Astronomy & Astrophysics manuscript no. output
June 17, 2026
©ESO 2026

Ubiquitous CO faintness and the prevalence of CO-dark gas in compact emission-line galaxies

Siqi Liu [1], Xucheng Yin[1], A-Li Luo[2,3,4]⋆, Wei Zhang [2], Ping Yan[5], and Qiu Bo[6]⋆⋆

[1] School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing 100083, China.
[2] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China
[3] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 101408, China.
[4] University of Chinese Academy of Sciences, Nanjing, Jiangsu, 211135, China
[5] Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210008, China
[6] Shunde Innovation School, University of Science and Technology Beijing, Beijing, China

June 17, 2026

ABSTRACT

Context. Green Pea and Blueberry galaxies are compact, low-metallicity starbursts that provide nearby analogs of high-redshift star-forming systems.
Aims. We investigate whether their weak CO emission reflects genuinely small molecular reservoirs or CO-dark gas where independent tracers are available.
Methods. We compiled 60 strong [O iii]λ5007 compact emission-line galaxies with partially overlapping gas-tracer coverage: ^{12}CO(1–0) observations from the Delingha 13.7 m telescope, H i 21 cm data from Arecibo, GBT, and FAST (ten detections and 41 upper limits), and Herschel/PACS [C ii] $\lambda158\,\mu$m data (five detections and four upper limits after excluding merger systems). Quantitative conclusions are drawn only from detected or physically constraining subsets. Molecular masses were estimated using a metallicity-dependent CO-to-H_2 conversion factor and the [C ii]-to-H_2 calibration for low-metallicity dwarf galaxies. Two CO upper limits are deep enough to constrain the expected molecular reservoir. We also removed major mergers and examined the large-scale environment using the SDSS DR7 VoidFinder catalog.
Results. For the five [C ii]-detected sources, we find $M_{\mathrm{H_2}}^{[\mathrm{CII}]} \sim 10^{9.0}$–$10^{9.7}\,M_\odot$ and a median molecular gas depletion time of $\tau_{\mathrm{dep,H_2}} \sim$ 0.16 Gyr, shorter than the $\sim$ 1 Gyr typical of main-sequence galaxies. H i-detected sources have a median atomic depletion time of $\tau_{\mathrm{dep,HI}} \sim 2.6$ Gyr. The detected subsets show high gas fractions and short depletion times consistent with rapid gas cycling, but the [C ii] result remains limited by small-number statistics. About 39% of the volume-matched sample lies in cosmic voids.
Conclusions. CO alone misses a substantial fraction of the star-forming fuel in these low-metallicity starbursts. Combined H i, [C ii], and deep CO constraints are needed to characterize their interstellar medium and CO-dark molecular component.

Key words. Compact galaxies – Interstellar molecules – Molecular gas – Interstellar atomic gas – Star formation

1. Introduction

Strong [O iii]λ5007 compact emission-line galaxies are low-metallicity starburst systems characterized by large [O iii] equivalent widths, compact optical morphologies, high ionization parameters, elevated star formation rates (SFRs), and high specific SFRs (sSFRs) (Cardamone et al. 2009; Yang et al. 2017; Jiang et al. 2019; Liu et al. 2023a). In the literature, the terminology is not fully uniform: Blueberry galaxies generally refer to the lowest-redshift and lowest-mass analogs ($z \lesssim 0.05$), classical Green Pea galaxies (GPs) occupy the intermediate redshift regime ($z \sim 0.1$–0.35), and the LAMOST strong-[O iii] catalogs extend the same selection to somewhat higher redshift systems sometimes called Purple Grapes (Yang et al. 2017; Liu et al. 2023a). Because these groups share the defining physical properties of compactness, strong nebular emission, low metallicity, and bursty star formation, we refer to the combined set as compact emission-line galaxies throughout this work, while retaining the original survey labels where needed.

The Schmidt-Kennicutt (S-K) law describes the tight correlation between star formation surface density and molecular gas surface density, observed using various tracers (Kennicutt et al. 2007; Leroy et al. 2008; Bigiel et al. 2008; Kennicutt & Evans 2012; Genzel et al. 2012; Leroy et al. 2013; Kumari et al. 2020). This relationship underscores the critical role of molecular gas in fueling star formation. However, in low-metallicity environments such as Green Pea galaxies, the standard CO-to-H_2 conversion factor (α_{CO}) is expected to be significantly elevated above the Galactic value of 4.3 $M_\odot\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$ due to reduced dust shielding and enhanced CO photodissociation (Bolatto et al. 2013; Amorín et al. 2016; Accurso et al. 2017), complicating the interpretation of molecular gas masses derived from CO observations.

Despite their rich cold gas content, current measurements of molecular hydrogen (H_2) in external galaxies, especially at the low-stellar-mass end ($M_\star < 10^9\,M_\odot$), remain scarce. Green Pea galaxies offer an ideal laboratory for probing molecular gas in low-mass galaxies due to their relatively low stellar masses and intense star formation activity. A key unresolved issue is the composition of the gas that fuels star formation in these systems, including the interconversion between different gas phases

⋆ Corresponding author: lal@nao.cas.cn
⋆⋆ Corresponding author: qiubo@ustb.edu.cn

Article published by EDP Sciences, to be cited as https://doi.org/10.1051/0004-6361/202660286

(atomic, molecular, and ionized). The role of the large-scale cosmic environment in shaping the interstellar medium of these compact emission-line galaxies remains comparatively underexplored. Previous investigations by Brunker et al. (2022) indicated that while a majority of such galaxies reside in low-density wall structures, a nonnegligible fraction (∼ 20%) are found within cosmic voids. Recent studies have continued to expand the census of compact metal-poor star-forming systems and their gas content (e.g., Breneman et al. 2025; Khasnovis et al. 2025; Chandola et al. 2024). Here, we include an environmental analysis as supporting context, while discussing the possible connection with gas retention and environmental gas-removal mechanisms in Sect. 4.4.

Observational data can provide insights into the M_{H_2}/M_{HI} ratio at the low-mass end, investigating whether these galaxies have a lower molecular-to-atomic gas ratio compared to more massive galaxies (Saintonge & Catinella 2022). Additionally, studies have suggested the presence of substantial CO-dark H_2 gas in low-metallicity galaxies, traceable via the [C II]λ158 μm emission line (Madden et al. 2020). Combining CO-traced and CO-dark H_2 gas enables a comprehensive assessment of the total H_2 gas content, molecular gas mass fraction, and gas depletion timescales.

To address these questions, we present a multiwavelength analysis of the cold interstellar medium (ISM) in a parent compilation of 60 compact emission-line galaxies selected from existing optical catalogs and followed up in one or more cold-gas tracers. The strength of this dataset is its multi-tracer perspective rather than the size of any single detected subsample. It integrates: (1) ^{12}CO(1–0) non-detections from the Delingha 13.7 m telescope, including two integrations deep enough to provide physically constraining upper limits on CO-traced H_2 mass; (2) H I 21 cm data from Arecibo, GBT, and FAST, including ten detections and 41 upper limits from Kanekar et al. (2021), Khasnovis et al. (2025), and Chandola et al. (2024);

(3) [C II] 158 μm observations from Herschel/PACS for nine sources (five detections after excluding mergers, non-catalog sources, and the marginal case HAE0840).

A critical improvement over previous work is the adoption of a metallicity-dependent α_{CO} following Amorín et al. (2016), which yields conversion factors ∼ 5–25× higher than the Galactic value for our targets. This approach provides more physically motivated upper limits on the CO-bright molecular gas mass. We interpret our findings within the framework of gas scaling relations from the xCOLD GASS survey (Saintonge et al. 2017) and the review by Saintonge & Catinella (2022), quantifying how compact emission-line galaxies deviate from the norms of the general galaxy population in the local Universe.

The structure of this paper is as follows. In Sect. 2, we describe our galaxy sample and the multiwavelength observational data, including the systematic morphological inspection used to exclude merger systems. In Sect. 3, we detail the derivation of key physical parameters, including SFRs, stellar masses, gas-phase metallicities, and estimates of atomic and molecular gas masses. In Sect. 4, we present scaling relations involving gas and stellar content, comparing our sample with reference datasets from xGASS and xCOLD GASS, and examine how atomic and molecular gas fractions influence the star-forming main sequence (SFMS). We summarize our findings and conclude in Sect. 5.

In this study, we adopted the Wilkinson Microwave Anisotropy Probe (WMAP) 9 cosmology (Hinshaw et al. 2013) with $H_0 = 69.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_m = 0.287$, and a flat Universe. We used AB magnitudes (Oke & Gunn 1983), a Chabrier initial mass function (IMF; Chabrier 2003), and Bruzual & Charlot (2003) stellar population synthesis (SPS) models.

2. Sample and observations

In this section, we describe the construction of our sample and the multiwavelength observations employed in our analysis. The foundational sample consists of spectroscopically confirmed compact emission-line galaxies identified through the LAMOST (Wang et al. 1996; Su & Cui 2004; Luo et al. 2012; Liu et al. 2022, 2023a) and SDSS (Cardamone et al. 2009) surveys. We combined three complementary datasets to assemble a sample of 60 galaxies with observations in at least one of H I, CO, and [C II].

2.1. Parent sample, selection, and tracer coverage

The parent optical catalogs used in this work select compact, strong [O III] emitters rather than a homogeneous gas-selected sample. The Cardamone et al. (2009) Green Pea sample was identified from SDSS imaging and spectroscopy based on their distinctive green broadband colors produced by strong [O III]λ5007 emission; the Yang et al. (2017) Blueberry sample extends this compact, high-excitation population to lower stellar masses and redshifts ($z \lesssim 0.05$); and the LAMOST catalogs of Liu et al. (2022) and Liu et al. (2023a) further broaden the redshift and luminosity coverage of strong [O III] emitters identified through fiber spectroscopy. Together, these catalogs define a parent population united by compact morphology, high ionization, and elevated specific SFRs, but they do not constitute a volume- or mass-complete census. The final compilation used here spans $0.012 \leq z \leq 0.213$, $\log(M_\star/M_\odot) \approx 6.0$–9.4, 12 + log(O/H) ≈ 7.15–8.30, and SFRs from $\sim 4 \times 10^{-3}$ to $70\,M_\odot\,\mathrm{yr}^{-1}$.

The 60 galaxies are not observed in all three gas tracers. Table 1 summarizes the tracer coverage and shows explicitly that the detected samples are much smaller than the parent optical compilation. We therefore treated the [C II] detections, the H I detections, and the deepest CO non-detections as complementary constraints rather than a statistically complete 60-object molecular-gas sample.

Table 1: Gas-tracer coverage

Tracer	Covered	Detections	Upper limits
H I 21 cm	51	10	41
^{12}CO(1–0)	14	0	14 (2 constraining)
[C II] 158 μm	9	5	4

Notes: The categories overlap and therefore should not be summed. “Constraining upper limits” denote non-detections deep enough to place the gas mass below or comparable to the molecular mass expected from independent tracers or scaling relations.

Figure 2 summarizes the parameter space and tracer coverage of the final sample. The present work is biased toward nearby, gas-rich, or highly star-forming systems because cold-gas follow-up is observationally expensive, especially for low-metallicity dwarf galaxies. The H I-observed subset is concentrated at $z < 0.1$ because of radio frequency coverage and sensitivity, while the [C II] subset is drawn from a single Herschel program. Conclusions based only on the five [C II] detections are therefore interpreted as case studies of compact starbursts rather than population averages.

2.2. H I 21 cm observations

Our H I data were drawn from three sources: the Arecibo/GBT sample of Kanekar et al. (2021), the GBT Blueberry follow-up of Khasnovis et al. (2025), and the FAST Blueberry observations of Chandola et al. (2024). These literature samples partly overlap and include systems that are excluded from the present analysis because of merger morphology or because they fall outside the final compact-[O III] gas-follow-up compilation.

After applying the same merger exclusions and overlap removal used throughout the paper, the final H I subset contains 51 unique sources: ten detections and 41 upper limits. The upper limits are retained in the gas-fraction and depletion-time diagrams as censored constraints, but they were not treated as detections when computing medians or detection rates.

We also examined data from the blind CRAFTS survey conducted with FAST (Nan 2006; Nan et al. 2011; Li & Pan 2016). Two sources (J0847+2944 and J0856+3111) fall within the CRAFTS footprint, but tentative signals were found to be inconsistent with the expected radial velocities and were treated as non-detections likely caused by radio frequency interference. We retained these two sources in our sample catalog for completeness; however, due to the shallow nature of the blind survey and the likely RFI contamination, their H I upper limits are not physically constraining and were excluded from subsequent statistical scaling relations to avoid biasing the results.

In total, our sample includes 51 unique sources with H I data: ten detections and 41 upper limits. The H I detection rate is 20% (10/51).

2.3. CO observations from PMO

We were awarded observing time for ^{12}CO(1–0) observations of Green Pea galaxies using the 13.7-m millimeter-wave telescope at the Delingha Millimeter-wave Observatory, operated by the Purple Mountain Observatory (hereafter PMO), under observing programs 23A005, 23B006, 25B010, and 25B011. Observations were conducted using the upper sideband receiver centered at 115 GHz, targeting the ^{12}CO(1–0) transition. The ON–OFF spectral line observing mode was employed: for each target, one beam was pointed at the source while another simultaneously observed a nearby off-source sky position, enabling efficient subtraction of atmospheric and instrumental background.

The backend system consists of a fast Fourier transform spectrometer (FFTS) with a total bandwidth of 1 GHz and 16 384 spectral channels, yielding a channel spacing of 61 kHz (corresponding to a velocity resolution of $\sim$0.16 km s^{-1} at 115 GHz). For each source, we co-added individual integrations meeting the quality criteria $T_{\rm sys} < 220$ K and rms noise $\sigma < 0.7$ K per channel. No significant ^{12}CO(1–0) emission was detected in any of the stacked spectra; therefore, we adopted the 3σ noise level as the upper limit on the line flux.

To validate our observational sensitivity and data reduction pipeline, we observed the control source J1101+4513, for which Gao et al. (2019) reported a detection. This source was not selected from our parent compact-[O III] gas-follow-up sample and is therefore excluded from our statistical analysis. Applying identical stacking and calibration procedures—including the same antenna temperature-to-flux density conversion factor of 24.7 Jy K^{-1}—we successfully recovered the ^{12}CO(1–0) emission in the stacked spectrum, confirming that the reduction pipeline can recover a known CO line when present.

In total, 14 sources in our sample have CO observations. Because of the intrinsic faintness of CO in these low-metallicity environments, most non-detections only demonstrate that CO is below the sensitivity of a given observation. We define a CO upper limit as physically constraining when the inferred $M_{\rm H_2}^{\rm CO}$ limit, after applying the metallicity-dependent $\alpha_{\rm CO}$, is comparable to or lower than the molecular mass expected from independent tracers or scaling relations. For the PMO setup and the metallicities of our targets, this typically requires $\sigma_{T_{\rm CO}} \lesssim 4$ mK and integration times on the order of 9 hr or longer. Consequently, only two sources (J1200+2719 and J1509+4543) have sufficiently deep integrations (> 9 hr) to provide constraining upper limits on $M_{\rm H_2}^{\rm CO}$. While this limits the statistical power of the CO data, these two exceptionally deep non-detections serve as valuable case studies for testing the limits of CO-traced molecular gas in compact starbursts.

We reduced the data obtained with the PMO 13.7-m telescope using the `CLASS` software package (Pety 2005), which is part of the `GILDAS` suite[1]. All individual scans were visually inspected, and those exhibiting severe baseline curvature or anomalous noise characteristics were excluded from further analysis. A first-order (linear) baseline was subsequently fitted to and subtracted from the line-free portions of each spectrum.

Following correction for the main beam efficiency ($\eta_{\rm mb}$), antenna temperatures ($T_{\rm A}^*$) were converted to main beam brightness temperatures ($T_{\rm MB}$) using the relation $T_{\rm MB} = T_{\rm A}^*/\eta_{\rm mb}$. Individual spectra were then co-added, with weights proportional to the inverse square of their rms noise levels to optimize the signal-to-noise ratio (S/N) in the final stacked spectrum.

The ^{12}CO(1–0) line luminosity was computed following the prescription of Bolatto et al. (2008):

$$\frac{L'_{\rm CO}}{\rm K\ km\ s^{-1}\ pc^2} = 2453 \left(\frac{S_{\rm CO}\,\Delta v}{\rm Jy\ km\ s^{-1}}\right)\left(\frac{D_L}{\rm Mpc}\right)^2 (1+z)^{-1}, \quad (1)$$

where $S_{\rm CO}\,\Delta v$ is the velocity-integrated CO flux density, and D_L is the luminosity distance. This expression was applied uniformly to the PMO measurements and upper limits.

2.4. Herschel PACS spectroscopy

The Photodetector Array Camera and Spectrometer (PACS; Poglitsch et al. 2010) aboard the Herschel Space Observatory (Pilbratt et al. 2010) provides [C II]$\lambda158\,\mu$m spectroscopy in its red channel (103–190 μm, $R \approx 1500$). We used observations from the program "Merger driven vs. Cold-flow driven Star Formation in Typical High-redshift Galaxies" (PI: Shim, H.; Proposal ID: OT2_hshim_2).

The original Herschel program targeted 18 compact emission-line galaxies. We used Level 2.0 pipeline products with spectral flat-fielding and de-glitching applied. For each target, we used the spectrum labeled `c1`, generated by the point-source extraction task `extractCentralSpectrum`, which applies the central-to-total (CTT) flux conversion for unresolved sources.

We performed a complete reanalysis of all [C II] spectra using Monte Carlo error estimation and Gaussian line fitting (see Appendix C and Table A.1). Detections are defined as sources with integrated line S/N ≥ 3.5. For non-detections, we report 3σ upper limits assuming a velocity dispersion of $\sigma_v = 50$ km s^{-1}.

Of the 18 observed sources, four were excluded as merger systems (HAE1624 ≡ J1624−0022, HAE1557 ≡ J1557+2320, HAE1622 ≡ J1622+1636, and HAE1457_2 ≡ J1457+3051; see Sect. 2.6). Moreover, five additional sources were observed in the same program but excluded from our parent GP catalog

[1] `https://www.iram.fr/IRAMFR/GILDAS/`

(HAE1523, HAE1628, HAE1457_1, HAE0303, and HAE0842) were also excluded. This left us with nine sources in our [C ii] analysis: five detections and four non-detections. HAE0840 (S/N = 3.2) is conservatively classified as a non-detection given its marginal signal and elevated reduced χ^2.

2.5. Optical spectroscopy

To determine the gas-phase metallicities and characterize the ionization conditions of our sample, we used optical spectroscopy from three facilities. For the majority of our sample, we relied on archival spectra from the Sloan Digital Sky Survey (SDSS; Cardamone et al. 2009). For sources lacking SDSS coverage, we used optical spectra from the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Wang et al. 1996; Luo et al. 2012) and the MMT Observatory (Yang et al. 2017). Specifically, three sources in our sample (J0847+2944, J0856+3111, and J1136+3427) rely on LAMOST spectra for metallicity measurements, while the remaining sources use either SDSS or MMT data.

The availability of optical spectra enables the construction of the [N ii] Baldwin–Phillips–Terlevich (BPT) diagnostic diagram (Baldwin et al. 1981) to verify the star-forming nature of our sources. We measured emission-line fluxes from 38 SDSS spectra and eight LAMOST spectra via Gaussian profile fitting, supplemented by SDSS `galSpecLine` pipeline measurements where available. In total, 45 of 60 sources have measured [N ii]/Hα and [O iii]/Hβ ratios. Excluding J0109−0106, whose anomalously low [O iii]/Hβ ratio likely reflects a poor spectral fit, 44 sources have reliable BPT classifications. As shown in Fig. 1, all 44 sources fall firmly in the star-forming region, well below both the Kauffmann et al. (2003) empirical boundary and the Kewley et al. (2001) theoretical maximum starburst line, with log ([N ii]/Hα) ranging from −2.53 to −0.97. Their loci are consistent with the broader GP populations of Jiang et al. (2019) and Liu et al. (2023a), confirming the absence of significant active galactic nucleus (AGN) contamination.

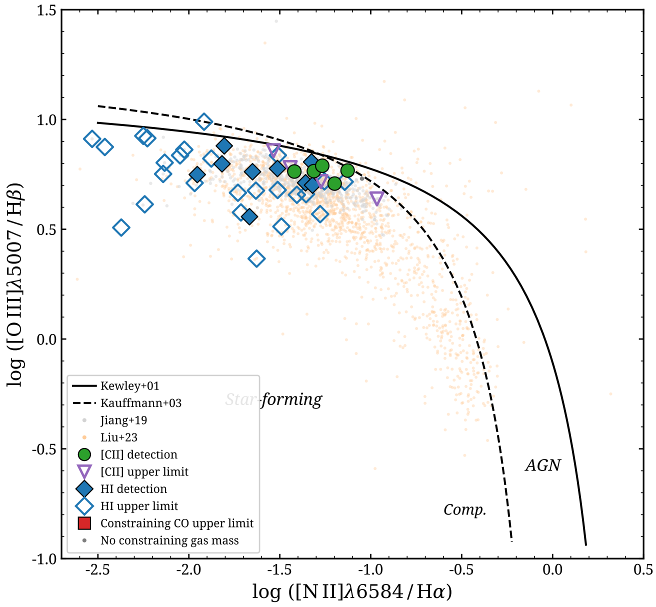

Fig. 1: [N ii]-BPT diagnostic diagram. The solid and dashed curves show the Kewley et al. (2001) theoretical maximum starburst line and the Kauffmann et al. (2003) empirical demarcation, respectively. The small gray and light orange points in the background show the Green Pea samples of Jiang et al. (2019) and Liu et al. (2023a), respectively. The symbols identify the gas-measurement status used throughout this paper: the filled green circles denote [C ii] 158 μm detections, the open purple downward triangles denote [C ii] upper limits, the filled blue diamonds denote H i 21 cm detections, the open blue diamonds denote H i upper limits, the filled red squares mark constraining CO upper limits, and the small gray point marks the source without a constraining gas-mass measurement. All 44 sources with reliable line ratios (excluding J0109−0106; see text) occupy the star-forming region.

2.6. Morphological inspection and merger exclusion

A critical concern raised in previous analyses of compact emission-line galaxies is the potential contamination by merger or strongly interacting systems (e.g., Chary & Shim 2015). To address this, we performed a systematic visual inspection of all candidate sources (see Fig. D.1 in Appendix D) [2] using false-color cutout images from the DESI Legacy Imaging Surveys (Dey et al. 2019), which provide deep *grz*-band photometry with sub-arcsecond resolution. Each source was classified into one of three categories: (i) isolated compact, showing a single, unresolved or barely resolved source with no companions within $\sim 10''$; (ii) tail or asymmetry, displaying tidal features, asymmetric morphology, or a faint companion but with the primary source remaining dominant; and (iii) merger or strong interaction, exhibiting clear double nuclei, extended tidal tails, or comparably bright companions at consistent photometric redshifts.

Sources in category (iii) were excluded from the final sample. In total, we removed 15 merger systems: six from the Kanekar et al. (2021) subsample (J1134+5006, J1148+2546, J1345+0442, J1451−0056, J1455+3808, J1624−0022), three from the Herschel [C ii] subsample (J1557+2320, J1622+1636, J1457+3051), and six from Khasnovis et al. (2025) (J0130+1746, J0238−0048, J1132+2459, J1149+4525, J1321+4727, J1359+5726).

We note that 12 sources in the retained sample exhibit minor tidal features or asymmetries (category ii) but were kept in our analysis as their primary emission remains dominated by a single compact source. These are flagged in Appendix D.

By restricting our analysis to morphologically isolated systems, we ensured that the derived gas properties reflect the intrinsic ISM conditions of compact starbursts rather than merger-driven enhancements.

2.7. Location in sample parameter space

Figure 2 summarizes the parameter space and gas-tracer coverage of the final 60-object compilation. In the $M_\star$–SFR plane (panel a), our galaxies span $\log(M_\star/M_\odot) \approx 6.0$–$9.4$ and $\log(\mathrm{SFR}/M_\odot\,\mathrm{yr}^{-1}) \approx -2.4$ to 1.8, covering the parameter space between the low-mass Blueberry regime and the more massive Green Pea regime. The majority of our sources (39 out of 60) lie above the SFMS of Speagle et al. (2014) at $z = 0.04$, consistent with the starburst nature of the GP population. In the mass–metallicity plane (panel b), our sample falls systematically below the T_e-based MZR of Curti et al. (2020) (52 out of 59 sources with metallicity measurements), in agreement with the well-established metal deficiency of compact emission-line

[2] except J0357+1808, for which DESI Legacy Imaging Surveys coverage is unavailable; we used the SDSS cutout instead

galaxies at fixed stellar mass. Panels (c) and (d) make the local and heterogeneous nature of the compilation explicit: H I coverage is concentrated at $z < 0.1$, while the [C II] subset is small and drawn from the Herschel program. These distributions illustrate the selection effects discussed in Sect. 2.1: the gas-follow-up sample is weighted toward nearby, low-mass systems with high sSFRs and should not be interpreted as a statistically complete 60-object gas survey.

3. Physical parameters

To characterize the interstellar medium and star formation properties of our sample, we derived key physical parameters through a multiwavelength approach. This section summarizes the stellar, star formation, metallicity, and radiation-field quantities used in the gas analysis.

3.1. Physical parameter estimation

We derived key physical properties using multiwavelength data and established astrophysical diagnostics. Stellar masses ($M_\star$) were obtained via spectral energy distribution (SED) fitting with CIGALE (Burgarella et al. 2005; Boquien et al. 2019). The photometric input spans the ultraviolet to mid-infrared, including GALEX far-ultraviolet (FUV) and near-ultraviolet (NUV) bands (Bianchi et al. 2017), SDSS *ugriz* bands, and AllWISE $W1$–$W4$ (Wright et al. 2010, 2019). Our fiducial model configuration includes: a delayed-τ star formation history, Bruzual & Charlot (2003) SPS models, a Chabrier (2003) IMF, nebular emission (continuum and lines), a Calzetti-like dust attenuation law modified for starburst galaxies, and the Casey (2012) dust emission model. We also tested an extended configuration incorporating the SKIRTOR AGN torus model (Stalevski et al. 2012, 2016) for sources with elevated reduced chi-squared values ($\chi^2_{\rm red} > 5$). In all cases the best-fit AGN fraction is consistent with zero and the fit quality does not improve, confirming that AGN contamination is negligible in our sample and that the residual $\chi^2_{\rm red}$ likely arises from strong nebular emission-line contributions to the broad-band photometry.

As an independent SFR indicator is robust against dust extinction, we also computed $\mathrm{SFR_{NUV,corr}}$ following the hybrid UV+IR prescription of Durbala et al. (2020):

$$\log_{10}(\mathrm{SFR_{NUV,corr}}) = \log_{10}(\nu L_{\nu,\rm corr}) - 43.17, \qquad (2)$$

where $\nu L_{\nu,\rm corr} = \nu L_{\nu,\rm NUV} + 2.26\,\nu L_\nu(22\,\mu\mathrm{m})$ is in erg s^{-1} and SFR is in $M_\odot$ yr^{-1}. For three sources (J0159+0722, J1402+3033, and J1400+1951) that lack both GALEX and WISE counterparts within a 5″ radius, we estimated the SFR from the Hα luminosity measured in their SDSS spectra. We corrected the Hα fluxes for dust extinction using the observed Balmer decrement (Hα/Hβ), assuming an intrinsic ratio of 2.86 and the Cardelli et al. (1989) extinction law. The dust-corrected luminosities were then converted to SFRs using the Kennicutt (1998) calibration adjusted to a Chabrier (2003) IMF (SFR = $5.5\times10^{-42} L_{\mathrm{H}\alpha,\rm corr}$). The SDSS fiber for J1400+1951 captured a background quasar at $z = 2.085$. Therefore, for this target, we measured the Hα flux from its LAMOST spectrum (z=0.0532), which successfully targeted the galaxy. Due to wavelength-dependent flux calibration issues in the LAMOST spectrum, we applied an absolute flux calibration anchored to the SDSS r-band photometry and report the uncorrected SFR as a lower limit.

Gas-phase metallicities were determined using optical spectroscopy from SDSS or LAMOST. For galaxies with a $\geq 2\sigma$ detection of the [O III]λ4363 auroral line, we applied the direct T_e-method (Peimbert & Torres-Peimbert 1984; Izotov et al. 2006). For the remaining sources, we adopted the strong-line calibration of Pettini & Pagel (2004) based on the N2 index. The metallicities span $12 + \log(\mathrm{O/H}) = 7.46$–$8.30$, corresponding to ~ 0.06–$0.41\,Z_\odot$.

For the three sources relying on LAMOST spectra (J0847+2944, J0856+3111, and J1136+3427), we derived metallicities using the direct T_e-method, where the [O III]λ4363 line is detected. These LAMOST-based metallicities are flagged with a dagger symbol (†) in Table A.1 and Table A.2.

3.2. Radiation field and hydrogen density

To constrain the physical conditions of the photodissociation regions (PDRs) in our sample, we estimated the interstellar radiation field intensity (G_0, in units of the Habing field) and the hydrogen number density ($n_{\rm H}$) using a simplified diagnostic method based on observed optical emission-line ratios and SFR surface densities. These parameters are crucial for understanding the state of the neutral ISM and for validating the use of [C II] as a molecular gas tracer.

3.2.1. Simplified diagnostic method

For the hydrogen number density, we used the [S II]λ6717/λ6731 doublet ratio as a classical electron density diagnostic (Osterbrock & Ferland 2006). At $T_e \sim 10^4$ K, this ratio is sensitive to n_e in the range $\sim 10^{1.5}$–$10^{3.5}$ cm^{-3} and well matched to the expected conditions in compact emission-line galaxies. We adopted $n_{\rm H} \approx n_e$ as a first-order approximation for the ionized zone, noting that the actual hydrogen density in PDR layers may be higher by a factor of a few. Of the 60 sources in our sample, 45 have SDSS spectra available, and 33 yield valid [S II] ratios within the diagnostic range (0.44–1.44). The remaining sources either lack SDSS spectral coverage or have [S II] lines with insufficient S/N.

For the radiation field intensity, we estimated G_0 from the SFR surface density following the approach of Wolfire et al. (2003):

$$G_0 \approx \frac{\Sigma_{\rm SFR}}{\Sigma_{\rm SFR,MW}} \times G_{0,\rm MW}, \qquad (3)$$

where $\Sigma_{\rm SFR,MW} \approx 2.5\times10^{-3}\,M_\odot\,\mathrm{yr^{-1}\,kpc^{-2}}$ and $G_{0,\rm MW} \approx 1.7$. We adopted a characteristic effective radius of $r_{\rm eff} = 0.5$ kpc, typical for Green Pea galaxies (Cardamone et al. 2009; Izotov et al. 2011). This method yields G_0 estimates for 44 sources with available SFR measurements.

The uncertainties in these diagnostic estimates are dominated by different terms. For $n_{\rm H}$, we propagated the [S II] line-flux uncertainties through the doublet ratio and adopt a floor of 0.2 dex to account for the conversion from electron density in the ionized gas to characteristic hydrogen density near the PDR interface. For G_0, the formal SFR uncertainties are smaller than the systematic uncertainty introduced by assuming a common effective radius. Allowing $r_{\rm eff}$ to vary by a factor of two changes $\log G_0$ by 0.6 dex. We therefore treated individual G_0 values as order-of-magnitude diagnostics and used them primarily to check whether the sample lies within the validity range of the [C II]-to-H_2 calibration.

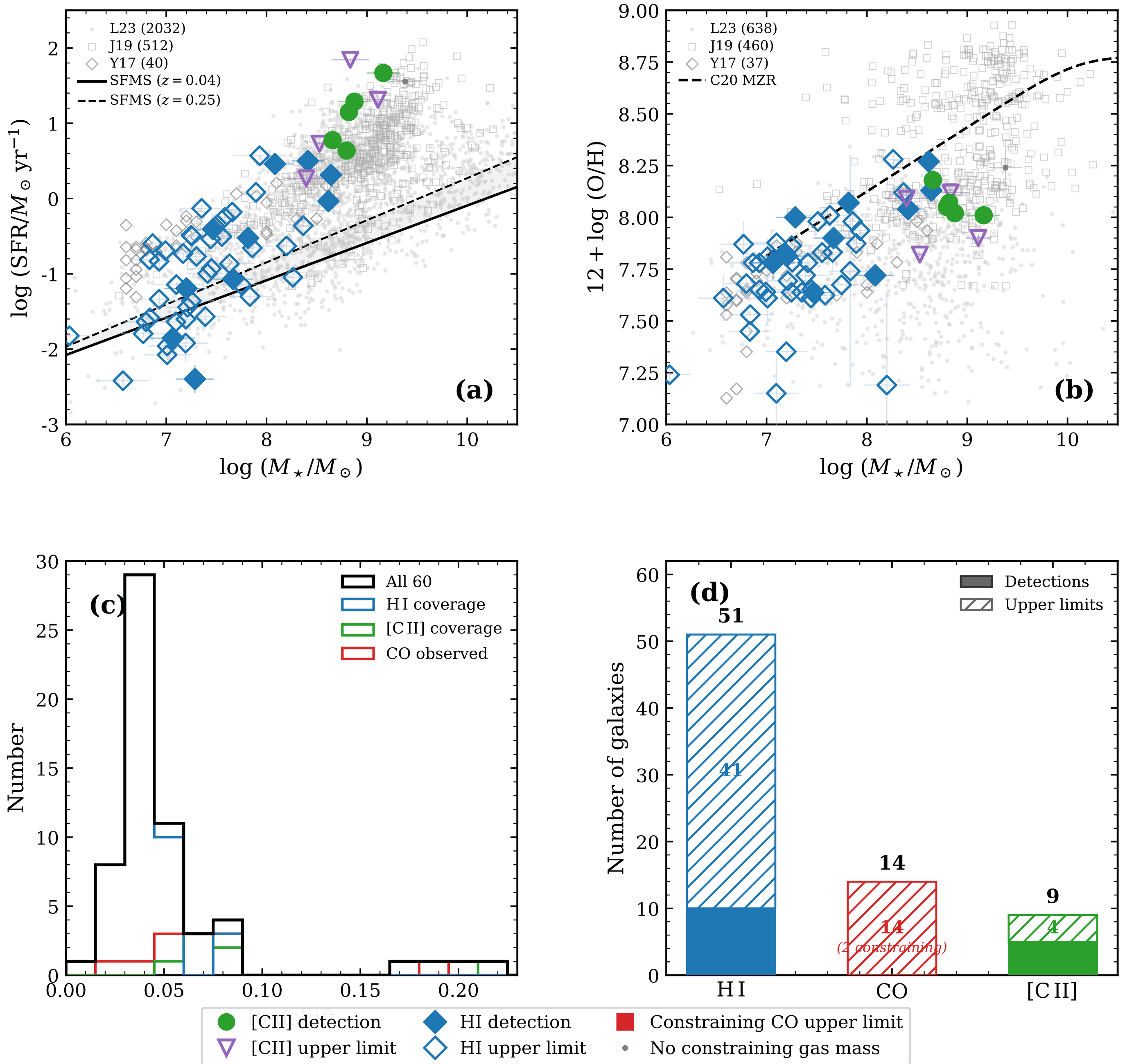

Fig. 2: Parameter space and gas-tracer coverage of the final 60-object compilation. The small gray dots, open squares, and open diamonds in the background show the parent Green Pea samples of Liu et al. (2023a), Jiang et al. (2019), and Yang et al. (2017), respectively. (a) Stellar mass vs. SFR. The solid and dashed black lines show the SFMS from Speagle et al. (2014) at $z = 0.04$ and $z = 0.25$, respectively, with the shaded region indicating the enclosed area. (b) Stellar mass vs. T_e-based gas-phase metallicity. The dashed curve shows the T_e-based mass–metallicity relation of Curti et al. (2020). The error bars indicate the 1σ uncertainties on stellar mass and metallicity. (c) Redshift distribution of the full sample and of the H I, CO, and [C II] observed subsets. (d) Gas-tracer coverage summary. The solid-filled and hatched portions of each bar indicate detections and upper limits, respectively. Numbers above the bars give the total covered sample size. The CO bar shows that only two non-detections are deep enough to provide constraining CO-based M_{H_2} limits. The point symbols follow the detection and upper-limit code defined in Fig. 1.

3.2.2. Results

The derived radiation field intensities and hydrogen number densities are presented in Fig. 3. We find $\log G_0$ values ranging from 0.5 to 4.8, with a median of ~ 2.2 (Fig. 3a), spanning nearly four orders of magnitude and reflecting the diverse star-forming conditions across our sample. The subsample with [C II] or CO observations, which consists of the most actively star-forming galaxies, occupies the upper end of the G_0 distribution ($\log G_0 \gtrsim 3$), consistent with their elevated SFR surface densities. The derived $\log n_H$ values fall between 1.3 and 3.1 cm^{-3}, with a median of $\sim 2.1\,cm^{-3}$ (Fig. 3b), representing moderate to high gas densities comparable to those found in local starburst nuclei and high-redshift star-forming environments (Steidel et al. 2016; Tadaki et al. 2019).

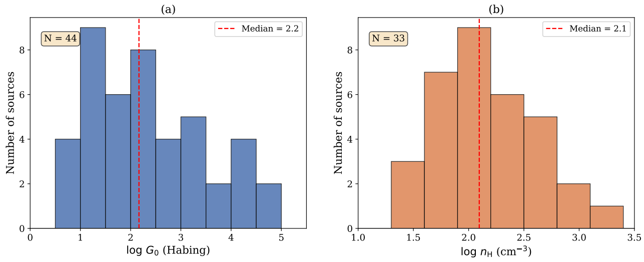

Fig. 3: Distribution of the derived PDR physical parameters for our sample: (a) interstellar radiation field intensity ($\log G_0$) in Habing units, estimated from the SFR surface density ($N = 44$), and (b) hydrogen number density ($\log n_H$) in cm^{-3}, derived from the [S II]$\lambda6717/\lambda6731$ doublet ratio ($N = 33$). The dashed red lines indicate the median values. The typical uncertainties are ~ 0.2–0.3 dex for $\log n_H$ and ~ 0.6 dex for $\log G_0$, the latter dominated by the assumed emitting radius.

As shown in Fig. 4, the derived values for our sample are plotted in the $\log G_0$–$\log n_H$ plane. Of the 33 sources with both parameters available, 25 (76%) fall within the parameter space ($1 \leq \log n_H \leq 4$, $1 \leq \log G_0 \leq 4$) explored by the Madden et al. (2020) [C II]-to-H_2 calibration. This calibration is based on Cloudy models of low-metallicity dwarf galaxies in which the H I and H_2 contributions are separated. The subsample with [C II] or CO observations (green squares) primarily occupies the upper-right region of the diagram, reflecting their higher SFR surface densities and moderately elevated gas densities. The eight sources that lie outside the Madden model range, primarily above $\log G_0 = 4$, are among the most extreme starbursts in our sample; their [C II]-based M_{H_2} estimates should therefore be treated with additional caution.

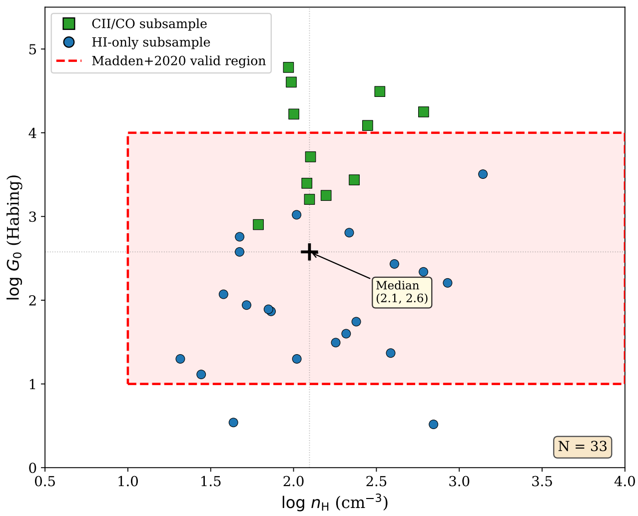

Fig. 4: Distribution of the sample in the $\log G_0$–$\log n_H$ plane ($N = 33$). The green squares denote sources with [C II] or CO observations (Table A.1), while the blue circles denote H I-only sources (Table A.2). The dashed red rectangle indicates the valid parameter space ($1 \leq \log G_0 \leq 4$, $1 \leq \log n_H \leq 4$) for the Madden et al. (2020) [C II]-to-H_2 calibration. The black cross marks the median position.

3.3. Molecular gas mass estimation

We estimated the molecular gas content using two complementary tracers: ^{12}CO(1–0) emission for CO-bright gas and [C II]$\lambda158\,\mu$m emission as a proxy for CO-dark molecular gas. The following subsections describe the two estimators and the assumptions that enter each conversion.

3.3.1. CO-based upper limits with metallicity-dependent α_{CO}

For sources observed with the PMO 13.7 m telescope, no ^{12}CO(1–0) detections were achieved. The molecular gas mass is related to the CO luminosity via

$$M_{H_2}^{CO} = \alpha_{CO}\, L'_{CO(1-0)}. \tag{4}$$

Given the low-metallicity nature of our targets (12 + log(O/H) $\sim$ 7.5–8.3), adopting the Galactic conversion factor $\alpha_{CO} = 4.3\,M_\odot\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$ would significantly underestimate the molecular gas mass. We therefore adopted the metallicity-dependent prescription of Amorín et al. (2016):

$$\alpha_{CO} = 4.3\left(\frac{Z}{Z_\odot}\right)^{-1.5} M_\odot\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}, \tag{5}$$

where $Z/Z_\odot = 10^{[12+\log(O/H)]-8.69}$ and $12 + \log(O/H)_\odot = 8.69$ (Asplund et al. 2009). This prescription was calibrated on blue compact dwarf galaxies (BCDs), which share many physical properties with GPs, and yields α_{CO} values ranging from ~ 17 to ~ 300 for our targets.

We note that this metallicity dependence is broadly consistent with other prescriptions in the literature (Schruba et al. 2012; Genzel et al. 2012; Bolatto et al. 2013), although the exact power-law index remains uncertain. This uncertainty is considered when comparing CO upper limits with [C II]-based molecular masses.

3.3.2. [C II]-based molecular gas masses

For the CO-dark component, we used [C II]$\lambda158\,\mu$m line luminosities derived from Herschel PACS spectroscopy. The line luminosity is computed following Solomon et al. (1992):

$$L_{[CII]} = 1.04 \times 10^{-3}\, S_{[CII]}\,\Delta v\, \nu_{rest}\,(1+z)^{-1}\, d_L^2, \tag{6}$$

where $L_{[CII]}$ is in $L_\odot$, $S_{[CII]}\,\Delta v$ is in jansky kilometer per second, ν_{rest} is in gigahertz, and d_L is in megaparsec.

To infer the total molecular gas mass from [C II], we adopted the model-based scaling relation calibrated for local dwarf galaxies by Madden et al. (2020):

$$M_{H_2}^{total} = 10^{2.12}\left[\frac{L_{[CII]}}{L_\odot}\right]^{0.97} M_\odot. \tag{7}$$

The Madden et al. (2020) calibration combines observations of low-metallicity dwarf galaxies with grids of Cloudy models spanning different radiation fields, densities, metallicities, and cloud depths. In those models, the hydrogen phase structure (H I and H_2) is self-consistently computed as a function of cloud depth, and the total H_2 mass is calibrated against the total observed [C II] luminosity (which includes contributions from both the H I and H_2 layers of the PDR). The applicability of Eq. (7) to our targets is supported by the n_H and G_0 values derived in Sect. 3.2, which mostly fall within the parameter space of the Madden et al. (2020) models ($n_H = 10$–10^4 cm^{-3},

$1 < \log G_0 < 4$). Sources outside this range were retained but flagged as having more uncertain [C ii]-based masses.

As an independent cross-check, we also computed molecular gas masses using the empirical calibration of Zanella et al. (2018), who find $\alpha_{\rm [CII]} \equiv M_{\rm H_2}/L_{\rm [CII]} \approx 30\,M_\odot/L_\odot$ with a scatter of 0.2 dex. The two methods yield consistent results within the uncertainties, providing confidence in our [C ii]-based gas mass estimates.

4. Results and discussion

Our multiwavelength analysis reveals that compact emission-line galaxies exhibit extreme ISM conditions and elevated star formation activity. To place these systems in a broader galactic context, we compare their gas reservoirs and star formation properties with those of the xCOLD GASS survey (Saintonge et al. 2017; Catinella et al. 2018).

4.1. Scaling relations

Scaling relations provide a powerful framework for characterizing the cold gas content of galaxies and identifying systematic deviations from typical evolutionary pathways (Roberts & Haynes 1994). In Fig. 5, we place our sample within this context by comparing their atomic and molecular gas fractions ($f_{\rm gas} = M_{\rm gas}/M_\star$) and gas depletion times ($\tau_{\rm dep} = M_{\rm gas}/{\rm SFR}$) against the scaling relations defined by xGASS (Catinella et al. 2018) and xCOLD GASS (Saintonge et al. 2017).

In these figures, GP sources are classified into three categories based on their [C ii] coverage: the green circles denote [C ii]-based H_2 detections (five sources), the purple downward triangles indicate [C ii] 3σ upper limits (four sources), and the blue diamonds represent sources with H i or CO observations but without [C ii] coverage (51 sources). The error bars reflect the propagated uncertainties from flux measurements, distance estimates, and the metallicity-dependent $\alpha_{\rm CO}$ conversion. Reference samples from xGASS (Catinella et al. 2018) and xCOLD GASS (Saintonge et al. 2017) are shown in gray, with binned median relations computed using survival analysis to account for censored data. The blue shaded regressions are shown only as visual summaries of monotonic trends in the reference samples; our interpretation relies on the location of the GP points relative to the binned medians and upper limits rather than on the fitted slopes alone.

4.1.1. Gas mass fractions

In the top row of Fig. 5, the five [C ii]-based GP detections (green circles) primarily occupy the upper envelope of the reference molecular-gas-fraction distribution at fixed stellar mass, with several lying ~0.5–1 dex above the binned medians. The [C ii] upper limits (purple triangles) are broadly consistent with this upper envelope. In contrast, the deepest CO-based upper limits fall below the [C ii]-inferred molecular masses, highlighting the difficulty of using CO as a complete molecular gas tracer in these low-metallicity systems.

Both atomic and molecular gas fractions exhibit a clear decline with increasing stellar mass, stellar surface mass density ($\mu_\star$), and rest-frame $NUV - r$ color. Among these parameters, $NUV - r$ serves as the strongest single predictor of H i gas fraction, consistent with prior studies (Zhang et al. 2009; Catinella et al. 2010; Liu et al. 2023b).

The tightest empirical constraint on molecular gas fraction in the reference sample arises from its correlation with sSFR. Using the xCOLD GASS calibration summarized by Saintonge & Catinella (2022), the [C ii]-detected GPs are consistent with, or modestly above, the local relation by up to ~0.2–0.4 dex. Given the small number of [C ii] detections, we interpret this as suggestive evidence for high molecular gas fractions in individual compact starbursts rather than as a statistically significant population offset.

4.1.2. Depletion times

The gas depletion timescale ($\tau_{\rm dep} \equiv M_{\rm gas}/{\rm SFR}$) quantifies the timescale over which star formation can be sustained at the current rate, assuming no inflows or recycling. For the ten H i-detected sources, we find a median atomic gas depletion time of $\tau_{\rm dep,HI} = 2.6$ Gyr, with a wide range from 0.43 to 240 Gyr reflecting the diversity of the sample. The extremely long depletion times (> 10 Gyr) are found in sources with very low SFRs (e.g., J0840+5333 with an SFR = 0.004 $M_\odot\,{\rm yr}^{-1}$).

For the five [C ii]-detected sources, the median molecular gas depletion time is $\tau_{\rm dep,H_2} = 0.16$ Gyr, with a range of 0.12–0.39 Gyr. These values are notably shorter than the ~ 1 Gyr typical of main-sequence galaxies (Saintonge et al. 2017), indicating highly efficient star formation in these compact starbursts.

4.2. The main sequence

We present the distribution of our galaxies in the stellar mass versus SFR plane and their position relative to the SFMS in Fig. 6. Following Speagle et al. (2014), we adopted the time-dependent main-sequence relation evaluated at the median redshift of our sample ($z \approx 0.04$, corresponding to a cosmic age of $t \approx 13.3$ Gyr):

$$\log\left[{\rm SFR_{MS}}\,(M_\odot\,{\rm yr}^{-1})\right] = (0.84 - 0.026\,t)\,\log(M_*/M_\odot) - (6.51 - 0.11\,t), \qquad (8)$$

where t is the age of the Universe in gigayears.

In our sample, compact emission-line galaxies generally have elevated SFRs compared to the main sequence, typically lying ~ 0.4–1 dex or more above it. This places many of them in the starburst regime. In the left panel, the H i-rich GP sources (blue diamonds) exhibit high atomic gas fractions ($M_{\rm HI}/M_* \sim 1$–10) at low stellar masses ($M_* \sim 10^{7.5}$–$10^9\,M_\odot$), consistent with gas-rich dwarf galaxies. The center panel shows that [C ii]-detected sources (green circles) have high molecular gas fractions ($M_{\rm H_2}/M_* \sim 0.3$–3), while the right panel indicates short molecular gas depletion times ($\tau_{\rm dep,H_2} \sim 10^8$–$10^9$ yr). The [C ii] upper limits (purple triangles) are consistent with these trends. Taken together, these detected subsets are consistent with rapid gas cycling in compact starbursts, plausibly driven by intense radiation fields, high gas pressures, and enhanced feedback mechanisms (Lisenfeld et al. 2019).

4.3. The CO-dark gas fraction

A key finding of this work is that CO emission alone appears to significantly underestimate the molecular gas content in the compact emission-line galaxies studied here. Although our sample of deep CO observations is limited by the demanding integration times required, the available data provide compelling individual case studies. Specifically, for the sources with the deepest CO integrations (> 9 hr) or those with both CO upper limits

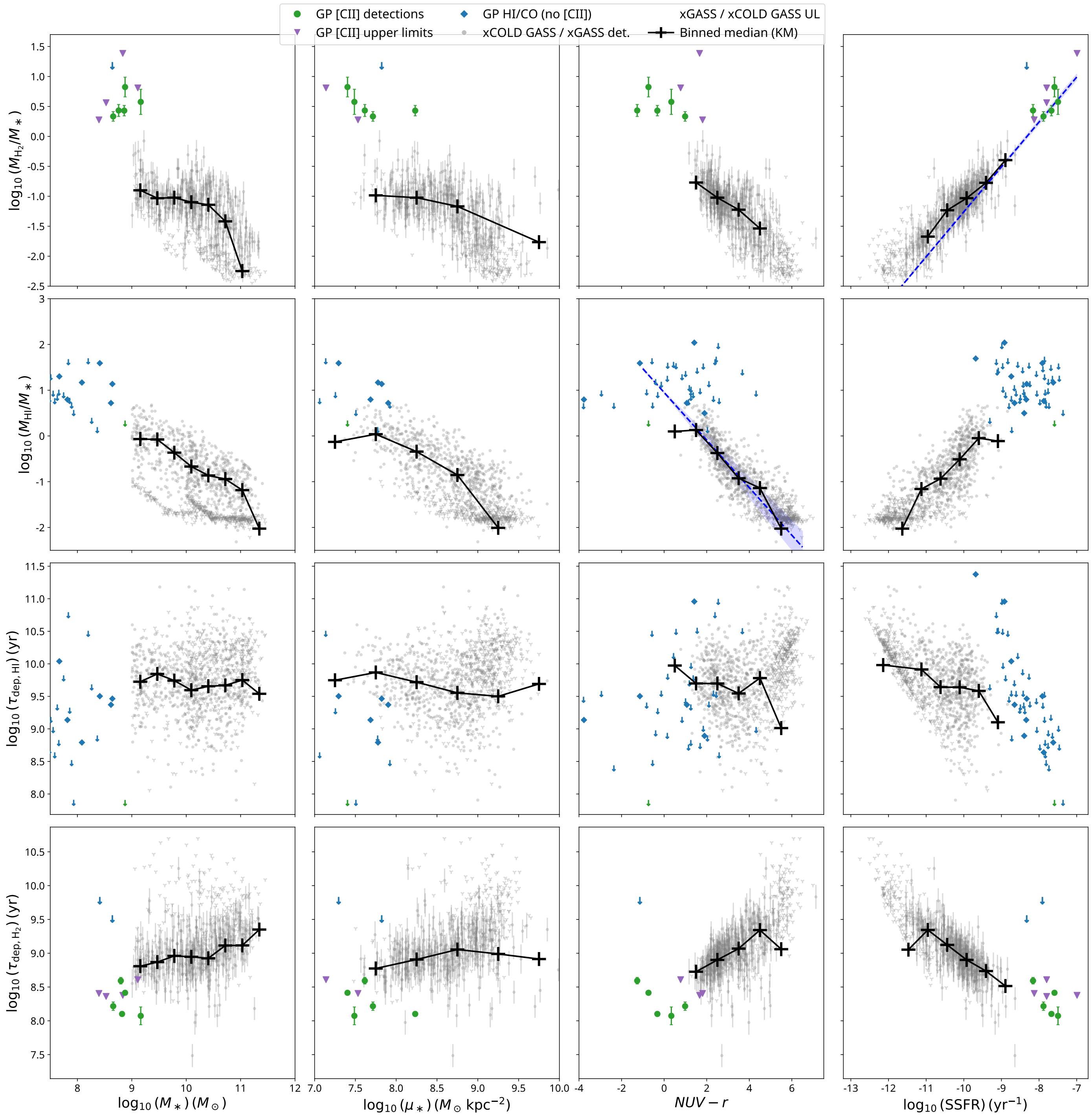

Fig. 5: Scaling relations for cold gas content in local galaxies. From top to bottom: Molecular gas fraction (M_{H_2}/M_*), atomic gas fraction (M_{HI}/M_*), atomic gas depletion time ($\tau_{dep,HI}$), and molecular gas depletion time (τ_{dep,H_2}). Each quantity is plotted against stellar mass (M_*), stellar mass surface density (μ_*), rest-frame $NUV - r$ color, and sSFR. GP sources are classified into three categories: green circles for [C II]-based H_2 detections, purple downward triangles for [C II] 3σ upper limits, and blue diamonds for sources with H I or CO data but no [C II] coverage. Reference samples from xGASS (rows 2, 3) and xCOLD GASS (rows 1, 4) are shown in gray, with small circles for detections and Y-crosses for upper limits. the black pluses indicate binned medians computed via Kaplan–Meier survival analysis to account for censored data. The dashed blue lines in selected panels show illustrative linear fits to the reference scaling relations where the reference sample displays a clear monotonic trend, with shaded regions indicating the 1σ uncertainty. The error bars on GP data points reflect propagated measurement uncertainties. Note: The stellar mass surface density of J0852+1216 was computed using $R_{50} = 1.5''$ from an alternate SDSS photometric detection (the primary detection has an unmeasured Petrosian radius). For J1735+5703, we adopt the photometry from the SDSS spectroscopic match ($r_{cModel} = 17.69$ mag, $R_{50} = 0.84''$) rather than a faint background source initially crossmatched in the photometric catalog.

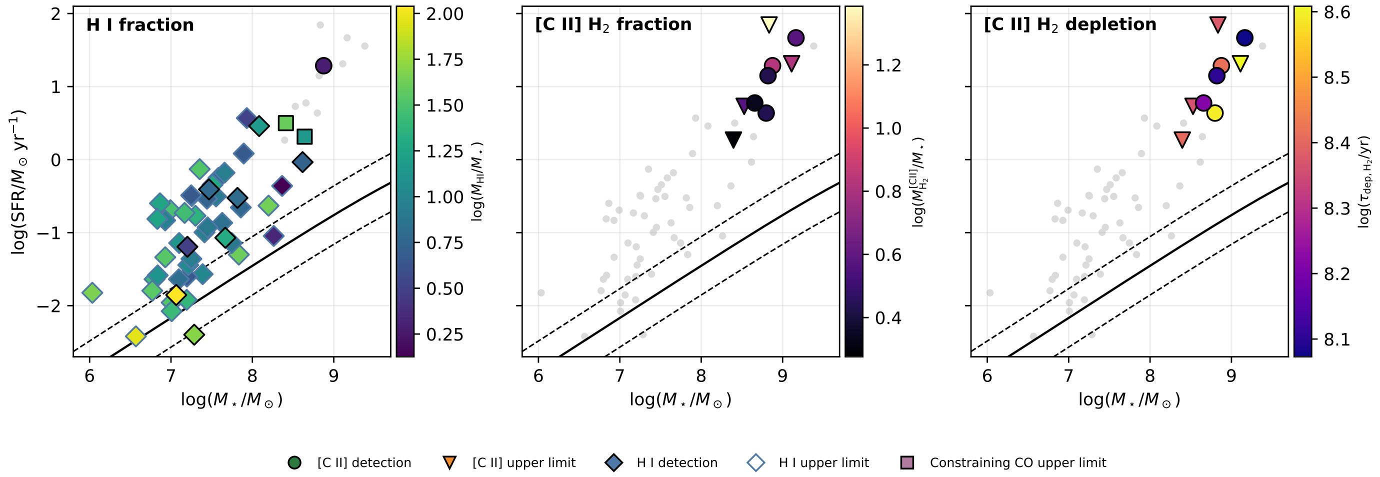

Fig. 6: Distribution of the sample galaxies in the SFR–M_* plane, color-coded by (left) atomic gas fraction $\log(M_{\rm HI}/M_*)$, (center) molecular gas fraction $\log(M_{\rm H_2}/M_*)$, and (right) molecular gas depletion time $\log(\tau_{\rm dep,H_2}/{\rm yr})$. In each panel, the solid black curve shows the SFMS from Speagle et al. (2014) at $z = 0.04$, with the dashed lines marking the ±0.4 dex scatter. The green circles denote [C ii] detections, the purple downward triangles denote [C ii] upper limits, and the blue diamonds denote sources with H i or CO observations. Left panel: H i detections and upper limits (filled and open diamonds), with the filled color encoding $\log(M_{\rm HI}/M_*)$. Center and right panels: Upper limits on the molecular gas quantity, indicated by the downward arrows. The background plus signs and Y-shaped markers show the xGASS/xCOLD GASS detections and upper limits, respectively, color-coded by the same quantity as the GP sources.

and [C ii] detections, the inferred [C ii]-based molecular gas masses are higher than the CO-based limits obtained with the metallicity-dependent $\alpha_{\rm CO}$ of Amorín et al. (2016). Other low-metallicity prescriptions, including the more extreme conversions discussed by Madden et al. (2020), would raise the CO-based limits and should be considered when placing conservative upper bounds on CO-bright gas. They do not remove the qualitative tension that [C ii] traces substantial molecular material in systems where CO remains undetected, but they emphasize that the inferred CO-dark fraction is prescription dependent.

This finding is consistent with theoretical expectations for low-metallicity environments, where reduced dust shielding allows UV photons to photodissociate CO molecules throughout much of the molecular cloud, leaving a large reservoir of "CO-dark" H_2 that is traced by [C ii] but invisible to CO observations (Wolfire et al. 2010; Bolatto et al. 2013; Madden et al. 2020). We caution that [C ii] emission can also arise from ionized gas and atomic PDR surfaces, not exclusively from molecular gas. The Madden et al. (2020) calibration explicitly separates the H i and H_2 contributions within its Cloudy framework, but applying a single characteristic model to an unresolved galaxy can still overestimate $M_{\rm H_2}$ if the molecular material is distributed among smaller clumps with different densities, radiation fields, or filling factors. Our [C ii]-based molecular masses should therefore be interpreted as model-dependent estimates of the molecular gas associated with PDRs, with systematic uncertainties that are likely larger than the formal line-flux errors.

4.4. Cosmic environment and local density

To investigate the influence of large-scale structure on the gas reservoirs of compact emission-line galaxies, we crossmatched our sample with the VoidFinder catalog derived from the SDSS DR7 volume-limited main galaxy sample (Douglass et al. 2023). The VoidFinder algorithm (El-Ad & Piran 1997; Sutter et al. 2012) classifies the cosmic web into voids and wall structures based on the spatial distribution of galaxies. For the 56 sources in our sample located within the reliable redshift range of the SDSS DR7 volume-limited catalog ($z \leq 0.107$), we find that 22 galaxies (39.3%) reside strictly within cosmic voids, while the remaining 34 (60.7%) are located in wall structures.

This void fraction is higher than the $\sim$ 20% reported for the general population of similar compact starbursts by Brunker et al. (2022). We interpret this enhancement cautiously because our sample requires targeted H I or molecular gas observations and is therefore not a volume-complete census. Nevertheless, a low-density environment may help galaxies retain cold gas by reducing the importance of external removal processes such as ram-pressure stripping and high-speed encounters (Gunn & Gott 1972; Moore et al. 1996). This provides a plausible context for the high gas fractions observed in part of the sample, but it does not by itself prove that environment is the dominant driver.

Panel (a) of Fig. B.1 in Appendix B shows the distribution of the normalized distance to the nearest void center ($d_{\rm void}/R_{\rm eff}$). All void galaxies have $d_{\rm void}/R_{\rm eff} < 1$ by definition, with a median of 0.78, indicating that they reside well within the void interiors rather than at the void edges. The wall galaxies span a wider range of $d_{\rm void}/R_{\rm eff}$, with a significant fraction located near void boundaries ($1 < d_{\rm void}/R_{\rm eff} < 2$).

To further quantify the local environmental density, we computed the projected third-nearest-neighbor surface density (Σ_3) using SDSS DR18 spectroscopic galaxies within a line-of-sight velocity window of $\pm 1000\,{\rm km\,s^{-1}}$. Panels (b) and (c) of Fig. B.1 in Appendix B reveal a clear contrast in local density: compact emission-line galaxies in voids exhibit a median local density of $\Sigma_3 = 0.32\,{\rm Mpc^{-2}}$, whereas those in walls have a median density of $\Sigma_3 = 0.87\,{\rm Mpc^{-2}}$. The pronounced isolation of the void subsample suggests that internal processes can sustain vigorous star formation without requiring a major external interaction, although deeper imaging and spatially resolved gas kinematics are needed to test this interpretation directly. This finding is further corroborated by our rigorous morphological inspection (Sect. 2.6), which excluded systems with major merger signatures.

Figure 7 presents the redshift versus normalized void distance for the full sample. The void galaxies (cyan circles) are cleanly separated below the $d/R_{\rm eff} = 1$ boundary, while wall galaxies (red squares) scatter to much larger distances. The four sources beyond the SDSS DR7 volume-limited redshift boundary ($z > 0.107$) are all classified as wall galaxies due to their large distances from identified voids, although we note that the void catalog is incomplete at these redshifts.

The sky distribution of void and wall galaxies is presented in Appendix B.

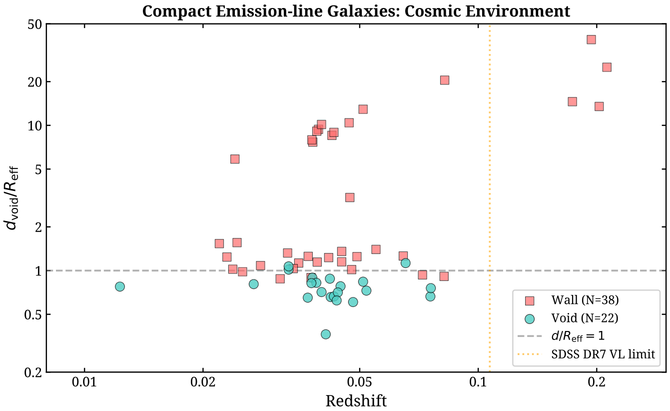

Fig. 7: Redshift vs. normalized distance to the nearest void center ($d_{\rm void}/R_{\rm eff}$) for the full sample of 60 compact emission-line galaxies, shown on logarithmic axes. The cyan circles denote void residents ($d/R_{\rm eff} < 1$), and the red squares denote wall galaxies. The horizontal dashed line marks the void boundary at $d/R_{\rm eff} = 1$. The vertical dotted line indicates the upper redshift limit of the SDSS DR7 volume-limited sample ($z = 0.107$).

5. Conclusions

We assembled a sample of 60 compact emission-line galaxies (Green Peas and Blueberries), carefully excluding systems with morphological signatures of major mergers based on systematic inspection of DESI Legacy Imaging Surveys data. By combining deep ^{12}CO(1–0) observations, archival H I 21 cm data, and [C II]$\lambda 158\,\mu$m spectroscopy from Herschel/PACS, we characterize their atomic and molecular gas reservoirs. Our main conclusions are as follows:

1. CO faintness is ubiquitous. No ^{12}CO(1–0) emission is detected in any of our targets, including two sources with exceptionally deep integration times exceeding 9 hr with the PMO 13.7 m telescope. Even with the metallicity-dependent $\alpha_{\rm CO}$ of Amorín et al. (2016), which yields conversion factors $\sim$ 5–25$\times$ the Galactic value, the physically constraining CO-based upper limits remain below the [C II]-based molecular gas mass estimates. While the number of deep CO observations is limited, these case studies highlight the challenge of using CO as a molecular gas tracer in these systems.

C II reveals substantial molecular gas. Among the five [C II]-detected non-merger sources, we measure molecular gas masses of $M_{\rm H_2}^{\rm [CII]} \sim 10^{9.0}$–$10^{9.7}\,M_\odot$, with a median molecular gas depletion time of $\tau_{\rm dep,H_2} \sim 0.16$ Gyr—significantly shorter than the $\sim$ 1 Gyr typical of main-sequence galaxies.

2. Atomic gas properties are diverse. The ten H I-detected sources span a wide range of atomic gas masses ($M_{\rm HI} \sim 10^{7.7}$–$10^{10}\,M_\odot$) and depletion times ($\tau_{\rm dep,HI} \sim 0.43$–240 Gyr), with a median of 2.6 Gyr. The H I detection rate of 20% (10/51) is consistent with previous surveys of low-mass star-forming galaxies.
3. Rapid gas cycling in the detected subsets. The combination of short depletion times, elevated gas fractions, and high sSFRs in the H I- and [C II]-detected subsets is consistent with rapid gas transformation and consumption through star formation. Because the [C II] result is based on five detections, this conclusion should be read as evidence from individual compact starbursts rather than as a population-average measurement.
4. The multi-tracer approach is essential. Our results underscore that CO alone fails to trace the bulk of the star-forming fuel in low-metallicity starbursts. A complete characterization of the ISM requires the integration of multiple tracers—namely H I, [C II], and deep CO upper limits—to account for the dominant CO-dark molecular gas component.
5. High void fraction and isolation. Environmental analysis using the SDSS DR7 VoidFinder catalog reveals that nearly 40% of the compact emission-line galaxies in our volume-limited subsample reside in cosmic voids. Their local surface densities (Σ_3) are significantly lower than those in galaxy walls. This isolation may help preserve cold gas reservoirs by reducing external gas removal, although a causal connection requires spatially resolved follow-up.

Future progress will require spatially resolved observations of the multiphase ISM—including H i interferometry, [C ii] mapping with JWST or ALMA, and high-sensitivity CO imaging—to directly probe the spatial distribution, kinematics, and physical conditions of atomic, molecular, and ionized gas components in these analogs of high-redshift star-forming systems. Such observations would test whether the short depletion times inferred here correspond to compact central reservoirs or more extended multiphase gas.

Data availability

The LAMOST spectra used in this work are available in electronic form at the CDS via anonymous ftp to `cdsarc.u-strasbg.fr` (130.79.128.5) or via `http://cdsweb.u-strasbg.fr/cgi-bin/qcat?J/A+A/`.

Acknowledgements. This work is supported by the National Key R&D Program of China (grant Nos. 2025YFF0510602, 2021YFA1600401, and 2021YFA1600400), the National Natural Science Foundation of China (NSFC; grant Nos. 12273075, 12090041, and 12090040), the Basic and Applied Basic Research Foundation of Guangdong Province (No. 2025A1515012126), and the Fundamental Research Funds for the Central Universities (grant No. FRF-TP-26-048). This work made use of the data from LAMOST (Large Sky Area Multi-Object Fiber Spectroscopic Telescope, also known as the Guoshoujing Telescope) (https://cstr.cn/31118.02.LAMOST). LAMOST is a Chinese national mega-science facility, operated by National Astronomical Observatories, Chinese Academy of Sciences. This publication makes use of data products from the 13.7m telescope of the Qinghai observation station of the Purple Mountain Observatory and the millimeter wave radio astronomy database. We are grateful to Ji-Xian Sun, Bing-Gan Ju, and the staff at the Qinghai Station of the Purple Mountain Observatory for their continuous support during the observations and data reduction. S.L. also thanks Zheng Zheng for his assistance in understanding the CRAFTS data. The Herschel PACS spectroscopic data used in this work are based on observations obtained from the Herschel Science Archive, under the program "Merger driven vs. Cold-flow driven Star Formation in Typical High-redshift Galaxies: [C ii] Line Observation of Local Analogs of High-redshift Galaxies" (PI: Shim, H.; Proposal ID: OT2_hshim_2). These data were accessed via the SPG, version 14.2.0, and are publicly available through the European Space Agency at `https://doi.org/10.5270/esa-gsw1416`. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5–26555. The Legacy 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). We gratefully acknowledge the use of data from these surveys for our morphological classification.

References

Accurso, G., Saintonge, A., Catinella, B., et al. 2017, MNRAS, 470, 4750
Amorín, R., Muñoz-Tuñón, C., Aguerri, J. A. L., & Planesas, P. 2016, A&A, 588, A23
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5
Bianchi, L., Shiao, B., & Thilker, D. 2017, The Astrophysical Journal Supplement Series, 230, 24
Bigiel, F., Leroy, A., Walter, F., et al. 2008, AJ, 136, 2846
Bolatto, A. D., Leroy, A. K., Rosolowsky, E., Walter, F., & Blitz, L. 2008, ApJ, 686, 948
Bolatto, A. D., Wolfire, M., & Leroy, A. K. 2013, ARA&A, 51, 207
Boquien, M., Burgarella, D., Roehlly, Y., et al. 2019, A&A, 622, A103
Breneman, J. A., McQuinn, K. B. W., Menchaca, R., et al. 2025, ApJ, 991, 191
Brunker, S. W., Salzer, J. J., Trevaskis, A., et al. 2022, ApJ, 926, 145
Bruzual, G. & Charlot, S. 2003, Monthly Notices of the Royal Astronomical Society, 344, 1000
Bruzual, G. & Charlot, S. 2003, MNRAS, 344, 1000
Burgarella, D., Buat, V., & Iglesias-Páramo, J. 2005, MNRAS, 360, 1413
Cardamone, C., Schawinski, K., Sarzi, M., et al. 2009, MNRAS, 399, 1191
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Casey, C. M. 2012, MNRAS, 425, 3094
Catinella, B., Saintonge, A., Janowiecki, S., et al. 2018, MNRAS, 476, 875
Catinella, B., Schiminovich, D., Kauffmann, G., et al. 2010, MNRAS, 403, 683
Chabrier, G. 2003, PASP, 115, 763
Chandola, Y., Tsai, C.-W., Saikia, D. J., et al. 2024, ApJ, 977, L8
Chary, R. & Shim, H. 2015, in Astronomical Society of the Pacific Conference Series, Vol. 499, Revolution in Astronomy with ALMA: The Third Year, ed. D. Iono, K. Tatematsu, A. Wootten, & L. Testi, 75
Curti, M., Mannucci, F., Cresci, G., & Maiolino, R. 2020, MNRAS, 491, 944
Dey, A., Schlegel, D. J., Lang, D., et al. 2019, AJ, 157, 168
Douglass, K. A., Veyrat, D., Martens, B., et al. 2023, ApJS, 265, 24
Durbala, A., Finn, R. A., Crone Odekon, M., et al. 2020, AJ, 160, 271
El-Ad, H. & Piran, T. 1997, ApJ, 491, 421
Gao, Y., Xiao, T., Li, C., et al. 2019, ApJ, 887, 172
Genzel, R., Tacconi, L. J., Combes, F., et al. 2012, ApJ, 746, 69
Gunn, J. E. & Gott, J. Richard, I. 1972, ApJ, 176, 1
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Izotov, Y. I., Guseva, N. G., & Thuan, T. X. 2011, ApJ, 728, 161
Izotov, Y. I., Stasińska, G., Meynet, G., Guseva, N. G., & Thuan, T. X. 2006, A&A, 448, 955
Jiang, T., Malhotra, S., Rhoads, J. E., & Yang, H. 2019, ApJ, 872, 145
Kanekar, N., Ghosh, T., Rhoads, J., et al. 2021, ApJ, 913, L15
Kauffmann, G., Heckman, T. M., Tremonti, C., et al. 2003, MNRAS, 346, 1055
Kennicutt, Robert C., J. 1998, ARA&A, 36, 189
Kennicutt, R. C. & Evans, N. J. 2012, ARA&A, 50, 531
Kennicutt, Jr., R. C., Calzetti, D., Walter, F., et al. 2007, ApJ, 671, 333
Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121
Khasnovis, A., Kanekar, N., Malhotra, S., & Rhoads, J. 2025, ApJ, 991, L21
Kumari, N., Irwin, M. J., & James, B. L. 2020, A&A, 634, A24
Leroy, A. K., Walter, F., Brinks, E., et al. 2008, AJ, 136, 2782
Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, AJ, 146, 19
Li, D. & Pan, Z. 2016, Radio Science, 51, 1060
Lisenfeld, U., Xu, C. K., Gao, Y., et al. 2019, A&A, 627, A107
Liu, S., Luo, A. L., Yang, H., et al. 2022, ApJ, 927, 57
Liu, S., Luo, A.-L., Zhang, W., et al. 2023a, ApJS, 267, 16
Liu, S., Luo, A. L., Zhang, W., et al. 2023b, Research in Astronomy and Astrophysics, 23, 065006
Luo, A. L., Zhang, H.-T., Zhao, Y.-H., et al. 2012, Research in Astronomy and Astrophysics, 12, 1243
Madden, S. C., Cormier, D., Hony, S., et al. 2020, A&A, 643, A141
Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler, A. 1996, Nature, 379, 613
Nan, R. 2006, Science in China: Physics, Mechanics and Astronomy, 49, 129
Nan, R., Li, D., Jin, C., et al. 2011, International Journal of Modern Physics D, 20, 989
Oke, J. B. & Gunn, J. E. 1983, ApJ, 266, 713
Osterbrock, D. E. & Ferland, G. J. 2006, Astrophysics of gaseous nebulae and active galactic nuclei
Peimbert, M. & Torres-Peimbert, S. 1984, in Astrophysics and Space Science Library, Vol. 112, Planetary Nebulae, ed. D. R. Flower, 155
Pettini, M. & Pagel, B. E. J. 2004, MNRAS, 348, L59
Pety, J. 2005, in SF2A-2005: Semaine de l'Astrophysique Francaise, ed. F. Casoli, T. Contini, J. M. Hameury, & L. Pagani, 721
Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1
Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2
Roberts, M. S. & Haynes, M. P. 1994, ARA&A, 32, 115
Saintonge, A. & Catinella, B. 2022, ARA&A, 60, 319
Saintonge, A., Catinella, B., Tacconi, L. J., et al. 2017, ApJS, 233, 22
Schruba, A., Leroy, A. K., Walter, F., et al. 2012, AJ, 143, 138
Solomon, P. M., Downes, D., & Radford, S. J. E. 1992, ApJ, 398, L29
Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJS, 214, 15
Stalevski, M., Fritz, J., Baes, M., Nakos, T., & Popović, L. Č. 2012, MNRAS, 420, 2756
Stalevski, M., Ricci, C., Ueda, Y., et al. 2016, MNRAS, 458, 2288
Steidel, C. C., Strom, A. L., Pettini, M., et al. 2016, ApJ, 826, 159
Su, D.-Q. & Cui, X.-Q. 2004, Chinese J. Astron. Astrophys., 4, 1
Sutter, P. M., Lavaux, G., Wandelt, B. D., et al. 2012, ApJ, 761, 187
Tadaki, K.-i., Iono, D., Hatsukade, B., et al. 2019, ApJ, 876, 1
Wang, S.-G., Su, D.-Q., Chu, Y.-Q., Cui, X., & Wang, Y.-N. 1996, Appl. Opt., 35, 5155
Wolfire, M. G., Hollenbach, D., & McKee, C. F. 2010, ApJ, 716, 1191
Wolfire, M. G., McKee, C. F., Hollenbach, D., & Tielens, A. G. G. M. 2003, ApJ, 587, 278
Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., et al. 2010, AJ, 140, 1868
Wright, E. L., Eisenhardt, P. R. M., & Mainzer, A. K. e. a. 2019, AllWISE Source Catalog
Yang, H., Malhotra, S., Rhoads, J. E., & Wang, J. 2017, ApJ, 847, 38
Zanella, A., Daddi, E., Magdis, G., et al. 2018, MNRAS, 481, 1976
Zhang, W., Li, C., Kauffmann, G., et al. 2009, MNRAS, 397, 1243

Appendix A: Gas-measurement tables

Table A.1: Multiwavelength gas measurements

Source	z	$\int S_{\rm HI}\,dV$ (Jy km s^{-1})	$M_{\rm HI}$ ($10^8\,M_\odot$)	$\int S_{\rm [CII]}\,dV$ (Jy km s^{-1})	$\log L_{\rm [CII]}$ ($L_\odot$)	$M_{\rm H_2}^{\rm [CII]}$ ($10^8\,M_\odot$)	12+ log(O/H)	SFR ($M_\odot$ yr^{-1})	M_* ($10^8\,M_\odot$)	$\sigma_{T_{\rm CO}}$ (mK)	$t_{\rm int}$ (hr)	$M_{\rm H_2}^{\rm CO}$ ($10^8\,M_\odot$)	Ref.
[C ii] detections (S/N ≥ 3.5):													
J0852+1216	0.0756	<0.050	<14.0	300.7 ± 14.5	$7.82^{+0.02}_{-0.02}$	50.1 ± 2.3	8.02±0.02	19.3 ± 0.6	7.52 ± 2.83	3.67	3.5	–	K21+H/P
J0905+4410	0.0654	–	–	132.2 ± 15.8	$7.33^{+0.05}_{-0.06}$	17.0 ± 1.9	8.05±0.01	4.34 ± 0.24	6.28 ± 1.27	3.78	3.7	–	H/P
J1429+0643	0.1735	–	–	60.1 ± 18.2	$7.85^{+0.11}_{-0.16}$	55.0 ± 16.7	8.01±0.01	46.5 ± 1.7	14.6 ± 5.5	3.10	4.0	–	H/P
J1528+3956	0.0645	–	–	77.4 ± 11.2	$7.08^{+0.06}_{-0.07}$	9.8 ± 1.4	8.18±0.05	5.96 ± 0.20	4.55 ± 0.53	4.45	4.0	–	H/P
J1735+5703	0.0472	–	–	266.4 ± 20.8	$7.35^{+0.03}_{-0.04}$	17.8 ± 1.4	8.07±0.03	14.1 ± 0.2	6.62 ± 1.22	–	–	–	H/P
[C ii] non-detections (S/N < 3.5):													
J0240-0828	0.0822	–	–	< 60.0	< 7.19	< 12.3	7.82 ± 0.01	5.36 ± 0.30	3.37 ± 0.21	–	–	–	H/P
J0333+0017	0.1938	–	–	< 73.5	< 8.04	< 83.2	7.90±0.02	20.5 ± 3.2	12.9 ± 3.7	–	–	–	H/P
J0840+1805	0.0722	–	–	< 30.0	< 6.77	< 4.7	8.09 ± 0.05	1.85 ± 0.31	2.49 ± 1.28	4.07	4.2	–	H/P
J1541+4536	0.2029	–	–	< 136.5	< 8.35	< 166.0	8.12±0.02	69.7 ± 2.1	6.83 ± 2.95	5.00	2.1	–	H/P
H i or CO sources without [C ii] coverage:													
J0847+2944	0.0492	<0.093	<27.9	–	–	–	7.74 ± 0.56†	0.050 ± 0.005*	0.682 ± 0.336	23.28	0.5	–	FAST/L23
J0856+3111	0.0758	<0.239	<67.6	–	–	–	7.19 ± 0.45†	0.234 ± 0.028*	1.58 ± 0.84	7.30	4.2	–	FAST/L23
J1108+2238	0.0238	0.155(016)	4.1(0.4)	–	–	–	8.07±0.02	0.299 ± 0.031	0.656 ± 0.127	7.30	5.0	–	K21
J1200+2719	0.0819	0.310(028)	100.6(9.0)	–	–	–	8.04±0.01	3.16 ± 0.32	2.58 ± 0.96	3.64	9.2	<188.9	K21
J1319+0050	0.0477	0.234(025)	21.7(2.3)	–	–	–	8.27±0.04	0.924 ± 0.112	4.14 ± 0.67	6.14	6.1	–	K21
J1509+4543	0.0481	0.543(081)	60.0(8.9)	–	–	–	8.13±0.02	2.06 ± 0.09	4.39 ± 0.59	3.75	14.8	<66.0	K21
J1545+0858	0.0377	0.302(019)	17.8(1.1)	–	–	–	7.72±0.01	2.88 ± 0.09	1.21 ± 0.52	6.37	6.9	–	K21
J1550+2320‡	0.2123	–	–	–	–	–	8.24 ± 0.09	35.9 ± 2.9	24.1 ± 9.3	35.46	0.1	–	L23

Notes: Physical properties and multiwavelength gas measurements for the Green Pea galaxy subsample with [C ii] or CO observations. Columns: (1) source identifier; (2) spectroscopic redshift; (3) integrated H i 21 cm flux density; (4) atomic gas mass; (5) integrated [C ii]λ158 μm flux density from Gaussian fitting; (6) [C ii] line luminosity; (7) molecular gas mass derived from [C ii]; (8) gas-phase metallicity; (9) SFR; (10) stellar mass; (11) CO brightness temperature rms; (12) CO integration time; (13) molecular gas mass upper limit from CO; (14) data source. Upper limits are denoted by "<". Parenthetical uncertainties refer to the last significant digits. [C ii] detections are defined by S/N ≥ 3.5, and [C ii] upper limits are 3σ values computed as $3\,\sigma_{\rm rms}\,\sqrt{2\pi}\,\sigma_v$ with $\sigma_v = 50\,{\rm km\,s^{-1}}$. HAE0840 (≡ J0840+1805; S/N = 3.2) is conservatively classified as a non-detection given its marginal signal and elevated reduced χ^2. J1550+2320 was observed by Herschel/PACS but no usable [C ii] line constraint is available. ‡ J1550+2320 is retained only to document a very shallow ancillary PMO pointing; it has no gas-mass constraint and is excluded from all gas-mass statistics and scaling-relation fits. † Only LAMOST spectrum available, measured with the direct-T_e method. * SFR derived with NUV photometry only. References: K21: Kanekar et al. (2021); FAST: this work (CRAFTS); H/P: Herschel/PACS (OT2_hshim_2); L23: Liu et al. (2023a).

Table A.2: H I-only measurements

Source	z	$\int S_{\rm HI}\, dV$ (Jy km s^{-1})	$M_{\rm HI}$ ($10^8\, M_\odot$)	12+ log(O/H)	SFR ($M_\odot$ yr^{-1})	M_* ($10^8\, M_\odot$)	Ref.
Khasnovis et al. (2025) — GBT:							
J0024-0103	0.0393	0.127(016)	9.3(1.2)	7.90±0.03	0.085 ± 0.002[*]	0.465 ± 0.213	Kh25
J0046+3002	0.0425	<0.034	<3.0	7.64±0.04	0.203 ± 0.013	0.0982 ± 0.0706	Kh25
J0109-0106	0.0389	<0.056	<4.2	7.98±0.04	0.221 ± 0.093	0.721 ± 0.148	Kh25
J0159+0722	0.0241	<0.121	<3.4	7.61±0.12	0.011 ± 0.002[¶]	0.102 ± 0.046	Kh25
J0801+3823	0.0376	<0.046	<3.1	8.12±0.02	0.436 ± 0.004	2.32 ± 0.83	Kh25
J0822+1042	0.0369	<0.026	<1.7	7.68±0.05	0.023 ± 0.005[*]	0.0628 ± 0.0315	Kh25
J0825+1846[a]	0.0379	<0.034	<2.3	7.78±0.02[Y]	0.101 ± 0.014[*]	0.258 ± 0.124	Kh25, Ch24
J0840+5333	0.0314	0.204(020)	9.5(1.0)	8.00±0.05	0.0040 ± 0.0017[*]	0.193 ± 0.085	Kh25
J0926+4504[a]	0.0422	<0.030	<2.5	7.72±0.03	0.027 ± 0.008[*]	0.245 ± 0.123	Kh25, Ch24
J1146+0342	0.0447	<0.031	<2.9	8.01±0.06	0.136 ± 0.020[*]	0.426 ± 0.248	Kh25
J1201+1732	0.0440	0.138(012)	12.7(1.1)	7.78±0.08	0.014 ± 0.005[*]	0.116 ± 0.059	Kh25
J1204+3538	0.0339	<0.048	<2.7	7.81 ± 0.09	0.0084 ± 0.0025[*]	0.102 ± 0.053	Kh25
J1220+4915	0.0123	<0.071	<0.5	7.24 ± 0.04	0.0150 ± 0.0002	0.0108 ± 0.0053	Kh25
J1248+4123	0.0269	<0.096	<3.4	7.61 ± 0.11	0.0038 ± 0.0007[*]	0.0370 ± 0.0211	Kh25
J1359-0103	0.0244	<0.136	<3.9	8.28 ± 0.06	0.090 ± 0.027	1.83 ± 0.35	Kh25
J1402+3033	0.0252	<0.070	<2.1	7.87 ± 0.11	0.016 ± 0.002[¶]	0.0589 ± 0.0302	Kh25
J1444+0409[a]	0.0388	<0.041	<3.0	7.83±0.02	0.313 ± 0.054	0.358 ± 0.059	Kh25, Ch24
J2247+0407	0.0378	<0.041	<2.9	7.65 ± 0.04	0.046 ± 0.010[*]	0.085 ± 0.0453	Kh25
Chandola et al. (2024) — FAST:							
J0146+0319	0.047	–	<2.2	7.624±0.003[Y]	0.562 ± 0.017	0.383 ± 0.124	Ch24
J0216+1715	0.0400	–	<0.7	7.351±0.005[Y]	0.025 ± 0.008[*]	0.157 ± 0.076	Ch24
J0238+0124	0.0510	–	<2.2	7.65±0.01[Y]	0.118 ± 0.026[*]	0.282 ± 0.143	Ch24
J0357+1808	0.0380	–	<2.7	7.783±0.007[Y]	0.317 ± 0.079[**]	0.180 ± 0.047	Ch24
J0820+5431	0.0390	–	<0.9	7.53±0.005[Y]	0.026 ± 0.006[*]	0.0687 ± 0.0345	Ch24
J0827+1059	0.0450	–	<1.5	7.610±0.006[Y]	0.291 ± 0.009	0.277 ± 0.142	Ch24
J0837+1823	0.0420	–	<1.7	7.877±0.005[Y]	0.072 ± 0.006[*]	0.126 ± 0.072[b]	Ch24
J0922+6324	0.0400	–	<1.8	7.693±0.013[Y]	0.036 ± 0.007[*]	0.163 ± 0.080	Ch24
J1026+0426	0.0430	–	0.5	7.813±0.006[Y]	0.064 ± 0.009[*]	0.159 ± 0.081	Ch24
J1032+4919	0.0450	–	<7.4	7.638±0.009[Y]	0.740 ± 0.084	0.225 ± 0.112	Ch24
J1035+1400	0.0410	–	<3.8	7.825±0.014	0.012 ± 0.007[*]	0.156 ± 0.081	Ch24
J1113+0301	0.0230	–	<0.8	7.780±0.003[Y]	0.147 ± 0.029	0.0848 ± 0.0279	Ch24
J1123+2050	0.0328	–	<2.7	7.936±0.019[Y]	3.70 ± 0.11	0.854 ± 0.516	Ch24
J1132+0809	0.0510	–	1.7	7.637±0.005[Y]	0.391 ± 0.111[**]	0.293 ± 0.121	Ch24
J1136+3427	0.0350	–	<0.8	7.15 ± 0.18[†]	0.023 ± 0.007[*]	0.125 ± 0.062	Ch24
J1139+0040	0.0417	–	<4.3	7.674±0.003[Y]	0.072 ± 0.004[*]	0.558 ± 0.275	Ch24
J1323-0132	0.0220	–	<1.3	7.78±0.01	0.252 ± 0.022	0.0732 ± 0.0454	Ch24
J1347+0755	0.0437	–	<6.2	7.98±0.01[Y]	0.452 ± 0.089	0.324 ± 0.156	Ch24
J1355+4651	0.0280	–	<1.2	7.449±0.008[Y]	0.154 ± 0.006	0.0680 ± 0.0327	Ch24
J1400+1951	0.0550	–	<1.5	7.863±0.002[Y]	0.0437 ± 0.0007[#]	0.178 ± 0.093	Ch24
J1509+3731	0.0330	–	<3.6	7.873±0.042[Y]	1.21 ± 0.05	0.783 ± 0.058	Ch24
J1556+4806	0.0520	–	<4.0	7.831±0.089[Y]	0.661 ± 0.068	0.455 ± 0.093	Ch24
J1602+1445	0.0370	–	<4.5	–[§]	0.170 ± 0.058	0.200 ± 0.053	Ch24
J1608+3528	0.0330	–	<4.2	7.824±0.033[Y]	0.186 ± 0.035	0.147 ± 0.044	Ch24
J2320+1225	0.0430	–	<0.8	7.638±0.009[Y]	0.324 ± 0.005	0.177 ± 0.066	Ch24

Notes: Physical properties and H I measurements for the Green Pea galaxy subsample observed only in the 21 cm line. Columns: (1) source identifier; (2) spectroscopic redshift; (3) integrated H I 21 cm flux density; (4) atomic gas mass; (5) gas-phase metallicity; (6) SFR; (7) stellar mass; (8) data source. Upper limits are 3σ values and are denoted by "<". Parenthetical uncertainties refer to the last significant digits.

[a] Source observed in both the Kh25 and Ch24 surveys; the more sensitive measurement is listed under the primary reference.

[Y] Metallicity from Yang et al. (2017) using the direct-T_e method with MMT or SDSS spectra.

[C] Metallicity from Chandola et al. (2024) using the direct-T_e method.

[†] Only LAMOST spectrum available, measured with the direct-T_e method.

[§] Only LBT spectrum available given Yang et al. (2017), however, without metallicity provided.

[*] SFR derived with NUV photometry only.

[**] SFR derived with IR photometry only.

[¶] SFR derived with Hα luminosity with SDSS spectrum.

[#] SFR derived with Hα luminosity with LAMOST spectrum.

[b] Fitted using GALEX FUV/NUV and Pan-STARRS1 *grizy* photometry (no SDSS or AllWISE data available).

References: Kh25: Khasnovis et al. (2025); Ch24: Chandola et al. (2024).

Appendix B: Cosmic environment distributions

In Sect. 4.4, we quantified the large-scale cosmic environment of our sample using the SDSS DR7 VoidFinder catalog and local density estimates. Figure B.1 presents the detailed distributions of these environmental metrics. Panel (a) shows that void galaxies are deeply embedded within void interiors ($d_{\rm void}/R_{\rm eff} < 1$), while wall galaxies are often found near void boundaries. Panels (b) and (c) illustrate the pronounced difference in local density, with void galaxies residing in significantly underdense regions compared to their wall counterparts.

Figure B.2 shows the sky distribution of our sample in equatorial coordinates, color-coded by their VoidFinder classification. The void and wall galaxies are distributed across the SDSS spectroscopic footprint without obvious spatial segregation, indicating that the elevated void fraction reported in Sect. 4.4 is not driven by a localized overdensity or survey edge effect.

Appendix C: Herschel PACS spectra

In Fig. C.1, we present the Herschel/PACS [C II]λ158 μm spectra for the non-merger subsample, along with the best-fit Gaussian models used to measure the line fluxes. The S/N and detection status for each source are listed in Table A.1.

Appendix D: DECaLS morphological cutouts

We present DESI Legacy Imaging Surveys false-color (*grz*) cutout images for all 60 sources in our final sample (Fig. D.1), used for the morphological classification described in Sect. 2.6. Each panel is $21'' \times 21''$ (80 pixels × 80 pixels) centered on the SDSS coordinates. Sources with minor tidal features or asymmetries (category ii) are marked with an asterisk.

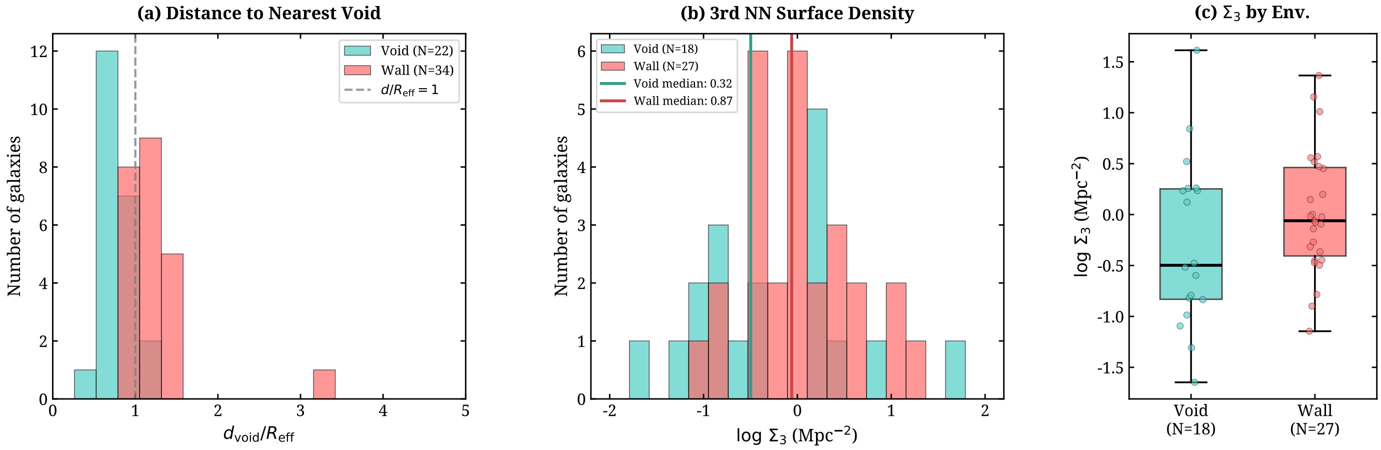

Fig. B.1: Cosmic environment of compact emission-line galaxies based on the SDSS DR7 VoidFinder catalog. (a) Distribution of the normalized distance to the nearest void center ($d_{\rm void}/R_{\rm eff}$), where $R_{\rm eff}$ is the effective radius of the nearest void. The dashed line marks $d/R_{\rm eff} = 1$, the boundary between void interiors and wall structures. (b) Distribution of the projected third-nearest-neighbor surface density (Σ_3) for void (cyan) and wall (red) galaxies. The vertical lines indicate the median Σ_3 for each subsample. (c) Box plot comparing log Σ_3 for void ($N = 18$) and wall ($N = 27$) galaxies, with individual data points overlaid. Void galaxies exhibit systematically lower local densities, with a median Σ_3 approximately 2.7 times lower than wall galaxies.

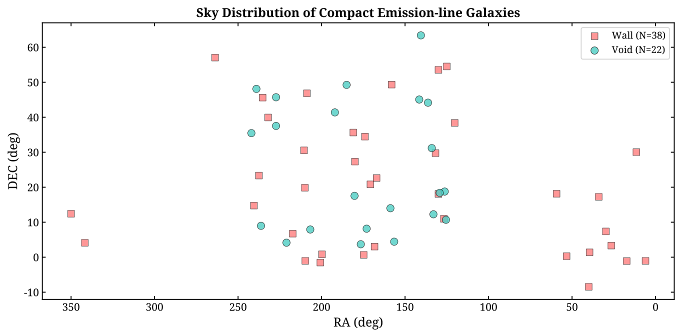

Fig. B.2: Sky distribution of the 60 compact emission-line galaxies in our sample, color-coded by their VoidFinder classification: cyan circles for void galaxies ($N = 22$) and red squares for wall galaxies ($N = 38$). The sample spans the SDSS spectroscopic footprint, and the two populations are spatially intermixed across the survey area.

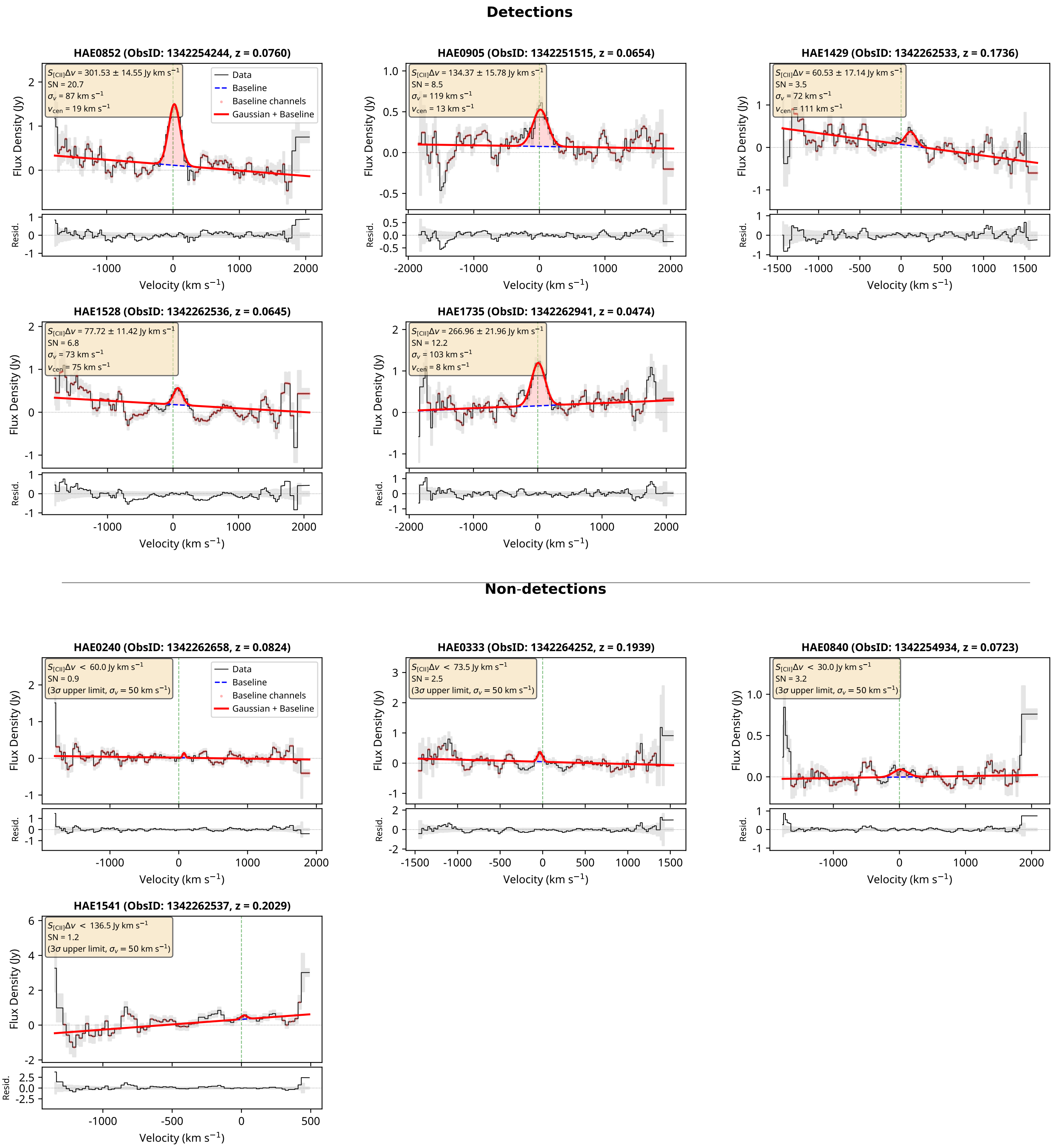

Fig. C.1: Herschel/PACS [C II]$\lambda158\,\mu$m spectra for the non-merger GP galaxy subsample. Upper section: Five detected sources (S/N $\geq$ 3.5). Lower section: Four non-detected sources (including HAE0840 with S/N = 3.2, conservatively classified as a non-detection). In each panel, the black histogram shows the observed spectrum, the red curve shows the best-fit Gaussian model plus baseline, and the gray shading indicates the 1σ flux uncertainties. The dashed blue line marks the fitted linear baseline, and the vertical dashed green line indicates the systemic velocity ($v = 0\,\mathrm{km\,s^{-1}}$). The lower subpanel of each source shows the fit residuals. For detections, the inset text box reports the integrated flux ($S_{[\mathrm{CII}]}\Delta v$), S/N, velocity dispersion (σ_v), and central velocity offset (v_{cen}). For non-detections, 3σ upper limits on $S_{[\mathrm{CII}]}\Delta v$ are reported assuming $\sigma_v = 50\,\mathrm{km\,s^{-1}}$, computed as $3\,\sigma_{\mathrm{rms}}\sqrt{2\pi}\,\sigma_v$, where σ_{rms} is the baseline noise level. All measurements are listed in Table A.1.

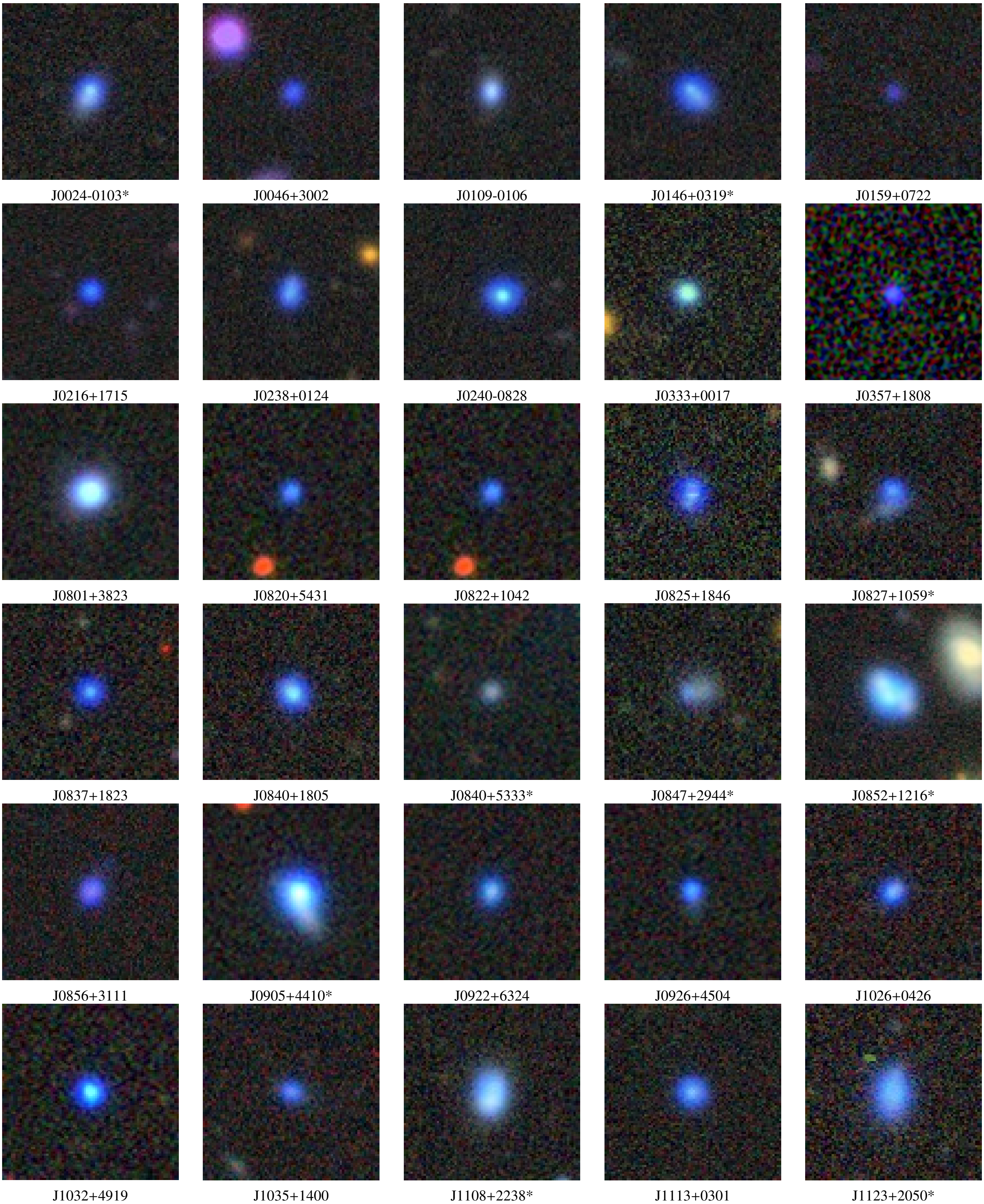

Fig. D.1: DESI Legacy Imaging Surveys false-color (*grz*) cutout images for the final sample (Page 1 of 2). (Note that J0357+1808 is obtained from SDSS, which is not covered in DeCALS.) Each panel is 21″ × 21″ (80 pixels × 80 pixels) centered on the SDSS coordinates. Sources with minor tidal features or asymmetries (category ii) are marked with an asterisk.

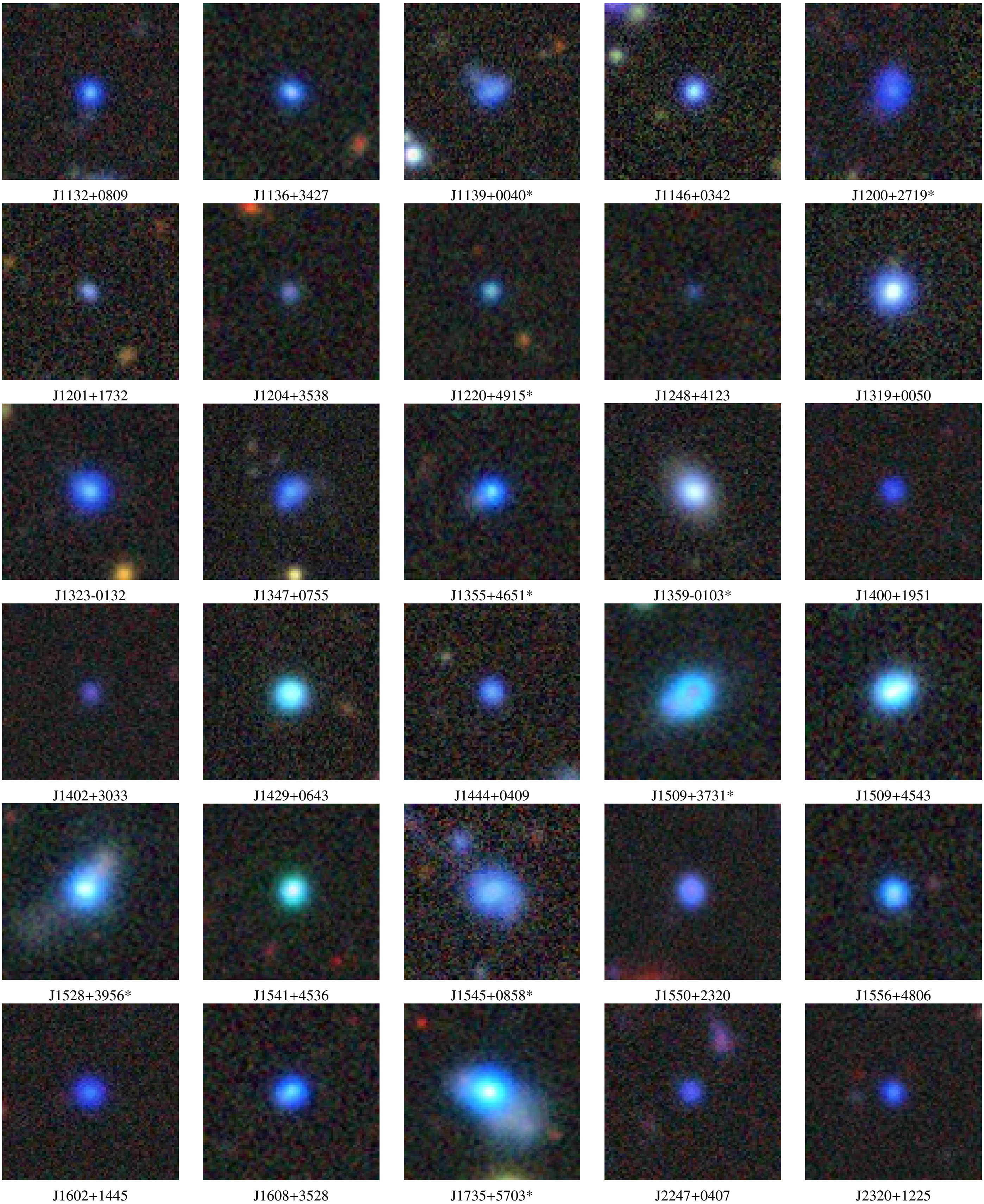

Fig. D.1: DESI Legacy Imaging Surveys false-color (*grz*) cutout images for the final sample (Page 2 of 2).