3$ | $3 < z < 5$ |
|---|---|---|---|---|---|
| $8.5 \leq \log(M_{g,*}/M_{\odot}) \leq 10$ | Corrected | $-2.36 \pm 0.17$ | $-2.27 \pm 0.10$ | $-2.52 \pm 0.14$ | $-1.89 \pm 0.33$ |
| | Uncorrected | $-2.47 \pm 0.37$ | $-2.18 \pm 0.19$ | $-1.91 \pm 0.65$ | $-0.77 \pm 0.77$ |
| $\log(M_{g,*}/M_{\odot}) > 10$ | Corrected | $-1.61 \pm 0.06$ | $-1.56 \pm 0.05$ | $-1.61 \pm 0.06$ | $-1.66 \pm 0.33$ |
| | Uncorrected | $-1.66 \pm 0.08$ | $-1.58 \pm 0.06$ | $-1.57 \pm 0.13$ | $-1.62 \pm 0.67$ |
| Combined | Corrected | $-1.87 \pm 0.05$ | $-1.70 \pm 0.05$ | $-1.88 \pm 0.05$ | $-1.81 \pm 0.24$ |
| | Uncorrected | $-1.93 \pm 0.12$ | $-1.69 \pm 0.06$ | $-1.67 \pm 0.08$ | $-1.57 \pm 0.24$ |

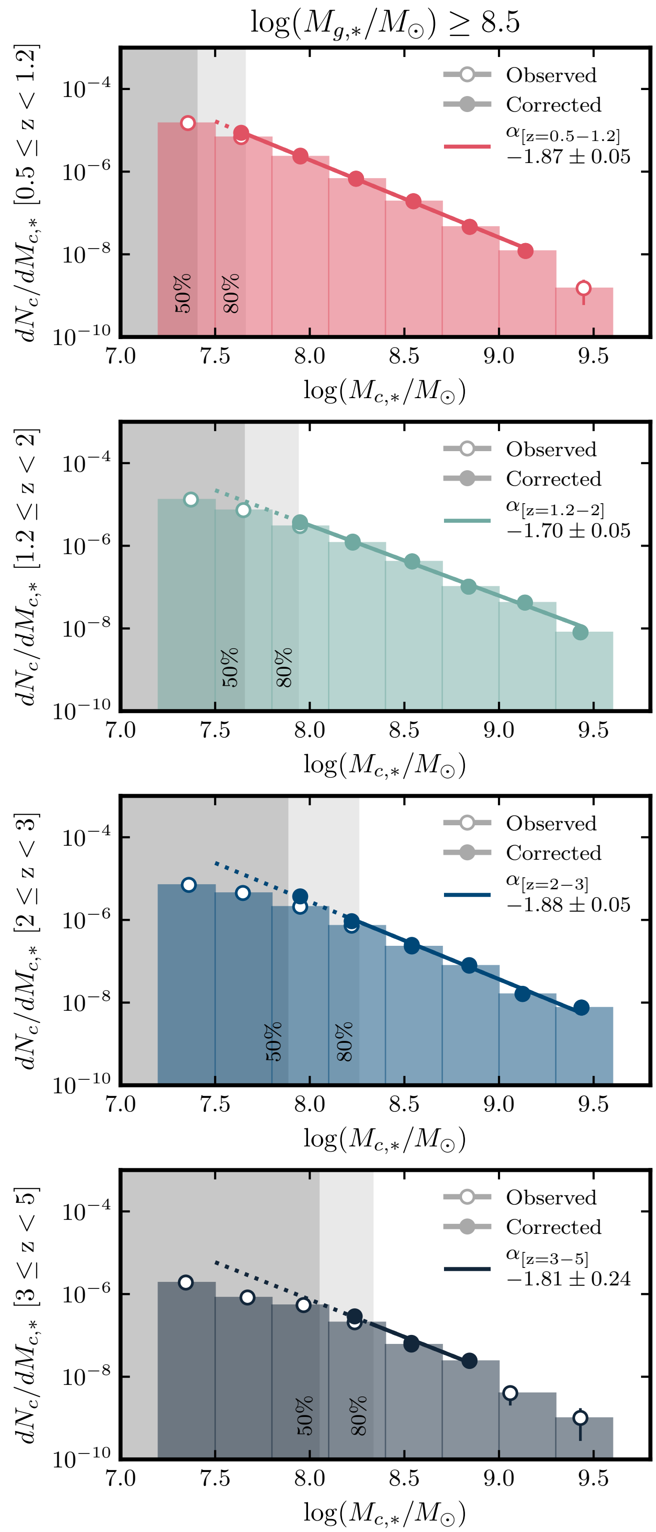


Figure 3.10: Similar to Fig. 3.9 but for the combined cSMF. Before combining, a correction factor is applied to each cSMF to account for any completeness differences between the galaxy mass bins. The darker and lighter gray region denote where the mass completeness is 50% and 80% complete, respective.

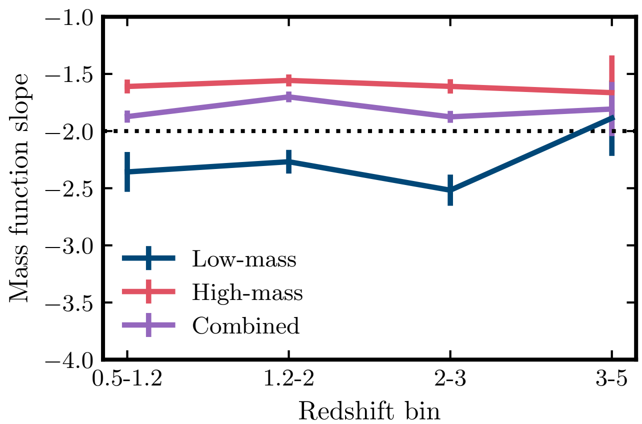


Figure 3.11: The redshift evolution of the cSMF based on different galaxy mass bins.

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 $\alpha_{[\mathrm{high-mass}]} \sim -1.6$ compared to $\alpha_{[\mathrm{low-mass}]} \sim -2.3$. 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.

## 3.7 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 Dessauges-Zavadsky and Adamo (2018); Claeyssens et al. (2025). In the former case, Dessauges-Zavadsky and Adamo (2018) found a slope of -1.7 for their sample at $1 < z < 3.5$, while Claeyssens 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 $0.7 < z < 5.5$. These results are often interpreted as being consistent with a scenario where clumps form predominantly *in-situ*, with broadly similar formation conditions across redshifts.

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 $\alpha_{[\mathrm{high-mass}]} \sim -1.6$ while galaxies with masses of $8.5 < \log(M_{g,*}/M_\odot) < 10$ have steeper slopes of $\alpha_{[\mathrm{low-mass}]} \sim -2.3$, with uncertainties of $\pm 0.1$ dex. We find that there is little redshift evolution in the slopes for both mass bins. Similarly, 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 possible 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. We then examine how clump evolution can flatten the cSMFs.

### 3.7.1 Clump formation

The prevalent theory for clump formation is through violent disk instability, with disk gas mass (and subsequently gas fraction) being a key parameter for disk instabilities (e.g., Dekel et al. 2009b; Ceverino et al. 2010). Several simulations suggest that varying gas fractions play a role in shaping the cSMF. For example, Renaud et al. (2024) found higher gas fractions lead to flatter cSMF slopes due to the formation of massive clumps with increasing gas fraction; a trend also observed in Fensch and Bournaud (2021).

We further explore this relation using our observed slopes. We first limit the discussion to galaxies with 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 stellar mass and SFR, following the relation described in Wang et al. (2022). Figure 3.12 therefore shows the relationship between increasing gas fraction and the cSMF slope, as well as values reported in 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 Renaud et al. (2024), we find little correlation between gas fraction and the cSMF slope.

However, Fensch and 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. Figure 3.12 shows the cSMF slopes 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.

The result suggests that gas fraction can flatten the slope of the cSMF, likely by enabling the formation of more massive clumps. However, gas fraction can not explain the steeper slopes found in low-mass galaxies in our study. In particular, Wang et al. (2022) found increasing molecular gas fraction with decreasing stellar masses, albeit their lowest mass bin is at $\log(M_{g,*}/M_{\odot}) \sim 10.2$. If this trend holds at lower galaxy masses, it would suggest flatter slopes for our lower mass bin of $8.5 < \log(M_{g,*}/M_{\odot}) < 10$, which are not observed. This discrepancy likely indicates that despite having higher molecular gas fraction, the total gas mass in lower mass galaxies may be smaller and subsequently limit their ability to form massive clumps compared to more massive galaxies. This is suggested from Toomre analyses, where the scale for fragmentation within a differential rotating

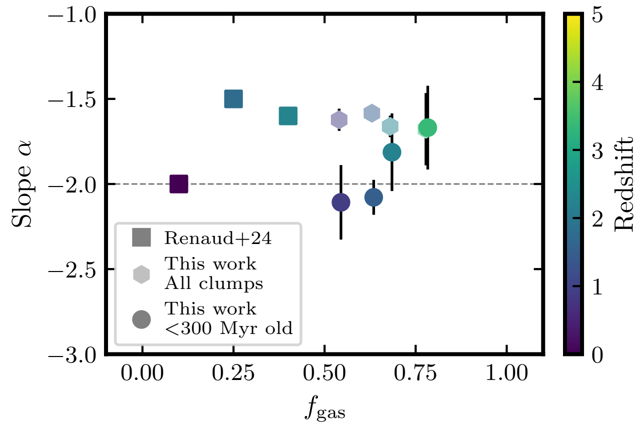


Figure 3.12: The relationship between the cSMF slope and gas fraction. The data from Renaud et al. (2024) represents simulations of isolated galaxies at varying $f_{\rm gas}$, and are analogous of the Milky-Way at different redshifts, with mass of $\sim 10^{10.7}$ solar masses. In comparison, we plot the slope of cSMF for galaxies between $\log(M_{g,*}/M_{\odot}) \geq 10$ for the different redshift bins and their corresponding $f_{\rm gas}$, derived using the relation from Wang et al. (2022). The lightened circle markers denote the cSMF slopes derived using all clumps, while the darker circle markers only account for clumps that are <300 Myr old. The darker circle markers have been shifted by 0.005 in $f_{\rm gas}$ for visual clarity.

disk is given as $\lambda_c \sim G\Sigma/\Omega^2$ (e.g., Binney and Tremaine, 2008, Chapter 6.2.3). Here, $G$ is the gravitational constant, $\Sigma$ is the gas density, and $\Omega$ is the angular circular velocity. As shown in Dekel et al. (2009b), this reduces to obtain the maximum scale for clumps, with $\lambda_c \equiv R_c \propto \delta R_d$, and the mass of clumps as $M_c \propto \delta^2 M_d$. Here, $\delta$ represents the disk mass fraction relative to the total mass, $R_d$ and $M_d$ are the radius and mass of the gas disk, respectively, suggesting that massive clumps are likely to be hosted in massive disks.

We note some caveats for this discussion. The simulated galaxies in 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 $f_{\rm gas}$ also span between $0.1-0.4$. Our galaxies span between $0.5 < z < 5$, and are estimated to be more gas-rich, with $0.5 < f_{\rm gas} < 0.7$. While our galaxies are different compared to those in Renaud et al. (2024), we emphasize that this comparison aims to investigate how the slope of the cSMF might change for galaxies of similar masses but with increasing $f_{\rm gas}$. While the result suggests a trend between increasing gas fractions and flatter cSMF slope for massive galaxies, this relationship cannot account for the differences in the slope across galaxy mass bins. This points to other factors that can affect the shape of the clump mass distribution.

#### Can *ex-situ* clumps affect the cSMF?

In this section, we now consider how clumps that are of *ex-situ* origins can 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., Mandelker et al. 2014, 2017). Observational studies, including Ribeiro et al. (2017); 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.

We note that while we find a slight redshift evolution in the slopes of the cSMF for galaxies with

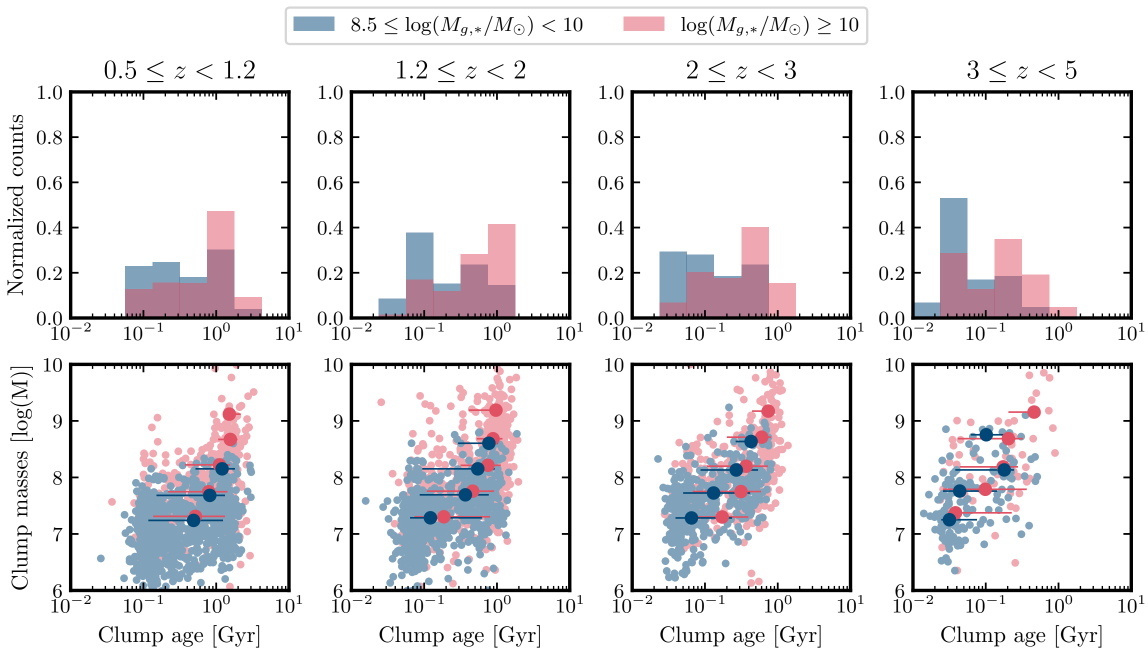


Figure 3.13: The top panel shows the normalized distribution of clump ages, while the bottom panels show the distribution of clump ages with respect to their masses. The error bars denote the 16th, 50th and 84th percentile.

$8.5 < \log(M_{g,*}/M_{\odot}) < 10$, this should be taken cautiously as the slopes have larger uncertainties at higher redshifts. However, a visual inspection of Figure 3.3 shows a lack of disky 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 (Estrada-Carpenter et al., 2025). Simulations have shown that proto-clumps can form along gas filaments accreting onto the host galaxies (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.

### 3.7.2 Clump evolution

In the previous discussions, we investigate how the slopes of the cSMF are affected by formation mechanisms. Here, we now consider how the shape of the cSMF is affected as clumps evolve. We investigate the evolution of clump populations by comparing clump mass and age, using the inferred ages to define samples of young and old clumps. The top of Figure 3.13 shows the distribution of clump ages in low- and high-mass galaxies. The bottom panel shows the relationship between clump

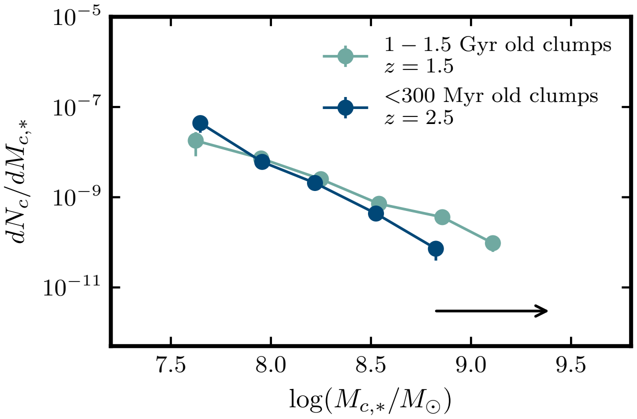


Figure 3.14: The evolution of the normalized clump mass function for young clumps (<300 Myr old) over approximately 1.5 Gyr of cosmic time. As clumps evolve, we find a relative decrease in the number of lower mass clumps, while massive clumps are more likely to persist. The change in the mass function also suggests that clumps continue to grow in stellar mass as they evolve. 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.

mass, their age, and the host galaxy mass. Some clumps have an estimated age exceeding a gigayear despite our UV-based selection. These higher ages suggest that their star formation rate history is more extended as the age here is defined as the lookback time at which 75% of the mass is created.

If clump populations evolve without mass-dependent effects, the shape of the cSMF should remain similar between young and old samples. This scenario is tested by selecting a population of young clumps ($<300$ Myr) at $z \sim 2.5$ and compare its cSMF to a population of older clumps ($>1$ Gyr and $<1.5$ Gyr) at $z \sim 1.5$. 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 $z \sim 1.5$. Lastly, we limit the analyses to clumps found in galaxies of similar masses between $10^{10}$ and $10^{11}$ solar masses to mitigate the dependence of the slope of the mass function on galaxy masses. We focus on this higher galaxy mass bin because it provides the larger clump sample. Therefore, the two clump populations have nearly identical host-galaxy mass distributions, with a mean of $\log(M_*/M_{\odot}) \simeq 10.5$ and a scatter of $\sim 0.3$ dex.

Figure 3.14 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 is consistent with a scenario in which lower-mass clumps are preferentially disrupted. Indeed, feedback processes, such as supernovae and radiation pressure, limit clump survivability, particularly at lower clump masses (e.g., Moody et al., 2014; Mandelker et al., 2017; Ginzburg et al., 2021; Dekel et al., 2022). Other mechanisms such as clump-clump mergers can also flatten of the slope of the mass function by removing smaller mass clumps to create bigger ones (e.g., Behrendt et al. 2019).

It is important to address a few caveats. We note that the interpretation of Figure 3.14 assumes that all clumps have similar star formation rates. However, if the star formation rate of a clump is correlated to its stellar mass (similar to a star-forming main sequence), this could also flatten the relation over time. Additionally, our UV-based clump selection preferentially identifies actively star-forming clumps, which may bias against older, low-mass clumps. Such selection effects could contribute to the flattening of the cSMF for older clumps. A more detailed characterization of the clump detection completeness as a function of both mass and age will therefore be beneficial to fully quantify the impact of this bias. Given these caveats, the observed evolution of the clump mass function nevertheless suggests that evolutionary trends are present in the clump population, motivating further work that will be required to disentangle intrinsic physical processes from observational selection effects.

## 3.8 Chapter 3 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 ($f_{\rm clumpy}$) 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 ($\alpha_{\rm [high-mass]} \sim -1.6$) relative to less massive galaxies ($\alpha_{\rm [low-mass]} \sim -2.3$). 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, this change in slope indicates that the relative

abundance of low-mass clumps decreases with time. Such evolution may reflect a combination of clump growth through mergers, and/or the preferential disruption of lower-mass clumps.

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 initial conditions for clump formation and the processes that shape their subsequent evolution.* 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 processes such as clump mergers and stellar feedback that may preferentially remove 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 program 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 partnership 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 - Chapter 3

## 3.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 Tan et al. (2022). In brief, we normalize the stellar mass maps to the F444W images, and injected Gaussian noise sampled from the pixel-to-pixel noise of the F444W image. We then used STATMORPH 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 §3.4.2, 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 3.A.1 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.

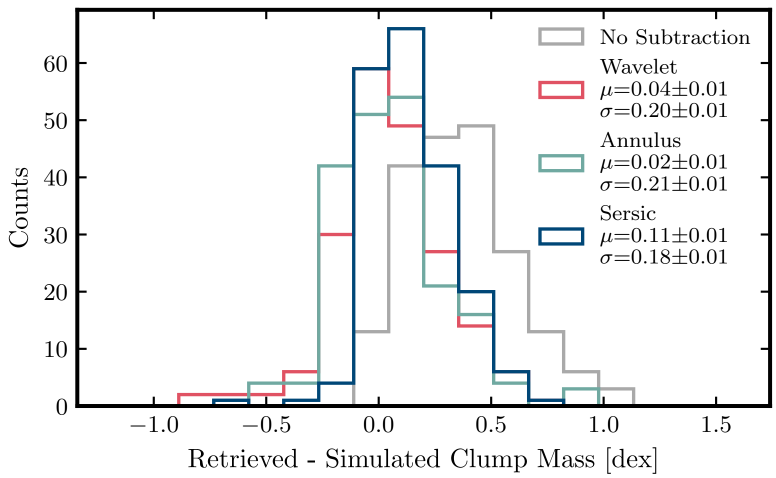


Figure 3.A.1: A comparison of different background subtraction methods for retrieving clump masses. In particular, we tested three different methods to estimate the underlying stellar disk; by fitting the disk as a 2D Sérsic profile, using wavelet transformation to construct a smooth disk (see text for 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.

# Chapter 4

# Metal-poor gas accretion leading to clump formation in cosmic noon SFGs

"You'll have bad times, but that'll always wake you up to the good stuff you weren't paying attention to."
— Sean, Good Will Hunting

---

*The people involved in the analyses for this chapter include Adam Muzzin, Jialu Chen, Ben Forrest, Gillian Wilson, Sunna Withers, Katherine Myers, Nelson Nunes and the CANUCS collaboration. In this chapter, I expand on the findings of Chapter 2 by compiling and observing the gas-phase metallicity of a larger sample of* $\sim$ *300 galaxies. I investigate how stellar properties of clumps relate to the global gas-phase metallicity, showing that younger, star-forming clumps are more common in galaxies with lower gas-phase metallicities.*


## Abstract

Recent studies probing the gas-phase metallicity of a small sample of $< 30$ star-forming galaxies (SFGs) have found that clumpy galaxies are metal-poor compared to nonclumpy galaxies, with clumps exhibiting a metallicity offset from the host gas disk, suggesting ongoing gas accretion events. In this work, we expand on these studies by comparing the gas-phase metallicity of a sample of $\sim$300 clumpy and nonclumpy SFGs between $0.6 < z < 2.6$. We find that clumpy and nonclumpy galaxies generally have lower and higher gas-phase metallicities compared to the mass-metallicity relation, respectively. Using a smaller subset of $\sim$ 30 galaxies observed in the JWST CANUCS fields, we further investigate the relationship between the resolved stellar properties of clumps and the integrated gas-phase metallicity of the host galaxies. In particular, clumps in galaxies with metallicity above the mass-metallicity relation (i.e., $\Delta Z > 0$) are generally older and have higher SFRs, compared to clumps whose host galaxies have $\Delta Z < 0$. We do not find a significant mass difference between these two clump populations. Finally, we compute the merger statistic using the Gini-M20 morphological parameters and find that most clumpy galaxies are not mergers based on

their stellar mass maps. The results suggest that the clumpy nature of cosmic noon galaxies is linked to gas accretion events that dilute metals and form star-forming clumps.

## 4.1 Introduction

It is well known that high-redshift star-forming galaxies generally host kiloparsec-scale clumpy structures (e.g., Wuyts et al., 2012; Guo et al., 2012). In fact, clumpy galaxies are ubiquitous at higher $z$, with redshift evolution similar to that of the cosmic star formation density, peaking near cosmic noon (e.g., Shibuya et al., 2016; Sattari et al., 2023; de la Vega et al., 2025; Sok et al., 2025b). Observations with the *James Webb Space Telescope* of distant galaxies further show that clumpy substructures persist at $z > 5$, thanks to improved sensitivity and spatial resolution probed through gravitational lensing. Indeed, we are able to probe clumpy structures down to parsec scales (e.g., Messa et al., 2024; Adamo et al., 2024; Mowla et al., 2024; Fujimoto et al., 2025; Claeyssens et al., 2023, 2025; Bradač et al., 2025). Such a unique and prominent mode of clumpy star formation at higher redshifts suggests that galaxy assembly history may be inherently clumpy, with even Milky Way-like galaxies possibly hosting clump formation (e.g., Mowla et al. 2024; Tan et al. 2024). Understanding how clumps are formed and whether these mechanisms change over cosmic time is therefore paramount for understanding galaxy evolution.

Despite the abundance of star-forming clumps, their origin and evolution are relatively unconstrained. The common framework for clump formation is violent disk instability (VDI), in which clumps form due to gravitational instabilities in gas-rich regions (e.g., Toomre, 1964; Dekel et al., 2009b). The gas-rich nature of these disks is attributed to the continuous accretion of cosmological gas streams (e.g., Kereš et al., 2005, 2009; Dekel et al., 2009b,a; van de Voort and Schaye, 2012; Nelson et al., 2013). The accreting gas is expected to be metal-poor (van de Voort and Schaye, 2012), in which a metallicity offset between clumps and host gas disks is expected as observed in simulations (Ceverino et al., 2016). Observational constraints for clump formation via VDI generally probe the Toomre $Q$ parameter of the gas disks (e.g., Förster Schreiber et al., 2011b; Genzel et al., 2011; Girard et al., 2018; Fujimoto et al., 2025). However, this argument relies on simplified assumptions about disk structure, and future studies may also need to consider more formalisms of the Toomre stability theory to better capture the complexity of disk stability (e.g., Romeo et al., 2010; Romeo and Wiegert, 2011; Bacchini et al., 2024).

Another pathway for clump formation is through mergers, in which gas is compressed by tidal forces arising from galaxy interactions (e.g., Renaud et al., 2015; Nakazato et al., 2024). Direct observational support comes from kinematic studies of smaller samples of galaxies, including Puech (2010); Menéndez-Delmestre et al. (2013); Calabrò et al. (2019). Indirect inferences about clump formation resulting from mergers are also made by using the redshift evolution of the merger rate to explain the redshift evolution of the fraction of clumpy galaxies (Guo et al., 2015; de la Vega et al., 2025). However, clumps may further be of *ex-situ* origins, being satellites, as observed in simulations (e.g., Mandelker et al., 2014, 2017) and suggested in observations (e.g., Ribeiro et al., 2017; Zanella et al., 2019).

Addressing the mechanisms driving clump formation requires observational data that link clump formation to a physical process, such as gas inflow. Recently, in Sok et al. (2025a), we showed that in a comparable sample of approximately 30 star-forming galaxies, those that are clumpy have a metallicity offset of $\sim -0.1$ dex compared to nonclumpy galaxies. Similarly, resolved metallicity maps for approximately 20 SFGs obtained with JWST NIRISS slitless grism data show that clumps typically have a similar metallicity offset of $-0.1$ dex relative to the host galaxy disk (Estrada-Carpenter et al., 2025). While the sample size of these studies is small ($< 30$), these findings suggest that clump formation likely occurs due to lower-metallicity gas accretion events.

Building upon these previous studies, this chapter aims to investigate the relationship between clump formation and gas-phase metallicity for a statistically significant sample by compiling existing spectroscopic data from the Large Early Galaxy Astrophysics Census (LEGA-C) survey (van der Wel et al., 2016) and MOSFIRE Deep Evolution Field (MOSDEF) survey (Kriek et al., 2015). We further obtain robust gas-phase metallicity constraints for $\sim$70 SFGs in the Canadian NIRISS Unbiased Cluster Survey (CANUCS, Willott et al. 2022) fields with new MOSFIRE observations. Combined with the unparalleled multi-wavelength imaging of the CANUCS data, we examine the relationship between the global gas-phase metallicity of host galaxies and the resolved properties of star-forming clumps, derived from spectral energy distribution (SED) fitting. Such a multi-wavelength approach for studying clumps enables us to directly test the physical processes driving clump formation at cosmic noon.

The chapter is organized as follows. Section 4.2 describes the data presented in this work. We describe the methodology in Section 4.3. More specifically, the section explains how clumpy and nonclumpy galaxies are identified, and how the stellar properties of the clumps are estimated. Section 4.5 explores the gas-phase metallicity of SFGs and its relation to galaxy clumpiness. Section 4.6 expands on this by examining the relationship between metallicity offset with respect to the mass-metallicity relation and properties of star-forming clumps. We discuss the implications for clump formation mechanisms in Section 4.7 and summarize our work in Section 4.8.

Throughout the chapter, the term metallicity represents the gas-phase oxygen abundance. All magnitudes are expressed in the AB system (Oke and Gunn, 1983). We adopt a $\Lambda$CDM cosmological model of the universe with $\Omega_\lambda = 0.7$, $\Omega_\mathrm{M} = 0.3$, and a Hubble constant of $H_0 = 70$ km/s/Mpc.

## 4.2 Data

To provide an in-depth analysis of the relationship between galaxy clumpiness and gas-phase metallicity across a range of redshifts, we make use of several major extragalactic surveys. In particular, we combine high-quality emission line measurements from the Large Early Galaxy Astrophysics Census (LEGA)-C survey (van der Wel et al., 2016) and MOSFIRE Deep Evolution Field (MOSDEF) survey (Kriek et al., 2015), which together span a redshift range between $z \sim 0.6$ to $z \sim 2.6$. These datasets are chosen for their large galaxy samples and availability of ancillary imaging, enabling us to investigate the connection between galaxy morphology and their gas-phase metallicity over a broad redshift range.

In addition to these publicly available datasets, we include new near-infrared spectroscopic observations of galaxies in the Canadian NIRISS Unbiased Cluster Survey (CANUCS, Willott et al. 2022) fields obtained with the MOSFIRE instrument (McLean et al., 2012) on the 10-meter Keck I telescope. At the time of these observations, CANUCS was among the first few extragalactic fields with deep, multi-wavelength JWST/NIRCam imaging. These enable robust measurements of the resolved properties of clumps such as stellar masses and ages. In the following sections, we describe the spectroscopic and photometric data used in this work, starting with the LEGA-C and MOSDEF surveys, followed by our targeted observations of the CANUCS fields.

### 4.2.1 The LEGA-C Survey

We utilize the third data release of LEGA-C (van der Wel et al., 2021). LEGA-C is a European Southern Observatory (ESO) Public Spectroscopic survey of approximately 3500 galaxies at $0.6 < z < 1$, observed with the Very Large Telescope/Visible Multi-Object Spectrograph (VIMOS; Le Fèvre et al. 2003) at a spectral resolution of $R \sim 3500$. Galaxies are located in the Cosmic Evolution Survey (COSMOS) field (Scoville et al., 2007), and were selected as star-forming galaxies from the UltraVISTA (McCracken et al., 2012) K-band catalogue presented by Muzzin et al. (2013).

We further cross-match the LEGA-C emission line flux catalogue with the deconvolved ground-based imaging of the COSMOS field (Sok et al., 2022). The deconvolved imaging is available for over 20,000 star-forming galaxies between $0.5 < z < 2$, with wavelength coverage spanning from the Subaru $B_j$-band to the UltraVISTA $Ks$-band. In this chapter, we use these images to construct the rest-frame $U$ surface brightness maps for the LEGA-C sample and classify them as being clumpy or nonclumpy (see §4.3).

### 4.2.2 The MOSDEF Survey

We also include data from the MOSDEF survey, a 47-night observational program with MOSFIRE that targets $\sim 1500$ galaxies with $1.37 < z < 3.80$. Specifically, the survey provided medium-resolution rest-frame optical spectra for these galaxies, including bright emission lines such as H$\alpha$ and [N II] $\lambda$6584. The galaxies are located in three of the CANDELS (Grogin et al., 2011; Koekemoer et al., 2011) fields (AEGIS; Davis et al. 2007, COSMOS; Scoville et al. 2007, and GOODS-N; Giavalisco et al. 2004). Similarly, targets for the MOSDEF survey were selected from the multi-wavelength photometric and spectroscopic catalogues of the 3D-HST survey (Brammer et al., 2012; Skelton et al., 2014).

As with the LEGA-C data, we classify galaxies in the MOSDEF survey as clumpy or nonclumpy using the photometric imaging products from the 3D-HST catalogue. In particular, we use the available HST imaging, spanning F475W to F160W, to construct their rest-frame $U$-band brightness map, following the procedures described in §4.3. Finally, we note that the imaging products related to the LEGA-C and MOSDEF surveys mentioned in this chapter are used solely to classify galaxy morphology. We do not attempt to extract stellar properties of clumps here, either due to these having limited angular resolution (in the case of LEGA-C), or lack of sufficient filter coverage for SED fitting (in the case of MOSDEF).

### 4.2.3 CANUCS x MOSFIRE Observations

The primary focus of this study is to understand the relationship between the properties of star-forming clumps and the integrated gas-phase metallicity of their host galaxies. To this end, we use the JWST CANUCS fields, comprising JWST NIRCam and NIRISS broad-band and medium-band imaging. These images reach $3\sigma$ AB magnitude depths of 27–30 in a $0.3''$-diameter aperture. A detailed list of available filters in each field is provided in Sarrouh et al. (2025). The addition of medium-band imaging enables more accurate modelling of the strong nebular emission observed in high-redshift galaxies (e.g., Withers et al. 2023), leading to better photometric redshift determinations and stellar mass estimates. Relying solely on broad-band JWST data often leads to underestimated stellar masses (e.g., Sarrouh et al. 2024). Building on our previous work (Sok et al., 2025b), which investigated the clump mass and age distributions, we now extend the analysis by inferring gas-phase metallicities from the [N II] $\lambda6584$/H$\alpha$ ratio. This allows us to explore the overall relationship between clump properties and the global properties of their host galaxies, such as gas-phase metallicities.

### Observations and Data Reduction

We obtain four nights on the Keck I telescope with MOSFIRE to perform deep spectroscopic observations of galaxies in the CANUCS fields. We target galaxies whose H$\alpha$ and [N II] $\lambda6584$ emission lines fall within the wavelength coverage of the H ($1.450 - 1.826$ $\mu m$) and K ($1.450 - 1.826$ $\mu m$) bands, respectively. In total, we create 17 (7 H and 10 K) masks using the MAGMA slit mask design software. We adopt a slit width of $0''.7$, corresponding to a spectral resolution of $R \sim 3600$.

The primary targets are selected from the catalogue presented in Sarrouh et al. (2025). We determine slit-placement priorities for each mask based on several criteria. In particular, we target only star-forming galaxies and prioritize those that are massive and have higher SED-based H$\alpha$ flux estimates. The priority on stellar masses ensures that both H$\alpha$ and [N II] $\lambda6584$ can be observed with high S/N, particularly at higher redshifts where metallicity is expected to be lower. The H$\alpha$ estimates are obtained from the SED-derived star formation rate, with a dust correction applied based on their SED-inferred $A_v$. As most of our primary targets have no prior spectroscopic observations, we use the photometric redshift to ensure that both H$\alpha$ and [N II] $\lambda6584$ fall within the observation window of the $H$ and $K$ bands. During the first two nights, we include galaxies with stellar masses greater than $10^9 M_\odot$, and SED-based H$\alpha$ fluxes greater than $1 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$, maximizing the number of galaxies per mask with longer exposure time. However, our observing strategy changed after the first two nights to target galaxies with $F_{\mathrm{H}\alpha} \geq 5 \times 10^{-18}$ erg s$^{-1}$ cm$^{-2}$.

Each MOSFIRE mask has an on-source integration time of approximately 2 hours, and masks were executed from February 2024 to January 2025. We follow a similar observing strategy as the MOSDEF survey, specifically using an ABA$'$B$'$ dither pattern with nods of $+1''.5$, $-1''.2$, $+1''.2$, $-1''.5$, and individual exposures of 2/3 minutes for the H/K bands. In total, we place 201 primary targets on slit, and successfully measured spectroscopic redshifts for 141 of them (based on the detection of H$\alpha$). The overall success rate is lower, partly because the first few masks observed include targets with lower SED-derived H$\alpha$ fluxes.

We use the custom IDL pipeline of the MOSDEF survey, described in Kriek et al. (2015), to reduce and produce two-dimensional (2D) science spectra for each slit on the mask. The one-dimensional (1D) science and error spectra are extracted based on the optimal extraction algorithm (Horne, 1986). Figure 4.1 shows examples of the reduced spectra. We do not apply any correction for slit losses, as this would be a multiplicative factor applied to both H$\alpha$ and [N II] $\lambda6584$ fluxes, assuming weak metallicity gradients. However, we note that the metallicity gradient may differ from one galaxy to another.

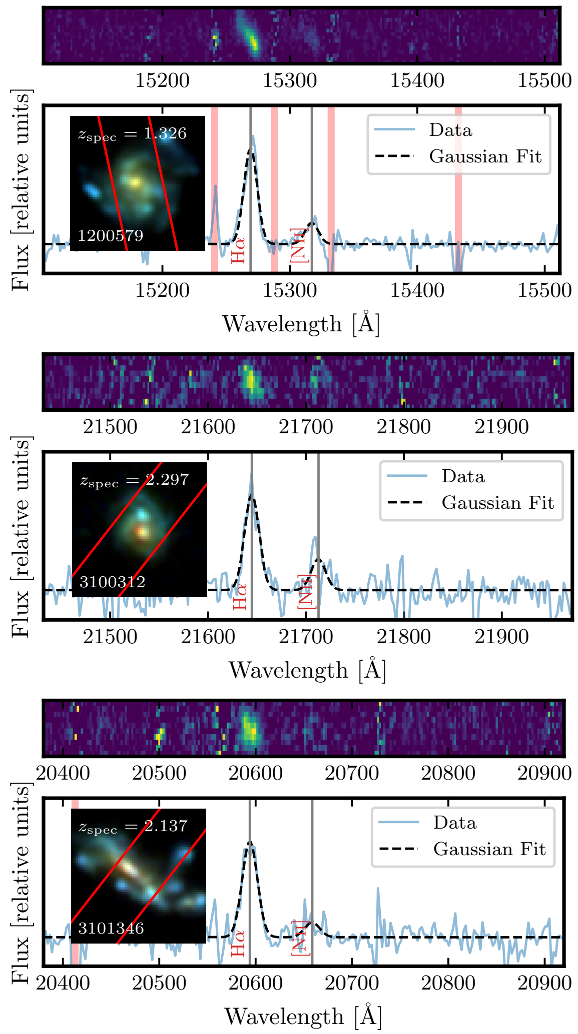


Figure 4.1: Examples of the reduced two dimensional spectrum and the extract spectrum. The inset figure shows the composite image of the galaxy, with the red line denoting the MOSFIRE slit position and width.

### 4.2.4 Sample Selection

For each emission line flux catalogue, we select a sample of galaxies using several criteria. The criteria selection includes,

1. S/N $\geq 3$ for [O II] $\lambda 3726$, [O III] $\lambda 4959$, [O III] $\lambda 5007$, and H$\beta$ to ensure a robust estimate of the gas-phase metallicity for the LEGA-C sample. We additionally require S/N $\geq 3$ for $H\gamma$ to apply a dust correction based on the Balmer decrement.

2. S/N $\geq 3$ for H$\alpha$ and [N II] $\lambda 6584$ to ensure a robust estimate of the metallicity in both the MOSDEF and CANUCS samples. No dust correction is applied here since H$\alpha$ and [N II] $\lambda 6584$ are close in wavelength.

3. We exclude emission lines strongly affected by skylines in both the MOSDEF and CANUCS samples. For CANUCS, this is done by visual inspection, while for MOSDEF we adopt their

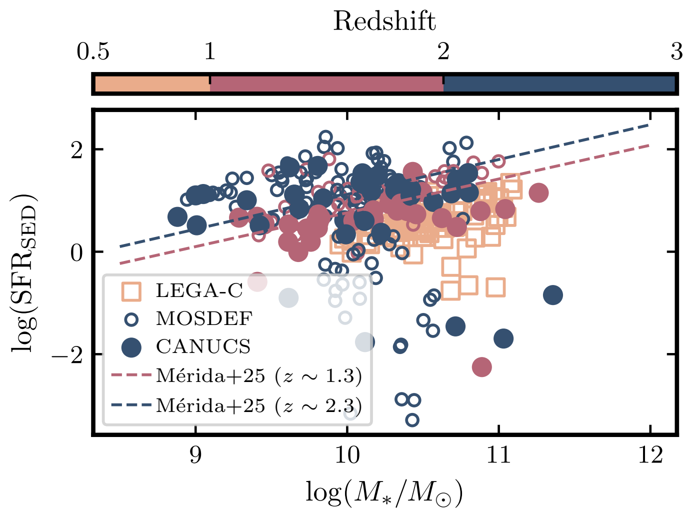


Figure 4.2: The mass and SFR distribution of galaxies from LEGA-C, MOSDEF and CANUCS, color coded by redshift. Both stellar mass and SFR are derived from Dense Basis. The star-forming main sequences from Mérida et al. (2025) are shown as the dashed lines.

skyline flag criterion, SLFLAG $> 0.2$.

4. We also exclude potential AGN. This corresponds to excluding galaxies with FLAG_SPEC $\geq 1$ (van der Wel et al., 2021) and galaxies flagged as mass-extinction AGNs using the cut from Juneau et al. (2011) in the LEGA-C sample. For MOSDEF and CANUCS, we exclude galaxies with $\log([\text{N II}]\,\lambda6584/\text{H}\alpha) > -0.4$ (Cid Fernandes et al., 2010).

These criteria yield a sample of robust emission lines for 94 LEGA-C, 125 MOSDEF, and 57 CANUCS galaxies. Figure 4.2 shows the distribution of galaxies based on their stellar masses and star formation rates. These are derived by fitting the SED using Dense Basis (Iyer and Gawiser, 2017). For the LEGA-C anbd MOSDEF galaxies, we performed the fit using the photometric catalogue from Muzzin et al. (2013) and Skelton et al. (2014) with the spectroscopic redshift.

## 4.3 Methodology

### 4.3.1 Finding clumpy galaxies

This section describes the steps for determining whether a galaxy hosts star-forming clumps. This process begins by binning pixels together to achieve sufficient signal-to-noise (S/N) for spectral energy distribution (SED) fitting. Specifically, we fit the SED of the Voronoi bins to estimate and reconstruct the rest-frame $U$ (∼3600 Å) surface brightness. The rest-frame $U$-band brightness map is then used to identify clumpy regions within galaxies. Galaxies are then classified as clumpy or non-clumpy based on the presence of such areas.

#### Spatially binning pixels

Fitting the spectral energy distribution in an unresolved manner has been shown to lead to overestimation of the stellar masses (e.g., Sorba and Sawicki 2015, 2018). To mitigate this, we apply a spatially resolved SED fitting technique to obtain robust estimates of the physical properties of star-forming clumps and their host galaxies. However, since individual pixels often have signal-to-noise ratios (S/N) that are unsuitable for SED fitting, we first spatially bin pixels together to improve the signal-to-noise ratio (S/N) of bins while retaining spatial information. This binning step is performed using the Voronoi tessellation technique, VORBIN (Cappellari and Copin, 2003).

Similar to our previous work (Sok et al., 2022, 2025b), we require each Voronoi bin to achieve a minimum S/N of 5. For each survey, we select the band probing the reddest rest-frame wavelength to define the bins. This ends up being the deconvolved $K_s$ for LEGA-C, F160W for MOSDEF and F444W for CANUCS. However, this choice does not affect the identification of clumps, since the spatial redistribution of flux is performed using the rest-frame UV images after fitting the binned SEDs (Section 4.3.1). This procedure preserves the UV morphology when mapping the fitted quantities back to the native pixel grid, under the assumption that all pixels within a given bin share the same colors.

These criteria yield a sample comprising 94 LEGA-C, 125 MOSDEF, and 69 CANUCS galaxies. Figure 4.2 shows the distribution of galaxies based on their stellar masses and star formation rates. The star formation rates for the LEGA-C and MOSDEF samples are based on the UV + IR fluxes. For CANUCS, the SFRs are obtained from the SED fitting using DENSE BASIS.

#### Creating $U_{\rm rest}$ maps

For each galaxy, we then calculate the fluxes for each Voronoi bin in each filter. We determine the flux for each bin by summing up the pixel values. The flux uncertainty is estimated as the quadrature sum of the statistical and systematic uncertainties. The statistical uncertainty is determined to be the aperture flux uncertainty, while the systematic uncertainty accounts for potential biases during data reduction and calibration. In practice, the flux uncertainty for a given Voronoi bin is assumed to be the same as the flux uncertainty within a circular aperture encompassing the same area. The systematic uncertainty is considered to be a multiplicative factor of 5% of the flux. This is similar to what is used in PIXEDFIT (Abdurro'uf et al., 2021), another Python-based spatially resolved SED-fitting tool.

We use EAZY-PY (Brammer et al., 2008) to estimate rest-frame $U$ luminosity for each bin, adopting the recalibrated rest-frame $U$ filter from Maíz Apellániz (2006). To this end, we set the redshift to each galaxy's spectroscopic redshift. As binning degrades the angular resolution, particularly for the bins at the outskirts of the galaxies where S/N is lowest, we recover spatial resolution by "dezonifying" each map (Cid Fernandes et al., 2013). For a given derived quantity of a Voronoi bin $Q_{\rm bin}$, the pixel within is distributed as,

$$Q'_i = Q_{\rm bin} \frac{F_i}{F_{\rm bin}}, \tag{4.1}$$

where $i$ denotes the pixel, $F_i$ is the flux value at that pixel, and $F_{\rm bin}$ is the total flux of the Voronoi bin. This therefore preserves the total value for each bin while reintroducing spatial resolution.

#### Clumpy/nonclumpy classification

In this Section, we describe how galaxies are classified as clumpy and nonclumpy. The classification is based on the $U_{\rm rest}$ surface brightness map of the galaxy. To this end, we construct a normalized

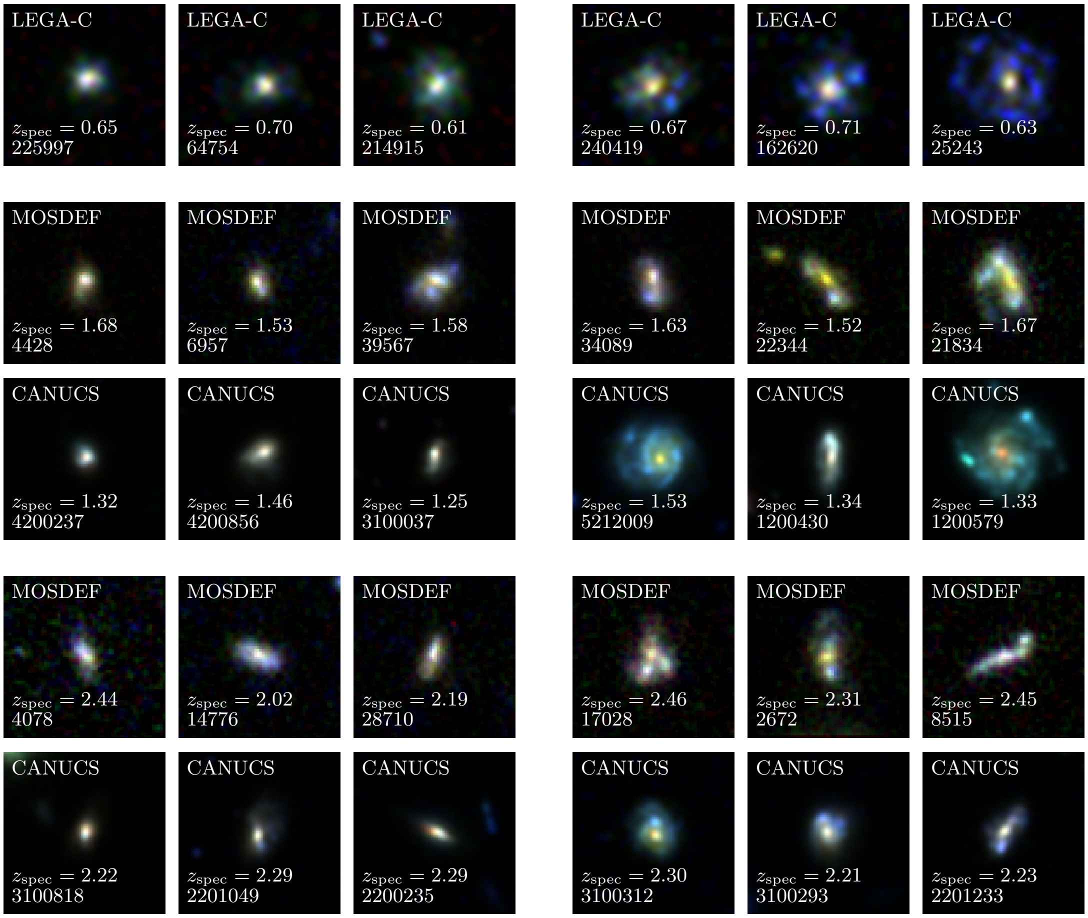


Figure 4.3: Examples of the classification of clumpy and nonclumpy galaxies. We show galaxies with increasing fractional $U_{\rm rest}$ luminosity from left to right, with the three left- and right-most galaxies being classified as nonclumpy and clumpy, respectively. We note that some galaxies host clumps are yet classified as nonclumpy because their clumps do not meet the adopted threshold as defined in Section 4.3.1.

Each row further separates galaxies by the different redshift intervals sampled by the different surveys. Each cutout is $3.6'' \times 3.6''$.

light profile from the $U_{\rm rest}$ map using the prescription described in Wuyts et al. (2012). We refer the reader to their paper and to Sok et al. (2022) for a detailed explanation of the methodology. Briefly, galaxies with no clumps (i.e., a smooth disk) will have a constructed normalized $U_{\rm rest}$ light profile that is smoothly decreasing in intensity with galactocentric radius. Since clumps are regions of enhanced star formation, clumps appear as a bump on an otherwise smooth normalized light profile in the $U_{\rm rest}$.

The construction of the normalized light profile begins with defining elliptical parameters for each galaxy. The center of mass of the galaxies is adopted from the stellar mass map, assuming it traces the gravitational potential. An elliptical aperture is then used to obtain the curve of growth centred on the centre of mass. A half-light radius $R_e$ is then calculated as the effective radius at which half the light of the galaxy is contained within that radius. We then calculate the mean $U_{\rm rest}$ surface brightness within the half-light brightness as $\Sigma_e$. The galactocentric distance of each pixel within the $U_{\rm rest}$ brightness map is then normalized by the half-light radius. The $U_{\rm rest}$ flux of each pixel is also normalized by $\Sigma_e$. This renormalization maps each pixel into a new parameter space, where Wuyts et al. (2012) defined a region within this parameter space to select clumpy pixels from the $U_{\rm rest}$ brightness map. The dividing line for the clumpy regime was determined by visually analyzing a number of normalized maps from different galaxies. In this chapter, clumpy galaxies are defined as those galaxies with clumps contributing to at least 8% of the total $U_{\rm rest}$ luminosity. This threshold, compared to the 5% adopted in Sok et al. (2025b), is chosen to isolate galaxies dominated by actively star-forming clumps, thereby increasing sensitivity to metallicity dilution expected from recent gas accretion. The thresholding value is broadly consistent with other studies that identified clumps and clumpy galaxies (e.g., Guo et al. 2015; Wuyts et al. 2013; Sok et al. 2025a). Examples of the classification of galaxies are shown in Figure 4.3.

### 4.3.2 Properties of Star-Forming Clumps

For star-forming clumps in galaxies located in the CANUCS field, we make use of the available deep medium- and broad-band JWST imaging to extract their physical properties, including stellar masses, star formation rates and ages. This next section describes how individual clumps are identified within clumpy galaxies and how we determine their physical properties. We note that this sample is a subset of the clumpy galaxy sample presented in Sok et al. (2025b). We therefore refer the reader to that chapter for a more detailed description of the methodology; however, we briefly describe it here.

#### Finding clumps in galaxies

Clumps are identified in a similar process as described in Kalita et al. (2025a), where the $U_{\rm rest}$ image is passed through a high-pass filter to create a high-contrast image. In practice, the low-pass image of the galaxy is determined by decomposing the image into six different wavelets using SCARLET (Melchior et al., 2018). Each wavelet scale probes a different spatial scale. We obtain a smooth version of the image by reconstructing the third wavelet scale and higher. Subtracting the image from its smoothed image results in an image containing only smaller-scale structures. A peak-finding algorithm is then used to locate clump candidates. We set a minimum $S/N > 3$ within an aperture of $0''.2$ to remove spurious detection of clumps. Figure 4.4 illustrates this and highlights the detected clumps for one galaxy.

#### Inferring physical properties from SED

Star-forming clumps are generally superimposed within the disk of their host galaxies. Background subtraction is typically performed to mitigate flux contamination from the disk. In previous studies, disk subtraction is commonly performed using aperture photometry, where the disk background flux is estimated from an annulus aperture (e.g., Guo et al. 2018; Sattari et al. 2023). We follow a similar methodology to subtract the disk. Specifically, we use an aperture size of $0''.2$ centred on the clumps, with an annuli aperture of $\sim 0''.3 - 0''.35$, which is found to reliably extract clump masses from the disk (Sok et al., 2025b).

The background-subtracted flux is then obtained for each filter to extract physical information of the clumps with SED fitting. The SED is fitted using DENSE BASIS (Iyer and Gawiser, 2017;

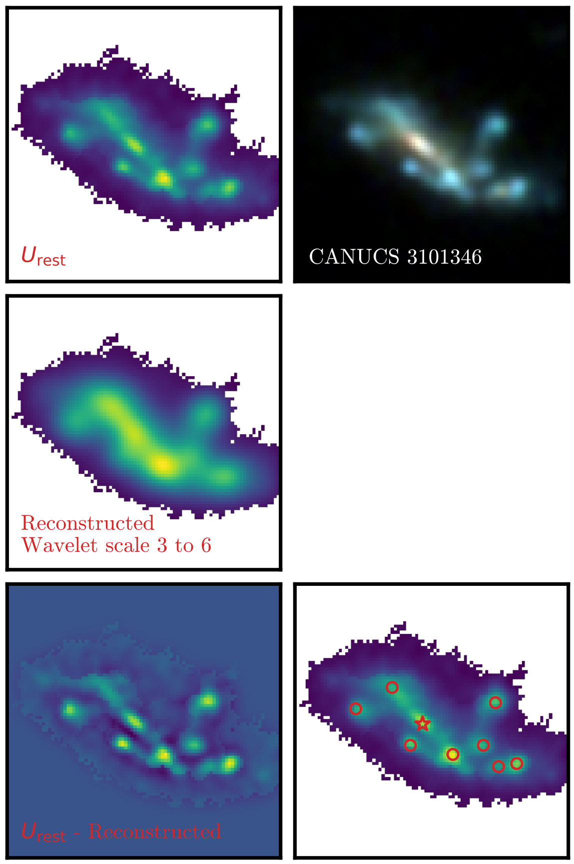


Figure 4.4: An example showing how clumps are identified. Clumps are detected based on the high-contrast image. This high-contrast image is obtained by subtracting the rest-frame bright map by its reconstructed smooth image (see text). The reconstructed image is shown in the middle left panel. The bottom left panel shows the high-contrast image, while the bottom right shows the detected clumps (circle) and bulge (star).

Iyer et al., 2019). The method describes the star formation history (SFH) using a tuple ($M_*$, SFR, $t_X$). In the current analysis, $t_X$ are a set of three lookback times describing when the galaxy formed equally spaced quantiles of its total mass (i.e., $t_{25}$, $t_{50}$, $t_{75}$). We adopt a Chabrier initial mass function (Chabrier, 2003) and Calzetti attenuation law (Calzetti et al., 2000) to model the stellar populations. The priors include flat priors for $\log(M_*)$ with range of (7,12), log(sSFR) with range of (-14,-7), $\log(Z/Z_\odot)$ with range of (-1.5,0.25), and $A_v$ with range of (0,4). Dense Basis performs SED fitting using an amortized brute-force Bayesian approach, in which an atlas of model SEDs is precomputed from random draws of parameters sampled from the prior distributions.

We extract the stellar mass, star formation rate (averaged over the past 100 Myr), and age for each clump. The clump age is defined as the lookback time at which 75% of the stellar mass has formed ($t_{75}$), as inferred from the reconstructed star formation history. This metric closely tracks the mass-weighted age while remaining more sensitive to recent star formation. We adopt the posterior medians as the fiducial values and quote the 16th–84th percentile ranges as uncertainties.

## 4.4 Gas-Phase Metallicity

This Section describes how we infer gas-phase metallicity from strong emission lines across different surveys. There exist several metallicity diagnostics that use flux ratios of strong emission lines have been empirically and/or theoretically calibrated (for a review see Kewley et al. 2019). In this work, we utilize both the $R23$ and $N2$ indices to infer gas-phase metallicity.

For the LEGA-C samples, metallicities are derived using the $R23$ strong line calibration from Kobulnicky and Kewley (2004). The $R23$ index requires multiple strong emission lines to obtain the relevant flux ratios to infer gas-phase metallicities,

$$R23 = \frac{[\mathrm{O\,II}]\,\lambda3726 + [\mathrm{O\,III}]\,\lambda4959 + [\mathrm{O\,III}]\,\lambda5007}{\mathrm{H}\beta}, \tag{4.2}$$

$$O23 = \frac{[\mathrm{O\,III}]\,\lambda4959 + [\mathrm{O\,III}]\,\lambda5007}{[\mathrm{O\,II}]\,\lambda3726}, \tag{4.3}$$

The $R23$ index is sensitive to the ionization parameter, but this sensitivity can be corrected for using the $O23$ index. The typical intrinsic scatter for an individual metallicity measurement is $\sim$0.1 dex (Kobulnicky and Kewley, 2004).

For the MOSDEF and CANUCS samples, we use the $N2$ index to infer metallicity. The $N2$ index is commonly used in higher-redshift observation due to the close proximately of H$\alpha$ and [N II] $\lambda$6584, and is defined by the line ratio,

$$N2 = \frac{[\mathrm{N\,II}]\,\lambda6584}{\mathrm{H}\alpha}. \tag{4.4}$$

To this end, we follow the line calibration from Pettini and Pagel (2004), which is given as,

$$12 + \log(\mathrm{O/H}) = 8.9 + 0.57 \times N2. \tag{4.5}$$

The intrinsic scatter of the relation is approximately 0.18 dex.

We note that different line diagnostics provide different gas-phase metallicities. While there exist several ways to calibrate one line diagnostic to another (e.g., Kewley and Ellison, 2008; Teimoorinia et al., 2021), we do not apply any such calibration in this chapter, as the primary focus here is on the

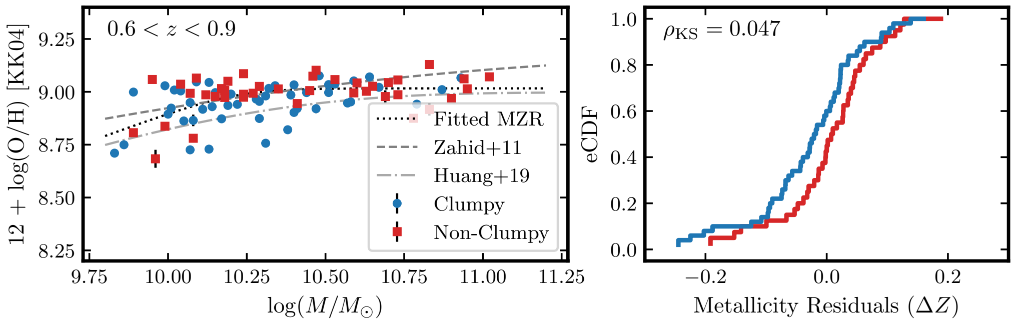


Figure 4.5: The mass-metallicity relation for the LEGA-C sample using the $R23$ index. The dotted line shows the best-fit mass-metallicity relation. The right panel shows the empirical cumulative distribution function (eCDF) of the metallicity residual, defined to be $\Delta Z = Z - \text{MZR}$. We find that the eCDF of clumpy is generally shifted to lower $\Delta Z$ values compared to nonclumpy galaxies, indicating that clumpy galaxies have metallicity lower than the MZR.

relative metallicity offset between clumpy and nonclumpy galaxies. Finally, throughout this work, metallicity is expressed as the ratio of the number of oxygen atoms to hydrogen atoms and quoted as $12 + \log(\text{O/H})$.

## 4.5 Do clumpy galaxies have lower metallicity compared to the MZR?

Previous work has found that clumps and their host galaxies have lower metallicity than the host disk and other nonclumpy galaxies, respectively (e.g., Sok et al., 2025a; Estrada-Carpenter et al., 2025). In particular, we showed that for a controlled sample of galaxies with similar masses, star formation rates and colours, clumpy galaxies are systematically more metal-diluted than their nonclumpy counterpart (Sok et al., 2025a). In this Section, we further investigate the metallicity offset between clumpy and nonclumpy galaxies across a broader sample. We remind the reader that clumpy galaxies are defined as galaxies with star-forming clumps that contribute to at least 8% of the total $U_{\text{rest}}$ luminosity of the host galaxy.

In these analyses, we do not apply any conversion factor between the $R23$ and $N2$ indices as we are mainly interested in the relative offset in gas-phase metallicity between clumpy and non-clumpy galaxies for each sample/redshift bin. We first examine the metallicity offset in the LEGA-C sample. The left panel of Figure 4.5 shows the gas-phase metallicity inferred from the $R23$ index. We also plot the mass-metallicity relation from Zahid et al. (2011) and Huang et al. (2019), which inferred metallicity using the same line diagnostic for galaxies at a similar redshift. In general, we find that our metallicity range is consistent with previous measurements.

Following our previous analyses presented in Sok et al. (2025a), we also show the metallicity offset between clumpy and nonclumpy galaxies as the distribution of $\Delta Z$, defined to be the offset between individual metallicity values to the mass-metallicity relation. To this end, we fit the mass-metallicity relation using the parametrization given in Curti et al. (2020),

$$12 + \log(\text{O/H}) = Z_0 - \frac{\gamma}{\beta} \log\left[1 + \left(\frac{M}{M_0}\right)^{-\beta}\right], \tag{4.6}$$

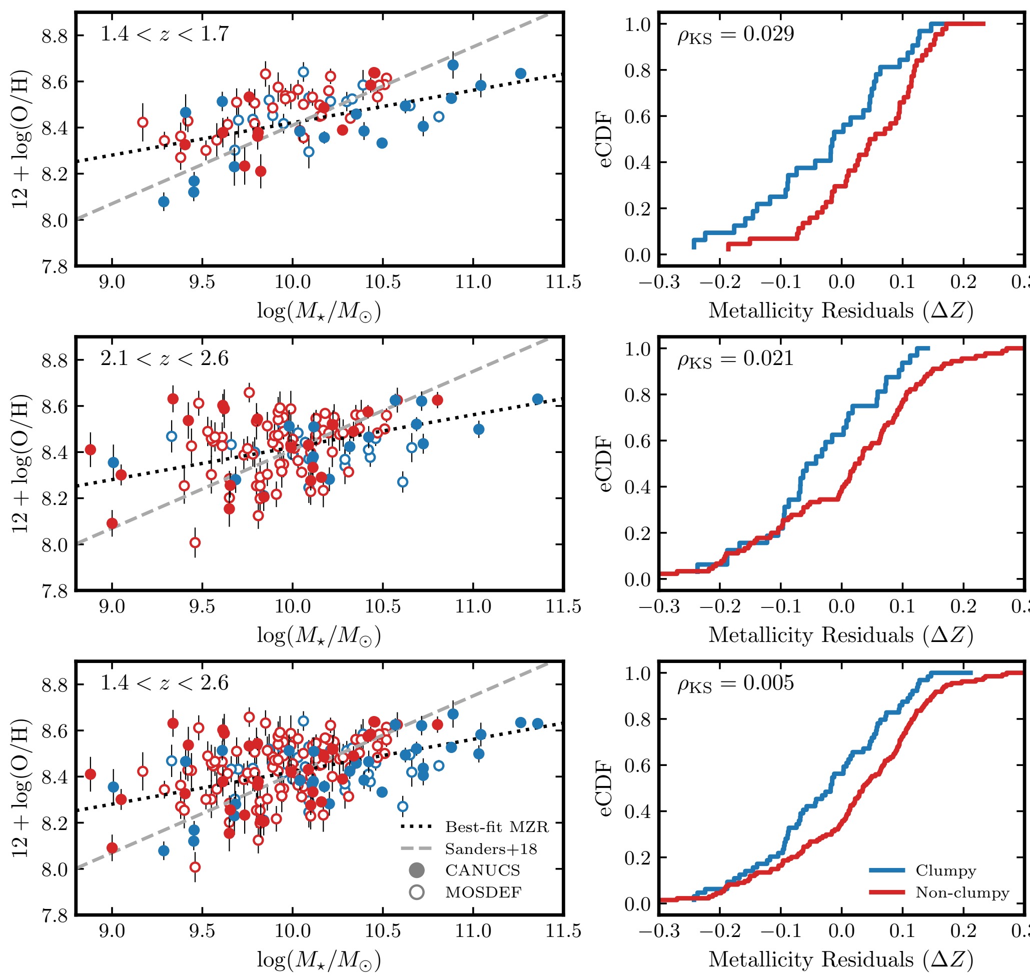


Figure 4.6: Similar to Figure 4.5, but now we show the mass-metallicity relation for the MOSDEF and CANUCS sample using the N2 index. The best-fitted mass-metallicity relation is determined by combining all galaxies between $1.4 < z < 2.6$ shown in the bottom panels.

where $Z_0$ represents the asymptotic metallicity at the high-mass end, and $M_0$ is the characteristic turnover mass above which the relation begins to flatten. At masses below $M_0$, the MZR follows a power-law behaviour with slope $\gamma$, while $\beta$ controls the sharpness of the transition between these regimes. The right panels of Figure 4.5 therefore show the empirical cumulative distribution function (eCDF) constructed from $\Delta Z$ for both clumpy and nonclumpy galaxies. We find that the eCDF of $\Delta Z$ for nonclumpy galaxies is generally shifted to the right of clumpy galaxies, suggesting that the typical values of $\Delta Z$ for nonclumpy galaxies are usually higher compared to clumpy galaxies.

Similarly, the left panels of Figure 4.6 show the distribution of the gas-phase metallicity inferred from the $N2$ index. We plot the best-fitting mass-metallicity relation, obtained by combining the sample of galaxies with $1.4 < z < 2.6$. We also plot the mass-metallicity relation from Sanders et al. (2018) to compare with our derived gas-phase metallicity. Their mass-metallicity relation was obtained from a sample of 260 SFGs in the MOSDEF sample at $z \sim 2.3$. We note that the slope of our mass-metallicity relation is shallower than that in the MOSDEF survey; however, using their best-fitted mass-metallicity relation does not affect the main findings of this study. Finally, we perform a two-sample Kolmogorov-Smirnov (KS) test to determine whether the eCDF of $\Delta Z$ for nonclumpy galaxies is systematically lower than that of clumpy galaxies. The resulting p-values $\rho_{KS}$ are reported in the right panels of both Figure 4.5 and 4.6, where a value of $\rho_{\rm KS} < 0.05$ indicates less than a five percent probability that the observed difference in eCDFs arises by chance. In general, we find that $\rho_{\rm KS} < 0.05$, indicating that the two eCDFs are statistically different, and that clumpy galaxies intrinsically have lower metallicities than nonclumpy galaxies. These results are consistent with what we reported in Sok et al. (2025a).

It is worth noting potential selection biases in the underlying galaxy sample and how these may affect the result. The CANUCS data consist of clumpy massive galaxies (above $10^{10.5}\ M_\odot$), with no nonclumpy counterparts. These clumpy galaxies also have lower metallicity compared to the mass-metallicity relation, and can likely bias our analyses. To assess whether this has any effect on our results, we perform the same analyses but within a selected mass range of $\log(M_*/M_\odot) < 10.5$. This gives a $\rho_{KS} = 0.03$ for the combined CANUCS and MOSDEF sample between $1.4 < z < 2.6$.

## 4.6 Stellar Population of Clumps

As shown in the previous section, clumpy galaxies are generally metal-poor compared to nonclumpy galaxies. The recent work of Estrada-Carpenter et al. (2025) further showed that clump populations consist of those with metallicities similar to the host disk and those that are lower. The galaxies that host lower metallicity clumps are generally found to have lower metallicities compared to those galaxies that host clumps with no metallicity offsets. These observations suggest that clump formation, at least those with lower gas-phase metallicity, is the result of a near-pristine gas accretion event.

An interesting question is how the properties of star-forming clumps differ between those above and below the mass-metallicity relation. The stellar masses and ages of clumps have been shown to provide insights into their origins (e.g., Ribeiro et al. 2017; Zanella et al. 2019; Sok et al. 2025b). In this section, we therefore examine clump properties as a function of their host galaxy's offset from the MZR. Here, we omit the LEGA-C samples because the deconvolved ground-based images have lower resolution. The following analyses therefore present only the derived properties for the CANUCS samples, including both HST and JWST imaging. Similar to Sok et al. (2025b), we only compute the stellar properties of clumps found in galaxies that are not highly magnified ($\mu < 2$) and that are not edge on (axial ratio of $q > 0.2$).

We divide the sample into clumps hosted by galaxies above the MZR ($\Delta Z > 0$) and those below it ($\Delta Z < 0$). In Figure 4.7, we present the distributions of clump stellar masses, ages and SED-derived SFR for these two samples. To improve statistics, we combine the two redshift bins, resulting in a sample of approximately 120 clumps between $1.3 < z < 2.6$. We find that clumps whose host galaxies lie below the MZR are generally less massive and younger, compared to clumps whose hosts lie above the MZR. Furthermore, clumps found in galaxies below the MZR also have higher SFRs (SED-derived) compared to those found above the MZR.

To quantify these differences, we perform the KS test to determine whether the mass and age distributions of clumps in galaxies below the MZR are smaller and younger than those above the MZR. For clump masses, the KS test returns a p-value greater than 0.05, suggesting that the masses are not statistically different. In contrast, both the age and SFR distributions yield p-values less

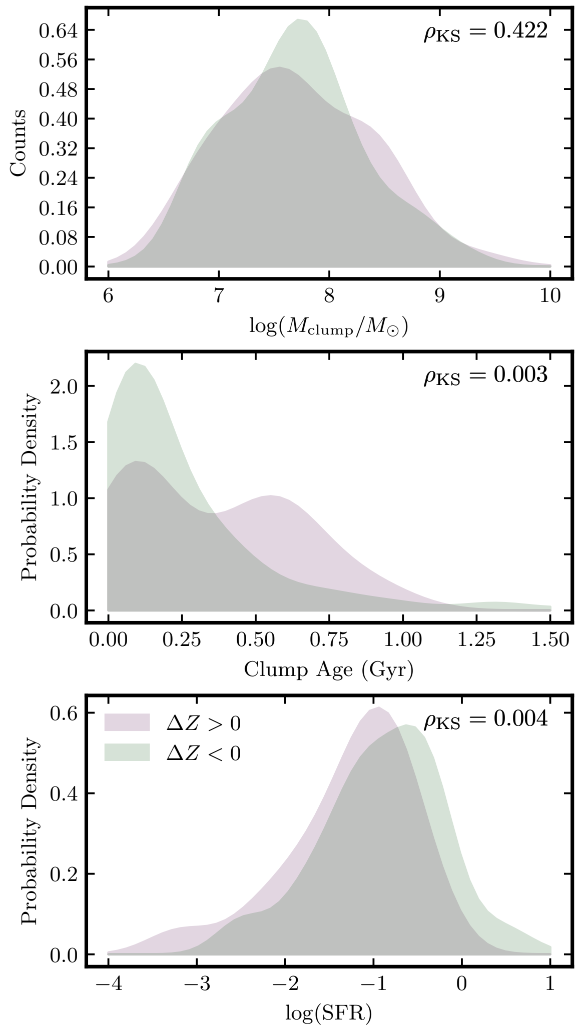


Figure 4.7: The kernel density estimate of clump masses, ages, and SED-derived SFR for clumps in galaxies with $\Delta Z > 0$ (purple) and $\Delta Z < 0$ (green). In general, we find that clumps in galaxies that lie below the mass-metallicity relation are younger and highly star-forming compared to clumps in galaxies that lie above the mass-metallicity relation.

than 0.05, implying that clumps below the MZR are statistically younger and exhibit more active star formation than those above the relation.

## 4.7 Discussion

The mass-metallicity relation (and similarly, the fundamental metallicity relation) provides an empirical framework for connecting galaxy evolution to physical processes such as star formation, the infall of metal-poor gas, and the outflow of enriched material. In particular, scatter in the gas-phase metallicity around the relation is often interpreted as being driven by the accretion of metal-poor gas and/or mergers. In this Section, we discuss whether mergers or accretion from cosmological streams drive the resulting metallicity offset and clump formation.

Generally, metal-poor gas accretion events are suggested to lead to the dilution of the gas-phase metallicity (e.g., Sánchez Almeida et al., 2014; Ceverino et al., 2016; Sok et al., 2025a; Estrada-Carpenter et al., 2025). On the other hand, mergers are also suggested to affect the scatter of the mass-metallicity relation. In interacting galaxies, the metallicity dilution and enrichment can be driven by the inflow of metal-poor or metal-rich gas already present within the merging galaxies. In the former case, metallicity dilution of around 0.1 dex is often attributed to major mergers (with a mass-ratio larger than $\mu > 1/3$), while weaker amplitudes in metallicity offsets are expected for minor mergers (Bustamante et al., 2018, 2020). This interpretation is often inferred from observing an increase in metal dilution with respect to decreasing projected separation between galaxy pairs (e.g., Omori and Takeuchi, 2022). However, the outcome of the enrichment process also likely depends on the initial metallicities of the interacting galaxies. For example, Michel-Dansac et al. (2008) found that the secondary member of minor mergers exhibits metallicities ∼0.2 dex higher compared to the mass–metallicity relation, possibly due to accretion of enriched gas from the more massive companion.

We further assess the merger scenario by comparing the observability timescales for the metallicity dilution and mergers. For a 1:1 merger scenario, the metallicity dilution phase can persist for $\sim 0.5 - 1.5$ Gyr (Montuori et al., 2010; Grønnow et al., 2015; Bustamante et al., 2018), whereas the merger observability timescale based on G–M20 morphology is 0.2–0.5 Gyr (Lotz et al., 2008, 2010). Noting that the age of the clumps is comparable ($<$ 500 Myr, Figure 4.7) to the merger observability timescale, it implies that if mergers were the primary driver for the observed metallicity offset in clumpy galaxies, we would expect to detect corresponding merger signatures in galaxies hosting clumps that are younger than 500 Myr.

We use a statistical morphological classification based on the Gini coefficient ($G$) versus $M_{20}$ (e.g., Lotz et al. 2004) to determine potential mergers. The Gini coefficient quantifies the inequality of a galaxy's pixel flux distribution, $M_{20}$ measures how the brightest 20% of the light is spatially distributed. We compute these nonparametric morphological quantities based on the stellar mass maps of galaxies using STATSMORPH (Rodriguez-Gomez et al., 2019), where mergers have a merger statistic that satisfies $S_{\rm merger}(G, M_{20}) > 0$. We adopt the updated merger statistics of Tan et al. (2024), which were calibrated against visual classifications in the CANUCS fields to identify mergers between $0.3 < z < 5$.

Out of the 69 galaxies in CANUCS with robust metallicities, only 27 of them are classified as clumpy. These clumpy galaxies are plotted in the $G - M_{20}$ parameter space in Figure 4.8. We find that only one galaxy is classified as being a merger. Additionally, the majority of clumpy galaxies with clumps younger than 500 Myr are not classified as mergers. This suggests that the majority of the galaxies in this sample (at least for the CANUCS sample) are likely experiencing metal-poor gas accretion from a cosmological stream rather than mergers.

We note that merger classification based solely on $S_{\rm merger}(G, M_{20})$ is generally not complete (e.g., Faisst et al. 2025), especially for systems in the final coalescence stage. Resolved kinematic observations are therefore needed to identify such systems. While such information is not available for the current sample, studies of <20 star-forming galaxies at $z \sim 0.8$ show that the kinematics of both clumpy and nonclumpy galaxies are broadly similar and consistent with rotation-dominated systems (Sok et al., 2025a). Nevertheless, mergers-driven clumps may be more critical at $z > 6$ (e.g., de la Vega et al. 2025), or in an overdense environment (Sattari et al., 2023).

In this interpretation, Figure 4.7 then suggests a scenario in which the accretion of metal-poor

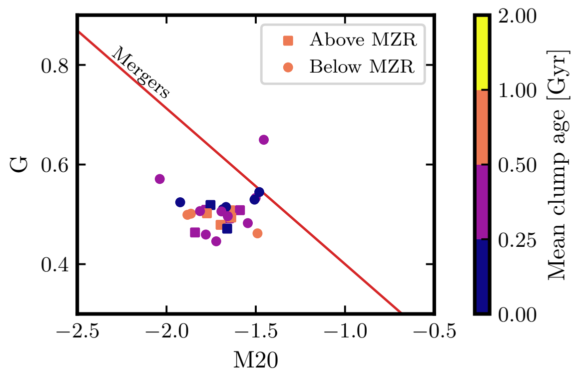


Figure 4.8: The Gini-M20 relation for clumpy galaxies in the CANUCS field. The red line denotes the dividing line between interacting and non-interacting systems.

gas drives metallicity dilution. Star formation then occurs within the accreted gas, as indicated by the presence of younger, highly star-forming clumps in galaxies that lie below the mass–metallicity relation. In contrast, the possible absence of younger, less massive clumps in galaxies that lie above the mass-metallicity relation can be explained by feedback. As stars form and feedback mechanisms shut off star formation, clumps are then subjected to mass-dependent disruptive processes, where lower-mass clumps are preferentially destroyed (e.g., Ginzburg et al. 2021; Dekel et al. 2022).

## 4.8 Chapter 4 Conclusion

The formation of star-forming clumps is often linked to violent disk instabilities, where instabilities in disks of high-$z$ galaxies arise from the continuous accretion of cosmological gas. Until recently, direct observational constraints on the link between clump formation and physical processes, such as gas inflow, have been limited. In particular, independent studies have begun to report evidence for gas inflow and metallicity dilution in clumpy galaxies, albeit for a small sample of $< 30$ galaxies (Sok et al., 2025a; Estrada-Carpenter et al., 2025).

In this study, we compile spectroscopic and photometric data from the LEGA-C and MOSDEF surveys, combined with new MOSFIRE observations of galaxies in the CANUCS fields, to investigate the relationship between gas-phase metallicity and galaxy morphology. We infer metallicities using the $R23$ and $N2$ indices in a sample of $\sim 300$ galaxies between $0.6 < z < 2.6$. The main findings are summarized as follows:

- Across all redshift bins, clumpy galaxies exhibit, on average, lower metallicities compared to non-clumpy galaxies, relative to the mass-metallicity relation (MZR).
- For clumps in the CANUCS field, we perform SED fitting to extract their masses, SFRs, and ages in clumpy galaxies. We find no significant mass differences between clumps in galaxies lying above or below the mass–metallicity relation (MZR).
- However, galaxies with metallicity below the MZR generally host younger and higher star-forming clumps, compared to galaxies with clumps that lie above the MZR.
- Based on the $S_{\mathrm{merger}}(G, M_{20})$ classification of the stellar mass maps, we find that only one

galaxy is consistent with being a merger. This indicates that lower metallicity observed in clumpy galaxies are more likely driven by metal-poor gas accretion events rather than by mergers.

Together, the results of this study support a scenario in which gas accretion events dilute metal content in galaxies, as reflected in the young stellar populations of clumpy regions. Furthermore, clump populations in galaxies with metallicity above the MZR are generally older and show less star formation than those below the MZR, suggesting possibly episodic accretion events. Future work incorporating resolved kinematic studies will be essential to disentangle the relative roles of gas inflow and mergers in driving metallicity dilution and clump formation.

## Acknowledgments

Some of the data presented herein were obtained at Keck Observatory, which is a private 501(c)3 non-profit organization operated as a scientific partnership among the California Institute of Technology, the University of California, and the National Aeronautics and Space Administration. The Observatory was made possible by the generous financial support of the W. M. Keck Foundation.

The authors wish to recognize and acknowledge the very significant cultural role and reverence that the summit of Maunakea has always had within the Native Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain.

# Chapter 5

# Conclusions

"The universe has many horrors yet to throw at us. This is not the end of our struggle. This is just the beginning of our crusade to save Humanity."
— Emperor of Mankind, Warhammer 40K Lore

## 5.1 Summary

Across cosmic time, star formation occurs with scales ranging from a few parsecs, as seen in star clusters in the nearby Universe, to sub-kiloparsec scales in clumpy star-forming regions near cosmic noon. These sub-kiloparsec clumps are intriguing because of their abundance at higher redshifts and their absence in the local Universe. Notably, the fraction of star-forming galaxies that host clumps seems to follow the star formation rate density of the Universe (e.g., Shibuya et al. 2016; de la Vega et al. 2025; Sok et al. 2025b), indicating that clumps are closely tied to the mass assembly of galaxies. Understanding the mechanisms that drive clump formation is therefore essential for constraining the physics of galaxy evolution.

Section 1.2 stages the questions asked in this dissertation by highlighting previous results on the nature of clumps. Clumps are suggested to form due to violent disk instability. At the same time, major mergers may play a secondary role, as indicated by the redshift evolution of the fraction of clumpy galaxies and merger rates (e.g., Guo et al., 2015; Nakazato et al., 2024; de la Vega et al., 2025). Although we have a broad understanding of the origins of clumps and their evolution, a coherent picture of the detailed physics driving these processes remains elusive, mainly due to a lack of observational data that can directly test them. Section 1.3 then sets the aims for this work, namely to address these long-standing questions by examining the relationship between the resolved properties of clumps and global properties of the host galaxy. This dissertation is not possible without the data from ground-based spectroscopic observations with GNIRS (Sok et al., 2025a) and MOSFIRE (Sok et al. in prep.), but also from high-resolution JWST CANUCS imaging (Willott et al., 2022) of thousands of star-forming galaxies.

The formation of clumps is commonly suggested to be driven by violent disk instability as the result of gas accretion (e.g., Dekel et al. 2009b; van de Voort and Schaye 2012; Ceverino et al. 2016), where the gas contents are generally assumed to be metal-poor. This framework for clump formation is motivated by other studies that found that galaxies at cosmic noon are generally gas-rich (Tacconi et al., 2010; Daddi et al., 2010). In Chapter 2, we conduct a pilot study to provide one of the first observational constraints on whether a physical process, such as an inflow of metal-poor gas, drives clump formation. In measuring the gas-phase metallicity for a sample of $\sim 30$ galaxies, we find evidence that clump formation in clumpy galaxies is likely the result of metal-poor accretion events, as suggested by lower metallicities observed in clumpy galaxies compared to non-clumpy galaxies. We further find that the kinematics of clumpy galaxies are not significantly different from those of nonclumpy galaxies. Particularly, both clumpy and nonclumpy galaxies are consistent with being rotationally dominated. Several interpretation is drawn from this, where (1) this suggests that mergers may not be responsible for clump formation and (2) whatever processes drive clump formation act on short timescales.

Chapter 3 further examines clump formation and evolution. In particular, we identify clumpy galaxies and their clumps in the CANUCS field and obtain resolved clump masses and ages via SED fitting. We find that the aggregated clump stellar mass function (cSMF) has a power-law slope of $\alpha = -2$, consistent with a picture where clumps formed *in-situ* through turbulent-induced gravitational instabilities (e.g., Dessauges-Zavadsky and Adamo 2018; Claeyssens et al. 2025). When separating by mass, we find that the slope of the cSMF is flatter for massive galaxies ($> 10^{10}$ solar masses) and shallower for less massive galaxies ($< 10^{10}$ solar masses). This result suggests that massive galaxies generally form more massive clumps, most likely due to their larger disks (e.g., from Toomre analyses, as shown by Dekel et al. 2009b) and stronger rotationally supported disks compared to less massive systems (Orr and Rennehan, 2024). Additionally, the initial mass function depends on the gas fraction (as inferred from the empirical relation described in Wang et al. 2022), where massive clumps are preferably formed at higher gas fraction. The slope of the clump stellar mass function becomes progressively flatter with increasing clump age, which may reflect the combined effects of processes such as feedback, disruption, or clump–clump interaction. However, it should be noted here that this could also be driven by selection effects, where selecting based on the UV biases the identification of low-mass, older clumps.

Finally, Chapter 4 expands on the work of Chapter 2, by compiling and investigating the relationship between the gas-phase metallicity of galaxies and their morphologies of a sample of $\sim 300$. Similar to the previous work (Sok et al., 2025a), we find that clumpy galaxies generally have lower metallicity than nonclumpy galaxies. A subset of these galaxies is in the CANUCS fields, enabling us to extract their stellar masses, ages and SFR distributions (Sok et al., 2025b). These galaxies are not mergers, as suggested by their $G$ vs $M_{20}$ morphological parameter (e.g., Lotz et al. 2004). We also compare clump populations in galaxies with gas-phase metallicity below the mass-metallicity relation $\Delta Z < 0$ to those above $\Delta Z > 0$. We find that galaxies with $\Delta Z < 0$ host younger clumps that are still actively forming stars.

Overall, the findings presented in this work paint a consistent picture that depicts the rise and fall of clumps across cosmic time. Clump formation is attributed to a metal-poor gas accretion event, leading to the formation of new clumps. The properties of clumps also seem to depend on those of the host gas disk, with a higher gas fraction favouring clump formation. The mass distribution of clumps further dictates their survivability and evolution.

## 5.2 Outlook for the future

A wise man once told me that science is never done, and answering questions will lead to new questions that need to be addressed. Therefore, new questions should be left for another day. The findings above undoubtedly shed light on the mechanisms driving clump formation and their evolution. Nevertheless, more work is needed to unambiguously determine mechanisms driving clump formation. In particular, I highlight two avenues of research that can provide fruitful progress in the field.

### 5.2.1 Linking Clump Formation to their Fuel Supply

The findings in Chapter 3 show that the nature of clumps is inevitably tied to their gas supply. In particular, the rise and fall of clumps seemingly follow the redshift evolution of the molecular gas density (see Fig. 3.6). However, observational constraints on the molecular gas content of clumpy galaxies remain limited. Future work should therefore focus on probing the molecular gas of both clumps and their host galaxies to address the fuel that drives clump formation.

Molecular gas is composed mainly of $H_2$, which is difficult to detect directly. As a result, mapping molecular gas in galaxies relies on indirect tracers such as CO, which are accessible with ALMA observations (e.g., Tacconi et al. 2020). Combined with lensing fields from CANUCS, ALMA observations of low-$J$ CO transitions will trace the resolved molecular gas density, providing direct constraints on the fuel supply to clumps. In particular, the data will enable tests of how clump formation is linked to disk dynamics, as inferred from the CO lines, and whether more and more massive clumps are formed in a rotationally-supported gas-rich disk. Moreover, when combined with existing resolved stellar properties of clumps, such observations will provide crucial connections between total gas fraction, clump masses and the gravitational boundness of clumps (as inferred by their dynamical and free-fall timescales). These are needed to link the baryonic cycle from gas accretion to star formation to stellar feedback. Of course, obtaining resolved CO maps with ALMA is time-intensive and requires substantial allocations of observing time.

### 5.2.2 Is It Actually Just Raining Clumps?

The prevailing theory of clump formation is the accretion of low-metallicity gas onto a gas disk, which leads to violent disk instability. Chapter 3 shows that the clump stellar mass function is consistent with predominantly *in-situ* formation. The findings from Chapter 2 and Chapter 4 suggest that clump origins are linked to a gas accretion event. At the same time, recently resolved metallicity maps of a sample of $\sim 20$ clumpy galaxies near $z \simeq 1$ obtained with JWST NIRISS grism observations surprisingly show that some clumps also have low metallicity compared to the disks in which they reside (Estrada-Carpenter et al., 2025). Such a finding is unexpected for clumps formed *in-situ* as gas mixing is expected to equilibrate the metallicity, and instead suggests that some clumps may be *ex-situ* gas clumps that are "raining" in. Together, these results highlight that the detailed physics of clump formation remains unresolved.

Discriminating between clumps of *in-situ* and *ex-situ* is not straightforward from visual classification alone (e.g., Faisst et al. 2025). A more powerful approach is to examine their kinematics using high-resolution integral-field spectroscopy (IFS). As suggested by Noguchi (1999), the kinematics of clumps are expected to be coplanar if they formed via fragmentation, whereas clumps induced by mergers or raining in will exhibit random orientations. Resolved IFS data also enable simultaneous measurements of metallicity and kinematic offsets, as demonstrated by Morishita et al. (2025). Extending such observations for a larger sample will provide the necessary data to distinguish clumps of *in-situ* origin and those of *ex-situ* origin, whether merger-induced or accreting.

Unfortunately, resolved IFS observations are also time-intensive. Forthcoming instruments and next-generation telescopes are poised to make such studies more feasible. In particular, instruments such as the Gemini Infrared Multi-Object Spectrograph (GIRMOS, Sivanandam et al. 2024) will offer a novel technique for simultaneously correcting atmospheric aberrations across multiple objects. Combined with the imminent arrival of 30+ meter ground-based telescopes, we are also entering an exciting period for the study of galaxy evolution. The increased photon-collecting area of these facilities, together with adaptive–optics–assisted spectroscopy, will enable the characterization of resolved properties of star-forming clumps across statistically significant samples of galaxies.

# Bibliography


Abdurro'uf, Yen-Ting Lin, Po-Feng Wu, and Masayuki Akiyama. Introducing piXedfit: A Spectral Energy Distribution Fitting Code Designed for Resolved Sources. ApJS, 254(1):15, May 2021. doi: 10.3847/1538-4365/abebe2.

R. G. Abraham, N. R. Tanvir, B. X. Santiago, R. S. Ellis, K. Glazebrook, and S. van den Bergh. Galaxy morphology to I=25 mag in the Hubble Deep Field. MNRAS, 279(3):L47–L52, April 1996. doi: 10.1093/mnras/279.3.L47.

A. Adamo, K. Hollyhead, M. Messa, J. E. Ryon, V. Bajaj, A. Runnholm, S. Aalto, D. Calzetti, J. S. Gallagher, M. J. Hayes, J. M. D. Kruijssen, S. König, S. S. Larsen, J. Melinder, E. Sabbi, L. J. Smith, and G. Östlin. Star cluster formation in the most extreme environments: insights from the HiPEEC survey. MNRAS, 499(3):3267–3294, December 2020. doi: 10.1093/mnras/staa2380.

Angela Adamo, Larry D. Bradley, Eros Vanzella, Adélaïde Claeyssens, Brian Welch, Jose M Diego, Guillaume Mahler, Masamune Oguri, Keren Sharon, Abdurro'uf, Tiger Yu-Yang Hsiao, Xinfeng Xu, Matteo Messa, Augusto E. Lassen, Erik Zackrisson, Gabriel Brammer, Dan Coe, Vasily Kokorev, Massimo Ricotti, Adi Zitrin, Seiji Fujimoto, Akio K. Inoue, Tom Resseguier, Jane R. Rigby, Yolanda Jiménez-Teja, Rogier A. Windhorst, Takuya Hashimoto, and Yoichi Tamura. Bound star clusters observed in a lensed galaxy 460 Myr after the Big Bang. *arXiv e-prints*, art. arXiv:2401.03224, January 2024. doi: 10.48550/arXiv.2401.03224.

Dominic Adams, Vihang Mehta, Hugh Dickinson, Claudia Scarlata, Lucy Fortson, Sandor Kruk, Brooke Simmons, and Chris Lintott. Galaxy Zoo: Clump Scout: Surveying the Local Universe for Giant Star-forming Clumps. ApJ, 931(1):16, May 2022. doi: 10.3847/1538-4357/ac6512.

Oscar Agertz, Romain Teyssier, and Ben Moore. Disc formation and the origin of clumpy galaxies at high redshift. MNRAS, 397(1):L64–L68, July 2009. doi: 10.1111/j.1745-3933.2009.00685.x.

Liyualem Ambachew, Deanne B. Fisher, Karl Glazebrook, Marianne Girard, Danail Obreschkow, Roberto Abraham, Alberto Bolatto, Laura Lenkić, and Ivana Damjanov. Stellar masses of clumps in gas-rich, turbulent disc galaxies. MNRAS, 512(2):3079–3097, May 2022. doi: 10.1093/mnras/stac622.

Jacqueline Antwi-Danso, Casey Papovich, Joel Leja, Danilo Marchesini, Z. Cemile Marsan, Nicholas S. Martis, Ivo Labbé, Adam Muzzin, Karl Glazebrook, Caroline M. S. Straatman, and Kim-Vy H. Tran. Beyond UVJ: Color Selection of Galaxies in the JWST Era. ApJ, 943(2):166, February 2023. doi: 10.3847/1538-4357/aca294.

C. Bacchini, C. Nipoti, G. Iorio, F. Roman-Oliveira, F. Rizzo, P. E. Mancera Piña, A. Marasco, A. Zanella, and F. Lelli. A 3D view on the local gravitational instability of cold gas discs in star-forming galaxies at $0 \lesssim z \lesssim 5$. A&A, 687:A115, July 2024. doi: 10.1051/0004-6361/202449925.

M. Behrendt, M. Schartmann, and A. Burkert. The possible hierarchical scales of observed clumps in high-redshift disc galaxies. MNRAS, 488(1):306–323, September 2019. doi: 10.1093/mnras/stz1717.

James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. 2008.

Frédéric Bournaud, Bruce G. Elmegreen, and Debra Meloy Elmegreen. Rapid Formation of Exponential Disks and Bulges at High Redshift from the Dynamical Evolution of Clump-Cluster and Chain Galaxies. ApJ, 670(1):237–248, November 2007. doi: 10.1086/522077.

Frédéric Bournaud, Leila C. Powell, Damien Chapon, and Romain Teyssier. Star formation in galaxy mergers: ISM turbulence, dense gas excess, and scaling relations for disks and starbusts. In Nicholas H. Brummell, A. Sacha Brun, Mark S. Miesch, and Yannick Ponty, editors, *Astrophysical Dynamics: From Stars to Galaxies*, volume 271 of *IAU Symposium*, pages 160–169, August 2011. doi: 10.1017/S1743921311017571.

Frédéric Bournaud, Valentin Perret, Florent Renaud, Avishai Dekel, Bruce G. Elmegreen, Debra M. Elmegreen, Romain Teyssier, Philippe Amram, Emanuele Daddi, Pierre-Alain Duc, David Elbaz, Benoit Epinat, Jared M. Gabor, Stéphanie Juneau, Katarina Kraljic, and Emeric Le Floch'. The Long Lives of Giant Clumps and the Birth of Outflows in Gas-rich Galaxies at High Redshift. ApJ, 780(1):57, January 2014. doi: 10.1088/0004-637X/780/1/57.

Maruša Bradač, Jon Judež, Chris Willott, Gregor Rihtaršič, Nicholas S. Martis, Anishya Harshan, Giordano Felicioni, Yoshihisa Asada, Guillaume Desprez, Douglas Clowe, Anthony H. Gonzalez, Christine Jones, Brian C. Lemaux, Vladan Markov, Lamiya Mowla, Gaël Noirot, Annika H. G. Peter, Andrew Robertson, Ghassan T. E. Sarrouh, Marcin Sawicki, Tim Schrabback, and Roberta Tripodi. Star Formation under a Cosmic Microscope: Highly magnified $z = 11$ galaxy behind the Bullet Cluster. *arXiv e-prints*, art. arXiv:2509.20446, September 2025. doi: 10.48550/arXiv.2509.20446.

Larry Bradley, Brigitta Sipőcz, Thomas Robitaille, Erik Tollerud, Zè Vinícius, Christoph Deil, Kyle Barbary, Tom J Wilson, Ivo Busko, Axel Donath, Hans Moritz Günther, Mihai Cara, P. L. Lim, Sebastian Meßlinger, Simon Conseil, Zach Burnett, Azalee Bostroem, Michael Droettboom, E. M. Bray, Lars Andersen Bratholm, Adam Ginsburg, William Jamieson, Geert Barentsen, Matt Craig, Brett M. Morris, Marshall Perrin, Shivangee Rathi, Sergio Pascual, and Iskren Y. Georgiev. astropy/photutils: 2.0.2, October 2024. URL `https://doi.org/10.5281/zenodo.13989456`.

Gabriel B. Brammer, Pieter G. van Dokkum, and Paolo Coppi. EAZY: A Fast, Public Photometric Redshift Code. ApJ, 686(2):1503–1513, October 2008. doi: 10.1086/591786.

Gabriel B. Brammer, Pieter G. van Dokkum, Marijn Franx, Mattia Fumagalli, Shannon Patel, Hans-Walter Rix, Rosalind E. Skelton, Mariska Kriek, Erica Nelson, Kasper B. Schmidt, Rachel Bezanson, Elisabete da Cunha, Dawn K. Erb, Xiaohui Fan, Natascha Förster Schreiber, Garth D. Illingworth, Ivo Labbé, Joel Leja, Britt Lundgren, Dan Magee, Danilo Marchesini, Patrick McCarthy, Ivelina Momcheva, Adam Muzzin, Ryan Quadri, Charles C. Steidel, Tomer Tal, David Wake, Katherine E. Whitaker, and Anna Williams. 3D-HST: A Wide-field Grism Spectroscopic Survey with the Hubble Space Telescope. ApJS, 200(2):13, June 2012. doi: 10.1088/0067-0049/200/2/13.

Sebastián Bustamante, Martin Sparre, Volker Springel, and Robert J. J. Grand. Merger-induced metallicity dilution in cosmological galaxy formation simulations. MNRAS, 479(3):3381–3392, September 2018. doi: 10.1093/mnras/sty1692.

Sebastián Bustamante, Sara L. Ellison, David R. Patton, and Martin Sparre. Galaxy pairs in the Sloan Digital Sky Survey - XIV. Galaxy mergers do not lie on the fundamental metallicity relation. MNRAS, 494(3):3469–3480, May 2020. doi: 10.1093/mnras/staa1025.

Marcello Cacciato, Avishai Dekel, and Shy Genel. Evolution of violent gravitational disc instability in galaxies: late stabilization by transition from gas to stellar dominance. MNRAS, 421(1):818–831, March 2012. doi: 10.1111/j.1365-2966.2011.20359.x.

Antonello Calabrò, Emanuele Daddi, Jérémy Fensch, Frédéric Bournaud, Anna Cibinel, Annagrazia Puglisi, Shuowen Jin, Ivan Delvecchio, and Chiara D'Eugenio. Merger induced clump formation in distant infrared luminous starburst galaxies. A&A, 632:A98, December 2019. doi: 10.1051/0004-6361/201935778.

Daniela Calzetti, Lee Armus, Ralph C. Bohlin, Anne L. Kinney, Jan Koornneef, and Thaisa Storchi-Bergmann. The Dust Content and Opacity of Actively Star-forming Galaxies. ApJ, 533(2):682–695, April 2000. doi: 10.1086/308692.

N. Cantale, F. Courbin, M. Tewes, P. Jablonka, and G. Meylan. Firedec: a two-channel finite-resolution image deconvolution algorithm. A&A, 589:A81, May 2016a. doi: 10.1051/0004-6361/201424003.

Nicolas Cantale, Pascale Jablonka, Frédéric Courbin, Gregory Rudnick, Dennis Zaritsky, Georges Meylan, Vandana Desai, Gabriella De Lucia, Alfonso Aragón-Salamanca, Bianca M. Poggianti, Rose Finn, and Luc Simard. Disc colours in field and cluster spiral galaxies at $0.5 \lesssim z \lesssim 0.8$. A&A, 589:A82, May 2016b. doi: 10.1051/0004-6361/201525801.

Michele Cappellari and Yannick Copin. Adaptive spatial binning of integral-field spectroscopic data using Voronoi tessellations. MNRAS, 342(2):345–354, June 2003. doi: 10.1046/j.1365-8711.2003.06541.x.

David Carton, Jarle Brinchmann, Thierry Contini, Benoît Epinat, Hayley Finley, Johan Richard, Vera Patrício, Joop Schaye, Themiya Nanayakkara, Peter M. Weilbacher, and Lutz Wisotzki. First gas-phase metallicity gradients of $0.1 \lesssim z \lesssim 0.8$ galaxies with MUSE. MNRAS, 478(4):4293–4316, August 2018. doi: 10.1093/mnras/sty1343.

Antonio Cava, Daniel Schaerer, Johan Richard, Pablo G. Pérez-González, Miroslava Dessauges-Zavadsky, Lucio Mayer, and Valentina Tamburello. The nature of giant clumps in distant galaxies probed by the anatomy of the cosmic snake. *Nature Astronomy*, 2:76–82, November 2018. doi: 10.1038/s41550-017-0295-x.

Daniel Ceverino, Avishai Dekel, and Frederic Bournaud. High-redshift clumpy discs and bulges in cosmological simulations. MNRAS, 404(4):2151–2169, June 2010. doi: 10.1111/j.1365-2966.2010.16433.x.

Daniel Ceverino, Jorge Sánchez Almeida, Casiana Muñoz Tuñón, Avishai Dekel, Bruce G. Elmegreen, Debra M. Elmegreen, and Joel Primack. Gas inflow and metallicity drops in star-forming galaxies. MNRAS, 457(3):2605–2612, April 2016. doi: 10.1093/mnras/stw064.

Gilles Chabrier. Galactic Stellar and Substellar Initial Mass Function. PASP, 115(809):763–795, July 2003. doi: 10.1086/376392.

R. Cid Fernandes, G. Stasińska, M. S. Schlickmann, A. Mateus, N. Vale Asari, W. Schoenell, and L. Sodré. Alternative diagnostic diagrams and the 'forgotten' population of weak line galaxies in the SDSS. MNRAS, 403(2):1036–1053, April 2010. doi: 10.1111/j.1365-2966.2009.16185.x.

R. Cid Fernandes, E. Pérez, R. García Benito, R. M. González Delgado, A. L. de Amorim, S. F. Sánchez, B. Husemann, J. Falcón Barroso, P. Sánchez-Blázquez, C. J. Walcher, and D. Mast. Resolving galaxies in time and space. I. Applying STARLIGHT to CALIFA datacubes. A&A, 557:A86, September 2013. doi: 10.1051/0004-6361/201220616.

Adélaïde Claeyssens, Angela Adamo, Johan Richard, Guillaume Mahler, Matteo Messa, and Miroslava Dessauges-Zavadsky. Star formation at the smallest scales: a JWST study of the clump populations in SMACS0723. MNRAS, 520(2):2180–2203, April 2023. doi: 10.1093/mnras/stac3791.

Adélaïde Claeyssens, Angela Adamo, Matteo Messa, Miroslava Dessauges-Zavadsky, Johan Richard, Ivan Kramarenko, Jorryt Matthee, and Rohan P. Naidu. Tracing star formation across cosmic time at tens of parsec-scales in the lensing cluster field Abell 2744. MNRAS, 537(3):2535–2558, March 2025. doi: 10.1093/mnras/staf058.

Charlie Conroy and James E. Gunn. FSPS: Flexible Stellar Population Synthesis. Astrophysics Source Code Library, record ascl:1010.043, October 2010.

Lennox L. Cowie, Esther M. Hu, and Antoinette Songaila. Faintest Galaxy Morphologies From HST WFPC2 Imaging of the Hawaii Survey Fields. AJ, 110:1576, October 1995. doi: 10.1086/117631.

G. Cresci, F. Mannucci, R. Maiolino, A. Marconi, A. Gnerucci, and L. Magrini. Gas accretion as the origin of chemical abundance gradients in distant galaxies. Nature, 467(7317):811–813, October 2010. doi: 10.1038/nature09451.

Mirko Curti, Filippo Mannucci, Giovanni Cresci, and Roberto Maiolino. The mass-metallicity and the fundamental metallicity relation revisited on a fully $T_e$-based abundance scale for galaxies. MNRAS, 491(1):944–964, January 2020. doi: 10.1093/mnras/stz2910.

E. Daddi, F. Bournaud, F. Walter, H. Dannerbauer, C. L. Carilli, M. Dickinson, D. Elbaz, G. E. Morrison, D. Riechers, M. Onodera, F. Salmi, M. Krips, and D. Stern. Very High Gas Fractions and Extended Gas Reservoirs in $z = 1.5$ Disk Galaxies. ApJ, 713(1):686–707, April 2010. doi: 10.1088/0004-637X/713/1/686.

Nicolò Dalmasso, Antonello Calabrò, Nicha Leethochawalit, Benedetta Vulcani, Kristan Boyett, Michele Trenti, Tommaso Treu, Marco Castellano, Maruša Bradač, Benjamin Metha, and Paola Santini. The rate and contribution of mergers to mass assembly from NIRCam observations of galaxy candidates up to 13.3 billion years ago. MNRAS, 533(4):4472–4484, October 2024. doi: 10.1093/mnras/stae2064.

M. Davis, P. Guhathakurta, N. P. Konidaris, J. A. Newman, M. L. N. Ashby, A. D. Biggs, P. Barmby, K. Bundy, S. C. Chapman, A. L. Coil, C. J. Conselice, M. C. Cooper, D. J. Croton, P. R. M. Eisenhardt, R. S. Ellis, S. M. Faber, T. Fang, G. G. Fazio, A. Georgakakis, B. F. Gerke, W. M. Goss, S. Gwyn, J. Harker, A. M. Hopkins, J. S. Huang, R. J. Ivison, S. A. Kassin, E. N. Kirby, A. M. Koekemoer, D. C. Koo, E. S. Laird, E. Le Floc'h, L. Lin, J. M. Lotz, P. J. Marshall, D. C. Martin, A. J. Metevier, L. A. Moustakas, K. Nandra, K. G. Noeske, C. Papovich, A. C. Phillips, R. M. Rich, G. H. Rieke, D. Rigopoulou, S. Salim, D. Schiminovich, L. Simard, I. Smail, T. A. Small, B. J. Weiner, C. N. A. Willmer, S. P. Willner, G. Wilson, E. L. Wright, and R. Yan. The All-Wavelength Extended Groth Strip International Survey (AEGIS) Data Sets. ApJ, 660(1): L1–L6, May 2007. doi: 10.1086/517931.

Alexander de la Vega, Bahram Mobasher, Faezeh Manesh, Niloofar Sharei, Nima Chartab, and Zahra Sattari. The Fraction of Clumpy Galaxies in JWST Surveys over $2 < z < 12$. *arXiv e-prints*, art. arXiv:2508.14972, August 2025.

Mithi A. de los Reyes, Chun Ly, Janice C. Lee, Samir Salim, Molly S. Peeples, Ivelina Momcheva, Jesse Feddersen, Daniel A. Dale, Masami Ouchi, Yoshiaki Ono, and Rose Finn. The Relationship between Stellar Mass, Gas Metallicity, and Star Formation Rate for H$\alpha$-Selected Galaxies at z $\approx$ 0.8 from the NewH$\alpha$ Survey. AJ, 149(2):79, February 2015. doi: 10.1088/0004-6256/149/2/79.

Gabriella De Lucia, Lizhi Xie, Fabio Fontanot, and Michaela Hirschmann. Gas accretion regulates the scatter of the mass-metallicity relation. MNRAS, 498(3):3215–3227, November 2020. doi: 10.1093/mnras/staa2556.

A. Dekel, Y. Birnboim, G. Engel, J. Freundlich, T. Goerdt, M. Mumcuoglu, E. Neistein, C. Pichon, R. Teyssier, and E. Zinger. Cold streams in early massive hot haloes as the main mode of galaxy formation. Nature, 457(7228):451–454, January 2009a. doi: 10.1038/nature07648.

Avishai Dekel, Re'em Sari, and Daniel Ceverino. Formation of Massive Galaxies at High Redshift: Cold Streams, Clumpy Disks, and Compact Spheroids. ApJ, 703(1):785–801, September 2009b. doi: 10.1088/0004-637X/703/1/785.

Avishai Dekel, Nir Mandelker, Frederic Bournaud, Daniel Ceverino, Yicheng Guo, and Joel Primack. Clump survival and migration in VDI galaxies: an analytical model versus simulations and observations. MNRAS, 511(1):316–340, March 2022. doi: 10.1093/mnras/stab3810.

Avishai Dekel, Offek Tziperman, Kartick C. Sarkar, Omri Ginzburg, Nir Mandelker, Daniel Ceverino, and Joel Primack. Conditions for clump survival in High-z disc galaxies. MNRAS, 521(3):4299–4322, May 2023. doi: 10.1093/mnras/stad855.

Miroslava Dessauges-Zavadsky and Angela Adamo. First constraints on the stellar mass function of star-forming clumps at the peak of cosmic star formation. MNRAS, 479(1):L118–L122, September 2018. doi: 10.1093/mnrasl/sly112.

Miroslava Dessauges-Zavadsky, Daniel Schaerer, Antonio Cava, Lucio Mayer, and Valentina Tamburello. On the Stellar Masses of Giant Clumps in Distant Star-forming Galaxies. ApJ, 836(2): L22, February 2017. doi: 10.3847/2041-8213/aa5d52.

P. Di Matteo, F. Bournaud, M. Martig, F. Combes, A. L. Melchior, and B. Semelin. On the frequency, intensity, and duration of starburst episodes triggered by galaxy interactions and mergers. A&A, 492(1):31–49, December 2008. doi: 10.1051/0004-6361:200809480.

Qiao Duan, Christopher J. Conselice, Qiong Li, Duncan Austin, Thomas Harvey, Nathan J. Adams, Kenneth J. Duncan, James Trussler, Leonardo Ferreira, Lewi Westcott, Honor Harris, Rogier A. Windhorst, Benne W. Holwerda, Thomas J. Broadhurst, Dan Coe, Seth H. Cohen, Xiaojing Du, Simon P. Driver, Brenda Frye, Norman A. Grogin, Nimish P. Hathi, Rolf A. Jansen, Anton M. Koekemoer, Madeline A. Marshall, Mario Nonino, Rafael Ortiz, III, Nor Pirzkal, Aaron Robotham, Russell E. Ryan, Jake Summers, Jordan C. J. D'Silva, Christopher N. A. Willmer, and Haojing Yan. Galaxy mergers in the epoch of reionization – I. A JWST study of pair fractions, merger rates, and stellar mass accretion rates at $z = 4.5$–$11.5$. MNRAS, 540(1):774–805, June 2025. doi: 10.1093/mnras/staf638.

Debra Meloy Elmegreen, Bruce G. Elmegreen, Douglas S. Rubin, and Meredith A. Schaffer. Galaxy Morphologies in the Hubble Ultra Deep Field: Dominance of Linear Structures at the Detection Limit. ApJ, 631(1):85–100, September 2005. doi: 10.1086/432502.

Dawn K. Erb. A Model for Star Formation, Gas Flows, and Chemical Evolution in Galaxies at High Redshifts. ApJ, 674(1):151–156, February 2008. doi: 10.1086/524727.

Vicente Estrada-Carpenter, Marcin Sawicki, Gabe Brammer, Guillaume Desprez, Roberto Abraham, Yoshihisa Asada, Maruša Bradač, Kartheik G. Iyer, Nicholas S. Martis, Jasleen Matharu, Lamiya Mowla, Adam Muzzin, Gaël Noirot, Ghassan T. E. Sarrouh, Victoria Strait, and Chris J. Willott. When, where, and how star formation happens in a galaxy pair at cosmic noon using CANUCS JWST/NIRISS grism spectroscopy. MNRAS, 532(1):577–591, July 2024. doi: 10.1093/mnras/stae1368.

Vicente Estrada-Carpenter, Marcin Sawicki, Roberto Abraham, Yoshihisa Asada, Maruša Bradač, Gabe Brammer, Guillaume Desprez, Kartheik G. Iyer, Nicholas S. Martis, Adam Muzzin, Gaël Noirot, Gregor Rihtaršič, Ghassan T. E. Sarrouh, Chris J. Willott, Jeremy Favaro, Vladan Markov, Rosa M. Mérida, Katherine Myers, and Visal Sok. Metal-Poor Star-Forming Clumps in Cosmic Noon Galaxies: Evidence for Gas Inflow and Chemical Dilution Using JWST NIRISS. *arXiv e-prints*, art. arXiv:2508.00985, August 2025. doi: 10.48550/arXiv.2508.00985.

Andreas L. Faisst, Lilan Yang, M. Brinch, C. M. Casey, N. Chartab, M. Dessauges-Zavadsky, N. E. Drakos, S. Gillman, G. Gozaliasl, C. C. Hayward, O. Ilbert, P. Jablonka, A. Kaminsky, J. S. Kartaltepe, A. M. Koekemoer, V. Kokorev, E. Lambrides, D. Liu, C. Maraston, C. L. Martin, A. Renzini, B. E. Robertson, D. B. Sanders, Z. Sattari, N. Scoville, C. M. Urry, A. P. Vijayan, J. R. Weaver, H. B. Akins, N. Allen, R. C. Arango-Toro, O. R. Cooper, M. Franco, F. Gentile, S. Harish, M. Hirschmann, A. A. Khostovan, C. Laigle, R. L. Larson, M. Lee, Z. Liu, A. S. Long, G. Magdis, R. Massey, H. J. McCracken, J. McKinney, L. Paquereau, J. Rhodes, R. M. Rich, M. Shuntov, J. D. Silverman, M. Talia, S. Toft, and J. A. Zavala. COSMOS-Web: The Role of

Galaxy Interactions and Disk Instabilities in Producing Starbursts at z < 4. ApJ, 980(2):204, February 2025. doi: 10.3847/1538-4357/ada566.

J. Fensch, F. Renaud, F. Bournaud, P. A. Duc, O. Agertz, P. Amram, F. Combes, P. Di Matteo, B. Elmegreen, E. Emsellem, C. J. Jog, V. Perret, C. Struck, and R. Teyssier. High-redshift major mergers weakly enhance star formation. MNRAS, 465(2):1934–1949, February 2017. doi: 10.1093/mnras/stw2920.

Jérémy Fensch and Frédéric Bournaud. The role of gas fraction and feedback in the stability and evolution of galactic discs: implications for cosmological galaxy formation models. MNRAS, 505 (3):3579–3589, August 2021. doi: 10.1093/mnras/stab1489.

Leonardo Ferreira, Christopher J. Conselice, Elizaveta Sazonova, Fabricio Ferrari, Joseph Caruana, Clár-Bríd Tohill, Geferson Lucatelli, Nathan Adams, Dimitrios Irodotou, Madeline A. Marshall, Will J. Roper, Christopher C. Lovell, Aprajita Verma, Duncan Austin, James Trussler, and Stephen M. Wilkins. The JWST Hubble Sequence: The Rest-frame Optical Evolution of Galaxy Structure at 1.5 < z < 6.5. ApJ, 955(2):94, October 2023. doi: 10.3847/1538-4357/acec76.

David B. Fisher, Karl Glazebrook, Roberto G. Abraham, Ivana Damjanov, Heidi A. White, Danail Obreschkow, Robert Basset, Georgios Bekiaris, Emily Wisnioski, Andy Green, and Alberto D. Bolatto. Connecting Clump Sizes in Turbulent Disk Galaxies to Instability Theory. ApJ, 839(1): L5, April 2017. doi: 10.3847/2041-8213/aa6478.

Daniel Foreman-Mackey, David W. Hogg, Dustin Lang, and Jonathan Goodman. emcee: The MCMC Hammer. PASP, 125(925):306, March 2013. doi: 10.1086/670067.

N. M. Förster Schreiber, R. Genzel, N. Bouché, G. Cresci, R. Davies, P. Buschkamp, K. Shapiro, L. J. Tacconi, E. K. S. Hicks, S. Genel, A. E. Shapley, D. K. Erb, C. C. Steidel, D. Lutz, F. Eisenhauer, S. Gillessen, A. Sternberg, A. Renzini, A. Cimatti, E. Daddi, J. Kurk, S. Lilly, X. Kong, M. D. Lehnert, N. Nesvadba, A. Verma, H. McCracken, N. Arimoto, M. Mignoli, and M. Onodera. The SINS Survey: SINFONI Integral Field Spectroscopy of z ~2 Star-forming Galaxies. ApJ, 706(2): 1364–1428, December 2009. doi: 10.1088/0004-637X/706/2/1364.

N. M. Förster Schreiber, A. E. Shapley, D. K. Erb, R. Genzel, C. C. Steidel, N. Bouché, G. Cresci, and R. Davies. Constraints on the Assembly and Dynamics of Galaxies. I. Detailed Rest-frame Optical Morphologies on Kiloparsec Scale of z ~2 Star-forming Galaxies. ApJ, 731(1):65, April 2011a. doi: 10.1088/0004-637X/731/1/65.

N. M. Förster Schreiber, A. E. Shapley, R. Genzel, N. Bouché, G. Cresci, R. Davies, D. K. Erb, S. Genel, D. Lutz, S. Newman, and et al. Constraints on the Assembly and Dynamics of Galaxies. II. Properties of Kiloparsec-scale Clumps in Rest-frame Optical Emission of z ~2 Star-forming Galaxies. ApJ, 739(1):45, September 2011b. doi: 10.1088/0004-637X/739/1/45.

N. M. Förster Schreiber, A. E. Shapley, R. Genzel, N. Bouché, G. Cresci, R. Davies, D. K. Erb, S. Genel, D. Lutz, S. Newman, K. L. Shapiro, C. C. Steidel, A. Sternberg, and L. J. Tacconi. Constraints on the Assembly and Dynamics of Galaxies. II. Properties of Kiloparsec-scale Clumps in Rest-frame Optical Emission of z ~2 Star-forming Galaxies. ApJ, 739(1):45, September 2011c. doi: 10.1088/0004-637X/739/1/45.

N. M. Förster Schreiber, A. Renzini, C. Mancini, R. Genzel, N. Bouché, G. Cresci, E. K. S. Hicks, S. J. Lilly, Y. Peng, A. Burkert, C. M. Carollo, A. Cimatti, E. Daddi, R. I. Davies, S. Genel, J. D. Kurk, P. Lang, D. Lutz, V. Mainieri, H. J. McCracken, M. Mignoli, T. Naab, P. Oesch, L. Pozzetti, M. Scodeggio, K. Shapiro Griffin, A. E. Shapley, A. Sternberg, S. Tacchella, L. J. Tacconi, S. Wuyts, and G. Zamorani. The SINS/zC-SINF Survey of z $\sim$ 2 Galaxy Kinematics: SINFONI Adaptive Optics-assisted Data and Kiloparsec-scale Emission-line Properties. ApJS, 238(2):21, October 2018a. doi: 10.3847/1538-4365/aadd49.

N. M. Förster Schreiber, A. Renzini, C. Mancini, R. Genzel, N. Bouché, G. Cresci, E. K. S. Hicks, S. J. Lilly, Y. Peng, A. Burkert, C. M. Carollo, A. Cimatti, E. Daddi, R. I. Davies, S. Genel, J. D. Kurk, P. Lang, D. Lutz, V. Mainieri, H. J. McCracken, M. Mignoli, T. Naab, P. Oesch, L. Pozzetti, M. Scodeggio, K. Shapiro Griffin, A. E. Shapley, A. Sternberg, S. Tacchella, L. J. Tacconi, S. Wuyts, and G. Zamorani. The SINS/zC-SINF Survey of z $\sim$ 2 Galaxy Kinematics: SINFONI Adaptive Optics-assisted Data and Kiloparsec-scale Emission-line Properties. ApJS, 238(2):21, October 2018b. doi: 10.3847/1538-4365/aadd49.

S. Fujimoto, M. Ouchi, K. Kohno, F. Valentino, C. Giménez-Arteaga, G. B. Brammer, L. J. Furtak, M. Kohandel, M. Oguri, A. Pallottini, J. Richard, A. Zitrin, F. E. Bauer, M. Boylan-Kolchin, M. Dessauges-Zavadsky, E. Egami, S. L. Finkelstein, Z. Ma, I. Smail, D. Watson, T. A. Hutchison, J. R. Rigby, B. D. Welch, Y. Ao, L. D. Bradley, G. B. Caminha, K. I. Caputi, D. Espada, R. Endsley, Y. Fudamoto, J. González-López, B. Hatsukade, A. M. Koekemoer, V. Kokorev, N. Laporte, M. Lee, G. E. Magdis, Y. Ono, F. Rizzo, T. Shibuya, K. Shimasaku, F. Sun, S. Toft, H. Umehata, T. Wang, and H. Yajima. Primordial rotating disk composed of at least 15 dense star-forming clumps at cosmic dawn. *Nature Astronomy*, August 2025. doi: 10.1038/s41550-025-02592-w.

Gemini Observatory and AURA. Gemini IRAF: Data reduction software for the Gemini telescopes. Astrophysics Source Code Library, record ascl:1608.006, August 2016.

R. Genzel, S. Newman, T. Jones, N. M. Förster Schreiber, K. Shapiro, S. Genel, S. J. Lilly, A. Renzini, L. J. Tacconi, N. Bouché, A. Burkert, G. Cresci, P. Buschkamp, C. M. Carollo, D. Ceverino, R. Davies, A. Dekel, F. Eisenhauer, E. Hicks, J. Kurk, D. Lutz, C. Mancini, T. Naab, Y. Peng, A. Sternberg, D. Vergani, and G. Zamorani. The Sins Survey of z ~2 Galaxy Kinematics: Properties of the Giant Star-forming Clumps. ApJ, 733(2):101, June 2011. doi: 10.1088/0004-637X/733/2/101.

M. Giavalisco, H. C. Ferguson, A. M. Koekemoer, M. Dickinson, D. M. Alexander, F. E. Bauer, J. Bergeron, C. Biagetti, W. N. Brandt, S. Casertano, C. Cesarsky, E. Chatzichristou, C. Conselice, S. Cristiani, L. Da Costa, T. Dahlen, D. de Mello, P. Eisenhardt, T. Erben, S. M. Fall, C. Fassnacht, R. Fosbury, A. Fruchter, J. P. Gardner, N. Grogin, R. N. Hook, A. E. Hornschemeier, R. Idzi, S. Jogee, C. Kretchmer, V. Laidler, K. S. Lee, M. Livio, R. Lucas, P. Madau, B. Mobasher, L. A. Moustakas, M. Nonino, P. Padovani, C. Papovich, Y. Park, S. Ravindranath, A. Renzini, M. Richardson, A. Riess, P. Rosati, M. Schirmer, E. Schreier, R. S. Somerville, H. Spinrad, D. Stern, M. Stiavelli, L. Strolger, C. M. Urry, B. Vandame, R. Williams, and C. Wolf. The Great Observatories Origins Deep Survey: Initial Results from Optical and Near-Infrared Imaging. ApJ, 600(2):L93–L98, January 2004. doi: 10.1086/379232.

S. Gillman, A. L. Tiley, A. M. Swinbank, U. Dudzevičiūtė, R. M. Sharples, Ian Smail, C. M. Harrison, Andrew J. Bunker, Martin Bureau, M. Cirasuolo, Georgios E. Magdis, Trevor Mendel, and John P. Stott. The evolution of gas-phase metallicity and resolved abundances in star-forming galaxies at z ≈ 0.6-1.8. MNRAS, 500(3):4229–4247, January 2021. doi: 10.1093/mnras/staa3400.

Omri Ginzburg, Marc Huertas-Company, Avishai Dekel, Nir Mandelker, Gregory Snyder, Daniel Ceverino, and Joel Primack. The nature of giant clumps in high-z discs: a deep-learning comparison of simulations and observations. MNRAS, 501(1):730–746, February 2021. doi: 10.1093/mnras/staa3778.

Omri Ginzburg, Avishal Dekel, Nir Mandelker, and Mark R. Krumholz. The evolution of turbulent galactic discs: gravitational instability, feedback, and accretion. MNRAS, 513(4):6177–6195, July 2022. doi: 10.1093/mnras/stac1324.

M. Girard, M. Dessauges-Zavadsky, D. Schaerer, J. Richard, K. Nakajima, and A. Cava. Physics of a clumpy lensed galaxy at z = 1.6. A&A, 619:A15, November 2018. doi: 10.1051/0004-6361/201833533.

Rachel Gledhill, Victoria Strait, Guillaume Desprez, Gregor Rihtaršič, Maruša Bradač, Gabriel Brammer, Chris J. Willott, Nicholas Martis, Marcin Sawicki, Gaël Noirot, Ghassan T. E. Sarrouh, and Adam Muzzin. CANUCS: An Updated Mass and Magnification Model of A370 with JWST. ApJ, 973(2):77, October 2024. doi: 10.3847/1538-4357/ad684a.

Norman A. Grogin, Dale D. Kocevski, S. M. Faber, Henry C. Ferguson, Anton M. Koekemoer, Adam G. Riess, Viviana Acquaviva, David M. Alexander, Omar Almaini, Matthew L. N. Ashby, Marco Barden, Eric F. Bell, Frédéric Bournaud, Thomas M. Brown, Karina I. Caputi, Stefano Casertano, Paolo Cassata, Marco Castellano, Peter Challis, Ranga-Ram Chary, Edmond Cheung, Michele Cirasuolo, Christopher J. Conselice, Asantha Roshan Cooray, Darren J. Croton, Emanuele Daddi, Tomas Dahlen, Romeel Davé, Duília F. de Mello, Avishai Dekel, Mark Dickinson, Timothy Dolch, Jennifer L. Donley, James S. Dunlop, Aaron A. Dutton, David Elbaz, Giovanni G. Fazio, Alexei V. Filippenko, Steven L. Finkelstein, Adriano Fontana, Jonathan P. Gardner, Peter M. Garnavich, Eric Gawiser, Mauro Giavalisco, Andrea Grazian, Yicheng Guo, Nimish P. Hathi, Boris Häussler, Philip F. Hopkins, Jia-Sheng Huang, Kuang-Han Huang, Saurabh W. Jha, Jeyhan S. Kartaltepe, Robert P. Kirshner, David C. Koo, Kamson Lai, Kyoung-Soo Lee, Weidong Li, Jennifer M. Lotz, Ray A. Lucas, Piero Madau, Patrick J. McCarthy, Elizabeth J. McGrath, Daniel H. McIntosh, Ross J. McLure, Bahram Mobasher, Leonidas A. Moustakas, Mark Mozena, Kirpal Nandra, Jeffrey A. Newman, Sami-Matias Niemi, Kai G. Noeske, Casey J. Papovich, Laura Pentericci, Alexandra Pope, Joel R. Primack, Abhijith Rajan, Swara Ravindranath, Naveen A. Reddy, Alvio Renzini, Hans-Walter Rix, Aday R. Robaina, Steven A. Rodney, David J. Rosario, Piero Rosati, Sara Salimbeni, Claudia Scarlata, Brian Siana, Luc Simard, Joseph Smidt, Rachel S. Somerville, Hyron Spinrad, Amber N. Straughn, Louis-Gregory Strolger, Olivia Telford, Harry I. Teplitz, Jonathan R. Trump, Arjen van der Wel, Carolin Villforth, Risa H. Wechsler, Benjamin J. Weiner, Tommy Wiklind, Vivienne Wild, Grant Wilson, Stijn Wuyts, Hao-Jing Yan, and Min S. Yun. CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. ApJS, 197(2):35, December 2011. doi: 10.1088/0067-0049/197/2/35.

Asger E. Grønnow, Kristian Finlator, and Lise Christensen. Merging galaxies produce outliers from the fundamental metallicity relation. MNRAS, 451(4):4005–4017, August 2015. doi: 10.1093/mnras/stv1232.

Yicheng Guo, Mauro Giavalisco, Henry C. Ferguson, Paolo Cassata, and Anton M. Koekemoer. Multi-wavelength View of Kiloparsec-scale Clumps in Star-forming Galaxies at z ~2. ApJ, 757(2):120, October 2012. doi: 10.1088/0004-637X/757/2/120.

Yicheng Guo, Henry C. Ferguson, Eric F. Bell, David C. Koo, Christopher J. Conselice, Mauro Giavalisco, Susan Kassin, Yu Lu, Ray Lucas, Nir Mandelker, Daniel H. McIntosh, Joel R. Primack, Swara Ravindranath, Guillermo Barro, Daniel Ceverino, Avishai Dekel, Sandra M. Faber, Jerome J. Fang, Anton M. Koekemoer, Kai Noeske, Marc Rafelski, and Amber Straughn. Clumpy Galaxies in CANDELS. I. The Definition of UV Clumps and the Fraction of Clumpy Galaxies at $0.5 < z < 3$. ApJ, 800(1):39, February 2015. doi: 10.1088/0004-637X/800/1/39.

Yicheng Guo, David C. Koo, Yu Lu, John C. Forbes, Marc Rafelski, Jonathan R. Trump, Ricardo Amorín, Guillermo Barro, Romeel Davé, S. M. Faber, Nimish P. Hathi, Hassen Yesuf, Michael C. Cooper, Avishai Dekel, Puragra Guhathakurta, Evan N. Kirby, Anton M. Koekemoer, Pablo G. Pérez-González, Lihwai Lin, Jeffery A. Newman, Joel R. Primack, David J. Rosario, Christopher N. A. Willmer, and Renbin Yan. Stellar Mass-Gas-phase Metallicity Relation at $0.5 \leq z \leq 0.7$: A Power Law with Increasing Scatter toward the Low-mass Regime. ApJ, 822(2):103, May 2016. doi: 10.3847/0004-637X/822/2/103.

Yicheng Guo, Marc Rafelski, Eric F. Bell, Christopher J. Conselice, Avishai Dekel, S. M. Faber, Mauro Giavalisco, Anton M. Koekemoer, David C. Koo, Yu Lu, Nir Mandelker, Joel R. Primack, Daniel Ceverino, Duilia F. de Mello, Henry C. Ferguson, Nimish Hathi, Dale Kocevski, Ray A. Lucas, Pablo G. Pérez-González, Swara Ravindranath, Emmaris Soto, Amber Straughn, and Weichen Wang. Clumpy Galaxies in CANDELS. II. Physical Properties of UV-bright Clumps at $0.5 \leq z < 3$. ApJ, 853(2):108, February 2018. doi: 10.3847/1538-4357/aaa018.

Dávid Guszejnov, Philip F. Hopkins, and Michael Y. Grudić. Universal scaling relations in scale-free structure formation. MNRAS, 477(4):5139–5149, July 2018. doi: 10.1093/mnras/sty920.

Kevin N. Hainline, Benjamin D. Johnson, Brant Robertson, Sandro Tacchella, Jakob M. Helton, Fengwu Sun, Daniel J. Eisenstein, Charlotte Simmonds, Michael W. Topping, Lily Whitler, Christopher N. A. Willmer, Marcia Rieke, Katherine A. Suess, Raphael E. Hviding, Alex J. Cameron, Stacey Alberts, William M. Baker, Stefi Baum, Rachana Bhatawdekar, Nina Bonaventura, Kristan Boyett, Andrew J. Bunker, Stefano Carniani, Stephane Charlot, Jacopo Chevallard, Zuyi Chen, Mirko Curti, Emma Curtis-Lake, Francesco D'Eugenio, Eiichi Egami, Ryan Endsley, Ryan Hausen, Zhiyuan Ji, Tobias J. Looser, Jianwei Lyu, Roberto Maiolino, Erica Nelson, Dávid Puskás, Tim Rawle, Lester Sandles, Aayush Saxena, Renske Smit, Daniel P. Stark, Christina C. Williams, Chris Willott, and Joris Witstok. The Cosmos in Its Infancy: JADES Galaxy Candidates at $z > 8$ in GOODS-S and GOODS-N. ApJ, 964(1):71, March 2024. doi: 10.3847/1538-4357/ad1ee4.

Yunkun Han and Zhanwen Han. A Comprehensive Bayesian Discrimination of the Simple Stellar

Population Model, Star Formation History, and Dust Attenuation Law in the Spectral Energy Distribution Modeling of Galaxies. ApJS, 240(1):3, January 2019. doi: 10.3847/1538-4365/aaeffa.

Philip F. Hopkins. A general theory of turbulent fragmentation. MNRAS, 430(3):1653–1693, April 2013. doi: 10.1093/mnras/sts704.

Philip F. Hopkins, Eliot Quataert, and Norman Murray. Stellar feedback in galaxies and the origin of galaxy-scale winds. MNRAS, 421(4):3522–3537, April 2012. doi: 10.1111/j.1365-2966.2012.20593.x.

K. Horne. An optimal extraction algorithm for CCD spectroscopy. PASP, 98:609–617, June 1986. doi: 10.1086/131801.

Chi Huang, Hu Zou, Xu Kong, Johan Comparat, Zesen Lin, Yulong Gao, Zhixiong Liang, Timothee Delubac, Anand Raichoor, Jean-Paul Kneib, Donald P. Schneider, Xu Zhou, Qirong Yuan, and Matthew A. Bershady. The Mass-Metallicity Relation at z $\sim$ 0.8: Redshift Evolution and Parameter Dependency. ApJ, 886(1):31, November 2019. doi: 10.3847/1538-4357/ab4902.

Marc Huertas-Company, Yicheng Guo, Omri Ginzburg, Christoph T. Lee, Nir Mandelker, Maxwell Metter, Joel R. Primack, Avishai Dekel, Daniel Ceverino, Sandra M. Faber, David C. Koo, Anton Koekemoer, Gregory Snyder, Mauro Giavalisco, and Haowen Zhang. Stellar masses of giant clumps in CANDELS and simulated galaxies using machine learning. MNRAS, 499(1):814–835, November 2020. doi: 10.1093/mnras/staa2777.

Shigeki Inoue and Takayuki R. Saitoh. Natures of a clump-origin bulge: a pseudo-bulge like but old metal-rich bulge. MNRAS, 422(3):1902–1913, May 2012. doi: 10.1111/j.1365-2966.2011.20338.x.

Kartheik Iyer and Eric Gawiser. Reconstruction of Galaxy Star Formation Histories through SED Fitting:The Dense Basis Approach. ApJ, 838(2):127, April 2017. doi: 10.3847/1538-4357/aa63f0.

Kartheik G. Iyer, Eric Gawiser, Sandra M. Faber, Henry C. Ferguson, Jeyhan Kartaltepe, Anton M. Koekemoer, Camilla Pacifici, and Rachel S. Somerville. Nonparametric Star Formation History Reconstruction with Gaussian Processes. I. Counting Major Episodes of Star Formation. ApJ, 879(2):116, July 2019. doi: 10.3847/1538-4357/ab2052.

Kartheik G. Iyer, Camilla Pacifici, Gabriela Calistro-Rivera, and Christopher C. Lovell. The Spectral Energy Distributions of Galaxies. *arXiv e-prints*, art. arXiv:2502.17680, February 2025. doi: 10.48550/arXiv.2502.17680.

Stéphanie Juneau, Mark Dickinson, David M. Alexander, and Samir Salim. A New Diagnostic of Active Galactic Nuclei: Revealing Highly Absorbed Systems at Redshift >0.3. ApJ, 736(2):104, August 2011. doi: 10.1088/0004-637X/736/2/104.

Boris S. Kalita, John D. Silverman, Emanuele Daddi, Wilfried Mercier, Luis C. Ho, and Xuheng Ding. Near-IR clumps and their properties in high-z galaxies with JWST/NIRCam. MNRAS, 537 (1):402–418, February 2025a. doi: 10.1093/mnras/staf031.

Boris S. Kalita, Tomoko L. Suzuki, Daichi Kashino, John D. Silverman, Emanuele Daddi, Luis C. Ho, Xuheng Ding, Wilfried Mercier, Andreas L. Faisst, Kartik Sheth, Francesco Valentino, Annagrazia Puglisi, Toshiki Saito, Darshan Kakkad, Olivier Ilbert, Ali Ahmad Khostovan, Zhaoxuan

Liu, Takumi Tanaka, Georgios Magdis, Jorge A. Zavala, Qinghua Tan, Jeyhan S. Kartaltepe, Lilan Yang, Anton M. Koekemoer, Jed McKinney, Brant E. Robertson, Shuowen Jin, Christopher C. Hayward, Michaela Hirschmann, Maximilien Franco, Marko Shuntov, Ghassem Gozaliasl, Aidan Kaminsky, and R. Michael Rich. Clumps as multiscale structures in cosmic noon galaxies. MNRAS, 536(3):3090–3111, January 2025b. doi: 10.1093/mnras/stae2781.

Jeyhan S. Kartaltepe, Caitlin Rose, Brittany N. Vanderhoof, Elizabeth J. McGrath, Luca Costantin, Isabella G. Cox, L. Y. Aaron Yung, Dale D. Kocevski, Stijn Wuyts, Henry C. Ferguson, Micaela B. Bagley, Steven L. Finkelstein, Ricardo O. Amorín, Brett H. Andrews, Pablo Arrabal Haro, Bren E. Backhaus, Peter Behroozi, Laura Bisigello, Antonello Calabrò, Caitlin M. Casey, Rosemary T. Coogan, M. C. Cooper, Darren Croton, Alexander de la Vega, Mark Dickinson, Adriano Fontana, Maximilien Franco, Andrea Grazian, Norman A. Grogin, Nimish P. Hathi, Benne W. Holwerda, Marc Huertas-Company, Kartheik G. Iyer, Shardha Jogee, Intae Jung, Lisa J. Kewley, Allison Kirkpatrick, Anton M. Koekemoer, James Liu, Jennifer M. Lotz, Ray A. Lucas, Jeffrey A. Newman, Camilla Pacifici, Viraj Pandya, Casey Papovich, Laura Pentericci, Pablo G. Pérez-González, Jayse Petersen, Nor Pirzkal, Marc Rafelski, Swara Ravindranath, Raymond C. Simons, Gregory F. Snyder, Rachel S. Somerville, Elizabeth R. Stanway, Amber N. Straughn, Sandro Tacchella, Jonathan R. Trump, Jesús Vega-Ferrero, Stephen M. Wilkins, Guang Yang, and Jorge A. Zavala. CEERS Key Paper. III. The Diversity of Galaxy Structure and Morphology at z = 3-9 with JWST. ApJ, 946(1):L15, March 2023. doi: 10.3847/2041-8213/acad01.

Daichi Kashino, John D. Silverman, David Sanders, Jeyhan Kartaltepe, Emanuele Daddi, Alvio Renzini, Giulia Rodighiero, Annagrazia Puglisi, Francesco Valentino, Stéphanie Juneau, Nobuo Arimoto, Tohru Nagao, Olivier Ilbert, Olivier Le Fèvre, and Anton M. Koekemoer. The FMOS-COSMOS Survey of Star-forming Galaxies at z $\sim$ 1.6. VI. Redshift and Emission-line Catalog and Basic Properties of Star-forming Galaxies. ApJS, 241(1):10, March 2019. doi: 10.3847/1538-4365/ab06c4.

Jr. Kennicutt, Robert C. The Global Schmidt Law in Star-forming Galaxies. ApJ, 498(2):541–552, May 1998. doi: 10.1086/305588.

Dušan Kereš, Neal Katz, David H. Weinberg, and Romeel Davé. How do galaxies get their gas? MNRAS, 363(1):2–28, October 2005. doi: 10.1111/j.1365-2966.2005.09451.x.

Dušan Kereš, Neal Katz, Mark Fardal, Romeel Davé, and David H. Weinberg. Galaxies in a simulated $\Lambda$CDM Universe - I. Cold mode and hot cores. MNRAS, 395(1):160–179, May 2009. doi: 10.1111/j.1365-2966.2009.14541.x.

Lisa J. Kewley and Sara L. Ellison. Metallicity Calibrations and the Mass-Metallicity Relation for Star-forming Galaxies. ApJ, 681(2):1183–1204, July 2008. doi: 10.1086/587500.

Lisa J. Kewley, David C. Nicholls, and Ralph S. Sutherland. Understanding Galaxy Evolution Through Emission Lines. ARA&A, 57:511–570, August 2019. doi: 10.1146/annurev-astro-081817-051832.

Henry A. Kobulnicky and Lisa J. Kewley. Metallicities of $0.3<z<1.0$ Galaxies in the GOODS-North Field. ApJ, 617(1):240–261, December 2004. doi: 10.1086/425299.

Anton M. Koekemoer, S. M. Faber, Henry C. Ferguson, Norman A. Grogin, Dale D. Kocevski, David C. Koo, Kamson Lai, Jennifer M. Lotz, Ray A. Lucas, Elizabeth J. McGrath, Sara Ogaz, Abhijith Rajan, Adam G. Riess, Steve A. Rodney, Louis Strolger, Stefano Casertano, Marco Castellano, Tomas Dahlen, Mark Dickinson, Timothy Dolch, Adriano Fontana, Mauro Giavalisco, Andrea Grazian, Yicheng Guo, Nimish P. Hathi, Kuang-Han Huang, Arjen van der Wel, Hao-Jing Yan, Viviana Acquaviva, David M. Alexander, Omar Almaini, Matthew L. N. Ashby, Marco Barden, Eric F. Bell, Frédéric Bournaud, Thomas M. Brown, Karina I. Caputi, Paolo Cassata, Peter J. Challis, Ranga-Ram Chary, Edmond Cheung, Michele Cirasuolo, Christopher J. Conselice, Asantha Roshan Cooray, Darren J. Croton, Emanuele Daddi, Romeel Davé, Duilia F. de Mello, Loic de Ravel, Avishai Dekel, Jennifer L. Donley, James S. Dunlop, Aaron A. Dutton, David Elbaz, Giovanni G. Fazio, Alexei V. Filippenko, Steven L. Finkelstein, Chris Frazer, Jonathan P. Gardner, Peter M. Garnavich, Eric Gawiser, Ruth Gruetzbauch, Will G. Hartley, Boris Häussler, Jessica Herrington, Philip F. Hopkins, Jia-Sheng Huang, Saurabh W. Jha, Andrew Johnson, Jeyhan S. Kartaltepe, Ali A. Khostovan, Robert P. Kirshner, Caterina Lani, Kyoung-Soo Lee, Weidong Li, Piero Madau, Patrick J. McCarthy, Daniel H. McIntosh, Ross J. McLure, Conor McPartland, Bahram Mobasher, Heidi Moreira, Alice Mortlock, Leonidas A. Moustakas, Mark Mozena, Kirpal Nandra, Jeffrey A. Newman, Jennifer L. Nielsen, Sami Niemi, Kai G. Noeske, Casey J. Papovich, Laura Pentericci, Alexandra Pope, Joel R. Primack, Swara Ravindranath, Naveen A. Reddy, Alvio Renzini, Hans-Walter Rix, Aday R. Robaina, David J. Rosario, Piero Rosati, Sara Salimbeni, Claudia Scarlata, Brian Siana, Luc Simard, Joseph Smidt, Diana Snyder, Rachel S. Somerville, Hyron Spinrad, Amber N. Straughn, Olivia Telford, Harry I. Teplitz, Jonathan R. Trump, Carlos Vargas, Carolin Villforth, Cory R. Wagner, Pat Wandro, Risa H. Wechsler, Benjamin J. Weiner, Tommy Wiklind, Vivienne Wild, Grant Wilson, Stijn Wuyts, and Min S. Yun. CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey—The Hubble Space Telescope Observations, Imaging Data Products, and Mosaics. ApJS, 197(2):36, December 2011. doi: 10.1088/0067-0049/197/2/36.

Mariska Kriek, Alice E. Shapley, Naveen A. Reddy, Brian Siana, Alison L. Coil, Bahram Mobasher, William R. Freeman, Laura de Groot, Sedona H. Price, Ryan Sanders, Irene Shivaei, Gabriel B. Brammer, Ivelina G. Momcheva, Rosalind E. Skelton, Pieter G. van Dokkum, Katherine E. Whitaker, James Aird, Mojegan Azadi, Marc Kassis, James S. Bullock, Charlie Conroy, Romeel Davé, Dušan Kereš, and Mark Krumholz. The MOSFIRE Deep Evolution Field (MOSDEF) Survey: Rest-frame Optical Spectroscopy for ~1500 H-selected Galaxies at $1.37 < z < 3.8$. ApJS, 218 (2):15, June 2015. doi: 10.1088/0067-0049/218/2/15.

Pavel Kroupa. On the variation of the initial mass function. MNRAS, 322(2):231–246, April 2001. doi: 10.1046/j.1365-8711.2001.04022.x.

Mark R. Krumholz, Christopher F. McKee, and Joss Bland-Hawthorn. Star Clusters Across Cosmic Time. ARA&A, 57:227–303, August 2019. doi: 10.1146/annurev-astro-091918-104430.

Ivanna Langan, Johannes Zabl, Nicolas F. Bouché, Michele Ginolfi, Gergö Popping, Ilane Schroetter, Martin Wendt, Joop Schaye, Leindert Boogaard, Jonathan Freundlich, Johan Richard, Jorryt Matthee, Wilfried Mercier, Thierry Contini, Yucheng Guo, and Maxime Cherrey. MusE GAs FLOw and Wind (MEGAFLOW) IX. The impact of gas flows on the relations between the mass, star formation rate, and metallicity of galaxies. MNRAS, 521(1):546–557, May 2023. doi: 10.1093/mnras/stad357.

Oliver Le Fèvre, Michel Saisse, Dario Mancini, Sylvie Brau-Nogue, Oreste Caputi, Louis Castinel, Sandro D'Odorico, Bianca Garilli, Markus Kissler-Patig, Christian Lucuix, Guido Mancini, Alain Pauget, Giovanni Sciarretta, Marco Scodeggio, Laurence Tresse, and Gianpaolo Vettolani. Commissioning and performances of the VLT-VIMOS instrument. In Masanori Iye and Alan F. M. Moorwood, editors, *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes*, volume 4841 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 1670–1681, March 2003. doi: 10.1117/12.460959.

Joel Leja, Adam C. Carnall, Benjamin D. Johnson, Charlie Conroy, and Joshua S. Speagle. How to Measure Galaxy Star Formation Histories. II. Nonparametric Models. ApJ, 876(1):3, May 2019. doi: 10.3847/1538-4357/ab133c.

R. C. Livermore, T. A. Jones, J. Richard, R. G. Bower, A. M. Swinbank, T. T. Yuan, A. C. Edge, R. S. Ellis, L. J. Kewley, Ian Smail, K. E. K. Coppin, and H. Ebeling. Resolved spectroscopy of gravitationally lensed galaxies: global dynamics and star-forming clumps on $\sim$100 pc scales at $1 < z < 4$. MNRAS, 450(2):1812–1835, June 2015. doi: 10.1093/mnras/stv686.

J. M. Lotz, A. Koekemoer, D. Coe, N. Grogin, P. Capak, J. Mack, J. Anderson, R. Avila, E. A. Barker, D. Borncamp, G. Brammer, M. Durbin, H. Gunning, B. Hilbert, H. Jenkner, H. Khandrika, Z. Levay, R. A. Lucas, J. MacKenty, S. Ogaz, B. Porterfield, N. Reid, M. Robberto, P. Royle, L. J. Smith, L. J. Storrie-Lombardi, B. Sunnquist, J. Surace, D. C. Taylor, R. Williams, J. Bullock, M. Dickinson, S. Finkelstein, P. Natarajan, J. Richard, B. Robertson, J. Tumlinson, A. Zitrin, K. Flanagan, K. Sembach, B. T. Soifer, and M. Mountain. The Frontier Fields: Survey Design and Initial Results. ApJ, 837(1):97, March 2017. doi: 10.3847/1538-4357/837/1/97.

Jennifer M. Lotz, Joel Primack, and Piero Madau. A New Nonparametric Approach to Galaxy Morphological Classification. AJ, 128(1):163–182, July 2004. doi: 10.1086/421849.

Jennifer M. Lotz, Patrik Jonsson, T. J. Cox, and Joel R. Primack. Galaxy merger morphologies and time-scales from simulations of equal-mass gas-rich disc mergers. MNRAS, 391(3):1137–1162, December 2008. doi: 10.1111/j.1365-2966.2008.14004.x.

Jennifer M. Lotz, Patrik Jonsson, T. J. Cox, and Joel R. Primack. The effect of mass ratio on the morphology and time-scales of disc galaxy mergers. MNRAS, 404(2):575–589, May 2010. doi: 10.1111/j.1365-2966.2010.16268.x.

Piero Madau and Mark Dickinson. Cosmic Star-Formation History. ARA&A, 52:415–486, August 2014. doi: 10.1146/annurev-astro-081811-125615.

J. Maíz Apellániz. A Recalibration of Optical Photometry: Tycho-2, Strömgren, and Johnson Systems. AJ, 131(2):1184–1199, February 2006. doi: 10.1086/499158.

Nir Mandelker, Avishai Dekel, Daniel Ceverino, Dylan Tweed, Christopher E. Moody, and Joel Primack. The population of giant clumps in simulated high-z galaxies: in situ and ex situ migration and survival. MNRAS, 443(4):3675–3702, October 2014. doi: 10.1093/mnras/stu1340.

Nir Mandelker, Avishai Dekel, Daniel Ceverino, Colin DeGraf, Yicheng Guo, and Joel Primack. Giant clumps in simulated high- z Galaxies: properties, evolution and dependence on feedback. MNRAS, 464(1):635–665, January 2017. doi: 10.1093/mnras/stw2358.

F. Mannucci, G. Cresci, R. Maiolino, A. Marconi, and A. Gnerucci. A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies. MNRAS, 408(4): 2115–2127, November 2010. doi: 10.1111/j.1365-2966.2010.17291.x.

R. A. Marino, F. F. Rosales-Ortega, S. F. Sánchez, A. Gil de Paz, J. Vílchez, D. Miralles-Caballero, C. Kehrig, E. Pérez-Montero, V. Stanishev, J. Iglesias-Páramo, A. I. Díaz, A. Castillo-Morales, R. Kennicutt, A. R. López-Sánchez, L. Galbany, R. García-Benito, D. Mast, J. Mendez-Abreu, A. Monreal-Ibero, B. Husemann, C. J. Walcher, B. García-Lorenzo, J. Masegosa, A. Del Olmo Orozco, A. M. Mourão, B. Ziegler, M. Mollá, P. Papaderos, P. Sánchez-Blázquez, R. M. González Delgado, J. Falcón-Barroso, M. M. Roth, G. van de Ven, and CALIFA Team. The O3N2 and N2 abundance indicators revisited: improved calibrations based on CALIFA and $T_e$-based literature data. A&A, 559:A114, November 2013. doi: 10.1051/0004-6361/201321956.

Alec Martin, Yicheng Guo, Xin Wang, Anton M. Koekemoer, Marc Rafelski, Harry I. Teplitz, Rogier A. Windhorst, Anahita Alavi, Norman A. Grogin, Laura Prichard, Ben Sunnquist, Daniel Ceverino, Nima Chartab, Christopher J. Conselice, Y. Sophia Dai, Avishai Dekel, Jonathan P. Gardner, Eric Gawiser, Nimish P. Hathi, Matthew J. Hayes, Rolf A. Jansen, Zhiyuan Ji, David C. Koo, Ray A. Lucas, Nir Mandelker, Vihang Mehta, Bahram Mobasher, Kalina V. Nedkova, Joel Primack, Swara Ravindranath, Brant E. Robertson, Michael J. Rutkowski, Zahra Sattari, Emmaris Soto, and L. Y. Aaron Yung. UV-bright Star-forming Clumps and Their Host Galaxies in UVCANDELS at $0.5 \leq z \leq 1$. ApJ, 955(2):106, October 2023. doi: 10.3847/1538-4357/aced3e.

Lucio Mayer, Floor van Donkelaar, Matteo Messa, Pedro R. Capelo, and Angela Adamo. In Situ Formation of Star Clusters at $z > 7$ via Galactic Disk Fragmentation: Shedding Light on Ultracompact Clusters and Overmassive Black Holes Seen by JWST. ApJ, 981(2):L28, March 2025. doi: 10.3847/2041-8213/adadfe.

H. J. McCracken, B. Milvang-Jensen, J. Dunlop, M. Franx, J. P. U. Fynbo, O. Le Fèvre, J. Holt, K. I. Caputi, Y. Goranova, F. Buitrago, J. P. Emerson, W. Freudling, P. Hudelot, C. López-Sanjuan, F. Magnard, Y. Mellier, P. Møller, K. K. Nilsson, W. Sutherland, L. Tasca, and J. Zabl. UltraVISTA: a new ultra-deep near-infrared survey in COSMOS. A&A, 544:A156, August 2012. doi: 10.1051/0004-6361/201219507.

Ian S. McLean, Charles C. Steidel, Harland W. Epps, Nicholas Konidaris, Keith Y. Matthews, Sean Adkins, Theodore Aliado, George Brims, John M. Canfield, John L. Cromer, Jason Fucik, Kristin Kulas, Greg Mace, Ken Magnone, Hector Rodriguez, Gwen Rudie, Ryan Trainor, Eric Wang, Bob Weber, and Jason Weiss. MOSFIRE, the multi-object spectrometer for infra-red exploration at the Keck Observatory. In Ian S. McLean, Suzanne K. Ramsay, and Hideki Takami, editors, *Ground-based and Airborne Instrumentation for Astronomy IV*, volume 8446 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 84460J, September 2012. doi: 10.1117/12.924794.

P. Melchior, F. Moolekamp, M. Jerdee, R. Armstrong, A. L. Sun, J. Bosch, and R. Lupton. SCARLET: Source separation in multi-band images by Constrained Matrix Factorization. *Astronomy and Computing*, 24:129, July 2018. doi: 10.1016/j.ascom.2018.07.001.

Karín Menéndez-Delmestre, Andrew W. Blain, Mark Swinbank, Ian Smail, Rob J. Ivison, Scott C. Chapman, and Thiago S. Gonçalves. Mapping the Clumpy Structures within Submillimeter Galaxies Using Laser-Guide Star Adaptive Optics Spectroscopy. ApJ, 767(2):151, April 2013. doi: 10.1088/0004-637X/767/2/151.

Wilfried Mercier, Boris Sindhu Kalita, Marko Shuntov, Rafael C. Arango-Toro, Olivier Ilbert, Laurence Tresse, Yohan Dubois, Clotilde Laigle, Hossein Hatamnia, Nicolas McMahon, Andreas Faisst, Isa Cox, Maxime Trebitsch, Leo Michel-Dansac, Si-Yue Yu, Michaela Hirschmann, Marc Huertas-Company, Arianna Long, Anton Koekemoer, Grégoire Aufort, Joseph Lewis, Ghassem Gozaliasl, R. Michael Rich, Jason Rhodes, Henry Joy McCracken, Caitlin Casey, Jeyhan Kartaltepe, Brant Robertson, Maximilien Franco, Daizhong Liu, Hollis Akins, Natalie Allen, and Sune Toft. Clumpiness of galaxies revealed in the near-infrared with COSMOS-Web. *arXiv e-prints*, art. arXiv:2506.13881, June 2025. doi: 10.48550/arXiv.2506.13881.

Rosa M. Mérida, Marcin Sawicki, Kartheik G. Iyer, Gaël Noirot, Chris J. Willott, Maruša Bradač, Guillaume Desprez, Nicholas S. Martis, Adam Muzzin, Gregor Rihtaršič, Ghassan T. E. Sarrouh, Jeremy Favaro, Gaia Gaspar, Anishya Harshan, and Jon Judež. Probing the Star Formation Main Sequence down to $10^7 M_\odot$ at $1 < z < 9$. *arXiv e-prints*, art. arXiv:2509.22871, September 2025. doi: 10.48550/arXiv.2509.22871.

Matteo Messa, Miroslava Dessauges-Zavadsky, Johan Richard, Angela Adamo, David Nagy, Françoise Combes, Lucio Mayer, and Harald Ebeling. Multiply lensed star forming clumps in the A521-sys1 galaxy at redshift 1. MNRAS, 516(2):2420–2443, October 2022. doi: 10.1093/mnras/stac2189.

Matteo Messa, Miroslava Dessauges-Zavadsky, Angela Adamo, Johan Richard, and Adélaïde Claeyssens. Properties of the brightest young stellar clumps in extremely lensed galaxies at redshifts 4 to 5. MNRAS, 529(3):2162–2179, April 2024. doi: 10.1093/mnras/stae565.

U. Meštrić, E. Vanzella, A. Zanella, M. Castellano, F. Calura, P. Rosati, P. Bergamini, A. Mercurio, M. Meneghetti, C. Grillo, G. B. Caminha, M. Nonino, E. Merlin, G. Cupani, and E. Sani. Exploring the physical properties of lensed star-forming clumps at $2 \lesssim z \lesssim 6$. MNRAS, 516(3): 3532–3555, November 2022. doi: 10.1093/mnras/stac2309.

Leo Michel-Dansac, Diego G. Lambas, M. Sol Alonso, and Patricia Tissera. The mass-metallicity relation of interacting galaxies. MNRAS, 386(1):L82–L86, May 2008. doi: 10.1111/j.1745-3933.2008.00466.x.

M. Montuori, P. Di Matteo, M. D. Lehnert, F. Combes, and B. Semelin. The dilution peak, metallicity evolution, and dating of galaxy interactions and mergers. A&A, 518:A56, July 2010. doi: 10.1051/0004-6361/201014304.

Christopher E. Moody, Yicheng Guo, Nir Mandelker, Daniel Ceverino, Mark Mozena, David C. Koo, Avishai Dekel, and Joel Primack. Star formation and clumps in cosmological galaxy simulations with radiation pressure feedback. MNRAS, 444(2):1389–1399, October 2014. doi: 10.1093/mnras/stu1534.

Takahiro Morishita, Massimo Stiavelli, Eros Vanzella, Pietro Bergamini, Kristan Boyett, Marco Chiaberge, Claudio Grillo, Nicha Leethochawalit, Matteo Messa, Guido Roberts-Borsani, Piero Rosati, and Anowar J. Shajib. Metallicity Scatter Originating from Subkiloparsec Starbursting Clumps in the Core of a Protocluster at z = 7.88. ApJ, 985(1):83, May 2025. doi: 10.3847/1538-4357/adc4c3.

Lamiya Mowla, Kartheik Iyer, Yoshihisa Asada, Guillaume Desprez, Vivian Yun Yan Tan, Nicholas Martis, Ghassan Sarrouh, Victoria Strait, Roberto Abraham, Maruša Bradač, Gabriel Brammer, Adam Muzzin, Camilla Pacifici, Swara Ravindranath, Marcin Sawicki, Chris Willott, Vince Estrada-Carpenter, Nusrath Jahan, Gaël Noirot, Jasleen Matharu, Gregor Rihtaršič, and Johannes Zabl. The Firefly Sparkle: The Earliest Stages of the Assembly of A Milky Way-type Galaxy in a 600 Myr Old Universe. *arXiv e-prints*, art. arXiv:2402.08696, February 2024. doi: 10.48550/arXiv.2402.08696.

K. L. Murata, M. Kajisawa, Y. Taniguchi, M. A. R. Kobayashi, Y. Shioya, P. Capak, O. Ilbert, A. M. Koekemoer, M. Salvato, and N. Z. Scoville. Evolution of the Fraction of Clumpy Galaxies at 0.2 $< z <$ 1.0 in the COSMOS Field. ApJ, 786(1):15, May 2014. doi: 10.1088/0004-637X/786/1/15.

Adam Muzzin, Danilo Marchesini, Mauro Stefanon, Marijn Franx, Bo Milvang-Jensen, James S. Dunlop, J. P. U. Fynbo, Gabriel Brammer, Ivo Labbé, and Pieter van Dokkum. A Public $K_s$ -selected Catalog in the COSMOS/ULTRAVISTA Field: Photometry, Photometric Redshifts, and Stellar Population Parameters. ApJS, 206(1):8, May 2013. doi: 10.1088/0067-0049/206/1/8.

Yurina Nakazato, Daniel Ceverino, and Naoki Yoshida. A merger-driven scenario for clumpy galaxy formation in the epoch of reionization: Physical properties of clumps in the FirstLight simulation. *arXiv e-prints*, art. arXiv:2402.08911, February 2024. doi: 10.48550/arXiv.2402.08911.

Dylan Nelson, Mark Vogelsberger, Shy Genel, Debora Sijacki, Dušan Kereš, Volker Springel, and Lars Hernquist. Moving mesh cosmology: tracing cosmological gas accretion. MNRAS, 429(4):3353–3370, March 2013. doi: 10.1093/mnras/sts595.

Matthew Newville, Till Stensitzki, Daniel B. Allen, Michal Rawlik, Antonino Ingargiola, and Andrew Nelson. Lmfit: Non-Linear Least-Square Minimization and Curve-Fitting for Python. Astrophysics Source Code Library, record ascl:1606.014, June 2016.

Masafumi Noguchi. Early Evolution of Disk Galaxies: Formation of Bulges in Clumpy Young Galactic Disks. ApJ, 514(1):77–95, March 1999. doi: 10.1086/306932.

Gaël Noirot, Marcin Sawicki, Roberto Abraham, Maruša Bradač, Kartheik Iyer, Thibaud Moutard, Camilla Pacifici, Swara Ravindranath, and Chris J. Willott. Across the green valley with HST grisms: colour evolution, crossing time-scales, and the growth of the red sequence at z = 1.0-1.8. MNRAS, 512(3):3566–3588, May 2022. doi: 10.1093/mnras/stac668.

J. B. Oke and J. E. Gunn. Secondary standard stars for absolute spectrophotometry. ApJ, 266:713–717, March 1983. doi: 10.1086/160817.

Antonija Oklopčić, Philip F. Hopkins, Robert Feldmann, Dušan Kereš, Claude-André Faucher-Giguère, and Norman Murray. Giant clumps in the FIRE simulations: a case study of a massive high-redshift galaxy. MNRAS, 465(1):952–969, February 2017. doi: 10.1093/mnras/stw2754.

Kiyoaki Christopher Omori and Tsutomu T. Takeuchi. The bimodality in the mass-metallicity relation in SDSS-MaNGA galaxy pairs. A&A, 660:A145, April 2022. doi: 10.1051/0004-6361/202142858.

Matthew E. Orr and Douglas Rennehan. How the Cookie Crumbles: A Model for Star-forming Clumps in High-redshift Disk Galaxies. *arXiv e-prints*, art. arXiv:2410.23337, October 2024. doi: 10.48550/arXiv.2410.23337.

B. E. J. Pagel, M. G. Edmunds, D. E. Blackwell, M. S. Chun, and G. Smith. On the composition of H II regions in southern galaxies - I. NGC 300 and 1365. MNRAS, 189:95–113, October 1979. doi: 10.1093/mnras/189.1.95.

Chien Y. Peng, Luis C. Ho, Chris D. Impey, and Hans-Walter Rix. Detailed Structural Decomposition of Galaxy Images. AJ, 124(1):266–293, July 2002. doi: 10.1086/340952.

Max Pettini and Bernard E. J. Pagel. [OIII]/[NII] as an abundance indicator at high redshift. MNRAS, 348(3):L59–L63, March 2004. doi: 10.1111/j.1365-2966.2004.07591.x.

A. J. Pickles. A Stellar Spectral Flux Library: 1150-25000 Å. PASP, 110(749):863–878, July 1998. doi: 10.1086/316197.

Marc Postman, Dan Coe, Narciso Benítez, Larry Bradley, Tom Broadhurst, Megan Donahue, Holland Ford, Or Graur, Genevieve Graves, Stephanie Jouvel, Anton Koekemoer, Doron Lemze, Elinor Medezinski, Alberto Molino, Leonidas Moustakas, Sara Ogaz, Adam Riess, Steve Rodney, Piero Rosati, Keiichi Umetsu, Wei Zheng, Adi Zitrin, Matthias Bartelmann, Rychard Bouwens, Nicole Czakon, Sunil Golwala, Ole Host, Leopoldo Infante, Saurabh Jha, Yolanda Jimenez-Teja, Daniel Kelson, Ofer Lahav, Ruth Lazkoz, Dani Maoz, Curtis McCully, Peter Melchior, Massimo Meneghetti, Julian Merten, John Moustakas, Mario Nonino, Brandon Patel, Enikö Regös, Jack Sayers, Stella Seitz, and Arjen Van der Wel. The Cluster Lensing and Supernova Survey with Hubble: An Overview. ApJS, 199(2):25, April 2012. doi: 10.1088/0067-0049/199/2/25.

M. Puech. Clumpy galaxies at z ~0.6: kinematics, stability and comparison with analogues at other redshifts. MNRAS, 406(1):535–547, July 2010. doi: 10.1111/j.1365-2966.2010.16689.x.

Dávid Puskás, Sandro Tacchella, Charlotte Simmonds, Kevin Hainline, Francesco D'Eugenio, Stacey Alberts, Santiago Arribas, William M. Baker, Andrew J. Bunker, Stefano Carniani, Stéphane Charlot, Qiao Duan, Daniel J. Eisenstein, Zhiyuan Ji, Benjamin D. Johnson, Gareth C. Jones, Roberto Maiolino, William McClymont, Marcia Rieke, Pierluigi Rinaldi, Brant Robertson, Hannah Übler, Christina C. Williams, Christopher N. A. Willmer, Chris Willott, and Joris Witstok. Constraining the major merger history of z ~3-9 galaxies using JADES: dominant in situ star formation. MNRAS, 540(3):2146–2175, July 2025. doi: 10.1093/mnras/staf813.

F. Renaud, F. Bournaud, K. Kraljic, and P. A. Duc. Starbursts triggered by intergalactic tides andinterstellar compressive turbulence. MNRAS, 442:L33–L37, July 2014. doi: 10.1093/mnrasl/slu050.

Florent Renaud, Frédéric Bournaud, and Pierre-Alain Duc. A parsec-resolution simulation of the Antennae galaxies: formation of star clusters during the merger. MNRAS, 446(2):2038–2054, January 2015. doi: 10.1093/mnras/stu2208.

Florent Renaud, Oscar Agertz, and Alessandro B. Romeo. From giant clumps to clouds. IV. Extreme star-forming clumps on top of universal cloud scaling relations in gas-rich galaxies. A&A, 687: A91, July 2024. doi: 10.1051/0004-6361/202449721.

B. Ribeiro, O. Le Fèvre, P. Cassata, B. Garilli, B. C. Lemaux, D. Maccagni, D. Schaerer, L. A. M. Tasca, G. Zamorani, E. Zucca, R. Amorín, S. Bardelli, N. P. Hathi, A. Koekemoer, and J. Pforr. The VIMOS Ultra-Deep Survey: A major merger origin for the high fraction of galaxies at $2 < z < 6$ with two bright clumps. A&A, 608:A16, November 2017. doi: 10.1051/0004-6361/201630057.

G. Rihtaršič, M. Bradač, G. Desprez, A. Harshan, G. Noirot, V. Estrada-Carpenter, N. S. Martis, R. G. Abraham, Y. Asada, G. Brammer, K. G. Iyer, J. Matharu, L. Mowla, A. Muzzin, G. T. E. Sarrouh, M. Sawicki, V. Strait, C. J. Willott, R. Gledhill, V. Markov, and R. Tripodi. CANUCS: Constraining the MACS J0416.1-2403 strong lensing model with JWST NIRISS, NIRSpec, and NIRCam. A&A, 696:A15, April 2025. doi: 10.1051/0004-6361/202451117.

Vicente Rodriguez-Gomez, Gregory F. Snyder, Jennifer M. Lotz, Dylan Nelson, Annalisa Pillepich, Volker Springel, Shy Genel, Rainer Weinberger, Sandro Tacchella, Rüdiger Pakmor, Paul Torrey, Federico Marinacci, Mark Vogelsberger, Lars Hernquist, and David A. Thilker. The optical morphologies of galaxies in the IllustrisTNG simulation: a comparison to Pan-STARRS observations. MNRAS, 483(3):4140–4159, March 2019. doi: 10.1093/mnras/sty3345.

M. Romano, P. Cassata, L. Morselli, G. C. Jones, M. Ginolfi, A. Zanella, M. Béthermin, P. Capak, A. Faisst, O. Le Fèvre, D. Schaerer, J. D. Silverman, L. Yan, S. Bardelli, M. Boquien, A. Cimatti, M. Dessauges-Zavadsky, A. Enia, S. Fujimoto, C. Gruppioni, N. P. Hathi, E. Ibar, A. M. Koekemoer, B. C. Lemaux, G. Rodighiero, D. Vergani, G. Zamorani, and E. Zucca. The ALPINE-ALMA [CII] survey. The contribution of major mergers to the galaxy mass assembly at $z \sim 5$. A&A, 653: A111, September 2021. doi: 10.1051/0004-6361/202141306.

Alessandro B. Romeo and Joachim Wiegert. The effective stability parameter for two-component galactic discs: is $Q^{-1} \approx Q^{-1}{}_{stars} + Q^{-1}{}_{gas}$? MNRAS, 416(2):1191–1196, September 2011. doi: 10.1111/j.1365-2966.2011.19120.x.

Alessandro B. Romeo, Andreas Burkert, and Oscar Agertz. A Toomre-like stability criterion for the clumpy and turbulent interstellar medium. MNRAS, 407(2):1223–1230, September 2010. doi: 10.1111/j.1365-2966.2010.16975.x.

Barbara S. Ryden. The Intrinsic Shape of Spiral Galaxies in the 2MASS Large Galaxy Atlas. ApJ, 641(2):773–784, April 2006. doi: 10.1086/500497.

Brett Salmon, Dan Coe, Larry Bradley, Rychard Bouwens, Marusa Bradač, Kuang-Han Huang, Pascal A. Oesch, Daniel Stark, Keren Sharon, Michele Trenti, Roberto J. Avila, Sara Ogaz, Felipe Andrade-Santos, Daniela Carrasco, Catherine Cerny, William Dawson, Brenda L. Frye, Austin Hoag, Traci Lin Johnson, Christine Jones, Daniel Lam, Lorenzo Lovisari, Ramesh Mainali, Matt Past, Rachel Paterno-Mahler, Avery Peterson, Adam G. Riess, Steven A. Rodney, Russel E. Ryan, Irene Sendra-Server, Victoria Strait, Louis-Gregory Strolger, Keiichi Umetsu, Benedetta Vulcani, and Adi Zitrin. RELICS: The Reionization Lensing Cluster Survey and the Brightest High-z Galaxies. ApJ, 889(2):189, February 2020. doi: 10.3847/1538-4357/ab5a8b.

Edwin E. Salpeter. The Luminosity Function and Stellar Evolution. ApJ, 121:161, January 1955. doi: 10.1086/145971.

J. Sánchez Almeida, A. B. Morales-Luis, C. Muñoz-Tuñón, D. M. Elmegreen, B. G. Elmegreen, and J. Méndez-Abreu. Metallicity Inhomogeneities in Local Star-forming Galaxies as a Sign of Recent Metal-poor Gas Accretion. ApJ, 783(1):45, March 2014. doi: 10.1088/0004-637X/783/1/45.

J. Sánchez Almeida, B. G. Elmegreen, C. Muñoz-Tuñón, D. M. Elmegreen, E. Pérez-Montero, R. Amorín, M. E. Filho, Y. Ascasibar, P. Papaderos, and J. M. Vílchez. Localized Starbursts in Dwarf Galaxies Produced by the Impact of Low-metallicity Cosmic Gas Clouds. ApJ, 810(2): L15, September 2015. doi: 10.1088/2041-8205/810/2/L15.

Ryan L. Sanders, Alice E. Shapley, Mariska Kriek, William R. Freeman, Naveen A. Reddy, Brian Siana, Alison L. Coil, Bahram Mobasher, Romeel Davé, Irene Shivaei, Mojegan Azadi, Sedona H. Price, Gene Leung, Tara Fetherolf, Laura de Groot, Tom Zick, Francesca M. Fornasini, and Guillermo Barro. The MOSDEF Survey: A Stellar Mass-SFR-Metallicity Relation Exists at z $\sim$ 2.3. ApJ, 858(2):99, May 2018. doi: 10.3847/1538-4357/aabcbd.

Ryan L. Sanders, Alice E. Shapley, Tucker Jones, Naveen A. Reddy, Mariska Kriek, Brian Siana, Alison L. Coil, Bahram Mobasher, Irene Shivaei, Romeel Davé, Mojegan Azadi, Sedona H. Price, Gene Leung, William R. Freeman, Tara Fetherolf, Laura de Groot, Tom Zick, and Guillermo Barro. The MOSDEF Survey: The Evolution of the Mass-Metallicity Relation from z = 0 to z 3.3. ApJ, 914(1):19, June 2021. doi: 10.3847/1538-4357/abf4c1.

Ghassan T. E. Sarrouh, Adam Muzzin, Kartheik G. Iyer, Lamiya Mowla, Sunna Withers, Nicholas S. Martis, Roberto Abraham, Yoshihisa Asada, Maruša Bradač, Gabriel B. Brammer, Guillaume Desprez, Vince Estrada-Carpenter, Jasleen Matharu, Gaël Noirot, Marcin Sawicki, Victoria Strait, Chris J. Willott, and Johannes Zabl. Exposing Line Emission: The Systematic Differences of Measuring Galaxy Stellar Masses with JWST NIRCam Medium versus Wide Band Photometry. ApJ, 967(1):L17, May 2024. doi: 10.3847/2041-8213/ad43e8.

Ghassan T. E. Sarrouh, Yoshihisa Asada, Nicholas S. Martis, Chris J. Willott, Kartheik G. Iyer, Gaël Noirot, Adam Muzzin, Marcin Sawicki, Gabriel Brammer, Guillaume Desprez, Gregor Rihtaršič, Johannes Zabl, Roberto Abraham, Maruša Bradač, René Doyon, Jacqueline Antwi-Danso, Samantha Berek, Westley Brown, Vince Estrada-Carpenter, Jeremy Favaro, Giordano Felicioni, Ben Forrest, Gaia Gaspar, Katriona M. L. Gould, Rachel Gledhill, Anishya Harshan, Nusrath Jahan, Naadiyah Jagga, Jon Judež, Danilo Marchesini, Vladan Markov, Jasleen Matharu, Shannon MacFarland, Maya Merchant, Rosa M. Mérida, Lamiya Mowla, Katherine Myers, Kiyoaki C. Omori, Camilla Pacifici, Swara Ravindranath, Luke Robbins, Victoria Strait, Visal Sok, Vivian Yun Yan Tan, Roberta Tripodi, Gillian Wilson, and Sunna Withers. CANUCS/Technicolor Data Release 1: Imaging, Photometry, Slit Spectroscopy, and Stellar Population Parameters. *arXiv e-prints*, art. arXiv:2506.21685, June 2025.

Zahra Sattari, Bahram Mobasher, Nima Chartab, Daniel D. Kelson, Harry I. Teplitz, Marc Rafelski, Norman A. Grogin, Anton M. Koekemoer, Xin Wang, Rogier A. Windhorst, Anahita Alavi, Laura Prichard, Ben Sunnquist, Jonathan P. Gardner, Eric Gawiser, Nimish P. Hathi, Matthew J. Hayes, Zhiyuan Ji, Vihang Mehta, Brant E. Robertson, Claudia Scarlata, L. Y. Aaron Yung, Christopher J. Conselice, Y. Sophia Dai, Yicheng Guo, Ray A. Lucas, Alec Martin, and Swara Ravindranath. Fraction of Clumpy Star-forming Galaxies at $0.5 \leq z \leq 3$ in UVCANDELS: Dependence on Stellar Mass and Environment. ApJ, 951(2):147, July 2023. doi: 10.3847/1538-4357/acd5d6.

C. Schreiber, K. Glazebrook, T. Nanayakkara, G. G. Kacprzak, I. Labbé, P. Oesch, T. Yuan, K. V. Tran, C. Papovich, L. Spitler, and C. Straatman. Near infrared spectroscopy and star-formation histories of $3 \leq z \leq 4$ quiescent galaxies. A&A, 618:A85, October 2018. doi: 10.1051/0004-6361/201833070.

N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J. P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, E. Schinnerer, D. Schminovich, P. Shopbell, Y. Taniguchi, and N. D. Tyson. The Cosmic Evolution Survey (COSMOS): Overview. ApJS, 172(1):1–8, September 2007. doi: 10.1086/516585.

Takatoshi Shibuya, Masami Ouchi, Mariko Kubo, and Yuichi Harikane. Morphologies of ~190,000 Galaxies at z = 0-10 Revealed with HST Legacy Data. II. Evolution of Clumpy Galaxies. ApJ, 821(2):72, April 2016. doi: 10.3847/0004-637X/821/2/72.

Takatoshi Shibuya, Noriaki Miura, Kenji Iwadate, Seiji Fujimoto, Yuichi Harikane, Yoshiki Toba, Takuya Umayahara, and Yohito Ito. Galaxy morphologies revealed with Subaru HSC and super-resolution techniques. I. Major merger fractions of $L_{UV}$ 3-15 $L^*_{UV}$ dropout galaxies at $z \sim 4$-7. PASJ, 74(1):73–91, February 2022. doi: 10.1093/pasj/psab111.

J. D. Silverman, D. Kashino, D. Sanders, J. S. Kartaltepe, N. Arimoto, A. Renzini, G. Rodighiero, E. Daddi, J. Zahid, T. Nagao, L. J. Kewley, S. J. Lilly, N. Sugiyama, I. Baronchelli, P. Capak, C. M. Carollo, J. Chu, G. Hasinger, O. Ilbert, S. Juneau, M. Kajisawa, A. M. Koekemoer, K. Kovac, O. Le Fèvre, D. Masters, H. J. McCracken, M. Onodera, A. Schulze, N. Scoville, V. Strazzullo, and Y. Taniguchi. The FMOS-COSMOS Survey of Star-forming Galaxies at z~1.6. III. Survey Design, Performance, and Sample Characteristics. ApJS, 220(1):12, September 2015. doi: 10.1088/0067-0049/220/1/12.

Suresh Sivanandam, Scott Chapman, Adam Muzzin, Scott Christie, Paul Hickson, Alan McConnachie, Simon Thibault, Jenny Atwood, André Anthony, Marcin Sawicki, Jennifer Dunn, Jae-Joon Lee, Kim Venn, Ruben Diaz, Stephen Goodsell, Shawn Barbod, Martin Tschimmel, Mark Barnet, Gavin Hay, Scott Roberts, Gaetano Sivo, Masen Lamb, Shaojie Chen, Saugata Dutt, Patrick Nkwari, Olivier Lardière, Uriel Conod, Sarik Jeram, and Julia Scharwächter. The final design of the Gemini Infrared Multi-Object Spectrograph: Gemini's new adaptive optics facility instrument. In Julia J. Bryant, Kentaro Motohara, and Joël. R. D. Vernet, editors, *Ground-based and Airborne Instrumentation for Astronomy X*, volume 13096 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 130960K, July 2024. doi: 10.1117/12.3020305.

Rosalind E. Skelton, Katherine E. Whitaker, Ivelina G. Momcheva, Gabriel B. Brammer, Pieter G. van Dokkum, Ivo Labbé, Marijn Franx, Arjen van der Wel, Rachel Bezanson, Elisabete Da Cunha, Mattia Fumagalli, Natascha Förster Schreiber, Mariska Kriek, Joel Leja, Britt F. Lundgren, Daniel Magee, Danilo Marchesini, Michael V. Maseda, Erica J. Nelson, Pascal Oesch, Camilla Pacifici, Shannon G. Patel, Sedona Price, Hans-Walter Rix, Tomer Tal, David A. Wake, and Stijn Wuyts. 3D-HST WFC3-selected Photometric Catalogs in the Five CANDELS/3D-HST Fields: Photometry, Photometric Redshifts, and Stellar Masses. ApJS, 214(2):24, October 2014. doi: 10.1088/0067-0049/214/2/24.

Visal Sok, Adam Muzzin, Pascale Jablonka, Z. Cemile Marsan, Vivian Y. Y. Tan, Leo Alcorn, Danilo Marchesini, and Mauro Stefanon. Finite-resolution Deconvolution of Multiwavelength Imaging of 20,000 Galaxies in the COSMOS Field: The Evolution of Clumpy Galaxies over Cosmic Time. ApJ, 924(1):7, January 2022. doi: 10.3847/1538-4357/ac2f40.

Visal Sok, Adam Muzzin, Pascale Jablonka, Vivian Yun Yan Tan, Z. Cemile Marsan, Danilo Marchesini, Gillian Wilson, and Leo Y. Alcorn. An Indication of Gas Inflow in Clumpy Star-forming Galaxies near z $\sim$ 1: Lower Gas-phase Metallicities in Clumpy Galaxies Compared to Nonclumpy Galaxies. ApJ, 979(1):14, January 2025a. doi: 10.3847/1538-4357/ad96bb.

Visal Sok, Adam Muzzin, Vivian Yun Yan Tan, Yoshihisa Asada, Maruša Bradač, Nicholas S. Martis, Gaël Noirot, Ghassan T. E. Sarrouh, Marcin Sawicki, Chris J. Willott, Sunna Withers, Samantha C. Berek, and Katherine Myers. The Stellar Mass and Age Distributions of Star-Forming Clumps at $0.5 < z < 5$ in JWST CANUCS: Implications for Clump Formation and Destruction. *arXiv e-prints*, art. arXiv:2509.25363, September 2025b.

R. Sorba and M. Sawicki. Missing stellar mass in SED fitting: spatially unresolved photometry can underestimate galaxy masses. MNRAS, 452(1):235–245, September 2015. doi: 10.1093/mnras/stv1235.

Robert Sorba and Marcin Sawicki. Spatially unresolved SED fitting can underestimate galaxy masses: a solution to the missing mass problem. MNRAS, 476(2):1532–1547, May 2018. doi: 10.1093/mnras/sty186.

Emmaris Soto, Duilia F. de Mello, Marc Rafelski, Jonathan P. Gardner, Harry I. Teplitz, Anton M. Koekemoer, Swara Ravindranath, Norman A. Grogin, Claudia Scarlata, Peter Kurczynski, and Eric Gawiser. Physical Properties of Sub-galactic Clumps at $0.5 \leq Z \leq 1.5$ in the UVUDF. ApJ, 837(1):6, March 2017. doi: 10.3847/1538-4357/aa5da3.

Jean-Luc Starck, Jalal Fadili, and Fionn Murtagh. The Undecimated Wavelet Decomposition and its Reconstruction. *IEEE Transactions on Image Processing*, 16(2):297–309, February 2007. doi: 10.1109/TIP.2006.887733.

L. J. Tacconi, R. Genzel, R. Neri, P. Cox, M. C. Cooper, K. Shapiro, A. Bolatto, N. Bouché, F. Bournaud, A. Burkert, F. Combes, J. Comerford, M. Davis, N. M. Förster Schreiber, S. Garcia-Burillo, J. Gracia-Carpio, D. Lutz, T. Naab, A. Omont, A. Shapley, A. Sternberg, and B. Weiner. High molecular gas fractions in normal massive star-forming galaxies in the young Universe. Nature, 463(7282):781–784, February 2010. doi: 10.1038/nature08773.

Linda J. Tacconi, Reinhard Genzel, and Amiel Sternberg. The Evolution of the Star-Forming Interstellar Medium Across Cosmic Time. ARA&A, 58:157–203, August 2020. doi: 10.1146/annurev-astro-082812-141034.

Valentina Tamburello, Alireza Rahmati, Lucio Mayer, Antonio Cava, Miroslava Dessauges-Zavadsky, and Daniel Schaerer. Clumpy galaxies seen in H $\alpha$: inflated observed clump properties due to limited spatial resolution and sensitivity. MNRAS, 468(4):4792–4800, July 2017. doi: 10.1093/mnras/stx784.

Vivian Yun Yan Tan, Adam Muzzin, Z. Cemile Marsan, Visal Sok, Leo Y. Alcorn, Jasleen Matharu, Heath Shipley, Danilo Marchesini, Kalina V. Nedkova, Nicholas Martis, Arjen van der Wel, and Katherine E. Whitaker. Resolved Stellar Mass Maps of Galaxies in the Hubble Frontier Fields: Evidence for Mass Dependency in Environmental Quenching. ApJ, 933(1):30, July 2022. doi: 10.3847/1538-4357/ac7051.

Vivian Yun Yan Tan, Adam Muzzin, Ghassan T. E. Sarrouh, Jacqueline Antwi-Danso, Visal Sok, Naadiyah Jagga, Roberto Abraham, Yoshihisa Asada, Guillaume Desprez, Kartheik Iyer, Nicholas S. Martis, Rosa M. Mérida, Lamiya A. Mowla, Gaël Noirot, Kiyoaki Christopher Omori, Marcin Sawicki, Roberta Tripodi, and Chris J. Willott. Resolved mass assembly and star formation in Milky Way Progenitors since $z = 5$ from JWST/CANUCS: From clumps and mergers to well-ordered disks. *arXiv e-prints*, art. arXiv:2412.07829, December 2024. doi: 10.48550/arXiv.2412.07829.

Hossen Teimoorinia, Mansoureh Jalilkhany, Jillian M. Scudder, Jaclyn Jensen, and Sara L. Ellison. A reassessment of strong line metallicity conversions in the machine learning era. MNRAS, 503 (1):1082–1095, May 2021. doi: 10.1093/mnras/stab466.

A. Toomre. On the gravitational stability of a disk of stars. ApJ, 139:1217–1238, May 1964. doi: 10.1086/147861.

Paul Torrey, Mark Vogelsberger, Lars Hernquist, Ryan McKinnon, Federico Marinacci, Robert A. Simcoe, Volker Springel, Annalisa Pillepich, Jill Naiman, Rüdiger Pakmor, Rainer Weinberger, Dylan Nelson, and Shy Genel. Similar star formation rate and metallicity variability time-scales drive the fundamental metallicity relation. MNRAS, 477(1):L16–L20, June 2018. doi: 10.1093/mnrasl/sly031.

Freeke van de Voort and Joop Schaye. Properties of gas in and around galaxy haloes. MNRAS, 423 (4):2991–3010, July 2012. doi: 10.1111/j.1365-2966.2012.20949.x.

A. van der Wel, K. Noeske, R. Bezanson, C. Pacifici, A. Gallazzi, M. Franx, J. C. Muñoz-Mateos, E. F. Bell, G. Brammer, S. Charlot, P. Chauké, I. Labbé, M. V. Maseda, A. Muzzin, H. W. Rix, D. Sobral, J. van de Sande, P. G. van Dokkum, V. Wild, and C. Wolf. The VLT LEGA-C Spectroscopic Survey: The Physics of Galaxies at a Lookback Time of 7 Gyr. ApJS, 223(2):29, April 2016. doi: 10.3847/0067-0049/223/2/29.

Arjen van der Wel, Rachel Bezanson, Francesco D'Eugenio, Caroline Straatman, Marijn Franx, Josha van Houdt, Michael V. Maseda, Anna Gallazzi, Po-Feng Wu, Camilla Pacifici, Ivana Barisic, Gabriel B. Brammer, Juan Carlos Munoz-Mateos, Sarah Vervalcke, Stefano Zibetti, David Sobral, Anna de Graaff, Joao Calhau, Yasha Kaushal, Adam Muzzin, Eric F. Bell, and Pieter G. van Dokkum. The Large Early Galaxy Astrophysics Census (LEGA-C) Data Release 3: 3000 High-quality Spectra of $K_s$-selected Galaxies at z > 0.6. ApJS, 256(2):44, October 2021. doi: 10.3847/1538-4365/ac1356.

Floor van Donkelaar, Lucio Mayer, Pedro R. Capelo, Tomas Tamfal, Thomas R. Quinn, and Piero Madau. Stellar cluster formation in a Milky Way-sized galaxy at z > 4 - I. The proto-globular cluster population and the imposter amongst us. MNRAS, 522(2):1726–1735, June 2023. doi: 10.1093/mnras/stad946.

Maria L. van Loon, Peter D. Mitchell, and Joop Schaye. Explaining the scatter in the galaxy mass-metallicity relation with gas flows. MNRAS, 504(4):4817–4828, July 2021. doi: 10.1093/mnras/stab1254.

E. Vanzella, M. Castellano, P. Bergamini, T. Treu, A. Mercurio, C. Scarlata, P. Rosati, C. Grillo, A. Acebron, G. B. Caminha, M. Nonino, T. Nanayakkara, G. Roberts-Borsani, M. Bradac, X. Wang, G. Brammer, V. Strait, B. Vulcani, U. Meštrić, M. Meneghetti, F. Calura, Alaina Henry, A. Zanella, M. Trenti, K. Boyett, T. Morishita, A. Calabrò, K. Glazebrook, D. Marchesini, S. Birrer, L. Yang, and T. Jones. Early Results from GLASS-JWST. VII. Evidence for Lensed, Gravitationally Bound Protoglobular Clusters at z = 4 in the Hubble Frontier Field A2744. ApJ, 940(2):L53, December 2022. doi: 10.3847/2041-8213/ac8c2d.

E. Vanzella, M. Messa, A. Adamo, F. Loiacono, M. Oguri, K. Sharon, L. D. Bradley, P. Bergamini, M. Meneghetti, A. Claeyssens, B. Welch, M. Bradac, A. Zanella, A. Bolamperti, F. Calura, T. Y-Y. Hsiao, E. Zackrisson, M. Ricotti, L. Christensen, J. M. Diego, F. E. Bauer, X. Xu, S. Fujimoto, C. Grillo, M. Lombardi, P. Rosati, T. Resseguier, A. Zitrin, A. Bik, J. Richard, Abdurro'uf, R. Bhatawdekar, D. Coe, B. Frye, A. K. Inoue Y. Jimenez-Teja, C. Norman, J. R. Rigby, M. Trenti, and T. Hashimoto. The z = 9.625 Cosmic Gems Galaxy was a "Compact Blue Monster" Propelled by Massive Star Clusters. *arXiv e-prints*, art. arXiv:2507.18699, July 2025. doi: 10.48550/arXiv.2507.18699.

Tsan-Ming Wang, Benjamin Magnelli, Eva Schinnerer, Daizhong Liu, Ziad Aziz Modak, Eric Faustino Jiménez-Andrade, Christos Karoumpis, Vasily Kokorev, and Frank Bertoldi. $A^3$COSMOS: A census on the molecular gas mass and extent of main-sequence galaxies across cosmic time. A&A, 660:A142, April 2022. doi: 10.1051/0004-6361/202142299.

Stefan Waterval, Carlo Cannarozzo, and Andrea V. Macciò. Gas accretion at high redshift: cold flows all the way. MNRAS, 537(3):2726–2751, March 2025. doi: 10.1093/mnras/staf198.

Benjamin J. Weiner, Christopher N. A. Willmer, S. M. Faber, Jason Melbourne, Susan A. Kassin, Andrew C. Phillips, Justin Harker, A. J. Metevier, N. P. Vogt, and D. C. Koo. A Survey of Galaxy Kinematics to z~1 in the TKRS/GOODS-N Field. I. Rotation and Dispersion Properties. ApJ, 653(2):1027–1048, December 2006. doi: 10.1086/508921.

Chris J. Willott, René Doyon, Loic Albert, Gabriel B. Brammer, William V. Dixon, Koraljka Muzic, Swara Ravindranath, Aleks Scholz, Roberto Abraham, Étienne Artigau, Maruša Bradač, Paul Goudfrooij, John B. Hutchings, Kartheik G. Iyer, Ray Jayawardhana, Stephanie LaMassa, Nicholas Martis, Michael R. Meyer, Takahiro Morishita, Lamiya Mowla, Adam Muzzin, Gaël Noirot, Camilla Pacifici, Neil Rowlands, Ghassan Sarrouh, Marcin Sawicki, Joanna M. Taylor, Kevin Volk, and Johannes Zabl. The Near-infrared Imager and Slitless Spectrograph for the James Webb Space Telescope. II. Wide Field Slitless Spectroscopy. PASP, 134(1032):025002, February 2022. doi: 10.1088/1538-3873/ac5158.

E. Wisnioski, N. M. Förster Schreiber, M. Fossati, J. T. Mendel, D. Wilman, R. Genzel, R. Bender, S. Wuyts, R. L. Davies, H. Übler, K. Bandara, A. Beifiori, S. Belli, G. Brammer, J. Chan, R. I. Davies, M. Fabricius, A. Galametz, P. Lang, D. Lutz, E. J. Nelson, I. Momcheva, S. Price, D. Rosario, R. Saglia, S. Seitz, T. Shimizu, L. J. Tacconi, K. Tadaki, P. G. van Dokkum, and E. Wuyts. The KMOS$^{3D}$ Survey: Data Release and Final Survey Paper. ApJ, 886(2):124, December 2019. doi: 10.3847/1538-4357/ab4db8.

Emily Wisnioski, Karl Glazebrook, Chris Blake, Ted Wyder, Chris Martin, Gregory B. Poole, Rob Sharp, Warrick Couch, Glenn G. Kacprzak, Sarah Brough, Matthew Colless, Carlos Contreras, Scott Croom, Darren Croton, Tamara Davis, Michael J. Drinkwater, Karl Forster, David G. Gilbank, Michael Gladders, Ben Jelliffe, Russell J. Jurek, I. Hui Li, Barry Madore, Kevin Pimbblet, Michael Pracy, David Woods, and H. K. C. Yee. The WiggleZ Dark Energy Survey: high-resolution kinematics of luminous star-forming galaxies. MNRAS, 417(4):2601–2623, November 2011. doi: 10.1111/j.1365-2966.2011.19429.x.

Sunna Withers, Adam Muzzin, Swara Ravindranath, Ghassan T. E. Sarrouh, Roberto Abraham, Yoshihisa Asada, Maruša Bradač, Gabriel Brammer, Guillaume Desprez, Kartheik Iyer, Nicholas Martis, Lamiya Mowla, Gaël Noirot, Marcin Sawicki, Victoria Strait, and Chris J. Willott. Spectroscopy from Photometry: A Population of Extreme Emission Line Galaxies at $1.7 \lesssim z \lesssim 6.7$ Selected with JWST Medium Band Filters. ApJ, 958(1):L14, November 2023. doi: 10.3847/2041-8213/ad01c0.

Stijn Wuyts, Natascha M. Förster Schreiber, Reinhard Genzel, Yicheng Guo, Guillermo Barro, Eric F. Bell, Avishai Dekel, Sandra M. Faber, Henry C. Ferguson, Mauro Giavalisco, Norman A. Grogin, Nimish P. Hathi, Kuang-Han Huang, Dale D. Kocevski, Anton M. Koekemoer, David C. Koo, Jennifer Lotz, Dieter Lutz, Elizabeth McGrath, Jeffrey A. Newman, David Rosario, Amelie Saintonge, Linda J. Tacconi, Benjamin J. Weiner, and Arjen van der Wel. Smooth(er) Stellar Mass Maps in CANDELS: Constraints on the Longevity of Clumps in High-redshift Star-forming Galaxies. ApJ, 753(2):114, July 2012. doi: 10.1088/0004-637X/753/2/114.

Stijn Wuyts, Natascha M. Förster Schreiber, Erica J. Nelson, Pieter G. van Dokkum, Gabe Brammer, Yu-Yen Chang, Sandra M. Faber, Henry C. Ferguson, Marijn Franx, Mattia Fumagalli, Reinhard Genzel, Norman A. Grogin, Dale D. Kocevski, Anton M. Koekemoer, Britt Lundgren, Dieter Lutz, Elizabeth J. McGrath, Ivelina Momcheva, David Rosario, Rosalind E. Skelton, Linda J. Tacconi, Arjen van der Wel, and Katherine E. Whitaker. A CANDELS-3D-HST synergy: Resolved Star Formation Patterns at $0.7 < z < 1.5$. ApJ, 779(2):135, December 2013. doi: 10.1088/0004-637X/779/2/135.

H. J. Zahid, L. J. Kewley, and F. Bresolin. The Mass-Metallicity and Luminosity-Metallicity Relations from DEEP2 at z ~0.8. ApJ, 730(2):137, April 2011. doi: 10.1088/0004-637X/730/2/137.

H. Jabran Zahid, Gabriel I. Dima, Rolf-Peter Kudritzki, Lisa J. Kewley, Margaret J. Geller, Ho Seong Hwang, John D. Silverman, and Daichi Kashino. The Universal Relation of Galactic Chemical Evolution: The Origin of the Mass-Metallicity Relation. ApJ, 791(2):130, August 2014. doi: 10.1088/0004-637X/791/2/130.

A. Zanella, E. Le Floc'h, C. M. Harrison, E. Daddi, E. Bernhard, R. Gobat, V. Strazzullo, F. Valentino, A. Cibinel, J. Sánchez Almeida, M. Kohandel, J. Fensch, M. Behrendt, A. Burkert, M. Onodera, F. Bournaud, and J. Scholtz. A contribution of star-forming clumps and accreting satellites to the mass assembly of z $\sim$ 2 galaxies. MNRAS, 489(2):2792–2818, October 2019. doi: 10.1093/mnras/stz2099.