Title

DESI Emission-line Galaxies: Unveiling the Diversity of [O ii] Profiles and Its Links to Star Formation and Morphology

Journal

The Astrophysical Journal, 977(2)

ISSN

0004-637X

Authors

Lan, Ting-Wen
Prochaska, J Xavier
Moustakas, John
et al.

Publication Date

2024-12-01

DOI

10.3847/1538-4357/ad8f33

The Astrophysical Journal, 977:225 (19pp), 2024 December 20
https://doi.org/10.3847/1538-4357/ad8f33

DESI Emission-line Galaxies: Unveiling the Diversity of [O II] Profiles and Its Links to Star Formation and Morphology

Ting-Wen Lan[1], J. Xavier Prochaska[2,3,4], John Moustakas[5], Małgorzata Siudek[6,7], J. Aguilar[8], S. Ahlen[9], D. Bianchi[10], D. Brooks[11], T. Claybaugh[8], S. Cole[12], K. Dawson[13], A. de la Macorra[14], P. Doel[11], J. E. Forero-Romero[15,16], E. Gaztañaga[6,17,18], S. Gontcho A Gontcho[8], G. Gutierrez[19], J. Guy[8], K. Honscheid[20,21,22], R. Kehoe[23], T. Kisner[8], A. Lambert[8], M. Landriau[8], A. Meisner[24], R. Miquel[25,26], A. Muñoz-Gutiérrez[14], J. A. Newman[27], C. Poppett[8,28,29], F. Prada[30], G. Rossi[31], E. Sanchez[32], M. Schubnell[33,34], H. Seo[35], D. Sprayberry[24], G. Tarlé[34], B. A. Weaver[24], and H. Zou[36]

[1] Graduate Institute of Astrophysics and Department of Physics, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 10617, Taiwan; twlan@ntu.edu.tw
[2] Department of Astronomy and Astrophysics, University of California, Santa Cruz, 1156 High Street, Santa Cruz, CA 95065, USA
[3] Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), WPI, The University of Tokyo Institutes for Advanced Study (UTIAS), The University of Tokyo, Kashiwa, Chiba, Kashiwa 277-8583, Japan
[4] Division of Science, National Astronomical Observatory of Japan, 2-21-1, Osawa, Mitaka, Tokyo 181-8588, Japan
[5] Department of Physics and Astronomy, Siena College, 515 Loudon Road, Loudonville, NY 12211, USA
[6] Institute of Space Sciences, ICE-CSIC, Campus UAB, Carrer de Can Magrans s/n, 08913 Bellaterra, Barcelona, Spain
[7] Instituto Astrofísica de Canarias, C/ Via Lactea s/n, E38205 La Laguna, Spain
[8] Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[9] Physics Dept., Boston University, 590 Commonwealth Avenue, Boston, MA 02215, USA
[10] Dipartimento di Fisica "Aldo Pontremoli," Università degli Studi di Milano, Via Celoria 16, I-20133 Milano, Italy
[11] Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[12] Institute for Computational Cosmology, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK
[13] Department of Physics and Astronomy, The University of Utah, 115 South 1400 East, Salt Lake City, UT 84112, USA
[14] Instituto de Física, Universidad Nacional Autónoma de México, Cd. de México C.P. 04510, Mexico
[15] Departamento de Física, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio Ip, CP 111711, Bogotá, Colombia
[16] Observatorio Astronómico, Universidad de los Andes, Cra. 1 No. 18A-10, Edificio H, CP 111711 Bogotá, Colombia
[17] Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain
[18] Institute of Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building, Portsmouth, PO1 3FX, UK
[19] Fermi National Accelerator Laboratory, PO Box 500, Batavia, IL 60510, USA
[20] Center for Cosmology and AstroParticle Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA
[21] Department of Physics, The Ohio State University, 191 West Woodruff Avenue, Columbus, OH 43210, USA
[22] The Ohio State University, Columbus, 43210 OH, USA
[23] Department of Physics, Southern Methodist University, 3215 Daniel Avenue, Dallas, TX 75275, USA
[24] NSF NOIRLab, 950 N. Cherry Avenue, Tucson, AZ 85719, USA
[25] Institució Catalana de Recerca i Estudis Avançats, Passeig de Lluís Companys, 23, 08010 Barcelona, Spain
[26] Institut de Física d'Altes Energies (IFAE), The Barcelona Institute of Science and Technology, Campus UAB, 08193 Bellaterra Barcelona, Spain
[27] Department of Physics & Astronomy and Pittsburgh Particle Physics, Astrophysics, and Cosmology Center (PITT PACC), University of Pittsburgh, 3941 O'Hara Street, Pittsburgh, PA 15260, USA
[28] Space Sciences Laboratory, University of California, Berkeley, 7 Gauss Way, Berkeley, CA 94720, USA
[29] University of California, Berkeley, 110 Sproul Hall #5800 Berkeley, CA 94720, USA
[30] Instituto de Astrofísica de Andalucía (CSIC), Glorieta de la Astronomía, s/n, E-18008 Granada, Spain
[31] Department of Physics and Astronomy, Sejong University, Seoul, 143-747, Republic of Korea
[32] CIEMAT, Avenida Complutense 40, E-28040 Madrid, Spain
[33] Department of Physics, University of Michigan, Ann Arbor, MI 48109, USA
[34] University of Michigan, Ann Arbor, MI 48109, USA
[35] Department of Physics & Astronomy, Ohio University, Athens, OH 45701, USA
[36] National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing, 100012, People's Republic of China

Received 2024 August 10; revised 2024 September 30; accepted 2024 November 4; published 2024 December 16

Abstract

We study the [O II] profiles of emission-line galaxies (ELGs) from the Early Data Release of the Dark Energy Spectroscopic Instrument (DESI). To this end, we decompose and classify the shape of [O II] profiles with the first two eigenspectra derived from principal component analysis. Our results show that DESI ELGs have diverse line profiles, which can be categorized into three main types: (1) narrow lines with a median width of $\sim$50 km s^{-1}, (2) broad lines with a median width of $\sim$80 km s^{-1}, and (3) two redshift systems with a median velocity separation of $\sim$150 km s^{-1}, i.e., double-peak galaxies. To investigate the connections between the line profiles and galaxy properties, we utilize the information from the COSMOS data set and compare the properties of ELGs, including star formation rate (SFR) and galaxy morphology, with the average properties of reference star-forming galaxies with similar stellar mass, sizes, and redshifts. Our findings show that, on average, DESI ELGs have a higher SFR and more asymmetrical/disturbed morphology than the reference galaxies. Moreover, we uncover a relationship between the line profiles, the excess SFR, and the excess asymmetry parameter, showing that DESI ELGs with broader [O II] line profiles have more disturbed morphology and higher SFR than the reference star-forming galaxies. Finally, we discuss possible physical mechanisms giving rise to the observed relationship and the implications of our findings on the galaxy clustering measurements, including the halo occupation distribution modeling of DESI ELGs and the observed excess velocity dispersion of the satellite ELGs.

Unified Astronomy Thesaurus concepts: Galaxy spectroscopy (2171); Emission line galaxies (459); Principal component analysis (1944); Galaxy mergers (608)

1. Introduction

Galaxies that produce strong emission lines from the H II star-forming regions, i.e., emission-line galaxies (ELGs), have become one of the key observational targets for large cosmological surveys. Via the strong emission-line features, one can detect and measure the redshifts of ELGs across a wide range of cosmic time with accessible observational resources and use those galaxies as tracers of the 3D large-scale structure of the Universe. Previous cosmological surveys, including the WiggleZ Dark Energy Survey (M. J. Drinkwater et al. 2010) and the Extended Baryon Oscillation Spectroscopic Survey in the Sloan Digital Sky Survey-IV (SDSS-IV eBOSS; K. S. Dawson et al. 2016) projects, have demonstrated that cosmological parameters, such as the expansion rate of the Universe and the dark energy equation of state, can be constrained via the baryonic acoustic oscillations (e.g., D. J. Eisenstein et al. 2005), signals detected from the ELG clustering measurements (e.g., C. Blake et al. 2011; A. Raichoor et al. 2021). The success of these programs has motivated the ongoing and upcoming surveys, including the Dark Energy Spectroscopic Instrument (DESI; M. Levi et al. 2013; DESI Collaboration et al. 2016a, 2016b, 2022), the Prime-Focus Spectrograph (PFS; M. Takada et al. 2014) and Euclid (Euclid Collaboration et al. 2022) to select ELGs as one of the main targets and collect tens of millions of spectroscopic measurements of ELGs for their cosmological programs.

Among the emission lines produced by H II regions, the [O II] $\lambda\lambda$3726, 3729 doublet has been considered as one of the key transitions for ground-based spectroscopic measurements (e.g., J. Comparat et al. 2015) owing to its detectability in optical and near-infrared wavelengths from the local Universe to redshift $\sim$2, the cosmic epoch at the peak of the star formation rate (SFR) density (see P. Madau & M. Dickinson 2014, for a review). The [O II]$\lambda\lambda$3726, 3729 doublet can be used to not only pinpoint the redshifts of galaxies but also obtain information on galaxy physical properties, including SFR (line luminosity), electron density (doublet line ratio), and gas kinematics (line width; e.g., R. C. Kennicutt 1998; J. Moustakas et al. 2006; D. E. Osterbrock & G. J. Ferland 2006; S. A. Kassin et al. 2012; L. J. Kewley et al. 2019). However, extracting such information from spectra requires a spectral resolution of $R \sim 3300$ to resolve the [O II]$\lambda\lambda$3726, 3729 doublet feature (J. Comparat et al. 2013). This limits the exploration and characterization of the [O II] line properties of ELGs with the spectra from previous cosmological surveys, such as WiggleZ ($R \sim 1300$, M. J. Drinkwater et al. 2010) and SDSS-IV eBOSS ($R \sim 2000$, K. S. Dawson et al. 2016). On the other hand, the DESI survey has the required spectral resolution (DESI Collaboration et al. 2022) and will compile a large ELG catalog with resolved [O II] lines from redshift 0.6 to redshift 1.6, a redshift region rarely probed to date. With this large spectroscopic data set, it is crucial to understand the underlying galaxy population selected as ELGs in the DESI survey and characterize their physical properties to fully utilize this big data set and maximize the scientific returns for cosmology and galaxy science.

In this work, we explore diversity in the DESI ELG galaxy population based on their [O II] emission-line profiles and investigate the relationships between the line profiles and the physical properties of the galaxies. To this end, we make use of a principal component analysis (PCA; see I. Jolliffe 2002, for a review), a dimensional reduction technique, to obtain the key eigenspectra, which carry physical information of the line profiles. We use them to describe line profiles and classify galaxies. We then obtain the physical properties, such as stellar mass (M_*), SFR, and morphology parameters, of a subset of ELGs in the COSMOS field (N. Scoville et al. 2007) and examine the connections between the line profiles and galaxy properties. This enables us to reveal possible underlying mechanisms driving the [O II] line profiles.

The structure of this paper is as follows. The data sets used in this analysis are described in Section 2. We present the PCA results in Section 3 and explore the relations between the line profiles and galaxy properties in Section 4. We discuss our results and their implications in Section 5, and summarize in Section 6. Throughout this paper, we adopt a flat ΛCDM cosmology with $h = 0.694$ and $\Omega_m = 0.287$ (WMAP9; G. Hinshaw et al. 2013). We use AB magnitudes corrected with Galactic extinction (D. J. Schlegel et al. 1998).

2. Data Sets

2.1. DESI ELG Spectra

The DESI survey (DESI Collaboration et al. 2022) is a Stage IV project, primarily designed for probing the cosmological parameters of the Universe (M. Levi et al. 2013; DESI Collaboration et al. 2016a, 2024a). DESI consists of bright-time and dark-time science targets (E. Chaussidon et al. 2023; A. P. Cooper et al. 2023; C. Hahn et al. 2023; A. Raichoor et al. 2023; R. Zhou et al. 2023) with [O II] ELGs (A. Raichoor et al. 2023) being one of the key targets in the program. In order to obtain the desired redshift range of DESI ELGs, a specific target selection scheme is developed for the sources detected in the DESI Legacy Imaging Surveys (A. Dey et al. 2019), which include the g, r, z bands and near-infrared images from the Wide-field Infrared Survey Explorer (WISE) satellite (E. L. Wright et al. 2010), via the `Tractor` algorithm (D. Lang et al. 2016) with large galaxies masked (J. Moustakas et al. 2023). To ensure the efficiency of DESI operations, dedicated pipelines are developed for target selection (A. D. Myers et al. 2023), survey operations (E. F. Schlafly et al. 2023), and fiber assignment (A. Raichoor et al. 2024, in preparation).

Spectroscopic data is collected by the DESI instrumentation installed on the 4 m Mayall Telescope at Kitt Peak National Observatory (DESI Collaboration et al. 2022; T. N. Miller et al. 2023; J. H. Silber et al. 2023). The instrument has a $3^\circ\!.2$ diameter field of view covered with 5020 robotic fiber positioners. The wavelength coverage ranges from 3600 to 9800 Å, with spectral resolution from ∼2000 in the shortest wavelength to ∼5000 in the longest wavelength. The raw spectra are processed, reduced, and calibrated by an automatic pipeline (J. Guy et al. 2023). The redshifts of sources are then obtained via an algorithm called `Redrock` (S. Bailey et al. 2024, in preparation; see also A. Brodzeller et al. 2023; A. Anand et al. 2024), with the redshift accuracy and precision for all types of extragalactic sources, including ELGs, being tested and validated against redshifts obtained from spectra with long exposure times and confirmed via visual inspection (D. M. Alexander et al. 2023; T.-W. Lan et al. 2023).

In this analysis, we make use of the DESI spectra of ELGs observed during the Survey Validation (SV) phase (DESI Collaboration et al. 2024a) to explore the [O II]$\lambda\lambda$3726, 3729 emission-line profiles. The DESI SV campaign consists of two key phases: the SV1 and the One-Percent Survey. The SV1 observations include targets selected from a wider parameter space in order to test and finalize the target selection scheme of the DESI main survey. The One-Percent Survey then adopted the finalized target selection scheme and obtained the observations that cover a sky area ∼of 140 deg^2, about 1% of the DESI final footprint (DESI Collaboration et al. 2024a).

We focus on the ELGs observed in the One-Percent Survey to inform the properties of ELGs of the DESI main survey. The ELGs are selected within a color–magnitude space defined with $g-r$, $r-z$ colors, g-band magnitude, and g-band fiber magnitude for preferentially including star-forming galaxies at $1.1 < z < 1.63$ (A. Raichoor et al. 2023).

We first make use of the data from the Early Data Release (DESI Collaboration et al. 2024b) which includes ∼280,000 spectra of ELGs with robust redshift measurements that pass the redshift quality selection for ELGs, $\log_{10}\mathrm{SNR}(F_{\mathrm{O\,II}}) > 0.9 - 0.2 \times \log_{10}\Delta\chi^2$, where $\mathrm{SNR}(F_{\mathrm{O\,II}})$ is the signal-to-noise ratio of the [O II] emission lines measured via a double-Gaussian fitting (A. Raichoor et al. 2023) and $\Delta\chi^2$ is the difference between the χ^2 values from the second best-fit model and the first best-fit model provided by `Redrock` (S. Bailey et al. 2024, in preparation). This selection yields an ELG sample with redshift purity (the fraction of ELGs with `Redrock` redshifts and visually inspected redshift differences smaller than ∼1000 km s^{-1}) greater than 99% (T.-W. Lan et al. 2023). We further select ELGs with $0.6 < z < 1.63$ and with spectral types identified by `Redrock` as galaxies. This selection yields a sample of ∼260,000 ELGs. In addition, using line information available at different redshifts, we remove possible active galactic nuclei (AGN) contamination. The selections remove approximately 4% of DESI ELGs. The detailed selections are described in Appendix A. Approximately 250,000 ELGs pass the above selection criteria. We note that this sample includes ELGs with LOw Priority (LOP) and Very-LOw priority (VLO) selections (see A. Raichoor et al. 2023, for details) for the main DESI survey. Approximately 78% of the sample are LOP ELGs, the fiducial ELG sample for DESI clustering measurements.

2.2. The COSMOS Data Sets

To explore the physical properties of DESI ELGs in the context of the whole star-forming galaxy population across redshifts, we utilize the COSMOS2020 catalog (J. R. Weaver et al. 2022), which includes approximately 1.7 million sources with photometric redshifts (z_{photo}) and physical properties of galaxies estimated based on multiwavelength broadband deep imaging data sets. The COSMOS field was covered by the DESI One-Percent Survey. Therefore, we can compile a sample with a few thousand DESI ELGs with properties derived from deep COSMOS data sets.

Catalog Selection. In the COSMOS2020 data release, there are two source catalogs, the `CLASSIC` catalog and the `FARMER` catalog, which are constructed based on the SExtractor (E. Bertin & S. Arnouts 1996) and the `Tractor` (D. Lang et al. 2016) algorithms, respectively for source detection and flux measurements. The primary difference between the two methods is that in the `CLASSIC` catalog, the fluxes of sources are measured with the aperture photometry method, while in the `FARMER` catalog, the fluxes of sources are measured with the forced photometric best-fit models of source light distribution. As shown in J. R. Weaver et al. (2022), for sources brighter than 24 mag in the Hyper Suprime-Cam z band, these two methods yield consistent measurements with magnitude deviation lower than 0.05 magnitude. However, the `FARMER` catalog only includes sources within the UltraVISTA footprint (H. J. McCracken et al. 2012), which consists of about half the amount of sources in the CLASSIC catalog with the full COSMOS field coverage. To maximize the number of the matched DESI ELGs within the COSMOS catalog, we use the measurements provided by the `CLASSIC` catalog.

Photometric Redshifts and Physical Properties. The COSMOS2020 catalog also includes photometric redshifts and the estimated physical properties of galaxies, including the stellar mass (M_*) and SFR, based on the `LePhare` (O. Ilbert et al. 2006) and `EAZY` (G. B. Brammer et al. 2008) codes. These two algorithms adopt different sets of galaxy spectral energy distribution templates and different star formation histories. For the details, we refer the reader to J. R. Weaver et al. (2022) and K. M. L. Gould et al. (2023).

In this work, we adopt the photometric redshifts and galaxy properties from the `EAZY` algorithm. This choice is motivated by the fact that for the DESI ELGs in the COSMOS2020 footprint and with $|z_{\mathrm{DESI}} - z_{\mathrm{photo}}| < 0.1$, the SFR based on the H$\alpha$ emission-line luminosity (R. C. Kennicutt 1998; R. C. Kennicutt et al. 2009) from the `EAZY` algorithm has a higher Spearman's correlation coefficient ($\rho_s \simeq 0.72$), with [O II] emission-line luminosity, directly measured from the DESI spectra, than the estimated SFR provided by the EAZY ($\rho_s \simeq 0.68$) and `LePhare` ($\rho_s \simeq 0.66$) codes. More specifically, we use the best-fit z_{photo}, the stellar mass, and the SFR-based Hα emission-line luminosity from the `EAZY` algorithm applied to the `CLASSIC` catalog. We note that the COSMOS2020 catalog provides dust-reddened Hα luminosity. We correct the dust effect by using $A_{\mathrm{H}\alpha} = 0.81\,A_V$ (M. Kriek & C. Conroy 2013) and adopt the Hα intrinsic luminosity and SFR relation from R. C. Kennicutt et al. (2009) for the Chabrier initial mass function (G. Chabrier 2003). All the stellar mass and the SFR of galaxies used in the work are based on the above method. We emphasize that the goal of this research is to explore the relationship between the [O II] line profiles and the physical properties of galaxies and to compare the physical properties of DESI ELGs with the properties of the overall star-forming galaxy population at similar redshifts. Identifying the most accurate estimates of the physical properties of DESI ELGs is beyond the scope of this paper.

Morphological Parameters. Besides the photometric redshifts and physical properties, we make use of the nonparametric diagnostics of galaxy morphology provided by the

Zurich Estimator of Structural Type catalog (ZEST; COSMOS Team 2007; C. Scarlata et al. 2007). The catalog includes

1. Asymmetry (A), which quantifies the rotational symmetry of galaxy light distribution;
2. Concentration (C), which quantifies the concentration of galaxy light distribution;
3. Gini coefficient (G), which quantifies the uniformity of the light distribution,
4. M20, which is the second-order moment of the brightest 20% pixels, and
5. R80, which is the length of the best-fit ellipse including 80% of total light along the semimajor axis in units of arcsec (″), reflecting the observed size of the galaxy

(e.g., R. G. Abraham et al. 2003; J. M. Lotz et al. 2004; C. J. Conselice 2014, for a review). These parameters are derived from Hubble Space Telescope (HST) Advanced Camera for Surveys (ACS) F814W images (A. M. Koekemoer et al. 2007; A. Leauthaud et al. 2007; N. Scoville et al. 2007) for galaxies with $I_{\rm AB} < 24$, a depth sufficient to include all the DESI ELGs (A. Raichoor et al. 2023). The F814W filter covers the rest-frame near-ultraviolet to optical wavelengths of DESI ELGs. At $z \sim 0.6$, the filter covers the range of 4400 to 5900Å, and at $z \sim 1.6$, it covers the range of 2700 to 3600 Å. The average width of the point-spread function for the HST ACS images is $\sim$0.″1 (A. M. Koekemoer et al. 2007), which corresponds to $\sim$0.67 kpc at $z = 0.6$ and $\sim$0.86 kpc at $z = 1.6$.

Combining the information from these three catalogs,[37] we produce two samples of DESI ELGs with $|z_{\rm DESI} - z_{\rm photo}| < 0.1$:

1. *DESI-COSMOS*. The first one, DESI-COSMOS, is the combination of the DESI ELGs and COSMOS2020 `CLASSIC` catalog, which is done by crossmatching sources with a 0.″75 radius. This yields a sample with $\sim$4500 sources having photo-z, stellar mass, and SFR measurements.
2. *DESI-ZEST*. The second sample, DESI-ZEST, includes the morphological information from the ZEST catalog by matching the DESI-COSMOS sources with the ZEST sources with a 0.″75 radius. This DESI-ZEST sample is a subset of the DESI-COSMOS sample, consisting of about 2200 galaxies.

We note that there are rare cases ($\sim$0.3%) that one ELG has two or more galaxies within a 0.″75 radius from the COSMOS2020 or the ZEST catalogs. For such cases, we adopt the properties of the most massive and brightest (in F814W filter) galaxy.

In order to compare the properties of DESI ELGs and that of the overall star-forming galaxies at the similar redshifts, we construct reference samples by selecting star-forming galaxies from the COSMOS2020 catalog that satisfy

$$\log_{10} {\rm SFR} > -0.65 + 0.65 \times (\log_{10} M_* - 10) + 1.07 \times (z - 0.1). \quad (1)$$

This functional form is adopted from J. Moustakas et al. (2013) for separating star-forming and quiescent galaxies with a modification on the zero-point to better include star-forming galaxies based on the adopted SFR. We refer to this star-forming galaxy sample from the full COSMOS2020 catalog as COSMOS-SFG. We also crossmatch this COSMOS-SFG sample with the ZEST catalog with a 0.″75 radius to obtain a sample of approximately 100,000 star-forming galaxies with morphological parameters based on HST ACS images. This sample is referred to as COSMOS-ZEST. These samples will be used in Section 4.

3. PCA Decomposition and Classification

The basis of the PCA (I. Jolliffe 2002, for a review) is to calculate the eigenvalues of the correlation matrix for a given data set and the corresponding eigenvectors. By doing so, one can identify key eigenvectors that carry the bulk of the variance of the data set. Via PCA, each original data vector can be described by the sum of the mean vector and a linear combination of the eigenvectors.

In this analysis, our data matrix consists of DESI ELG spectra around the [O II] emission-line region ($3722 < \lambda < 3735$ Å). Via the PCA decomposition, each ELG spectrum can be described as

$$f_{\lambda,i} = \langle \mu \rangle + \sum_{j=1}^{m} {\rm Coeff}_{ij} \times |\xi_j\rangle, \quad (2)$$

where $f_{\lambda,i}$ is the ith ELG spectrum, $\langle \mu \rangle$ is the mean spectrum of the whole ELG spectra, $|\xi_j\rangle$ is the jth eigenvector (eigenspectrum), and ${\rm Coeff}_{ij}$ is the coefficient for the ith ELG spectrum and the jth eigenspectrum. In this manner, one describes each spectrum with only a few coefficients and may explore and classify [O II] profiles based on their values. This approach has been adopted in previous studies for exploring various types of astronomical data sets, such as quasar spectra, galaxy spectra, and galaxy physical properties (e.g., C. W. Yip et al. 2004; N. Suzuki 2006; M. Taghizadeh-Popp et al. 2012). In this work, we use the weighted PCA algorithm, which takes into account the measurement uncertainty and missing data, developed by L. Delchambre (2015).[38]

3.1. Spectral Processing

In order to perform PCA, we shift all the ELG spectra to their rest frame based on their `Redrock` redshifts and project the spectra onto a common wavelength grid with a pixel size of 0.5 Å ($\sim$40 km s^{-1}). This pixel size corresponds to the original pixel size of DESI spectra (0.8 Å) in the rest frame of a source at redshift 0.6, the lower bound of the redshift in our analysis. We estimate the continuum of each spectrum using the median value of pixels around the [O II]$\lambda\lambda$3726, 3729 emission lines ($3715 < \lambda < 3722$ Å and $3735 < \lambda < 3742$ Å) and subtract the estimated continuum from the spectrum. We then normalize each spectrum by the maximum value of the pixels with an SNR > 3 between 3726 and 3731 Å of the spectrum. By doing so, the results of the PCA are more sensitive to the shape variation of the [O II]$\lambda\lambda$3726, 3729 emission lines, the focus of this study, than the flux difference between spectra. Finally, we remove spectra with any bad pixels and with the median SNR of the [O II] regions between 3726 and 3731 Å lower than 2. The final sample includes 229,996 ELGs.

---

[37] Except for $\sim$0.4% of DESI ELGs within the locations of the stellar masks of the COSMOS2020 catalog, the rest of the DESI ELGs in the COSMOS field have matched counterparts within the 0.″75 radius.

[38] We use the package at https://github.com/jakevdp/wpca developed by Jake VanderPlas.

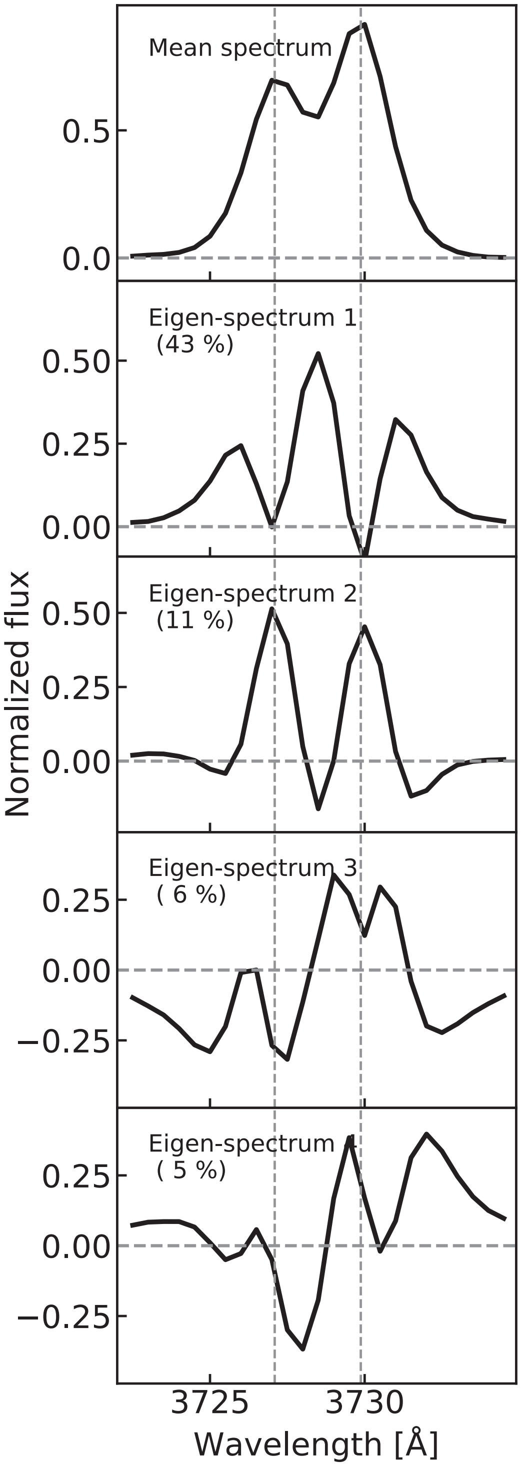

Figure 1. Mean spectrum and PCA eigenspectra of DESI ELGs. The top panel shows the mean spectrum, and the second to fifth panels show the eigenspectra ordered by their explained variance listed in each panel. The two vertical dashed lines indicate the wavelengths of the [O II] doublet.

3.2. Eigenspectra

The mean spectrum and the eigenspectra obtained by PCA are shown in Figure 1. The top panel shows the mean spectrum of the ELGs with clear [O II]$\lambda\lambda$3726, 3729 emission lines. The lower four panels show the first four eigenspectra, respectively.

1. *First eigenspectrum*. The first eigenspectrum explains 43% of the variance of the [O II] line region. It has three peaks at 3726 Å, 3728.5 Å, and 3731 Å, which do not coincide with the [O II]$\lambda\lambda$3726, 3729 lines. The [O II] $\lambda\lambda$3726, 3729 lines in the first eigenspectrum are nearly zero. By modulating flux off the line center, this eigenspectrum broadens the line width of the emission with positive coefficients and reduces the line width with negative coefficients.
2. *Second eigenspectrum*. The second eigenspectrum explains 11% of the variance. It has two peaks that coincide with the wavelengths of [O II]$\lambda\lambda$3726, 3729. With a positive value coefficient, this component increases the [O II]$\lambda\lambda$3726, 3729 line strength, and with a negative coefficient, it decreases the strength. Therefore, the second eigenspectrum modulates the amplitude and peakedness of the [O II]$\lambda\lambda$3726, 3729 emission lines.
3. *Third and fourth eigenspectra*. The third and fourth eigenspectra both explain only $\sim$5% of the variance. They both have positive values around the [O II] 3730 line and negative values around the [O II] 3727 line. Therefore, these two components can account for the asymmetry of the line profiles induced by both physical and nonphysical signals, such as the variation of the line ratio between [O II] λ3726 and [O II]λ3729 (e.g., M. Kaasinen et al. 2017) and the residuals of the sky lines.

The first two eigenspectra explain $\sim$54% (43%, 11%) of the total variances of the ELG spectra. In the following, we focus on the first two eigenspectra and their application in characterizing the DESI ELG [O II]$\lambda\lambda$3726, 3729 line profiles, while we plan to investigate the application of the third and fourth eigenspectra in future studies.

3.3. Coefficients

Using only the first two eigenspectra, the [O II]$\lambda\lambda$3726, 3729 line profile of each ELG can be described by two coefficients. Figure 2 shows the distribution of Coefficient 1 (Coeff_1) (x-axis) and Coefficient 2 (Coeff_2) (y-axis). We find that the [O II] $\lambda\lambda$3726, 3729 line profiles of DESI ELGs cover the parameter space with both Coeff_1 and Coeff_2, ranging from positive to negative values. Based on the (Coeff_1, Coeff_2) values, we classify ELGs into four regions using two empirical selection boundaries, $\mathrm{Coeff}_2 = 0.15 \times \mathrm{Coeff}_1 - 0.25$ and $\mathrm{Coeff}_1 = 0$, and show example spectra in Figure 3:

1. *Narrow region (N)*. $\mathrm{Coeff}_1 < 0$ and $\mathrm{Coeff}_2 > 0.15 \times \mathrm{Coeff}_1 - 0.25$. In this region, the doublet of [O II] $\lambda\lambda$3726, 3729 is narrow and is clearly separated.
2. *Broad region (B)*. $\mathrm{Coeff}_1 > 0$ and $\mathrm{Coeff}_2 > 0.15 \times \mathrm{Coeff}_1 - 0.25$. With the increase of Coeff_1 values, the [O II]$\lambda\lambda$3726, 3729 line width becomes broader, and the two [O II] lines gradually blend together with increasing Coeff_1.
3. *Double-peak*[39] *region (D)*. $\mathrm{Coeff}_1 > 0$ and $\mathrm{Coeff}_2 < 0.15 \times \mathrm{Coeff}_1 - 0.25$. The combination of a positive value of Coeff_1 and a negative value of Coeff_2 yields [O II] line profiles with three emission-line peaks with the central peak at $\sim$3728.5 Å, as shown in the lower-right panels of Figure 3. These profiles are consistent with the

---

[39] We note that the [O II] profiles have three peaks, which, in fact, consist of two components with an overlapping [O II] emission-line doublet. For transition with a single line, the emission-line profiles are double peak. Following the literature, we use "double-peak" galaxies to describe these systems.

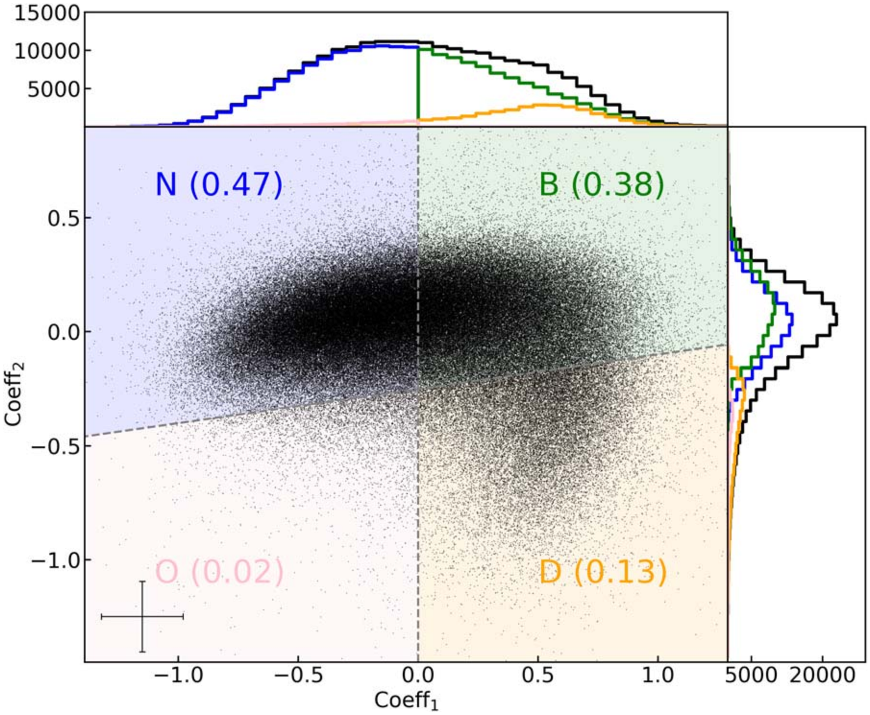

Figure 2. Distribution of ELGs in the PCA Coeff$_1$ (x-axis) and Coeff$_2$ (y-axis) space. ELGs are separated into four regions, narrow (N), broad (B), double-peak (D), and outlier (O), by the two dashed lines (Coeff$_2 = 0.15 \times$ Coeff$_1 - 0.25$ and Coeff$_1 = 0$) and with background colors being blue, green, orange, and pink respectively. The value shown in each region indicates the fraction of total sources in the region, and the top and right panels show the number distributions of the four types as a function of Coeff$_1$ and Coeff$_2$, respectively. The median values of the uncertainties of Coeff$_1$ and Coeff$_2$ are 0.17 and 0.16, respectively, as shown in the lower-left corner. The uncertainties of the PCA coefficients of each spectrum are estimated by perturbing the spectrum with the observed spectral noise 1000 times and obtaining the corresponding coefficients and the standard deviations.

[O II] lines from two components with velocity offsets. The two velocity component spectral features are also observed in other emission lines, such as Hβ and [O III], for low-z ELGs with double-peak emission lines. Within the selection, there are $\simeq$30,000 double-peak galaxies identified from the DESI EDR data set, which is currently the largest double-peak galaxy catalog at $z > 0.6$.

4. *Outlier region (O).* Coeff$_1 < 0$ and Coeff$_2 < 0.15 \times$ Coeff$_1 - 0.25$. In this region, the [O II] line profiles are affected by residuals of strong sky emission lines, as shown in the lower-left panels of Figure 3.

These four regions are highlighted in blue, green, orange, and pink for narrow, broad, double-peak, and outlier emission lines, respectively, in Figure 2. The histograms on the top and right panels show the number distributions of Coeff$_1$ and Coeff$_2$, respectively. We find that $\sim$47% of ELGs are in the narrow region, $\sim$38% of ELGs are in the broad region, $\sim$13% are in the double-peak region, and $\sim$2% are in the outlier region. We classify the ELGs into four regions, and emphasize that while based on the coefficients, their distribution is smooth and continuous.

3.4. Coefficients versus Galaxy Observed Properties

We now explore and describe the relationships between the observed properties of ELGs and (Coeff$_1$, Coeff$_2$) values. In the upper panels of Figure 4, we first show the median values of galaxy observables in each Coeff$_1$ and Coeff$_2$ bin, including the number of ELGs, their redshifts, the observed WISE 1 band magnitude, and the observed g-band magnitude from left to right, respectively.

Number. The first upper panel from the left shows the number distribution of DESI ELGs as a function of Coeff$_1$ and Coeff$_2$. As shown in Figure 2, most of the ELGs are in the narrow and broad regions, with a small fraction of ELGs in the double-peak region.

Redshifts. The second upper panel shows the median redshifts as a function of Coeff$_1$–Coeff$_2$. As can be seen, there is a redshift dependence between the median redshifts and Coeff$_1$. ELGs with narrow [O II]$\lambda\lambda$3726, 3729 systems (Coeff$_1 < 0$) have lower median redshifts than [O II]$\lambda\lambda$3726, 3729 with Coeff$_1 > 0$. Additionally, ELGs with Coeff$_1$ and Coeff$_2$ values near the edge of the entire distribution tend to have median redshifts $z \sim 1.5$. This is due to the fact that the line profiles of those galaxies are affected by sky emission-line residuals and therefore have relatively extreme Coeff$_1$ and Coeff$_2$ values.

Observed WISE 1 Band Magnitude. The third upper panel of Figure 4 shows the median observed WISE 1 band magnitude of ELGs. There is a trend indicating that ELGs with larger Coeff$_1$ values and lower Coeff$_2$ values are, on average, brighter in the WISE 1 band. This trend is not driven by the redshift correlation shown in the second panel. It shows an opposite correlation where high-z ELGs with Coeff$_1 > 0$ tend to have a brighter observed magnitude than ELGs at low redshifts with Coeff$_1 < 0$. The WISE 1 band (3.4 μm) corresponds to $\sim$1.7 μm in the rest frame of ELGs at $z \sim 1$, being sensitive to the stellar mass of galaxies. Therefore, the relationship between the coefficients and WISE 1 mag links the relationship between the coefficients and stellar mass.

Observed g-band Magnitude. The last upper panel of Figure 4 shows the median g-band observed magnitude. A trend can be observed, showing the brightness of ELGs in the g band increases

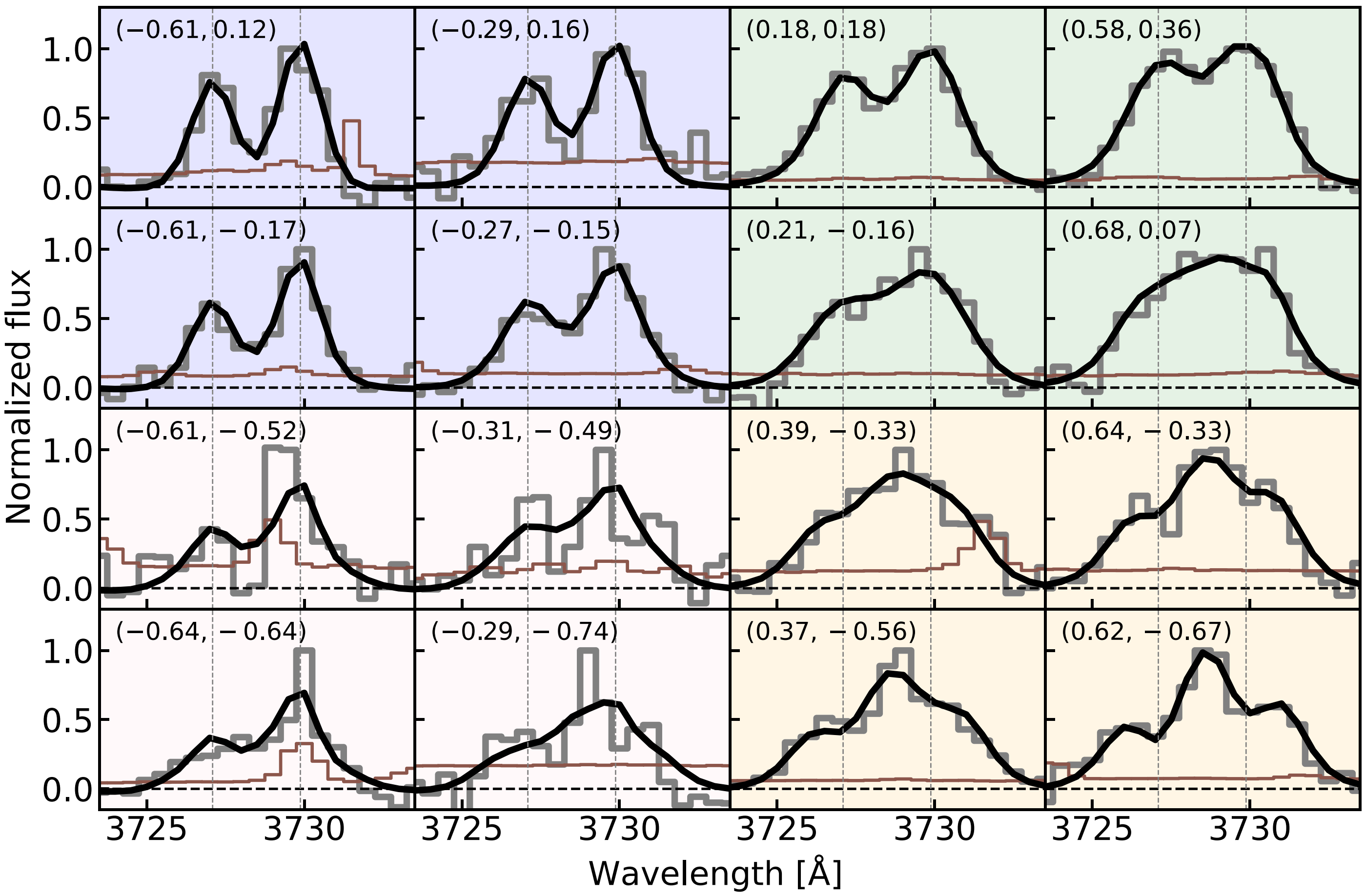

Figure 3. Example spectra of DESI ELGs as a function of Coeff_1 and Coeff_2. The values shown on the top left corner in each panel are (Coeff_1, Coeff_2) values of the ELG. The gray spectrum is the observed spectrum, the black curve is the PCA reconstructed spectrum based on the first two eigenspectra, and the uncertainty array is in brown. The background colors of the panels indicate the locations of the ELGs in the Coeff_1 and Coeff_2 space, as shown in Figure 2.

with the Coeff_1 values. ELGs with $\mathrm{Coeff}_1 \sim 1$, on average, are brighter than the rest of the ELGs. Similarly to the WISE 1 band distribution, this trend is not driven by the redshift correlation shown in the second panel. We also find that the relationship between the coefficients and *g*-band magnitude differs from the relationship between the coefficients and WISE 1 band magnitude. DESI ELGs tend to have a bright median *g*-band magnitude around $\mathrm{Coeff}_1 > 0.5$ and $\mathrm{Coeff}_2 \simeq 0$, while DESI ELGs tend to have a bright median WISE 1 band magnitude around $\mathrm{Coeff}_1 > 0.5$ but extends to $\mathrm{Coeff}_2 \sim -1$. Considering the *g* band covers the ultraviolet wavelength in the rest frame of ELGs, this difference suggests that the relationship between the coefficients and SFR is not entirely driven by the relationship between the coefficients and stellar mass.

In addition to the photometric properties of DESI ELGs, we extract the [O II] line information by performing spectral fitting analysis with two models. The first one, F_s, is a single redshift model describing [O II]$\lambda\lambda$3726, 3729 lines with two Gaussian profiles with a single redshift,

$$F_s(x;\, A_s,\, \mu_1,\, \sigma_s) = A_s \times e^{\frac{-(x-\mu_1)^2}{2\sigma_s^2}} + 1.33\, A_s \times e^{\frac{-(x-\mu_1-2.783)^2}{2\sigma_s^2}}, \tag{3}$$

where x is the wavelength (Å), A_s is the amplitude of the [O II] λ3726 line, μ_1 is the center wavelength (3727.092 Å) of the [O II]λ3726 line, σ_s is the line width (Å) of the Gaussian profile. The [O II]λ3729 line is fitted by the second term where we adopt a fixed line ratio [O II]3729/[O II]3726 = 1.33, the ratio between the two lines from the mean spectrum. The second model, F_d, is a two redshift model describing the [O II] $\lambda\lambda$3726, 3729 lines with two sets of two Gaussian profiles, F_s, with a wavelength offset $\Delta\lambda$,

$$F_d(x;\, A_{d_1},\, \mu_1,\, \sigma_{d_1},\, R,\, \Delta\lambda,\, \sigma_{d_2}) = F_s(x;\, A_{d_1},\, \mu_1,\, \sigma_{d_1}) + F_s(x;\, R \times A_{d_1},\, \mu_1 + \Delta\lambda,\, \sigma_{d_2}), \tag{4}$$

where R describes the amplitude difference in terms of ratio between the [O II]λ3726 lines from two redshifts. With these two models, we obtain the best-fit parameters and the χ^2 of the two models for all the ELG [O II]$\lambda\lambda$3726, 3729 profiles. The lower panels of Figure 4 summarize the fitting results.

p-value. We first quantify the performance of the two models for describing the [O II]$\lambda\lambda$3726, 3729 profiles with an F-test by following D. Maschmann et al. (2020). We estimate $f_{\mathrm{stat}} = \frac{(\chi_s^2 - \chi_d^2)\,/\,(N_s - N_d)}{\chi_d^2\,/\,N_d}$, where χ_s^2 and χ_d^2 are the χ^2 values of the best fits from the single and two redshift models and N_s and N_d are the corresponding degree of freedoms. We convert the f_{stat} values into *p*-values for this hypothesis test, with the null hypothesis being that the F_d model does not provide a better fit than the F_s model. In other words, lower *p*-values indicate a higher probability of rejecting the null hypothesis.

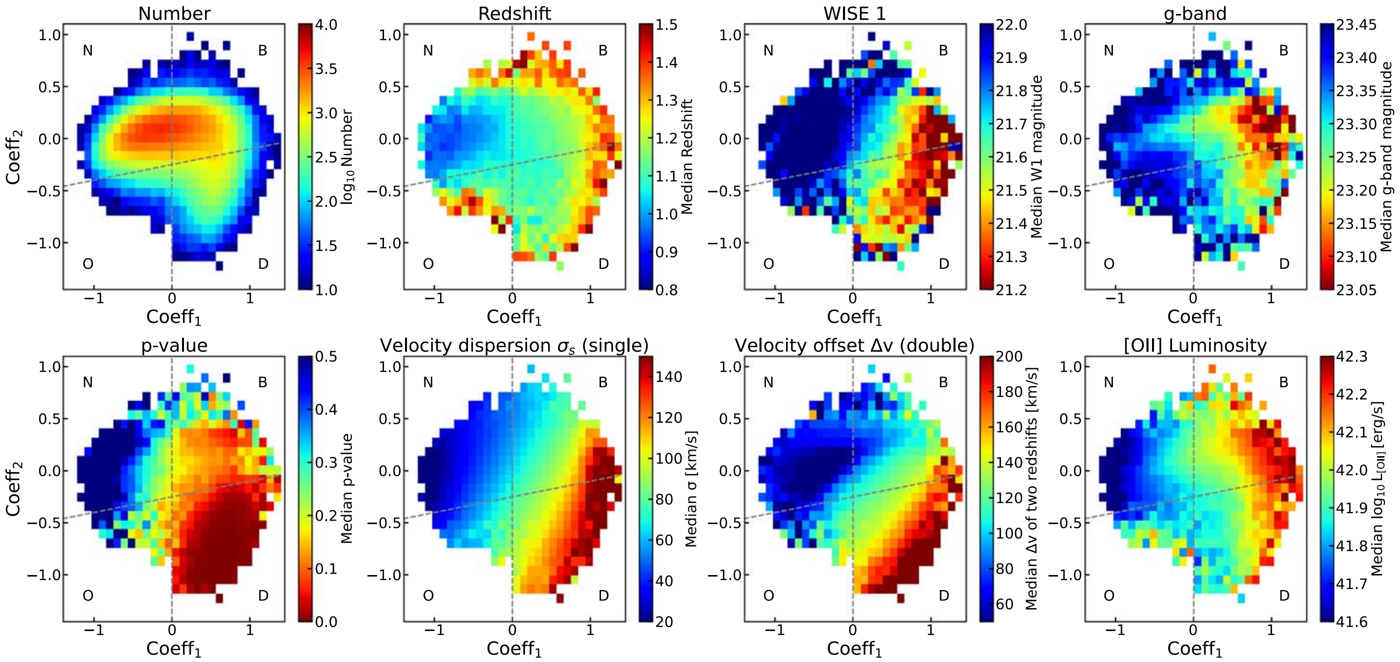

Figure 4. Distributions of observed properties (upper panels) and spectral properties (lower panels) of ELGs as a function of Coeff$_1$ (x-axis) and Coeff$_2$ (y-axis). In the upper panels, the left panel shows the number of sources in each Coeff$_1$ and Coeff$_2$ bin. The second to the fourth panels show the median values of the redshifts, WISE 1 mag, and g-band magnitude, respectively. The lower panels from left to right show the median values of the p-value, with the null hypothesis being that the F_d model does not provide a better fit than the F_s model, velocity dispersion σ_s, velocity offset (Δv), and [O II] luminosity respectively. We note that for clarity, pixels with a number of sources lower than 10, which occupy ∼0.7% of the total number of sources, are not displayed.

The first lower panel of Figure 4 from the left shows the median p-values, indicating that the p-values depend on Coeff$_1$ and Coeff$_2$. For systems with negative Coeff$_1$ values, the p-values are around 0.5, suggesting that the single redshift model is sufficient to describe the two [O II] lines. The median p-values decrease with the Coeff$_1$ values. This indicates that broader line profiles tend to require two components for describing the line profiles, especially the line profiles with three emission-line peaks in the double-peak region, which have median p-values lower than 0.05 (a commonly adopted threshold for rejecting the null hypothesis).

Line Velocity Properties. The second and the third lower panels of Figure 4 show the median velocity dispersion of the emission lines estimated from the line width parameter σ_s of the single redshift model and the median velocity separation estimated from the $\Delta\lambda$ parameter of the two redshift models, respectively. We report the velocity dispersion with the spectral resolution effect being corrected. Both parameters increase from the top left corner ($\mathrm{Coeff}_1 < 0$ and $\mathrm{Coeff}_2 > 0$) to the bottom right corner ($\mathrm{Coeff}_1 > 0$ and $\mathrm{Coeff}_2 < 0$). The median velocity dispersion for DESI ELGs in the narrow region, broad region, and double-peak region are $\sim 50\,\mathrm{km\,s^{-1}}$, $\sim 80\,\mathrm{km\,s^{-1}}$, and $\sim 100\,\mathrm{km\,s^{-1}}$, respectively. For the systems that require two redshifts, velocity dispersion from the single redshift model partially reflects the velocity offset of the redshift difference shown in the third panel. The median velocity offsets for DESI ELGs in the broad and double-peak regions are $\sim 100\,\mathrm{km\,s^{-1}}$ and $\sim 150\,\mathrm{km\,s^{-1}}$.

[O II] *Luminosity*. Finally, the last lower panel of Figure 4 shows the median total luminosity ($\mathrm{erg\,s^{-1}}$) of the two [O II] lines estimated with the best fit of the two redshift models. There is a correlation between [O II] luminosity and Coeff$_1$, indicating that ELGs with broader line profiles have, on average, higher [O II] luminosity, a trend similar to the g-band magnitude. This trend is detected across the entire redshift range probed in this work. Considering [O II] luminosity as a tracer of the SFR of galaxies (e.g., R. C. Kennicutt 1998), this result indicates that the SFR of ELGs correlates with the emission-line profiles—the broader the line, the higher the SFR.

4. Relationship between the Line Profiles and Galaxy Physical Properties

4.1. SFR and Stellar Mass

With the observed and spectral properties of DESI ELGs being explored, in order to better understand the underlying relationship, we now explore the physical properties of DESI ELGs and their connections to the [O II]$\lambda\lambda$3726, 3729 line profiles, using the DESI-COSMOS sample as described in Section 2. Figure 5 shows the median values of ELG stellar mass (left panel) and SFR (right panel) in the Coeff$_1$−Coeff$_2$ parameter space. The colors reflect the values of stellar mass and SFR. Being consistent with the trend observed in the median values of WISE 1 band magnitude and g-band magnitude shown in Figure 4, the stellar mass and SFR increase with Coeff$_1$, demonstrating the overall relationship between the line profiles, stellar mass, and SFR.

Given that the stellar mass and SFR of star-forming galaxies correlate with each other, we further investigate the correlation between Coeff$_1$ and SFR with a fixed range of stellar mass of ELGs. The results are shown in Figure 6 for three redshift bins from $0.6 < z < 1$ (left), $1 < z < 1.25$ (middle), and $1.25 < z < 1.63$ (right). The upper panels show the stellar mass distributions of ELGs with different Coeff$_1$ ranges. The vertical color dashed line shows the median stellar mass for each Coeff$_1$ bin, indicating that ELGs with higher Coeff$_1$ values, on average, have higher stellar mass.

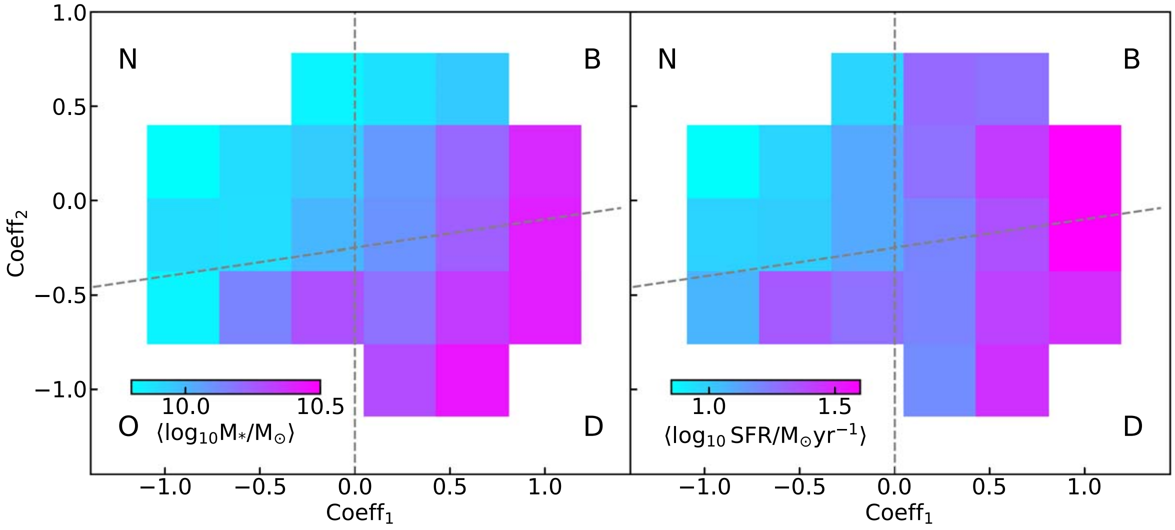

Figure 5. Stellar mass and SFR of ELGs as a function of Coeff$_1$ (x-axis) and Coeff$_2$ (y-axis). The colors in the left and right panels indicate the median stellar mass and SFR of ELGs, respectively. For clarity, pixels with the number of sources lower than five, which occupy ∼0.6% of the total number of sources, are not displayed.

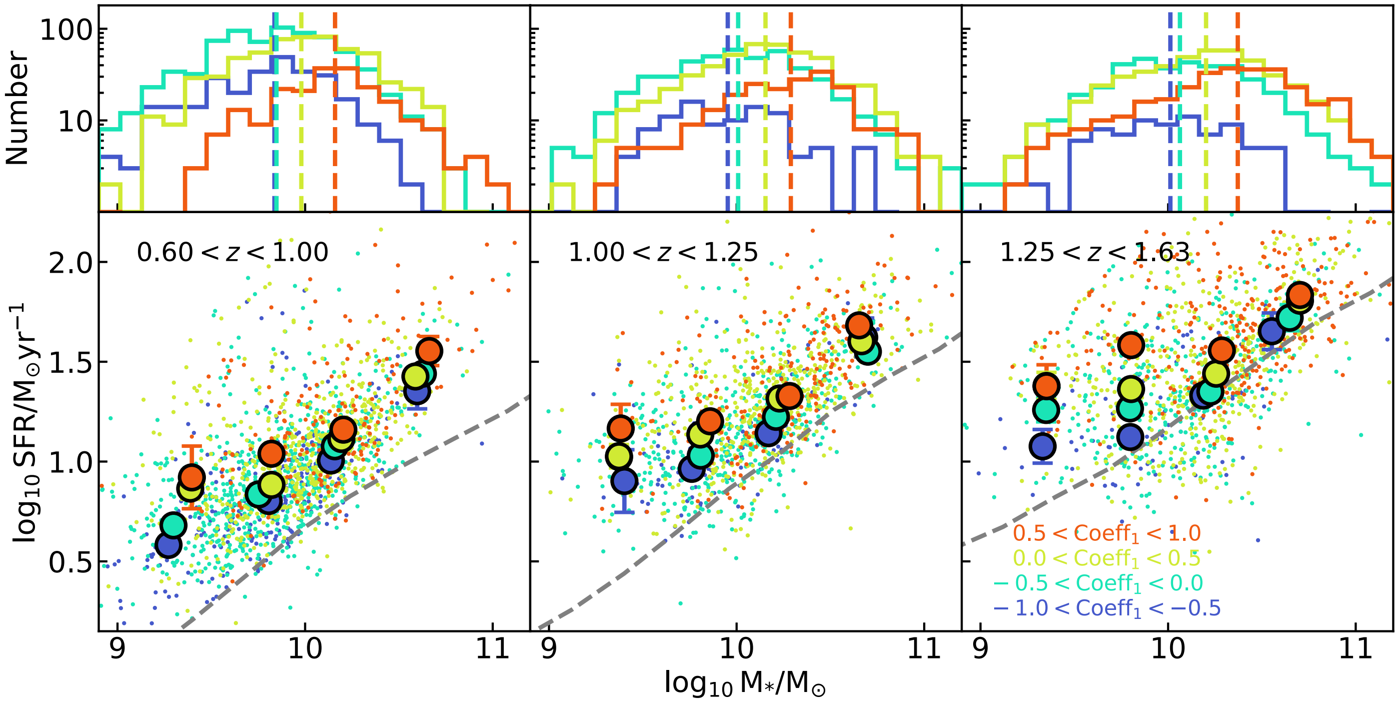

Figure 6. Correlations between stellar mass, SFR, and Coeff$_1$. The upper panels show the number distributions of ELGs with different Coeff$_1$ values indicated by blue ($-1 < \mathrm{Coeff}_1 < -0.5$), cyan ($-0.5 < \mathrm{Coeff}_1 < 0$), green ($0 < \mathrm{Coeff}_1 < 0.5$), and orange ($0.5 < \mathrm{Coeff}_1 < 1$). The vertical color dashed lines reflect the median stellar mass of the ELGs within the corresponding Coeff$_1$ bin. The three panels from left to right show the results for ELGs at three redshift bins from $z = 0.6$ to $z = 1.63$. The lower panels show the SFR as a function of stellar mass with colors indicating Coeff$_1$. The small data points are individual ELGs and the large data points with uncertainty are the median SFR as a function of stellar mass for each Coeff$_1$ bin. The uncertainties are estimated based on bootstrapping the sample 1000 times. The gray dashed lines in the lower panels are the median relationship based on galaxies in the COSMOS-SFG sample within the same redshift regions.

The lower panels of Figure 6 show the SFR of ELGs as a function of stellar mass with colors indicating their Coeff$_1$ values. The data points with uncertainties are the median values of $\log_{10}$ SFR for the four Coeff$_1$ bins. The uncertainties are estimated by bootstrapping the sample 1000 times. We find that with a fixed stellar mass, the SFR increases with Coeff$_1$. This trend is observed across the stellar mass and redshift ranges of the DESI ELG sample. We also calculate the median SFR of the overall star-forming galaxy population at the same redshift range from the COSMOS-SFG sample defined in Section 2.2. The gray dashed line in each panel shows the corresponding median trend. As can be seen, the majority of DESI ELGs have an SFR higher than the median SFR of the overall star-forming galaxy population. This result indicates that DESI preferentially selects star-forming galaxies with an SFR higher than the main-sequence galaxies, a conclusion also reached by S. Yuan et al. (2023) using DESI ELGs with $0.8 < z < 1.1$ based on a similar analysis.

To summarize, Figure 6 demonstrates that while there is a general trend between the line profiles and the stellar mass and SFR, which are coupled together due to the overall stellar mass and SFR correlation, there is an additional correlation between the SFR and line profiles when considering ELGs with similar stellar mass. This trend is consistent with the correlation between gas velocity dispersion and SFR of star-forming galaxies observed at different redshifts from the local Universe (e.g., D. R. Law et al. 2022), $z \sim 1$ (e.g., Y. Mai et al. 2024) to $z \sim 2$ (e.g., H. Übler et al. 2019). In the following, we explore the relationship between this excess SFR with respect to the overall star-forming galaxy population and line profiles and investigate possible mechanisms behind this relationship.

4.2. Excess SFR, Galaxy Morphology, and Line Profiles

To explore how the excess SFR of ELGs correlates with the line profiles and inform the underlying mechanisms, in addition to the stellar mass and SFR, we use the DESI-ZEST sample, which includes five morphological properties derived from HST ACS F814W images from the ZEST catalog (C. Scarlata et al. 2007) in the analysis: asymmetry (A), Gini coefficient (G), concentration (C), and the second-order moment of the brightest 20% pixels (M20), and R80 parameter reflecting the galaxy size of ELGs. In Appendix B, we summarize the average relationships between the line profiles and the morphological properties.

We calculate the differences between the physical and morphological parameters of DESI ELGs and that of overall star-forming galaxies from the COSMOS-ZEST sample at similar redshifts, stellar mass, and sizes. This analysis allows us to remove the correlations driven by these three parameters. More specifically, for each DESI ELG, we use at least 10 COSMOS-ZEST galaxies as references with redshift difference $|\Delta z| < 0.025$, stellar mass difference $|\Delta \log_{10} M_*/M_\odot| < 0.05$, and size difference $|\Delta \mathrm{R80}| < 0.''025$. If less than 10 COSMOS-ZEST galaxies are found, we increase the conditions until there are at least 10 galaxies within the selection. Finally, for each ELG, we calculate the difference between the ELG parameter values, SFR, A, G, M20, and C, and the median values of their references as $\Delta \log_{10}$ SFR, ΔA, ΔG, ΔM20, and ΔC, respectively.

Figure 7 shows the normalized probability density distributions of the excess values of the parameters. The top three parameters are the controlled parameters, stellar mass, redshift, and size (R80). By construction, the distributions are narrow and center around 0. The lower five parameters are the morphological parameters, G, M20, C, A, and the SFR values. We find that while there are mild positive excesses of G and M20 parameters, and on average, lower C parameters in comparison to the reference galaxies, the most significant difference is the asymmetric parameter A. The majority ($\sim$85%) of ELGs have asymmetric parameter A values higher than the reference galaxies with $\Delta A > 0$, a trend similar to the trend of the SFR with $\sim$86% ELGs with $\Delta \log_{10}$ SFR > 0.

To first explore the correlations between parameters, Table 1 shows the Spearman's correlation coefficients (ρ_s) between $\Delta \log_{10}$ SFR and the spectral PCA Coeff$_1$ and Coeff$_2$ and the excess values of morphological parameters with the uncertainties estimated by bootstrapping the sample 1000 times. We order the parameters according to their values of Spearman's correlation coefficients. The results show that

1. the $\Delta \log_{10}$ SFR correlates with ΔA with the highest absolute $\rho_s \simeq 0.31 \pm 0.02$ value among the parameters;
2. the second parameter with high $\rho_s \simeq 0.16 \pm 0.02$ value is the line profile coefficient, Coeff$_1$. We note that ρ_s between M_* and $\Delta \log_{10}$SFR is -0.13 ± 0.02. This demonstrates that the correlation between Coeff$_1$ and $\Delta \log_{10}$ SFR is not driven by stellar mass.
3. ΔM_{20}, Coeff$_2$, ΔG, and ΔC parameters also have some correlations with $\Delta \log_{10}$ SFR.

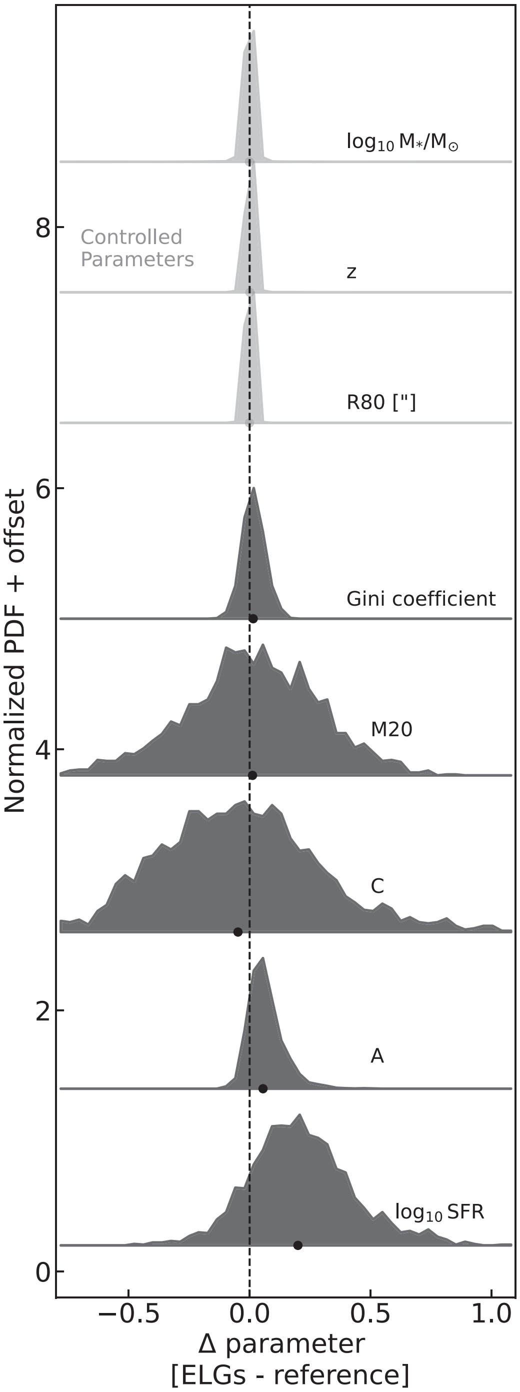

Figure 7. Normalized distributions of the differences between the properties of DESI ELGs and that of the reference galaxies. The top three properties are $\log_{10} M_*/M_\odot$, redshift, and galaxy sizes (R80), which are controlled parameters. The lower five properties are Gini coefficients, M20, concentration C, galaxy asymmetry A, and $\log_{10}$ SFR. Distributions are shifted in the y-axis for clarity. The black data points indicate the median values of the distributions.

These results demonstrate that $\Delta \log_{10}$ SFR of DESI ELGs correlates with both the morphological structures and the spectral profiles of galaxies.

To better understand the interplay between the shape of galaxies and the line profiles, Table 2 shows the ρ_s between Coeff$_1$ and Coeff$_2$ and the excess values of morphological

Table 1
Spearman's Correlation Coefficients between $\Delta\mathrm{Log}_{10}\mathrm{SFR}/M_\odot\ \mathrm{yr}^{-1}$ and Other Parameters

	ΔA	Coeff$_1$	ΔM_{20}	Coeff$_2$	ΔG	ΔC	M_*
$\Delta\log_{10}\frac{\mathrm{SFR}}{M_\odot\mathrm{yr}^{-1}}$	0.31 ± 0.02	0.16 ± 0.02	0.12 ± 0.02	0.05 ± 0.02	0.05 ± 0.02	-0.09 ± 0.02	-0.13 ± 0.02

Table 2
Spearman's Correlation Coefficients between PCA Coefficients and Morphological Parameters

	ΔA	ΔM_{20}	ΔC	ΔG
Coeff$_1$	0.08 ± 0.02	0.02 ± 0.02	0.01 ± 0.02	0.00 ± 0.02
Coeff$_2$	0.07 ± 0.02	0.01 ± 0.02	0.05 ± 0.02	0.05 ± 0.02

parameters with the uncertainties estimated by bootstrapping the sample 1000 times. As can be seen, Coeff$_1$ and Coeff$_2$ correlate the most with the ΔA ($>3\sigma$ detection), and ρ_s with other morphological parameters are consistent with no detection ($<3\sigma$). In other words, based on the ρ_s values of different parameter pairs, we find that there is a relationship between $\Delta\log_{10}$SFR, ΔA, Coeff$_1$, and Coeff$_2$.

We estimate the median ΔA and $\Delta\log_{10}$SFR as a function of Coeff$_1$ and Coeff$_2$ to quantify the relationship. The results are shown in the upper panels of Figure 8, indicating that both median ΔA (left panel) and median $\Delta\log_{10}$SFR (right panel) increase from negative Coeff$_1$ to positive Coeff$_1$ and from negative Coeff$_2$ to positive Coeff$_2$.

We also separate the Coeff$_1$–Coeff$_2$ parameter space into nine regions based on the general classification shown in Figure 2:

1. *Narrow region*. N1 (Coeff$_1 < -0.65$), N2 ($-0.65 <$ Coeff$_1 < -0.35$), and N3 ($-0.35 <$ Coeff$_1 < 0$);
2. *Broad region*. B1 ($0 <$ Coeff$_1 < 0.35$), B2 ($0.35 <$ Coeff$_1 < 0.65$), B3 (Coeff$_1 > 0.65$);
3. *Double-peak region*. D1 ($0 <$ Coeff$_1 < 0.35$), D2 ($0.35 <$ Coeff$_1 < 0.65$), D3 (Coeff$_1 > 0.65$).

We calculate the fraction of ELGs with high ΔA and $\Delta\log_{10}$SFR in each region. The lower panels of Figure 8 summarize the results. The lower-left panel shows the fractions for $\Delta A > 0.1$. We find that the fractions of ELGs with high ΔA increase from the narrow region (Coeff$_1 < 0$) to the broad region (Coeff$_1 > 0$). The fractions can increase by a factor of 3 from ∼0.1 to ∼0.3. Moreover, for ELGs with Coeff$_1 > 0$, the fraction is higher for ELGs in the broad region (Coeff$_2 \sim 0$) than in the double-peak region (Coeff$_2 \leqslant -0.4$). The lower-right panel shows the same measurements with $\Delta\log_{10}\mathrm{SFR} > 0.35$. The overall trends of $\Delta\log_{10}$SFR are consistent with ΔA, showing that ELGs with line profiles in the broad region (B) have relatively higher asymmetry morphology and higher SFR than their reference star-forming galaxies with similar stellar mass, sizes, and at the same redshifts. In addition, we perform the same analysis for a control sample, which consists of star-forming galaxies from the COSMOS-ZEST sample selected to have similar stellar mass, redshifts, and sizes as the properties of DESI ELGs with $|\Delta\log_{10}M_*/M_\odot| < 0.05$, $|\Delta z| < 0.025$, and $|\Delta \mathrm{R80}| < 0.025$, respectively. The color boxes in the lower panels of Figure 8 show the results, indicating that the DESI ELGs have preferentially higher ΔA and $\Delta\log_{10}$SFR than typical star-forming galaxies.

Finally, we quantify the median $\Delta\log_{10}$ SFR as a function of ΔA and line profiles with the results shown in Figure 9. The blue and green data points are measurements from ELGs with Coeff$_1 < -0.35$ and Coeff$_1 > 0.35$, respectively. The gray data points show the measurements in between. Here, we only include ELGs within the narrow and broad regions with Coeff$_2 > 0.15 \times$ Coeff$_1 - 0.25$. The measurements indicate that there is a general correlation between $\Delta\log_{10}$ SFR and ΔA, which has the same slope but different offsets for galaxies with different Coeff$_1$. With similar ΔA values, DESI ELGs with broader [O II] lines have, on average, 0.1 dex higher $\Delta\log_{10}$SFR than those with narrow [O II] lines. This again illustrates the connection between the [O II] line profiles, ΔA, and $\Delta\log_{10}$SFR of DESI ELGs.

5. Discussion

5.1. The Physical Mechanisms Driving the Relationship between SFR, Asymmetry, and Line Profiles

Utilizing data from the DESI spectroscopic observations and the COSMOS multiwavelength deep images, we find a relationship between three parameters of DESI ELGs, excess SFR ($\Delta\log_{10}$SFR), the asymmetry of galaxy shape (ΔA), and the [O II] line profiles. In the following, we discuss two scenarios that can possibly explain the observed relationship.

Merger Scenario. The merging of galaxies has been considered one of the crucial processes that can transform star-forming galaxies into passive ones. Theoretical works have shown that when two gas-rich galaxies merge, gas flows into the centers of the galaxies and triggers excess star formation activities (e.g., P. F. Hopkins et al. 2008; J. M. Lotz et al. 2008; D. R. Patton et al. 2020; C. Bottrell et al. 2023). These merger-induced star formation activities have also been supported by observational studies. Various observational approaches have been used to identify merging galaxies. For example, making use of data from the SDSS surveys, one can identify galaxy pairs with a range of impact parameters and use such a sample to explore the difference of the SFR compared with the SFR of controlled isolated galaxies with similar mass (e.g., S. L. Ellison et al. 2008, 2010; P. S. Behroozi et al. 2015). Another approach is to use the morphological parameters of galaxies derived from imaging data sets to identify possible shape features induced by merging events, such as tidal structures and/or asymmetry (e.g., J. M. Lotz et al. 2008; H. M. Yesuf et al. 2021). Both simulations and observations have shown that merging events at the closest separation enhances the SFR by approximately a factor of 2 with a fixed stellar mass.

Our results can be explained by this galaxy merger scenario. The excess asymmetry value ΔA indicates that a sizable fraction of DESI ELGs are disturbed galaxies, e.g., there are approximately 25% of DESI ELGs with $\Delta A > 0.1$, while there are only 7% of the controlled sample with $\Delta A > 0.1$ (lower panels of Figure 8). Figure 10 shows the HST ACS F814W images of randomly selected DESI ELGs as a function of ΔA. For most of the galaxies with $\Delta A > 0.05$, two or multi-clumpy

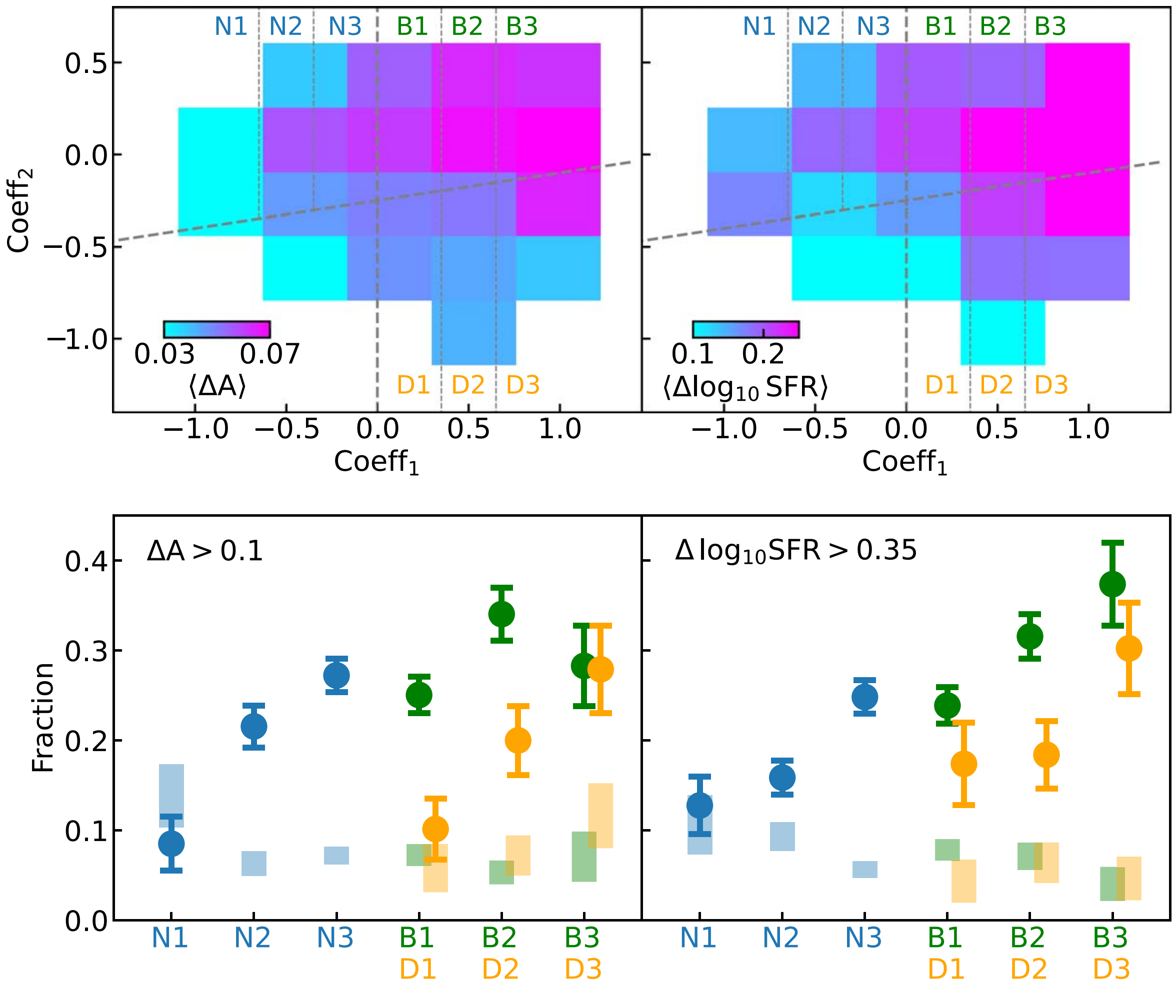

Figure 8. Correlation between ΔA, $\Delta\log_{10}$ SFR, and line profiles. The upper panels show the median ΔA (left) and $\Delta\log_{10}$ SFR$/M_{\odot}$ yr^{-1} (right) as a function of Coeff$_1$ and Coeff$_2$. The lower panels show the fractions of ELGs with $\Delta A > 0.1$ (left) and $\Delta\log_{10}$ SFR > 0.35 (right) within nine Coeff$_1$–Coeff$_2$ regions from narrow (blue), broad (green), to double-peak (orange), respectively. The color boxes show the same measurements for a control sample consisting of star-forming galaxies selected to have similar stellar mass, redshifts, and sizes as the DESI ELGs. The uncertainties are based on bootstrapping the sample 1000 times.

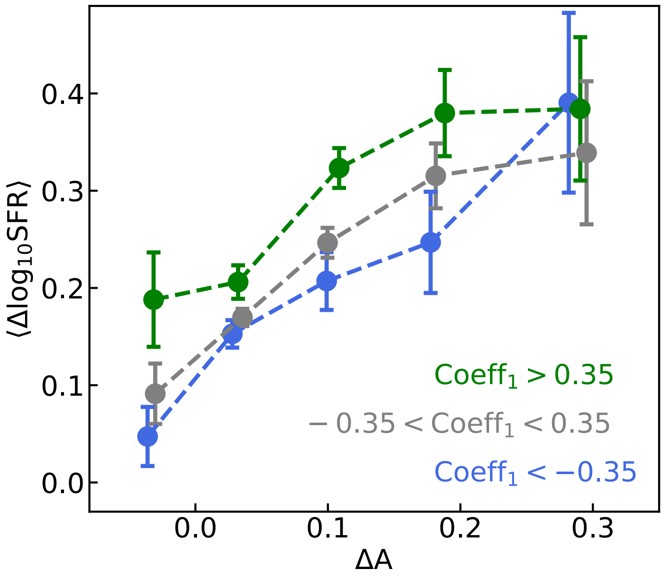

Figure 9. Median $\Delta\log_{10}$ SFR as a function of ΔA and Coeff$_1$. The blue, gray, and green data points show the median $\Delta\log_{10}$ SFR as a function of ΔA for ELGs with Coeff$_1 < -0.35$, $-0.35 <$ Coeff$_1 < 0.35$, and Coeff$_1 > 0.35$, respectively. These measurements only include the ELGs in the narrow and broad regions with uncertainties based on bootstrapping the sample 1000 times.

structures can be observed at small scales ($<1''$) with disturbed and possibly tidal features. The correlation between Coeff$_1$ and ΔA values provides further support for this merger scenario—when two interacting galaxies are at a close distance, the emission lines of the two galaxies are included in a single DESI fiber (12 kpc in diameter at $z \sim 1$) with a moderate line-of-sight velocity difference Δv which will produce a [O II]$\lambda\lambda$3726, 3729 line profile broader than a typical isolated star-forming galaxy with similar stellar mass. The three observational properties of DESI ELGs, the line profile, the asymmetry of galaxy shape, and the excess SFR, can all be fit within the merger scenario simultaneously.

Disk Instability Scenario. Another possibility is that the multi-clumpy structures shown in Figure 10 are star-forming regions formed due to disk instability (e.g., A. Dekel et al. 2009; N. Mandelker et al. 2014) in a single galaxy. Previous studies (e.g., K. L. Murata et al. 2014; Y. Guo et al. 2015; A. Martin et al. 2023; Z. Sattari et al. 2023) have shown that the clumpy fraction, the ratio between the number of star-forming galaxies with at least one off-center clump and the total number of star-forming galaxies, roughly peaks between $z \sim 1$ and $z \sim 2$, overlapping the redshift region of the DESI ELGs, and correlates with SFR with a fixed stellar mass. Bright star-forming clumpy structures can lead to asymmetric light

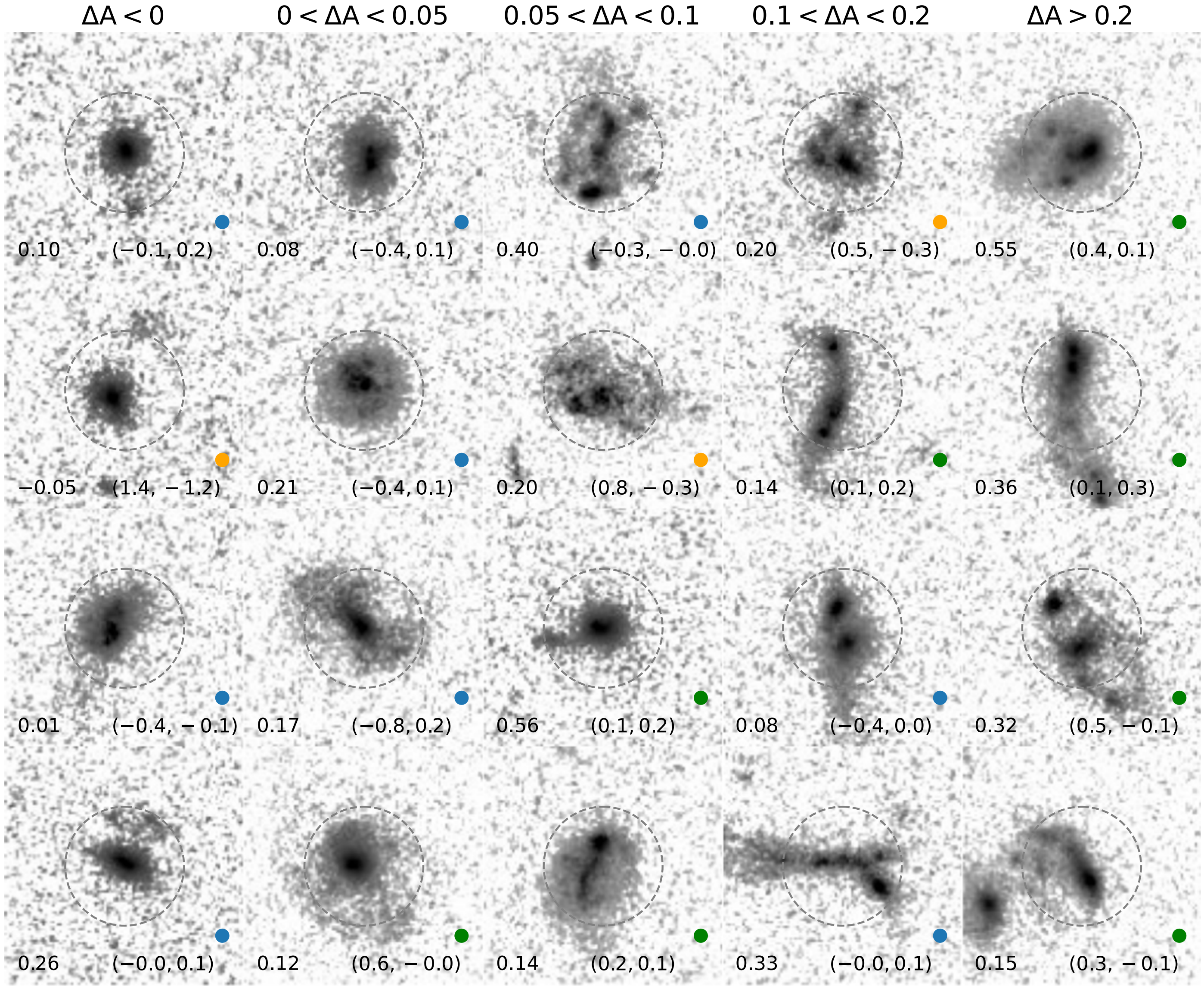

Figure 10. Example images (3″ × 3″) of DESI ELGs with different ΔA values. Each column shows the HST ACS F814W images of four randomly selected DESI ELGs within the ΔA selection, as shown at the top of the column. The ΔA values increase from left to right, respectively. The values listed on the lower-left corner of the images are the $\Delta\log_{10}$ SFR of the ELGs and the values on the right are (Coeff$_1$, Coeff$_2$). The color dots on the lower-right corner indicate the spectral regions of the ELGs, with blue, green, and orange for narrow, broad, and double-peak regions, respectively. The gray circle indicates the size of the DESI fiber (1″.5 in diameter).

distribution with possible associated emission lines (e.g., D. B. Fisher et al. 2017; P. Oliva-Altamirano et al. 2018) that contribute to the emission-line profiles. Moreover, the disk instability is expected to drive turbulence, which in turn enhances gas velocity dispersion in galaxies (e.g., N. J. Goldbaum et al. 2016). Therefore, this scenario can also produce correlations between ΔSFR, ΔA, and line profiles as observed for DESI ELGs.

The origins of clumpy structures observed in high-redshift galaxies are still under active investigation. Some studies have shown that the clump size and gas kinematics relationship of some clumpy galaxies is consistent with the prediction of the disk instability scenario (e.g., D. B. Fisher et al. 2017). However, B. Ribeiro et al. (2017) have shown that galaxies with only two major clumps tend to have clump mass being inconsistent with the expected mass based on disk instability. The authors argue that galaxies with two major clumps can be ongoing galaxy mergers. In addition, D. M. Elmegreen et al. (2021) have shown that the observed appearance of major mergers at high redshifts is similar to the observed appearance of clumpy isolated disk galaxies. These studies indicate the complexity of distinguishing clumps formed by disk instability and galaxy mergers. In the DESI ELG sample, while there are galaxies having only two major clumps, some galaxies have more than two clumps, as can be seen in Figure 10. Therefore, it is possible that DESI ELGs include both types of systems.

While estimating the contributions of these two scenarios in the overall DESI ELG population is beyond the scope of this work, one can combine multiwavelength deep images with radio observations for spatially resolved gas properties to identify tidal features, characterize the properties of the clumpy structures and resolve the nature of these galaxies.

5.2. The Origins of Double-peak Galaxies

Our results show that while ELGs in the double-peak region tend to have higher velocity offsets between two velocity components than ELGs in the broad region, ELGs in the double-peak region, especially in the D1 and D2 regions, have on average lower ΔA and $\Delta \log_{10}$ SFR than ELGs located in the broad (B) region with $\mathrm{Coeff}_1 > 0$ and $\mathrm{Coeff}_2 \simeq 0$, as shown in Figure 8.

We propose that this difference might be due to the fact that instead of tracing two galaxies during the merging process or galaxies with violent disk instability, there is a higher fraction of ELGs in the double-peak region, with the velocity components reflecting the rotating disks of the galaxies without undergoing SFR enhancement events. One possible way to explore these two scenarios is to investigate the galaxy inclination and ΔA relation of ELGs. The expectation for the rotating disk scenario is that those galaxies tend to be more edge-on with low ΔA. On the other hand, the morphology of two merging galaxies with close separation or galaxies with two major clumps can resemble the morphology of edge-on galaxies. Therefore, merging/clumpy galaxies also tend to look like edge-on galaxies but with higher ΔA than the rotating disks.

In Figure 11, we examine the distribution of ΔA and galaxy inclination, using the elongation parameter, the ratio between the lengths of semimajor and semiminor axes of galaxies, from HST images as a proxy (A. Leauthaud et al. 2007). We only select ELGs with $10 < \log_{10} M_*/M_\odot < 10.5$ to reduce possible effects associated with galaxy mass. The top panel shows the normalized number distributions of the elongation parameter of DESI ELGs in the B1 and B2 regions and D1 and D2 regions, indicated by green and orange, respectively. The middle panel shows the distribution of ΔA as a function of the elongation parameter, and the bottom panel shows the fraction of ELGs with $\Delta A > 0.1$ as a function of the elongation parameter for the two types of ELGs. We find that ELGs in the B1 and B2 regions and D1 and D2 regions both have a broad range of elongation parameters. The middle and bottom panels show that for ELGs in the D1 and D2 regions, the fraction of sources with $\Delta A > 0.1$ is consistently $\sim$0.1 across all elongation parameter values. In contrast, the fraction of ELGs in the B1 and B2 regions with $\Delta A > 0.1$ increases with the elongation parameter from 0.2 to 0.4. These observed trends are consistent with the above-proposed scenario that mergers/clumpy galaxies tend to have high elongation and high ΔA, while rotating disks tend to have high elongation but with low ΔA. However, we note that the above results are suggestive, and spatially resolved kinematics information is needed to conclusively identify the origins of the double-peak emission lines observed in DESI ELGs.

Previous studies have also suggested that double-peak emission lines observed in galaxies are associated with rotating disks in the local Universe. For example, D. Maschmann et al. (2020) found $\sim$5000 double-peak ELGs from the SDSS spectroscopic data set. Based on simulations, D. Maschmann et al. (2023) concluded that the double-peak emission lines can originate from the central rotating bars of galaxies or minor mergers. Y.-M. Chen et al. (2016) also found that rotating galaxy disks can explain the double-peak emission lines observed in SDSS disk star-forming galaxies.

5.3. DESI ELG Selection for Galaxy Properties

We now explore the key selection criteria for DESI ELGs, which preferentially include galaxies with high SFR and asymmetric morphology. The DESI ELG selections for the main survey are summarized in Table 2 of A. Raichoor et al. (2023), which includes

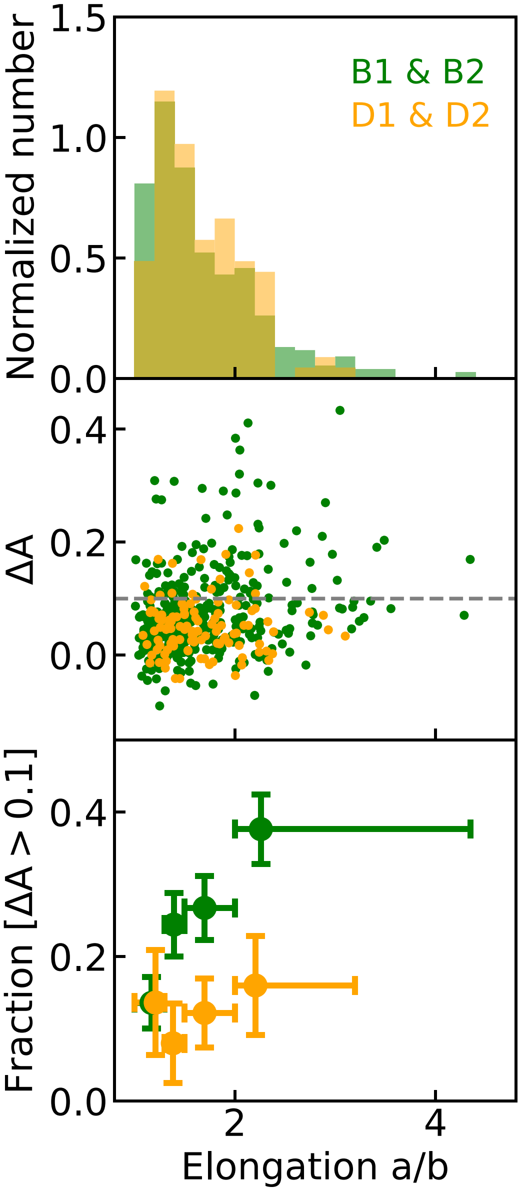

Figure 11. Relation between ΔA and galaxy elongation of ELGs with $10 < \log_{10} M_*/M_\odot < 10.5$. Top: normalized number distributions of ELGs in the B1 and B2 regions (green) and in the D1 and D2 regions (orange) as a function of galaxy elongation. Middle: ΔA distribution as a function of galaxy elongation for the two ELG samples. Bottom: fractions of ELGs with $\Delta A > 0.1$ as a function of galaxy elongation for the two ELG samples. The uncertainties are estimated by bootstrapping the sample 1000 times.

1. $g > 20$,
2. g band fiber magnitude (gfib) < 24.1, and
3. $g - r$ and $r - z$ color boundaries.

To investigate how these selection conditions affect the galaxy properties, we use the COSMOS-ZEST galaxy sample and crossmatch the sample with the photometric catalog[40] from the DESI Legacy Imaging Surveys (A. Dey et al. 2019) to obtain the photometric information used in the DESI target selections. We then select star-forming galaxies with $\log_{10} M_*/M_\odot > 9$, $1.1 < z_{\mathrm{photo}} < 1.63$ (the parameter space that covers the main DESI ELGs), and $g > 20$. For each star-forming galaxy, we follow the

---

[40] DR10 https://www.legacysurvey.org/dr10/catalogs/.

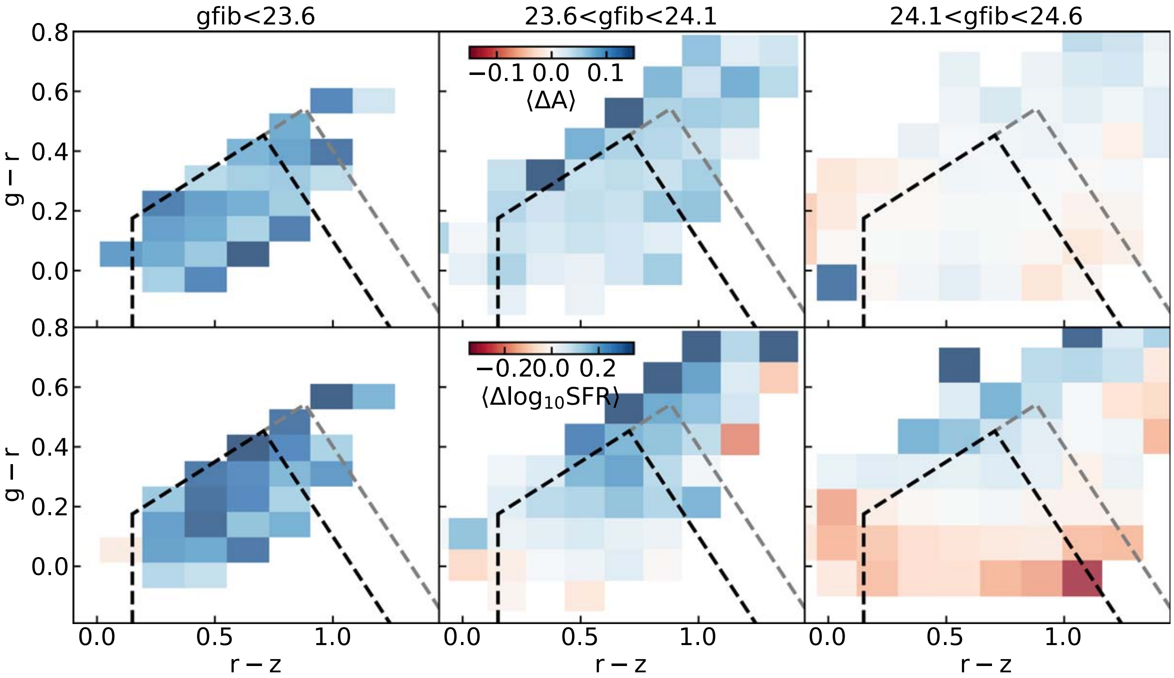

Figure 12. Median ΔA values and $\Delta\log_{10}$ SFR of star-forming galaxies as a function of $g-r$, $r-z$, and g-band fiber magnitude (gfib). Galaxies are selected with $\log_{10} M_*/M_\odot > 9$, $1.1 < z_{\rm photo} < 1.63$, and $g > 20$ from the COSMOS2020 catalog with HST ACS morphology measurements. The upper panels are median ΔA values as a function of galaxy colors with gfib < 23.6 (left), 23.6 < gfib < 24.1 (middle), and 24.1 < gfib < 24.6 (right). The lower panels are median $\Delta\log_{10}$ SFR as a function of galaxy colors with gfib < 23.6 (left), 23.6 < gfib < 24.1 (middle), and 24.1 < gfib < 24.6 (right).

same procedure for estimating ΔA and $\Delta\log_{10}$SFR as the DESI ELG sample by searching for at least 10 galaxies as references with redshift difference $|\Delta z| < 0.025$, stellar mass difference $|\Delta \log_{10} M_*/M_\odot| < 0.05$, and size difference $|\Delta \mathrm{R80}| < 0\farcs025$ and increasing the conditions if less than 10 galaxies are found. The ΔA and $\Delta\log_{10}$SFR are the differences between the values of the galaxies and the median values of the reference galaxies.

With ΔA and $\Delta\log_{10}$SFR information, we calculate the median ΔA and $\Delta\log_{10}$SFR as a function of $g-r$, $r-z$ colors, and gfib. The upper panels of Figure 12 show the median ΔA values across the $g-r$ and $r-z$ color space with gfib < 23.6 (left panel), $23.6 <$ gfib < 24.1 (middle panel), and $24.1 <$ gfib < 24.6 (right panel). The black dashed lines are the boundaries for the DESI ELG LOP sample and the gray dashed lines are the extended region for the VLO sample. As can be seen, galaxies that are brighter in gfib preferentially have higher ΔA values. A similar trend is observed for $\Delta\log_{10}$SFR as shown in the lower panels of Figure 12, indicating the median $\Delta\log_{10}$SFR values of galaxies. We note that while there is no strong correlation between ΔA and galaxy colors, a correlation between $\Delta\log_{10}$SFR and galaxy colors can be observed in the middle and right panels. Therefore, we conclude that the g-band fiber magnitude selection is the primary factor for selecting galaxies with more disturbed morphology and higher SFR than the overall star-forming galaxy population at the same redshifts, and the ($g-r$, $r-z$) color selections are an additional factor modulating the SFR. Note that while gfib ~ 24.1 reaches the limiting magnitude of the Legacy Surveys (A. Dey et al. 2019) with larger uncertainties, we have performed the same analysis with g-, r-, z-band images with depths of ∼27–28 mag from Hyper Suprime-Cam (H. Aihara et al. 2018, 2019) in the COSMOS2020 catalog, which yields consistent results.

5.4. Implications for DESI ELG Clustering Measurements

Recent studies have shown in order to reproduce the DESI ELG clustering properties, especially the small-scale signals ($\leqslant 0.2\,\mathrm{Mpc}\,h^{-1}$), and obtain physically motivated ELG-halo connection models, additional parameters are required in the halo occupation distribution modeling (e.g., H. Gao et al. 2023; A. Rocher et al. 2023; S. Yuan et al. 2023; H. Gao et al. 2024). These results indicate that the abundance of satellite ELGs in the halos depends on the properties of the central ELGs—a phenomenon called “conformity” (see R. H. Wechsler & J. L. Tinker 2018, for a review).

By exploring the properties of ELGs at $0.8 < z < 1.1$ and their morphology with the COSMOS data set, S. Yuan et al. (2023) postulate that such conformity signals are driven by galaxy-merging-induced star formation in both central and satellite galaxies. Our findings of the relationship between SFR, galaxy morphology, and the line profiles are consistent with this scenario. In addition, the small-scale enhancement of the clustering amplitude of DESI ELGs is aligned with the results of simulations focusing on galaxy mergers (e.g., A. R. Wetzel et al. 2009).

In addition to the behavior of spatial clustering properties of DESI ELGs, A. Rocher et al. (2023) find that DESI ELGs have an unexpected clustering property in velocity space. They find that, on average, the velocity dispersion of the satellite DESI ELGs is larger than the velocity dispersion of dark matter particles by ∼30%. We argue that this observed property is likely associated with spectral profiles of DESI ELGs having two velocity components along the sightlines. As shown in Figure 4, for two redshift systems, the median value of the line-of-sight velocity offsets is $\sim 150\,\mathrm{km\,s^{-1}}$. However, the current `Redrock` pipeline typically determines the redshift in the middle of the two velocity components. In this case, this can possibly introduce a $\sim 75\,\mathrm{km\,s^{-1}}$ velocity offset of the central galaxy redshifts systematically and adds additional velocity dispersion in the relative velocities between centrals and satellites.

Here, we consider a simple case to quantify possible signals. Assuming that the dark matter halo mass of ELGs is $\approx 10^{12}\,M_\odot$ (e.g., A. Rocher et al. 2023) with the corresponding velocity dispersion of dark matter particles being $\sim$100 km s^{-1}, adding a systematic 75 km s^{-1} offset in the velocity measurements will increase the estimated velocity dispersion by 25%, which is similar to the results in A. Rocher et al. (2023). This demonstrates that this redshift determination effect can be responsible for at least part of the apparent excess velocity dispersion of the satellite galaxies and it needs to be taken into account for the clustering measurements in velocity space.

6. Conclusions

By performing PCA on ∼230,000 spectra of ELGs at $0.6 < z < 1.63$ from the DESI Early Data Release, we decomposed [O II] profiles based on the derived PCA eigenspectra and explored the diversity of [O II] line profiles in low-dimensional coefficient space. We further utilized the physical and morphological properties of galaxies in the COSMOS field and investigated the relationship between the [O II] line profiles and the properties of the galaxies. Our main findings are summarized as follows:

1. We find that the first two eigenspectra, which correspond to the line width (first) and the peakedness (second) of the [O II] doublet lines, can explain ∼54% of the total variance of the [O II] line profiles. Using the coefficients of these two eigenspectra, we show that DESI ELGs can be classified into at least three types with narrow [O II] lines, broad [O II] lines, and two redshift systems, demonstrating the diversity of the [O II] line profiles.
2. Combining PCA results with the galaxy's physical properties from the COSMOS2020 catalog, we find that ELGs with broader line profiles tend to have higher stellar mass and SFR. Moreover, by fixing stellar mass, we find that ELGs with broader line profiles have higher median SFR. We also find that DESI ELGs preferentially have higher SFR than the average SFR of the star-forming galaxies at similar redshifts. These trends are observed across the entire redshift range of DESI ELGs from $z \sim 0.6$ to 1.63.
3. We include the morphological properties of DESI ELGs derived from HST ACS images and quantify the enhancement of various properties, including SFR and shape properties, of DESI ELGs with respect to the reference star-forming galaxies with similar stellar mass, sizes, and redshifts. We find that $\Delta\log_{10}$SFR correlates the most with ΔA and Coeff$_1$ with Spearman's correlation coefficients of 0.31 and 0.16, respectively. Moreover, Coeff$_1$ has the highest correlation coefficient with ΔA ($\rho_s = 0.08$ with $\sim 4\sigma$ detection) than with other morphological parameters.
4. The median ΔA and $\Delta\log_{10}$ SFR and the fraction of high ΔA and high $\Delta\log_{10}$ SFR as a function of Coeff$_1$ and Coeff$_2$ both show that ELGs with broad line profiles have higher ΔA and $\Delta\log_{10}$ SFR than ELGs with narrow or double-peak line profiles. This result reveals an underlying relationship between ΔA, $\Delta\log_{10}$ SFR, and line profiles.
5. We show that ELGs with high ΔA have on average ∼0.2 dex enhancement of the SFR than ELGs with low ΔA and the Coeff$_1$ can further contribute to the enhancement of SFR by ∼0.1 dex.

Finally, we argue that this inter-relationship between the line profiles, and physical and morphological properties of DESI ELGs can be naturally explained by both the galaxy merger and disk instability star-forming clumps scenarios.

The results of this work show that the large DESI spectroscopic data set opens a new window for statistically investigating the galaxy's physical properties and how the mechanism drives galaxy evolution at $1 < z < 1.6$. The combination of the DESI data and upcoming imaging data sets provided by space telescopes, such as Euclid (Euclid Collaboration et al. 2022) and the Roman Space Telescope (R. Akeson et al. 2019), will further offer a large galaxy sample with both kinematics and morphological information of galaxies for galaxy evolution science.

Understanding how the physical properties of galaxies link to the properties of dark matter halos and the large-scale environments is crucial for obtaining precise cosmological measurements. Recent galaxy clustering results of DESI ELGs indicate that the standard galaxy–halo connection models are insufficient to describe small-scale clustering measurements in both spatial and velocity space. As demonstrated in this work, this is possibly due to the nature of DESI ELGs and the redshift determination effect due to their diverse line profiles. It will be informative to perform clustering measurements as a function of emission-line profiles and obtain a better understanding of small-scale clustering properties of the DESI ELG population. This approach can be used to utilize the hidden spectral information from the ongoing and upcoming cosmological spectroscopic surveys, including PFS (M. Takada et al. 2014), Euclid (Euclid Collaboration et al. 2022), the Roman Space Telescope (R. Akeson et al. 2019), and the next-generation surveys, e.g., the DESI-II and Spec-5 experiments (D. J. Schlegel et al. 2022), which use ELGs as primary tracers for the large-scale structure of the Universe.

Acknowledgments

We thank the anonymous referee for the constructive report. We also want to thank Anand Raichoor and Chiara Circosta for their comments and suggestions for the early version of this paper and John Weaver for his help with using the COSMOS2020 catalog. T.W.L. thanks Chian-Chou Chen for the discussions and suggestions. T.W.L. was supported by the National Science and Technology Council (MOST 111-2112-M-002-015-MY3, NSTC 113-2112-M-002-028-MY3), the Ministry of Education, Taiwan (MOE Yushan Young Scholar grant NTU-110VV007, NTU-111V1007-2, NTU-112V1007-3, NTU-113V1007-4), National Taiwan University research grant (NTU-CC-111L894806, NTU-CC-112L893606, NTU-CC-113L891806, NTU-111L7318, NTU-112L7302). J.M. gratefully acknowledges funding support for this work from the U.S. Department of Energy, Office of Science, Office of High Energy Physics under Award Number DE-SC0020086. M.S. acknowledges support by the Polish National Agency for Academic Exchange (Bekker grant BPN/BEK/2021/1/00298/DEC/1), the State Research Agency of the Spanish Ministry of Science and Innovation under the grants "Galaxy Evolution with Artificial Intelligence" (PGC2018-100852-A-

I00) and "BASALT" (PID2021-126838NB-I00). This work was partially supported by the European Union's Horizon 2020 Research and Innovation program under the Maria Skłodowska-Curie grant agreement (No. 754510).

This material is based upon work supported by the U.S. Department of Energy (DOE), Office of Science, Office of High-Energy Physics, under contract No. DE-AC02-05CH11231, and by the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility under the same contract. Additional support for DESI was provided by the U.S. National Science Foundation (NSF), Division of Astronomical Sciences under Contract No. AST-0950945 to the NSF's National Optical-Infrared Astronomy Research Laboratory; the Science and Technology Facilities Council of the United Kingdom; the Gordon and Betty Moore Foundation; the Heising-Simons Foundation; the French Alternative Energies and Atomic Energy Commission (CEA); the National Council of Humanities, Science and Technology of Mexico (CONAHCYT); the Ministry of Science and Innovation of Spain (MICINN), and by the DESI Member Institutions: https://www.desi.lbl.gov/collaborating-institutions.

The DESI Legacy Imaging Surveys consist of three individual and complementary projects: the Dark Energy Camera Legacy Survey (DECaLS), the Beijing-Arizona Sky Survey (BASS), and the Mayall *z*-band Legacy Survey (MzLS). DECaLS, BASS, and MzLS together include data obtained, respectively, at the Blanco telescope, Cerro Tololo Inter-American Observatory, NSF's NOIRLab; the Bok telescope, Steward Observatory, University of Arizona; and the Mayall telescope, Kitt Peak National Observatory, NOIRLab. NOIRLab is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation. Pipeline processing and analyses of the data were supported by NOIRLab and the Lawrence Berkeley National Laboratory. Legacy Surveys also uses data products from the Near-Earth Object Wide-field Infrared Survey Explorer (NEOWISE), a project of the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Legacy Surveys was supported by the Director, Office of Science, Office of High Energy Physics of the US Department of Energy; the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility; the US National Science Foundation, Division of Astronomical Sciences; the National Astronomical Observatories of China, the Chinese Academy of Sciences, and the Chinese National Natural Science Foundation. LBNL is managed by the Regents of the University of California under contract with the US Department of Energy. The complete acknowledgments can be found at https://www.legacysurvey.org/.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the US National Science Foundation, the US Department of Energy, or any of the listed funding agencies.

The authors are honored to be permitted to conduct scientific research on Iolkam Du'ag (Kitt Peak), a mountain with particular significance to the Tohono O'odham Nation.

Data Availability

All data points shown in the figures are available at Zenodo doi:10.5281/zenodo.13148442.

Software: Astropy (Astropy Collaboration et al. 2013, 2018, 2022), `NumPy` (C. R. Harris et al. 2020), `SciPy` (P. Virtanen et al. 2020), `matplotlib` (J. D. Hunter 2007), `AstroML` (J. T. Vanderplas et al. 2012), `Kapteyn Package` (J. P. Terlouw & M. G. R. Vogelaar 2014).

Appendix A
Removing AGN

It is possible that a fraction of DESI ELGs have actively accreting supermassive black holes (AGN), which can produce [O II] lines. To eliminate the complexity due to the AGN contributions, we utilize available line information obtained via the `FastSpecFit`[41] algorithm developed by J. Moustakas et al. (2024, in preparation). We identify DESI ELGs that satisfy *any* of the following conditions and remove them from our analysis:

$$\log_{10} F_{[\mathrm{O\,III}]\lambda5007}/F_{\mathrm{H}_\beta} > 0.2 \text{ and line width}_{[\mathrm{O\,III}]\lambda5007}\ \mathrm{FWHM} > 300\ \mathrm{km\ s^{-1}}, \tag{A1}$$

$$\mathrm{S/N}(F_{\mathrm{Mg\,II}\lambda2796}) > 2, \tag{A2}$$

$$\mathrm{S/N}(F_{[\mathrm{Ne\,V}]\lambda3426}) > 2. \tag{A3}$$

The first condition (A1) follows the condition used in N. L. Zakamska et al. (2003) and R. Reyes et al. (2008) for selecting Type II quasars at $0.4 < z < 0.9$ where only [O III] λ5007 and H_β lines are accessible in the optical wavelength coverage. In N. L. Zakamska et al. (2003), the selections are $\log_{10} F_{[\mathrm{O\,III}]\lambda5007}/F_{\mathrm{H}_\beta} > 0.3$ and line $\mathrm{width}_{[\mathrm{O\,III}]5007}$ FWHM > 400 km s^{-1}. In our condition, we reduce the line ratio by 0.1 dex and the line width by 100 km s^{-1} so that our selection is conservative.

The second condition, the signal-to-noise cut of the Mg II emission line flux (Equation (A2)), is used to remove ELG spectra with possible broad Mg II emission lines originating from quasars, e.g., the missing quasar population in D. M. Alexander et al. (2023).

The third condition, the signal-to-noise cut of the [Ne V] λ3426 emission line flux (A3), is motivated by the fact that the existence of the [Ne V]λ3426 line requires relatively energetic radiation, and AGN has been considered as the main source for producing the required radiation (e.g., N. Maddox 2018; N. J. Cleri et al. 2023).

Approximately 4.3% of the total sources in the original DESI ELG sample meet the above conditions. Individually, Equations (A1), (A2), and (A3) contribute ∼0.7%, ∼0.9%, and ∼3.0% respectively. The final 4.3% includes sources in common.

Appendix B
Morphology Parameters as a Function of Line Profiles

Here, we summarize the relationships between the morphological parameters and the line profiles. Figure 13 shows the median values of the morphological parameters, galaxy size (R80), Gini coefficient (*G*), the second-order moment of the brightest 20% pixels (M20), concentration (*C*), and asymmetry (*A*) from left to right respectively, as a function of Coeff_1 and Coeff_2. These results are based on the DESI-ZEST sample.

Trends can be observed from Figure 13. First, DESI ELGs in the double-peak region tend to have larger sizes than DESI ELGs in other regions, as shown in the first panel. This is

---

[41] https://fastspecfit.readthedocs.io/en/latest/index.html

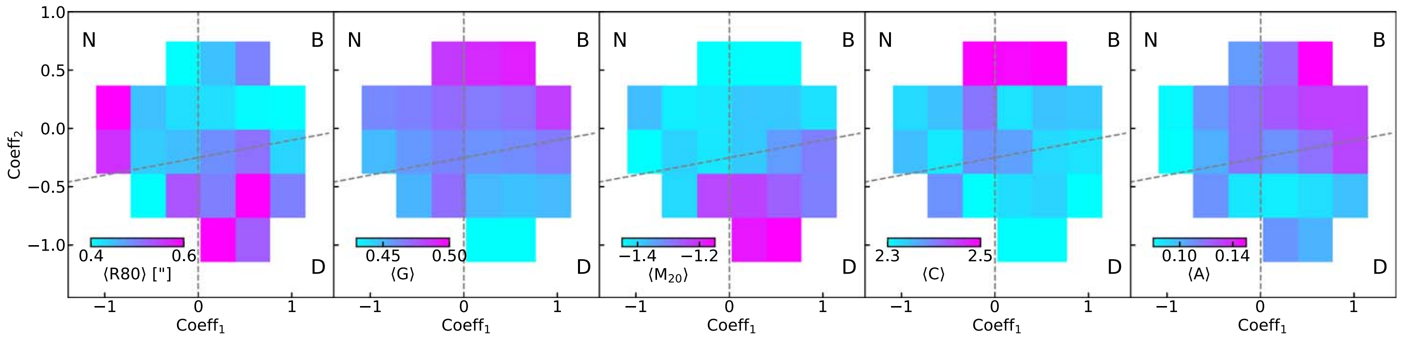

Figure 13. Morphological parameters as a function of Coeff_1 and Coeff_2. Median values of galaxy size (R80 in units of arcsecs), Gini coefficient (*G*), second-order moment of the brightest 20% pixels (M20), concentration (*C*), and asymmetry (*A*) are shown from left to right, respectively.

consistent with the fact that DESI ELGs in the double-peak region also have higher stellar mass. Correlating with R80, the M20 values also tend to be larger in the double-peak region. We note that the DESI ELGs with $\mathrm{Coeff}_1 \sim -1$ have, on average, lower redshifts and therefore have larger observed sizes. Second, the Gini coefficient (*G*) and asymmetry parameter (*A*) behave similarly with higher median values in the broad region, indicating that DESI ELGs in the broad region tend to have relatively disturbed morphology. Finally, the concentration *C* anticorrelates with M20, having the highest median values for DESI ELGs with high Coeff_2. These internal correlations between parameters are consistent with the results summarized in C. Scarlata et al. (2007). We note that while these morphological parameters all have some dependencies on the line profiles, we focus on the asymmetry parameter in this work given that the excess asymmetry parameter ΔA yields the highest Spearman's correlation coefficients with $\Delta\log_{10}$ SFR and Coeff_1 and Coeff_2 as shown in Section 4.

ORCID iDs

Ting-Wen Lan https://orcid.org/0000-0001-8857-7020
J. Xavier Prochaska https://orcid.org/0000-0002-7738-6875
John Moustakas https://orcid.org/0000-0002-2733-4559
Małgorzata Siudek https://orcid.org/0000-0002-2949-2155
S. Ahlen https://orcid.org/0000-0001-6098-7247
D. Bianchi https://orcid.org/0000-0001-9712-0006
D. Brooks https://orcid.org/0000-0002-8458-5047
S. Cole https://orcid.org/0000-0002-5954-7903
K. Dawson https://orcid.org/0000-0002-0553-3805
A. de la Macorra https://orcid.org/0000-0002-1769-1640
J. E. Forero-Romero https://orcid.org/0000-0002-2890-3725
E. Gaztañaga https://orcid.org/0000-0001-9632-0815
S. Gontcho A Gontcho https://orcid.org/0000-0003-3142-233X
J. Guy https://orcid.org/0000-0001-9822-6793
R. Kehoe https://orcid.org/0000-0002-7101-697X
T. Kisner https://orcid.org/0000-0003-3510-7134
M. Landriau https://orcid.org/0000-0003-1838-8528
A. Meisner https://orcid.org/0000-0002-1125-7384
R. Miquel https://orcid.org/0000-0002-6610-4836
J. A. Newman https://orcid.org/0000-0001-8684-2222
F. Prada https://orcid.org/0000-0001-7145-8674
E. Sanchez https://orcid.org/0000-0002-9646-8198
H. Seo https://orcid.org/0000-0002-6588-3508
D. Sprayberry https://orcid.org/0000-0001-7583-6441
G. Tarlé https://orcid.org/0000-0003-1704-0781
H. Zou https://orcid.org/0000-0002-6684-3997

References

Abraham, R. G., van den Bergh, S., & Nair, P. 2003, ApJ, 588, 218
Aihara, H., AlSayyad, Y., Ando, M., et al. 2019, PASJ, 71, 114
Aihara, H., Arimoto, N., Armstrong, R., et al. 2018, PASJ, 70, S4
Akeson, R., Armus, L., Bachelet, E., et al. 2019, arXiv:1902.05569
Alexander, D. M., Davis, T. M., Chaussidon, E., et al. 2023, AJ, 165, 124
Anand, A., Guy, J., Bailey, S., et al. 2024, AJ, 168, 124
Astropy Collaboration, Price-Whelan, A. M., Lim, P. L., et al. 2022, ApJ, 935, 167
Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123
Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33
Behroozi, P. S., Zhu, G., Ferguson, H. C., et al. 2015, MNRAS, 450, 1546
Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393
Blake, C., Kazin, E. A., Beutler, F., et al. 2011, MNRAS, 418, 1707
Bottrell, C., Yesuf, H. M., Popping, G., et al. 2024, MNRAS, 527, 6506
Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, ApJ, 686, 1503
Brodzeller, A., Dawson, K., Bailey, S., et al. 2023, AJ, 166, 66
Chabrier, G. 2003, PASP, 115, 763
Chaussidon, E., Yèche, C., Palanque-Delabrouille, N., et al. 2023, ApJ, 944, 107
Chen, Y.-M., Gu, Q.-S., Tremonti, C. A., Shi, Y., & Jin, Y. 2016, MNRAS, 459, 3861
Cleri, N. J., Yang, G., Papovich, C., et al. 2023, ApJ, 948, 112
Comparat, J., Kneib, J.-P., Bacon, R., et al. 2013, A&A, 559, A18
Comparat, J., Richard, J., Kneib, J.-P., et al. 2015, A&A, 575, A40
Conselice, C. J. 2014, ARA&A, 52, 291
Cooper, A. P., Koposov, S. E., Allende Prieto, C., et al. 2023, ApJ, 947, 37
COSMOS team 2007, COSMOS Zurich Structure & Morphology Catalog, IPAC, doi:10.26131/IRSA160
Dawson, K. S., Kneib, J.-P., Percival, W. J., et al. 2016, AJ, 151, 44
Dekel, A., Sari, R., & Ceverino, D. 2009, ApJ, 703, 785
Delchambre, L. 2015, MNRAS, 446, 3545
DESI Collaboration, Abareshi, B., Aguilar, J., et al. 2022, AJ, 164, 207
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024a, AJ, 167, 62
DESI Collaboration, Adame, A. G., Aguilar, J., et al. 2024b, AJ, 168, 58
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016a, arXiv:1611.00036
DESI Collaboration, Aghamousa, A., Aguilar, J., et al. 2016b, arXiv:1611.00037
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Drinkwater, M. J., Jurek, R. J., Blake, C., et al. 2010, MNRAS, 401, 1429
Eisenstein, D. J., Zehavi, I., Hogg, D. W., et al. 2005, ApJ, 633, 560
Ellison, S. L., Patton, D. R., Simard, L., et al. 2008, AJ, 135, 1877
Ellison, S. L., Patton, D. R., Simard, L., et al. 2010, MNRAS, 407, 1514
Elmegreen, D. M., Elmegreen, B. G., Whitmore, B. C., et al. 2021, ApJ, 908, 121
Euclid Collaboration, Scaramella, R., Amiaux, J., et al. 2022, A&A, 662, A112
Fisher, D. B., Glazebrook, K., Abraham, R. G., et al. 2017, ApJL, 839, L5
Gao, H., Jing, Y. P., Gui, S., et al. 2023, ApJ, 954, 207
Gao, H., Jing, Y. P., Xu, K., et al. 2024, ApJ, 961, 74
Goldbaum, N. J., Krumholz, M. R., & Forbes, J. C. 2016, ApJ, 827, 28
Gould, K. M. L., Brammer, G., Valentino, F., et al. 2023, AJ, 165, 248
Guo, Y., Ferguson, H. C., Bell, E. F., et al. 2015, ApJ, 800, 39
Guy, J., Bailey, S., Kremin, A., et al. 2023, AJ, 165, 144

Hahn, C., Wilson, M. J., Ruiz-Macias, O., et al. 2023, AJ, 165, 253
Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Natur, 585, 357
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2008, ApJS, 175, 356
Hunter, J. D. 2007, CSE, 9, 90
Ilbert, O., Arnouts, S., McCracken, H. J., et al. 2006, A&A, 457, 841
Jolliffe, I. 2002, Principal Component Analysis (Berlin: Springer), 338
Kaasinen, M., Bian, F., Groves, B., et al. 2017, MNRAS, 465, 3220
Kassin, S. A., Weiner, B. J., Faber, S. M., et al. 2012, ApJ, 758, 106
Kennicutt, R. C., Jr. 1998, ARA&A, 36, 189
Kennicutt, R. C., Jr., Hao, C.-N., Calzetti, D., et al. 2009, ApJ, 703, 1672
Kewley, L. J., Nicholls, D. C., & Sutherland, R. S. 2019, ARA&A, 57, 511
Koekemoer, A. M., Aussel, H., Calzetti, D., et al. 2007, ApJS, 172, 196
Kriek, M., & Conroy, C. 2013, ApJL, 775, L16
Lan, T.-W., Tojeiro, R., Armengaud, E., et al. 2023, ApJ, 943, 68
Lang, D., Hogg, D. W., & Mykytyn, D., 2016 The Tractor: Probabilistic astronomical source detection and measurement, Astrophysics Source Code Library, ascl:1604.008
Law, D. R., Belfiore, F., Bershady, M. A., et al. 2022, ApJ, 928, 58
Leauthaud, A., Massey, R., Kneib, J.-P., et al. 2007, ApJS, 172, 219
Levi, M., Bebek, C., Beers, T., et al. 2013, arXiv:1308.0847
Lotz, J. M., Jonsson, P., Cox, T. J., & Primack, J. R. 2008, MNRAS, 391, 1137
Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163
Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415
Maddox, N. 2018, MNRAS, 480, 5203
Mai, Y., Croom, S. M., Wisnioski, E., et al. 2024, MNRAS, 533, 3878
Mandelker, N., Dekel, A., Ceverino, D., et al. 2014, MNRAS, 443, 3675
Martin, A., Guo, Y., Wang, X., et al. 2023, ApJ, 955, 106
Maschmann, D., Halle, A., Melchior, A.-L., et al. 2023, A&A, 670, A46
Maschmann, D., Melchior, A.-L., Mamon, G. A., et al. 2020, A&A, 641, A171
McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al. 2012, A&A, 544, A156
Miller, T. N., Doel, P., Gutierrez, G., et al. 2024, AJ, 168, 95
Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50
Moustakas, J., Kennicutt, R. C., & Tremonti, C. A. 2006, ApJ, 642, 775
Moustakas, J., Lang, D., Dey, A., et al. 2023, ApJS, 269, 3
Murata, K. L., Kajisawa, M., Taniguchi, Y., et al. 2014, ApJ, 786, 15
Myers, A. D., Moustakas, J., Bailey, S., et al. 2023, AJ, 165, 50
Oliva-Altamirano, P., Fisher, D. B., Glazebrook, K., et al. 2018, MNRAS, 474, 522
Osterbrock, D. E., & Ferland, G. J. 2006, in Astrophysics of Gaseous Nebulae and Active Galactic Nuclei, ed. D. E. Osterbrock & G. J. Ferland (2nd ed.; Sausalito, CA: Univ. Science Books), 2006
Patton, D. R., Wilson, K. D., Metrow, C. J., et al. 2020, MNRAS, 494, 4969
Raichoor, A., de Mattia, A., Ross, A. J., et al. 2021, MNRAS, 500, 3254
Raichoor, A., Moustakas, J., Newman, J. A., et al. 2023, AJ, 165, 126
Reyes, R., Zakamska, N. L., Strauss, M. A., et al. 2008, AJ, 136, 2373
Ribeiro, B., Le Fèvre, O., Cassata, P., et al. 2017, A&A, 608, A16
Rocher, A., Ruhlmann-Kleider, V., Burtin, E., et al. 2023, JCAP, 2023, 016
Sattari, Z., Mobasher, B., Chartab, N., et al. 2023, ApJ, 951, 147
Scarlata, C., Carollo, C. M., Lilly, S., et al. 2007, ApJS, 172, 406
Schlafly, E. F., Kirkby, D., Schlegel, D. J., et al. 2023, AJ, 166, 259
Schlegel, D. J., Ferraro, S., Aldering, G., et al. 2022, arXiv:2209.03585
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
Scoville, N., Abraham, R. G., Aussel, H., et al. 2007, ApJS, 172, 38
Scoville, N., Aussel, H., Brusa, M., et al. 2007, ApJS, 172, 1
Silber, J. H., Fagrelius, P., Fanning, K., et al. 2023, AJ, 165, 9
Suzuki, N. 2006, ApJS, 163, 110
Taghizadeh-Popp, M., Heinis, S., & Szalay, A. S. 2012, ApJ, 755, 143
Takada, M., Ellis, R. S., Chiba, M., et al. 2014, PASJ, 66, R1
Terlouw, J. P., & Vogelaar, M. G. R., 2016 Kapteyn Package, Astrophysics Source Code Library, ascl:1611.010
Übler, H., Genzel, R., Wisnioski, E., et al. 2019, ApJ, 880, 48
Vanderplas, J. T., Connolly, A. J., Ivezić, Ž., & Gray, A. 2012, in Proc. CIDU, 47
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Weaver, J. R., Kauffmann, O. B., Ilbert, O., et al. 2022, ApJS, 258, 11
Wechsler, R. H., & Tinker, J. L. 2018, ARA&A, 56, 435
Wetzel, A. R., Cohn, J. D., & White, M. 2009, MNRAS, 394, 2182
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Yesuf, H. M., Ho, L. C., & Faber, S. M. 2021, ApJ, 923, 205
Yip, C. W., Connolly, A. J., Szalay, A. S., et al. 2004, AJ, 128, 585
Yuan, S., Wechsler, R. H., Wang, Y., et al. 2023, arXiv:2310.09329
Zakamska, N. L., Strauss, M. A., Krolik, J. H., et al. 2003, AJ, 126, 2125
Zhou, R., Dey, B., Newman, J. A., et al. 2023, AJ, 165, 58