13th Serbian Conference on Spectral Line Shapes in Astrophysics
Belgrade, Serbia, August 23-27, 2021

REPRESENTATION AND CHARACTERIZATION OF BROAD-LINE AGN SPECTRA BASED ON MANIFOLD LEARNING

Isidora Jankov,
Dragana Ilić
Andjelka Kovačević

MATΦ
University of Belgrade
Faculty of Mathematics

E-mail: isidora_jankov@matf.bg.ac.rs

TALK OUTLINE

1 INTRO TO MANIFOLD LEARNING

2 ALGORITHM: ROBUST LLE

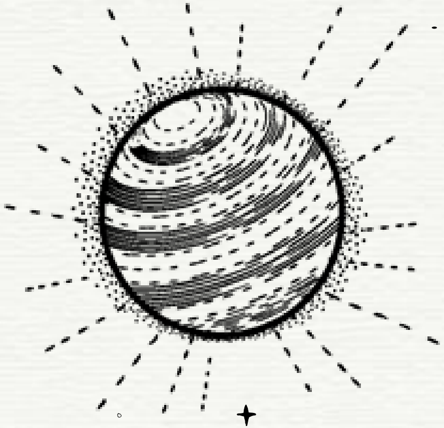
3 APLICATION TO BROAD-LINE AGN DATASETS

4 RESULTS & CONCLUSIONS

1

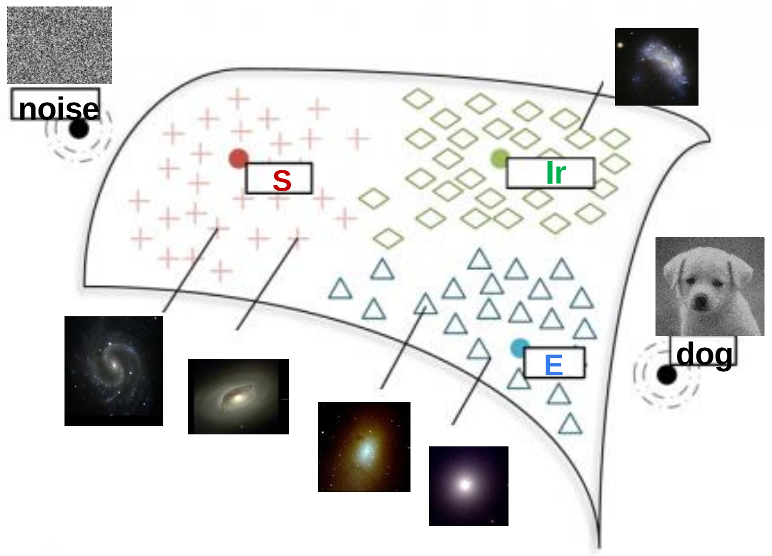
INTRO TO MANIFOLD LEARNING

Overview of
basic concepts



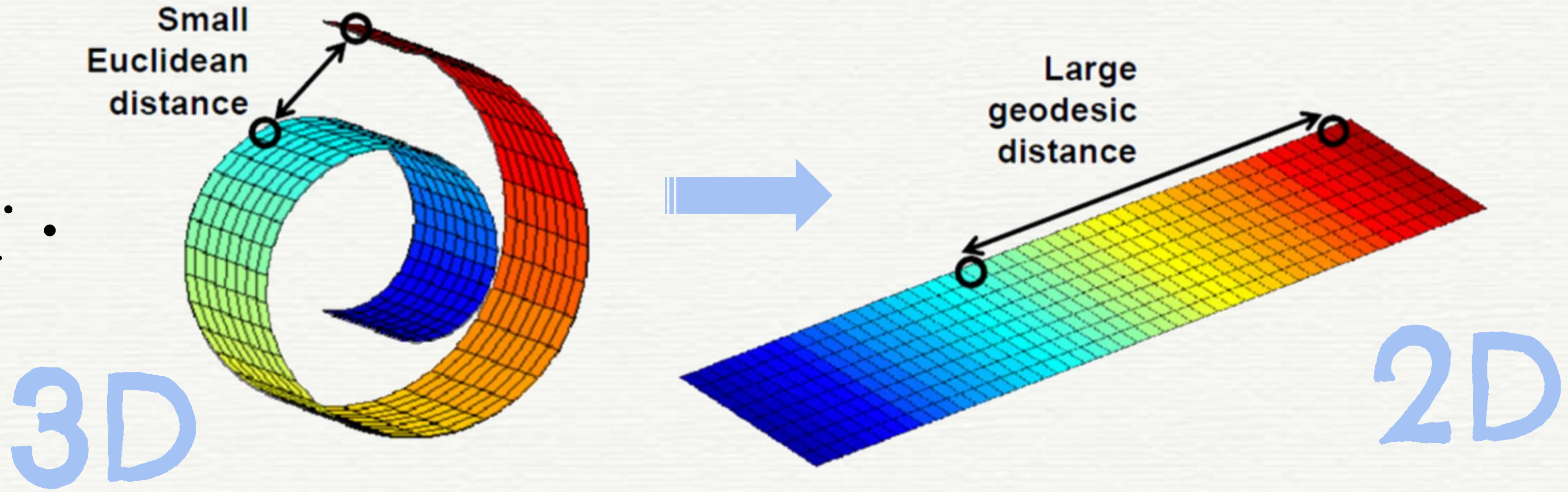
What is a manifold?

- **Definition**: Manifold is a topological space that locally resembles Euclidian space, but may vary widely in global properties.
- **Example**: surface of planet Earth!
- **Manifold hypothesis**: Real world high dimensional data lies near low-dimensional manifolds embedded within the high-dimensional space (e.g. Fefferman et al., 2016; Carlsson, 2009).



Manifold learning

A class of unsupervised machine learning techniques which can learn the geometry of the manifold embedded in the high-dimensional space and project it to a lower dimensional space, while preserving original relationships between the points - **nonlinear dimensionality reduction.**



Useful for...

Data visualization

Reduction of high-dimensional data to 2D/3D allows its visual inspection.

Unsupervised classification

..of galaxies, stars & AGN from their spectra, light curves, photometric and spectroscopic parameters.

Pre-processing

Dimensionality reduction before applying supervised ML to save computational resources and lower the impact of curse of dimensionality.

2

ALGORITHM:
ROBUST LLE

Locally Linear Embedding and its robust variant

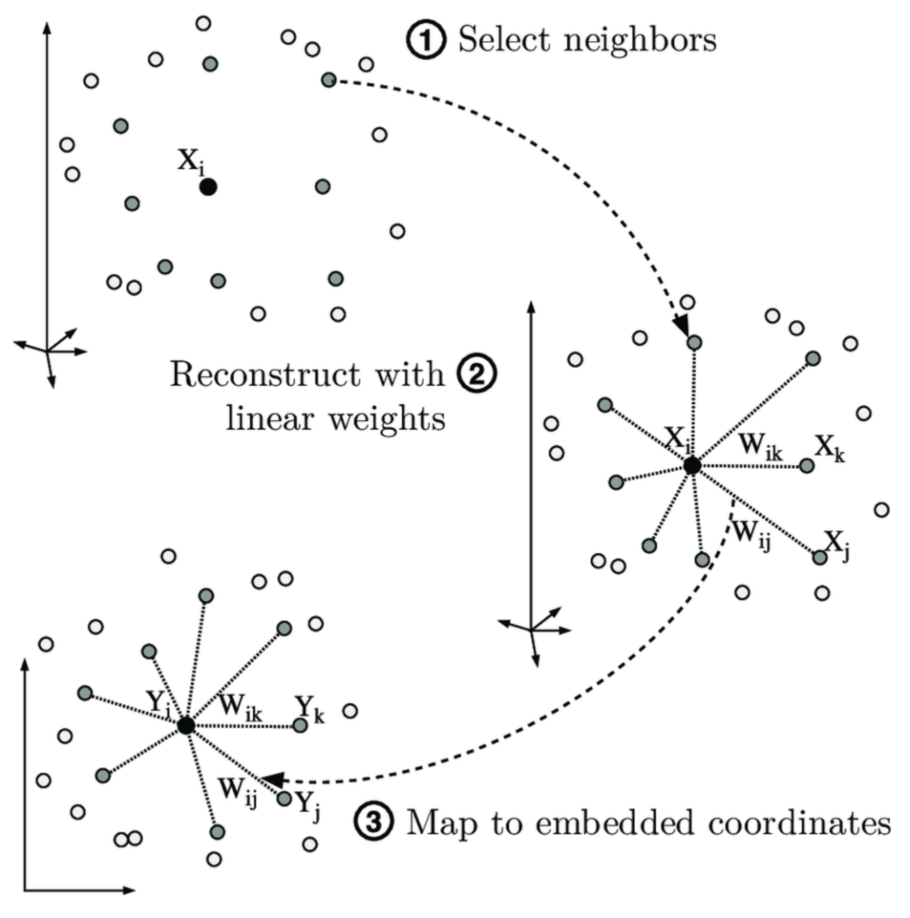
Locally Linear Embedding
(LLE, Roweis & Saul 2000)

Assumtions:

- Data lies on a manifold
- Manifold is a union of patches, each having a locally linear structure

Algorithm:

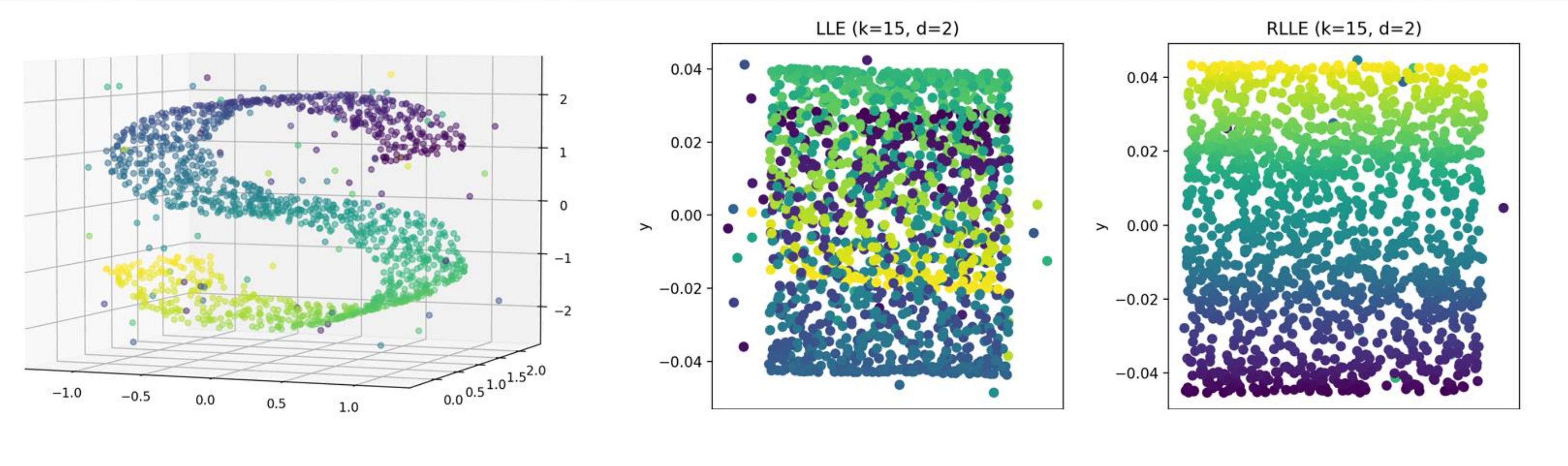
- Characterizes the local geometry of these patches by linear coefficients that reconstruct each data point from its neighbors.
- Finds the **projection** in the lower dimensional space where the coefficients are preserved for every point, i.e. **local geometry is preserved**.
- **Free parameters**: output dim. (**d**) & num. of nearest neighbors (**k**)



Source: https://cs.nyu.edu/~roweis/lle/algorithm.html

Robust LLE

A variant where the outliers are removed using methods of robust statistics (Chang & Yeung, 2006).

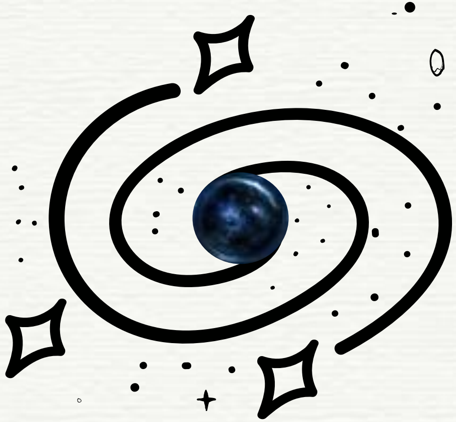


We are developing a python code for robust outlier detection based on Robust LLE and plan to make it open-source.

3

APPLICATION TO BROAD-LINE AGN DATASETS

Motivation & Data Analysis



Motivation

- Prepare for the **Big Data Era** - utilize novel statistical tools for data analysis in astronomical context.
- Explore **multidimensional parameter space** of AGN.
- Search for **key correlations** - a toolbox for understanding AGN physics.
- Expand on large body of work concerning linear dimensionality reduction of AGN data and physical interpretation of obtained projections (e.g., Boroson & Green 1992, Sulentic+2000, Marziani+2001, Marziani & Sulentic 2014, Shen & Ho 2014), but now using **non-linear methods.**
- Compare projections obtained using only **measured spectral parameters** and projections obtained from **raw spectra.**

1 PROCESS, 2 DATASETS

1 — Data selection

1. AGN optical spectral parameters (from Liu+2019 broad-line AGN SDSS DR7 catalog)
2. AGN optical spectra (matching SDSS spectra - only best quality)

2 — Outlier removal

Outliers removed using our python code based on RobustLLE algorithm (Chang & Yeung, 2006)

3 — Selecting free parameters

1. Number of nearest neighbors (k)
2. Output dimension (d)

As described in Jankov+2020.

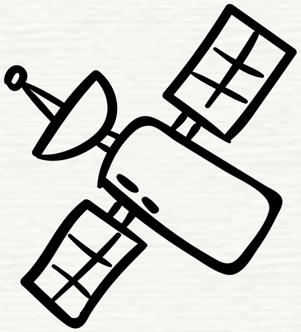
4 — Applying LLE to data

Obtained **3D projections** of two datasets:

→ AGN optical spectral parameters
→ AGN optical spectra

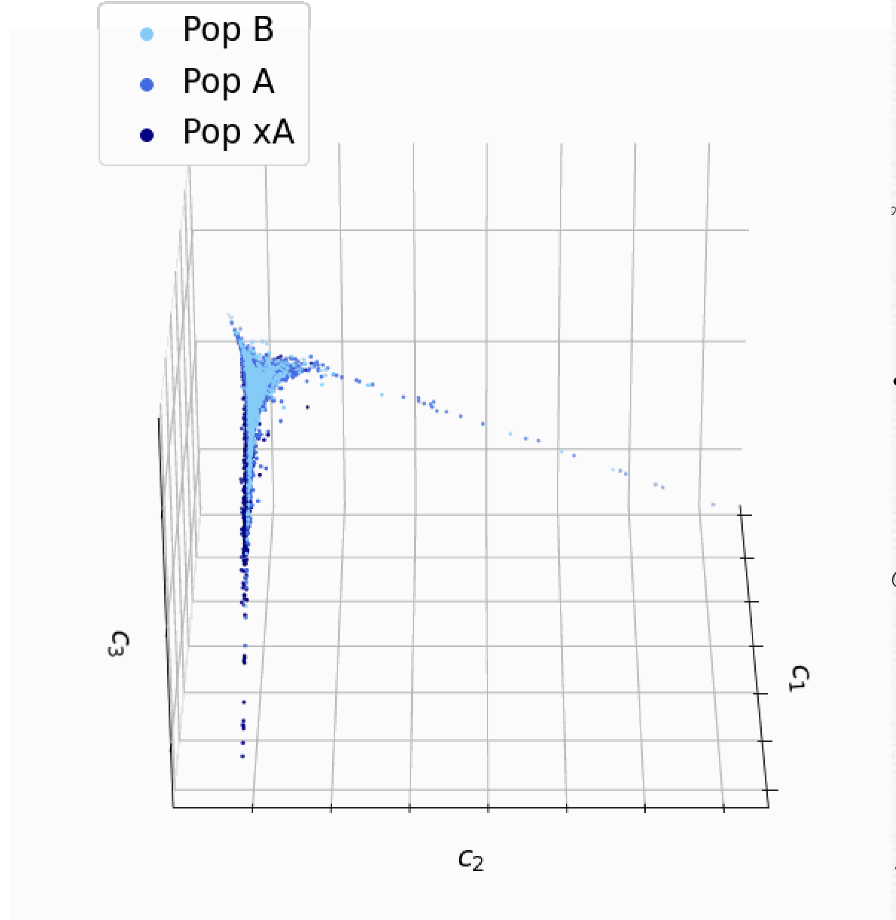
4

RESULTS & CONCLUSIONS

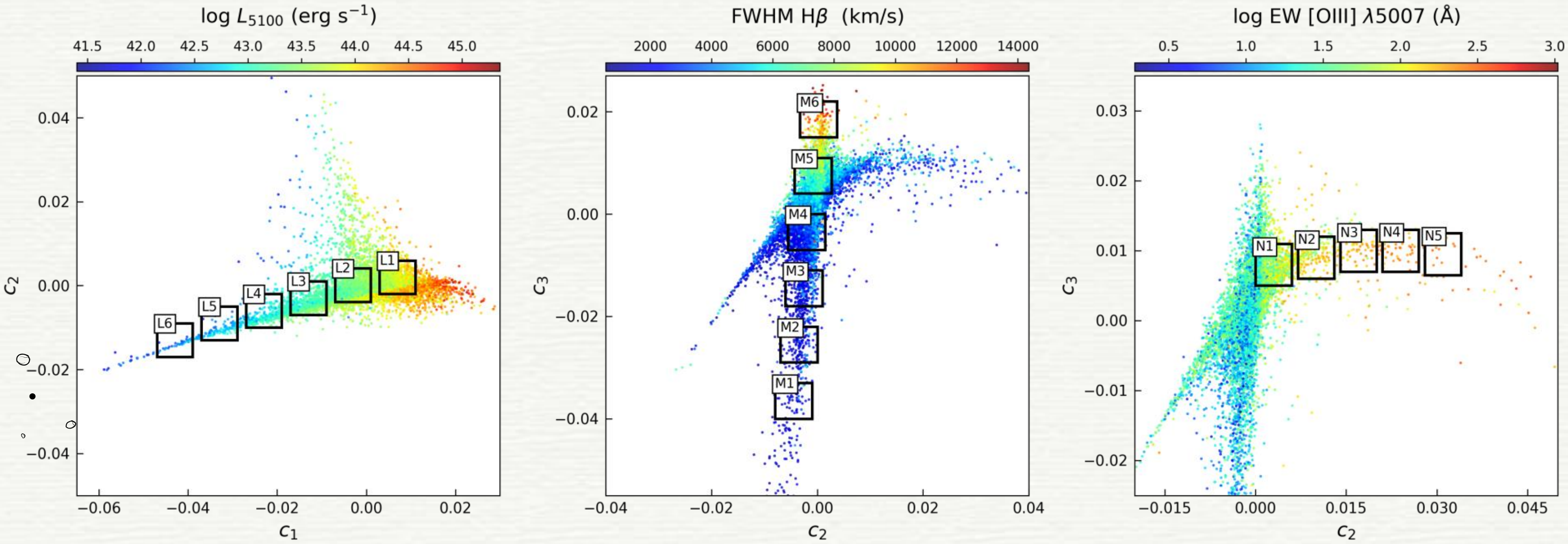


Results:
Parameter LLE

- Robust LLE applied to a sample of **7834 objects**
- **Parameters:**
  - EW, FWHM, L of broad & narrow emission lines
  - Continuum luminosity (L_{5100})
  - EW ratio of FeII and Hβ
- Dimensionality reduced from **19D to 3D**
- **Quasar populations** (Sulentic+2000, Marziani & Sulentic 2014) clearly distinguished in the projection



Parameter LLE: Quasar branches



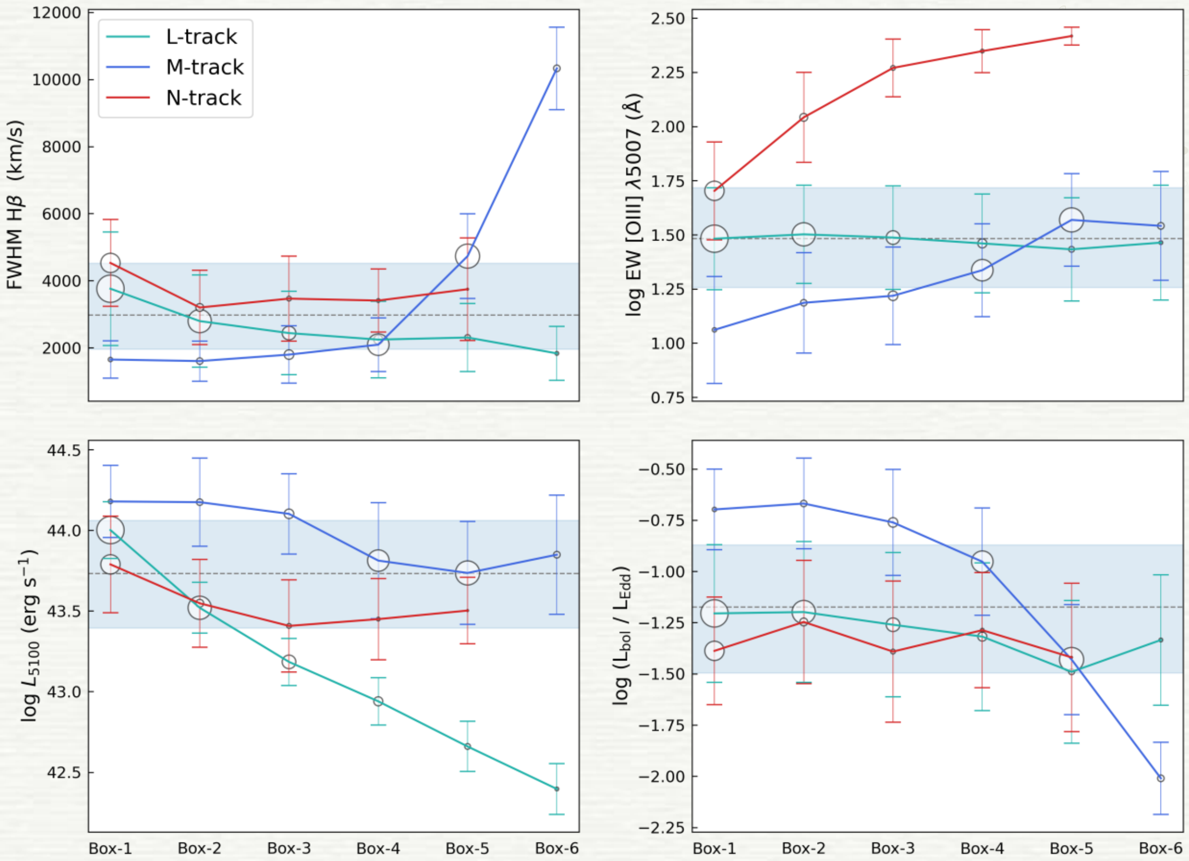
Luminosity (L) branch

Main sequence (M) branch

Narrow line (N) branch

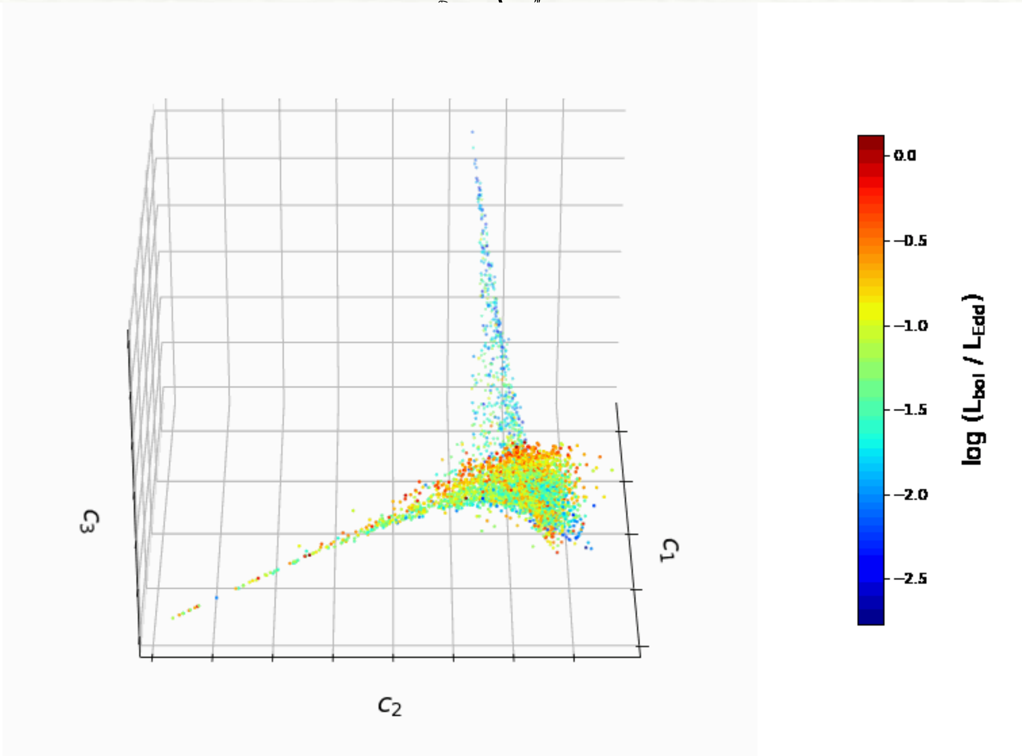
Parameter LLE:
Quasar branches

- Summary statsitics of the main trends in three quasar branches (tracks)
- M-branch: satisfies the main sequence correlations
- N-branch: distinct from the main sequence
- Tip of the N-branch: **extreme [OIII]λ5007 objects** (e.g., Ludwig+2009)
- L-branch: luminosity+redshift correlations

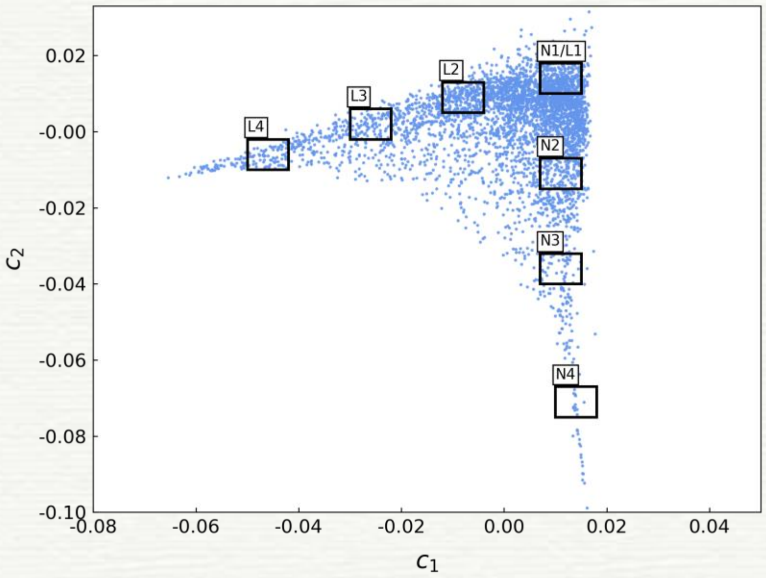


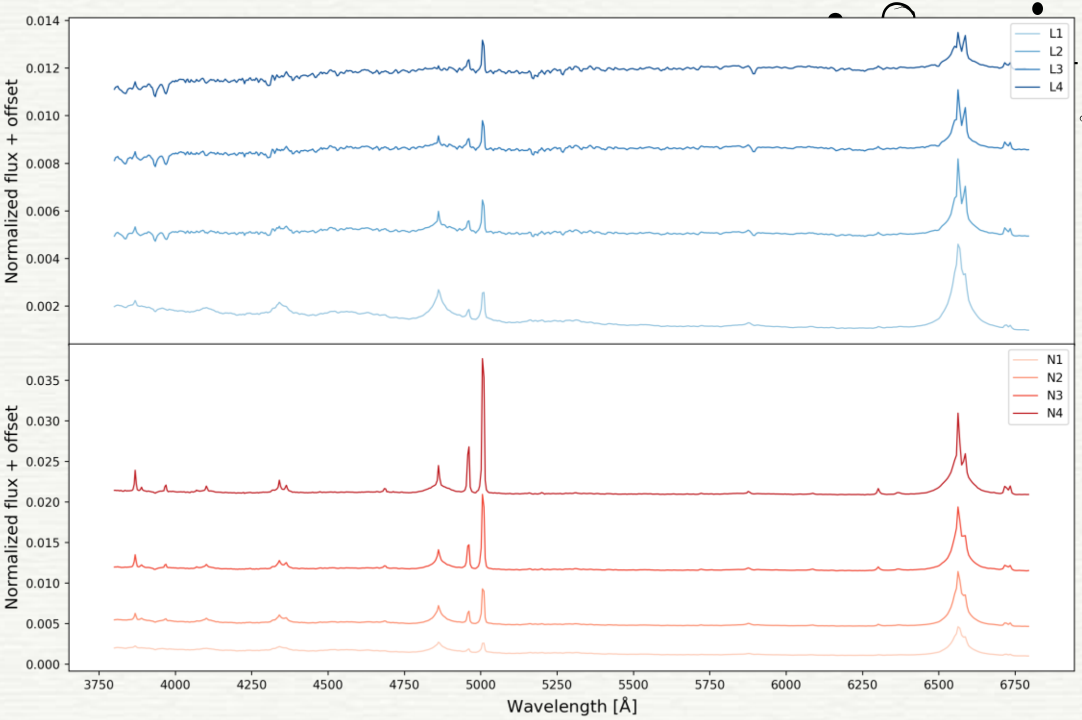
Preliminary results: Spectral LLE

- Robust LLE applied to a sample of **4022 spectra** of broad-line AGN (SDSS DR7)
- This is a subsample with best quality spectra from the Liu+2019 sample used in Parameter LLE.
- Dimensionality reduced from **650-D to 3-D!**
- Objects with different accretion rate (Eddington ratio) separated in the projection.



Preliminary results: Spectral LLE

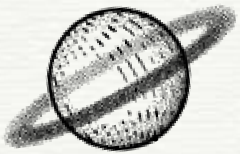




Conclusions

Parameter LLE

- Robust LLE applied to 19 measured optical spectral parameters of low-z broad-line SDSS quasars (Liu+2019 catalog)
- Successfully identified different quasar branches, clearly establishing the main sequence (M-branch), but also identifying new ones (N-branch).
- Obtained a 3-D projection of the quasar main sequence and other optical correlates for the SDSS quasars from this catalog, for the first time.



Spectral LLE

- Robust LLE applied to corresponding SDSS spectra, reducing dimensionality from 650 down to 3.
- Obtained projections were probed by calculating mean spectra from different regions.
- Clear trends observed along branches (e.g. narrow line emission)
- At least one of the branches is manifestation of the branches obtained from Parameter LLE.

References

Boroson, T. A., Green, R. F. : 1992, *ApJS*, **80**, 109

Carlsson, G. : 2009, *Bull. Amer. Math. Soc. (N.S.)*, **2**, 255-308

Chang, H., Yeung, D. : 2006, *Pattern Recognition*, **39**, 1053-1065

Fefferman, C., Mitter, S., Narayanan, H. : 2016, *Journal of the American Mathematical Society*, **29**, 983-1049

Jankov, I., Ilić D., Kovačević, A. : 2020, Publ. Astron. Obs. Belgrade, **99**, 291

Liu, H. et al. : 2019, *ApJS*, **243**, 21

Ludwig, R. R. et al. : 2009, *ApJ*, **706**, 995-1007

Marziani, P. et al. : 2001, *ApJ*, **558**, 553-560

Marziani, P., Sulentic, J. W. : 2014, *MNRAS*, **442**, 1211

Roweis, S. T., Saul, L. K. : 2000, *Science*, **290**, 2323-2326

Shen, Y., Ho, L. C. : 2014, *Nature*, **513**, 210-213

Sulentic, J. W. et al. : 2000, *Annu. Rev. Astron. Astrophys.*, **38**, 521-571

Thank you!