Statistical Quantification and Parametrization of Galaxies with SPICA

Tsutomu T. TAKEUCHI

Division of Particle and Astrophysical Science,
Nagoya University, Japan

Annual Meeting of the ASJ Spring, Tsukuba, 16-19 March, 2020

SubWG for the Nearby Galaxies and the Galaxy:

Hiroyuki KANEKO,
Junichi BABA (NAOJ),
Koichiro NAKANISHI （NAOJ, Sokendai),
Kentaro MOTOHARA,
Fumi EGUSA,
Itsuki SAKON (University of Tokyo),
Tsutmu T. TAKEUCHI (Nagoya University),
Hanae INAMI (Hiroshima University),
Yoshimasa WATANABE (Nihon University)

1. Introduction

1.1 What is the galaxy evolution?

Galaxies evolve in various aspects:

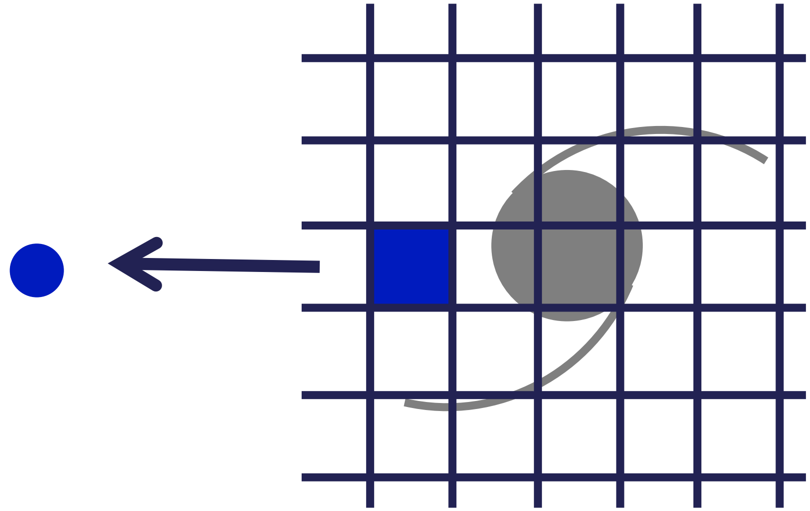
$$\boxed{x = x(T|T > t)}$$

$$\begin{aligned}
\mathrm{SFR}(t) &= f_1(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
M_*(t) &= f_2(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
M_{\mathrm{mol}}(t) &= f_3(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
M_{\mathrm{HI}}(t) &= f_4(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
M_{\mathrm{dust}}(t) &= f_5(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
M_{\mathrm{halo}}(t) &= f_6(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
\delta_{\mathrm{gal}}(t) &= f_7(\mathrm{SFR}, M_*, M_{\mathrm{mol}}, M_{\mathrm{HI}}, M_{\mathrm{dust}}, M_{\mathrm{halo}}, \delta_{\mathrm{gal}}, \ldots) \\
&\vdots
\end{aligned}$$

This is the formal and ultimate goal of the studies on galaxy evolution, but clearly it is a substantially complicated problem. It is time to define the evolution of galaxies with more objective point of view.

1.2 "Evolution" of the ISM

This situation is similar in the studies on the ISM in galaxies.



The ISM evolves drastically, depending on various local physical conditions and processes.

⇒ The governing equations would be also very complex and highly entangled with each other.

2. Example: Classification in Luminosity Space

2.1 Data: VIMOS Public Extragalactic Redshift Survey (VIPERS)

Final public release of complete VIPERS galaxy catalogue of ~90,000 redshifts (PDR-2)

- 18 November 2016 -

Go to PDR-2 data download page

For the press: final science release information page

http://vipers.inaf.it/

VIPERS products

- **Redshifts**
- **Spectra**
- **Photometry from CFHTLS *ugriz* (optical), *GALEX* FUV and NUV (ultraviolet), and additional *ZYJHK* and *K*s (near-infrared) bands (and other properties derived from SED fitting).**

- **Stellar mass from Hyperzmass code**
- **Line-estimated metallicity for subsamples with high-S/N spectra**

- **Density field reconstructed from galaxy distribution**
- **Cluster/group membership**

- **Ancillary data (Herschel, WISE, etc.)**

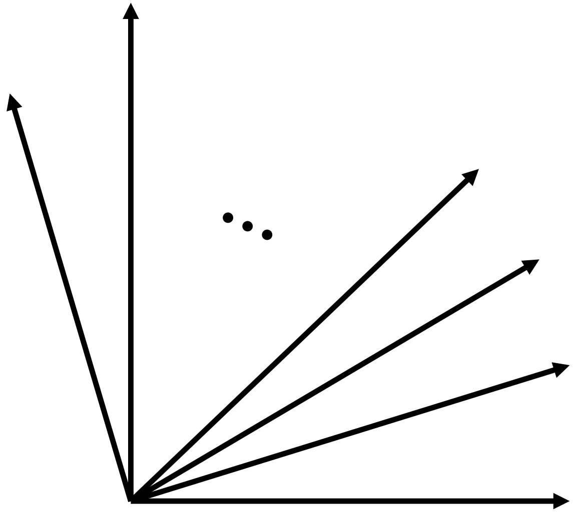
2.2 Classification in multiwavelength luminosity space

Construct a subsample with high-S/N data: 52,114

Redshift range: $0.4 < z < 1.3$

Twelve rest-frame magnitudes (FUV, NUV, *u*, *g*, *r*, *i*, *z*, *B*, *V*, *J*, *H*, *K*s) and redshift, normalized around unity.

⇒ The feature space is 13-dimensional, perhaps impossible to identify any feature by intuition.



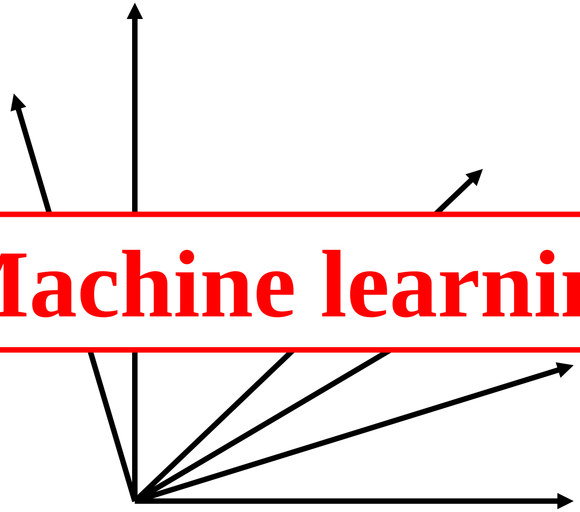
2.2 Classification in multiwavelength luminosity space

Construct a subsample with high-S/N data: 52,114

Redshift range: 0.4 < *z* < 1.3

Twelve rest-frame magnitudes (FUV, NUV, *u*, *g*, *r*, *i*, *z*, *B*, *V*, *J*, *H*, *K*s) and redshift, normalized around unity.

⇒ The feature space is 13-dimensional, perhaps impossible to identify any feature by intuition.



Machine learning!

Unsupervised machine learning in luminosity space

Fisher Expectation-Maximization (FEM) algorithm (Bouveryron & Brunet 2012)

1. **Assign initial cluster (class) centers by *k*-means++**
2. **Execute FEM**
   i. **E-step: calculate the complete log-likelihood under the current value of the Gaussian mixture model**
   ii. **F-step: boundary is chosen to maximize the distances between groups, and to minimize the internal scatters**
   iii. **M-step: parameters of Gaussian functions are optimized by maximizing the conditional expectations of the complete log-likelihood**
   iv. **Back to 2.i (E-step) until the result converges.**

Unsupervised machine learning in luminosity space

Fisher Expectation-Maximization (FEM) algorithm (Bouveryron & Brunet 2012)

1. **Assign initial cluster (class) centers by *k*-means++**
2. **Execute FEM**
   1. **E-step: calculate the complete log-likelihood under the current value of the Gaussian mixture model**
   2. **F-step: boundary is chosen to maximize the distances between groups, and to minimize the internal scatters**
   3. **M-step: parameters of Gaussian functions are optimized by maximizing the conditional expectations of the complete log-likelihood**
   4. **Back to 2.i (E-step) until the result converges.**

Unsupervised machine learning in luminosity space

Fisher Expectation-Maximization (FEM) algorithm (Bouveryron & Brunet 2012)

1. Assign initial cluster (class) centers by *k*-means++
2. Execute FEM
   i. E-step: calculate the complete log-likelihood under the current value of the Gaussian mixture model
   ii. **F-step: boundary is chosen to maximize the distances between groups, and to minimize the internal scatters**
   iii. M-step: parameters of Gaussian functions are optimized by maximizing the conditional expectations of the complete log-likelihood
   iv. Back to 2.i (E-step) until the result converges.

Unsupervised machine learning in luminosity space

Fisher Expectation-Maximization (FEM) algorithm (Bouveryron & Brunet 2012)

1. Assign initial cluster (class) centers by *k*-means++
2. Execute FEM
   i. E-step: calculate the complete log-likelihood under the current value of the Gaussian mixture model
   ii. F-step: boundary is chosen to maximize the distances between groups, and to minimize the internal scatters
   iii. M-step: parameters of Gaussian functions are optimized by maximizing the conditional expectations of the complete log-likelihood
   iv. Back to 2.i (E-step) until the result converges.

Unsupervised machine learning in luminosity space

Fisher Expectation-Maximization (FEM) algorithm (Bouveryron & Brunet 2012)

1. Assign initial cluster (class) centers by *k*-means++
2. Execute FEM
   i. E-step: calculate the complete log-likelihood under the current value of the Gaussian mixture model
   ii. F-step: boundary is chosen to maximize the distances between groups, and to minimize the internal scatters
   iii. M-step: parameters of Gaussian functions are optimized by maximizing the conditional expectations of the complete log-likelihood
   iv. Back to 2.i (E-step) until the result converges.

k-means++ algorithm

1. **Choose one center uniformly at random from sample data points.**

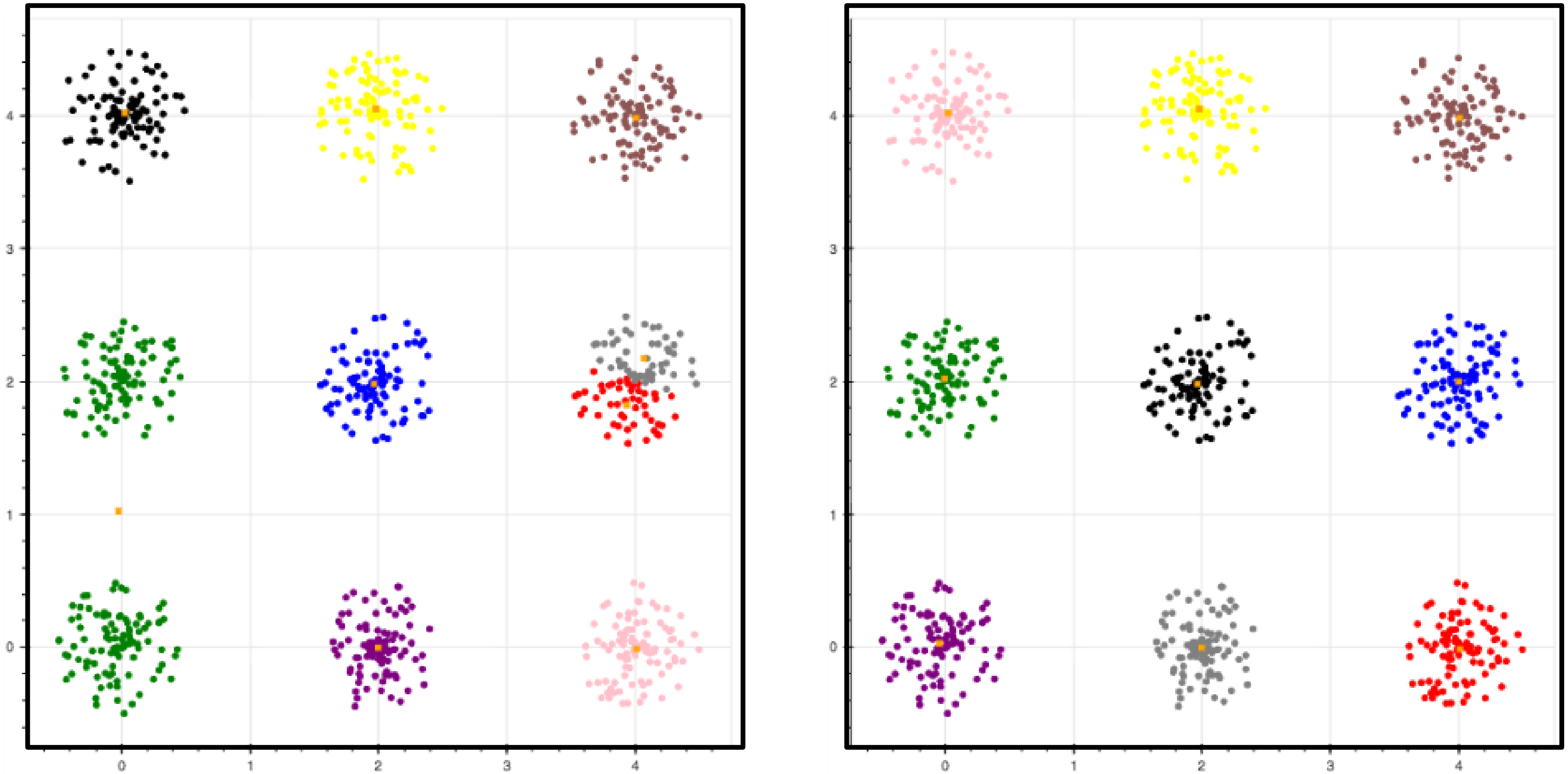
2. **For each data point x, compute $D(x)$, the distance between x and the nearest center that has already been chosen.**

3. **Choose one new data point at random as a new center, using a weighted probability distribution where a point x is chosen with probability proportional to $D(x)^2$.**

4. **Repeat Steps 2 and 3 until k centers have been chosen.**

5. **Now that the initial centers have been chosen, proceed using standard k-means clustering.**

k-means++ algorithm



***k*-means** **_k_-means++**

https://qiita.com/NoriakiOshita/items/cbd46d907d196efe64a3

Gaussian mixture model

For *m* variables and *n* Gaussians,

$$p(\boldsymbol{x}) = \sum_{k=1}^{K} \pi_k \mathcal{N}(\boldsymbol{x} | \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k)$$

$$\mathcal{N}(\boldsymbol{x} | \boldsymbol{\mu}_k, \boldsymbol{\sigma}_k) = \frac{1}{(2\pi)^{\frac{m}{2}} (\det \boldsymbol{\sigma}_k)^{\frac{1}{2}}} \exp\left[-\frac{1}{2} (\boldsymbol{x} - \boldsymbol{\mu}_k)^{\mathrm{T}} \boldsymbol{\sigma}_k (\boldsymbol{x} - \boldsymbol{\mu}_k) \right]$$

x : [x_1, x_2, x_3, … x_m]

μ_k : mean vector (1 × *m*) of *k*-th Gaussian

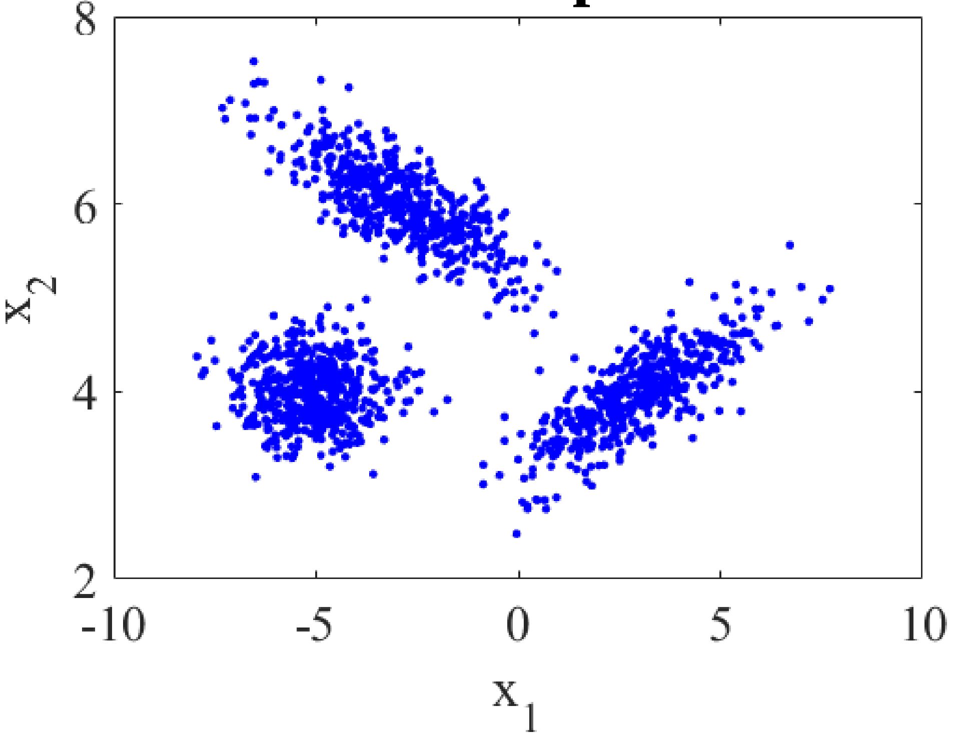
Σ_k : covariance matrix (*m* × *m*) of *k*-th Gaussian

π_k : mixing coefficient

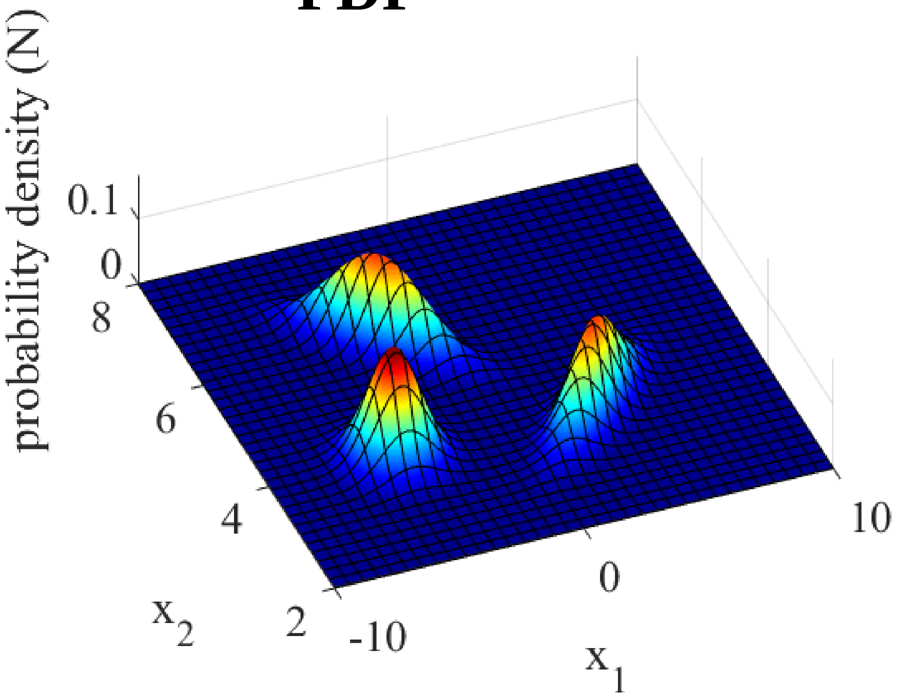
$$\sum_{k=1}^{n} \pi_k = 1$$

Gaussian mixture model: example

Scatter plot



PDF

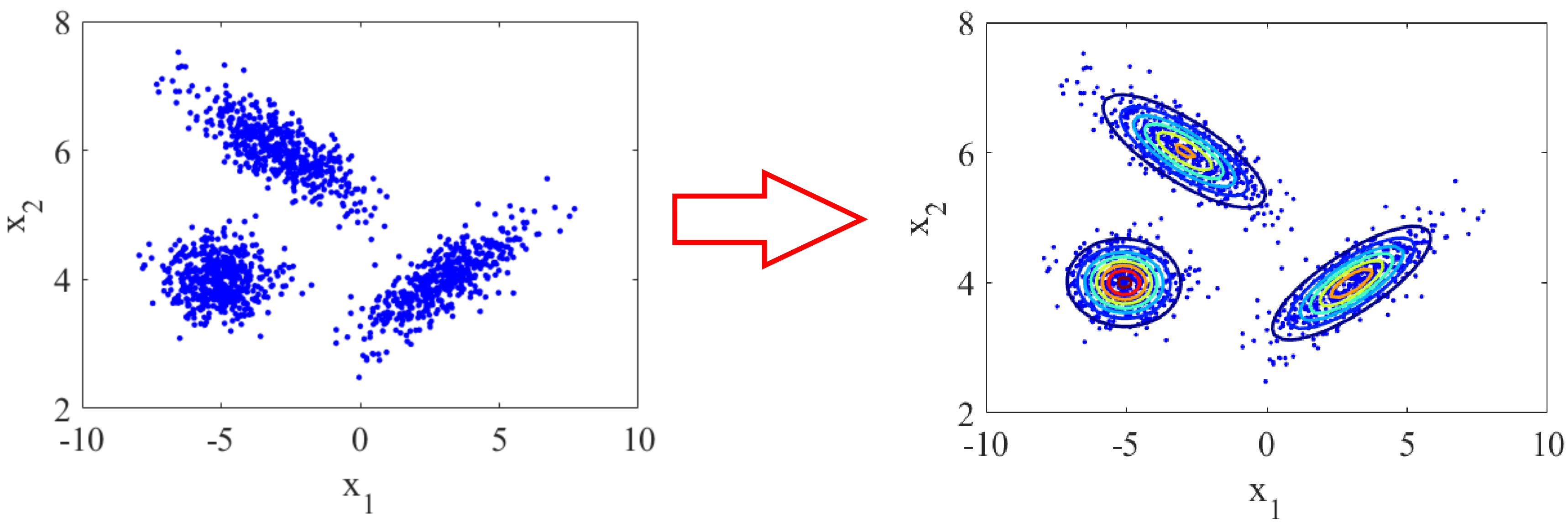


Estimation of the parameters for Gaussian mixture model

Given a dataset x : [x_1, x_2, x_3, … x_m], we estimate

μ_k : mean vector,
σ_k : covariance matrix, and
π_k : mixing coefficient

via maximum likelihood method.



Estimation of the parameters for Gaussian mixture model

When N data are divided into K clusters, a log-likelihood function is

$$\ln p(\boldsymbol{x}|\boldsymbol{\pi},\boldsymbol{\mu},\boldsymbol{\sigma}) = \ln \prod_{n=1}^{N}\left[\sum_{k=1}^{K}\boldsymbol{\pi}_k \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_k,\boldsymbol{\sigma}_k)\right]$$

$$= \sum_{n=1}^{N}\ln\left[\sum_{k=1}^{K}\boldsymbol{\pi}_k \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_k,\boldsymbol{\sigma}_k)\right]$$

We calculate a contribution function $\boldsymbol{\gamma(z_{nj})}$

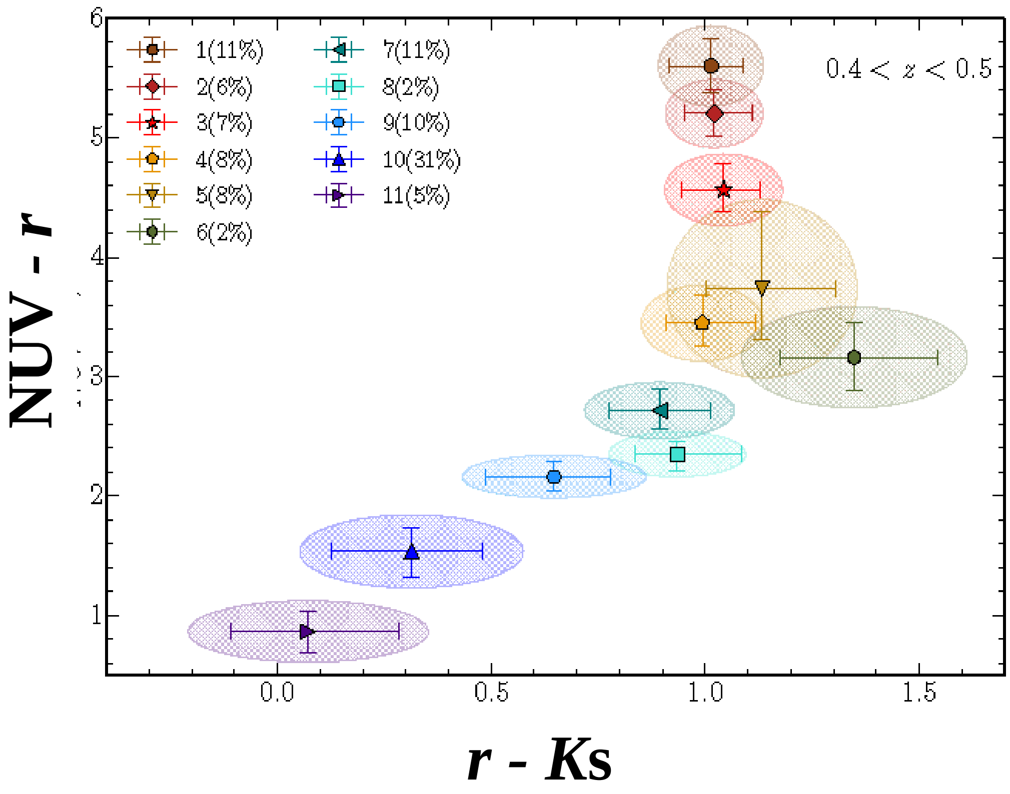
$$\gamma\left(z_{nj}\right) = \frac{\boldsymbol{\pi}_j \mathcal{N}\left(\boldsymbol{x}_n|\boldsymbol{\mu}_j,\boldsymbol{\sigma}_j\right)}{\sum_{k=1}^{K}\boldsymbol{\pi}_k \mathcal{N}(\boldsymbol{x}_n|\boldsymbol{\mu}_k,\boldsymbol{\sigma}_k)}$$

(z : latent variable, $0 \leq z \leq 1$)

2.3 Result

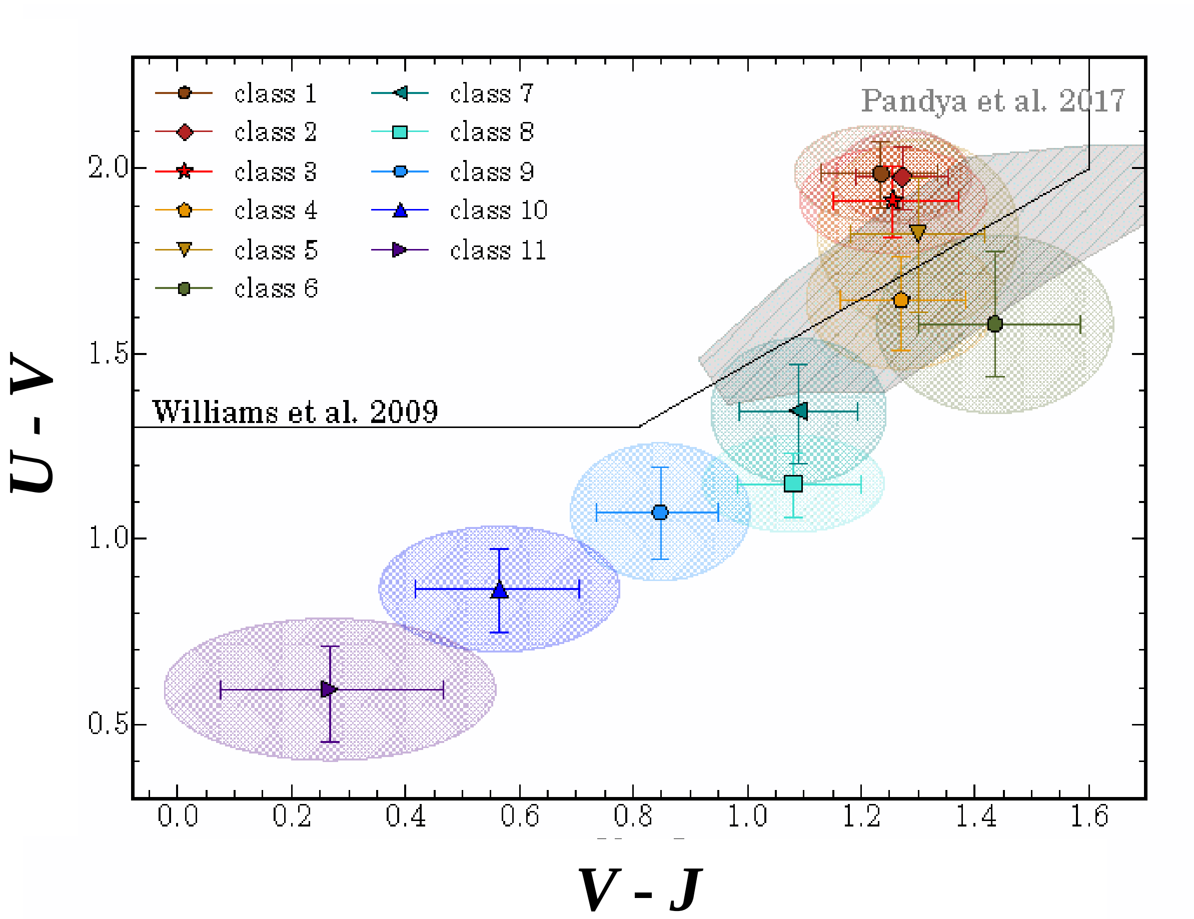
The FEM classification separates VIPERS galaxies into twelve classes.

***N.B.* Twelfth class is the AGN and not included in the further discussion.**



The *UVJ* diagram

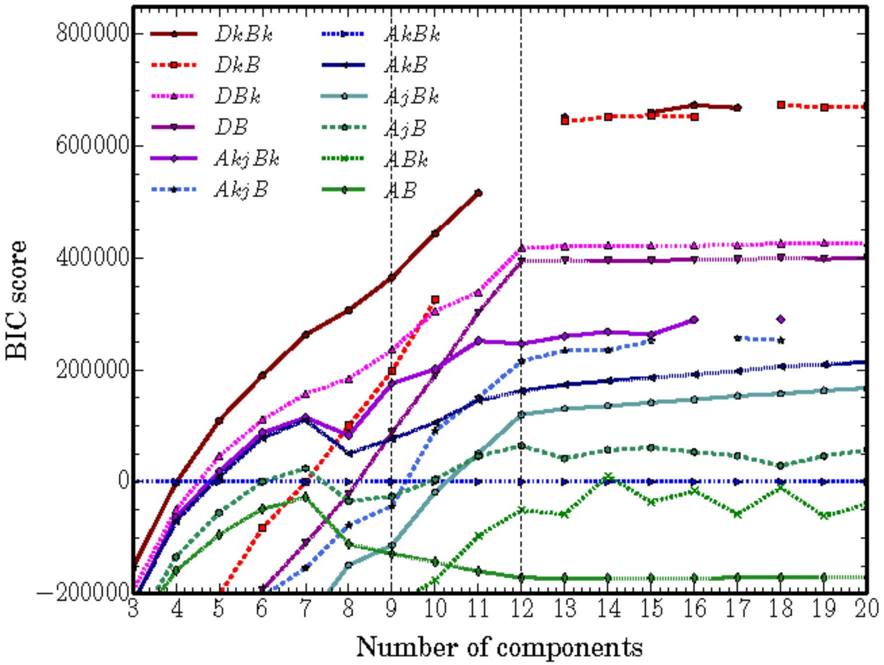
Williams et al. (2009) and subsequent authors proposed that the *UVJ* diagram provide a separation between passive and star-forming galaxies.



Selection of best model and optimal number of classes

The number of classes is knot known a priori: a major difficulty in unsupervised algorithm.

In this work, the best model and the range of possible classes are chosen based on three statistical model-selection criteria: AIC (Akaike 1974), BIC (Schwarz 1978), and ICL (Baudry 2012).



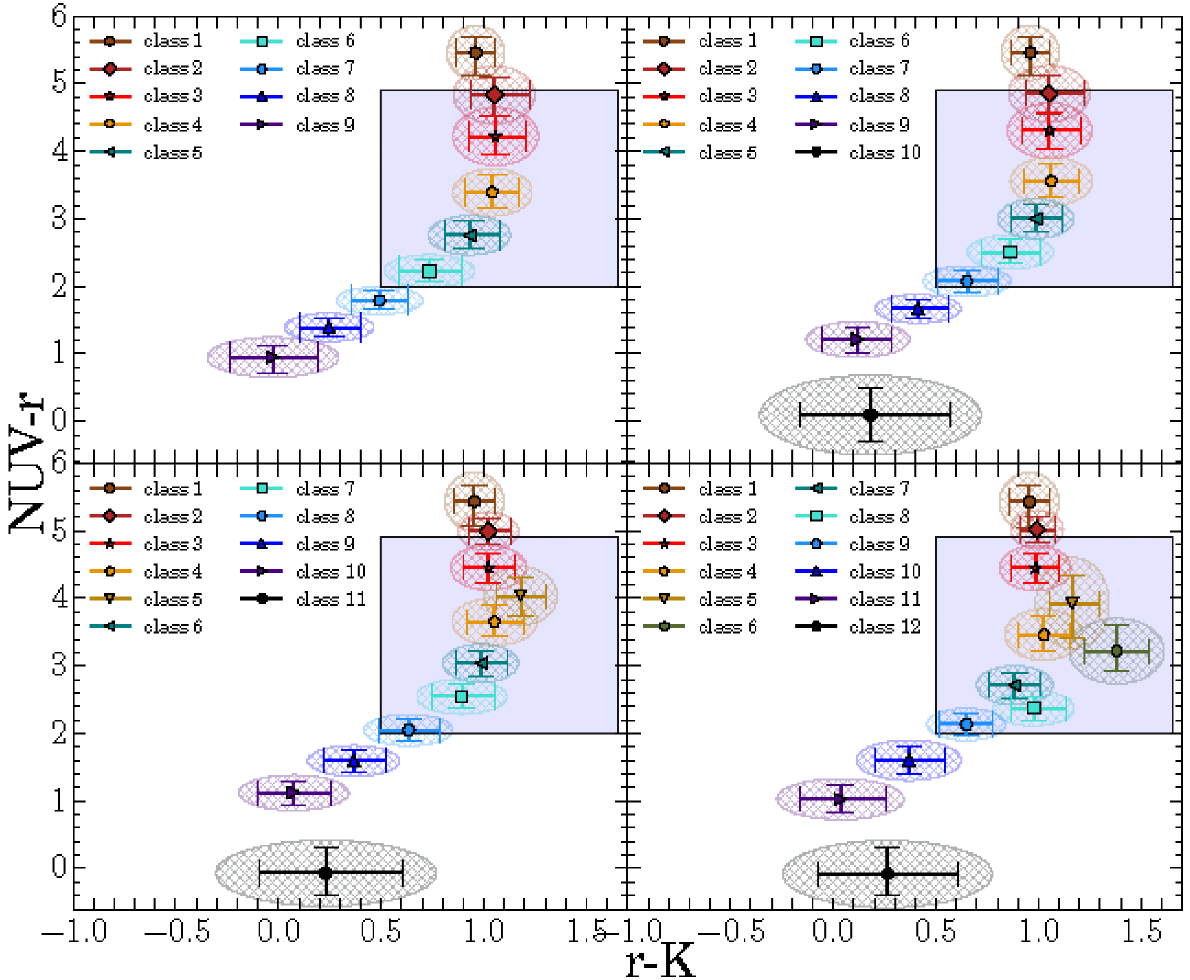
Siudek et al. (2018)

Selection of best model and optimal number of classes

The result can be intuitively verified on a NUV*rK* color-color diagram.

9 classes

10 classes



11 classes

12 classes

Siudek et al. (2018)

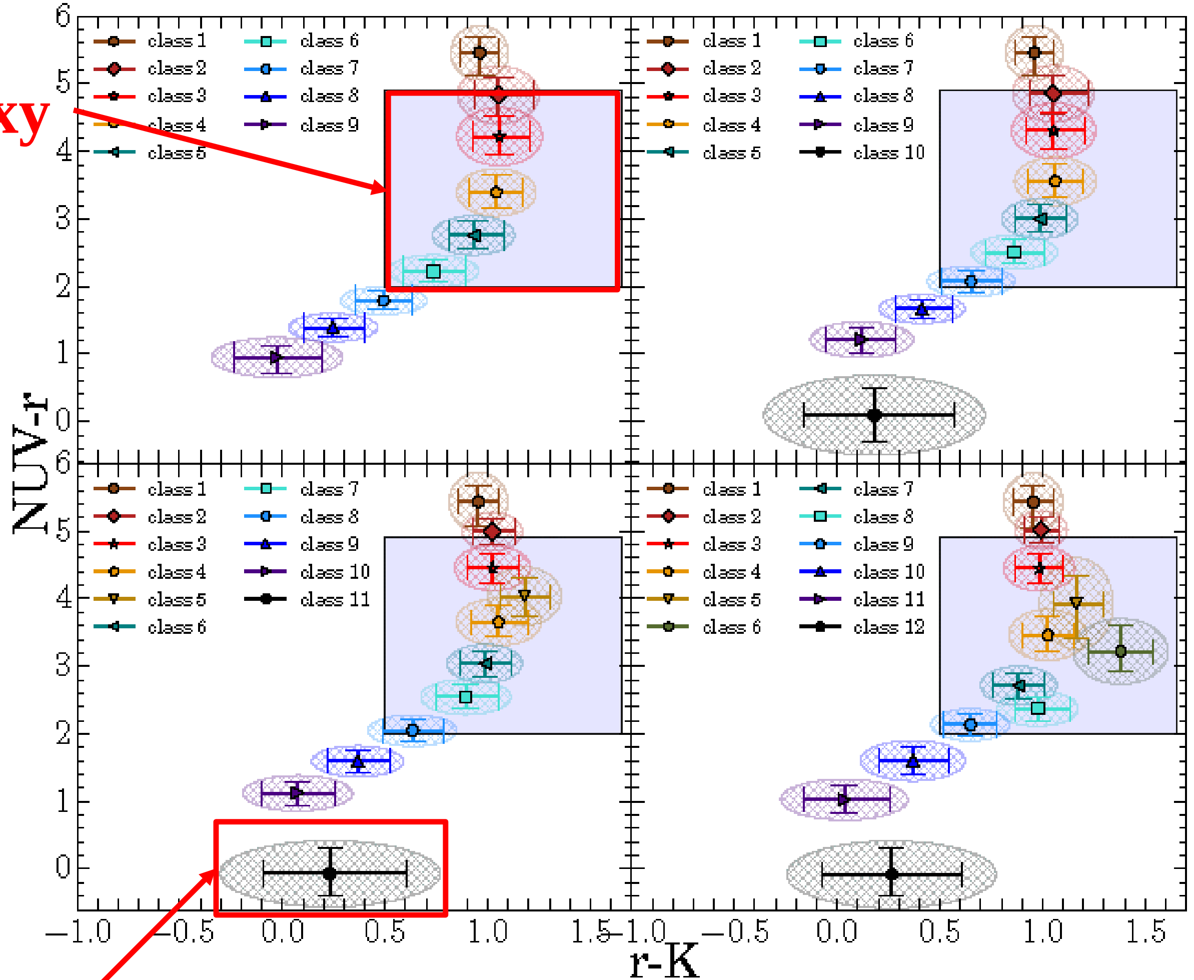
Selection of best model and optimal number of classes

The result can be intuitively verified on a NUV*rK* color-color diagram.

Dusty galaxy

9 classes

10 classes



11 classes

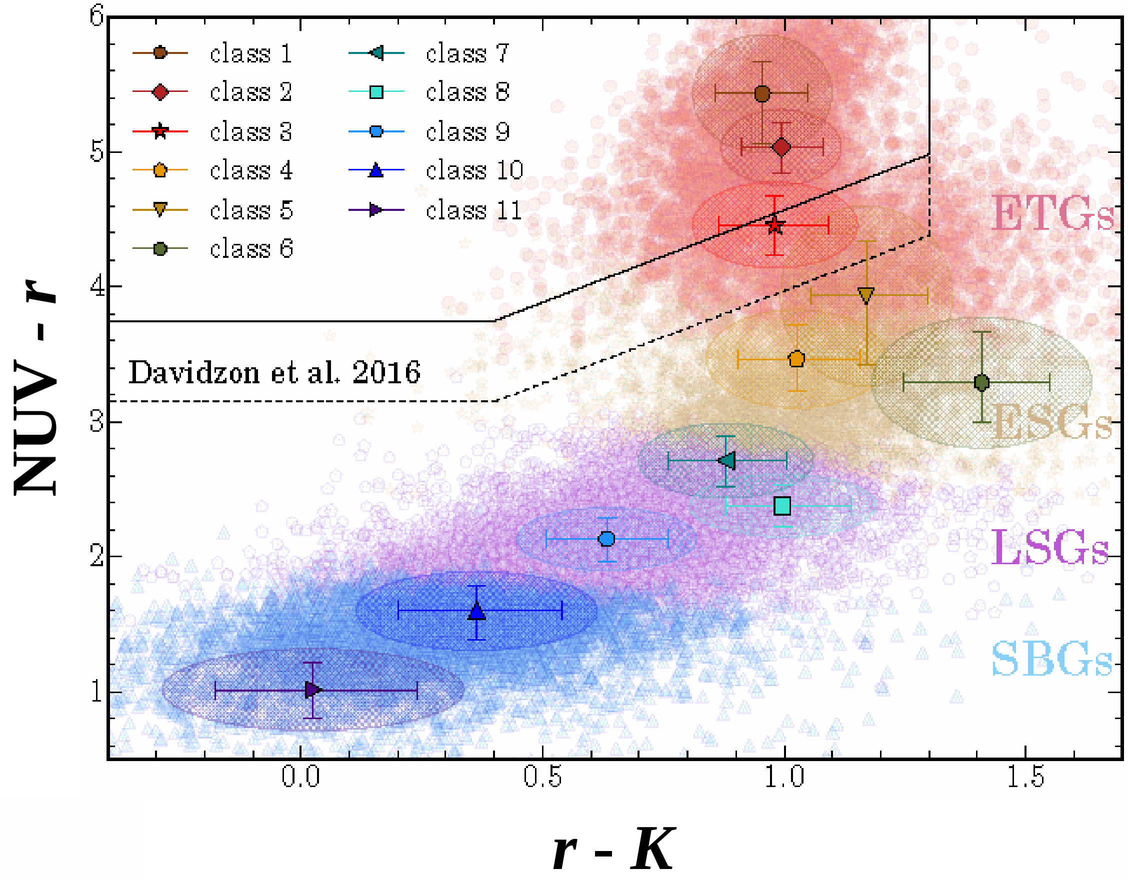
12 classes

High-*z* AGNs

Siudek et al. (2018)

The NUV*rK* diagram

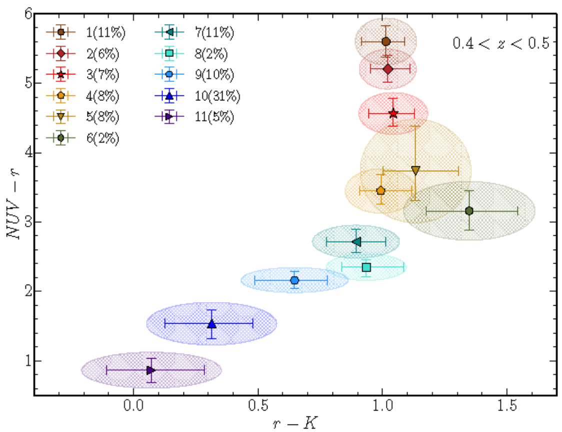
The NUV*rK* diagram (Arnouts et al. 2013) is similar to the *UVJ* plane but allows for a better separation between passive and star-forming galaxies.



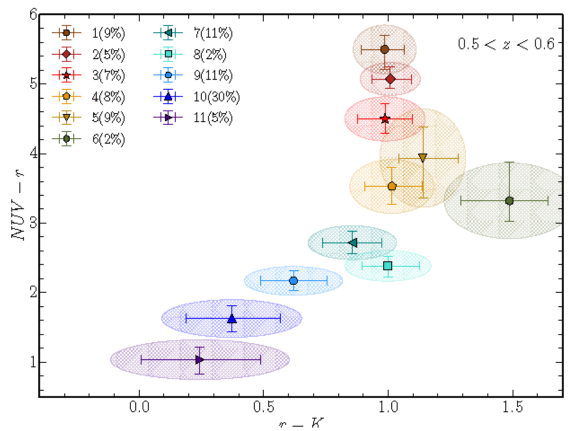
The redshift (in)dependence at $0.4 < z < 1.0$

The color evolution of galaxies is clearly visible, but the classification is almost unchanged with redshifts.

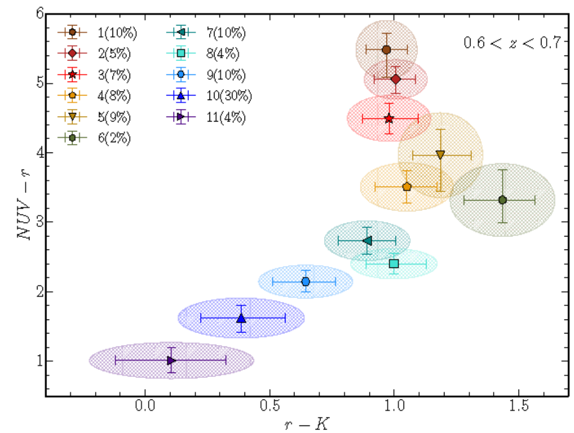
$0.4 < z < 0.5$



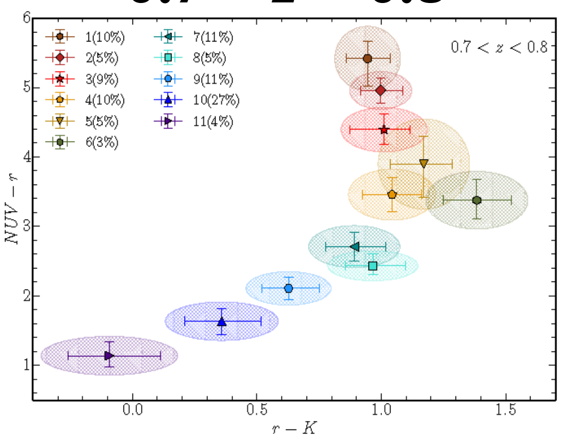
$0.5 < z < 0.6$



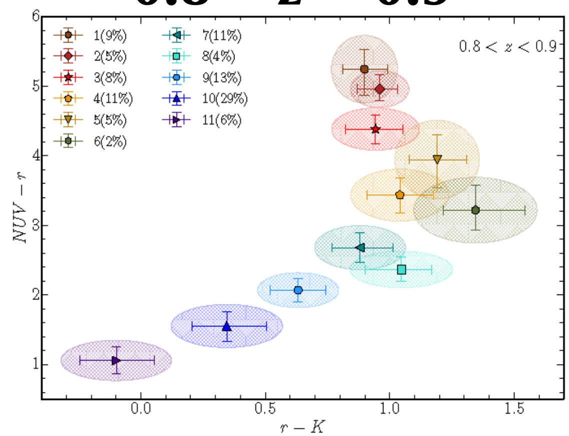
$0.6 < z < 0.7$



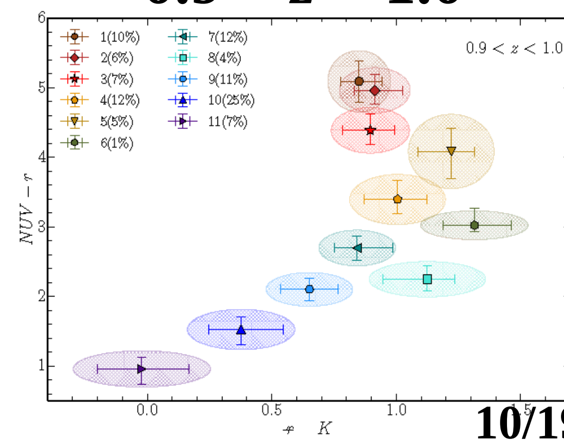
$0.7 < z < 0.8$



$0.8 < z < 0.9$

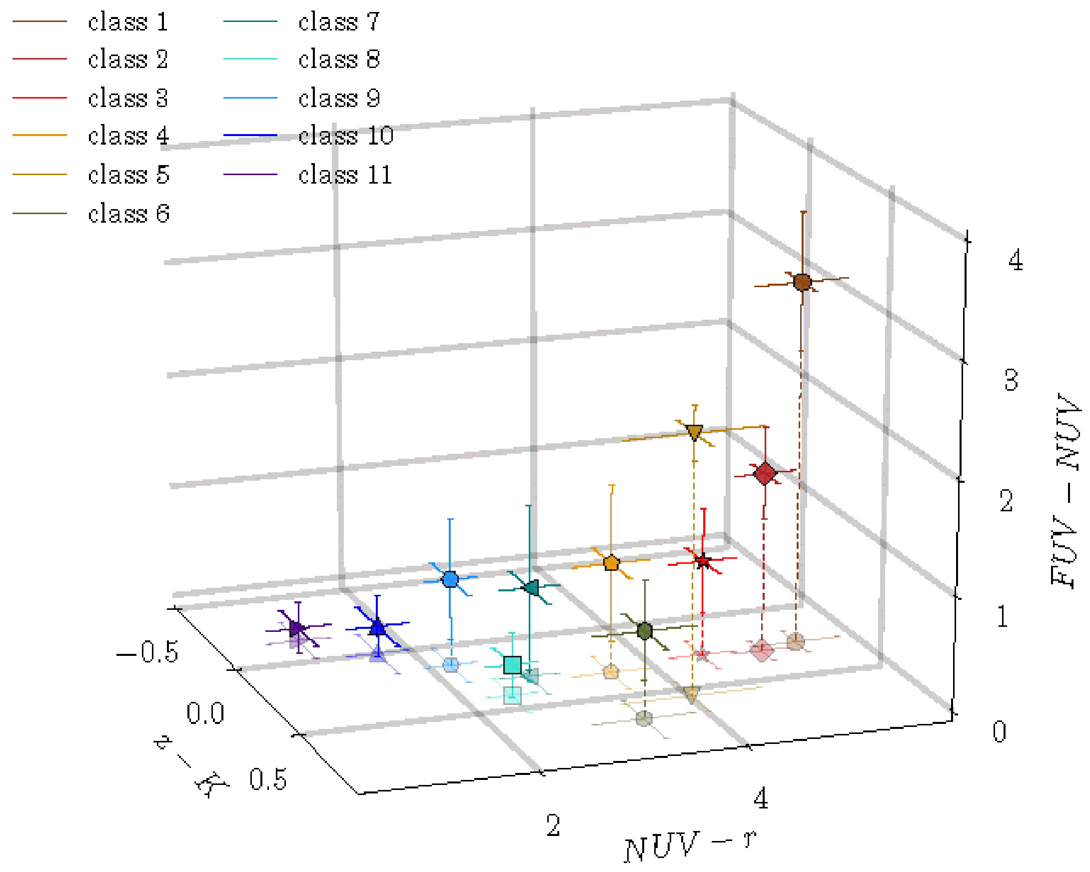


$0.9 < z < 1.0$



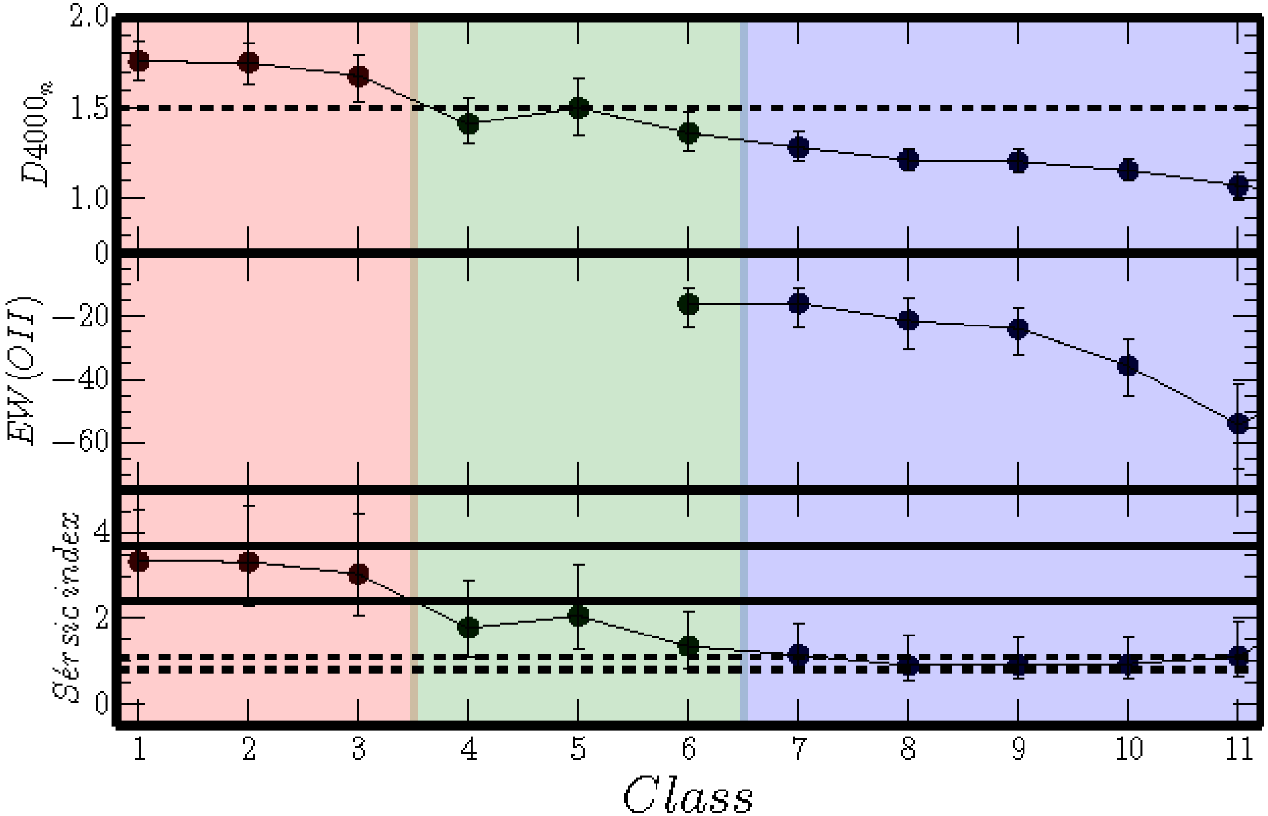
The NUV*rK* 3D diagram

The red galaxies are divided into three subclasses, separated by the FUV - NUV color.



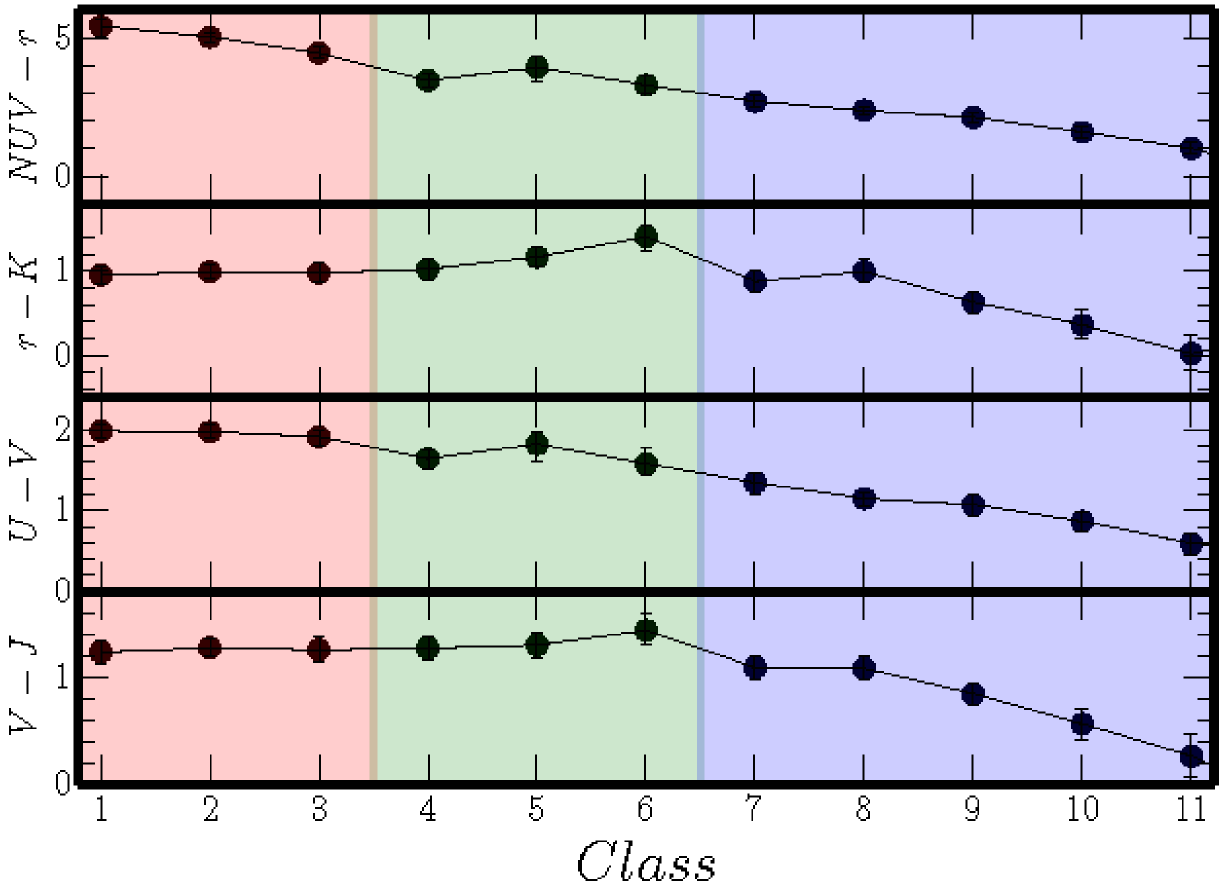
2.4 Global properties of classified galaxies

Spectral and morphological features



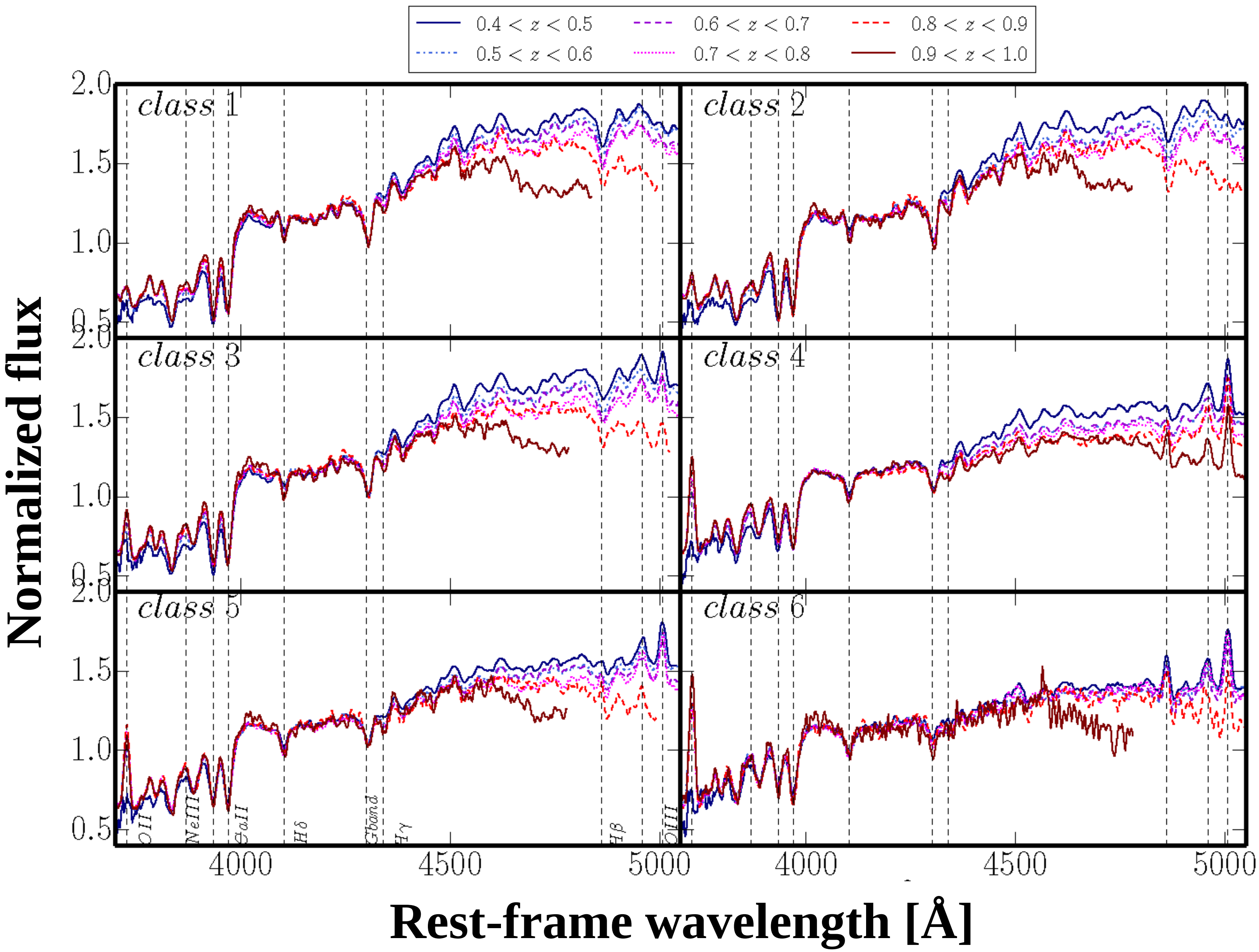
D4000: indicator of galaxy age (relatively new SFH)
EW[OII]: emission from SF region
Sérsic index: shape parameter indicating early (~ 4) or late (~ 1) type

Colors

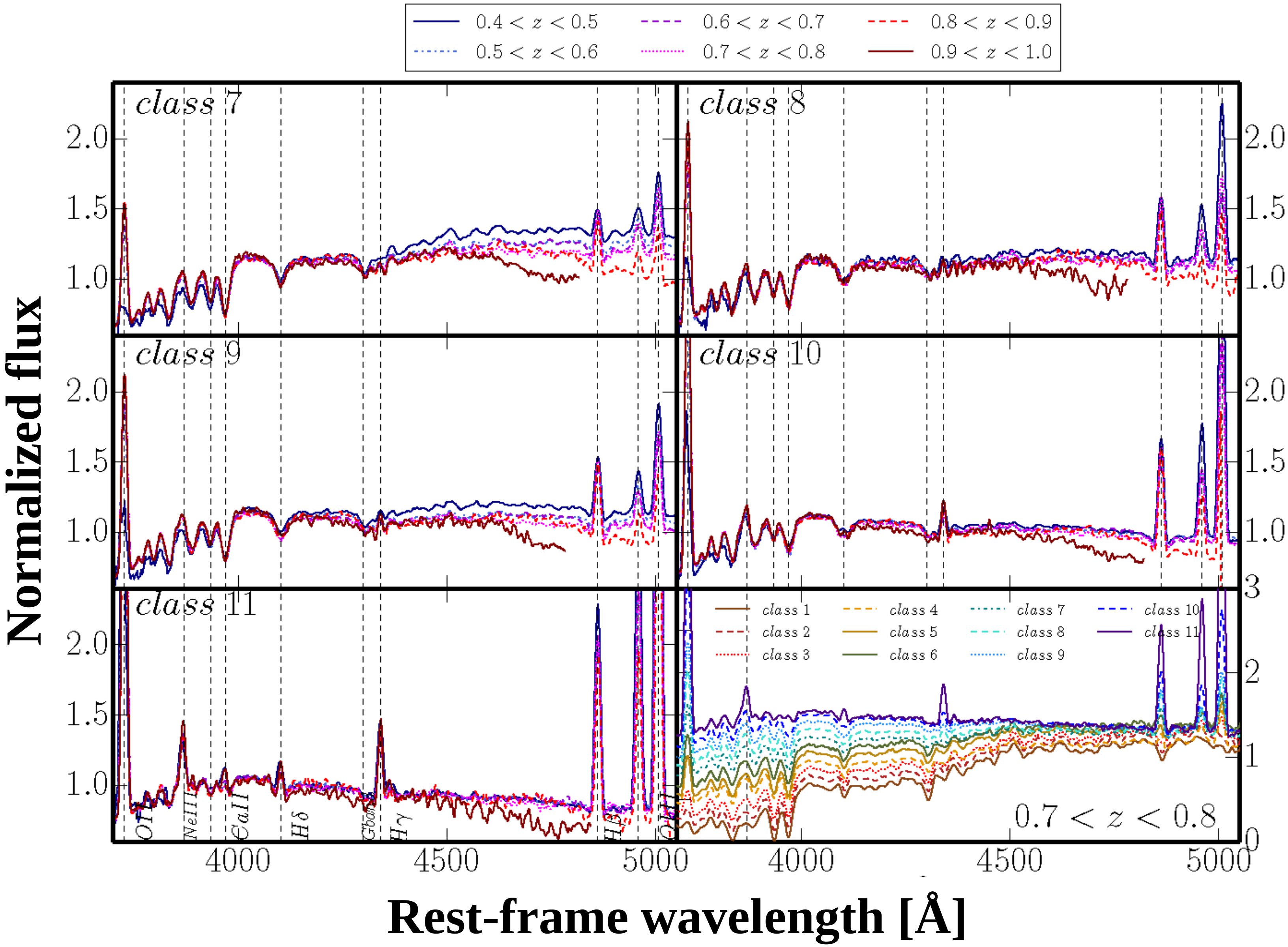


A turnover exists in some colors.

Stacked spectra of class 1-6



Stacked spectra of class 7-11



Summary of spectral features in classified galaxies

Class 1-3 (red)

Strong absorption features, D4000 break, red continuum

Class 4-6 (green)

Balmer absorption, gradually stronger emission lines

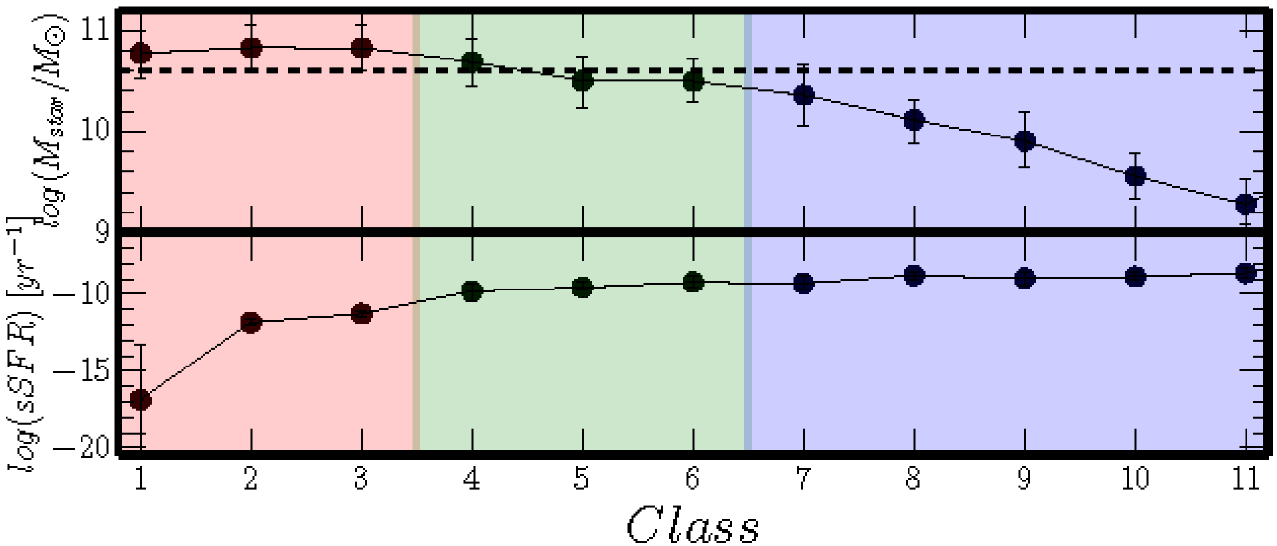
Class 7-11 (blue)

Strong emission lines, blue continuum

Redshift dependence

In each class, higher-*z* galaxies have gradually bluer continuum, significant at longer wavelengths

Stellar mass and specific SFR (sSFR)

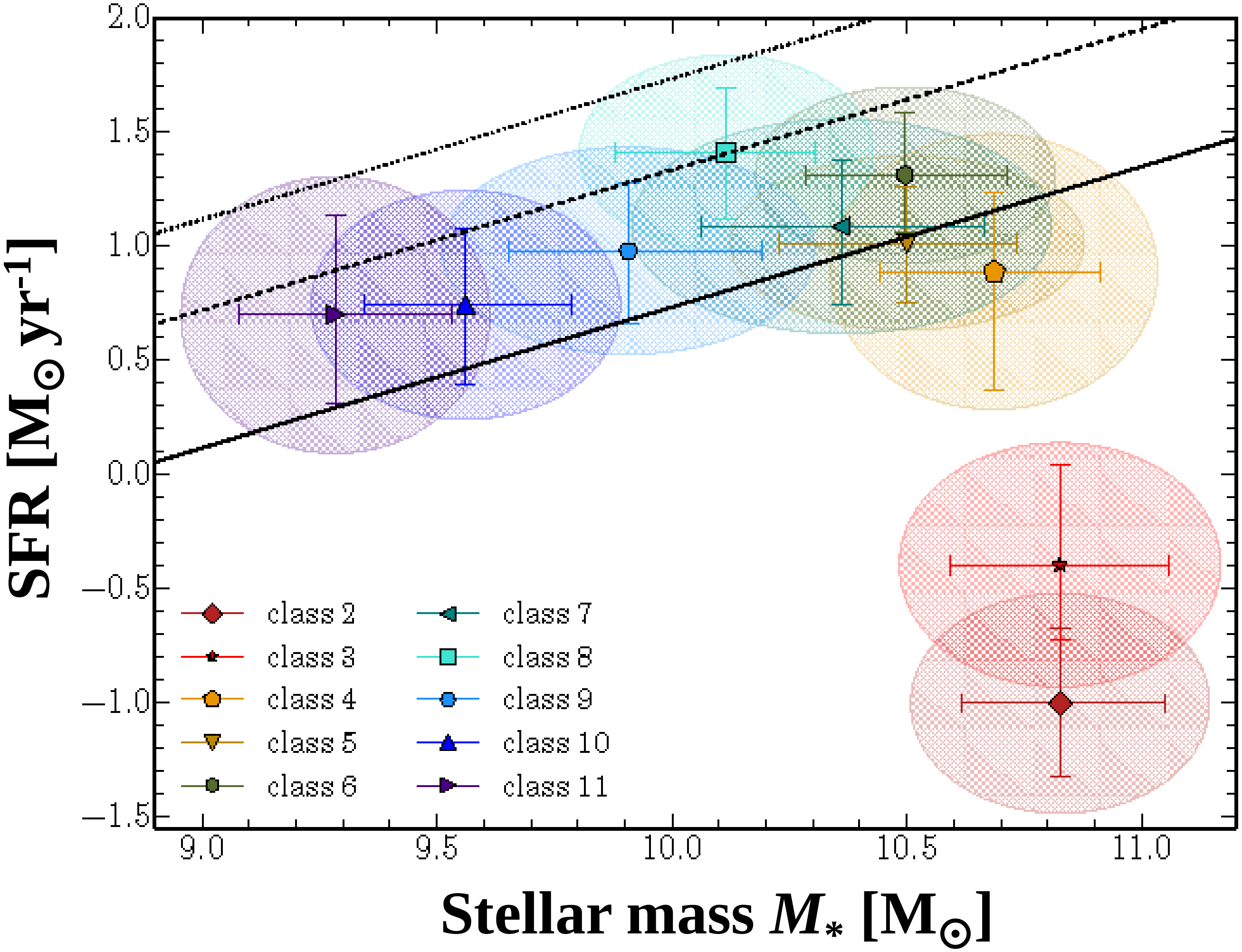


sSFR: SFR per stellar mass [yr^{-1}]

$$\mathrm{sSFR} = \frac{\mathrm{SFR}}{M_*}$$

Indicator of the SF activity normalized with the galaxy size.

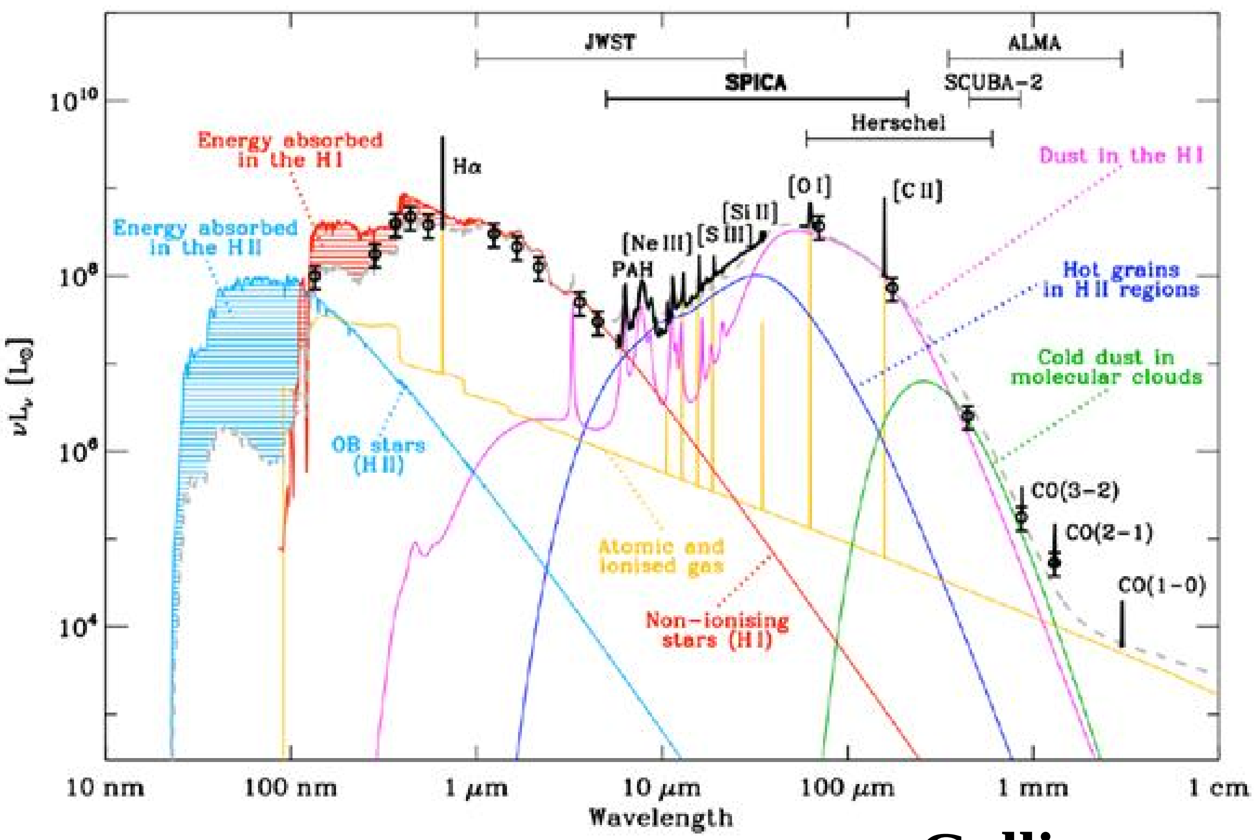
Main sequence of star-forming galaxies



The main sequence is naturally reproduced from this method. **Note that this is free from color selection bias, which introduced a significant confusion in this diagram.**

3. Application to SPICA Galaxies

3.1 Evolutionary Sequence at Infrared



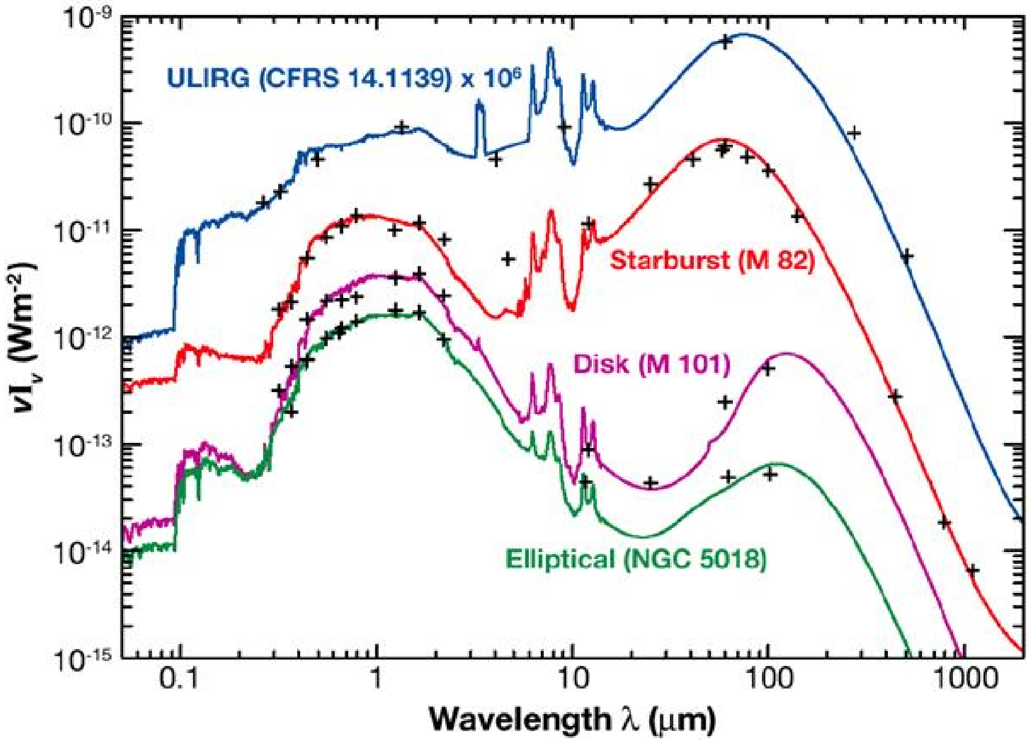
Galliano et al. (2008)

The IR regime provides large amount of additional information to the SEDs of galaxies.

3.1 Evolutionary Sequence at Infrared

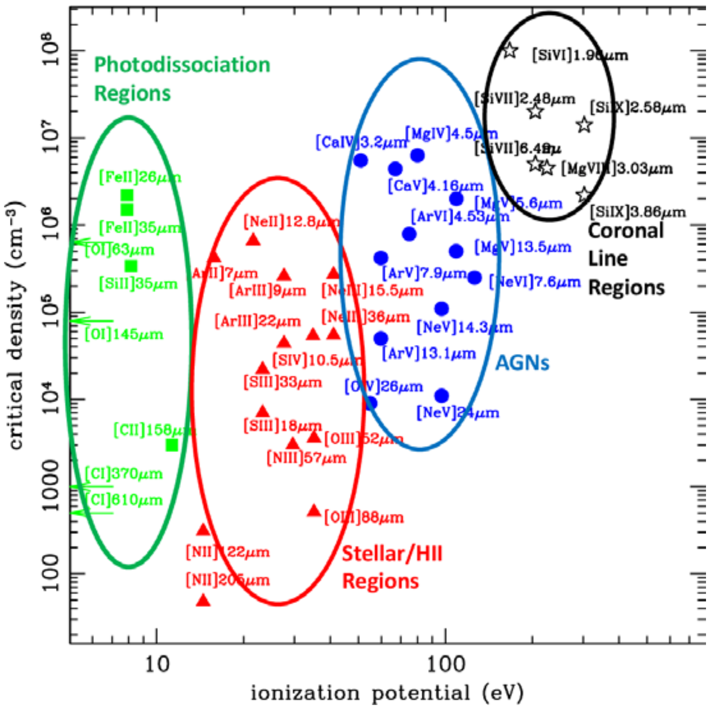
We can perform a similar analysis for the galaxy population detected by SPICA, including IR emission lines/bands!

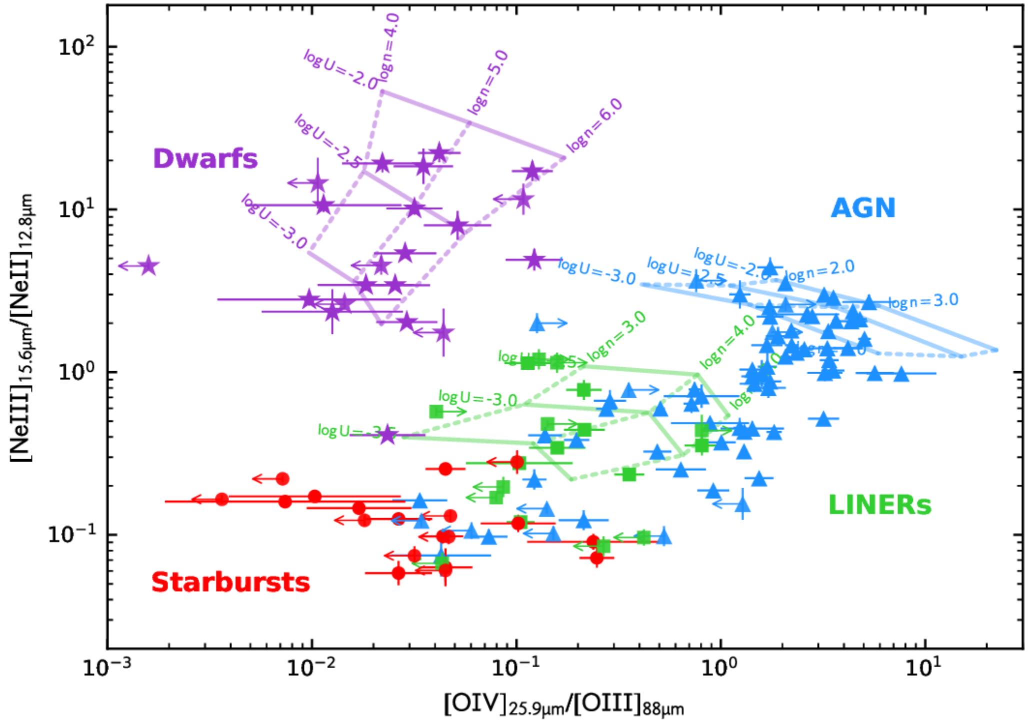
⇒ It will reveal the obscured side of galaxy evolution as a sequence of various physical properties.



Galliano et al. (2004)

3.2 Evolutionary Census of the ISM in Galaxies

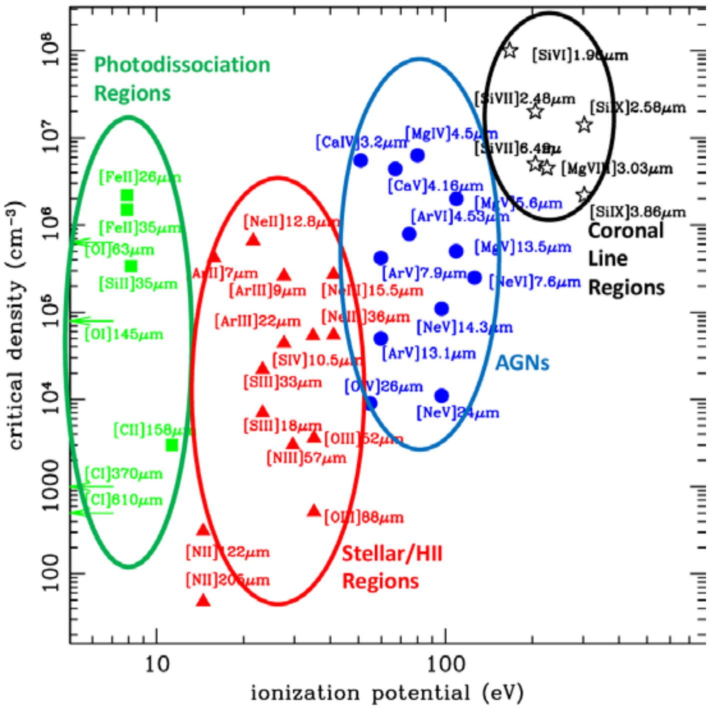


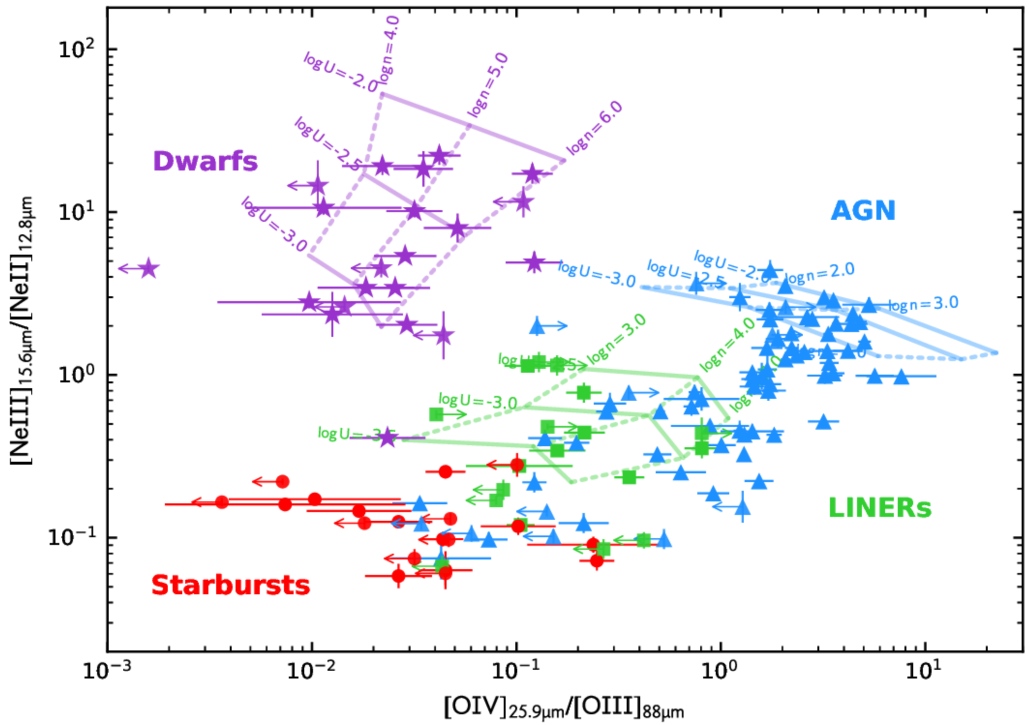


Spinoglio et al. (2017)

Various IR diagnostics have been proposed. We can do this business all at once by the unsupervised machine learning!

3.2 Evolutionary Census of the ISM in Galaxies





Spinoglio et al. (2017)

Various IR diagnostics have been proposed. We can do this business all at once by the unsupervised machine learning!

We will have a “submanifold” in the high-dimensional feature space, which shows the sequence of the ISM evolution.

4. Conclusions

It is time to reconsider what the galaxy evolution is. We propose an objective method to discover the sequence of galaxy evolutionary properties “galaxy manifold” with unsupervised machine learning. This can also be applied to explore the evolutionary aspects of the ISM in galaxies.

1. **We performed an unsupervised machine learning classification directly to the twelve-dimensional luminosity-redshift space of VIPERS galaxies.**

2. **Our classification yielded twelve galaxy classes (+ one AGN class). Class 1-3 corresponds to red passive galaxies, 4-6 to “green” galaxies, and 7-11 to blue star-forming galaxies, respectively. This forms a submanifold in the 13-dim feature space: galaxy manifold.**

4. Conclusions

3. **Classes 1-3 show stronger absorption lines and D4000 and Sérsic index *n* ~ 4 (spheroid-like), indicating old stellar population. Classes 7-11 show strong [OII] emission and *n* ~ 1 (disk-like), consistent with active SF.**

4. **The star-forming galaxy main sequence is clearly reproduced by classes 7-11, with class 1-3 as quiescent non-star forming galaxies. Note that, in contrast to previous studies, our result is not affected by the complicated color selection and robust.**

5. **This will be very useful to explore the evolutionary track of galaxies at IR, since the IR spectral features provide large amount of additional information.**

4. Conclusions

6. Various IR line/band diagnostics have ever been proposed. The proposed method can do this type of analysis at once objectively. This will open a new window to the systematic exploration of the physical properties in the galaxy ISM.