Scientific justification

CELGs@4MOST: Compact emission-line galaxies studied with S-PLUS and 4MOST

Compact emission line galaxies (CELGs) observed in the nearby Universe are thought to be the analogues of the much more common galaxies at high redshift, representing a unique opportunity to study in detail the building blocks of larger structures in the early Universe, shedding light on how low-mass systems merge, grow, and evolve into the more massive galaxies we observe today. Despite their scientific importance, these compact galaxies are systematically underrepresented in major spectroscopic surveys. Their compact morphology causes them to frequently be misclassified as stars in photometric surveys, leading to their exclusion from galaxy-targeted programs. This has created a significant selection bias in our census of the local Universe, leaving a key population of galaxies almost entirely unexplored spectroscopically. To date, these objects have not been included in any 4MOST survey programs.

These local galaxies are very compact and blue, resembling low-mass ($10^8 - 10^9$ $M_\odot$), metal-poor blue compact dwarf (BCD) galaxies, green peas (GPs) (Cardamone et al., 2009), little blue dots in the Hubble Space Telescope Frontier Fields (Elmegreen et al., 2017), and blueberries (Yang et al., 2017). Although they may be fainter than the original GP and blueberry populations and are expected to exhibit somewhat lower star-formation rates, their overall similarity in morphology and colours allows us to place them within the broader evolutionary sequence connecting nearby star-forming galaxies with their high-redshift counterparts. From a cosmological perspective, low-mass, star-forming galaxies are among the most promising candidates for being sources of escaping Lyman continuum (LyC) photons, thought to be responsible for the reionisation of the Universe after the cosmic Dark Ages. We now have a unique opportunity to obtain 4MOST spectroscopy for a statistically significant population of CELG candidates identified in the Southern Photometric Local Universe Survey (Mendes de Oliveira et al., 2019) imaging data. These candidates resemble other extreme compact systems and will allow us to trace the physical processes that drive the formation and evolution of compact starbursts across cosmic time.

Our program leverages a unique sample from S-PLUS, which provides imaging in 12 photometric bands over 3000 square degrees. The S-PLUS narrow-band filters allow for the efficient identification of CELGs, (we refer to them as S-PLUS CELGs), specifically candidates at $z \approx 0.3$ where the strong [O III]λ5007 emission is redshifted into the J0660 filter (Hα at rest frame). Here we propose to obtain low-resolution spectra for a sample of 523 high-confidence candidates, in order to build the first statistically significant, spectroscopically confirmed catalogue of these objects with 4MOST and S-PLUS.

Scientific Goals

Our proposed program (**CELGs@4MOST**), will provide a dataset with three primary scientific goals:

1. **Spectroscopic Confirmation and Redshift Catalog:** We will obtain definitive spectroscopic redshifts, confirming the nature of our candidates and building a robust catalog of compact emission line galaxies selected from **S-PLUS**.
2. **Characterize Ionization Sources:** By using key diagnostic emission lines (e.g., Hβ, [OIII], Hα, [NII]) in the BPT diagram, together with the [OIII]λ5007/[OII]λ3727 ratio (O32), we will constrain the source and level of the ionization in these systems.
3. **Determine Physical Properties:** With 4MOST LRS spectra, we will (i) derive star-formation rates (SFRs) from dust-corrected Hα fluxes, (ii) estimate gas-phase metallicities using strong-line methods (e.g., **O3N2**, **N2**, **O32**) and/or the direct T_e method, (iii) obtain stellar masses via full spectral fitting with population-synthesis codes (e.g., **Starlight**), and (iv) resolve multi-component emission line profiles, allowing us to determine the integrated kinematical properties of these galaxies associated with their starburst mechanisms.

This analysis is crucial to disentangle the presence of different gas components, such as multiple velocity systems, each of which may trace distinct physical and chemical conditions within the interstellar medium. Such observations will provide crucial insight into whether the strong bursts of star formation in these systems are triggered by minor mergers and/or interactions. With these observations, we will be able to test whether S-PLUS CELGs share the same complex kinematical structures as GPs and blueberries, thereby assessing their role as nearby analogues of high-redshift star-forming systems and providing crucial insight into the mechanisms driving their extreme starbursts, as well as their evolutionary link to compact starbursts across cosmic time.

Our goals are achievable even with a partial dataset. The current number of spectroscopically confirmed galaxies of this type is extremely small, limited to a handful of objects from our own pilot observations. Obtaining spectra for even 10% of our proposed sample would constitute a groundbreaking dataset, dramatically increasing the number of confirmed objects.

Technical justification

We are requesting observations with the **4MOST Low-Resolution Spectrograph (LRS)** for a sample of **523 candidate compact emission line galaxies**. All targets will be observed with a single exposure of **1200 seconds (20 minutes)**, regardless of their **r-band magnitude**.

Our sample spans a range of r-band magnitudes ($r \approx 18.5$ to 20.5 mag), which we divide into three bins for analysis. For each bin, we calculated the **signal-to-noise ratio (S/N)** in the middle-case scenario. All calculations assume a 20-minute exposure:

- **Bright Bin ($r < 19.0$, $N = 12$):** A **continuum S/N ∼ 9.5** at **6000 Å** is expected, enabling detailed spectroscopic analysis across the continuum and emission lines.
- **Intermediate Bin ($19.0 \leqslant r < 20.0$, $N = 181$):** The **continuum S/N ∼ 6** at **6000 Å**, sufficient for robust emission line measurements and redshift confirmation.
- **Faint Bin ($20.0 \leqslant r \leqslant 20.5$, $N = 330$):** The **continuum S/N ∼ 3.5** at **6000 Å**, allowing basic characterization and identification of key emission features.

This uniform exposure strategy maximizes efficiency and ensures that all targets receive at least a baseline spectroscopic characterization. The total **fiber-hours** required for the observations are calculated as follows:

- **Bright Bin:** $12 \times 20\,\text{min} = \mathbf{4}$ **fiber-hours**.
- **Intermediate Bin:** $181 \times 20\,\text{min} = \mathbf{60.33}$ **fiber-hours**.
- **Faint Bin:** $330 \times 20\,\text{min} = \mathbf{110}$ **fiber-hours**.

The **total request is approximately** 174.33 **fiber-hours**. If only the **bright and intermediate bins** are observed, the request reduces to 64.33 **fiber-hours**. Even a partial completion focusing on the brightest objects will deliver important new insights and substantially increase the number of spectroscopically confirmed CELGs in the local Universe.

Suitability for 4MOST

This program is uniquely suited for 4MOST. Our targets are distributed over the 3000 square degrees of the S-PLUS footprint (see Fig 1). The vast field of view and high multiplexing capability of 4MOST are essential to observe such a widely distributed sample efficiently. Building a statistically significant sample with single-object or small-field spectrographs would be prohibitively time-consuming and impractical. The LRS mode is ideal, as it covers the full wavelength range from 370 to 950 nm needed for our primary science goals.

List of investigators and their affiliations

PI: Ricardo Demarco, Institute of Astrophysics, Facultad de Ciencias Exactas, Universidad Andrés Bello, Sede Concepción, Talcahuano, Chile.

Co-I: Claudia Mendes de Oliveira, University of São Paulo, Instituto de Astronomia, Geofísica e Ciências Atmosféricas, São Paulo, Brazil.

Co-I: Chris Haines, Instituto de Astronomia y Ciencias Planetarias, Universidad de Atacama, Copiapó, Chile.

References

Cardamone et al. Galaxy zoo green peas: discovery of a class of compact extremely star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 399(3):1191–1205, nov 2009. ISSN 1365-2966. doi: 10.1111/j.1365-2966.2009.15383.x. URL http://dx.doi.org/10.1111/j.1365-2966.2009.15383.x.

Elmegreen et al. Little blue dots in the hubble space telescope frontier fields: Precursors to globular clusters? *The Astrophysical Journal Letters*, 851(2):L44, dec 2017. doi: 10.3847/2041-8213/aaa0ce. URL https://doi.org/10.3847/2041-8213/aaa0ce.

C. Mendes de Oliveira et al. The Southern Photometric Local Universe Survey (S-PLUS): improved SEDs, morphologies, and redshifts with 12 optical filters. *Monthly Notices of the Royal Astronomical Society*, 489 (1):241–267, oct 2019. doi: 10.1093/mnras/stz1985.

Yang et al. Blueberry galaxies: The lowest mass young starbursts. *The Astrophysical Journal*, 847(1):38, sep 2017. doi: 10.3847/1538-4357/aa8809. URL https://doi.org/10.3847/1538-4357/aa8809.

Target selection criteria

Our target selection is designed to produce a clean and reproducible sample of CELG candidates at $z \approx 0.3$. Candidates were selected from the 3000 square degrees of the Southern Photometric Local Universe Survey (S-PLUS) Data Release 4.

The primary selection criterion identifies objects with a strong flux excess in the J0660 narrow-band filter, which is sensitive to redshifted [OIII]λ5007 emission. We selected sources exhibiting a magnitude contrast of at least 0.5 mag in the J0660 filter relative to the continuum, estimated from the adjacent r and i broad-band filters. Figure 3 shows examples of selected galaxies, illustrating their photometric spectral energy distributions and compact on-sky appearance.

To ensure sample purity, we applied a series of rigorous cuts. We used the S-PLUS catalog parameters to select sources: (1) classified as galaxies (stellar probability $< 25\%$); (2) with photometric redshifts consistent with our goal ($0.25 <$ photo_z < 0.35); and (3) with high-quality photometry (SEX_FLAGS $\leqslant 3$) in a magnitude range of $15 < r < 20.5$. To specifically isolate compact objects, we required a near-circular morphology (elongation between 0.75 and 1.25) and a small angular size (FWHM < 2.5 pixels in the r band). Further cuts based on aperture photometry were used to remove artifacts and objects contaminated by nearby sources.

The robustness of this selection process was validated by recent spectroscopic follow-up of 6 candidates with Gemini/GMOS(using the B600 grating). We achieved a 100% success rate, confirming all targets as compact star-forming galaxies at the intended redshift (with one of the galaxies having also an AGN contribution mixed with the star-forming component - last one in Figure 3). This pilot study demonstrates the high purity and homogeneity of our final sample of 523 candidates. The on-sky distribution of this sample is presented in Figure 1.

Data reduction will use the standard 4MOST pipeline for basic calibration; we will develop a Python module for simultaneous continuum fitting and emission line modelling (delivery within 12 months after data release).

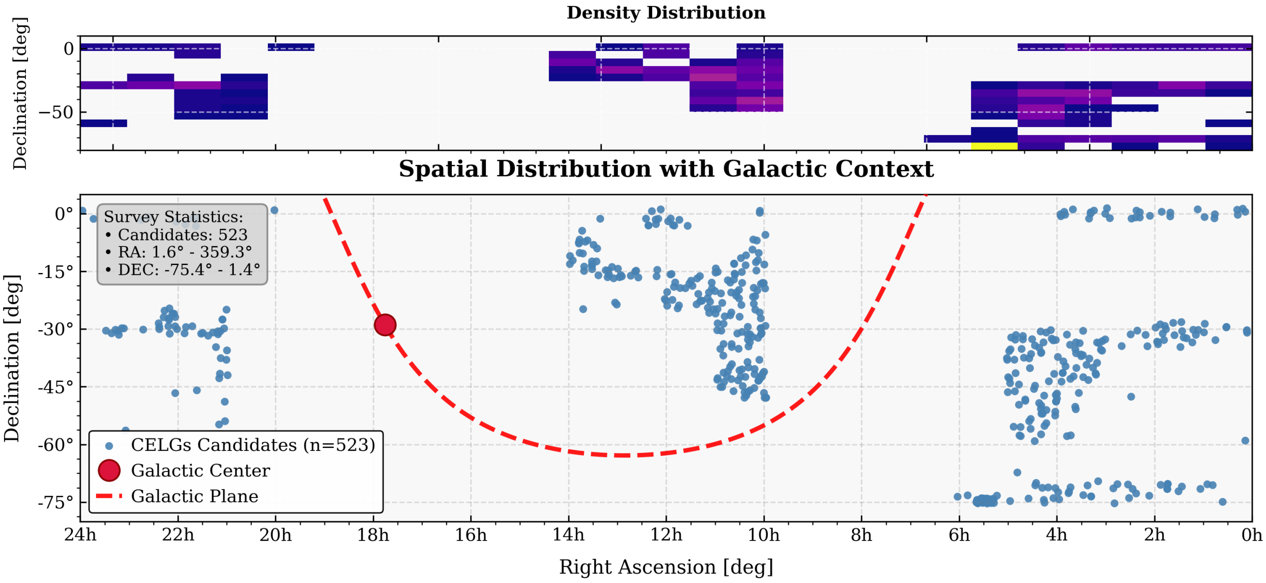

Figure 1: Spatial distribution of compact emission line galaxy candidates identified in the S-PLUS DR4 survey. The figure shows the density of objects in right ascension (RA) and declination (DEC), with blue dots marking the positions of all 523 emission line candidates. The galactic plane is indicated by a red dashed line, and the Galactic center is marked in red. This distribution illustrates the wide coverage and dispersion of our sample across the S-PLUS footprint.

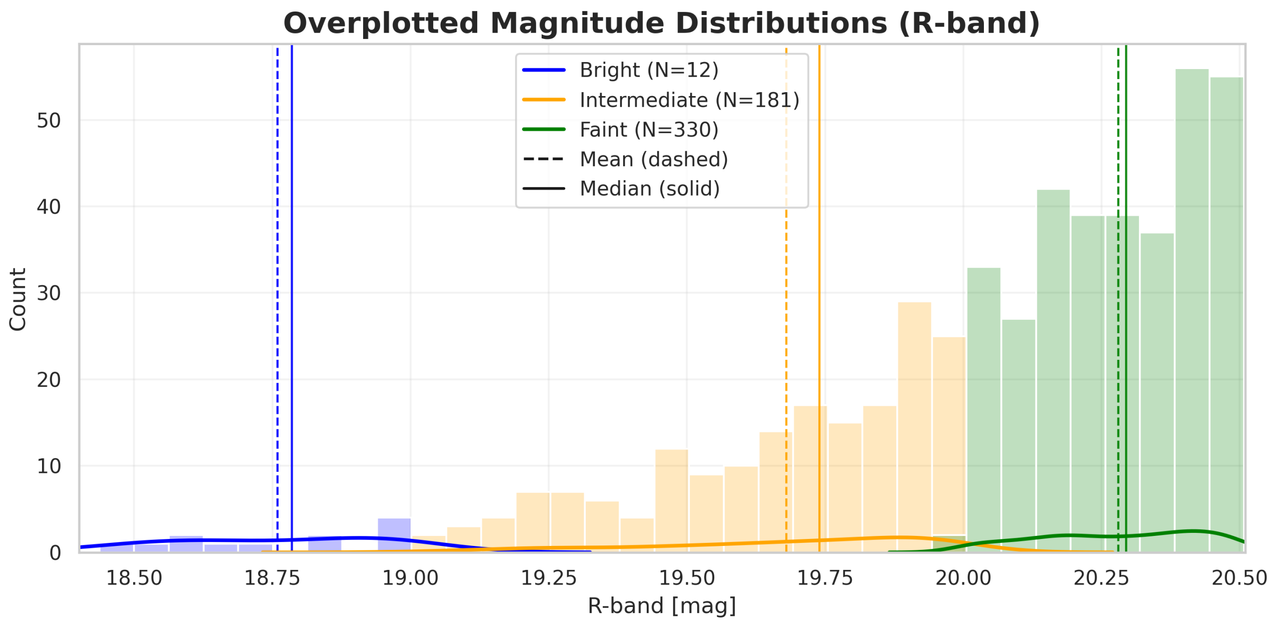

Figure 2: R-band magnitude distribution of the sample, marked by magnitude bin used for exposure-time allocation. The three bins and their sample sizes are: Bright ($r < 19.0$; $N = 12$), Intermediate ($19.0 \leqslant r < 20.0$; $N = 181$), Faint ($20.0 \leqslant r \leqslant 20.5$; $N = 330$). Vertical lines indicate representative central values for each bin. These counts were used to compute the total requested fiberhours (see main text).

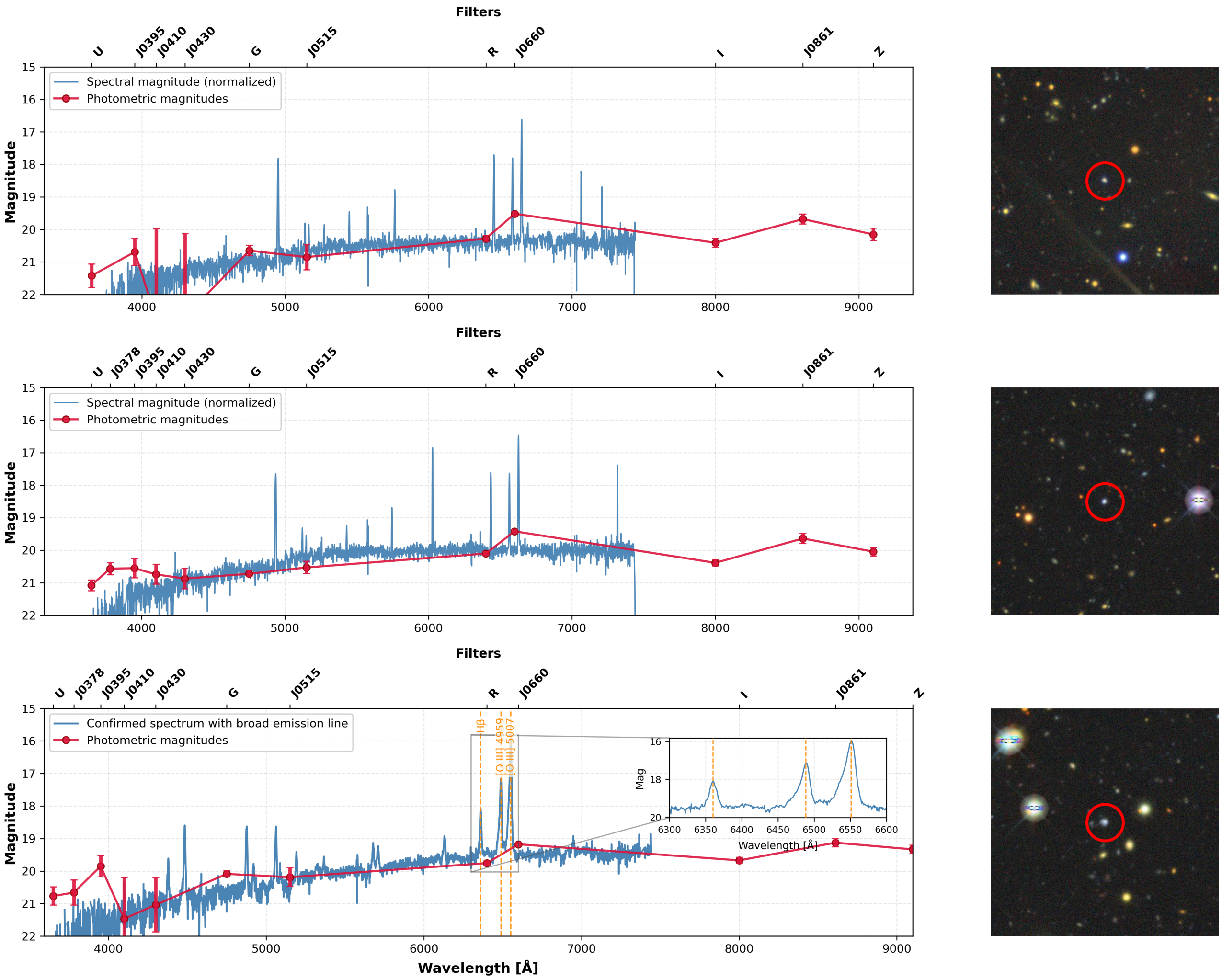

Figure 3: Examples of three galaxies selected for the sample, for which recently we obtained Gemini spectroscopy. In the right-hand panel we show the RGB Legacy image composed of the g, r, and i bands; on the left we display the photometric points (x-axis: central wavelength of each filter; y-axis: magnitude) with the (rough preliminary flux-calibrated) spectrum overplotted, illustrating how the emission lines fall into the Hα filter—for objects at $z \approx 0.3$ this is primarily [O III]. In the third (lower) panel we present one of the galaxies that exhibits broader emission lines; the inset highlights a zoom on the Hβ and [O III] lines, correctly labeled and shifted to the observed frame.