A Machine-Learning Framework for Optical–X-ray AGN Studies

Building a unified spectral representation for prediction, anomaly detection and population studies

Noemi Lery Borrelli

Mara Salvato, Massimo Brescia , Johannes Buchner, Raphael Shirley, William Roster, Brivael Laloux, Hattie Starck

Napoli, Italy

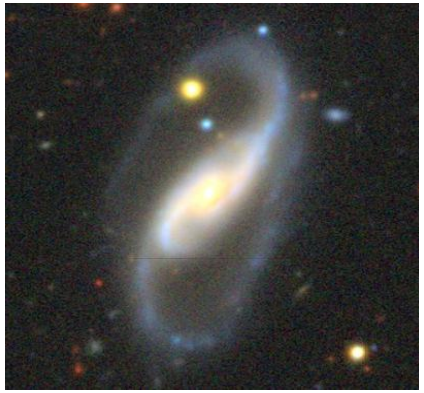
Garching, Germany

Università degli Studi di Napoli Federico II

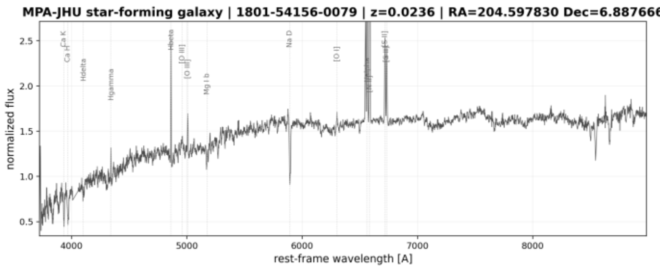
Max-Planck-Institut für extraterrestrische Physik

What is an AGN?

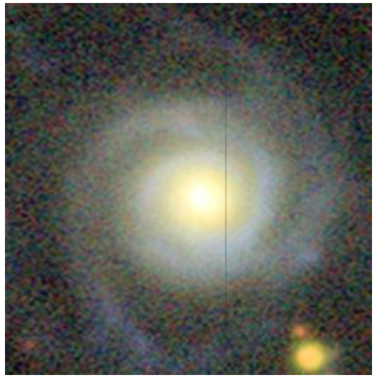
Normal galaxy



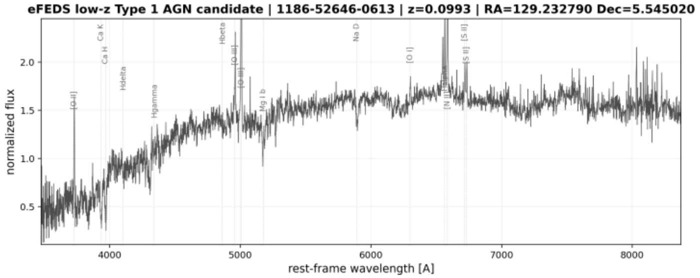
Host-dominated emission



Seyfert (Active galaxy)

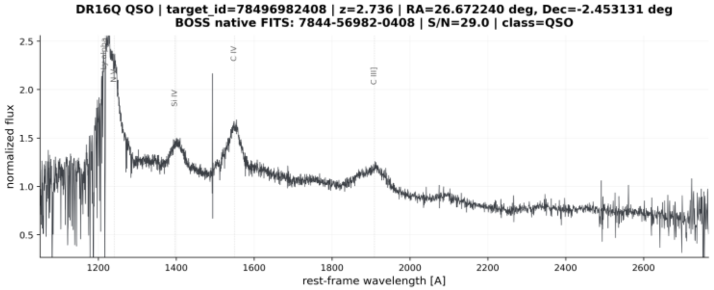


Host + nuclear contribution



Quasar (Active galaxy)

Nucleus-dominated emission



Increasing nuclear dominance relative to the host galaxy

- Active galactic nuclei (or AGN) are supermassive black holes accreting matter at the centres of galaxies.
- The host-to-nucleus ratio is a continuum, not a binary choice — and it varies with luminosity, obscuration and viewing geometry

Goals

Represent galaxies of every type, Seyferts, and quasars, within a single machine-learning model

1. Learn a unified optical spectral representation *(Latent Space)*

Train a self-supervised ML model on galaxies, AGN, and quasars

2. Freeze the encoder

Use the latent embeddings as compact spectral features

Train downstream models

- **Predict X-ray properties,**
- **Classify optical sources**

Latent space analysis

- **Explore source populations**
- **Anomaly detection** *(identify rare sources, transient spectral states, and unusual optical–X-ray combinations)*

One unified optical representation for several scientific applications

Dataset: *eROSITA Final Equatorial Depth Survey (eFEDS)*

eFEDS is the science sample; larger optical surveys teach the representation what spectra can look like.

Why eFEDS?

- **12,847 X-ray-selected AGN** with optical spectra
- Reliable spectroscopic redshifts up to $z \sim 5.8$
- Wide range of AGN types

For the same sources, we have:

- **Optical spectroscopy**
- **X-ray information**
- **Multi-wavelength context**
- **Host galaxy properties**

eFEDS: Aydar et al. 2025, 2026; Brunner et al. 2022; Liu et al. 2022; Salvato et al. 2022

MPA-JHU: Brinchmann et al. 2004; Kauffmann et al. 2003

DR16Q: Lyke et al. 2020

Dataset: *eROSITA Final Equatorial Depth Survey (eFEDS)*

eFEDS is the science sample; larger optical surveys teach the representation what spectra can look like.

Why eFEDS?

- **12,847 X-ray-selected AGN** with optical spectra
- Reliable spectroscopic redshifts up to $z \sim 5.8$
- Wide range of AGN types

For the same sources, we have:

- **Optical spectroscopy**
- **X-ray information**
- **Multi-wavelength context**
- **Host galaxy properties**

Why is pre-training needed?

- Broad variety of spectral populations
- Sample not large enough to train the model from scratch

eFEDS is used after the model has learned the main optical spectral regimes.

eFEDS: Aydar et al. 2025, 2026; Brunner et al. 2022; Liu et al. 2022; Salvato et al. 2022

MPA-JHU: Brinchmann et al. 2004; Kauffmann et al. 2003

DR16Q: Lyke et al. 2020

Dataset: *eROSITA Final Equatorial Depth Survey (eFEDS)*

eFEDS is the science sample; larger optical surveys teach the representation what spectra can look like.

Why eFEDS?

- **12,847 X-ray-selected AGN** with optical spectra
- Reliable spectroscopic redshifts up to $z \sim 5.8$
- Wide range of AGN types

For the same sources, we have:

- **Optical spectroscopy**
- **X-ray information**
- **Multi-wavelength context**
- **Host galaxy properties**

Why is pre-training needed?

- Broad variety of spectral populations
- Sample not large enough to train the model from scratch

eFEDS is used after the model has learned the main optical spectral regimes.

Pre-training sources

MPA–JHU
~900k SDSS galaxies
(local galaxies + narrow AGN)

DR16Q
~ 750k confirmed quasars
broad-line AGN + high-z

↓

PRE-TRAINING

↓

COMMON OPTICAL LATENT

↓

eFEDS ADAPTATION + X-RAY

↓

prediction | anomalies | population studies

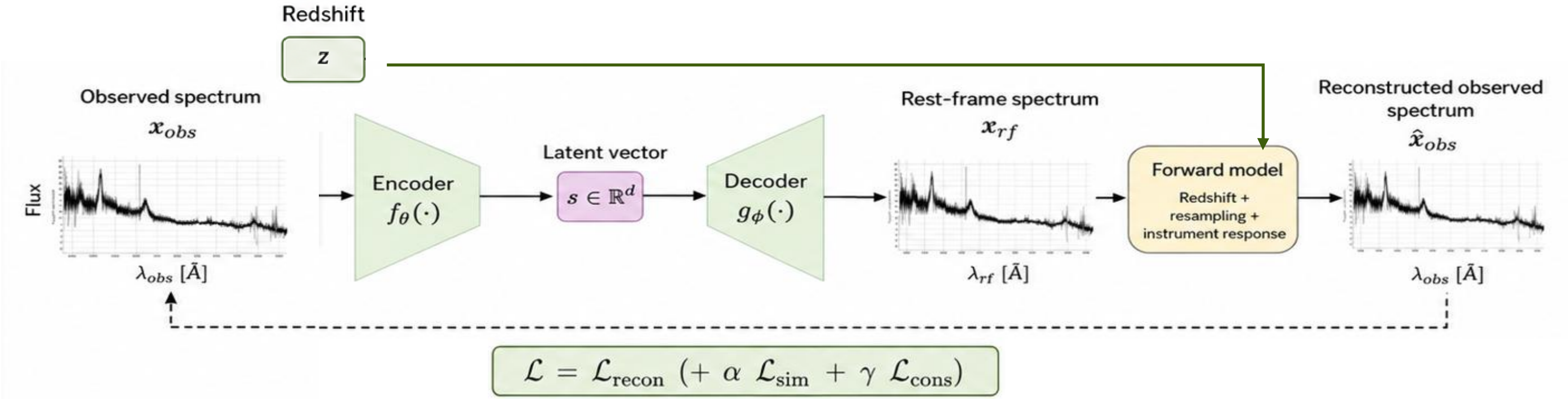
eFEDS: Aydar et al. 2025, 2026; Brunner et al. 2022; Liu et al. 2022; Salvato et al. 2022

MPA-JHU: Brinchmann et al. 2004; Kauffmann et al. 2003

DR16Q: Lyke et al. 2020

SPENDER as a starting architecture

Goal: *encode intrinsic spectral properties while reducing redshift-driven structure in the latent space.*



Core idea

- **Input:** observed-frame spectrum x_{OBS}
- **Encoder:** compresses it into a latent vector s
- **Decoder:** reconstructs an intrinsic **rest-frame** spectrum x_{RF}
- **Forward model:** applies known z, resampling and instrument response to compare with the observed spectrum

Redshift is handled explicitly after decoding, rather than being left entirely to the latent representation.

1. Similarity loss

Links latent distance to rest-frame spectral distance

$$| x_1' - x_2' | \Rightarrow | s_1 - s_2 |$$

→ encourages spectra of **physically similar galaxies** to cluster in latent space.

2. Consistency loss

Penalises latent changes under redshift augmentation δz

$$s(x) \approx s(x_{aug})$$

→ Forces the encoder to be invariant to z

SPENDER architecture: *Melchior et al. 2022, Liang et al. 2023*

One model, many spectral regimes: challenges and design choices

Main Challenges

1. Heterogeneous spectral populations

2. Wide redshift range

3. Different instruments

One model, many spectral regimes: challenges and design choices

Challenge

1. Heterogeneous spectral populations

Strategy

1. Progressive Curriculum Learning

Gradual introduction of the spectral complexity

Phase 1

MPA-JHU · SDSS

Galaxy warm-start: broad diversity of galaxy spectra

→

Phase 2

DR16Q · BOSS

Add broad-line quasars, high z and UV spectral features

→

Phase 3

MPA-JHU AGN · SDSS

Bridge host-dominated galaxies and active nuclei

→

Phase 4

Mixed Training

Balanced: Galaxies + AGN + QSO

→

eFEDS

Adapt the optical representation to the X-ray selected sample

Each phase resumes from the previous last checkpoint.

One model, many spectral regimes: challenges and design choices

Challenge

1. Heterogeneous spectral populations

Strategy

1. Progressive Curriculum Learning

Gradual introduction of the spectral complexity

Phase 1
MPA-JHU · SDSS
Galaxy warm-start: broad diversity of galaxy spectra

→

Phase 2
DR16Q · BOSS
Add broad-line quasars, high z and UV spectral features

→

Phase 3
MPA-JHU AGN · SDSS
Bridge host-dominated galaxies and active nuclei

→

Phase 4
Mixed Training
Balanced: Galaxies + AGN + QSO

→

eFEDS
Adapt the optical representation to the X-ray selected sample

Each phase resumes from the previous last checkpoint.

2. Wide redshift range

- The same rest-frame features enter different observed windows

2. Wide rest-frame modelling

- Common rest-frame wavelength grid
- Adaptive normalization windows

3. Different instruments

- SDSS and BOSS have different sampling and wavelength coverage
- A model can reconstruct both while still encoding the instrument

3. Instrument homogenisation

- Old run: SDSS and BOSS native grids → latent space partly instrument-driven
- Current run: BOSS spectra resampled onto the SDSS observed grid

Reconstructions

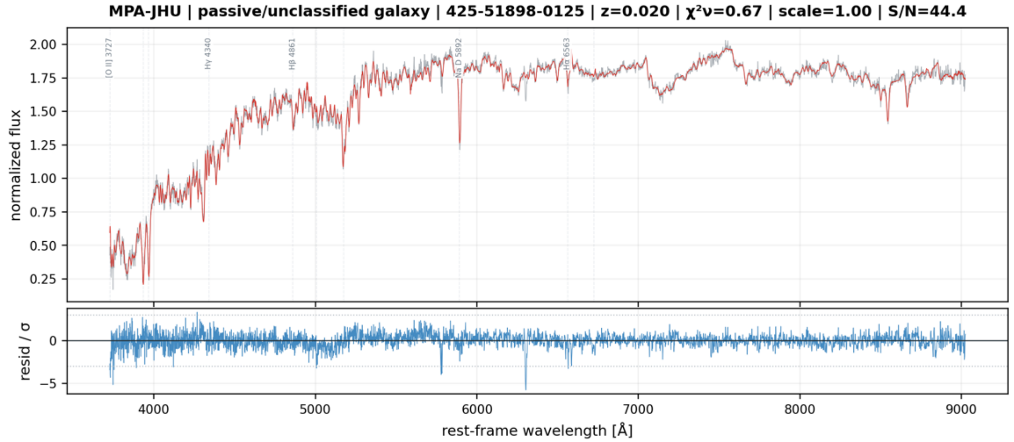
1. A single model reconstructs all major spectral populations

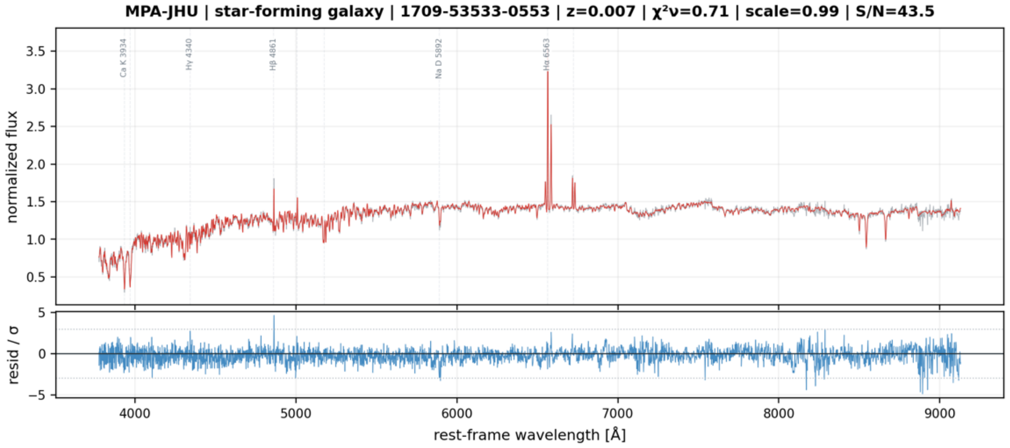
Galaxies, AGN and quasars from $z\sim 0$ to $z \sim 6$

Pre-training

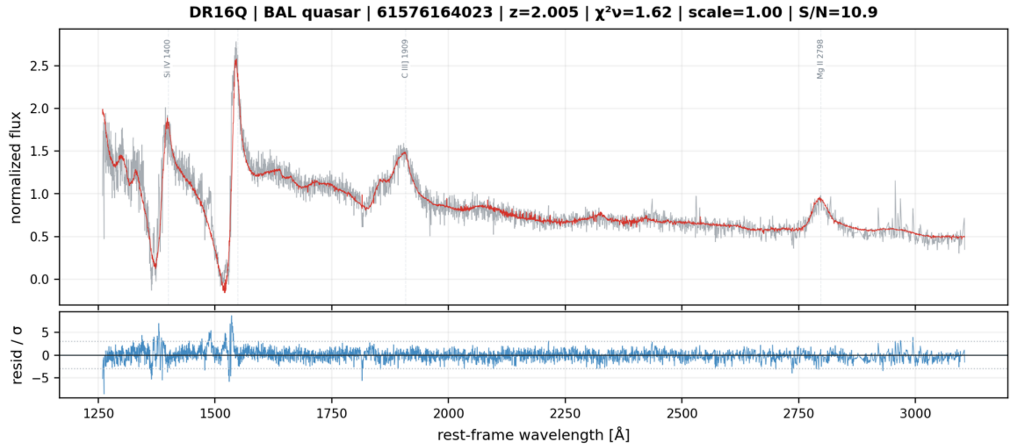
observed spectrum — reconstruction

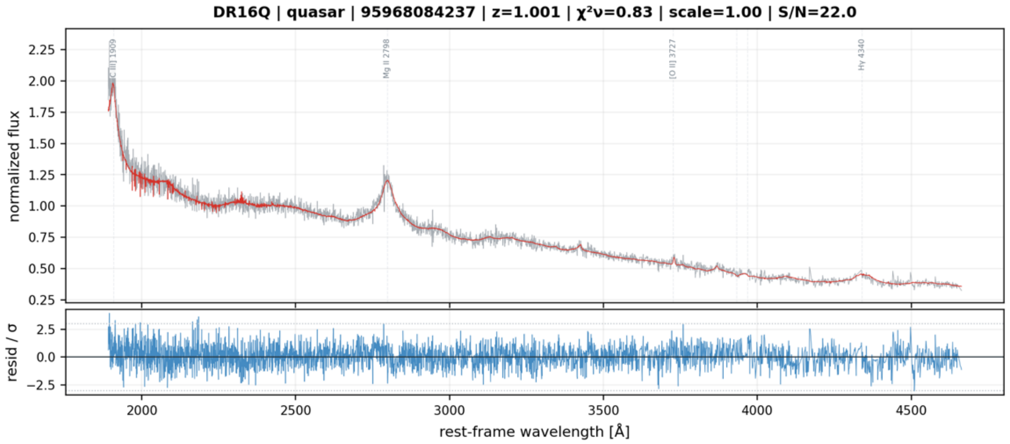
MPA-JHU





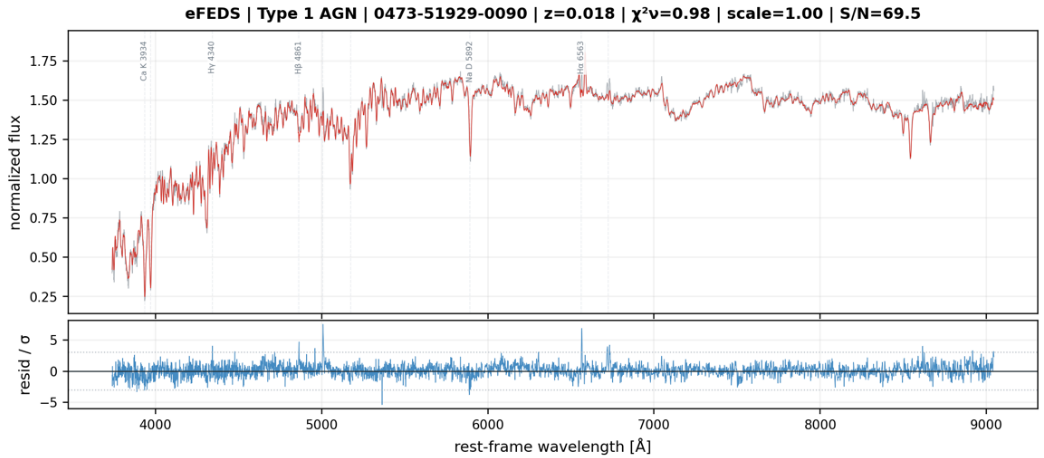
DR16Q



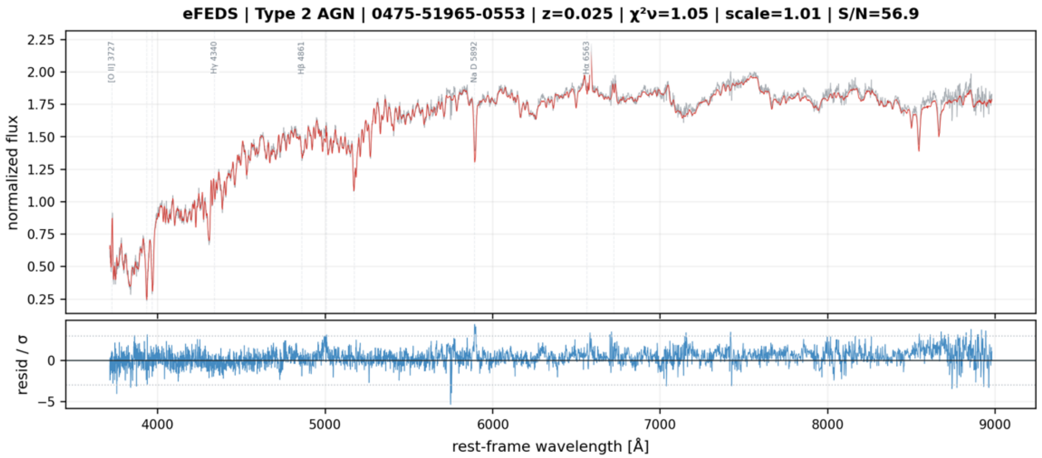


Reconstructions: eFEDS

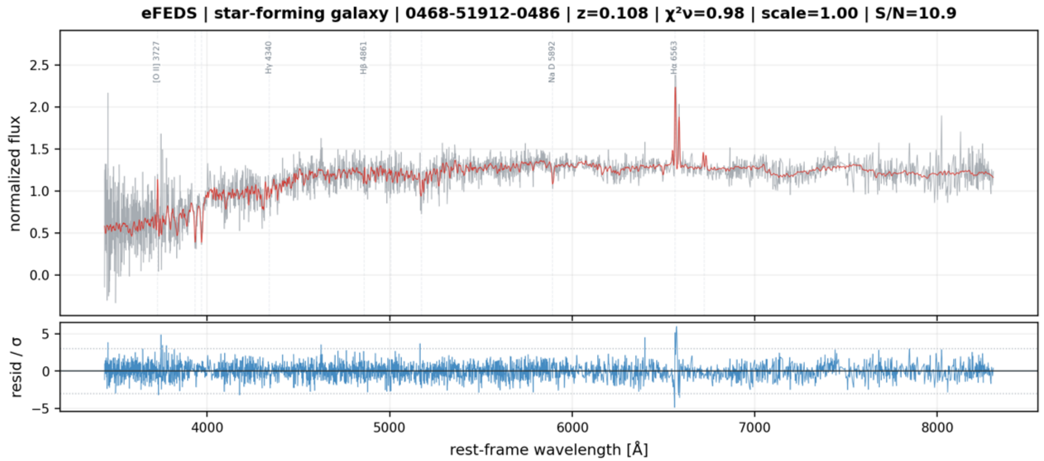
Type 1 AGN



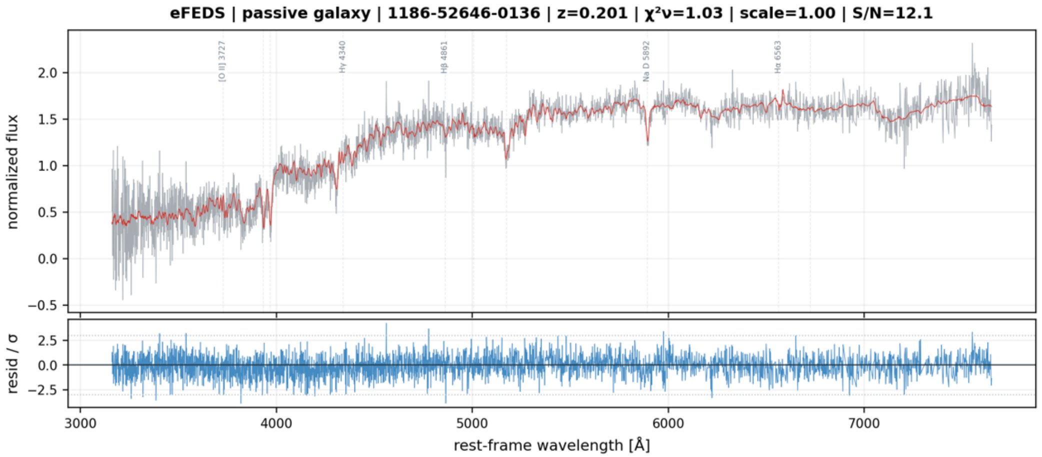
Type 2 AGN



Star forming

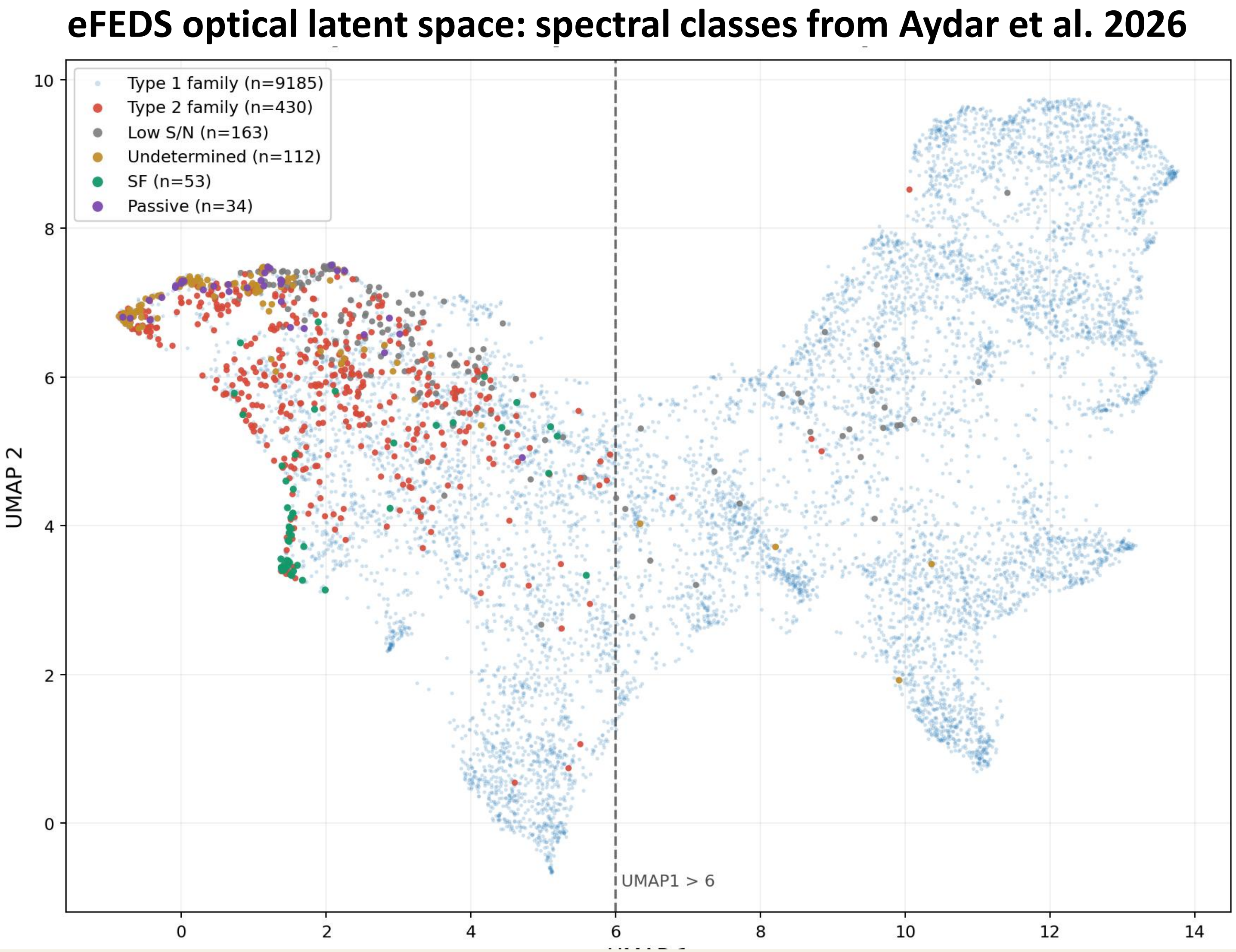


Passive



*A **single model** can reconstruct very **different spectral polulations**, from low-redshift galaxies to high-redshift quasars*

Optical latent space separates AGN spectral families

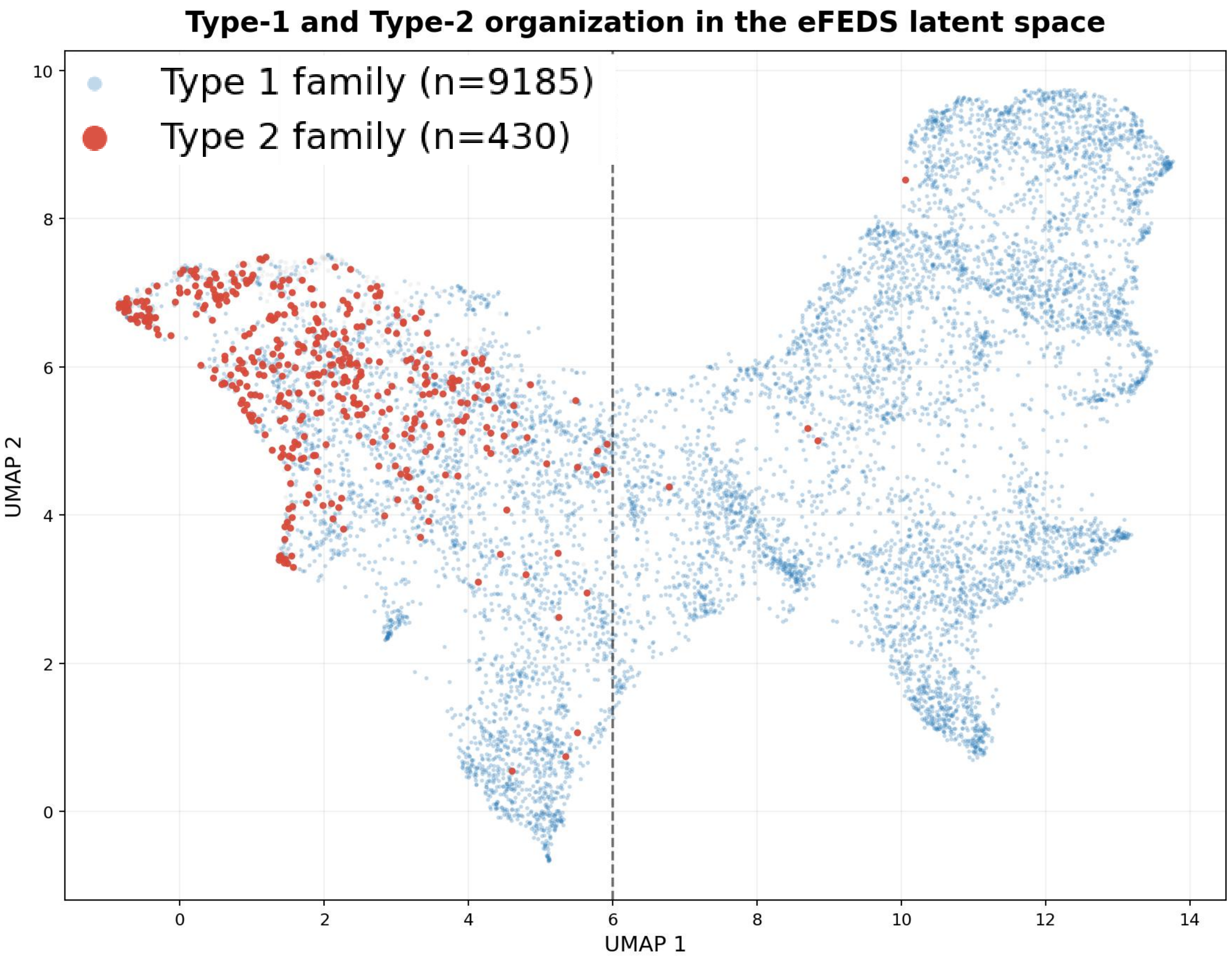


- eFEDS optical spectra encoded into a 24-dimensional latent vector via a spectral autoencoder.
- UMAP projection of the latent space reveals a **structured organization by optical spectral class** (Aydar et al. 2026 classification).
- The right wing (UMAP1 > 6) is strongly dominated by Type-1 / QSO-like sources.

➢ This separation was not imposed by labels during training — **classes are used only afterward**, to interpret the structure.

The latent space shows a clear physical organization

The Type-1 wing is almost pure



Type 2 family highlighted (red) against Type 1 (blue) and other/uncertain (grey) — dashed line marks UMAP1 = 6.

Right wing (UMAP1 > 6)

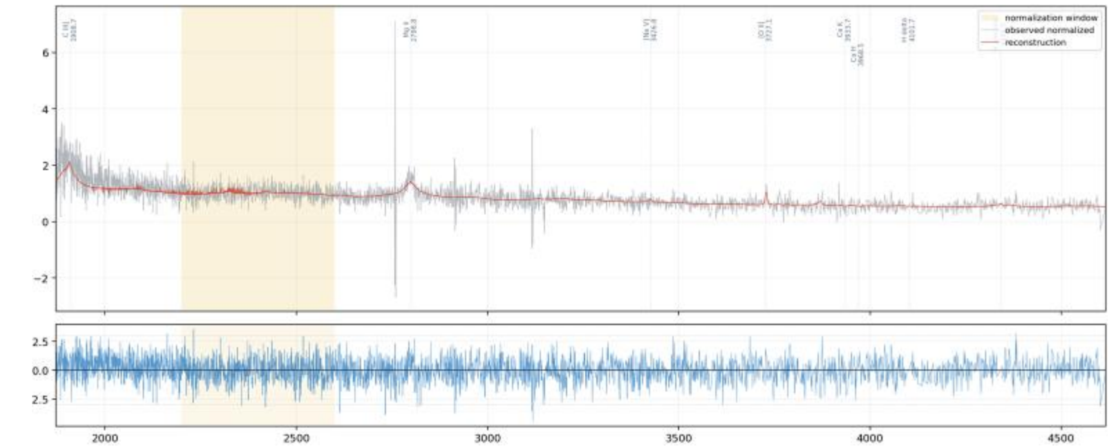
Type 1 family	**99.4% (5,902)**
QSO-like (SDSS pipeline)	**99.3% (5,893)**
Median redshift	**z ≈ 1.48**
Type 2 candidates	**0.07% (4)**
Non-Type-1 total	**0.6% (34)**

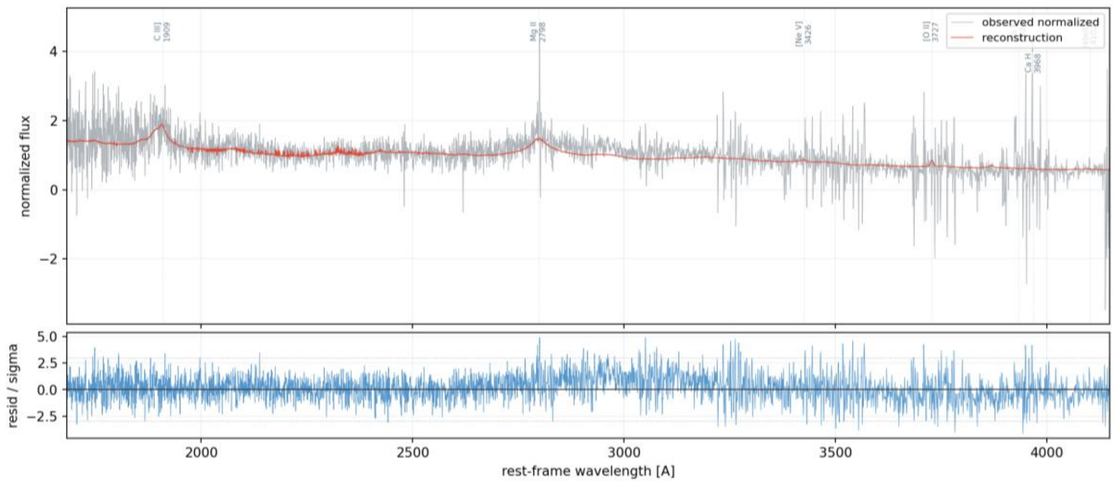
Among the non-Type-1 sources

• Low S/N	26
• Undetermined	4
• Type 2 candidates	4
• SF / passive	0

Most non-Type-1 objects in this region are Low S/N or Undetermined

Two “Undetermined” sources identified as quasar candidates





eFEDS J092255.8+012154

Undetermined — S/N=3.9, z=1.02

Narrow lines used for BPT classification may be too weak or fall outside the observed coverage — but the spectrum shows a broad Mg II line and a blue continuum, the signature of a quasar.

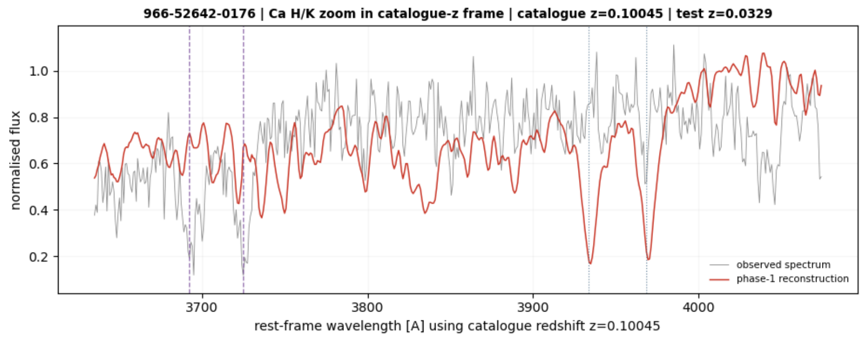
eFEDS J093519.1-001242

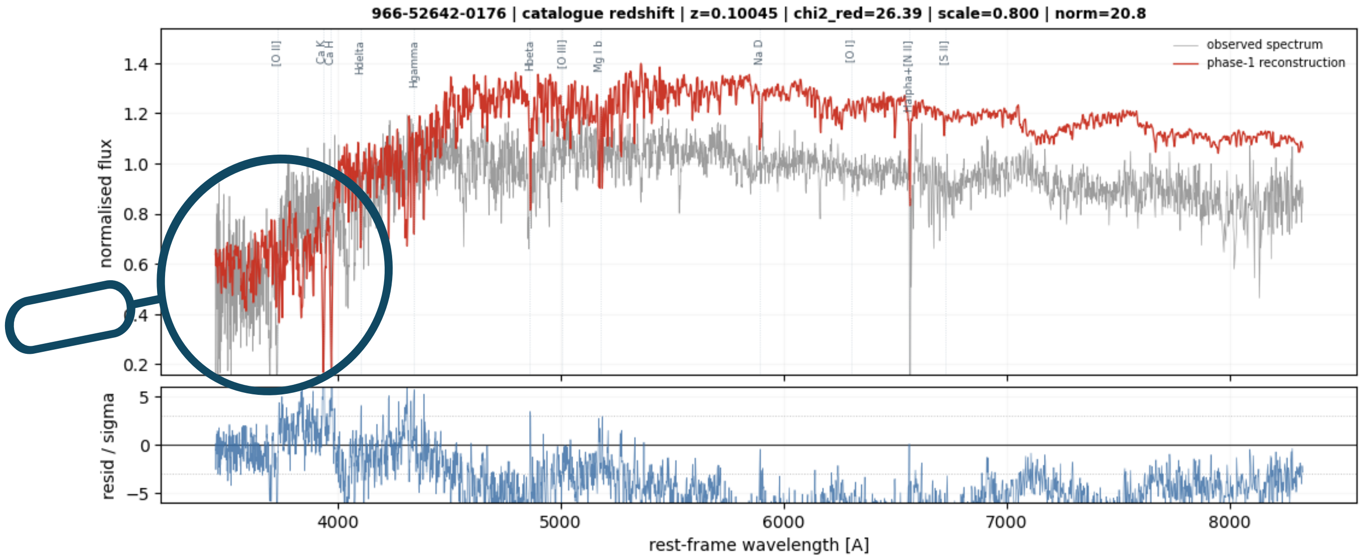
Undetermined — S/N=4.3, z=1.25

C III] and Mg II both show a clear excess above the continuum despite the noise.

The latent space can provide useful ***classification information*** *when standard optical diagnostics are limited.*

Reconstructions: *Inspection of the problematic reconstructions*



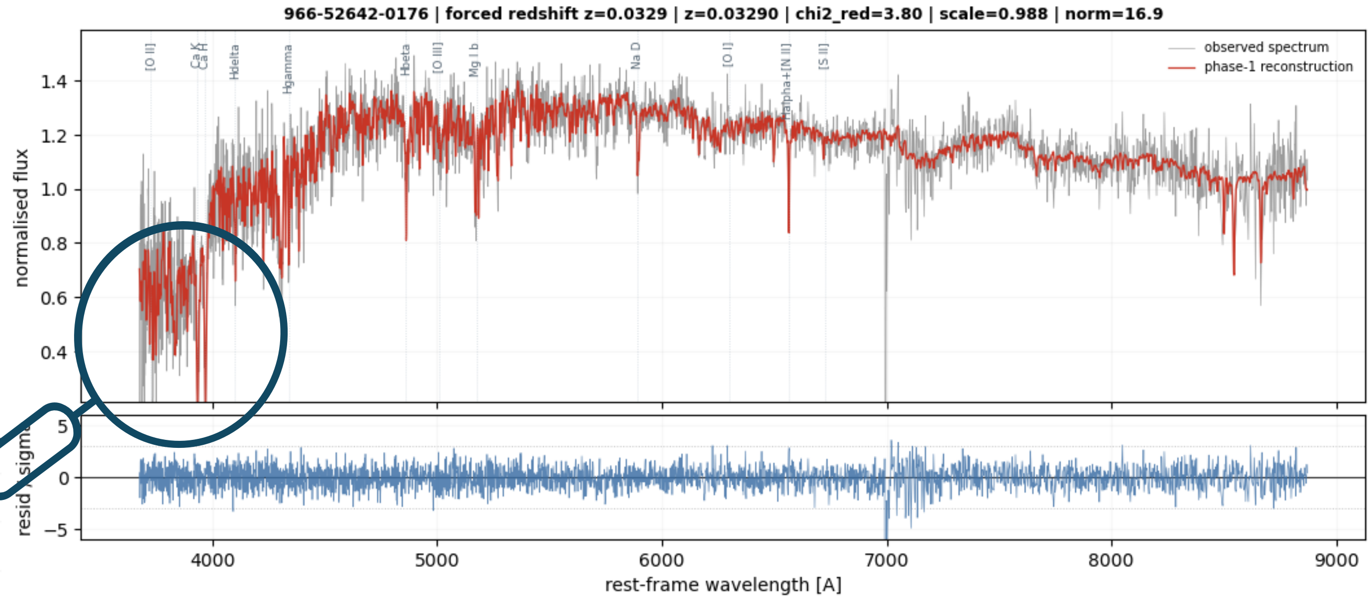


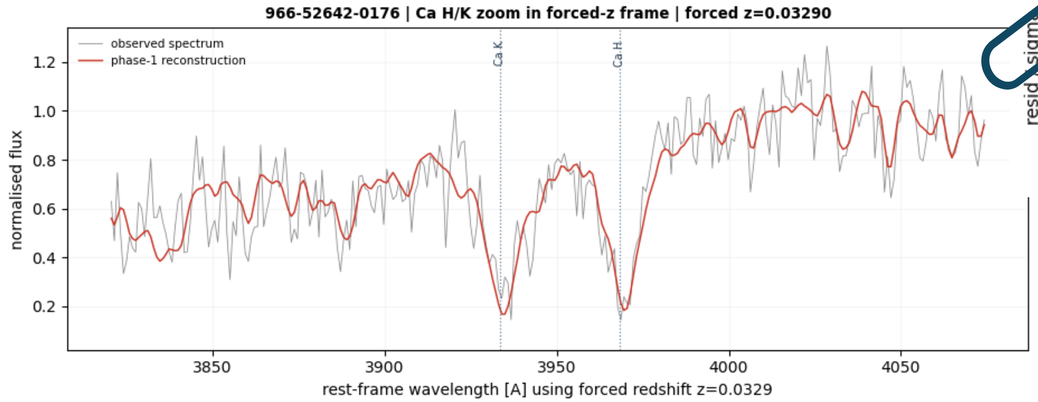
Wrong Redshift

Catalogue redshift is wrong → Bad reconstruction

Reconstructions: *Inspection of the problematic reconstructions*

What happens if we give the model the correct redshift?

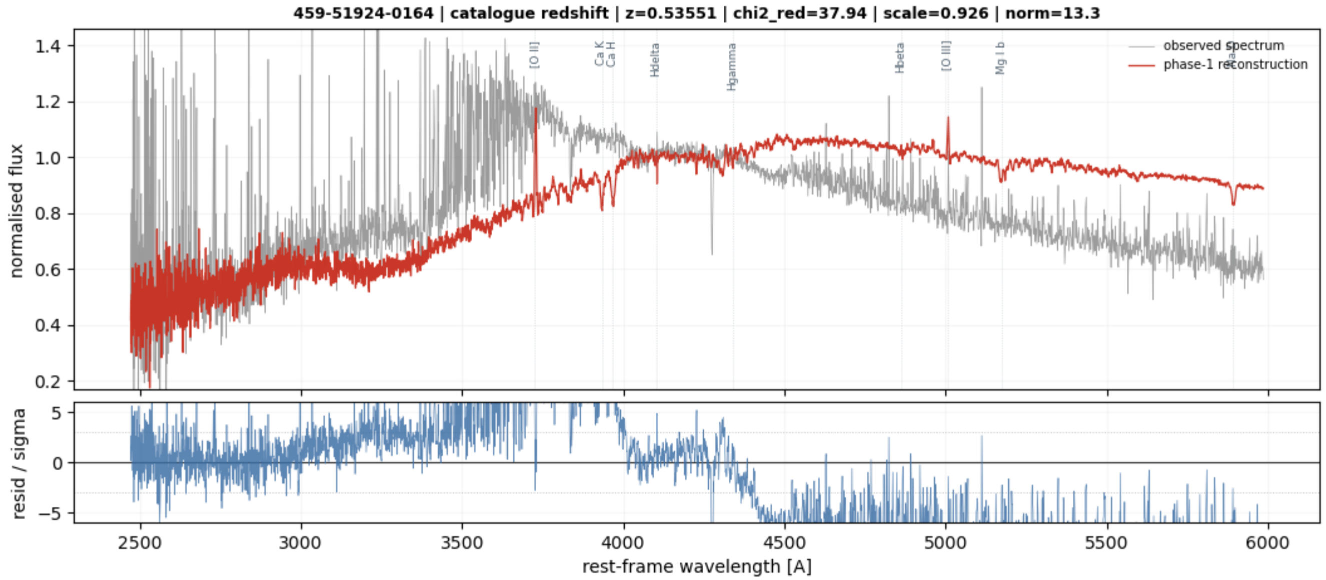




The same model reconstructs the spectrum well once the redshift is corrected

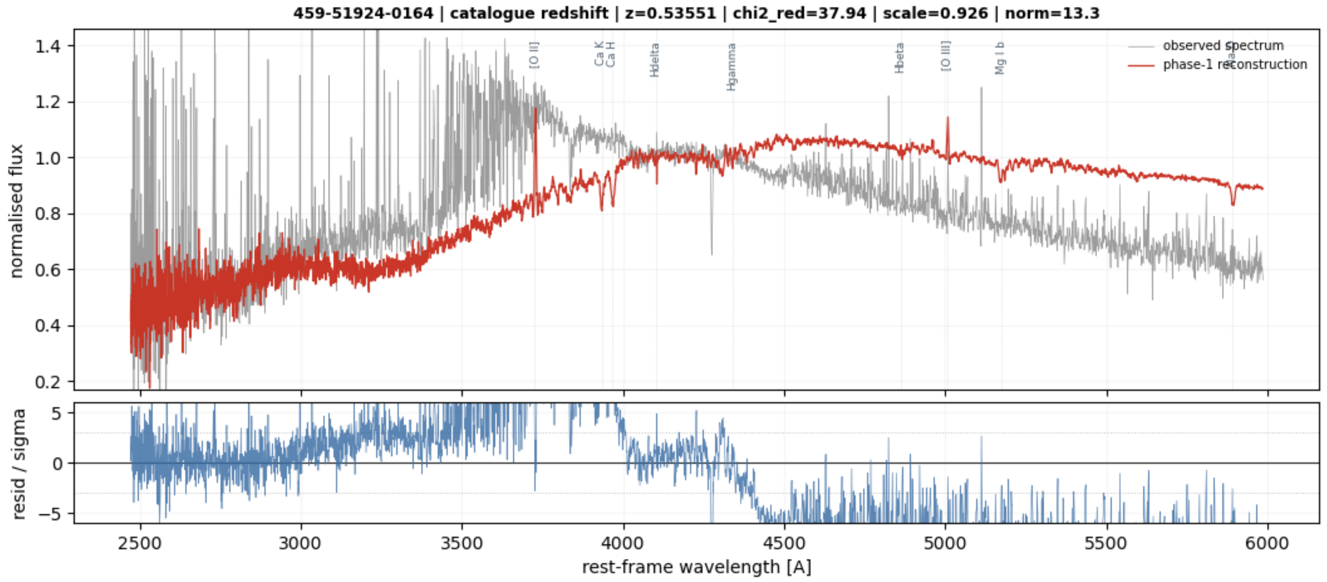
Problematic reconstructions can identify incorrect catalogue redshifts

Reconstructions: *Inspection of the problematic reconstructions*



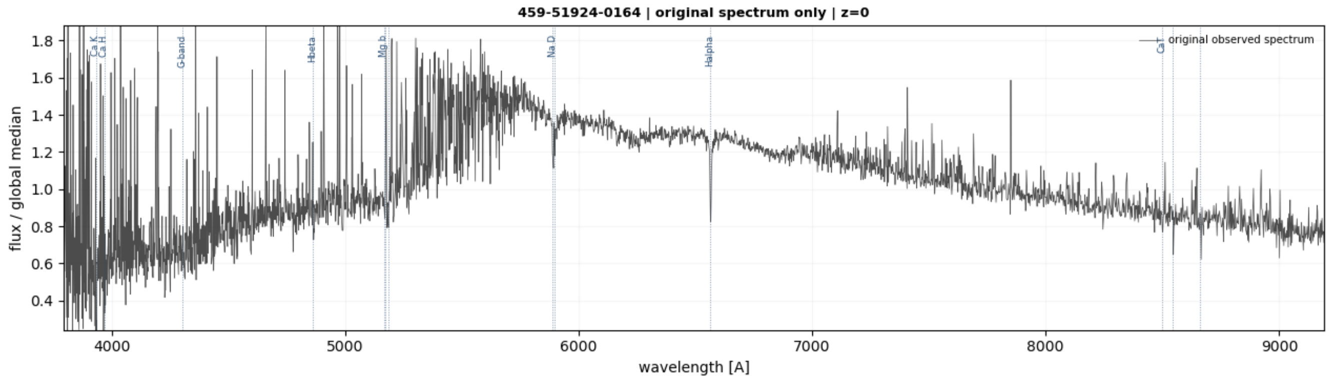
- This source is classified as a galaxy at $z = 0.53$, but its spectral shape and line pattern doesn't look galaxy-like at that redshift.

Reconstructions: *Inspection of the problematic reconstructions*



- This source is classified as a galaxy at $z = 0.53$, but its spectral shape and line pattern doesn't look galaxy-like at that redshift.

- When the spectrum is interpreted at $z = 0$, absorption features coincide with Na D and Hα, suggesting that the source may be a stellar contaminant



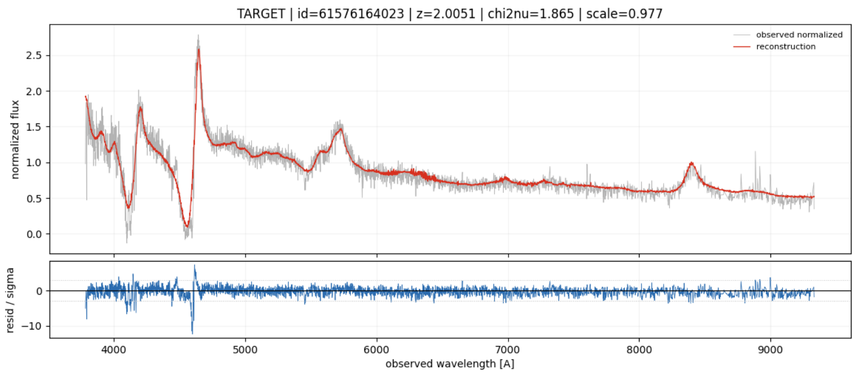
Reconstruction failures can flag catalog contaminants and misclassifications

Pattern Recognition

- **Nearest neighbours of a BAL-quasar target in the phase-2 latent space (DR16Q)**

Broad Absorption Line quasars (BAL quasars): *a class characterised by broad, blueshifted absorption features produced by high-velocity outflows.*

- Selection one target BAL-quasar

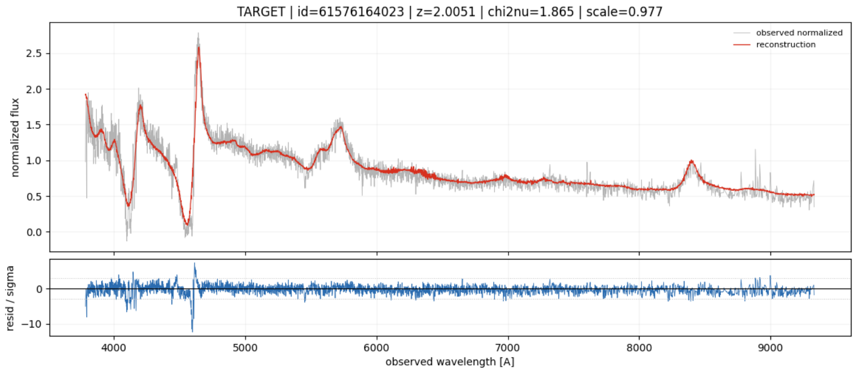


- Computed its nearest KNN in the final latent space
- Inspected neighbour spectra

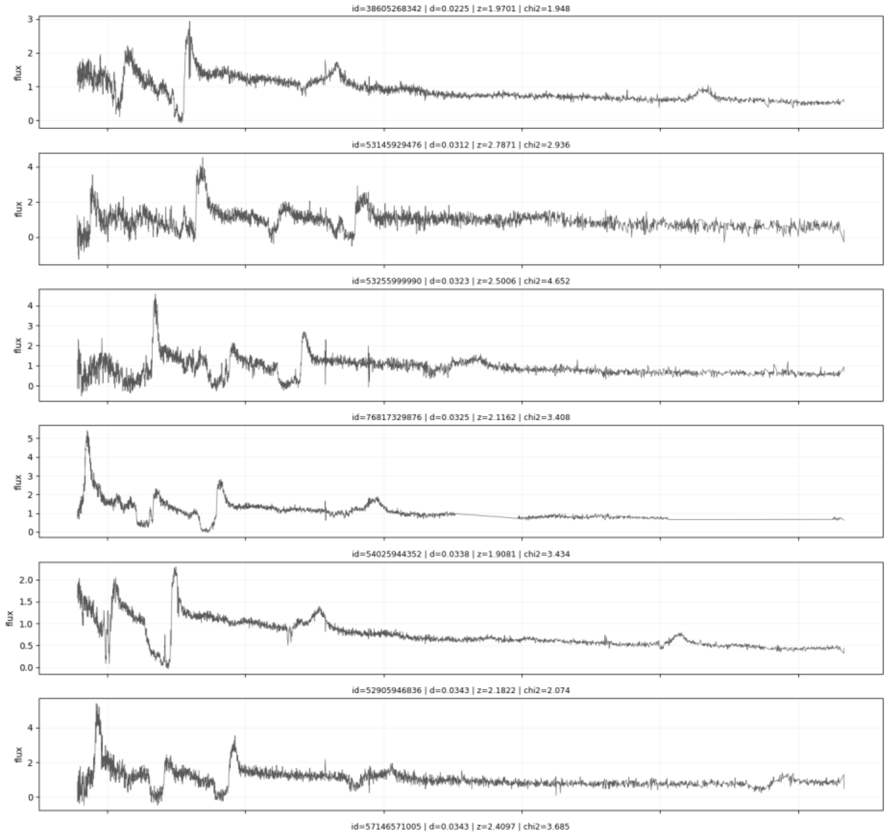
Pattern Recognition

- **Nearest neighbours of a BAL-quasar target in the phase-2 latent space (DR16Q)**

  - Selection one target BAL-quasar in phase 2



KNN



Results

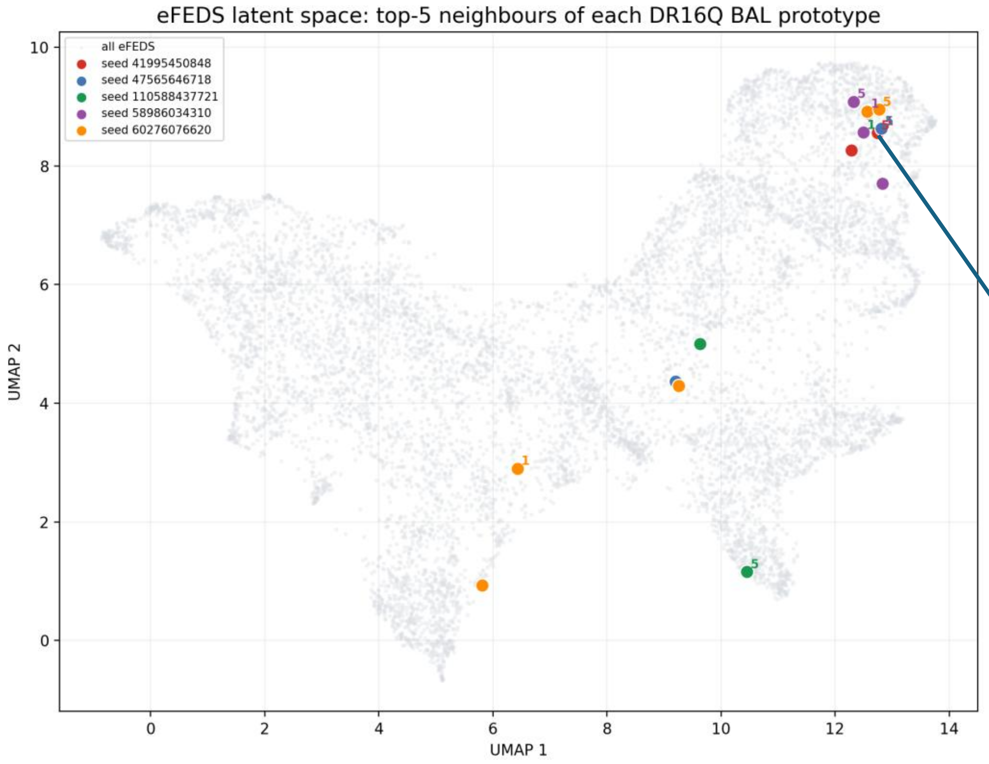
- The closest objects are all BAL-like, or BAL quasars

! No BAL labels were used to create this grouping

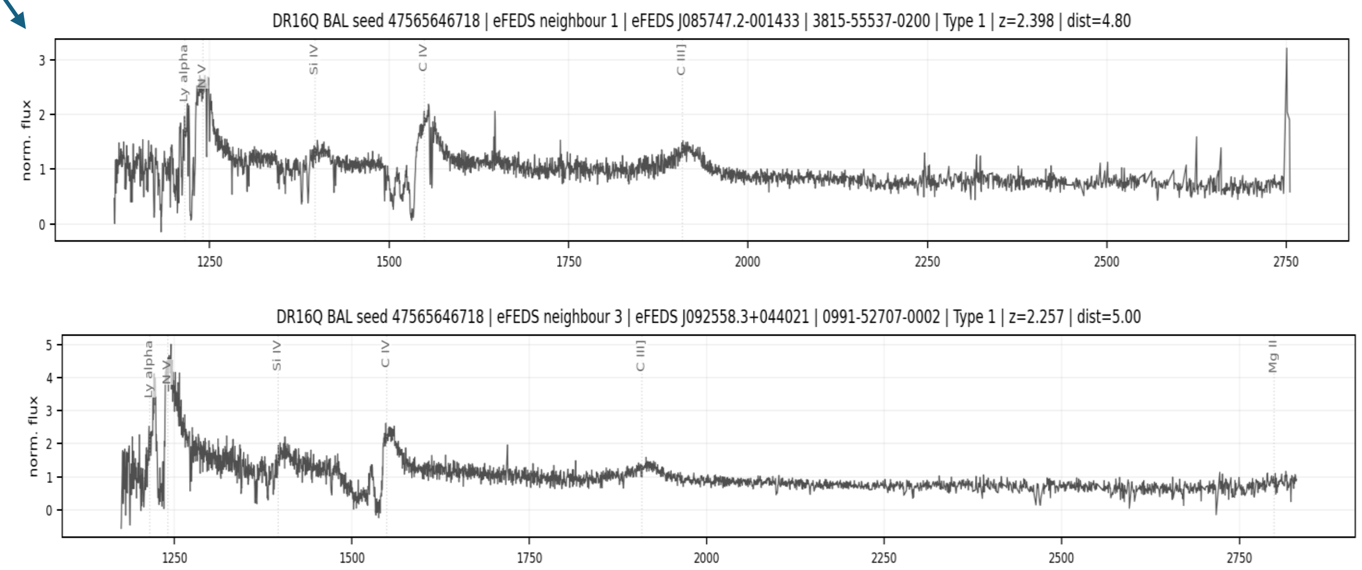
The latent space has learned the main spectral pattern associated with BAL quasars

Pattern Recognition

- **Searching for BAL-like candidates in eFEDS**



- Input prototypes: 5 DR16Q BAL quasars
- Search space: eFEDS optical latent space



Conclusions

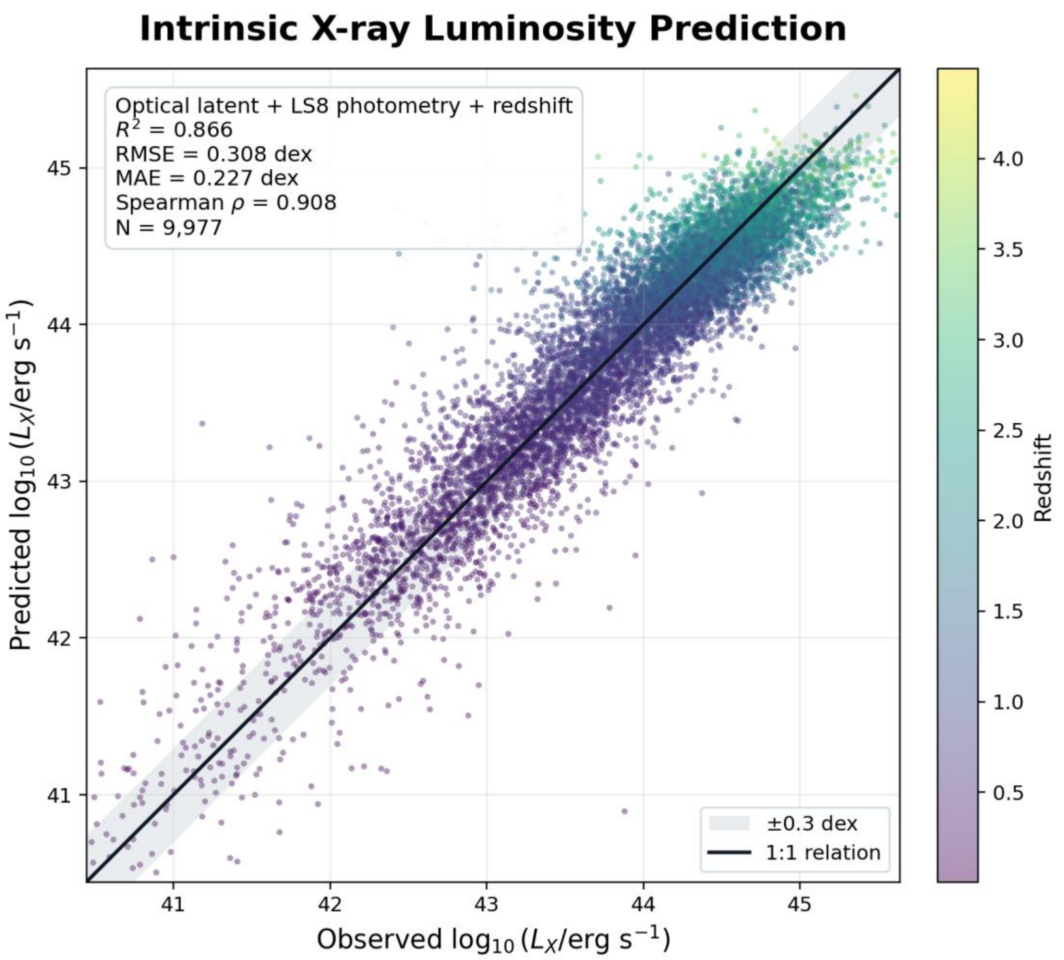
- A single model can represent galaxies, AGN and quasars across a wide range of spectral types and redshifts.
- The latent space shows a clear physical organisation of optical spectral classes.
- Poor reconstructions can reveal catalogue problems, such as wrong redshifts or contaminants.
- Nearest-neighbour searches recover rare spectral patterns, such as BAL quasars, without using class labels.
- Next: connect the optical latent representation with the X-ray information in eFEDS.

Thanks ☺!

noemilery.borrelli@unina.it

Preliminary X-ray downstream test

- **Target: Intrinsic X-ray luminosity**



Input features:

- Optical latent vector, 24 dimensions (spender eFEDS)
- Redshift z
- Dereddened LS8 photometryInput features:
  - Fluxes: g, r, z, W1, W2
  - Colours: g-r, r-z, W1-W2, ...

Best model:

HistGradientBoostingRegressor

Performance:

R^2 = 0.866
RMSE = 0.308 dex
MAE = 0.227 dex
Spearman ρ = 0.908

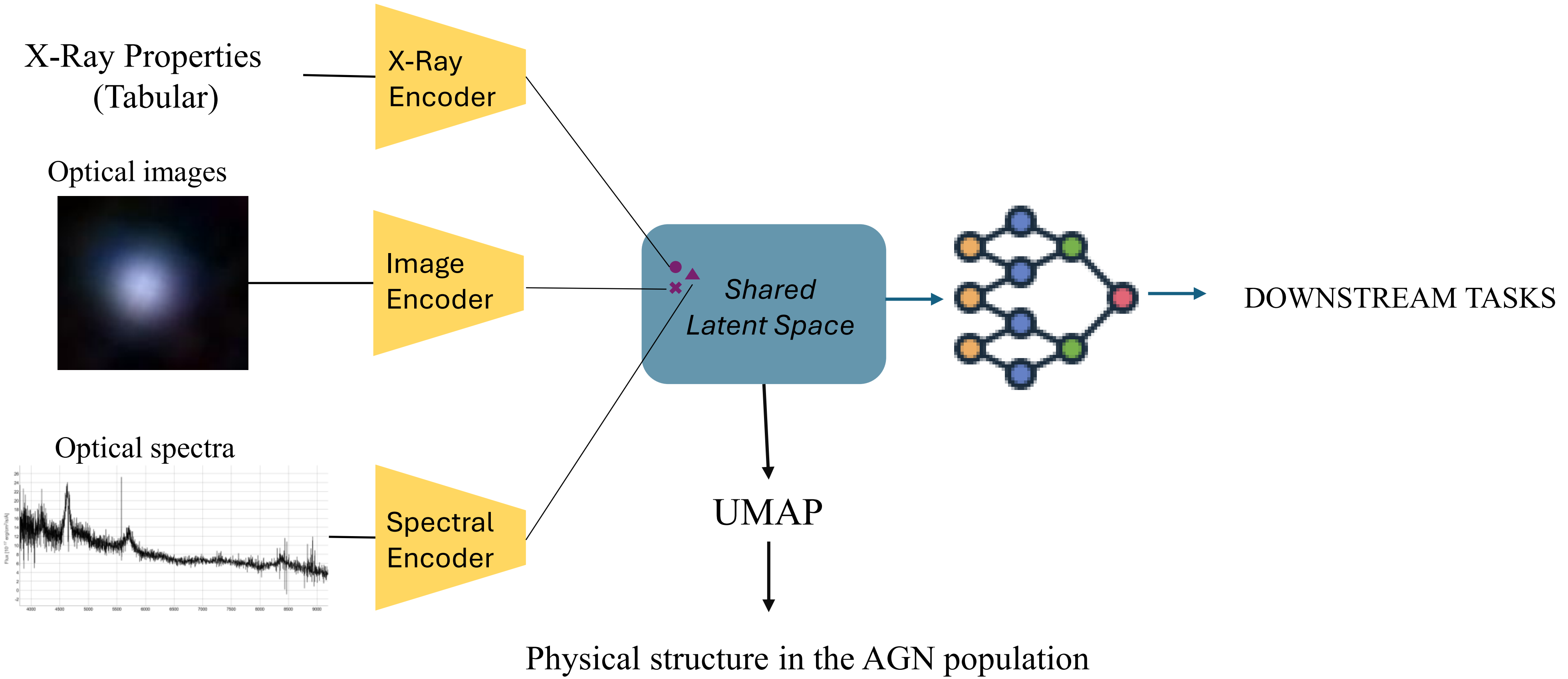
Backup — Abaut the model

Long-term idea: a multimodal representation model for AGN

Goal: *learn a **shared representation of AGN** combining different observational modalities.*

Step 1 — Encoder pretraining

Step 2 — Joint multimodal training (eFEDS–SDSS AGN catalog)



Spectral Encoder: autoencoder SPENDER *(Liang et al. 2023)*

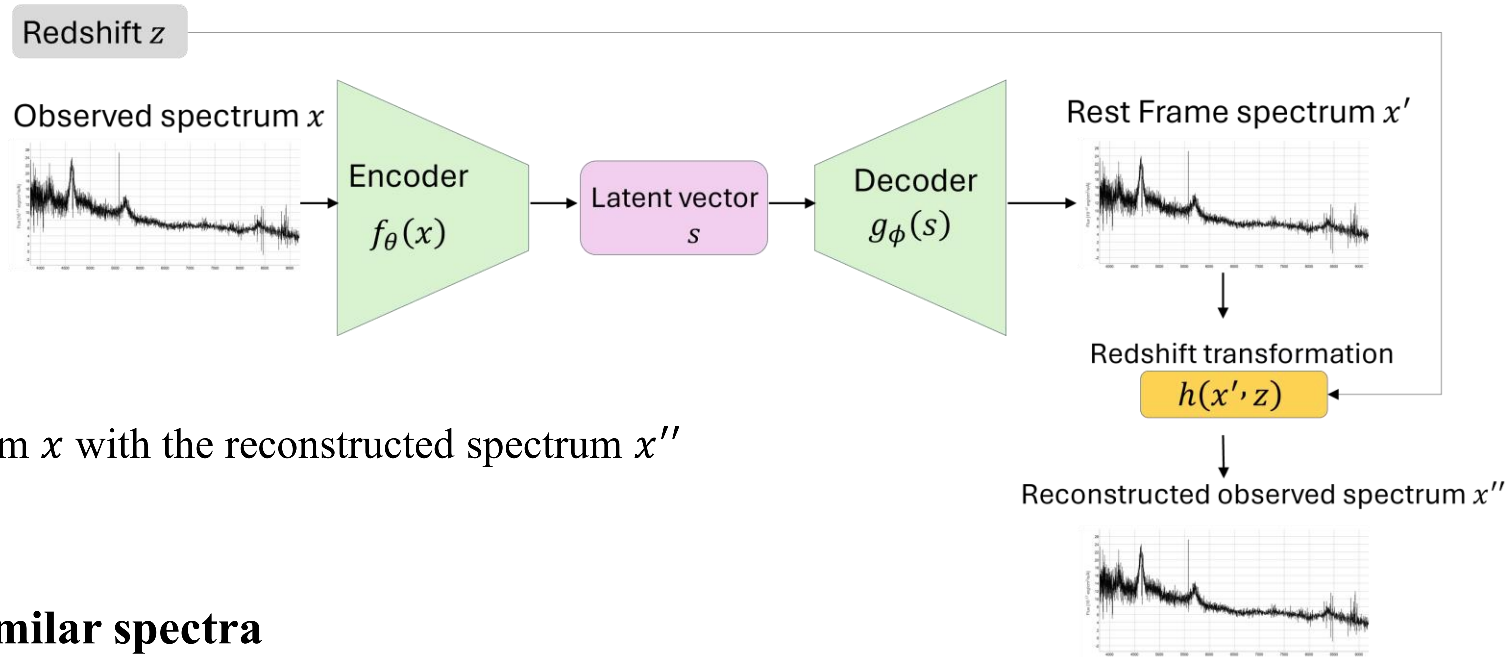
Why SPENDER?

- it works directly on **observed-frame spectra**
- it aims to capture **intrinsic spectral properties**
- this is more coherent with a future **multimodal model with images**

Original spender architecture is not enough

Goal
Learn a latent space of galaxy spectra that captures **intrinsic physical properties**, independent of observational effects such as redshift.

Basic architecture (Melchior, 2022)

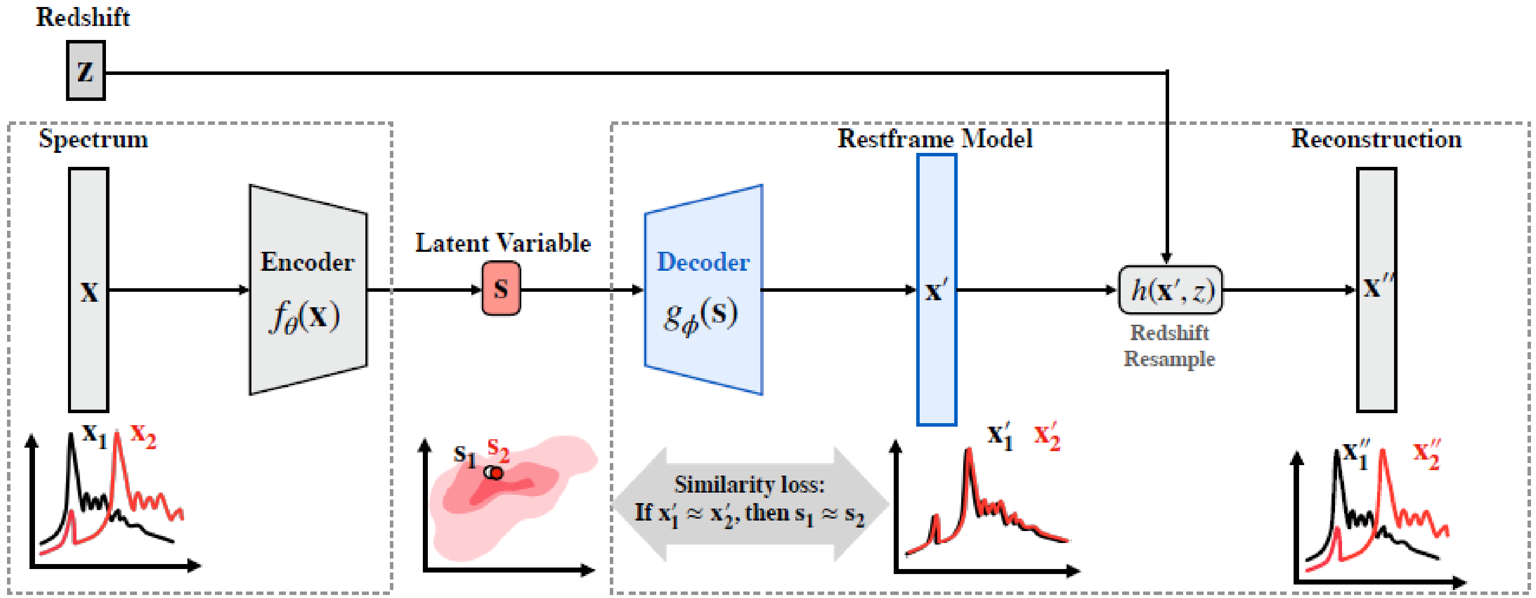


Reconstruction loss: compares the observed spectrum x with the reconstructed spectrum x''

Problem :

- different latent vectors can produce **very similar spectra**
- the same galaxy observed at different redshifts can be mapped to **different latent positions**

Spectral Encoder: autoencoder SPENDER *(Liang et al. 2023)*



Improvements (Liang 2023)

To remove degeneracies and enforce physical consistency, the model introduces additional training constraints.

1. Similarity loss

If two spectra have **similar rest-frame reconstructions ,** their latent representations must also be close:

$$| x_1' - x_2' | \Rightarrow | s_1 - s_2 |$$

→ encourages spectra of **physically similar galaxies** to cluster in latent space.

2. Consistency loss

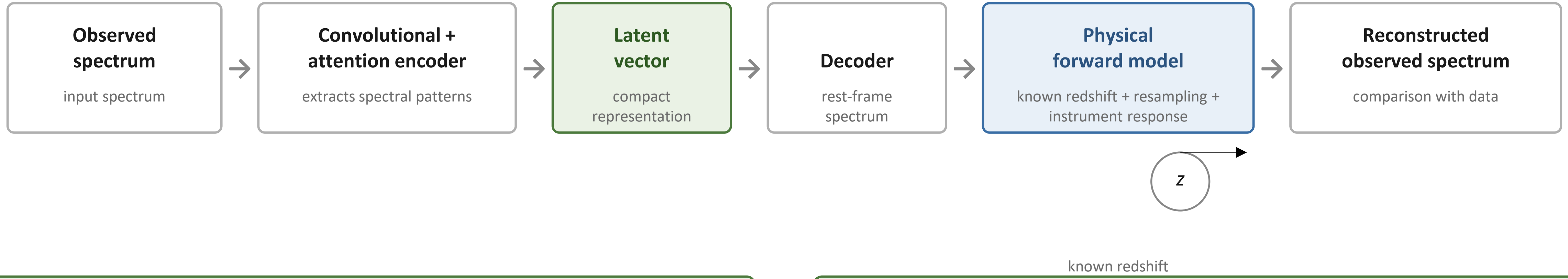
Artificially redshifted versions of the same spectrum are generated during training. The model enforces

$$s(x) \approx s(x_{aug})$$

→the latent representation becomes **invariant to redshift and observational effects**.

SPENDER as a starting architecture

Separate intrinsic spectral information from known observational transformations



Attention

- learns which spectral features are informative
- can recognise them at different observed wavelengths

Explicit redshift handling

- decoder produces the intrinsic rest-frame spectrum
- known observational effects are applied afterwards

Key idea: the latent space should describe intrinsic spectral properties, not known observational shifts.

SPENDER architecture: Melchior et al. 2022; Liang et al. 2023

Common Model Architecture (All Phases)

Input spectrum — *SDSS observed wavelength grid* → **Normalizer** — *deterministic, score-based* → **Conv Encoder** — *convolutional* → **Latent z** — dz = 24 → **MLP Decoder** — *24→64→256→1024→Nrest* → **Reconstruction** — *rest-frame, 650–9500 Å* →

dz = 24	**16,642**	**650–9500 Å**	**7×10^{-5}**
LATENT DIMENSION	REST-FRAME PIXELS (NREST)	REST-FRAME DECODER RANGE	$\Delta\mathrm{LOG}_{10}\Lambda$ (LOG-SPACED GRID)

- Decoder hidden widths: 24 → 64 → 256 → 1024 → Nrest, expanding the compact latent code into a full rest-frame spectrum.
- The normalizer is inside the model wrapper and is applied deterministically before encoding (see Normalization, slides 7–8).

Model Size

20,559,554

TOTAL STATE ELEMENTS

20,535,070

TRAINABLE PARAMETERS

24,484

NON-TRAINABLE BUFFER ELEMENTS

Non-trainable wavelength-grid buffers

Encoder observed-frame grid	3,921 elements
Decoder rest-frame grid	16,642 elements
Normalizer observed-frame grid	3,921 elements

How parameters are counted

$$N_{par} = \Sigma_i \Pi_j s_{ij}$$

Dense layer: $N_{par} = m \cdot n + m$ (weights + bias)

1D conv: $N_{par} = C_{out} \cdot C_{in} \cdot K + C_{out}$

Redshift and Rest-Frame Modeling

$$\lambda_{rest} = \lambda_{obs} / (1 + z)$$

Each object's redshift links its observed-frame spectrum to the common rest-frame decoder grid.

- The same physical line appears at different observed wavelengths for different redshifts — a common rest-frame grid aligns physical features (optical galaxy lines, UV quasar lines, continuum breaks, broad absorption structures) across the whole sample.
- The wide redshift range means incomplete rest-frame coverage: not every object covers the same rest-frame wavelengths.
- This is handled by two mechanisms working together: inverse-variance masking (missing pixels simply don't contribute to the loss) and adaptive normalization windows (next slide).

Normalization: Removing Flux Scale, Keeping Shape

PURPOSE

Remove the arbitrary absolute flux scale (distance, fibre aperture, observing conditions, calibration) while preserving spectral shape and relative line/continuum information.

$$f\,norm = f\,raw\,/\,ni$$

$$ivar\,norm = ivar\,raw \cdot ni^2$$

Inverse variance scales as n^2 because σ_norm = σ_raw / n.

SCORE-BASED ADAPTIVE NORMALIZER — deterministic & label-agnostic

- Uses only: redshift, observed wavelength coverage, valid-pixel count per window, median in-window S/N, distance from spectral edges, robust fallback statistics.
- Does NOT use object class, BPT class, QSO label, eFEDS label, or instrument label.

Why multiple candidate windows? No single rest-frame continuum window is available for all galaxies, AGN, and quasars across the full redshift range — low-z galaxies have good optical windows, high-z quasars often need UV windows instead.

Normalizer: Candidate Windows & Settings

Phase / sample	Candidate rest-frame windows (Å)
Phase 1, Phase 3	4200–4400, 4500–4700, 5100–5400, 5500–5800, 6000–6200, 6800–7000
Phase 2, Phase 4, eFEDS	1450–1480, 1700–1800, 1975–2050, 2200–2600, 3000–3400, 4200–4400, 4500–4700, 5100–5400, 5500–5800, 6000–6200, 6800–7000

20
MIN. VALID PIXELS / WINDOW

0.15
MIN. SCORE THRESHOLD

0.80
TARGET COVERAGE FRACTION

50
TARGET # PIXELS

5
TARGET MEDIAN S/N

100 Å
TARGET DISTANCE FROM EDGE

Robust fallback statistics are enabled throughout.

All Five Phases at a Glance

Phase	Input sample	Size used	Role
Phase 1	MPA-JHU galaxies + low-z AGN	883,570 spectra	Stable low-z optical baseline
Phase 2	DR16Q quasars (BOSS+SDSS)	421,058 train / 46,705 valid	Add high-z, UV quasar features
Phase 3	MPA-JHU AGN-transition	142,556 selected / 141,835 used	Reinforce low-z AGN & transition objects
Phase 4	Mixed: DR16Q + MPA-JHU AGN + bg.	≈149,085 + repeats (58/27/15%)	Joint consolidation of all populations
eFEDS	eFEDS X-ray selected sample	9,985 selected (6,989/1,500/1,496)	Fine-tune to the X-ray sample — final model

Each phase resumes training from the previous phase's final checkpoint. Full per-phase settings (epochs, learning rate, loss schedule) follow on the next five slides.

Phase 1 — MPA-JHU Pre-Training

883,570

SPECTRA (SN-SELECTED)

SN ≥ 3

CONTINUUM-QUALITY CUT

dz = 24

LATENT DIM (FIXED)

FINAL ROLE

- Learn a stable low-redshift optical galaxy/AGN representation.
- Initialize the curriculum with spectra dominated by optical continuum and common galaxy emission/absorption features.
- Keep the latent dimension fixed at 24 — this constraint holds for every later phase.
- No removal based only on BPT class: star-forming, passive/unclassifiable, composite, AGN, and low-S/N emission-line classes are all included if continuum quality is acceptable.

KEY TRAINING SETTINGS

- Input grid: SDSS observed-frame grid.
- Raw batches with model-internal normalization.
- Score-based optical normalization windows.
- Reconstruction + similarity + consistency losses, SPENDER-like cyclic similarity schedule.
- Consistency activated after warmup and kept active.
- Final checkpoint: long clean run + short refinement (reconstruction + consistency only, similarity off) to stabilize reconstruction before Phase 2.

Phase 2 — DR16Q Quasar Training

421,058 / 46,705
TRAIN / VALID SPECTRA

300
EPOCHS

1e-4 → 1e-5
LEARNING RATE (COSINE)

FINAL ROLE

- Add broad-line quasars and high-redshift UV spectral features.
- Make the representation sensitive to quasar spectral diversity.
- Keep compatibility with the Phase-1 architecture and latent size.
- 467,763 successful spectra total; BOSS spectra interpolated onto the SDSS observed grid; SN_MEDIAN_ALL ≥ 3; redshift priority ZSYS > ZFIT > ZDR16Q > ZPCA > Z.

KEY TRAINING SETTINGS

- Resume from the final Phase-1 checkpoint. Batch size 256.
- UV + optical adaptive normalization windows.
- Similarity active after 20 warmup epochs, off in the last 10.
- Consistency: 20 warmup epochs, 30-epoch ramp.
- Redshift augmentation local within z-bins of width 0.5.
- Lyα/Lyβ masking loop (SpenderQ-like): 30-epoch warmup, then 3× 90-epoch iterations, masking pixels where flux falls below reconstruction by >1.1σ (Lyα) / 0.8σ (Lyβ) / 3.0σ (redward); bin width 4–16 Å.

The Lyα/Lyβ mask loop: reconstruct → flag deviant forest pixels → zero their inverse variance → repeat.

Phase 3 — MPA-JHU AGN-Transition Reinforcement

141,835
SPECTRA IN FINAL BATCHES

180
EPOCHS

1e-5 → 3e-6
LEARNING RATE (COSINE)

FINAL ROLE

- Reintroduce and reinforce low-redshift active/transition spectra after the high-redshift quasar phase.
- Prevent the representation from becoming too quasar-dominated.
- Improve the handling of Composite, AGN, and LINER-like spectra.
- Uses BPT classes Composite, AGN, and low-S/N LINER, with continuum-quality selection; 142,556 selected in the manifest.

KEY TRAINING SETTINGS

- Resume from the final Phase-2 checkpoint. Batch size 256.
- Score-based optical normalization (same window set as Phase 1).
- No Lyα mask loop — this is a low-redshift optical sample.
- Similarity: 10-epoch warmup, active until epoch 170, then off for the last 10 epochs.
- Consistency: 10-epoch warmup, 50-epoch ramp.
- Final choice is the global-redshift configuration, not the earlier narrow-bin exploratory version.

Phase 4 — Unified Mixed Training

58 / 27 / 15%

DR16Q / AGN / BACKGROUND

200

EPOCHS

2.5e-6 → 8e-7

LEARNING RATE (COSINE)

FINAL ROLE

- Consolidate the final pre-eFEDS representation.
- Jointly expose the model to galaxies, low-z AGN, and high-z quasars.
- Reduce catastrophic forgetting across the curriculum.
- Effective per-epoch exposure: DR16Q 421,058 (repeat ×1), MPA-JHU AGN 198,656 (repeat ×2, 99,328 train spectra), reduced MPA-JHU background ≈105,000 (repeat ×1).

KEY TRAINING SETTINGS

- Resume from the best Phase-3 global-redshift checkpoint. Batch size 256.
- UV + optical adaptive normalization.
- Similarity: 20-epoch warmup, two complete cycles over the active interval, off in the final 10 epochs.
- Consistency: 20-epoch warmup, 30-epoch ramp.
- Lyα masking applied to DR16Q-like spectra only.
- Reduced background: 150,000 selected, 149,085 spectra in final batches.

eFEDS Fine-Tuning

9,985
SELECTED EFEDS SPECTRA

6,989 / 1,500 / 1,496
TRAIN / VALID / TEST

8,386 / 1,599
BOSS- / SDSS-ORIGIN

70
EPOCHS

FINAL ROLE

- Adapt the Phase-4 optical representation to the eFEDS X-ray selected spectroscopic sample.
- Retain the same latent dimensionality and architecture (dz = 24) — no architectural change.
- This produces the FINAL latent space used for the eFEDS UMAP and downstream X-ray exploratory tests.
- Selected catalog spans both origins: 8,386 BOSS-origin + 1,599 SDSS-origin spectra, all placed on the common SDSS grid.

KEY TRAINING SETTINGS

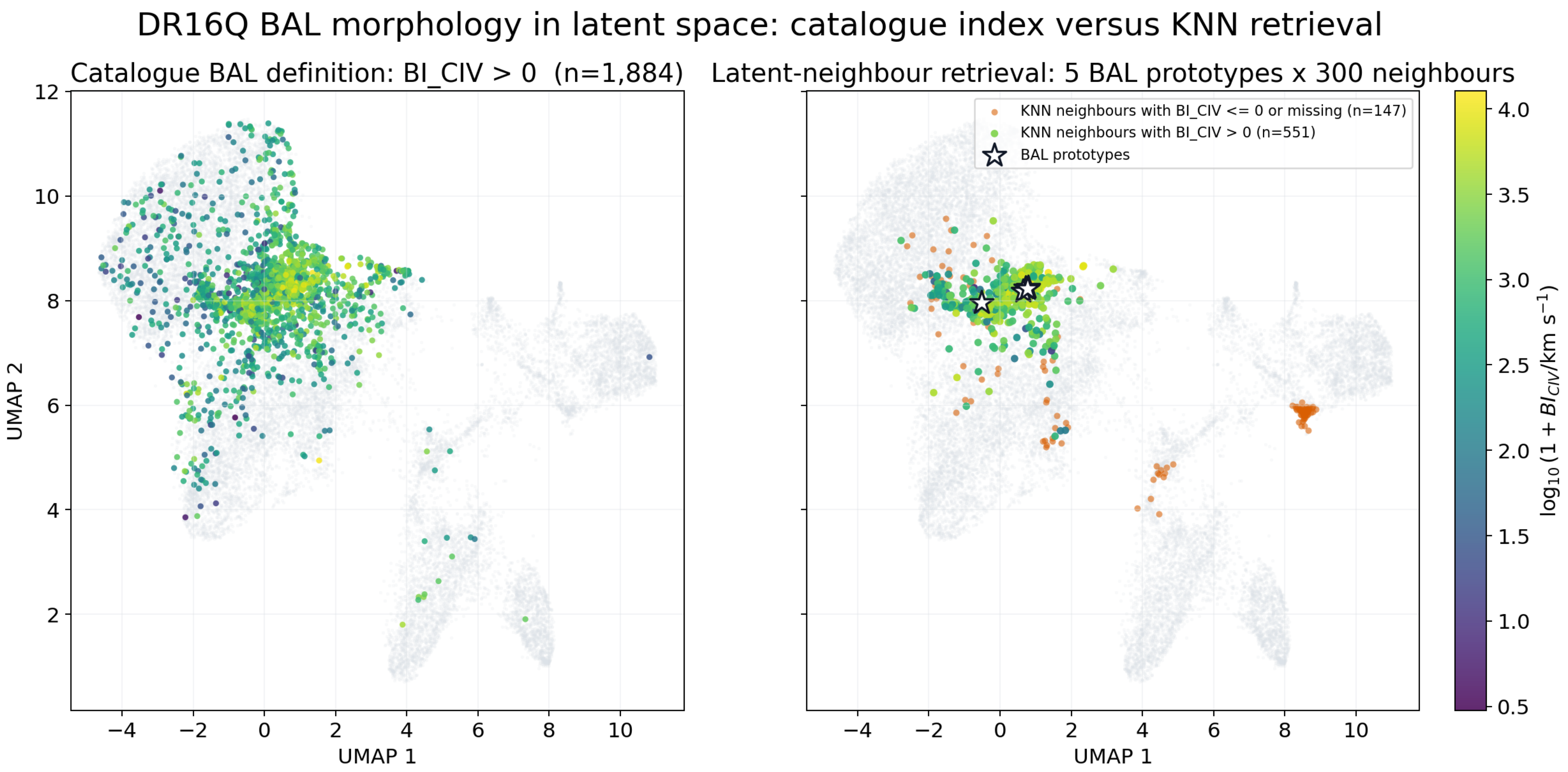
- Resume from the final Phase-4 checkpoint. Batch size 256.
- Learning rate $5\times10^{-7} \rightarrow 1\times10^{-7}$ (cosine) — the gentlest schedule in the whole curriculum.
- UV + optical adaptive normalization (same as Phase 2/4).
- Similarity: 10-epoch warmup, off in the last 10 epochs.
- Consistency: 10-epoch warmup, 20-epoch ramp.
- Redshift augmentation over the full $0 < z < 5$ range, local bin width 0.5.
- No additional Lyα mask loop at this stage.

Backup — Latent space vs. physical properties

One-vs-rest UMAP diagnostics: optical type, X-ray obscuration, and off-region Type 2 outliers

Pattern Recognition

- **Nearest neighbours of a BAL-quasar target in the phase-2 latent space**

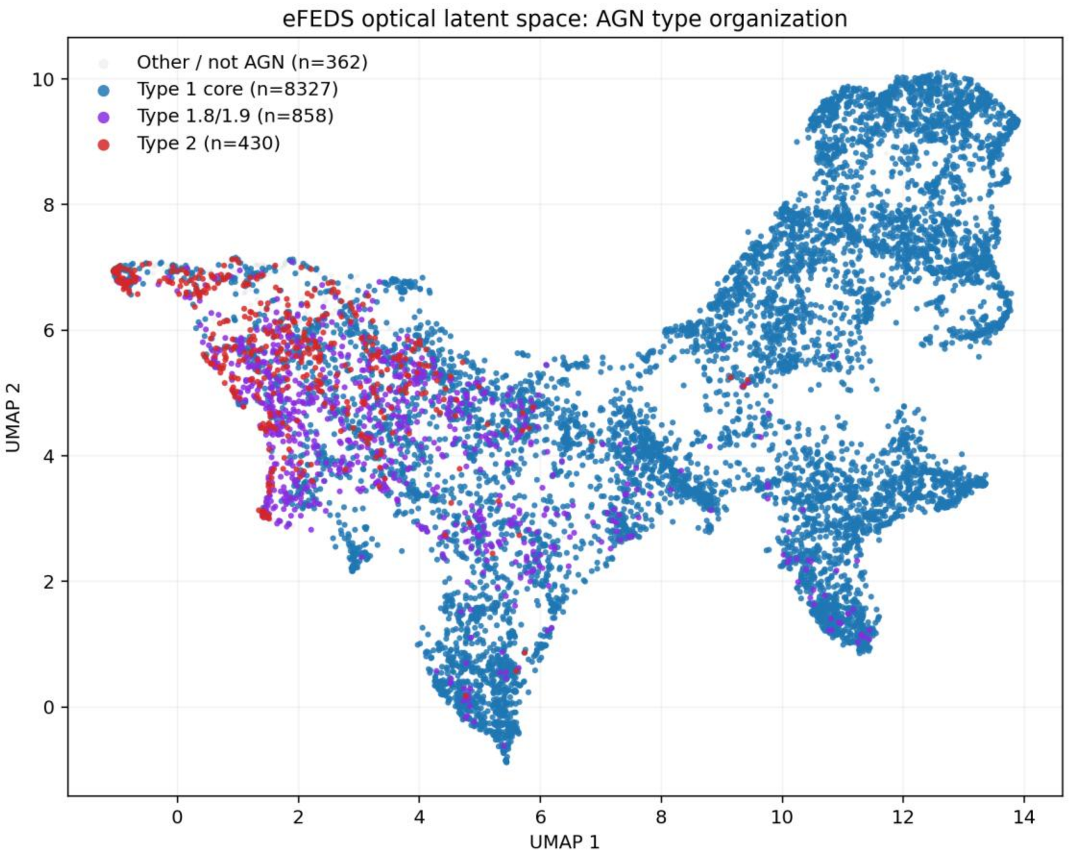


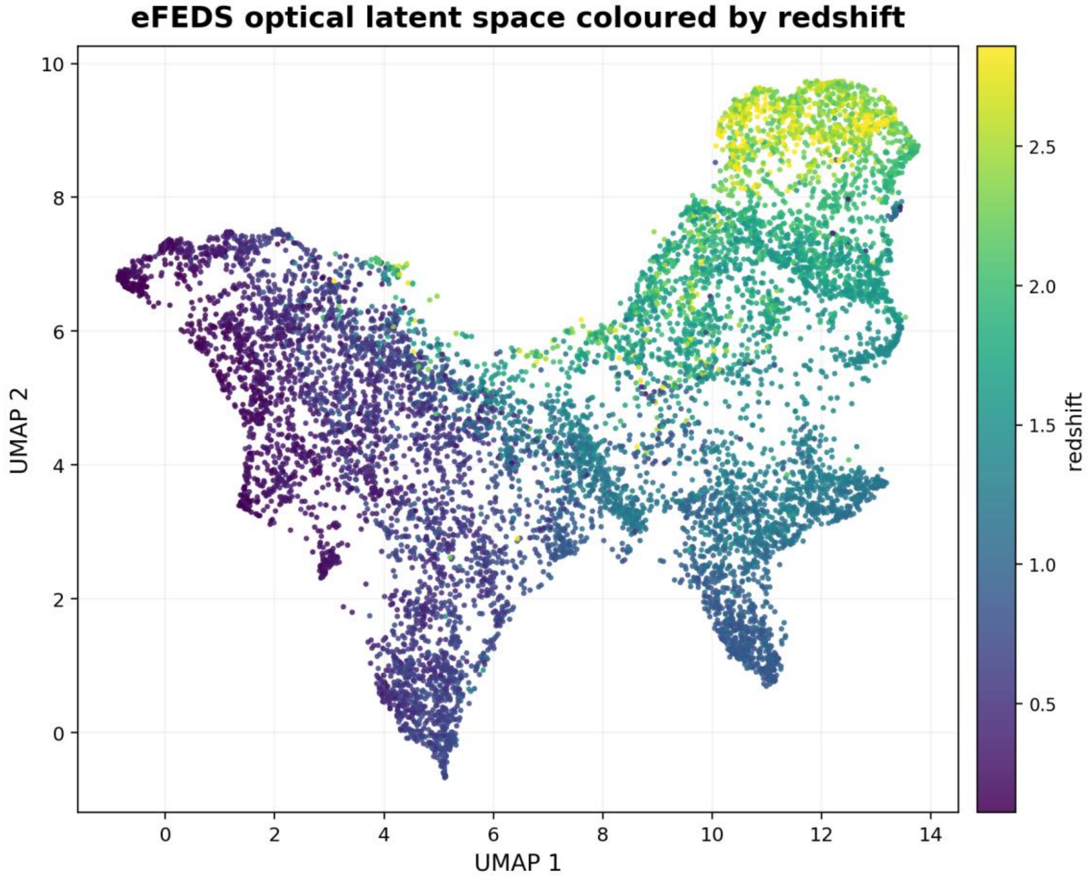
Backup — Latent space vs. physical properties

One-vs-rest UMAP diagnostics: optical type, X-ray obscuration, and off-region Type 2 outliers

LATENT EXPLORATION USING UMAP

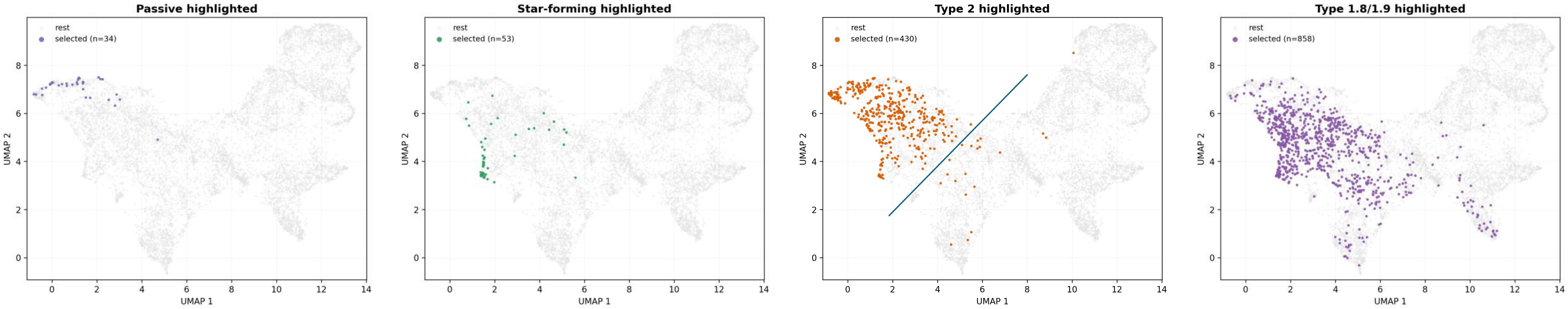
- **UMAP projection of the latent space**





Aydar et al. 2026 for the classification labels

The latent space encodes a continuous AGN activity sequence



Passive (n=34) — extreme tip

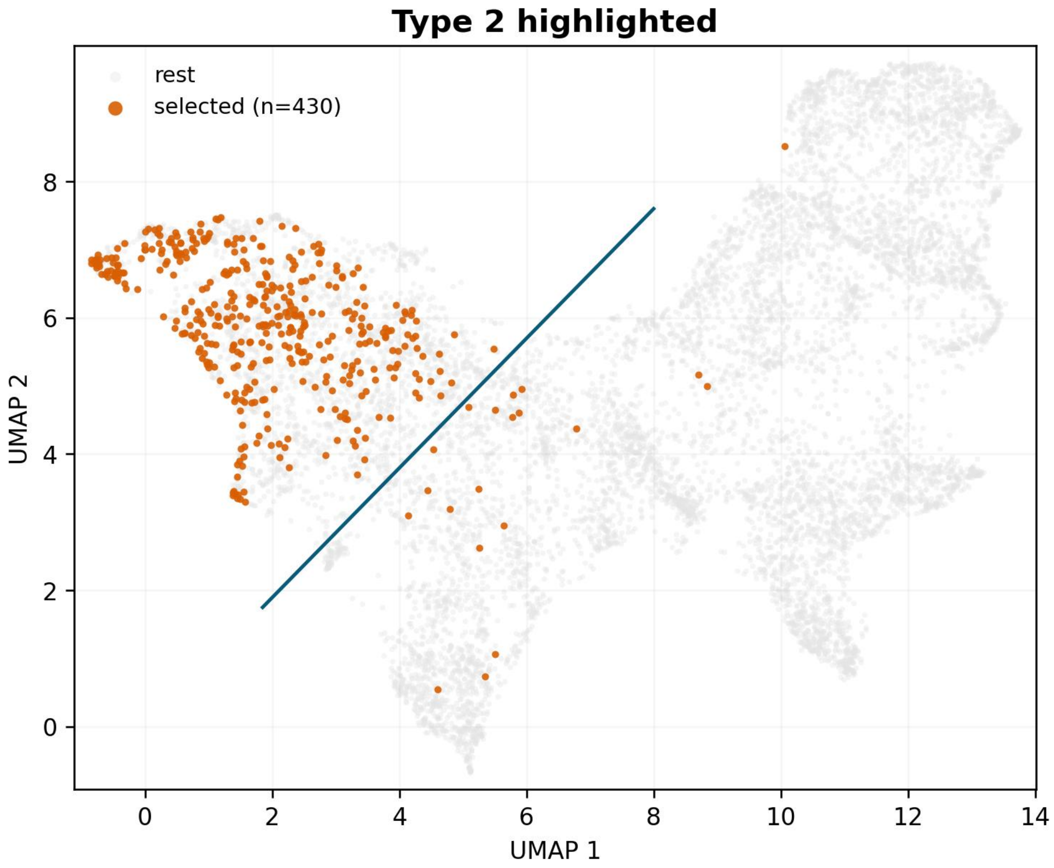
Star-forming (n=53) — same region

Type 2 (n=430) — broader head

Type 1.8/1.9 (n=858) — bridges to Type 1

- Passive → star-forming → Type 2 → Type 1.8/1.9 → Type 1 core occupy progressively larger, shifted regions of the map.
- No optical label was used to build the latent space — this spatial gradient emerged purely from spectral reconstruction.

A handful of “Type 2” sources land in the Type-1 region

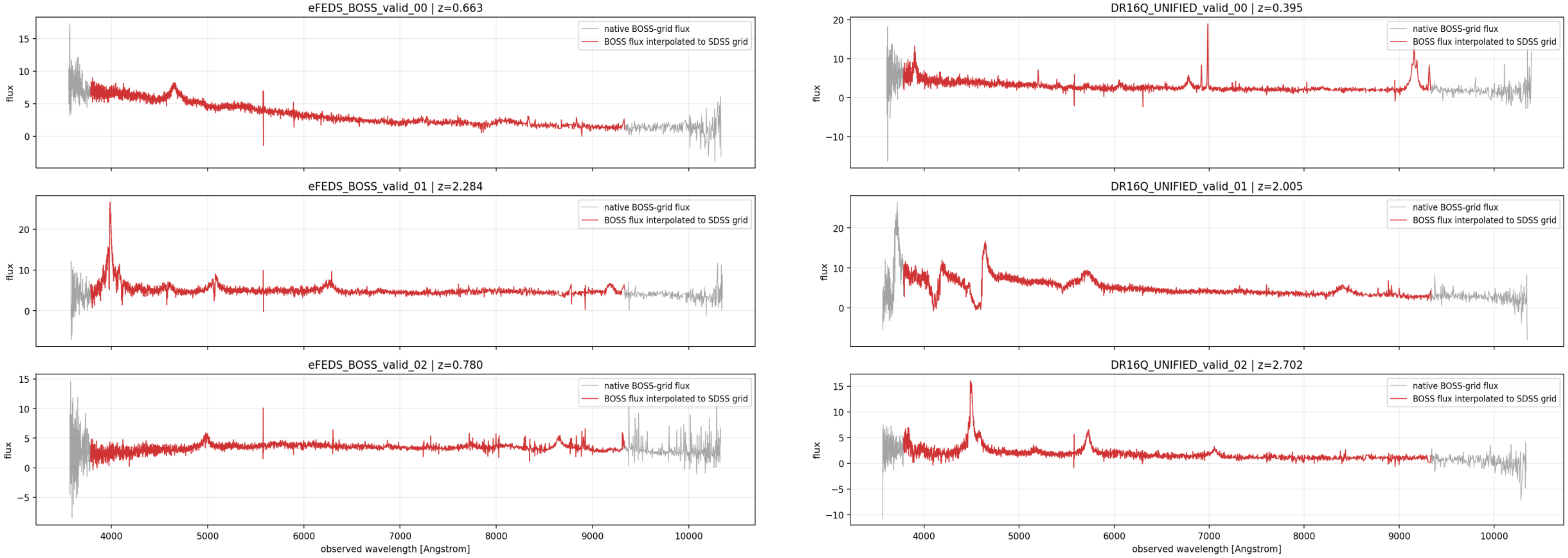


- 6 Type-2-labelled sources fall well outside the main Type 2 region (diagonal line), inside the Type-1-dominated wing.
- 4 of 6 have very low S/N (0.4–2.1): the “Type 2” label itself is unreliable there — likely noise-driven misclassification, not a real outlier.
- 2 of 6 have usable S/N (3.4 and 8.4). The S/N=8.4 case shows a systematic reconstruction offset (χ^2=10.1, not noise)

Backup / appendix — see next slide for the individual spectra

Flux interpolation examples

Grey is native BOSS grid; red is the same spectrum on the SDSS grid.



The interpolation mostly removes the BOSS-only blue/red edges