Synergies between low- and intermediate-redshift galaxy population classifications revealed with unsupervised machine learning

Sebastian Turner[1,*], Malgorzata Siudek[2,3,4,**], A. Pollo[4,5], S. Salim[6], K. Malek[4,7], I. K. Baldry[1], et al.

[1] Astrophysics Research Institute, Liverpool John Moores University, UK; [2] Institut de F'isica d'Altes Energies, Barcelona Institute of Science and Technology, Spain; [3] Centre for Theoretical Physics, Polish Academy of Sciences, Warsaw, Poland; [4] National Centre for Nuclear Research, Warsaw, Poland; [5] Astronomical Observatory of the Jagiellonian University, Krakow, Poland; [6] Department of Astronomy, Indiana University, USA; [7] Aix Marseille Univ. CNRS, CNES, LAM Marseille, France

* s.turner1@2012.ljmu.ac.uk * @sebturne * astro.ljmu.ac.uk/~aststurn ** msiudek@ifae.es

Motivation & Abstract

The diversity of galaxies in the Universe reflects the varying balance of processes that influence their evolution. We aim to understand the cosmological evolution of subpopulations of galaxies given by unsupervised machine learning. We derive two samples of galaxies: one at intermediate redshifts from VIPERS, and one at low redshifts from GSWLC. We characterise both samples with the same 13 features. The simplicity of the features facilitates comparisons with other studies. Our clustering approach, which incorporates dimensionality reduction, partitions the samples into 11 and 12 clusters respectively. The partitions are broadly similar, with some subtle differences among clusters containing quiescent galaxies. We also speculate on a possible cluster of post-starburst galaxies.

VIPERS

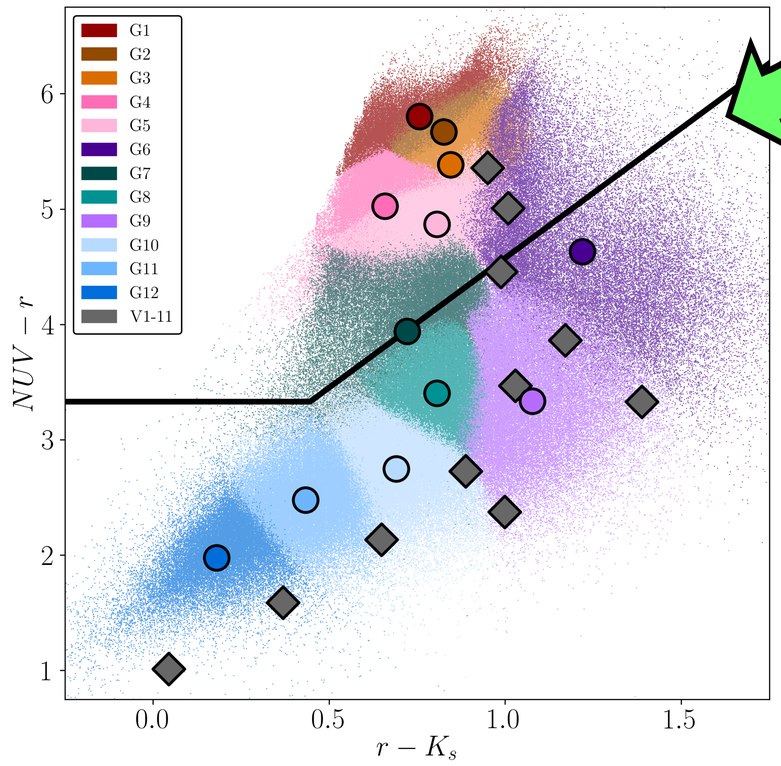
Our sample of **~50000 galaxies** at **intermediate redshifts** ($0.4 < z < 1.3$) is from VIPERS. The **input features** to the clustering are **12 broad-band absolute magnitudes** and **spectroscopic redshifts**. The absolute magnitudes span the full wavelength coverage of VIPERS: ultraviolet to infrared, and are determined via a LePhare SED fit. The spectroscopic redshifts are included to account for any cosmological evolution of galaxies within the redshift range of the survey. FEM found **11 clusters** of galaxies in this sample.

GSWLC

Our sample of **~600000 galaxies** at **low redshifts** ($z < 0.3$) is from GSWLC, itself based on SDSS. Clustering in this sample used the **same input features as** the **intermediate-redshift sample**. The absolute magnitudes were, however, determined via a **different SED fit** (CIGALE), and **without input infrared photometry**. We aim to constrain how these differences may have influenced the clustering. FEM found **12 clusters** of galaxies in this sample.

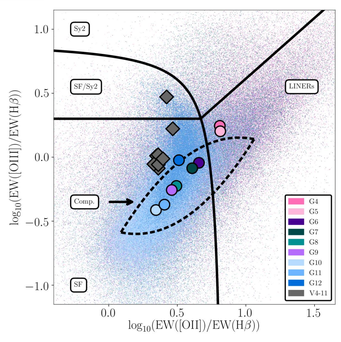
The $NUVrK_s$ plane

The colour-colour plane below offers a clear view of the **clustering results**, which are **broadly consistent**. The black line distinguishes quiescent galaxies (above) and star-forming galaxies (below). The offset of results reflects the reduction in star formation of the galaxy population as a whole with time.



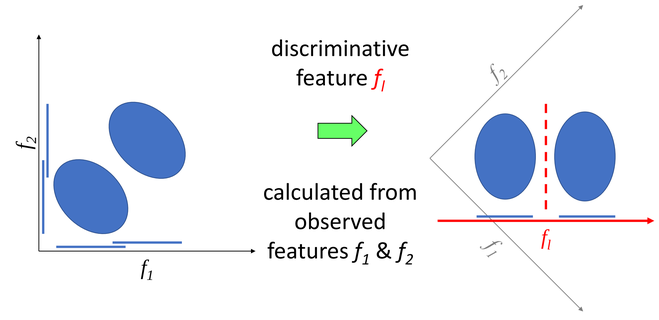
Post Starburst Cluster?

Cluster **G4** galaxies may be **post-starburst**. Their **LINER-like emission** spectra suggest recent star formation or AGN. Their **relatively high Sérsic indices and environmental densities** are also linked to starbursts or AGN via **mergers**. This cluster appears unique to the low-redshift sample.



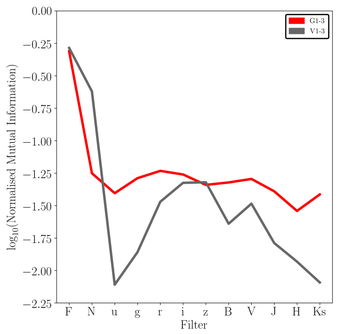
Clustering

We apply the Fisher Expectation-Maximisation **(FEM) clustering algorithm**. It uses **dimensionality reduction to model clusters** in a **discriminative latent subspace** of the input feature space using **Gaussian density functions**. This ensures that only distinguishing information encoded in the input features is used to model the clusters. We use the Bayesian Information Criterion to evaluate clustering.



Red Clusters

Both samples exhibit three **tightly-spaced clusters** at the red ends of their distributions in the $NUVrK_s$ plane. In order to discern **which input features distinguish these quiescent clusters**, we calculate mutual information between each of the features and the cluster labels.



We find that for both samples, it is the **ultraviolet absolute magnitudes** that distinguish these clusters. Hence, we appear to be capturing variation in the **last remnants of star formation activity** in these otherwise quiescent galaxies.

References

FEM: Bouveyron, C. & Brunet, C., 2012, S&C, 22, 301 / *VIPERS:* Scodeggio, M., et al., 2018, A&A, 609, A84 / Garilli, B., et al., 2014, A&A, 562, A23 / Guzzo, L., et al., 2014, A&A, 566, A108 / *GSWLC:* Salim, S., et al., 2018, ApJ, 859, 11 / Salim, S., et al. 2016, ApJ SS, 227, 2 / *Clustering in VIPERS:* **Siudek, M.**, et al., 2018, A&A, 617, A70 / *Clustering in VIPERS with photometric redshifts:* **Siudek, M.**, et al., 2019. Submitted to A&A / *Clustering in GSWLC:* **Turner, S., Siudek, M.**, et al., in prep