University Library



A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Dalmasso, Nicolò

Title:
The Galaxy–Dark Matter Halo Connection at Cosmic Dawn

Date:
2025

Persistent Link:
https://hdl.handle.net/11343/363978

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.

The Galaxy–Dark Matter Halo Connection at Cosmic Dawn

by

Nicolò Dalmasso

ORCID: 0000-0002-1850-4050

A thesis submitted in total fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Science
School of Physics
THE UNIVERSITY OF MELBOURNE

December 2025

THE UNIVERSITY OF MELBOURNE

Abstract

Faculty of Science
School of Physics

Doctor of Philosophy

by Nicolò Dalmasso
ORCID: 0000-0002-1850-4050

Understanding galaxy formation and evolution during cosmic dawn ($z > 6$) remains one of the most challenging frontiers in astrophysics. While theoretical models predict that the first galaxies formed within dark matter halos of specific masses, observational constraints on the relationship between early galaxies and their host halos remain limited. Key open questions include: How do galaxies populate dark matter halos at these early epochs? What drives the observed luminosity function evolution? How do environmental effects and mergers contribute to early galaxy assembly? These questions are fundamental to understanding reionization, early star formation processes, and the emergence of cosmic structure.

Galaxy clustering measurements provide a unique window into the galaxy-halo connection by quantifying the spatial distribution of galaxies and their correlation with underlying dark matter structure. Through clustering analysis, we can constrain halo occupation statistics, measure galaxy bias evolution, and infer characteristic halo masses hosting observed galaxy populations. However, traditional clustering techniques face significant challenges at high redshifts due to small sample sizes, non-uniform survey depths, and complex selection effects that can bias correlation function measurements.

This work introduces a novel approach for measuring the two-point correlation function of galaxies in narrow beam surveys with varying depths, addressing the systematic biases introduced by the heterogeneous depth strategy characteristic of modern deep surveys. We develop a Monte Carlo source injection and recovery methodology using the GLACiAR2 code to generate reference random catalogs that accurately account for spatially varying completeness, depth variations, and photometric scatter effects. Our technique replaces traditional uniform random point distributions over the field of view with realistic source recovery simulations that incorporate spectral energy distribution modeling and the full photometric pipeline. We validate this approach by demonstrating consistency between results from uniform-depth sub-regions and heterogeneous survey areas, extending reliable clustering measurements to $z \sim 8$ using the combined HST CANDELS GOODS datasets with sample sizes as small as $N_{\rm g} = 160$ sources.

Applying this validated methodology to deep JWST/NIRCam observations from the JADES survey, we analyze $N_{\rm g} \simeq 7000$ photometrically-selected Lyman Break Galaxies to extend galaxy clustering measurements to $z \sim 11$, achieving the first measurement of galaxy clustering at less than 500 million years since the Big Bang. Using power-law correlation function fitting, we measure galaxy bias $b = 9.6 \pm 1.7$ for sources at $z = 10.6$ with $M_{\rm F200W} < -17.3$, indicating host dark matter halos of approximately $M_{\rm h} \sim 10^{10.6}\, M_{\odot}$.

Our HOD modeling of an expanded sample ($N_{\rm g} \simeq 6500$) reveals accelerated evolution of galaxy-halo relationships, with characteristic halo masses decreasing by over an order of magnitude from $M_{\rm h} \sim 10^{11.45}\, M_{\odot}$ at $z = 5.5$ to $M_{\rm h} \sim 10^{10.12}\, M_{\odot}$ at $z = 10.6$, while effective bias increases from $b_{\rm g}^{\rm eff} = 5.64^{+0.10}_{-0.13}$ to $b_{\rm g}^{\rm eff} = 8.13^{+0.04}_{-0.02}$ for a sample of candidates with $M_{\rm F200W} < -17.0$. We demonstrate that satellite galaxy fractions peak during the epoch of reionization ($z \sim 7-8$), reaching $\sim 28\%$ before declining to $< 1\%$ by $z \sim 5-6$.

To complement clustering analysis, we investigate galaxy merger rates through morphological analysis of high-resolution JWST/NIRCam imaging in the Abell 2744 field, extending merger studies to redshift $z \sim 9$ (lookback time ~ 13.3 billion years). Using quantitative morphological parameters (Gini, M_{20}, and Asymmetry) applied to 675 galaxy candidates ($M_{\rm U} \in [-26.6, -17.9]$), we identify 64 merger candidates, yielding a merger fraction $f_{\rm m} = 0.11 \pm 0.04$ that shows no evidence of redshift evolution even at the highest redshifts considered.

This extends the constant merger fraction trend previously observed at $z \lesssim 6$ well into the epoch of reionization. We find no significant correlation between specific star formation rates and merger status, with deviations typically falling within $1 - 1.5\sigma$ of the null hypothesis. The redshift-independent merger fraction supports a scenario where increasing dark matter halo merger rates at higher redshifts are balanced by

decreasing dynamical timescales, providing independent constraints on galaxy assembly mechanisms during the first billion years complementary to clustering analysis.

To ensure robust spectroscopic measurements, we conduct a flux calibration intercomparison between JWST's NIRISS Wide-Field Slitless Spectroscopy and NIRSpec Multi-Object Spectroscopy modes. Analysis of 12 spectroscopically confirmed galaxies ($z \sim 1-3$) in the Abell 2744 field reveals systematic differences dependent on source morphology. Compact objects show comparable flux between instruments (within 1σ), while extended sources exhibit consistently higher NIRISS measurements due to NIRSpec slit losses. Monte Carlo analysis quantifies that NIRSpec recovers $\geq 63\%$ of NIRISS flux for sources with slit coverage $> 15\%$ or $R_{\rm e} < 1$ kpc, dropping to $24-63\%$ for larger or poorly covered sources.

Critical for galaxy characterization, we demonstrate that Hα λ6563/[O III] λ5007 emission line ratios can vary by up to ~ 0.3 dex between instruments, directly impacting metallicity and star formation rate derivations. These systematic differences highlight the necessity of considering instrumental effects when combining multi-instrument spectroscopic datasets, providing essential calibration framework for high-redshift galaxy studies utilizing both JWST spectrographs.

This work establishes the first robust observational framework for clustering analysis at $z > 8$, providing fundamental constraints on early galaxy formation models and reionization scenarios. Our results directly inform theoretical predictions about dark matter halo occupation, early galaxy assembly, and environmental effects during cosmic dawn. The methodological advances presented here enable future studies with larger JWST datasets, while our empirical constraints on the galaxy-halo connection provide crucial benchmarks for next-generation cosmological simulations and semi-analytical models of galaxy formation.

Declaration of Authorship

I, Nicolò Dalmasso, declare that this thesis titled, "The Galaxy–Dark Matter Halo Connection at Cosmic Dawn" and the work presented in it are my own. I confirm that:

- The thesis comprises only my original work towards the Doctor of Philosophy except where indicated in the preface;
- due acknowledgement has been made in the text to all other material used; and
- the thesis is fewer than 100,000 words in length,, exclusive of tables, maps, bibliographies and appendices as approved by the Research Higher Degrees Committee.

Signed: Nicolò Dalmasso

Date: 11/09/2025

Preface

This thesis is presented across seven chapters: an opening introduction, five chapters detailing original research contributions, and a final summary and conclusions chapter. While the majority of the work is my own, there are certain aspects that have been developed through collaborations. Any contributions from external sources are appropriately acknowledged within the text.

- Chapter 1 establishes the foundation for this thesis by introducing our current cosmological framework and fundamental concepts in galaxy formation and evolution. It examines galaxy clustering techniques and their measurements, highlighting dark matter's role in shaping universal structure. The chapter concludes with a review of the key telescopes that provided the observational data central to this research, equipping readers with the necessary background for the original research presented in subsequent chapters.

- Chapter 2 is published as:
  Nicolò Dalmasso, Michele Trenti, Nicha Leethochawalit
  "Galaxy clustering measurements out to redshift ~ 8 from Hubble Legacy Fields", 19 December 2023, Monthly Notices of the Royal Astronomical Society, Volume 528, Issue 1, February 2024, Pages 898–908.

- Chapter 3 is published as:
  Nicolò Dalmasso, Nicha Leethochawalit, Michele Trenti, Kristan Boyett
  "Galaxy clustering at cosmic dawn from JWST/NIRCam observations to redshift z$\sim$11", 21 August 2024, Monthly Notices of the Royal Astronomical Society, Volume 533, Issue 2, September 2024, Pages 2391–2398.

- Chapter 4 is currently pending publication as:
  Nicolò Dalmasso, Giovanni Ferrami, Nicha Leethochawalit, Emanuele M. Ventura, Michele Trenti
  "Accelerated evolution of galaxy host halo masses during Cosmic Dawn from deep JWST clustering".
  The material presented here is the final submitted version pending publication on "Monthly Notices of the Royal Astronomical Society".

- Chapter 5 is published as:
  Nicolò Dalmasso, Antonello Calabrò, Nicha Leethochawalit, Benedetta Vulcani, Kristan Boyett, Michele Trenti, Tommaso Treu, Marco Castellano, Maruša Bradač, Benjamin Metha, Paola Santini
  "The rate and contribution of mergers to mass assembly from NIRCam observations of galaxy candidates up to 13.3 billion years ago", 3 September 2024, Monthly Notices of the Royal Astronomical Society, Volume 533, Issue 4, October 2024, Pages 4472–4484.

- Chapter 6 is published as:
  Nicolò Dalmasso, Peter J. Watson, Tommaso Treu, Michele Trenti, Benedetta Vulcani, Themiya Nanayakkara, Maruša Bradač, Tucker Jones, Kristan Boyett, Xin Wang, Sara Mascia, Laura Pentericci
  "Quantifying Spectroscopic Flux Variations Between JWST NIRISS and NIRSpec: Slit Losses in Emission Line Measurements of z∼1-3 Galaxies", 24 October 2025, Monthly Notices of the Royal Astronomical Society, Volume 544, Issue 2, December 2025, Pages 1915–1925.

- Chapter 7 synthesizes the key scientific discoveries of this doctoral research, presenting a cohesive conclusion that integrates the major theoretical and observational findings. It offers critical insights into galactic structures and their development across cosmic time, while highlighting implications for future research directions in the field.

Acknowledgements

Thanks to Nicolò.

This is for myself. The effort, the joy, the pain, the doubts, the certainties, the distance, the loneliness, the sadness, the happiness, the relief, and the uncertainties that lie ahead.

It may sound selfish, but it's truly how I feel. It's important to acknowledge that I did it. Of course, I had all the support I needed from friends, colleagues, family, but in the end, I did it. This is something nobody will ever take away from me.

Thanks to my parents, Simona and Piero, who have supported me all the way since the very beginning, allowing my brother and me to attend University even at the cost of financial sacrifices. They gave us the chance to follow our aspirations and find our own path.

Thanks to my supervisor, Michele, for accepting me as a PhD student from the other side of the world, enabling me to move countries and complete this chapter of my life.

Thanks Gio and Manu. Thanks to the new friends I've made here, inside and outside the office, for the memories we've created and shared.

Thanks to Compa, Pino, and Silvia. Friends made in tough times are the ones with whom you share true values. Never forget Via Pietro Giuria 1.

Till next time, stay safe.

"You don't stop when you don't feel like it anymore but you stop when you are done."
- Nicolò Dalmasso

Contents

Abstract i

Declaration of Authorship iv

Preface v

Acknowledgements vii

List of Figures xi

List of Tables xvi

1 Introduction 1
- 1.1 The Cosmological Framework . . . 1
- 1.2 Early Universe . . . 2
- 1.3 Dark Matter Halos and Cosmic Web . . . 3
- 1.4 First light . . . 4
- 1.5 Cosmic noon and modern Universe . . . 5
- 1.6 Galaxy Clustering . . . 7
  - 1.6.1 Theoretical Foundations and Observational Methodologies . . . 7
  - 1.6.2 Evolution of High-Redshift Clustering Studies . . . 8
  - 1.6.3 Physical Implications for Galaxy Formation . . . 10
- 1.7 Observational Facilities into the Early Universe . . . 12
  - 1.7.1 The Hubble Space Telescope Legacy . . . 12
    - 1.7.1.1 Mission . . . 12
    - 1.7.1.2 Instrumentation . . . 12
    - 1.7.1.3 Legacy . . . 13
  - 1.7.2 The James Webb Space Telescope Era . . . 14
    - 1.7.2.1 Mission . . . 14
    - 1.7.2.2 Instrumentation . . . 14
    - 1.7.2.3 First Results . . . 15
- 1.8 Thesis structure . . . 16

2 Galaxy clustering measurements out to redshift ∼8 from Hubble Legacy Fields 17

2.1 Abstract . . . 17
2.2 Introduction . . . 18
2.3 Data Selection . . . 21
2.4 Clustering estimation . . . 22
2.5 Clustering analysis methods . . . 25
2.5.1 Random point catalog generation . . . 26
2.5.1.1 Uniform random points . . . 26
2.5.1.2 Random points from source recovery map . . . 26
2.5.2 Spatial dependence of the clustering signal and random points methods comparison . . . 27
2.6 Clustering Results . . . 29
2.6.1 Angular correlation function and Bias . . . 29
2.6.2 Duty Cycle . . . 33
2.7 Summary and perspectives . . . 35

3 Galaxy clustering at cosmic dawn from JWST/NIRCam observations to redshift z∼11 39
3.1 Abstract . . . 39
3.2 Introduction . . . 40
3.3 Data selection . . . 42
3.4 Clustering estimation . . . 44
3.4.1 ACF fitting model . . . 44
3.4.2 Random points catalog . . . 46
3.5 Results and Discussion . . . 47
3.5.1 Angular correlation function and Galaxy bias . . . 47
3.5.2 Galaxy bias magnitude dependence . . . 49
3.5.2.1 ACF measured for the bias evolution with magnitude . . . 50
3.5.3 Abundance matching and Duty Cycle . . . 50
3.6 Summary and perspectives . . . 53

4 Accelerated evolution of galaxy host halo masses during Cosmic Dawn from deep JWST clustering 54
4.1 Abstract . . . 54
4.2 Introduction . . . 55
4.3 Data sets and sample selection . . . 57
4.3.1 JADES GOODS-South Data . . . 57
4.3.2 Random points catalog . . . 58
4.4 Clustering Analysis . . . 59
4.4.1 Angular Correlation Function (ACF) estimation . . . 59
4.4.2 Halo Occupation Distribution (HOD) model . . . 60
4.5 Dark Matter Halo Clustering estimations . . . 64
4.5.1 HOD model fitting . . . 64
4.5.2 Satellite fraction, Effective bias and Dark Matter Halo mass . . . 67
4.6 Summary and Perspectives . . . 70

5 The rate and contribution of mergers to mass assembly from NIRCam observations of galaxy candidates up to 13.3 billion years ago 72
5.1 Abstract . . . 72

5.2 Introduction . . . 73
5.3 Observational data and sample selection . . . 76
5.4 Analysis . . . 78
5.4.1 Morphological Parameter Definition . . . 78
5.4.2 Identifying Mergers . . . 80
5.5 Results and Discussion . . . 81
5.5.1 Mergers characterization . . . 81
5.5.2 Merger fraction versus redshift . . . 83
5.5.3 Wavelength dependency study . . . 88
5.5.4 sSFR of interacting and non-interacting systems . . . 88
5.6 Summary and perspectives . . . 93

6 Quantifying Spectroscopic Flux Variations between JWST NIRISS and NIRSpec: Slit Losses in Emission Line Measurements of z 1-3 Galaxies 95
6.1 Abstract . . . 95
6.2 Introduction . . . 96
6.3 Data sets and sample selection . . . 98
6.4 Slit loss analysis-characterization . . . 101
6.4.1 Characterization of Emission Line Fluxes and Equivalent Widths . 101
6.4.2 Flux Characterization in Emission Line Maps . . . 108
6.5 Summary and Conclusions . . . 111

7 Summary and Conclusions 115
7.1 Galaxy clustering measurement at z∼8 and beyond . . . 116
7.2 High redshift galaxies study from NIRCam imaging . . . 117
7.3 JWST NIRISS-NIRSpec Performance Analysis . . . 117

Bibliography 119

List of Figures

2.1 Scientific and root-mean-square maps for GOODS-N and GOODS-S . . . 23
2.2 ACFs results comparison between different generated random points catalogs . . . 28
2.3 ACFs results from galaxies contained in GOODS-N and GOODS-S . . . 30
2.4 ACFs results from galaxies contained in XDF . . . 31
2.5 Galaxy bias evolution in respect to redshift $b(z)$ . . . 31
2.6 Correletion length comparison with previous work in respect to redshift $r_0(z)$ . . . 35
2.7 Comparison of ACFs obtained considering the deepest region in the survey against the entire area . . . 36
2.8 Duty cycle evolution in respect to redshift $\epsilon_{DC}(z)$ . . . 37

3.1 Measured ACFs in the GOODS-S field at six different mean redshifts. The black dots denote the ACF measurements with associated 1σ uncertainties, the blue solid lines represent a the power law best-fit with parameters listed in the legend box. . . . 48
3.2 Galaxy bias evolution as a function of redshift z. Circles represent the results obtained in this work. Other symbols are previous bias measurements: stars Overzier et al. [197], squares Barone-Nugent et al. [17], downward triangles Harikane et al. [117], diamonds Qiu et al. [219] and upward triangles Dalmasso et al. [66]. Green dashed lines represent the theoretical bias evolution at a fixed dark matter halo mass M_h $[M_\odot]$. Note that our observations are substantially deeper than other studies, hence the inferred bias is lower. . . . 49
3.3 Galaxy bias evolution as a function of absolute magnitude M_{F200W} for three different redshift bins (blue, orange and grey for $z = 5, 6, 7$ respectively. Circles represent the results obtained in this work. Other symbols are literature measurements: stars Overzier et al. [197], squares Barone-Nugent et al. [17], downward triangles Harikane et al. [117], diamonds Qiu et al. [219] and upward triangles Dalmasso et al. [66]. Colored dashed lines represent the theoretical bias evolution model as a function of absolute magnitude M_{F200W}, derived from Trenti et al. [268]. . . . 51
3.4 ACFs measured in the GOODS-S field for each subsample are depicted in Fig.3.3. Solid lines represent the best-fit power-law function (Eq.3.4) with $\beta = 0.6$ held constant as in the main analysis. The mean redshift and absolute magnitude are indicated in the upper right corner of each panel. . . . 52

4.1 Measured ACFs in the GOODS-S field at six mean redshifts from bootstrap resampling with 1σ uncertainties derived from the covariance matrix. Blue line represent the HOD best-fit model ($\omega_{\rm mod}(\theta)$), with orange and red lines showing the one-halo ($\omega_{\rm 1h}$) and two-halo ($\omega_{\rm 2h}$) terms. Shaded areas indicate 1σ uncertainties. Mean redshift and absolute UV magnitude thresholds (F200W band) are shown in the upper right of each panel. 63
4.2 Normalized redshift distributions of the parent galaxy samples, with fine structure resulting from a combination of photometric redshift selection efficiency and sample variance in excess of Poisson noise due to clustering. 64
4.3 Evolution of clustering parameters as a function of redshift: satellite fraction $f_{\rm sat}$ (top panel), effective galaxy bias $b_{\rm g}^{\rm eff}$ (middle panel), and characteristic halo mass $M_{\rm h}[M_\odot]$ (bottom panel). Results from this work are shown as solid red squares, while previous HOD modelling studies are color-coded according to the legend (Bhowmick et al. [28], Harikane et al. [117, 118], Paquereau et al. [199], Shuntov et al. [249]). 69

5.1 NIRCam scientific map in the F150W filter. The red region represents the area where the magnification is $\mu \geq 2.0$. The galaxies detected in this region of the survey are excluded from the parent sample. 77
5.2 Illustrative representations from various redshift ranges, featuring extracted cutouts alongside corresponding morphological metrics, ID and photometric redshift. These snapshots are derived from a composite `hres` map (20mas per pixel), integrating data from F150W and F200W bands. The upper row highlights non-merging galaxies, while the lower row presents merger candidates. Each cutout stamp is 2.2" on a side, while the horizontal bar is 0.5" wide and the white contour represents the 2σ segmentation map boundary. 80
5.3 Visual representation of the classification method adopted in this study for five different z-bins (see Sec.5.4.2). Galaxies located above the horizontal dashed line (representing Eq.5.4) are labelled with blue triangles and are classified in the Silver sample. Among these, the subset to the right of the vertical dashed line (representing Eq.5.5) is highlighted in red, representing the galaxies in the Gold sample. We used high-resolution reduction of the F150W and F200W NIRCam imaging (see Sec.5.3). . . . 81
5.4 Demography for the galaxy (merger) samples versus redshift. Points correspond to those in Fig.5.3. Black squares indicate the total number of candidates considered in each redshift bin (bold number in column (2) in Tab.5.1). Blue triangles represent galaxies in the Silver sample, while red triangles denote those classified in the Gold sample. 82
5.5 Rest frame U-magnitude (left) and stellar mass (right) as a function of photometric redshift for the Gold sample (red) and for all other, non-Gold galaxies (grey). Solid lines represent the corresponding median evolution in respect to the mean redshift in each bin 83
5.6 The Gold sample and non-Gold galaxies obtained in all the different redshift bins. Note that the central region is not populated as we excluded regions with magnification $\mu \geq 2.0$, Fig.5.1. 84

5.7 Merger fraction of the Silver sample (left panels) and of the Gold sample (right panels) as a function of redshift. The definition of the fraction follows Eq.5.6. In the top panels, we show a further separation based on the rest frame magnitude, with bright galaxies classified as having U-band magnitudes below $M_U = -20.1$, and faint galaxies lying above this threshold. In the bottom panels, we separate galaxies into a high mass sample ($M_* > 10^{8.5} M_\odot$) and a low mass sample ($M_* < 10^{8.5} M_\odot$). 85

5.8 Comparison of the merger fraction (f_m) evolution with respect to redshift in our work against literature measurements from: Cibinel et al. [52], Conselice & Arnold [56], Conselice et al. [59], Lotz et al. [166], Man et al. [172], Romano et al. [234], Ventou et al. [274]. Note that each study implements their own catalogs, data selection, and methodologies for identifying mergers, leading to a non homogeneous comparison that may be affected substantially by systematic effects. 86

5.9 Mergers classification method presented in Sec.5.4.2. Each row shows results for a different redshift bin, with each panel within each row showing results when a different filter from NIRCam is used for determining the morphological parameters. Galaxies located above the horizontal dashed line (representing Eq.5.4) are labelled with blue triangles and are classified in the Silver sample. Among these, the subset to the right of the vertical dashed line (representing Eq.5.5) is highlighted in red, representing the galaxies in the Gold sample. 89

5.10 Demography for the galaxy (merger) samples versus redshift. Points correspond to those in Fig.5.9. Black squares indicate the total number of candidates considered in each redshift bin. Blue triangles represent galaxies in the Silver sample, while red triangles denote those classified in the Gold sample. 89

5.11 Merger fraction evolution in respect to redshift, using different filters on the NIRCam instrument for the identification. The left panels show the Silver sample and the right panels represent the Gold sample. In each panel, we also show the results for the subsamples of bright and faint galaxies, defined by a cut in magnitude at $M_U = -20.1$. 90

5.12 As in Fig.5.11, but splitting the parent sample into two subsamples based on stellar mass, with high-mass galaxies defined to be those with stellar masses greater than $M_* = 10^{8.5} M_\odot$, and low-mass galaxies having stellar masses below this value. 91

5.13 Excess specific star formation rate (sSFR) evolution as a function of redshift for both the complete set of Gold sample merger and non-Gold galaxies (see Eq.5.8). The shaded areas indicate uncertainty ranges of $\sigma = (1, 1.5, 2)$. The measurement is statistically consistent with no redshift evolution. 92

6.1 NIRISS imaging cutouts are presented using a drizzled stack of the F115W, F150W, and F200W filters, processed to better highlight the light associated with each galaxy candidate. Solid white contours represent the NIRISS segmentation maps, rectangles show the NIRSpec slits assigned to the target (solid magenta and dash-dot white). The ID and redshift z refer to measurements from NIRISS, using the catalog presented in Watson et al. [278] and a 0.6 arcsec scale bar is included. The "A" parameter, shown in the bottom right corner, represents the fractional area of the segmentation map that is encompassed by the slit. Objects are classified as either "compact" or "extended" based on the classification criteria presented in Sec.6.3. 99

6.2 Example Analysis: Spectroscopically Confirmed Candidate NIRISS ID 2744 at z=1.86. Top left: NIRISS image cutout (0.6 arcsec scale bar shown) with the NIRISS segmentation map (solid white contour) and NIRSpec slits assigned to the target (solid magenta and dash-dot white). Top right: Combined spectral data showing the original NIRSpec spectrum (gray), the NIRISS spectrum (blue), and the NIRSpec spectrum after convolution to match the NIRISS resolution (red), with identified emission lines marked. Bottom panels: Detailed emission line fits for all the line matched between the two instruments. The shaded red regions in each panel indicate an example of the continuum estimation regions computed for a single iteration of the MC sampling. 102

6.3 Probability density distributions of $\text{flux}_{\text{ratio}}$ (top row) and EW_{ratio} (bottom row) for the emission lines of the source presented in Fig.6.2. The distributions are shown in logarithmic scale for visualization purposes and represent the marginalized probability densities obtained by varying the continuum estimation regions. In red the 50th and from left to right in black 16th and 84th percentiles. 103

6.4 All panels show results for the observed sources from three different JWST programs, color-coded by subsample (compact or extended). Error bars indicate 1σ uncertainties calculated using the methodology described in Sec.6.4.1. In these panels, we see that measurements for compact sources from NIRISS and NIRSpec are in good agreement. However, for more extended objects (shown on the left side of panel c), the NIRSpec measurements are systematically lower, as expected. 106

6.5 Same format as Fig.6.4, but showing the logarithmic EW_{ratio} as a function of various quantities. Unlike the dichotomy seen in $\text{flux}_{\text{ratio}}$, both compact and extended sources display similar scatter in EW_{ratio} with no strong trends. 107

6.6 NIRISS imaging cutouts from the corrisponding emission line maps for extended galaxies. Solid white contours represent NIRISS segmentation maps, while rectangles indicate NIRSpec slit positioning. Magenta (og: original) denotes the fixed position from the original observation from the corresponding JWST program. The black one represents the slit corresponding to the most probable $\text{flux}_{\text{ratio}}$ (50th percentile) obtained from the MC analysis conducted on the associated emission line maps. For each cutout, the corresponding histogram displays the probability density of the $\text{flux}_{\text{ratio}}$ (Eq.6.1a) in logarithmic scale derived from the MC analysis and color-coded the same way with associated 1σ uncertainties in red. 111

6.7 $\text{flux}_{\text{ratio}}$ as a function of slit coverage area and effective radius R_e for extended galaxies.Results are color-coded by measurement method: "precise spectra fit" refers to measurements from the analysis in Sec.6.4.1, while "emission line map" indicates results from the analysis in Sec.6.4.2 (also presented in Tab.6.2). 112
6.8 Probability density histogram of the Hα λ6563 to $[O\,\text{III}]$ λ5007 emission line ratio for four extended sources. Multiple measurement methods are shown. NIRISS measurements: "seg map" represents the flux ratio measured within the segmentation map across both emission line maps. NIRSpec measurements: "og slit" shows flux measurements using the original assigned slit as the observed region, while "50th MC" indicates the $50th$ percentile from MC sampling (Sec.6.4.2). "Spectra Fit" for both instruments denotes the emission line ratio derived from precise emission line fitting detailed in Sec.6.4.1. 113

List of Tables

2.1 Galaxy candidates at different redshift 22
2.2 Summary of clustering parameters obtained from ACF analysis 34

3.1 Galaxy clustering parameters obtained from ACF analysis with power-law model with a fixed $\beta = 0.6$. (1) Redshift bin. (2) Mean redshift. (3) Number of galaxies in the redshift bin. (4) Number of galaxies selected according to the criteria (Sec.3.3). (5) Best-fit value of the amplitude A_w. (6) Best-fit value of the correlation length r_0 in units of h^{-1}Mpc. (7) Galaxy bias. (8) Dark matter halo mass (log scale) in units of $M_\odot$. (9) Duty cycle. (10) Reduced χ^2 value. 47

4.1 Magnitude limits of our galaxy samples, defined as the value where injection-recovery simulations achieve 50% source detection efficiency. Absolute magnitudes are in the rest-frame UV; apparent magnitudes are in F444W. 58
4.2 Summary of the clustering results from the HOD model best fit. Columns: (1) Mean redshift, (2) Number of galaxies, (3) Average UV absolute magnitude of the sample in the F200W band , (4) Best-fit value of $M_{\rm min}$, (5) Best-fit value of $M_{\rm sat}$, (6) Fraction of galaxy satellites in the dark matter halo (Eq.4.17), (7) Effective galaxy bias (Eq.4.18), (8) Mean dark matter halo mass $M_{\rm h}$ (Eq.4.19). 65

5.1 Summary of the galaxies used in this work for mergers identification, obtained using JWST NIRCam data from the GLASS Collaboration. The sources are selected as discussed in Sec.5.3. (1) Redshift bin. (2) Number of galaxies detected.. (3) Mean redshift of the subsample. (4) Absolute U Magnitude range of the subsample. (5) Stellar mass range. Quantities in columns (3), (4) and (5) are referred to the subsamples used in this work in **bold** (2). Both distributions of magnitude and stellar mass in respect to redshift are presesented in Fig. 5.5. 76

6.1 This table summarizes the galaxies used in this study for slit loss characterization, as discussed in Sec.6.3. Columns (8)-(12) list the emission lines analyzed in this work, with an "X" indicating lines that are matched by both the NIRISS and NIRSpec instruments and for which we successfully fitted the emission line to calculate the integrated flux. Column descriptions: (1) JWST Program; (2) Galaxy RA [deg]; (3) Galaxy DEC [deg]; (4) Associated NIRISS ID (Watson et al. [278]); (5) Associated NIRSpec ID (Mascia et al. [175], Price et al. [218]); (6) Spectroscopic redshift from NIRISS; (7) Apparent AB magnitude from NIRSpec in the F200W filter. 99

6.2 Emission line $\text{flux}_{\text{ratio}}$ in Eq.6.1a obtained from different emission line maps of extended galaxies. Column descriptions: (1) Associated NIRISS ID; (2) Area of the galaxy covered by the slit; $\text{flux}_{\text{ratio}}$ from emission line maps (3)-(6). 110

Chapter 1

Introduction

1.1 The Cosmological Framework

Modern cosmology emerged in the early 20th century with Albert Einstein's revolutionary theory of General Relativity in 1915 (Einstein [82]), which established a mathematical framework for understanding gravity. The theory's equations revealed a dynamic Universe, prompting Einstein to introduce a cosmic repulsive force, the cosmological constant Λ needed to maintain what was then believed to be a static Universe model (Einstein [83]).

During this transformative period, Dutch astronomer Jacobus Cornelius Kapteyn demonstrated that the Sun exists within a rotating stellar system (Kapteyn [140]), challenging earlier geocentric perspectives. The subsequent "Great Debate" between Heber Doust Curtis and Harlow Shapley further refined our cosmic understanding, establishing that Earth occupies an unremarkable position within the Milky Way, rather than residing at its center (Shapley & Curtis [244]). A pivotal breakthrough came through Henrietta Swan Leavitt's work on the period-luminosity relationship of Cepheid variables (Leavitt & Pickering [155]). Building upon this foundation, Edwin Powell Hubble discovered in 1929 what would later be named after him—the Hubble constant (Hubble [130]). This constant, H_0, quantifies the Universe's expansion rate by measuring the recession velocity of galaxies relative to their distances, demonstrating that the Universe expands uniformly when assumed homogeneous and isotropic. The Hubble law establishes the fundamental relationship between proper distance D (in Mpc) and recessional velocity v (in $\mathrm{km \cdot s^{-1}}$):

$$v = H_0 \cdot D. \tag{1.1}$$

The Universe's temporal evolution is characterized by a scale factor $a(t)$, related to the Hubble constant through:

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)}, \tag{1.2}$$

where $\dot{a}(t)$ represents the time derivative of $a(t)$. The scale factor equals zero at the Universe's birth (t_u) and unity at present time (t_0). This scale factor directly relates to the cosmological redshift z:

$$a(t) = \frac{1}{1+z}. \tag{1.3}$$

The redshift z measures the fractional change in a photon's wavelength from emission to observation:

$$z = \frac{\lambda_{obs}}{\lambda_{em}} - 1, \tag{1.4}$$

where λ_{em} and λ_{obs} denote the emitted and observed wavelengths, respectively. As space expands, light waves traverse greater distances, increasing their wavelength and shifting toward the red end of the electromagnetic spectrum.

1.2 Early Universe

The Big Bang, occurring approximately 13.8 billion years ago, marks the origin of space and time and initiated the Universe's expansion from an extremely hot, dense state. In its earliest moments, the Universe existed as a plasma of elementary particles and radiation at temperatures exceeding 10^{12} K. As it expanded and cooled, these particles combined through Big Bang nucleosynthesis during the first few minutes, producing a primordial composition of roughly 75% hydrogen and 25% helium by mass (Cyburt et al. [64], Olive et al. [193], Planck Collaboration et al. [215], Schramm & Turner [241]).

For the next several hundred thousand years, the Universe remained in a hot, ionized state where photons were continuously scattered by free electrons, making it opaque to electromagnetic radiation. During this epoch, matter and radiation were tightly coupled, with radiation pressure preventing the gravitational collapse of density perturbations on scales smaller than the horizon (Peebles [208], Silk [250]).

Approximately 380,000 years after the Big Bang, the epoch of recombination occurred when the Universe had cooled to roughly 3000 K, allowing electrons to combine with nuclei and form neutral hydrogen and helium atoms. This process decoupled matter from radiation, allowing photons to travel freely for the first time and producing the Cosmic Microwave Background (CMB) radiation we observe today (Dicke et al. [76], Penzias & Wilson [211], Planck Collaboration et al. [215], Smoot et al. [252]).

The CMB exhibits subtle temperature fluctuations of order $\Delta T/T \sim 10^{-5}$ that correspond to primordial density perturbations present at the time of last scattering (Planck Collaboration et al. [215]). These nearly gaussian, adiabatic fluctuations, likely generated during cosmic inflation, represent the initial conditions from which all cosmic structure would subsequently grow through gravitational instability.

1.3 Dark Matter Halos and Cosmic Web

The small density perturbations imprinted in the CMB give an early snapshot of the cosmic web that grewthrough gravitational instability, providing the gravitational scaffolding upon which all subsequent structure formation would depend (Blumenthal et al. [30], Peebles [210]). While baryonic matter remained coupled to the radiation field through Thomson scattering until recombination, dark matter overdensity growth started after the Big Bang.

In the linear regime, small density perturbations evolve according to linear perturbation theory, where the density contrast $\delta(t) = (\rho - \bar{\rho})/\bar{\rho}$ grows as:

$$\delta(t) = \delta_i \cdot D(t), \tag{1.5}$$

where $D(t)$ is the linear growth factor and depends on the cosmological parameters (Peebles [209]). In a matter-dominated universe, $D(t) \propto a(t)$, where $a(t)$ is the scale factor, allowing perturbations to grow linearly with cosmic time.

As density contrasts approach $\delta \sim 1$, linear theory breaks down and regions begin to decouple from the Hubble flow. According to the spherical collapse model, a region reaches maximum expansion when $\delta_{linear} \approx 1.06$ and subsequently collapses to form a virialized dark matter halo when $\delta_{linear} \approx 1.69$ (Gunn & Gott [112], Peebles [209]).

Structure formation proceeds hierarchically, with smaller halos forming first and then merging to create progressively larger systems. The Press-Schechter formalism provides a statistical description of the halo mass function:

$$n(M, z)\, dM = \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}}{M^2} \frac{\delta_c}{\sigma(M, z)} \left| \frac{d \ln \sigma(M, z)}{d \ln M} \right| \exp\left(-\frac{\delta_c^2}{2\sigma^2(M, z)} \right) dM, \tag{1.6}$$

where $\sigma(M, z)$ represents the variance of density fluctuations smoothed on mass scale M, and $\delta_c \approx 1.69$ is the critical overdensity for collapse (Press & Schechter [217]). This framework successfully predicts the abundance of halos across a wide range of masses and redshifts.

The spatial distribution of dark matter halos is not random but reflects the underlying dark matter density field. This relationship is quantified through the bias parameter $b(M, z)$, defined as $\delta_h = b(M, z)\delta_m$, where δ_h and δ_m are the halo and matter overdensities, respectively (Mo & White [181], Sheth et al. [247]). More massive halos are more strongly biased ($b > 1$), meaning they preferentially form in regions of higher dark matter density, leading to stronger clustering.

This hierarchical assembly creates the cosmic web composed by a network of filaments, sheets, and voids that characterizes the large-scale structure of the Universe (Bond et al. [33], Springel et al. [256]). Dark matter halos serve as the gravitational wells within this web, capturing baryonic matter and enabling the formation of the first stars and galaxies. The concordance between cold dark matter (CDM) predictions and observed galaxy clustering provides strong evidence for this paradigm and allows precise constraints on cosmological parameters (Percival et al. [212], Tegmark et al. [261]).

1.4 First light

With dark matter halos providing the gravitational scaffolding, the Universe was now poised for the formation of its first luminous objects. Following recombination, the cosmos entered the "Dark Ages"—a period lasting from $z \sim 1100$ to $z \sim 60 - 65$ during which no stars had yet formed to illuminate the Universe (Barkana & Loeb [14], Loeb & Barkana [162]). During this epoch, the Universe was filled primarily with neutral hydrogen and helium gas at temperatures of ∼10-100 K, with structure formation proceeding solely through the gravitational growth of dark matter density perturbations.

As dark matter halos reached sufficient masses ($M_{halo} \sim 10^5 - 10^6\, M_\odot$), they began to efficiently accrete and shock-heat the primordial gas (Abel et al. [1], Tegmark et al. [260]). In these early "minihalos," baryonic matter could cool through molecular hydrogen (H_2) formation, the only available coolant in the metal-free primordial gas. When hydrogen number densities reached $\sim 10^4$ cm^{-3} and temperatures dropped below ~ 200 K, gravitational collapse accelerated, eventually triggering the formation of the first stars—Population III stars—at redshifts as high as $z \sim 65$ (Naoz et al. [185]). These primordial stellar births mark the end of the Dark Ages and herald the beginning of cosmic reionization, with the most observable early star formation likely occurring around $z \sim 20 - 30$ (Bromm et al. [44], Yoshida et al. [289]).

These primordial stars were fundamentally different from modern stellar populations. Formed from metal-free gas, they were typically very massive ($10 - 300\, M_\odot$), extremely hot and luminous, and short-lived (lifetimes $\sim 3 - 10$ Myr) (Heger et al. [123], Schaerer

[238]). Their intense ultraviolet radiation ($h\nu > 13.6$ eV) began to ionize the surrounding hydrogen, creating the first HII regions and marking the end of the Dark Ages (Madau et al. [171], Wise et al. [285]).

The death of these massive Population III stars through energetic supernovae had profound consequences for subsequent star formation. These explosions dispersed the first heavy elements (metals) into the intergalactic medium, fundamentally altering the cooling properties of gas in nearby halos (Gnedin [109], Machacek et al. [169]). Metals enabled more efficient cooling through fine-structure transitions of carbon, oxygen, and silicon, allowing star formation in smaller halos and leading to the transition from Population III to Population II star formation.

As dark matter halos continued to grow hierarchically, reaching masses $M_{halo} \gtrsim 10^8\, M_\odot$ by $z \sim 10$, they could retain gas heated by stellar feedback and support sustained star formation (Dekel & Birnboim [74], Trenti et al. [267]). These systems represent the first true galaxies—gravitationally bound collections of stars, gas, and dark matter. Early galaxy formation followed the hierarchical "bottom-up" paradigm, where small protogalaxies formed first and subsequently merged to build progressively larger systems (Kauffmann et al. [141], White & Rees [281]).

Observational evidence for this early epoch comes from deep surveys with the *Hubble Space Telescope* and *James Webb Space Telescope*, which have detected galaxies at $z > 10$, corresponding to the first few hundred million years after the Big Bang (Naidu et al. [183], Oesch et al. [189]). These early galaxies typically exhibit compact morphologies ($R_e < 1$ kpc), high specific star formation rates, low metallicities ($Z \lesssim 0.1 Z_\odot$), and irregular structures consistent with ongoing mergers and accretion (Bouwens et al. [36], Robertson [229]).

The emergence of the first stars and galaxies initiated the epoch of reionization, a phase transition that would reshape the thermal and ionization state of the intergalactic medium and profoundly influence all subsequent cosmic evolution.

1.5 Cosmic noon and modern Universe

The first stars and galaxies established during the epoch of first light initiated a cascade of processes that would fundamentally transform the Universe. The intense ultraviolet radiation from early stellar populations drove the epoch of reionization ($z \sim 6 - 20$), gradually ionizing the neutral hydrogen in the intergalactic medium and making the Universe transparent to UV photons (Fan et al. [90], Planck Collaboration et al. [215]).

This phase transition marked a critical milestone in cosmic evolution, enabling the formation of more complex galactic systems and establishing the ionized medium that persists today.

As the Universe continued to expand and evolve, it entered what is often termed "cosmic noon" ($z \sim 1 - 3$), an epoch characterized by peak star formation activity and rapid galaxy assembly (Madau & Dickinson [170]). During this period, star formation rates reached their maximum values, with the cosmic star formation rate density peaking at $z \sim 2$ before gradually declining toward the present day (Behroozi et al. [20], Hopkins et al. [126]). This era witnessed the formation of the most massive galaxies and the establishment of fundamental scaling relations that persist in the local Universe.

Galaxy mergers emerged as a dominant driver of cosmic evolution during and after cosmic noon, fundamentally shaping both individual galaxies and the large-scale structure of the Universe (Toomre & Toomre [265]). These interactions are typically classified as major or minor mergers, where larger galaxies accrete smaller companions through a process often termed galactic cannibalism (Ostriker & Hausman [194]).

At the scale of individual galaxies, mergers drive profound morphological and dynamical transformations. Gravitational interactions create distinctive features such as tidal tails and bridges, while compressing gas reservoirs and often triggering intense starbursts that can elevate star formation rates by up to two orders of magnitude (Barnes & Hernquist [16], Mihos & Hernquist [180]). These events frequently result in the formation of elliptical galaxies, as the ordered stellar motions in spiral disks are randomized through violent relaxation (Lynden-Bell [167]). Mergers also promote chemical enrichment through enhanced supernova rates and mixing of stellar populations from different progenitor galaxies.

On cosmological scales, the hierarchical merging of dark matter halos deepens gravitational wells and reinforces the nodes and filaments of the cosmic web. This process influences galaxy demographics and morphologies across different environments, while energetic outflows from merger-driven starbursts and active galactic nuclei can significantly impact the intergalactic medium through feedback processes (Hopkins et al. [127], Springel et al. [256]).

Observational studies reveal that merger rates were substantially higher in the early Universe, with merger fractions declining from $z \sim 3$ to the present (Conselice et al. [58], Le Fèvre et al. [152]). Approximately 5% of nearby galaxies are currently experiencing mergers (Conselice et al. [58], Schmidt et al. [240]), while merger fractions at higher redshifts can exceed 20% (Lotz et al. [165], Man et al. [172]). Major mergers contribute more significantly to stellar mass growth than minor mergers, and over cosmic time,

mergers can more than double the mass of the most massive galaxies (Rodriguez-Gomez et al. [232]).

Throughout cosmic history, galaxy interactions have triggered bursts of star formation and fueled the growth of supermassive black holes, which through energetic feedback processes regulate subsequent star formation and galaxy evolution (Di Matteo et al. [75], Hopkins et al. [126]). The ongoing assembly of larger structures such as galaxy clusters and superclusters continues today, as galaxies and gas accrete along cosmic web filaments (Kravtsov & Borgani [146]).

From the initial quantum fluctuations preserved in the cosmic microwave background, the Universe has evolved into the intricate network of galaxies, clusters, and cosmic web filaments observed today. This remarkable transformation spanning from the first stars at cosmic dawn to the complex galactic ecosystems of the modern Universe represents a fundamental consequence of gravitational instability, dark matter physics, and the complex interplay of astrophysical processes that continue to shape cosmic evolution.

1.6 Galaxy Clustering

Galaxy clustering represents one of the most powerful probes of cosmology and galaxy formation, providing a direct observational window into the connection between luminous galaxies and the underlying dark matter structure that governs cosmic evolution. The spatial distribution of galaxies encodes fundamental information about the initial conditions of the Universe, the physics of structure formation, and the complex interplay between baryonic and dark matter processes throughout cosmic history.

1.6.1 Theoretical Foundations and Observational Methodologies

The theoretical foundation for galaxy clustering studies was established through the pioneering work of Peebles [209], who formalized the two-point correlation function as a statistical tool to quantify the excess probability of finding galaxy pairs compared to random distributions. This framework provided the mathematical basis for connecting observable galaxy clustering patterns to the underlying cosmic web structure. The correlation function $\xi(r)$ quantifies the deviation from randomness in galaxy positions, where positive values indicate clustering and negative values suggest avoidance or voids.

The critical conceptual breakthrough came with the recognition that galaxy clustering could provide direct constraints on the masses of dark matter halos hosting observed galaxies. This insight emerged from the pronounced dependence of halo clustering on

halo mass, as demonstrated by theoretical work including Bardeen et al. [13], Mo & White [181], and subsequent developments. The bias framework, initially developed by Kaiser [139], established the direct link between observable galaxy surveys and cosmological simulations, enabling quantitative comparisons between theory and observation. The galaxy bias parameter b represents the systematic deviation in galaxy clustering compared to the underlying matter distribution, with higher bias values indicating that galaxies preferentially inhabit more massive, and therefore more clustered, dark matter halos. This relationship forms the cornerstone of modern galaxy clustering analysis, as it provides the bridge between observable galaxy properties and the invisible dark matter scaffolding that determines large-scale structure.

Galaxy clustering measurements employ several complementary observational approaches, each with distinct advantages and limitations. Spectroscopic redshift surveys provide the most precise clustering analyses through three-dimensional reconstructions of galaxy positions, enabling detailed studies of the correlation function in real space. Early foundational work by Davis & Huchra [70] and later large-scale efforts such as those described by Colless et al. [53] established the methodological framework that continues to guide modern surveys. Photometric surveys trade precision in individual distance measurements for wider sky coverage and the ability to include significantly fainter galaxies. These surveys rely on angular clustering measurements, which relate sky-projected clustering patterns to real-space correlations through the Limber approximation Limber [159]. This approach has proven particularly valuable for high-redshift studies where spectroscopic observations become increasingly challenging due to observational limitations.

1.6.2 Evolution of High-Redshift Clustering Studies

The systematic study of high-redshift galaxy clustering emerged with deep observations from the Hubble Space Telescope, marking a new era in our understanding of early cosmic structure. Arnouts et al. [9] established the fundamental methodology by measuring clustering evolution from local galaxies to $z \sim 4.5$ using Lyman Break Galaxies (LBGs), revealing the characteristic evolutionary patterns that would define the field. This pioneering work demonstrated that galaxy bias increases systematically with redshift beyond $z \gtrsim 1.5$, providing the first direct observational evidence that high-redshift galaxies preferentially inhabit increasingly massive dark matter halos relative to the cosmic mean.

Building upon these foundations, the studies by Giavalisco et al. [107], Giavalisco & Dickinson [106], and Adelberger et al. [6] extended systematic measurements to intermediate redshifts $z \sim 3-4$, establishing the angular correlation function methodology that would become standard for high-redshift studies. These investigations revealed the fundamental relationship between galaxy luminosity, clustering strength, and host halo mass, providing the empirical foundation for theoretical models of galaxy formation. The discovery that UV-bright LBGs trace dark matter halos with characteristic masses in the range $10^{11} - 10^{12} M_{\odot}$ established a crucial anchor point for understanding early galaxy formation processes.

The field experienced a significant methodological advancement with the introduction of halo occupation distribution (HOD) frameworks to high-redshift clustering analysis, pioneered by Lee et al. [156]. By decomposing the correlation function into one-halo and two-halo terms, this approach enabled direct constraints on the galaxy-halo connection, revealing how central and satellite galaxy populations occupy dark matter halos as functions of mass and redshift. The HOD formalism, further developed through theoretical work by Ma & Fry [168], Peacock & Smith [207], Seljak [243], Kravtsov et al. [147], and Zheng et al. [292], provided the crucial theoretical bridge between observable clustering patterns and predictions from cosmological N-body simulations. This framework enabled researchers to move beyond simple power-law fits to correlation functions and extract detailed information about galaxy occupation statistics within dark matter halos.

Concurrent developments by Ouchi et al. [195] and Overzier et al. [197] systematically extended clustering measurements to higher redshifts, demonstrating the continued increase in galaxy bias with cosmic time and establishing methodological frameworks capable of probing the epoch of reionization. These studies revealed that characteristic halo masses hosting UV-bright galaxies remain relatively constant across cosmic time, while bias increases due to the evolving dark matter clustering amplitude as predicted by hierarchical structure formation models. The work was complemented by theoretical advances in understanding halo bias evolution from Sheth et al. [247] and Tinker et al. [263], which provided the theoretical framework for interpreting observed clustering trends.

As observations pushed toward cosmic dawn, the exploration of galaxy clustering during the epoch of reionization required increasingly innovative approaches to overcome fundamental observational challenges including small sample sizes, cosmic variance effects, and the rarity of bright galaxies at early times. Barone-Nugent et al. [17] achieved the first systematic clustering measurements at $z > 7$ by combining multiple survey fields

with careful systematic control, establishing that galaxy bias continues to increase toward cosmic dawn and that the most luminous early galaxies inhabit the most massive available dark matter halos. This work was extended by comprehensive studies from Harikane et al. [117], who provided systematic measurements across $z \sim 4-7$, establishing the robustness of bias evolution trends and constraining the range of halo masses hosting early galaxies.

The field continued to mature through comparative investigations by Hatfield et al. [121] and Qiu et al. [219], which explored the sensitivity of results to different methodologies and sample selection criteria. These studies demonstrated that while individual measurements may vary due to cosmic variance and systematic effects, the overall trend of increasing bias with redshift remains robust across different analytical approaches. Modern clustering studies also began incorporating more sophisticated HOD modeling approaches, as demonstrated by Ishikawa et al. [131], Harikane et al. [118], and Shuntov et al. [248], extending the framework to higher redshifts and fainter galaxy populations.

A critical methodological challenge emerged as deep surveys began employing heterogeneous observing strategies that could introduce systematic biases in clustering measurements. Dalmasso et al. [66] (see Ch.2) developed novel approaches to address non-uniform survey depth and completeness through sophisticated random catalog generation using artificial source injection and recovery simulations. This methodology, applied to Hubble Legacy Field data, enabled the highest redshift clustering measurement achieved pre-JWST while properly accounting for the systematic effects that arise from mosaic observations with varying exposure depths. The most comprehensive analysis of the pre-JWST era was provided by Harikane et al. [119], who measured clustering across $z \sim 2-7$ using samples exceeding four million galaxy candidates, establishing the systematic nature of bias evolution across cosmic time and demonstrating consistency between clustering-based halo mass estimates and independent constraints from abundance matching techniques and gravitational lensing observations.

1.6.3 Physical Implications for Galaxy Formation

Galaxy clustering studies have revealed fundamental insights into how galaxies populate dark matter halos across cosmic time, providing crucial constraints on galaxy formation efficiency and environmental effects. The systematic increase in galaxy bias with redshift, combined with relatively stable characteristic halo masses for UV-selected samples, demonstrates that early galaxies represent increasingly rare peaks in the primordial density field. This evolution reflects the hierarchical nature of structure formation, where

identical halo masses correspond to progressively more biased environments at earlier cosmic epochs.

Theoretical frameworks including the Press-Schechter formalism (Press & Schechter [217]) and its extensions provide the mathematical foundation for connecting observed clustering patterns to halo mass functions, enabling direct constraints on star formation efficiency and duty cycles. These studies have established that UV-bright galaxies during reionization preferentially inhabit halos with masses corresponding to the transition scale where atomic hydrogen cooling becomes efficient and robust star formation can proceed. The integration of clustering measurements with theoretical models has been further advanced by Trenti et al. [267] and Trenti et al. [268], who developed frameworks connecting galaxy bias evolution to star formation efficiency and demonstrated how magnitude-limited surveys probe increasingly massive objects at higher redshifts due to luminosity distance effects.

Abundance matching techniques, pioneered by Vale & Ostriker [271] and Martínez et al. [174], combine clustering measurements with luminosity functions to constrain the duty cycle representing the fraction of dark matter halos hosting detectable galaxies at any given time. The luminosity function evolution studies by Bouwens et al. [34], Bouwens et al. [36], and Finkelstein et al. [95] have provided essential constraints on the number density evolution of high-redshift galaxies, which when combined with clustering measurements, reveal the efficiency of galaxy formation processes. Comprehensive analyses by Mason et al. [176] and Harikane et al. [119] have demonstrated how the interplay between luminosity function evolution and clustering strength constrains both the duty cycle and the characteristic halo masses hosting star-forming galaxies across cosmic time.

Theoretical predictions from models including those by Wyithe et al. [288] suggested that duty cycles should approach unity at high redshifts as most massive halos begin forming stars efficiently, a prediction that has been largely confirmed by observational studies combining clustering and luminosity function measurements. However, significant scatter in these relationships reflects the complex physics of early star formation, including feedback processes, metal enrichment, and the role of reionization in suppressing star formation in low-mass halos Mason et al. [177]. The work of Harikane et al. [117] and Harikane et al. [119] has been particularly influential in establishing the robustness of these relationships across different galaxy selection methods and survey strategies.

These measurements provide direct observational constraints on star formation efficiency within dark matter halos and reveal the environmental dependence of early galaxy formation processes. The evolution of duty cycles from moderate values at intermediate redshifts to near-unity values during reionization demonstrates the transition from relatively inefficient to highly efficient star formation in the most massive early halos. This

evolution, supported by both clustering and luminosity function studies, provides crucial insights into the physical processes governing the transformation of primordial gas into the first generations of stars and galaxies that illuminated the early Universe.

1.7 Observational Facilities into the Early Universe

This section briefly introduces the two space telescopes whose observation campaigns formed the basis of my doctoral research.

1.7.1 The Hubble Space Telescope Legacy

1.7.1.1 Mission

Named after Edwin Hubble, the Hubble Space Telescope (HST) (Smith et al. [251]) is one of NASA's Great Observatories. Launched in 1990, HST is a large, space-based observatory designed to capture high-resolution images and collect data across ultraviolet, visible, and near-infrared wavelengths. Its primary scientific goals include investigating the structure and composition of the Universe, and studying the formation and evolution of galaxies, stars, and planets.

The need for a space telescope arose from the limitations of ground-based observatories. Earth's atmosphere distorts and absorbs incoming light, especially in the ultraviolet and infrared, which hinders clear and detailed observations. By operating above the atmosphere, HST eliminates these distortions, enabling sharper and more accurate imaging. Furthermore, HST is not subject to the day-night cycle or weather interruptions, allowing for continuous and consistent monitoring of astronomical phenomena.

These capabilities have been crucial for advancing our understanding of fundamental questions, such as the Universe's expansion rate, the nature of black holes, the formation of stars and galaxies, and the search for life on other planets.

1.7.1.2 Instrumentation

Advanced Camera for Surveys (ACS): Provides wide-field, high-sensitivity visible-light imaging, enabling deep surveys of the Universe, mapping dark matter, detecting distant objects, and studying galaxy cluster evolution (Stark [257]).

Cosmic Origins Spectrograph (COS): Specializes in ultraviolet spectroscopy, ideal for examining stars, quasars, and the intergalactic medium, and probing galaxy and planet formation (Hirschauer [125]).

Near Infrared Camera and Multi-Object Spectrometer (NICMOS): Sensitive to infrared light, revealing details about distant galaxies, stellar nurseries, and planetary systems obscured in other wavelengths (Viana & et al. [275]).

Space Telescope Imaging Spectrograph (STIS): Combines imaging and spectroscopy across ultraviolet to near-infrared, used for studying black holes, massive stars, the intergalactic medium, and exoplanet atmospheres (Rickman & Brown [221]).

Wide Field Camera 3 (WFC3): Extends HST's reach across ultraviolet, visible, and infrared, and serves as the primary instrument for high-resolution imaging of both nearby and distant galaxies (Dressel [79]).

1.7.1.3 Legacy

Measuring the Expansion Rate of the Universe: By observing distant galaxies and using Type Ia supernovae as standard candles, HST has refined the value of the Hubble constant H_0, which quantifies the Universe's expansion rate (Freedman et al. [101], Planck Collaboration et al. [215], Riess et al. [225]).

Deep Field Observations: HST's deep field campaigns have provided unprecedented views into the early Universe, revealing thousands of galaxies at various evolutionary stages and offering key insights into galaxy formation and evolution (Ferguson et al. [91], Giavalisco et al. [108], Williams et al. [282]).

Probing Dark Matter and Dark Energy: Through gravitational lensing and studies of the Universe's expansion, HST has mapped dark matter distributions and contributed to our understanding of dark energy's effects (Massey et al. [178], Riess et al. [226], Schmidt et al. [239]).

Unveiling Early Galaxy Evolution: HST's sharp imaging has enabled astronomers to trace galaxy evolution, directly observe mergers, and identify primeval galaxies, providing visual evidence of the dynamic processes shaping the cosmic web (Abraham et al. [2], Conselice [55], Conselice et al. [58], Duncan et al. [80], Hammer et al. [116], Lotz et al. [164]).

Pushing the Frontier of Early Universe Studies: HST has allowed astronomers to observe the Universe as early as 500 million years after the Big Bang (Bhatawdekar et al. [27], Oesch et al. [190]). Recent results suggest that the formation of the first stars and

galaxies occurred even earlier than previously thought, highlighting the need for next-generation telescopes to probe the Universe's earliest epochs (Boylan-Kolchin & Weisz [41], Hashimoto et al. [120], Mason et al. [177]).

HST's legacy is defined by its transformative impact on nearly every field of astrophysics, from refining cosmological parameters to revealing the structure and history of galaxies across cosmic time.

1.7.2 The James Webb Space Telescope Era

1.7.2.1 Mission

The James Webb Space Telescope (JWST) (Gardner et al. [103]), named after NASA's second administrator James E. Webb, is the most advanced infrared space observatory to date. Launched in 2021 as the successor to HST, JWST is designed to study the Universe at near- and mid-infrared wavelengths, enabling astronomers to observe older, more distant, and fainter objects than ever before. Its primary scientific goals include probing the first light after the Big Bang, tracing the formation and evolution of galaxies, stars, and planetary systems, and exploring the potential for life on exoplanets.

Operating primarily in the infrared (0.6–28 μm), JWST can penetrate cosmic dust clouds that obscure visible light, revealing the birthplaces of stars and planets and allowing the study of the earliest galaxies and stars. Its location at the second Lagrange point (L2), 1.5 million kilometers from Earth, provides a stable, cold environment, minimizing thermal interference and maximizing sensitivity for faint infrared signals.

1.7.2.2 Instrumentation

Near Infrared Camera (NIRCam): JWST's primary imager, covering $\lambda = 0.6 - 5\ \mu m$, equipped with coronagraphs to block starlight and reveal faint objects such as exoplanets. NIRCam is essential for detecting distant galaxies, observing star formation, and studying exoplanetary systems (Rieke et al. [222, 224]).

Near Infrared Spectrograph (NIRSpec): Operates in the $\lambda = 0.6 - 5.3\ \mu m$ range, capable of observing up to 100 objects simultaneously using a microshutter array. NIRSpec provides multi-object, integral field, and high-resolution spectroscopy for detailed studies of galaxies, star-forming regions, and exoplanet atmospheres (Böker et al. [32], Jakobsen et al. [133], Rigby et al. [228]).

Mid-Infrared Instrument (MIRI): Functions as both a camera and spectrograph in the $\lambda = 5 - 28\ \mu m$ range. MIRI enables detailed imaging and spectroscopy of distant galaxies, star and planet formation regions, and cold objects like comets and Kuiper Belt objects (Rieke et al. [223], Wright et al. [286, 287]).

Fine Guidance Sensor/Near InfraRed Imager and Slitless Spectrograph (FGS/NIRISS): The FGS provides precise pointing and stabilization, while NIRISS supports exoplanet spectroscopy, galaxy structure studies, and first light investigations with modes for wide-field imaging and slitless spectroscopy in the $\lambda = 0.6 - 5\ \mu m$ range (Albert et al. [8], Doyon et al. [78], Willott et al. [284]).

1.7.2.3 First Results

JWST's early observations have already had a significant impact:

Measuring the Universe's Expansion Rate: By resolving Cepheid variables in crowded fields, JWST has confirmed Hubble's measurements of the Hubble constant, suggesting that systematic errors are not the main source of the "Hubble tension." This persistent discrepancy may point to new physics, such as exotic dark energy, modified gravity, or unknown particles (Freedman [99], Freedman & Madore [100], Riess et al. [227]).

Galaxy Evolution: JWST's infrared imaging and spectroscopy, particularly with NIRCam, have enabled the detection and characterization of some of the earliest galaxies in the Universe. These observations allow astronomers to determine the ages and properties of stars in galaxies less than a billion years after the Big Bang, offering unprecedented insight into early galaxy formation and evolution (Boyett et al. [38], Carniani et al. [50], Naidu et al. [183]).

Probing the First Light and Reionization: JWST is uniquely suited to study the epoch of reionization and the emergence of the first stars and galaxies, thanks to its sensitivity to redshifted infrared light from the early Universe (Bouwens et al. [37], Finkelstein et al. [96], Robertson [229], Tacchella et al. [259]).

JWST's capabilities mark a new era in observational cosmology, opening the infrared window to the earliest epochs of cosmic history and providing transformative insights into the processes that shaped the modern Universe (Naidu et al. [184]).

1.8 Thesis structure

This thesis presents an investigation of galaxy formation during cosmic dawn through multiple interconnected studies. The first major component focuses on galaxy clustering analysis, beginning with foundational HST observations (Ch.2) that establish the methodological framework for measuring clustering in heterogeneous surveys. This approach is then extended to JWST data (Ch.3), achieving the first galaxy clustering measurements at $z \sim 11$ during the first 500 million years after the Big Bang. In Ch.4, sophisticated halo occupation distribution modeling applied to JWST clustering measurements reveals how galaxy–halo relationships evolved across cosmic time, with characteristic halo masses changing by more than an order of magnitude during the first billion years.

The second component investigates galaxy assembly through morphological analysis, with Ch.5 examining merger rates extending to $z \sim 9$ and providing independent constraints on assembly mechanisms during reionization. Ch.6 shifts to a distinct focus on instrumental calibration and spectroscopic analysis, establishing critical protocols for inter-instrument comparison between JWST's NIRISS and NIRSpec modes and quantifying systematic effects on emission line measurements. To conclude, Ch.7 provides an overview of the thesis, highlighting the main results and implications for each topic studied.

This work establishes the first robust observational framework for galaxy clustering analysis at $z > 8$, extending measurements to $z \sim 11$ and providing fundamental constraints on early galaxy formation models and reionization scenarios. By validating theoretical predictions against empirical observations from $z \sim 5 - 6$ through cosmic dawn, our results directly inform models of dark matter halo occupation, early galaxy assembly, and environmental effects during the first billion years. The methodological advances presented here enable future studies with larger JWST datasets, while our empirical constraints on the galaxy–halo connection provide crucial benchmarks for next-generation cosmological simulations.

Chapter 2

Galaxy clustering measurements out to redshift ∼8 from Hubble Legacy Fields

This chapter is a reproduction of Dalmasso et al. [66]. The author led both the data analysis and writing with the collaboration of the co-authors. This reproduction has been reformatted to fit the style of the thesis, changes include:

- **Figures and tables have been formatted to display correctly on the thesis**
- **Some bibliographic records have been updated**

2.1 Abstract

In this chapter we present a novel approach for measuring the two-point correlation function of galaxies in narrow pencil beam surveys with varying depths. Our methodology is utilized to expand high-redshift galaxy clustering investigations up to $z \sim 8$ by analyzing a comprehensive sample consisting of $N_g = 160$ Lyman break galaxy candidates obtained through optical and near-infrared photometric data within the CANDELS GOODS datasets from the Hubble Space Telescope (HST) Legacy Fields. For bright sources with $M_{UV} < -19.8$, we determine a galaxy bias of $b = 9.33 \pm 4.90$ at $\overline{z} = 7.7$ and a correlation length of $r_0 = 10.74 \pm 7.06\ h^{-1}Mpc$. We obtain similar results for the XDF, with a galaxy bias measurement of $b = 8.26 \pm 3.41$ at the same redshift for a slightly fainter sample with a median luminosity of $M_{UV} = -18.4$. By comparing with

dark-matter halo bias and employing abundance matching, we deduce a characteristic halo mass of $M_h \sim 10^{11.5} M_\odot$ and a duty cycle close to unity. To validate our approach for variable-depth datasets, we replicate the analysis in a region with near-uniform depth using a standard two-point correlation function estimator, yielding consistent outcomes. Our study not only provides a valuable tool for future utilization in James Webb Space Telescope (JWST) datasets but also suggests that the clustering of early galaxies continues to increase with redshift beyond $z \gtrsim 8$, potentially contributing to the existence of protocluster structures observed in early JWST imaging and spectroscopic surveys at $z \gtrsim 8$.

2.2 Introduction

Studying galaxy clustering, i.e. quantifying the non-uniform angular distribution of galaxies in the sky, is a fundamental tool to investigate how the structures we observe today formed and evolved during the course of the history of the Universe (e.g.,Mo & White [181]). In particular, clustering measurements at high redshift offer insight to link the rapid build-up of the galaxy luminosity function (LF) to the assembly of dark-matter halos (e.g., Cooray [60, 61], Lee et al. [156], Vale & Ostriker [272]).

A complementary approach to investigate this is through abundance matching (Vale & Ostriker [271]), this technique involves comparing the number of galaxies with luminosity exceeding a specified threshold L_g to the amount of halos with masses surpassing a given threshold M_h under two assumptions: each halo or sub-halo is hosting an individual galaxy and the luminosity and the halo mass functions are monotonically related (e.g.,Martínez et al. [174], Neyrinck et al. [187]). Through the combination of clustering measurements and abundance matching analysis, the galaxy duty cycle (ϵ_{DC}) can be constrained (Martínez et al. [174]). Also referred as occupation fraction this parameter is defined as the fraction of UV-bright Lyman break galaxies (LBGs) that reside within a dark matter halo, and it is an important quantity to constrain average star formation efficiency within dark matter halos.

Over the past two decades, clustering measurements have been performed out to $z \lesssim 7$. Using data from Hubble Deep Field North, Arnouts et al. [9] quantified the redshift evolution of clustering between redshift $0 \leq z \leq 4.5$ for photometrically selected LBGs. The angular correlation function (ACF) derived showed a clear decrease in amplitude from $z = 0$ up to a redshift $z \sim 1 - 1.5$ followed by a slow increase in amplitude at higher redshift, with the same trend found in the correlation length. These two behaviours translate into an increasing galaxy bias $b(z)$ at $z \gtrsim 1.5$ showing that a typical minimum halo mass of $M_h \gtrsim 10^{12} h^{-1} M_\odot$ is needed for galaxies with $b(z \simeq 4) \simeq 3$. Galaxy bias, denoted as b, represents the systematic deviation or clustering of galaxies compared to

the underlying matter distribution in the universe. It serves as an indicator of the non-uniform spatial distribution of galaxies, influenced by dark matter distribution and the large-scale structure of the universe. An increasing trend in galaxy bias implies stronger galaxy clustering and deviation from the matter distribution in larger cosmic structures. As we move to higher redshifts, it has been deduced that the minimum dark matter halo mass must fall within the range of $M_h \sim 10^{11.0-11.5} M_\odot$ for galaxies situated at a redshift of approximately $z \simeq 3 - 3.5$ (Adelberger et al. [6], Giavalisco & Dickinson [106]). The increase of $b(z)$ with z continues at $z \geq 3.5$ and the characteristic halo mass remains close to $M_h \sim 5 \cdot 10^{11} M_\odot$ at $z \simeq 3.5 - 6$ (Ouchi et al. [195], Overzier et al. [197]).

Using LBGs candidates at $3.5 \leq z \leq 5.5$, Lee et al. [156] investigated a more realistic model for the ACF with a function composed of two terms, a correlation between two halos on relatively large separations ($w_{2h}(\theta)$) and a one-halo term associated to presence of multiple galaxies in a single halo ($w_{1h}(\theta)$). The latter dominates the correlation signal at very small separations (comparable to the typical virial radius of a halo with $M_h \gtrsim 10^{12}$ $M_\odot$).

In their work, Wyithe et al. [288] derived a theoretical estimate of $\epsilon_{DC}(z) \simeq 0.10 - 0.15$ for $z \gtrsim 6$ using a combination of clustering and abundance matching analyses. This approach aimed to reconcile the observed discrepancies in total stellar mass and associated star formation rates.

This theoretical prediction has been subsequently tested by Barone-Nugent et al. [17] with new galaxy clustering measurements out to $z = 7.2$ using LBGs from the GOOD-S/CANDELS and XDF Hubble legacy fields observations (Bouwens et al. [34]). Their analysis of the evolution of galaxy bias with redshift shows an increasing trend, reaching $b = 9.7^{+2.0}_{-2.5}$ for their sample of UV bright LBGs at redshift $z = 7.2$. This translate to a minimum dark matter halo mass of $M_h \gtrsim 10^{11.0} M_\odot$.

Using a comparable sample of LBGs derived from the Hubble Legacy deep imaging, Harikane et al. [117] conducted a study on the ACF up to redshift $z = 6.8$, focusing on UV bright candidates ($m_{UV} < 27.6$). Through their measurements, they deduced a range of acceptable halo masses for galaxies residing within the redshift interval of $z \sim 6 - 7$ to be approximately $M_h \sim (1 - 20) \cdot 10^{11} M_\odot$, in accordance with what had been found previously in the high redshift regime ($z \sim 5.8$) (e.g.,Barone-Nugent et al. [17], Hamana et al. [115], Hildebrandt et al. [124], Lee et al. [156]).

In contrast to Harikane et al. [117]'s redshift $z = 5.8$ findings of a galaxy bias of $b = 4.7^{+1.0}_{-1.3}$, both Hatfield et al. [121] and Qiu et al. [219] report slightly higher galaxy bias values for galaxies with similar redshift and absolute magnitude when estimating halo mass, measuring $b = 6.0 \pm 1.1$ and $b = 6.1^{+1.6}_{-2.0}$ respectively. These differences may be due to computation method variations, but clustering trends with luminosity remain consistent. Both studies find significantly higher typical host halo mass and galaxy bias in their samples compared to lower luminosity ones.

In their subsequent study (Harikane et al. [119]), the authors focus on measuring the rest-UV luminosity functions and angular correlation functions across a wide range of redshifts, specifically $z \sim 2-7$. Their findings indicate an overall increasing trend in galaxy bias, compared to their previous work (Harikane et al. [117]). As a result of this increasing galaxy bias, there is a corresponding rise in the required host halo mass threshold, notably shifting for luminous UV galaxies towards values approximating $M_h \gtrsim 10^{11.37} M_\odot$ around redshift $z \sim 6$, with this shift becoming more pronounced for even brighter galaxy candidates.

Further progress at higher redshift was precluded both by the limited sample sizes at $z \sim 8$ and by the heterogeneous nature of surveys carried out in the last ten years, which often employed a "wedding-cake" depth strategy (i.e. multiple layers of exposure times). This strategy is very efficient for characterising the luminosity function across a wide dynamic range (Bouwens et al. [34]), but it makes the measurement of the two-point correlation function (which can be used to contrain galaxy clustering) more challenging. The two-point correlation function encodes the excess probability of finding two objects, in our case galaxies, separated by a given distance. The excess is evaluated relative to a uniform distribution of sources (Landy & Szalay [150]).

With a heterogeneous survey including regions at different depths, special care must be placed to avoid introduction of systematic uncertainty and artificial clustering signal simply because some areas have an excess of galaxies due to deeper observations. In this chapter we investigate the impact of non-homogeneous survey depth on the two point correlation function measurement and propose a method to conduct an accurate analysis by careful generation of the reference random sample for use in the two point correlation function estimator. Rather than simply generating a uniform random distribution over the geometry of the survey and taking into account a binary mask to model geometry and presence of foreground bright sources that mask faint high-redshift objects, we generate the reference random sample through a Monte Carlo simulation that places artificial sources with realistic spectral energy distributions and recovers them through the full photometric pipeline.

The primary objective of this study is to showcase a novel approach through the analysis of data obtained from the Hubble Legacy Fields (Bouwens et al. [36]). This study has two main goals: firstly, to validate our methodology by comparing it with traditional two-point correlation techniques using substantial lower redshift datasets. Secondly, to extend our measurements to high redshifts ($z = 8$) by combining surveys with varying depths, given the challenge of obtaining large high-redshift samples independently.

Our novel method should be ideal for applications to upcoming catalogs from the increasing number of JWST photometric surveys that are characterising the galaxy population at $z > 8$ (Hainline et al. [114], Paris et al. [200]).

This chapter is organized as follows. In Sec.2.3 we describe the data selection for the

LBGs candidates. Variables estimation and the theory behind the ACF is presented in Sec.2.4. In Sec.2.5 we address the different methods to generate the points catalog needed to proceed in the measurements. Our findings and numerical results are discussed in Sec.2.6 . In Sec.2.7 we presents our summary and conclusions.
In the research presented in this chapter we assume, when relevant, the cosmological parameters determined by the Planck Collaboration Ade et al. [4]: $(\Omega_M, \Omega_\lambda, h, \sigma_8) = (0.315, 0.685, 0.673, 0.828)$. Magnitudes are in the AB system Oke & Gunn [191].

2.3 Data Selection

This study utilises the most recent data released by the HST, specifically the comprehensive collection known as the Hubble Legacy Fields[1]. In our exploration of this extensively researched field, we have opted to make use of publicly available catalogs, such as the VizieR Online Data Catalog[2] (Bouwens et al. [36]).
Our ultimate objective is to assemble a sizable galaxy dataset for conducting clustering analysis. To assess the influence of varying depth on source recovery and detect any potential systematic biases, we generated χ^2 images by utilizing data from the Y_{105}, J_{125}, JH_{140}, and H_{160} filters. These χ^2 images will be fundamental in the development of random point catalogs, as further detailed in the subsequent sections.
We began by creating a parent sample sourced from the VizieR Online Data Catalog, involving the selection of galaxies within the redshift range $z \in [4.5, 8.5]$. We subsequently divided this primary sample into four redshift bins, each with a width of $\Delta z = 1.0$, with the mean redshift within each bin being $\overline{z}$=5.1, 6.1, 6.8, and 7.7, respectively. To refine the analysis and identify sources in contiguous regions, we obtained the latest versions of the science and weight images from the Hubble Legacy Fields Project (HST Legacy field data, Bouwens et al. [35], Whitaker et al. [280]). Specifically, Version 2.0 was used for the GOODS-S field, while Version 2.5 was employed for the GOODS-N field, with both images having a scale of 60 milliarcseconds per pixel. Subsequently, we converted the weight images into root mean square (RMS) images.
Considering that the GOODS-S field contains the XDF/HUDF patch, which involves much deeper observations that may introduce systematic uncertainties in the estimation of the correlation function beyond our ability to correct for them, we took an additional step to separate the XDF/HUDF region out. Furthermore, we excluded the HUDFP1 and HUDFP2 parallel ultra-deep regions within the GOODS-S field due to the absence of the required B_{435} filter for $z \sim 5$ selection in the HUDFP1 region, as well as the

[1] https://archive.stsci.edu/prepds/hlf/
[2] https://ui.adsabs.harvard.edu/link_gateway/2021yCat..51620047B/doi:10.26093/cds/vizier.51620047

Number of sources					
This work	Bouwens-2021	$\overline{z} = 5.1$	$\overline{z} = 6.1$	$\overline{z} = 6.8$	$\overline{z} = 7.7$
XDF	HUDF/XDF	187 (**187**)	146 (**146**)	124 (**124**)	44 (**44**)
GOODS-S	ERS	223 (**108**)	96 (**51**)	69 (**36**)	15 (**13**)
	C. GOODS-S	709 (**297**)	291 (**123**)	129 (**50**)	51 (**30**)
GOODS-N	C. GOODS-N	1092 (**507**)	372 (**192**)	259 (**150**)	108(**73**)
TOT		2211 (**1099**)	905 (**512**)	581 (**360**)	218 (**160**)

TABLE 2.1: Summary of the LBG sources used in this work for clustering measurements in three main fields (XDF, GOODS-S and GOODS-N). The sources are extracted from Bouwens et al. [36] as discussed in Sec.2.3, and the table includes for each redshift bin the parent catalog source number and in **bold** between parentheses the objects selected for analysis.

presence of multiple bright stars in the HUDFP2 region. Additionally, the selection area in the GOODS-S field was trimmed to the common overlap area in the F125W, F140W, and F160W filters to ensure a uniform selection of sources. The XDF area is not included in the RMS map of GOODS-S see right panels of Fig.2.1 therefore it will be analysed separately.

The next step involved filtering the catalog by applying an absolute magnitude cut of $M_{UV} < -19.8$ in the two GOODS regions in order to obtain a magnitude-limited (uniform) sub-sample for conducting a consistent analysis across the entire survey area and redshift range. This absolute magnitude limit was chosen to ensure source detection above the nominal 5σ depth limit for all fields and redshifts. We decided to not apply any magnitude (absolute or apparent) cut to the XDF field since the whole XDF region is internally uniform. Additionally, applying the same absolute magnitude cut as the GOODS region would have resulted in the number of candidates that are too few to conduct a meaningful analysis. Therefore, our investigation encompasses all XDF sources, irrespective of their brightness. For reference, the average absolute magnitudes of the samples within the XDF area range from $\overline{M}_{UV} = -18.6$ to $\overline{M}_{UV} = -18.1$. Details of the data selection for each redshift bin considered in this study are summarized in Tab.2.1.

For a comprehensive description of the dataset (including the determination of galaxy luminosity functions), we refer the reader to Bouwens et al. [36].

2.4 Clustering estimation

First, we derive the observed ACF indicated as $w_{obs}(\theta)$, that measures the excess probability of finding two objects, galaxies in our case, at an angular separation θ. We use

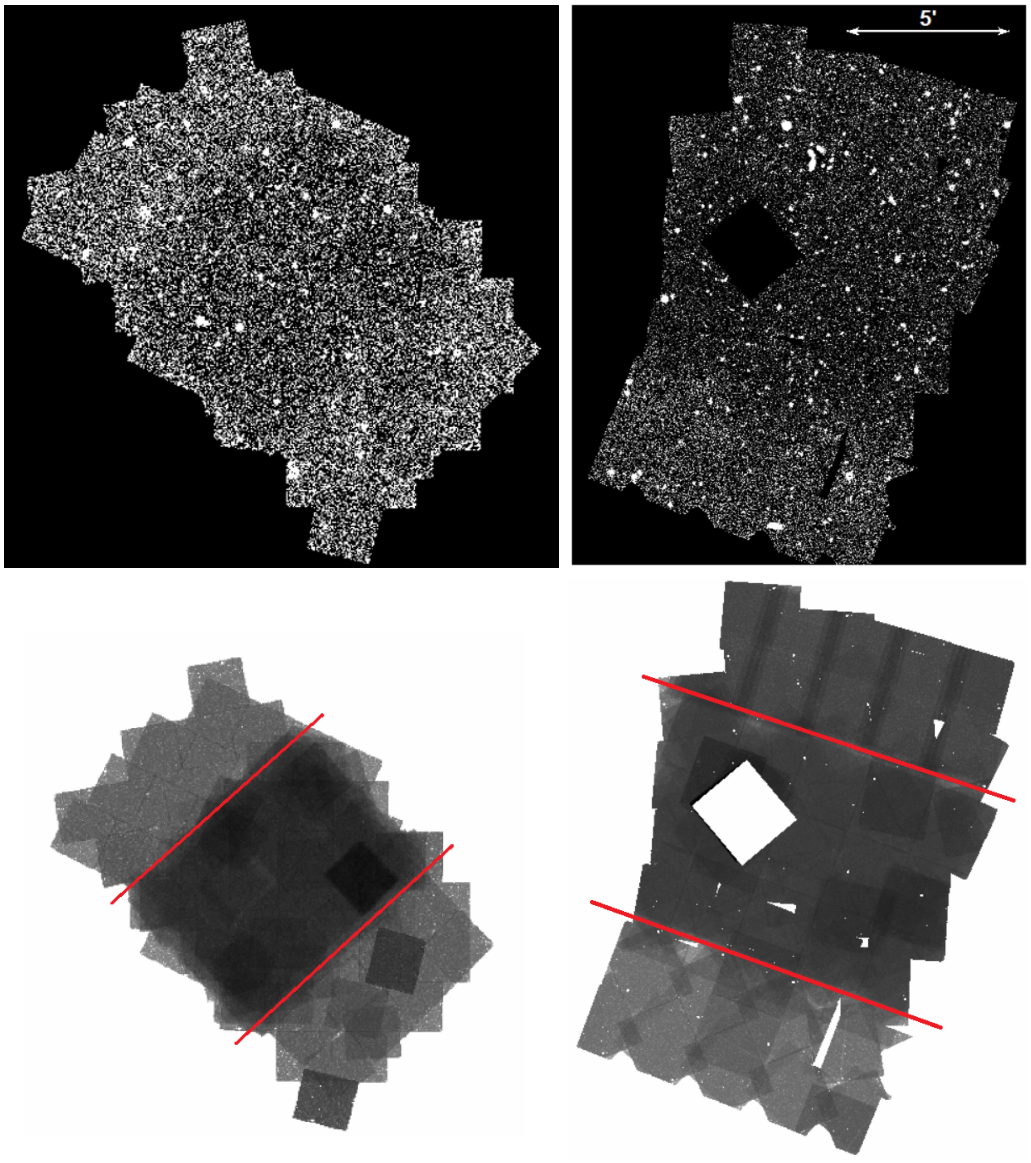

FIGURE 2.1: **Left**: GOODS-North maps. **Right**: GOODS-South maps. The top row shows the science images while the bottom has RMS maps in logarithmic scale and divided in sections by red lines: **Top**, **Central** and **Bottom** to show the differences in depth due a wedding-cake depth strategy of the survey (see Grogin et al. [110]).

the ACF estimator defined by Landy & Szalay [150]:

$$w_{obs}(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}, \tag{2.1}$$

where $DD(\theta)$ is the number of pairs of observed galaxies with angular separations $(\theta \pm \delta\theta)$. $RR(\theta)$ is the analogous quantity determined from a random catalog with the same geometry and selection properties as the observed catalog. $DR(\theta)$ is the number of random-galaxy cross pairs.

For the functional form of ACF, we assume a power law parameterization $w(\theta) = A_w\theta^{-\beta}$, where $w(\theta)$ is our estimator. The finite size of this survey leads to an underestimation

of the measured quantities, which can be rectified by incorporating a constant known as the integral constraint (IC; see Peacock & Nicholson [205]):

$$w_{\rm true}(\theta) = w_{\rm obs}(\theta) + IC, \tag{2.2}$$

estimated as:

$$\begin{aligned} IC &= \frac{1}{\Omega^2}\int_1\int_2 w_{\rm true}(\theta)d\Omega_1 d\Omega_2 \\ &= \frac{\sum_i RR(\theta_i)w_{\rm true}(\theta)}{\sum_i RR(\theta_i)} = \frac{\sum_i RR(\theta_i)A_w\theta_i^{-\beta}}{\sum_i RR(\theta_i)}, \end{aligned} \tag{2.3}$$

where $w_{\rm true}(\theta)$ is the intrinsic ACF and $w_{\rm obs}(\theta)$ is the measurement within the survey area. For consistency with previous work (e.g.,Barone-Nugent et al. [17], Lee et al. [156], Overzier et al. [197]), we fix $\beta = 0.6$ and use linear binning with 12.5" width with the maximum separation at $\theta = 250$".

When the correlation slope β is fixed, the quantity IC/A_w only depends on the size and the shape of the survey area and can be determined from Eq.2.3. The dimensionless parameter A_w is the only variable that needs to be fit as:

$$w_{\rm model}(\theta) = A_w(\theta^{-\beta} - IC/A_w), \tag{2.4}$$

by maximizing the likelihood expressed as:

$$\mathcal{L} = \prod_{i\in\text{fields}} \frac{1}{\sigma_{obs}(\theta)_i\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{w_{obs}(\theta)_i - w_m(\theta)_i}{\sigma_{obs}(\theta)_i}\right)^2}, \tag{2.5}$$

where $w_{obs}(\theta)_i$ and $w_m(\theta)_i$ represent the observed and modelled ACF measurements while $\sigma_{obs}(\theta)_i$ the uncertainties in field i. Errors in the ACF, for both fields, are estimated using bootstrap resampling (Ling et al. [161]).

To estimate the galaxy bias, we approximate the real-space correlation function with a power law $\xi(r) = (r/r_0)^{-\gamma}$ with $\beta \equiv \gamma - 1$, where the coefficients are related to the dimensionless A_w coefficient by the Limber transform (Adelberger et al. [6]):

$$A_w = \frac{r_0^\gamma B[1/2, 1/2(\gamma-1)]\int_0^\infty dz N(z)^2 f^{1-\gamma} g(z)^{-1}}{[\int_0^\infty dz N(z)]^2}. \tag{2.6}$$

Here f is the transverse comoving distance, i.e. $f \equiv (1+z)D_A(\theta)$ where D_A the angular diameter distance. $N(z)$ is the redshift distribution of the dropouts that takes into account efficiency and completeness of source identification, $B(t_1, t_2) = \Gamma(t_1)\Gamma(t_2)/\Gamma(t_1 + t_2)$ is the beta function and $g(z) \equiv c/H(z)$ is the comoving distance.

After finding A_w by maximizing the likelihood Eq.2.5 we are able to invert Eq.2.6 in order to obtain the associated correlation length r_0. From the real-space correlation

function $\xi(r)$ we define the galaxy bias b through the ratio between the galaxy variance $\sigma_{8,g}$ and the linear matter fluctuation $\sigma_8(z)$ both in a sphere of comoving $8h^{-1}cMpc$ radius:

$$b = \frac{\sigma_{8,g}}{\sigma_8(z)}, \tag{2.7}$$

assuming $\sigma_8(0) = 0.828$ from the cosmology model and the galaxy variance expressed as:

$$\sigma^2_{8,g} = \frac{72(r_0/8h^{-1}cMpc)^\gamma}{(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma}. \tag{2.8}$$

2.5 Clustering analysis methods

To conduct measurements of galaxy clustering, it is necessary to generate a catalog of random points for the calculation of the observed ACF estimator w_{obs}, as defined in Eq.2.1. Previous studies adopted a simpler approach when generating the random points catalogs, mainly because they were based on surveys with relatively uniform depth.
In such cases, the primary concern is modeling the survey geometry and accounting for obstructions caused by bright foreground sources. However, the GOODS-CANDELS survey exhibits non-uniform depth, as shown in Fig.2.1. Consequently, in order to evaluate the pair count $RR(\theta)$ for the random points, we investigate the impact of the random catalog generation on the ACF measurement. This investigation is motivated by the observations presented in Fig.2.1, where the survey fields consist of sub-fields with varying individual depths. Moreover, distinct patterns are discernible in small regions, characterized by non-uniform depths (e.g., refer to the bottom right image illustrating the RMS map of GOODS-South in its bottom and top regions). All these factors contribute to the heterogeneity of the survey.
Since the survey's completeness depends on the depth of observations, there may be an excess or lack of sources in the observed catalog at specific locations and preferred angular separations associated with the patterns in the RMS maps. In this section, we introduce two approaches for clustering analysis in the presence of non-uniform-depth data and evaluate their respective strengths. Our analysis focuses on candidate galaxies selected through photometry in the CANDELS data over GOODS at an average redshift of $z = 5.1$ (Leethochawalit et al. [157]), in both cases the random points catalog will be of dimension $N_r = 20N_d$ where N_d is the number of data in the analysis at a certain redshift value.
Specifically, we explore two different aspects that are fundamental for galaxy clustering measurements, the first one being the generation of the random points catalog. The second one is robustness of the analysis with respect to the use of different regions of the survey.

2.5.1 Random point catalog generation

2.5.1.1 Uniform random points

This represent the standard approach of generating random point by sampling a uniform distribution throughout the area covered by the survey, in our case the GOODS field. To take into account the survey geometry (and presence of bright foreground sources and/or regions with missing data), the uniform random catalog is filtered through a binary map (reject or accept) constructed considering all pixels in the image that have a well-behaved RMS map value and are not part of the segmentation maps of (bright) sources. In practice, for faint and relatively rare galaxies at high redshift, one can use the full segmentation map without noticeable difference (Barone-Nugent et al. [17]).

2.5.1.2 Random points from source recovery map

This second method, which is introduced in this chapter, aims to incorporate the varying depth of observations, including subtle features arising from mosaicing and dithering, as observed in the RMS map depicted in Fig. 2.1. In order to improve the fidelity of our source detection and completeness modeling in the random catalog, we qualitatively account for these data characteristics by implementing an artificial source injection and recovery process within the region covered by the ACF observations. The procedure utilizes the publicly available source injection and recovery code, `GLACiAR2` (Leethochawalit et al. [157]), and employs a hit-and-miss Monte Carlo technique, assuming a model luminosity function for the LBG population. This code is typically employed to compute completeness in determining the luminosity function and, as such, realistically simulates all pertinent factors that influence the likelihood of identifying galaxies in observations, including non-uniform depth and photometric scatter in the Lyman break and/or photometric redshift selection algorithm.

For the GOODS field, we injected a total of $N_g = 500$ galaxies within each M_{UV} and redshift bin. These galaxies possess Sérsic light profiles with $n = 1$, random inclinations and ellipticities, and a physical size of $r = 1.25$ kpc at $z = 5$. At other redshifts, the size of the injected galaxies was scaled as $1/(1 + z)$. They exhibit a power-law intrinsic spectral energy distribution (SED) with a Lyman break, where the power-law UV slope is randomly drawn from a Gaussian distribution with a mean of $\beta_{\rm UV} = -2.2$ and a standard deviation of $\sigma = 0.4$. The magnitudes of the galaxies in each photometric band were determined based on their input spectral energy distributions. These artificial sources were then randomly injected into the science image using a 2D uniform distribution of positions. The process of source extraction followed the same steps utilized for the real data sample. The magnitudes, redshifts, and positions of the recovered galaxies

were recorded and employed for subsequent analysis.
Since the catalog of recovered artificial sources was generated assuming a flat input luminosity function $\Phi(M)$, we performed a Monte Carlo hit-and-miss procedure to select the final sample of random points for clustering analysis. The probability function for selection was defined as:

$$p(M) = \frac{\Phi(M)}{\Phi(M_{\rm lim})}, \tag{2.9}$$

where $\Phi^* = 10^{-2.81}$, $M^* = -20.68$, and $\alpha = -1.58$ were used for the Schechter luminosity function (Leethochawalit et al. [157]).

2.5.2 Spatial dependence of the clustering signal and random points methods comparison

Surveys covering a wide area may show variations in depth, resulting in a layered structure resembling a 'wedding cake', as illustrated in Fig.2.1 for the CANDELS survey. Specifically, the GOODS fields consist of three distinct regions, each characterised by an approximately constant exposure time (i.e., map RMS values).
To study the impact of non-uniform exposure, we perform clustering analysis using both methods to generate the random catalog, as described in the previous sections, on the three separate regions: upper, middle and lower. The visual results are presented side by side in Fig.2.2.

The top row of Fig.2.2 illustrates the results obtained from analyzing the complete data sample using both random point catalog generation methods. It is evident that utilizing a uniform random point distribution in the ACF measurement leads to an overestimation when compared to the best-fit power law.
This behavior becomes apparent in the top-left corner of Fig.2.2 and manifests at large angular separations ($\theta \sim 100$"). It arises due to the introduction of artificial clustering signals induced by non-uniform depth variations across the survey area, with fluctuations of the order of one WFC3 field of view (i.e., $\gtrsim 100$"). These variations can be visually correlated with the typical fluctuations in the RMS maps shown in Fig.2.1. By employing a uniformly generated random point catalog, these variations are not taken into account.
In contrast, this issue is alleviated by adopting our proposed improved procedure for generating the random point catalog using a Monte Carlo recovery process (top-right of Fig.2.2), although it does not entirely eliminate the artificial clustering signals.
Interestingly, employing the latter approach yields a larger uncertainty in the correlation

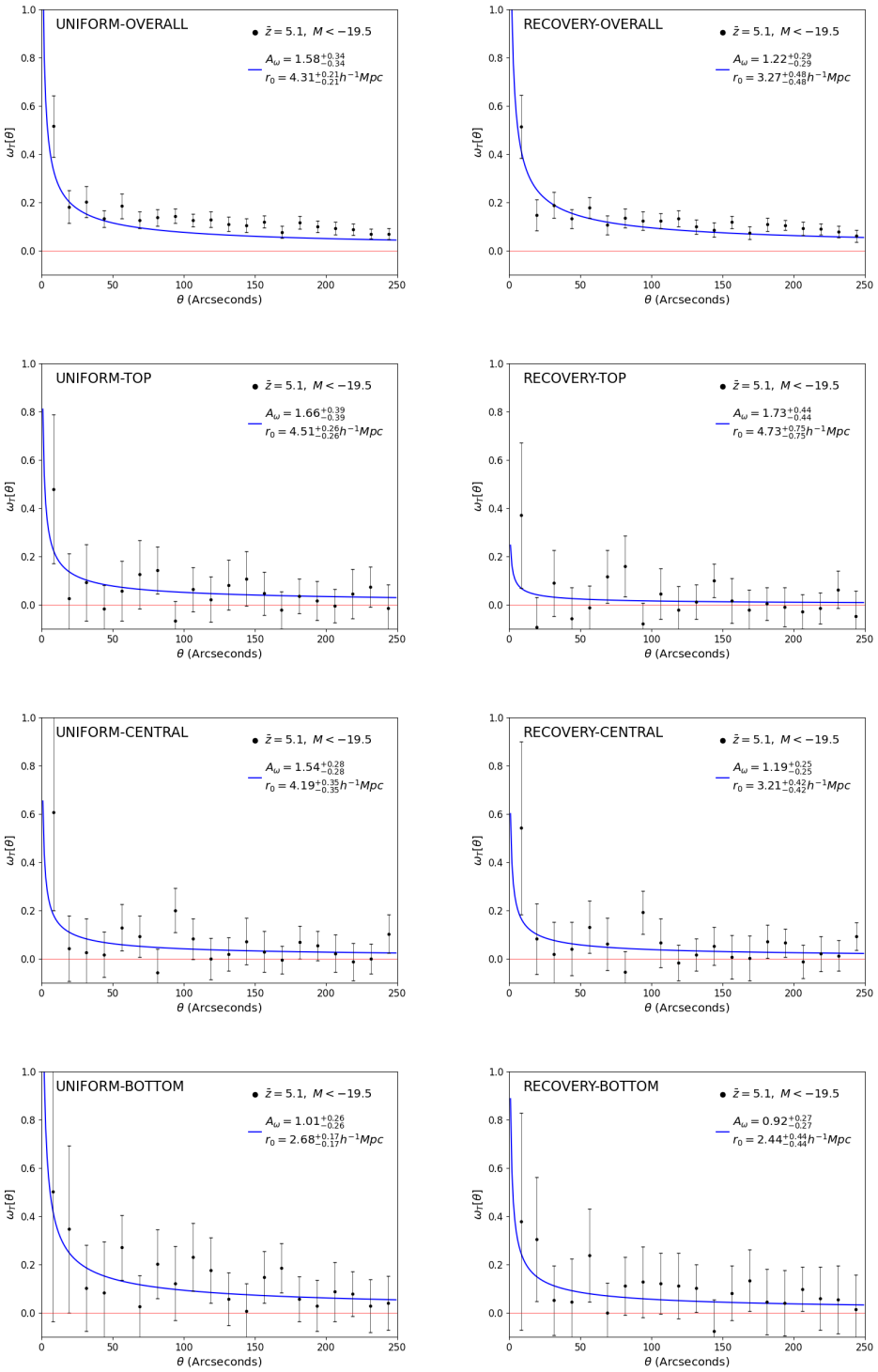

FIGURE 2.2: ACFs for LBGs at $\overline{z} = 5.1$ with $M_{UV} < -19.5$. On the left column we can see results obtained using a catalog of random points generated by a uniform distribution Sec.2.5.1.1, while on the right the analysis conducted with a random point recovery catalog Sec.2.5.1.2. From top to bottom we can see the results produced using data from: **Overall**, **Top**, **Central** and **Bottom** region of the survey in Fig.2.1. From this figure we can see that the central region is the one that better represent the analysis on the overall survey, this is because is the most deep as we can see in the bottom row of Fig.2.1.

length r_0 compared to the uniform random catalog counterpart. The difference amounts to a factor of two, and we consider this estimation to be more representative of the true errors in the measurement.
Examining the analysis conducted on the individual sub-regions, we observe a consistent trend in terms of error estimations for r_0. In all cases, the estimations tend to be $1.5-2\sigma$ larger compared to the results obtained from the uniformly generated sample.
By dividing the field into sub-regions of approximately uniform depth, we avoid significant overestimation of the ACF. This outcome arises from the fact that considering spatially confined regions leads to reduced variations in depth among the individual pointings superimposed in the mosaic. However, each region yields ACF and r_0 estimations that differ from one another at a level surpassing the 1σ associated uncertainty.

2.6 Clustering Results

2.6.1 Angular correlation function and Bias

In Fig.2.3 we present the ACF estimated from the candidate galaxies in the combined GOODS-N and GOODS-S catalog, while Fig.2.4 displays the ACF estimated from the XDF region. In both figures, we provide the measurements and associated uncertainties for each separation bin, along with the best-fitting model (indicated by the blue line) and its associated parameters A_w and r_0 reported as sub-plot legends. For comparison we present the results for all the redshift bins considered in the analysis $\overline{z} = 5.1$, $\overline{z} = 6.1$, $\overline{z} = 6.8$ and $\overline{z} = 7.7$ in separate panels.
As discussed in the data selection process, Sec.2.3, we are accounting for a cut in absolute magnitude in both GOODS-N and GOODS-S considering only candidates with $M_{UV} < -19.8$, while for XDF we did not consider any constraint in the magnitude (neither absolute or apparent). This means that the average absolute magnitudes of the samples are lower for this last field, ranging from $\overline{M}_{UV} = -18.6$ to $\overline{M}_{UV} = -18.1$ (without any apparent trend in redshift). The results of the ACF analysis for all fields are summarized in Tab. 2.2, where the left part reports the parameters obtained from the combined GOODS calatog and the right the results from the analysis on XDF.
First, we find that a power-law fit presents an overall good description of the data at all redshifts considered, including for our highest redshift sample at $\overline{z} = 7.7$ in both Fig.2.3 and Fig.2.4 (albeit the uncertainty is larger). From lower to higher redshift, if we consider a fixed angular scale, the ACF is increasing. As a result the two clustering parameters A_w and r_0 are increasing accordingly. Qualitatively, the XDF angular correlation function shape and redshift evolutionary trends are similar, but the increase in

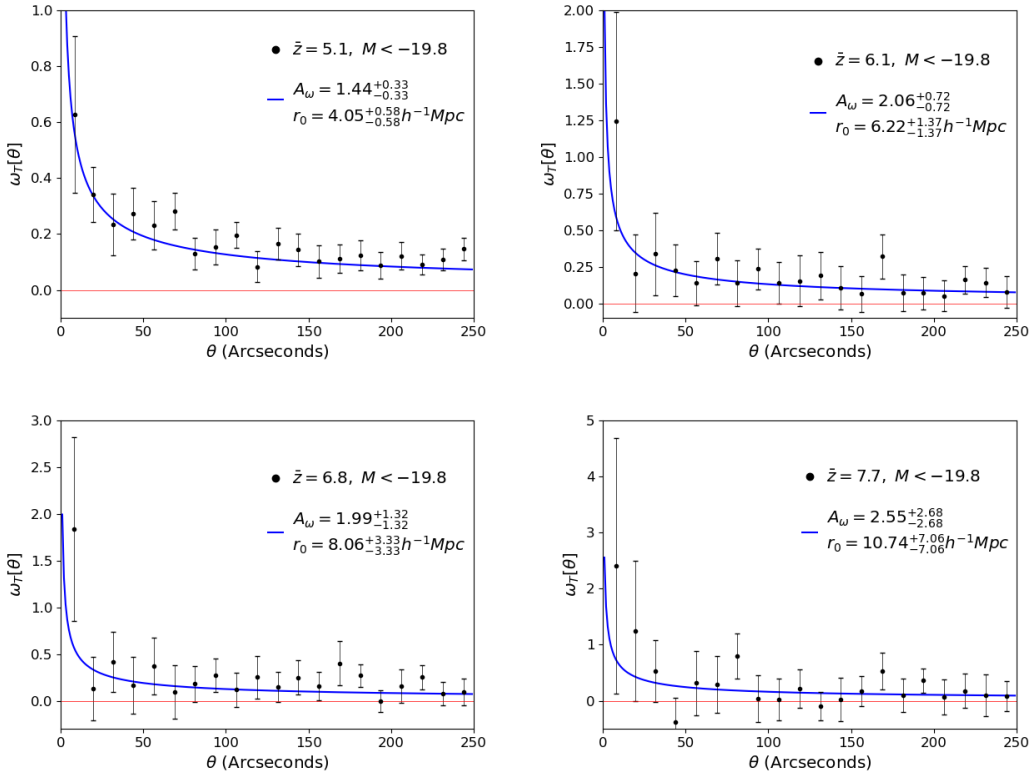

FIGURE 2.3: ACFs of LBGs in the GOODS catalog at four different redshifts. The results presented were obtained with random points generated from a recovery process Sec.2.5.1.2 and all the galaxies contained in the GOODS-N and GOODS-S region. The parameter A_w is a dimensionless coefficient obtained by fitting the ACF estimator Eq.2.1, is linked by Eq.2.6 to the correlation length r_0. Both these parameters present an increasing trend with redshift as excepted from theory and previous findings on this topic, we can see how the power law model well represent the ACF measurements.

clustering strength is more pronounced for this small-area field.

Fig.2.5 shows the evolution of the galaxy bias, obtained at the four different redshifts studied from the ACF, overlaid on dashed curves that correspond to the bias of dark-matter halos at constant mass (derived following Sheth & Tormen [245]). The comparison shows that the galaxy bias increases with redshift tracing approximately constant halo mass $M_h \sim 10^{11.4}$ $M_\odot$ for our large-area magnitude-limited GOODS sample. The results we obtain are fully consistent with theory (e.g., Sheth & Tormen [245]) and with previous observational studies (e.g.,Barone-Nugent et al. [17], Harikane et al. [117, 119], Hatfield et al. [121], Lee et al. [156], Overzier et al. [197], Qiu et al. [219]).

At a redshift of $z \sim 6$, our investigation of a luminous sample of LBGs unveils a galaxy bias of $b = 6.0 \pm 1.1$, which closely resembles the outcomes reported by other groups in the same redhisft regime (e.g.,Harikane et al. [117], Hatfield et al. [121]). In their study,

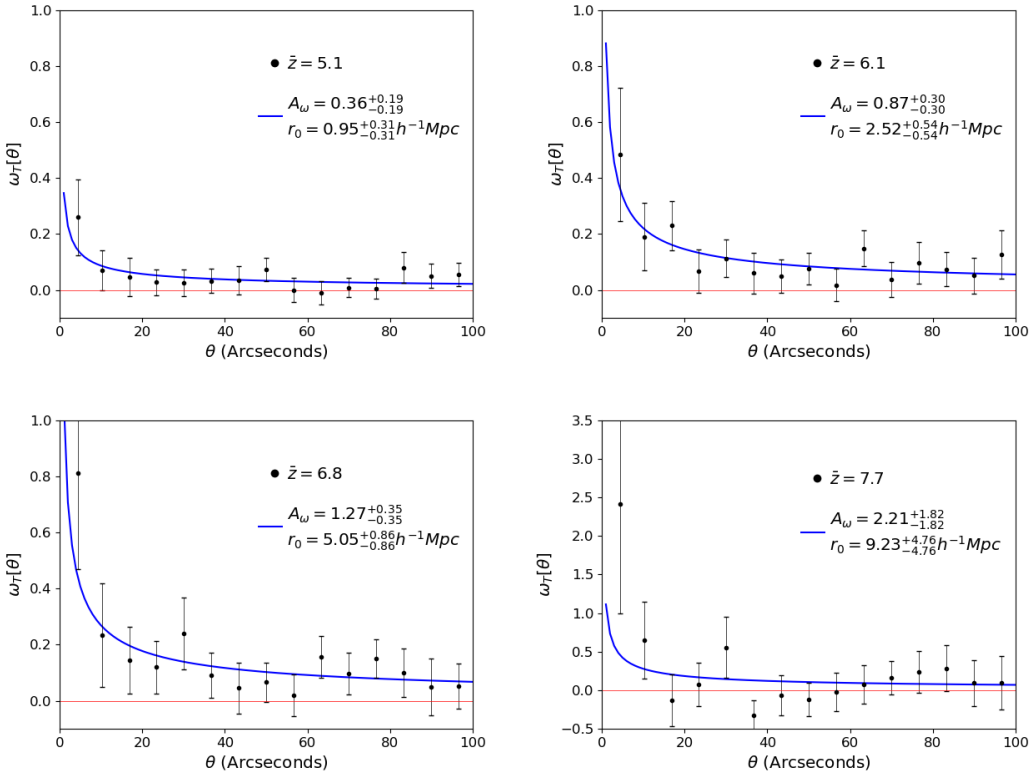

FIGURE 2.4: **Top**: ACF for LBGs in GOODS-S deep field XDF catalog. The results presented were obtained with random points generated from a recovery process Sec.2.5.1.2 and the entire set of data from Tab.2.1 without applying a cut in magnitude. The parameter A_w is a dimensionless coefficient obtained by fitting the ACF estimator Eq.2.1, is linked by Eq.2.6 to the correlation length r_0. In this figure we can see the same increasing trend of the clustering parameters as observed in Fig.2.3, values are lower because of the smaller area of XDF and the reduced pool of available candidates.

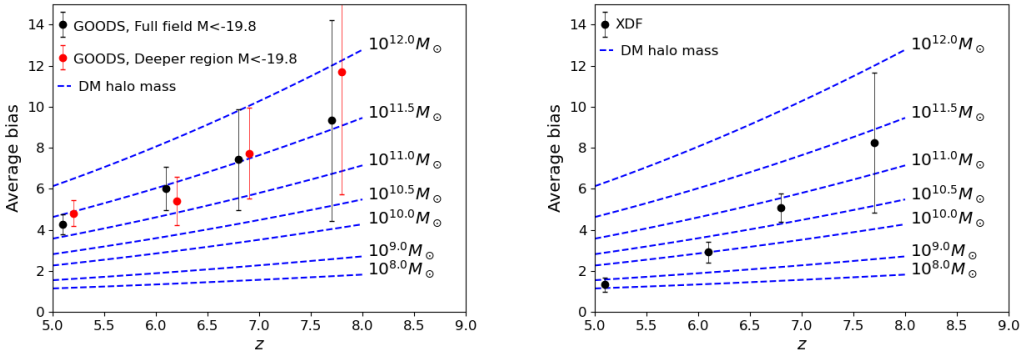

FIGURE 2.5: Bias evolution as a function of redshift from $5.1 \leq z \leq 7.7$. Dotted lines represents the dark matter halo bias from the Sheth & Tormen [245] mass function. On the left panel the red dots are slightly shifted for better visualisation.

they determine a minimum halo mass of $M_h \sim 10^{11.5} M_\odot$ using a simple abundance model, consistent with the predictions made in this work at the same redshift with respect to the halo mass function depicted in the left panel of Fig.2.5. Furthermore, Qiu et al. [219] presents the galaxy bias obtained at a redshift of $z \sim 7$ through clustering measurements with a similar sample. The resulting value is $b = 8.6^{+2.2}_{-2.8}$, which is consistent within the measurement uncertainties with the value found in our study at an average redshift of $z = 6.8$.

In the recent investigation conducted by Harikane et al. [119], they provide a summary of their findings based on the ACF of over $N = 4 \cdot 10^6$ galaxy candidates at $z \sim 2 - 7$. The far-right panel of the figure presents the results obtained for luminous UV galaxies at a redshift of $z = 5.9$. According to their analysis, the halo mass of these galaxies is estimated to be approximately $M_h \sim 10^{11.05} M_\odot$, with an uncertainty consistent with the findings of Barone-Nugent et al. [17] at a similar redshift of $z \sim 6$.

From the right panel of Fig. 2.5, we see that the XDF field shows a smaller bias compared to GOODS at all redshifts. This is expected because the sources are on average fainter as the data is deeper and no magnitude cut is applied to retain a sufficient number of galaxies for clustering analysis. However, in this case there is a strong redshift evolution in the inferred characteristic halo mass of the population, which at face value seems to evolve from $M_h \sim 10^8 M_\odot$ at $z = 5.1$ to $M_h \sim 10^{11.5} M_\odot$ at $z = 7.7$. Given the small area of the field, we consider this trend most likely to originate from cosmic variance in the clustering measurement itself, which is not captured by the error bars in the figure (covering only statistical uncertainty). Incidentally this trend was also present in the XDF analysis of Barone-Nugent et al. [17] and we have here extended it again from $z \lesssim 7$ to $z \lesssim 8$.

To provide a visual comparison with previous investigations on the topic, we present in Fig.2.6 the comparison in between the correlation lengths r_0. Through a comprehensive analysis of low redshift data ($z \lesssim 7.0$), our work exhibits a similar trend to that presented in other independent studies conducted on a comparable sample of LBGs that are bright. We find that the characteristic correlation length increases as we move towards higher redshifts. Our findings for each redshift bin align with previous redshift measurements reported in the literature (e.g.,Adelberger et al. [6], Barone-Nugent et al. [17], Harikane et al. [117, 119], Hatfield et al. [121], Lee et al. [156], Qiu et al. [219]), with consistency observed within the 1σ uncertainty range.

An important novelty of our analysis is the extension of this trend to higher redshifts ($z = 7.7$), specifically towards the mid-point of the epoch of reionization. We identify a correlation length of $r_0 = 10.74^{+7.06}_{-7.06}\ h^{-1}Mpc$, which aligns with the aforementioned trend. However, it is worth noting that the large error bars are primarily a consequence of the limited quantity of LBGs available in our sample at such high redshift. In the

coming years, this limitation is expected to improve significantly thanks to the utilization of the JWST in new surveys (e.g.,Hainline et al. [114], Paris et al. [200]).
As demonstrated in Sec.2.5.2, our analysis reveals that the clustering parameters are not significantly overestimated when studying similar-depth sub-regions. To ensure consistency, we extend our investigation to include the entire field while focusing on measurements obtained solely from the central region of both the GOODS-N and GOODS-S catalogs (see Fig.2.1). In Fig.2.7, we present the results, wherein the blue fitting curve represents the central region data, and we overlay the best-fit model obtained using the complete sample of galaxies labeled in green.
Tabulated in central column of Tab.2.2 are the results for both A_w and r_0, along with the associated galaxy bias, for all redshifts considered. Comparison of these measurements reveals consistency within the typical uncertainties of 1σ, which can also be visually observed in the left panel of Fig.2.5. It is worth noting that the galaxy bias b, when derived solely from the central region, marginally exceeds the value obtained from the overall analysis. This slight difference can be attributed to the fact that the central region represents the deepest portion among the three sub-regions, thus effectively capturing the general behavior exhibited by the entire field.

2.6.2 Duty Cycle

To study star formation efficiency, we need to link the dark matter halo mass function and the luminosity function. We can do this by introducing the duty cycle (ϵ_{DC}), the fraction of dark matter halos hosting UV-bright LBGs. We can estimate the duty cycle by matching objects at the same comoving density (Vale & Ostriker [271]).
This procedure is conducted by matching the number of galaxies with luminosity greater than a given L_g to the number of halos with mass greater than a minimum dark matter halo mass M_h assuming only a galaxy in each dark-matter halo (Martínez et al. [174]):

$$\epsilon_{DC}(z) \int_{M_h}^{+\infty} n(M_h, z) dM_h = \int_{L_g}^{+\infty} \Phi(L, z) dL, \tag{2.10}$$

where $\Phi(L, z)$ if the LF at redshift z and $n(M_h, z)$ is the halo MF Sheth & Tormen [245]. By taking the intersection between the bias inferred from abundance matching, using a Schechter LF: $\Phi^* = 10^{-0.44}$,$M^* = -20.19$ and $\alpha = -1.81$ at $\overline{z} = 7.7$ Bouwens et al. [34], and from clustering analysis, we derive the value of $\epsilon_{DC}(z)$. In Fig.2.8 we present the duty cycle $\epsilon_{DC}(z)$ evolution with the redshift.
We can see that at low redshift the duty cycle is contained below the unit value, while at $z > 6.5$ the duty cycle tends towards unity ($\epsilon = 1$), meaning that all the dark matter halos are hosting UV-bright LBGs with a mean absolute magnitude of $M_{UV} = -19.8$.

GOODS overall region			
$\overline{z}$	A_w	r_0	b
5.1	$1.44^{+0.33}_{-0.33}$	$4.05^{+0.58}_{-0.58}$	$4.27^{+0.49}_{-0.49}$
6.1	$2.06^{+0.72}_{-0.72}$	$6.22^{+1.37}_{-1.37}$	$6.03^{+1.06}_{-1.06}$
6.8	$1.99^{+1.32}_{-1.32}$	$8.06^{+3.33}_{-3.33}$	$7.42^{+2.45}_{-2.45}$
7.7	$2.55^{+2.68}_{-2.68}$	$10.74^{+7.06}_{-7.06}$	$9.33^{+4.90}_{-4.90}$
GOODS central region			
$\overline{z}$	A_w	r_0	b
5.1	$1.66^{+0.43}_{-0.43}$	$4.69^{+0.75}_{-0.75}$	$4.81^{+0.62}_{-0.62}$
6.1	$1.81^{+0.79}_{-0.79}$	$5.44^{+1.49}_{-1.49}$	$5.42^{+1.18}_{-1.18}$
6.8	$2.09^{+1.19}_{-1.19}$	$8.50^{+3.03}_{-3.03}$	$7.73^{+2.21}_{-2.21}$
7.7	$3.35^{+3.41}_{-3.41}$	$14.24^{+9.07}_{-9.07}$	$11.69^{+5.96}_{-5.96}$
XDF			
$\overline{z}$	A_w	r_0	b
5.1	$0.36^{+0.19}_{-0.19}$	$0.95^{+0.31}_{-0.31}$	$1.33^{+0.35}_{-0.35}$
6.1	$0.87^{+0.30}_{-0.30}$	$2.52^{+0.54}_{-0.54}$	$2.92^{+0.50}_{-0.50}$
6.8	$1.27^{+0.35}_{-0.35}$	$5.05^{+0.86}_{-0.86}$	$5.10^{+0.69}_{-0.69}$
7.7	$2.21^{+1.82}_{-1.82}$	$9.23^{+4.76}_{-4.76}$	$8.26^{+3.41}_{-3.41}$

TABLE 2.2: Galaxy clustering parameters obtained from ACF analysis with a fixed $\beta = 0.6$. For the combined GOODS analysis only the bright LBGs are considered, with M_{UV} <-19.8, while for XDF we have not applied a magnitude cut for lack in candidates, see Tab.2.1. **Note**: the correlation length r_0 is expressed in $h^{-1}Mpc$ units.

We note that the duty cycle estimated from the XDF field favours lower values of ϵ compared to the broader GOODS fields. This is a reflection of the increase in galaxy bias presented in the right-hand panel of Fig. 2.5.

The values found in this work are consistent with those of the XDF analysis conducted in Barone-Nugent et al. [17], they present the bias measured for XDF, HUDF92 and HUDF91 range from $b \sim 2$ to $b \sim 4$ in between $z \sim 4$ and $z \sim 6$. These three results do not present any particular pattern, but show very large variations in individual measurements, supporting our interpretation of sample variance being the most likely explanation of the XDF bias and duty cycle measurements.

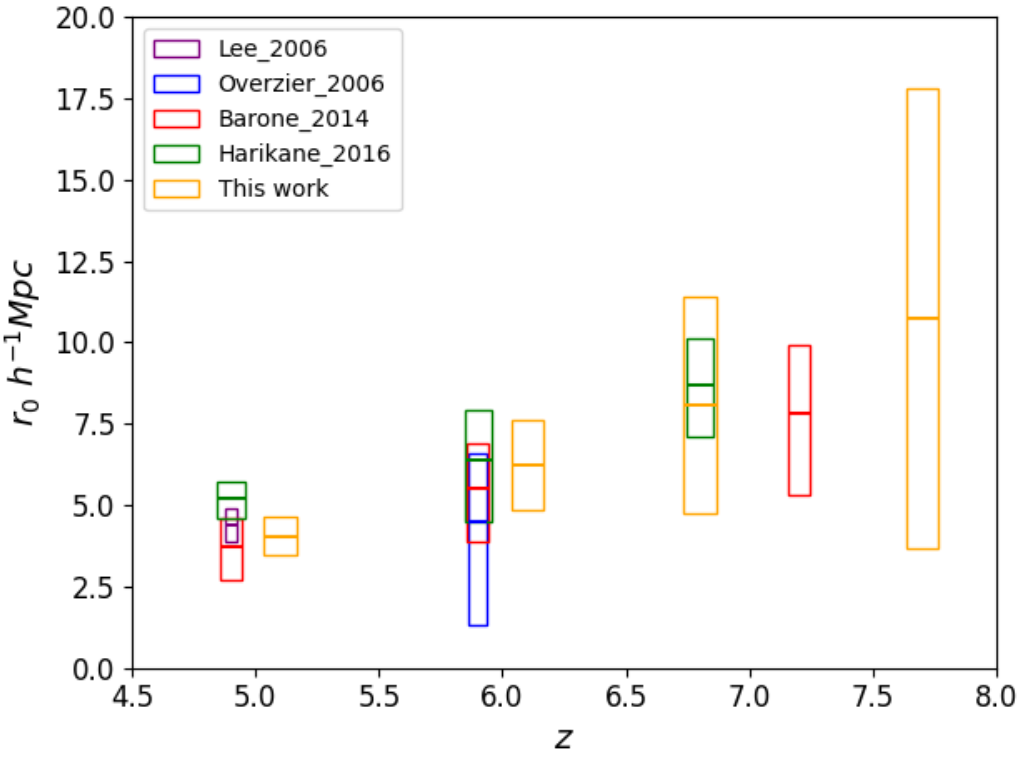

FIGURE 2.6: Comparison of the correlation length r_0 with previous studies conducted on a sample of UV bright LBGs at similar redshifts Barone-Nugent et al. [17], Harikane et al. [117], Lee et al. [156], Overzier et al. [197]. The height of the boxes represents the corresponding error bars, while the thick horizontal line within each box represents the peak measurement of the respective study. **Note**: the width of the boxes does not convey any physical information; it is solely used for enhanced visualization purposes.

2.7 Summary and perspectives

The primary goal of this study was to extend the investigation of galaxy clustering up to a redshift of $z \sim 8$. To achieve the necessary critical number of galaxies for such a study, we combined multiple surveys of varying depth and introduced a novel method to address the systematic bias in correlation measurements within regions of varying depth. To achieve our goal, we utilized the most recent data from the HST Legacy Fields, focusing specifically on three distinct fields: GOODS-N, GOODS-S, and XDF. Our analysis covered a range of mean redshifts, namely $z = 5.1$, $z = 6.1$, $z = 6.8$, and $z = 7.7$. By exploring these fields comprehensively, we aimed to provide an exhaustive analysis that surpasses previous studies in this field.

Here, we summarize the main aspects addressed and results obtained in this work:

• In order to gain a better understanding of various approaches to the analysis of the ACF, we conducted a preliminary study. This involved considering the different depths of the catalogs used and incorporating random points generated through a recovery process, as described in Sec.2.5. The purpose of this investigation was to explore the impact

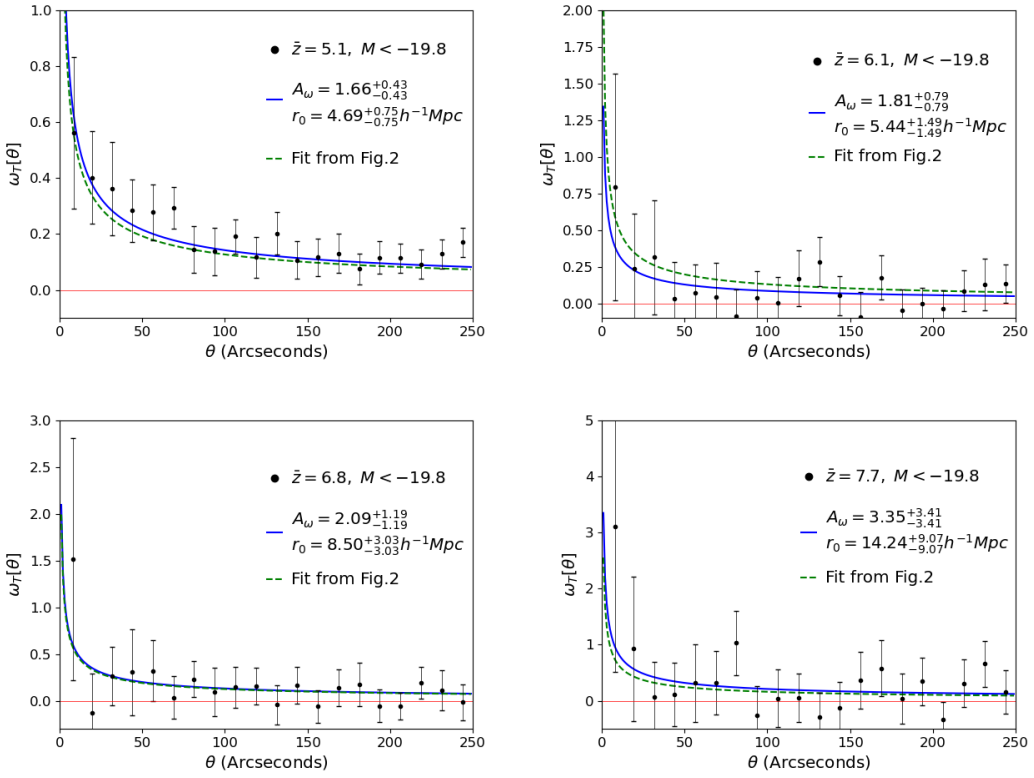

FIGURE 2.7: ACF for LBGs in GOODS catalog. The results presented were obtained with random points generated from a recovery process Sec.2.5.1.2 and all the galaxies contained in the GOODS-N and GOODS-S central region (bottom row Fig.2.1). The blue fitting line return the dimensionless parameter A_w linked by Eq.2.6 to the correlation length r_0 , while the green dashed line represent the fit obtained by considering the entire combined GOODS catalog.

of these factors on the ACF analysis and assess their implications for our research.

• To apply our innovative ACF estimation procedure, we utilized a combined GOODS catalog as well as the XDF field. The corresponding ACF measurements are presented in Fig.2.3 for the combined GOODS catalog and Fig.2.4 for the XDF field. At a mean redshift of $z = 7.7$, we obtained novel measurements of the galaxy bias. For the combined GOODS catalog, the galaxy bias was determined to be $b = 9.33^{+4.90}_{-4.90}$, while for the XDF field, the galaxy bias was found to be $b = 8.26^{+3.41}_{-3.41}$. These measurements represent important contributions to our understanding of galaxy clustering at this high redshift.

• In order to validate our results, we conducted a thorough comparison with previous studies on the topic, and we found agreement between our findings and those from the literature, as demonstrated in Fig.2.6. To further strengthen our analysis, we performed a supplementary validation by focusing on the central deep region of the GOODS catalog. Our results from this validation analysis were consistent with our main study, providing additional support for the robustness and reliability of our method. This validation analysis is presented in Fig.2.7. Together, these comparisons and supplementary

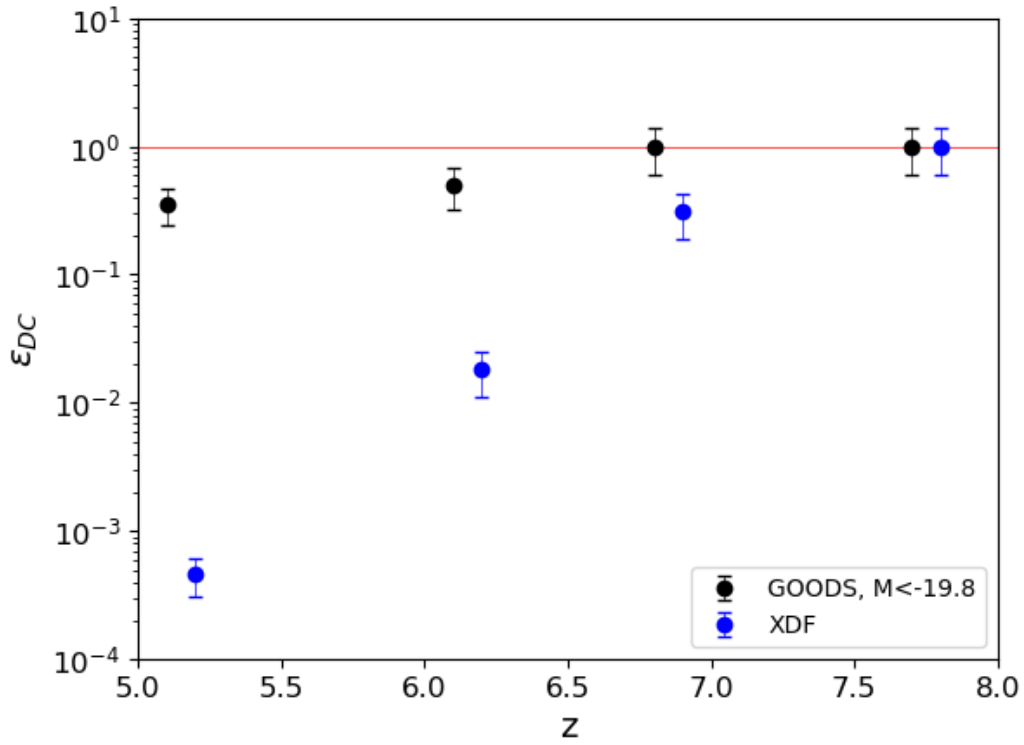

FIGURE 2.8: Duty cycle ϵ_{DC} analysis conducted on the two different samples at the four redshift considered in this work. For both $\overline{z} = 6.8$ and $\overline{z} = 7.7$ we found the upper limit represented by $\epsilon_{DC} = 1.0$.
The results from XDF, in blue, are slightly shifted in redshift to the right for clarity.

analyses reinforce the credibility and significance of our study.

• In Fig.2.5, we have summarized all the ACF results and depicted them graphically alongside the trend of dark matter halo bias from the model by Sheth & Tormen [245], considering different mass values. Notably, all three results at $z = 7.7$ consistently suggest a dark matter halo mass in the range of $M_h = 10^{11.4\sim11.5} M_\odot$. This finding indicates that as we move towards higher redshift, LBGs located within larger dark matter halos exhibit higher UV luminosity. Moreover, these results are consistent with previous studies on high-redshift galaxies (e.g.,Adelberger et al. [6], Barone-Nugent et al. [17], Harikane et al. [117, 119], Hatfield et al. [121], Lee et al. [156], Qiu et al. [219]). Our findings suggest that galaxies with greater mass and luminosity display higher levels of clustering compared to previous studies at similar redshifts. This implies that these galaxies are hosted by more massive dark matter halos, providing valuable insights into the relationship between galaxy properties and their underlying dark matter structures.

• We investigate the duty cycle of galaxies within dark matter halos using an abundance matching framework. We confirm the results from a previous study (Barone-Nugent et al. [17]), where our measured duty cycle approaches unity above $z > 6.5$ (see Fig.2.8).

Furthermore, we extended the duty cycle trend to $z \sim 8$ with our novel determination of the ACF at $z = 7.7$.

The extension of the clustering analysis to a redshift of $z = 7.7$ was conducted using a sample size of $N_g = 160$ objects derived from surveys with varying depths. This serves as a reference point for future forecasts of galaxy clustering studies at redshifts $z \gtrsim 8$ using the JWST. With the current preliminary determinations of galaxy luminosity functions (e.g.,Donnan et al. [77], Leethochawalit et al. [157]) and ongoing photometric surveys (e.g.,Hainline et al. [114], Paris et al. [200]), it is anticipated that high-quality photometric samples will become available within the next 12-24 months. These advancements will enable the study of the connection between galaxies and their associated dark matter halos during the earliest stages of cosmic reionization. This prospect holds great promise for furthering our understanding of the formation and evolution of galaxies in the early Universe.

Chapter 3

Galaxy clustering at cosmic dawn from JWST/NIRCam observations to redshift z∼11

This chapter is a reproduction of Dalmasso et al. [67]. The author lead both the data analysis and writing with the collaboration of the co-authors. This reproduction has been reformatted to fit the style of the thesis, changes include:

- **Figures and tables have been formatted to display correctly on the thesis**
- **Some bibliographic records have been updated**

3.1 Abstract

We report measurements of the galaxy two-point correlation function at cosmic dawn, using photometrically-selected sources from the JWST Advanced Deep Extragalactic Survey (JADES). The JWST/NIRCam dataset comprises approximately $N_g \simeq 7000$ photometrically-selected Lyman Break Galaxies (LBGs), spanning in the redshift range $5 \leq z < 11$. The primary objective of this study is to extend clustering measurements beyond redshift $z > 10$, finding a galaxy bias $b = 9.6 \pm 1.7$ for the sample at $\overline{z} = 10.6$. The result suggests that the observed sources are hosted by dark matter halos of approximately $M_h \sim 10^{10.6}$ $M_\odot$, in broad agreement with theoretical and numerical modelling of early galaxy formation during the epoch of reionization. Furthermore, the JWST JADES dataset enables an unprecedented investigation of clustering of dwarf

galaxies two orders of magnitude fainter than the characteristic L_* luminosity (i.e. with $M_{F200W} \simeq -15.8$) during the late stages of the epoch of reionization at $z \sim 6$. By analyzing clustering as a function of luminosity, we find that $b(M_{F200W})$ aligns with previous results for brighter galaxies and then decreases with M_{F200W}, as theoretically expected for fainter candidates. These initial results demonstrate the potential for further quantitative characterisation of the interplay between assembly of dark matter and light during cosmic dawn that the growing samples of JWST observations are enabling.

3.2 Introduction

The Universe we observe today formed from primordial fluctuations emerged during an inflationary epoch, subsequently amplified by gravitational instability shaping matter distribution. Gas dissipated over time, gravitating towards dark matter halos' central regions, intricately linking the growth and spatial arrangement of galaxies to their dark matter halos (Bullock et al. [46], Peebles [209], Persic & Salucci [213]).

The conceptualization of the galaxy–halo connection coincided with the realization that the spatial distribution of galaxies can yield insights into their formative properties. Early studies such as Peebles [209] focused on the two-point statistics of galaxies, as well as investigations on signal strength sensitivity to measure galaxy mass and luminosity (e.g., Bahcall & Soneira [12], Davis & Peebles [71]).

Recognizing that measuring these clustering properties could furnish information about the masses of the dark matter halos housing these galaxies was a pivotal insight due to the pronounced dependence of halo clustering on halo mass (e.g., Bardeen et al. [13], Jenkins et al. [134], Klypin et al. [145], Mo & White [181], Sheth & Tormen [246], Tinker et al. [263]).

A key tool to quantify clustering is the two-point correlation function ($\xi(r)$; Peebles [209]), which quantifies the excess probability of finding a galaxy at a given separation distance "r" compared to a random distribution. Given that precise redshift are not always available (e.g. in the case of photometrically selected samples), this function is often replaced by the Angular Two-Point Correlation Function (ACF), which expresses the excess as a function of the angular separation denoted by (θ) instead. The ACF can then be mapped to an two-point correlation function in physical space based on the (estimated) redshift distribution of the sample, leading to a characteristic correlation length (r_0; Davis & Peebles [71]). This parameter r_0 represents the typical separation distance at which galaxies exhibit significant correlation or clustering. The value of r_0 may vary depending on the specific sample of galaxies, the redshift range under investigation, and the cosmological model applied.

The correlation length is closely tied to determining another crucial parameter: galaxy

bias (b; Peebles [209]), representing the systematic deviation or clustering of galaxies compared to the underlying matter distribution in the Universe. A rising trend in galaxy bias suggests stronger clustering and in-turn this implies that the galaxy population is hosted in more massive dark-matter halos.

Clustering, when combined with information on the galaxy luminosity function, can also provide insight on the typical occupation fraction of dark-matter halos that host luminous galaxies.This fraction is known as the duty cycle (ϵ_{DC}). The duty cycle is typically derived by combining galaxy clustering measurements with a method called abundance matching (e.g., Vale & Ostriker [271]) by comparing the count of galaxies with luminosities surpassing a specified threshold, denoted as L_g, to the count of dark matter halos with masses exceeding a defined value M_h.

Past studies established early in the history of high redshift galaxy observations that with higher UV luminosities tend to be hosted by more massive dark matter halos (e.g., Adelberger et al. [6], Foucaud et al. [98], Giavalisco & Dickinson [106], Lee et al. [156], Ouchi et al. [195, 196]). The results have continued to be confirmed as a well-established trend over the last decade (e.g., Barone-Nugent et al. [17], Bian et al. [29], Dalmasso et al. [66], Harikane et al. [117, 119], Hatfield et al. [121], Qiu et al. [219], Savoy et al. [237]). Lyman Break Galaxies (LBGs) after the end of reionization at redshift $4.0 < z < 5.0$ show clear evidence for a luminosity-dependent bias, observed to the UV luminosity detection limit of past surveys $M_{UV} \lesssim -16.7$ (e.g., Harikane et al. [117, 119], Lee et al. [156]). This pattern persists when extending the redshift range to $5.0 < z < 6.0$, with absolute magnitudes ranging from $-22 \lesssim M_{UV} \lesssim -18$. The findings range from $b = 5.5^{+0.2}_{-0.4}$ for fainter objects to $b = 7.7 \pm 0.1$ for brighter ones, further corroborating the trend of galaxy bias with luminosity (e.g., Barone-Nugent et al. [17], Dalmasso et al. [66], Harikane et al. [117, 119], Hatfield et al. [121], Qiu et al. [219]). These findings align with expectations from basic theoretical modelling of the evolution of galaxy bias with luminosity based on assuming some efficiency in converting DM to stars, which may be calibrated at a reference redshift (e.g., Trenti et al. [267], Trenti et al. [268]).

Progress in clustering and bias measurement at high redshift is primarily limited by the challenges in building high quality samples of galaxies with robust redshifts. Furthermore, care must be applied when observations do not have uniform depth, as a spatially varying completeness can mimic a physical clustering signal. In Dalmasso et al. [66], we introduced a novel approach to galaxy clustering measurements to address the issue of non-uniform catalogs in depth and completeness. These two traits stem from the fact that during the observation of a target sky area with the narrow fields of view of infrared cameras on Hubble (and JWST), mosaicing is needed. As a result certain regions are observed multiple times from overlapping pointings and dither patterns. This results in some areas having greater depth than others in the final survey mosaic (due to overlap of individual frames). In Dalmasso et al. [66] we demonstrated that these systematic

sources of error can be effectively mitigated by generating a reference catalog of random points derived from artificial source recovery simulations. This allowed us to account effectively for map heterogeneities arising from multiple overlapping observations.
With this method Dalmasso et al. [66] measured a galaxy bias of $b = 9.3 \pm 4.9$ at redshift $z = 7.7$ using the Hubble Space Telescope (HST) Legacy Field data (Bouwens et al. [36]), which is the highest redshift clustering measurement reported so far.
Thanks to the increased depth, spectral coverage and field of view of JWST/NIRCam compared to Hubble/WFC3, this study aims to extend the clustering redshift frontier further. In addition, we also aim to take advantage of the increased sensitivity to study the luminosity dependence of clustering reaching down to $M_{F200W} \lesssim -14.5$, approaching faint dwarf galaxies. To conduct this analysis, we utilize photometrically-selected sources from the JWST Advanced Deep Extragalactic Survey (JADES) detected by the NIRCam instrument.
The structure of this chapter is outlined as follows. In Sec.3.3, we elucidate our process for selecting data pertaining to the LBG candidates. In Sec.3.4, we delve into the theory and variables estimation to study the ACF. Our discoveries and numerical results are expounded upon in Sec.3.5. To conclude in Sec.3.6, we offer a summary of our work and present our conclusions. In this work we assume, when relevant, the cosmological parameters determined by the Planck Collaboration Ade et al. [5]: $(\Omega_M, \Omega_\lambda, h, \sigma_8) = (0.315, 0.685, 0.673, 0.828)$. Magnitudes are in the AB system Oke & Gunn [191].

3.3 Data selection

This work is based on the NIRCam imaging public data release over the GOODS South field by the JWST Advanced Deep Extragalactic Survey (JADES)[1] collaboration (Bunker et al. [47], Eisenstein et al. [84], Hainline et al. [114], Rieke et al. [224]).
Specifically, in this study we utilized images obtained from JADES Version 1.0, acquired through the NIRCam instrument exhibiting a pixel scale of 30 milliarcseconds. As for the galaxies survey used in this work we considered the candidates detected in the F200W band presented in the "NIRCam Photometry"[1] catalog as well as the scientific images, segmentation maps, root mean square (rms) all presented in Rieke et al. [224].
For reference, we summarize briefly the detection and photometry methods used to create the catalog, referring the reader to Rieke et al. [224] for more details. The detection image was created by stacking signal and noise images of specific filters from the NIRCam LW. Python-based tools like `Astropy` AstropyCollaboration [11], `Photutils`[2],

[1] https://archive.stsci.edu/hlsp/jades
[2] https://photutils.readthedocs.io/en/stable/citation.html

`scikit-image` van der Walt et al. [294], and `cupy` Okuta et al. [192] were utilized, drawing inspiration from `NoiseChisel` Akhlaghi & Ichikawa [7]. A detection catalog was generated using the `Photutils detect-sources` routine, followed by operations to reduce noise and separate blended objects. Deblending was performed iteratively, and satellites were identified within extended sources. A final segmentation map was defined, incorporating masks for bright stars and persistence. The segmentation map, generated through the described detection method, served as the basis for constructing the JADES photometric catalog. Photometric redshifts were measured using the template-fitting code `EAZY` Brammer et al. [43] with the templates from Hainline et al. [114] and PSF-matched Kron fluxes.

To create our parent catalog of galaxies we took the photometric catalog from Rieke et al. [224] and selected candidates within the redshift interval of $5.0 \leq z < 11.0$. This procedure resulted in the delineation of six discrete redshift bins, each characterized by an identical width of $\Delta z = 1.0$, along with an associated assignment of an average redshift value denoted as $\overline{z}$ for each bin namely $\overline{z} = 5.5$, 6.5, 7.4, 8.5, 9.3, and 10.6. Second, within each of these redshift bins, we select candidates that exhibited a signal-to-noise ratio exceeding $SNR \geq 5.0$ in the UV band. Third, to enhance the precision of our analysis and identify sources in adjacent spatial regions, we ensure that the galaxies are confined to the observable region in both rms and segmentation maps.

Lastly, capitalizing on the exceptional depth and sensitivity of JWST, we aimed to explore the cosmos at ever-diminishing magnitudes. Given that uncertainty in the estimation of completeness of a survey may affect the clustering estimation, especially if there are spatial variations across the field, we impose a magnitude threshold of $m_{F200W} < 31.1$ ($M_{F200W} < -15.5$ at $z \simeq 5.5$). The corresponding completeness ranges from 8% to 35% for a magnitude range of $30.1 < m_{F200W} \leq 31.1$. Therefore, we opted to include only the sub-regions with a completeness above 30%, ensuring a robust analysis in the faintest regime. For brighter objects, the completeness ranges from 60% to 73% for $28.6 < m_{F200W} \leq 30.1$, and from 73% to 92% for $26.6 < m_{F200W} \leq 28.6$.

The outcome of the data selection process is displayed in the initial four columns of Tab.3.1. These columns present the redshift bin (#1), the mean redshift designated as $\overline{z}$ (#2), the total count of galaxies within each redshift bin from the source catalog (#3), and the number of galaxies we selected in each bin for subsequent analysis based on the selection criteria explained above (#4).

3.4 Clustering estimation

3.4.1 ACF fitting model

We start by constructing an estimator w_{obs} for the observed ACF. This estimator measures the excess probability of encountering pairs of objects, specifically galaxies within our investigation, at a specific angular separation denoted as (θ) compared to a uniform random distribution. We follow Landy & Szalay [151], which define w_{obs} as follow:

$$w_{\rm obs}(\theta) = \frac{DD(\theta) - 2DR(\theta) + RR(\theta)}{RR(\theta)}, \tag{3.1}$$

In this equation, $DD(\theta)$ identifies the count of pairs of observed galaxies falling within an angular separation range of ($\theta \pm \delta\theta$). $RR(\theta)$ corresponds to a similar count but is derived from a random catalog that shares identical geometric and selection characteristics as the observed catalog. $DR(\theta)$ denotes the count of pairs that include one observed galaxy and one object from the random catalog.

The finite size of this survey results in an underestimation of the measured parameters, a bias that can be corrected by introducing a constant known as the integral constraint (IC; Peacock & Nicholson [206]):

$$w_{\rm true}(\theta) = w_{\rm obs}(\theta) + IC. \tag{3.2}$$

This correction constant is estimated as:

$$\begin{aligned} IC &= \frac{1}{\Omega^2}\int_1\int_2 w_{\rm true}(\theta)d\Omega_1 d\Omega_2 \\ &= \frac{\sum_i RR(\theta_i)w_{\rm true}(\theta)}{\sum_i RR(\theta_i)} = \frac{\sum_i RR(\theta_i)A_w\theta_i^{-\beta}}{\sum_i RR(\theta_i)}. \end{aligned} \tag{3.3}$$

In the above equations, $w_{\rm true}(\theta)$ represents the intrinsic ACF, $w_{\rm obs}(\theta)$ is the measured ACF within the survey area and $RR(\theta_i)$ is the random-random count of galaxies in a specific angular bin.

Owing to variations in the sizes of candidate subsamples within each redshift bin under examination in this analysis (refer to column #4 in Tab.3.1), a decision was made to partition the investigated angular range, denoted by $\theta \in (0, 250)$", into varying numbers of angular bins tailored for each redshift range considered. The guiding criterion adopted involves subdividing the angular range into a finite set of bins, ensuring an approximately uniform count of candidate pairs (DD) per bin. This binning strategy was adopted to ensure that Poisson noise and sample sizes remain consistent across different angular separations. The results are then plotted on a logarithmic scale, but this has no impact in the determination of the bin sizes and ranges.

Following best practices from prior work (e.g., Barone-Nugent et al. [17], Dalmasso et al. [66], Lee et al. [156], Overzier et al. [197]), we assume a fixed value of $\beta = 0.6$ and consider a maximum separation of $\theta = 250$". When the correlation slope β is held constant, the quantity IC/A_w is solely contingent on the size and configuration of the survey area, and it can be calculated using Eq.3.3. The single parameter available for optimization is A_w, which is fine-tuned by minimizing the χ^2 value. The fitting model is outlined as follows:

$$w_{\rm mod}(\theta) = A_w(\theta^{-\beta} - IC/A_w). \tag{3.4}$$

To assign uncertainties to the ACF measurements, we construct the normalized covariance matrix using the standard estimator:

$$C_{ij} = \frac{1}{N-1}\sum_{l=1}^{N}\left[\omega^l(\theta_i) - \overline{\omega}(\theta_i)\right]\left[\omega^l(\theta_j) - \overline{\omega}(\theta_j)\right]. \tag{3.5}$$

In this equation, the summation is over N independent realizations. However, we used results obtained from a bootstrap resampling analysis (e.g., Benoist et al. [21], Ling et al. [161]) extrapolated from the same parent sample, so they are not independent. When the covariance matrix is estimated from the data itself, a correction factor of $(N-1)^2/N$ must be taken into account, and the covariance matrix becomes:

$$C_{ij} = \frac{N_{\rm boot}-1}{N_{\rm boot}}\sum_{l=1}^{N_{\rm boot}}\left[\omega^l(\theta_i) - \overline{\omega}(\theta_i)\right]\left[\omega^l(\theta_j) - \overline{\omega}(\theta_j)\right], \tag{3.6}$$

where $N_{\rm boot}$ is the total number of resamplings performed, $\omega^l(\theta_i)$ is the measured cross-correlation function from each realization in the i-th bin, and $\overline{\omega}(\theta_i)$ is the mean of the cross-correlation function in the same bin. We considered the associated uncertainty in each bin to be the square root of the corresponding diagonal element in the covariance matrix.

To derive the galaxy bias, we approximate the real-space correlation function using a power-law expression $\xi(r) = (r/r_0)^{-\gamma}$, with γ defined as $\gamma \equiv \beta + 1$, i.e., we fix γ to 1.6. We derive r_0 using the amplitude A_w through the Limber transform Adelberger et al. [6]:

$$A_w = \frac{r_0^{\gamma} B[1/2, 1/2(\gamma-1)] \int_0^{\infty} dz N(z)^2 f^{1-\gamma} g(z)^{-1}}{[\int_0^{\infty} dz N(z)]^2}. \tag{3.7}$$

In this context, f stands for the transverse comoving distance, given by $f \equiv (1+z)D_A(\theta)$, where D_A is the angular diameter distance. $N(z)$ represents the redshift distribution of dropouts, factoring in source identification efficiency and completeness. The term $B(t_1, t_2) = \Gamma(t_1)\Gamma(t_2)/\Gamma(t_1+t_2)$ denotes the beta function, $g(z) \equiv c/H(z)$ is the comoving distance and r_0 represent the correlation length measured in $h^{-1}Mpc$.

From the real-space correlation function $\xi(r)$, we define the galaxy bias $b(z)$ as the ratio

between the galaxy variance $\sigma_{8,g}$ and the linear matter fluctuation $\sigma_8(z)$, both within a sphere of comoving radius $8h^{-1}Mpc$ Peebles [209]:

$$b(z) = \frac{\sigma_{8,g}}{\sigma_8(z)}, \tag{3.8}$$

assuming $\sigma_8(0) = 0.828$ from the cosmology model and the galaxy variance expressed as:

$$\sigma^2_{8,g} = \frac{72(r_0/8h^{-1}Mpc)^\gamma}{(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma}. \tag{3.9}$$

3.4.2 Random points catalog

In Eq.3.1, $RR(\theta)$ is the pair count of uniform random points (galaxies) generated in a representative and realistic way that emulates any selection bias of the parent sample. To generate the random points catalog we followed the recovery procedure explained in detail in Dalmasso et al. [66]. In short, we used the injection-recovery code `GLACiAR2` Leethochawalit et al. [158] to inject galaxies in bins of redshifts ranging from $z_{min} = 5.0$ to $z_{max} = 11.0$ with an increment of 0.5 and UV magnitudes ranging from $m_{min} = -24.4$ to $m_{max} = 32.$ with an increment of 0.5 mag into the JADES images. We injected $N = 3200$ galaxies in each redshift-magnitude bin at random positions (roughly one galaxies per 30 arcsec^2). The galaxies have disk-type light profiles with Sersic index $n = 1$ and random inclinations and ellipticities. The SEDs of the injected galaxies are randomly drawn from the JAGUAR mock catalogue Williams et al. [283] of the same redshift bin. The galaxies are recovered in a similar fashion to the real data as described in Rieke et al. [224] where the detection image is the signal-to-noise image created from F227W, F335M, F56W, F410M and F444W images. The final random catalogues are those that are recovered to be either isolated or minimally overlapping with existing objects.

As the random catalog of synthetic sources was generated based on the assumption of a flat input luminosity function $\Phi(m)$, we carried out a Monte Carlo hit-and-miss procedure to establish the final set of random points for our clustering analysis. The procedure consisted of assigning to each element in the random catalog of galaxy a number between zero and one and compare that with the probability of finding that galaxy in an observation. The probability function governing the selection process was formulated as follows:

$$p(m) = \frac{\Phi(m)}{\Phi(m_{\rm lim})}, \tag{3.10}$$

where m is the apparent magnitude of the considered galaxy and m_{lim} is the limit apparent magnitude considered of $m_{F200W} = 31.1$ (see Sec.3.3). We used the Schechter luminosity function parameters evolution with redshift from Bouwens et al. [36].

GOODS-S Full field									
$z-bin$ (1)	$\overline{z}$ (2)	$N_{g,tot}$ (3)	$N_{g,sel}$ (4)	A_w (5)	$r_0/(\mathrm{h}^{-1}\mathrm{Mpc})$ (6)	b (7)	$log(M_h/M_\odot)$ (8)	ϵ_{DC} (9)	χ^2 (10)
$5 \leq z < 6.0$	5.5	3449	2075	0.6 ± 0.3	1.5 ± 0.5	2.1 ± 0.6	$9.5^{+0.2}_{-0.2}$	$0.006^{+0.005}_{-0.003}$	0.6
$6 \leq z < 7.0$	6.5	3880	2786	0.8 ± 0.3	2.2 ± 0.6	3.1 ± 0.5	$9.9^{+0.3}_{-0.4}$	$0.011^{+0.023}_{-0.008}$	0.1
$7 \leq z < 8.0$	7.4	1472	1041	0.9 ± 0.4	3.1 ± 0.7	4.1 ± 0.8	$10.1^{+0.4}_{-0.5}$	$0.012^{+0.040}_{-0.010}$	0.8
$8 \leq z < 9.0$	8.5	982	631	1.8 ± 0.4	3.5 ± 0.7	5.2 ± 0.7	$10.15^{+0.3}_{-0.3}$	$0.012^{+0.032}_{-0.009}$	0.9
$9 \leq z < 10.0$	9.3	426	163	1.9 ± 0.6	4.8 ± 1.3	7.2 ± 1.5	$10.47^{+0.3}_{-0.4}$	$0.045^{+0.230}_{-0.039}$	0.7
$10 \leq z < 11.0$	10.6	207	55	2.5 ± 1.1	6.8 ± 1.4	9.6 ± 1.7	$10.60^{+0.3}_{-0.4}$	$0.075^{+0.540}_{-0.067}$	0.2

TABLE 3.1: Galaxy clustering parameters obtained from ACF analysis with power-law model with a fixed $\beta = 0.6$. (1) Redshift bin. (2) Mean redshift. (3) Number of galaxies in the redshift bin. (4) Number of galaxies selected according to the criteria (Sec.3.3). (5) Best-fit value of the amplitude A_w. (6) Best-fit value of the correlation length r_0 in units of h^{-1}Mpc. (7) Galaxy bias. (8) Dark matter halo mass (log scale) in units of $M_\odot$. (9) Duty cycle. (10) Reduced χ^2 value.

3.5 Results and Discussion

3.5.1 Angular correlation function and Galaxy bias

Fig.3.1 presents the ACF estimated using LBG candidates from the GOODS-S field. The lower right panel features the newly obtained measurement at redshift $z = 10.6$. Also, within the same figure we present results spanning the redshift range $5.0 \leq z < 10.0$ to show the evolution of the ACF with redshift and to compare our new measurements from JWST data with those reported in the literature and based on other facilities. Each panel also reports two clustering parameters: A_w, the coefficient obtained from fitting the angular correlation function using Eq.3.4, and the correlation length r_0 derived from the fit (both with respective uncertainties at 1σ). The parameters obtained are tabulated in columns (5) and (6) in Tab.3.1.

With a simple functional model (Eq.3.4) to fit our measurements, we get overall agreement with the data, even at high redshifts. Interestingly, the measurement and fit remains of acceptable quality even when number of sources is low, as exemplified by the $N_g = 55$ candidates at redshift $z = 10.6$ (bottom right panel of Fig.3.1). This determination at redshift $z = 10.6$ marks the first measurement of galaxy clustering at less than 500 Myr since the Big Bang based on sources substantially fainter than previously reported with $M_{F200W} \lesssim -17.3$. The galaxy bias measurement can also be used to determine the characteristic dark matter halo mass that hosts the observed galaxies. For a fixed redshift, this is done by computing the theoretical dark matter halo bias $b(M_h)$ as a function of halo mass M_h as described by Sheth & Tormen [246] and by numerically inverting the relationship to derive $M_h(b)$. Using the results presented in Fig.3.1 and reported in column (6) in Tab.3.1 we obtain the corresponding dark matter halo

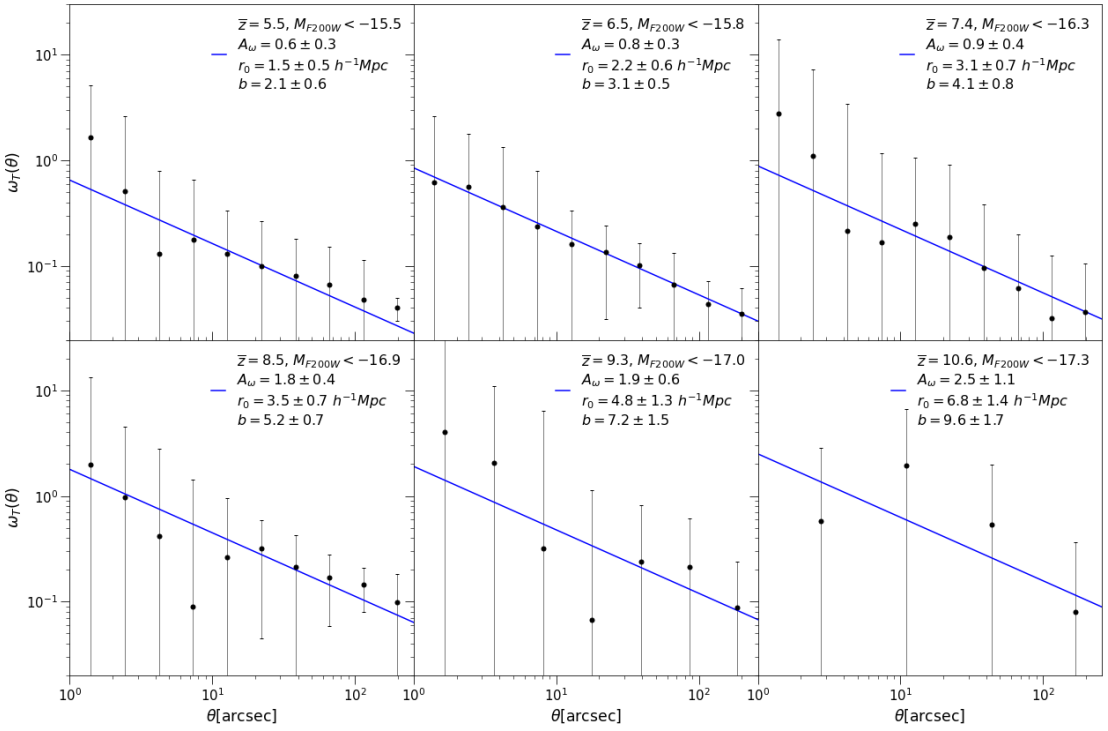

FIGURE 3.1: Measured ACFs in the GOODS-S field at six different mean redshifts. The black dots denote the ACF measurements with associated 1σ uncertainties, the blue solid lines represent a the power law best-fit with parameters listed in the legend box.

masses in column (8). Our results confirm the trends in redshift evolution of clustering reported in the literature, specifically the correlation length and bias steadily increase with redshift. In Tab.3.1, the last column displays the values of the reduced χ^2 to offer information on the single power law fit quality.

In Fig.3.2, the evolution of galaxy bias is presented for ACF measurements obtained with galaxy candidates brighter than $M_{F200W} \lesssim -17.0$ at $z < 10$. Fig. 3.2 also reports bias measurements in the literature, based primarily on Hubble Space Telescope and/or Subaru observations (Barone-Nugent et al. [17], Dalmasso et al. [66], Harikane et al. [117], Overzier et al. [197], Qiu et al. [219]). It should be noted that a direct comparison is not possible, as the JWST observations are substantially deeper. However, there is qualitative consistency since the bias we measure is lower than the previous measurements for brighter sources at fixed redshift. Intrinsically, our observations show a clear trend of increasing bias with redshift and a slight increase characteristic halo mass as well, which we interpret as a result of the evolution of the luminosity distance, which makes a magnitude limited survey probe brighter (and therefore more massive objects) at increasing redshift (e.g., Mason et al. [176], Trenti et al. [268]).

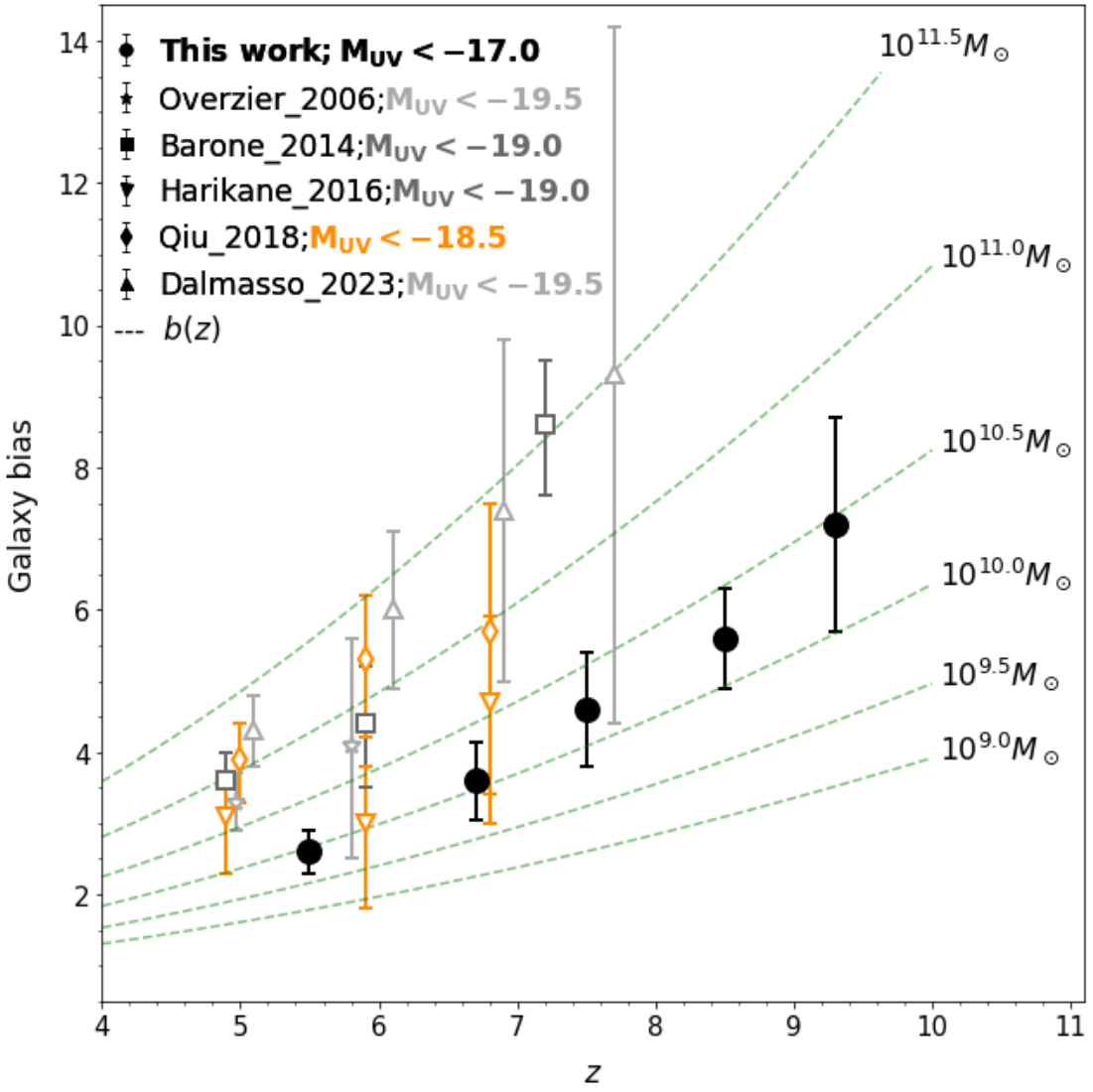

FIGURE 3.2: Galaxy bias evolution as a function of redshift z. Circles represent the results obtained in this work. Other symbols are previous bias measurements: stars Overzier et al. [197], squares Barone-Nugent et al. [17], downward triangles Harikane et al. [117], diamonds Qiu et al. [219] and upward triangles Dalmasso et al. [66]. Green dashed lines represent the theoretical bias evolution at a fixed dark matter halo mass M_h $[M_\odot]$. Note that our observations are substantially deeper than other studies, hence the inferred bias is lower.

3.5.2 Galaxy bias magnitude dependence

Interestingly, our sample is sufficiently large to explore luminosity dependence on galaxy bias by building subsamples of sources selected at different magnitude intervals for sources at redshifts $z \sim 5$, $z \sim 6$, and $z \sim 7$ (sources at higher redshift are too few to enable a split in magnitude subsamples). The magnitude bin size for each redshift selection has been chosen to ensure a balanced distribution of sources in each subsample, resulting in the average M_{F200W} for each bin shown in Fig.3.3.

For comparison in Fig.3.3 we also report results from prior studies on galaxy clustering. Reassuringly, with a consistent luminosity selection, our bias determination for the brightest samples are overlapping with prior measurements reported in the literature (Barone-Nugent et al. [17], Dalmasso et al. [66], Harikane et al. [117], Overzier et al. [197], Qiu et al. [219]). This agreement gives both confidence in our analysis, and also confirms that previous photometric samples of candidate galaxies at the end of the epoch of reionization have the same clustering properties as the higher quality JADES samples (which benefit from deeper data for improved photometric selection).
Fig.3.3 shows a decrease of bias with decreasing galaxy luminosity out to $M_{F200W} \sim -15.5$. This observed trend is effectively captured by the theoretical model for the bias evolution $b(M_{F200W})$ derived from Trenti et al. [268] using the Gamma Ray Burst no-metal bias scenario, which traces star formation and therefore gives a halo mass-UV magnitude relationship we can plot against our data.

As an additional check for the robustness of estimated errors, we considered a larger catalog of random points ($N_r^{'} = 20 \cdot N_r$), finding that the results are consistent with each other, with a 2% deviation from results obtained using a smaller random catalog. In passing, we note that if we were to extend our analysis to fainter sources in the magnitude regime around $M_{F200W} \sim -14.0$, we would derive a galaxy bias of $b \simeq 7$ at redshift $z = 5$, i.e. a significant increase over the decreasing trend presented in Fig. 3.3. Such high value would be inconsistent with theoretical predictions for clustering of dwarf galaxies. We attribute the high measurement to spatially-dependent systematic error estimation of completeness for the faintest sources detectable by JWST that are not fully captured by our source recovery simulations because of the extreme faintness of such sources, which results in very low completeness. This provides supports for our choice to limit the analysis to sources with brightness $m_{F200W} < 31.1$ ($M_{F200W} < -15.5$ at $z \simeq 5.5$) in order to achieve robust clustering measurements.

3.5.2.1 ACF measured for the bias evolution with magnitude

In this subsection, we present as Fig.3.4 the detailed set of plots used in our analysis of the ACF measurements in order to investigate potential dependencies in galaxy bias evolution relative to the magnitude of the sample analyzed. The corresponding discussion and interpretation forms part of Sec. 3.5.2.

3.5.3 Abundance matching and Duty Cycle

To explore the connection between dark mass and star formation at high redshift, we can resort to abundance matching and duty cycle modelling. This process entails statistically

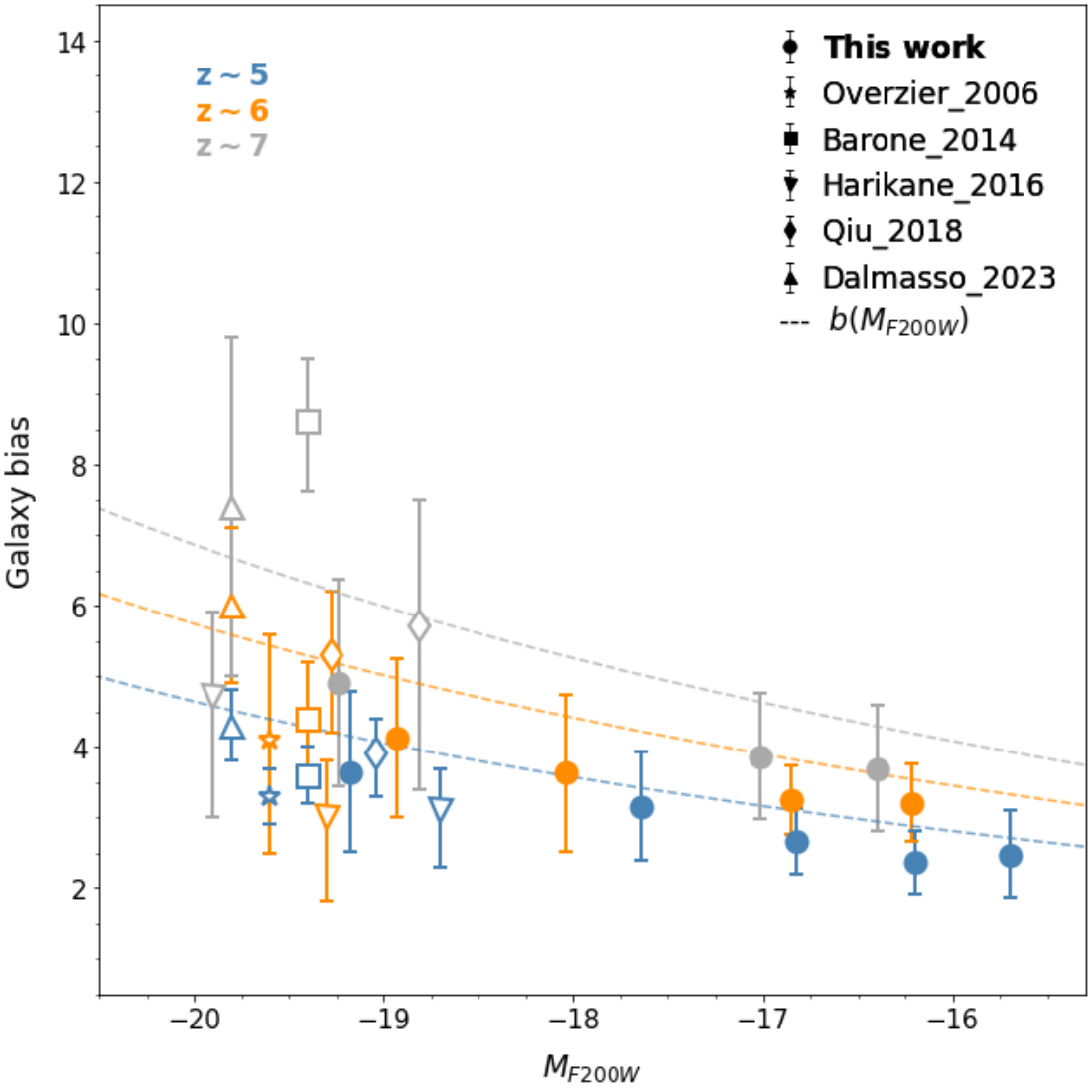

FIGURE 3.3: Galaxy bias evolution as a function of absolute magnitude M_{F200W} for three different redshift bins (blue, orange and grey for $z = 5, 6, 7$ respectively. Circles represent the results obtained in this work. Other symbols are literature measurements: stars Overzier et al. [197], squares Barone-Nugent et al. [17], downward triangles Harikane et al. [117], diamonds Qiu et al. [219] and upward triangles Dalmasso et al. [66]. Colored dashed lines represent the theoretical bias evolution model as a function of absolute magnitude M_{F200W}, derived from Trenti et al. [268].

correlating the mass of dark matter halos with the luminosity (or stellar mass) of galaxies in a given sample. This methodology is based on the premise that the most luminous galaxy inhabits the most massive halo, followed by the second most massive galaxy residing in the subsequent most massive halo, and so forth. By doing so, it is possible to infer the intricate relationship between galaxy formation and the underlying dark matter structure (but see Ren et al. [220]). This relationship serves as the basis for calculating the duty cycle ϵ_{DC}, which is defined as the fraction of time a dark matter halo of a certain mass hosts a UV-bright LBGs (e.g.,Martínez et al. [174], Vale & Ostriker [271]).

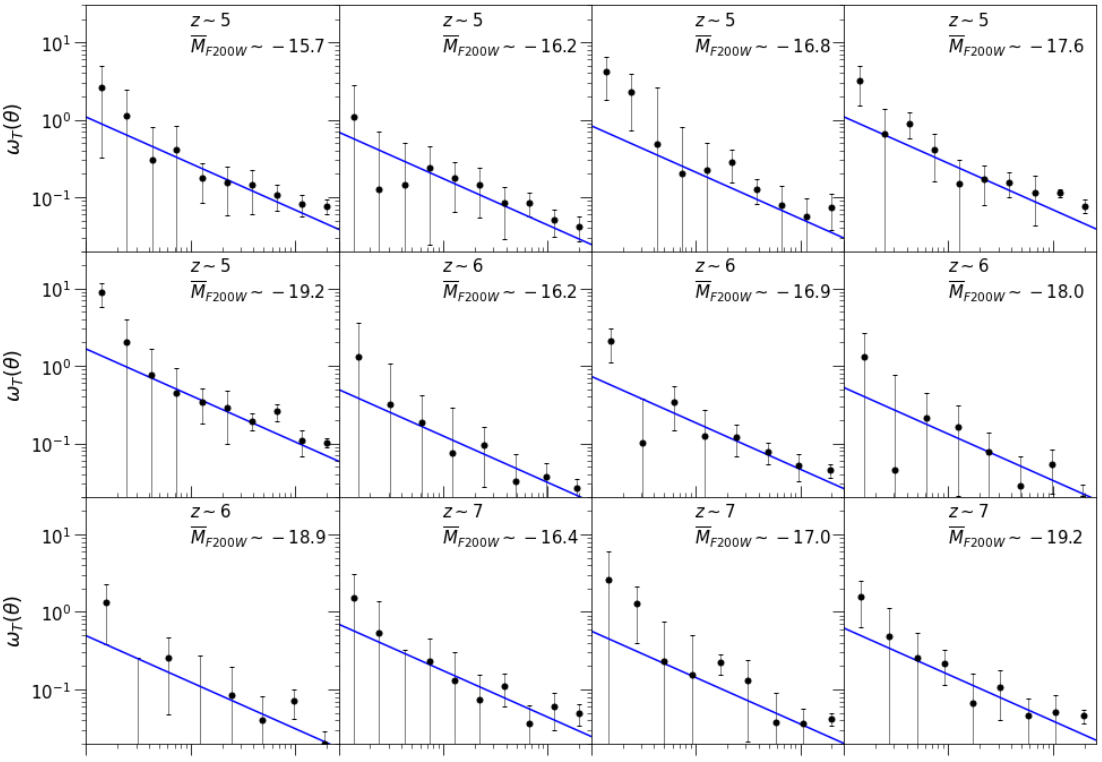

FIGURE 3.4: ACFs measured in the GOODS-S field for each subsample are depicted in Fig.3.3. Solid lines represent the best-fit power-law function (Eq.3.4) with $\beta = 0.6$ held constant as in the main analysis. The mean redshift and absolute magnitude are indicated in the upper right corner of each panel.

The fundamental equation for conducting abundance matching is:

$$\epsilon_{DC} \int_{M_h}^{+\infty} n(M_h, z) dM_h = \int_{L_g}^{+\infty} \Phi(L, z) dL. \tag{3.11}$$

Here, $\Phi(L, z)$ represents the Luminosity Function (LF) at redshift z, $n(M_h, z)$ stands for the Dark Matter Halo Mass Function (MF) Sheth & Tormen [246] and (M_h,L_g) are respectively the the minimum dark matter halo mass and the minimum luminosity of the galaxy. At a fixed value of L_g, Eq.3.11 leads to the determination of a one-to-one relationship between ϵ_{DC} and M_h that satisfies the equation constraint. The degeneracy between the two quantities can be resolved through clustering (bias) measurements, which independently define the relationship between L_g and M_h. This is because bias increases with halo mass and the luminosity of galaxies, as more massive halos tend to host more luminous galaxies, thus affecting the overall distribution of halos Sheth & Tormen [246]. When it comes to the luminosity function parameters, we employed the parameters that exhibit evolution with redshift, as presented in Bouwens et al. [36]. The obtained duty cycle values are presented in the last column of Tab.3.1.

In our findings, we observe a growing trend linked to the redshift of the sample, peaking at a maximum duty cycle value of $\epsilon_{DC} = 0.075^{+0.540}_{-0.067}$. This observation, observed at our

highest redshift of $z = 10.6$, aligns with the theoretical predictions outlined in Wyithe et al. [288]. The authors emphasized that for $z > 4$, a low duty cycle ($\epsilon_{DC} \lesssim 0.15$) is necessary to reconcile the observed star formation rate with the relatively low observed total stellar mass. Within the redshift range examined in our study, we find duty cycle values below this theoretical threshold, confirming the existence of a previously unidentified population of UV-faint galaxies for $z > 4$.

3.6 Summary and perspectives

In this study, we utilised the methodology outlined and discussed in Dalmasso et al. [66] to carry out galaxy clustering measurements well into the epoch of reionization using faint sources (M_{F200W}¡-17.3 at z=10.6) using JWST data from the guaranteed time observation program JWST Advanced Deep Extragalactic Survey (JADES)(Bunker et al. [47], Eisenstein et al. [84], Hainline et al. [114], Rieke et al. [224]) with the NIRCam instrument. Our main results and conclusions are as follows:

• We derived a galaxy bias measurement of $b = 9.6 \pm 1.7$ from a sample of $N_g = 55$ candidates at redshift $z = 10.6$ with a cut in absolute magnitude of $M_{F200W} < -17.3$.

• We studies self-consistently the clustering evolution from $5.0 \leq z < 10$. This allowed us to delineate the evolution of galaxy bias, revealing an increasing trend as anticipated by theory. Specifically, we report a value of $b = 2.1 \pm 0.6$ at $z = 5.5$ and $b = 7.2 \pm 1.5$ at redshift $z = 9.3$ (see Tab.3.1).

• We studied the relationship between galaxy bias and magnitude (Fig.3.3). Our results at the faintest magnitudes are in line with theoretical predictions out to $M_{F200W} \sim -15.5$ at $z \sim 5$, and they present an evolution consistent with previous studies carried out for brighter sources.

Through this study, we have demonstrated the viability of conducting clustering measurements on LBG samples observed during the epoch of reionization, even as the candidates become increasingly faint and rare, thanks to the power of JWST infrared instrumentation. As photometric catalogs extend to even fainter magnitudes and the pool of available candidates expands from the growing number of JWST observations, the initial results reported here have the potential to blossom into a quantitative dataset that can be compared to more sophisticated halo occupation models (e.g., see Harikane et al. [119]) to understand how dark matter halos transform gas into stars the earliest times in the Universe history.

Chapter 4

Accelerated evolution of galaxy host halo masses during Cosmic Dawn from deep JWST clustering

The material presented in this chapter is still a work in progress but has already been structured in the format of a scientific publication. The author led both the data analysis and writing, in collaboration with the co-authors. This version has been reformatted to match the thesis style. The following changes have been made:

- **Figures and tables have been adapted to display properly within the thesis layout**
- **Some bibliographic references have been updated**

4.1 Abstract

We present the deepest clustering analysis of early galaxies to date, analyzing $N_{\rm g} \simeq 6500$ photometrically-selected Lyman Break Galaxies from JWST's Advanced Deep Extragalactic Survey (JADES) to reveal how galaxies and dark matter evolved during cosmic dawn ($5 \leq z < 11$). Using halo occupation distribution (HOD) modeling of the two-point angular correlation function, we trace the galaxy-halo relationships across the first billion years of cosmic history. Our analysis reveals that galaxies at $z = 10.6$ reside in dark matter halos over an order of magnitude less massive ($M_{\rm h} \sim 10^{10.12} M_{\odot}$) than their counterparts at $z = 5.5$ ($M_{\rm h} \sim 10^{11.45} M_{\odot}$), while exhibiting correspondingly

higher effective bias values ($b_{\rm g}^{\rm eff} = 8.13^{+0.04}_{-0.02}$ compared to $5.64^{+0.10}_{-0.13}$). Correspondingly, the satellite galaxy fraction hints at a declining trend with decreasing redshift, reaching $< 1\%$ by $z \sim 5-6$. However, the significant systematic and random uncertainties in the data-model comparison prevent us from drawing robust conclusions on the evolution - if any - of the satellite fraction during the epoch of reionization. These results provide the first view of the coevolution between galaxies and dark matter evolved at redshift $\gtrsim 10$, offering additional and independent constraints on early galaxy formation models tuned to reproducing luminosity function evolution.

4.2 Introduction

Observations of unexpectedly luminous, massive galaxies at high-redshifts obtained with the James Webb Space Telescope (JWST, e.g., Arrabal Haro et al. [10], Boyett et al. [38], Carniani et al. [50], Labbé et al. [149], Naidu et al. [183], Napolitano et al. [186]) have highlighted some discrepancies with the predictions of the ΛCDM framework on galaxy formation at early times Padmanabhan & Loeb [198], or tuning of the baryonic processes (e.g., Boylan-Kolchin [39], Ferrara et al. [93], Gelli et al. [105], Mason et al. [177], Yung et al. [290]). These results suggest a more rapid assembly of baryons than the ΛCDM model typically predicts, indicating an early Universe with substantial galaxy masses. Wide-field JWST surveys, along with spectroscopic analyses of massive galaxy candidates, are essential for assessing this discrepancy and determining if these high-redshift galaxies align with ΛCDM expectations.

A fundamental approach for testing cosmological models involves investigating the relation between luminous galaxies and the dark matter halo that hosts them by observing the projected spatial arrangement of galaxies, as summarised by the two-point correlation function (e.g., Bahcall & Soneira [12], Davis & Peebles [71], Peebles [209]). Such analyses are based on galaxy clustering statistics, which explore the spatial distribution and correlations among galaxies. This, in turn, offers essential information regarding the large-scale structure of the Universe and the processes that drive galaxy formation and evolution Groth & Peebles [111].

A significant application of clustering analysis centers on Lyman Break Galaxies (LBGs), star-forming galaxies identifiable by their dust-unaffected rest-frame UV spectral energy distribution (SED). These galaxies offer insights into the relationship between the properties of galaxies and their host halos, probing essential aspects of galaxy evolution. LBGs exhibit strong clustering patterns, with correlation lengths similar to present-day bright spiral galaxies (e.g., Adelberger et al. [6], Giavalisco & Dickinson [106], Giavalisco et al. [107], Porciani & Giavalisco [216]). This clustering indicates that LBGs are biased

tracers of the mass density field, as they predominantly reside in massive dark matter halos, which display stronger clustering than less massive halos (e.g., Bardeen et al. [13], Mo & White [181], Sheth & Tormen [246]).

Previous studies have demonstrated a clear link between halo mass and galaxy properties, highlighted by the correlation between LBG clustering strength and UV luminosity. It has been found in past studies that brighter LBGs exhibit larger correlation lengths (e.g., Adelberger et al. [6], Giavalisco & Dickinson [106], Ouchi et al. [195]) implying that brighter galaxies are generally hosted within more massive halos.

Studying the angular correlation function (ACF), which quantifies the increased probability of finding a galaxy at a given angular separation relative to a random distribution, provides valuable insights into substructure and the halo occupation distribution (HOD). However, obtaining reliable measurements with high signal-to-noise (S/N) ratios across a broad luminosity range is crucial, as most satellite galaxies within dark matter halos are typically less luminous than their central counterparts Giavalisco & Dickinson [106], Hamana et al. [115].

Distinct galaxy populations, categorized by luminosity, mass, morphology, and intrinsic characteristics, demonstrate unique clustering patterns influenced by selection biases that favor particular halo masses. At fixed redshift, more massive halos generally exhibit enhanced clustering compared to their less massive counterparts, resulting in galaxies hosted in massive halos appearing more clustered than those in lower mass halos. Furthermore, biases inherent in clustering methodologies and the construction of the ACF affect clustering measurements therefore large samples of galaxies that cover a broad range of luminosities are essential for precise measurements of clustering segregation. It is also crucial to account for survey depth and completeness as disparities in observational coverage across different regions may lead to inflated clustering estimates at greater separations if a uniform distribution of galaxies is assumed as outlined in Dalmasso et al. [66]. These techniques have been employed to study a simple model of galaxy clustering at high redshift, based on a power-law fit to the ACF, using the first data release of JWST Advanced Deep Extragalactic Survey (JADES) Dalmasso et al. [67].

In this study, we present clustering measurements of LBGs at high redshift ($5 \leq z < 11$), based on Version 2.0 of the JADES data release, which includes approximately twice the number of photometrically identified high-redshift galaxies compared to previous releases. Crucially, this expanded sample enables us to extend HOD modeling to the faint end of the luminosity function ($M_{UV} \leq -17.0$) and to higher redshifts than previously accessible, thereby probing galaxy-halo connections in unexplored regions of parameter

space. We model the observed ACF using a HOD framework to disentangle the contributions from central and satellite galaxies to the clustering signal. We further investigate the evolution of the average halo occupation to constrain the satellite fraction at cosmic dawn and assess the consistency of these results with ΛCDM predictions for early galaxy formation.

The structure of this paper is outlined as follows: in Sec.4.3, we describe data reduction and the creation of the parent sample; in Sec.4.4, we present the ACF fits of the HOD model used in this study; in Sec.4.5, we discuss our results; and in Sec.4.6, we offer a summary of our key findings. We adopt the cosmological parameters determined by the Planck Collaboration Planck Collaboration et al. [214]: $(\omega_{\rm m}, \Omega_\lambda, h, \sigma_8) = (0.3075, 0.6925, 0.6774, 0.8159)$. Magnitudes are reported in the AB system Oke & Gunn [191].

4.3 Data sets and sample selection

4.3.1 JADES GOODS-South Data

In this study, we analyze imaging data from the public release targeting the Great Observatories Origins Deep Survey-South (GOODS-South) field, provided by the JADES[1] collaboration Bunker et al. [47], D'Eugenio et al. [65], Eisenstein et al. [85, 86], Hainline et al. [114], Rieke et al. [224]. This data set, captured by the NIRCam instrument, spans 67.7 arcminutes^2 across the Deep and Medium programs, utilizing nine filters (F090W, F115W, F150W, F200W, F277W, F335M, F356W, F410M, F444W) covering a wavelength range of $0.8 - 5.0\mu m$ with a spatial resolution of 0.03 arcsec/pixel. A detailed description of the JADES Version 2.0 data release, which includes the galaxy candidates studied here, can be found in Eisenstein et al. [86].

Our analysis focuses on galaxies within a redshift range of $5 \leq z < 11$, divided into six bins with a width of $\Delta z = 1.0$ with the resulting average redshifts $\overline{z} = 5.5$, 6.5, 7.4, 8.5, 9.3 and 10.6. We select candidates based on a signal-to-noise ratio (SNR) of at least 5.0 in the F200W NIRCam band, ensuring that they are confined within observable regions on both root mean square (rms) and segmentation maps to improve the accuracy of our results.

To reduce uncertainties caused by varying depths and multiple exposures in different regions, which could impact clustering estimates, we estimate the completeness of the

[1] https://archive.stsci.edu/hlsp/jades

z	$M_{\mathrm{UV,50\%}}$	$m_{\mathrm{AB,50\%}}$
5.5	−15.50	31.10
6.5	−15.50	31.37
7.5	−15.50	31.59
8.5	−15.50	31.78
9.5	−16.50	30.95
10.5	−19.67	27.93

TABLE 4.1: Magnitude limits of our galaxy samples, defined as the value where injection-recovery simulations achieve 50% source detection efficiency. Absolute magnitudes are in the rest-frame UV; apparent magnitudes are in F444W.

survey ($\mathscr{C}$). As in Dalmasso et al. [67], we focus on the faintest observable galaxies and apply a magnitude threshold of $M_{UV} < -17.0$ for all cataloged candidates.

Detection completeness was measured using the injection-recovery tool `GLACiAR2` Leethochawalit et al. [157] as a function of magnitude and redshift. Completeness decreases from ∼80% at bright magnitudes ($M_{\mathrm{UV}} < -18$) to ∼50% around $M_{\mathrm{UV}} \sim -15.5$ to -16.5 at intermediate redshifts ($z \sim 5$–9), with a notable shift to brighter absolute magnitudes at $z = 10.5$. The 50% completeness limits for each redshift bin are presented in Tab.4.1.

4.3.2 Random points catalog

To perform clustering analysis, it is essential to create a complementary catalog of simulated galaxies. These simulated galaxies play a key role in defining the ACF estimator, particularly in the calculation of RR(θ), the pair count of random points. These random points are generated to mimic the observational conditions, reflecting any selection biases in the parent sample.

Due to variations in observational depth, survey completeness is not uniform across the field. As a result, certain portions of the imaging field may exhibit an apparent excess or deficit of sources, potentially influenced by angular patterns correlated with the rms maps. To correct for these effects, we perform an artificial source injection and recovery process, generating random points within the ACF measurement area to model the completeness variation.

The random point catalog generation follows the recovery procedure detailed in Dalmasso et al. [66]. Briefly, we used the injection-recovery tool `GLACiAR2` Leethochawalit et al. [157] to inject galaxies into JADES images. These galaxies were distributed across redshift bins ranging from $z_{\mathrm{min}} = 4.5$ to $z_{\mathrm{max}} = 13.0$ in steps of 0.5, and UV magnitudes from $M_{\mathrm{min}} = -25$ to $M_{\mathrm{max}} = -13$ in 0.5 mag decrements. For each redshift-magnitude bin, we injected $N = 3200$ galaxies at random positions, corresponding to approximately

one galaxy per 30 arcsec^2. The injected galaxies were modeled with disk-like light profiles (Sersic index $n = 1$), random inclinations, and ellipticities. The spectral energy distributions (SEDs) were drawn randomly from the JAGUAR mock catalog Williams et al. [283], matching the redshift bin. Galaxy recovery followed the same process used for the real data Rieke et al. [224], utilizing a detection image constructed from the F227W, F335M, F356W, F410M, and F444W bands.

To finalize the random points catalog, we selected recovered candidates and employed a Monte Carlo hit-and-miss method. Each simulated galaxy in the random catalog was assigned a probability of detection, defined by:

$$\mathrm{p}(M) = \frac{\Phi(M)}{\Phi(M_{\mathrm{lim}})}, \tag{4.1}$$

where M is the galaxy's absolute magnitude, and M_{lim} is the limiting magnitude set at $M_{UV} = -17.0$. The luminosity function parameters evolve with redshift, following the Schechter profile described by Bouwens et al. [36].

4.4 Clustering Analysis

4.4.1 Angular Correlation Function (ACF) estimation

The ACF is characterized by the observable $\omega_{\mathrm{obs}}(\theta)$ as defined by Landy & Szalay [150]:

$$\omega_{\mathrm{obs}}(\theta) = \frac{\mathrm{DD}(\theta) - 2\mathrm{DR}(\theta) + \mathrm{RR}(\theta)}{\mathrm{RR}(\theta)}, \tag{4.2}$$

which measures the excess probability of finding pairs of objects in the parent sample at a given angular separation compared to a random distribution of candidates within the same survey area.

In Eq.4.2 the term $\mathrm{DD}(\theta)$ represents the number of pairs of galaxies within an angular separation range of $(\theta \pm \delta\theta)$ (data-data pairs), $\mathrm{DR}(\theta)$ denotes pairs formed by one observed galaxy and one randomly generated galaxy (data-random pairs), and $\mathrm{RR}(\theta)$ corresponds to the pair count from the randomly generated catalog (random-random pairs). In this study, we consider an angular range of $\theta \in (0.2, 850)$ arcsec, using a variable number of equally spaced angular bins.

The ACF, as defined above, is typically underestimated due to the finite size of the observational survey. To correct for this bias, we introduce a coefficient known as the

integral constraint (IC) Groth & Peebles [111], Peacock & Nicholson [205]:

$$\mathrm{IC} = \frac{\sum_i \mathrm{RR}(\theta_i)\omega_{\mathrm{model}}(\theta)}{\sum_i \mathrm{RR}(\theta_i)}, \tag{4.3}$$

where $\omega_{\mathrm{model}}(\theta)$ is the best-fit model ACF and $\mathrm{RR}(\theta_i)$ is the pair count of randomly generated galaxies within a specific angular bin. This coefficient is used to correct the observed ACF, linking the observed and true ACF measurements as follows:

$$\omega_{\mathrm{true}}(\theta) = \omega_{\mathrm{obs}}(\theta) + \mathrm{IC}. \tag{4.4}$$

To accurately account for this coefficient in the ACF estimation, we considered $\mathrm{RR}(\theta_i)$ up to the maximum angular separation covered by our survey area. This ensures that, at large angular separations, the IC approaches zero preventing any bias in our measurements. Notably, both the IC and $\omega_{\mathrm{model}}(\theta)$ are determined simultaneously during the model fitting process, as outlined in Sec.4.4.2.

To quantify the uncertainties in the ACF measurements, we construct the normalized covariance matrix using the standard estimator:

$$C_{ij} = \frac{1}{N_{\mathrm{boot}} - 1} \sum_{l=1}^{N_{\mathrm{boot}}} \left[\omega^l(\theta_i) - \overline{\omega}(\theta_i)\right] \left[\omega^l(\theta_j) - \overline{\omega}(\theta_j)\right], \tag{4.5}$$

where N_{boot} is the total number of resamplings performed, $\omega^l(\theta_i)$ is the cross-correlation function measured from each realization in the i-th bin, and $\overline{\omega}(\theta_i)$ is the mean of the cross-correlation function in the same bin. The uncertainty associated with each bin is taken to be the square root of the corresponding diagonal element in the covariance matrix.

4.4.2 Halo Occupation Distribution (HOD) model

To interpret galaxy clustering, we apply the HOD formalism, which describes how galaxies populate dark matter halos (Ma & Fry [168], Peacock & Smith [207], Seljak [243]). The underlying assumption of the model is that all dark matter halos are spherical with a density distribution that depends only on their mass. In addition, we assume that the average number of galaxies residing in each halo depends on the halo mass[2].

The galaxy population is divided between central and satellite galaxies. Central galaxies reside at the center of the dark matter halo, while satellites reside at the center of a

[2]The outcome of these assumptions on galaxy clustering have been well tested on both numerical simulations and observations in the local/low redshift universe.

sub-halo and orbit around central galaxies inside a larger host halo. The mean number of galaxies $\langle N \rangle$ residing in a halo of mass $M_{\rm h}$ is then the sum of the average of central and satellite galaxies,

$$\langle N(M_{\rm h}) \rangle = \langle N_{\rm c}(M_{\rm h}) \rangle + \langle N_{\rm s}(M_{\rm h}) \rangle. \tag{4.6}$$

We adopt expressions of the number of central galaxies $\langle N_{\rm c} \rangle$ and the number of satellite galaxies $\langle N_{\rm s} \rangle$ motivated by N-body and smoothed particle hydrodynamics simulations (e.g., Garel et al. [104], Kravtsov et al. [147], Zheng et al. [292]) and defined as

$$\langle N_{\rm c}(M_{\rm h}) \rangle = \frac{1}{2} \left[1 + {\rm erf} \left(\frac{\log {\rm M_h} - \log {\rm M_{min}}}{\sqrt{2} \sigma_{\log {\rm M_h}}} \right) \right] \tag{4.7}$$

and

$$\langle N_{\rm s}(M_{\rm h}) \rangle = \langle N_{\rm c}(M_{\rm h}) \rangle \left(\frac{M_{\rm h}}{M_{\rm sat}} \right)^{\alpha} , \tag{4.8}$$

respectively. The values of $M_{\rm min}$ and $\sigma_{\log M_{\rm h}}$ define the inflection point and width of the sigmoid function describing $\langle N_{\rm c} \rangle$. The values of $M_{\rm sat}$ and α represent the amplitude and slope of a power law factor that determines the average number of satellites. We fix $\sigma_{\log M_{\rm h}} = 0.2$ following previous studies (e.g., Conroy et al. [54], Harikane et al. [117], Kravtsov et al. [147], Zheng et al. [292]).

In this paper, we are interested in testing whether the angular clustering resulting from the HOD formalism is compatible with the distribution of galaxies over a large range of luminosities at $z \gtrsim 5$, and comparing the resulting halo-luminosity relation to theoretical predictions of the ΛCDM paradigm. In order to do so, we need to relate the average number of galaxies, determined by the values of $M_{\rm min}$, $M_{\rm sat}$ and α, to the measured ACF. This is achieved by estimating the ACF at a angular distance θ from the galaxy power spectrum $P_{\rm g}$ via the Limber approximation Bartelmann & Schneider [18] projected over a the normalized redshift distribution $\mathcal{N}(z)$ of the observed sample

$$\omega(\theta) = \int {\rm d}z \mathcal{N}(z) \frac{{\rm d}z}{{\rm d}r} \int {\rm d}k \frac{k}{2\pi} P_{\rm g}(k, z) J_0[\theta r(z) k] \, , \tag{4.9}$$

where $r(z)$ is the radial comoving distance and J_0 is the zeroth-order Bessel function of the first kind.

The galaxy power spectrum receive contributions by the 1-halo term and 2-halo term expressed as:

$$P_{\rm g}(k, z) = P_{\rm g}^{\rm 1h}(k, z) + P_{\rm g}^{\rm 2h}(k, z). \tag{4.10}$$

The contribution to the power spectrum coming from pairs of galaxies contained within the same halo is

$$P_{\rm g}^{\rm 1h}(k, z) = P_{\rm g}^{\rm cs}(k, z) + P_{\rm g}^{\rm ss}(k, z) \, , \tag{4.11}$$

and is split into the contributions of the central-satellite and satellite-satellite galaxies pairs, defined as

$$P_{\rm g}^{\rm cs}(k,z) = \frac{2}{n_{\rm g}^2}\int {\rm d}M_{\rm h}\langle N_{\rm c}N_{\rm s}(M_{\rm h})\rangle\frac{{\rm d}n}{{\rm d}M_{\rm h}}u(k,M_{\rm h},z) \tag{4.12}$$

and

$$P_{\rm g}^{\rm ss}(k,z) = \frac{1}{n_{\rm g}^2}\int {\rm d}M_{\rm h}\langle N_{\rm s}(N_{\rm s}-1)(M_{\rm h})\rangle\frac{{\rm d}n}{{\rm d}M_{\rm h}}u^2(k,M_{\rm h},z)\,, \tag{4.13}$$

respectively. On the other hand, the contributions to the galaxy power spectrum introduced by pairs residing in different halos is

$$P_{\rm g}^{\rm 2h}(k,z) = \left[\frac{1}{n_{\rm g}}\int {\rm d}M_{\rm h}\langle N(M_{\rm h})\rangle\frac{{\rm d}n}{{\rm d}M_{\rm h}}b_{\rm h}(M_{\rm h},z)u(k,M_{\rm h},z)\right]^2\,, \tag{4.14}$$

where $b_{\rm h}$ the halo bias factor from Tinker et al. [263] in the linear regime. Here, $u(k,M_{\rm h},z)$ is the Fourier transform of the dark matter halo NFW density profile normalized by its mass (e.g., Cooray & Sheth [62]), adopting the concentration parameter relation from Correa et al. [63]. Finally, $n_{\rm g}(z)$ is the galaxy number density obtained weighting the halo mass function from Behroozi et al. [19] on the average occupation number

$$n_{\rm g}(z) = \int {\rm d}M_{\rm h}\frac{{\rm d}n}{{\rm d}M_{\rm h}}\langle N(M_{\rm h})\rangle\,, \tag{4.15}$$

where the halo mass function $\frac{{\rm d}n}{{\rm d}M_{\rm h}}$ provides the comoving number density of halos as a function of halo mass and redshift, calibrated using N-body simulations and validated at high redshifts (Behroozi et al. [19], Tinker et al. [262]). The quantity $\bar{n}_{\rm g}$ is the integral of $n_{\rm g}(z)$ over comoving volume averaged over the redshift distribution of the sample

$$\overline{n}_{\rm g} = \frac{\int {\rm d}z\frac{{\rm d}V(r)}{{\rm d}z}\mathcal{N}(z)n_{\rm g}(z)}{\int {\rm d}z\frac{{\rm d}V(r)}{{\rm d}z}\mathcal{N}(z)}\,. \tag{4.16}$$

To analyze the relation between the distribution of galaxies within their host halos, we also define the following average quantities: the average satellite fraction

$$f_{\rm sat}(z) = \frac{1}{\overline{n}_{\rm g}}\int {\rm d}M_{\rm h}\frac{{\rm d}n}{{\rm d}M_{\rm h}}\langle N_{\rm s}(M_{\rm h})\rangle\,, \tag{4.17}$$

the effective bias

$$b_{\rm g}^{\rm eff}(z) = \frac{1}{\overline{n}_{\rm g}}\int {\rm d}M_{\rm h}\frac{{\rm d}n}{{\rm d}M_{\rm h}}\langle N(M_{\rm h})\rangle b_{\rm h}(M_{\rm h},z)\,, \tag{4.18}$$

and the average halo mass

$$\langle M_{\rm h}\rangle(z) = \frac{1}{\overline{n}_{\rm g}}\int {\rm d}M_{\rm h}\frac{{\rm d}n}{{\rm d}M_{\rm h}}M_{\rm h}\langle N(M_{\rm h})\rangle\,. \tag{4.19}$$

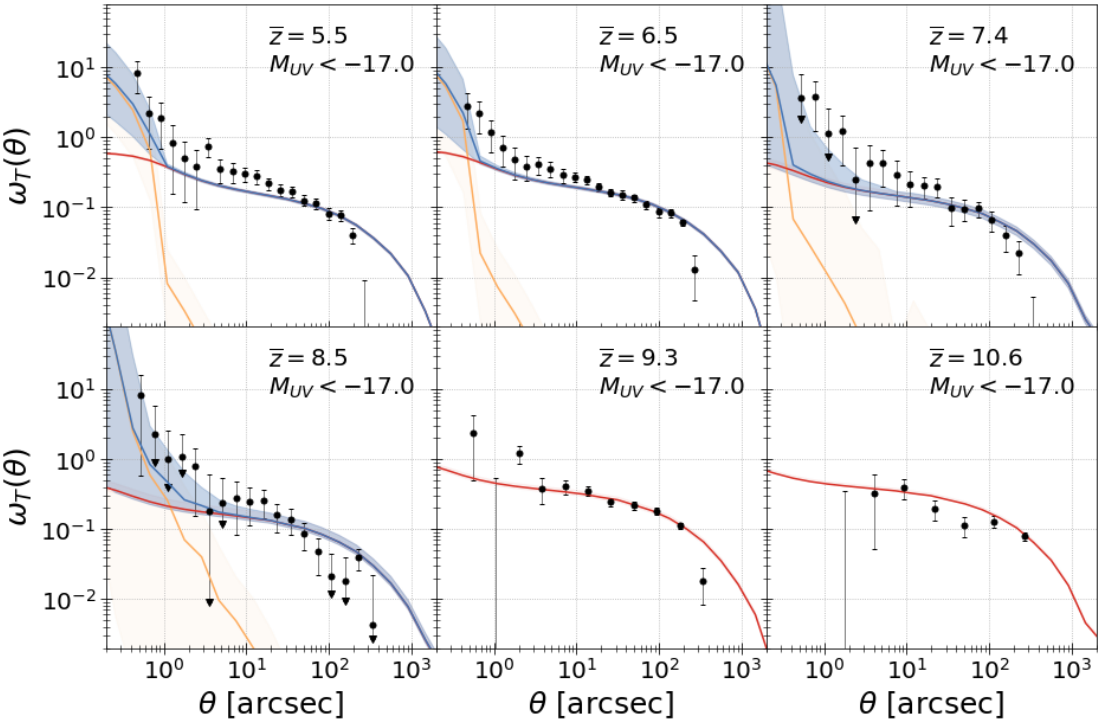

FIGURE 4.1: Measured ACFs in the GOODS-S field at six mean redshifts from bootstrap resampling with 1σ uncertainties derived from the covariance matrix. Blue line represent the HOD best-fit model ($\omega_{\rm mod}(\theta)$), with orange and red lines showing the one-halo ($\omega_{\rm 1h}$) and two-halo ($\omega_{\rm 2h}$) terms. Shaded areas indicate 1σ uncertainties. Mean redshift and absolute UV magnitude thresholds (F200W band) are shown in the upper right of each panel.

We expect these average quantities to be weakly dependent on the approximation of a linear evolution of the halo bias factor adopted in our HOD model. In fact, while the one-halo term, and therefore the satellite fraction, is primarily determined by the ACF below 5 arcsec and the two-halo term is well constrained by the observed ACF between 10^2 and 10^3 arcsec, the non-linear bias has only a percent-level effect on angular clustering at $z \ll 1$van den Bosch et al. [293], and contributes to two-halo term primarily around $\theta \sim 10$ arcsec at $3 < z < 5$ (Jose et al. [137, 138]). Nonetheless, the impact of the linear bias approximation on galaxy clustering remains untested at the high redshifts and low luminosities probed in this study. We defer to future work an in-depth comparison between our results and dedicated simulations of high-z galaxies that can fully account for non-linear bias effects.

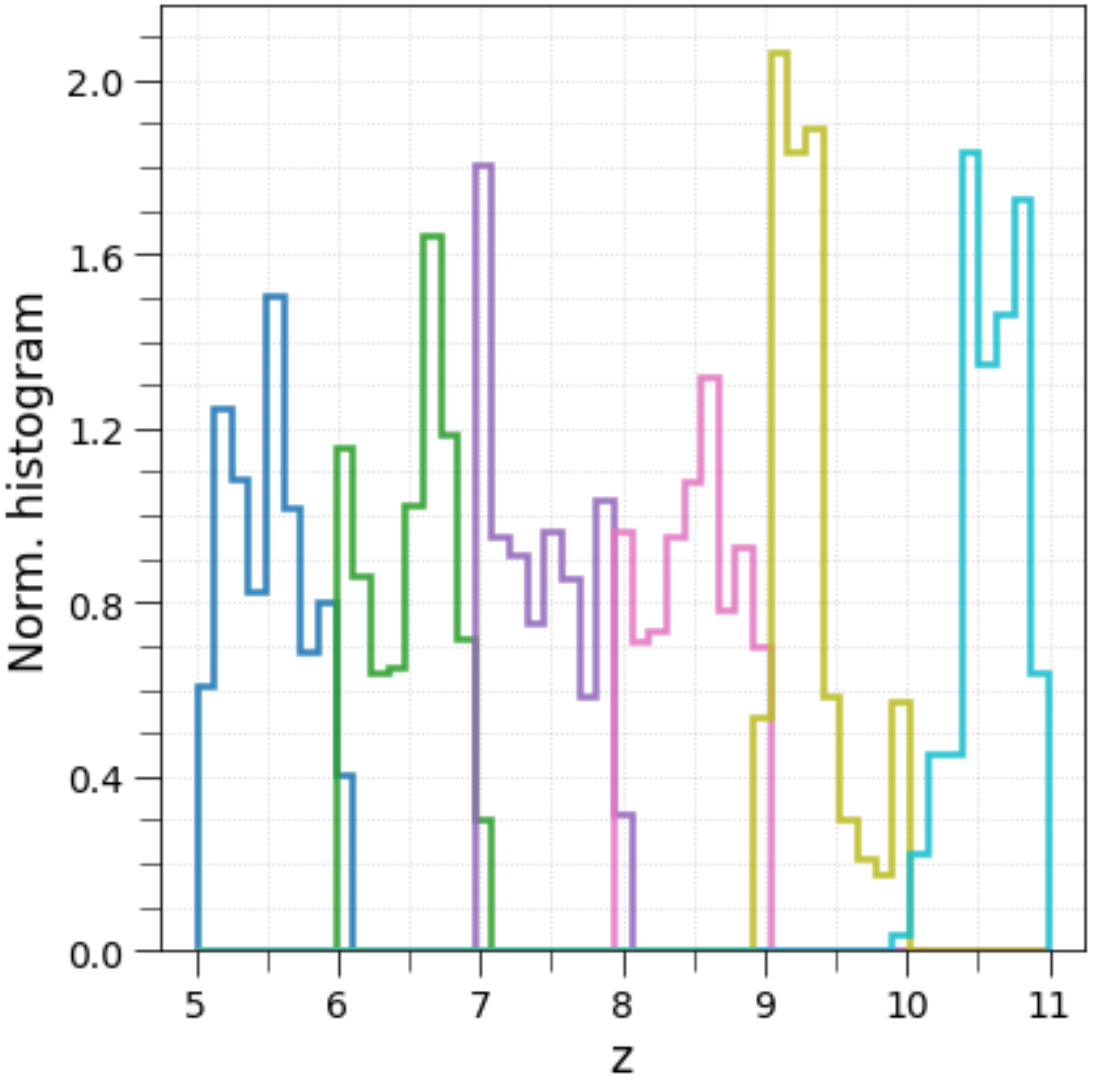

FIGURE 4.2: Normalized redshift distributions of the parent galaxy samples, with fine structure resulting from a combination of photometric redshift selection efficiency and sample variance in excess of Poisson noise due to clustering.

4.5 Dark Matter Halo Clustering estimations

4.5.1 HOD model fitting

To fit the ACF results with our HOD model, we employ a Markov Chain Monte Carlo (MCMC) method, allowing two parameters ($M_{\rm min}$, $M_{\rm sat}$) to vary freely during the process. Following findings from previous works we decided to fix $\alpha = 1$ for the analysis (i.e., Conroy et al. [54], Harikane et al. [117, 118], Ishikawa et al. [131], Kravtsov et al. [147], Zehavi et al. [291], Zheng et al. [292]).

HOD clustering measurements							
$\overline{z}$	$N_{\rm g}$	$< M_{UV} >$	$\log(M_{\rm min}/M_\odot)$	$\log(M_{\rm sat}/M_\odot)$	$\log(f_{\rm sat})$	$b_{\rm g}^{\rm eff}$	$\log(M_{\rm h}/M_\odot)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
5.5	1686	-17.90	$11.57^{+0.06}_{-0.08}$	$14.58^{+0.52}_{-0.36}$	$-3.13^{+0.32}_{-0.49}$	$5.64^{+0.10}_{-0.13}$	$11.45^{+0.02}_{-0.04}$
6.5	2937	-17.87	$11.16^{+0.05}_{-0.05}$	$14.18^{+0.86}_{-0.42}$	$-2.99^{+0.38}_{-0.82}$	$6.41^{+0.15}_{-0.15}$	$11.18^{+0.03}_{-0.04}$
7.4	909	-17.77	$10.61^{+0.14}_{-0.19}$	$12.82^{+2.14}_{-1.39}$	$-2.17^{+1.23}_{-2.05}$	$5.92^{+0.31}_{-0.36}$	$10.65^{+0.08}_{-0.10}$
8.5	712	-17.88	$10.05^{+0.24}_{-0.30}$	$10.72^{+3.42}_{-1.95}$	$-0.54^{+0.52}_{-3.17}$	$6.20^{+0.18}_{-0.15}$	$10.37^{+0.05}_{-0.02}$
9.3	459	-17.82	$8.95^{+0.64}_{-0.65}$	–	–	$6.95^{+0.13}_{-0.08}$	$10.23^{+0.02}_{-0.04}$
10.6	218	-17.81	$8.82^{+0.61}_{-0.56}$	–	–	$8.13^{+0.04}_{-0.02}$	$10.12^{+0.08}_{-0.06}$

TABLE 4.2: Summary of the clustering results from the HOD model best fit. Columns: (1) Mean redshift, (2) Number of galaxies, (3) Average UV absolute magnitude of the sample in the F200W band , (4) Best-fit value of $M_{\rm min}$, (5) Best-fit value of $M_{\rm sat}$, (6) Fraction of galaxy satellites in the dark matter halo (Eq.4.17), (7) Effective galaxy bias (Eq.4.18), (8) Mean dark matter halo mass $M_{\rm h}$ (Eq.4.19).

The likelihood function adopted is defined as:

$$\ln(\mathscr{L}) = -\frac{1}{2}\sum_{i,j}\left[\omega(\theta_i) - \omega_{\rm m}(\theta_i)\right] C_{ij}^{-1}\left[\omega(\theta_j) - \omega_{\rm m}(\theta_j)\right], \tag{4.20}$$

where $\omega(\theta)$ represents the corrected ACF measurements from Eq.4.4, $\omega_{\rm m}(\theta)$ are the model predictions from the HOD in Eq.4.9, and C_{ij}^{-1} is the inverse covariance matrix from Eq.4.5.

For the MCMC fitting, we applied flat priors of $\log(M_{\rm min}/M_\odot), \log(M_{\rm sat}/M_\odot) \in [8, 16]$ for both free parameters. We ran 32 walkers for up to 2000 steps, starting from literature-based initial values (i.e., Harikane et al. [117, 118], Paquereau et al. [199], Shuntov et al. [248, 249], Zheng et al. [292]), with convergence determined by $30\tau < N_{\rm iter}$ and $\Delta\tau/\tau < 15\%$ Paquereau et al. [199].

In Fig.4.1, we present a comparison between the measured ACFs of LBGs candidates in the GOODS-S field from JADES data release Version 2.0 and the ACFs predicted using our HOD formalism. The blue line represents our best-fit model based on the 50th percentile of the MCMC-derived parameters, while the shaded region encompasses the model predictions for the 16th and 84th percentile ranges of ($M_{\rm min}$, $M_{\rm sat}$). The one-halo term $\omega_{1h}(\theta)$ and two-halo term $\omega_{2h}(\theta)$ are represented by yellow and red solid lines with their associated uncertainty regions, respectively. Tab.4.2 summarizes the best-fit HOD parameters from our MCMC analysis along with three key clustering metrics: the satellite fraction $f_{\rm sat}$, effective galaxy bias $b_{\rm g}^{\rm eff}$, and dark matter halo mass $M_{\rm h}[M_\odot]$.

In all six panels of Fig.4.1 we observe a non-optimal fitting of the correlation function by the MCMC analysis with the implemented HOD model. These systematic discrepancies

show under-prediction at intermediate scales ($\theta \sim 3-10$ arcsec) and over-prediction at large scales ($\theta \gtrsim 100$ arcsec). We attribute the data-modeling mismatch to a combination of intrinsic limitations of the HOD modeling and of specific assumptions in our approach, such as fixing a subset of the free parameters due to the limited number of datapoints available at these high redshifts. Furthermore, an element that may impact the fit quality is the adoption of a linear evolution for the halo bias factor, following Tinker et al. [263]. This linear approximation is thought to be well justified in the low to intermediate redshift regime $z < 5$ for relatively bright sources. However, its impact on clustering studies during the epoch of reionization, and for the low-luminosity (sub L_*) galaxies observed by JWST remains untested. Therefore this may contribute significantly to the observed model-data discrepancies and will warrant detailed further investigations. Importantly, the quality of the fit in Fig.4.1 is not unique to our analysis but represents a common issue in HOD modeling of high-redshift galaxy clustering. In fact, similar systematic deviations between model predictions and observations are evident in multiple studies at these epochs including Harikane et al. [117, 118], Paquereau et al. [199] and Shuntov et al. [249]. These examples demonstrate that limited data points and inherent model assumptions produce comparable fitting difficulties across independent analyses of high-z galaxies.

Furthermore, for the highest redshift samples considered here ($z > 9$, last two panels in Fig.4.1), there were not enough galaxies in the sample to measure the two point correlation function at small angular separations. Therefore, we performed the fit using only the two-halo term $\omega_{2h}(\theta)$, constraining solely $M_{\rm min}$.

Our analysis utilizes a sample of LBGs spanning a wide redshift range during cosmic dawn ($5 \leq z < 11$) that probes the faint end of the luminosity function from $M_{\rm UV} < -17.0$ at various redshift cf. normalized redshift distribution in Fig.4.2). While direct one-to-one comparisons with existing literature are challenging due to our unique sample characteristics, we can identify general evolutionary trends in the HOD parameters.

As shown in Tab.4.2 the free parameters in our HOD model evolve with redshift within our measurement uncertainties. Specifically our findings align with established trends in the literature, we observe $M_{\rm min}$ decreasing with increasing redshift, consistent with previous studies by Hamana et al. [115], Conroy et al. [54], and Harikane et al. [117, 118]. Similarly, $M_{\rm sat}$ shows a decreasing trend with redshift, in agreement with Hamana et al. [115] and Conroy et al. [54].

4.5.2 Satellite fraction, Effective bias and Dark Matter Halo mass

Fig.4.3 presents a comparison of the clustering parameters $f_{\rm sat}$, $b_{\rm g}^{\rm eff}$, and $M_{\rm h}[M_\odot]$ derived from our HOD estimations with results from previous studies employing HOD modelling (Bhowmick et al. [28], Harikane et al. [117, 118], Paquereau et al. [199], Shuntov et al. [249]).

The effective galaxy bias measurements demonstrate a clear increasing trend with redshift, rising from $b_{\rm g}^{\rm eff} = 5.64^{+0.10}_{-0.13}$ at $z = 5.5$ to $b_{\rm g}^{\rm eff} = 8.13^{+0.04}_{-0.02}$ at $z = 10.6$. We note an apparent local dip at $z = 7.4$ where $b_{\rm g}^{\rm eff} = 5.92^{+0.31}_{-0.36}$, but this measurement is consistent with the neighboring redshift bins within $\sim 1.5 - 2\sigma$, likely reflecting statistical scatter due to cosmic variance and the systematic uncertainties in HOD modeling discussed above. The overall increasing trend is in agreement with theoretical predictions from hierarchical structure formation models Mo & White [181], Sheth & Tormen [246], which predict that galaxies of a given luminosity should reside in increasingly biased (rarer) regions of the cosmic density field at higher redshifts.

Our measurements are broadly consistent with previous observational studies employing HOD modeling Harikane et al. [117, 118], Paquereau et al. [199], Shuntov et al. [249], though we note that our bias values for galaxies with $M_{\rm UV} < -17.0$ appear somewhat higher than those reported for brighter samples at similar redshifts. For instance, Harikane et al. [117] and Shuntov et al. [249] find lower bias for galaxies with $M_{\rm UV} < -19.5$ and $M_{\rm UV} < -19.1$, respectively. However, multiple factors complicate direct comparison. First, while Harikane et al. [117] reports clustering measurements up to $z \sim 7$, overlapping with the lowest redshift bins of our work, their strongest constraints are obtained at $z = 4$–6. Since galaxy bias changes significantly with redshift and luminosity at $z > 5$, even small offsets in the effective redshift and/or luminosity distribution of the samples can lead to appreciable differences in the derived bias values. Additionally, at the faint end of the luminosity function probed by our sample, the relationship between UV luminosity and halo mass may differ from expectations based on brighter galaxies, particularly if star formation efficiency or dust obscuration varies systematically with halo mass at high redshift. A comprehensive understanding of the luminosity-halo mass relation across the full range of UV magnitudes at $z > 6$ will require larger survey volumes and joint constraints from clustering and abundance matching, confirming that high-redshift galaxy populations trace progressively more biased environments as we probe earlier cosmic epochs.

For completeness, we also applied a standard power-law fit to the ACFs analyzed in this work, using the same methodology adopted in previous non-HOD clustering studies. This analysis yields lower effective bias values, ranging from $b_g = 3.15 \pm 0.21$ at $z = 5.5$

to $b_g = 8.14 \pm 1.13$ at $z = 9.3$, consistent with earlier power-law–based measurements (i.e. Barone-Nugent et al. [17], Dalmasso et al. [66], Dalmasso et al. [67], Harikane et al. [117], Overzier et al. [197], Qiu et al. [219]). Therefore, the offset between the power-law–derived galaxy bias and the HOD-derived effective galaxy bias reflects the known systematic differences between the two estimators rather than any discrepancy in the underlying data.

Our analysis also shows a significant decrease of the characteristic dark matter halo mass $M_{\rm h}[M_\odot]$ with redshift. The halo mass decreases by more than one order of magnitude from $\log_{10}(M_{\rm h}/M_\odot) = 11.45^{+0.02}_{-0.04}$ during the final stages of cosmic reionization at $z = 5.5$ to $\log_{10}(M_{\rm h}/M_\odot) = 10.12^{+0.08}_{-0.06}$ at $z = 10.6$, during cosmic dawn.

In this scenario, galaxies at high redshift reside in progressively less massive dark matter halos compared to similar galaxy populations at lower redshift. Correspondingly, the light-to-mass ratio (relative to the dark matter halo mass) increases by approximately an order of magnitude between $z \sim 6$ and $z \sim 10$ considering that our sample has a consistent selection based on a luminosity-limit to $M_{UV} < -17.0$ and that the average luminosity has limited evolution across redshift. This clearly indicates enhanced star formation efficiency, as the alternative of reduced halo occupation at earlier cosmic epochs is unlikely to apply.

While this study does not aim to provide a physical explanation for this evolution, it offers robust clustering-based measurements that are independent of luminosity function analyses, complementing the trends reported by Hamana et al. [115], Conroy et al. [54], Harikane et al. [117, 118] and Shuntov et al. [249]. These trends are consistent with hierarchical structure formation, where the galaxy-halo relationship evolves as cosmic structure assembles over time.

The redshift evolution of the satellite galaxy fraction $f_{\rm sat}$ (top panel of Fig.4.3) reveals changes in galaxy clustering properties throughout the early Universe. At the final stages of cosmic reionization ($z \sim 5-6$), our HOD modeling indicates that satellite galaxies constitute less than 1% of the total galaxy population across UV magnitudes down to $M_{\rm UV} \simeq -17.0$.

At the highest redshift probed in our sample ($z \sim 7-8$), our measurements suggest an enhancement in satellite populations, with $f_{\rm sat}$ reaching values around $\sim 28\%$, though we caution that these measurements carry substantial uncertainties. This potential increase would suggest a significant growth in the number of satellite galaxies hosted within individual dark matter halos during the period when the Universe was still (partially) neutral.

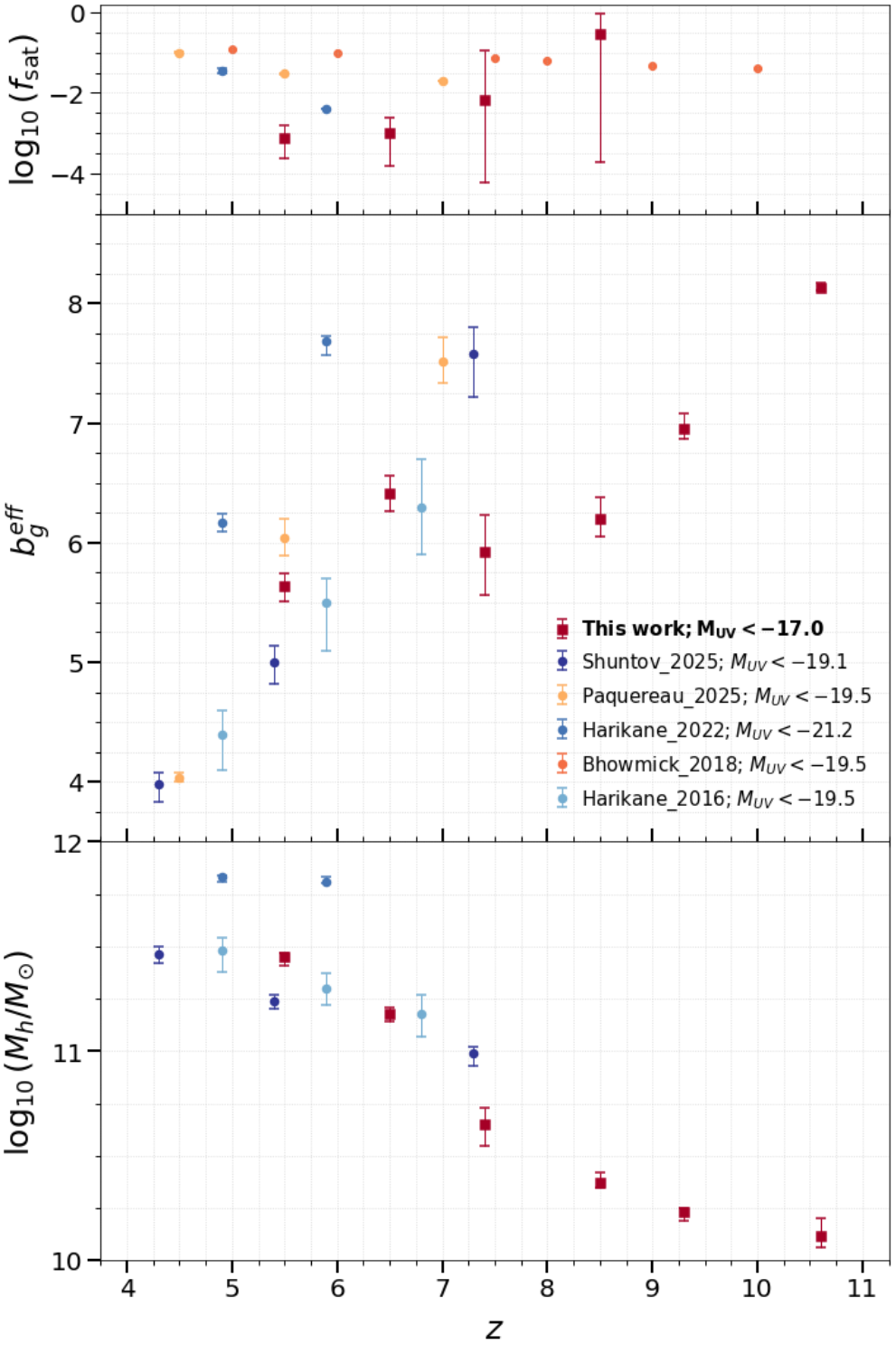

FIGURE 4.3: Evolution of clustering parameters as a function of redshift: satellite fraction $f_{\rm sat}$ (top panel), effective galaxy bias $b_{\rm g}^{\rm eff}$ (middle panel), and characteristic halo mass $M_{\rm h}[M_\odot]$ (bottom panel). Results from this work are shown as solid red squares, while previous HOD modelling studies are color-coded according to the legend (Bhowmick et al. [28], Harikane et al. [117, 118], Paquereau et al. [199], Shuntov et al. [249]).

Interestingly, this trend differs from the monotonic predictions of Bhowmick et al. [28], where the BlueTides simulation predicted consistently low satellite fractions during the early stages of reionization. The elevated measurements at $z \sim 7-8$, if confirmed, may reflect the interplay between structure formation and reionization feedback processes not fully captured in current theoretical models. However, we note that our analysis assumes the continued validity of the HOD framework at these high redshifts, where the significantly higher cosmic density and ongoing reionization may introduce systematic effects not accounted for in standard HOD parameterizations.

Should this evolution be confirmed by future studies, it would be well compatible with a boosted efficiency of early star formation in low-mass halos These galaxies would cluster in dense, multi-galaxy systems during cosmic dawn before environmental processes reshaped the galaxy population.

4.6 Summary and Perspectives

In this study we analyzed the clustering of $N_{\rm g} \simeq 6500$ LBGs at redshifts $5 \leq z < 11$ from the GOODS-South survey conducted by the JADES Collaboration Bunker et al. [47], D'Eugenio et al. [65], Eisenstein et al. [85, 86], Hainline et al. [114], Rieke et al. [224] by adopting a HOD formalism. Leveraging JWST's unprecedented sensitivity, this analysis extends current knowledge on the interplay between galaxies and dark matter halos beyond bright sources to the faint end of the luminosity function and to higher redshifts than previously accessible.

Our major results and conclusions are summarized as follows:

• We found that the redshift evolution of the free HOD parameters is consistent with previous studies with analogous samples. Specifically, both $M_{\rm min}[M_\odot]$ and $M_{\rm sat}[M_\odot]$ show a decreasing trend with redshift, suggesting that high-redshift galaxies are hosted by significantly less massive dark matter halos, as also indicated by the $M_{\rm h}[M_\odot]$ measurements (e.g.,Conroy et al. [54], Hamana et al. [115], Harikane et al. [117, 118], Paquereau et al. [199], Shuntov et al. [249]).

• We studied high-redshift LBGs in previously unexplored regions of the luminosity-redshift parameter space, obtaining new results on galaxy clustering measurements and presenting the galaxy bias evolution from redshift $z = 5.5$ with $b_{\rm g}^{\rm eff} = 5.64^{+0.10}_{-0.13}$ to $z = 10.6$ with a galaxy bias measurement of $b_{\rm g}^{\rm eff} = 8.13^{+0.04}_{-0.02}$, supporting previous results that show an increasing trend in galaxy bias with redshift (e.g.,Barone-Nugent et al. [17], Dalmasso et al. [66], Dalmasso et al. [67], Harikane et al. [118], Overzier et al. [197], Paquereau et al. [199], Park et al. [202], Qiu et al. [219], Sheth & Tormen [246], Shuntov

et al. [249]). Notably, our measurement at $z = 10.6$ represents the highest redshift constraint to date for LBG clustering.

• We measured rapid change with redshift of the characteristic dark matter halo mass $M_{\rm h}$, decreasing by more than one order of magnitude from $\log_{10}(M_{\rm h}/M_{\odot}) = 11.45^{+0.02}_{-0.04}$ at $z = 5.5$ to $\log_{10}(M_{\rm h}/M_{\odot}) = 10.12^{+0.08}_{-0.06}$ at $z = 10.6$. This evolutionary trend shows that galaxy populations with similar average UV magnitudes are hosted by increasingly less massive dark matter halos at higher redshifts, with the stellar light-to-halo mass ratio rising by roughly a factor of ten from $z \sim 6$ to $z \sim 10$.

• We found an increasing evolution of the galaxy satellite fraction $f_{\rm sat}$ with respect to redshift. While satellite galaxies comprise less than 1% of the population at $z \sim 5{-}6$, our measurements suggest this fraction may increase during reionization, reaching values as high as $\sim 28\%$ at $z \sim 8.5$, though with substantial uncertainties at the highest redshifts. If confirmed, this would suggest highly efficient galaxy formation in dense environments during cosmic dawn, with early galaxies preferentially forming in multi-galaxy systems before environmental processes reduced satellite fractions at later times.

By probing the relationship between galaxies and dark matter across cosmic dawn, this work provides new insights into the physical processes that may have governed the earliest phases of galaxy assembly, the potential role of environmental effects in shaping primordial galaxy populations, and the transition from the initial conditions of structure formation to the mature cosmic web we observe today.

Chapter 5

The rate and contribution of mergers to mass assembly from NIRCam observations of galaxy candidates up to 13.3 billion years ago

This chapter is a reproduction of Dalmasso et al. [68]. The author lead both the data analysis and writing with the collaboration of the co-authors. This reproduction has been reformatted to fit the style of the thesis, changes include:

- **Figures and tables have been formatted to display correctly on the thesis**
- **Some bibliographic records have been updated**

5.1 Abstract

We present an analysis of the galaxy merger rate in the redshift range $4.0 < z < 9.0$ (i.e. about 1.5 to 0.5 Gyr after the Big Bang) based on visually identified galaxy mergers from morphological parameter analysis. Our dataset is based on high-resolution NIRCam JWST data (a combination of F150W and F2000W broad-band filters) in the low-to-moderate magnification ($\mu < 2$) regions of the Abell 2744 cluster field. From a parent set of 675 galaxies ($M_U \in [-26.6, -17.9]$), we identify 64 merger candidates from

the Gini, M_{20} and Asymmetry morphological parameters, leading to a merger fraction $f_m = 0.11 \pm 0.04$. There is no evidence of redshift evolution of f_m even at the highest redshift considered, thus extending well into the epoch of reionization the constant trend seen previously at $z \lesssim 6$. Furthermore, we investigate any potential redshift dependent differences in the specific star formation rates between mergers and non-mergers. Our analysis reveals no significant correlation in this regard, with deviations in the studied redshift range typically falling within $(1 - 1.5)\sigma$ from the null hypotesis that can be attributed to sample variance and measurement errors. Finally, we also demonstrate that the classification of a merging system is robust with respect to the observed (and equivalently rest-frame) wavelength of the high-quality JWST broad-band images used. This preliminary study highlights the potential for progress in quantifying galaxy assembly through mergers during the epoch of reionization, with significant sample size growth expected from upcoming large JWST infrared imaging datasets.

5.2 Introduction

Interactions between galaxies critically influence their physical properties and play a core role in their mass accumulation and growth, driving morphological transformations and often resulting in mergers (Toomre & Toomre [265]). The role of interactions and mergers is pivotal to our understanding of galaxy evolution, including how their impact may change over cosmic time. One key measurement is the fraction of the galaxy population undergoing a merger, denoted as merger fraction (f_m), and whether there is any dependence with redshift or on galaxy properties, such as stellar mass or luminosity.
Direct imaging studies of the galaxy merger fraction broadly fall into two main methodologies; (i) examining the frequency of galaxies in close proximity to estimate upcoming mergers ("close pair identification method", e.g., Barnes [15], Le Fèvre et al. [153], Patton et al. [204], Schmidt et al. [240]) and (ii) measuring the morphology of individual sources to identify ongoing/recently completed mergers ("morphological parameters method", e.g., Conselice & Arnold [56], Conselice et al. [58, 59], Lotz et al. [165]).
Over the last two decades, investigations beyond the local Universe have predominantly utilized Hubble Space Telescope (HST) imaging. Lotz et al. [166] used the All-Wavelength Extended Groth Strip International Survey (AEGIS) to study galaxy morphologies and merger frequencies over the redshift range $0.2 < z < 1.2$ and determined a merger fraction consistent with no redshift evolution; measuring $f_m = 0.10 \pm 0.02$ using a parent sample with $I_{F814W} < 25.0$. This was supported by Conselice et al. [59], who identified a consistent merger fraction of $f_m \sim 0.10$ for galaxies between redshifts $0.6 < z < 1.2$ and with stellar mass $M_* > 10^{10} M_\odot$, using both Extended Groth Strip (EGS) (Davis et al. [72]) and The Cosmic Evolution Survey (COSMOS) (Scoville et al.

[242]) surveys.

In contrast, other studies focusing on massive galaxies ($M_* > 10^{10} M_\odot$) lying on the star-forming main sequence (MS) at $0.2 < z < 2.0$ identified an increase in the merger fraction with increasing redshift; from $f_m \sim 0.03$ at $z \sim 0.5$ to $f_m \sim 0.13$ at $z \sim 1.5$ Cibinel et al. [52], Speagle et al. [255]. Additionally, these investigations found a dependence of the merger fraction on galaxy properties, with the likelihood of galaxies merging depending significantly on their distance from the MS. Analyzing massive galaxies with $M_* \geq 10^{10.8} M_\odot$ selected from the Ks-band catalogs of UltraVISTA/COSMOS and CANDELS/3DHST, Man et al. [172] observed an increasing trend in the merger fraction with redshift for $z < 0.65$. The merger fraction then stabilizes at approximately $f_m \sim 0.10$ up to a redshift of $z \sim 2.4$.

Kim et al. [144] also determined a redshift evolution in the merger fraction in a mass-selected galaxy sample ($9.0 < \log(M_*/M_\odot) < 11.5$) in the North Ecliptic Pole-Wide field using the morphological parameters of the candidates to detect mergers. Within their sample, they obtained a merger fraction of $f_m \in (0.10, 0.20)$ at $z < 0.6$, consistent with Conselice et al. [59], with a marginally higher fraction of $f_m \sim 0.20$ for galaxies in the redshift range $4.0 < z < 6.0$, consistent with the high mass sample ($10^9 M_\odot$) from Conselice & Arnold [56] at the same epoch and with the same merger identification method.

Moreover, Ventou et al. [273, 274] examined the evolution of the merger fraction out to $z = 6$ using the close pair identification method with data collected from MUSE (the Multi-Unit Spectroscopic Explorer) and determined that the evolution of the merger fraction is best described by two distinct behaviors. In the local Universe, they observed a significant decline in the merger fraction with cosmic time reaching the lower limit of $f_m = 0.009 \pm 0.02$ at redshift $z = 0.05$ De Propris et al. [73]. However, for redshift greater than $z \sim 1$, the major merger fraction remained approximately constant at a value of around $f_m \sim 0.1$. This fraction is determined by including the total number of galaxies in the sample, the number of close pairs consistent with major mergers (i.e. 1:6 mass ratio between secondary and primary galaxies) and those without companions due to spatial resolution limits (Ventou et al. [273]). These findings are not only consistent with observational measurements, such as those in Conselice et al. [59], but also with simulations, including ILLUSTRIS Snyder et al. [253] and EMERGE O'Leary et al. [188], especially when considering a sample of galaxies with a stellar mass $M_* \geq 10^{9.5} M_\odot$. Importantly, none of these studies suggest an increase in the merger fraction beyond $z > 2$.

In contrast, a result indicating a higher merger fraction in the redshift range $4.4 \leq z < 4.6$ and $5.1 \leq z < 5.9$ for primary galaxies in the mass range of $9 \lesssim \log(M_*/M_\odot) \lesssim 11$ was found in Romano et al. [234]. The study involved analyzing the [CII] emission line from a sample of 118 star-forming galaxies to obtain morpho-kinematic information

from the ALMA Large Program to INvestigate [CII] at Early times (ALPINE) survey (Béthermin et al. [25], Faisst et al. [88], Le Fèvre et al. [154]). A major merger fraction of $f_m = 0.44^{+0.11}_{-0.16}$ at $z \sim 4.5$ and $f_m = 0.34^{+0.10}_{-0.13}$ at $z \sim 5.5$ is reported, consistent with the morphological study conducted by Conselice & Arnold [56], which found a merger fraction of $f_m = 0.23 \pm 0.05$ at $z = 3.8$ and $f_m = 0.19^{+0.05}_{-0.06}$ at $z = 4.8$ for galaxies with a stellar mass $M_* > 10^{9-10} M_\odot$.

Significantly, investigations into the merger fraction using conventional methodologies, primarily relying on ground-based observatories and the Hubble Space Telescope, have been confined to redshifts $z < 6$ due to the limited availability of high-resolution deep near-infrared (NIR) imaging and wavelength filters crucial for directly examining the rest-optical structures of high-redshift galaxies. However, with the advent of the James Webb Space Telescope (JWST), we can now expand the study of morphological characteristics and parameters of galaxies to higher redshifts, leveraging the enhanced sensitivity of the Near Infrared Camera (NIRCam) instrument. Recent studies have shown that the morphological parameters and the overall morphology of galaxies at the time of reionization ($z > 7$) do not vary significantly between optical and UV rest-frame observations (e.g., Tohill et al. [264], Treu et al. [270], Vulcani et al. [277]). This has been made possible thanks to the enhanced resolution and broader wavelength range provided by the JWST NIRCam observational filters compared to those of the HST.

Notably in recent years, the research to understand galaxy mergers has transcended standard methodologies of identifying mergers by employing random forest (RF) and Machine Learning classifications trained on simulated JWST images from CEERS and the IllustrisTNG simulation, achieving accuracies of approximately 60% Rodriguez-Gomez et al. [233], Rose et al. [235], Snyder et al. [254]. These alternative methods have been utilized to measure the merger fraction and observe redshift evolution, revealing a transition from $f_m \sim 0.03$ at $z \sim 0.7$ to $f_m \sim 0.40$ at $z \sim 3.7$ Rose et al. [235]. This highlights significant deviations from studies using standard methods and underscores that the identification method of galaxy mergers at high redshifts holds the potential to introduce substantial systematic effects.

In this study we conduct a comprehensive study of galaxy mergers in the epoch of reionization by leveraging the deep NIRCam imaging from the GLASS-JWST ERS program (Treu et al. [269]), UNCOVER JWST-GO-2561 (Bezanson et al. [26]) and DDT-2756 (PI Wenlei Chen). These new observations provide us with the opportunity to extend the morphological studies out to redshift $z \sim 9$ (lookback time of about 13.3 Gyr), and offer a more extensive dataset in comparison to prior studies that relied on HST data. Our study focuses on the low-magnification regions in the outskirts of galaxy cluster Abell 2744, where we employ morphological statistical parameters to assess the prevalence of merger systems across a wide range of redshifts. Our objective is to extend the investigation of galaxy mergers to the high redshift, a domain that has not been

NIRCam observations				
$z-bin$	N_g	$\overline{z}$	M_U	$\log_{10}(M_*/M_\odot)$
(1)	(2)	(3)	(4)	(5)
$4.0 \leq z < 5.0$	845(**319**)	4.50	$[-26.6, -17.9]$	$[6.9, 11.4]$
$5.0 \leq z < 6.0$	1333(**186**)	5.34	$[-24.7, -18.3]$	$[7.2, 10.9]$
$6.0 \leq z < 7.0$	386(**107**)	6.47	$[-23.1, -18.5]$	$[7.3, 9.9]$
$7.0 \leq z < 8.0$	414(**30**)	7.47	$[-23.6, -18.7]$	$[7.6, 10.7]$
$8.0 \leq z < 9.0$	250(**33**)	8.50	$[-25.6, -19.3]$	$[7.5, 12.4]$

TABLE 5.1: Summary of the galaxies used in this work for mergers identification, obtained using JWST NIRCam data from the GLASS Collaboration. The sources are selected as discussed in Sec.5.3. (1) Redshift bin. (2) Number of galaxies detected.. (3) Mean redshift of the subsample. (4) Absolute U Magnitude range of the subsample. (5) Stellar mass range. Quantities in columns (3), (4) and (5) are referred to the subsamples used in this work in **bold** (2). Both distributions of magnitude and stellar mass in respect to redshift are presesented in Fig. 5.5.

thoroughly explored before, encompassing both bright and faint galaxies out to $z \sim 9$. The structure of this chapter is organized as follows: Sec.5.3 provides a comprehensive overview of the data selection process for the galaxy candidates. In Sec.5.4, we outline the galaxy morphological parameters and merger criteria that we will employ to identify interacting systems. In Sec.5.5 we determine the merger fraction and then examine and discuss any redshift evolution within our sample and whether there are any galaxy characteristics differences between mergers and non-mergers (e.g. the sSFR). Finally, in Sec.5.6, we provide a summary of our discoveries. In App.5.5.3, we present a study exploring the influence of observation filter wavelengths on the analysis of morphological parameters.

In this work we assume, when relevant, a standard cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}Mpc^{-1}}$, $\Omega_\mathrm{m} = 0.3$, $\Omega_\Lambda = 0.7$. and a Kroupa et al. [148] Initial Mass Function (IMF). Magnitudes are in the AB system Oke & Gunn [191].

5.3 Observational data and sample selection

The NIRCam imaging data used in this study were obtained through three public programs focused on the foreground galaxy cluster Abell 2744 and its immediate surroundings: (i) GLASS JWST-ERS-1324 Treu et al. [269], (ii) UNCOVER JWST-GO-2561 Bezanson et al. [26], and (iii) the Director's Discretionary Time Program 2756 (PI Wenlei Chen). We use the publicly available galaxy and photometric catalogs for the combined footprint of these three programs, as provided by Merlin et al. [179], Paris et al. [201] with the photometric redshifts estimated by fitting the full HST+JWST photometry to Bruzual & Charlot [45] templates by means of the ZPHOT code (Fontana et al. [97]), using the same technique adopted in Santini et al. [236]. Additionally, we utilize the

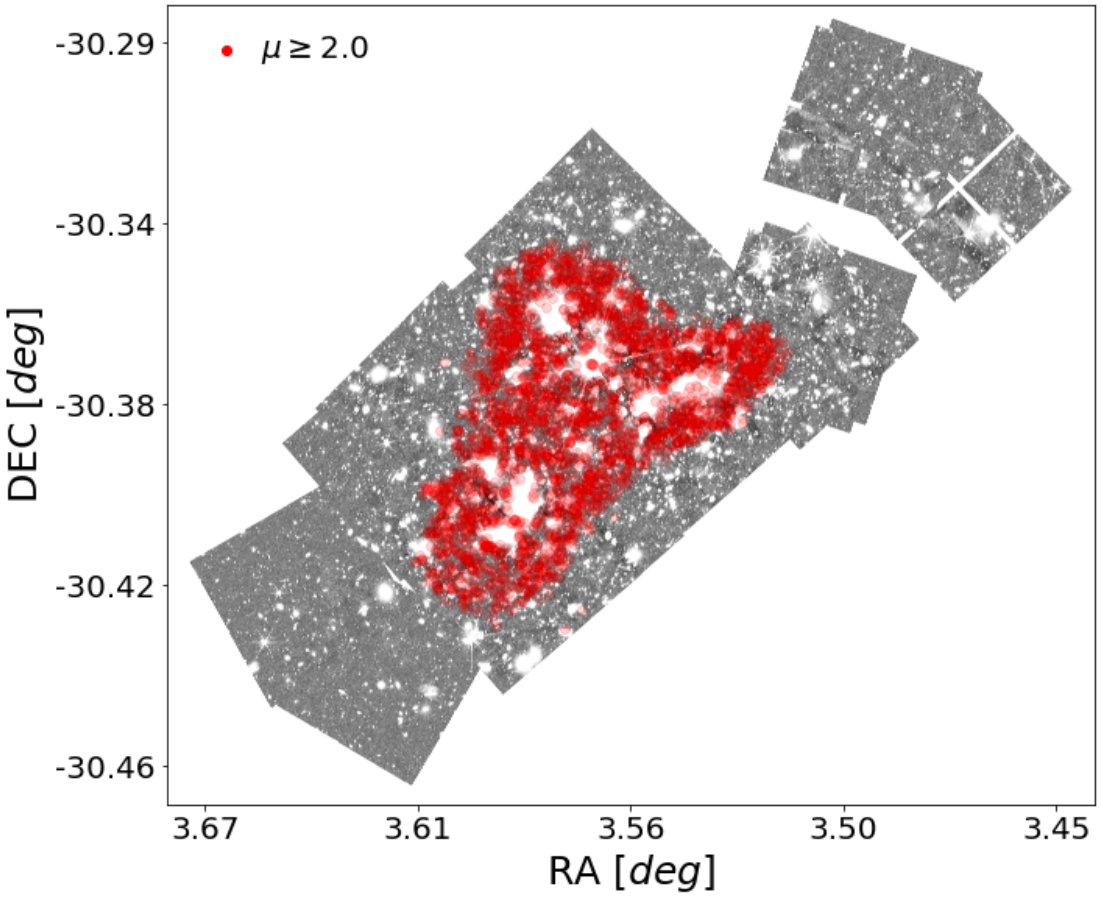

FIGURE 5.1: NIRCam scientific map in the F150W filter. The red region represents the area where the magnification is $\mu \geq 2.0$. The galaxies detected in this region of the survey are excluded from the parent sample.

high-resolution reduction (hereafter referred to as `hres`, with 1 pixel = 0.02") of the F150W and F200W NIRCam filters (with filter windows of 1300-1700 Å and 1700-2300 Å respectively) by Brammer [42] over the entire survey area, as illustrated in Fig.5.1. Because we focus on a high redshift sample of galaxies, we restrict the Paris et al. [201] source catalog to candidates with a photometric redshift in the range of $4 \leq z < 9$. We set an upper bound at $z = 9$ due to the low number of candidates at such a high redshift, limiting any statistical constraints we can place on our population analysis. This provides an initial catalogue of 3228 galaxies. To ensure we can robustly measure the morphological parameters, we imposed both a signal-to-noise ratio (SNR) and a star class requirements, demanding that candidate galaxies in the observed NIRCam imaging bands exhibit a total SNR exceeding $SNR = 7.0$ in each band and that the value of the stellarity index `class_star` measured by the `SExtractor` code (Bertin & Arnouts [24]) is not consistent with that of a point source, i.e., we require that `class_star`< 0.9. Additionally, for our morphological analysis we aim to ensure that the parameters we

measure reflect the true shape of the galaxies and are not deformed by gravitational lensing, which is important in our study because our NIRCam footprint includes the Abell 2744 galaxy cluster. We considered the impact of gravitational lensing and opted to exclude regions with excessively high magnification coefficients. Galaxies were deemed suitable if their median magnification, computed by Bergamini et al. [22], fell below $\mu < 2.0$. Fig.5.1 illustrates the region excluded by this criterion, highlighted in red.
Finally, to ensure the completeness of our sample, we implement an apparent magnitude cutoff, making a conservative assumption using the limiting AB magnitude at 5σ for NIRCam in the F150W band (rest-frame UV 1500-3000Å for our sample), which was set at $m_{AB} = 28.87$. Absolute UV magnitudes at 1500Å and stellar masses were calculated by fitting the photometry with Bruzual & Charlot [45] templates and assuming a delayed exponentially declining SFH, as done in Santini et al. [236].
Tab.5.1 shows the number of candidates detected by NIRCam (total: 3228) in each redshift bin and the number of targets after the application of the selection criteria used in this work (total: 675).
Given that the angular-diameter distance is relatively constant at high z, to conduct morphological parameters estimations for each galaxy we consider a cutout from the composite high-resolution map (with 1 pixel = 0.02") of 110px on a side corresponding to 2.2" on a side (our science images are drizzled on a 0.02"/pixel scale).

5.4 Analysis

5.4.1 Morphological Parameter Definition

In this work we employ three widely recognized quantitative morphological statistics, employed for characterization of mergers (Conselice [55]): the Gini coefficient (G), the second-order moment of brightness (M_{20}), and the asymmetry (A). The definitions are given below for convenience of the reader and to set the notation[1].
The Gini coefficient, proposed as a morphological parameter by Abraham et al. [3], quantifies the inequality in the distribution of pixel intensities within a galaxy image. It is computed as:

$$G = \frac{1}{\bar{X} n(n-1)} \sum_{i}^{n} (2i - n - 1) X_i, \tag{5.1}$$

where X_i represents the intensity of the ith pixel, n is the total number of pixels assigned to the galaxy from the segmentation map, and $\bar{X}$ is the mean intensity. By definition, this parameter falls within the range of $[0, 1]$, where a value of 0 implies that the galaxy

[1]The morphological parameters were measured using `JWSTmorph`, a publicly accessible code available on the GitHub repository: https://github.com/Anthony96/JWSTmorph.git

exhibits a uniform distribution in terms of intensity, while a value of 1 indicates that one pixel possesses all the flux. For galaxy mergers, it is anticipated that this morphological parameter will be higher compared to non-mergers. This expectation arises from the fact that the process of galaxy mergers often results in a more concentrated distribution of light within the merged system. This concentration occurs due to interactions between galaxies and the central concentration of stellar material, contributing to a higher value of this parameter (e.g., Lotz et al. [165]).

The second-order moment of brightness M_{20}, introduced by Lotz et al. [164], measures the compactness and concentration of the brightest 20% of a galaxy's light. It is defined as:

$$M_{20} = \log_{10}\left(\frac{\sum_i M_i}{M_{tot}}\right), \text{ with } \sum_i f_i < 0.2 f_{\rm tot} \tag{5.2}$$

where $f_{\rm tot}$ is the total flux of the galaxy pixels identified by the segmentation map, M_i is the second-order moment of brightness for each pixel, and $M_{\rm tot} = \sum_i^n M_i = \sum_i^n f_i \left[(x_i - x_c)^2 + (y_i - y_c)^2\right]$, with f_i being the single pixel intensities, x_i and y_i are the pixel coordinates, while x_c and y_c correspond to the galaxy center where $\rm M_{tot}$ is minimized. Typical values for this parameter are typically found in the range of $[-3, 0]$. Increasing values within this range are correlated with a greater number of off-centered bright features associated with the galaxy under study. The merger process tends to make the central region of the merging galaxies more concentrated and brighter. Consequently, this concentration has a significant impact on the M_{20} parameter. Non-merging galaxies, on the other hand, typically exhibit a smoother and less concentrated distribution of bright pixels, resulting in lower M_{20} values. These galaxies may lack the distinct, compact, and bright regions that are characteristic of merging systems, M_{20} is expected to be higher for galaxy mergers compared to non-mergers.

The asymmetry parameter (A), proposed by Abraham et al. [2] and Conselice et al. [57], is calculated as:

$$A = \frac{\Sigma\,|I - I_\pi|}{\Sigma I} - A_{bkg}, \tag{5.3}$$

where I is the original cutout image, I_π is the image rotated by 180 degrees (or π radians), and $A_{\rm bkg}$ denotes the asymmetry of the background. By construction, the asymmetry parameter tends to be higher for galaxy mergers compared to non-mergers. This is because mergers usually disrupt structures, introduce asymmetry and irregularities into a galaxy's appearance.

These morphological parameters are measured using `hres` image cutouts of the candidates selected in Sec.5.3 using `JWSTmoprh`. The segmentation maps, which are used to assign the pixels to a galaxy, are derived as in Treu et al. [269] through the *photutils* package[2], using a threshold flux for detection of 2σ above the background.

[2] https://photutils.readthedocs.io/en/stable/citation.html

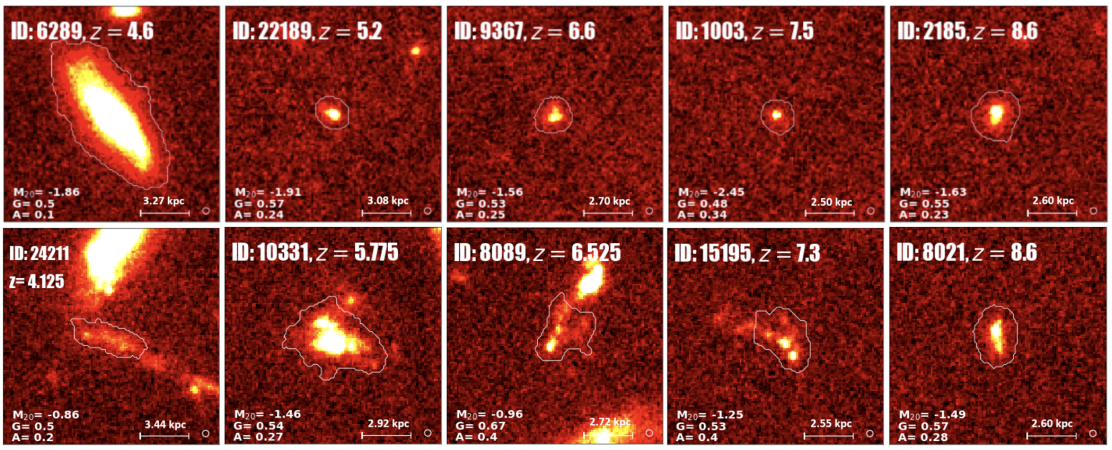

FIGURE 5.2: Illustrative representations from various redshift ranges, featuring extracted cutouts alongside corresponding morphological metrics, ID and photometric redshift. These snapshots are derived from a composite `hres` map (20mas per pixel), integrating data from F150W and F200W bands. The upper row highlights non-merging galaxies, while the lower row presents merger candidates. Each cutout stamp is 2.2" on a side, while the horizontal bar is 0.5" wide and the white contour represents the 2σ segmentation map boundary.

5.4.2 Identifying Mergers

The morphological parameters described above can be used to classify whether a galaxy is a merger, and Conselice et al. [58] and Lotz et al. [165] set out merger criteria using the two equations:

$$f(G, M_{20}) = G + 0.14M_{20} > 0.33 \tag{5.4}$$

$$A \geq 0.35 \tag{5.5}$$

These two equations have been derived from HST observations and have been validated for lower redshifts, extending only up to $z \sim 1.2$. It is worth noting that this redshift range falls considerably short of our intended investigation, which targets galaxies at $z \geq 4$. Nevertheless, recent research Treu et al. [270], Vulcani et al. [277] has revealed minimal variations in the underlying parameters with respect to redshift. In light of these findings, we proceed to employ these equations in our examination of our galaxy sample. To facilitate classification, we introduce two distinct samples:

• Silver: galaxies that satisfy the first criteria Eq.5.4 are classified as members of the Silver sample.

• Gold: galaxies that satisfy both criteria equations Eq.5.4 and Eq.5.5 are classified as members of the Gold sample.

Note that according to the definition above, the candidates classified as mergers in the

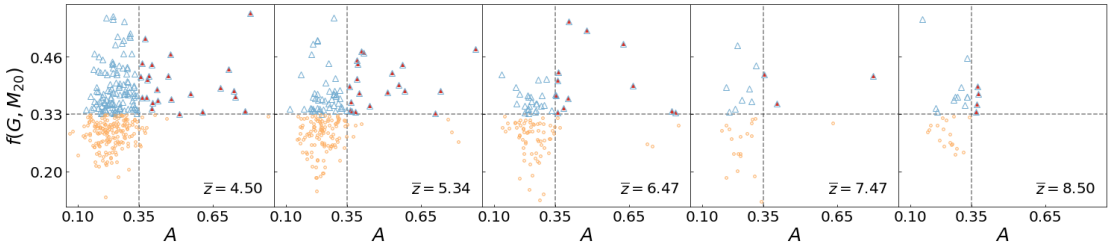

FIGURE 5.3: Visual representation of the classification method adopted in this study for five different z-bins (see Sec.5.4.2). Galaxies located above the horizontal dashed line (representing Eq.5.4) are labelled with blue triangles and are classified in the Silver sample. Among these, the subset to the right of the vertical dashed line (representing Eq.5.5) is highlighted in red, representing the galaxies in the Gold sample. We used high-resolution reduction of the F150W and F200W NIRCam imaging (see Sec.5.3).

Gold sample are a subset of those in the Silver sample. The difference between these two subsamples is in a requirement on the asymmetry parameter A. Physically, this means that in addition to considering the flux and brightness distribution from G and M_{20}, we also account for the pixel non-uniformity caused by the disruptive merging process, indicating a more dispersed galaxy.

By employing these classification criteria, we are able to categorize mergers based on their adherence to specific conditions, enabling a more systematic analysis of merger populations in the field of astrophysics Lotz et al. [165].

In Fig.5.2, we offer a representative example of galaxies categorized as mergers and non-mergers in each redshift bin. The upper row showcases a non-merger candidate, while the lower row presents a merger candidate, the associated morphological parameters are presented as a reference.

5.5 Results and Discussion

5.5.1 Mergers characterization

We measure the morphological parameters for the 675 galaxies in our parent sample, and use these to define a Silver and Gold sub-sample of mergers based on our merger criteria. We measure these parameters from the `hres` imaging, and note that we find consistent values for the structural parameters and merger classifications when using alternative filters (see App.5.5.3). Our classification outcomes are visually depicted in Fig.5.3, where the two dashed lines represent the threshold values derived from Eq.5.4 and 5.5. Each data point on the graph corresponds to an individual galaxy from our sample, chosen based on the criteria detailed in Sec.5.3 and summarized in Tab.5.1.

Fig.5.4 provides the respective count of galaxies falling into each region defined by the

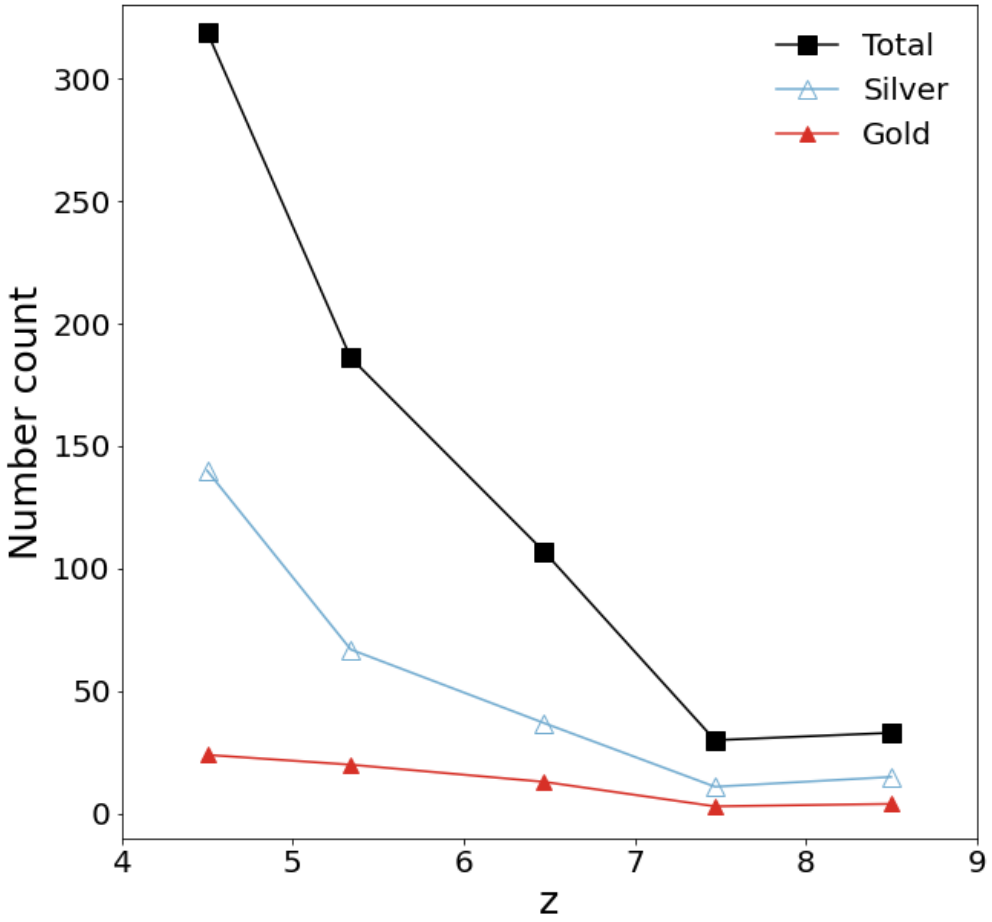

FIGURE 5.4: Demography for the galaxy (merger) samples versus redshift. Points correspond to those in Fig.5.3. Black squares indicate the total number of candidates considered in each redshift bin (bold number in column (2) in Tab.5.1). Blue triangles represent galaxies in the Silver sample, while red triangles denote those classified in the Gold sample.

dashed lines in Fig.5.3. The black squares represents the total sample of galaxies that meet the selection criteria outlined in Sec.5.3, the blue triangles designates the Silver sample, while the red triangles represents the Gold sample.

To offer a more comprehensive characterization of the candidate sample selected for this study and the identified mergers, we additionally present two figures. Fig.5.5 shows the galaxy absolute magnitude in U-band (left panel) and stellar mass (right panel) as a function of redshift for the Gold sample and the non-Gold galaxies (all the galaxies excluded from the Gold sample). It is apparent that the two distinct categories of galaxies exhibit a comparable range in terms of both magnitude and mass, so we can state that mergers are not typically brighter or more massive than isolated galaxies (to 2σ) and this inference is valid across the whole redshift range of our analysis. Furthermore, for a spatial perspective and to better visualize the distribution of classified mergers, Fig.5.6 displays the F150W image in (RA, DEC) coordinates. Overlaid on the image are galaxies classified as mergers (Gold sample) and those that do not fall into this category.

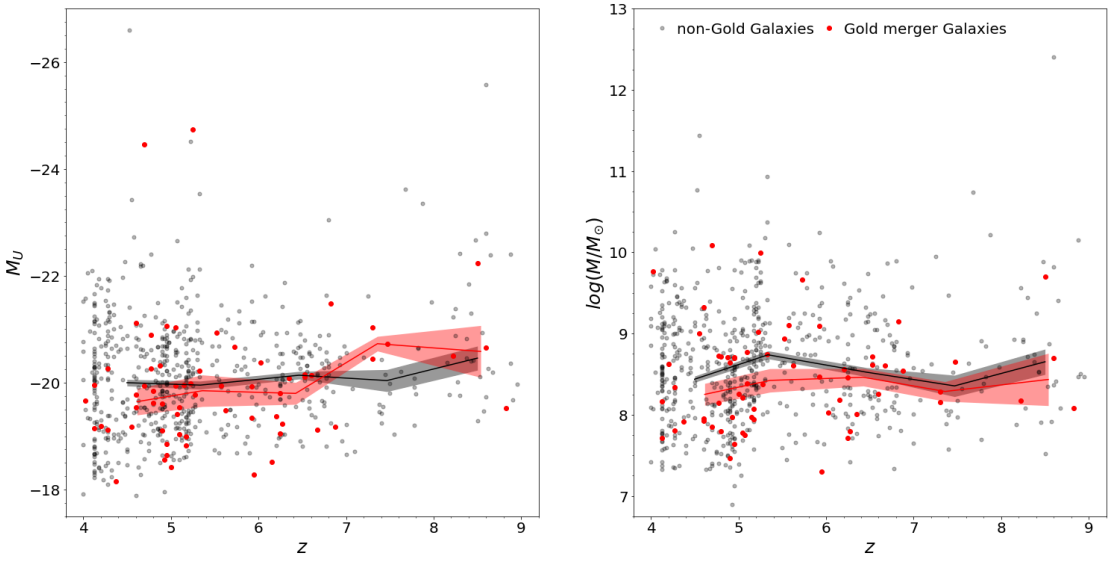

FIGURE 5.5: Rest frame U-magnitude (left) and stellar mass (right) as a function of photometric redshift for the Gold sample (red) and for all other, non-Gold galaxies (grey). Solid lines represent the corresponding median evolution in respect to the mean redshift in each bin

of the two different samples of mergers and non-mergers. Shaded regions represent the standard error corresponding to these median values.

This visualization is based on high-resolution data and covers the entire redshift range, specifically $z \in [4.0, 9.0]$. From Fig.5.6 we can see that there is no peak in the spatial distribution of the merger galaxies contained in the Gold sample; they are more or less uniformly distributed within the area enclosed in the foreground of the Abell 2744 cluster field.

5.5.2 Merger fraction versus redshift

In each redshift bin, we define the merger fraction $f_{m,i}$ as the number of mergers (N_m) over the total number of galaxies ($N_{\rm tot}$) in the i-th redshift bin:

$$f_{m,i} = \left(\frac{N_m}{N_{\rm tot}}\right)_i . \tag{5.6}$$

To explore the potential trend of merger fraction with magnitude, we implement a magnitude cutoff to divide the sample into bright and faint galaxies. We set a magnitude threshold at $M_U = -20.1$, the median of all galaxies presented in the left panel of Fig.5.5. We apply a similar approach when creating two sub-samples by introducing a stellar mass cutoff for the candidates, with a threshold of $\log(M/M_\odot) = 8.5$. This threshold corresponds to the median of all galaxies presented in the right panel of Fig.5.5. In Fig.5.7, we depict the merger fraction across the redshift range, considering both the Gold and Silver samples examined in this study. We present the merger fraction f_m

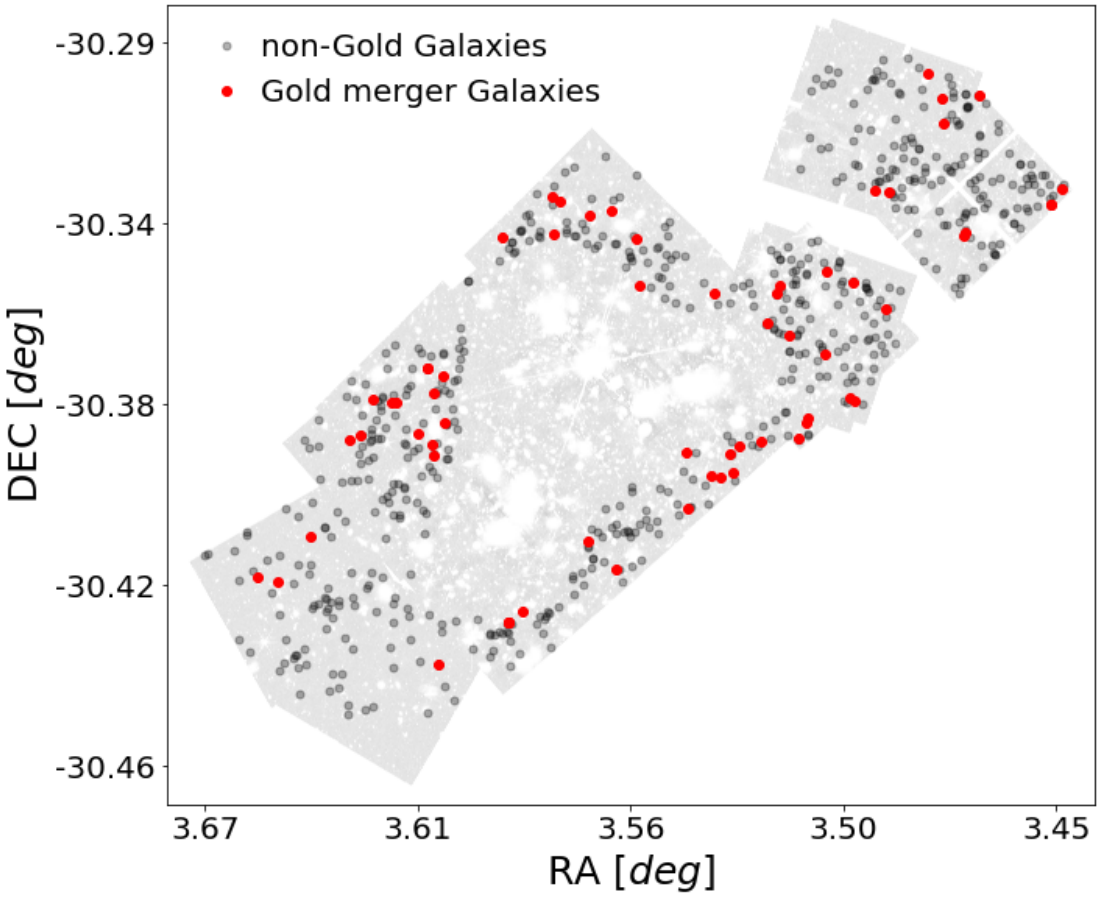

FIGURE 5.6: The Gold sample and non-Gold galaxies obtained in all the different redshift bins. Note that the central region is not populated as we excluded regions with magnification $\mu \geq 2.0$, Fig.5.1.

in the top two panels when considering the total sample of galaxies, alongside two subsamples created by implementing a magnitude cutoff to distinguish them. The two lower panels show the same quantity, but this time, the sample is segregated based on stellar mass.

The distinction in defining these two samples has a noticeable impact on the overall merger fraction across all the redshifts considered in our study. This effect is visually evident in Fig.5.4 when we look at the number of galaxies in each group. When we consider our entire sample (blue regions in Fig.5.7) we observe minimal variation with redshift in the measured merger fraction with redshift for our Silver and Gold criteria. Specifically, for the Silver sample, the merger fraction is measured to be $f_m = 0.39 \pm 0.06$ when considering the entire redshift range. In contrast, the Gold sample exhibits a substantially lower of $f_m = 0.11 \pm 0.04$. The choice between the Silver and Gold samples has a significant impact on the derived merger fraction values. While the Silver sample, with its more permissive criteria, yields higher merger fractions across the redshift range, the Gold sample, which employs stricter criteria, results in consistently lower and more

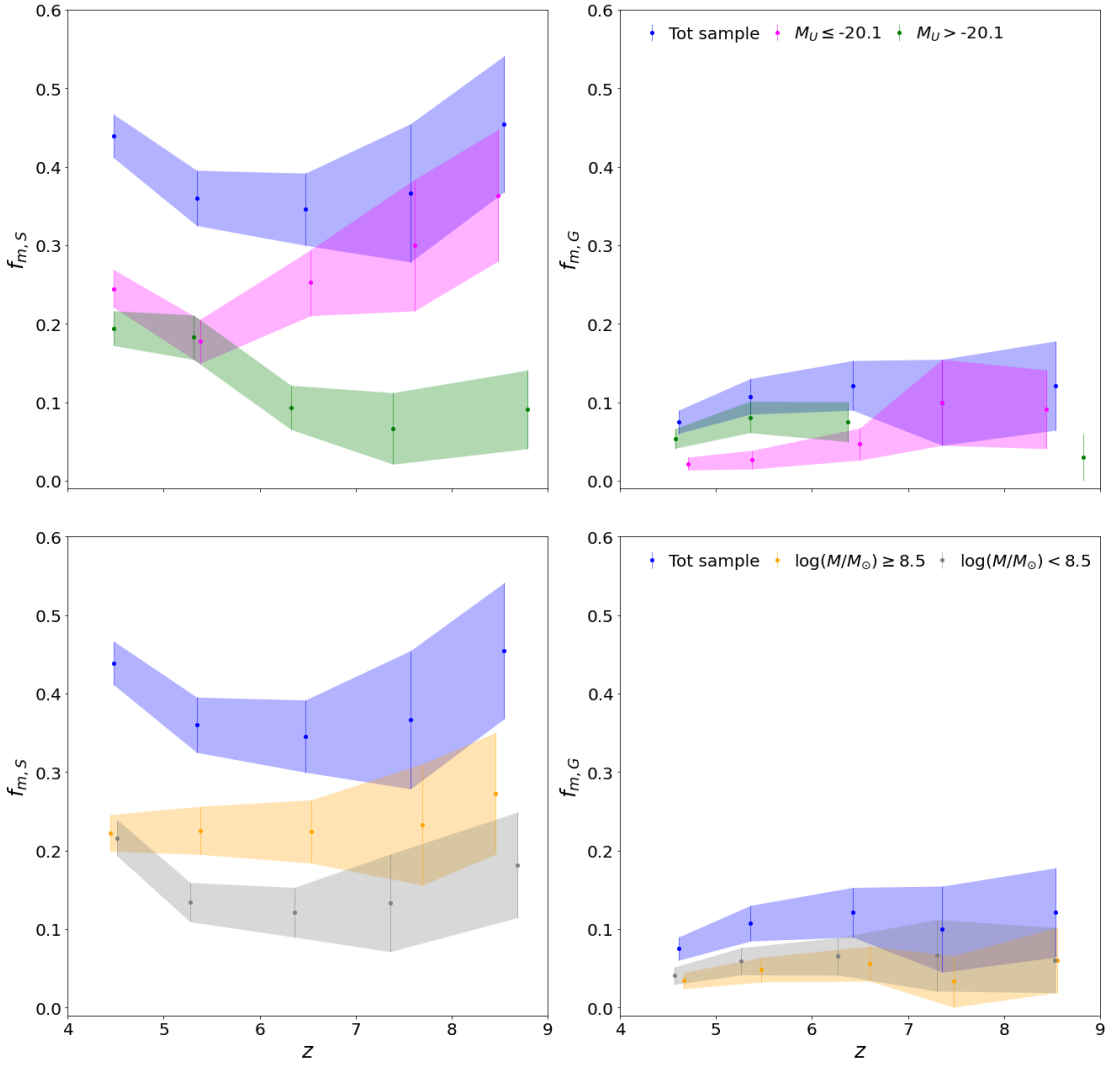

FIGURE 5.7: Merger fraction of the Silver sample (left panels) and of the Gold sample (right panels) as a function of redshift. The definition of the fraction follows Eq.5.6. In the top panels, we show a further separation based on the rest frame magnitude, with bright galaxies classified as having U-band magnitudes below $M_U = -20.1$, and faint galaxies lying above this threshold. In the bottom panels, we separate galaxies into a high mass sample ($M_* > 10^{8.5} M_\odot$) and a low mass sample ($M_* < 10^{8.5} M_\odot$).

stable merger fractions.

Earlier investigations (e.g. Conselice et al. [59], Jogee et al. [136], Kim et al. [144], Lin et al. [160], López-Sanjuan et al. [163], Lotz et al. [165, 166], Man et al. [172]) on galaxies at $z \lesssim 3$ consistently observed that the merger fraction tends to exhibit a relatively stable pattern around the value of $f_m \sim 0.10$, showing no significant overarching trend in relation to redshift. Studies conducted at higher redshifts (e.g.,Ventou et al. [273, 274]) also suggest that beyond a certain redshift threshold, likely below $z < 1$, the merger fraction remains constant around a value of $f_m \sim 0.10$ and does not display significant fluctuations. Our assessments of merger fractions based on Gold-criteria identification reveal no discernible redshift evolution, with an f_m value consistent with findings from

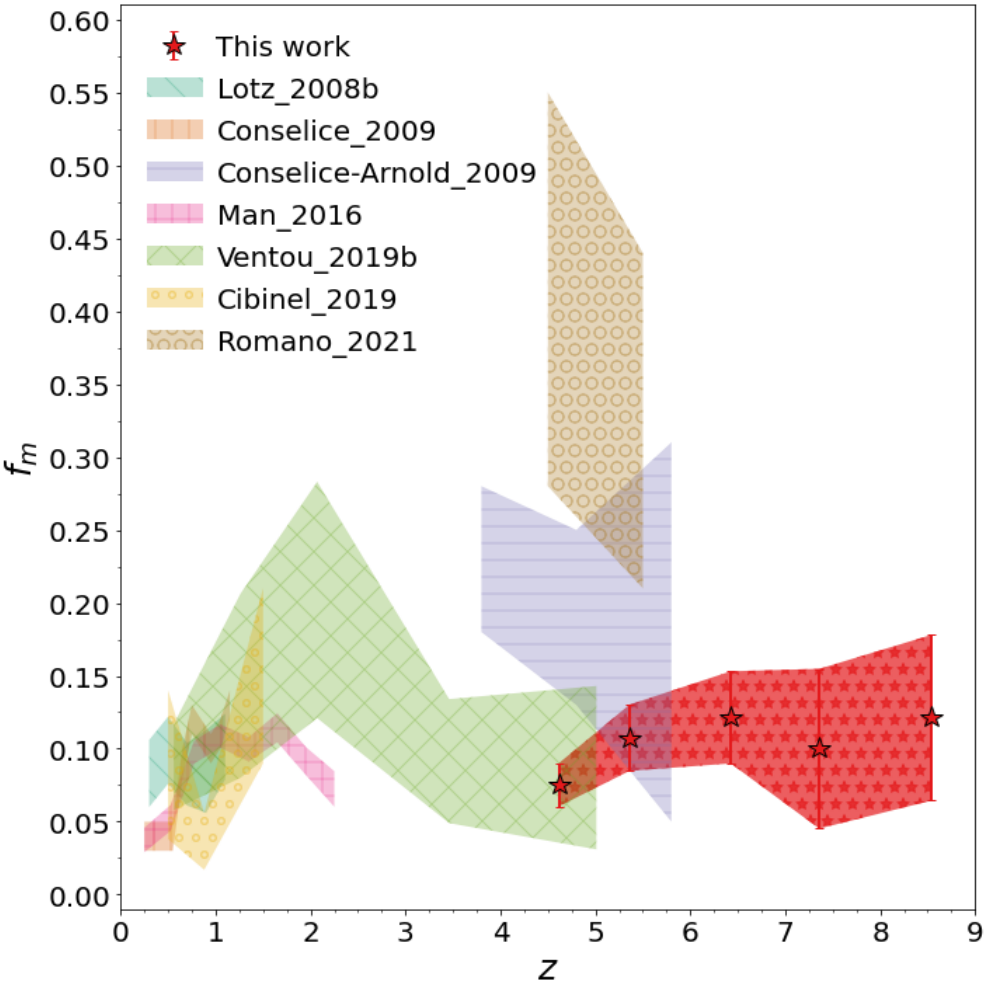

FIGURE 5.8: Comparison of the merger fraction (f_m) evolution with respect to redshift in our work against literature measurements from: Cibinel et al. [52], Conselice & Arnold [56], Conselice et al. [59], Lotz et al. [166], Man et al. [172], Romano et al. [234], Ventou et al. [274]. Note that each study implements their own catalogs, data selection, and methodologies for identifying mergers, leading to a non homogeneous comparison that may be affected substantially by systematic effects.

lower redshift investigations.

Fig.5.8 is a comparative plot to show how the redshift range $0 < z < 9$ is populated by the results obtained with different methods and data selections for studying the evolution of the merger fraction with redshift. In particular, Lotz et al. [166] conducted a morphological study on galaxies at low redshift $0.2 < z < 1.2$ with luminosity $L_{IR} > 10^{11} L_{\odot}$. In a similar redshift range and also using morphological analysis, Conselice et al. [59] focused on galaxies with stellar mass $M_* > 10^{10} M_{\odot}$, extending the study to $4 < z < 6$ in Conselice & Arnold [56]. Using a pair statistic approach from Man et al. [172] with the UltraVISTA/COSMOS Ks-band selected catalog, the authors measured the major merger fraction for galaxies with stellar mass $M_* \geq 10^{10.8} M_{\odot}$ within the redshift range $0.2 < z < 3$. Similarly, Cibinel et al. [52] used a combination of close pairs and morphological analysis in a sample of low redshift galaxies ($0.2 \leq z \leq 2.0$) with stellar mass

exceeding $M_* > 10^{10} M_\odot$. Ventou et al. [274] analyzed close pairs from cosmological simulations in the redshift range $0.2 < z < 6$ for galaxies with stellar mass $M_* > 10^{9.5} M_\odot$. In Romano et al. [234], a combination of morphological and kinematic measurements was employed, using emission lines of [CII] from a sample of galaxies with $4.4 < z < 5.9$ and stellar mass $9 \lesssim \log(M_*/M_\odot) \lesssim 11$. The results presented in this study are shown in Fig.5.8, with a red star representing the Gold sample.

Our main results for the merger fraction, considering the primary Gold sample, differ from some studies at intermediate redshift ranges ($4 < z < 6$) that derive a higher merging fraction, but use a different method for classification. Specifically, studies based on kinematic properties from emission spectra of galaxies, such as Romano et al. [234], show a merger fraction of $f_m = 0.44^{+0.11}_{-0.16}$ at $z \sim 4.5$ and $f_m = 0.34^{+0.10}_{-0.13}$ at $z \sim 5.5$. Interestingly, our Silver sample results, where our selection criteria are less stringent (Sec.5.4.2), agree with these studies, leading to a merger fraction of $f_m \sim 0.39 \pm 0.06$ for intermediate redshift. Given the intrinsic uncertainties and limitations of morphological selection of merger candidates, a NIRSpec follow-up would be highly beneficial to understand the differences in the Gold versus Silver samples, and contribute to a robust and accurate merger fraction measurement.

Interestingly, we measure a merger fraction which is nearly redshift independent. To investigate the physical nature of this result, we can consider a simple model, where the merger fraction f_m is expressed as a function of the merger rate:

$$f_m(z) = \Gamma_m(z) \cdot \tau_m(z), \tag{5.7}$$

where $\Gamma_m \equiv dN_m/dt$ is the merger rate, representing the number of mergers per galaxy per unit of time. τ_m is the typical timescale in which a close pair will eventually merge into a single system, assuming that all close pairs will undergo a merging event. In studying the merger rate of dark matter halos, Fakhouri & Ma [89] discuss how this quantity exhibits an almost universal behavior, showing a proportionality with the redshift of $\Gamma_{DMH} \propto (1+z)^\beta$ with $\beta = 2 - 2.3$ up to high redshift $z \leq 7$. This has been derived also for galaxy merger events, finding the same redshift proportionality as for dark matter halos with $\beta = 1.8$ out to $z \leq 3$ in Hopkins et al. [128] and $\beta = 2.43 - 2.87$ with $z \leq 8$ in Rodriguez-Gomez et al. [231]. The expected increase in Γ_m with increasing redshift from the dark matter halo dynamics is balanced by a decrease in τ_m as galaxies become smaller at higher redshift and thus the dynamical timescale decreases. This trend has been quantified by previous work Romano et al. [234], Snyder et al. [253]) which derive $\tau_m \propto (1+z)^{-2}$. As a result, the expression in Eq. 5.7 becomes approximately constant with redshift, in agreement with our results and previous measurements, including Conselice et al. [59], Mantha et al. [173], Ventou et al. [273].

5.5.3 Wavelength dependency study

In this subsection, we investigate potential correlations between the detection of galaxy mergers and variations in the merger fraction f_m. This examination takes into account the various wavelengths used for candidate identification, utilizing a selection of bands, specifically F200W, F277W, F356W, and F444W, in conjunction with the `hres` band used as the reference in the main study.

Fig.5.9 provides a visualization of the morphological parameters within the same redshift bin and band scheme. Furthermore, Fig.5.10 offers a graphical representation of the number of galaxies within the total, Silver, and Gold samples. From these figures, it becomes apparent that there is no significant discrepancy in the number of galaxies detected in each band when examining the same redshift bin. This suggests that there is no pronounced trend in detection based on wavelength.

To further explore potential dependencies, we applied the same magnitude and stellar mass cuts that were considered in the primary study. The goal was to assess whether the wavelength used for the survey impacted the results. Fig.5.11 and Fig.5.12 present the outcomes of this supplementary analysis. As demonstrated in these figures, the merger fraction f_m does not show significant variations with the slope of the evolution trend with respect to redshift constrained within the range (0,0.02) for each band. The mean values of the merger fraction derived from the five different bands are: f_m[`hres`]= 0.11 ± 0.04, f_m[F200W]= 0.18 ± 0.04, f_m[F277W]= 0.09 ± 0.03, f_m[F356W]= 0.07 ± 0.04 and f_m[F444W]= 0.5 ± 0.03. Overall, most of the results agree within 1σ with the primary result obtained using the `hres` band of $f_m = 0.11 \pm 0.04$, while the remaining results fall within the $1.5 - 2\sigma$ range.

In summary, our analysis reveals that there is no strong increasing or decreasing trend in the merger fraction when considering different detection wavelengths. These results are consistent with the previous findings by Treu et al. [270], Vulcani et al. [277], which suggested that the morphology of Lyman Break Galaxies remains relatively consistent across different wavelengths, from the rest frame optical to the rest frame UV.

5.5.4 sSFR of interacting and non-interacting systems

In this section we aim to explore potential variations in the impact of galaxy mergers on specific star formation activity across distinct redshifts.

To derive the star formation rates, we use `BAGPIPES` Carnall et al. [49] to fit the photometries of the galaxies. `BAGPIPES`' model galaxy spectra were generated with a Kroupa et al. [148] initial mass function, a Calzetti et al. [48] dust attenuation law, a BPASS Eldridge & Stanway [V2.2.1 87] stellar population, and nebular emission lines from `CLOUDY`

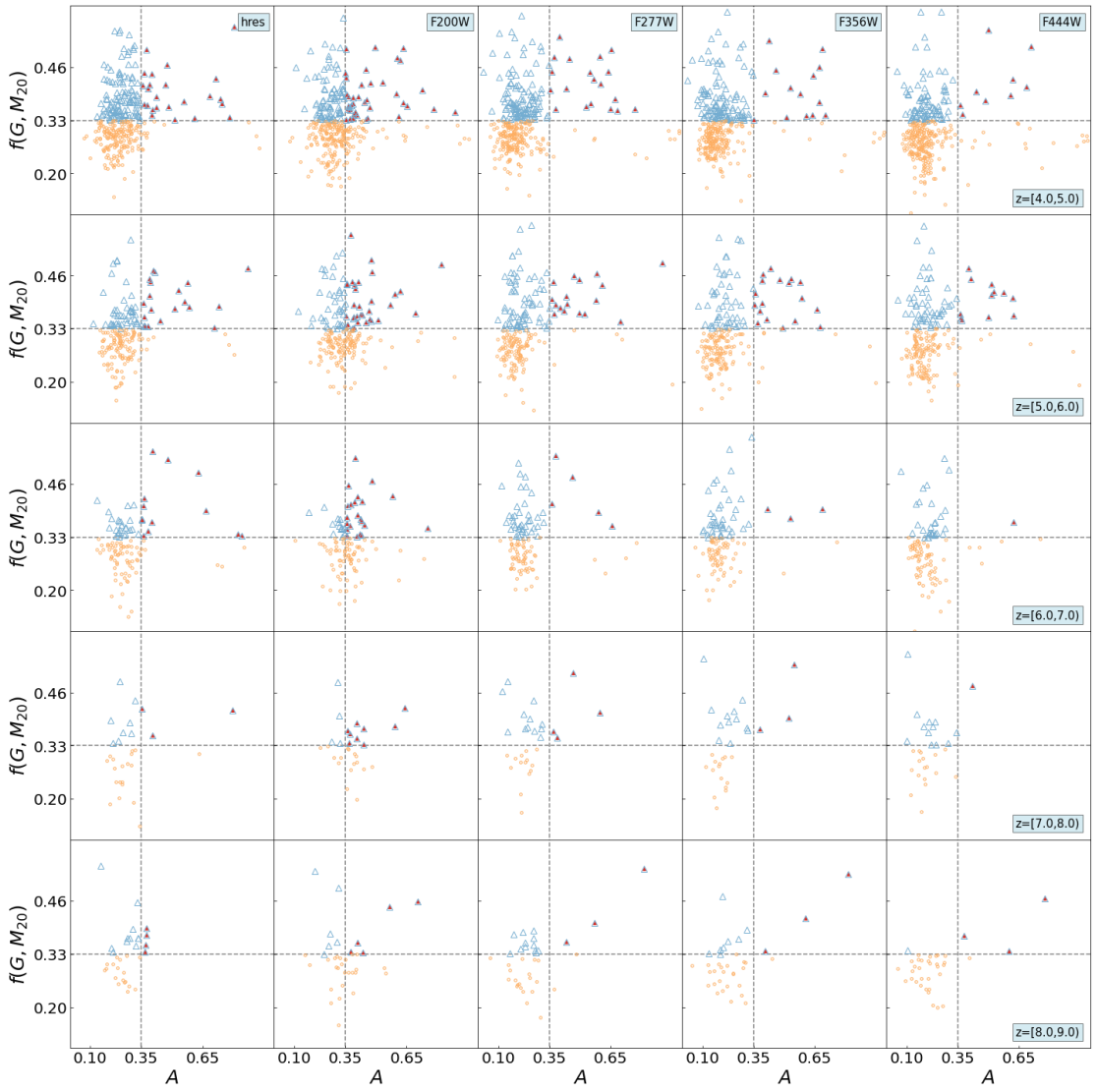

FIGURE 5.9: Mergers classification method presented in Sec.5.4.2. Each row shows results for a different redshift bin, with each panel within each row showing results when a different filter from NIRCam is used for determining the morphological parameters. Galaxies located above the horizontal dashed line (representing Eq.5.4) are labelled with blue triangles and are classified in the Silver sample. Among these, the subset to the right of the vertical dashed line (representing Eq.5.5) is highlighted in red, representing the galaxies in the Gold sample.

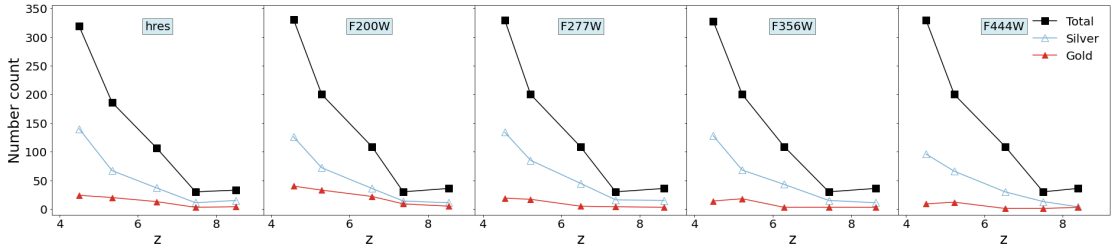

FIGURE 5.10: Demography for the galaxy (merger) samples versus redshift. Points correspond to those in Fig.5.9. Black squares indicate the total number of candidates considered in each redshift bin. Blue triangles represent galaxies in the Silver sample, while red triangles denote those classified in the Gold sample.

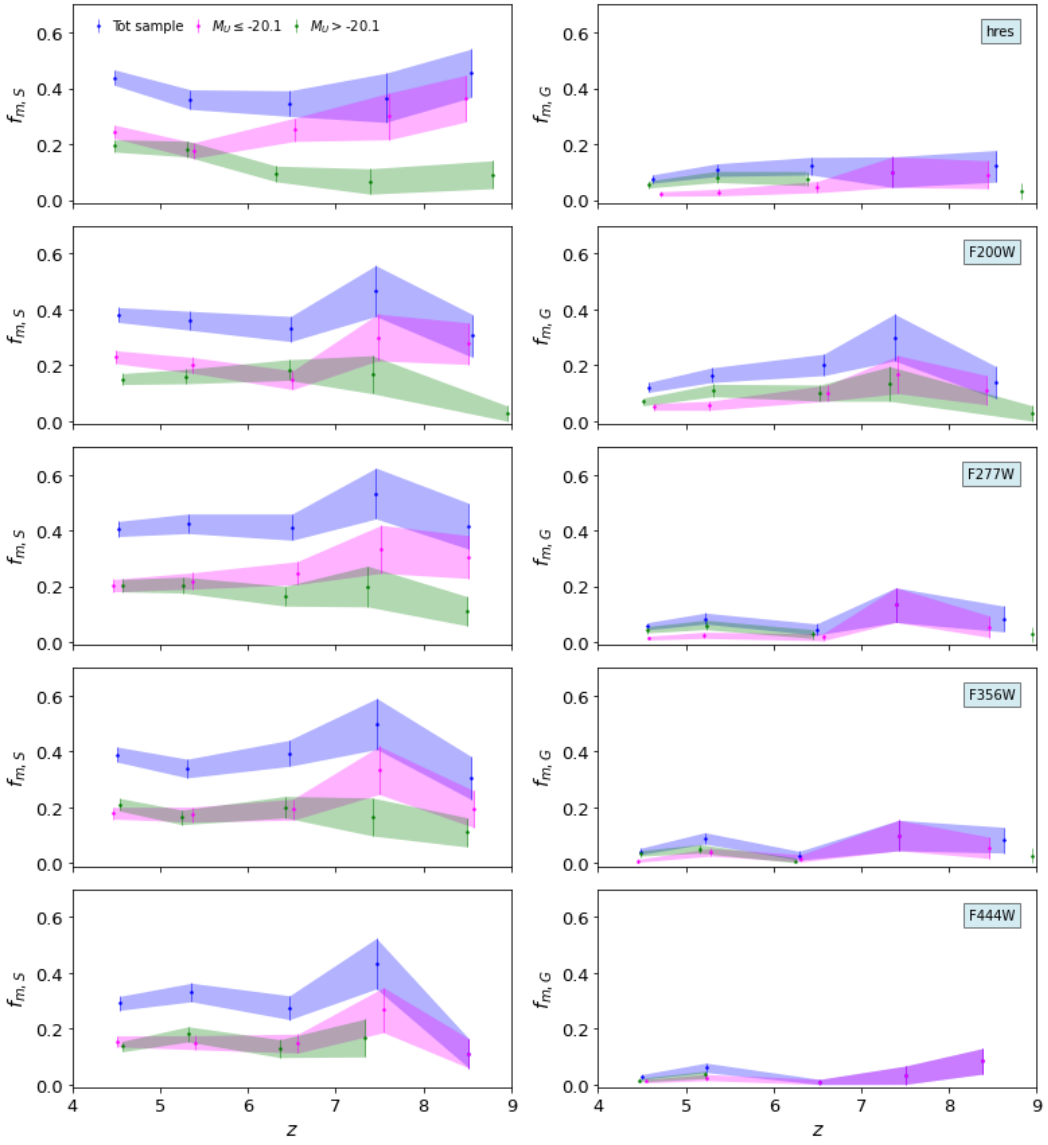

FIGURE 5.11: Merger fraction evolution in respect to redshift, using different filters on the NIRCam instrument for the identification. The left panels show the Silver sample and the right panels represent the Gold sample. In each panel, we also show the results for the subsamples of bright and faint galaxies, defined by a cut in magnitude at $M_U = -20.1$.

photoionization code Ferland et al. [92]. For the fit, we set the metallicity range to [0,1.2] $Z_\odot$, the dust attenuation A_V to [0,3] mag, the logarithm of the ionization parameter U to [-4,-1], and the logarithm of the total mass of stars formed in solar mass to [6,13]. We assumed two star formation histories: log-normal and non-parametric forms Iyer et al. [132].

We utilize a metric representing the excess of specific star formation rate (sSFR) for galaxies classified as mergers (Gold sample)in comparison to non-mergers (non Gold

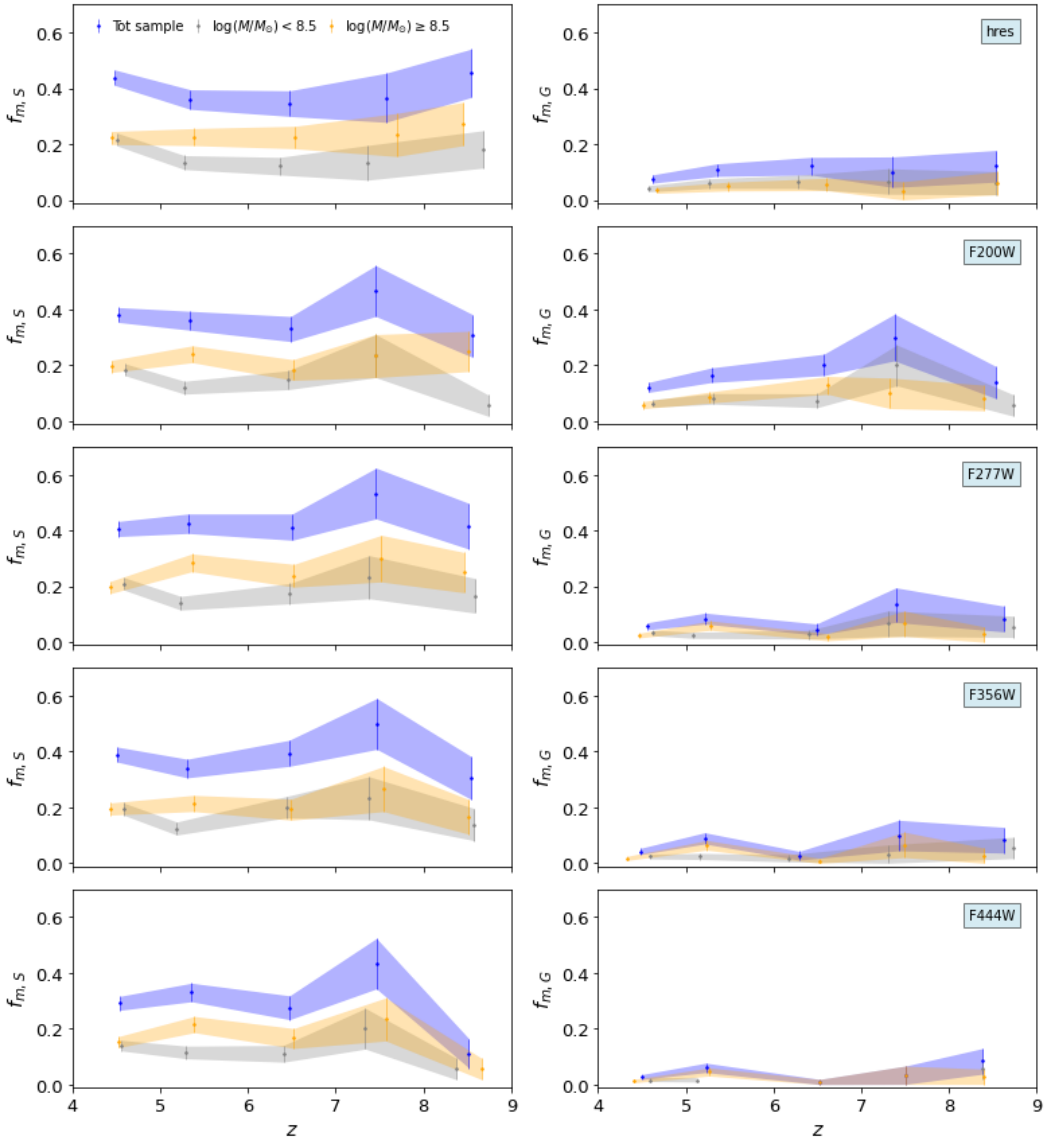

FIGURE 5.12: As in Fig.5.11, but splitting the parent sample into two subsamples based on stellar mass, with high-mass galaxies defined to be those with stellar masses greater than $M_* = 10^{8.5} M_\odot$, and low-mass galaxies having stellar masses below this value.

sample). This excess is calculated in the i-th redshift bin using the equation:

$$\Delta \log(\mathrm{sSFR}_i) = \log\left(\frac{< \mathrm{sSFR}_{\mathrm{merger}} >_i}{< \mathrm{sSFR}_{\text{non-merger}} >_i}\right), \tag{5.8}$$

the associated standard uncertainty from error propagation is used. To present results across our redshift range $4 \leq z < 9$ we decided to partition the sample of objects to ensure a roughly uniform count of mergers in each redshift bin.

According to the definition, values of $\Delta \log(\mathbf{sSFR}) > 0$ indicate that galaxies identified as mergers exhibit a higher rate of star formation per unit mass compared to galaxies classified as non-mergers. Therefore, these measurements would suggest that

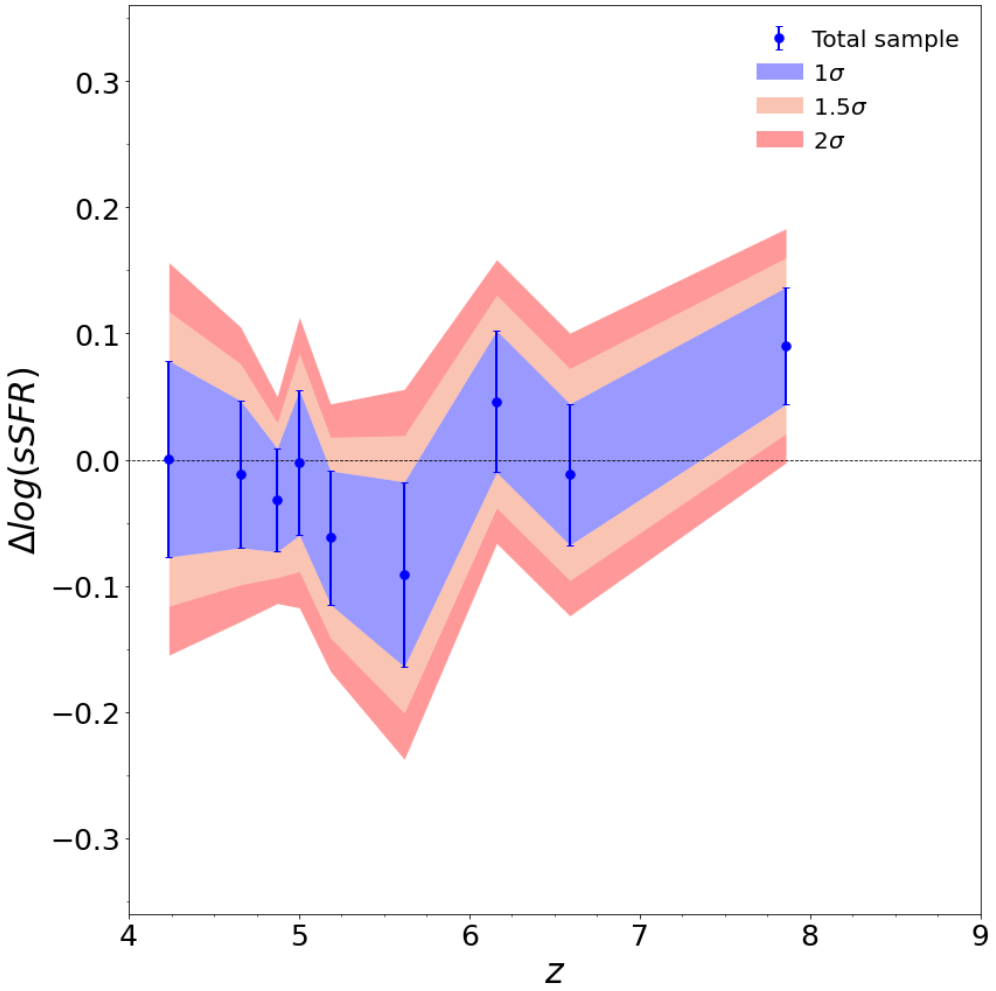

FIGURE 5.13: Excess specific star formation rate (sSFR) evolution as a function of redshift for both the complete set of Gold sample merger and non-Gold galaxies (see Eq.5.8). The shaded areas indicate uncertainty ranges of $\sigma = (1, 1.5, 2)$. The measurement is statistically consistent with no redshift evolution.

those candidates are currently experiencing heightened levels of star formation relative to a reference or baseline. Conversely, data points falling below the zero baseline ($\Delta \log(\mathbf{sSFR}) > 0$) indicate merger events where the sSFR is lower than what we observe in galaxies classified as non-mergers.

In Fig.5.13 we show the evolution of the estimator defined in Eq.5.8 with respect to redshift. The fluctuations around the zero line fall within the 1σ range for the majority of the data points, while the rest fall within the 2σ range. These findings suggest that there is no strong indication of any notable increase or decrease in the sSFR evolution for mergers compared to non-mergers.

5.6 Summary and perspectives

In this study, we focus on high-z galaxies in the outskirts (low magnification regions) of the foreground galaxy cluster Abell 2744, observed as part of the GLASS-JWST program ERS-1324 Treu et al. [269], UNCOVER JWST-GO-2561 Bezanson et al. [26] and DDT-2756 (PI Wenlei Chen).

Our main objective was to use high-resolution (1px = 0.02") F150W and F200W imaging to conduct an in-depth analysis of the evolution of the merger fraction f_m with respect to redshift, extending from $z = 4$ up to $z \simeq 9$. We accomplished this by investigating various subsamples, which were derived based on magnitude and mass thresholds. Our key findings are as follows:

• We investigated morphological parameters for high-redshift galaxies, categorizing them into two distinct merger subsamples: Silver and Gold. We employed the criteria outlined in Eq.5.4 and Eq.5.5, as defined in prior research on this topic. Our results are depicted in Fig.5.3, illustrating the distribution of morphological parameters, and in Fig.5.4, which provides a visual representation of the number of candidates in each sample, correlated with the mean redshift of the candidates within the selected bins.

• To obtain a more comprehensive understanding of the merger population in comparison to non-interacting galaxies, we have provided visual representations of the magnitude and mass distributions in Fig.5.5. These distributions have been segregated between candidates identified as Gold mergers and all galaxies included in the study, adhering to the selection criteria outlined in Sec.5.3. It is worth noting that no distinct trends have emerged from our analysis. We have observed that mergers are present across the entire range of magnitudes and masses at different redshifts. Notably, the distributions of these two categories exhibit similarities, as indicated by the medians of both magnitude and mass.

• Fig.5.7 displays the evolution of the merger fraction of two separate subsamples for various magnitude and mass ranges. The overall sample, represented by the blue contours, shows relatively small variations in the merger fraction, making it challenging to discern a significant trend. Notably, for the Gold sample, the values remain stable across the entire range of redshifts under investigation. The calculated mean value, along with its associated error, is $f_m = 0.11 \pm 0.04$. This outcome aligns with previous research at lower redshifts Conselice et al. [59], Jogee et al. [136], Kim et al. [144], Lin et al. [160], López-Sanjuan et al. [163], Lotz et al. [166], Man et al. [172], Ventou et al. [273, 274], all of which indicate a merger fraction around $f_m \sim 0.10$ with lack of evidence for redshift evolution. We presented a simple argument based on dark matter halo merger rates and dynamical timescales of galaxies (which determine for how long a merger remains visible) to support a scenario where the merger fraction is almost redshift independent.

• Our study aimed to investigate potential differences in the star formation rate between

two categories: mergers and non-mergers. As illustrated in Fig.5.13, we do not observe a clear and robust trend in the sSFR for mergers compared to non-merger galaxies, with redshift as the independent variable. Therefore, we conclude the sSFR is independent of whether a galaxy is a merger or not. In turn, this seems to suggest that smooth accretion (or minor mergers) are the main drivers in setting the SSFR for a star forming galaxy.

• There is no apparent correlation between the identification of mergers according to the wavelengths of the observation filters utilized to establish the initial sample. Consequently, the merger fraction does not display any specific pattern in this aspect maintaining a level compatible with the main result of this work of $f_m = 0.11 \pm 0.04$ across the examined redshift range. This result extends the findings obtained at lower redshifts, where Treu et al. [270], Vulcani et al. [277] argue that morphological parameters, which form the basis for classifying mergers and non-mergers and subsequently calculating the merger fraction, do not exhibit a strong dependence on the redshift at which galaxies are observed.

To go beyond the preliminary results for the merger fraction during the initial stages of the epoch of reionization, further data is required. To improve statistical robustness and address potential cosmic variance effect Trenti & Stiavelli [266] as well as any weak-magnification lensing bias, a promising prospective strategy involves examining galaxies and images across diverse fields (e.g., COSMOS Scoville et al. [242], EGS Davis et al. [72], PRIMER Dunlop et al. [81]), in addition to those considered in this work. This will enable a thorough exploration of the galactic environment, incorporating factors such as metallicity and spatial density and open opportunities to quantitative comparison to predictions from numerical simulations and theoretical modelling of galaxy assembly in the first billion years.

Chapter 6

Quantifying Spectroscopic Flux Variations between JWST NIRISS and NIRSpec: Slit Losses in Emission Line Measurements of z 1-3 Galaxies

This chapter is a reproduction of Dalmasso et al. [69]. The author lead both the data analysis and writing with the collaboration of the co-authors. This reproduction has been reformatted to fit the style of the thesis, changes include:

- **Figures and tables have been formatted to display correctly on the thesis**
- **Some bibliographic records have been updated**

6.1 Abstract

We analyze JWST NIRISS and NIRSpec spectroscopic observations in the Abell 2744 galaxy cluster field. From approximately 120 candidates, we identify 12 objects with at least a prominent emission lines among [*O* II] λ3727, Hβ λ4861, [*O* III] λ4959, [*O* III] λ5007, and Hα λ6563 that are spectroscopically confirmed by both instruments. Our key findings reveal systematic differences between the two spectrographs based on source morphology and shutter aperture placement. Compact objects show comparable or higher

integrated flux in NIRSpec relative to NIRISS (within 1σ uncertainties), while extended sources consistently display higher flux in NIRISS measurements. This pattern reflects NIRSpec's optimal coverage for compact objects while potentially undersampling extended sources. Quantitative analysis demonstrates that NIRSpec recovers at least 63% of NIRISS-measured flux when the slit covers $> 15\%$ of the source or when $R_e < 1$kpc. For lower coverage or larger effective radii, the recovered flux varies from 24% to 63%. When studying the H$\alpha\,\lambda6563/[O\,\textsc{iii}]\,\lambda5007$ emission line ratio, we observe that measurements from these different spectrographs can vary by up to $\sim$0.3 dex, with significant implications for metallicity and star formation rate characterizations for individual galaxies. These results highlight the importance of considering instrumental effects when combining multi-instrument spectroscopic data and demonstrate that source morphology critically influences flux recovery between slit-based and slitless spectroscopic modes in JWST observations.

6.2 Introduction

Spectroscopic observations are fundamental to astronomical research, enabling detailed investigations of celestial objects' physical properties, chemical compositions, and dynamical characteristics. Since the 19th century, when astronomers first used prisms and diffraction gratings to disperse starlight, advances in instrumentation have greatly expanded the range and sophistication of spectroscopic techniques available to researchers.

Modern astronomical instrumentation has developed three primary spectroscopic techniques that exemplify the sophisticated technological landscape of contemporary research: integral-field, slitless, and slit-based spectroscopy. Integral-field spectroscopy (IFS) provides the unique capability to obtain spatially resolved spectral data across extended objects. However, it presents significant technical challenges, including complex data reduction processes and substantial computational demands for processing the resulting three-dimensional data cubes (Böker et al. [31]). The achievable spectral resolution with IFS is determined by the specific instrument design and can be comparable to that of traditional slit-based spectrographs, rather than being inherently limited by the technique itself.

Slitless spectroscopy, as implemented by the Wide Field Slitless Spectroscopy (WFSS) mode of the Near Infrared Imager and Slitless Spectrograph (NIRISS) aboard the James Webb Space Telescope (JWST, Gardner et al. [102]), offers a comprehensive approach by dispersing light across the entire field of view, facilitating simultaneous spectral mapping of multiple objects. This method provides a broad observational perspective, capturing spectral information without the constraints of pre-selection, though it introduces the

challenge of potential spectral overlap between sources and higher backround noise. The WFSS mode is particularly advantageous for multi-object observations and will serve as our primary focus when examining NIRISS capabilities in this work.

In contrast, slit-based spectroscopy, represented by the Near Infrared Spectrograph (NIRSpec, Jakobsen et al. [133]), employs a more targeted observational strategy. While NIRSpec offers Fixed Slits Spectroscopy (FS), Multi-Object Spectroscopy (MOS) and Integral Field Unit (IFU) capabilities, this study will specifically examine the MOS mode employing a Micro-Shutter Assembly (MSA) narrow slit mask. This technique significantly reduces sky background contamination and enhances overall observational sensitivity. The precision of MSA-based spectroscopy allows for highly selective spectral observations, strategically isolating specific astronomical sources through carefully controlled shutter configurations. For an in-depth analysis of the MOS technique, we refer the reader to Ferruit et al. [94].

A fundamental limitation of slit-based spectroscopy is the occurrence of slit losses. This effect arises when a fraction of the incoming radiation from a target source does not propagate through the entrance slit and is consequently excluded from the final recorded spectrum. Slit losses originate from a combination of geometric misalignment, the intrinsic source morphology, and the wavelength-dependent optical response of the instrument.

They are strongly dependent on the spatial extent and surface brightness distribution of the source, the point spread function (PSF) at each wavelength, and the on-sky alignment of the source relative to the slit. These losses introduce systematic uncertainties in the flux calibration, particularly for resolved or partially resolved sources, and must be accounted for either through empirical correction strategies or forward modeling of the instrumental response. Empirical corrections are generally based on available photometry or simple assumptions about the surface brightness distribution of the source. However, such corrections are not necessarily accurate for emission lines maps which are often not a simple scaling of the continuum surface brightness distribution and can be rather irregular.

In this study, we perform an initial exploration of the effects of slit losses using actual JWST NIRISS and NIRSpec data. In practice, we compare flux measurement performance between JWST's NIRISS WFSS and NIRSpec MSA instruments operating in multi-object spectroscopy modes. This analysis examines how instrumental differences affect spectroscopic accuracy across varying observational parameters. We have explicitly limited our investigation to these multi-object observation modes, excluding NIRSpec's IFU capability as it falls beyond the scope of this work. The study quantifies performance disparities between these complementary spectroscopic approaches,

with particular emphasis on slitloss effects in NIRSpec MSA relative to NIRISS WFSS observations.

This work fulfills one of the key deliverables of The GLASS JWST Early Release Science Program (GLASS-JWST-ERS)[1] by providing a direct comparison between the two spectrographs. Our analysis is constrained by the unique characteristics of our dataset, which is based on the catalog release containing redshifts and emission line fluxes determined from JWST/NIRISS Wide-Field Slitless Spectroscopy observations in the Abell 2744 cluster field Watson et al. [278]. This catalog represents another important deliverable of the GLASS-JWST program and provides the foundation for our comparative spectroscopic study. We emphasize that, in the spirit of the Early Release Science Programs, this is just the first step in this investigation. A valuable future step will be to extend our investigations to larger samples of galaxies with both NIRSpec and NIRISS investigations. However, this next step will likely require some time, owing to the inherent complexity of the NIRISS dataset and data reduction, making this initial exploration more valuable.

The structure of this paper is outlined as follows: in Sec.6.3, we describe data reduction and the creation of the parent sample; in Sec.6.4, we present two distinct analyses: precise emission line fitting and the study of emission line maps, with corresponding discussion of both approaches; and in Sec.6.5, we offer a summary of our key findings. When needed we adopt the cosmological parameters determined by the Planck Collaboration Planck Collaboration et al. [214]: $(\omega_{\rm m}, \Omega_\lambda, h, \sigma_8) = (0.3075, 0.6925, 0.6774, 0.8159)$. Magnitudes are reported in the AB system Oke & Gunn [191].

6.3 Data sets and sample selection

The NIRSpec spectral data used in this study were obtained through three public programs focused on the foreground galaxy cluster Abell 2744 and its immediate surroundings: (i) GLASS-JWST-ERS Program (PID 1324, PI T. Treu)[1], (ii) JWST-DDT Program (PID 2756, PI W. Chen), both with publicly available fully reduced spectra from Mascia et al. [175], and (iii) UNCOVER-JWST Program (PID 2561, PI I. Labbé)[2], specifically the latest data release DR4, detailed in Price et al. [218].

The NIRSpec observations from GLASS utilized the high-resolution configuration with R2700 gratings, providing spectral resolutions of R$\sim$1400-3600 across the 0.7-5.0 μm range. In contrast, the DDT and UNCOVER programs employed the low-resolution R100 prism mode, delivering more moderate spectral resolutions of R$\sim$30-300 across the

[1] https://archive.stsci.edu/hlsp/glass-jwst
[2] https://jwst-uncover.github.io

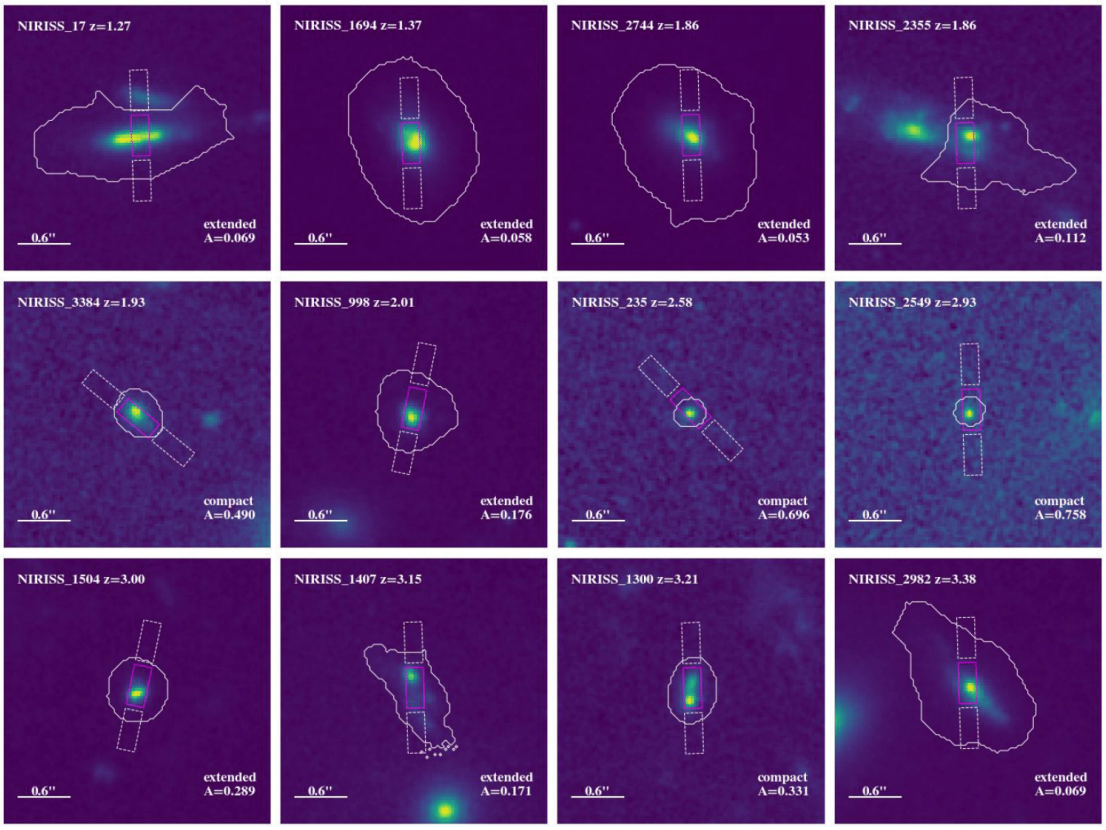

FIGURE 6.1: NIRISS imaging cutouts are presented using a drizzled stack of the F115W, F150W, and F200W filters, processed to better highlight the light associated with each galaxy candidate. Solid white contours represent the NIRISS segmentation maps, rectangles show the NIRSpec slits assigned to the target (solid magenta and dash-dot white). The ID and redshift z refer to measurements from NIRISS, using the catalog presented in Watson et al. [278] and a 0.6 arcsec scale bar is included. The "A" parameter, shown in the bottom right corner, represents the fractional area of the segmentation map that is encompassed by the slit. Objects are classified as either "compact" or "extended" based on the classification criteria presented in Sec.6.3.

Selected Candidates											
JWST Program	RA	DEC	NIRISS ID	NIRSpec ID	z_{spec}	m_{AB}	[O II] λ3727	Hβ λ4861	[O III] λ4959	[O III] λ5007	Hα λ6563
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
GLASS[1324]	3.58625	-30.41654	17	410067	1.27	23.3	-	-	X	X	X
GLASS[1324]	3.57674	-30.39360	1694	20021	1.37	22.71	-	-	X	X	X
GLASS[1324]	3.57018	-30.38372	2744	20008	1.86	22.81	X	X	X	X	X
GLASS[1324]	3.58697	-30.38699	2355	20025	1.86	24.41	-	X	X	X	X
UNCOVER[2561]	3.59938	-30.37594	3384	26882	1.93	26.35	-	-	X	X	X
DDT[2756]	3.59852	-30.40177	998	20015	2.01	24.79	-	-	X	X	X
UNCOVER[2561]	3.58490	-30.41102	235	8313	2.58	27.73	-	-	X	X	-
GLASS[1324]	3.60686	-30.38555	2549	90155	2.93	27.9	-	X	X	X	-
DDT[2756]	3.60711	-30.39562	1504	90012	3.00	25.01	-	-	X	X	-
GLASS[1324]	3.57497	-30.39678	1407	341568	3.15	24.2	-	-	X	X	-
GLASS[1324]	3.60780	-30.39805	1300	341060	3.21	25.63	-	-	X	X	-
GLASS[1324]	3.57786	-30.38118	2982	360005	3.38	23.0	X	-	-	-	-

TABLE 6.1: This table summarizes the galaxies used in this study for slit loss characterization, as discussed in Sec.6.3. Columns (8)-(12) list the emission lines analyzed in this work, with an "X" indicating lines that are matched by both the NIRISS and NIRSpec instruments and for which we successfully fitted the emission line to calculate the integrated flux. Column descriptions: (1) JWST Program; (2) Galaxy RA [deg]; (3) Galaxy DEC [deg]; (4) Associated NIRISS ID (Watson et al. [278]); (5) Associated NIRSpec ID (Mascia et al. [175], Price et al. [218]); (6) Spectroscopic redshift from NIRISS; (7) Apparent AB magnitude from NIRSpec in the F200W filter.

0.6-5.3 μm wavelength range. We highlight an important distinction in NIRSpec data processing across our survey datasets. Data from GLASS-JWST-ERS and JWST-DDT programs retain path loss and barshadow corrections within their reduction pipeline, while UNCOVER-JWST DR4 data omits these corrections. The path loss correction addresses flux attenuation resulting from geometric aperture effects and diffraction losses, with particular significance for sources displaced from optimal slit positioning. The barshadow correction accounts for photon losses due to microshutter bar occlusion in MOS observations, providing necessary flux calibration for point source measurements after one-dimensional spectral extraction.

The NIRISS spectral data analyzed in this work were acquired through the publicly available GLASS-JWST-ERS Program, utilizing the catalog presented in Watson et al. [278]. These observations were obtained with the WFSS mode, using the F115W, F150W, and F200W blocking filters to cover a wavelength range $\approx 1.0 - 2.2\,\mu$m at a spectral resolution of $R \approx 150$ at 1.4 μm. The wavelength coverage of these filters does not overlap, and the throughput is only greater than 50% of the peak value within the windows 1013-1283 Å, 1330-1671 Å, and 1751-2226 Å respectively. In the F115W and F150W filters, approximately 1.5 hours were dedicated to direct imaging, with a further 3 hours in each of the two orthogonal grism orientations. Total exposure times in the F200W filter were half the duration. Galaxies were detected using a stack of the direct imaging in all three filters, drizzled to a pixel scale of 0.03 arcsec/pixel. This was performed without an additional diffuse background subtraction step, and with a threshold 1σ above the existing background after filtering with a matched convolution kernel, in order to better capture the extended intracluster light present in this field. The NIRISS candidate identifications referenced throughout our analysis correspond to those in Watson et al. [278], and for further details regarding the data reduction process, we direct readers to that work.

The entire sample of 1D spectra used in this work was extracted from the 2D data using the optimal extraction technique described by Horne [129]. Specifically, for NIRISS, the extraction profiles were obtained by summing the 2D spectra along the dispersion axis, as described in Watson et al. [278]. For NIRSpec, a gaussian profile was fitted to the signal along the dispersion axis to perform the extraction, following the procedures detailed in Mascia et al. [175], Price et al. [218].

For our analysis, we focused on strong emission lines widely used as diagnostics of galaxy properties: [O II] λ3727, Hβ λ4861, [O III] λ4959, [O III] λ5007, and Hα λ6563[3]. Given the

[3]Note that at the spectral resolution of NIRISS, it is not possible to separate Hα λ6563 from either [N II] λ6548 or [N II] λ6583. Throughout this paper, what we refer to as Hα λ6563 more precisely denotes the blended complex of all three lines.

wavelength coverage of NIRISS, we identified the redshift interval $z \in [0.5, 4.9]$ as the range where any of these emission features would be observable.

Both the NIRISS and NIRSpec catalogs were filtered to select sources within this redshift range that exhibited at least two of the emission lines of interest, each detected at a signal-to-noise ratio $S/N > 5$. To ensure the robustness of our final sample, we conducted a detailed visual inspection of the spectra from both instruments, confirming the presence and strength of the emission lines and verifying their suitability for further analysis. We also examined the drizzled NIRISS detection image, as described above, to assess the morphological properties of each candidate galaxy.
Galaxies were then classified as compact if 50% of their flux distribution was contained within 80% of half the NIRSpec slit height (with a full slit height of 0.46 arcsec); otherwise, they were considered extended.
The complete sample of 12 candidate galaxies selected for this study is shown in Fig.6.1, with a summary of their properties and detected emission lines provided in Tab.6.1.

6.4 Slit loss analysis-characterization

6.4.1 Characterization of Emission Line Fluxes and Equivalent Widths

For each galaxy in our sample, we conducted a systematic analysis of strong emission lines (detailed in columns (8)-(12) of Tab.6.1) using spectral data from both NIRISS and NIRSpec instruments.

The spectral analysis involved fitting emission lines through χ^2 minimization to determine integrated fluxes from 1D spectra obtained from both NIRISS and NIRSpec instruments. We established flat continuum baselines by estimating levels from two regions surrounding each emission line of interest, incorporating Monte Carlo (MC) sampling of these continuum regions to account for uncertainties in continuum placement and obtain robust uncertainty estimates for the spectral line measurements.

The MC sampling consisted of m^n total iterations, where $m = 4$ denotes the number of continuum parameters varied (comprising the left and right mean distances from the emission line center, and the widths of both regions) and n represents the number of MC iterations. Each emission line was inspected separately by assigning an initial region (both left and right of the emission line peak) from where to start the sampling. This was done to ensure that the initial region captured the characteristics of the spectrum in the vicinity of the line to be fitted, to avoid arbitrarily and systematically assigning exploration regions where there is a clear drop in the spectrum or a feature that would

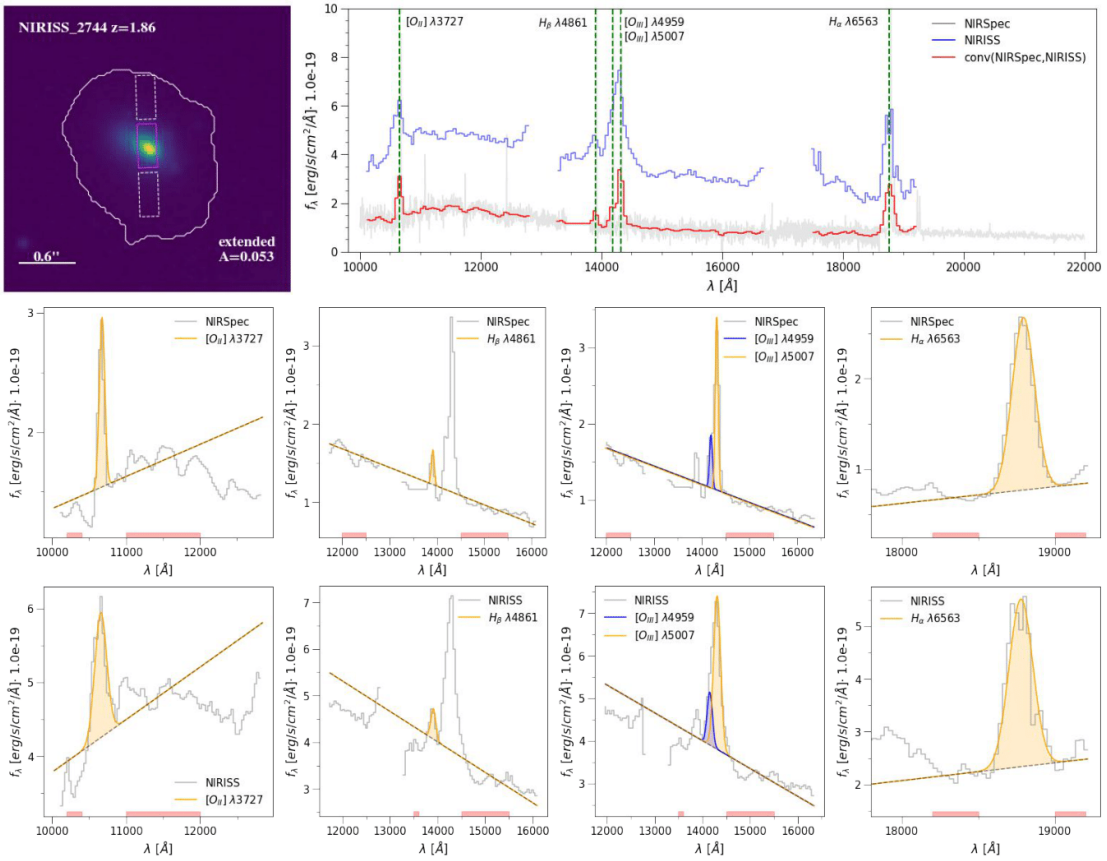

FIGURE 6.2: Example Analysis: Spectroscopically Confirmed Candidate NIRISS ID 2744 at z=1.86. Top left: NIRISS image cutout (0.6 arcsec scale bar shown) with the NIRISS segmentation map (solid white contour) and NIRSpec slits assigned to the target (solid magenta and dash-dot white). Top right: Combined spectral data showing the original NIRSpec spectrum (gray), the NIRISS spectrum (blue), and the NIRSpec spectrum after convolution to match the NIRISS resolution (red), with identified emission lines marked. Bottom panels: Detailed emission line fits for all the line matched between the two instruments. The shaded red regions in each panel indicate an example of the continuum estimation regions computed for a single iteration of the MC sampling.

distort our analysis. Although the initial range differed for each emission line, a fixed minimum width of 100 Å was established for all regions, below which the MC sampling could not go. This resampling methodology enabled exploration of the parameter space for possible continuum definitions and assessment of how these variations influence our line measurements.

Fig.6.2 shows one iteration from the m^n MC sampling, illustrating the analysis performed for a representative galaxy. The top-left panel shows the galaxy's imaging, while the top-right displays the extracted spectrum, with measurements from NIRISS (blue) and NIRSpec (grey). To ensure consistent spectral resolution between instruments, we convolved the higher-resolution NIRSpec MSA data to match the lower-resolution NIRISS

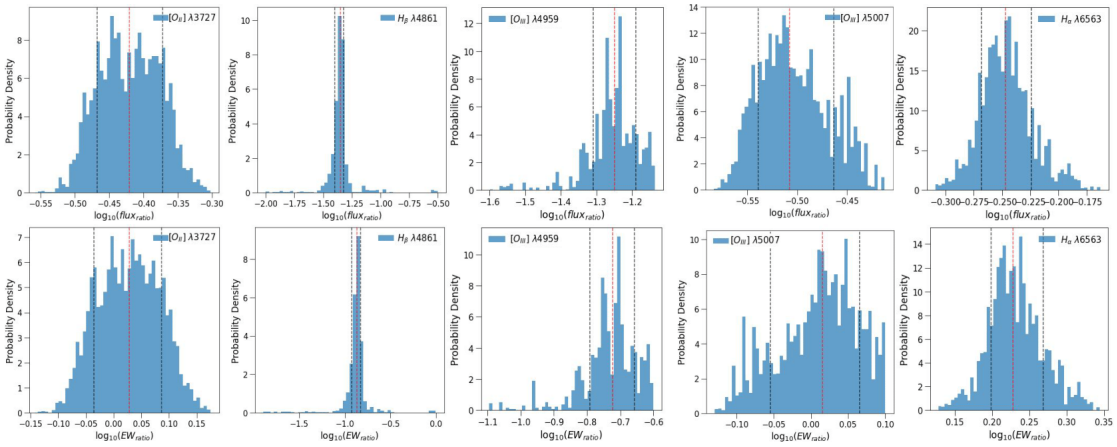

FIGURE 6.3: Probability density distributions of $\text{flux}_{\text{ratio}}$ (top row) and EW_{ratio} (bottom row) for the emission lines of the source presented in Fig.6.2. The distributions are shown in logarithmic scale for visualization purposes and represent the marginalized probability densities obtained by varying the continuum estimation regions. In red the 50th and from left to right in black 16th and 84th percentiles.

WFSS mode using a gaussian kernel with width corresponding to the quadrature difference between the resolution elements. The resulting convolved NIRSpec spectrum (red) was used throughout our comparative analysis to prevent resolution-dependent biases in line measurements and continuum features. The two subsequent rows illustrate a snapshot of the MC sampling of the continuum, highlighting the red-shaded regions on either side of the emission line used to define the continuum (in that precise snapshot), along with the corresponding emission line fit. The [O III] doublet was fitted simultaneously, with the flux ratio [O III] λ5007/[O III] λ4959 fixed at 3:1 according to quantum mechanical transition probabilities.

For each MC iteration and pair of matched emission lines, we measured the integrated flux and the equivalent width (EW) defining two ratios as:

$$\text{flux}_{\text{ratio}} = \frac{\text{flux}_{\text{NIRSpec}}}{\text{flux}_{\text{NIRISS}}} \tag{6.1a}$$

$$\text{EW}_{\text{ratio}} = \frac{\text{EW}_{\text{NIRSpec}}}{\text{EW}_{\text{NIRISS}}} \tag{6.1b}$$

Fig.6.3 presents the resulting probability distributions of the integrated flux and EW ratios for each emission line when considering different continua. We adopted the $50th$ percentile as our fiducial value, with 1σ uncertainties derived from the 16th and 84th percentiles of these distributions.

In both Fig.6.4 and Fig.6.5, we plot against emission line rest-frame wavelength (panel a), galaxy AB magnitude in the F200W filter (panel b), NIRSpec slit coverage as a fraction of the NIRISS segmentation map area (panel c) and the effective radius of the

galaxy R_e (panel d). In all the panels we visually differentiate galaxies categorized as compact (blue) and extended (red) and with different markers indicating their observation program (see Sec.6.3).

Our analysis in Fig.6.4 reveals a clear dichotomy between compact and extended sources. Compact sources show $\text{flux}_{\text{ratio}}$ measurements with remarkable consistency, where NIRSpec integrated fluxes are comparable to or slightly exceed NIRISS measurements, agreeing within 1σ uncertainties.

Extended sources exhibit systematically different behavior. NIRISS spectra consistently yield higher integrated flux measurements than NIRSpec observations, resulting in $\log_{10}(\text{flux}_{\text{ratio}}) < 0$. This persistent trend indicates a fundamental difference in how the two instruments measure flux for extended sources.

The systematic flux measurement differences observed between NIRISS and NIRSpec arise from their fundamentally distinct instrumental configurations, as shown in Fig.6.1. NIRSpec employs fixed-slit spectroscopy through a microshutter array (0.2×0.46 arcsec), confining observations to flux within narrow slits (magenta rectangles in Fig.6.1). While this configuration enables simultaneous multi-target observations with minimized background noise and maximized spectral resolution, it can systematically underestimate total flux for extended sources due to slit losses.

NIRISS, by contrast, utilizes slitless spectroscopy that disperses light across its entire field of view without spatial constraints. This approach captures the total emission within each galaxy's segmentation map (solid white contours in Fig.6.1), providing integrated flux measurements across the complete source morphology. The resulting systematic differences are particularly pronounced for extended galaxies, as demonstrated in panel (c) and (d) of Fig.6.4 , where larger sources show NIRISS flux measurements consistently exceeding those from NIRSpec.

These instrumental effects introduce systematic uncertainties that must be carefully considered when interpreting spectroscopic measurements. Recent studies have quantified these biases: He et al. [122] demonstrated that emission line fitting on 1D NIRISS spectra can systematically underestimate true line fluxes by approximately 30% due to morphological broadening effects inherent to slitless spectroscopy, while Jiang et al. [135] quantified significant systematic uncertainties from NIRSpec slit losses through direct comparison with IFU data. The comparable integrated flux measurements between instruments for compact sources further validates this interpretation, as these targets are well-suited to NIRSpec's spatially constrained aperture.

On the other hand our analysis of the EW_{ratio} presented in Fig.6.5 reveals a different pattern compared to the $\text{flux}_{\text{ratio}}$ analysis. While $\text{flux}_{\text{ratio}}$ showed a clear dichotomy

between compact and extended sources with persistent systematic offsets, the $\mathrm{EW_{ratio}}$ exhibit more intricate behavior.

For compact sources, EW ratios display substantial dispersion about around $\log_{10}(\mathrm{EW_{ratio}}) = 0$, with several points deviating from it. This pattern diverges from the $\mathrm{flux_{ratio}}$ behavior observed in compact sources, where measurements exhibited closer adherence to the null hypothesis of equivalent flux recovery. The amplified scatter suggests that continuum-normalized measurements introduce systematic variability for compact objects that does not affect absolute flux determinations.

Extended sources exhibit a different behavior compared to their $\mathrm{flux_{ratio}}$ counterparts. While the logarithmic $\mathrm{flux_{ratio}}$ values were consistently negative (indicating that NIRISS systematically measured higher fluxes), the $\mathrm{EW_{ratio}}$ measurements for extended sources are more scattered between approximately -1 and 0.5 in logarithmic scale. This difference indicates that continuum normalization mitigates the systematic flux differences observed between instruments, though morphological characteristics such as more or less clumpy regions within these extended sources likely contribute to the remaining scatter, as different instruments might sample these non-uniform structures differently. Several extended source measurements now appear consistent between instruments within uncertainties, which was not seen in $\mathrm{flux_{ratio}}$ analysis.

The wavelength panel (Fig.6.5) illustrates this distinction, where extended source $\mathrm{EW_{ratio}}$ cluster closer to zero than their corresponding $\mathrm{flux_{ratio}}$, especially when looking at longer restframe wavelengths. This suggests that while NIRSpec captures less total flux from extended sources due to slit limitations, the line to continuum ratio remains more comparable between instruments than absolute flux measurements would predict.

The $\mathrm{EW_{ratio}}$ relationship with magnitude, coverage area and effective radius shows reduced systematic offsets compared to $\mathrm{flux_{ratio}}$ measurements, indicating that although NIRSpec captures less total flux from extended sources, the line-to-continuum ratio remains more consistent between instruments.

Our results suggest that EW measurements provide more consistent cross-instrument comparisons than direct flux measurements, particularly for extended sources. This improvement likely occurs because EW normalization by local continuum compensates for aperture-dependent light collection differences between NIRISS and NIRSpec. For spectroscopic studies using these instruments, flux measurements require calibration adjustments when analyzing extended sources, while EW-based analyses offer more reliable line strength comparisons, though wavelength-dependent effects should still be considered.

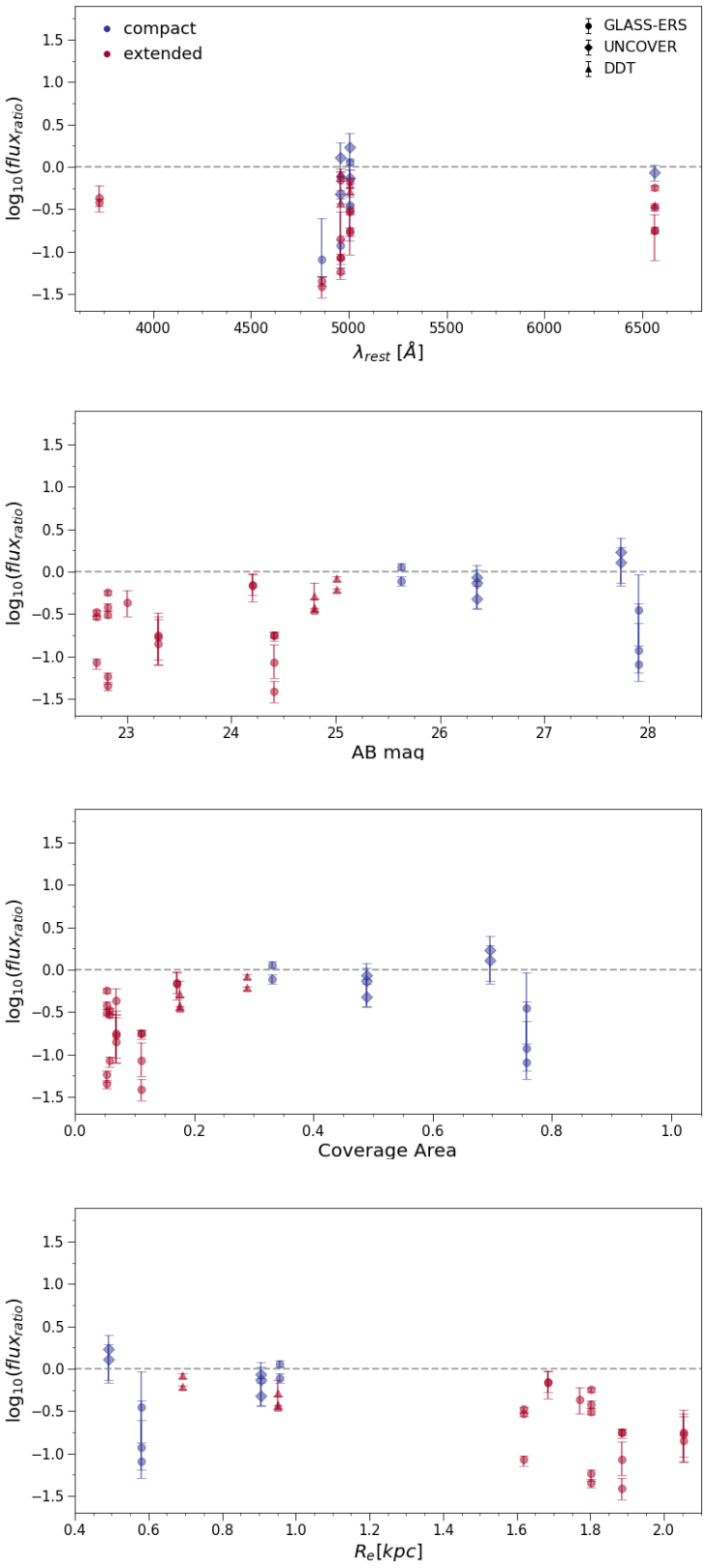

FIGURE 6.4: All panels show results for the observed sources from three different JWST programs, color-coded by subsample (compact or extended). Error bars indicate 1σ uncertainties calculated using the methodology described in Sec.6.4.1. In these panels, we see that measurements for compact sources from NIRISS and NIRSpec are in good agreement. However, for more extended objects (shown on the left side of panel c), the NIRSpec measurements are systematically lower, as expected.

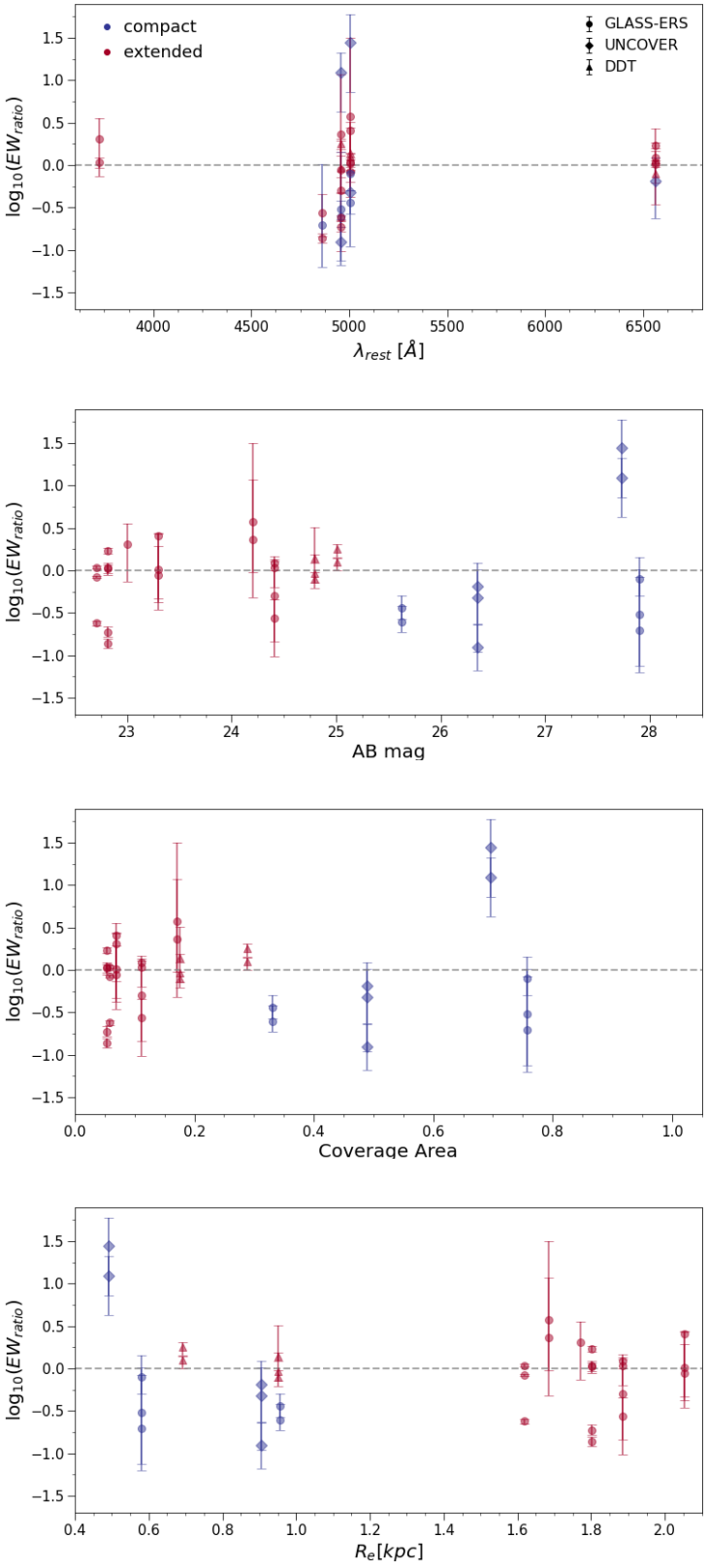

FIGURE 6.5: Same format as Fig.6.4, but showing the logarithmic $\mathrm{EW_{ratio}}$ as a function of various quantities. Unlike the dichotomy seen in $\mathrm{flux_{ratio}}$, both compact and extended sources display similar scatter in $\mathrm{EW_{ratio}}$ with no strong trends.

6.4.2 Flux Characterization in Emission Line Maps

To further investigate the flux measurement differences between the two instruments, we focused on the impact of NIRSpec slit positioning on extended sources (defined in Sec.6.3, see also the cutouts in Fig.6.1). For this analysis, we utilized the NIRISS emission line maps of [O II] λ3727, Hβ λ4861, [O III] λ5007, and Hα λ6563. It should be noted that separate emission line maps for [O III] λ4959 and [O III] λ5007 were not available; instead, a single map represents the combined doublet emission Brammer [42], Watson et al. [278].

We treated each galaxy under examination as a primary observation source, considering an optimal scenario for NIRSpec slit positioning. We used the galaxy coordinates (RA and DEC) from the catalog in Tab.6.1 as reference. To account for potential uncertainties in aligning the slit with the galaxy's center, we sampled points from a gaussian distribution centered on these coordinates with $\sigma = 0.12$ arcsec and further assumed precise NIRSpec slit positioning capability within a radius of $r = 0.1$ arcsec around the slit's central point. This approach ensured plausible observations targeting bright galactic regions around the nominal center rather than peripheral points near the edge of the galaxy's segmentation map. For our MC sampling, we generated $n = 10^4$ simulated NIRSpec slits, allowing the slit rotation to vary randomly between $[0, \pi]$.

Fig.6.6 presents the emission line map cutouts for each extended galaxy (solid white contour representing the NIRISS segmentation map), overlaid with the original NIRSpec slit (magenta) and the slit associated with the most probable ($50th$ percentile) $\text{flux}_{\text{ratio}}$ measurements from the MC sampling (black). Each imaging panel is accompanied by the corresponding $\text{flux}_{\text{ratio}}$ probability distribution obtained from the MC sampling, with color-coded vertical lines including the 1σ uncertainties. Numerical results from this analysis are presented in Tab.6.2.

The probability distribution panels clearly demonstrate that slit positioning is crucial for capturing the emitted flux from the observed galaxy. A more peripheral or uncentered positioning of the slit can significantly underestimate and compromise the spectral analysis, resulting in incorrect characterization of the source's redshift and brightness. This occurs because the slit captures only a fraction of the total flux, making measurements highly sensitive to the alignment between the slit and the galaxy's bright regions as observed in imaging data.

While the ideal approach would involve optimizing slit positioning and orientation for each individual galaxy, such a method is pragmatically unfeasible due to the substantial time constraints it would impose. Moreover, such an approach would fundamentally

contradict the primary design philosophy of NIRSpec, which is engineered to enable simultaneous spectral observations of multiple astronomical targets.

Despite these limitations, our findings establish a quantitative framework for comparing flux measurements between NIRISS and NIRSpec. Given the assumptions required to approximate a realistic observation in the scenario when a slit is assigned to a candidate with an unfixed position and orientation in NIRSpec it is expected that flux measurements for an extended galaxy will likely be underestimated compared to those obtained with NIRISS. The extent of this underestimation is proportional to the size of the source relative to the MSA slit.

In the two panels of Fig.6.7, we compare measurements of the $\text{flux}_{\text{ratio}}$ for extended sources, showing results from the emission line maps (green) and the spectral fit (magenta) as a function of both coverage area and effective radius R_e. We find that when the galaxy area coverage exceeds 15%, or for sources with $R_e < 1kpc$, the $\text{flux}_{\text{ratio}}$ consistently remains above 63%, with both spectrographs yielding results that agree within their respective uncertainties. For coverage below 15%, corresponding to more extended sources, the recovered flux displays greater variation, ranging from 24% to 63%. In this regime, emission line maps generally produce higher flux ratios compared to those derived from spectral fitting techniques.

We also examined the emission line ratio between Hα λ6563 and [*O* III] λ5007 for four extended galaxies presenting these emission lines NIRISS[1694], NIRISS[2355] and NIRISS[998] from Fig.6.6. We applied the previously detailed MC sampling method to investigate whether this line ratio maintains consistency across different galactic regions when observed using various methodologies, or if the different emission line maps exhibit intrinsic characteristics that cause the line ratio to vary rather than remain constant.

Fig.6.8 displays the probability distribution of the Hα λ6563/[*O* III] λ5007 line ratio across various measurement scenarios, including both emission line fitting using the 1D spectra (see Sec.6.4.1) and measurements derived from emission line maps for each galaxy using both spectrographs. The vertical lines representing different instruments and methods for measuring the Hα λ6563/[*O* III] λ5007 emission line ratio are scattered across the probability distribution, producing inconsistent measurements that vary by $\sim$ 0.3 dex. Across all four panels, the measurements obtained from the precise spectral fit described in Sec.6.4.1 underestimate the probability distribution, providing lower values for the emission line ratio. While we expected the ratio to be similar within a margin of error regardless of measurement method or instrument, our findings suggest that line ratios are significantly influenced by observational position.

$\text{flux}_{\text{ratio}}$ from MC Sampling					
NIRISS ID (1)	A (2)	$[O\,\text{II}]\,\lambda3727$ (3)	$\text{H}\beta\,\lambda4861$ (4)	$[O\,\text{III}]\,\lambda5007$ (5)	$\text{H}\alpha\,\lambda6563$ (6)
2744	0.053	$0.243^{+0.023}_{-0.047}$	$0.252^{+0.027}_{-0.048}$	$0.354^{+0.028}_{-0.073}$	-
1694	0.058	-	-	$0.584^{+0.055}_{-0.097}$	$0.411^{+0.035}_{-0.063}$
17	0.069	-	-	$0.509^{+0.193}_{-0.241}$	$0.425^{+0.157}_{-0.152}$
2355	0.112	-	$0.443^{+0.15}_{-0.21}$	$0.634^{+0.053}_{-0.176}$	$0.346^{+0.044}_{-0.039}$
1407	0.171	-	-	$0.475^{+0.192}_{-0.131}$	-
998	0.176	-	-	$0.708^{+0.051}_{-0.213}$	$0.712^{+0.056}_{-0.235}$
1504	0.289	-	-	$0.638^{+0.079}_{-0.161}$	-

TABLE 6.2: Emission line $\text{flux}_{\text{ratio}}$ in Eq.6.1a obtained from different emission line maps of extended galaxies. Column descriptions: (1) Associated NIRISS ID; (2) Area of the galaxy covered by the slit; $\text{flux}_{\text{ratio}}$ from emission line maps (3)-(6).

These observed inconsistencies in $\text{H}\alpha\,\lambda6563/[O\,\text{III}]\,\lambda5007$ line ratio measurements have critical implications for galactic parameter estimation. These variations could lead to substantial discrepancies in metallicity derivations depending on the instrument and methodology used (i.e., Khoram & Belfiore [143], Nagao et al. [182]), challenging the robustness of metallicity estimates, especially for dust-obscured systems where IR lines are often favored. For star formation analyses, the positive correlation between SFR and $H\alpha + H\beta$ luminosities means variations in line ratio measurements directly propagate into SFR uncertainties (Kennicutt [142]). This is further complicated by the finding that emission lines experience more attenuation than the continuum, with attenuation increasing with stellar mass, thus amplifying the impact of methodological differences on derived SFR and dust attenuation values (Villa-Vélez et al. [276]).

Our results suggest that variations in line ratios are influenced by instrumental factors such as slit positioning, as well as data reduction and post-processing elements including segmentation map definition, contamination levels, and background attenuation. These methodological differences can artificially inflate uncertainties in both metallicity and SFR derivations. These findings underscore the need for standardized measurement protocols and careful cross-calibration when combining multi-instrument datasets, particularly in JWST studies comparing NIRISS and NIRSpec results.

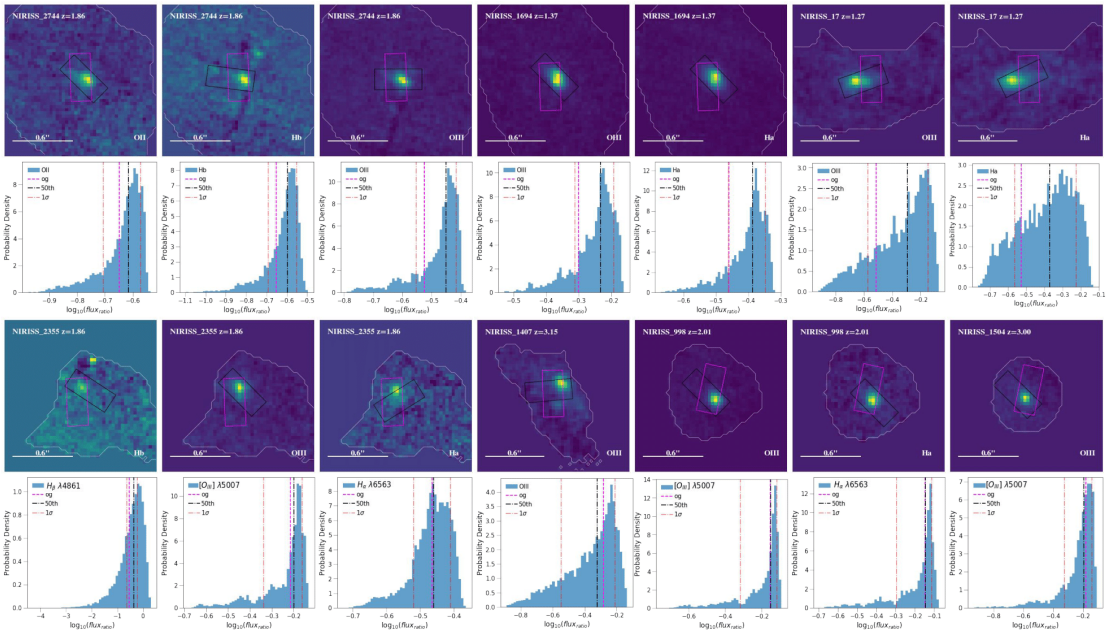

FIGURE 6.6: NIRISS imaging cutouts from the corrisponding emission line maps for extended galaxies. Solid white contours represent NIRISS segmentation maps, while rectangles indicate NIRSpec slit positioning. Magenta (og: original) denotes the fixed position from the original observation from the corresponding JWST program. The black one represents the slit corresponding to the most probable $\text{flux}_{\text{ratio}}$ (50th percentile) obtained from the MC analysis conducted on the associated emission line maps. For each cutout, the corresponding histogram displays the probability density of the $\text{flux}_{\text{ratio}}$ (Eq.6.1a) in logarithmic scale derived from the MC analysis and color-coded the same way with associated 1σ uncertainties in red.

6.5 Summary and Conclusions

This investigation examines 12 spectroscopically confirmed galaxies located in the Abell 2744 field, observed using both NIRISS and NIRSpec instruments on JWST. The dataset comprises observations from three JWST programs: GLASS-JWST-ERS (PID 1324), JWST-DDT (PID 2756), and UNCOVER-JWST (PID 2561). Our primary aim was to characterize spectroscopic measurement variations between these JWST spectrographs using both precise emission line fitting and Monte Carlo sampling-based analysis of emission line maps. Our key findings are:

- Analysis of strong emission line fluxes revealed no significant trend in flux ratios with respect to rest-frame wavelength or AB magnitude of the observed sources.
- We identified a distinct dichotomy in $\text{flux}_{\text{ratio}}$ between compact and extended sources. Compact objects showed consistent measurements between instruments, while extended galaxies exhibited systematically lower flux measurements with NIRSpec compared to NIRISS, attributable to the limited coverage of the NIRSpec MSA slit.

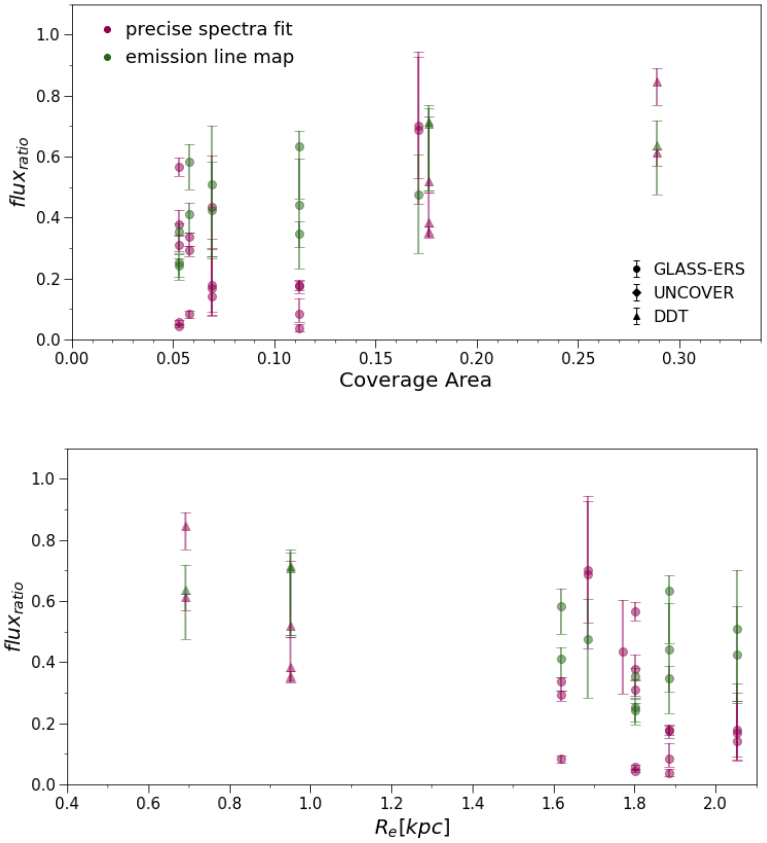

FIGURE 6.7: $\text{flux}_{\text{ratio}}$ as a function of slit coverage area and effective radius R_e for extended galaxies.Results are color-coded by measurement method: "precise spectra fit" refers to measurements from the analysis in Sec.6.4.1, while "emission line map" indicates results from the analysis in Sec.6.4.2 (also presented in Tab.6.2).

- EW ratios behaved differently than flux ratios. Compact sources showed greater scatter, but extended sources had better NIRISS-NIRSpec agreement for equivalent widths than for fluxes.

- Continuum normalization in EW calculations appears to mitigate the systematic offsets observed in direct flux comparisons, particularly for extended sources, suggesting that line-to-continuum ratios remain more consistent between instruments despite aperture differences.

- Our findings indicate that EW-based analyses may provide more reliable cross-instrument comparisons than direct flux measurements, especially for extended sources, though wavelength-dependent calibration effects should still be considered.

- Further investigation using emission line maps demonstrated that when the MSA slit covers at least 15% of a galaxy's segmentation map area or the effective radius of the galaxy is $R_e < 1kpc$, NIRSpec can recover a minimum of 63% of the flux measured by NIRISS.

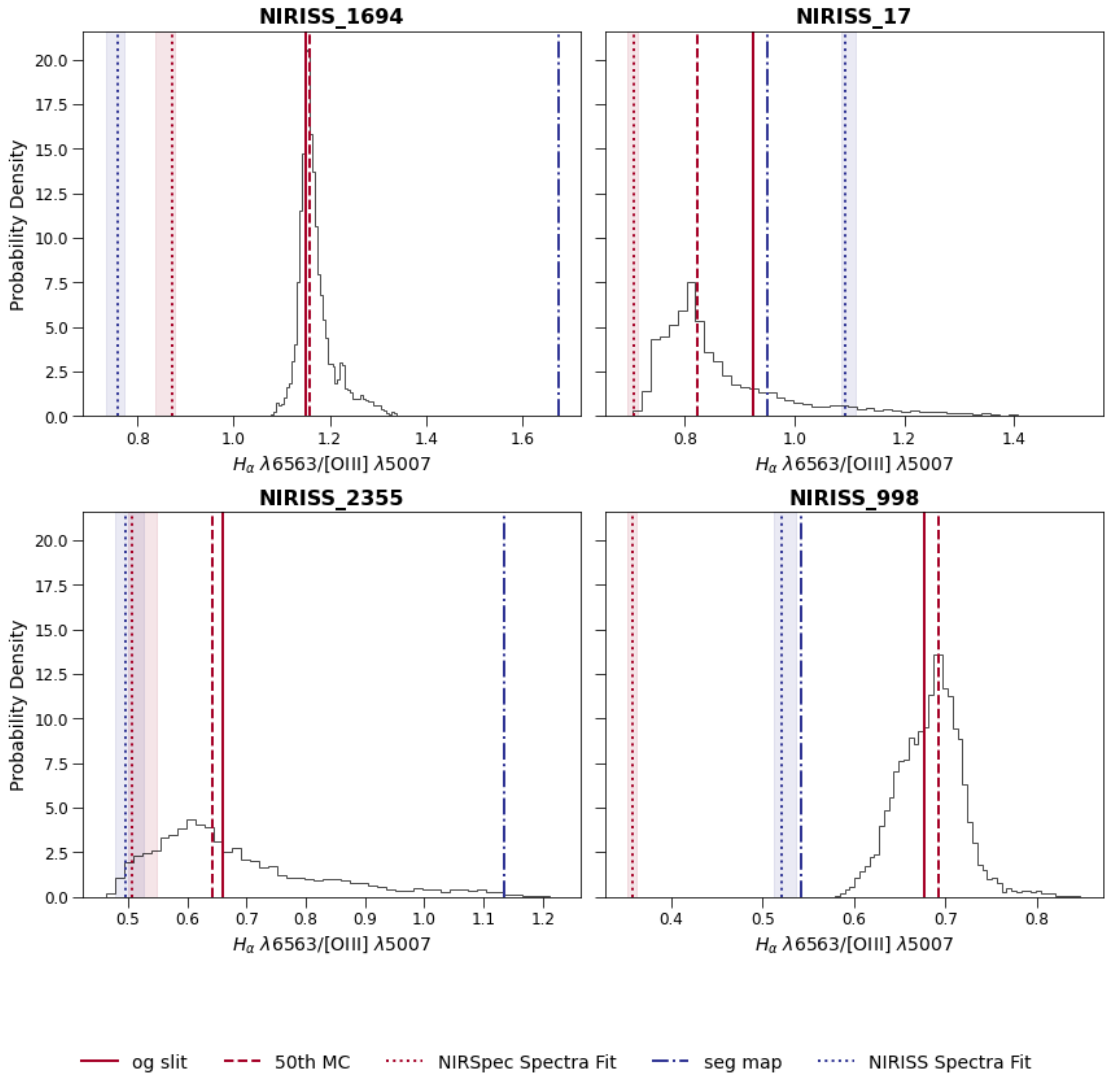

FIGURE 6.8: Probability density histogram of the Hα λ6563 to [O III] λ5007 emission line ratio for four extended sources. Multiple measurement methods are shown. NIRISS measurements: "seg map" represents the flux ratio measured within the segmentation map across both emission line maps. NIRSpec measurements: "og slit" shows flux measurements using the original assigned slit as the observed region, while "50th MC" indicates the 50*th* percentile from MC sampling (Sec.6.4.2). "Spectra Fit" for both instruments denotes the emission line ratio derived from precise emission line fitting detailed in Sec.6.4.1.

- Our analysis reveals that the Hα λ6563/[*O* III] λ5007 line ratio varies significantly ($\sim$ 0.3 dex) across measurement methods and instruments, potentially leading to discrepancies in metallicity and SFR estimates. This underscores the need for standardized protocols when interpreting spatially resolved galaxy properties from multi-instrument datasets.

Chapter 7

Summary and Conclusions

This doctoral thesis aims to apply galaxy clustering techniques to investigate galaxy populations and their complex relationships with dark matter halos (Wechsler & Tinker [279], Zheng et al. [292]). The primary motivation is to advance understanding of how galaxies trace dark matter halo evolution during the epoch of reionization and cosmic dawn (Park et al. [203], Robertson et al. [230]). Leveraging observations from the Hubble Space Telescope Legacy Fields, this work develops and implements a comprehensive analytical model to study galaxy clustering.

The field underwent a major transformation with the launch of the James Webb Space Telescope, which delivered discoveries that challenge the standard ΛCDM cosmological model, particularly in the context of early galaxy formation (Boylan-Kolchin [40], Gupta [113]). JWST revealed unexpectedly luminous and massive galaxies at redshifts $z \gtrsim 8$, creating tension with theoretical predictions and suggesting a more rapid assembly of baryonic matter than typically forecast by ΛCDM-based models (Casey et al. [51], Labbé et al. [149], Naidu et al. [183]). This discrepancy highlights potential gaps in understanding of massive galaxy formation at cosmic dawn (Steinhardt et al. [258]).

Motivated by these findings, the research was extended to incorporate JWST observations of the same regions as the Legacy Fields, allowing exploration of the faint end of the luminosity function and the evolution of dark matter halos in relation to galaxy luminosity. Through measurements of the Angular Correlation Function, the halo occupation distribution and galaxy substructure were investigated, focusing on achieving high signal-to-noise across a wide luminosity range (Berlind & Weinberg [23], Cooray & Sheth [62]). This work has provided significant insights into the fundamentals of galaxy formation and evolution in the early Universe, with particular emphasis on the epoch of reionization and cosmic dawn (Robertson et al. [230]).

To expand from the main focus of galaxy clustering measurements, the doctoral research also encompassed the study of galaxy mergers at high redshift using JWST/NIRCam imaging (Conselice et al. [58], Lotz et al. [165]). Additionally, a project examining measurements and performance differences between NIRISS WFSS and NIRSpec MSA was conducted, further broadening the scientific perspective and technical expertise in handling advanced observational datasets (Jakobsen et al. [133], Willott et al. [284]).

7.1 Galaxy clustering measurement at z∼8 and beyond

This research established a comprehensive framework for understanding galaxy clustering during cosmic dawn through three interconnected studies spanning from HST to JWST observations. The foundation was built with a novel methodology for measuring the two-point correlation function in narrow pencil beam surveys with varying depths (Ch.2). Initially applied to HST Legacy Fields data, this approach analyzed $N_g = 160$ Lyman break galaxy candidates in the CANDELS GOODS datasets, yielding the first clustering measurements extending to $z \sim 8$. At $z = 7.7$, bright sources ($M_{UV} < -19.8$) exhibited a galaxy bias of $b = 9.33 \pm 4.90$ and correlation length of $r_0 = 10.74 \pm 7.06$ h^{-1}Mpc, indicating host halos of $M_h \sim 10^{11.5} M_\odot$ with duty cycles approaching unity.

The advent of JWST dramatically expanded observational reach, enabling clustering analysis of much fainter galaxies deeper into the epoch of reionization. Applying the established methodology to JADES data revealed the viability of clustering measurements even as LBG candidates become increasingly rare at high redshift (Ch.3). With samples reaching $N_g \simeq 7000$ photometrically-selected sources, measurements extended beyond $z > 10$ for the first time. At $z = 10.6$, faint galaxies ($M_{F200W} < -17.3$) showed a galaxy bias of $b = 9.6 \pm 1.7$, corresponding to host halos of $M_h \sim 10^{10.6} M_\odot$. Crucially, this dataset enabled the first clustering analysis of dwarf galaxies during reionization, revealing a theoretically-predicted turning point in the bias-luminosity relation at $M_{F200W} \sim -16$.

Building on these results, an HOD analysis leveraged JWST's unprecedented capabilities to observe faint galaxy populations, employing a sample of $N_g \simeq 6500$ LBGs from JADES' second data release (Ch.4). This investigation determined an effective galaxy bias of $b_{\rm g}^{\rm eff} = 8.13^{+0.04}_{-0.02}$ at $z = 10.6$, representing the highest redshift clustering constraint to date. The HOD analysis revealed that characteristic halo masses decrease by over an order of magnitude from $\log_{10}(M_h/M_\odot) = 11.45^{+0.02}_{-0.04}$ at $z = 5.5$ to $\log_{10}(M_h/M_\odot) = 10.12^{+0.08}_{-0.06}$ at $z = 10.6$, while satellite fractions increase dramatically during cosmic dawn, reaching $\sim 28\%$ at $z \sim 8.5$ compared to $< 1\%$ at $z \sim 5-6$. This evolutionary picture demonstrates that similar UV luminosity populations inhabit progressively less massive

halos at higher redshifts, with stellar light-to-halo mass ratios rising by approximately a factor of ten from $z \sim 6$ to $z \sim 10$.

7.2 High redshift galaxies study from NIRCam imaging

We performed a comprehensive analysis of galaxy merger rates over the redshift range $4.0 < z < 9.0$, utilizing a methodology based on visual identification of mergers through detailed morphological parameter analysis (Ch.5). High-resolution JWST/NIRCam observations in the F150W and F200W broad-band filters were employed, focusing on low-to-moderate magnification ($\mu < 2$) regions within the Abell 2744 cluster field. These data were acquired from multiple programs: GLASS JWST-ERS-1324, UNCOVER JWST-GO-2561, and DDT-2756.

From a parent sample of 675 galaxies ($M_{UV} \in [-26.6, -17.9]$), we identified 64 merger candidates using Gini, M_{20}, and Asymmetry morphological parameters. This yielded a merger fraction of $f_m = 0.11 \pm 0.04$. Importantly, we observed no significant evolution in f_m across the entire redshift range, extending the previously reported constant trend at $z \lesssim 6$ into the epoch of reionization.

We also investigated potential correlations between redshift and differences in specific star formation rates between merging and non-merging galaxies. No significant relationship was found, with variations typically within 0.1 dex, attributable to sample variance and measurement uncertainties. Although preliminary, these results highlight the transformative potential of JWST infrared imaging for advancing our understanding of galaxy assembly through mergers during the epoch of reionization.

7.3 JWST NIRISS-NIRSpec Performance Analysis

This study aimed to quantify spectroscopic flux measurement differences between JWST's NIRISS WFSS (slitless) and NIRSpec MSA (slit-based) instruments by analyzing 12 spectroscopically confirmed galaxies in the Abell 2744 field, focusing on how source morphology and instrumental configurations affect flux recovery and emission line measurements (Ch.6). The analysis revealed a clear morphology-dependent dichotomy: compact objects showed comparable flux measurements between instruments (within 1σ uncertainties), while extended sources consistently exhibited higher flux in NIRISS measurements due to NIRSpec's slit losses. Quantitatively, NIRSpec recovered at least 63% of NIRISS-measured flux when the slit covered $> 15\%$ of the source or when the effective radius was $R_e < 1$ kpc, but only 24-63% for lower coverage or larger sources. While

flux ratios showed systematic offsets for extended sources, equivalent width measurements displayed better cross-instrument agreement, suggesting that continuum normalization mitigates aperture-dependent differences and that EW-based analyses provide more reliable cross-instrument comparisons. Critically, emission line ratio measurements (Hα λ6563/[O III] λ5007) varied by up to $\sim$0.3 dex between instruments and methods, with significant implications for metallicity and star formation rate characterizations, highlighting the need for standardized protocols when combining multi-instrument spectroscopic datasets for galaxy property derivations.

This work has established a robust observational foundation for understanding galaxy formation processes during the first billion years of cosmic history. By developing and validating novel clustering analysis methodologies capable of handling the complex observational challenges at $z > 8$, this thesis provides the essential framework needed to fully exploit the revolutionary capabilities of JWST and future wide-field surveys. The empirical constraints presented here—from the first clustering measurements at $z \sim 11$ to the quantification of evolving galaxy-halo relationships and merger processes—represent critical benchmarks that will guide theoretical models of early galaxy assembly and reionization physics. As JWST continues to unveil unprecedented samples of high-redshift galaxies with ever-increasing statistical power, and as complementary surveys like Euclid and the Roman Space Telescope come online, galaxy clustering is positioned to become the definitive observational tool for disentangling the complex interplay between dark matter halo properties and galaxy populations at cosmic dawn. The methodological advances and empirical frameworks established in this thesis will enable the astronomical community to transform these expanding datasets into fundamental insights about the emergence of galactic structure, ultimately resolving longstanding questions about how the first galaxies formed within the cosmic web and seeded the universe we observe today.

Bibliography

[1] Abel T., Bryan G. L., Norman M. L., 2002, Science, 295, 93

[2] Abraham R. G., Tanvir N. R., Santiago B. X., Ellis R. S., Glazebrook K., Bergh S. v. d., 1996, Monthly Notices of the Royal Astronomical Society, 279, L47

[3] Abraham R. G., van den Bergh S., Nair P., 2003, The Astrophysical Journal, 588, 218

[4] Ade P. A., et al., 2014a, Astronomy & Astrophysics, 571, A16

[5] Ade P. A., et al., 2014b, Astronomy & Astrophysics, 571, A16

[6] Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, The Astrophysical Journal, 619, 697

[7] Akhlaghi M., Ichikawa T., 2015, The Astrophysical Journal Supplement Series, 220, 1

[8] Albert L., et al., 2023, Publications of the Astronomical Society of the Pacific, 135, 075001

[9] Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, Monthly Notices of the Royal Astronomical Society, 310, 540

[10] Arrabal Haro P., et al., 2023, Nature Astronomy, 622, 707

[11] AstropyCollaboration 2022, The Astrophysical Journal, 935, 167

[12] Bahcall J. N., Soneira R. M., 1980, The Astrophysical Journals, 44, 73

[13] Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, The Astrophysical Journal, 304, 15

[14] Barkana R., Loeb A., 2001, Physics Reports, 349, 125

[15] Barnes J. E., 1988, The Astrophysical Journal, 331, 699

[16] Barnes J. E., Hernquist L., 1996, Astrophysical Journal, 471, 115

[17] Barone-Nugent R. L., et al., 2014, The Astrophysical Journal, 793, 17

[18] Bartelmann M., Schneider P., 2001, Physics Reportsp, 340, 291

[19] Behroozi P. S., Conroy C., Wechsler R. H., 2010, The Astrophysical Journal, 717, 379

[20] Behroozi P. S., Wechsler R. H., Conroy C., 2013, Astrophysical Journal, 770, 57

[21] Benoist C., Maurogordato S., da Costa L. N., Cappi A., Schaeffer R., 1996, The Astrophysical Journal, 472, 452

[22] Bergamini P., et al., 2023, The Astrophysical Journal, 952, 84

[23] Berlind A. A., Weinberg D. H., 2002, The Astrophysical Journal, 575, 587

[24] Bertin E., Arnouts S., 1996, Astronomy and Astrophysics Supplement, 117, 393

[25] Béthermin M., et al., 2020, Astronomy and Astrophysics, 643, A2

[26] Bezanson R., et al., 2022, arXiv e-prints, p. arXiv:2212.04026

[27] Bhatawdekar R., Conselice C. J., Margalef-Bentabol B., Duncan K., 2019, Monthly Notices of the Royal Astronomical Society, 486, 3805

[28] Bhowmick A. K., Campbell D., Di Matteo T., Feng Y., 2018, Monthly Notices of the Royal Astronomical Society, 480, 3177

[29] Bian F., et al., 2013, The Astrophysical Journal, 774, 28

[30] Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature Astronomy, 311, 517

[31] Böker T., et al., 2022, Astronomy and Astrophysics, 661, A82

[32] Böker T., et al., 2023, Publications of the Astronomical Society of the Pacific, 135, 038001

[33] Bond J. R., Kofman L., Pogosyan D., 1996, Nature Astronomy, 380, 603

[34] Bouwens R. J., et al., 2015, The Astrophysical Journal, 803, 34

[35] Bouwens R. J., et al., 2017, VizieR Online Data Catalog, p. J/ApJ/793/115

[36] Bouwens R. J., et al., 2021, The Astronomical Journal, 162, 47

[37] Bouwens R. J., et al., 2023, Monthly Notices of the Royal Astronomical Society, 523, 1036

[38] Boyett K., et al., 2024, Nature Astronomy, 8, 657

[39] Boylan-Kolchin M., 2023a, Nature Astronomy, 7, 731

[40] Boylan-Kolchin M., 2023b, Nature Astronomy, 7, 731

[41] Boylan-Kolchin M., Weisz D. R., 2021, Monthly Notices of the Royal Astronomical Society, 505, 2764

[42] Brammer G., 2023, grizli, Zenodo, doi:10.5281/zenodo.1146904

[43] Brammer G. B., van Dokkum P. G., Coppi P., 2008, The Astrophysical Journal, 686, 1503

[44] Bromm V., Coppi P. S., Larson R. B., 2002, Astrophysical Journal, 564, 23

[45] Bruzual G., Charlot S., 2003, Monthly Notices of the Royal Astronomical Society, 344, 1000

[46] Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, Monthly Notices of the Royal Astronomical Society, 321, 559

[47] Bunker A. J., et al., 2023, arXiv e-prints, p. arXiv:2306.02467

[48] Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, The Astrophysical Journal, 533, 682

[49] Carnall A., McLure R., Dunlop J., Davé R., 2018, Monthly Notices of the Royal Astronomical Society, 480, 4379

[50] Carniani S., et al., 2024, arXiv preprint arXiv:2405.18485

[51] Casey C. M., et al., 2024, The Astrophysical Journal, 965, 98

[52] Cibinel A., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, 5631

[53] Colless M., et al., 2001, Monthly Notices of the Royal Astronomical Society, 328, 1039

[54] Conroy C., Wechsler R. H., Kravtsov A. V., 2006, The Astrophysical Journal, 647, 201

[55] Conselice C. J., 2014, Annual Review of Astronomy and Astrophysics, 52, 291

[56] Conselice C. J., Arnold J., 2009, Monthly Notices of the Royal Astronomical Society, 397, 208

[57] Conselice C. J., Bershady M. A., Jangren A., 2000, The Astrophysical Journal, 529, 886

[58] Conselice C. J., Bershady M. A., Dickinson M., Papovich C., 2003, The Astronomical Journal, 126, 1183

[59] Conselice C. J., Yang C., Bluck A. F. L., 2009, Monthly Notices of the Royal Astronomical Society, 394, 1956

[60] Cooray A., 2005, Monthly Notices of the Royal Astronomical Society, 364, 303

[61] Cooray A., 2006, arXiv e-prints, pp astro–ph/0601090

[62] Cooray A., Sheth R., 2002, Physics Reports, 372, 1

[63] Correa C. A., Wyithe J. S. B., Schaye J., Duffy A. R., 2015, Monthly Notices of the Royal Astronomical Society, 452, 1217

[64] Cyburt R. H., Fields B. D., Olive K. A., Yeh T.-H., 2016, Reviews of Modern Physics, 88, 015004

[65] D'Eugenio F., et al., 2024, arXiv e-prints, p. arXiv:2404.06531

[66] Dalmasso N., Trenti M., Leethochawalit N., 2023, Monthly Notices of the Royal Astronomical Society, p. stad3901

[67] Dalmasso N., Leethochawalit N., Trenti M., Boyett K., 2024a, Monthly Notices of the Royal Astronomical Society, 533, 2391

[68] Dalmasso N., et al., 2024b, Monthly Notices of the Royal Astronomical Society, 533, 4472

[69] Dalmasso N., et al., 2025, Monthly Notices of the Royal Astronomical Society, 544, 1915

[70] Davis M., Huchra J., 1982, Astrophysical Journal, 254, 437

[71] Davis M., Peebles P. J. E., 1983, The Astrophysical Journal, 267, 465

[72] Davis M., et al., 2007, The Astrophysical Journall, 660, L1

[73] De Propris R., Conselice C. J., Liske J., Driver S. P., Patton D. R., Graham A. W., Allen P. D., 2007, The Astrophysical Journal, 666, 212

[74] Dekel A., Birnboim Y., 2006, Monthly Notices of the Royal Astronomical Society, 368, 2

[75] Di Matteo T., Springel V., Hernquist L., 2005, Nature Astronomy, 433, 604

[76] Dicke R. H., Peebles P. J. E., Roll P. G., Wilkinson D. T., 1965, The Astrophysical Journal, 142, 414

[77] Donnan C. T., et al., 2023, Monthly Notices of the Royal Astronomical Society, 518, 6011

[78] Doyon R., et al., 2012, in Clampin M. C., Fazio G. G., MacEwen H. A., Oschmann Jr. J. M., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8442, Space Telescopes and Instrumentation 2012: Optical, Infrared, and Millimeter Wave. p. 84422R, doi:10.1117/12.926578

[79] Dressel L., 2021, in , Vol. 13, WFC3 Instrument Handbook for Cycle 29 v. 13. p. 13

[80] Duncan K., et al., 2019, The Astrophysical Journal, 876, 110

[81] Dunlop J. S., et al., 2021, PRIMER: Public Release IMaging for Extragalactic Research, JWST Proposal. Cycle 1, ID. #1837

[82] Einstein A., 1915, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, pp 844–847

[83] Einstein A., 1917, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, pp 142–152

[84] Eisenstein D. J., et al., 2023a, arXiv e-prints, p. arXiv:2306.02465

[85] Eisenstein D. J., et al., 2023b, arXiv e-prints, p. arXiv:2306.02465

[86] Eisenstein D. J., et al., 2023c, arXiv e-prints, p. arXiv:2310.12340

[87] Eldridge J. J., Stanway E. R., 2009, Monthly Notices of the Royal Astronomical Society, 400, 1019

[88] Faisst A. L., et al., 2020, The Astrophysical Journals, 247, 61

[89] Fakhouri O., Ma C.-P., 2008, Monthly Notices of the Royal Astronomical Society, 386, 577

[90] Fan X., et al., 2006, The Astronomical Journal, 132, 117

[91] Ferguson H. C., et al., 2004, The Astrophysical Journall, 600, L107

[92] Ferland G. J., et al., 2017, Revista Mexicana de Astronomía y Astrofísica, 53, 385

[93] Ferrara A., Pallottini A., Dayal P., 2023, Monthly Notices of the Royal Astronomical Society, 522, 3986

[94] Ferruit P., et al., 2022, Astronomy and Astrophysics, 661, A81

[95] Finkelstein S. L., et al., 2015, The Astrophysical Journal, 810, 71

[96] Finkelstein S. L., et al., 2022, The Astrophysical Journal Letters, 940, L55

[97] Fontana A., D'Odorico S., Poli F., Giallongo E., Arnouts S., Cristiani S., Moorwood A., Saracco P., 2000, The Astronomical Journal, 120, 2206

[98] Foucaud S., McCracken H. J., Le Fèvre O., Arnouts S., Brodwin M., Lilly S. J., Crampton D., Mellier Y., 2003, Astronomy and Astrophysics, 409, 835

[99] Freedman W. L., 2021, The Astrophysical Journal, 919, 16

[100] Freedman W. L., Madore B. F., 2023, Journal of Cosmology and Astroparticle Physics, 2023, 050

[101] Freedman W. L., et al., 2001, The Astrophysical Journal, 553, 47

[102] Gardner J. P., et al., 2006, Space Science Reviews, 123, 485

[103] Gardner J. P., et al., 2023, Publications of the Astronomical Society of the Pacific, 135, 068001

[104] Garel T., Blaizot J., Guiderdoni B., Michel-Dansac L., Hayes M., Verhamme A., 2015, Monthly Notices of the Royal Astronomical Society, 450, 1279

[105] Gelli V., Mason C., Hayward C. C., 2024, The Astrophysical Journal, 975, 192

[106] Giavalisco M., Dickinson M., 2001, The Astrophysical Journal, 550, 177

[107] Giavalisco M., Steidel C. C., Adelberger K. L., Dickinson M. E., Pettini M., Kellogg M., 1998, The Astrophysical Journal, 503, 543

[108] Giavalisco M., et al., 2004, The Astrophysical Journall, 600, L93

[109] Gnedin N. Y., 2000, Astrophysical Journal, 535, 530

[110] Grogin N. A., et al., 2011, The Astrophysical Journal Supplement Series, 197, 35

[111] Groth E. J., Peebles P. J. E., 1977, The Astrophysical Journal, 217, 385

[112] Gunn J. E., Gott III J. R., 1972, Astrophysical Journal, 176, 1

[113] Gupta R. P., 2023, Monthly Notices of the Royal Astronomical Society, 524, 3385

[114] Hainline K. N., et al., 2023, arXiv e-prints, p. arXiv:2306.02468

[115] Hamana T., Ouchi M., Shimasaku K., Kayo I., Suto Y., 2004, Monthly Notices of the Royal Astronomical Society, 347, 813

[116] Hammer F., Flores H., Elbaz D., Zheng X. Z., Liang Y. C., Cesarsky C., 2005, Astronomy and Astrophysics, 430, 115

[117] Harikane Y., et al., 2016, The Astrophysical Journal, 821, 123

[118] Harikane Y., et al., 2022a, The Astrophysical Journal Supplement Series, 259, 20

[119] Harikane Y., et al., 2022b, The Astrophysical Journal Supplement Series, 259, 20

[120] Hashimoto T., et al., 2018, Nature Astronomy, 557, 392

[121] Hatfield P. W., Bowler R. A. A., Jarvis M. J., Hale C. L., 2018, Monthly Notices of the Royal Astronomical Society, 477, 3760

[122] He X., et al., 2024, The Astrophysical Journal Letters, 960, L13

[123] Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, Astrophysical Journal, 591, 288

[124] Hildebrandt H., Pielorz J., Erben T., van Waerbeke L., Simon P., Capak P., 2009, Astronomy and Astrophysics, 498, 725

[125] Hirschauer A. S., 2024, in , Vol. 17, COS Instrument Handbook v. 17.0. p. 17

[126] Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, The Astrophysical Journal Supplement Series, 163, 1

[127] Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, The Astrophysical Journal Supplement Series, 175, 356

[128] Hopkins P. F., et al., 2010, The Astrophysical Journal, 715, 202

[129] Horne K., 1986, Publications of the Astronomical Society of the Pacific, 98, 609

[130] Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

[131] Ishikawa S., et al., 2019, The Astrophysical Journal, 904

[132] Iyer K. G., Gawiser E., Faber S. M., Ferguson H. C., Kartaltepe J., Koekemoer A. M., Pacifici C., Somerville R. S., 2019, The Astrophysical Journal, 879, 116

[133] Jakobsen P., et al., 2022, Astronomy and Astrophysics, 661, A80

[134] Jenkins A., Frenk C., White S. D., Colberg J. M., Cole S., Evrard A. E., Couchman H., Yoshida N., 2001, Monthly Notices of the Royal Astronomical Society, 321, 372

[135] Jiang H., et al., 2024, The Astrophysical Journal, 972, 121

[136] Jogee S., et al., 2009, The Astrophysical Journal, 697, 1971

[137] Jose C., Lacey C. G., Baugh C. M., 2016, Monthly Notices of the Royal Astronomical Society, 463, 270

[138] Jose C., Baugh C. M., Lacey C. G., Subramanian K., 2017, Monthly Notices of the Royal Astronomical Society, 469, 4428

[139] Kaiser N., 1984, Astrophysical Journal, 284, L9

[140] Kapteyn J. C., 1922, The Astrophysical Journal, 55, 302

[141] Kauffmann G., White S. D. M., Guiderdoni B., 1993, Monthly Notices of the Royal Astronomical Society, 264, 201

[142] Kennicutt Jr. R. C., 1998, Annual Review of Astronomy and Astrophysics, 36, 189

[143] Khoram A. H., Belfiore F., 2025, Astronomy and Astrophysics, 693, A150

[144] Kim E., et al., 2021, Monthly Notices of the Royal Astronomical Society, 507, 3113

[145] Klypin A., Primack J. R., Holtzman J. A., 1996, The Astrophysical Journal, 466, 13

[146] Kravtsov A. V., Borgani S., 2012, Annual Review of Astronomy and Astrophysics, 50, 353

[147] Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, The Astrophysical Journal, 609, 35

[148] Kroupa P., et al., 2022, Monthly Notices of the Royal Astronomical Society, 517, 3613

[149] Labbé I., et al., 2023, Nature Astronomy, 616, 266

[150] Landy S. D., Szalay A. S., 1993a, The Astrophysical Journal, 412, 64

[151] Landy S. D., Szalay A. S., 1993b, The Astrophysical Journal, 412, 64

[152] Le Fèvre O., et al., 2000a, Monthly Notices of the Royal Astronomical Society, 311, 565

[153] Le Fèvre O., et al., 2000b, Monthly Notices of the Royal Astronomical Society, 311, 565

[154] Le Fèvre O., et al., 2020, Astronomy and Astrophysics, 643, A1

[155] Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1

[156] Lee K., Giavalisco M., Gnedin O. Y., Somerville R. S., Ferguson H. C., Dickinson M., Ouchi M., 2006, The Astrophysical Journal, 642, 63

[157] Leethochawalit N., Trenti M., Morishita T., Roberts-Borsani G., Treu T., 2022a, Monthly Notices of the Royal Astronomical Society, 509, 5836

[158] Leethochawalit N., Trenti M., Morishita T., Roberts-Borsani G., Treu T., 2022b, Monthly Notices of the Royal Astronomical Society, 509, 5836

[159] Limber D. N., 1953, Astrophysical Journal, 117, 134

[160] Lin L., et al., 2008, The Astrophysical Journal, 681, 232

[161] Ling E. N., Frenk C. S., Barrow J. D., 1986, Monthly Notices of the Royal Astronomical Society, 223, 21P

[162] Loeb A., Barkana R., 2001, Annual Review of Astronomy and Astrophysics, 39, 19

[163] López-Sanjuan C., et al., 2009, The Astrophysical Journal, 694, 643

[164] Lotz J. M., Primack J., Madau P., 2004, The Astronomical Journal, 128, 163

[165] Lotz J. M., Jonsson P., Cox T. J., Primack J. R., 2008a, Monthly Notices of the Royal Astronomical Society, 391, 1137

[166] Lotz J. M., et al., 2008b, The Astrophysical Journal, 672, 177

[167] Lynden-Bell D., 1967, Monthly Notices of the Royal Astronomical Society, 136, 101

[168] Ma C.-P., Fry J. N., 2000, The Astrophysical Journal, 543, 503

[169] Machacek M. E., Bryan G. L., Abel T., 2001, Astrophysical Journal, 548, 509

[170] Madau P., Dickinson M., 2014, Annual Review of Astronomy and Astrophysics, 52, 415

[171] Madau P., Rees M. J., Volonteri M., Haardt F., Oh S. P., 2004, Astrophysical Journal, 604, 484

[172] Man A. W. S., Zirm A. W., Toft S., 2016, The Astrophysical Journal, 830, 89

[173] Mantha K. B., et al., 2018, Monthly Notices of the Royal Astronomical Society, 475, 1549

[174] Martínez H. J., Zandivarez A., Merchán M. E., Domínguez M. J. L., 2002, Monthly Notices of the Royal Astronomical Society, 337, 1441

[175] Mascia S., et al., 2024, Astronomy and Astrophysics, 690, A2

[176] Mason C. A., Trenti M., Treu T., 2015, The Astrophysical Journal, 813, 21

[177] Mason C. A., Trenti M., Treu T., 2023, Monthly Notices of the Royal Astronomical Society, 521, 497

[178] Massey R., et al., 2007, Nature Astronomy, 445, 286

[179] Merlin E., et al., 2022, The Astrophysical Journal Letters, 938, L14

[180] Mihos J. C., Hernquist L., 1996, Astrophysical Journal, 464, 641

[181] Mo H. J., White S. D. M., 1996, Monthly Notices of the Royal Astronomical Society, 282, 347

[182] Nagao T., Maiolino R., Marconi A., 2006, Astronomy and Astrophysics, 459, 85

[183] Naidu R. P., et al., 2022, The Astrophysical Journall, 940, L14

[184] Naidu R. P., et al., 2024, arXiv e-prints, p. arXiv:2410.01874

[185] Naoz S., Noter S., Barkana R., 2006, Monthly Notices of the Royal Astronomical Society, 373, L98

[186] Napolitano L., et al., 2025, Astronomy and Astrophysics, 693, A50

[187] Neyrinck M. C., Hamilton A. J. S., Gnedin N. Y., 2004, in American Astronomical Society Meeting Abstracts. p. 148.02

[188] O'Leary J. A., Moster B. P., Naab T., Somerville R. S., 2021, Monthly Notices of the Royal Astronomical Society, 501, 3215

[189] Oesch P. A., et al., 2016, Astrophysical Journal, 819, 129

[190] Oesch P. A., Bouwens R. J., Illingworth G. D., Labbé I., Stefanon M., 2018, The Astrophysical Journal, 855, 105

[191] Oke J. B., Gunn J. E., 1983, The Astrophysical Journal, 266, 713

[192] Okuta R., Unno Y., Nishino D., Hido S., Loomis C., 2017, in Proceedings of Workshop on Machine Learning Systems (LearningSys) in The Thirty-first Annual Conference on Neural Information Processing Systems (NIPS). http://learningsys.org/nips17/assets/papers/paper_16.pdf

[193] Olive K. A., Steigman G., Walker T. P., 2000, Physics Reports, 333, 389

[194] Ostriker J. P., Hausman M. A., 1977, Astrophysical Journal, 217, L125

[195] Ouchi M., et al., 2004, The Astrophysical Journal, 611, 685

[196] Ouchi M., et al., 2005, The Astrophysical Journal, 635, L117

[197] Overzier R. A., Bouwens R. J., Illingworth G. D., Franx M., 2006, The Astrophysical Journal, 648, l5

[198] Padmanabhan H., Loeb A., 2023, The Astrophysical Journall, 953, L4

[199] Paquereau L., et al., 2025, arXiv e-prints, p. arXiv:2501.11674

[200] Paris D., et al., 2023a, arXiv preprint arXiv:2301.02179

[201] Paris D., et al., 2023b, The Astrophysical Journal, 952, 20

[202] Park J., et al., 2017, Monthly Notices of the Royal Astronomical Society, 472, 1995

[203] Park J., Mesinger A., Greig B., Gillet N., 2019, Monthly Notices of the Royal Astronomical Society, 484, 933

[204] Patton D. R., et al., 2002, The Astrophysical Journal, 565, 208

[205] Peacock J. A., Nicholson D., 1991a, Monthly Notices of the Royal Astronomical Society, 253, 307

[206] Peacock J. A., Nicholson D., 1991b, Monthly Notices of the Royal Astronomical Society, 253, 307

[207] Peacock J. A., Smith R. E., 2000, Monthly Notices of the Royal Astronomical Society, 318, 1144

[208] Peebles P. J. E., 1968, Astrophysical Journal, 153, 1

[209] Peebles P. J. E., 1980, The large-scale structure of the universe. Princeton University Press

[210] Peebles P. J. E., 1982, Astrophysical Journal Letters, 263, L1

[211] Penzias A. A., Wilson R. W., 1965, The Astrophysical Journal, 142, 419

[212] Percival W. J., Cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C., Szalay A. S., 2007, Monthly Notices of the Royal Astronomical Society, 381, 1053

[213] Persic M., Salucci P., 1992, Monthly Notices of the Royal Astronomical Society, 258, 14P

[214] Planck Collaboration et al., 2016, Astronomy and Astrophysics, 594, A13

[215] Planck Collaboration et al., 2020, Astronomy and Astrophysics, 641, A1

[216] Porciani C., Giavalisco M., 2002, The Astrophysical Journal, 565, 24

[217] Press W. H., Schechter P., 1974, The Astrophysical Journal, 187, 425

[218] Price S. H., et al., 2024, arXiv e-prints, p. arXiv:2408.03920

[219] Qiu Y., et al., 2018, Monthly Notices of the Royal Astronomical Society, 481, 4885

[220] Ren K., Trenti M., Mason C. A., 2019, The Astrophysical Journal, 878, 114

[221] Rickman E., Brown J., 2024, in , Vol. 24, STIS Instrument Handbook for Cycle 33 v. 24. p. 24

[222] Rieke M. J., Kelly D., Horner S., 2005, in Heaney J. B., Burriesci L. G., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 5904, Cryogenic Optical Systems and Instruments XI. pp 1–8, doi:10.1117/12.615554

[223] Rieke G. H., et al., 2015, Publications of the Astronomical Society of the Pacific, 127, 584

[224] Rieke M. J., et al., 2023, arXiv e-prints, p. arXiv:2306.02466

[225] Riess A. G., et al., 1998, The Astronomical Journal, 116, 1009

[226] Riess A. G., et al., 2016, The Astrophysical Journal, 826, 56

[227] Riess A. G., et al., 2023, The Astrophysical Journall, 956, L18

[228] Rigby J., et al., 2023, Publications of the Astronomical Society of the Pacific, 135, 048001

[229] Robertson B. E., 2022, Annual Review of Astronomy and Astrophysics, 60, 121

[230] Robertson B. E., Ellis R. S., Furlanetto S. R., Dunlop J. S., 2015, The Astrophysical Journal Letters, 802, L19

[231] Rodriguez-Gomez V., et al., 2015, Monthly Notices of the Royal Astronomical Society, 449, 49

[232] Rodriguez-Gomez V., et al., 2016, Monthly Notices of the Royal Astronomical Society, 458, 2371

[233] Rodriguez-Gomez V., et al., 2019, Monthly Notices of the Royal Astronomical Society, 483, 4140

[234] Romano M., et al., 2021, Astronomy and Astrophysics, 653, A111

[235] Rose C., et al., 2023, The Astrophysical Journal, 942, 54

[236] Santini P., et al., 2023, The Astrophysical Journall, 942, L27

[237] Savoy J., Sawicki M., Thompson D., Sato T., 2011, The Astrophysical Journal, 737, 92

[238] Schaerer D., 2002, Astronomy and Astrophysics, 382, 28

[239] Schmidt B. P., et al., 1998, The Astrophysical Journal, 507, 46

[240] Schmidt K. B., et al., 2013, Monthly Notices of the Royal Astronomical Society, 432, 285

[241] Schramm D. N., Turner M. S., 1998, Reviews of Modern Physics, 70, 303

[242] Scoville N., et al., 2007, The Astrophysical Journal Supplement Series, 172, 38

[243] Seljak U., 2000, Monthly Notices of the Royal Astronomical Society, 318, 203

[244] Shapley H., Curtis H. D., 1921, Bulletin of the National Research Council, 2, 171

[245] Sheth R. K., Tormen G., 1999a, Monthly Notices of the Royal Astronomical Society, 308, 119

[246] Sheth R. K., Tormen G., 1999b, Monthly Notices of the Royal Astronomical Society, 308, 119

[247] Sheth R. K., Mo H. J., Tormen G., 2001, Monthly Notices of the Royal Astronomical Society, 323, 1

[248] Shuntov M., et al., 2022, Astronomy and Astrophysics, 664, A61

[249] Shuntov M., et al., 2025, arXiv e-prints, p. arXiv:2503.14280

[250] Silk J., 1968, Astrophysical Journal, 151, 459

[251] Smith R. W., Hanle P. A., Kargon R. H., Tatarewicz J. N., 1993, The Space Telescope. A study of NASA, science, technology, and politics.

[252] Smoot G. F., et al., 1992, The Astrophysical Journall, 396, L1

[253] Snyder G. F., Lotz J. M., Rodriguez-Gomez V., Guimarães R. d. S., Torrey P., Hernquist L., 2017, Monthly Notices of the Royal Astronomical Society, 468, 207

[254] Snyder G. F., Rodriguez-Gomez V., Lotz J. M., Torrey P., Quirk A. C. N., Hernquist L., Vogelsberger M., Freeman P. E., 2019, Monthly Notices of the Royal Astronomical Society, 486, 3702

[255] Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, The Astrophysical Journals, 214, 15

[256] Springel V., et al., 2005, Nature Astronomy, 435, 629

[257] Stark D. V., 2024, in , Vol. 24, ACS Instrument Handbook for Cycle 33 v. 24.0. p. 24

[258] Steinhardt C. L., Capak P., Masters D., Speagle J. S., 2016, The Astrophysical Journal, 824, 21

[259] Tacchella S., et al., 2023, The Astrophysical Journal, 952, 74

[260] Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, Astrophysical Journal, 474, 1

[261] Tegmark M., et al., 2004, Physical Review D, 69, 103501

[262] Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, The Astrophysical Journal, 688, 709

[263] Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, The Astrophysical Journal, 724, 878

[264] Tohill C., Bamford S. P., Conselice C. J., Ferreira L., Harvey T., Adams N., Austin D., 2024, The Astrophysical Journal, 962, 164

[265] Toomre A., Toomre J., 1972, The Astrophysical Journal, 178, 623

[266] Trenti M., Stiavelli M., 2008, The Astrophysical Journal, 676, 767

[267] Trenti M., Stiavelli M., Bouwens R. J., Oesch P., Shull J. M., Illingworth G. D., Bradley L. D., Carollo C. M., 2010, The Astrophysical Journal Letters, 714, l202

[268] Trenti M., Perna R., Jimenez R., 2015, The Astrophysical Journal, 802, 103

[269] Treu T., et al., 2022, The Astrophysical Journal, 935, 110

[270] Treu T., et al., 2023, The Astrophysical Journall, 942, L28

[271] Vale A., Ostriker J. P., 2004, Monthly Notices of the Royal Astronomical Society, 353, 189

[272] Vale A., Ostriker J. P., 2008, Monthly Notices of the Royal Astronomical Society, 383, 355

[273] Ventou E., et al., 2017, A&A, 608, A9

[274] Ventou E., et al., 2019, Astronomy & Astrophysics, 631, A87

[275] Viana A., et al. 2009, Near Infrared Camera and Multi-Object Spectrometer Instrument Handbook for Cycle 17 v. 11.0

[276] Villa-Vélez J. A., Buat V., Theulé P., Boquien M., Burgarella D., 2021, Astronomy and Astrophysics, 654, A153

[277] Vulcani B., et al., 2023, The Astrophysical Journall, 948, L15

[278] Watson P. J., et al., 2025, arXiv e-prints, p. arXiv:2504.00823

[279] Wechsler R. H., Tinker J. L., 2018, Annual Review of Astronomy and Astrophysics, 56, 435

[280] Whitaker K. E., et al., 2019, The Astrophysical Journals, 244, 16

[281] White S. D. M., Rees M. J., 1978, Monthly Notices of the Royal Astronomical Society, 183, 341

[282] Williams R. E., et al., 1996, The Astronomical Journal, 112, 1335

[283] Williams C. C., et al., 2018, The Astrophysical Journals, 236, 33

[284] Willott C. J., et al., 2022, Publications of the Astronomical Society of the Pacific, 134, 025002

[285] Wise J. H., Abel T., Turk M. J., Norman M. L., Smith B. D., 2012, Monthly Notices of the Royal Astronomical Society, 427, 311

[286] Wright G. S., et al., 2015, Publications of the Astronomical Society of the Pacific, 127, 595

[287] Wright G. S., et al., 2023, Publications of the Astronomical Society of the Pacific, 135, 048003

[288] Wyithe J. S. B., Loeb A., Oesch P. A., 2014, Monthly Notices of the Royal Astronomical Society, 439, 1326

[289] Yoshida N., Omukai K., Hernquist L., Abel T., 2006, Astrophysical Journal, 652, 6

[290] Yung L. Y. A., et al., 2022, Monthly Notices of the Royal Astronomical Society, 515, 5416

[291] Zehavi I., et al., 2011, The Astrophysical Journal, 736, 59

[292] Zheng Z., et al., 2005, The Astrophysical Journal, 633, 791

[293] van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, Monthly Notices of the Royal Astronomical Society, 430, 725

[294] van der Walt S., et al., 2014, PeerJ, 2, e453