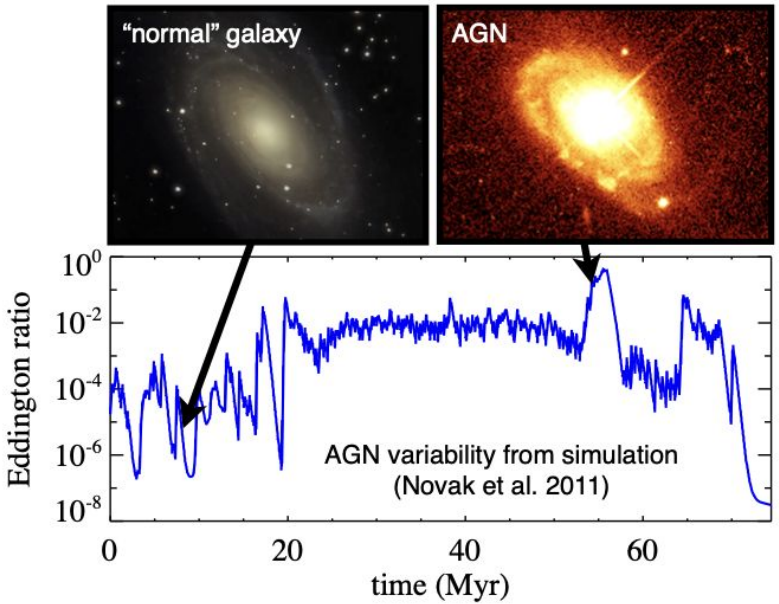
when Active Galactic Nuclei become rare gems

Stéphanie Juneau
NOIRLab

NOIRLab/NSF/AURA/J. da Silva

Context

- Most (all) massive galaxies have a supermassive black hole
- Active Galactic Nuclei are found in only a fraction of galaxies
  → BHs are only “active” some of the time (*flicker* on and off)

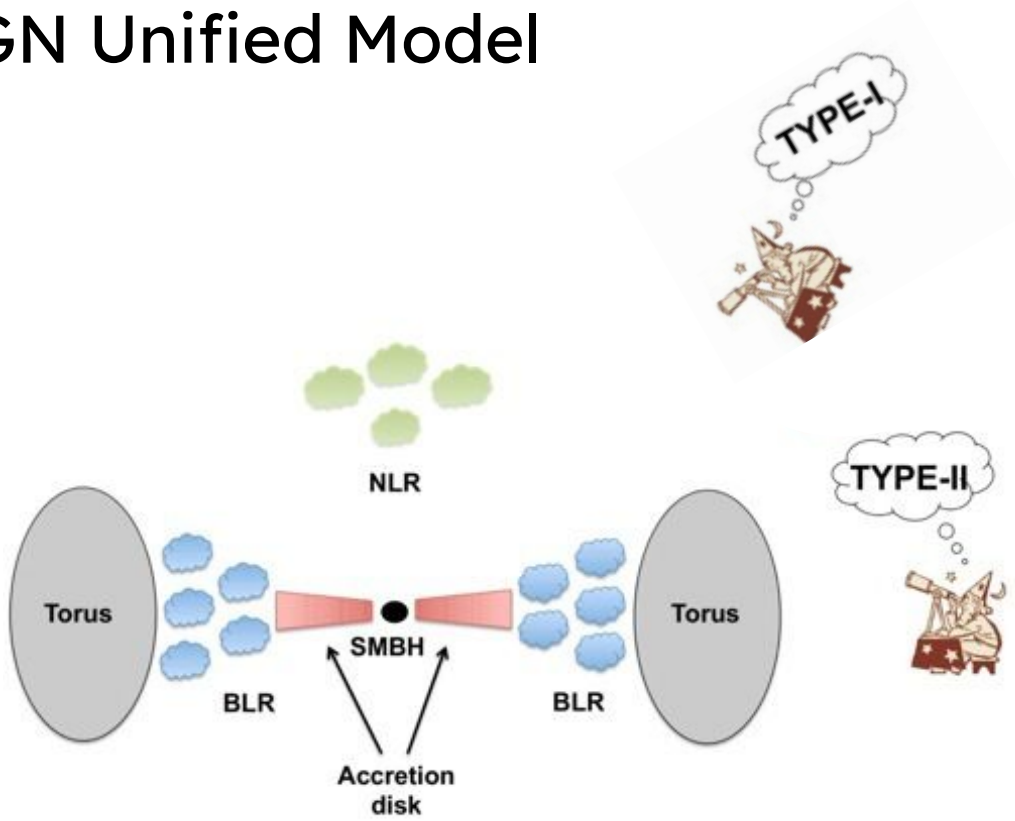


(Hickox et al. 2014)

Context

- Most (all) massive galaxies have a supermassive black hole
- Active Galactic Nuclei are found in only a fraction of galaxies
  → BHs are only “active” some of the time (*flicker* on and off)
- AGN are found in different “flavors”
  - *Active Galactic Nuclei: What’s in a name?* (Padovani+ 2017)
  - Due to orientation effects and/or evolutionary trends or phases

AGN Unified Model



This version: Claudio Ricci

Original versions:
Antonucci (1993); Urry & Padovani (1995)

Also see:
Netzer (2015); Thorne et al. (2022)

Context

- Most (all) massive galaxies have a supermassive black hole
- Active Galactic Nuclei are found in only a fraction of galaxies
  → BHs are only “active” some of the time (*flicker* on and off)
- AGN are found in different “flavors”
  - *Active Galactic Nuclei: What’s in a name?* (Padovani+ 2017)
  - Due to orientation effects and/or evolutionary trends or phases
- Open questions regarding supermassive black holes:
  - **Formation**: black hole seeds
  - **Growth**: accretion and/or mergers
  - **Feedback**: self-regulation + regulation of host galaxies

Approach

- Leverage large optical spectroscopy surveys
  - SDSS: ~1 million galaxies & quasars
  - DESI: ~1 million galaxies & quasars (EDR) and >10 million (DR1; 2025)

- **Part I:** Compare traditional AGN diagnostics to ML model trained on complete spectra (e.g., Portillo+20, Pat+22, Melchior+23, Liang+23a,23b, Böhm+23)

- **Part II:** Investigate DESI target selection: filling in a missing quasar population + finding rare gems (Juneau+24)

Probabilistic Autoencoder (PAE)

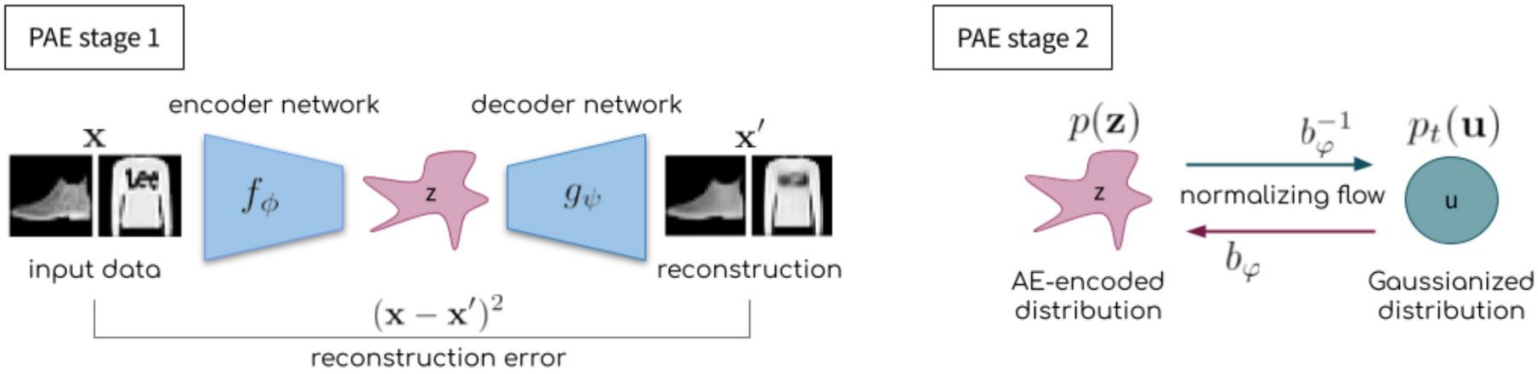


Figure 1. The two-step training of a Probabilistic Autoencoder (PAE) illustrated on an image data set. In the first stage, an autoencoder (AE) consisting of two neural networks is trained to minimize the reconstruction error after compressing (encoding) the data in a lower-dimensional latent space and decompressing (decoding) it back into the high-dimensional data space. In the second stage, a normalizing flow (NF) is trained to learn a bijective mapping from the AE-encoded latent space to a space in which the encoded data follows a Gaussian distribution.

https://github.com/VMBoehm/PytorchPAE

(Böhm & Seljak 2022; Böhm, Kim & Juneau 2023; also see a similar PAE approach by Melchior et al. 2023)

Probabilistic Autoencoder (PAE)

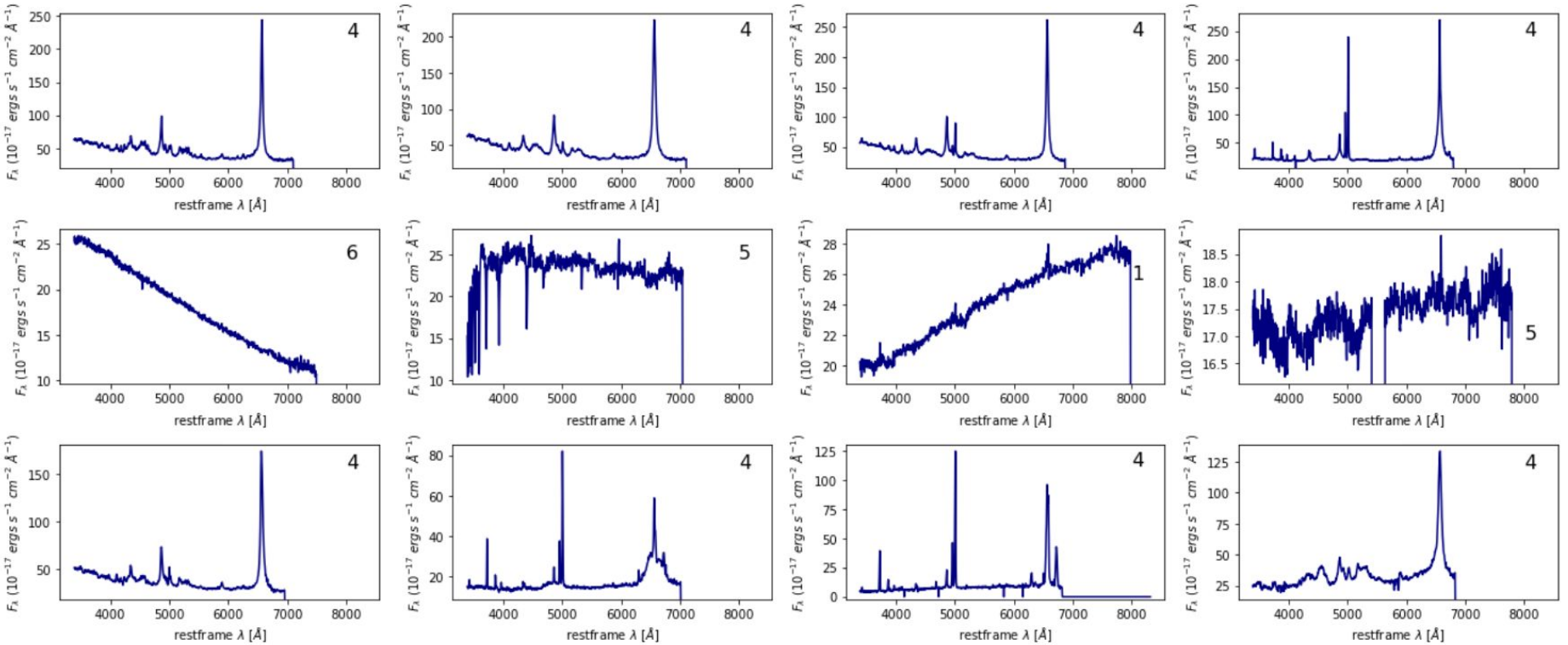
- Strategies to account for certain classes (e.g., QSO) being more rare than others (e.g., GALAXY)
  - Evaluate conditional probability given a label (class+subclass)

- Examples without (e.g., Pat+22, Liang+23a) and with (Böhm+23) conditional density estimation

https://github.com/VMBoehm/PytorchPAE

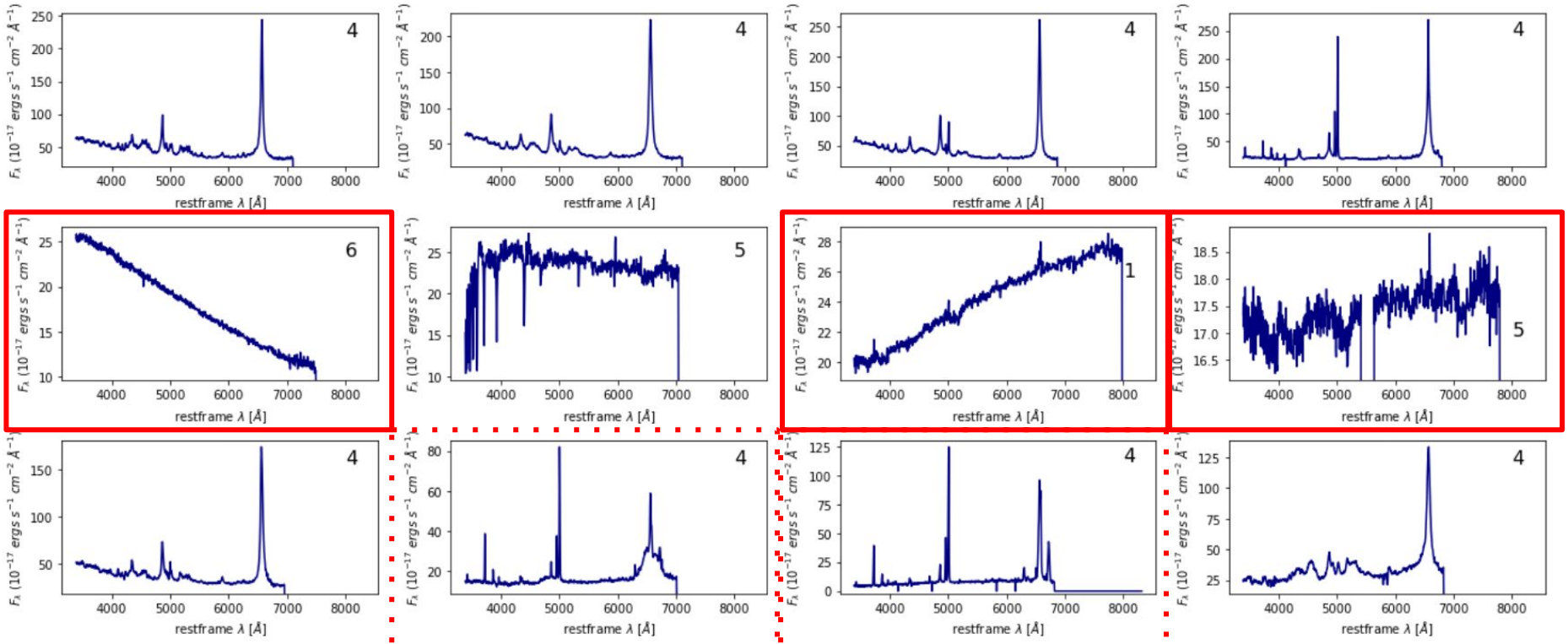
(Böhm & Seljak 2022; Böhm, Kim & Juneau 2023)

Example outlier SDSS spectra (non-conditional on labels)



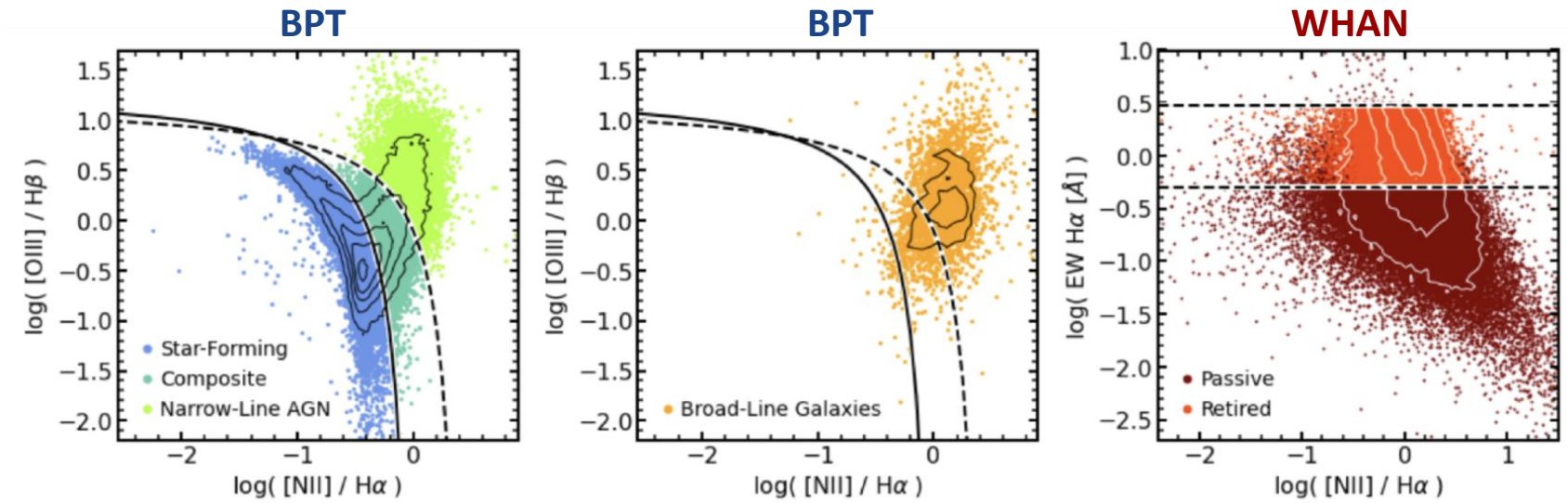
(Pat, Juneau, Böhm et al., 2022)

Example outlier SDSS spectra (non-conditional on labels)



(Pat, Juneau, Böhm et al., 2022)

Classification combining BPT & WHAN diagrams

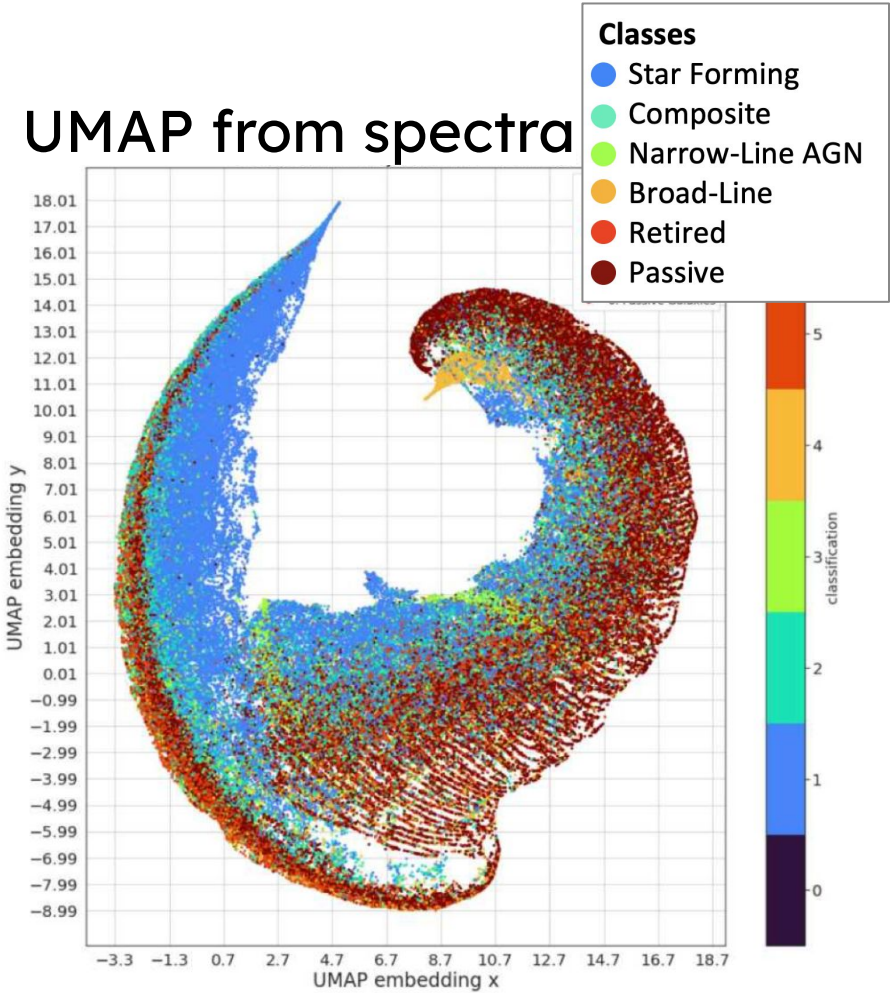


Classes
- Star Forming
- Composite
- Narrow-Line AGN
- Broad-Line
- Retired
- Passive

(Left) Star-Forming, Composite and Narrow-line AGN; **(Middle)** BROADLINE flag is set by SDSS pipeline → Broad-line AGN; **(Right)** Retired & Passive galaxies

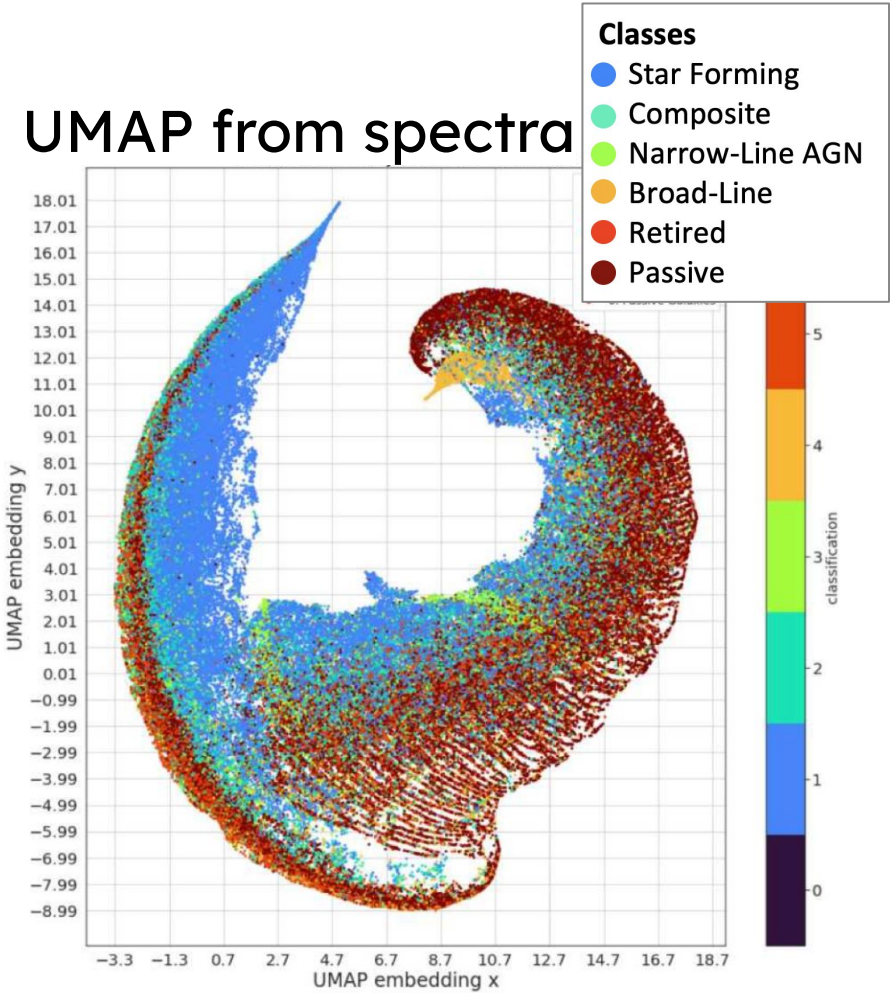
WHAN: Equivalent width of Hα vs. [NII]/Hα ratio (Cid-Fernandes et al. 2011)

UMAP from spectra

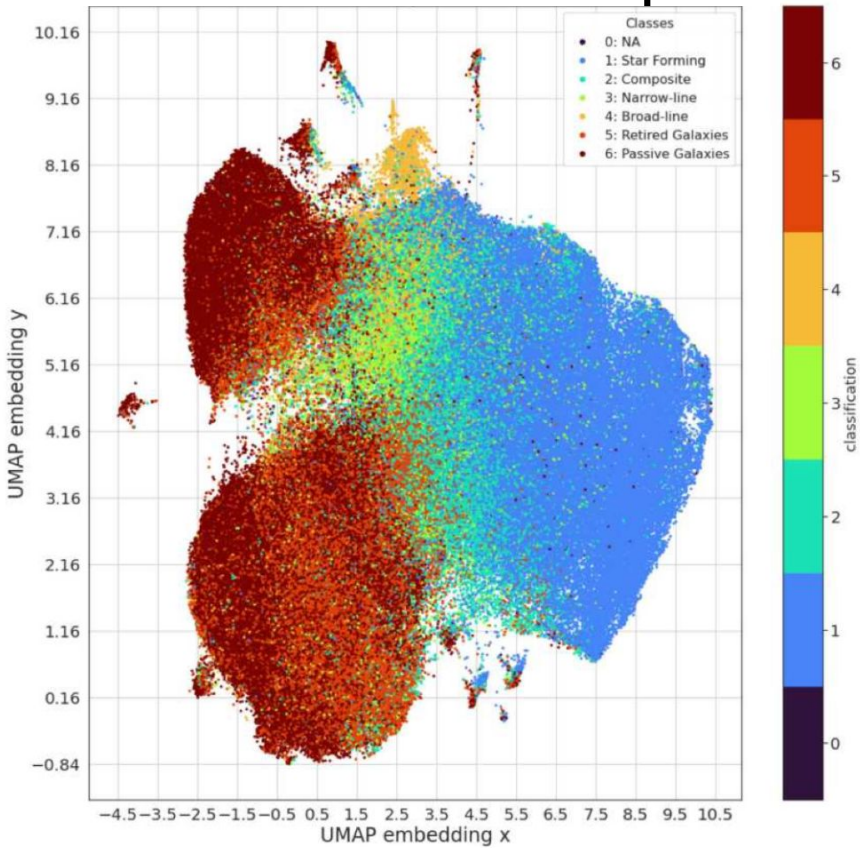


(Pat, Juneau, Böhm et al., 2022)

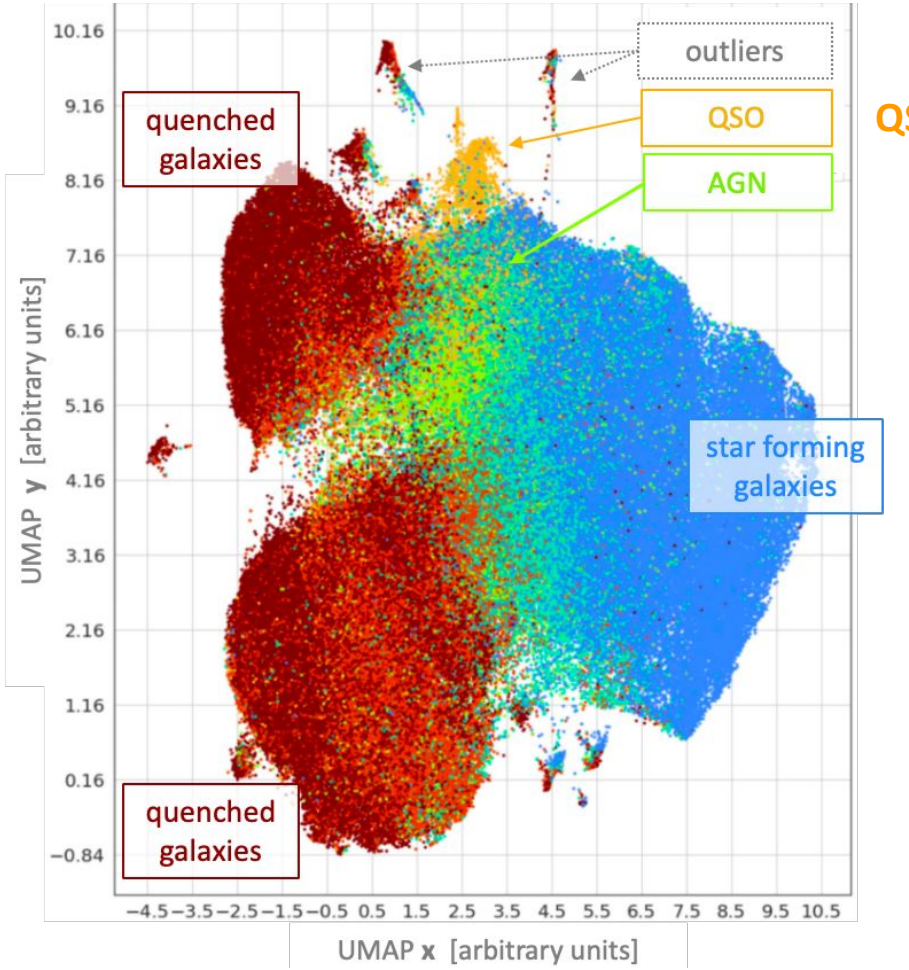
UMAP from spectra



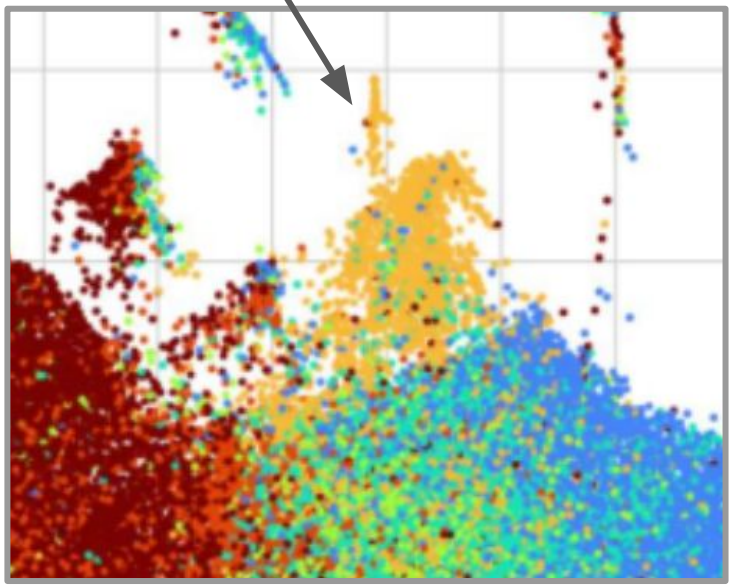
UMAP from AE latent space



(Pat, Juneau, Böhm et al., 2022)

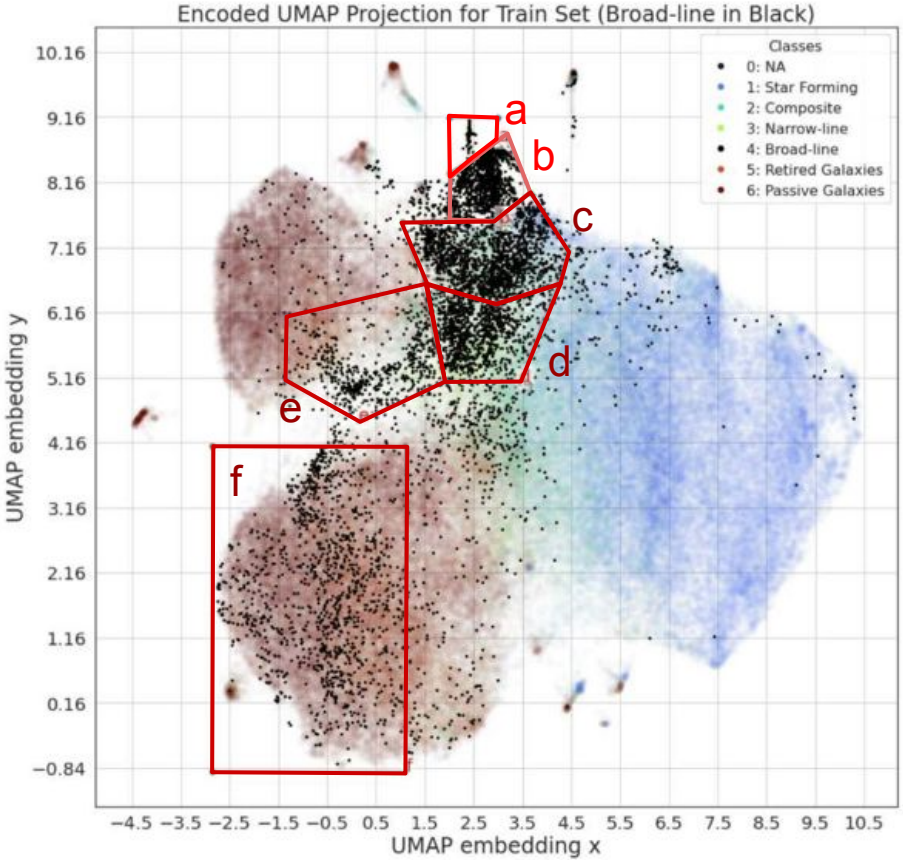


QSO (BROADLINE label)

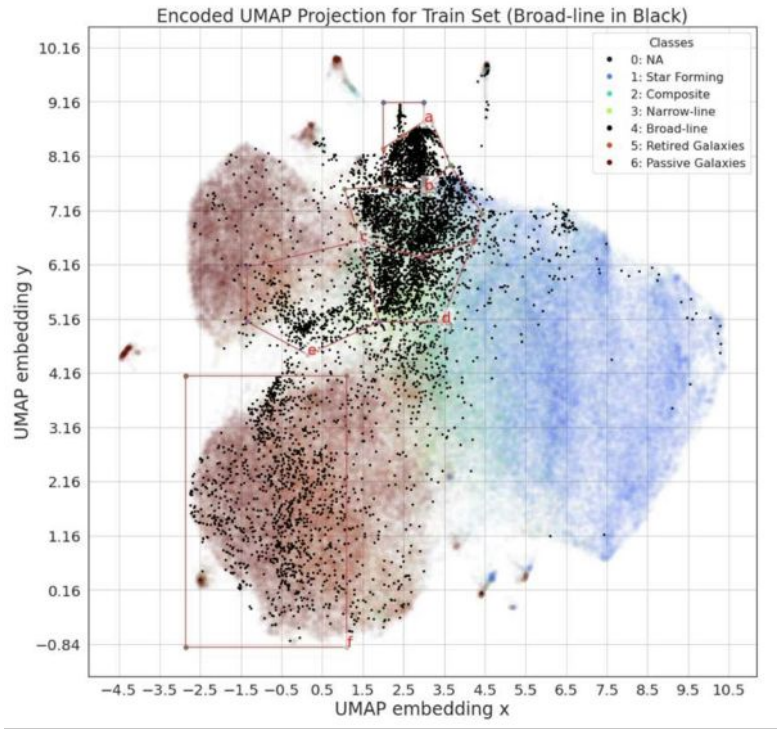


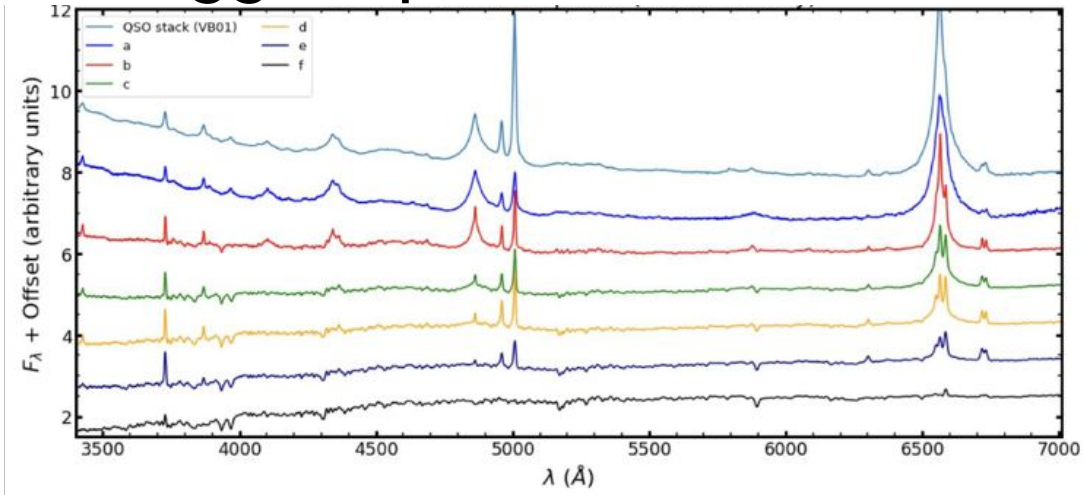
(Pat, Juneau, Böhm et al., 2022)

Trends among BROADLINE flagged spectra

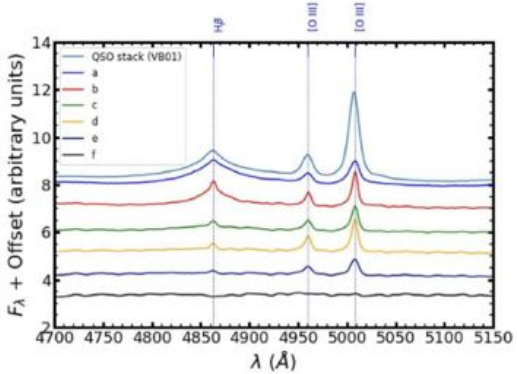


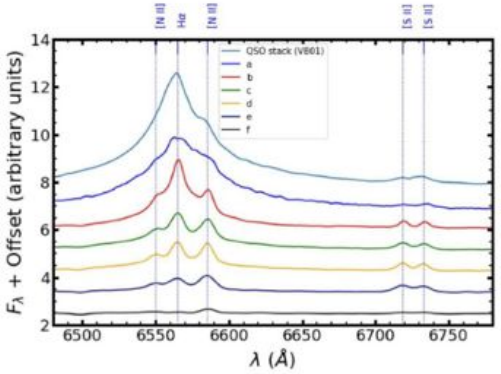
Trends among BROADLINE flagged spectra





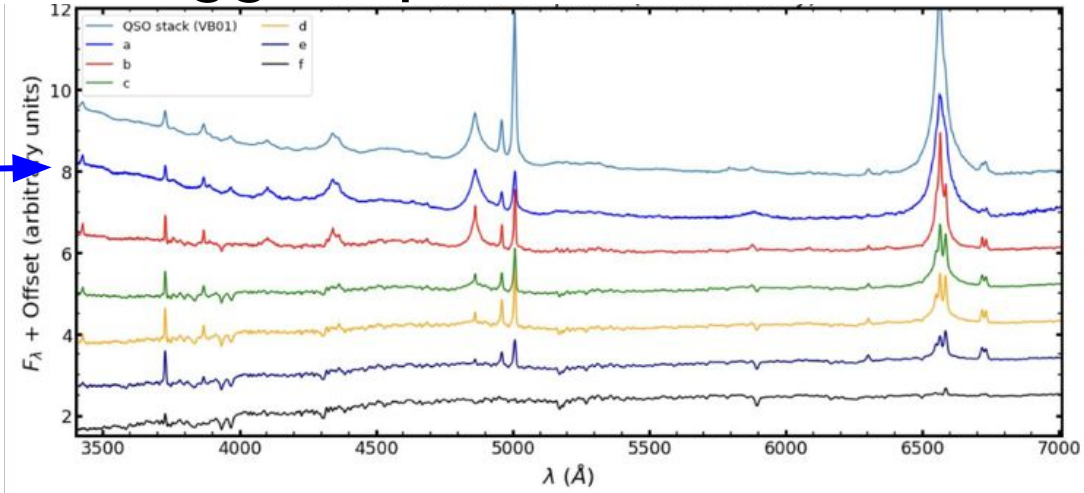
(a) Stacked spectra of broad-line sequence from UMAP (Figure 6)



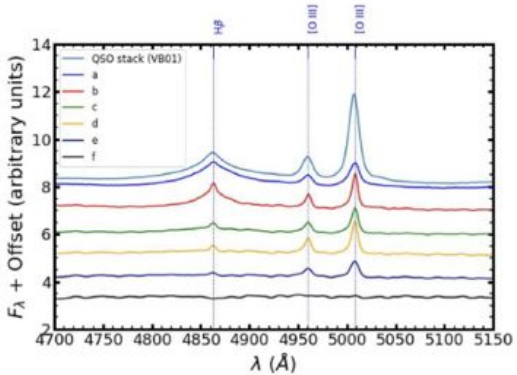


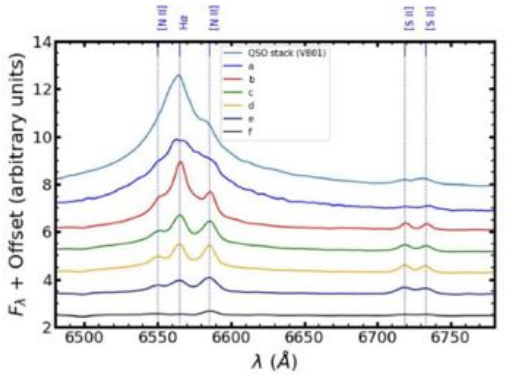
Trends among BROADLINE flagged spectra

The distinct branch consists of QSO spectra nearly identical to average SDSS quasar (VandenBerk 2001)



(a) Stacked spectra of broad-line sequence from UMAP (Figure 6)

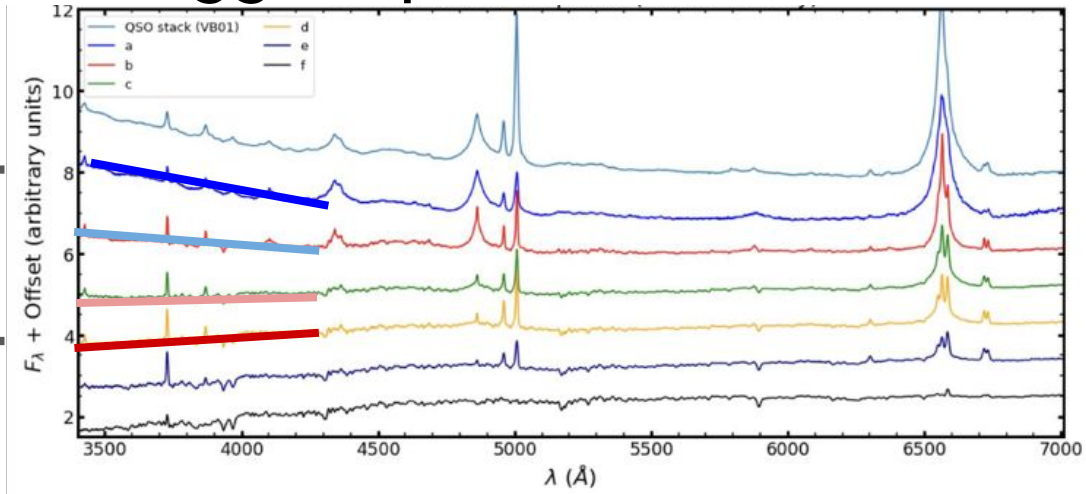




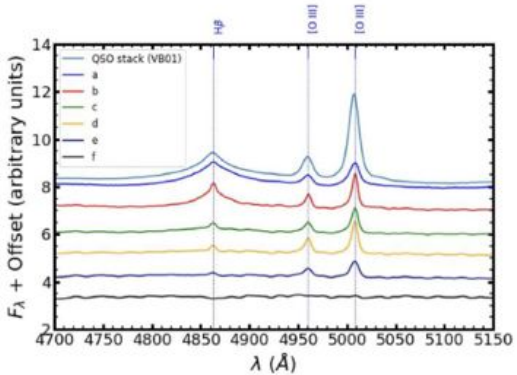
Trends among BROADLINE flagged spectra

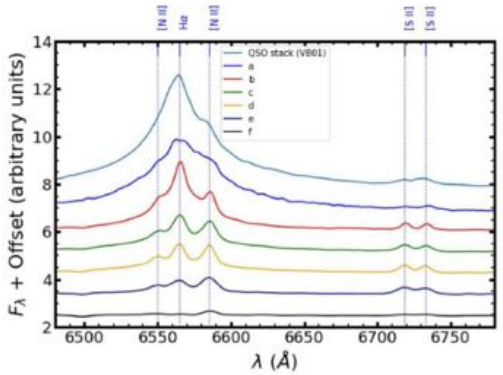
The distinct branch consists of QSO spectra nearly identical to average SDSS quasar (VandenBerk 2001)

Then a trend of reddening/obscuration

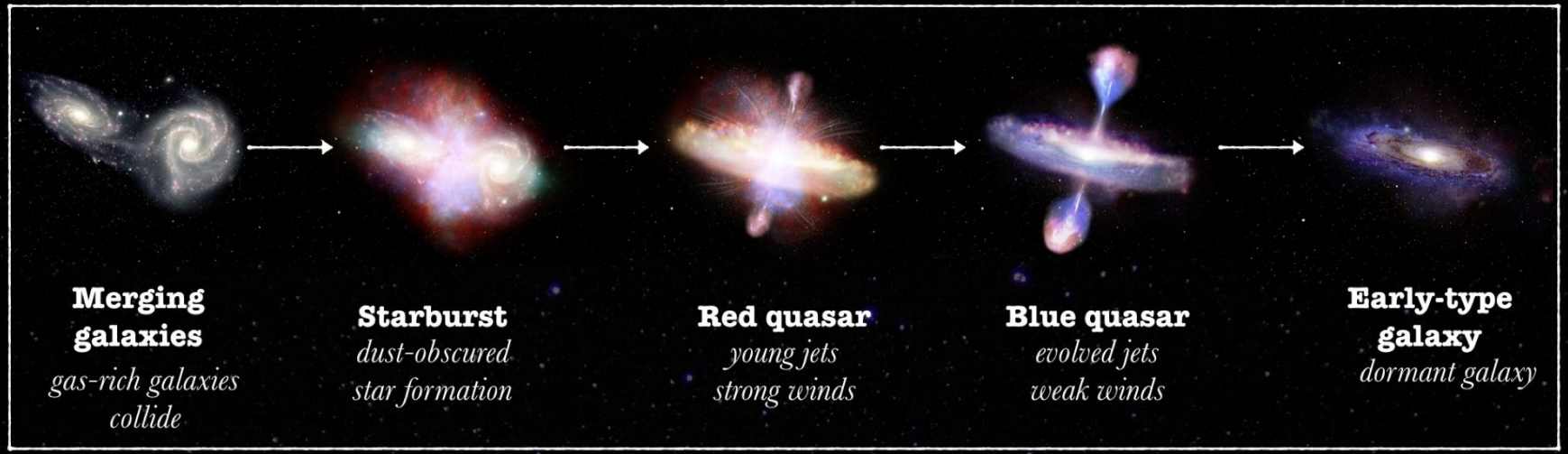


(a) Stacked spectra of broad-line sequence from UMAP (Figure 6)





Red quasars as a rare population (or in a special phase)



- DESI secondary program led by Vicky Fawcett (Fawcett et al. 2023)
- Construct a more complete and diverse census of quasars, particularly reddened and extremely reddened quasars.
- Test whether dust-obscured quasars represent a special (transition) phase in the evolution of quasars and the formation of massive galaxies.

Slide credits: R. Fawcett, J. Moustakas

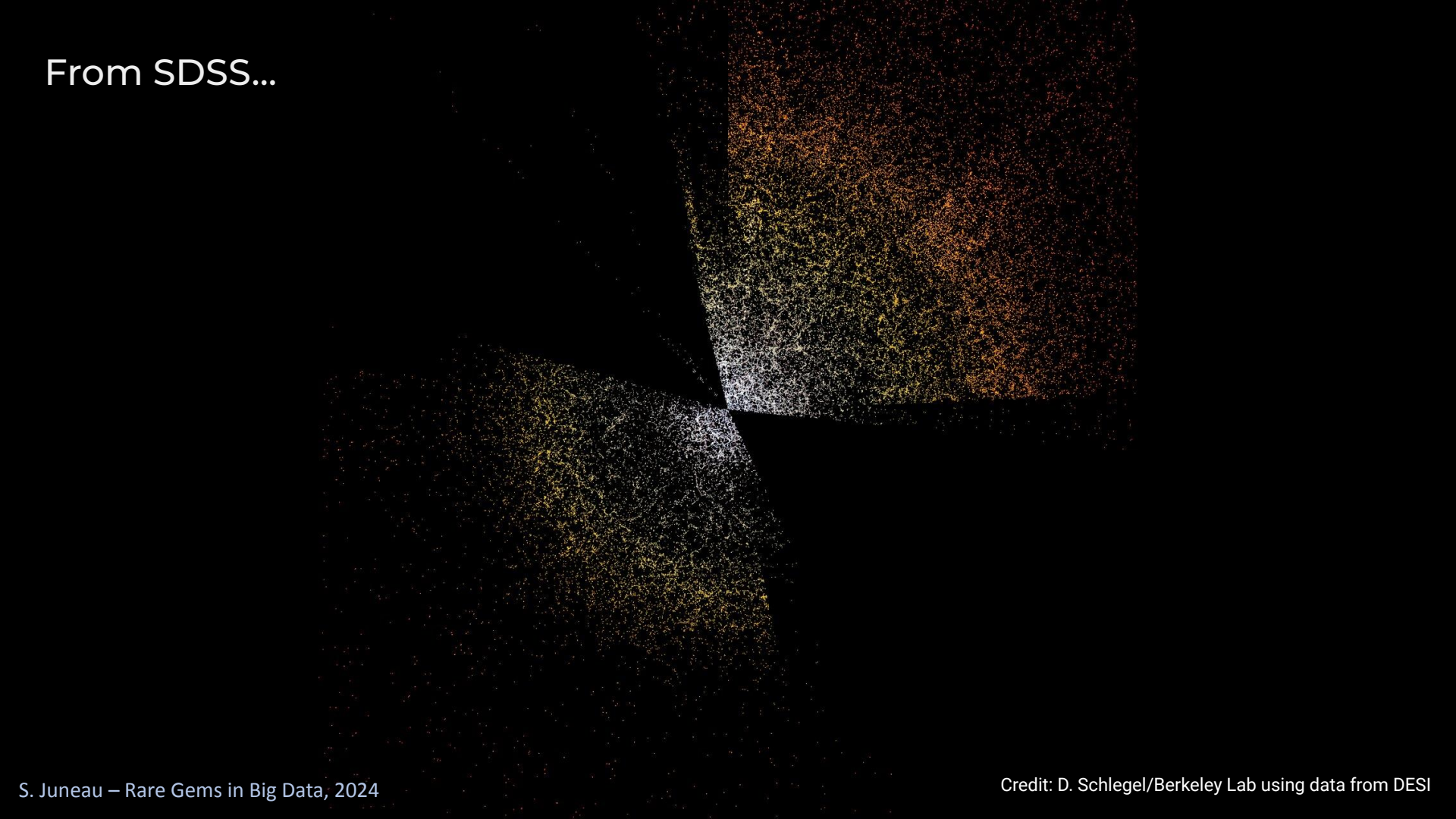
DESI (Dark Energy Spectroscopic Instrument)

Installed at the 4-meter Mayall telescope (Kitt Peak, AZ)

Will obtain spectra of 40 million galaxies & quasars in 5 years (2021-2026)

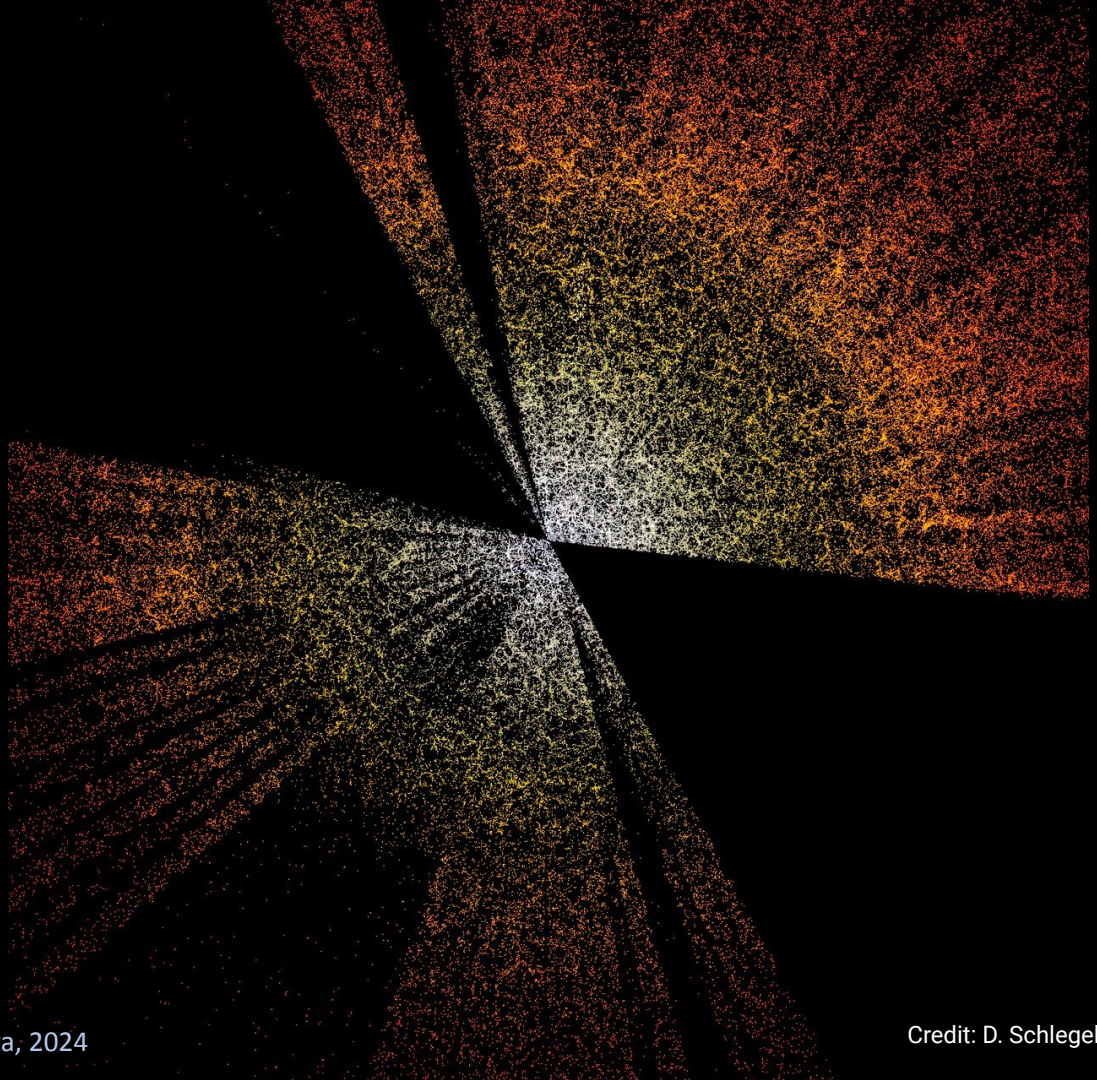
DESI is funded by the DOE Office of Science and several others as listed at desi.lbl.gov/acknowledgments/

From SDSS...



Credit: D. Schlegel/Berkeley Lab using data from DESI

...to DESI



Credit: D. Schlegel/Berkeley Lab using data from DESI

Embarking on a Quasar rescue mission

DESI Bright Galaxy Survey:

- Bright (r < 19.5) + Faint (r < 20.3 AB)
- Star/Galaxy separation (G – r > 0.6) removes stars (good) and some quasars (sad)

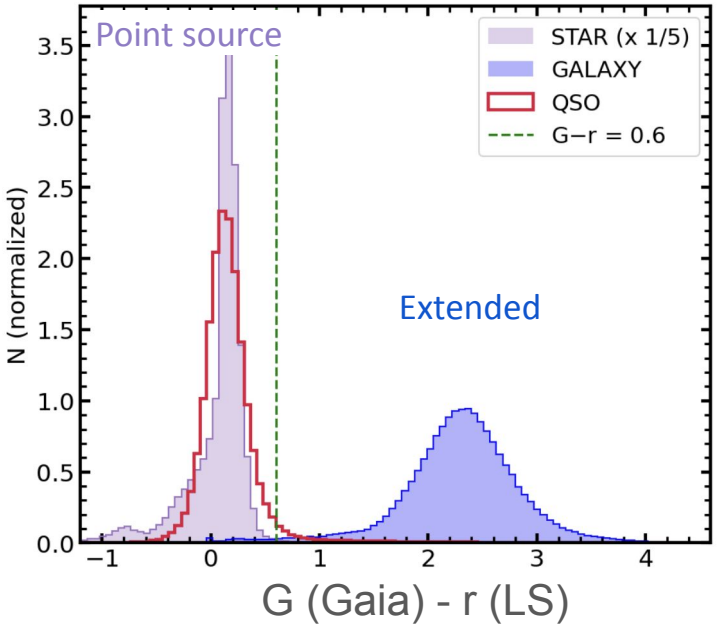
Mission:

- Put the quasars back in!
  (without increasing stellar contamination)

Strategy:

- Color cuts leveraging WISE bands probing AGN-heated dust (g, r, z, W1, W2)

Star/Galaxy separation



(Juneau et al. 2024; arXiv:2404.03621)

Embarking on a Quasar rescue

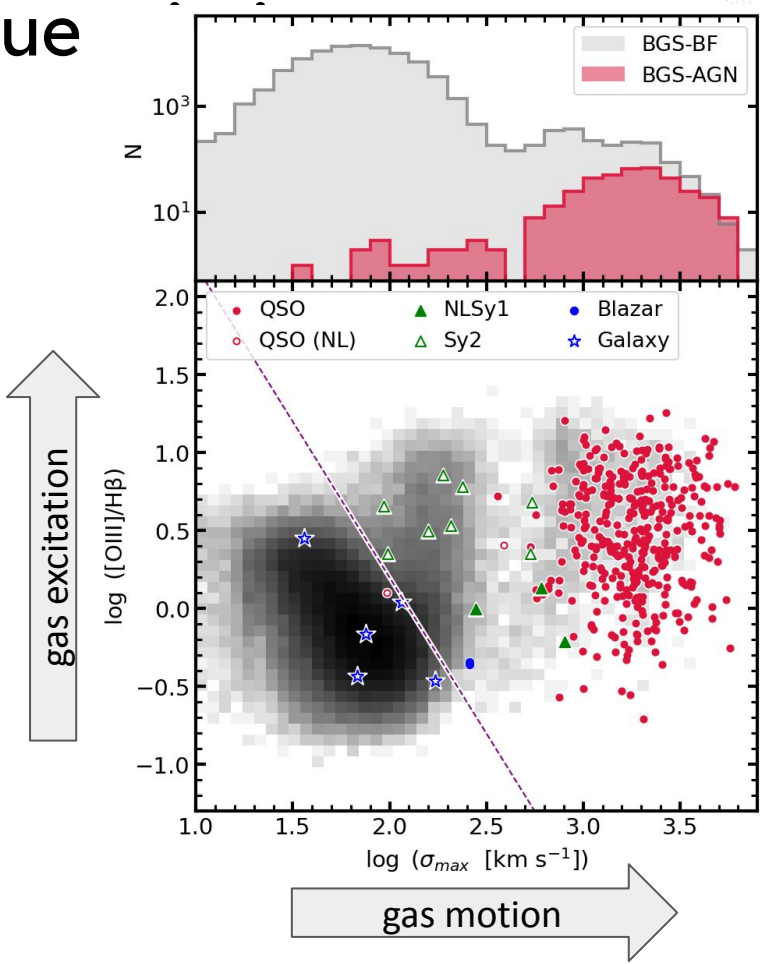
Final results

- As anticipated, BGS AGN targets are mostly Quasars but often with some host galaxy light → intermediate between the BGS sample and the QSO sample

- Also found: Type 2 Seyfert (Sy2) AGN, Narrow-Line Seyfert 1 (NLSy1) AGN, Blazars, BAL quasars, red quasars

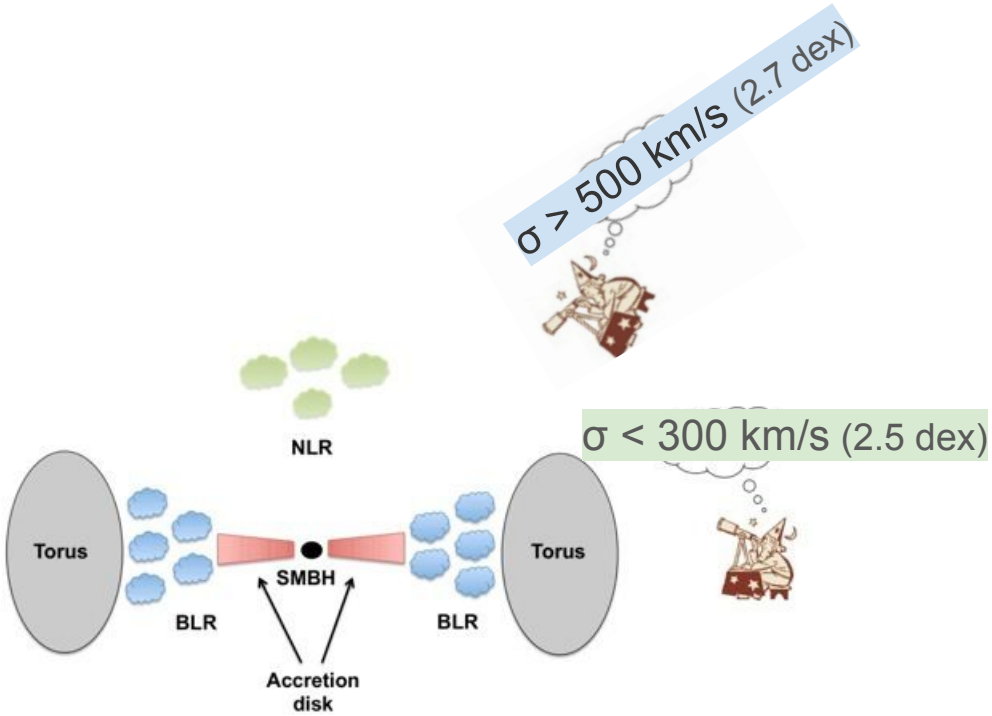
DESI DR1: >15,000 unique BGS AGN targets

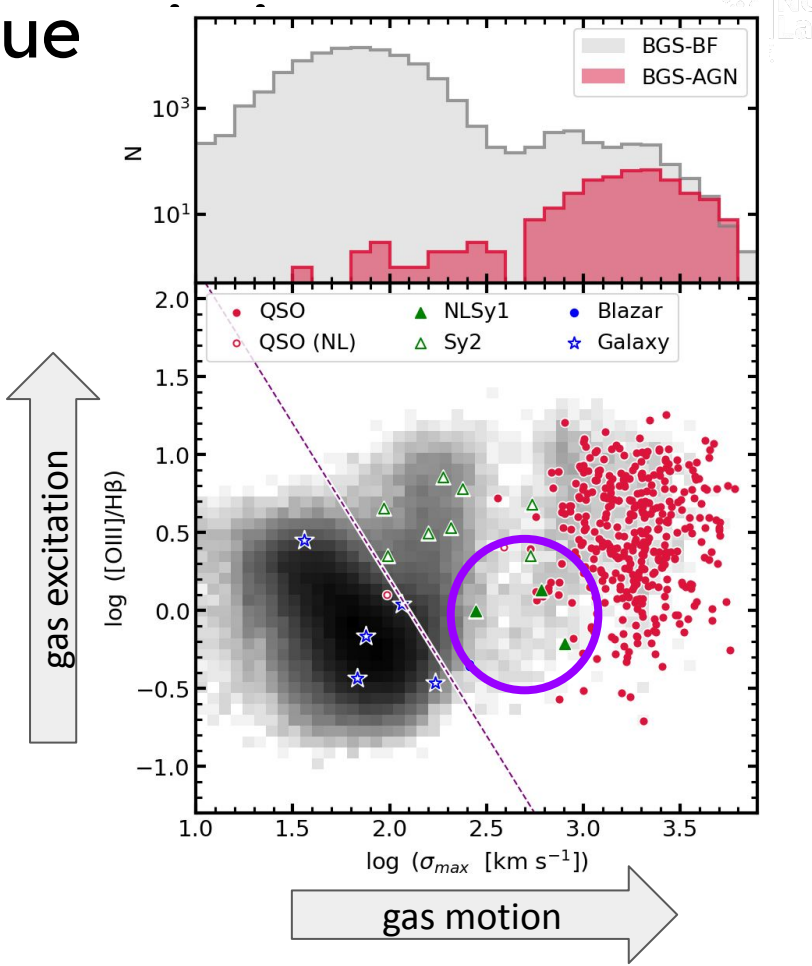
Will be part of AGN/QSO Catalog (millions)



(Juneau et al. 2024; arXiv:2404.03621)

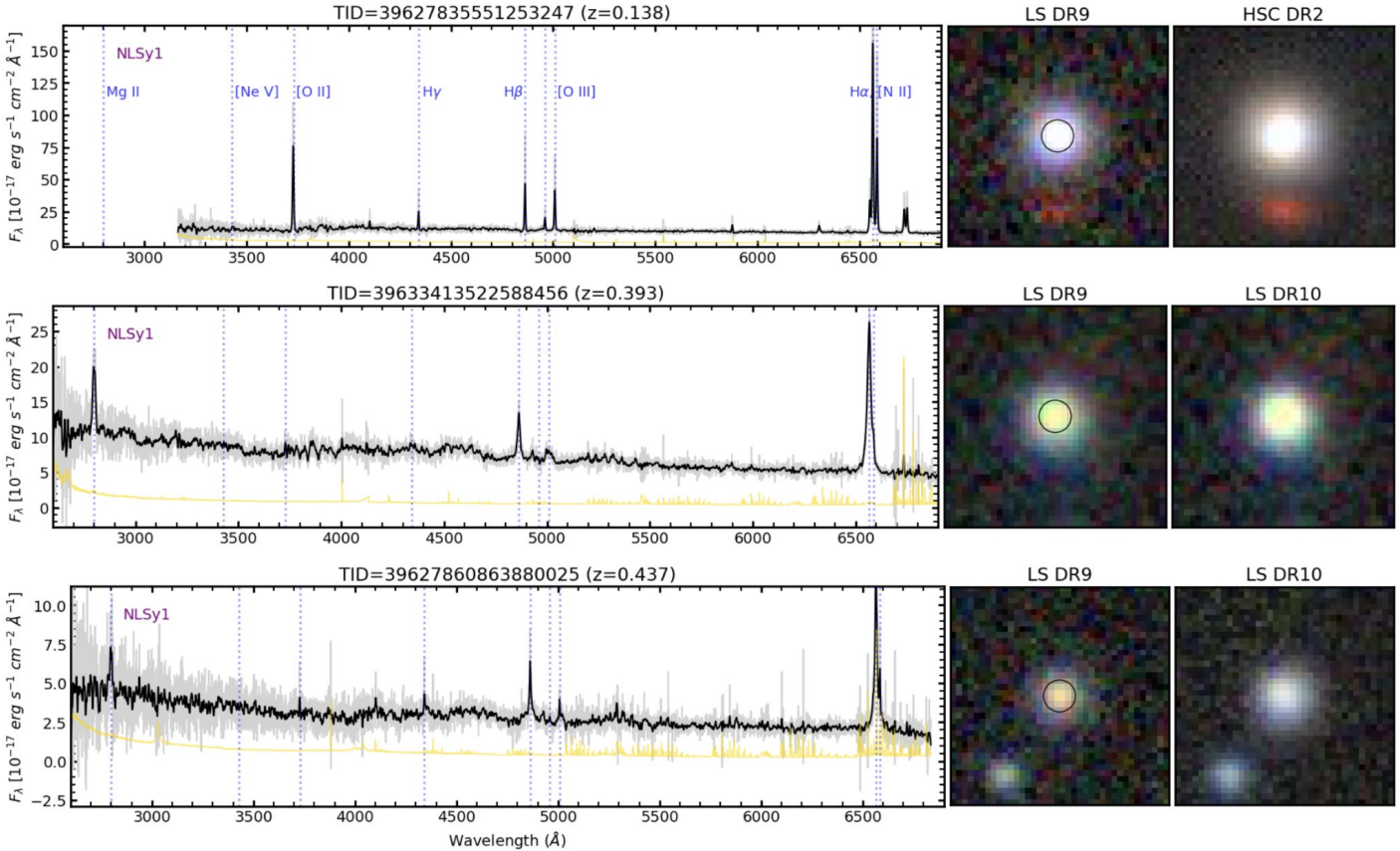
Embarking on a Quasar rescue





(Juneau et al. 2024; arXiv:2404.03621)

Spectra + images for NLSy1 = Narrow-Line Seyfert 1



Seyfert 1 → normally have *broad* lines

<u>Scenario:</u>

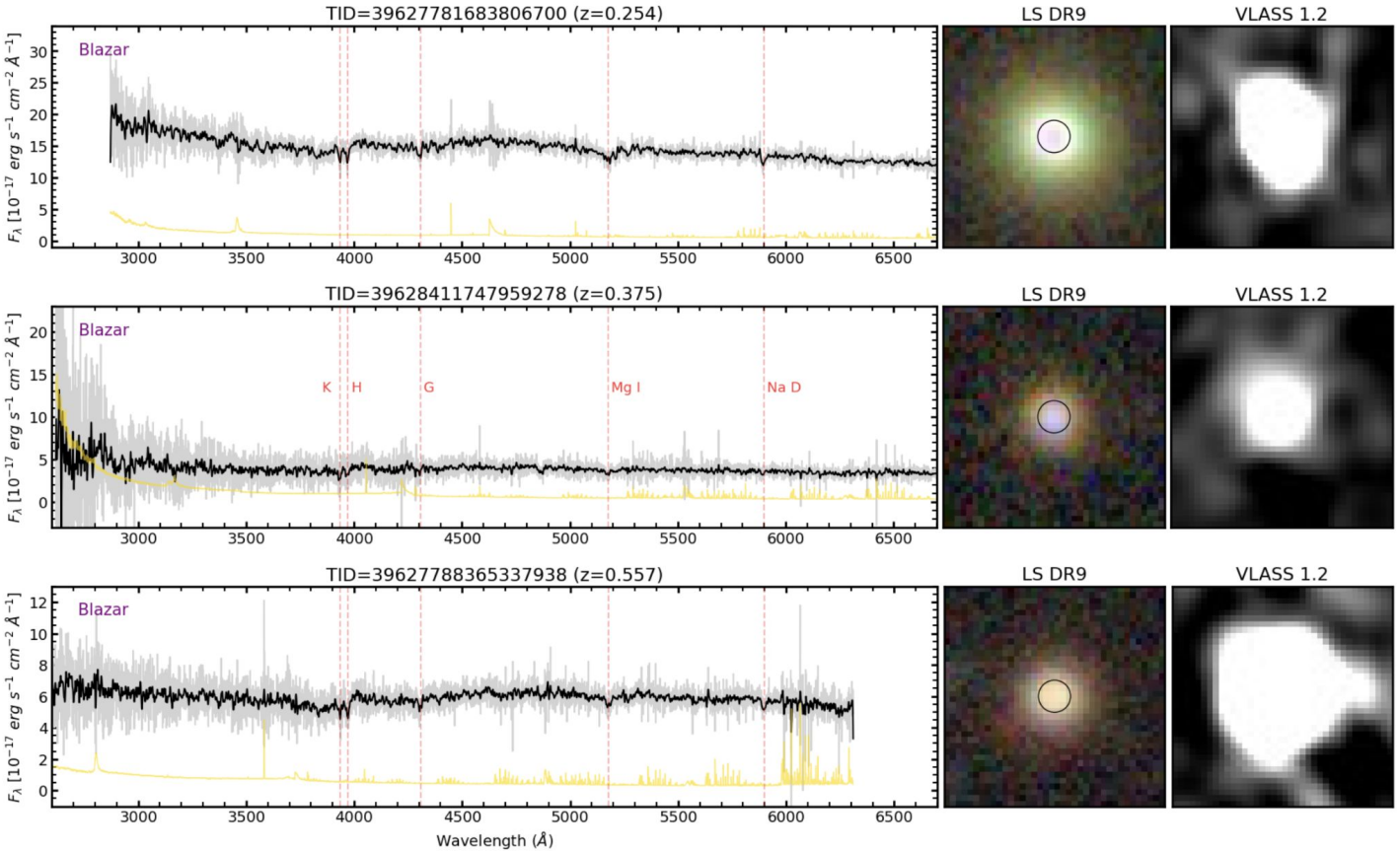
Young AGN with a low-mass black hole growing fast

→ analogs to high-redshift AGN found with JWST?

(topic: BH growth / BH seeds)

(Juneau et al. 2024; arXiv:2404.03621)

Spectra + images for Blazars



spectype=GALAXY

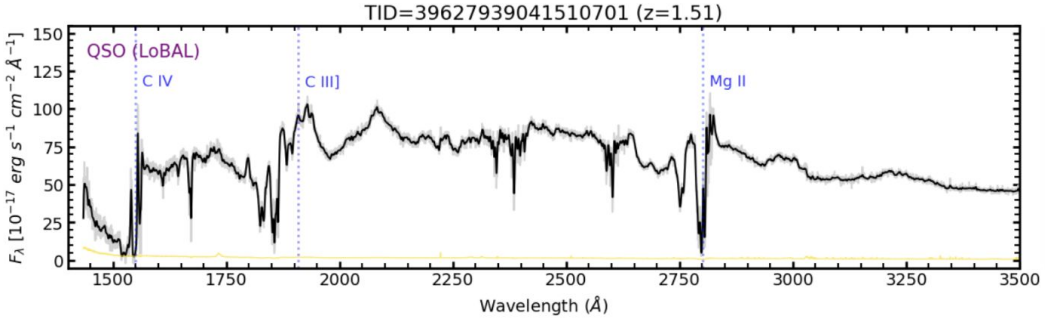
Stellar absorption lines but continuum rising toward the blue!

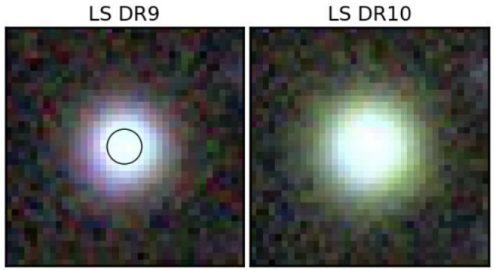
Obvious VLASS (3 GHz) counterparts

→ Looking toward radio jet
(rare orientation)

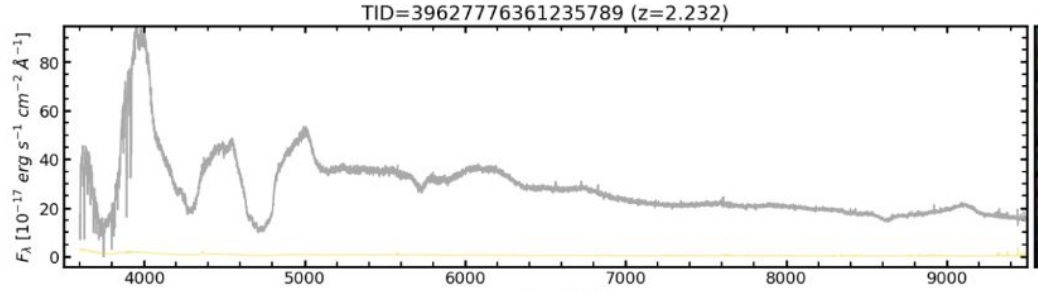
(Juneau et al. 2024; arXiv:2404.03621)

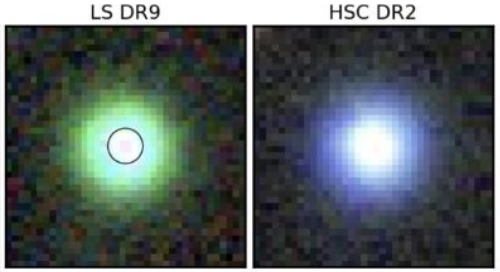
Spectra + images for BAL quasars



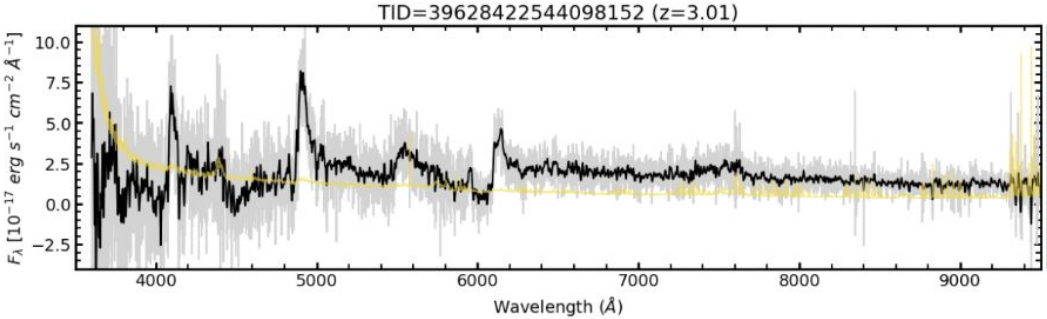


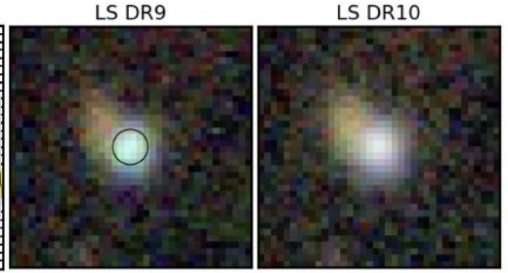
Missed by RR
(identified by VI)





Bright (r_{AB}<17.5)
PSF morphology
*also QSO target





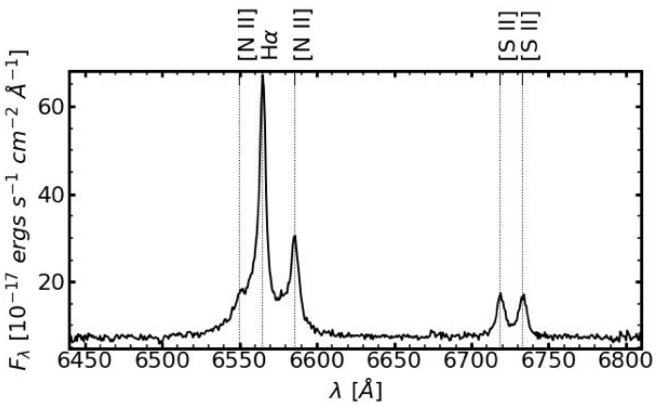
Faint (r_{AB}>17.5)

non-PSF morph.
due to projected
neighbor?

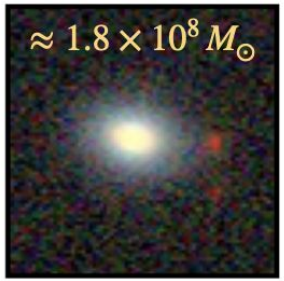
DARK ENERGY SPECTROSCOPIC INSTRUMENT
U.S. Department of Energy Office of Science

Active black holes in low-mass dwarf galaxies

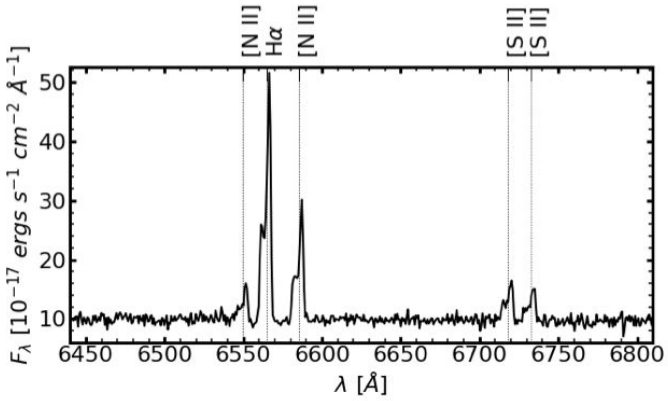
Ragadeepika Pucha (University of Utah)



Broad Balmer lines
→black hole mass



Dual-peaked lines;
binary black hole?



Research done using NOIRLab/Data Lab and NERSC

Slide credits: R. Pucha, J. Moustakas

Conclusions

- Probabilistic Autoencoders (PAE) can:
  - accurately reconstruct spectra (noiseless and inpainted over gaps)
  - find anomalies (with or without pre-assigned labels)
- Combining PAE + UMAP reveals interesting sequences worth studying further (Pat, Juneau et al. 2022)
  - Different populations (or evolutionary phases) of quasars
  - Similar to using VAE (Portillo et al. 2020) & SPENDER (Melchior et al. 2023 [**talk**], Liang et al. 2023a,b [**poster**])
- Developed target selection to recover *missing* quasars in the DESI Bright Galaxy Survey → found normal quasars but also some rare gems (Juneau et al. 2024)

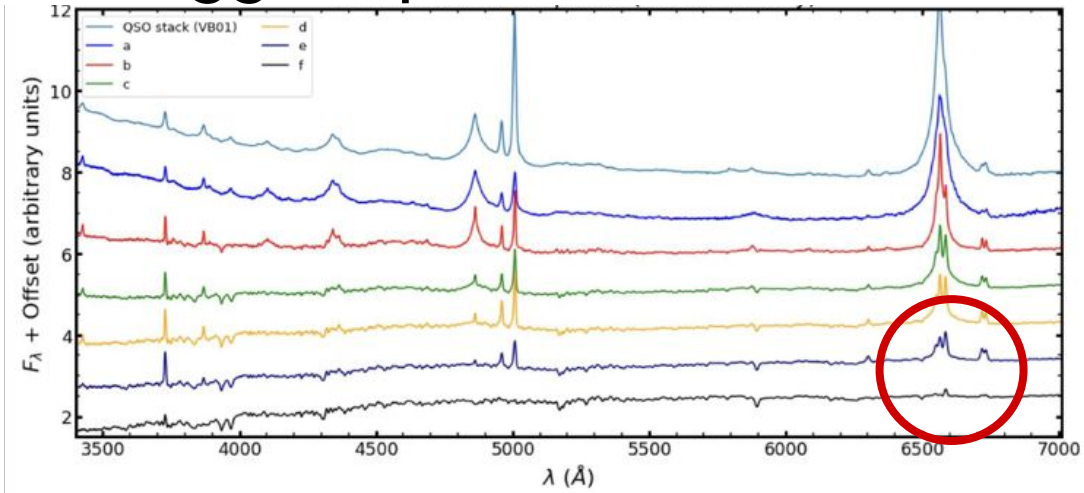
- **Discussion**: need to strategize about how to select targets but can also exploit some contaminants/failures as they can reveal interesting objects!

Trends among BROADLINE flagged spectra

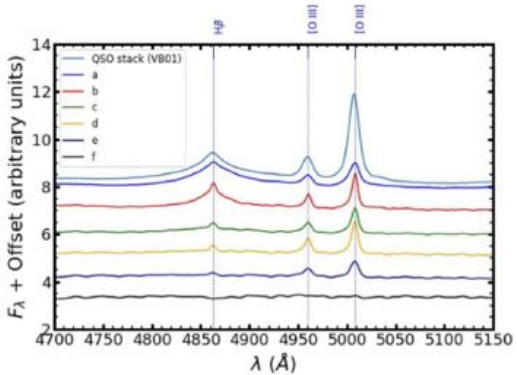
The distinct branch consists of QSO spectra virtually identical to average SDSS quasar (VandenBerk 2001)

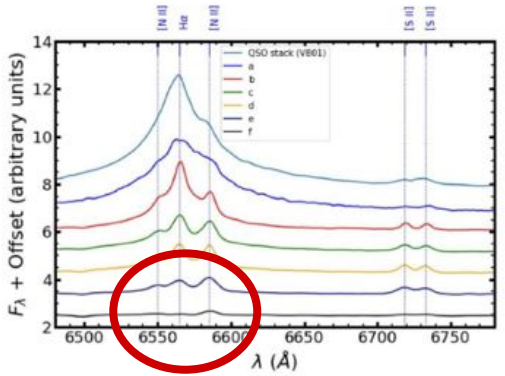
Then a trend of reddening/obscuration

Outliers don't have *true* broad (Balmer) lines and arise from SDSS definition (FWHM>400 km/s) picking up LINERs in massive galaxies

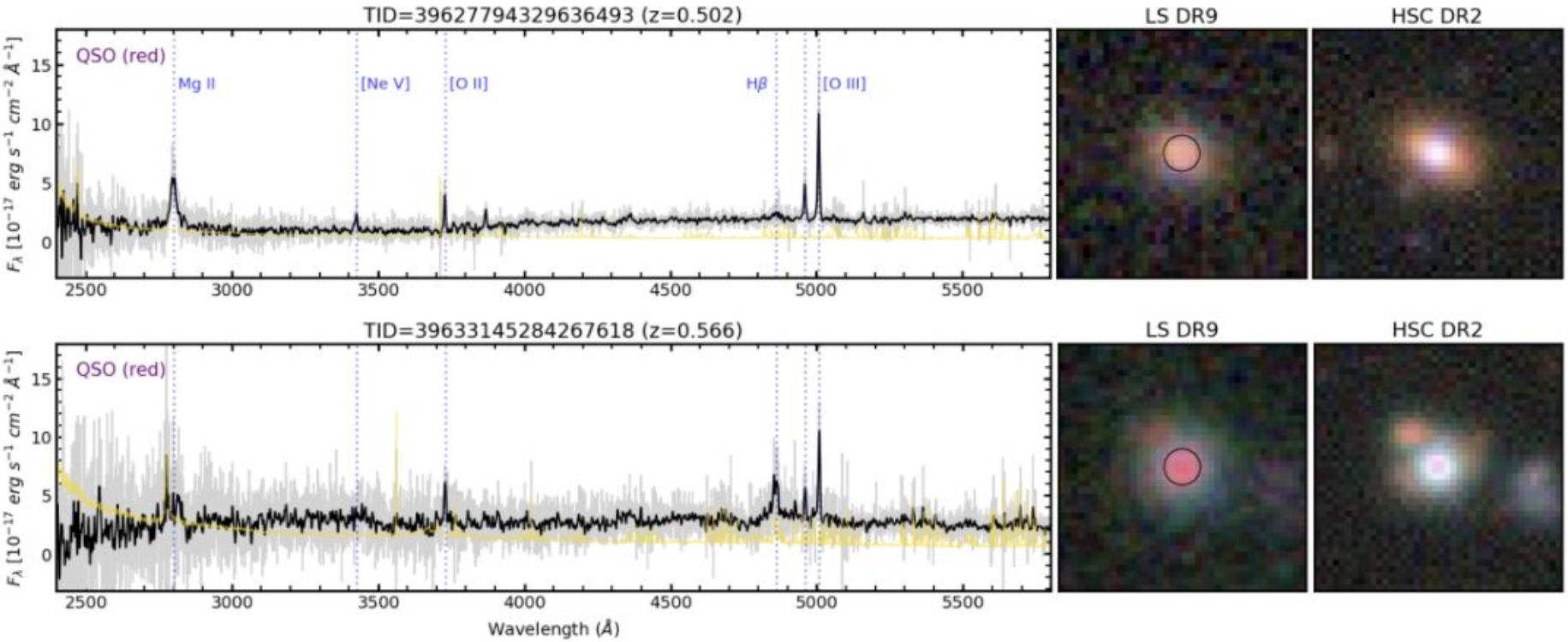


(a) Stacked spectra of broad-line sequence from UMAP (Figure 6)





Spectra + images for red quasars



(Juneau et al. 2024; arXiv:2404.03621)

Probabilistic Autoencoder (PAE)

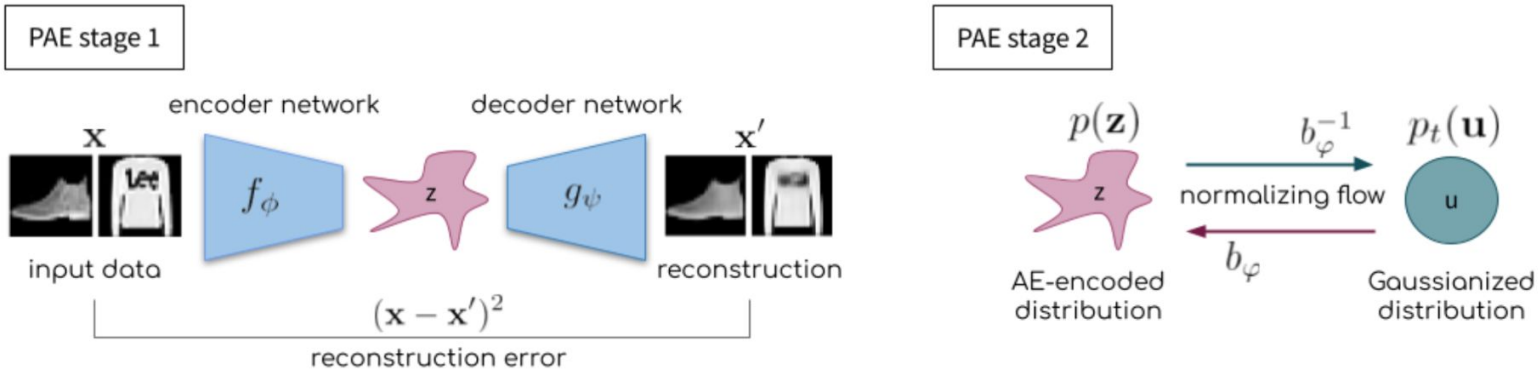


Figure 1. The two-step training of a Probabilistic Autoencoder (PAE) illustrated on an image data set. In the first stage, an autoencoder (AE) consisting of two neural networks is trained to minimize the reconstruction error after compressing (encoding) the data in a lower-dimensional latent space and decompressing (decoding) it back into the high-dimensional data space. In the second stage, a normalizing flow (NF) is trained to learn a bijective mapping from the AE-encoded latent space to a space in which the encoded data follows a Gaussian distribution.

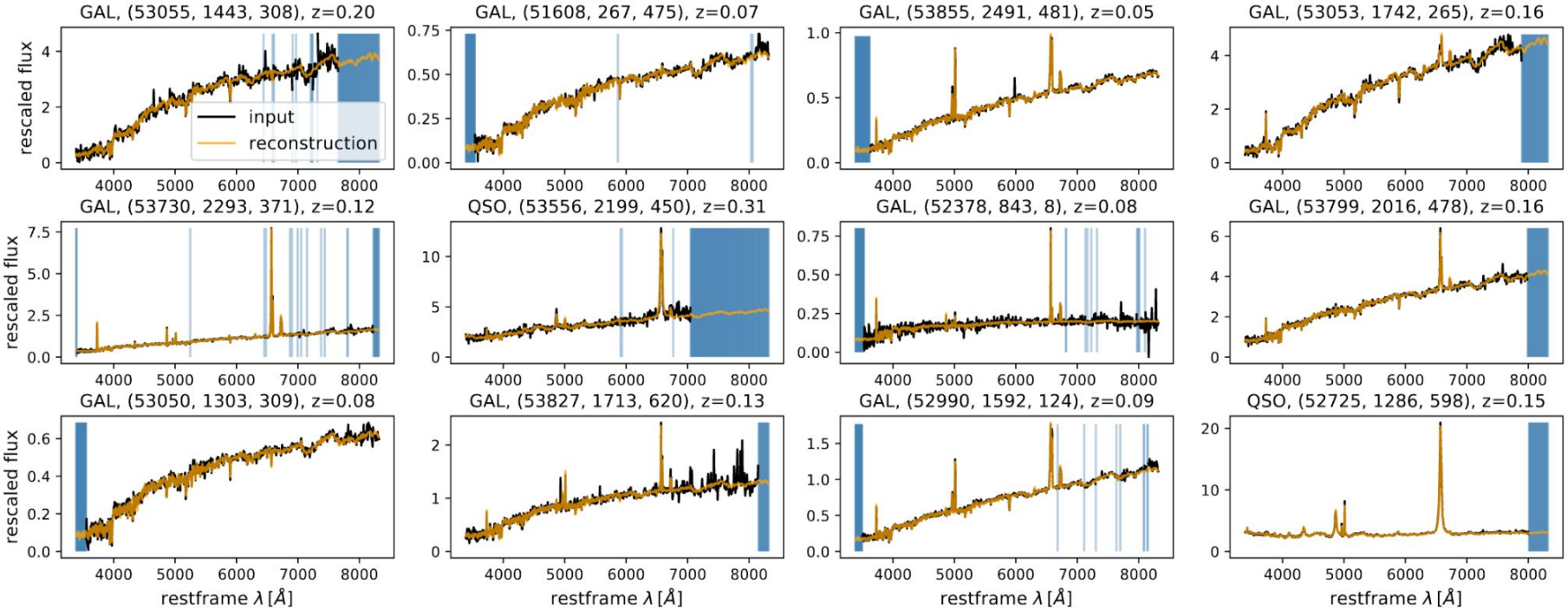
https://github.com/VMBoehm/PytorchPAE

(Böhm & Seljak 2022; Böhm, Kim & Juneau 2023)

Example PAE-reconstructed SDSS spectra

Black: observed spectra (noisy with masked regions)
Orange: denoised and inpainted reconstructed spectra

CLASS (MJD, PLATE, FIBERID)



(Böhm, Kim & Juneau 2023)

Probabilistic Autoencoder (PAE)

- Strategy to account for certain classes (e.g., QSO) being more rare than others (e.g., GALAXY)
  - Evaluation conditional probability given a label (class+subclass)

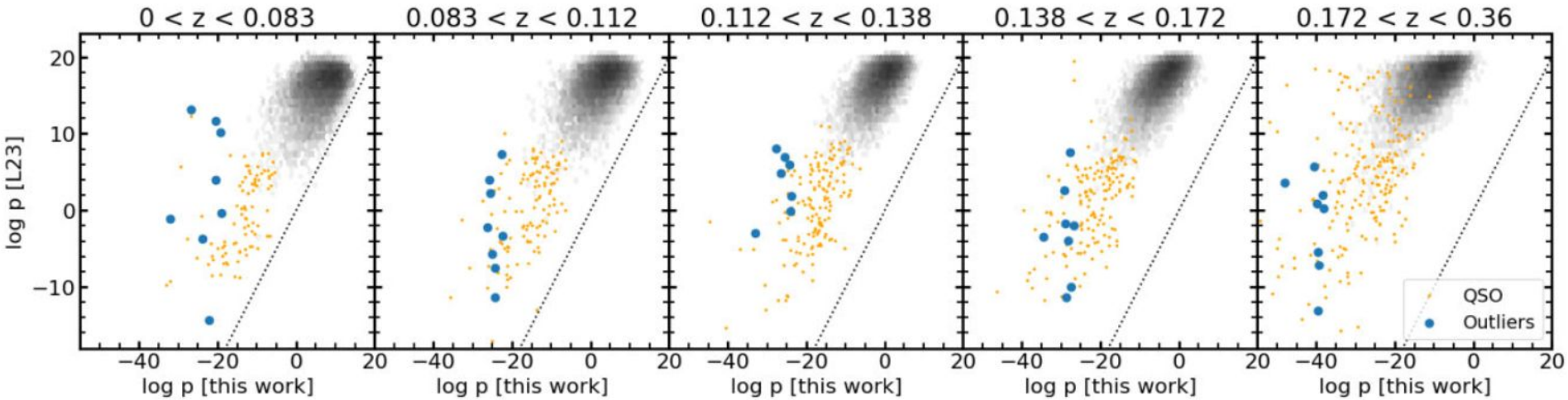
- Compare without (e.g., Pat+22, Liang+23a) and with (Böhm+23) conditional density estimation

https://github.com/VMBoehm/PytorchPAE

(Böhm & Seljak 2022; Böhm, Kim & Juneau 2023)

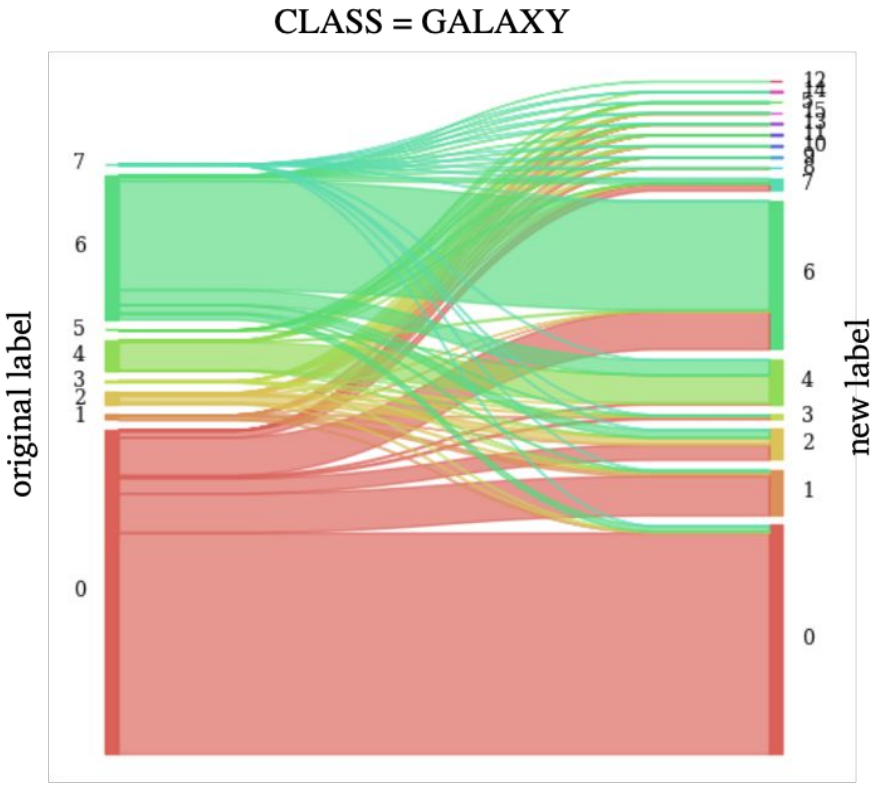
Comparing PAE with/without conditional density

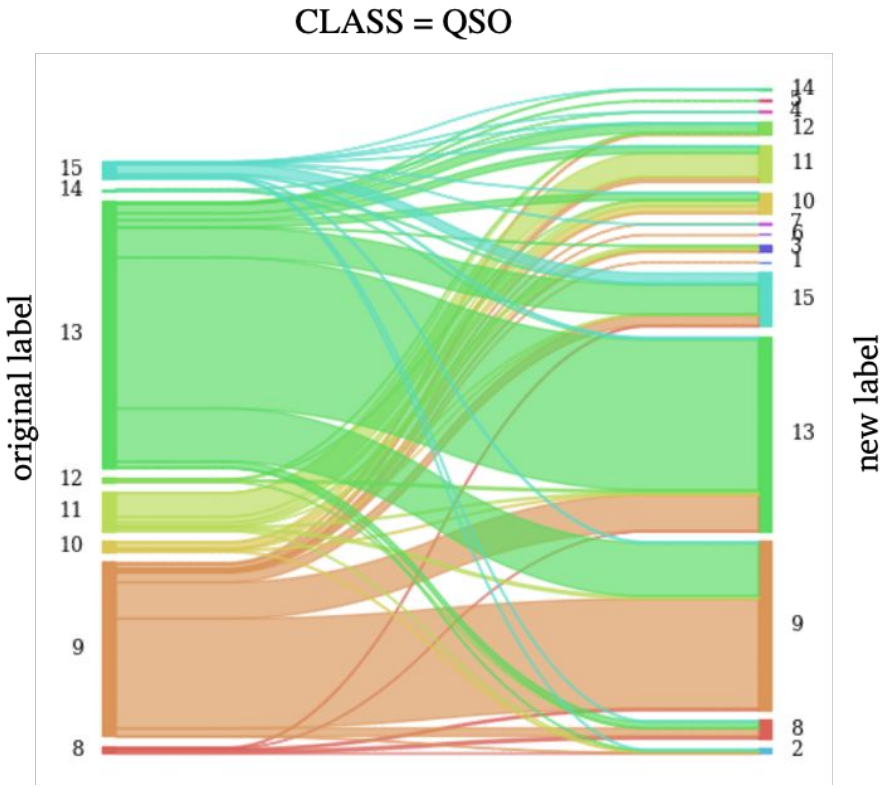
Black: quiescent galaxies
Blue: outliers in each redshift bin
Orange: QSO spectra (conditional vs. not)



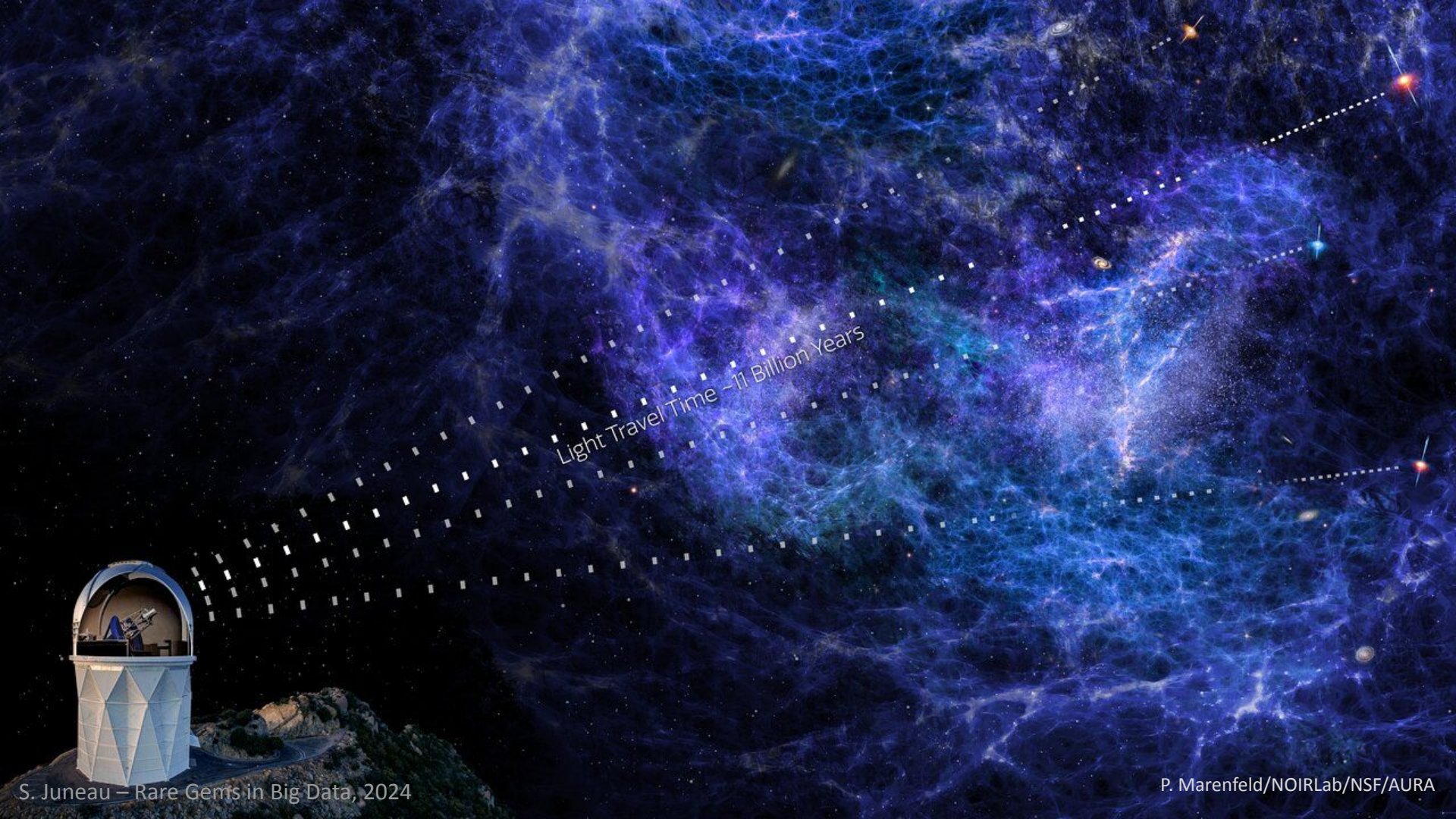
(Böhm, Kim & Juneau 2023)

Relabeling after one round of AE





(Böhm, Kim & Juneau 2023)



P. Marenfeld/NOIRLab/NSF/AURA

Embarking on a Quasar rescue mission

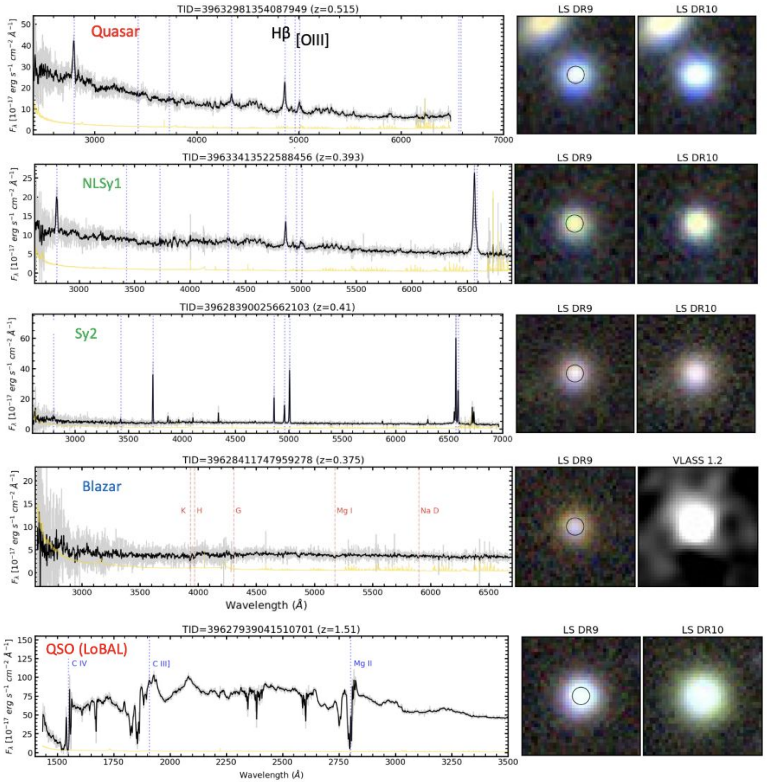
Final results

- As expected, BGS AGN targets are mostly Quasars but often with some host galaxy light → intermediate between the BGS sample and the QSO sample

- Also found: Type 2 Seyfert (Sy2) AGN, Narrow-Line Seyfert 1 (NLSy1) AGN, Blazars, BAL quasars, red quasars

DESI DR1: >15,000 unique BGS AGN targets

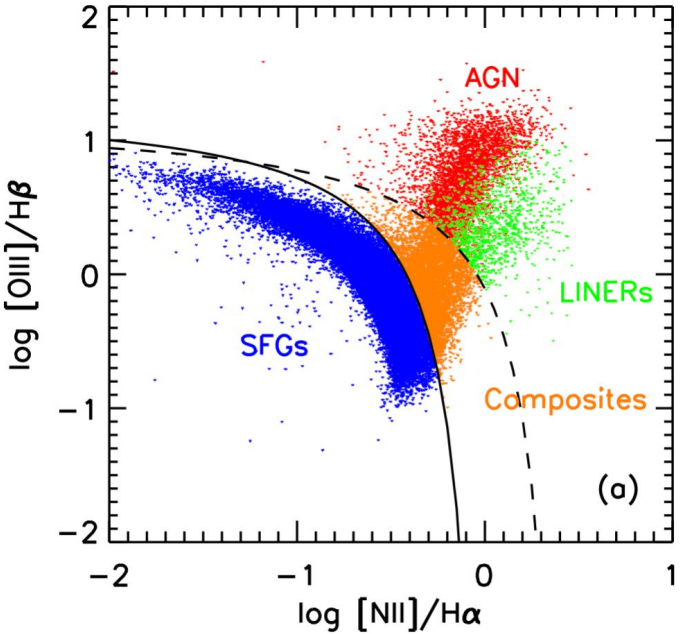
Will be part of AGN/QSO Catalog (millions)



(Juneau et al. 2024; arXiv:2404.03621)

Kinematic-Excitation (KEx) diagram

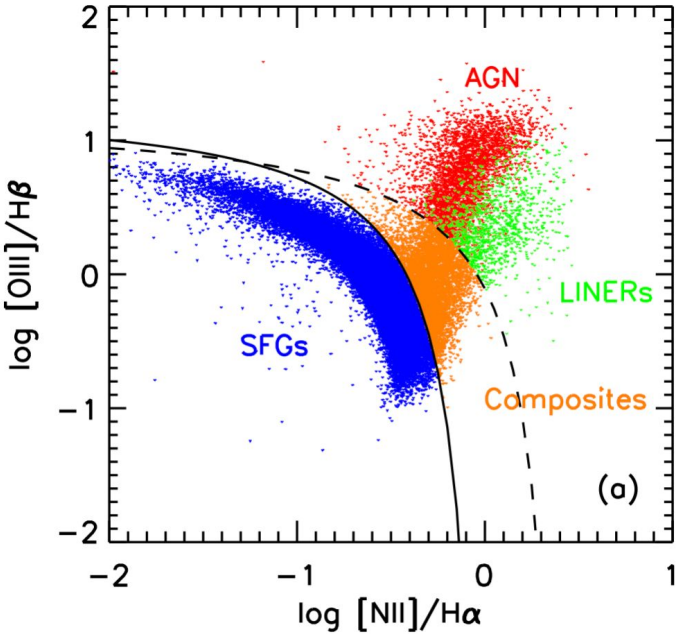
- Calibrated at low redshift with SDSS using BPT (below) and VO87 diagrams

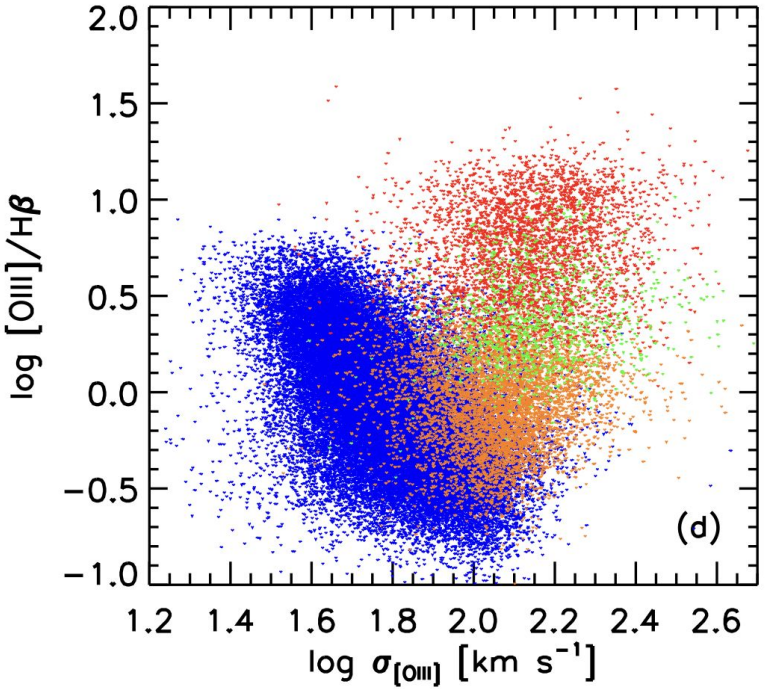


Zhang & Hao (2018; ApJ 856, 171)

Kinematic-Excitation (KEx) diagram

- Calibrated at low redshift with SDSS using BPT (left) and VO87 diagrams
- Uses the linewidth of [OIII]5007 as horizontal axis
- Can be applied to z~1 with optical spectra

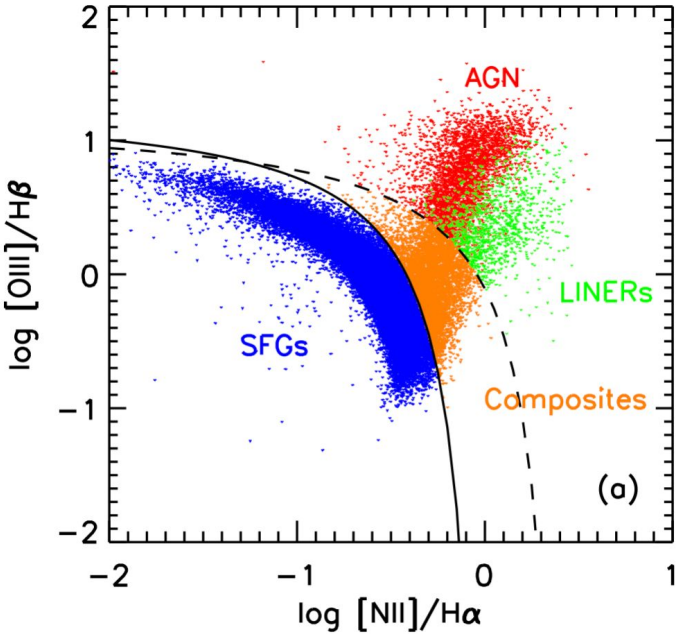


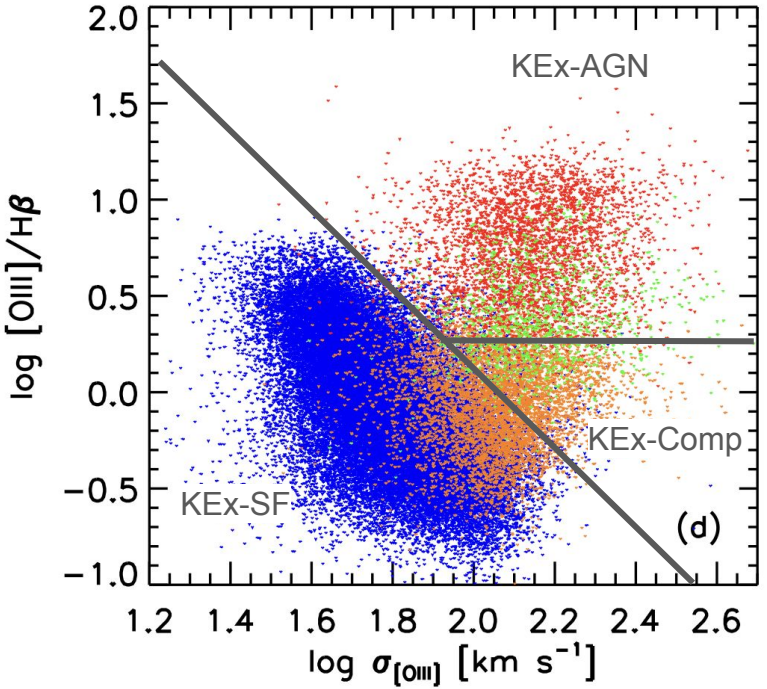


Zhang & Hao (2018; ApJ 856, 171)

Kinematic-Excitation (KEx) diagram

- Calibrated at low redshift with SDSS using BPT (left) and VO87 diagrams
- Uses the linewidth of [OIII]5007 as horizontal axis
- Can be applied to z~1 with optical spectra

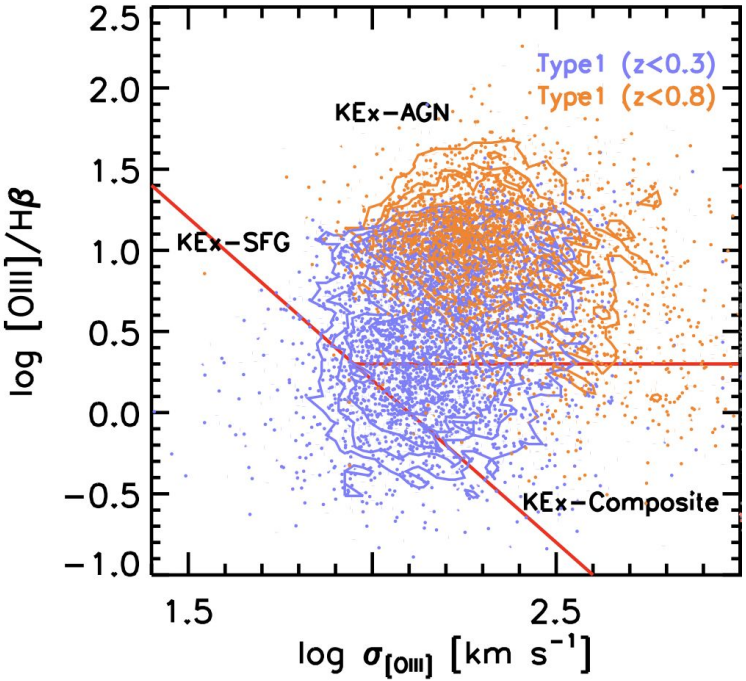




Zhang & Hao (2018; ApJ 856, 171)

Kinematic-Excitation (KEx) diagram

- But what about Type 1 (Broad-Line) AGN?



- Low-z sample (blue) reaches to lower AGN luminosities → extends into SF region
- High-z sample (red) biased toward higher AGN lum → cleaner separation

But…

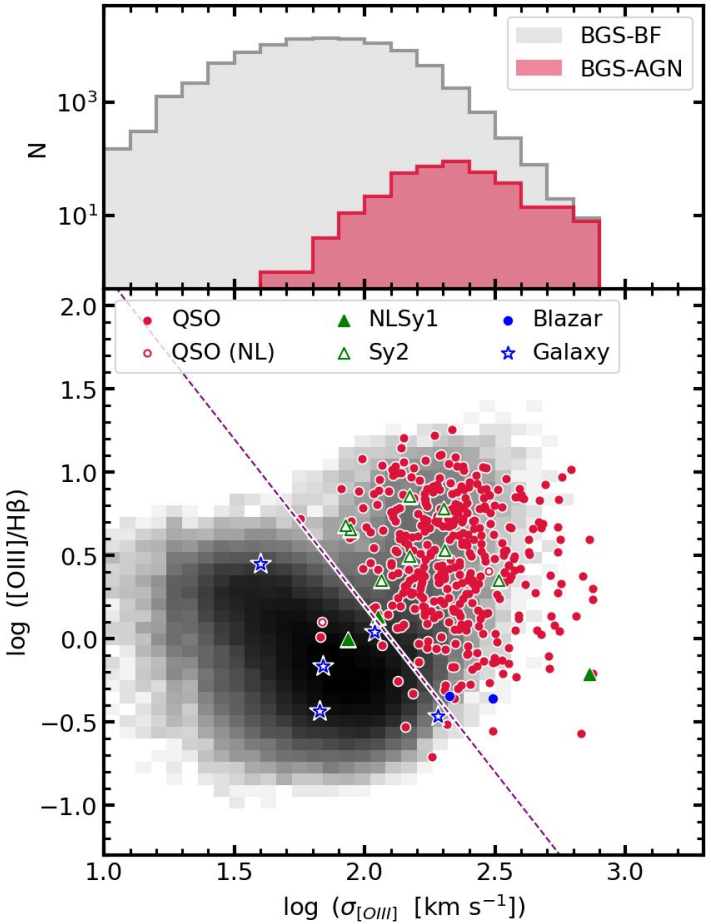
… we can split them further thanks to their broad linewidth in Hβ!

Zhang & Hao (2018; ApJ 856, 171)

KEx diagram

- [OIII]/Hβ is the ratio of the *narrow* line fluxes (caveat: fastspecfit assumes single Gaussian components for [OIII]5007 so no outflow)
- σ([OIII]) in km/s

→ no distinction b/w Type 1 vs. Type 2 AGN

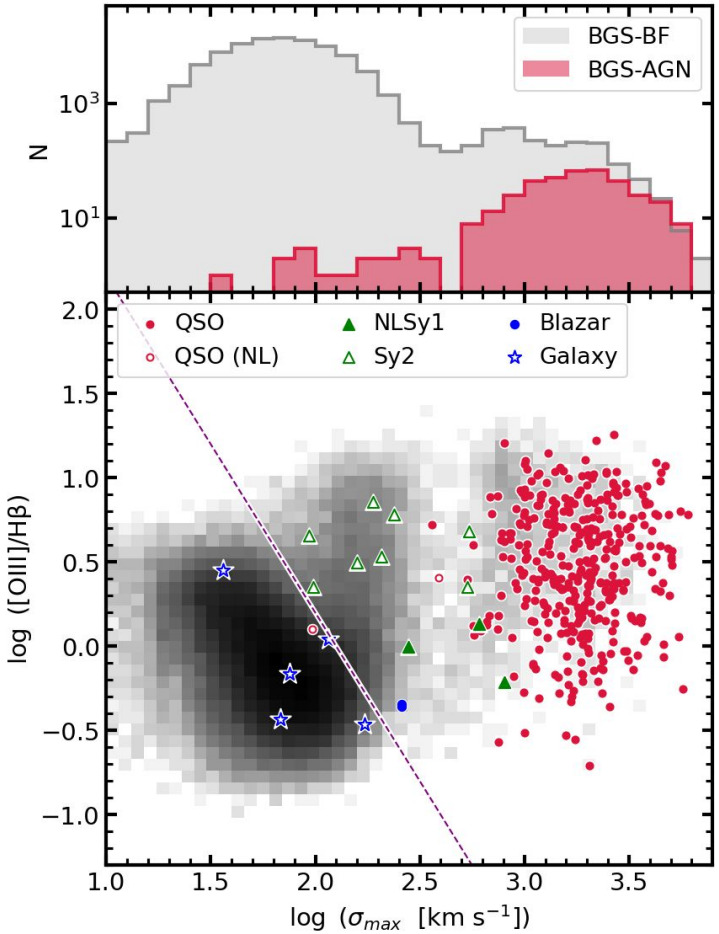


(Juneau et al. 2024; arXiv:2404.03621)

Modified KEx diagram

- [OIII]/Hβ is the ratio of the *narrow* line fluxes (caveat: fastspecfit assumes single Gaussian components for [OIII]5007 so no outflow)
- Replace σ([OIII]) with max{σ(MgII, Hβ, Hα)}

→ better visualize the Type 1 vs. Type 2 AGN dichotomy & relative proportions



(Juneau et al. 2024; arXiv:2404.03621)

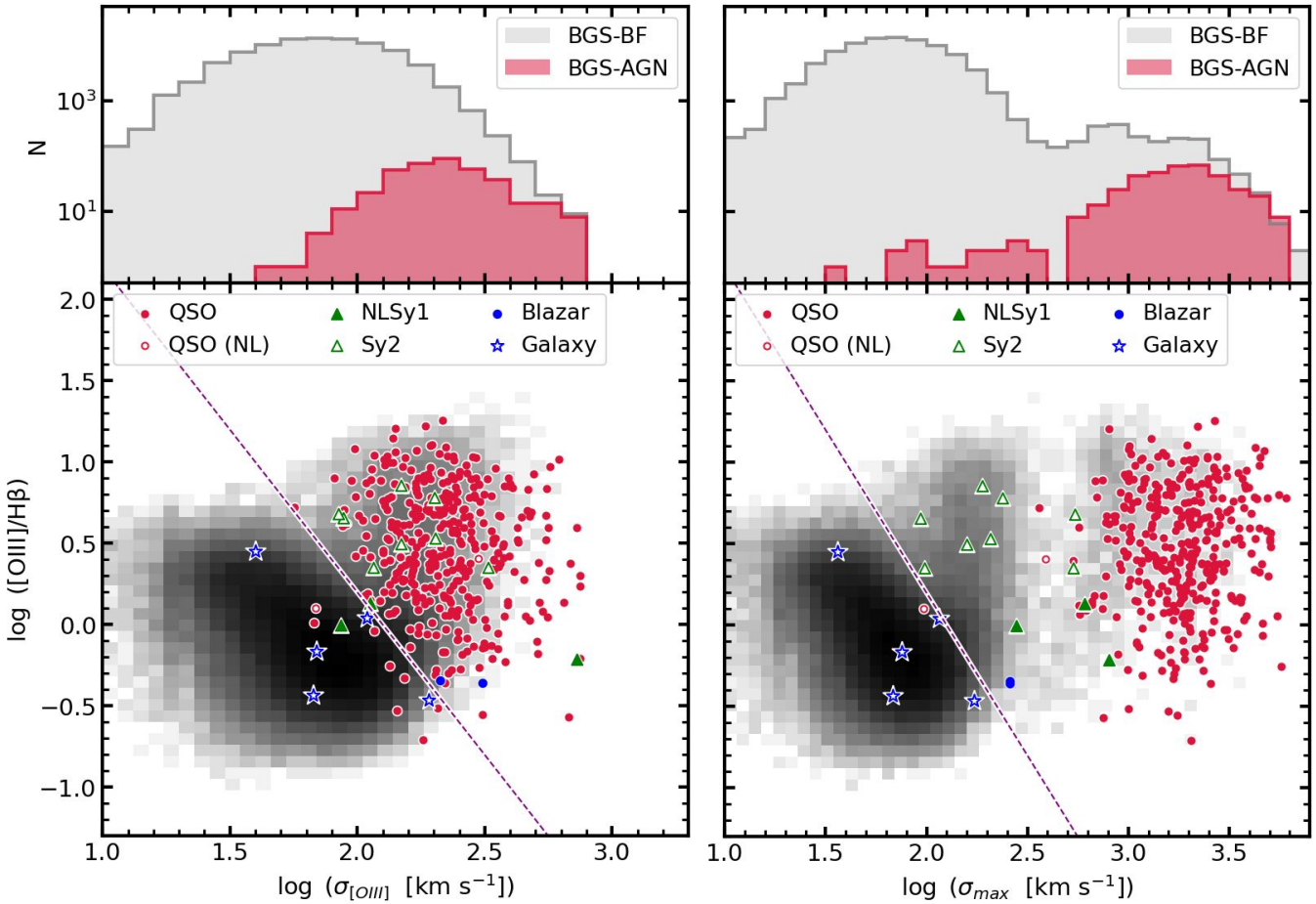
Modified KEx diagram

Side-by-side comparison

Left: σ([OIII])

Right: max{σ (MgII,Hβ,Ha)}

[OIII]/Hβ ratio is for the narrow lines only (i.e., excluding broad Hβ)



(Juneau et al. 2024; arXiv:2404.03621)