The Astrophysical Journal Supplement Series, 280:62 (34pp), 2025 October https://doi.org/10.3847/1538-4365/adfa70

A 13 Billion Year View of Galaxy Growth: Metallicity Gradient Evolution from the Local Universe to $z = 9$ with JWST and Archival Surveys

Zihao Li[1,2,3], Zheng Cai[1], Xin Wang[4,5,6], Zhaozhou Li[7], Avishai Dekel[7,8], Kartick C. Sarkar[7,9], Eduardo Bañados[10], Fuyan Bian[11], Aklant K. Bhowmick[12], Laura Blecha[13], Sarah E. I. Bosman[10,14], Jaclyn B. Champagne[15], Xiaohui Fan[15], Emmet Golden-Marx[16], Hyunsung D. Jun[17], Mingyu Li[1], Xiaojing Lin[1,15], Weizhe Liu[15], Fengwu Sun[18], Maxime Trebitsch[19], Fabian Walter[10], Feige Wang[20], Yunjing Wu[1], Jinyi Yang[20], Huanian Zhang[21], Shiwu Zhang[22], Mingyang Zhuang[23], and Siwei Zou[24,25]

[1] Department of Astronomy, Tsinghua University, Beijing 100084, People's Republic of China; zcai@tsinghua.edu.cn
[2] Cosmic Dawn Center (DAWN), Denmark
[3] Niels Bohr Institute, University of Copenhagen, Jagtvej 128, DK-2200, Copenhagen N, Denmark
[4] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, People's Republic of China; xwang@ucas.ac.cn
[5] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People's Republic of China
[6] Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, People's Republic of China
[7] Center for Astrophysics and Planetary Science, Racah Institute of Physics, The Hebrew University, Jerusalem, 91904, Israel
[8] Santa Cruz Institute for Particle Physics, University of California, Santa Cruz, CA 95064, USA
[9] Raman Research Institute, Sadashivanagar, C. V. Raman Ave., 560 080, Bangalore, India
[10] Max-Plank-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
[11] European Southern Observatory, Alonso de Cordova 3107, Casilla 19001, Vitacura, Santiago 19, Chile
[12] Department of Astronomy, University of Virginia, Charlottesville, VA 22904, USA
[13] Department of Physics, University of Florida, Gainesville, FL 32601, USA
[14] Institute for Theoretical Physics, Heidelberg University, Philosophenweg 12, D-69120, Heidelberg, Germany
[15] Steward Observatory, University of Arizona, 933 N Cherry Ave., Tucson, AZ 85721, USA
[16] INAF-Astronomical Observatory of Padova, vicolo dell'Osservatorio 5, 35122 Padova, Italy
[17] Department of Physics, Northwestern College, 101 7th St. SW, Orange City, IA 51041, USA
[18] Center for Astrophysics – Harvard & Smithsonian, 60 Garden St., Cambridge, MA 02138, USA
[19] LUX, Observatoire de Paris, Université PSL, Sorbonne Université, CNRS, 75014 Paris, France
[20] Department of Astronomy, University of Michigan, 500 S State St., Ann Arbor, MI 48109, USA
[21] Department of Astronomy, Huazhong University of Science and Technology, Wuhan 430074, People's Republic of China
[22] Zhejiang Lab, Hangzhou, Zhejiang 311121, People's Republic of China
[23] Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[24] Chinese Academy of Sciences South America Center for Astronomy, National Astronomical Observatories, CAS, Beijing 100101, People's Republic of China
[25] Departamento de Astronomía, Universidad de Chile, Casilla 36-D, Santiago, Chile
Received 2025 June 30; revised 2025 August 1; accepted 2025 August 5; published 2025 September 26

Abstract

Galaxy gas-phase metallicity gradients have been extensively studied over the past four decades, both in the local and in the high-redshift Universe, as they trace the baryon cycle and growth of galaxies. With the unprecedented spatial resolution and sensitivity of JWST, it is now possible to measure metallicity and its radial gradients out to redshifts as high as $z = 9$. Here, we present a sample of 455 spectroscopically confirmed galaxies from redshifts $1.7 \lesssim z \lesssim 9$ that are spatially resolved on subkiloparsec scales by deep JWST NIRCam or NIRISS Wide Field Slitless Spectroscopy. Synthesizing these new JWST observations with legacy observations from the literature, we observe that at redshift $z > 5$, galaxy centers are more metal rich, exhibiting negative metallicity gradients of ~ -0.4 dex kpc^{-1}. These gradients flatten over time, reaching near zero around $z \approx 2$, coinciding with the peak of the cosmic star formation rate. Beyond this point, the gradients become negative again at lower redshifts approaching $z = 0$. This evolution likely reflects transitions in galaxy formation modes: an inside-out growth phase dominated by intense central star formation with inefficient feedback and limited gas mixing during "cosmic dawn," enhanced gas mixing due to feedback-driven wind and gas accretion at "cosmic noon," and a later phase of slow evolution and reduced feedback toward the present day. These physical processes, including gas accretion and feedback, not only regulate star and galaxy formation on a cosmic scale but also shape the evolutionary pathways of individual galaxies over cosmic time.

Unified Astronomy Thesaurus concepts: Galaxy chemical evolution (580); Galaxy formation (595); High-redshift galaxies (734); Chemical enrichment (225); Metallicity (1031)

Materials only available in the online version of record: figure set, machine-readable table

1. Introduction

The process of star formation and the rate at which stars form across the Universe (referred to as star formation rate density (SFRD); P. Madau & M. Dickinson 2014) depend on a complex interplay of gas inflows, gas outflows, and other mechanisms operating at different epochs in cosmic history. As such, the dominant modes of galaxy formation are also expected to vary with time (C. Schreiber et al. 2015). These processes further influence the metal content in galaxies' interstellar mediums (ISM), referred to as metallicity, as well as the spatial distribution of metals within galaxies over time.

They provide insight into the overall growth and evolution of galaxies throughout cosmic history.

However, key physical processes that dominate galaxy formation and evolution at each epoch are still subject to debate. Traditionally, galaxy evolution has been understood as a gradual process, shaped by both internal dynamics and external interactions over long timescales. These include processes such as the gradual accretion of gas, the formation of stellar structures like bars, gravitational interactions with neighboring galaxies (galaxy harassment), and galaxy mergers (J. Kormendy & R. C. J. Kennicutt 2004; D. Kereš et al. 2005; P. F. Hopkins et al. 2009; M. Lang et al. 2014). Additionally, more dynamic and rapid events with shorter timescales, such as feedback from stellar winds and explosive supernovae (SNe) that last only a few million years, play a critical role in regulating star formation and the growth of galaxies (D. K. Erb 2015; K. El-Badry et al. 2016). Over the decades, various theoretical models and simulations have been developed to explain galaxy evolution (R. B. Larson 1974; S. J. Lilly et al. 2013; L. Magrini et al. 2016; P. F. Hopkins et al. 2023). Comprehensive models are required to account for the intricate processes driving the formation and distribution of chemical elements in galaxies, which are key to understanding their overall evolution (B. M. Tinsley 1980; R. L. Graf et al. 2024; C. Lyu et al. 2025; B. Tapia-Contreras et al. 2025).

Observing the cosmological evolution of metallicity gradients offers insights into the dominant processes at different epochs, as metal distribution is sensitive to gas transport and turbulence driven by these kinds of activity (P. Sharda et al. 2024). In the Milky Way, metallicity gradients have been extensively studied as a benchmark for understanding the chemical evolution of disk galaxies, providing a local reference point for interpreting extragalactic observations (M. Mollá et al. 2019; J. Lian et al. 2023; J. W. Johnson et al. 2024; R. L. Graf et al. 2025; B. Ratcliffe et al. 2025). Over the past four decades, significant advancements have been made in observing galaxies' metal content and metal distribution via multiple techniques such as the use of long-slit or multiobject fiber spectrographs (B. E. J. Pagel & M. G. Edmunds 1981; M. G. Edmunds & B. E. J. Pagel 1984; I. N. Evans 1986; J. M. Vilchez et al. 1988; G. A. Shields 1990; M. B. Vila-Costas & M. G. Edmunds 1992; L. van Zee et al. 1998; S. Considère et al. 2000; L. Magrini et al. 2007; D. S. N. Rupke et al. 2010a; K. Grasha et al. 2022), integral field unit observations (G. Cresci et al. 2010; J. Queyrel et al. 2012; A. M. Swinbank et al. 2012; T. Jones et al. 2013; P. Troncoso et al. 2014; D. Carton et al. 2015, 2018; I. T. Ho et al. 2015; N. Leethochawalit et al. 2016; E. Wuyts et al. 2016; F. Belfiore et al. 2017; N. M. Förster Schreiber et al. 2018; J. Lian et al. 2018; H. Poetrodjojo et al. 2018; L. Sánchez-Menguiano et al. 2018; F. Bresolin 2019; M. D. Thorp et al. 2019; Y. Zhuang et al. 2019; M. Curti et al. 2020a; A. Franchetto et al. 2021; S. Gillman et al. 2021; P. Lagos et al. 2022; G. Venturi et al. 2024; M. Ju et al. 2025; A. H. Khoram & F. Belfiore 2025; T. Li et al. 2025), slitless spectroscopy (T. Jones et al. 2015; X. Wang et al. 2017, 2020, 2022a; R. C. Simons et al. 2021; Z. Li et al. 2022; Y. Cheng et al. 2024), and the use of radio telescopes (L. Vallini et al. 2024). Those works have found a diversity of metallicity gradients across stellar masses and redshifts. They have reported negative metallicity gradients in local galaxies (e.g., I. T. Ho et al. 2015) as well as at higher redshifts (e.g., N. M. Förster Schreiber et al. 2018). Positive (inverted) gradients have also been observed, both in the local Universe (e.g., T. Li et al. 2025) and at high redshifts up to $z \sim 3$ (e.g., P. Troncoso et al. 2014; X. Wang et al. 2022a).

Various theoretical models have been proposed to explain certain behaviors of metallicity gradients. Negative gradients can arise in scenarios such as inside-out growth (S. Boissier & N. Prantzos 2000) and accretion disks with coplanar gas inflow (E. Wang & S. J. Lilly 2022; C. Lyu et al. 2025). P. Sharda et al. (2021b, 2021c) suggest that negative gradients originate under equilibrium conditions, modeled through a more complex framework that accounts for the competition of radial advection, metal production, and gas accretion. The inside-out scenario is supported by analyzing the radial distribution of stellar populations in low-redshift star-forming galaxies (E. Pérez et al. 2013; D. Goddard et al. 2017; N. Frankel et al. 2019), as well as by simulations (P. B. Tissera et al. 2016; F. Renaud et al. 2025).

On the other hand, flat or positive gradients can arise in the case of pristine gas accretion (G. Cresci et al. 2010; J. Molina et al. 2017; Z. Li et al. 2022), metal loss due to feedback-driven outflows (X. Wang et al. 2019; L. E. Porter et al. 2022; X. Sun et al. 2025), and mergers (D. S. N. Rupke et al. 2010a, 2010b; P. Torrey et al. 2012; P. B. Tissera et al. 2022). The analytical model by P. Sharda et al. (2021b) suggests galaxies with positive gradients are out of equilibrium with extreme outflow metal loss, pristine gas accretion, or the reaccretion of metal-enriched gas (J. Fu et al. 2013). Observations with a much-enriched circumgalactic medium (CGM) on large scales at cosmic noon (Z. Cai et al. 2017; S. Zhang et al. 2023) also reflect that metal enrichment could happen on a larger scale than previously expected.

However, a comprehensive understanding of how metallicity evolves over cosmic time remains elusive. This is largely because earlier observations often lacked the spatial resolution and sensitivity needed to map metal distributions within galaxies at scales smaller than a few hundred parsecs (subkiloparsec scales), which is crucial to resolve internal structures and small-scale physical processes, and at greater distances corresponding to redshifts $z \gtrsim 3$. With several years of JWST observations, a large spectroscopic sample of high-redshift galaxies has now been assembled, enabling a comprehensive investigation of metallicity gradients across cosmic time (X. Wang et al. 2022a; J. E. Birkin et al. 2023; S. Arribas et al. 2024; G. Venturi et al. 2024; M. Ju et al. 2025).

In this paper, we present a sample of 455 galaxies spanning redshifts $z = 1.7$ to 9 from JWST Wide Field Slitless Spectroscopy (WFSS) surveys, and measure the spatially resolved metallicities. Leveraging literature observations from archival surveys, we analyze the redshift evolution of metallicity gradients from $z = 0$ to $z = 9$.

This paper is organized as follows. In Section 2, we describe the observations and data reduction. The methods are described in Section 3. We present the results in Section 4 and discuss the physical implications of the observations. We summarize the main findings in Section 5. Throughout this article, we assume a standard flat ΛCDM cosmology with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Emission lines are indicated as follows: [O III] λ5008 ≔ [O III], [O II] $\lambda\lambda$3727, 3730 ≔ [O II], and [N II] λ6585 ≔ [N II], if presented without wavelength values.

2. Description of Data

2.1. Observations and Data Reduction

In this work, we analyze JWST observations from "A Spectroscopic Survey of Biased Halos in the Reionization Era" (ASPIRE; F. Wang et al. 2023; J. Yang et al. 2023), "First Reionization Epoch Spectroscopically Complete Observations" (FRESCO; P. A. Oesch et al. 2023), and "The Next Generation Deep Extragalactic Exploratory Public Survey" (NGDEEP; M. B. Bagley et al. 2024). By combining these surveys, we built a large statistical sample of galaxies that have spatially resolved grism spectroscopy at $z \sim 2$–8 that can be used for metallicity gradient studies with subkiloparsec resolution. The angular resolution reaches $0\farcs04 - 0\farcs14$ depending on the wavelength, corresponding to the physical size of $\sim$0.5 kpc in the redshift range $1.7 < z < 3.5$ and $\sim$0.7 kpc in the redshift range $5 < z < 9$.

The ASPIRE survey contains NIRCam F356W R-grism observations of 25 quasar fields, with direct imaging in the F356W, F200W, and F115W filters. FRESCO includes NIRCam F444W R-grism WFSS observations, with direct imaging in the F444W, F210M, and F182M filters. The NGDEEP observations include NIRISS WFSS R-grism and C-grism observations, with direct imaging in the F200W, F150W, and F115W NIRISS filters. The combined data products cover a spectral wavelength range $\lambda_{\rm obs} \in$ [1.0–2.2] $\cup$ [3.1, 5.0] μm, enabling redshift coverage $z \in$ [1.3, 3.5] $\cup$ [5.3, 9] for [O III] emitters.

The ASPIRE imaging is processed in the same way as described in F. Wang et al. (2023) and J. Yang et al. (2023). Briefly, the reduction is performed using version 1.8.3 of the JWST Calibration Pipeline with the reference files from version 11.16.15 of the standard Calibration Reference Data System. After Stages 2 and 3, the images are aligned to Gaia DR3 and drizzled to a common pixel scale of $0\farcs031$/pixel. For the FRESCO and NGDEEP samples, we use the public image products, including NIRCam and NIRISS, retrieved from the Dawn JWST Archive (DJA).[26] The DJA images are processed with the Grism Redshift & Line Analysis tool (GRIZLI; G. Brammer 2023).

We use JWST Calibration Pipeline `CALWEBB` Stage 1 to calibrate individual NIRCam WFSS exposures, with reference file `jwst_1090.pmap`. The $1/f$ noise is then subtracted using the routine described in F. Wang et al. (2023). The world coordinate system (WCS) information is assigned to each exposure with the `assign_wcs` step. The flat-field is done with `CALWEBB` Stage 2. We build the median backgrounds for the F356W and F444W filters based on all ASPIRE and FRESCO WFSS exposures, which are then scaled and subtracted from each exposure. We then measure the astrometric offsets between each of the short-wavelength images and the fully calibrated direct mosaic to align each grism exposure with the direct image.

The preprocessed NIRCam WFSS exposures are then processed with GRIZLI. We use the spectral tracing, grism dispersion, and sensitivity models described in F. Sun et al. (2023). The detection catalog for spectral extraction is built from the F356W (F444W) direct image for ASPIRE (FRESCO), and the continuum cross-contamination is subtracted by GRIZLI forward modeling using the reference image as the reference image for each grism exposure. The emission-line maps are drizzled from the 2D grism to the image plane, matching the WCS of the direct imaging, with `pixfrac = 1` and pixel size $= 0\farcs06$, to sample the NIRCam point-spread functions (PSFs; FWHM $= 0\farcs115$–$0\farcs145$) properly.

The NIRISS WFSS reduction is done similarly, with reference file `jwst_1090.pmap` and GRIZLI full end-to-end processing, including `CALWEBB` Stage 1 calibration from raw exposures, flat-fielding, $1/f$ noise subtraction, sky subtraction, and astrometry alignment. The contamination modeling and extraction of the spectra are performed with the latest GRIZLI NIRISS configuration file (J. Matharu & G. Brammer 2022). The emission-line maps are drizzled from the 2D grism to the image plane with `pixfrac = 1` and a pixel size of $0\farcs03$, for properly sampling the NIRISS PSFs (FWHM $= 0\farcs040$–$0\farcs066$).

2.2. Photometric Catalog

The sources in ASPIRE are identified using F356W as the detection image. We use `SourceXtractor++` (E. Bertin et al. 2020), a successor version of the original `SExtractor` (E. Bertin & S. Arnouts 1996), to construct the photometric catalog. The catalogs are extracted in the F115W, F200W, and F356W bands after matching the PSFs to those of F356W ($0\farcs1$). Fluxes measured in Kron apertures with ($k = 1.2$, $R = 1.7$) are corrected to total magnitudes using the ratio for Kron fluxes measured in larger apertures ($k = 1.5$, $R = 2.5$). The uncertainties for each source's photometry are measured by placing 1000 random apertures the size of the Kron aperture across the image and measuring the rms.

For the FRESCO and NGDEEP samples, we use the public catalogs retrieved from the DJA. The DJA catalogs are extracted with `SExtractor` (E. Bertin & S. Arnouts 1996). The photometry extracted in circular apertures with a diameter of $0\farcs5$ is corrected to the total fluxes within an elliptical Kron aperture. The DJA image reduction and photometry are detailed in F. Valentino et al. (2023).

2.3. Sample Selection

We apply the emitter finder algorithm detailed in F. Wang et al. (2023) to select galaxies with possible [O III] emission lines. Briefly, the algorithm first searches for peaks in 1D spectra and, assuming every peak as a candidate [O III] λ5008, measures the signal-to-noise ratio (SNR) at the expected position of [O III] λ4960. A good candidate has $\mathrm{SNR}_{[\mathrm{O\,III}]\lambda4960} > 2$ and a color excess in the corresponding image band. We note that there are $z > 5$ [O III] emitters in ASPIRE and FRESCO that only have [O III] λ5008 detection because [O III] λ4960 is too faint to be detected. We have ignored sources where [O III] λ4960 is not detected because they will decrease the SNR in our stack. We remove candidates with cross-contamination in the 2D spectra by visual inspection.

Since we only have [O III] and Hβ detection for our sample at $z > 5$, we use empirical metallicity calibrations (R3; R. L. Sanders et al. 2024) to convert the [O III]/Hβ ratio to metallicity (see more details in Section 3.3). The R3 calibration is double-valued at given line ratios, with a lower branch at $12 + \log(\mathrm{O/H}) \lesssim 7.9$ and a higher branch at $12 + \log(\mathrm{O/H}) \gtrsim 7.9$. The metallicity estimations are unreliable if high-metallicity galaxies are mixed with the low-metallicity sample. One solution is to use the mass–metallicity relation (MZR) to estimate the metallicity of galaxies and

[26] https://dawn-cph.github.io/dja/

ensure the applicability of choosing the lower-branch solution. A. Sarkar et al. (2025) provided an empirical mass–metallicity–redshift relation formulated as $12 + \log(\mathrm{O/H}) = 6.29 + 0.237 \times \log(M_*/M_\odot) - 0.06 \times (1 + z)$. From this relation, galaxies at $5 < z < 9$ with stellar masses $\log(M_*/M_\odot) \lesssim 9$ are expected to have metallicities $12 + \log(\mathrm{O/H}) \lesssim 7.9$.

From spectral energy distribution (SED) modeling (see details in Section 3.2), we find that most of our sample galaxies are below $\log(M_*/M_\odot) = 9$, with only a small fraction above this threshold—12.6% and 4.1% for ASPIRE and FRESCO samples, respectively. To eliminate the bias of metallicity estimation by massive galaxies, we have removed galaxies with $\log(M_*/M_\odot) > 9$ in our $z > 5$ sample, and the low-mass sample includes 284 and 47 galaxies in ASPIRE and FRESCO, respectively. Since there is still a considerable number ($N = 41$) of massive $\log(M_*/M_\odot) > 9$ galaxies in ASPIRE, we construct an additional massive sample with those galaxies, and we can measure their metallicity assuming the R3 upper-branch solution (see Appendix A). Since we only have two galaxies with $\log(M_*/M_\odot) > 9$ in FRESCO, which is insufficient to do a statistical study, we do not keep a $\log(M_*/M_\odot) > 9$ bin in FRESCO.

In the NGDEEP sample, we additionally require the detection of [O III], [O II], and Hβ, with the detection of either Hα or Hγ, to ensure the usage of both the R3 and [O II]/Hβ (R2; R. L. Sanders et al. 2024) indices and the usage of a reliable dust correction using the Balmer decrement. With the R2 calibration, we can resolve the double-branch ambiguity caused by the R3 calibration, since the [O II]/Hβ ratio increases monotonically with metallicity at $12 + \log(\mathrm{O/H}) \lesssim 8.7$ (R. Maiolino et al. 2008). We apply an SNR cut of 3 for both [O III] and [O II], which has been shown to yield reliable metallicity gradient measurements (X. Wang et al. 2020).

In our final galaxy sample, we incorporate 88 galaxies traced by [O III] with high SNR at $1.7 < z < 3.5$ in NGDEEP, 325 galaxies at $5.3 < z < 7$ in ASPIRE, and 42 galaxies at $7 < z < 9$ in FRESCO GOODS-S and GOODS-N. From line ratio diagnostics, these galaxies are negligibly impacted by active galactic nucleus (AGN) contamination (see Appendix I).

3. Methods

3.1. Stacking Emission Maps

We create median maps of [O III] and Hβ by stacking individual galaxies in two redshift bins. As the galaxies in our sample have been observed under uniform conditions and have stable PSFs, stacking gives us the advantage of not requiring individual emission lines to be detected with high SNR. The stacked maps allow us to explore regions with larger radii and lower surface brightness. We first weight each map by [O III] flux measured in the 1D spectrum, to avoid excessive weighting toward bright sources (E. J. Nelson et al. 2016; X. Wang et al. 2022b). To avoid introducing additional noise and systematics, we stack the galaxies without distorting the maps through rescaling, rotation, or deprojection. We use `DAOStarFinder` (L. Bradley et al. 2024) to fit the centroid of each galaxy's emission map using a Gaussian kernel, and then stack the centrally aligned emission maps. Several relevant works also measure centers based on emission lines (N. M. Förster Schreiber et al. 2009, 2018; E. Wuyts et al. 2016), although different methods may be applied. For example, N. M. Förster Schreiber et al. (2018) define the galaxy center as the morphological center of the outer isophote of the emission-line map. An alternative approach is to fit the center based on the continuum, but we do not find a significant difference in the stacked line profiles. We create median stacks by measuring the median value at each pixel. The uncertainties of the radial profiles are computed by bootstrapping the stacks. The physical sizes of the stacks are computed using the median redshift $z = 6.28$ and $z = 7.24$ in the ASPIRE and FRESCO stacks, respectively. In Figure 1, we show the stacked [O III] and Hβ maps in the first two panels, and the normalized radial profiles in the third panel.

3.2. SED Fitting

The SED modeling of our sample is performed with the Bayesian code `BEAGLE` (J. Chevallard & S. Charlot 2016) with the broadband photometry and flux of [O III] lines as inputs. We adopt a delayed-τ star formation history (SFH), the SMC dust attenuation law, and a Chabrier initial mass function (G. Chabrier 2003) with an upper limit of $100 M_\odot$. We assume a flat prior in log-space for the characteristic star formation timescale, τ, spanning 10^7 to $10^{10.5}$ yr, and for stellar masses from $10^4 M_\odot$ to $10^{12} M_\odot$. We set the effective optical depth in the V band in the range of $\hat{\tau}_V \in [0, 0.5]$ with a log-uniform prior. In FRESCO, the spectral fitting is performed using three JWST NIRCam bands (F182M, F210M, and F444W) and eight Hubble Space Telescope (HST) bands (F435W, F606W, F775W, F810W, F105W, F125W, F140W, and F160W). For the NGDEEP galaxy sample, we use eight JWST NIRCam bands (F090W, F182M, F210M, F277W, F335W, F356W, F410M, and F444W), three JWST NIRISS bands (F115W, F150W, and F200W), and eight HST bands (F435W, F606W, F775W, F810W, F105W, F125W, F140W, and F160W). For ASPIRE, we use three JWST NIRCam bands (F115W, F200W, and F356W).

We note that for the ASPIRE and FRESCO samples, we lack photometric bands (e.g., JWST MIRI) to fully constrain the SED shape redward of the 4000 Å break, which is sensitive to the age and dust properties of stellar populations. This limitation increases uncertainties in the derived physical properties, consistent with findings by Q. Li et al. (2024), who report that stellar mass and star formation rate (SFR) uncertainties can increase by ∼0.1 dex without MIRI coverage. Although the SED code gives small errors on stellar masses, they only represent model-based statistical uncertainty. There are also notable systematic uncertainties in the assumed models, including uncertainties in SFH, initial mass function, and dust attenuation laws. Stellar masses of young galaxies can differ significantly with different SFHs. Since young stars dominate the observed rest-UV and optical SED, they outshine potential older stellar populations that may be present. Consequently, the inferred early SFH, along with the estimated age and stellar mass of a given system, can be sensitive to the choice of SFH model. Parametric SFHs, like constant and delayed-τ SFHs, tend to underestimate the stellar masses compared with nonparametric SFHs, as they are not flexible enough to capture early star formation (L. Whitler et al. 2023; M. A. C. de los Reyes et al. 2025). Since many recent works on MZRs have adopted parametric SFHs to estimate stellar masses (J. Matthee et al. 2023; M. Curti et al. 2024; A. Sarkar et al. 2025), and our sample selection directly relies on the MZR provided by A. Sarkar et al. (2025), the adherence of the parametric SFH model can be more consistent with their formalism.

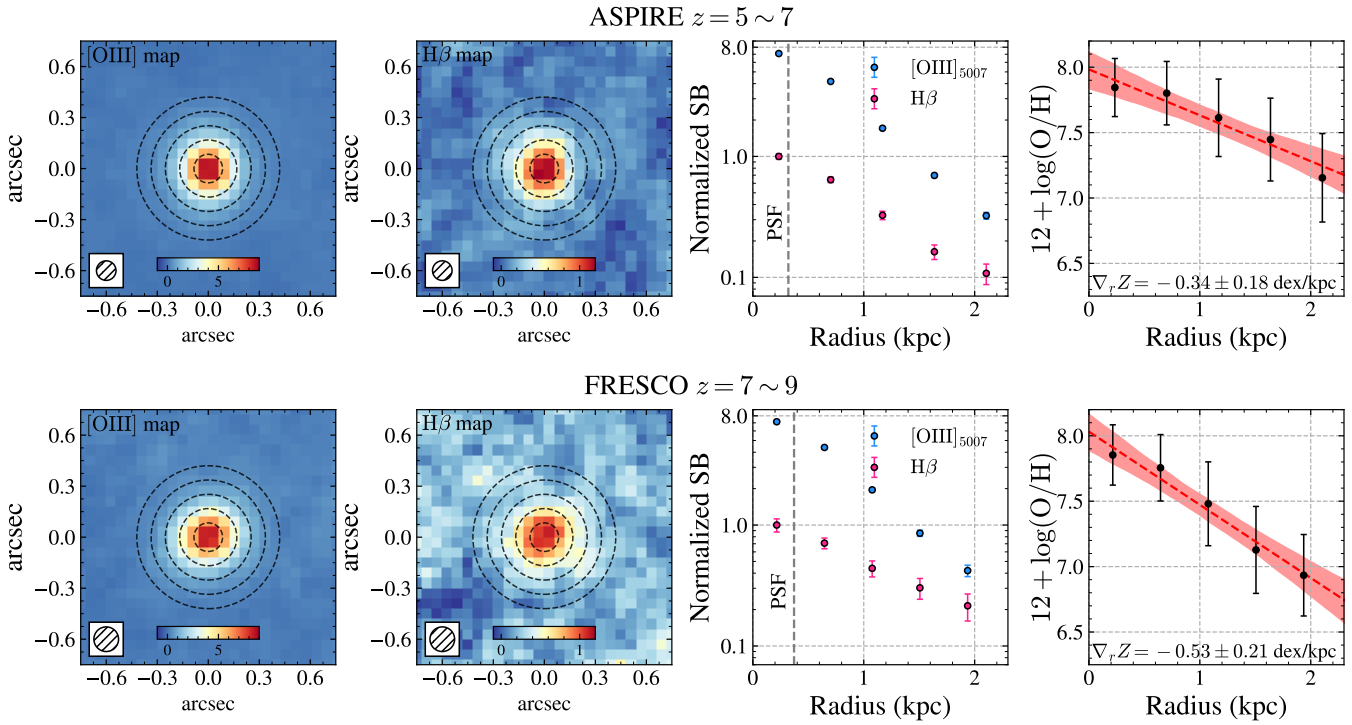


Figure 1. Stacking results of $5 < z < 7$ galaxies in ASPIRE (top) and $7 < z < 9$ galaxies in FRESCO (bottom). The first two columns show the median stack of emission maps ([O III], Hβ). As indicated by the color bar in the lower center of the line maps, the units correspond to surface brightness normalized by the peak Hβ flux. The black dashed annuli mark the region where we measure the line profiles. The FWHM of PSFs is shown on the bottom left in each line map. The third column shows the normalized surface brightness profiles of [O III] and Hβ measured within different annuli. The vertical dashed line marks the angular resolution (half of the PSF FWHM). The fourth column shows the fitted metallicity gradient. The black error bars include both statistical error and calibration error. The red shadow represents the 1σ confidence interval of the linear regression. Both samples in each redshift bin show steep gradients toward the center.

On the other hand, the choice of dust attenuation models also influences the fitted results. The empirical dust attenuation curves are derived based on nearby galaxies. The Calzetti law (D. Calzetti et al. 1994) was derived from a sample of local starburst galaxies, while the SMC law was measured with the Small Magellanic Cloud, showing a sharp rise at short wavelengths (J. C. Weingartner & B. T. Draine 2001). However, as the dust properties may evolve through cosmic time, the local templates may not be applicable for all redshifts. V. Markov et al. (2023) found that the estimated stellar masses in a galaxy sample at $z = 7$–8 can differ up to ∼0.3 dex depending on whether the Calzetti or SMC attenuation law is assumed. V. Markov et al. (2025) showed that the median dust attenuation slopes of galaxies at $z \sim 2$–3.5 approach the SMC slope, but gradually flatten with increasing redshift beyond $z > 3.5$. Our choice of the SMC dust law is appropriate for our $z = 1.7$–3.5 sample, but it may not provide the best fit for our $z > 5$ samples. A more flexible model may better characterize the dust properties at different redshifts (e.g., see V. Markov et al. 2023; A. de la Vega et al. 2025), but it is beyond the scope of this paper. Therefore, we caution potential uncertainties raised by different dust templates.

According to our SED modeling, the stellar masses in our final NGDEEP sample range from 10^7 to $10^{10} M_\odot$ with a median mass of $10^{8.43} M_\odot$, and the stellar masses in ASPIRE and FRESCO range from 10^7 to $10^{10} M_\odot$ with a median mass of $10^{8.16} M_\odot$ (Table 4 in Appendix L).

3.3. Metallicity Measurements

We jointly constrain metallicity, nebular dust extinction (A_v), and dereddened Hβ flux ($f_{\mathrm{H}\beta}$), by the Bayesian inference method for forward modeling (X. Wang et al. 2017; Z. Li et al. 2022). The likelihood function is defined as

$$\mathcal{L} \propto \exp\left(-\frac{1}{2} \cdot \sum_i \frac{(f_{\mathrm{EL}_i} - R_i \cdot f_{\mathrm{H}\beta})^2}{(\sigma_{\mathrm{EL}_i})^2 + (f_{\mathrm{H}\beta})^2 \cdot (\sigma_{R_i})^2}\right), \quad (1)$$

where f_{EL_i} and σ_{EL_i} represent the dereddened emission-line (e.g., [O III], [O II], Hγ, Hβ, or Hα) flux. The dereddened line fluxes are corrected from observed line ratios assuming the SMC dust law with the normalization A_v as a free parameter in the fitting. R_i corresponds to the expected line flux ratios between EL$_i$ and Hβ (i.e., R_i can be the Balmer decrement of Hγ/Hβ = 0.6, Hα/Hβ = 2.86, assuming Case B recombination, and the metallicity diagnostics of [O III]/Hβ and [O II]/Hβ), and σ_{R_i} is their intrinsic scatter. We perform Markov Chain Monte Carlo (MCMC) sampling using the package *emcee* (D. Foreman-Mackey et al. 2013b). Potential stellar absorption may impact the dust corrections. The typical equivalent width of Hβ is found to be EW(Hβ) ∼ 1–6 Å in local Hβ absorption line galaxies (R. M. McDermid et al. 2015). A. C. Carnall et al. (2023) reported EW(H$_\delta$) = 7.9 Å for a quiescent galaxy at $z = 4.658$. However, for star-forming galaxies dominated by young O and B stars, the equivalent widths of Balmer absorption lines are much weaker. Low-mass star-forming galaxies are found to exhibit high [O III] and Hβ equivalent widths both at $z \sim 2$ and at higher redshift $z \sim 6$ (A. van der Wel et al. 2011; J. Matthee et al. 2023), with Hβ emission reaching EW(Hβ) ∼ 100 Å. The Balmer absorption line strength contributes only to ∼1%–10% of the emission-

line strength and can be negligible. Several recent works also correct for dust by the observed Balmer emission-line ratios without taking stellar absorption into account (M. Curti et al. 2023; J. Matharu et al. 2023).

For the $z > 5$ galaxy sample, we adopt the metallicity diagnostics presented by R. L. Sanders et al. (2024). R. L. Sanders et al. (2024) built an empirical calibration using a sample of $z = 2$–9 galaxies with [O III] λ4363 detected using the direct T_e method, which better characterizes the properties of high-redshift galaxies. We also refer to D. Scholte et al. (2025) and P. Chakraborty et al. (2025), who provide similar T_e calibrations. One caveat for the [O III]/Hβ (R3) diagnostics is that it yields a double-branched solution if given a line ratio. For our NGDEEP sample, we have detections of both [O II] λ3727 and [O III] λ5007. With the additional [O II] λ3727 we can distinguish the two solutions using $[\mathrm{O\,II}]_{3727}/\mathrm{H}\beta$ (R2). However, in the ASPIRE and FRESCO samples, $[\mathrm{O\,II}]_{3727}$ is out of the spectral coverage. Thus we assume the lower-branch solution with $12 + \log(\mathrm{O/H}) \lesssim 7.9$, given the fact that most galaxies at similar redshifts of $z \geqslant 6$–7 have had their metal-poor nature verified using the T_e method (M. Curti et al. 2023; T. Jones et al. 2023; K. Nakajima et al. 2023; J. R. Trump et al. 2023; R. L. Sanders et al. 2024; P. Chakraborty et al. 2025).

We rely on the strong-line diagnostics from F. Bian et al. (2018) for sample galaxies at $1.7 < z < 3.5$, which was built from local analogs of $z \sim 2$ galaxies. As the R. L. Sanders et al. (2024) model is built from a sample of high-redshift galaxies with direct metallicity measurements, it is more applicable for our sample of relatively metal-poor $z > 5$ galaxies. However, this model does not well constrain potential higher-metallicity galaxies as its calibration is fitted with insufficient numbers of high-metallicity galaxies. Because of this issue with R. L. Sanders et al. (2024), we choose to use the F. Bian et al. (2018) model as it is robust for higher-metallicity galaxies and is widely used for galaxies at $z \approx 2$ (X. Wang et al. 2020).

In Appendix E, we test the systematic uncertainties arising from different strong-line calibrations and find that alternative methods (R. Maiolino et al. 2008; K. Nakajima et al. 2022; M. Curti et al. 2023) yield consistent metallicity gradient measurements with our fiducial measurements using R. L. Sanders et al. (2024) and F. Bian et al. (2018).

3.4. Gradient Measurements

To securely measure metallicity gradients in the NGDEEP sample, we apply Voronoi tessellation (M. Cappellari & Y. Copin 2003) to each galaxy's [O III] map to bin emission maps into subregions, each having SNR > 5 in [O III]. Voronoi tessellation is superior to averaging the signal in radial annuli because the former keeps the information of azimuthal variations in metallicity. After the binning, the number of resolution elements is $\gtrsim$50 in our NGDEEP sample. Increasing the SNR threshold leads to coarser binning, which averages out spatial variations, whereas lowering the threshold yields more bins with low SNRs, thereby increasing uncertainty in our measurements. An SNR cut of 5 has been tested to be an appropriate threshold for recovering robust gradient measurements, which balances both SNR and resolution elements (X. Wang et al. 2020). However, for our stacked sample, the azimuthal variations have already been averaged during stacking, so we simply average the flux in different annuli for $z > 5$ samples. We evenly divide the maps into five concentric annuli with radii from $r = 0$ to $r = 0\farcs42$. Due to the different redshifts of the ASPIRE and FRESCO samples, the physical sizes of the corresponding annuli are slightly different (see Figure 1).

Table 1
Metallicity Gradients from Median Stacks in Two $z > 5$ Bins

Survey	Redshift	$\nabla_r \log(\mathrm{O/H})$ (dex kpc^{-1})
ASPIRE	$5.3 < z < 7.0$	-0.34 ± 0.18
FRESCO	$7.0 < z < 9.0$	-0.53 ± 0.21

We use a simple linear least-squares method to fit the metallicity gradients with the following formula:

$$12 + \log(\mathrm{O/H}) = \theta_0 + \theta_1 r \quad (2)$$

where θ_0 is the intercept, θ_1 is the slope of the fitted line, and r is the galactocentric radius in kiloparsecs.

For the NGDEEP galaxies, we deproject the radius to the source plane. We assume the galaxies are circular disks, and the observed elasticity is due to projection effects. The inclination angle, i, is estimated by the axis ratio from Sérsic fitting: $\cos(i) = b/a$. The distance of each position to the center is then corrected by inclination. We note the assumption of disks may not apply for every galaxy at high redshift, especially given the recent discovery of fewer disklike galaxies toward high redshift (R. C. Simons et al. 2017; V. Pandya et al. 2024). While some works have corrected for projection (N. M. Förster Schreiber et al. 2018; X. Wang et al. 2020; Z. Li et al. 2022), others directly use the projected distance to measure gradients (M. Curti et al. 2020a; R. C. Simons et al. 2021). We find a $\sim$0.1 dex kpc^{-1} difference between the projected gradients and deprojected gradients, comparable to the 1σ statistical uncertainties. We use the deprojected results of the NGDEEP sample throughout this article to remain consistent with our previous work (X. Wang et al. 2020; Z. Li et al. 2022). The stacking method for the $z > 5$ sample averages the projection effects, so we do not correct for projection for the $z > 5$ sample. The systematic uncertainties of the individual and the stacked measurements are estimated by mock observations in Appendix G. We find good agreement between the stacked measurements and the median of the individual samples.

In the fourth panels of Figure 1, we show the measurements on median stacks, and the individual measurements of the NGDEEP sample are shown in Figure 25 in Appendix M. The metallicity gradients measured from the median stacks are reported in Table 1, and the individual measurements of the NGDEEP sample are reported in Table 4 in Appendix L.

We note that variation in R3 ratios could also be attributed to changes in the ionization parameter (U) (P. Garg et al. 2024), which potentially lead to a bias in metallicity measurements. In Appendix H, we show that the variations in R3 ratios are more significant than can be accounted for by changes in the ionization parameter, and potential variation in U only leads to an uncertainty of $\sim$0.1 dex kpc^{-1}. The steep gradients observed in Figure 1 are therefore robust.

4. Results and Discussion

4.1. Redshift Evolution of Metallicity Gradients

To have a more complete study of the evolution of the metallicity gradient, we have added existing measurements from a

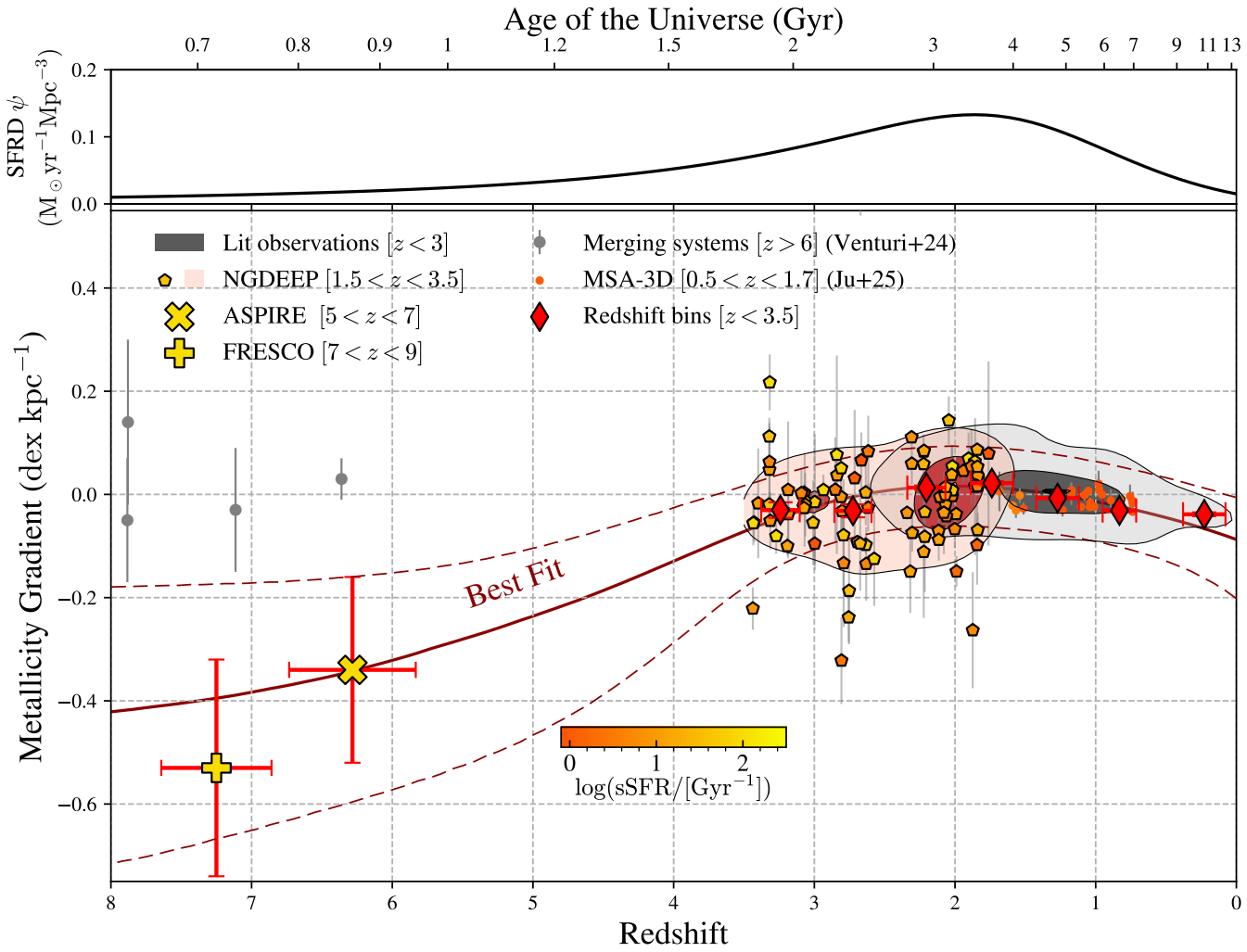


Figure 2. The redshift evolution of metallicity gradients. The thin red diamonds show the weighted mean in redshift bins at $z = [0, 3.5]$ from both this work and the literature. A "×" and "+" represent the median stacks of the ASPIRE sample at $z \approx 6$ and the FRESCO sample at $z \approx 7$. The NGDEEP sample at $z \approx 1$–3 is denoted by pentagons. The points from ASPIRE, FRESCO, and NGDEEP are color coded by specific SFR (sSFR). Gray contours encompass individual measurements at $z < 3$ in the literature using HST (X. Wang et al. 2017, 2020; R. C. Simons et al. 2021), and ground-based AO-assisted surveys (A. M. Swinbank et al. 2012; T. Jones et al. 2013; N. M. Förster Schreiber et al. 2018). We also include seeing-limited observations (L. van Zee et al. 1998; J. Queyrel et al. 2012; E. Wuyts et al. 2016; D. Carton et al. 2018; M. Curti et al. 2020a). The red contour only encompasses our observations at $1.7 < z < 3.5$, overlapping with literature observations at $z \sim 2$. Gray circles with error bars represent interacting galaxies at $z > 6$ recently observed with JWST NIRSpec (G. Venturi et al. 2024). Recent JWST MSA-3D observations (M. Ju et al. 2025) are noted in circles, color coded by sSFR. The dark-red line shows our best fit of all stacks, with dashed lines in the same color showing the 1σ confidence interval of the fit. In the upper panel, we plot the cosmic SFRD (P. Madau & M. Dickinson 2014), where the evolution of metallicity gradients also peaks at $z \sim 2$.

variety of sources, including a wealth of JWST observations (M. Ju et al. 2025), HST observations (X. Wang et al. 2017, 2020, 2022a; R. C. Simons et al. 2021), and ground-based adaptive optics (AO)-assisted observations (A. M. Swinbank et al. 2012; T. Jones et al. 2013; N. M. Förster Schreiber et al. 2018). These measurements are made mainly at $z \lesssim 2$. Additionally, we include samples obtained under seeing-limited conditions (L. van Zee et al. 1998; J. Queyrel et al. 2012; E. Wuyts et al. 2016; D. Carton et al. 2018; M. Curti et al. 2020a) for a more comprehensive analysis. Although different metallicity calibrators (e.g., R23, N2, and O3N2; R. Maiolino et al. 2008; M. Curti et al. 2020b) are used in different works, the systematic difference should not change the general trend in gradients (D. S. N. Rupke et al. 2010b; X. Wang et al. 2019), despite some potential impacts. Specifically, H. Poetrodjojo et al. (2021) found that metallicity gradients measured using different line diagnostics show deviations; however, different measurements generally correlate well with each other, with Spearman's rank coefficients greater than 0.6. Our final collected literature samples cover the redshift range $0 < z < 2.5$.

We note a difference between our $z > 6$ results and recent JWST NIRSpec observations of G. Venturi et al. (2024) (Figure 2). The results in G. Venturi et al. (2024) are based on the gradient measurement of four galaxies at $z > 6$, selected with the largest offset between UV/optical and far-infrared emission lines (S. Carniani et al. 2021). Such offsets may indicate nontypical dust geometry and/or galactic structures so that the dust and star-forming regions are not uniformly distributed, possibly induced by merger events (also see discussions in A. M. Garcia et al. 2025). Therefore, their sample differs from our main-sequence galaxy sample at $5 < z < 9$, which is unbiasedly selected based on NIRCam WFSS observations. These interacting systems are expected to undergo violent gas mixing (P. B. Tissera et al. 2022), and as such, the metal distribution is expected to be flatter than that of normal star-forming galaxies. These effects are likely to

explain the differences in our sample. Thus, we have not included these systems in our high-redshift sample for statistical analysis.

In Figure 2, we show the redshift evolution of the metallicity gradients from $z = 8$ to $z = 0$. We see that the cosmic evolution of the metallicity gradient shows ascending and descending phases before and after $z \approx 2$. Galaxies show steep negative gradients at the earliest times ($z > 5$). Such a steep gradient has been found in local extremely metal-poor galaxies (Y. Kashiwagi et al. 2021) and is indicative of inside-out growth and inefficient gas mixing. The metallicity gradients flatten over time until $z \sim 2$ and steepen again toward the present. To describe this trend, we first apply a linear regression analysis for individual galaxies. We divide the individual galaxies into two redshift bins, $1.75 < z < 3.5$ and $0 < z < 1.75$. The ascending phase with time can be described as $-0.046 \pm 0.012\,\mathrm{dex\,kpc^{-1}}/\delta z$ from $z = 3.5$ to $z = 1.75$, followed by a descending phase of $0.040 \pm 0.006\,\mathrm{dex\,kpc^{-1}}/\delta z$ from $z = 1.75$ to $z = 0$. We apply the likelihood ratio test to see if this evolution is significant compared with a nonevolving metallicity gradient. We find a p-value $= 1.3 \times 10^{-22}$, indicating a strong preference for the model with redshift evolution.

To better characterize the rising and descending phases for the metallicity gradients, we apply a double-power-law model:

$$\nabla_r \log(\mathrm{O/H})(z) = \gamma_0 \frac{(1+z)^{\gamma_1}}{1 + [(1+z)/\gamma_2]^{\gamma_3}} - \gamma_4. \quad (3)$$

Fitting the above function to our observed data (Appendix D), we obtain the following constraints: $\gamma_0 = 0.40^{+0.30}_{-0.19}$, $\gamma_1 = 0.27^{+0.16}_{-0.09}$, $\gamma_2 = 5.17^{+1.15}_{-0.59}$, $\gamma_3 = 5.23^{+2.42}_{-2.01}$, and $\gamma_4 = -0.50^{+0.20}_{-0.31}$. This fitting provides a first glimpse into the redshift evolution of the metallicity gradient with an ascending phase scaling as $\nabla_r \log(\mathrm{O/H})(z) \propto (1+z)^{-5.23}$ at $2 \lesssim z \lesssim 8$, and transforming to a slowly descending phase to the present day, scaling as $\nabla_r \log(\mathrm{O/H})(z) \propto (1+z)^{0.27}$. This analytical form suggests the transition happens around $z \approx 2.1$. Intriguingly, it is close to the peak of the SFRD (P. Madau & M. Dickinson 2014). Similar transitions at the same epoch include the break of the fundamental metallicity relation (M. Curti et al. 2024) at $z > 2$ and the break of MZRs of damped Lyα systems at $z > 2.6$ (P. Møller et al. 2013). They point to certain mode transitions about galaxy formation and chemical enrichment at cosmic noon.

Since galaxies are intrinsically smaller at higher redshift (T. Morishita et al. 2024), a fixed metallicity change implies steeper gradients in smaller systems. Given the correlation between galaxy size and stellar mass, expressing gradients relative to the effective radius R_e helps reduce dependence on mass and size (I. T. Ho et al. 2015; F. Bresolin 2019). In addition, the normalization of the effective radius helps reduce the impact from the beam-smearing effects of seeing-limited data (E. Wuyts et al. 2016). In Appendix B, we discuss the metallicity gradients measured with the effective radius, and the trend of redshift evolution is consistent.

4.2. Comparison with Hydrodynamical Simulations

In Figure 3, we compare our observations with simulations from MUGS (B. K. Gibson et al. 2013), MAGICC (B. K. Gibson et al. 2013), FIRE (X. Ma et al. 2017), FIRE-2 (X. Sun et al. 2025), TNG50 (Z. S. Hemler et al. 2021), EAGLE (P. B. Tissera et al. 2022), FOGGIE (A. Acharyya et al. 2025), and Illustris (A. M. Garcia et al. 2025). Those simulations adopt different settings for physical processes, including the treatment of feedback (see Appendix K).

The FOGGIE simulations show galaxies evolving from steep metallicity gradients at $z = 6$ to shallower gradients over time. They attribute the flattening to an accretion-dominated phase toward $z \sim 1$. While this evolutionary trend from $z > 5$ to $z \sim 2$ is qualitatively similar to our observations, the FOGGIE simulations predict gradients that are systematically steeper overall. A. Acharyya et al. (2025) noted that the SN feedback prescription in the FOGGIE simulations is insufficient to expel enough metal-rich gas, which may lead to the formation of steeper gradients compared with observations, especially at $z \sim 2$. The Illustris simulation shows a similar trend to FOGGIE, in which the galaxies evolve initially with steep gradients and flatten over time. A. M. Garcia et al. (2025) obtained the redshift evolution from the Illustris simulation as $-0.016 \pm 0.004\,\mathrm{dex\,kpc^{-1}}\delta z$. In addition, from the EAGLE (S. McAlpine et al. 2016), IllustrisTNG (D. Nelson et al. 2019), and SIMBA (R. Davé et al. 2019) simulations, A. M. Garcia et al. (2025) found similar redshift evolutions in the range of -0.015 to $-0.028\,\mathrm{dex\,kpc^{-1}}\delta z$. While these values are broadly consistent with our observed ascending phase, the simulations do not reproduce the peaking gradients around $z \sim 2$. A. M. Garcia et al. (2025) suggest that simulations with smooth stellar feedback may not efficiently mix metals, which could contribute to the discrepancy between simulations and observations around $z \sim 2$.

At $z \lesssim 3$, our observations align more closely with the EAGLE, FIRE, FIRE-2, and MAGICC simulations, whereas TNG50 predicts slightly more negative gradients and MUGS yields overly steep gradients at $z \gtrsim 1$. Differences in the evolution of metallicity gradients in these simulations can be interpreted as a result of the use of different feedback models. The feedback drives outflows that transport metal-enriched gas to larger radii or out of the galaxy, where it mixes with metal-poor gas in the ISM and CGM. This process redistributes the metals and results in a flat metallicity gradient. The more bursty feedback explicitly modeled in the FIRE simulation produces flat gradients at $z \gtrsim 1$ (X. Ma et al. 2017), while the smooth feedback in TNG50 allows for more negative gradients (Z. S. Hemler et al. 2021). MAGICC is a variation of the MUGS simulation, running with the same code, except for the enhanced thermal feedback energy from SNe and the inclusion of radiation from massive stars in the MAGICC simulation (B. K. Gibson et al. 2013). The enhanced feedback schemes in MAGICC produce flatter gradients at $z \sim 2$. However, the redshift evolution in the FIRE simulations is less evident, as they exhibit consistently flat gradients across all redshifts, likely due to the episodic nature of the strong feedback model used in FIRE (X. Ma et al. 2017; X. Sun et al. 2025). EAGLE simulations also suggest that galaxy mergers or interactions that trigger gas inflows and starbursts could flatten or invert metallicity gradients (P. B. Tissera et al. 2022). Therefore, the observed metallicity gradients at $z \sim 2$ generally prefer models with strong feedback as a consequence of bursty star formation, though the feedback strength should not result in persistently flat gradients. Thus, the evolution of the metallicity gradient indicates an evolving feedback strength through cosmic time. There is evidence that the energy and momentum injection from massive stars peaks at $z \approx 2$ (T. M. Heckman & P. N. Best 2023). At this redshift, outflows

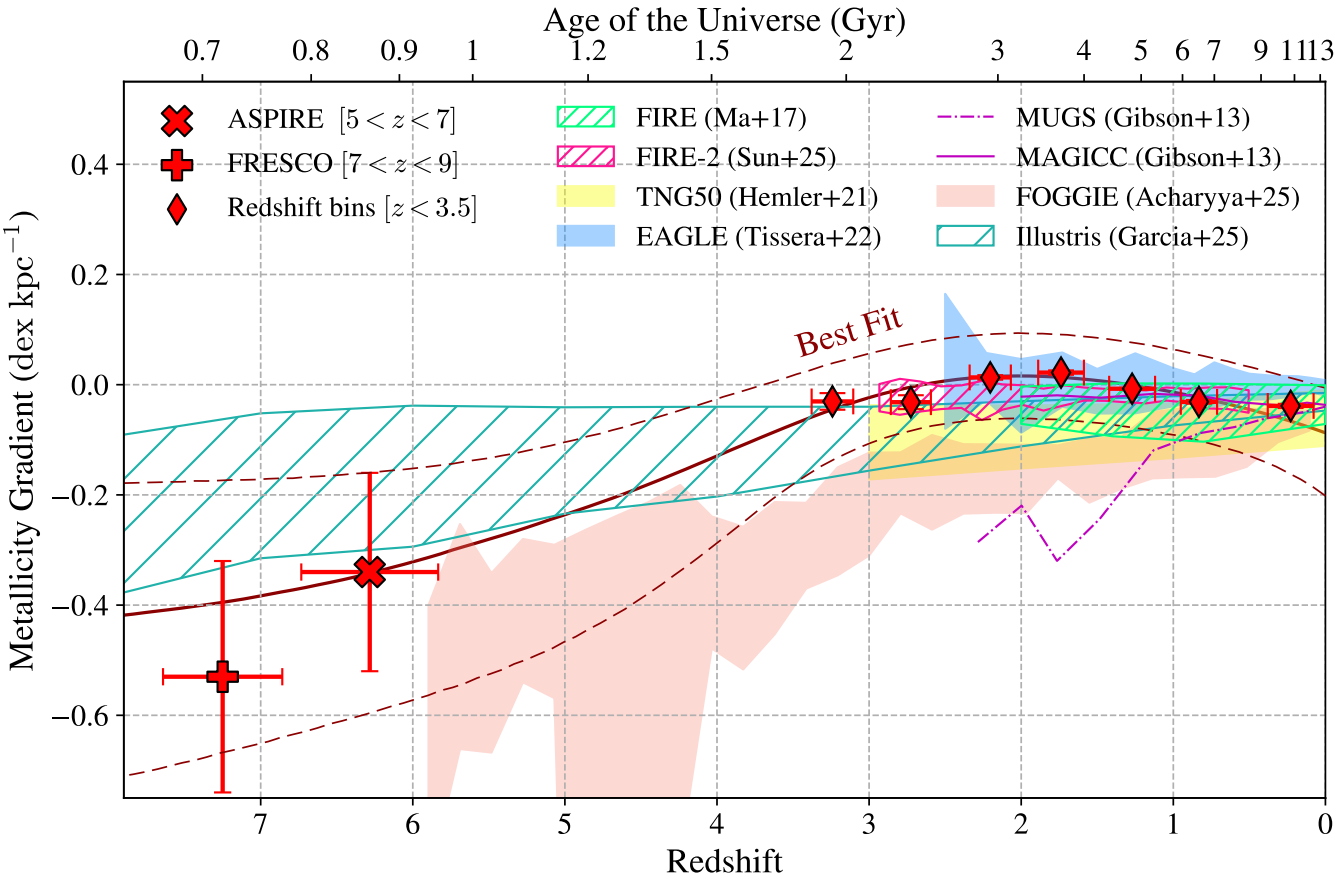


Figure 3. The comparison between the observed best-fit redshift evolution of metallicity gradients and predictions from different suites of cosmological simulations. The simulations including FIRE (X. Ma et al. 2017), FIRE-2 (X. Sun et al. 2025), TNG50 (Z. S. Hemler et al. 2021), EAGLE (P. B. Tissera et al. 2022), MUGS/MAGICC (B. K. Gibson et al. 2013), FOGGIE (A. Acharyya et al. 2025), and Illustris (A. M. Garcia et al. 2025) are marked in different colors.

are found to efficiently redistribute large proportions of metals from galaxies (R. L. Sanders et al. 2023), leading to flat/positive gradients (X. Wang et al. 2019). After this epoch, feedback strength should diminish as a consequence of decreased SFR (S. Yu et al. 2023), allowing negative gradients to develop.

In addition to outflow launched by feedback, inflows due to cosmic gas accretion from the intergalactic medium may also mix with the ISM gas and redistribute metal content. Cosmic accretion history in simulations suggests that cold-mode accretion dominates the total accretion rate at $z \sim 2$, and hot-mode accretion becomes increasingly important at lower redshifts (D. Kereš et al. 2009; F. van de Voort et al. 2011). At these lower redshifts, hot-mode accretion forms a shock shell around the massive halos $M_h \gtrsim 10^{12} M_\odot$ (Y. Birnboim & A. Dekel 2003; A. Dekel et al. 2009). As the cooling time is longer than the infall time, the shock-heated gas cannot cool and collapse inward efficiently, thereby limiting its ability to fuel star formation readily. However, cold-mode accretion is predicted to penetrate deep inside the halo without being shock-heated because the cooling time is shorter than the shock compression time. Such cold flows reduce the metallicity in the galactic center via filaments penetrating halos, which causes the flattening/inverting of the metallicity gradients at cosmic noon (G. Cresci et al. 2010; Z. Li et al. 2022).

The joint effect of feedback and gas accretion may contribute to the flat/positive gradients observed during cosmic noon. Another contributing factor is mergers and gravitational interactions, which can also drive gas inflows and enhance both star formation and feedback (B. Pérez-Díaz et al. 2024). In particular, stochastic gas accretion associated with these mergers may play an important role in low-mass galaxies, in the context of hierarchical structure formation. P. Torrey et al. (2012) found that gravitational tidal forces during galaxy mergers or close interactions can generate significant inflows, transporting metal-poor gas inward and redistributing metals within the disk. P. B. Tissera et al. (2022) showed that such merger-driven inflows are common in high redshifts, and can account for a notable fraction (∼20%) of the positive metallicity gradients at $z \lesssim 2$ observed in the EAGLE simulations.

In contrast, measurements at $z \sim$ 6–8 reveal steeper gradients compared to galaxies at $z \sim 2$, possibly indicating a reduced gas mixing at cosmic dawn allowed for the existence of steeper gradients in early galaxies. Recent works have revealed numerous instances of star-forming galaxies that deviate from the main sequence of the fundamental metallicity relation established at $z \lesssim 2$ (M. A. Lara-López et al. 2010). These deviations, characterized by metal deficiency at $z \gtrsim 7$ (K. E. Heintz et al. 2023; K. Nakajima et al. 2023), imply a continuous pristine gas inflow that effectively dilutes the metal abundances in these galaxies. We indeed observe a metallicity deficiency in the ASPIRE sample, which is consistent with the analytical model featuring intense gas inflows, as discussed in our companion work (Z. Li et al. 2025). With continuous gas inflow diluting galaxies, steep metallicity gradients may exist when star formation efficiency is high and local metal production is high at the center. Theoretical models (P. Sharda et al. 2021b; E. Wang & S. J. Lilly 2022) have suggested the importance of in situ metal production in building negative

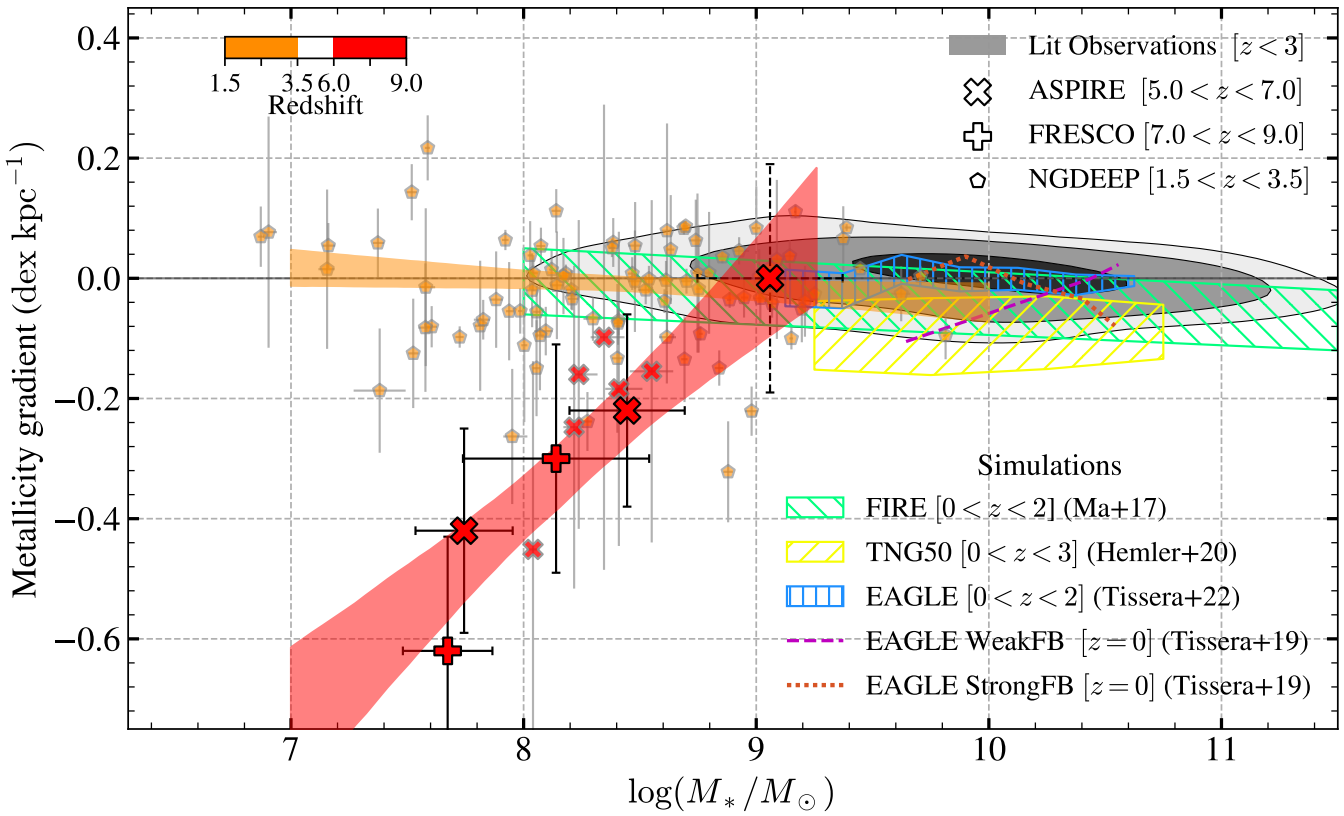


Figure 4. The mass dependence of metallicity gradients. Galaxies in NGDEEP, ASPIRE, and FRESCO are labeled with a pentagon, "×," and "+," respectively, color coded by redshift. The highest-mass bin in ASPIRE is marked with a dashed error bar; for this bin, the metallicity gradients were derived using the higher-branch solution. The red and orange shaded areas represent the 1σ confidence intervals for the linear fit to the stacked data at $z > 5$ and the individual NGDEEP measurements within $1.7 < z < 3.5$, respectively. The literature observations (A. M. Swinbank et al. 2012; T. Jones et al. 2013; E. Wuyts et al. 2016; X. Wang et al. 2017, 2020; D. Carton et al. 2018; N. M. Förster Schreiber et al. 2018; M. Curti et al. 2020a; R. C. Simons et al. 2021; M. Ju et al. 2025) are included as gray-filled regions with 1σ, 2σ, and 3σ levels of the density contours of the distribution. Green, yellow, and blue hatched areas represent the 1σ interval from the FIRE (X. Ma et al. 2017), TNG50 (Z. S. Hemler et al. 2021), and EAGLE (P. B. Tissera et al. 2022) simulations, respectively. The dotted orange and dashed magenta lines show the median metallicity gradients of EAGLE galaxies with strong feedback (StrongFB) and weak feedback (WeakFB) prescriptions (P. B. Tissera et al. 2019).

metallicity gradients. The observed negative metallicity gradients indicate that in situ star formation and metal production, driven by intense central starbursts, may dominate the metal distribution, while metal mixing by gas inflows and feedback-driven outflows likely plays a minor role.

At low redshift ($z \lesssim 1$), metallicity gradients tend to become negative again. The underlying mechanisms likely differ from those at $z > 5$, as star formation is significantly suppressed in hot halos at $z = 0$, due to inefficient cooling (F. van de Voort & J. Schaye 2012; C. A. Correa et al. 2018) and reduced gas accretion rates (C.-A. Faucher-Giguère et al. 2011; A. van der Wel et al. 2011). From the FIRE-2 simulation, R. L. Graf et al. (2024) suggest that as galaxies evolved, declining gas accretion (C. W. Trapp et al. 2022), reduced star formation (S. Yu et al. 2023) and feedback (M. E. Orr et al. 2022), and the settling of disks with reduced velocity dispersion and gas turbulence (J. C. Bird et al. 2021) led to a reduction in metal mixing, allowing negative gradients to develop and persist. Comparing observations in the local Universe and at $z \sim 2$ highlights a transition from strong turbulence, gas inflows, and feedback-driven mixing at high redshift to a more settled, secular evolution by $z \sim 0$.

4.3. Mass–Metallicity Gradient Relation

We explore how metallicity gradients change with galaxy mass, in terms of the mass–metallicity gradient relation (MZGR; also see M. Spolaor et al. 2009; P. Sharda et al. 2021a). We divide our FRESCO and ASPIRE samples into different mass bins and apply the same stacking method (Appendix A, Figures 9 and 10).

In Figure 4, we show the metallicity gradients as a function of stellar mass (MZGR), and compare the observations with the FIRE (X. Ma et al. 2017), TNG50 (Z. S. Hemler et al. 2021), and EAGLE (P. B. Tissera et al. 2022) simulations. The mass limits for JWST data in our work are deep enough to be used to compare the change of metallicity gradients over redshift.

We observe that they follow distinct log-linear relations at $z \sim 2$ and $z > 5$. For galaxies at $z > 5$, we find a notable positive correlation in MZGR quantified as $\nabla_r \log(\mathrm{O/H})_{5<z<9} = 0.30 \pm 0.19 \times \log(M_*/M_\odot) - 2.78 \pm 1.54$. This indicates that lower-mass galaxies exhibit steeper negative gradients. However, literature data (gray contours) suggest a weak negative mass dependence of the metallicity gradient within the redshift range $0 < z < 3.5$. Our observations from the NGDEEP sample at $1.7 < z < 3.5$ also reveal a marginal anticorrelation, quantified as $\nabla_r \log(\mathrm{O/H})_{1.7<z<3.5} = -0.020 \pm 0.018 \times \log(M_*/M_\odot) + 0.157 \pm 0.156$. This dependence on mass indicates the presence of different modes during those two periods.

A positive dependence on stellar mass is anticipated in secular galaxy evolution models, such that gradients flatten over time due to inside-out growth with the accumulation of stellar mass and metals (K. Pilkington et al. 2012). The

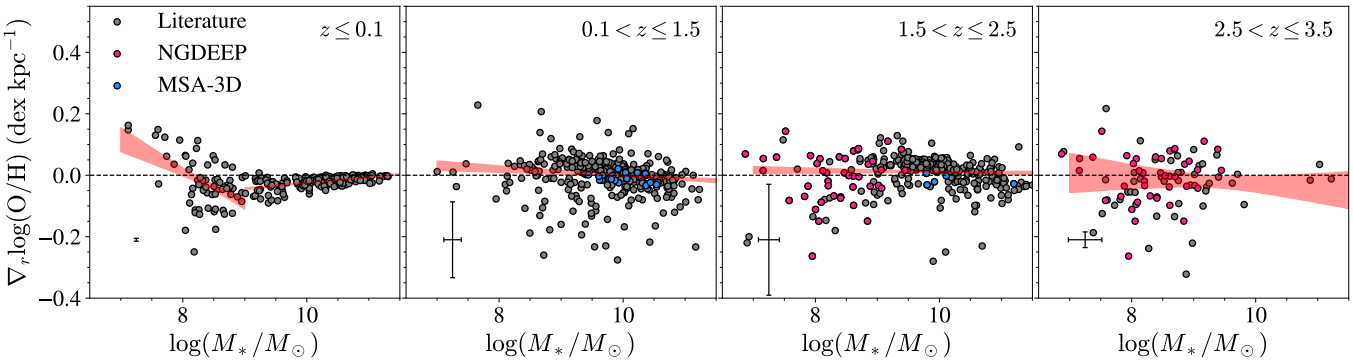


Figure 5. The MZGR for individual galaxies in four redshift bins. The red shadows represent the 1σ confidence interval of the linear regression. For $z \leqslant 0.1$ measurements, regressions are performed within two stellar mass bins, $\log(M_*/M_\odot) \leqslant 9$ and $\log(M_*/M_\odot) > 9$, while high-redshift measurements are performed across the entire mass range. The error bar on the bottom left of each figure represents the median uncertainty in stellar masses and gradients. The pre-JWST observations are shown as gray circles (A. M. Swinbank et al. 2012; T. Jones et al. 2013; I. T. Ho et al. 2015; L. Sánchez-Menguiano et al. 2016; E. Wuyts et al. 2016; D. Carton et al. 2018; N. M. Förster Schreiber et al. 2018; F. Bresolin 2019; M. Curti et al. 2020a; R. C. Simons et al. 2021; T. Li et al. 2025); NGDEEP and recent JWST MSA-3D (M. Ju et al. 2025) observations are highlighted in red and blue, respectively.

positive mass dependence was previously observed in local galaxies (I. T. Ho et al. 2015; F. Bresolin 2019), which is also predicted by EAGLE simulations at $z = 0$ with weak feedback (magenta dashed line in Figure 4; P. B. Tissera et al. 2019). W. M. Baker et al. (2025) also demonstrate the early signs of inside-out growth, with a galaxy at $z = 7.4$ that features a compact core and an extended star-forming disk. Our observed positive mass dependence further supports inside-out growth at cosmic dawn. Low-mass galaxies form earlier with concentrated cores, gradually building up outer disks over time. As they evolve, they accumulate stellar mass and metals at larger radii. The flattening of metallicity gradients in higher-mass galaxies highlights the dynamical nature of galaxy growth during cosmic dawn.

On the other hand, a negative relation at $z \sim 1$–2 has been observed by X. Wang et al. (2020), S. Gillman et al. (2021), and R. C. Simons et al. (2021), and they interpreted it as the effect of feedback-driven gas outflows, which is more pronounced in low-mass galaxies due to less restraint by lower gravitational potential (K. El-Badry et al. 2016; T. M. Heckman & P. N. Best 2023). FIRE simulations suggest that efficient gas outflow rates could efficiently mix metals, causing significant fluctuations in metallicity gradients (X. Ma et al. 2017). A similar negative relation is also found in EAGLE simulations with enhanced SN feedback (P. B. Tissera et al. 2019; see orange dotted line in Figure 4). From TNG50 simulations, A. M. Garcia et al. (2023) indicated that when the enrichment timescale is shorter than the gas mixing timescale, galaxies can maintain steep gradients. Additionally, if feedback processes effectively generate high-velocity gas outflows and turbulence, this can reduce the gas mixing timescale, leading to shallower gradients. The flattening of metallicity gradients toward low-mass galaxies arises as a result of stronger feedback effects with lower gravitational potential. The two distinct MZGRs at cosmic noon and cosmic dawn then indicate weak feedback at $z > 5$ and strong feedback at $z \sim 2$.

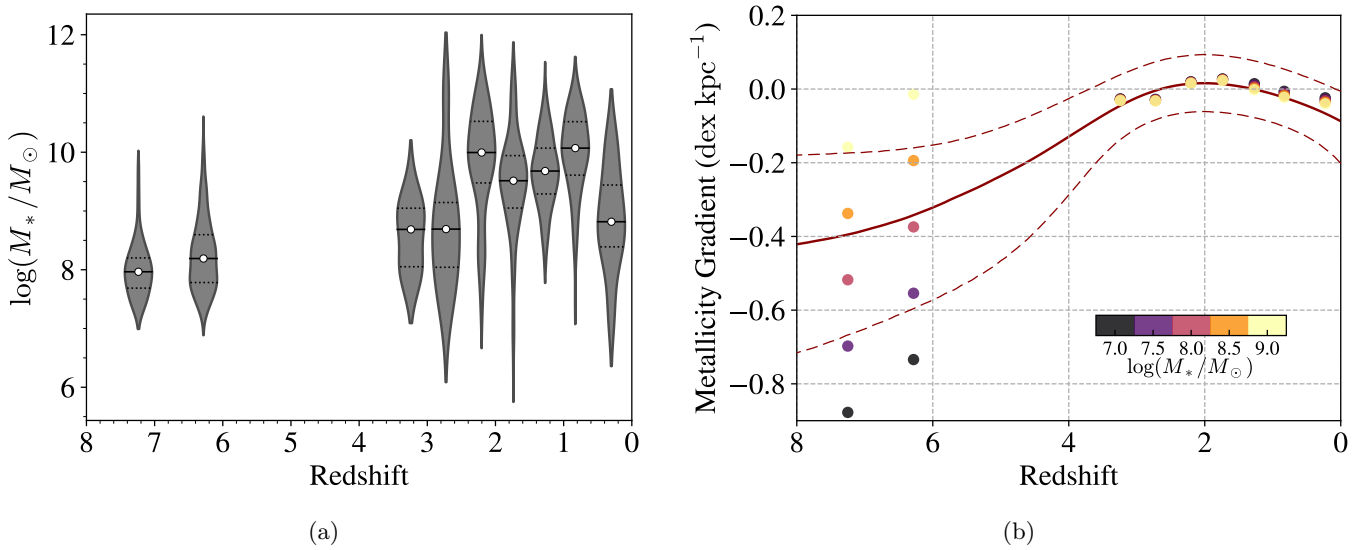
4.4. Redshift Evolution of MZGR

To further investigate whether the mass dependence of metallicity gradients evolves with redshift, we divide the sample of NGDEEP and literature observations into four redshift bins: $z \leqslant 0.1$, $0.1 < z \leqslant 1.5$, $1.5 < z \leqslant 2.5$, and $2.5 < z \leqslant 3.5$. Specifically, we have included the measurements from the local Universe at $z = 0$ reported by I. T. Ho et al. (2015), F. Bresolin (2019), and T. Li et al. (2025). We perform a linear regression on these measurements using the package LINMIX (B. C. Kelly 2007), with errors on both metallicity gradients and stellar masses taken into account. The regression is formulated as

$$\nabla_r \log(\mathrm{O/H})(M_*) = \alpha + \beta \log(M_*/M_\odot) + \mathrm{N}(0, \sigma), \quad (4)$$

where α is the intercept of the linear function and β is the slope of the linear function. N(0, σ^2) represents a normal distribution with σ representing the intrinsic scatter about the regression in units of dex kpc^{-1}. T. Li et al. (2025) recently reported an inverted MZGR for low-mass galaxies at $z = 0$ with $\log(M_*/M_\odot) < 9$, contrary to high-mass trends. This turnover pattern was also reported by F. Belfiore et al. (2017), H. Poetrodjojo et al. (2021), and A. H. Khoram & F. Belfiore (2025), although the exact turnover point remains uncertain, in the range of $\log(M_*/M_\odot) \sim 9 - 10$. Here we assume the turnover point at $\log(M_*/M_\odot) = 9$ and divide the $z = 0$ sample into low-mass ($\log(M_*/M_\odot) < 9$) and high-mass ($\log(M_*/M_\odot) \geqslant 9$) bins, which are fitted separately. For higher-redshift observations, no compelling evidence of such a turnover pattern has been reported so far. Therefore, we perform a linear fit across the entire mass range for the other three high-redshift bins.

In Figure 5, we show the fitting results from the four redshift bins. The fitted parameters are shown in Table 2. In the local Universe, massive galaxies ($\log(M_*/M_\odot) > 9$) show a positive MZGR, indicating secular evolution. In contrast, for the finding of T. Li et al. (2025), we see that low-mass galaxies ($\log(M_*/M_\odot) < 9$) exhibit a negative mass dependence. This suggests that in the local Universe, more-massive, evolved galaxies predominantly undergo secular evolution, whereas low-mass dwarf galaxies remain influenced by metal mixing and transport processes, such as outflows driven by ongoing star formation. At redshift $z \sim 2$, we find the slopes are consistent with the value previously reported by X. Wang et al. (2020) for a subsample in the redshift range $1.2 < z < 2.5$. However, unlike local ($z = 0$) galaxies, the $z \sim 2$ population exhibits a negative MZGR within the same mass range of $\log(M_*/M_\odot) \gtrsim 8$. This negative MZGR suggests enhanced gas mixing due to feedback-driven wind at cosmic noon, as we have discussed in Section 4.3.



(a) (b)

Figure 6. (a) A violin plot showing the mass distribution of our full sample at each redshift bin. The white circles represent the median stellar mass and redshift of each bin. The dotted lines represent 1σ intervals of the stellar mass distributions. (b) The redshift evolution of metallicity gradients with normalization by stellar masses. The dark-red lines represent the best fitting of the observed metallicity gradients, the same as those in Figure 2. The circles represent metallicity gradients normalized to a fixed stellar mass, with colors indicating stellar masses normalized relative to the observed MZGR at each redshift.

Table 2
Fitted MZGR Parameters in Equation (4) and the Spearman Coefficients (r, p)

Redshift	β (dex kpc^{-1})	α (dex kpc^{-1})	σ (dex kpc^{-1})	r, p
$z = 0$[a]	-0.100 ± 0.026	0.811 ± 0.218	0.063 ± 0.007	-0.511, 0.000
$z = 0$[b]	0.019 ± 0.002	-0.211 ± 0.019	0.013 ± 0.001	0.571, 0.000
$0.1 < z \leqslant 1.5$	-0.009 ± 0.004	0.084 ± 0.037	0.036 ± 0.002	-0.173, 0.002
$1.5 < z \leqslant 2.5$	-0.002 ± 0.003	0.032 ± 0.029	0.031 ± 0.002	0.027, 0.621
$2.5 < z \leqslant 3.5$	-0.010 ± 0.022	0.064 ± 0.196	0.121 ± 0.015	-0.060, 0.680
$5.0 \leqslant z \leqslant 9.0$[c]	0.304 ± 0.190	-2.784 ± 1.539	0.131 ± 0.125	0.949, 0.014

Notes.
[a] Fitted within the low-mass bin $\log(M_*/M_\odot) < 9$.
[b] Fitted within the high-mass bin $\log(M_*/M_\odot) \geqslant 9$.
[c] In this redshift range, the MZGR is fitted on median stacks (Figure 4).

We find a mild evolution of the MZGR slope β. The MZGR is positive at $z = 0$ for galaxies at $\log(M_*/M_\odot) > 9$, and it becomes negative at higher redshifts. The significance of the negative MZGR slope decreases with increasing redshift. At $z = 0$, the evolution of the MZGR slope coincides with the evolution of the metallicity gradient (Figure 2). This further indicates the evolution of feedback strength: With strong feedback at $z \sim 2$, the feedback promotes more positive gradients. Meanwhile, it has a stronger impact on low-mass galaxies, reflected in a more negative MZGR. At higher-redshift bins $2.5 < z \leqslant 3.5$, it starts to exhibit a less negative MZGR, closer to the prediction without efficient outflows, and finally, the MZGR is significantly positive at $5.0 \leqslant z \leqslant 9.0$.

In Figure 6(a), we show the mass distribution as a function of redshift in our sample. Since the galaxy sample is collected from a variety of works with different selection criteria, they have different stellar masses in different redshift bins. For example, the galaxies in the $0.5 < z < 2.5$ bins are more massive ($\approx$1 dex) than others, which corresponds to a -0.01 dex kpc^{-1} reduction in metallicity gradients using the MZGR derived above. To quantify the impact of different stellar masses on our observed redshift evolution, we normalize the gradients at each redshift bin to the same stellar masses based on the observed MZGR. In Figure 6(b), we compare the observed and mass-normalized metallicity gradients. We find that the overall trend of redshift evolution shown in Figure 2 remains consistent.

To investigate the intrinsic scatter for galaxies with different stellar masses, we divide the sample galaxies into low-mass ($M_* < 10^9 M_\odot$) and high-mass ($M_* > 10^9 M_\odot$) bins. At redshift $1.5 < z \leqslant 3.5$, the intrinsic scatter for low-mass galaxies is found to be $\sigma_{\rm MZGR} = 0.098 \pm 0.02$, while for massive galaxies, it is smaller, with $\sigma_{\rm MZGR} = 0.023 \pm 0.002$. In contrast, at lower redshifts ($0.01 < z \leqslant 1.5$), the scatter is reduced, with $\sigma_{\rm MZGR} = 0.045 \pm 0.008$ for low-mass galaxies and $\sigma_{\rm MZGR} = 0.037 \pm 0.002$ for massive galaxies. The increasing scattering of metallicity gradients toward the low-mass end was also found by X. Wang et al. (2020), who suggest more efficient feedback for low-mass galaxies, as shallower gravitational potentials are more susceptible to gas disruption (Y. Zhuang et al. 2019). In addition, the lowest-mass galaxies have the highest specific SFRs (sSFRs; X. Wang et al. 2020), which could be triggered by either cold-mode accretion (R. M. Woods et al. 2014) or merger-induced gas inflow (M. Martig & F. Bournaud 2008; J. P. Stott et al. 2013). The latter was found to be dominant in dwarf galaxies

(X. Wang et al. 2020). The combined effects of metal production in merger-induced starbursts and gas dilution could lead to diverse chemical structures, reflected by large scatter in metallicity gradients at $z \sim 2$. When we evenly divide the NGDEEP sample into high- and low-sSFR bins, we find that the high-sSFR subsample exhibits twice the intrinsic scatter compared to the low-sSFR subsample, reflecting the effects of those processes. The reduced scatters at $z < 1.5$ also suggest the impact of feedback on low-mass galaxies weakens at low redshifts.

4.5. Toy Models for Metallicity Gradients

To further understand how inflow and outflow take effect in our sample galaxies, we measure the gas fraction and metallicity in comparison with a simple analytical model (D. K. Erb 2008) (hereafter the Erb08 model) and the feedback-free starburst (FFB) model (A. Dekel et al. 2023, 2025; Z. Li et al. 2024). The Erb08 model assumes that the surface density of star formation is proportional to the surface density of the gas (Kennicutt–Schmidt (KS) law; M. Schmidt 1959; R. C. Kennicutt 1989) and that inflows and outflows regulate the gas required by sustained star formation. This model has established basic guidelines for the relation between mass, metallicity, and gas fraction at $z \lesssim 3$. Such a "bathtub" model is widely used to compare with observations (G. Cresci et al. 2010; K. Yabe et al. 2015; X. Wang et al. 2019; Y. Zhuang et al. 2019), although we acknowledge that it assumes no radial gas flows and constant star formation efficiency. We consider two simple cases for the Erb08 model, one with pure inflow (with inflow rate defined as inflow gas mass per star formation $f_i = \dot{M}_{\rm in}/{\rm SFR}$) and one with pure outflow (with outflow rate defined as outflow gas mass per star formation $f_o = \dot{M}_{\rm out}/{\rm SFR}$). The FFB scenario is proposed to characterize galaxies at $z \sim 10$ where the density of gas in star-forming clumps is above a threshold of $\sim 3 \times 10^3\,{\rm cm}^{-3}$ with gas metallicity $\lesssim 0.2 Z_\odot$, where $Z_\odot$ is the solar metallicity. The FFB model predicts efficient star formation in thousands of globular-cluster-like clouds occurring on a freefall timescale before the onset of effective stellar and SN feedback. This allows multiple generations of starbursts, each with a high star formation efficiency, separated by relatively quiescent periods. After each burst, SN feedback from massive stars will inevitably occur, yielding metals and enriching the galaxy. In Appendix J, we show that our $5 < z < 9$ galaxy sample satisfies the FFB criteria and is therefore suitable for applying this model.

We note that FFB is a simple analytical model aiming to illustrate the idea of efficient star formation with inefficient metal mixing. We refer the reader to more sophisticated models (e.g., J. Fu et al. 2013; M. Mollá et al. 2019; B. M. B. Henriques et al. 2020; P. Sharda et al. 2021a, 2021b, 2021c, 2024), which consider more physical processes such as metal diffusion, advection, and more detailed feedback processes. M. A. Bellardini et al. (2021) demonstrated that neglecting diffusion leads to pronounced gradient steepening, whereas larger diffusion coefficients promote gradient flattening. Our simplified model thus illustrates the conditions under which metal mixing is inefficient in the absence of mechanisms such as diffusion.

Here, we consider a case of the steady-state solution for gas metallicity involving both high inflow rates and high SFRs, along with the resultant outflows in the FFB scheme (A. Dekel et al. 2023; Z. Li et al. 2024). In reality, the stars of FFB galaxies are expected to form in multiple generations of bursts rather than a continuous process; therefore, this model should be considered a time-smoothed approximation. According to the conservation of gas mass among the ISM, gas inflow, SFR, and outflow ($M_{\rm ism}$, $M_{\rm in}$, SFR, and $M_{\rm out}$), and the conservation of the corresponding metal mass, we have

$$\dot{M}_{\rm ism} + \dot{M}_{\rm out} = \dot{M}_{\rm in} - (1 - f_{\rm sn}){\rm SFR} \tag{5}$$

$$Z_{\rm ism}\dot{M}_{\rm ism} + Z_{\rm out}\dot{M}_{\rm out} = Z_{\rm in}\dot{M}_{\rm in} - (Z_{\rm sf} - Z_{\rm sn} f_{\rm sn}){\rm SFR} \tag{6}$$

where $f_{\rm sn}$, $Z_{\rm sn}$, and $Z_{\rm sf}$ are the mass fraction and the metallicity of SN ejecta and the metallicity of gas clouds that form stars, respectively. The second equation assumes a steady-state solution without metallicity evolution, i.e., $\dot{Z}_{\rm ism} = \dot{Z}_{\rm out} = \ldots = 0$. Below, we estimate the metallicity and spherically averaged density profile, $\rho(r)$, of each component.

For a halo with an accretion rate of $\dot{M}_{\rm halo}$, the gas accretion rate and SFR are $\dot{M}_{\rm in} = f_b \dot{M}_{\rm halo}$ and ${\rm SFR} = \epsilon \dot{M}_{\rm in}$, respectively. In this equation, f_b is the cosmic baryonic fraction and ϵ is the star formation efficiency. The outflow rate is regulated by the mass loading factor, $\eta \equiv \dot{M}_{\rm out}/{\rm SFR}$. We introduce a parameter $f_{\rm out} = \eta/\eta_{\rm max} \leqslant 1$, where $\eta_{\rm max}$ is the maximum mass loading factor allowed for nonnegative $\dot{M}_{\rm ism}$, $\eta_{\rm max} = \epsilon^{-1} - 1 + f_{\rm sn}$. We then have $\dot{M}_{\rm out} = f_{\rm out}\,\eta_{\rm max}\,{\rm SFR}$ and $\dot{M}_{\rm ism} = (1 - f_{\rm out})\,\eta_{\rm max}\,{\rm SFR}$. We obtain the stellar mass and gas mass in the ISM by integrating over time: $M_{\rm star} = (1 - f_{\rm sn})\,\epsilon f_b M_{\rm halo}$ and $M_{\rm ism} = (1 - f_{\rm out})\,\eta_{\rm max}\,\epsilon f_b M_{\rm halo}$. Given the observed galaxy size R_e with a typical Sérsic index of $n_s = 1$, we estimate the density profile of stellar mass and ISM gas, $\rho_{\rm star}(r)$ and $\rho_{\rm ism}(r)$.

The model considers star formation occurring in the shock region with enhanced gas density at $R_{\rm shell} \simeq 2R_e$, where the inflow encounters the outflow (A. Dekel et al. 2023; Z. Li et al. 2024). We estimate the spherically averaged density of the inflow at radius r using the mass flux $\dot{M}'_{\rm in}(r)$ and inflow velocity $v_{\rm in}(r) \sim v_{\rm vir}$ (H. Aung et al. 2024): $\rho_{\rm in}(r) = \dot{M}'_{\rm in}(r)/4\pi r^2 v_{\rm in}(r)$. For $r > R_{\rm shell}$, we assume a constant inflow flux: $\dot{M}'_{\rm in}(r) = \dot{M}_{\rm in}$. Within $R_{\rm shell}$, the flux decreases due to star formation and conversion to ISM, approximated by $\dot{M}'_{\rm in}(r) = (\dot{M}_{\rm in} - {\rm SFR})(r^3/R^3_{\rm shell})$. After their birth, stars redistribute within the galaxy and launch outflows. We adopt the asymptotic solution for outflow density from R. A. Chevalier & A. W. Clegg (1985),

$$\rho_{\rm out}(r) = \begin{cases} 0.3\left(\dfrac{\dot{M}_{\rm out}}{\dot{E}_{\rm sn}}\right)^{1/2} \dfrac{\dot{M}_{\rm out}}{R^2_{\rm shell}} & {\rm for}\, r < R_{\rm shell}, \\ 0.05\left(\dfrac{\dot{M}_{\rm out}}{\dot{E}_{\rm sn}}\right)^{1/2} \dfrac{\dot{M}_{\rm out}}{r^2} & {\rm for}\, r > R_{\rm shell}, \end{cases} \tag{7}$$

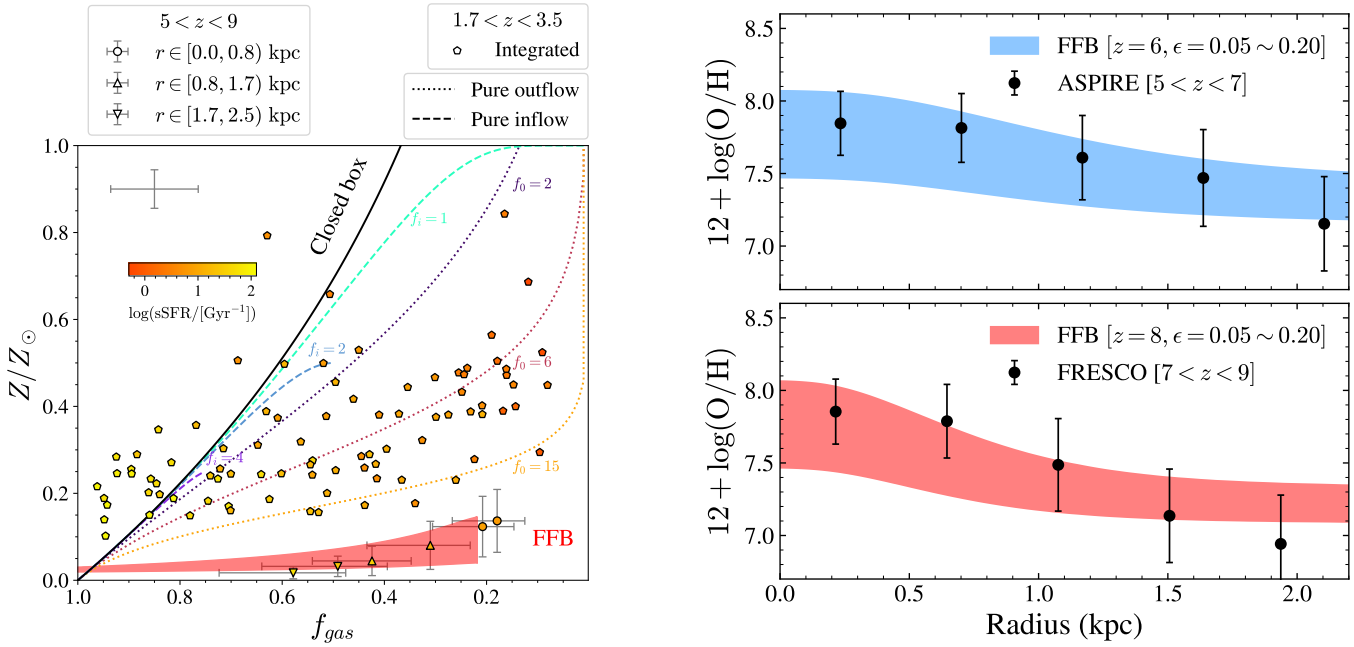
where $\dot{E}_{\rm sn}$ is the SN energy returned by a given SFR (Z. Li et al. 2024).

Assuming star formation from inflow gas without mixing (Z. Li et al. 2024), $Z_{\rm sf} = Z_{\rm in}$, and an outflow being launched from the ISM, $Z_{\rm ism} = Z_{\rm out}$, we get

$$Z_{\rm ism} = \frac{Z_{\rm in}(1 - \epsilon) + Z_{\rm sn} f_{\rm sn}\,\epsilon}{(1 - \epsilon) + f_{\rm sn}\,\epsilon}. \tag{8}$$

We then obtain the spherically averaged gas-phase metallicity profile as

$$Z(r) = \frac{Z_{\rm in}\rho_{\rm in}(r) + Z_{\rm ism}[\rho_{\rm out}(r) + \rho_{\rm ism}(r)]}{\rho_{\rm in}(r) + \rho_{\rm out}(r) + \rho_{\rm ism}(r)}, \tag{9}$$



(a) (b)

Figure 7. (a) The measured gas fraction and metallicity compared with the Erb08 model (D. K. Erb 2008) and the FFB model (A. Dekel et al. 2023; Z. Li et al. 2024). We consider Erb08 models with pure gas outflow ($f_i = 0$) and with pure inflow ($f_o = 0$). The curves for the Erb08 models are color coded by inflow/outflow rate relative to SFR, with f_i denoting the inflow model (dashed curves) and f_o denoting the outflow model (dotted curves). The solid black curve represents the closed-box model without outflow and inflow. The red shaded region shows the predictions of the FFB model (Z. Li et al. 2024) at redshift $z = 8$ with a star formation efficiency of $\epsilon = 0.05$–0.20 and a mass loading factor of 0.95, the maximum allowed value. The points in ASPIRE ($5 < z < 7$) and FRESCO ($7 < z < 9$) are estimated within different radial annuli (see Figure 14), while the pentagon points represent integrated measurements for the NGDEEP sample at $z = 1.7$–3.5. The symbols are color coded by sSFR. The gas fraction uncertainties of the ASPIRE and FRESCO galaxies include the uncertainty from unconstrained dust attenuation in estimating SFR maps. (b) The measured metallicity gradients at redshift $z = 5$–7 and $z = 7$–9 compared with the FFB model. The shaded regions represent the metallicity profiles predicted by the FFB model at redshifts 6 (blue) and 8 (red) with star formation efficiency $\epsilon = 0.05$–0.20, which is in line with our observations.

and compute the projected metallicity and gas fraction profiles ($\Sigma_{\rm gas}(r)/\Sigma_{\rm star+gas}(r)$, where Σ represents the surface density) accordingly (Figure 7). To mimic the observation, we compute the projected profiles for a galaxy sample with varying R_e. Our sample can be approximated by a lognormal distribution with a scatter of $\sigma_{\ln R_e} = 0.5$ and a median value of $R_e = 0.53$ kpc for $z = 5$–7 and 0.36 kpc for $z = 7$–9.

We adopt $\epsilon = 0.05$–0.2, $Z_{\rm in} = 0.01 Z_\odot$, $Z_{\rm sn} = 3 Z_\odot$, $f_{\rm sn}$=0.2, and $f_{\rm out} = 0.95$ (corresponding to $\eta = 18$–4), which reproduce the observation well. Equation (8) gives $Z_{\rm ism} = 0.04 Z_\odot$ for $\epsilon = 0.05$ and $0.15 Z_\odot$ for $\epsilon = 0.2$. The gas-phase metallicity is dominated by the enriched ISM in the central region and by the metal-poor inflows in the outskirts.

Different from Erb08, the FFB model introduces radially dependent inflows and outflows. Assuming steady winds driven by SNe in the FFB scheme, we consider a toy model for a steady solution of gas metallicity involving high inflow rates, SFRs, and resultant outflows as a time-average approximation. At high redshift, galaxies are fed by cold gas streams that penetrate deep into the halos without significant mixing until they join the central galaxy (A. Dekel et al. 2009). These metal-poor cold inflows dominate the gas mass outside the galaxy, leading to low metallicity at large radii. These radially dependent gas flows are similar to the accretion disk model proposed by E. Wang & S. J. Lilly (2022), and are consistent with an inside-out growth scenario.

We also estimate the gas fraction $f_{\rm gas}$ in our observation samples (see Appendix C) and compare it with predictions from different models. As shown in Figure 7(a), galaxies at cosmic noon can reasonably be predicted by the Erb08 models with a combination of inflow and outflow. We note that a few outliers in the NGDEEP sample with high sSFR exhibit metallicities higher than those predicted by the closed-box model. Since the model does not consider gas recycling, those points may be attributed to the reaccretion of metal-enriched gas, which increases the metallicity. D. Anglés-Alcázar et al. (2017) showed that in lower-mass galaxies, a larger fraction of the gas supply originates from wind recycling, with these systems reaccreting more gas relative to their stellar mass. As a result, the metallicity at high gas fractions can be higher than predicted by the closed-box model. This indicates efficient gas recycling in low-mass galaxies driven by high sSFR. The reaccretion of metal-enriched gas onto the galactic disk further contributes to flat/positive metallicity gradients (R. Schönrich & P. J. McMillan 2017).

To reproduce galaxies with both sufficiently low metallicities and low gas fractions, the Erb08 model requires a high outflow rate. An alternative is provided by the FFB model, which predicts lower metallicities ($Z < 0.1 Z_\odot$) at low gas fractions, as expected from the high star formation efficiency and the dense cold inflows with inefficient mixing. From Figure 7(b), our observed metallicity gradients in the ASPIRE and FRESCO samples are also in agreement with the predictions of the FFB models at redshifts of $z = 6$ and $z = 8$, respectively. The inefficient gas mixing mechanism in the FFB scenario permits the existence of steep gradients at sufficiently high redshift. This simple analytical model demonstrates that steep gradients can be maintained even in

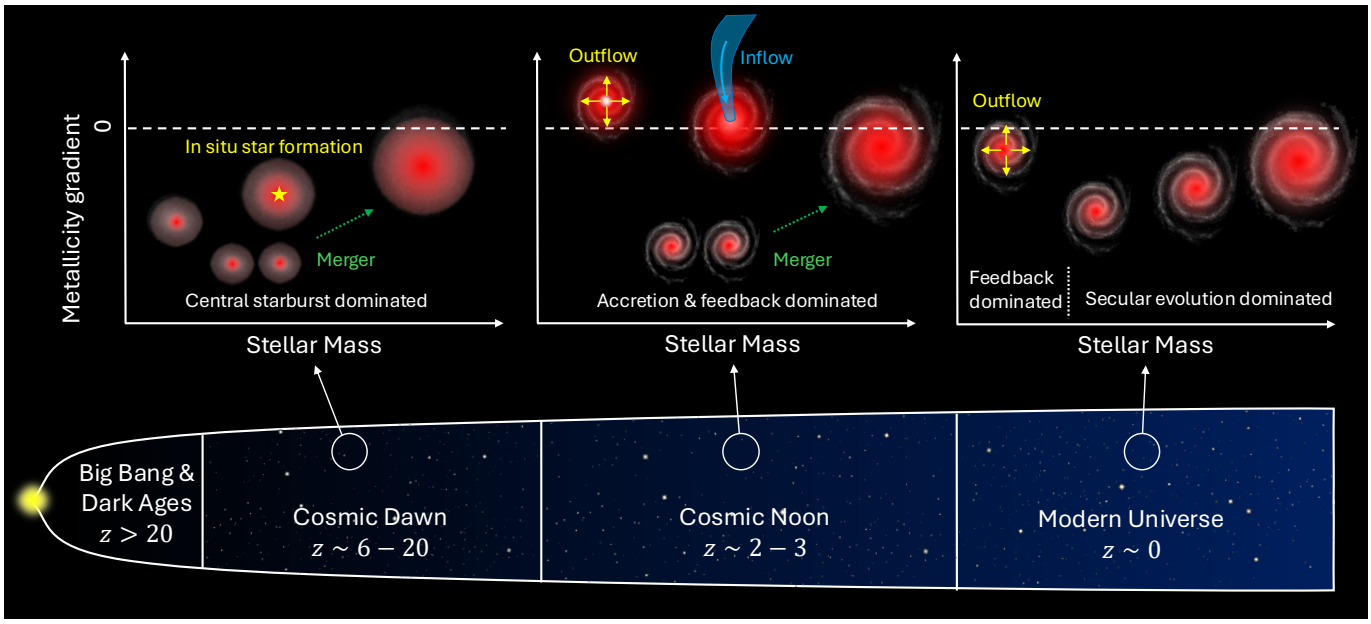


Figure 8. A sketch of how metallicity gradients are thought to evolve as galaxies grow across different cosmic epochs, based on our current knowledge. At cosmic dawn, the galaxies are observed with negative metallicity gradients, and they flatten with increasing stellar masses. This suggests the metal mixing and transport process may be inefficient, and coincides with the inside-out growth scenario with in situ metal production. At cosmic noon, the galaxies are observed with nearly flat metallicity gradients, and they have an anticorrelation with stellar masses. It suggests that metal mixing and transport mechanisms, including gas outflows, inflows, and radial transport, are efficient in redistributing metals in the ISM, which are sensitive to gravitational potentials. Mergers are also expected in the dense high-redshift Universe, which also contribute to gas mixing and flat gradients. In the local Universe, the MZGRs are observed to turn over at certain stellar masses, where higher-mass galaxies show positive MZGRs while lower-mass galaxies show the opposite trend. It suggests feedback-driven outflows may still be efficient for local dwarf galaxies, while in situ metal production is dominating high-mass galaxies.

the presence of inflows and outflows, provided that central in situ star formation dominates the galaxy formation and metal enrichment. In addition, we have estimated the sSFR maps of our $5 < z < 9$ samples based on SFRD and stellar mass density maps (Figure 14 in Appendix C). The sSFR profiles can be visualized in Figure 7(a). We find that the sSFR is lower in the central bins ($r \in [0.0, 0.8]$ kpc) and increases toward the outer regions ($r \in [1.7, 2.5]$ kpc). This provides additional evidence for inside-out growth, echoing the recently observed core–disk structure at $z = 7.4$ (W. M. Baker et al. 2025), where the sSFR was found to increase with radius. The central sSFR suppression likely traces early-forming, dense stellar cores, while an elevated outer disk sSFR indicates ongoing gas accretion and star formation in extended disks. These results across our $5 < z < 9$ sample suggest that inside-out growth may be a dominant pattern for galaxy formation at the epoch of reionization. Consistent with this picture, J. Matharu et al. (2024) also reported evidence for inside-out growth by comparing the Hα and stellar continuum profiles at $4.8 < z < 6.5$.

5. Summary

We summarize the results from the synthesis of our new JWST observations and literature observations in Figure 8. We find different phases in the redshift evolution of metallicity gradients, with a growing phase from sufficiently high redshifts $z > 5$, which flattens at $z \approx 2$ and transitions to a descending phase until the present day. The implications of these three phases are as follows:

1. The ascending phase may indicate rapid growth in inside-out mode, supported by continuous replenishment from cold gas accretion. This stage reflects the early buildup of galaxies in a Universe dominated by cold gas. The steep gradients align with inside-out growth and limited radial mixing. To explain this, we apply a toy model within the FFB scenario. In this framework, dense inflowing gas sustains efficient, localized star formation while the redistribution of metals is inefficient (A. Dekel et al. 2023; Z. Li et al. 2024). Such conditions preserve negative metallicity gradients, reflecting the early stages of galaxy assembly where internal chemical enrichment occurs earlier in the inner regions than in the outer disk. The observed positive correlation between metallicity gradient and stellar mass further supports the notion that massive galaxies grow more efficiently in an inside-out fashion during this period, linking their internal structure directly to their rapid mass accumulation history.
2. The metallicity gradients flatten over time toward $z \approx 2$, suggesting enhanced radial gas flows and mixing mechanisms. This epoch corresponds to the peak of the SFRD, and is driven by a combination of intense cold gas accretion, elevated SFR, and strong stellar/SN feedback (A. Dekel et al. 2009; T. M. Heckman & P. N. Best 2023; R. L. Sanders et al. 2023). The near-flat gradients observed at this time, along with large galaxy-to-galaxy scatter, reflect the turbulent internal dynamics triggered by gas accretion and feedback, which redistributes metals and flattens chemical abundances. The negative MZGR at this epoch signifies the dominant role of feedback in flattening gradients across stellar masses, especially in lower-mass systems (X. Ma et al. 2017).
3. The gradual decline in metallicity gradients observed at $z < 2$ coincides with the decrease in cold-mode accretion and a shift toward hot-mode accretion (F. van de Voort et al. 2011). This reduces the efficiency of star formation

and limits pristine gas inflows, leading to slower chemical evolution. The reduced star formation activity and feedback at later times allow the galactic disks to stabilize and steepen the metallicity gradients. This is reflected by the declining global SFRD and indicates the transition to a more gradual mode of galaxy growth driven by secular evolution. The positive MZGRs for massive galaxies also provide evidence. Although low-mass galaxies still show a negative MZGR (T. Li et al. 2025), feedback effects mainly impact young, low-mass systems in the local Universe, whereas high-mass galaxies tend to be more evolved and dynamically stable with steeper metallicity gradients.

These findings highlight the intricate connection between galaxy internal star formation and cosmic-scale galaxy formation across cosmic time. The synchronized evolution pattern between metallicity gradients and cosmic star formation density reveals that gas accretion, star formation, and feedback not only shape the evolutionary pathways of individual galaxies but also collectively drive the large-scale formation of the galaxy population in the Universe. Nevertheless, how these various physical processes interact to drive the observed evolution remains inconclusive and calls for further extensive investigation with both observations and simulations.

Acknowledgments

The authors thank the editors and anonymous referees for their careful reading and constructive comments, which greatly helped improve the manuscript. This work is based on observations made with the NASA/ESA/CSA James Webb Space Telescope. The JWST data presented in this article were obtained from the Mikulski Archive for Space Telescopes (MAST[27]) at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-03127 for JWST. The specific observations analyzed can be accessed via GO-2078 (F. Wang 2023), GO-1895 (P. Oesch & D. Magee 2023), and GO-2079 (N. Pirzkal 2023). The data has been processed using public software GRIZLI (G. Brammer 2023) and `Calwebb` (H. Bushouse et al. 2024). The photometric catalog can be downloaded from the DJA. The authors sincerely thank the FRESCO team (PI: Pascal Oesch) and the NGDEEP team (PI: Steven Finkelstein, Casey Papovich, and Nor Pirzkal) for developing their observing programs with a zero-exclusive-access period for the exquisite data. This work made use of the high-performance computing resources at Tsinghua University. This work uses images and photometry catalogs retrieved from the DJA. The DJA is an initiative of the Cosmic Dawn Center, which is funded by the Danish National Research Foundation under grant No. 140.

We thank Steven Finkelstein and Raymond Simons for helpful discussions regarding NGDEEP data and Pascal Oesch and Romain Meyer for helpful discussions regarding FRESCO data. Z.L. thanks Dandan Xu and Cheng Li for insightful discussions about the observational and theoretical context; Xiangcheng Ma, Guochao Sun, and Xunda Sun for valuable input about the FIRE simulations; Ayan Acharyya for helpful discussions on FOGGIE simulations and their theoretical interpretation; Alex Garcia for sharing results from EAGLE, Illustris, TNG, and SIMBA simulations; Piyush Sharda for helpful discussions with metallicity gradient models; Prerak Grag for help with the photoionization model; Zechang Sun for maintaining the high-performance computers at Tsinghua University; Mengting Ju for sharing results from JWST MSA-3D; and Zefeng Li and William Baker for helpful comments.

Z.L., Z.C., X.L., and Y.W. are supported by the National Science Foundation of China (grant No. 12073014), science research grants from the China Manned Space Project No. CMS-CSST-2021-A05, and the Tsinghua University Initiative Scientific Research Program (No. 20223080023). Z.L. also acknowledges that this work was conducted in part at the Cosmic Dawn Center with financial support from the center. X.W. is supported by the China Manned Space Program with grant No. CMS-CSST-2025-A06, the National Natural Science Foundation of China (grant 12373009), the CAS Project for Young Scientists in Basic Research grant No. YSBR-062, the Fundamental Research Funds for the Central Universities, and the Xiaomi Young Talents Program. X.W. also acknowledges work carried out, in part, at the Swinburne University of Technology, sponsored by the ACAMAR visiting fellowship. Z.L. acknowledges the European Union's Horizon 2020 program under the Marie Skłodowska-Curie Grant No. 101109759 ("CuspCore") and the Israel Science Foundation grant ISF 861/20. S.E.I.B. is supported by the Deutsche Forschungsgemeinschaft under Emmy Noether Grant No. BO 5771/1-1. F.W. acknowledges support from NSF grant AST-2308258. M.T. acknowledges support from the NWO grant 016.VIDI.189.162 ("ODIN"). S.Z. acknowledges support from the National Natural Science Foundation of China (No. 12303011).

Facility: JWST (NIRCam, NIRISS).

Software: GRIZLI (G. Brammer 2023), `Calwebb` (H. Bushouse et al. 2024), LINMIX (B. C. Kelly 2007), Astropy (Astropy Collaboration et al. 2022), SciPy (P. Virtanen et al. 2020), *emcee* (D. Foreman-Mackey et al. 2013a), `BEAGLE` (J. Chevallard & S. Charlot 2016), `SourceXtractor++` (E. Bertin et al. 2020), Photutils (L. Bradley et al. 2024).

Appendix A
Measuring Metallicity Gradients in Mass Bins

Galaxies grow into different sizes and masses, which are correlated through the mass–size relation (T. Shibuya et al. 2015; D. Langeroodi & J. Hjorth 2023). Metallicity gradients may be different for galaxies of different sizes. To investigate the impact of galaxy masses on the metallicity gradient, we divide the stacking into different mass bins at each redshift. For the ASPIRE sample, we utilize three mass bins: the most-massive bin includes galaxies with $\log(M_*/M_\odot) > 9.00$, while the remaining lower-mass galaxies ($\log(M_*/M_\odot) < 9.00$) are evenly split into two bins: a low-mass bin ($7.18 < \log(M_*/M_\odot) < 8.09$) and a high-mass bin ($8.09 < \log(M_*/M_\odot) < 9.00$). The FRESCO sample is less massive, and we evenly divide it into a low-mass bin ($7.30 < \log(M_*/M_\odot) < 7.95$) and a high-mass bin ($7.95 < \log(M_*/M_\odot) < 8.70$). For comparison between individual and stacked measurements, we also separate the NGDEEP sample into a low-mass bin ($6.87 < \log(M_*/M_\odot) < 8.44$) and a high-mass bin ($8.44 < \log(M_*/M_\odot) < 9.81$). Within each mass bin, the stellar mass range spans approximately ∼1 dex, corresponding to a change of $R_e \sim 0.1$–0.2 kpc as inferred from the mass–size relation (D. Langeroodi & J. Hjorth 2023). We apply the same median stacking method as used for galaxies in redshift bins. We apply the

[27] https://archive.stsci.edu/

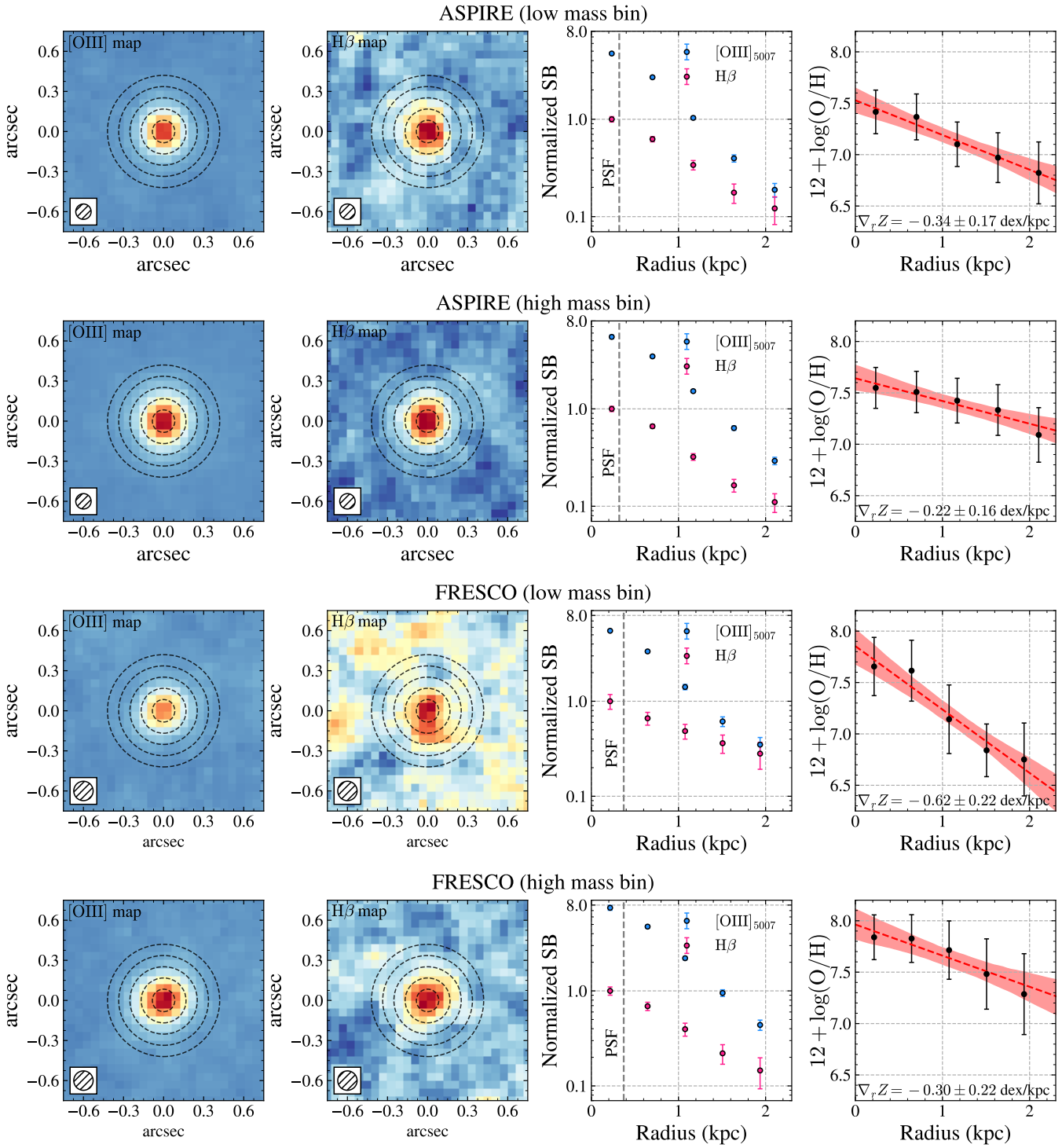


Figure 9. Stacking results in different mass bins. First two rows: results from ASPIRE galaxies in low-mass bin ($7.18 < \log(M_*/M_\odot) < 8.09$) and high-mass bin ($8.09 < \log(M_*/M_\odot) < 9.00$). Last two rows: results from FRESCO galaxies in low-mass bin ($7.30 < \log(M_*/M_\odot) < 7.95$) and high-mass bin ($7.95 < \log(M_*/M_\odot) < 8.70$).

same R3 method (R. L. Sanders et al. 2024) to measure metallicity gradients for high-mass bins and low-mass bins in the ASPIRE and FRESCO samples, shown in Figure 9. However, since the bin with $\log(M_*/M_\odot) > 9.00$ is massive enough that the MZR predicts these galaxies have higher metallicity, which is out of the lower-branch range of the R3 relation, we use the upper branch instead to measure their metallicity gradient, shown in Figure 10. For the NGDEEP sample, we use the same method with both R3 and R2 calibrations (F. Bian et al. 2018) to jointly constrain metallicity as for individual galaxies, shown in Figure 11. The stacks in mass bins provide a more representative characterization of the galaxy population with similar masses and sizes. From the stacked map of the NGDEEP sample, we measure the MZGR slope of $-0.019\,\mathrm{dex\,kpc^{-1}}$ using the two stacked points at low and high masses, which is consistent with the slope of

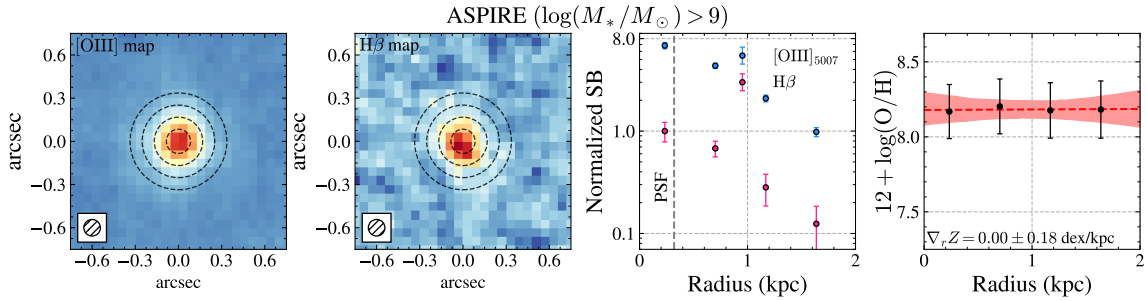


Figure 10. The same as Figure 9, but for the galaxies with the highest stellar masses ($\log(M_*/M_\odot) > 9.00$) in the ASPIRE sample. The metallicities are estimated using the upper-branch solution of the R3 relation in R. L. Sanders et al. (2024).

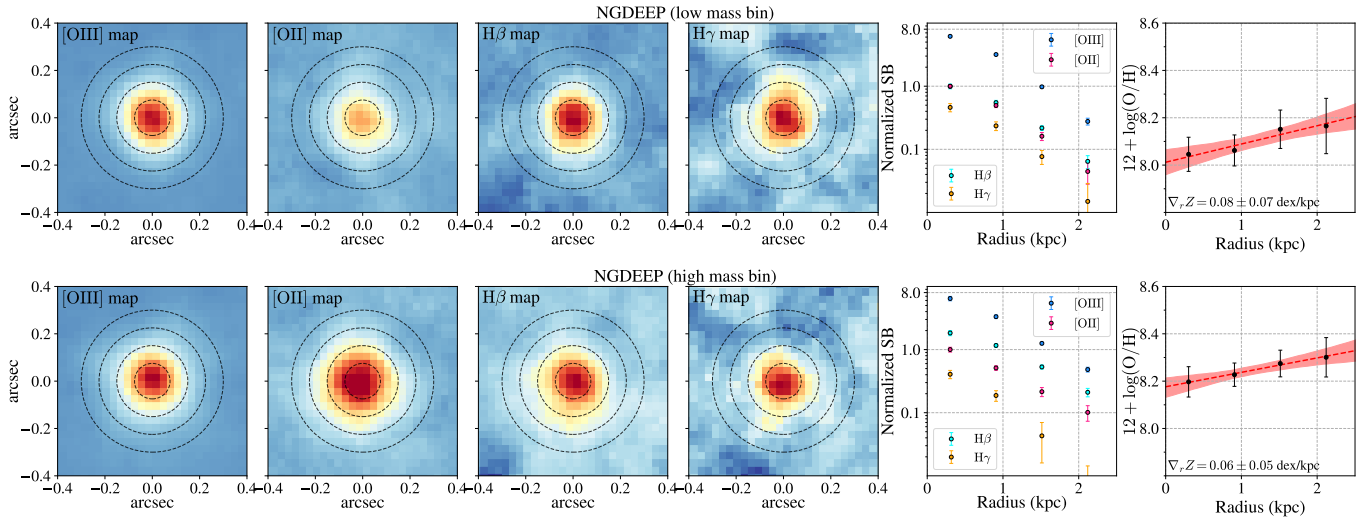


Figure 11. Stacking results from NGDEEP galaxies in low-mass bin ($6.87 < \log(M_*/M_\odot) < 8.44$) and high-mass bin ($8.44 < \log(M_*/M_\odot) < 9.81$).

$-0.020\,\mathrm{dex\,kpc^{-1}}$ obtained from a linear regression on individual measurements (Section 4.3).

Appendix B
Metallicity Gradients in Units of Effective Radius

As galaxies intrinsically have smaller sizes at higher redshift (T. Morishita et al. 2024), we expect steeper gradients for small galaxies if the changes in metallicity are the same. Galaxy size is also correlated with galaxy stellar mass, with massive galaxies having larger effective radii. Thus, the metallicity gradient measured with respect to the effective radius R_e should provide a quantity that is less dependent on stellar mass and size.

We use a Sérsic model convolved with a PSF to fit the galaxy morphology with GALFIT. With fitted R_e, we resample the emission maps to the pixel scale of $0.25\,R_e$ with flux conservation. Then we stack the galaxy emission maps using the median stacking method, before measuring the gradients on the stacks. In Figure 12, we show the stacked results from the ASPIRE and FRESCO samples. We still observe steep negative gradients in units of dex R_e^{-1}. This test eliminates possible bias from the variation of galaxy sizes.

The effective radii for the NGDEEP sample are measured on the F115W image. The metallicity gradients rescaled by R_e are shown in Figure 13. We also include literature observations at high redshift (D. Carton et al. 2018; R. C. Simons et al. 2021), and local observations from MaNGA (A. Franchetto et al. 2021), which provide metallicity gradients in the same units for comparison. The ascending phase at $1.7 < z < 3.5$ remains in the NGDEEP galaxies. The gradients at $z > 5$ are still steeper than galaxy populations at lower redshift despite their intrinsically smaller sizes. The redshift evolution of metallicity gradients measured with R_e is consistent with that measured with kiloparsecs.

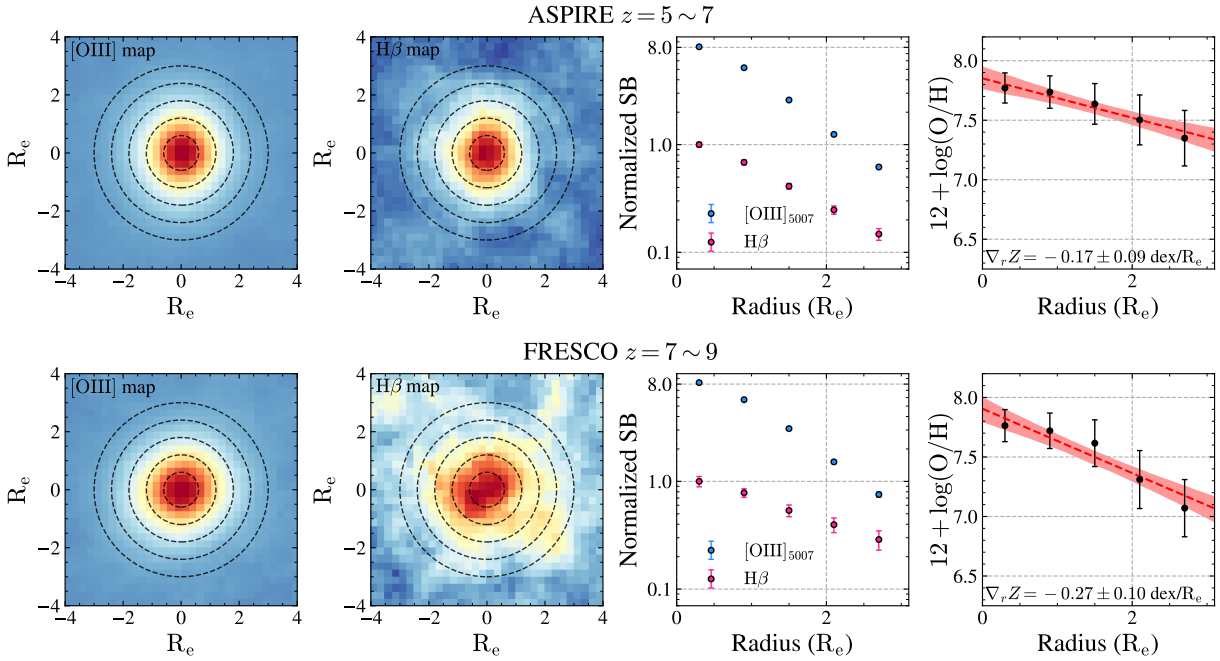


Figure 12. Stacking results in scale of effective radius (R_e).

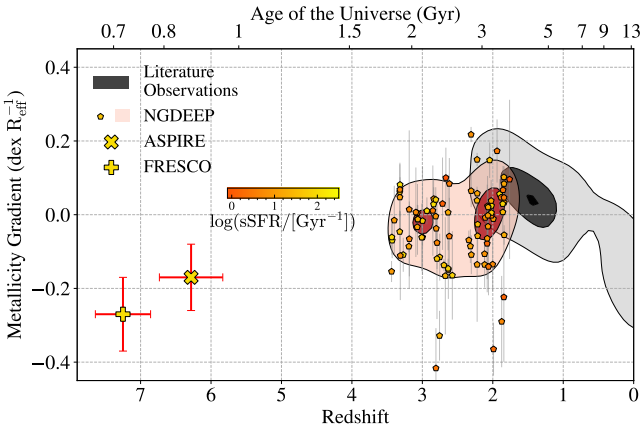


Figure 13. Metallicity gradients measured with respect to R_e. The symbols are the same as in Figure 2. The gray shadow encompasses literature observations (D. Carton et al. 2018; A. Franchetto et al. 2021; R. C. Simons et al. 2021).

Appendix C
Gas Fraction Estimation

The gas fraction is useful information when comparing our observations with analytical models as shown in Figure 4. In this section, we present our gas fraction estimation as follows. We estimate gas density by inverting the KS law (M. Schmidt 1959; R. C. Kennicutt 1989):

$$\Sigma_{\rm SFR} = 10^{-12} \kappa_s \Sigma_{\rm gas}^{1.4}, \quad \text{(C1)}$$

where $\Sigma_{\rm SFR}$ is in $M_\odot\ {\rm yr}^{-1}\ {\rm kpc}^{-2}$, and $\Sigma_{\rm gas}$ is in $M_\odot\ {\rm kpc}^{-2}$. κ_s is the burstiness parameter representing the deviation from the original KS relation (A. Ferrara et al. 2019). There is evidence that the KS relation is not universally applicable and may undergo an upward shift at higher redshifts (V. Markov et al. 2022; A. Pallottini et al. 2022). For example, E. Daddi et al. (2010) found that the normalization of $\Sigma_{\rm SFR}$ for highly star-forming galaxies is shifted by ∼0.9 dex from the conventional KS relation. Thus, highly star-forming galaxies are expected to have $\kappa_s > 1$. Thus, we adopt $\kappa_s = 10$ for the galaxy sample at $z \approx$ 1–3, and a higher value of $\kappa_s = 20$ for galaxies at $z \approx$ 6–7, due to more bursty star formation suggested by recent measurements (L. Vallini et al. 2021, 2024; V. Markov et al. 2022). Since κ_s for individual galaxies is vastly unconstrained, we add 50% uncertainty to the values we adopt to have a conservative estimation of the gas fraction.

We adopt the Balmer ratio Hα/Hβ = 2.86 in Case B recombination, where we estimate the SFR by using the Hβ luminosity and the following calibration (R. C. Kennicutt 1998):

$$\text{SFR} = 13.16 \times 10^{-42} \frac{L_{\rm H\beta}}{\rm erg}\ {\rm s}^{-1}\ (M_\odot\ {\rm yr}^{-1}). \quad \text{(C2)}$$

Since we cannot derive spatially resolved dust attenuation without other Balmer series, we cannot directly correct the Hβ flux for dust. Recent studies have found galaxies within a similar mass range at $z > 5$ to have a certain amount of dust, either through SED modeling (J. Matthee et al. 2023; S. Tacchella et al. 2023), or through the Balmer decrement (L. Sandles et al. 2024). The typical dust is expected to be $E(B-V) \sim 0.1$ corresponding to the uncertainty of 35% for the SFR derived from the uncorrected flux in Hβ. Since this uncertainty is not negligible, we add one-sided 35% uncertainties to the SFR maps. Meanwhile, for the NGDEEP sample, the Hβ fluxes have been corrected by dust using the Balmer decrement between Hα, Hβ, and Hγ (Equation (1)).

$\Sigma_{\rm star}$ is obtained by spatially resolved SED fitting to each pixel. We match the F115W and F200W bands with the PSF to the F356W band in the ASPIRE sample, and we match all photometric bands with the PSF to F444W in the FRESCO sample. After the PSF matching, both the images and emission maps have the same PSF. We use a median stack for all

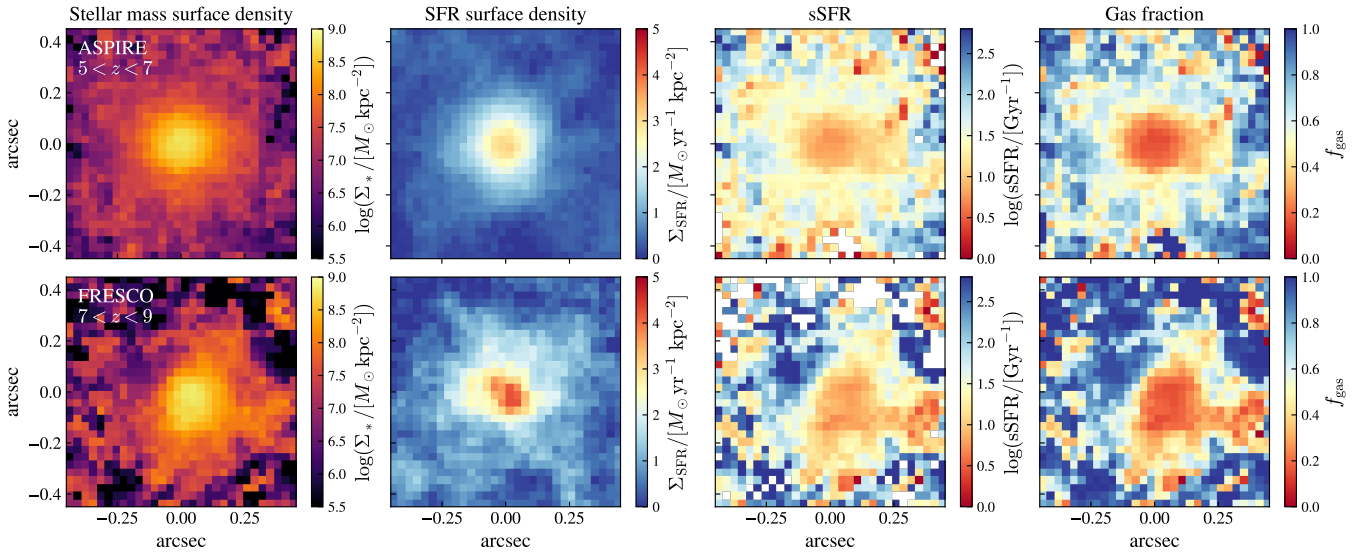


Figure 14. From left to right: stellar mass surface density, SFR surface density, sSFR, and gas fraction of ASPIRE (upper) and FRESCO (lower) samples.

cutouts in units of luminosity and convert back to flux density using the median redshifts $z = 6.28$ and $z = 7.24$ for ASPIRE and FRESCO, respectively. We then apply the SED fitting to each pixel, assuming the fixed redshifts $z = 6.28$ and $z = 7.24$ for the ASPIRE and FRESCO samples, respectively.

The gas surface density $\Sigma_{\rm gas}$ can be derived by the inverse KS law in Equation (C1). The gas fraction is then expressed as

$$f_{\rm gas} = \Sigma_{\rm gas}/(\Sigma_{\rm gas} + \Sigma_{\rm star}). \quad \text{(C3)}$$

The stellar mass surface density map, the SFR surface density map, and the derived gas fraction map are shown in Figure 14. The gas fraction we derive here has been used in Figure 7(a).

Appendix D
Fit Redshift Evolution of Metallicity Gradients

We perform MCMC sampling of the multidimensional parameter space with the package EMCEE (D. Foreman-Mackey et al. 2013b) to fit the double-power-law model of Equation (3). We apply a flat prior for the following parameters: $\gamma_0 \sim \mathcal{U}(-10, 10), \gamma_1 \sim \mathcal{U}(-1, 1), \gamma_2 \sim \mathcal{U}(0, 10), \gamma_3 \sim \mathcal{U}(-10, 10)$, and $\gamma_4 \sim \mathcal{U}(-10, 10)$. The sampling is performed with 32 walkers, 5000 iterations each, and with a burn-in period $n = 1000$. The posterior probability distribution for the parameters is shown in Figure 15. The median model and 1σ prediction interval are shown in Figures 2 and 3.

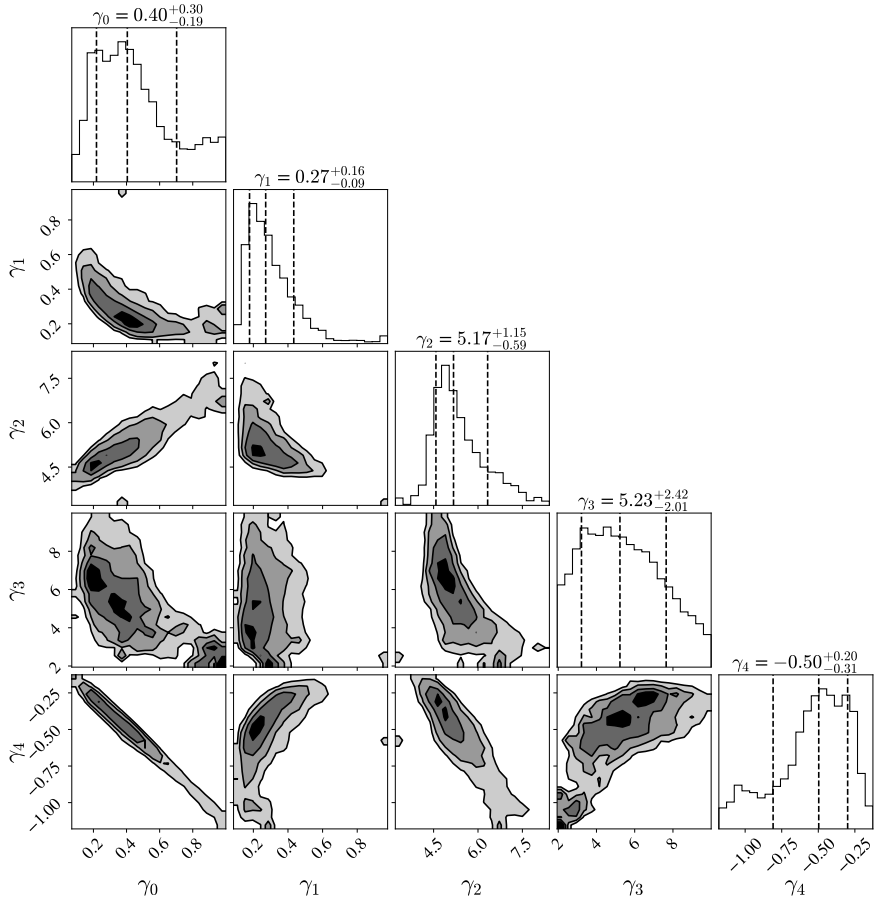


Figure 15. Posterior distribution of γ_0, γ_1, γ_2, γ_3, and γ_4 in Equation (3) from MCMC sampling. The values at the top of each column are the medians with 1σ intervals.

Appendix E
Systematics Using Different Emission-line Calibrations

To verify that our results are not significantly altered by different metallicity diagnostics, we compare our fiducial measurements using F. Bian et al. (2018) in the NGDEEP sample with other popular calibrations based on different samples and methods (R. Maiolino et al. 2008; M. Curti et al. 2020b; K. Nakajima et al. 2022). The calibrations in R. Maiolino et al. (2008) combine the direct electron temperature measurements from the Sloan Digital Sky Survey (SDSS) in the low-metallicity branch (T. Nagao et al. 2006) and the predictions of a photoionization model in the high-metallicity branch (L. J. Kewley & M. A. Dopita 2002). The calibrations in M. Curti et al. (2020b) are derived from a set of individual low-metallicity galaxies together with stacks of high-metallicity galaxies in the redshift range $0.027 < z < 0.25$, where auroral lines are detected in composite spectra, where the metallicities are self-consistently measured via the T_e method for both the high- and low-metallicity range. The calibrations in K. Nakajima et al. (2022) are derived from more metal-poor galaxies with auroral lines detected and selected from the HSC-SSP and SDSS catalogs and follow-up observations (T. Kojima et al. 2020; Y. Isobe et al. 2022), in addition to the high-metallicity galaxies used in M. Curti et al. (2020b).

Following the same procedures, we derive the metallicity maps, measure the metallicity gradients for each galaxy and the stacks, and compare those with the fiducial measurements. The comparisons are shown in Figure 16. We find no significant bias

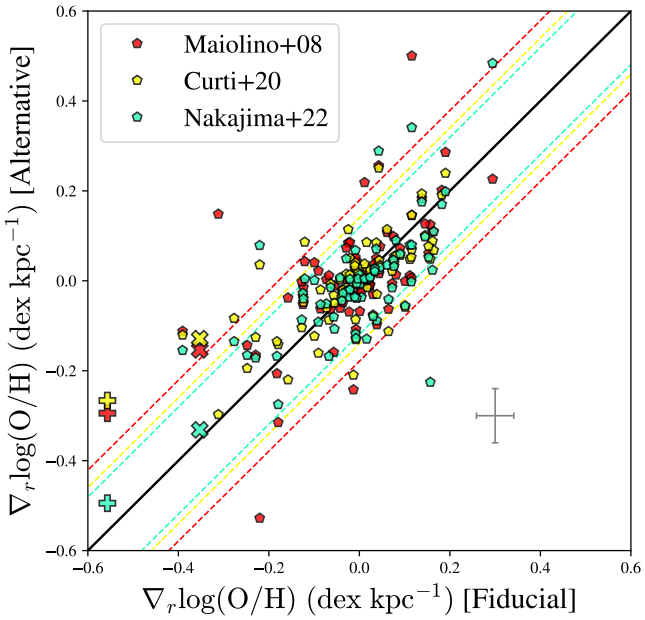


Figure 16. Comparisons between different strong-line calibrations for metallicity gradients. NGDEEP, ASPIRE, and FRESCO measurements are marked with pentagons, "×," and "+," respectively. The x-axis is the fiducial measurements with F. Bian et al. (2018) for NGDEEP galaxies and R. L. Sanders et al. (2024) for ASPIRE and FRESCO galaxies, and the y-axis is the measurements using alternative calibrations (R. Maiolino et al. 2008; M. Curti et al. 2020b; K. Nakajima et al. 2022). The diagonal black line shows equality, and the dashed lines show the standard deviation around equality for each calibration in the corresponding color.

using any of the three alternative calibrations, with all of the scatters around equality being within 1σ measurement uncertainties $\lesssim$0.1 dex kpc^{-1}. The different calibrations for two $z > 5$ stacks also introduce small uncertainties $\lesssim$0.1 dex kpc^{-1}. Thus, the choice of metallicity calibration does not alter our results on the evolution of the metallicity gradient.

Appendix F
Integrated Metallicity of $z > 5$ Sample

Here we measure the integrated metallicity of our stacks in Figure 1. We resample our 1D spectra to the rest frame on a common 1 Å wavelength grid with flux preserved using `spectres` (A. C. Carnall 2017). Following X. Wang et al. (2022b), to avoid the excessive weighting toward bright sources with stronger line fluxes, we normalize each spectrum by its measured [O III] flux. We take the median value of the normalized spectra at each wavelength grid, and the uncertainty is estimated by measuring the standard deviation from 1000 bootstrap realizations of the sample. The median stacked 1D rest-frame spectra of ASPIRE and FRESCO are shown in Figure 17. We fit [O III] and Hβ line fluxes using Gaussian profiles. We do not set line ratio constraints between [O III]$_{5007}$ and [O III]$_{4959}$. Assuming the lower-branch solution, the metallicities are measured using R3 calibrations from R. L. Sanders et al. (2024), the same methods we use for metallicity gradient measurements. We also compare the metallicities from different calibrations in K. Nakajima et al. (2022). We list the integrated line fluxes and metallicities in Table 3. We note that the K. Nakajima et al. (2022) calibration gives ∼0.1–0.2 dex high metallicities. More detailed analysis of the integrated metallicity and MZR of the ASPIRE and NGDEEP samples is presented in Z. Li et al. (2025) and X. He et al. (2025).

Table 3
Measurements from Stacked 1D Spectra in ASPIRE and FRESCO

Property	ASPIRE	FRESCO
log($R3$)	0.79 ± 0.02	0.78 ± 0.04
$F_{\rm O\,III_{4959}}$ [a]	34.17 ± 0.70	30.18 ± 1.64
$F_{\rm H_\beta}$ [a]	16.10 ± 0.60	16.51 ± 1.45
$12 + \log({\rm O/H})_{\rm Sanders+23}$ [b]	$7.61^{+0.07}_{-0.05}$	$7.57^{+0.18}_{-0.10}$
$12 + \log({\rm O/H})_{\rm Nakajima+22}$ [c]	$7.78^{+0.09}_{-0.06}$	$7.74^{+0.16}_{-0.10}$

Notes.
[a] Normalized as $F_{\rm O\,III_{5007}} = 100$.
[b] With R. L. Sanders et al. (2024) calibration assuming lower-branch solution.
[c] With K. Nakajima et al. (2022) calibration assuming lower-branch solution.

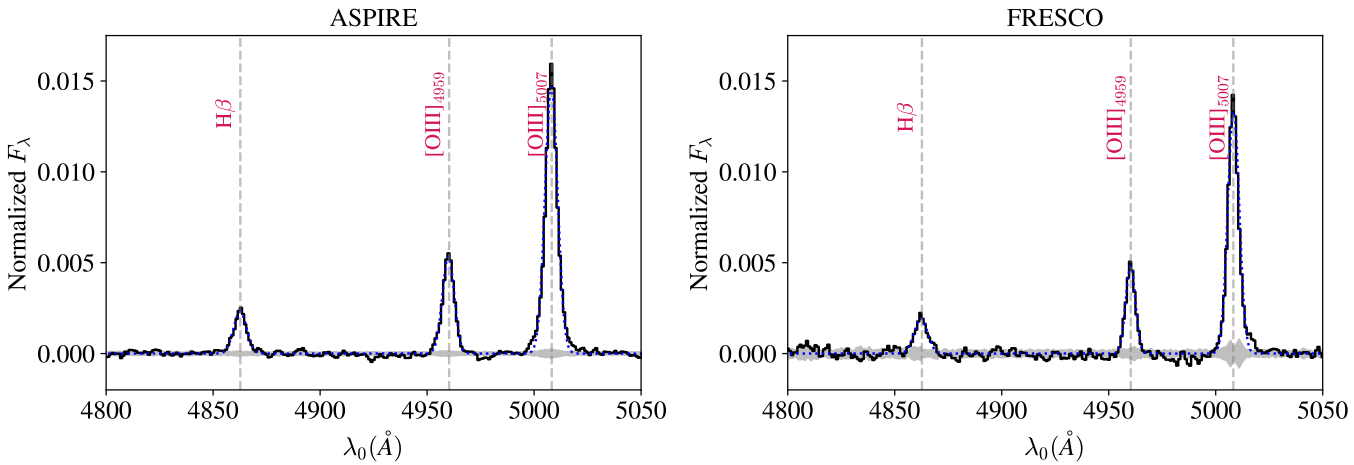


Figure 17. Median stacked 1D rest-frame spectra of galaxies in ASPIRE and FRESCO. The fluxes are normalized by [O III]$_{4959+5007}$ flux to avoid excessive weighting toward bright sources. The gray shadow represents the spectral uncertainty. The blue dotted lines are the best-fit Gaussian models.

Appendix G
Uncertainties in Gradient Measurements

To quantify how uncertainties in emission may impact our measurements in the NGDEEP sample, we do a mock test in the same manner as in Z. Li et al. (2022). We fit a normal model to the distribution of the effective radius R_e and lognormal models to the axis ratio b/a and $\mathrm{SNR}_{[\mathrm{O\,III}]}$ in the NGDEEP sample and randomly generate mock galaxies with 2D Sérsic surface brightness profiles given the distribution of R_e, b/a, and $\mathrm{SNR}_{[\mathrm{O\,III}]}$. We apply a flat gradient to all mock galaxies and use the same method discussed before to reconstruct the metallicity gradients. We measure the scatter of 200 mock galaxies and find the intrinsic scatter in the measurement to be $\approx$0.09 dex kpc^{-1}.

Since we use the stacking method for galaxies at $z > 5$ in ASPIRE and FRESCO, we also use the mock data set to test whether it reflects the median metallicity gradients by stacking emission maps. We start by randomly generating 300 galaxies with stellar masses following a log-norm distribution, which is fitted to our SED results. We then assign the effective radius and metallicity of each mock galaxy using the mass–size relation (D. Langeroodi & J. Hjorth 2023) and the MZR (A. Sarkar et al. 2025) measured at similar redshifts. We also consider the intrinsic scatter of these relations, so we add Gaussian noise to R_e and $12 + \log(\mathrm{O/H})$, with $\sigma_{\log(R_e/\mathrm{kpc})} = 0.25$ and $\sigma_{12+\log(\mathrm{O/H})} = 0.16$ (D. Langeroodi & J. Hjorth 2023; A. Sarkar et al. 2025). We fit a normal distribution to the observed axis ratios and then randomly assign the axis ratios to the Sérsic model of the emission maps using the normal distribution. Additionally, we find that $\mathrm{SNR}_{[\mathrm{O\,III}]}$ is not strongly correlated with other physical parameters, so we randomly assign $\mathrm{SNR}_{[\mathrm{O\,III}]}$ for these galaxies using a lognormal distribution that best represents the observed $\mathrm{SNR}_{[\mathrm{O\,III}]}$ distribution. Finally, we construct two different mock sets, one with metallicity gradients randomly chosen from a uniform distribution $\mathcal{U}(-0.6, 0)$, and one with all-zero gradients. We measure the median metallicity gradient from the stacked maps of [O III] and Hβ. The inferred metallicity gradients are $k = -0.34 \pm 0.13$ for the negative gradient data set and $k = -0.01 \pm 0.13$ for the flat gradient data set, as shown in Figure 18. We have found that the gradient measured from the stacked emission maps can effectively represent the median gradients of the mock sample. Therefore, the gradients observed for galaxies at $z \approx 6$–7 are strong representations of the populations.

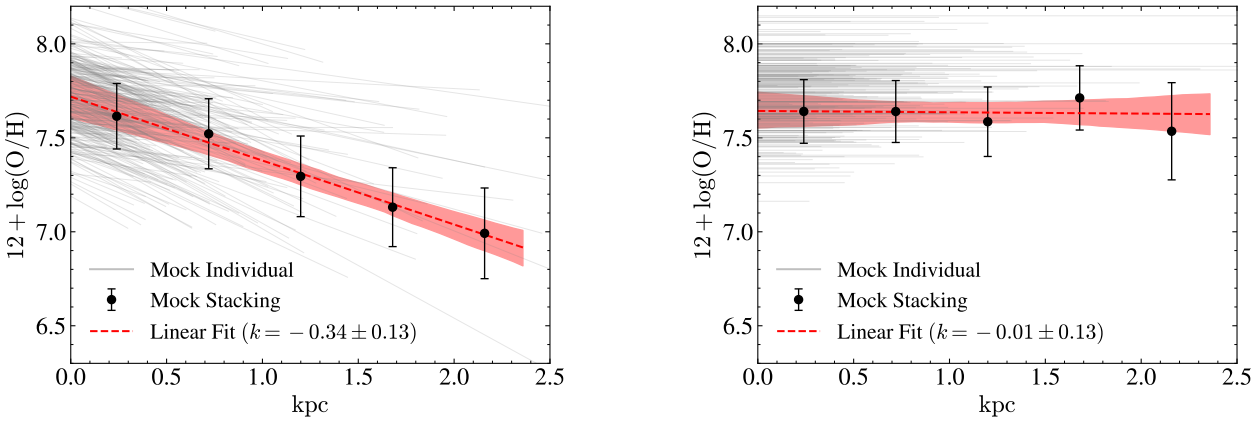


Figure 18. Metallicity gradients in the mock test. In the left panel, the metallicity gradients of the mock galaxies are randomly chosen from a uniform distribution $\mathcal{U}(-0.6, 0)$, while all the mock galaxies in the right panel have a zero gradient. Gray lines represent individual galaxies, with length representing twice the effective radius of the galaxy. Black points are measured from the stacked emission maps, with error bars showing the 1σ uncertainty. The red dashed line shows the linear fitting to the measured points, and the red shadowed region is its 1σ confidence interval.

Appendix H
The Impact on the Metallicity of the Variation of the Ionization Parameter (U)

The observed line ratios depend on both metallicity and the ionization parameter (U, the ratio of the number of ionizing photons to the gas number density). Galaxies at $z \sim 2$–9 typically have values of log(U) varying from -3.5 to -1.5 (A. L. Strom et al. 2018; N. A. Reddy et al. 2023; M. Tang et al. 2023; J. R. Trump et al. 2023). The varied log(U) can lead to a line ratio gradient, even if there is no gradient in metallicity.

To quantify the effects of varied log(U) on measured metallicity gradients, we examine the dependence of the observed line ratio on U using the recent photoionization model provided by P. Garg et al. (2024). We apply their two models that simulate galaxies at $z \approx 5$, one with low ionization (median log(U) $= -2.3$) and one with high ionization (median log(U) $= -0.86$). In Figure 19(a), we vary log(U) from the center to the outskirts, transitioning either from low to high or from high to low. In Figure 19(b), we observe that different log (U) can produce the maximum $\approx$0.1–0.2 dex in the R3 ratio. Meanwhile, our observed R3 ratios vary by $\approx$0.5 dex, larger than can be explained by the ionization parameter. As a result, different metallicities are required to reproduce the observed R3 ratios.

To further quantify the uncertainty induced by varied log (U), we consider the measurement of a galaxy out to a radius $R = 2$ kpc (as in our observations). Assuming a galaxy with uniform metallicity, but with an ionization parameter log(U) varying with radius, we expect to observe varied R3 ratios. Then, we use the calibration from R. L. Sanders et al. (2024) to derive the metallicity from such observed line ratios, and we obtain a spurious nonzero metallicity gradient—either positive or negative—corresponding to positive or negative gradients of log(U), respectively. In the right panel of Figure 19, we show such uncertainty as a function of galaxy metallicity $Z_{\rm gal}$. We see that the error increases with metallicity. For typical galaxies in our sample with $Z_{\rm gal} \sim 7.6$, the induced uncertainty is $\approx$0.1 dex kpc^{-1}, with the maximum variation of log(U) from log(U) $= -2.3$ to -0.86 (P. Garg et al. 2024). Compared with the metallicity gradient of $\nabla Z \approx -0.4$ dex kpc^{-1} in our JWST observations, the maximum uncertainty induced by log(U) falls within the range of the statistical error. Supporting this, H. Poetrodjojo et al. (2018) found no significant variation in ionization parameters with ionization parameter gradients $\lesssim$0.05 dex R_e^{-1}. If the gradient continues to 4 $R_e \sim 2$ kpc, the variation in log(U) is only $\lesssim$0.2 dex, much less than the extreme variation we assume $\sim$1.44 dex. Thus, possible variation in U does not dominate the variation of line ratios (also see G. Cresci et al. 2010), and we expect lower uncertainty $<$0.1 dex kpc^{-1} from varied log(U).

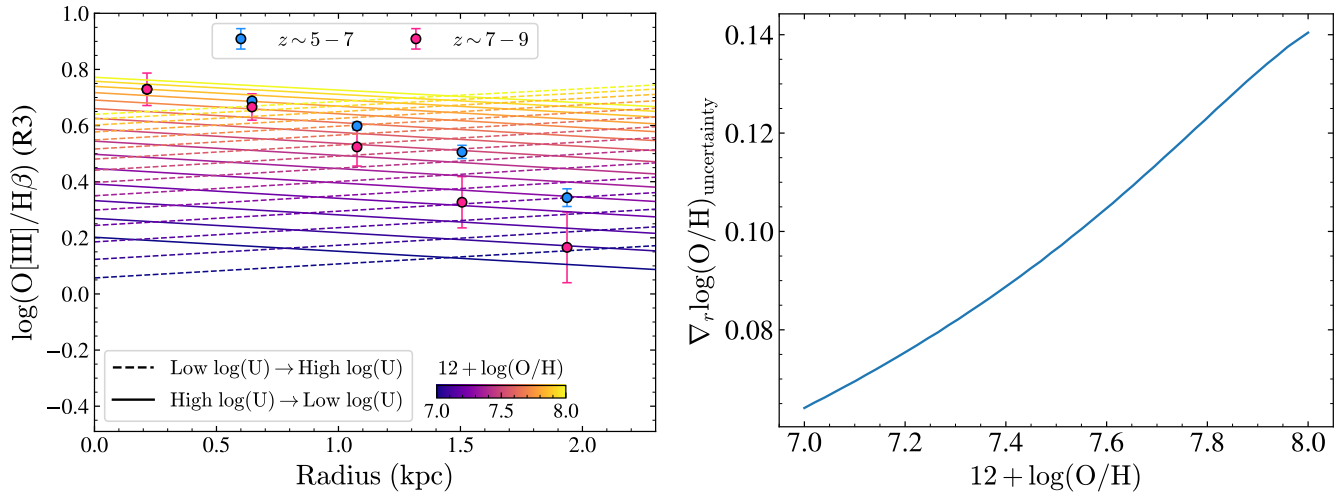


Figure 19. (a) The observed R3 ratio in two redshift bins (red and blue circles), and the modeled R3 ratio from P. Garg et al. (2024). The models are color coded with different metallicities in the range 12 + log(O/H) = 7–8. We show that the ionization parameter, U, varies from high in the center to low in the outskirts (solid line), and from low in the center to high in the outskirts (dashed line). (b) The uncertainty of the metallicity gradient due to the variation of ionization parameters.

Appendix I
AGN Contamination

Since galaxies with AGNs are powered by additional AGN ionization, the standard metallicity calibrations (F. Bian et al. 2018; R. L. Sanders et al. 2024) are not suitable for AGN-contaminated sources. As such, we remove sources with possible AGN contamination. We use the mass–excitation diagram (A. L. Coil et al. 2015) to separate AGNs from star-forming galaxies. In Figure 20, we have removed sources $\geqslant 2\sigma$ from the demarcation. As shown, all of our sample galaxies are classified as typical star-forming galaxies.

Nevertheless, we also note that AGNs may be more common in high-z galaxies. Y. Harikane et al. (2023) and R. Maiolino et al. (2024) found red AGNs with broad-line features take a fraction as high as 10%–20% at $z > 5$. Y. Harikane et al. (2023) and R. Maiolino et al. (2024) suggested an offset for high-z AGNs in the BPT diagram (and also in the mass–excitation diagram). R. Maiolino et al. (2024) provided several possibilities for the high-z AGN offset in the BPT diagram. One is that the narrow-line emission is dominated by star formation, instead of the AGN narrow-line region (NLR). R. Maiolino et al. (2025) further suggested a high covering factor of the broad-line region (BLR), prohibiting photons from escaping to produce NLR emissions.

An alternative explanation is that the NLR of high-z AGNs is characterized by low metallicities (Y. Harikane et al. 2023; R. Maiolino et al. 2024). With decreasing metallicity in the NLR, the emission-line ratios shift from the local AGN locus to be mixed with the star-forming locus. AGNs may be confused with low-mass, low-metallicity star-formation-dominated galaxies. Y. I. Izotov & T. X. Thuan (2008) show that AGN nonthermal radiation should contribute less than 10% to the total ionizing radiation in metal-poor ($12 + \log(\mathrm{O/H}) < 8$) AGN host galaxies, to reproduce the shift in the BPT diagram. In this case, the narrow-line emission from black holes should still be much weaker (<10%) than emissions from stellar radiation. From these models, AGN contamination in our sample should be small, and the [O III] fluxes should come predominantly from star formation.

Type-2 AGNs, however, have been observed to be similar to or more prevalent than type-1 AGNs at high redshift. A similar issue has been raised by J. Scholtz et al. (2025) that type-2 AGNs have offsets in the BPT diagram, and it is very difficult to distinguish them from star-forming galaxies using the standard BPT diagram unless high-ionization lines are detected. J. Scholtz et al. (2025) selected type-2 AGNs from a sample of galaxies with JWST/NIRSpec observations using high-ionization lines in combination with UV transitions. While they found that type-2 AGNs have weaker [N II]/Hα ratios, they have a similar [O III]/Hβ ratio to star-forming galaxies, and they measured the [O III]/Hβ ratio to be 4.55 ± 0.15 and 4.70 ± 0.14 from stacked spectra of type-2 AGNs and star-forming galaxies, respectively. Although it is difficult to distinguish type-2 AGNs solely based on [O III]/Hβ ratios, the central AGN should not significantly change the line ratios, as type-2 AGNs contribute a very similar amount of [O III] and Hβ emission to what star-forming galaxies contribute.

To further quantify the impact of AGNs on emission maps at $z > 5$, we have conducted a set of mock tests. We define the AGN fraction as the AGN contribution of [O III] luminosity to the total [O III] luminosity of the galaxy. As AGNs are point sources, their emissions are detected as PSF components in emission maps. We construct a mock galaxy at $z \sim 6$ with a Sérsic profile, and we add a central AGN component as a NIRCam F356W PSF. We assume

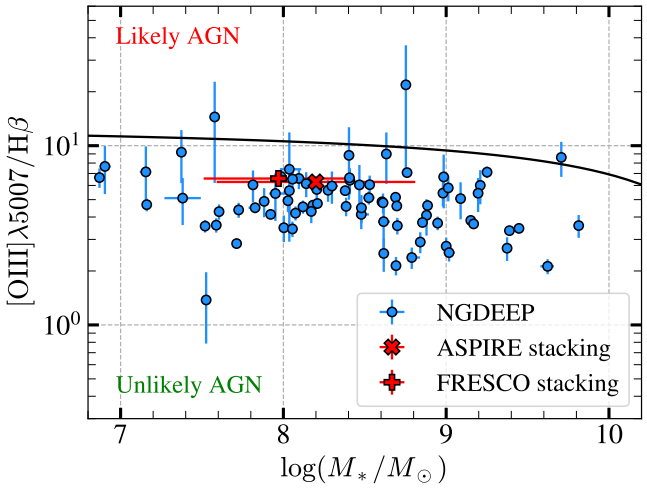


Figure 20. The mass–excitation diagram for our sample galaxies. The blue points show individual measurements in NGDEEP, and the red "×" and "+" represent the median stacked result in ASPIRE and FRESCO. The solid line is the demarcation scheme (A. L. Coil et al. 2015), where the points below the line are unlikely to be AGNs. Our sample galaxies in NGDEEP at $z \approx 1$–3 are safely excluded from being AGNs, and the ASPIRE and FRESCO stackings at $z \approx 6$–7 are also negligibly contaminated by possible AGNs.

two cases, one with a high [O III]/Hβ ratio as high as $([\mathrm{O\,III}]/\mathrm{H}\beta)_{\mathrm{AGN}} = 10$, as is expected at lower redshift (S. Juneau et al. 2014), and another with a similar ratio to that of a star-forming galaxy $([\mathrm{O\,III}]/\mathrm{H}\beta)_{\mathrm{AGN}} = 6$, as is expected from recent AGN observations at $z > 4$ (J. Scholtz et al. 2025). In Figure 21, we show the [O III] line, Hβ line, and line ratio profiles in galaxies with different AGN fractions. We find that an obvious sign of AGN in emission maps is a bump in the line ratio profiles at $r = 1.5$ kpc, induced by the wing of the PSF. We can also observe that when $([\mathrm{O\,III}]/\mathrm{H}\beta)_{\mathrm{AGN}} = 10$, higher than the line ratio from star formation, the AGN component contributes to a steeper [O III]/Hβ profile as a result of boosting [O III] emission in the center. However, if the AGN emission $([\mathrm{O\,III}]/\mathrm{H}\beta)_{\mathrm{AGN}} = 6$ is comparable to the emission from galaxies (e.g., as revealed in J. Scholtz et al. 2025), the change of [O III]/Hβ slope is quite small if the AGN fraction is below ∼50%. R. Maiolino et al. (2024, 2025) suggested that AGN NLR emission is subdominant, possibly due to either the high cover fraction in the BLR or the low-metallicity nature of the NLR. The AGN component should contribute to a small fraction of the total [O III] emissions from the host galaxy. M. Curti et al. (2024) also found no significant change in the MZR when including AGN candidates, assuming the metallicity calibration is still valid with the contribution of AGN ionization. As a result, our measured metallicity gradient based on [O III]/Hβ is less likely to be biased by AGN contamination.

Similar concerns are raised by J. R. Trump et al. (2011). The positive metallicity gradient observed at $z \sim 2$ is associated with more extended Hβ emission, resulting in a negative gradient of the R3 ratio. Since the metallicities fall on the upper branch of the R3 calibration, this corresponds to a positive gradient in metallicity. J. R. Trump et al. (2011) found more extended Hβ emission at $z \sim 2$. They attributed this to the consequence of enhanced [O III] emission boosted by the central AGN. Based on our earlier discussion, if AGNs are unlikely to dominate the flux profile, the observed trend can also be interpreted as a result of positive metallicity gradients, consistent with our results at $z \sim 2$.

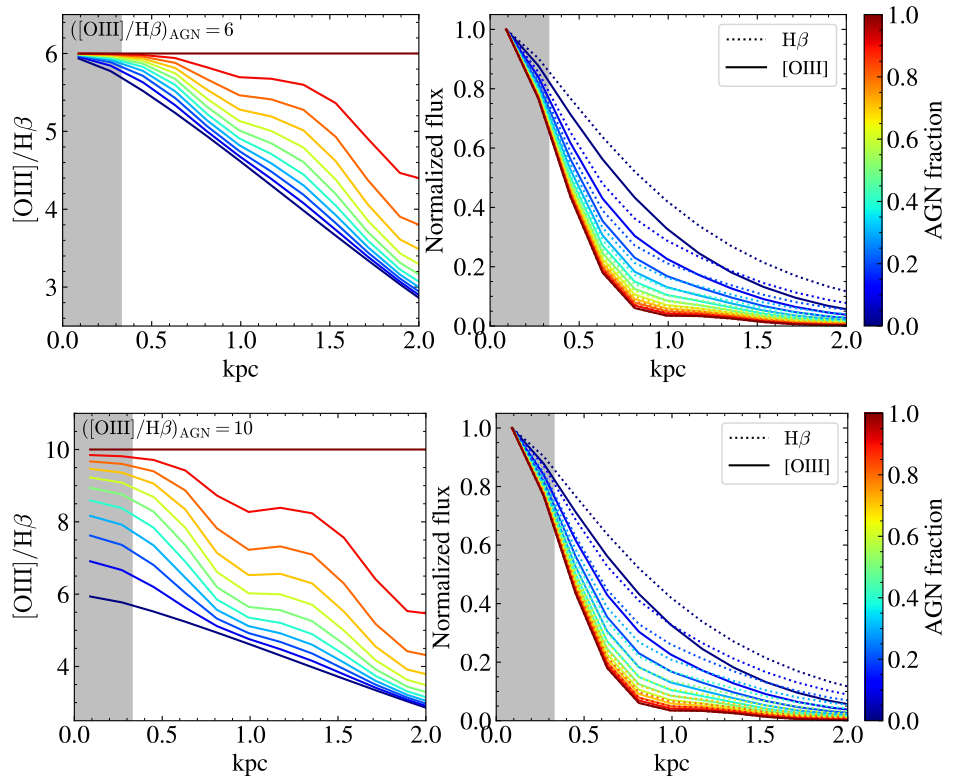


Figure 21. The [O III]/Hβ ratios (left) and their normalized profiles (right) with varying AGN fractions. The profiles are normalized by their central peak flux. In the top row, we assume a ratio of ([O III]/Hβ)$_{\rm AGN}$ = 6, indicative of AGN emission comparable to typical star-forming galaxies at $z > 5$. In the bottom row, we consider an enhanced [O III] emission from AGNs, with a ratio of ([O III]/Hβ)$_{\rm AGN}$ = 10.

Appendix J
Halo Mass–Redshift Distribution and Analytic Predictions

We estimate the halo masses using the empirical stellar-to-halo mass relation in M. Shuntov et al. (2022). It is worth noting that their highest-redshift bin extends only to $z = 5.5$, whereas we extend this relation for our $5 < z < 9$ sample. We should be cautious of this halo mass estimation, which may deviate at higher redshifts. As this is a qualitative comparison, we are less sensitive to accurate halo masses. In Figure 22, we show the halo mass versus redshift for the whole galaxy sample, in comparison with the analytic predictions of cold/hot accretion (A. Dekel & Y. Birnboim 2006; A. Dekel et al. 2009) and FFBs (A. Dekel et al. 2023). We find that at $z \gtrsim 2$, most of the sample galaxies are capable of sustaining cold streams. In contrast, at lower redshift $z \lesssim 2$, the inflow gas is expected to be shock-heated for massive galaxies, and only low-mass galaxies can be fed with cold streams. This suggests that cold-mode accretion plays a significant role at $z \sim 2$, contributing to flat metallicity gradients, while hot-mode accretion becomes dominant at lower redshifts, leading to negative gradients.

At higher redshift $z \gtrsim 6$, most of the galaxies can fall within the FFB regime. In massive halos above the green solid line, the mean gas density in the ISM is expected to reach the FFB threshold to boost global star formation in FFB mode. The green dashed line shows the model with higher density contrast ($c = 10$, Equation (37) in A. Dekel et al. 2023) between FFB clouds and the mean overdensity. For the less massive galaxies between the green solid and dashed lines, the gas clouds can reach the FFB threshold when their contraction

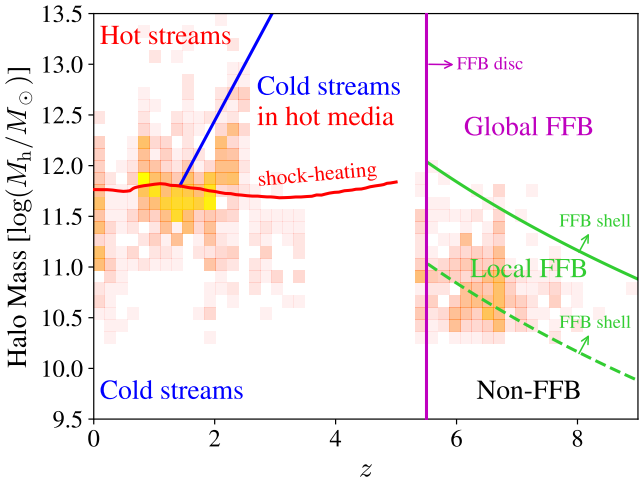


Figure 22. The 2D histogram of the halo mass and redshift distribution of all sample galaxies. The regimes predicted by analytic models of cold accretion, hot accretion (A. Dekel & Y. Birnboim 2006; A. Dekel et al. 2009), and FFBs (A. Dekel et al. 2023) are demarcated by lines in different colors.

reaches a higher density than the mean gas density, so that the FFB can happen locally. If the gas can further collapse into higher density, we would expect a lower halo mass threshold for the occurrence of local FFBs. As a result, local FFBs might happen in most of our sample galaxies, and this supports the steep negative gradients in the FFB scenario. This qualitative comparison hints at the transition of different gas accretion and star formation scenarios across redshifts, supporting our interpretation in the main text.

Appendix K
Degeneracies in Simulations

Simulations have shown that the radial distribution of metals is shaped by factors such as the gas-phase structure, feedback mechanisms, metal yields, and diffusion in simulations (S. Roca-Fàbrega et al. 2024). Below, we briefly discuss the impact of these factors.

1. *The gas-phase structure*. MUGS, FIRE, and FOGGIE all allow explicitly a multiphase ISM, while the ISM gas in EAGLE and TNG50 relies on the effective equation of state (eEOS, e.g., V. Springel & L. Hernquist 2003), resulting in different levels of metal mixing. The application of eEOS in TNG50 is not able to explicitly model small-scale (i.e., $<$100 pc) structures that pressurize the ISM, such as turbulence (Z. S. Hemler et al. 2021). The unresolved small-scale turbulence may serve to radially mix chemically enriched gas and flatten the metallicity gradients (P. Sharda et al. 2021c). Moreover, they are not affected by the continuous burst of feedback, and they go through more steady star formation, which consequently develops a steeper gradient (Z. S. Hemler et al. 2021). However, with continuous feedback, it can erase the metallicity gradient in simulations with a multiphase ISM (such as FIRE; X. Ma et al. 2017).

2. *Feedback strategy*. The different implementations of feedback (e.g., stellar and AGN feedback) in different simulations yield vastly different halo baryon fractions (Figure 9 of R. A. Crain & F. van de Voort 2023), particularly at the low-mass end, influencing the mixing and evacuation of gas to different extents. TNG50 simulations include the feedback of both stars and AGNs (Z. S. Hemler et al. 2021), but their subgrid prescriptions and parameterized winds result in smooth star formation and feedback. FIRE does not have AGN feedback, but its explicitly modeled strong stellar and radiative feedback will also play a significant role in the evacuation of gas from the CGM (X. Ma et al. 2017; X. Sun et al. 2025). For SN feedback, previous simulations have shown that a factor of $\approx$2.5 more powerful SN feedback could lead to a $\approx$0.2 dex kpc^{-1} higher metallicity gradient at $z \geqslant 1.5$ (B. K. Gibson et al. 2013). However, the SN feedback prescription in FOGGIE simulations is insufficient and unable to expel enough metal-rich gas (A. Acharyya et al. 2025). Different modes of feedback can also cause different extents of gas transport. I. Shimizu et al. (2019) found that kinetic-only feedback can more easily enrich the ISM and CGM than thermal-only feedback, leading to more metal-enriched outskirts and a shallower gradient. They also found that turnoff cooling also influences SN feedback efficiency. Thus, how we treat feedback and how we include different fractions of feedback models lead to different extents of metal enrichment.

3. *The metal yield and diffusion*. How the mass of stars converts to the mass of metals in the ISM and how the metal mixes with the ISM in simulations also influence the galaxy metallicity gradient. These two processes are described by the parameters of metal yield (the ratio between the metal mass and the stellar mass) and the yield reduction factor (ϕ_y, the fraction of metals mixed with the ISM). FIRE, MUGS, and FOGGIE use three different metal yields, with FIRE using Type II SNe (SNe II), Type Ia SNe (SNe Ia), and winds as sources (yields from S. E. Woosley & T. A. Weaver 1995, K. Iwamoto et al. 1999, and R. G. Izzard et al. 2004, respectively), MUGS only accounting for SNe II (C. M. Raiteri et al. 1996) and SNe Ia (F. K. Thielemann et al. 1986), and FOGGIE using a custom recipe for the metal enrichment. Different metal yields influence the absolute metals produced in the galaxies and, for example, can imprint on the normalization of MZRs (e.g., X. Ma et al. 2016; A. Marszewski et al. 2024). While the absolute amount of metal is less important on the metallicity gradient, the gradient relies on the relative abundance of metal at different radii. For the yield reduction factor, P. Sharda et al. (2021c) have shown that the galaxy metallicity gradient is inversely correlated with ϕ_y. In the FOGGIE simulation, they found the scatter of metallicity gradients may be related to this yield reduction factor (A. Acharyya et al. 2025). Note that the gas-phase metallicity gradient measures the radial change of the oxygen abundance relative to hydrogen in the ionized gas, and it is less sensitive to the different values of stellar yield assumed in the models. A factor of 10 increase in the yield reduction factor can only bring about a $\lesssim$0.2 dex change in gradient slopes according to analytical models (P. Sharda et al. 2021c). Thus, different cosmological simulation results can be used here for qualitative comparison.

These factors all influence the distribution of gas-phase metallicity in galaxies to different extents. Although the degeneracies inherent in these different simulation architectures complicate the ability to isolate individual effects based on limited observation constraints, understanding the relative importance of these factors helps improve our interpretation of galaxy formation and evolution processes. It is hard to quantify the exact mechanisms that drive the metallicity gradients we observe; we generally prefer stronger feedback at $z \sim 2$ and less efficient feedback at higher redshift $z > 5$. This nuanced understanding is essential for future studies aiming to reconcile simulation outcomes with observational data.

Appendix L
Measured Quantities of Individual Galaxies

In Table 4, we provide the measured quantities of individual galaxies.

Table 4
Measured Quantities of Individual Galaxies in Our Sample

Field	ID	R.A. (deg)	Decl. (deg)	Redshift	$\log(M_*/M_\odot)$	SFR ($M_\odot$ yr^{-1})	$12 + \log(\mathrm{O/H})$	$\nabla_r \log(\mathrm{O/H})$ (dex kpc^{-1})
ASPIRE	Zgrad-1	315.583439	−14.976233	6.66	$8.41^{+0.10}_{-0.06}$	$25.67^{+6.76}_{-3.23}$	$7.73^{+0.14}_{-0.18}$	-0.18 ± 0.26
ASPIRE	Zgrad-2	137.723621	−4.222954	6.64	$8.24^{+0.08}_{-0.03}$	$17.26^{+3.53}_{-1.06}$	$7.75^{+0.12}_{-0.17}$	-0.16 ± 0.26
ASPIRE	Zgrad-3	137.700518	−4.272600	6.21	$8.22^{+0.04}_{-0.05}$	$16.67^{+1.62}_{-1.70}$	$7.72^{+0.15}_{-0.20}$	-0.25 ± 0.27
ASPIRE	Zgrad-4	46.318882	−31.839629	6.67	$8.55^{+0.09}_{-0.09}$	$33.19^{+6.64}_{-5.98}$	$7.61^{+0.21}_{-0.21}$	-0.15 ± 0.28
ASPIRE	Zgrad-5	17.475148	−30.798016	6.71	$8.35^{+0.09}_{-0.06}$	$22.16^{+4.98}_{-2.75}$	$7.69^{+0.16}_{-0.19}$	-0.10 ± 0.39
ASPIRE	Zgrad-6	140.957012	4.049241	6.39	$8.04^{+0.03}_{-0.02}$	$10.98^{+0.69}_{-0.60}$	$7.31^{+0.28}_{-0.17}$	-0.45 ± 0.31
NGDEEP	00060	53.150753	−27.804927	2.04	$7.52^{+0.02}_{-0.02}$	$1.36^{+0.05}_{-0.06}$	$7.87^{+0.06}_{-0.05}$	0.14 ± 0.05
NGDEEP	00096	53.151582	−27.803721	2.12	$8.41^{+0.03}_{-0.02}$	$1.83^{+0.10}_{-0.13}$	$8.29^{+0.05}_{-0.05}$	-0.07 ± 0.02
NGDEEP	00173	53.148643	−27.801635	2.12	$8.10^{+0.03}_{-0.05}$	$1.52^{+0.09}_{-0.07}$	$7.98^{+0.04}_{-0.04}$	-0.09 ± 0.04
NGDEEP	00197	53.153745	−27.801119	2.02	$7.16^{+0.03}_{-0.03}$	$1.34^{+0.05}_{-0.07}$	$7.87^{+0.05}_{-0.04}$	0.05 ± 0.04
NGDEEP	00228	53.150438	−27.800622	2.62	$9.00^{+0.02}_{-0.02}$	$3.26^{+0.22}_{-0.20}$	$8.50^{+0.06}_{-0.08}$	0.08 ± 0.07
NGDEEP	00242	53.160060	−27.800122	2.22	$8.00^{+0.01}_{-0.01}$	$1.11^{+0.04}_{-0.04}$	$7.93^{+0.07}_{-0.07}$	-0.11 ± 0.13
NGDEEP	00310	53.156367	−27.799048	2.67	$9.37^{+0.01}_{-0.01}$	$1.29^{+0.07}_{-0.06}$	$8.52^{+0.04}_{-0.05}$	0.07 ± 0.05
NGDEEP	00317	53.145614	−27.798979	2.62	$9.62^{+0.04}_{-0.04}$	$11.65^{+0.99}_{-0.82}$	$8.61^{+0.04}_{-0.04}$	-0.03 ± 0.05
NGDEEP	00407	53.171345	−27.797899	1.99	$8.84^{+0.02}_{-0.02}$	$1.23^{+0.13}_{-0.13}$	$8.47^{+0.04}_{-0.05}$	-0.15 ± 0.03
NGDEEP	00413	53.152521	−27.797787	3.00	$7.58^{+0.04}_{-0.03}$	$2.08^{+0.08}_{-0.11}$	$8.05^{+0.07}_{-0.07}$	-0.01 ± 0.13
NGDEEP	00436	53.170460	−27.797382	2.08	$9.02^{+0.04}_{-0.03}$	$1.72^{+0.12}_{-0.12}$	$8.42^{+0.05}_{-0.07}$	-0.03 ± 0.04
NGDEEP	00466	53.168830	−27.796994	1.99	$8.61^{+0.01}_{-0.01}$	$4.08^{+0.19}_{-0.24}$	$8.05^{+0.05}_{-0.05}$	-0.04 ± 0.02
NGDEEP	00508	53.173491	−27.796419	2.75	$7.38^{+0.11}_{-0.11}$	$0.99^{+0.09}_{-0.08}$	$8.01^{+0.12}_{-0.10}$	-0.19 ± 0.10
NGDEEP	00536	53.148336	−27.796032	2.02	$8.61^{+0.03}_{-0.03}$	$1.56^{+0.11}_{-0.13}$	$8.34^{+0.05}_{-0.05}$	-0.00 ± 0.03
NGDEEP	00618	53.146372	−27.795289	2.63	$8.69^{+0.01}_{-0.01}$	$2.38^{+0.13}_{-0.11}$	$8.59^{+0.04}_{-0.05}$	-0.13 ± 0.07
NGDEEP	00619	53.143575	−27.795297	2.02	$8.18^{+0.02}_{-0.01}$	$7.02^{+0.33}_{-0.32}$	$8.09^{+0.07}_{-0.06}$	0.01 ± 0.01
NGDEEP	00683	53.148356	−27.794324	2.63	$7.73^{+0.02}_{-0.02}$	$3.21^{+0.13}_{-0.11}$	$8.14^{+0.11}_{-0.11}$	-0.10 ± 0.02
NGDEEP	00729	53.172439	−27.793993	2.80	$8.95^{+0.02}_{-0.01}$	$5.88^{+0.12}_{-0.11}$	$8.39^{+0.06}_{-0.07}$	-0.03 ± 0.03
NGDEEP	00745	53.152584	−27.793907	3.07	$9.71^{+0.01}_{-0.01}$	$15.44^{+0.59}_{-0.39}$	$8.34^{+0.03}_{-0.03}$	0.00 ± 0.02
NGDEEP	00760	53.165077	−27.793719	2.34	$8.89^{+0.01}_{-0.01}$	$6.19^{+0.16}_{-0.15}$	$8.35^{+0.04}_{-0.04}$	-0.04 ± 0.01
NGDEEP	00782	53.147407	−27.793410	2.79	$7.81^{+0.02}_{-0.02}$	$3.08^{+0.15}_{-0.29}$	$8.12^{+0.09}_{-0.08}$	-0.08 ± 0.11
NGDEEP	00784	53.170795	−27.793366	2.67	$7.71^{+0.02}_{-0.02}$	$2.79^{+0.25}_{-0.19}$	$7.93^{+0.09}_{-0.07}$	0.59 ± 0.05
NGDEEP	00789	53.172714	−27.793316	2.81	$8.88^{+0.02}_{-0.02}$	$2.08^{+0.32}_{-0.31}$	$8.36^{+0.07}_{-0.07}$	-0.32 ± 0.08
NGDEEP	00860	53.153127	−27.792447	3.32	$8.63^{+0.02}_{-0.02}$	$4.95^{+0.18}_{-0.23}$	$8.17^{+0.06}_{-0.05}$	0.05 ± 0.09
NGDEEP	00868	53.151284	−27.792470	1.84	$8.85^{+0.03}_{-0.03}$	$5.95^{+0.21}_{-0.20}$	$8.28^{+0.04}_{-0.04}$	0.04 ± 0.01
NGDEEP	00917	53.147644	−27.791848	1.87	$7.95^{+0.07}_{-0.04}$	$0.45^{+0.02}_{-0.02}$	$8.41^{+0.05}_{-0.06}$	-0.26 ± 0.11
NGDEEP	00928	53.146782	−27.791739	2.79	$8.40^{+0.03}_{-0.02}$	$1.16^{+0.09}_{-0.08}$	$7.94^{+0.10}_{-0.10}$	-0.13 ± 0.12
NGDEEP	00996	53.178202	−27.790838	1.76	$8.62^{+0.02}_{-0.02}$	$0.41^{+0.02}_{-0.02}$	$8.19^{+0.08}_{-0.11}$	0.08 ± 0.18
NGDEEP	01014	53.153116	−27.790565	3.32	$8.14^{+0.02}_{-0.02}$	$5.61^{+0.25}_{-0.45}$	$7.96^{+0.07}_{-0.08}$	0.11 ± 0.04
NGDEEP	01066	53.167405	−27.790198	2.76	$8.27^{+0.02}_{-0.02}$	$3.63^{+0.23}_{-0.15}$	$8.16^{+0.08}_{-0.08}$	-0.24 ± 0.05
NGDEEP	01145	53.171118	−27.789204	2.84	$6.90^{+0.04}_{-0.02}$	$0.81^{+0.07}_{-0.04}$	$7.93^{+0.08}_{-0.08}$	0.08 ± 0.19
NGDEEP	01173	53.141930	−27.788913	2.71	$9.09^{+0.03}_{-0.03}$	$1.48^{+0.09}_{-0.08}$	$8.28^{+0.08}_{-0.07}$	0.03 ± 0.13
NGDEEP	01222	53.169688	−27.788158	3.18	$9.21^{+0.01}_{-0.01}$	$7.56^{+0.25}_{-0.25}$	$8.29^{+0.05}_{-0.04}$	-0.04 ± 0.06
NGDEEP	01223	53.146252	−27.788128	2.00	$8.30^{+0.04}_{-0.04}$	$1.00^{+0.04}_{-0.04}$	$8.20^{+0.04}_{-0.04}$	-0.07 ± 0.04
NGDEEP	01227	53.143415	−27.788081	3.40	$8.75^{+0.01}_{-0.01}$	$5.86^{+0.19}_{-0.17}$	$8.05^{+0.05}_{-0.05}$	-0.02 ± 0.11
NGDEEP	01233	53.152356	−27.788058	2.31	$8.39^{+0.01}_{-0.02}$	$2.52^{+0.08}_{-0.10}$	$8.23^{+0.06}_{-0.06}$	0.06 ± 0.04
NGDEEP	01237	53.163104	−27.787936	2.22	$7.37^{+0.03}_{-0.02}$	$0.57^{+0.06}_{-0.04}$	$8.08^{+0.05}_{-0.05}$	0.06 ± 0.06
NGDEEP	01255	53.174319	−27.787763	2.07	$8.20^{+0.02}_{-0.02}$	$1.43^{+0.03}_{-0.02}$	$8.09^{+0.06}_{-0.05}$	-0.02 ± 0.03
NGDEEP	01291	53.177252	−27.787383	3.00	$9.81^{+0.01}_{-0.01}$	$4.68^{+0.25}_{-0.24}$	$8.41^{+0.06}_{-0.06}$	-0.10 ± 0.04
NGDEEP	01346	53.180819	−27.787398	1.94	$8.93^{+0.04}_{-0.02}$	$2.54^{+0.18}_{-0.17}$	$8.39^{+0.03}_{-0.04}$	0.05 ± 0.02
NGDEEP	01385	53.180809	−27.786328	2.69	$8.76^{+0.01}_{-0.01}$	$23.20^{+0.43}_{-0.64}$	$8.13^{+0.04}_{-0.04}$	-0.09 ± 0.01
NGDEEP	01482	53.169359	−27.785060	1.90	$6.87^{+0.01}_{-0.01}$	$0.75^{+0.03}_{-0.02}$	$7.71^{+0.08}_{-0.10}$	0.07 ± 0.05
NGDEEP	01489	53.149093	−27.785142	2.06	$9.07^{+0.02}_{-0.02}$	$2.79^{+0.20}_{-0.20}$	$8.39^{+0.05}_{-0.05}$	-0.04 ± 0.02
NGDEEP	01494	53.178699	−27.785025	2.11	$8.70^{+0.03}_{-0.02}$	$2.95^{+0.20}_{-0.18}$	$8.38^{+0.05}_{-0.05}$	-0.00 ± 0.02
NGDEEP	01544	53.142345	−27.784448	2.94	$8.04^{+0.02}_{-0.02}$	$8.36^{+0.25}_{-0.26}$	$8.09^{+0.10}_{-0.12}$	0.01 ± 0.03
NGDEEP	01583	53.144136	−27.783940	2.81	$8.48^{+0.05}_{-0.03}$	$2.28^{+0.29}_{-0.24}$	$8.26^{+0.09}_{-0.10}$	-0.00 ± 0.08
NGDEEP	01633	53.157671	−27.783403	2.63	$8.17^{+0.02}_{-0.02}$	$2.59^{+0.17}_{-0.12}$	$8.22^{+0.10}_{-0.12}$	0.00 ± 0.06
NGDEEP	01634	53.182599	−27.783504	2.07	$8.47^{+0.02}_{-0.02}$	$1.63^{+0.10}_{-0.10}$	$8.28^{+0.05}_{-0.05}$	0.01 ± 0.03
NGDEEP	01635	53.182278	−27.783395	2.07	$8.54^{+0.02}_{-0.02}$	$3.45^{+0.13}_{-0.13}$	$8.33^{+0.04}_{-0.05}$	-0.00 ± 0.02
NGDEEP	01641	53.167902	−27.783264	2.08	$7.88^{+0.03}_{-0.03}$	$0.70^{+0.03}_{-0.03}$	$8.20^{+0.06}_{-0.06}$	-0.04 ± 0.08
NGDEEP	01751	53.167504	−27.781909	2.07	$8.12^{+0.01}_{-0.01}$	$2.36^{+0.26}_{-0.23}$	$7.87^{+0.05}_{-0.04}$	0.01 ± 0.03

Table 4
(Continued)

Field	ID	R.A. (deg)	Decl. (deg)	Redshift	$\log(M_*/M_\odot)$	SFR ($M_\odot$ yr^{-1})	$12 + \log(\mathrm{O/H})$	$\nabla_r \log(\mathrm{O/H})$ (dex kpc^{-1})
NGDEEP	01774	53.165166	−27.781715	2.22	$8.69^{+0.01}_{-0.01}$	$5.49^{+0.34}_{-0.33}$	$8.15^{+0.05}_{-0.05}$	0.08 ± 0.01
NGDEEP	01786	53.151504	−27.781429	2.30	$8.41^{+0.02}_{-0.02}$	$3.05^{+0.15}_{-0.11}$	$8.12^{+0.05}_{-0.05}$	−0.07 ± 0.04
NGDEEP	01865	53.154815	−27.780452	2.57	$7.53^{+0.02}_{-0.02}$	$2.19^{+0.26}_{-0.23}$	$8.10^{+0.12}_{-0.11}$	−0.12 ± 0.09
NGDEEP	01874	53.157122	−27.780335	3.01	$7.94^{+0.02}_{-0.01}$	$3.63^{+0.11}_{-0.14}$	$8.05^{+0.09}_{-0.07}$	−0.05 ± 0.08
NGDEEP	01893	53.152903	−27.780194	1.85	$9.06^{+0.02}_{-0.02}$	$3.20^{+0.12}_{-0.13}$	$8.39^{+0.04}_{-0.04}$	0.02 ± 0.02
NGDEEP	01918	53.150506	−27.779833	3.07	$8.14^{+0.03}_{-0.03}$	$1.93^{+0.23}_{-0.16}$	$8.08^{+0.10}_{-0.10}$	−0.01 ± 0.09
NGDEEP	01966	53.157887	−27.779274	1.84	$8.62^{+0.04}_{-0.03}$	$0.61^{+0.05}_{-0.06}$	$8.32^{+0.05}_{-0.06}$	−0.10 ± 0.08
NGDEEP	01971	53.163741	−27.779153	3.27	$7.60^{+0.02}_{-0.03}$	$4.04^{+0.22}_{-0.26}$	$8.02^{+0.08}_{-0.08}$	−0.08 ± 0.03
NGDEEP	01986	53.176312	−27.778978	3.19	$9.15^{+0.01}_{-0.01}$	$20.62^{+0.71}_{-0.68}$	$8.39^{+0.05}_{-0.06}$	−0.10 ± 0.02
NGDEEP	02030	53.148379	−27.778614	3.32	$7.59^{+0.01}_{-0.01}$	$3.89^{+0.08}_{-0.11}$	$7.84^{+0.07}_{-0.06}$	0.22 ± 0.05
NGDEEP	02033	53.146435	−27.778375	1.86	$7.92^{+0.01}_{-0.01}$	$5.18^{+0.32}_{-0.24}$	$7.92^{+0.04}_{-0.03}$	0.06 ± 0.02
NGDEEP	02059	53.146068	−27.778118	1.86	$7.16^{+0.03}_{-0.04}$	$0.40^{+0.07}_{-0.05}$	$8.01^{+0.09}_{-0.12}$	0.02 ± 0.13
NGDEEP	02167	53.159582	−27.776803	3.43	$7.99^{+0.01}_{-0.01}$	$9.84^{+0.25}_{-0.23}$	$7.96^{+0.05}_{-0.04}$	−0.05 ± 0.04
NGDEEP	02180	53.144354	−27.776543	2.32	$8.06^{+0.02}_{-0.02}$	$1.56^{+0.07}_{-0.09}$	$8.14^{+0.08}_{-0.08}$	−0.15 ± 0.08
NGDEEP	02297	53.139253	−27.774915	1.84	$7.83^{+0.03}_{-0.04}$	$1.20^{+0.03}_{-0.04}$	$8.06^{+0.05}_{-0.04}$	−0.07 ± 0.03
NGDEEP	02304	53.167609	−27.774783	3.06	$8.04^{+0.02}_{-0.03}$	$1.93^{+0.19}_{-0.17}$	$8.08^{+0.11}_{-0.09}$	−0.02 ± 0.06
NGDEEP	02396	53.173541	−27.773773	1.88	$8.48^{+0.02}_{-0.03}$	$0.80^{+0.08}_{-0.06}$	$8.39^{+0.04}_{-0.05}$	0.05 ± 0.07
NGDEEP	02441	53.174137	−27.773047	3.44	$8.98^{+0.03}_{-0.03}$	$8.89^{+0.39}_{-0.32}$	$8.27^{+0.06}_{-0.05}$	−0.22 ± 0.04
NGDEEP	02446	53.148715	−27.773010	3.31	$9.20^{+0.01}_{-0.01}$	$10.61^{+0.30}_{-0.27}$	$8.28^{+0.06}_{-0.06}$	−0.05 ± 0.06
NGDEEP	02458	53.164201	−27.772965	2.84	$9.14^{+0.03}_{-0.02}$	$12.63^{+0.43}_{-0.38}$	$8.26^{+0.04}_{-0.04}$	0.04 ± 0.03
NGDEEP	02477	53.169766	−27.772581	3.32	$8.74^{+0.02}_{-0.03}$	$2.59^{+0.15}_{-0.12}$	$8.06^{+0.10}_{-0.09}$	0.06 ± 0.07
NGDEEP	02482	53.152918	−27.772575	1.84	$8.70^{+0.01}_{-0.01}$	$9.12^{+0.21}_{-0.46}$	$8.15^{+0.06}_{-0.05}$	0.09 ± 0.01
NGDEEP	02511	53.138215	−27.772110	2.81	$8.38^{+0.01}_{-0.01}$	$14.78^{+0.22}_{-0.19}$	$8.05^{+0.08}_{-0.05}$	0.05 ± 0.02
NGDEEP	02558	53.154498	−27.771493	2.22	$9.39^{+0.02}_{-0.01}$	$32.22^{+0.49}_{-1.05}$	$8.39^{+0.06}_{-0.10}$	0.08 ± 0.01
NGDEEP	02564	53.140975	−27.771279	2.22	$7.58^{+0.03}_{-0.03}$	$0.78^{+0.05}_{-0.04}$	$7.89^{+0.06}_{-0.06}$	−0.08 ± 0.11
NGDEEP	02578	53.144689	−27.771181	3.07	$9.25^{+0.02}_{-0.02}$	$22.67^{+0.41}_{-0.43}$	$8.27^{+0.03}_{-0.03}$	−0.03 ± 0.01
NGDEEP	02581	53.160020	−27.771086	1.84	$8.07^{+0.02}_{-0.02}$	$1.64^{+0.06}_{-0.06}$	$8.30^{+0.06}_{-0.06}$	0.05 ± 0.03
NGDEEP	02610	53.161804	−27.770720	3.32	$8.52^{+0.04}_{-0.03}$	$14.16^{+0.42}_{-0.45}$	$8.22^{+0.08}_{-0.08}$	−0.02 ± 0.02
NGDEEP	02649	53.152318	−27.770169	1.84	$9.45^{+0.02}_{-0.02}$	$10.07^{+0.35}_{-0.36}$	$8.35^{+0.04}_{-0.04}$	0.02 ± 0.01
NGDEEP	02685	53.173339	−27.769147	2.85	$8.80^{+0.02}_{-0.02}$	$1.95^{+0.11}_{-0.12}$	$8.14^{+0.08}_{-0.08}$	0.01 ± 0.10
NGDEEP	02724	53.169952	−27.768432	3.09	$9.01^{+0.01}_{-0.01}$	$5.13^{+0.12}_{-0.11}$	$8.14^{+0.08}_{-0.07}$	0.00 ± 0.05
NGDEEP	02743	53.165934	−27.767980	3.43	$8.06^{+0.02}_{-0.02}$	$9.06^{+0.81}_{-0.47}$	$8.08^{+0.07}_{-0.07}$	−0.06 ± 0.04
NGDEEP	02745	53.156388	−27.767910	2.02	$8.03^{+0.01}_{-0.01}$	$1.61^{+0.06}_{-0.07}$	$7.91^{+0.08}_{-0.08}$	0.04 ± 0.06
NGDEEP	02770	53.153851	−27.767378	2.31	$9.17^{+0.00}_{-0.01}$	$9.82^{+0.23}_{-0.13}$	$8.42^{+0.04}_{-0.04}$	0.11 ± 0.01
NGDEEP	02798	53.171741	−27.766728	3.19	$8.75^{+0.02}_{-0.02}$	$2.78^{+0.27}_{-0.21}$	$8.12^{+0.11}_{-0.12}$	0.01 ± 0.13
NGDEEP	02823	53.160498	−27.766138	2.67	$8.07^{+0.02}_{-0.02}$	$2.73^{+0.24}_{-0.21}$	$8.14^{+0.09}_{-0.10}$	−0.10 ± 0.09
NGDEEP	02897	53.172172	−27.763745	2.21	$8.21^{+0.03}_{-0.02}$	$4.65^{+0.26}_{-0.32}$	$7.96^{+0.04}_{-0.03}$	−0.03 ± 0.02

(This table is available in machine-readable form in the online article.)

Appendix M
Spectra and Metallicity Gradients of Individual Galaxies

Here we provide figures of spectra and metallicity gradient measurements for all individual galaxies in our sample, including six galaxies in the ASPIRE sample (Figures 23 and 24) and 88 galaxies in the NGDEEP sample (Figure 25).

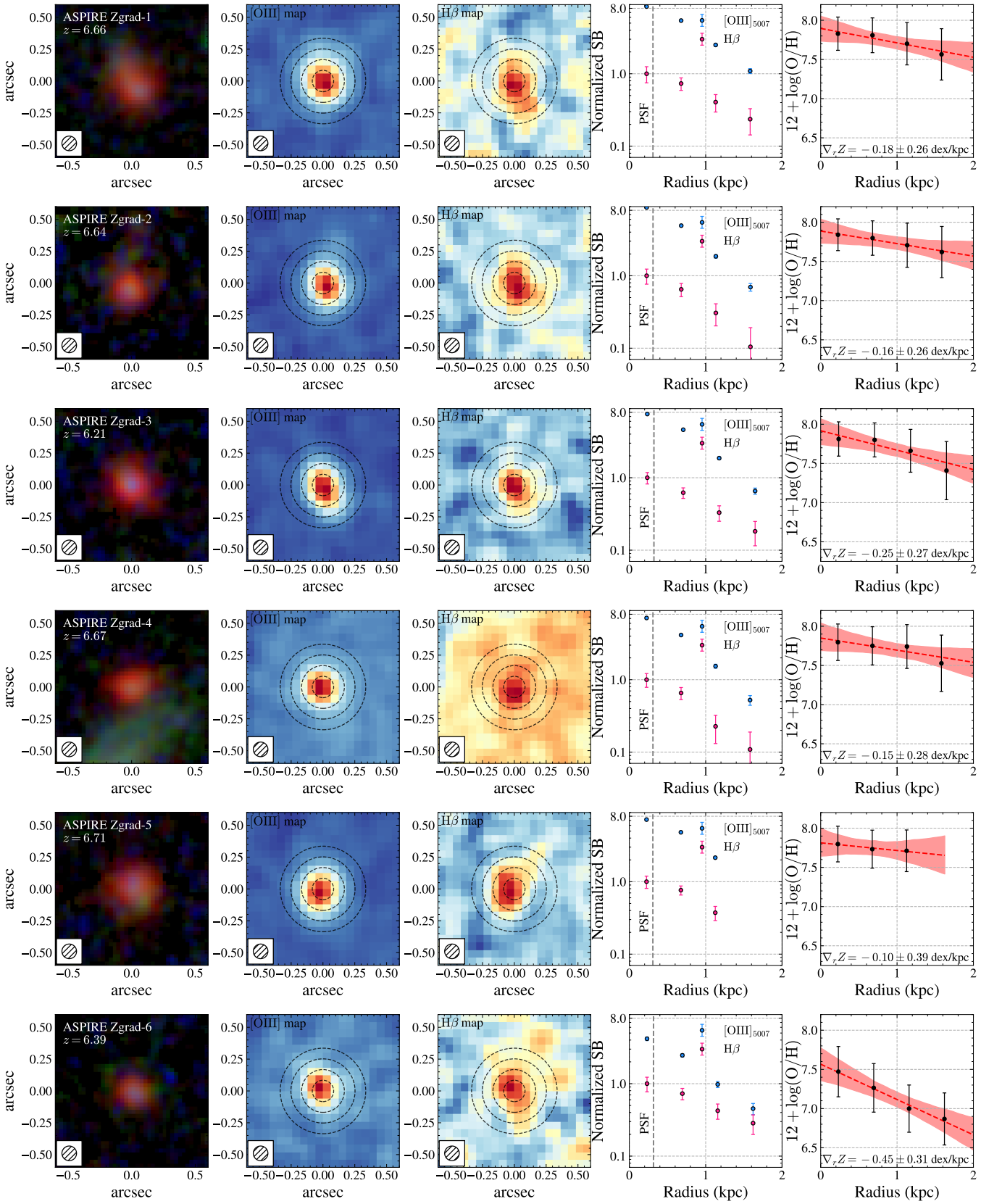


Figure 23. Metallicity gradient measurements of individual galaxies in ASPIRE. The rightmost four panels are the same as Figure 1, but for individual galaxies with the highest SNR. The first panel shows the false-color JWST NIRCam image (with F115W, F200W, and F356W) centered on each source.

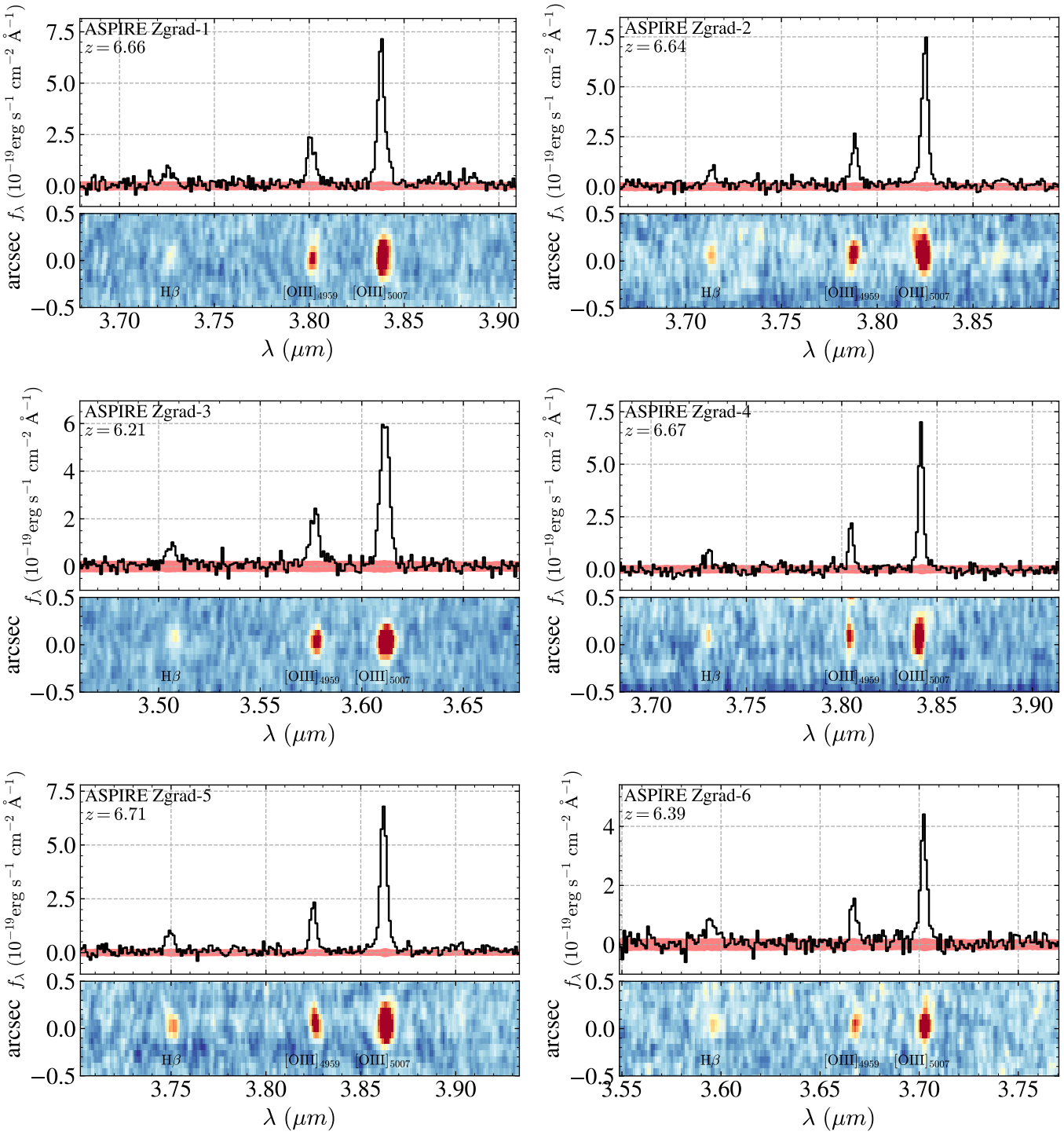


Figure 24. The 1D and 2D NIRCam F356W grism spectra of each source in Figure 23.

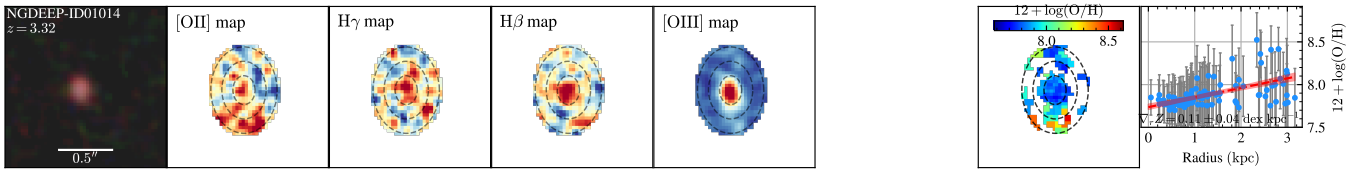


Figure 25. The false-color image, line map, metallicity map, and metallicity gradients for each source in the NGDEEP sample. (The complete figure set (88 images) is available in the online article.)

ORCID iDs

Zihao Li https://orcid.org/0000-0001-5951-459X
Zheng Cai https://orcid.org/0000-0001-8467-6478
Xin Wang https://orcid.org/0000-0002-9373-3865
Zhaozhou Li https://orcid.org/0000-0001-7890-4964
Avishai Dekel https://orcid.org/0000-0003-4174-0374
Kartick C. Sarkar https://orcid.org/0000-0002-7767-8472
Eduardo Bañados https://orcid.org/0000-0002-2931-7824
Fuyan Bian https://orcid.org/0000-0002-1620-0897
Aklant K. Bhowmick https://orcid.org/0000-0002-7080-2864
Laura Blecha https://orcid.org/0000-0002-2183-1087
Sarah E. I. Bosman https://orcid.org/0000-0001-8582-7012
Jaclyn B. Champagne https://orcid.org/0000-0002-6184-9097
Xiaohui Fan https://orcid.org/0000-0002-6336-3007
Emmet Golden-Marx https://orcid.org/0000-0001-5160-6713
Hyunsung D. Jun https://orcid.org/0000-0003-1470-5901
Mingyu Li https://orcid.org/0000-0001-6251-649X
Xiaojing Lin https://orcid.org/0000-0001-6052-4234
Weizhe Liu https://orcid.org/0000-0003-3762-7344
Fengwu Sun https://orcid.org/0000-0002-4622-6617
Maxime Trebitsch https://orcid.org/0000-0002-6849-5375
Fabian Walter https://orcid.org/0000-0003-4793-7880
Feige Wang https://orcid.org/0000-0002-7633-431X
Yunjing Wu https://orcid.org/0000-0003-0111-8249
Jinyi Yang https://orcid.org/0000-0001-5287-4242
Huanian Zhang https://orcid.org/0000-0002-0123-9246
Shiwu Zhang https://orcid.org/0000-0002-0427-9577
Mingyang Zhuang https://orcid.org/0000-0001-5105-2837
Siwei Zou https://orcid.org/0000-0002-3983-6484

References

Acharyya, A., Peeples, M. S., Tumlinson, J., et al. 2025, ApJ, 979, 129
Anglés-Alcázar, D., Faucher-Giguère, C.-A., Kereš, D., et al. 2017, MNRAS, 470, 4698
Arribas, S., Perna, M., Rodríguez Del Pino, B., et al. 2024, A&A, 688, A146
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Aung, H., Mandelker, N., Dekel, A., et al. 2024, MNRAS, 532, 2965
Bagley, M. B., Pirzkal, N., Finkelstein, S. L., et al. 2024, ApJL, 965, L6
Baker, W. M., Tacchella, S., Johnson, B. D., et al. 2025, NatAs, 9, 141
Belfiore, F., Maiolino, R., Tremonti, C., et al. 2017, MNRAS, 469, 151
Bellardini, M. A., Wetzel, A., Loebman, S. R., et al. 2021, MNRAS, 505, 4586
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Bertin, E., Schefer, M., Apostolakos, N., et al. 2020, in ASP Conf. Ser. 527, Astronomical Data Analysis Software and Systems XXIX, ed. R. Pizzo et al. (San Francisco, CA: ASP), 461
Bian, F., Kewley, L. J., & Dopita, M. A. 2018, ApJ, 859, 175
Bird, J. C., Loebman, S. R., Weinberg, D. H., et al. 2021, MNRAS, 503, 1815
Birkin, J. E., Hutchison, T. A., Welch, B., et al. 2023, ApJ, 958, 64
Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349
Boissier, S., & Prantzos, N. 2000, MNRAS, 312, 398
Bradley, L., Sipőcz, B., Robitaille, T., et al. 2024, astropy/photutils, v2.0.2, Zenodo, doi: 10.5281/zenodo.13989456
Brammer, G. 2023, grizli, v1.9.11, Zenodo, doi: 10.5281/zenodo.8370018
Bresolin, F. 2019, MNRAS, 488, 3826
Bushouse, H., Eisenhamer, J., Dencheva, N., et al. 2024, JWST calibration pipeline, v1.14.0, Zenodo, doi: 10.5281/zenodo.10870758
Cai, Z., Fan, X., Yang, Y., et al. 2017, ApJ, 837, 71
Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582
Cappellari, M., & Copin, Y. 2003, MNRAS, 342, 345
Carnall, A. C. 2017, arXiv:1705.05165
Carnall, A. C., McLure, R. J., Dunlop, J. S., et al. 2023, Natur, 619, 716
Carniani, S., Arribas, S., Bunker, A., et al. 2021, JWST Proposal, Cycle 1, 1893
Carton, D., Brinchmann, J., Contini, T., et al. 2018, MNRAS, 478, 4293
Carton, D., Brinchmann, J., Wang, J., et al. 2015, MNRAS, 451, 210
Chabrier, G. 2003, PASP, 115, 763
Chakraborty, P., Sarkar, A., Smith, R., et al. 2025, ApJ, 985, 24
Cheng, Y., Giavalisco, M., Simons, R. C., et al. 2024, ApJ, 964, 94
Chevalier, R. A., & Clegg, A. W. 1985, Natur, 317, 44
Chevallard, J., & Charlot, S. 2016, MNRAS, 462, 1415
Coil, A. L., Aird, J., Reddy, N., et al. 2015, ApJ, 801, 35
Considère, S., Coziol, R., Contini, T., & Davoust, E. 2000, A&A, 356, 89
Correa, C. A., Schaye, J., Wyithe, J. S. B., et al. 2018, MNRAS, 473, 538
Crain, R. A., & van de Voort, F. 2023, ARA&A, 61, 473
Cresci, G., Mannucci, F., Maiolino, R., et al. 2010, Natur, 467, 811
Curti, M., D'Eugenio, F., Carniani, S., et al. 2023, MNRAS, 518, 425
Curti, M., Maiolino, R., Cirasuolo, M., et al. 2020a, MNRAS, 492, 821
Curti, M., Maiolino, R., Curtis-Lake, E., et al. 2024, A&A, 684, A75
Curti, M., Mannucci, F., Cresci, G., & Maiolino, R. 2020b, MNRAS, 491, 944
Daddi, E., Elbaz, D., Walter, F., et al. 2010, ApJL, 714, L118
Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827
Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2
Dekel, A., Birnboim, Y., Engel, G., et al. 2009, Natur, 457, 451
Dekel, A., Mandelker, N., Li, Z., et al. 2025, arXiv:2506.11664
Dekel, A., Sarkar, K. C., Birnboim, Y., Mandelker, N., & Li, Z. 2023, MNRAS, 523, 3201
de la Vega, A., Kassin, S. A., Pacifici, C., et al. 2025, ApJ, 980, 168
de los Reyes, M. A. C., Asali, Y., Wechsler, R., et al. 2025, ApJ, 989, 91,
Edmunds, M. G., & Pagel, B. E. J. 1984, MNRAS, 211, 507
El-Badry, K., Wetzel, A., Geha, M., et al. 2016, ApJ, 820, 131
Erb, D. K. 2008, ApJ, 674, 151
Erb, D. K. 2015, Natur, 523, 169
Evans, I. N. 1986, ApJ, 309, 544
Faucher-Giguère, C.-A., Kereš, D., & Ma, C.-P. 2011, MNRAS, 417, 2982
Ferrara, A., Vallini, L., Pallottini, A., et al. 2019, MNRAS, 489, 1
Foreman-Mackey, D., Conley, A., Meierjurgen Farr, W., et al., 2013a emcee: The MCMC hammer, Astrophysics Source Code Library, ascl: 1303.002
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013b, PASP, 125, 306
Förster Schreiber, N. M., Genzel, R., Bouché, N., et al. 2009, ApJ, 706, 1364
Förster Schreiber, N. M., Renzini, A., Mancini, C., et al. 2018, ApJS, 238, 21
Franchetto, A., Mingozzi, M., Poggianti, B. M., et al. 2021, ApJ, 923, 28
Frankel, N., Sanders, J., Rix, H.-W., Ting, Y.-S., & Ness, M. 2019, ApJ, 884, 99
Fu, J., Kauffmann, G., Huang, M.-l., et al. 2013, MNRAS, 434, 1531
Garcia, A. M., Torrey, P., Bhagwat, A., et al. 2025, ApJ, 989, 147
Garcia, A. M., Torrey, P., Hemler, Z. S., et al. 2023, MNRAS, 519, 4716
Garg, P., Narayanan, D., Sanders, R. L., et al. 2024, ApJ, 972, 113
Gibson, B. K., Pilkington, K., Brook, C. B., Stinson, G. S., & Bailin, J. 2013, A&A, 554, A47
Gillman, S., Tiley, A. L., Swinbank, A. M., et al. 2021, MNRAS, 500, 4229
Goddard, D., Thomas, D., Maraston, C., et al. 2017, MNRAS, 466, 4731
Graf, R. L., Wetzel, A., Bailin, J., & Orr, M. E. 2024, arXiv:2410.21377
Graf, R. L., Wetzel, A., Bellardini, M. A., & Bailin, J. 2025, ApJ, 981, 47
Grasha, K., Chen, Q. H., Battisti, A. J., et al. 2022, ApJ, 929, 118
Harikane, Y., Zhang, Y., Nakajima, K., et al. 2023, ApJ, 959, 39
He, X., Zihao, W., Wang, X., et al. 2025, ApJ, submitted
Heckman, T. M., & Best, P. N. 2023, Galax, 11, 21
Heintz, K. E., Brammer, G. B., Giménez-Arteaga, C., et al. 2023, NatAs, 7, 1517
Hemler, Z. S., Torrey, P., Qi, J., et al. 2021, MNRAS, 506, 3024
Henriques, B. M. B., Yates, R. M., Fu, J., et al. 2020, MNRAS, 491, 5795
Ho, I. T., Kudritzki, R.-P., Kewley, L. J., et al. 2015, MNRAS, 448, 2030
Hopkins, P. F., Cox, T. J., Younger, J. D., & Hernquist, L. 2009, ApJ, 691, 1168
Hopkins, P. F., Gurvich, A. B., Shen, X., et al. 2023, MNRAS, 525, 2241
Isobe, Y., Ouchi, M., Suzuki, A., et al. 2022, ApJ, 925, 111
Iwamoto, K., Brachwitz, F., Nomoto, K., et al. 1999, ApJS, 125, 439
Izotov, Y. I., & Thuan, T. X. 2008, ApJ, 687, 133
Izzard, R. G., Tout, C. A., Karakas, A. I., & Pols, O. R. 2004, MNRAS, 350, 407
Johnson, J. W., Weinberg, D. H., Blanc, G. A., et al. 2024, ApJ, 988, 8
Jones, T., Ellis, R. S., Richard, J., & Jullo, E. 2013, ApJ, 765, 48
Jones, T., Sanders, R., Chen, Y., et al. 2023, ApJL, 951, L17
Jones, T., Wang, X., Schmidt, K. B., et al. 2015, AJ, 149, 107
Ju, M., Wang, X., Jones, T., et al. 2025, ApJL, 978, L39
Juneau, S., Bournaud, F., Charlot, S., et al. 2014, ApJ, 788, 88
Kashiwagi, Y., Inoue, A. K., Isobe, Y., et al. 2021, PASJ, 73, 1631
Kelly, B. C. 2007, ApJ, 665, 1489
Kennicutt, R. C., Jr. 1989, ApJ, 344, 685
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kereš, D., Katz, N., Fardal, M., Davé, R., & Weinberg, D. H. 2009, MNRAS, 395, 160
Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2
Kewley, L. J., & Dopita, M. A. 2002, ApJS, 142, 35
Khoram, A. H., & Belfiore, F. 2025, A&A, 693, A150
Kojima, T., Ouchi, M., Rauch, M., et al. 2020, ApJ, 898, 142

Kormendy, J., & Kennicutt, R. C. J. 2004, ARA&A, 42, 603
Lagos, P., Loubser, S. I., Scott, T. C., et al. 2022, MNRAS, 516, 5487
Lang, M., Holley-Bockelmann, K., & Sinha, M. 2014, ApJL, 790, L33
Langeroodi, D., & Hjorth, J. 2023, arXiv:2307.06336
Lara-López, M. A., Cepa, J., Bongiovanni, A., et al. 2010, A&A, 521, L53
Larson, R. B. 1974, MNRAS, 166, 585
Leethochawalit, N., Jones, T. A., Ellis, R. S., et al. 2016, ApJ, 820, 84
Li, Q., Conselice, C. J., Adams, N., et al. 2024, MNRAS, 531, 617
Li, T., Zhang, H.-X., Lyu, W., et al. 2025, A&A, 698, A208
Li, Z., Dekel, A., Sarkar, K. C., et al. 2024, A&A, 690, A108
Li, Z., Kakiichi, K., Christensen, L., et al. 2025, arXiv:2504.18616
Li, Z., Wang, X., Cai, Z., et al. 2022, ApJL, 929, L8
Lian, J., Bergemann, M., Pillepich, A., Zasowski, G., & Lane, R. R. 2023, NatAs, 7, 951
Lian, J., Thomas, D., Maraston, C., et al. 2018, MNRAS, 476, 3883
Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119
Lyu, C., Wang, E., Zhang, H., et al. 2025, ApJL, 981, L6
Ma, X., Hopkins, P. F., Faucher-Giguère, C.-A., et al. 2016, MNRAS, 456, 2140
Ma, X., Hopkins, P. F., Feldmann, R., et al. 2017, MNRAS, 466, 4780
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Magrini, L., Coccato, L., Stanghellini, L., Casasola, V., & Galli, D. 2016, A&A, 588, A91
Magrini, L., Vílchez, J. M., Mampaso, A., Corradi, R. L. M., & Leisy, P. 2007, A&A, 470, 865
Maiolino, R., Nagao, T., Grazian, A., et al. 2008, A&A, 488, 463
Maiolino, R., Risaliti, G., Signorini, M., et al. 2025, MNRAS, 538, 1921
Maiolino, R., Scholtz, J., Curtis-Lake, E., et al. 2024, A&A, 691, A145
Markov, V., Carniani, S., Vallini, L., et al. 2022, A&A, 663, A172
Markov, V., Gallerani, S., Ferrara, A., et al. 2025, NatAs, 9, 458
Markov, V., Gallerani, S., Pallottini, A., et al. 2023, A&A, 679, A12
Marszewski, A., Sun, G., Faucher-Giguère, C.-A., Hayward, C. C., & Feldmann, R. 2024, ApJL, 967, L41
Martig, M., & Bournaud, F. 2008, MNRAS, 385, L38
Matharu, J., & Brammer, G. 2022, Updated configuration files for JWST NIRISS slitless spectroscopy, v0.0, Zenodo, doi: 10.5281/zenodo.7447545
Matharu, J., Muzzin, A., Sarrouh, G. T. E., et al. 2023, ApJL, 949, L11
Matharu, J., Nelson, E. J., Brammer, G., et al. 2024, A&A, 690, A64
Matthee, J., Mackenzie, R., Simcoe, R. A., et al. 2023, ApJ, 950, 67
McAlpine, S., Helly, J. C., Schaller, M., et al. 2016, A&C, 15, 72
McDermid, R. M., Alatalo, K., Blitz, L., et al. 2015, MNRAS, 448, 3484
Molina, J., Ibar, E., Swinbank, A. M., et al. 2017, MNRAS, 466, 892
Mollá, M., Díaz, Á. I., Cavichia, O., et al. 2019, MNRAS, 482, 3071
Møller, P., Fynbo, J. P. U., Ledoux, C., & Nilsson, K. K. 2013, MNRAS, 430, 2680
Morishita, T., Stiavelli, M., Chary, R.-R., et al. 2024, ApJ, 963, 9
Nagao, T., Maiolino, R., & Marconi, A. 2006, A&A, 459, 85
Nakajima, K., Ouchi, M., Isobe, Y., et al. 2023, ApJS, 269, 33
Nakajima, K., Ouchi, M., Xu, Y., et al. 2022, ApJS, 262, 3
Nelson, D., Springel, V., Pillepich, A., et al. 2019, ComAC, 6, 2
Nelson, E. J., van Dokkum, P. G., Förster Schreiber, N. M., et al. 2016, ApJ, 828, 27
Oesch, P., & Magee, D. 2023, The JWST FRESCO Survey, STScI/MAST, doi:10.17909/GDYC-7G80
Oesch, P. A., Brammer, G., Naidu, R. P., et al. 2023, MNRAS, 525, 2864
Orr, M. E., Fielding, D. B., Hayward, C. C., & Burkhart, B. 2022, ApJ, 932, 88
Pagel, B. E. J., & Edmunds, M. G. 1981, ARA&A, 19, 77
Pallottini, A., Ferrara, A., Gallerani, S., et al. 2022, MNRAS, 513, 5621
Pandya, V., Zhang, H., Huertas-Company, M., et al. 2024, ApJ, 963, 54
Pérez, E., Cid Fernandes, R., González Delgado, R. M., et al. 2013, ApJL, 764, L1
Pérez-Díaz, B., Pérez-Montero, E., Fernández-Ontiveros, J. A., Vílchez, J. M., & Amorín, R. 2024, NatAs, 8, 368
Pilkington, K., Few, C. G., Gibson, B. K., et al. 2012, A&A, 540, A56
Pirzkal, N. 2023, NGDEEP-NIS Data Epoch 1, STScI/MAST, doi:10.17909/02WX-6J29
Poetrodjojo, H., Groves, B., Kewley, L. J., et al. 2018, MNRAS, 479, 5235
Poetrodjojo, H., Groves, B., Kewley, L. J., et al. 2021, MNRAS, 502, 3357
Porter, L. E., Orr, M. E., Burkhart, B., et al. 2022, MNRAS, 515, 3555
Queyrel, J., Contini, T., Kissler-Patig, M., et al. 2012, A&A, 539, A93
Raiteri, C. M., Villata, M., & Navarro, J. F. 1996, A&A, 315, 105
Ratcliffe, B., Khoperskov, S., Minchev, I., et al. 2025, A&A, 698, A267
Reddy, N. A., Topping, M. W., Sanders, R. L., Shapley, A. E., & Brammer, G. 2023, ApJ, 952, 167
Renaud, F., Ratcliffe, B., Minchev, I., et al. 2025, A&A, 694, A56
Roca-Fàbrega, S., Kim, J.-H., Primack, J. R., et al. 2024, ApJ, 968, 125
Rupke, D. S. N., Kewley, L. J., & Barnes, J. E. 2010a, ApJL, 710, L156
Rupke, D. S. N., Kewley, L. J., & Chien, L. H. 2010b, ApJ, 723, 1255
Sánchez-Menguiano, L., Sánchez, S. F., Pérez, I., et al. 2016, A&A, 587, A70
Sánchez-Menguiano, L., Sánchez, S. F., Pérez, I., et al. 2018, A&A, 609, A119
Sanders, R. L., Shapley, A. E., Jones, T., et al. 2023, ApJ, 942, 24
Sanders, R. L., Shapley, A. E., Topping, M. W., Reddy, N. A., & Brammer, G. B. 2024, ApJ, 962, 24
Sandles, L., D'Eugenio, F., Maiolino, R., et al. 2024, A&A, 691, A305
Sarkar, A., Chakraborty, P., Vogelsberger, M., et al. 2025, ApJ, 978, 136
Schmidt, M. 1959, ApJ, 129, 243
Scholte, D., Cullen, F., Carnall, A. C., et al. 2025, MNRAS, 540, 1800
Scholtz, J., Maiolino, R., D'Eugenio, F., et al. 2025, A&A, 697, A175
Schönrich, R., & McMillan, P. J. 2017, MNRAS, 467, 1154
Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, A&A, 575, A74
Sharda, P., Ginzburg, O., Krumholz, M. R., et al. 2024, MNRAS, 528, 2232
Sharda, P., Krumholz, M. R., Wisnioski, E., et al. 2021a, MNRAS, 504, 53
Sharda, P., Krumholz, M. R., Wisnioski, E., et al. 2021b, MNRAS, 502, 5935
Sharda, P., Wisnioski, E., Krumholz, M. R., & Federrath, C. 2021c, MNRAS, 506, 1295
Shibuya, T., Ouchi, M., & Harikane, Y. 2015, ApJS, 219, 15
Shields, G. A. 1990, ARA&A, 28, 525
Shimizu, I., Todoroki, K., Yajima, H., & Nagamine, K. 2019, MNRAS, 484, 2632
Shuntov, M., McCracken, H. J., Gavazzi, R., et al. 2022, A&A, 664, A61
Simons, R. C., Kassin, S. A., Weiner, B. J., et al. 2017, ApJ, 843, 46
Simons, R. C., Papovich, C., Momcheva, I., et al. 2021, ApJ, 923, 203
Spolaor, M., Proctor, R. N., Forbes, D. A., & Couch, W. J. 2009, ApJL, 691, L138
Springel, V., & Hernquist, L. 2003, MNRAS, 339, 289
Stott, J. P., Sobral, D., Smail, I., et al. 2013, MNRAS, 430, 1158
Strom, A. L., Steidel, C. C., Rudie, G. C., Trainor, R. F., & Pettini, M. 2018, ApJ, 868, 117
Sun, F., Egami, E., Pirzkal, N., et al. 2023, ApJ, 953, 53
Sun, X., Wang, X., Ma, X., et al. 2025, ApJ, 986, 179
Swinbank, A. M., Sobral, D., Smail, I., et al. 2012, MNRAS, 426, 935
Tacchella, S., Johnson, B. D., Robertson, B. E., et al. 2023, MNRAS, 522, 6236
Tang, M., Stark, D. P., Chen, Z., et al. 2023, MNRAS, 526, 1657
Tapia-Contreras, B., Tissera, P. B., Sillero, E., et al. 2025, A&A, 700, A69
Thielemann, F. K., Nomoto, K., & Yokoi, K. 1986, A&A, 158, 17
Thorp, M. D., Ellison, S. L., Simard, L., Sánchez, S. F., & Antonio, B. 2019, MNRAS, 482, L55
Tinsley, B. M. 1980, FCPh, 5, 287
Tissera, P. B., Machado, R. E. G., Sanchez-Blazquez, P., et al. 2016, A&A, 592, A93
Tissera, P. B., Rosas-Guevara, Y., Bower, R. G., et al. 2019, MNRAS, 482, 2208
Tissera, P. B., Rosas-Guevara, Y., Sillero, E., et al. 2022, MNRAS, 511, 1667
Torrey, P., Cox, T. J., Kewley, L., & Hernquist, L. 2012, ApJ, 746, 108
Trapp, C. W., Kereš, D., Chan, T. K., et al. 2022, MNRAS, 509, 4149
Troncoso, P., Maiolino, R., Sommariva, V., et al. 2014, A&A, 563, A58
Trump, J. R., Arrabal Haro, P., Simons, R. C., et al. 2023, ApJ, 945, 35
Trump, J. R., Weiner, B. J., Scarlata, C., et al. 2011, ApJ, 743, 144
Valentino, F., Brammer, G., Gould, K. M. L., et al. 2023, ApJ, 947, 20
Vallini, L., Ferrara, A., Pallottini, A., Carniani, S., & Gallerani, S. 2021, MNRAS, 505, 5543
Vallini, L., Witstok, J., Sommovigo, L., et al. 2024, MNRAS, 527, 10
van der Wel, A., Straughn, A. N., Rix, H. W., et al. 2011, ApJ, 742, 111
van de Voort, F., & Schaye, J. 2012, MNRAS, 423, 2991
van de Voort, F., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 2782
van Zee, L., Salzer, J. J., Haynes, M. P., O'Donoghue, A. A., & Balonek, T. J. 1998, AJ, 116, 2805
Venturi, G., Carniani, S., Parlanti, E., et al. 2024, A&A, 691, A19
Vila-Costas, M. B., & Edmunds, M. G. 1992, MNRAS, 259, 121
Vilchez, J. M., Pagel, B. E. J., Diaz, A. I., Terlevich, E., & Edmunds, M. G. 1988, MNRAS, 235, 633
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Wang, E., & Lilly, S. J. 2022, ApJ, 929, 95
Wang, F. 2023, A SPectroscopic survey of biased halos In the Reionization Era (ASPIRE): A JWST Quasar Legacy Survey, STScI/MAST, doi:10.17909/VT74-KD84

Wang, F., Yang, J., Hennawi, J. F., et al. 2023, ApJL, 951, L4
Wang, X., Jones, T., Vulcani, B., et al. 2022a, ApJL, 938, L16
Wang, X., Jones, T. A., Treu, T., et al. 2017, ApJ, 837, 89
Wang, X., Jones, T. A., Treu, T., et al. 2019, ApJ, 882, 94
Wang, X., Jones, T. A., Treu, T., et al. 2020, ApJ, 900, 183
Wang, X., Li, Z., Cai, Z., et al. 2022b, ApJ, 926, 70
Weingartner, J. C., & Draine, B. T. 2001, ApJ, 548, 296
Whitler, L., Stark, D. P., Endsley, R., et al. 2023, MNRAS, 519, 5859
Woods, R. M., Wadsley, J., Couchman, H. M. P., Stinson, G., & Shen, S. 2014, MNRAS, 442, 732
Woosley, S. E., & Weaver, T. A. 1995, ApJS, 101, 181
Wuyts, E., Wisnioski, E., Fossati, M., et al. 2016, ApJ, 827, 74
Yabe, K., Ohta, K., Akiyama, M., et al. 2015, ApJ, 798, 45
Yang, J., Wang, F., Fan, X., et al. 2023, ApJL, 951, L5
Yu, S., Bullock, J. S., Gurvich, A. B., et al. 2023, MNRAS, 523, 6220
Zhang, S., Cai, Z., Xu, D., et al. 2023, Sci, 380, 494
Zhuang, Y., Leaman, R., van de Ven, G., et al. 2019, MNRAS, 483, 1862