Production and potential escape of Lyman radiation in Extreme Emission-Line Galaxies from DESI

RICARDO AMORÍN
(IAA, GRANADA)

INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA
EXCELENCIA SEVERO OCHOA
CSIC

Escape of Lyman radiation from galactic labyrinths
7-11 April 2025, OAC, Kolymbari, Crete

"XTREM" Grant PID2023-147386NB-I00 funded by:
MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES
Cofinanciado por la Unión Europea
AGENCIA ESTATAL DE INVESTIGACIÓN

Production and potential escape of Lyman radiation in Extreme Emission-Line Galaxies from DESI

Ricardo Amorín

(IAA, Granada)

Main contributors

J.A. Fernandez-Ontiveros, A. Giménez-Alcázar, J.M. Vílchez, L. Bonatto, B. Perez-Díaz, E. Pérez-Montero, A. Hernán-Caballero, A. Lumbreras-Calle, P. Papaderos

Agenda

* Intro: Key questions & definitions: Extreme emission-line galaxies (EELGs) through cosmic time
* Our study:
- XTREM: A large statistically significant sample of EELGs at z<1. Properties and diagnostics
- Selection of candidates to strong Lya and LyC emitters:
  - Prediction from Cloudy models
  - Implications for high-z galaxies
- Conclusions and future

Instituto de Astrofísica de Andalucía · Excelencia Severo Ochoa · CSIC

Escape of Lyman radiation from galactic labyrinths
7-11 April 2025, OAC, Kolymbari, Crete

"XTREM" Grant PID2023-147386NB-I00 funded by:

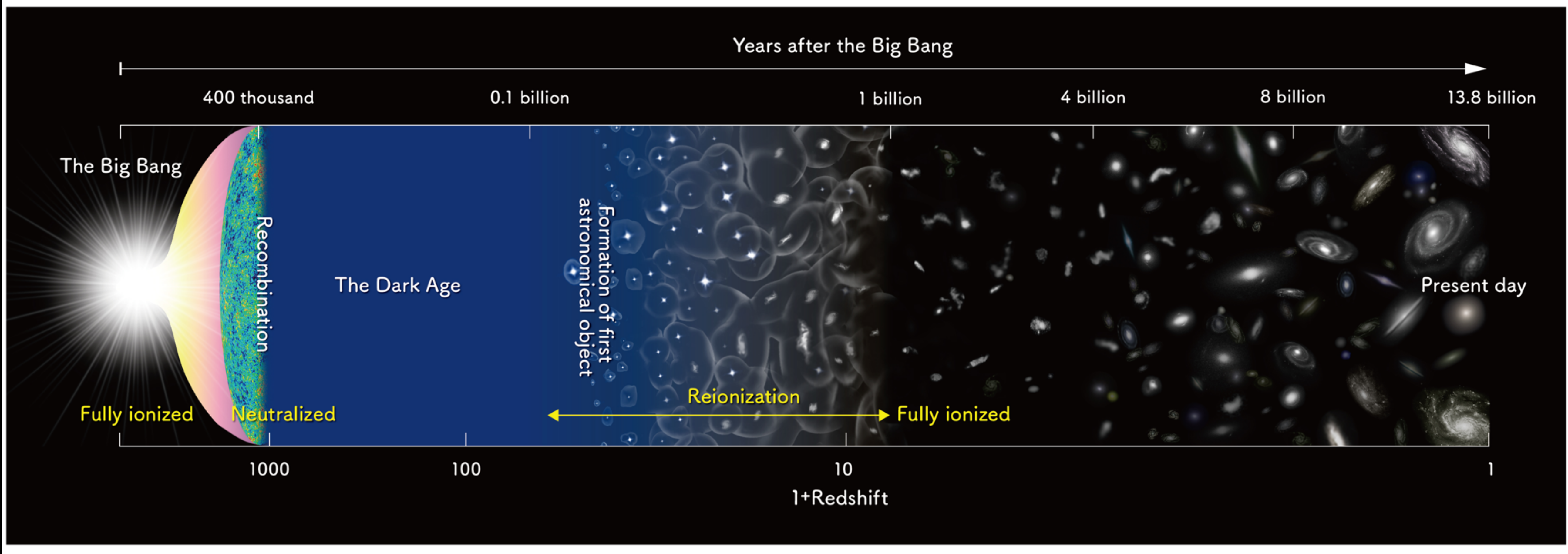
Ministerio de Ciencia, Innovación y Universidades · Cofinanciado por la Unión Europea · Agencia Estatal de Investigación

Key questions

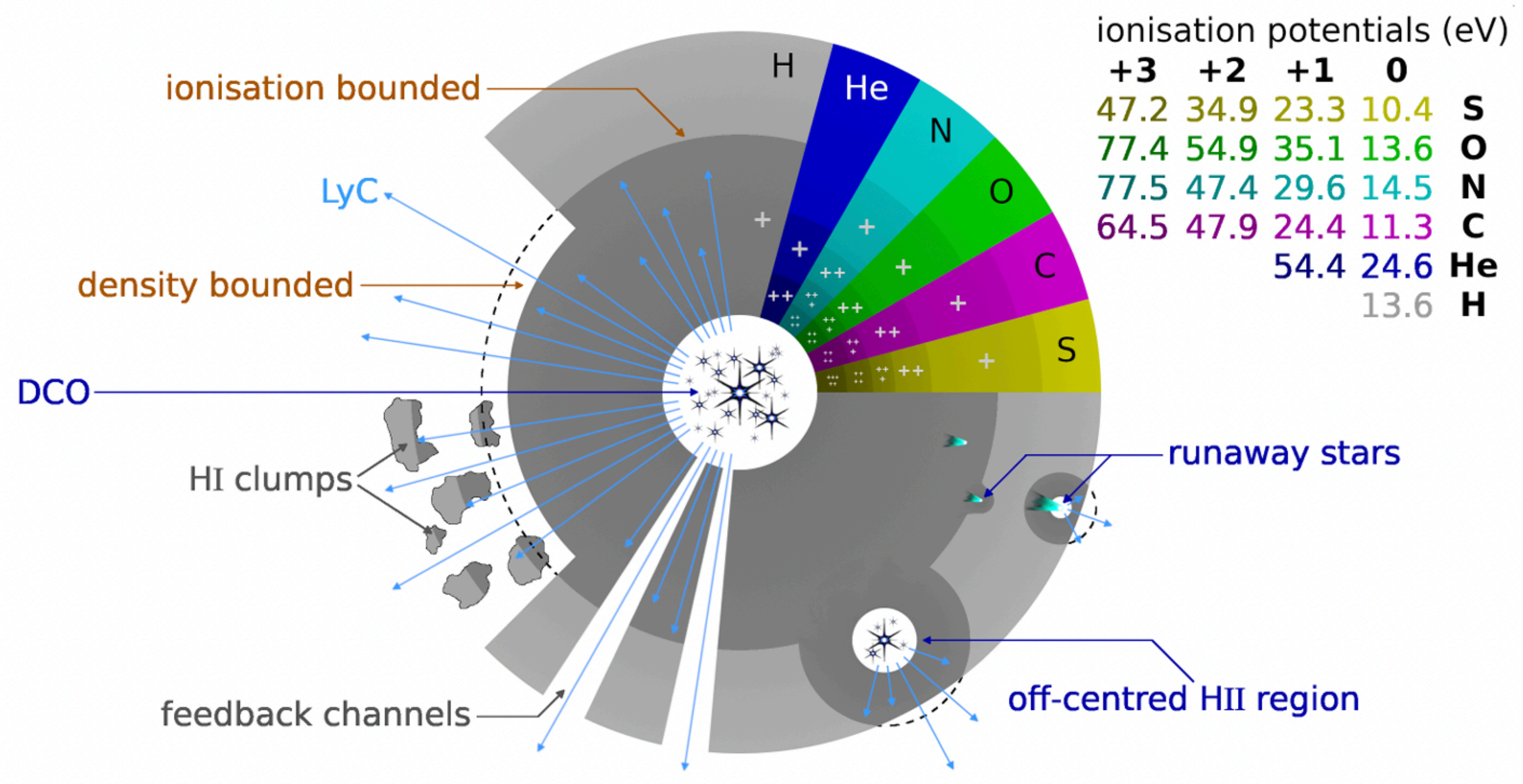
Escaping from galactic labyrinths

How do galaxies reionized the universe at z>6?
How LyC escape from galaxies?

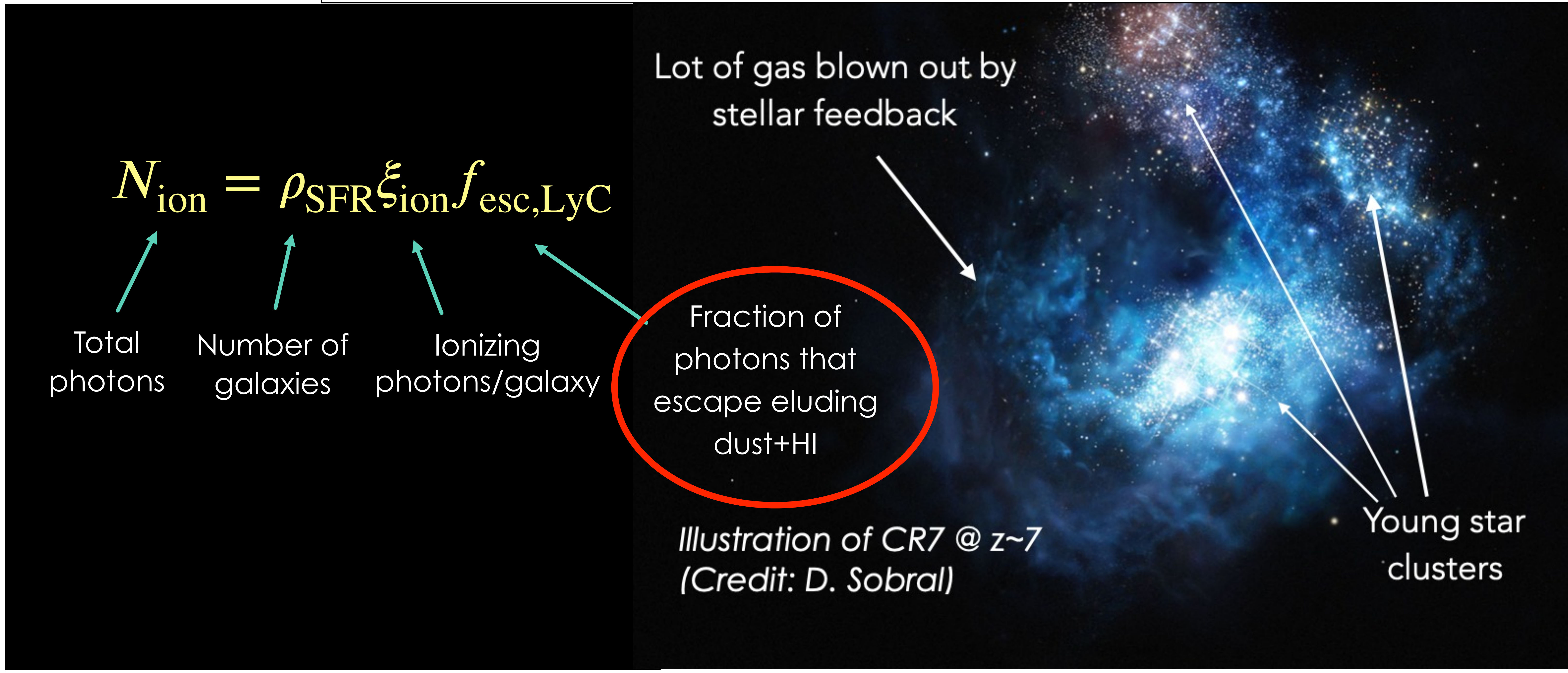
What physical properties and mechanisms favor
Lyman photon production and escape?



Simple models: Nakajima & Ouchi 13, Zackrisson+13; Jaskot & Oey 13, Ramambason+20



Enders+23

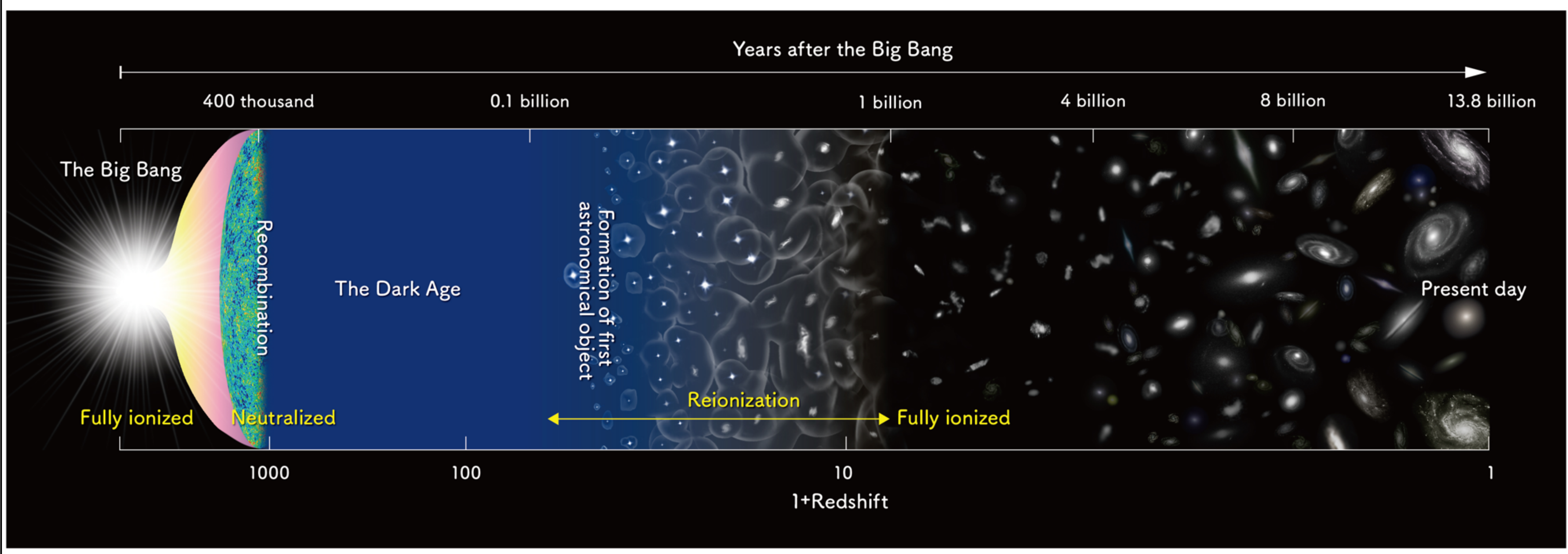


Key questions

Escaping from galactic labyrinths

How do galaxies reionized the universe at z>6?
How LyC escape from galaxies?

What physical properties and mechanisms favor Lyman photon production and escape?

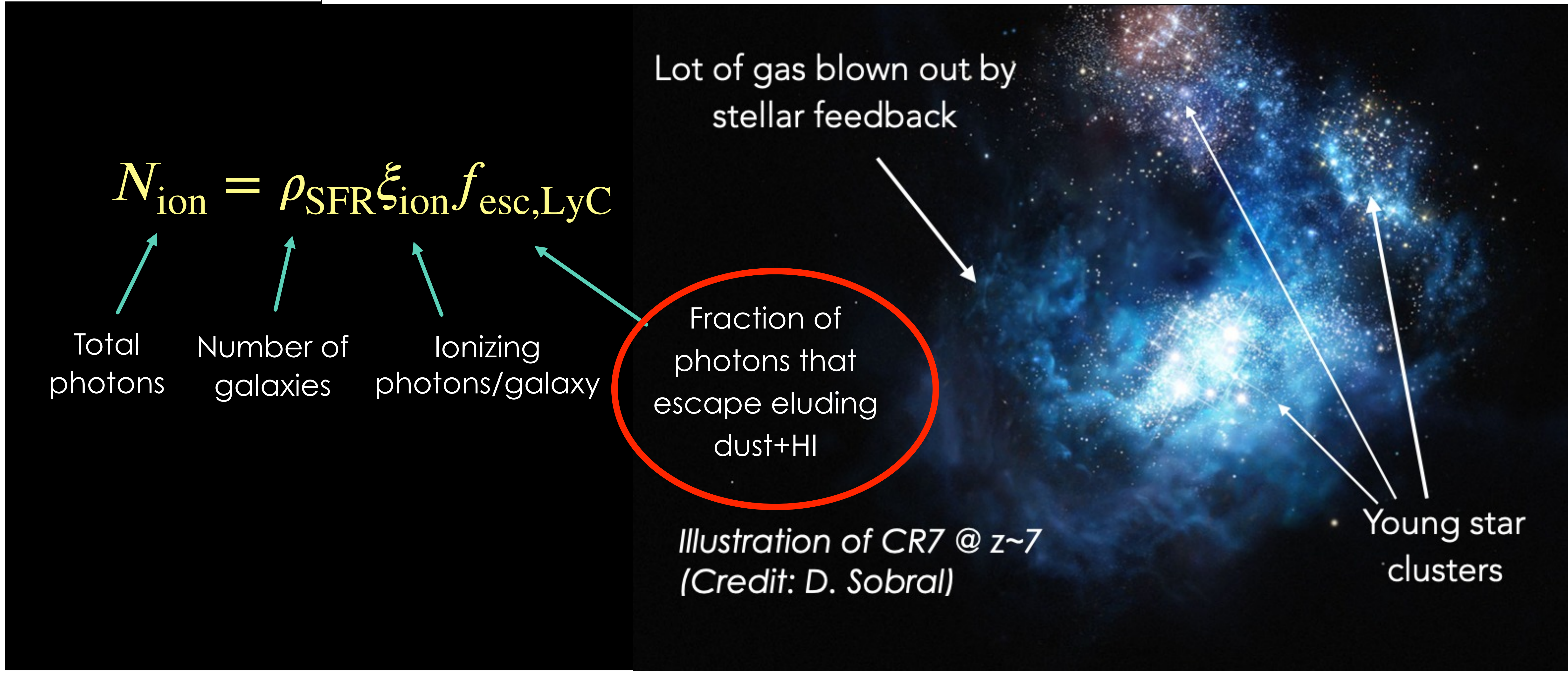


PROBLEM

At z>4, the strong drop-down of the IGM transmission makes impossible the *direct* detection of the LyC photons responsible for reionization, even for JWST or ELTs

THE LOW-Z ANALOG APPROACH

At z<4, LyC is directly observable with HST. Probing well-characterized analogs of reionization galaxies (LAEs, LBGs) may help to provide new insight on these questions



Key questions

Escaping from galactic labyrinths

What physical properties and mechanisms favor
Lyman photon production and escape?

THE LOW-Z ANALOG APPROACH: MAIN RESULTS

From super recent review of Anne Jaskot, ARA&A, 2025

See also e.g. Izotov+18, Flury+22, Saldana-Lopez+22, Chisholm+22

★ Growing samples of LyC emitting samples, especially at z~0.3 and z~2-3 from HST surveys and detailed multiwavength observations of smaller samples e.g. Low-z Lyman Continuum Survey (Flury+2), LaCOS (La Reste+25), Steidel+18, Siana+15, Vanzella+16, Rivera-Thorsen+19, Saha+20, Marques-Chaves+24,…

★ LyC emitters are diverse and multiple properties correlate with LyC escape, the most important are:

Neutral gas absorption
Dust attenuation
Nebular ionization
Concentrated star formation —> Feedback is key

(See e.g. Izotov+18, Verhamme+16; Flury+22b, Saldana-Lopez+22, Chisholm+22, Leclerq+24, Bait+24, Jaskot+24, Mascia+25)

★ Radiative feedback for the youngest starbursts with the highest LyC escape
★ SNe feedback also contributes and is possible crucial in not so young bursts

(See e.g. Komarova+19, Amorín+24, Bait+24, Jecmen & Oey 24; Carr+25)

★ Indirect diagnostics (e.g. O32, Lya, etc..) are promising but need to understand how the properties of LyC emitting galaxies evolve from low to high redshifts

Our favourite galactic labyrinth: Extreme emission line galaxies

Definitions and science motivations

Strong (and possibly short) episodes of ongoing star formation.

High equivalent widths (EWs $\gtrsim$200-2000A in [OIII]5007 and Hα) — Unusually large specific star formation rates (sSFRs) up to 10-100 Gyr^{-1} (e.g. van del Wel+11, Amorin+15)

Galaxies with low stellar masses ($\lesssim$ 10^{9}Msun), with low metallicities ($\lesssim$20%), high ionization conditions and little dust attenuatin / extinction—

Extremely rare at low redshifts (Green Pea - Blueberries- HII galaxies - BCD) — but increasingly common towards high-redshifts (e.g. Atek+14, Maseda+17 and others..)

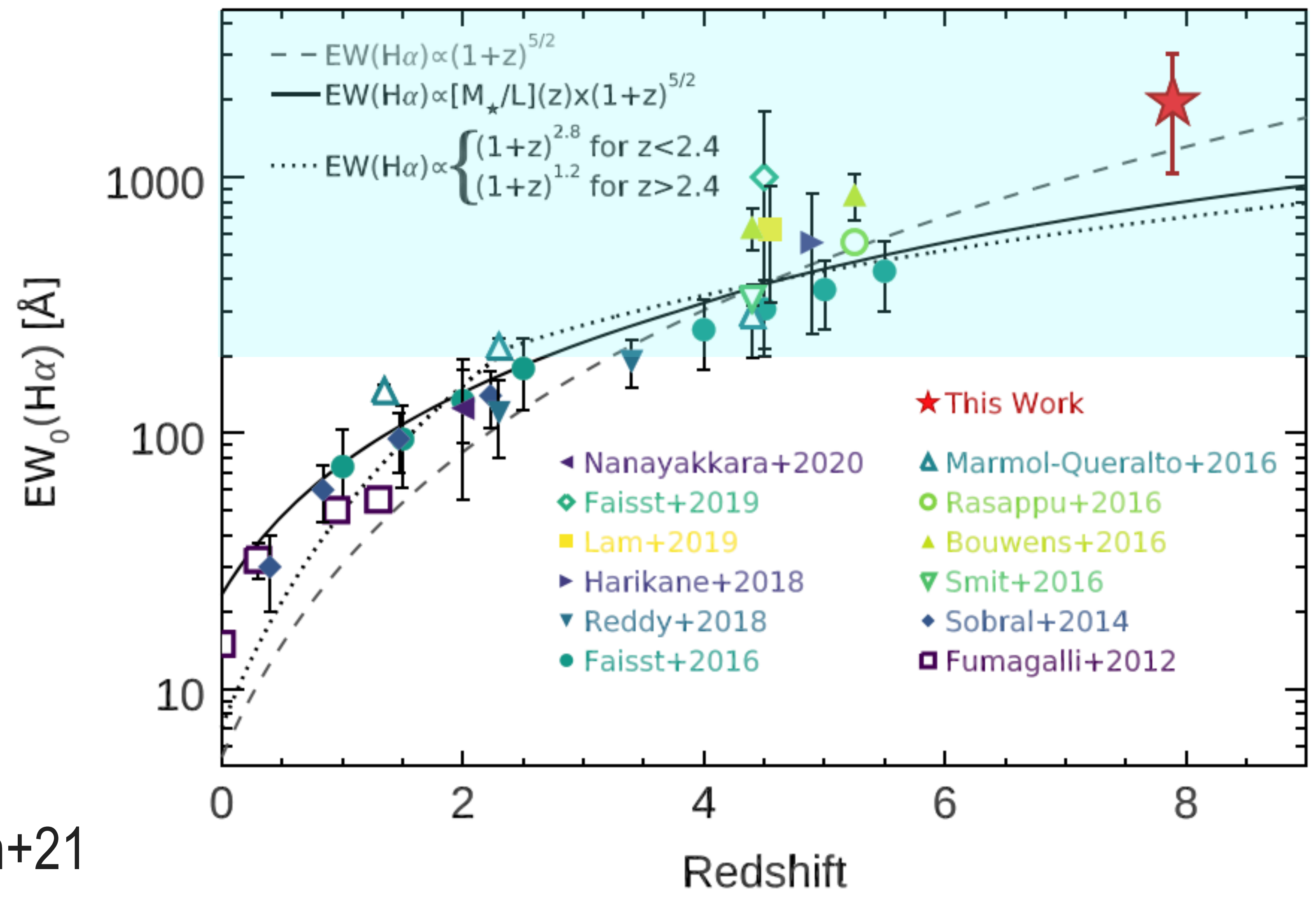
EELGs at z>4 are ubiquitous in JWST surveys conducted so far (e.g Llerena+24; Davis+24; Boyett+24, Atek+24, Trump+23; Cameron+23; Saxena+24, Heintz+24, Matthee+24; etc..)

— Tracers of major phases of galaxy growth in low-mass systems

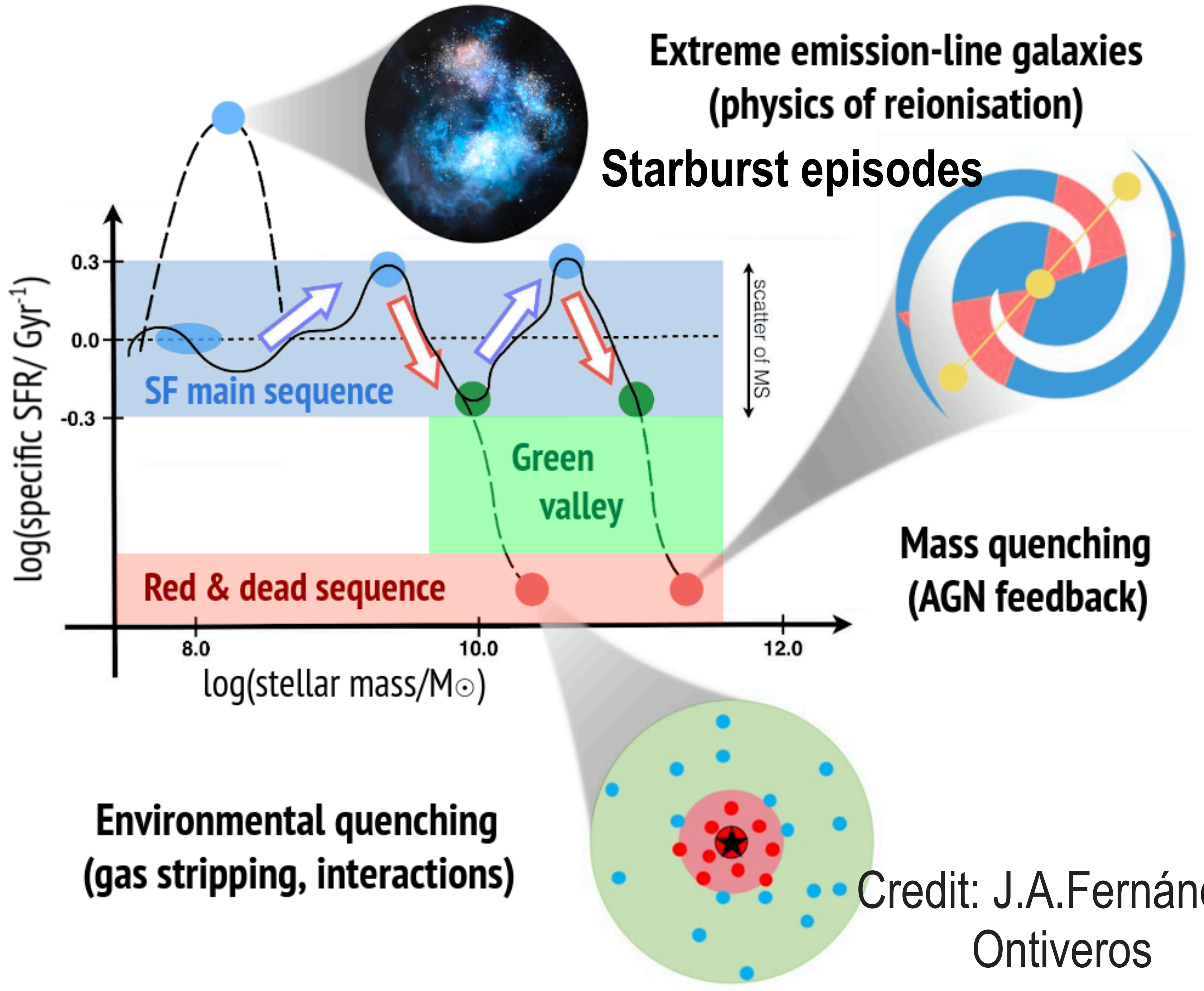
— Ideal laboratories to study star formation & feedback and chemical evolution

— EELGs are expected to be key contributors to cosmic reionization at z>5-6

— EELGs appear to have similar ISM properties across cosmic time (e.g. Schaerer+22)



Stefanon+21



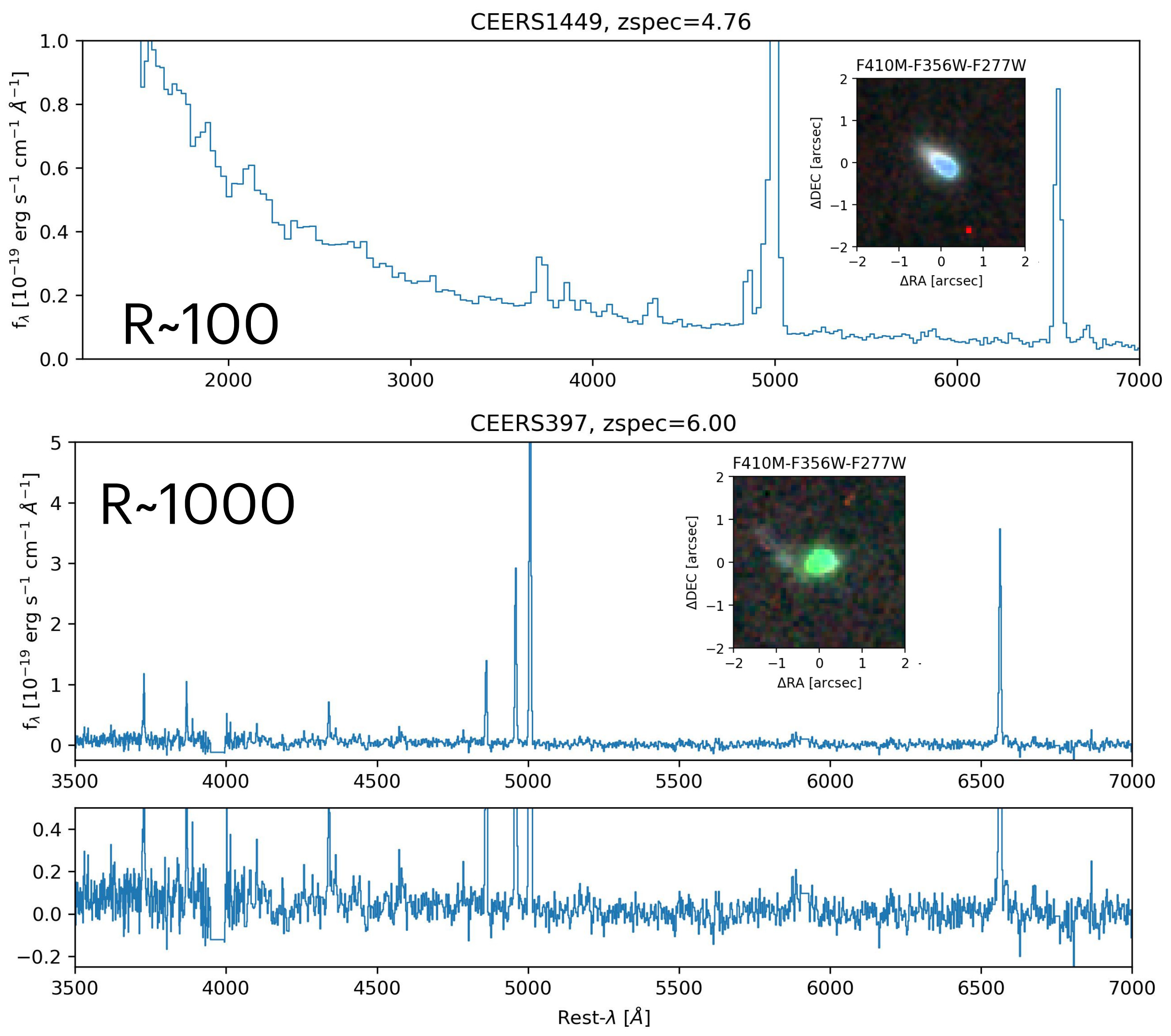
Credit: J.A.Fernández Ontiveros

Extreme emission line galxies

Definitions and science drivers

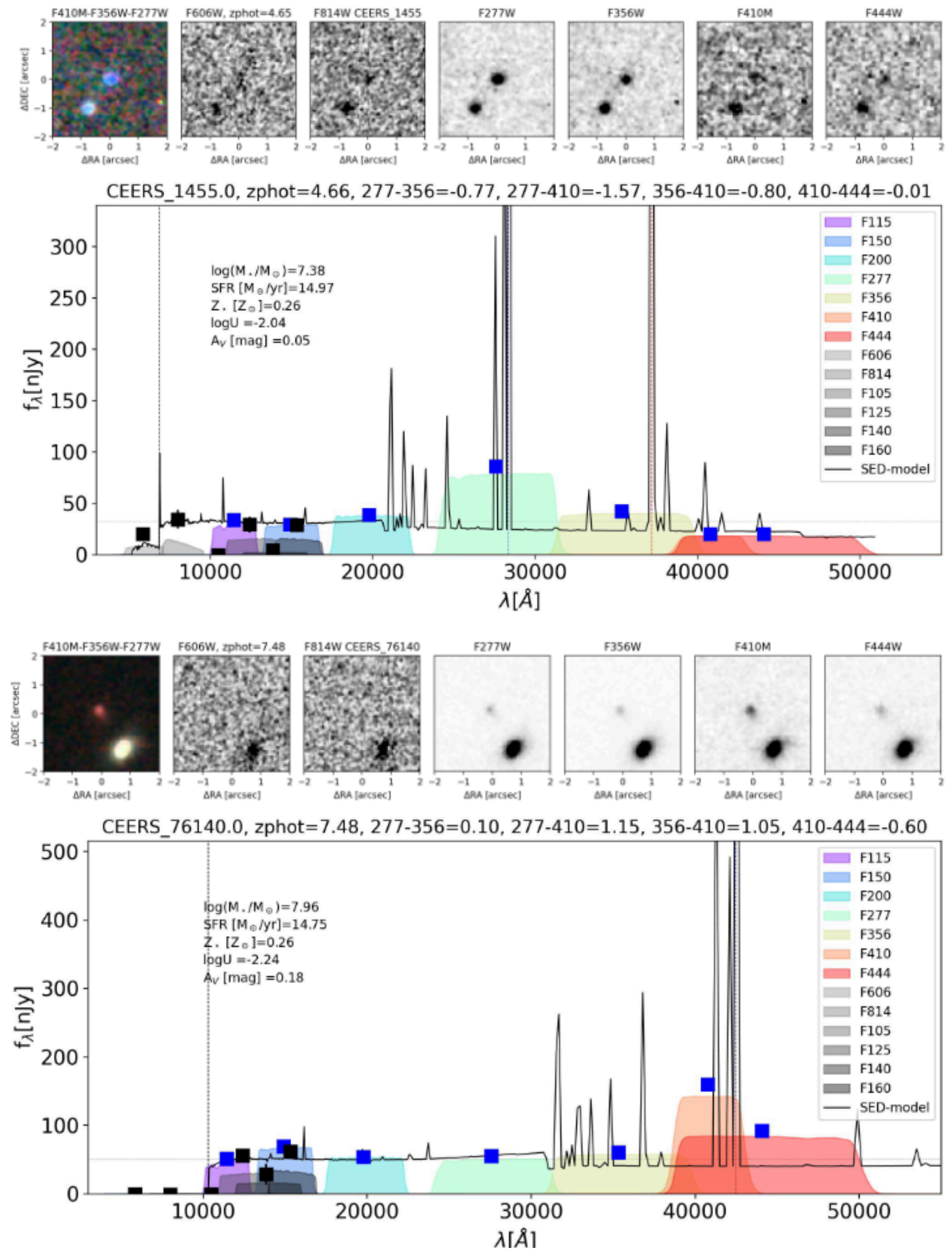
Many of the searches for EELGs select samples using either **narrow**- (e.g. Sobral+2013; Iglesias-Páramo+2022;Lumbreras-Calle+2022) **broad**- (e.g. van der Wel+2011;Onodera+2020; Kojima+2020; Chen+2023; Davis+2023) or **medium**-band **photometry** (e.g. Cohn+2018;Withers+2023; Simmonds+2023), and **slitless** (e.g. Atek+14;Maseda+2018; Kashino+2023) and **IFU** (e.g. del Moral-Castro+2024) **spectroscopy**.

Now feasible up from z~2 to z~12 with JWST



EELGs z~4-9 from JWST NIRCam & NIRSpec in CEERS

Llerena, Amorín & CEERS team (A&A, 2024)



Large samples are required to characterize EELG population and understand their physical properties at different redshifts

XTREM-DESI: EELG selection from ASK classification

Automatic Spectroscopic K-means clustering using EDR spectra

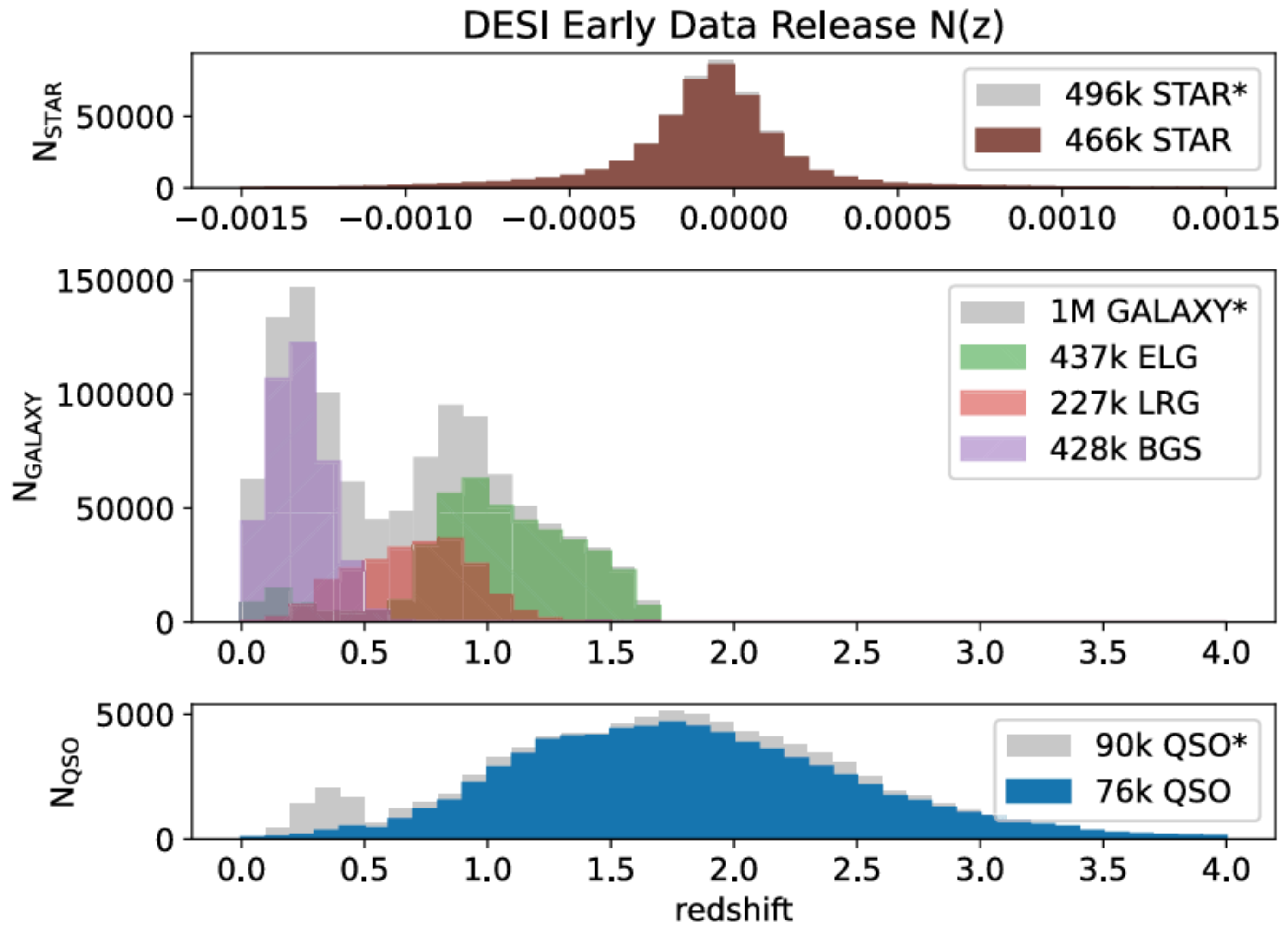
Inspired in the works by Sanchez Almeida et al. (2010, 2011, 2012) for SDSS-DR7 ASK spectral classification leading to a large sample of 2000 EELGs by Pérez-Montero, Amorín+2021

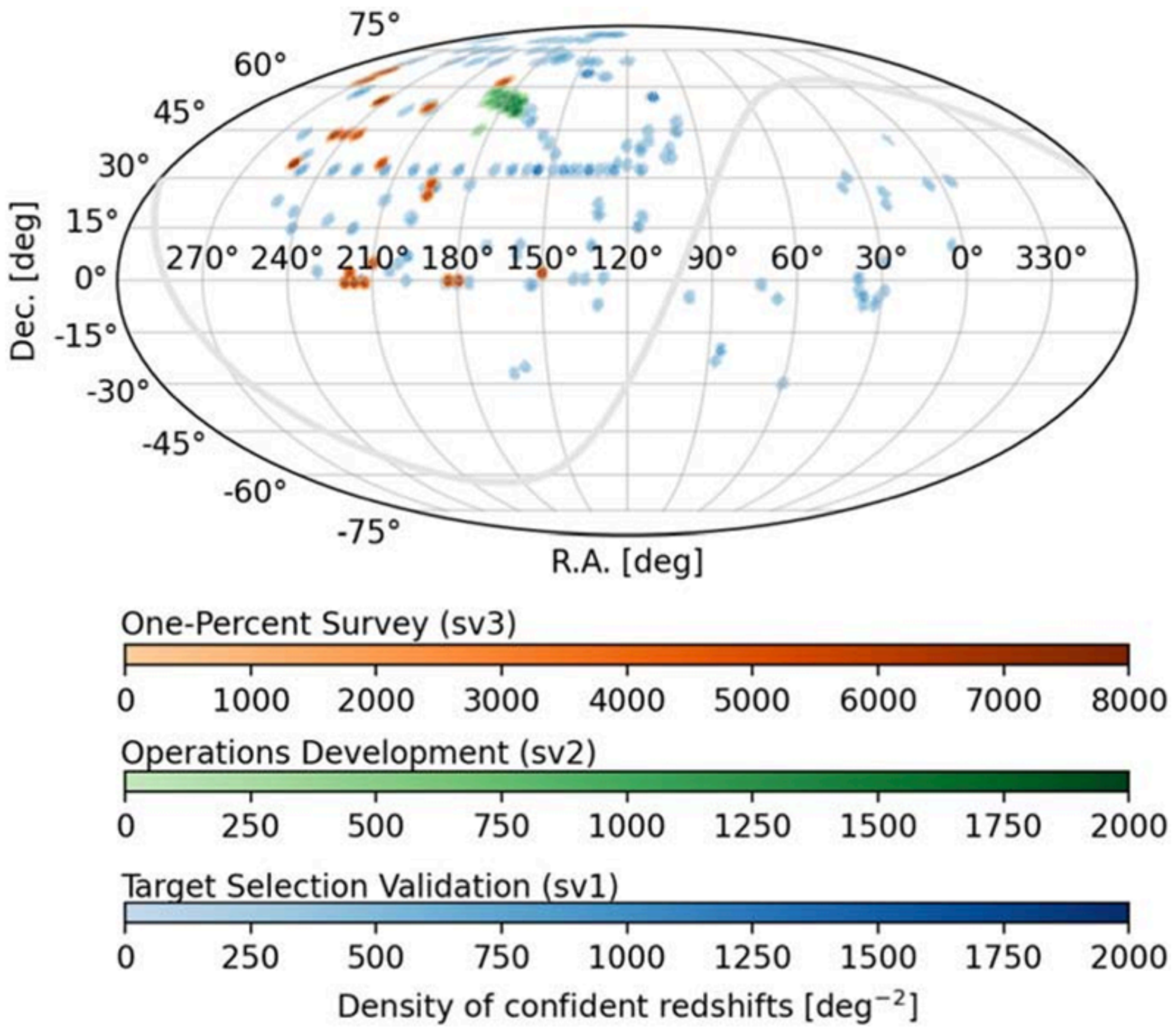
Idea: How to efficiently select rare populations, such as EELGs, without having to measure the whole release

Solution: Our K-means method is good to filter out outliers, i.e. galaxies that fall off the main spectral classes

1. Dataset selection and pre-processing: 2M spectra; ~550k Galaxy spectra (z<0.96)
2. Initialization method of ASK classification (250k spectra; z<0.25)
3. ASK classification using initial centroids as starting centers of classes
4. Quality assignation: assigning several classes to each galaxy

Amorín & Bonatto+25, in prep





XTREM-DESI: EELG selection from ASK classification

Automatic Spectroscopic K-means clustering using EDR spectra

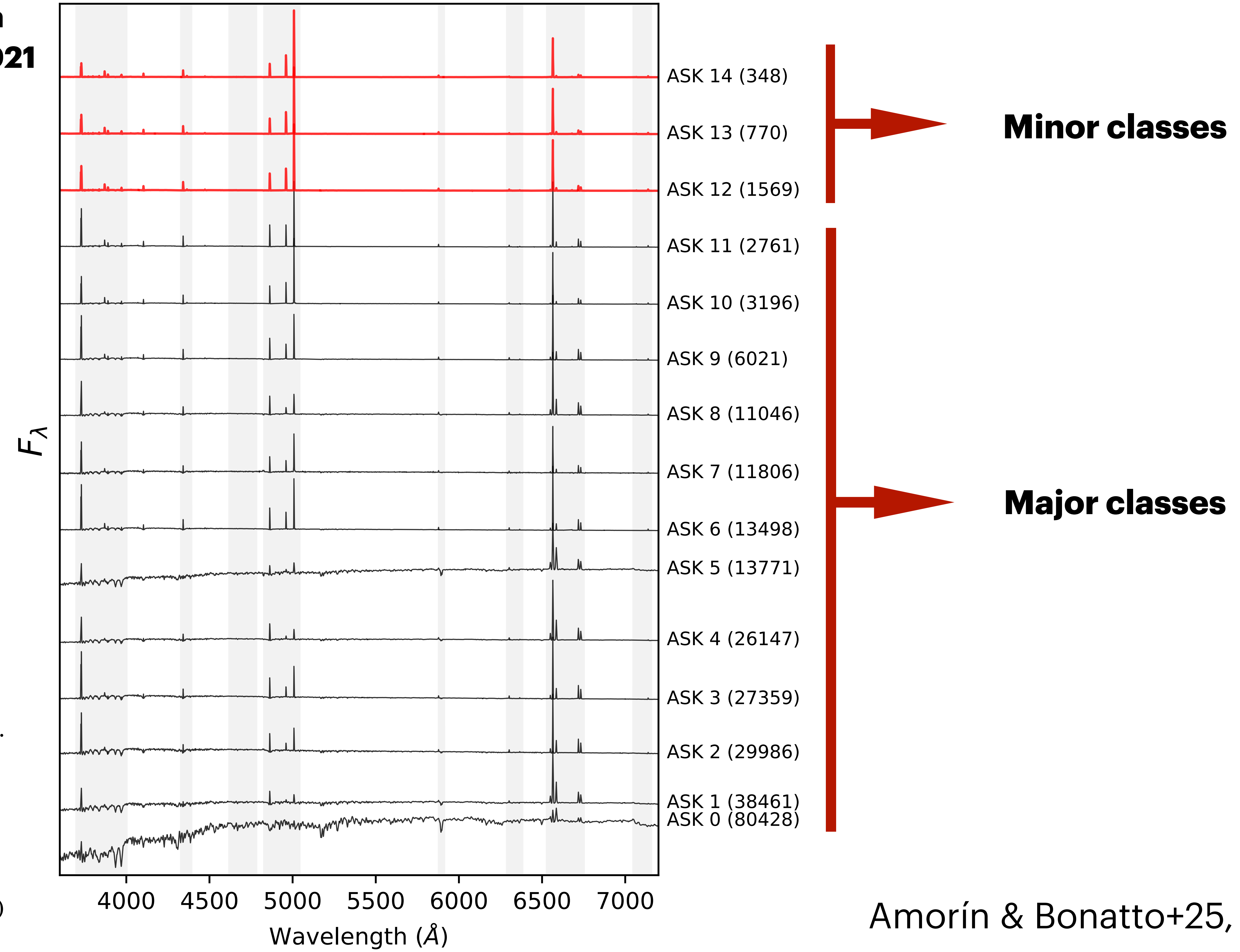
Inspired in the works by Sanchez Almeida et al. (2010, 2011, 2012) for SDSS-DR7 ASK spectral classification leading to a large sample of 2000 EELGs by Pérez-Montero, Amorín+2021

1. Dataset selection and pre-processing: 2M spectra; ~550k Galaxy spectra (z<0.96)
2. Initialization method of ASK classification (250k spectra; z<0.25)
3. ASK classification using initial centroids as starting centers of classes
4. Quality assignation: assigning several classes to each galaxy

Original k-means clusters: When the clusters were defined, the k-means algorithm partitioned the training dataset (redshift range 0-0.25) into k clusters.

Membership assignment: Each galaxy outside the redshift range must be evaluated against all k cluster centers to compute its distance and resulting merit function Q (c,**s**,**c_k**).

Interpretation: Each Q value represents the merit (or membership probability) of the galaxy belonging to one of the k clusters.



Amorín & Bonatto+25, in prep

XTREM-DESI: EELG selection from ASK classification

Automatic Spectroscopic K-means clustering using EDR spectra

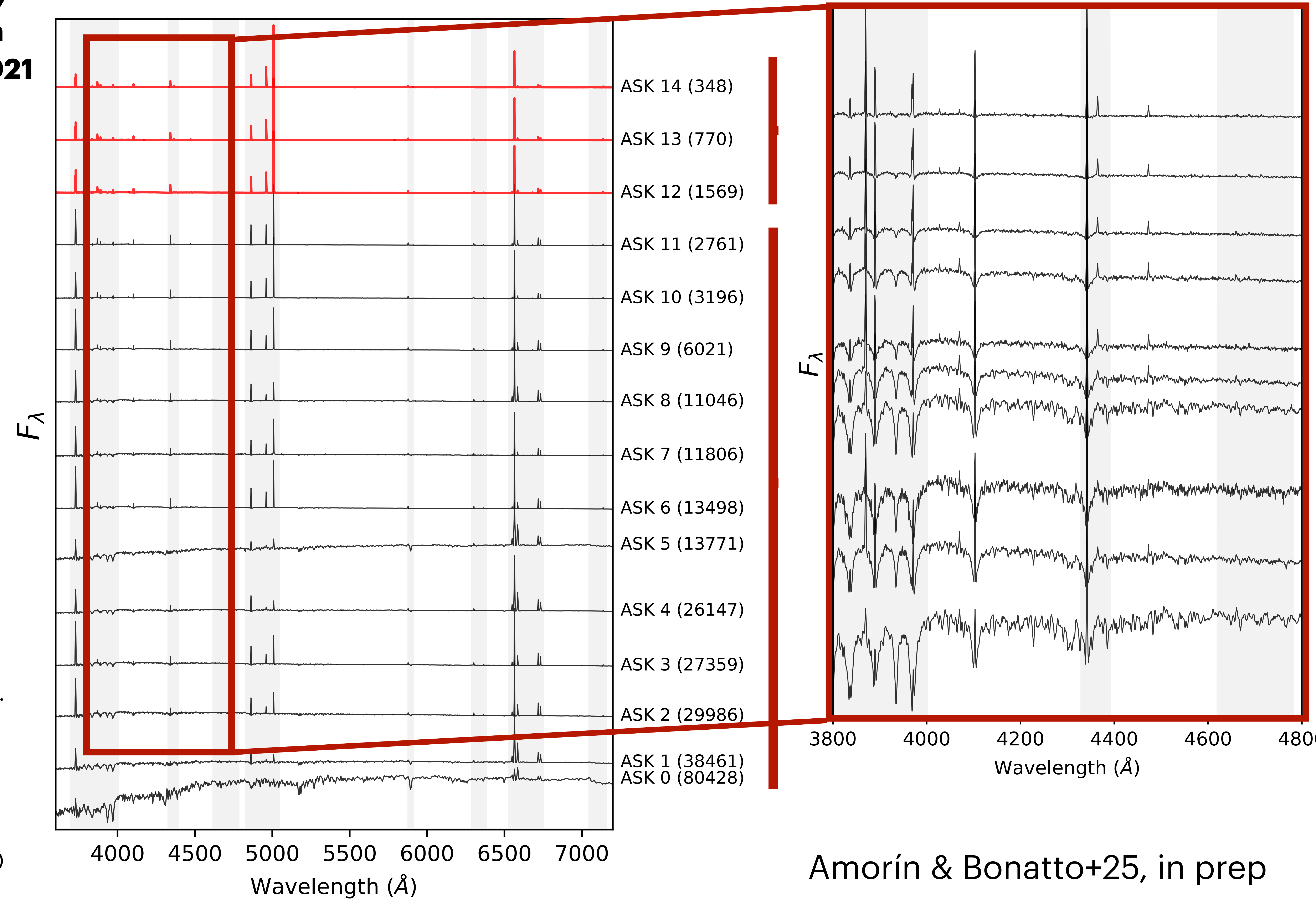
Inspired in the works by Sanchez Almeida et al. (2010, 2011, 2012) for SDSS-DR7 ASK spectral classification leading to a large sample of 2000 EELGs by Pérez-Montero, Amorín+2021

1. Dataset selection and pre-processing: 2M spectra; ~550k Galaxy spectra (z<0.96)
2. Initialization method of ASK classification (250k spectra; z<0.25)
3. ASK classification using initial centroids as starting centers of classes
4. Quality assignation: assigning several classes to each galaxy

Original k-means clusters: When the clusters were defined, the k-means algorithm partitioned the training dataset (redshift range 0-0.25) into k clusters.

Membership assignment: Each galaxy outside the redshift range must be evaluated against all k cluster centers to compute its distance and resulting merit function Q (c,**s**,**c_k**).

Interpretation: Each Q value represents the merit (or membership probability) of the galaxy belonging to one of the k clusters.



Amorín & Bonatto+25, in prep

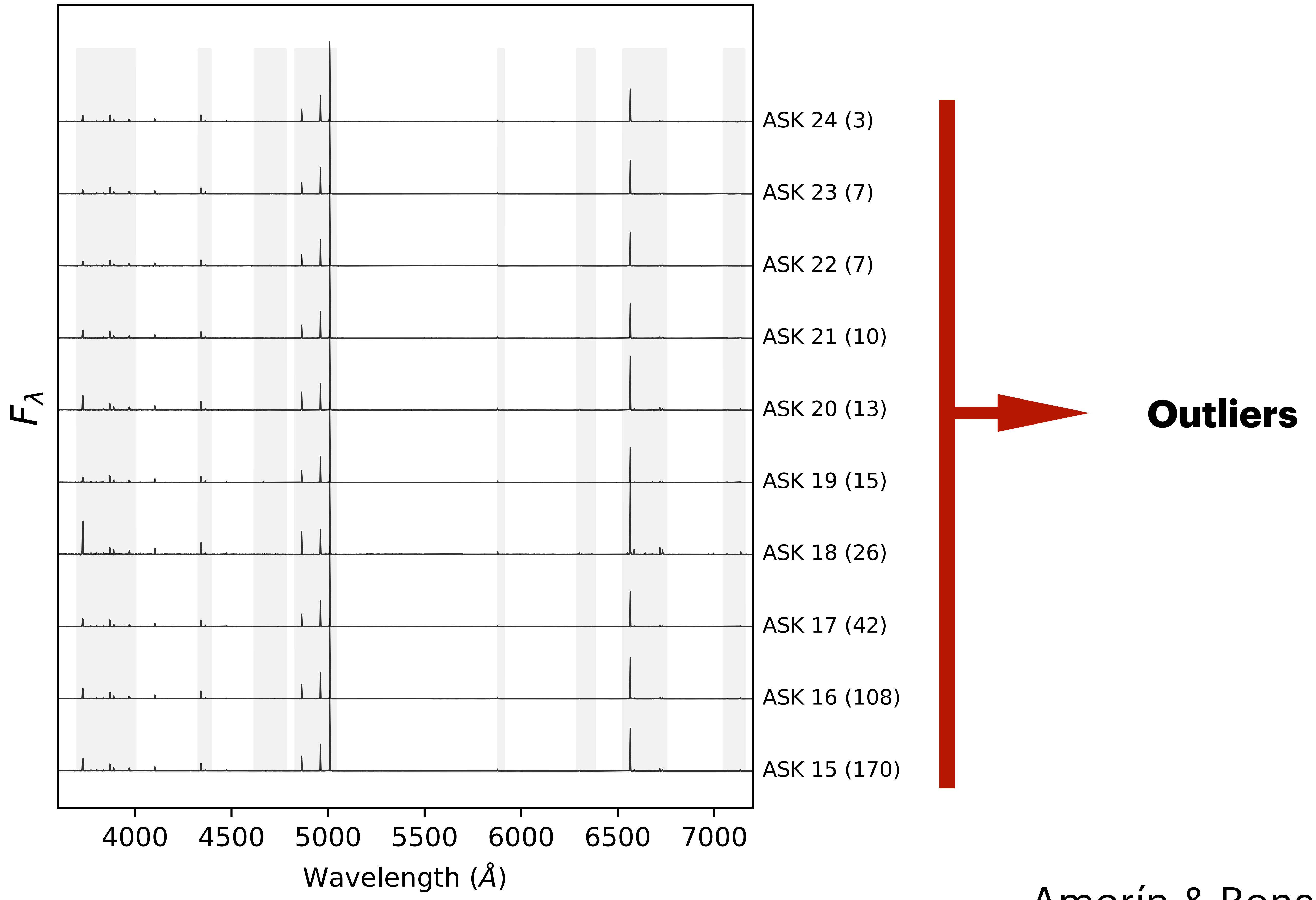
XTREM-DESI: EELG selection from ASK spectral analysis

Automatic Supervised K-means clustering method using EDR spectra

Inspired in the works by Sanchez Almeida et al. (2010, 2011, 2012) for SDSS-DR7 ASK spectral classification leading to a large sample of 2000 EELGs by Pérez-Montero, Amorín+2021

1. Dataset selection and pre-processing: 2M spectra; ~550k Galaxy spectra (z<0.96)
2. Initialization method of ASK classification (250k spectra; z<0.25)
3. ASK classification using initial centroids as starting centers of classes
4. Quality assignation: assigning several classes to each galaxy

OUR ASK METHOD IS AN EXCELLENT TOOL FOR DETECTION OF OUTLIERS, I.E. GALAXIES WITH RARE SPECTRAL FEATURES



Amorín & Bonatto+25, in prep

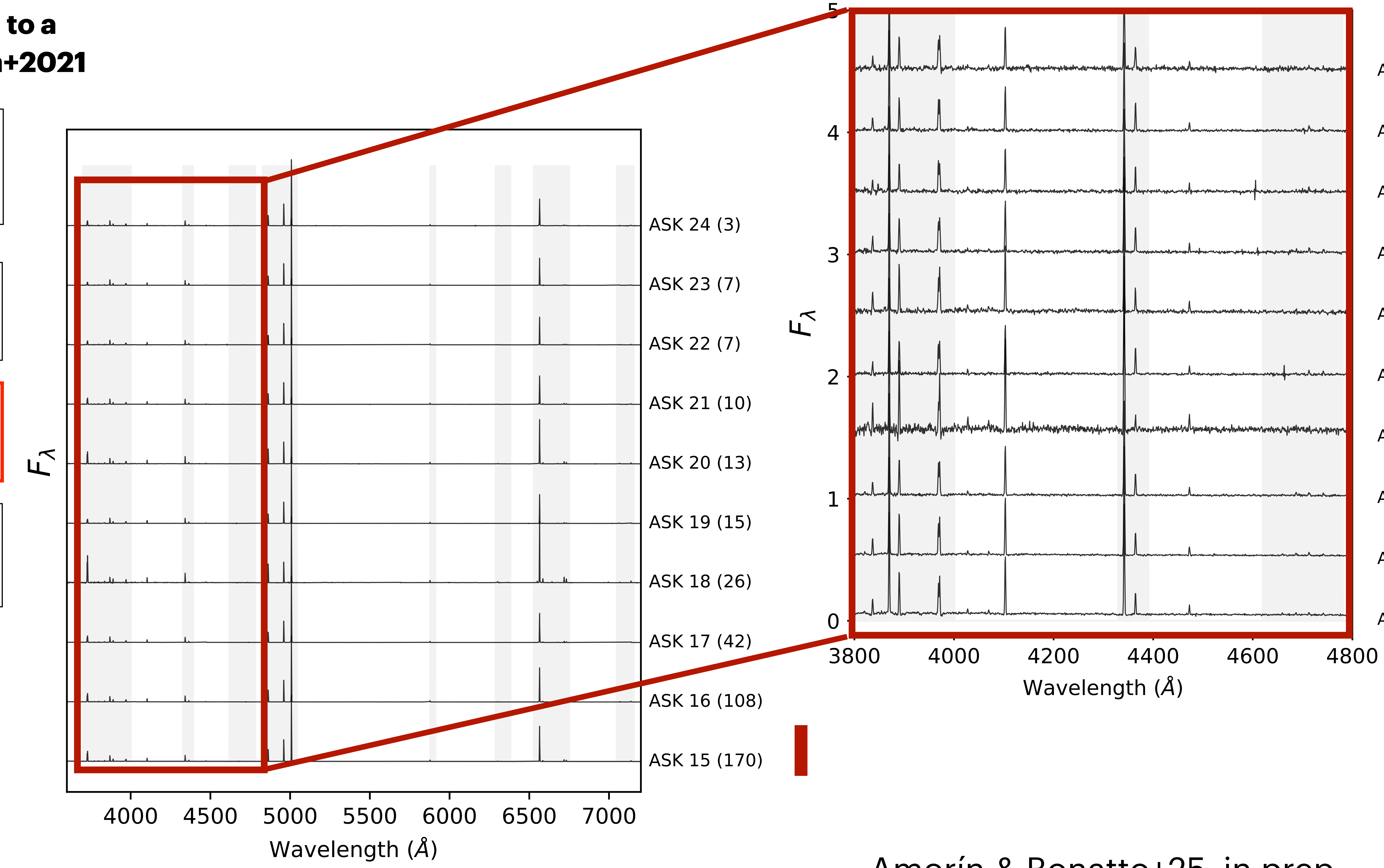
XTREM-DESI: EELG selection from ASK spectral analysis

Automatic Supervised K-means clustering method using EDR spectra

Inspired in the works by Sanchez Almeida et al. (2010, 2011, 2012) for SDSS-DR7 ASK spectral classification leading to a large sample of 2000 EELGs by Pérez-Montero, Amorín+2021

1. Dataset selection and pre-processing: 2M spectra; ~550k Galaxy spectra (z<0.96)
2. Initialization method of ASK classification (250k spectra; z<0.25)
3. ASK classification using initial centroids as starting centers of classes
4. Quality assignation: assigning several classes to each galaxy

OUR ASK METHOD IS AN EXCELLENT TOOL FOR DETECTION OF OUTLIERS, I.E. GALAXIES WITH RARE SPECTRAL FEATURES



Amorín & Bonatto+25, in prep

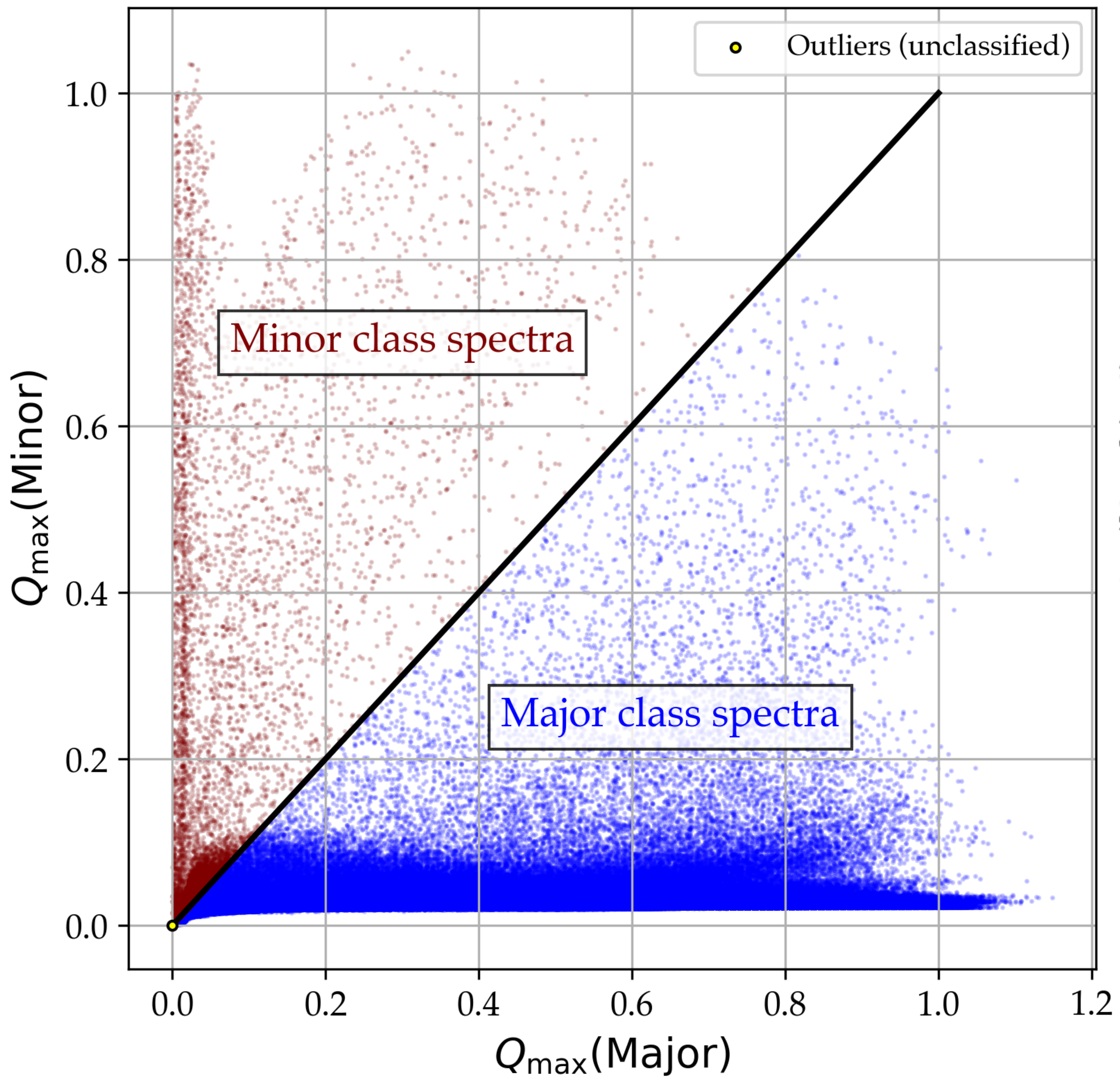
DESI ASK spectral analysis

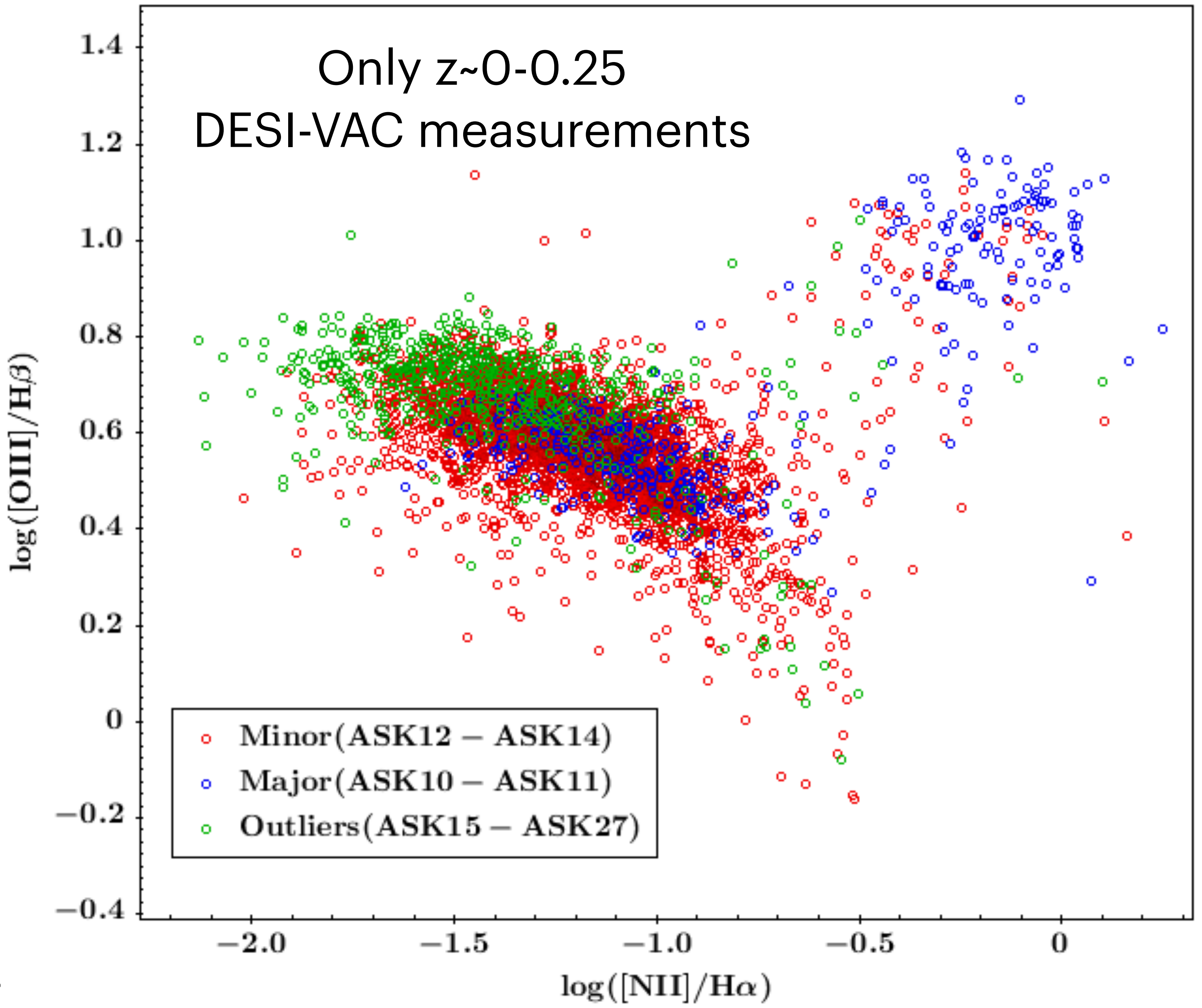
K-means clustering using EDR spectra

Inspired in the works by Sanchez Almeida et al. (2010, 2011, 2012) for SDSS-DR7 ASK spectral classification leading to a large sample of 2000 EELGs by Pérez-Montero, Amorín+2021

1. Dataset selection and pre-processing: 2M spectra; ~550k Galaxy spectra (z<0.96)
2. Initialization method of ASK classification (250k spectra; z<0.25)
3. ASK classification using initial centroids as starting centers of classes
4. Quality assignation: assigning several classes to each galaxy

OUR ASK METHOD IS AN EXCELLENT TOOL FOR DETECTION OF OUTLIERS, I.E. GALAXIES WITH RARE SPECTRAL FEATURES
—> $Q_{MAX}=0$





Amorín & Bonatto+25, in prep

The XTREM-20k: a sample of 20000 EELGs in DESI EDR

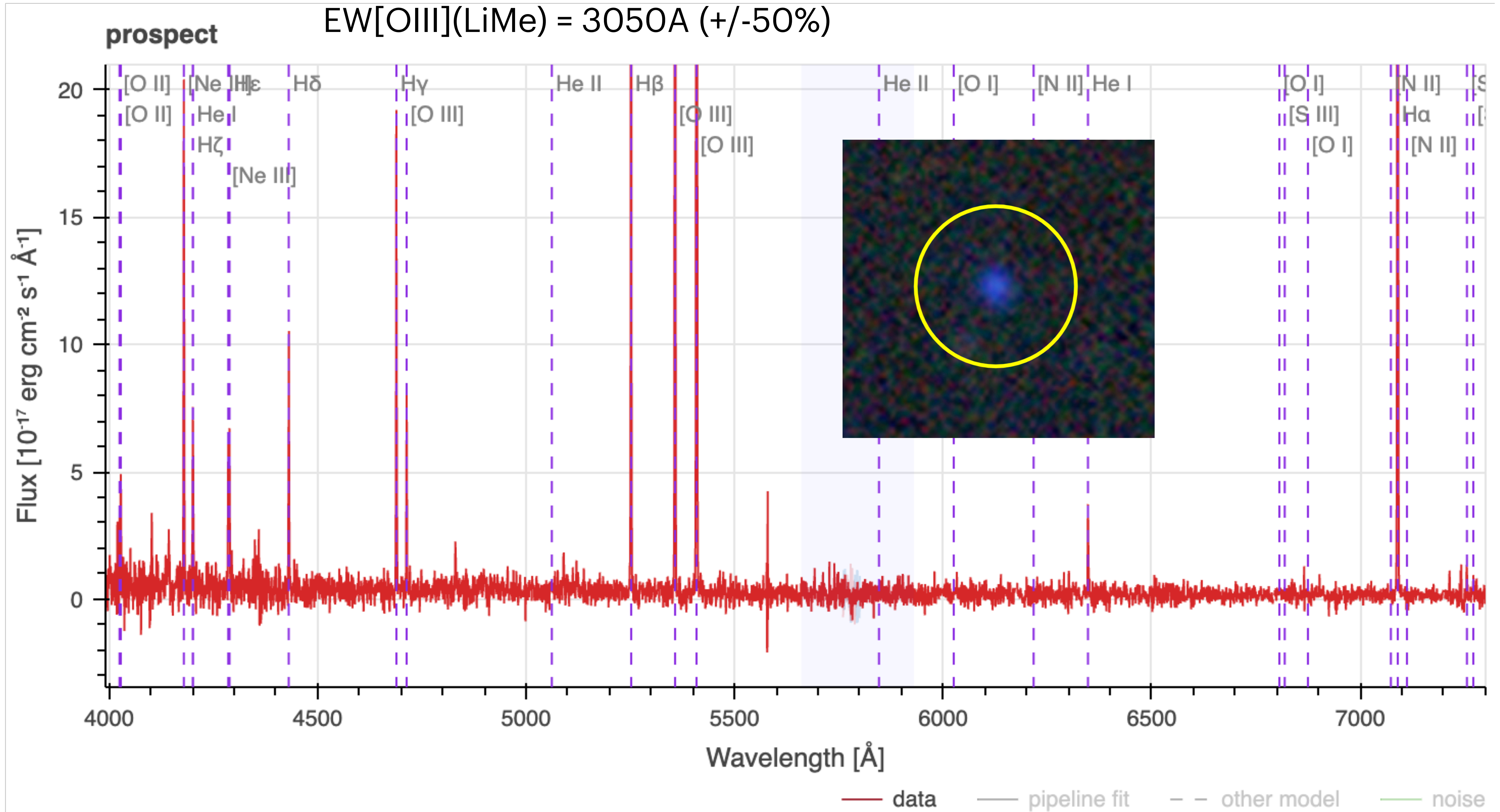
Compact low-mass starbursts

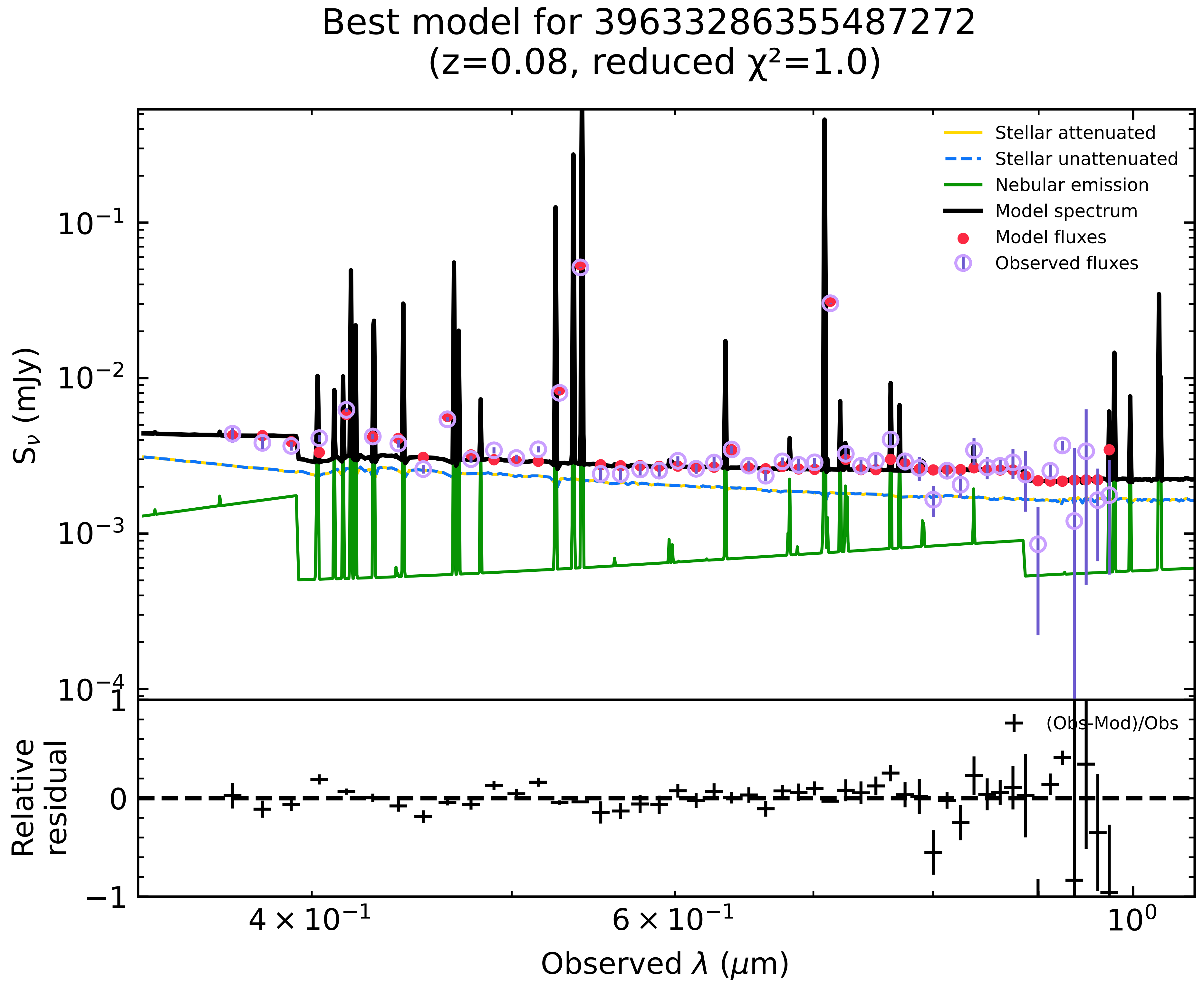
- A sample of ~20000 EELGs selected from our K-means method — ~10x larger than SDSS-DR17
- Emission line measurements using LiMe (Fernández et al. 2024)
    - Allows multi-gaussian line fitting
- Synthetic photometry from DESI spectra and SED fitting measurements using CIGALE
    - input model following Lumbreras+2022

$\log(M^*/M_\odot) = 6.8$, $M_u = -15.2$, $M_{W1} = -13.6$

EW([OIII] (CIGALE) = 1576A

EW[OIII](LiMe) = 3050A (+/-50%)

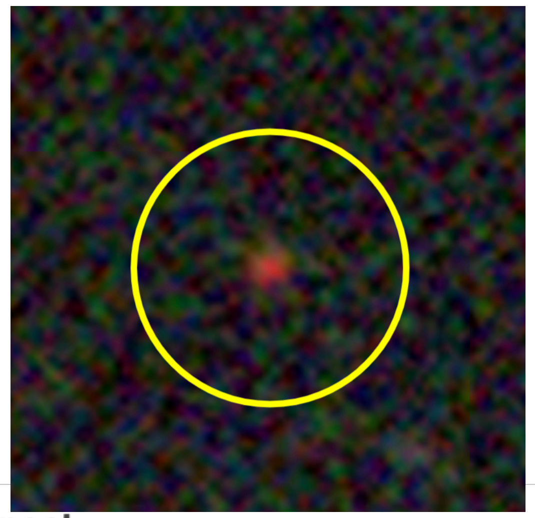




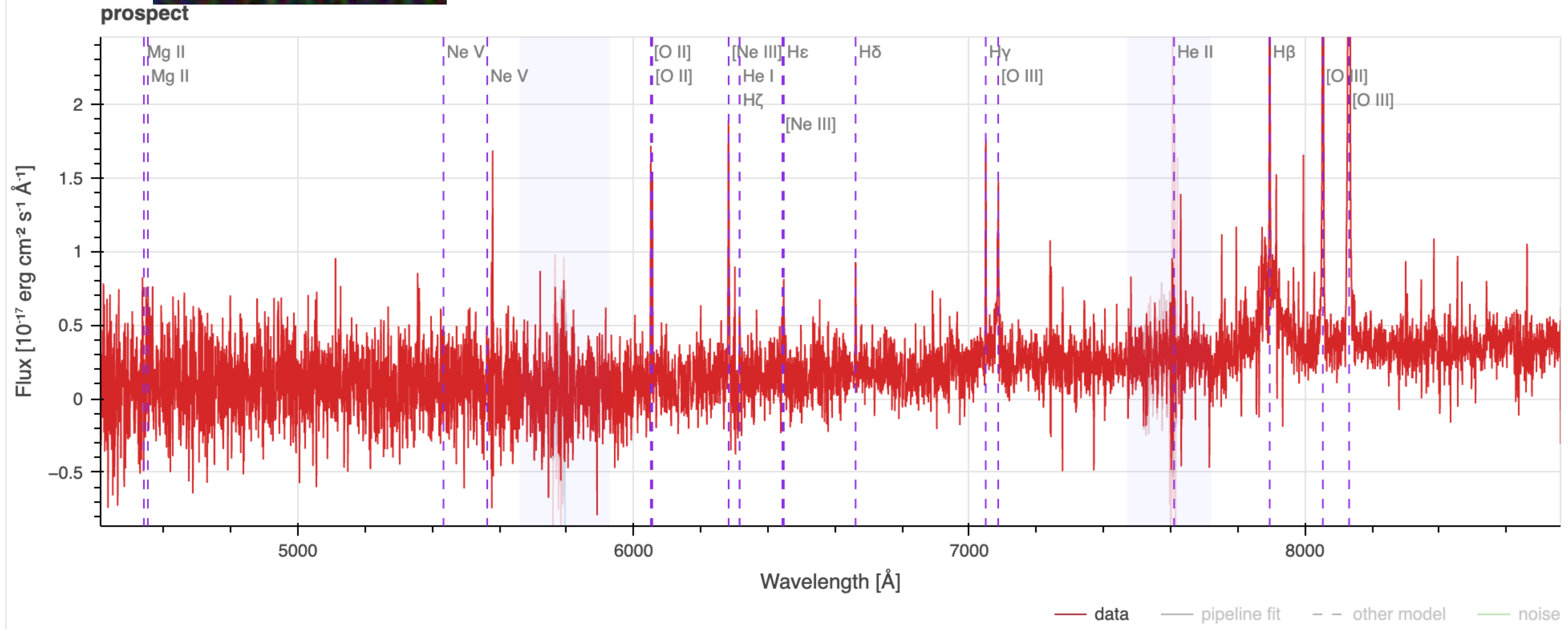
The XTREM-20k: a sample of 20000 EELGs in DESI EDR

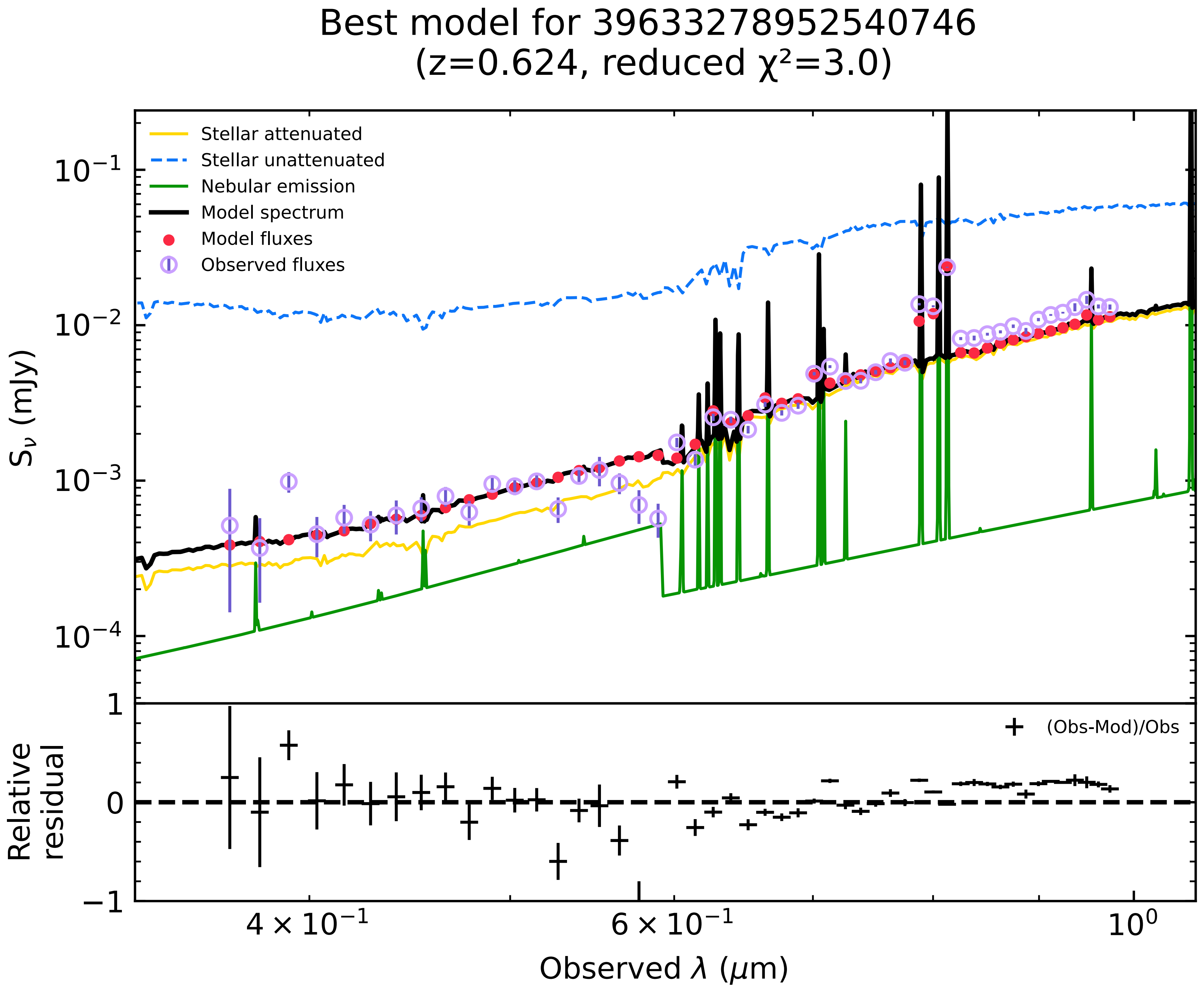
Rare (low-mass?) AGN types

- A sample of ~20000 EELGs selected from our K-means method — ~10x larger than SDSS-DR17
- Emission line measurements using LiMe (Fernández et al. 2024)
  - Allows multi-gaussian line fitting
- Synthetic photometry from DESI spectra and SED fitting measurements using CIGALE
  - input model following Lumbreras+2022



$\log(M^*/M_\odot)$ = 10.2 , M_u = -18.3 , M_{W1} = -22.1
EW([OIII] (CIGALE) = 220A
EW[OIII](LiMe) = 150A (+/-30%)



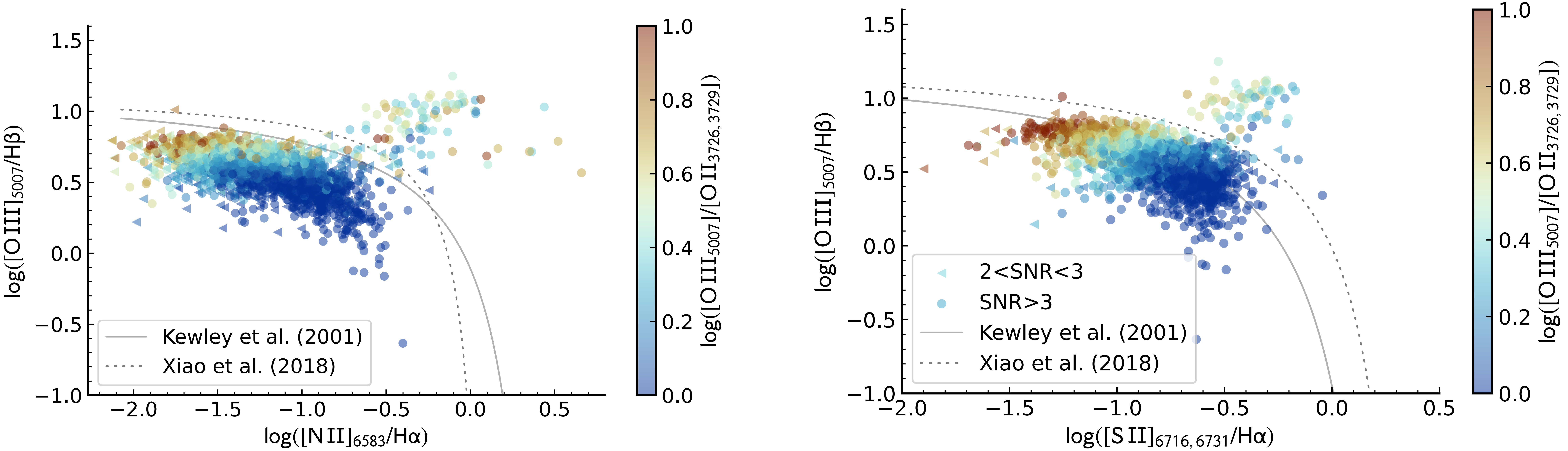


The XTREM-20k: a sample of 20000 EELGs in DESI EDR

Classic diagnostics

- A sample of ~20000 EELGs selected from our K-means method — ~10x larger than SDSS-DR17
- Emission line measurements using LiMe (Fernández et al. 2024)
  - Allows multi-gaussian line fitting
- Synthetic photometry from DESI spectra and SED fitting measurements using CIGALE
  - input model following Lumbreras+2022

BPT digrams for galaxies (only for z<0.4)



Amorín & Bonatto+25, in prep

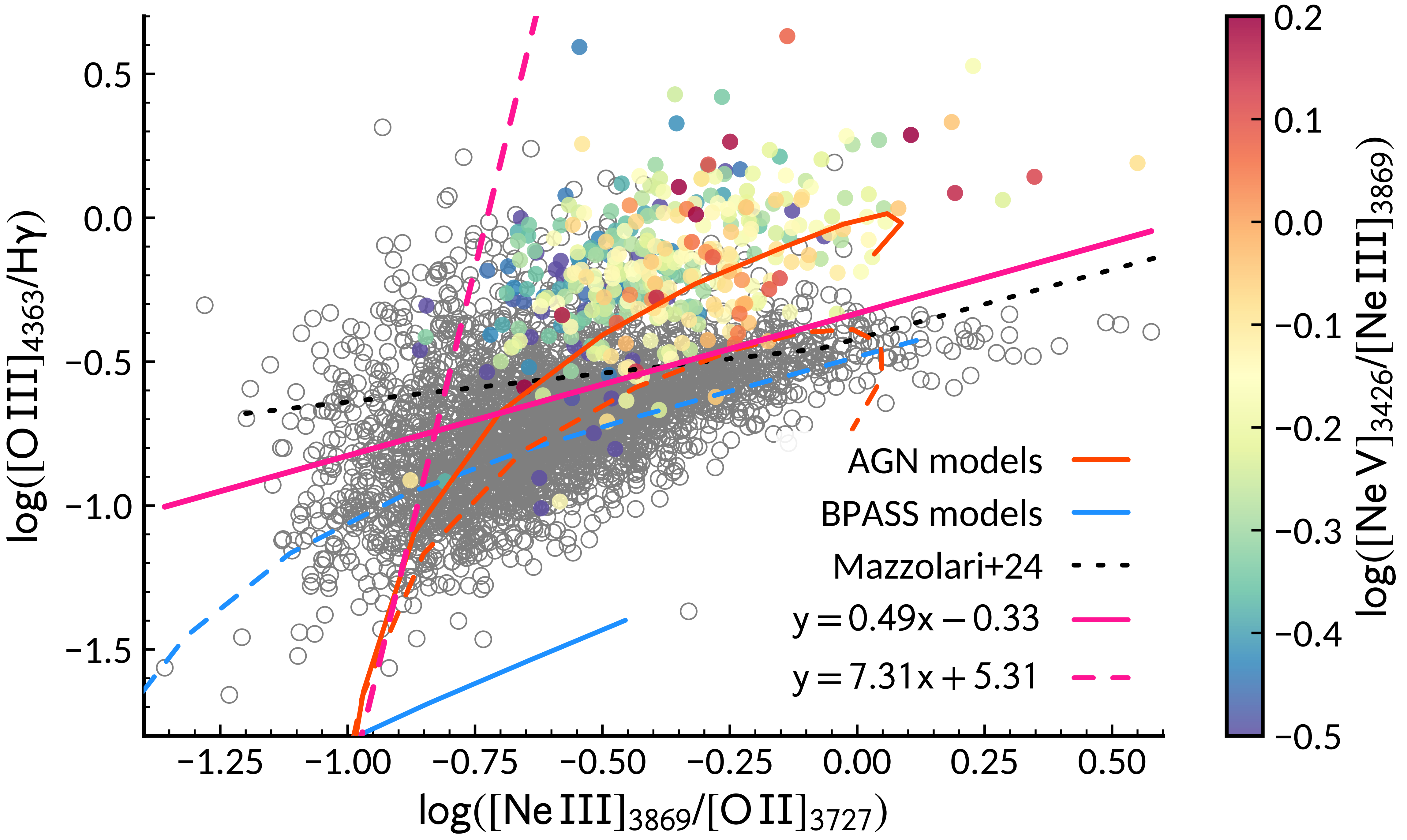
EELGs in DESI EDR

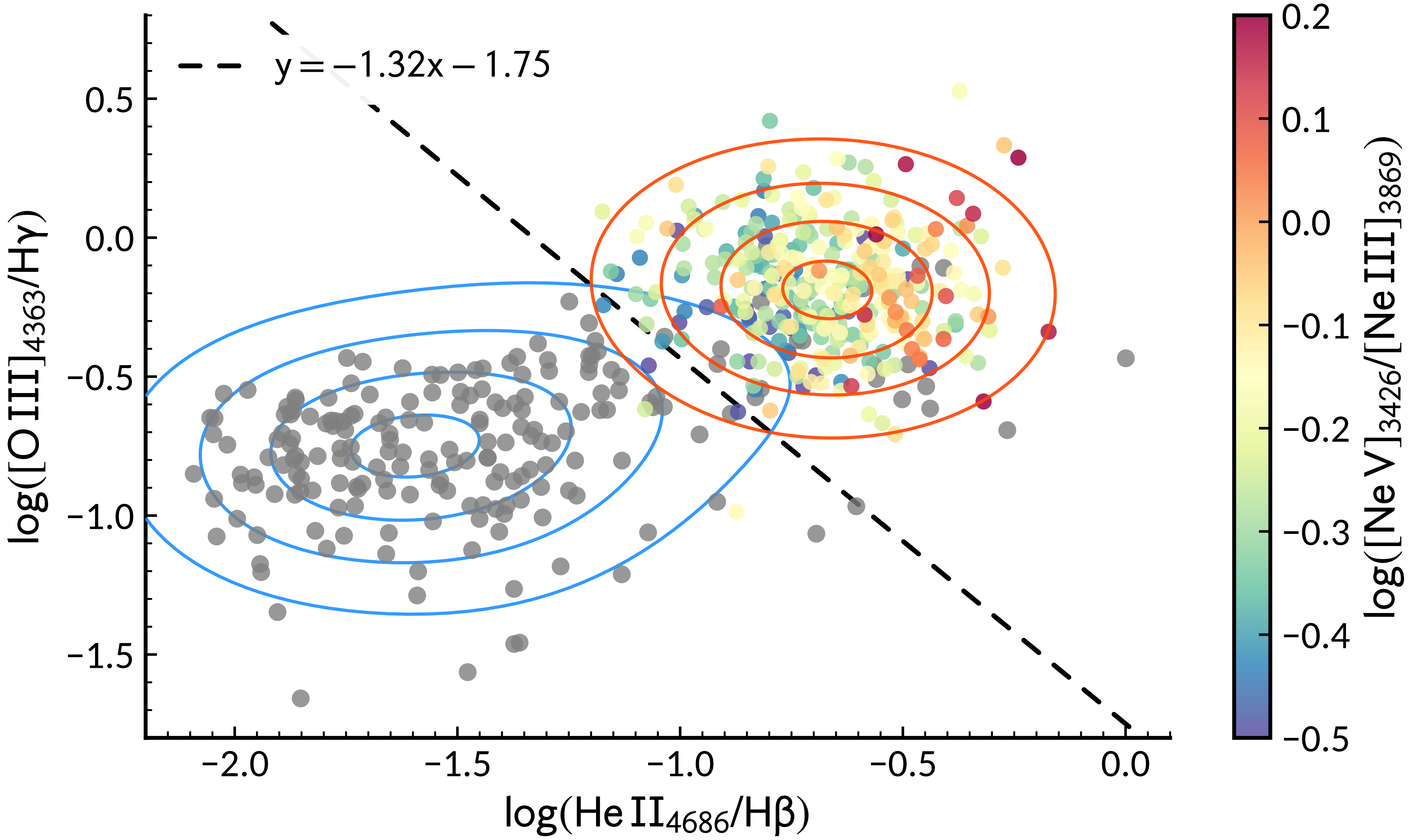
Disentangling hard radiation fields from NeV and HeII emitters

New diagnostics particularly useful for JWST high-z galaxies

e.g. Cleri+23; Mazzolari+24; Backhaus+25

OHNO and O3He2 digrams



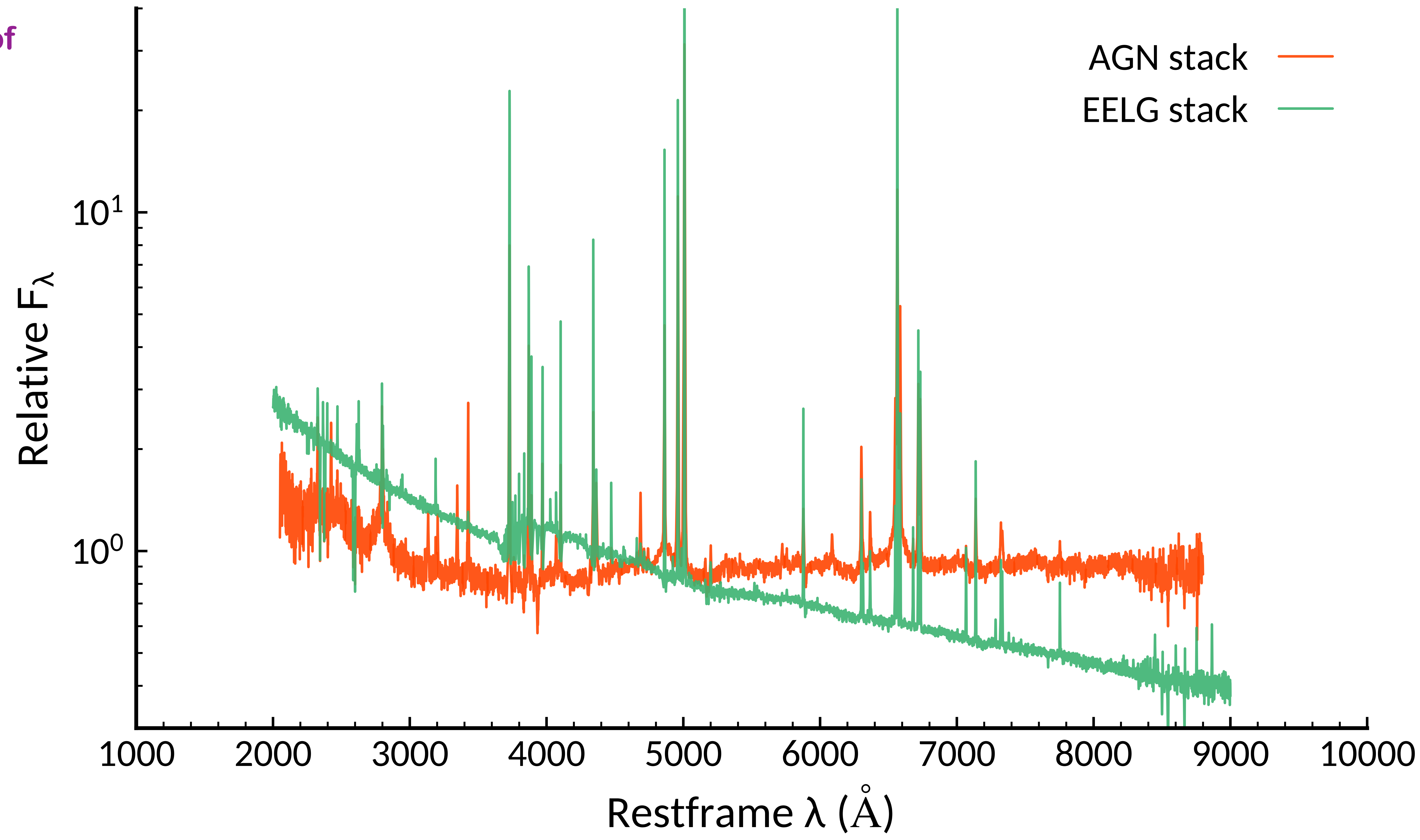


Fernandez-Ontiveros+, in prep
(See talk by Juan on Friday!)

The XTREM-20k: a sample of 20000 EELGs in DESI EDR

Median spectra

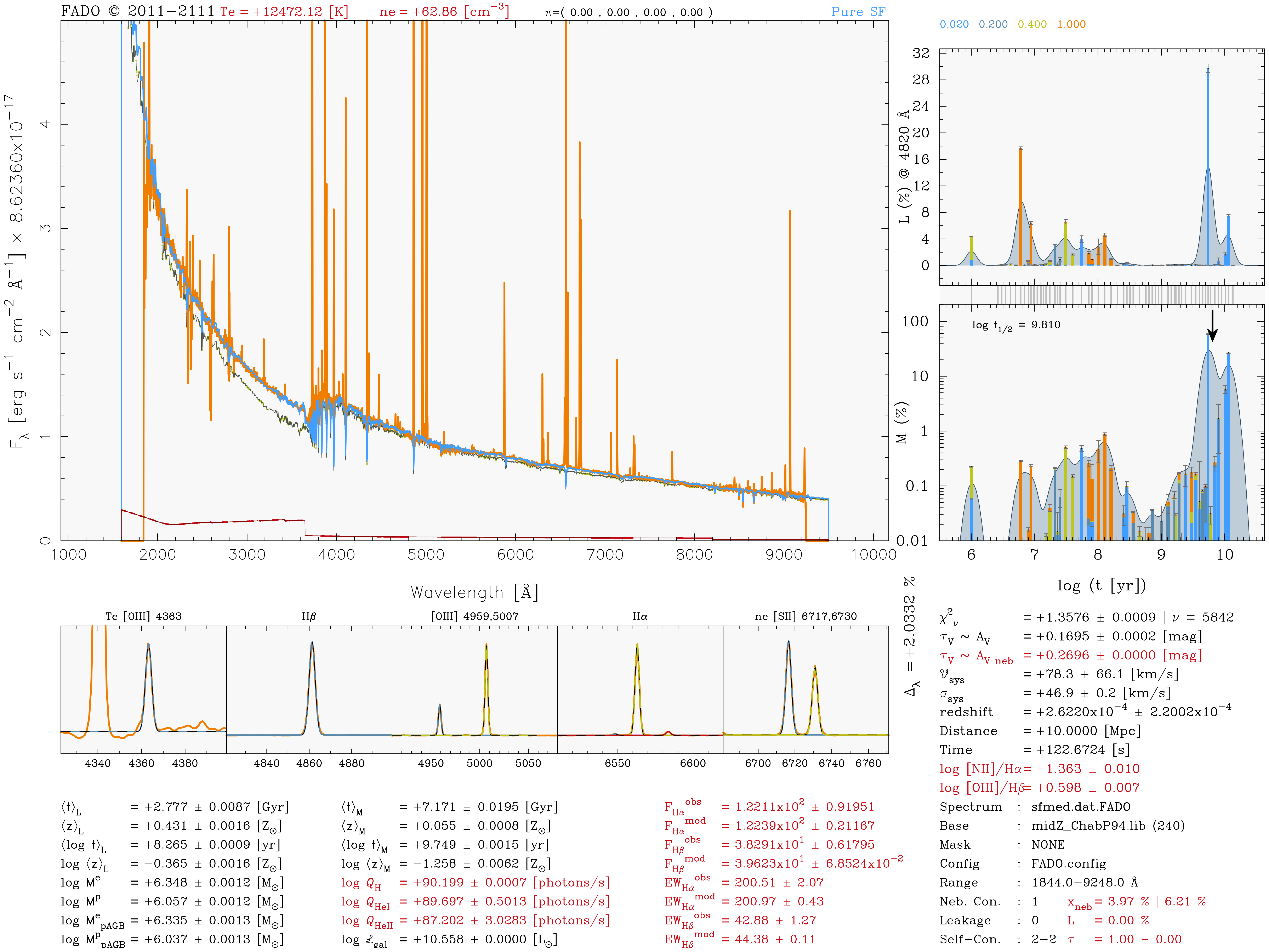
For a better description of the AGN-like EELGs see talk by Fernandez-Ontiveros on Friday!



The XTREM-20k: a sample of 20000 EELGs in DESI EDR

Median spectra: Spectral synthesis fitting with FADO

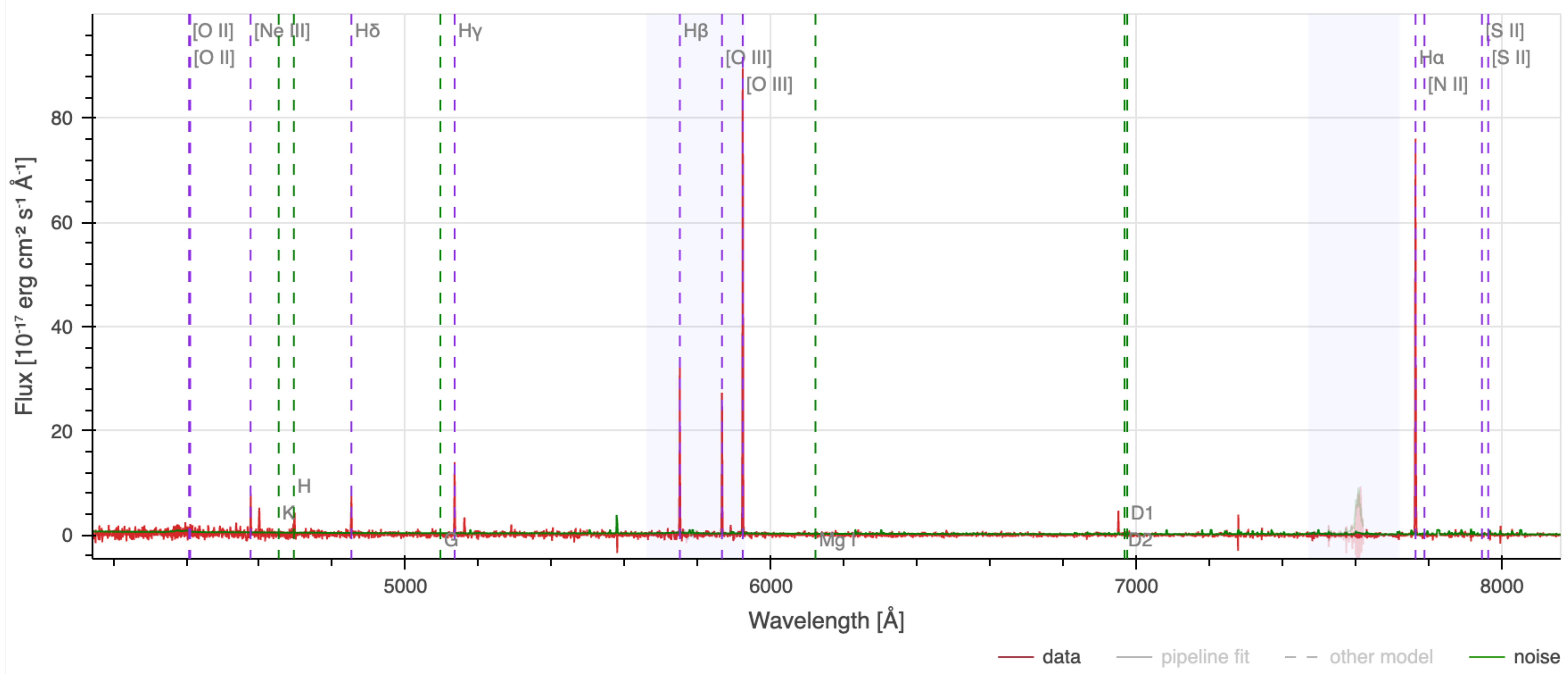
- Superb S/N in both lines and continuum allows detailed spectral synthesis
- FADO (Gomes & Papaderos 2016) is an powerful tool for such studies
- SF EELGs median properties consistent with average nearby HII/BCD/GPs in high S/N studies (e.g. Amorin+12; Breda+22; Fernandez+23)
    - Blue UV continuum
    - Significant contribution of nebular continuum
    - Bursty SFH



Lyman photon escape using XTREM-20k

Probing indirect method with our golden Te[OIII]4363 sample

- A sample of ~1500 EELGs at z~0.01 to 0.95 selected from XTREM-20k sample with SNR>3 in [OIII]4363 auroral line.
- ~75% belong to the "dark" survey —> fainter sources (~22-24 AB mag) —> low luminosity end: M_u ~ -13.5 to -18.5 mag
- Golden sample for metallicity study using direct method (Amorín+25, in prep)
  - T_e and N_e via PyNeb (Luridiana+15)
  - Direct metallicities (Perez-Montero+17)



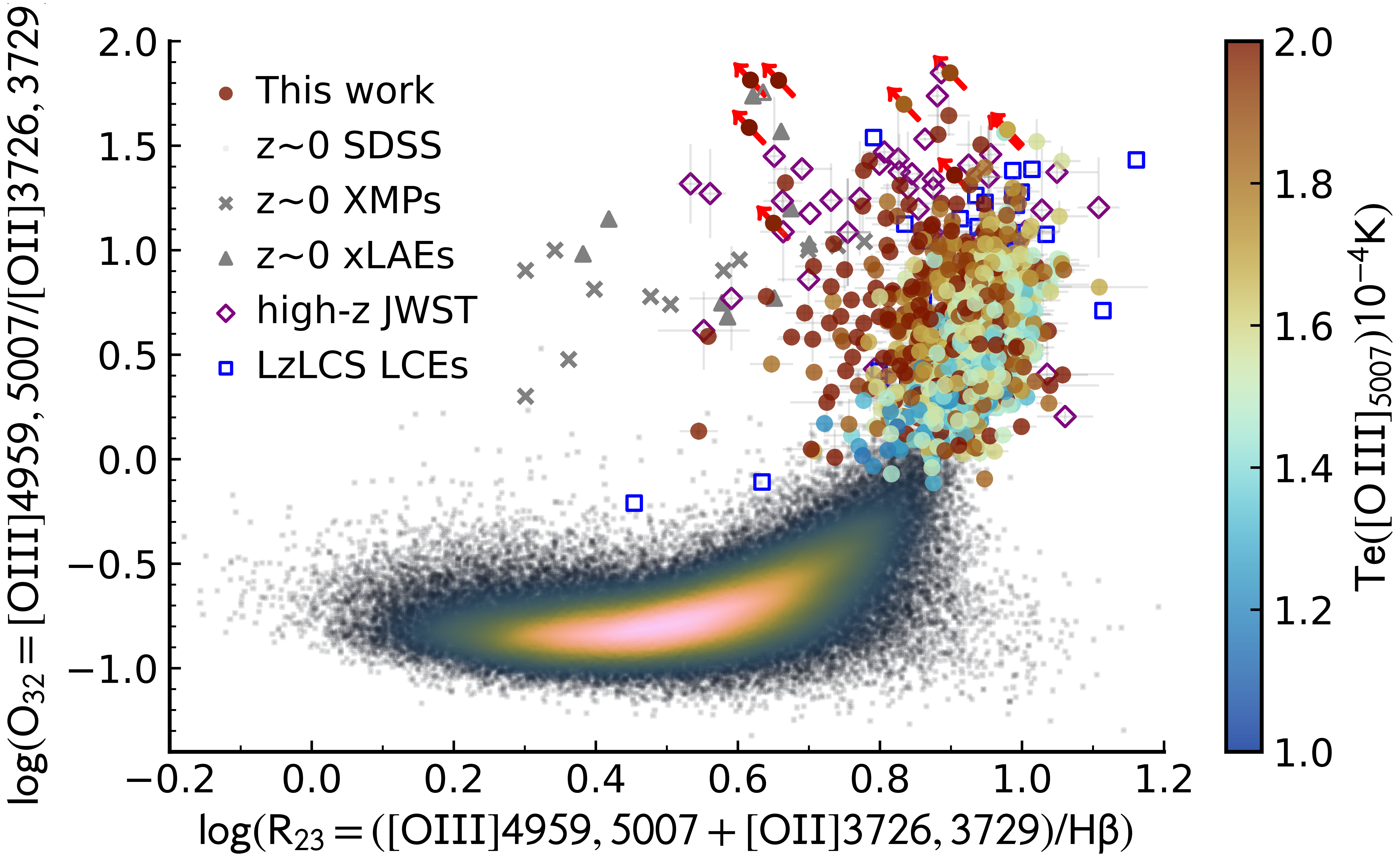
Gallery from deep Subaru HSC color tiles

- An unusual case (10 in 20k EELGs): very high Te and extreme O32, no low-excitation lines (Amorín+25, in prep)

XTREM ionization sample

Higher ionization at very low metallicity

- We find a unique sample of XMPs with unusually high ionization O32>10-70 and high Te
- Similar to the few local XMPs found as extreme LAE (xLAEs) by Izotov+24 from SDSS-DR16
- Excellent match to z>6 comparison samples from JWST (Cameron+23,24; Heintz+24; Saxena+24; Trump+23; Arrabal-Haro+24; Mowla+24; Willot+25; Cullen+25) — Selected to match metallicity range, with Te estimates
- Different from classic XMPs (IZw18, SBS0335-052, LeoP, etc..) typically show lower O3O2 and R23

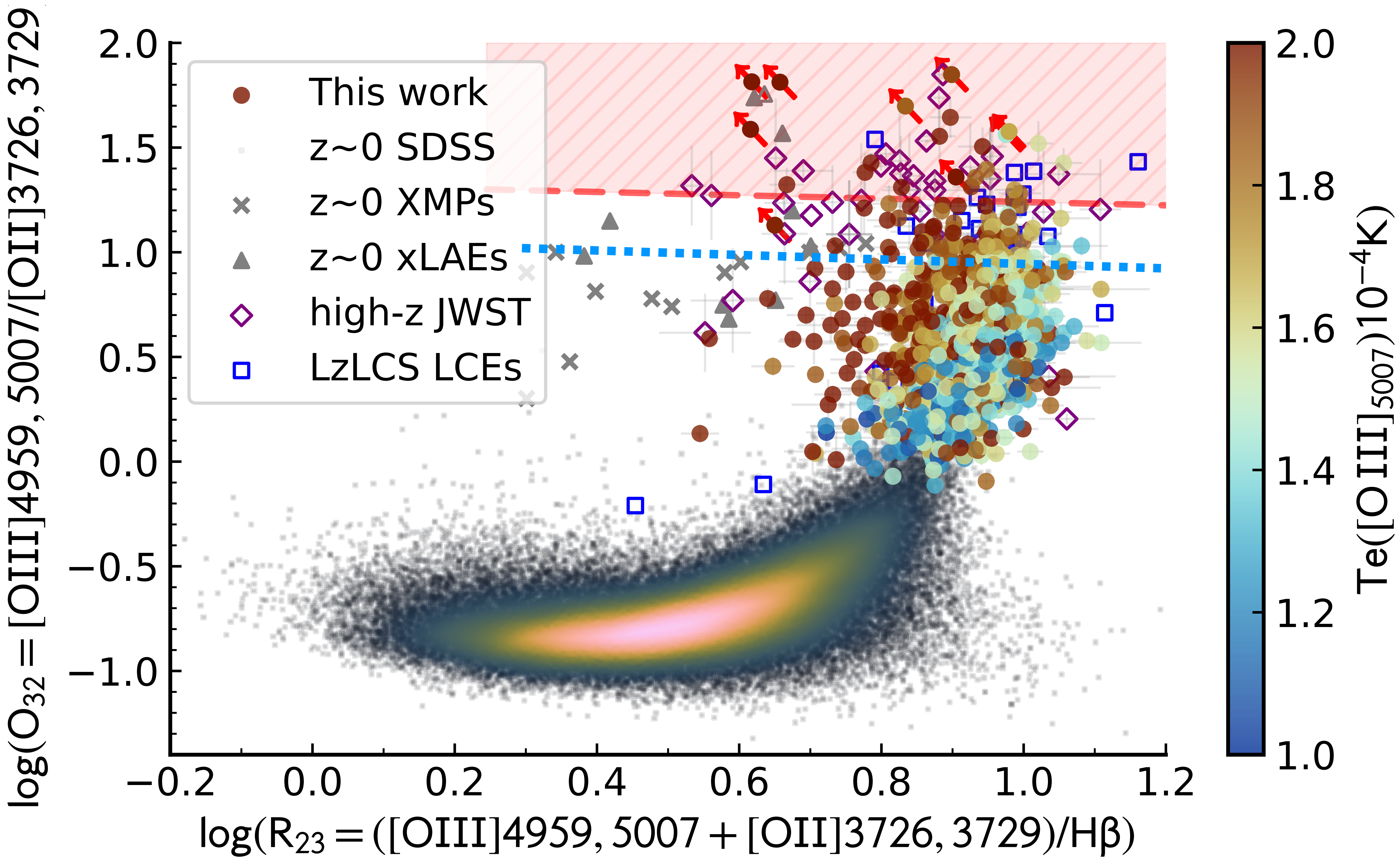


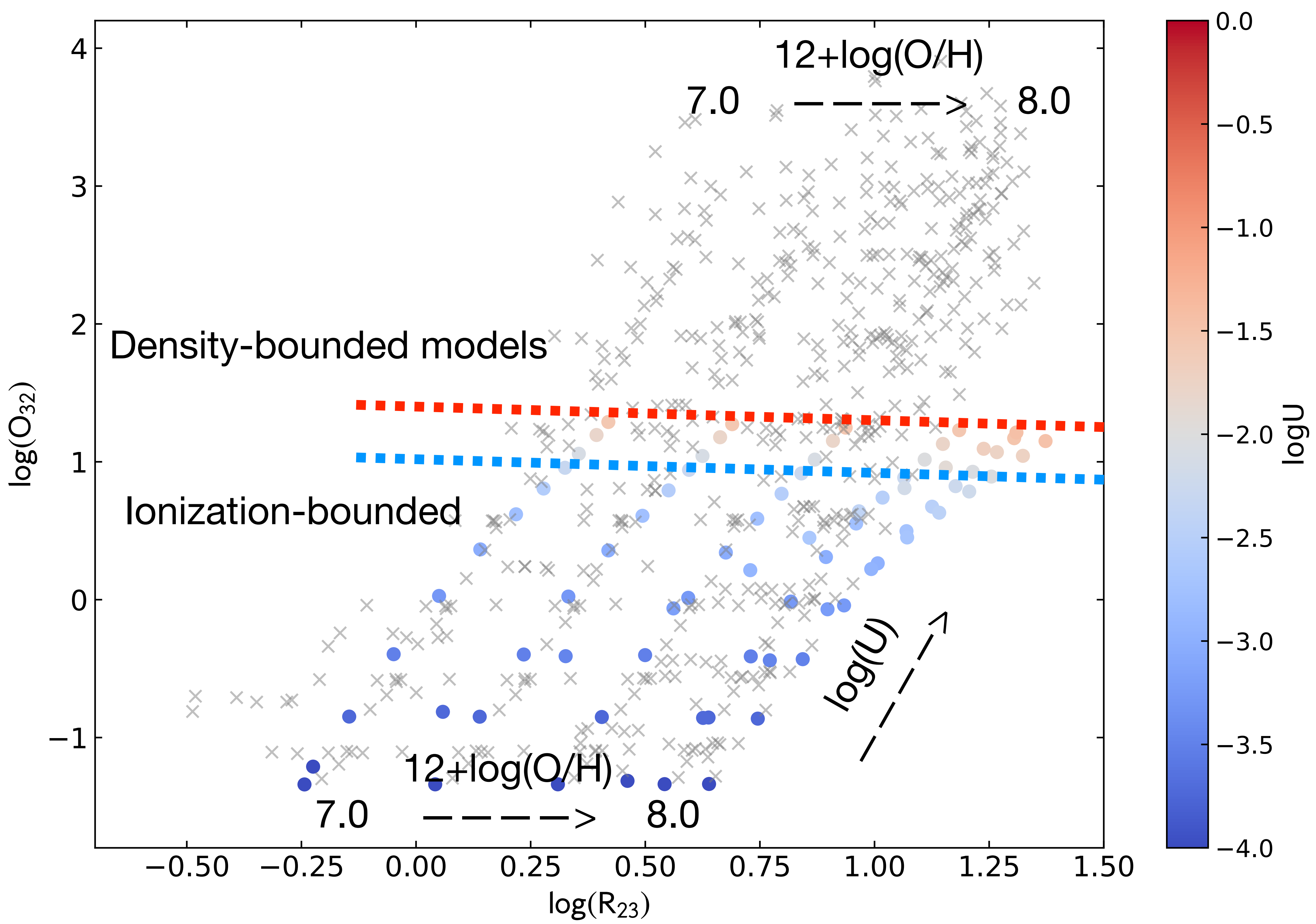
Amorín+25, in prep

XTREM ionization sample

Higher ionization at very low metallicity: density bounded conditions

- A clear trend in O32-R23 with Te is reproduced by Cloudy models with purely stellar BPASS models (Perez-Montero+2014;2021, 2023)
- Density-bounded modes are the only ones able to reproduce O32>10-20 with 12+log(O/H)<8 (Te>15000K) — Cf. Morriset+25
- No need of exotic sources with very hard radiation fields (models show BPASS incl. binaries with top-heavy IMF)



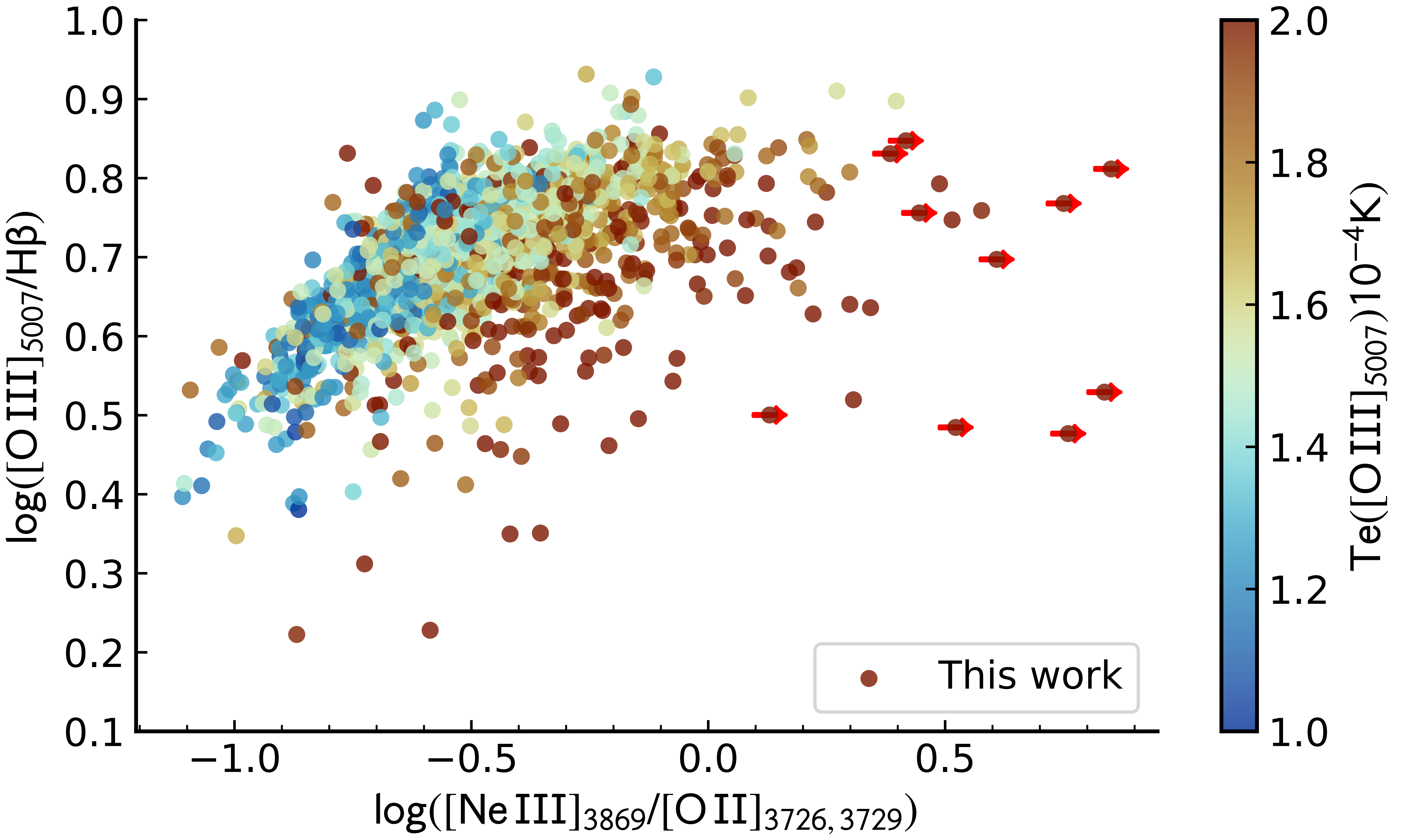


Amorín+25, in prep

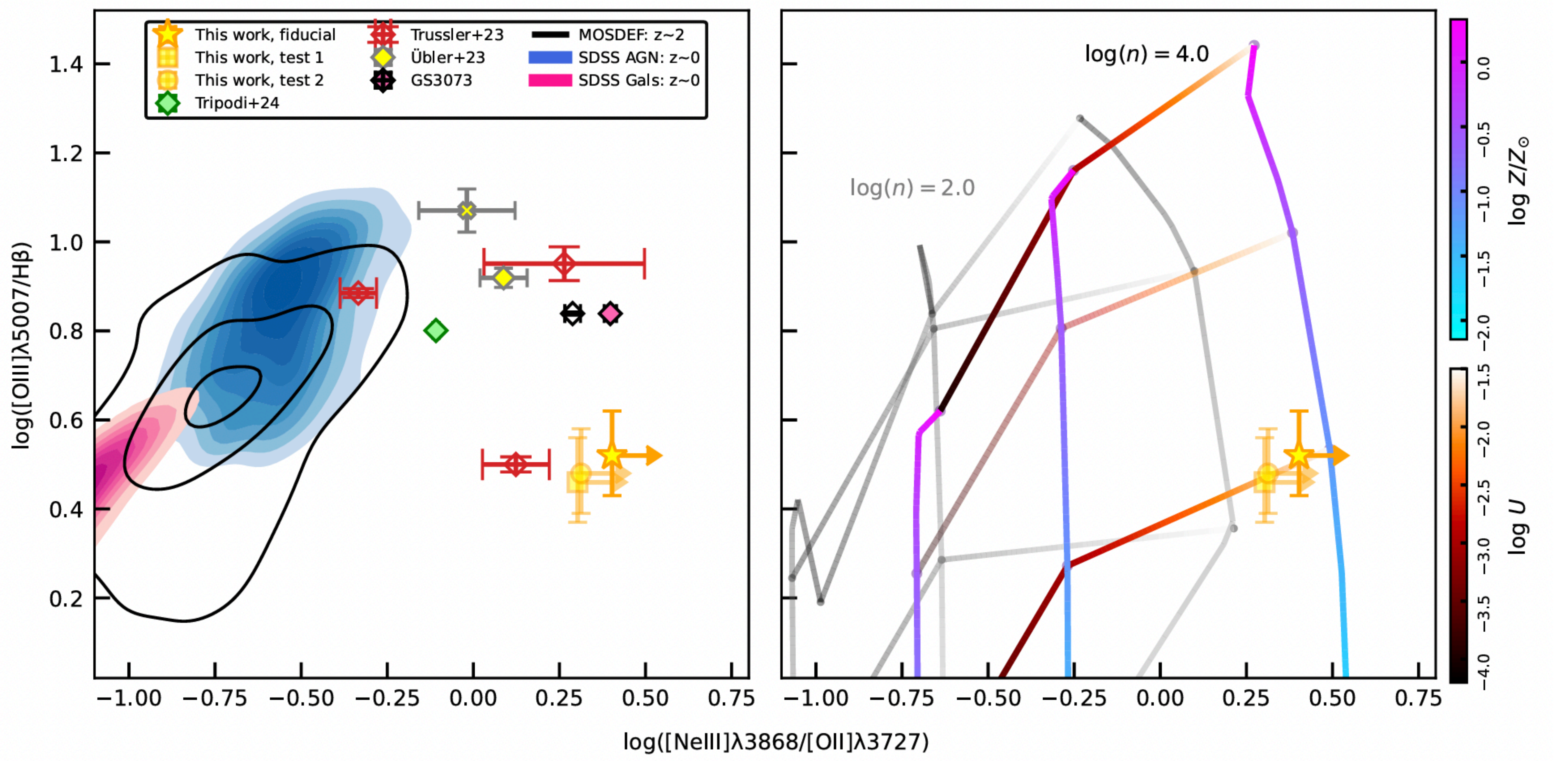
XTREM Te-based sample

Higher ionization at very low metallicity: candidates to strong LyC emitters

- Consistent results for the OHNO diagnostic



- Ionization-bounded models only explain the most extreme Ne3O2 and Te if assume AGN ionization at low metallicity (e.g. Tripodi+24)



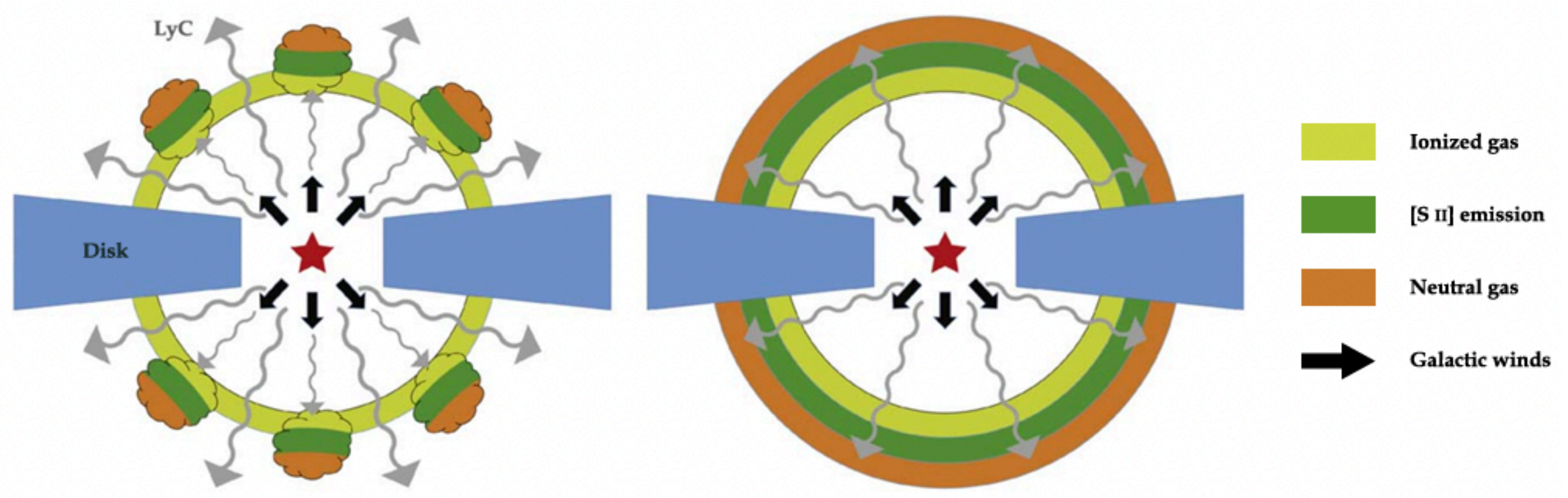
- However, density-bounded models with "normal" stellar ionization are claimed to work in some others (e.g., Cullen+25)

XTREM Te-based sample

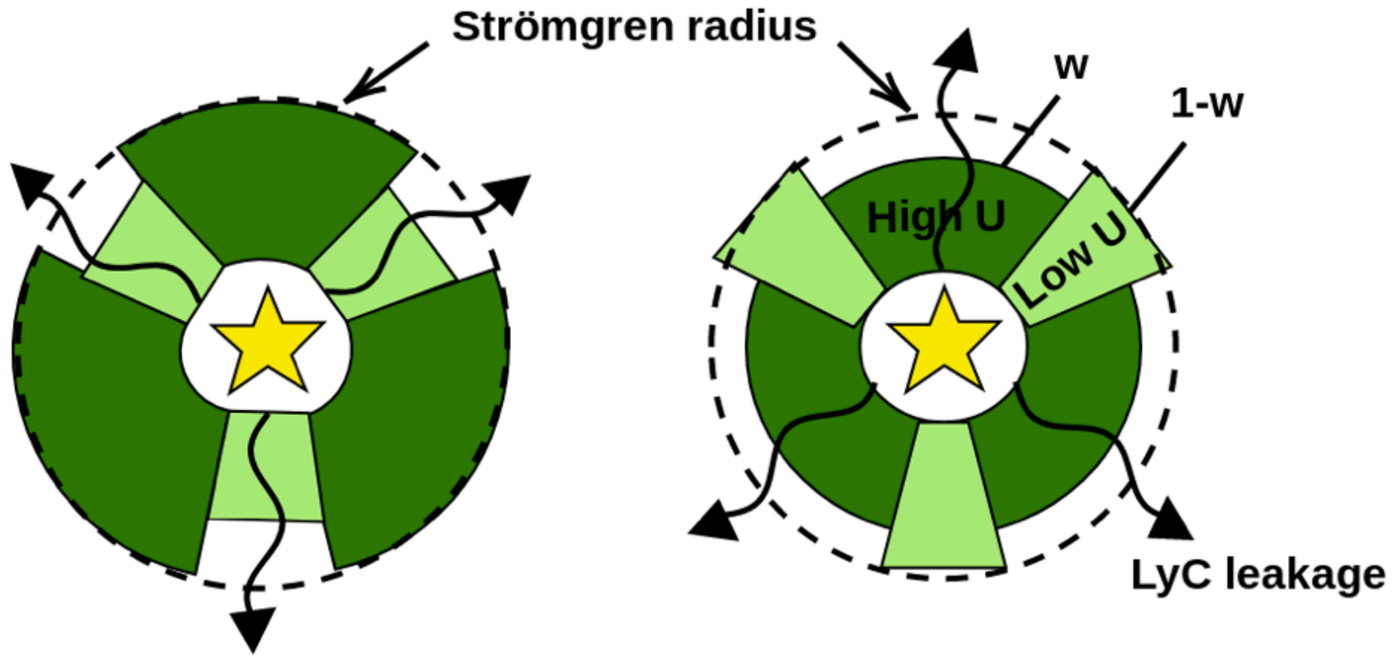
Higher ionization at very low metallicity: candidates to strong LyC emitters

- Selecting young bursts with high Te, low/moderate Ne and weak or non-detections of low-excitation lines ([NII], [SII], [OII]) appear optimal for candidates to strong LCE. Better to avoid restricting our samples by imposing high S/N cuts on these lines !

Cf. Wang+19,21 on [SII] deficiency



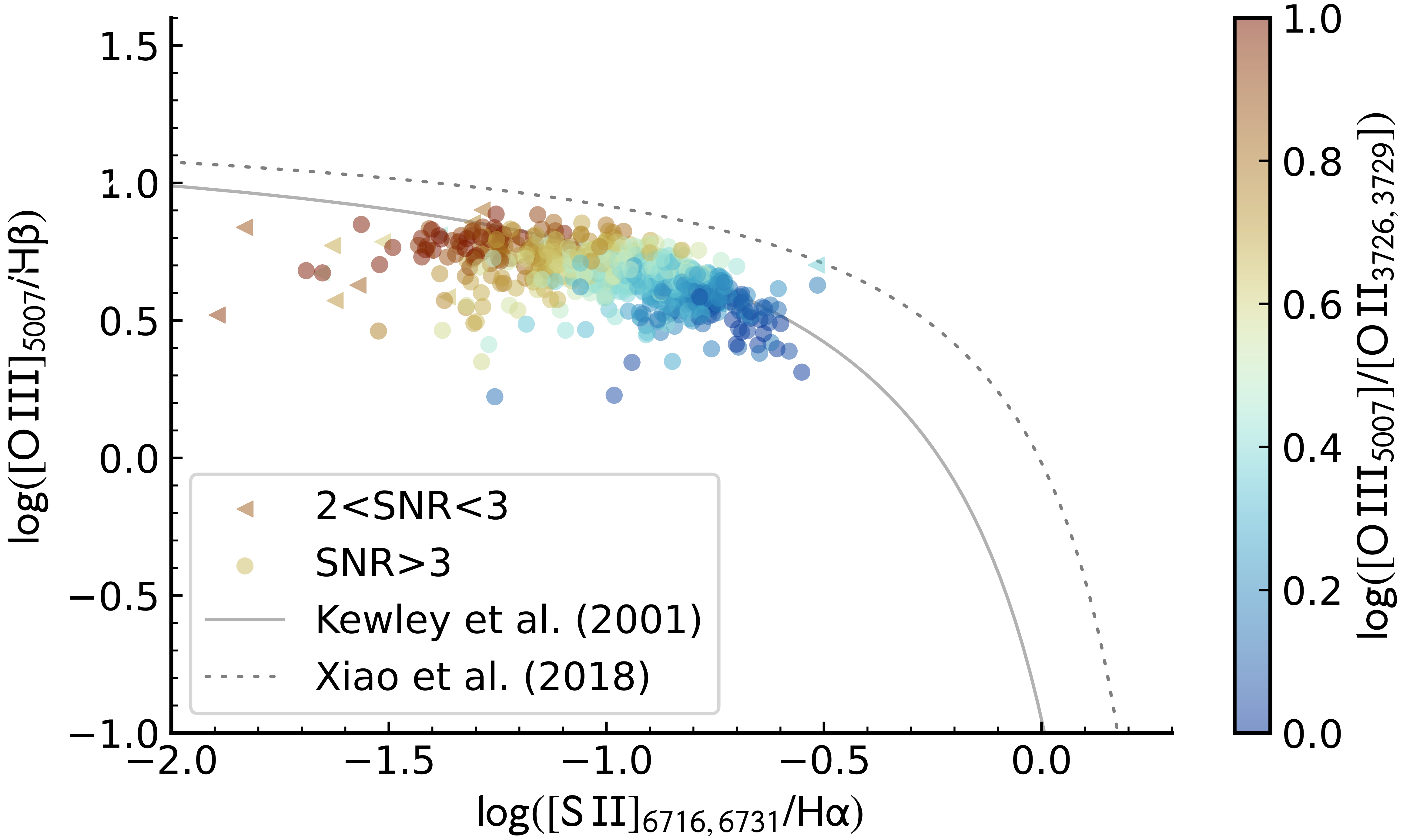
Ramambason et al. (2020)



Scenario (1)
Density-bounded channels/filaments

Scenario (2)
Density-bounded galaxy

Galaxies are far more complex than single HII regions
Picket fence models with holes, channels, filaments…



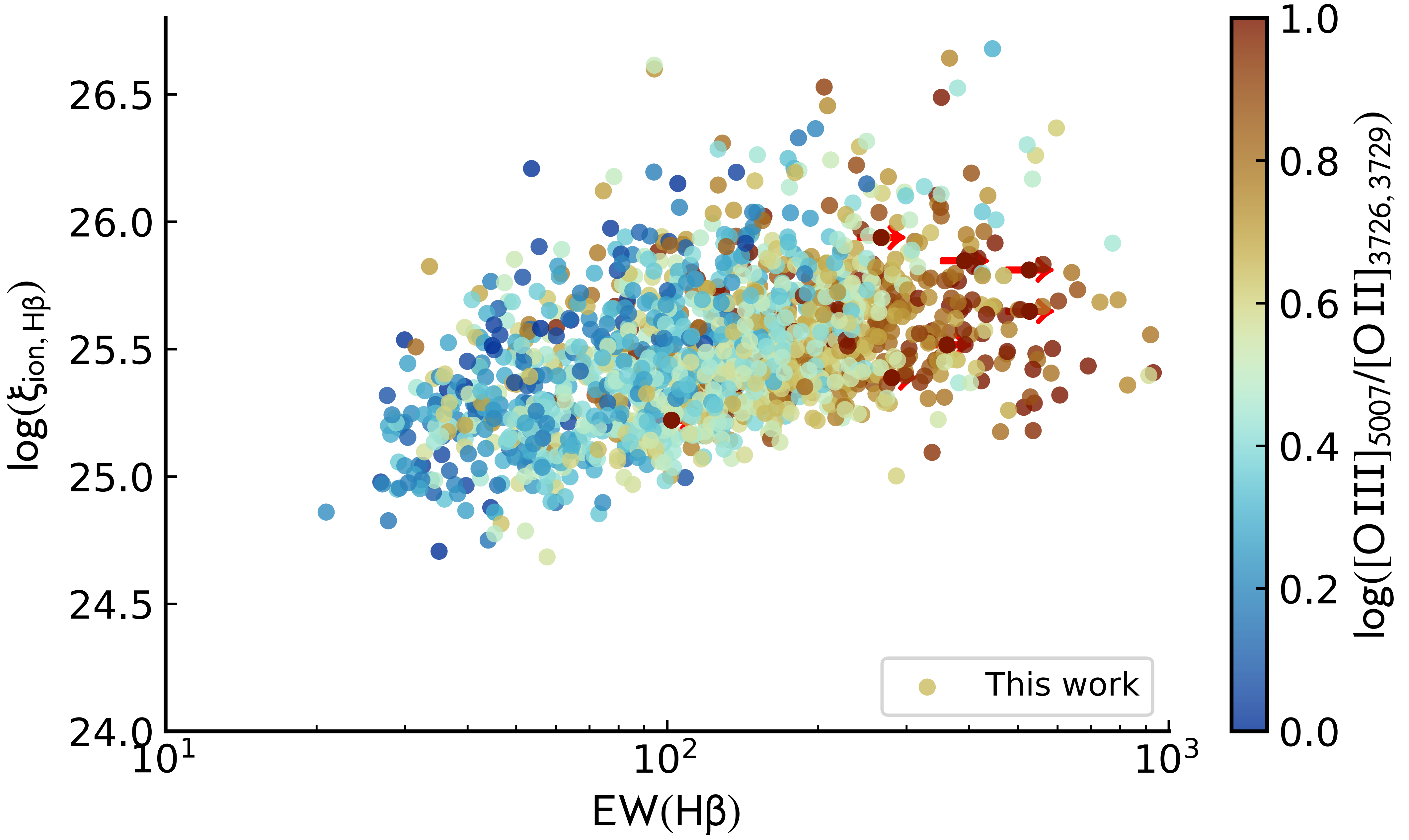
Amorín+25, in prep

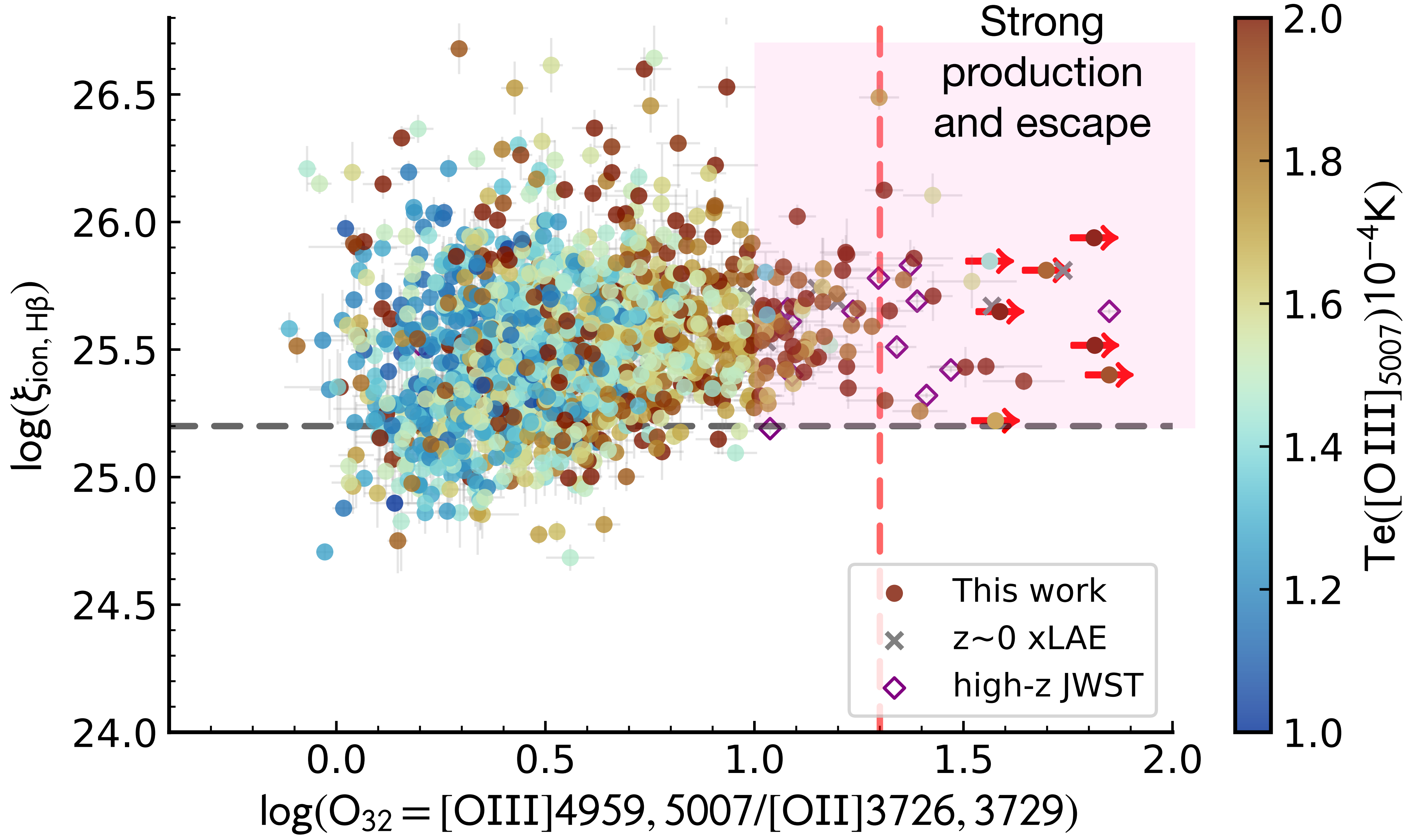
XTREM ionization sample

Probing the ionization photon efficiency

Our results

- We derive ξ_{ion} from $H\alpha$ and UV (1500) luminosity consistent with previous works (no LyC leakeage assumed) *See Antonio Gimenez Talk!*
- We find a clear increase of ξ_{ion} with EW. In agreement with previous works at low and high-z (e.g. Tang+19; Simmonds+24; Llerena+24)
- The highest $H\beta$ EWs (>100A) correspond to younger burst ages, higher ionization (O32, Ne3O2) and higher T_e — above EW($H\beta$)>300A the relation flattens (cf. Izotov+24)
- We find a scattered ξ_{ion} - O32 relation, with a **flat behaviour above 25.2 for EELGs with O32>10 AND Te>15-20kK (XMP regime)**

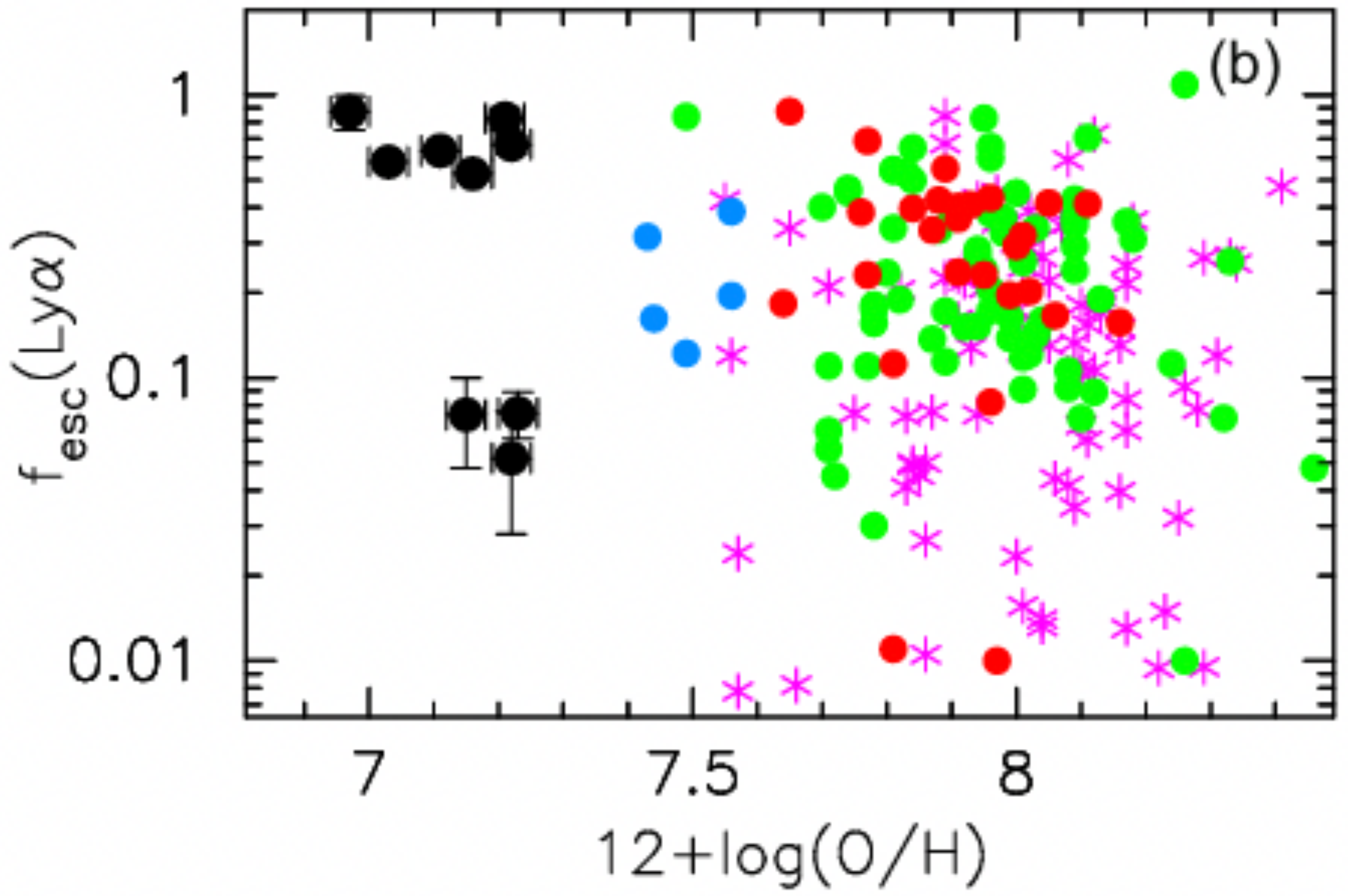


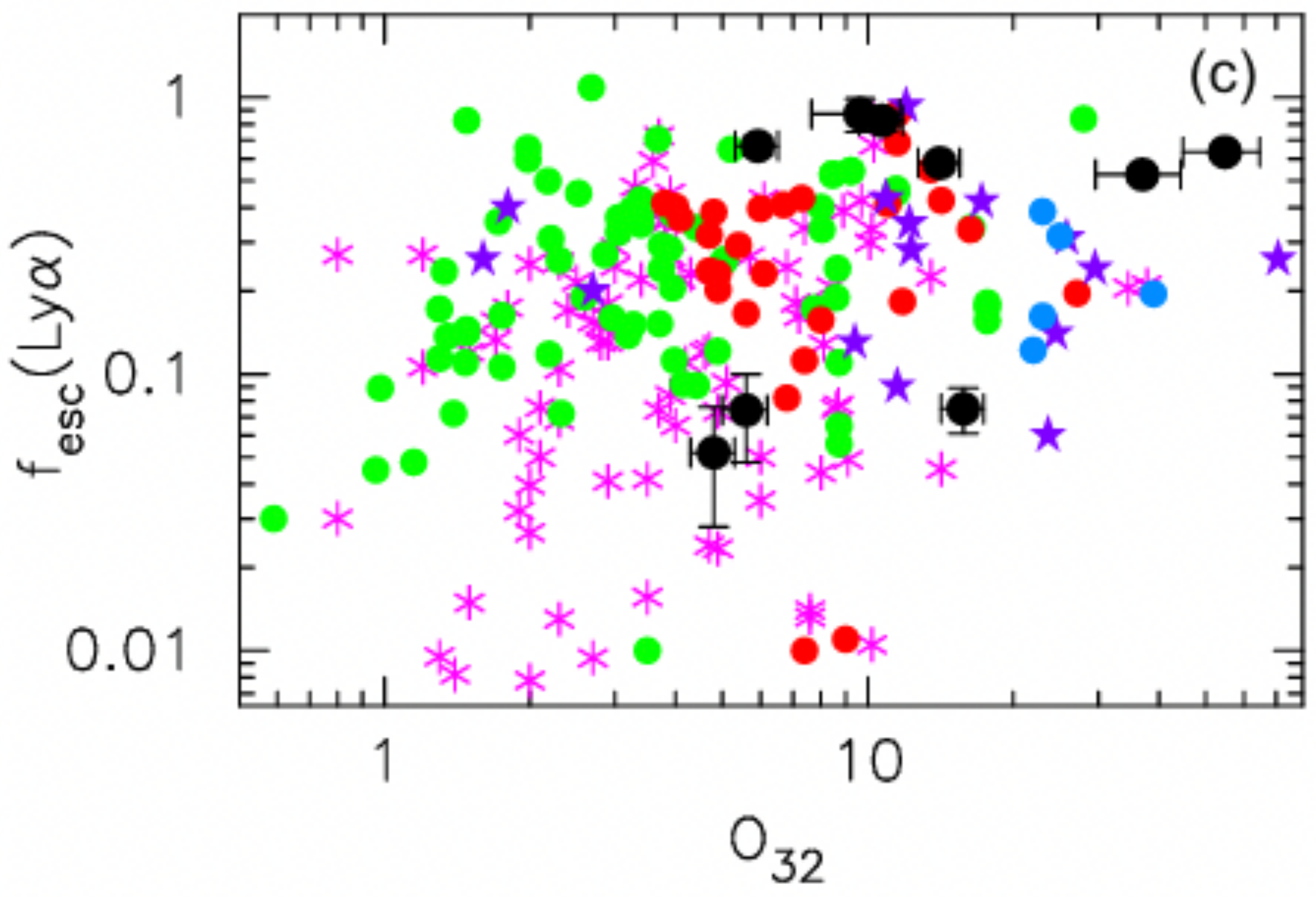


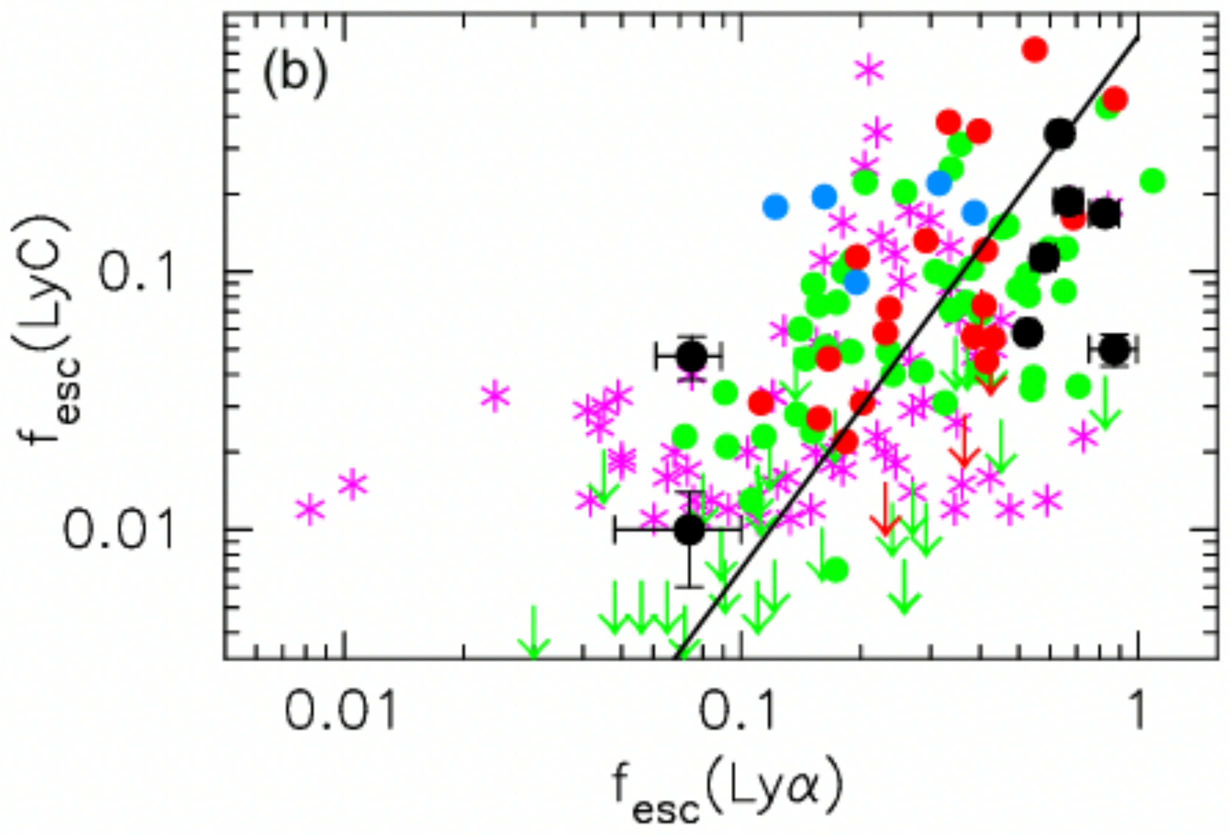
Amorín+25, in prep

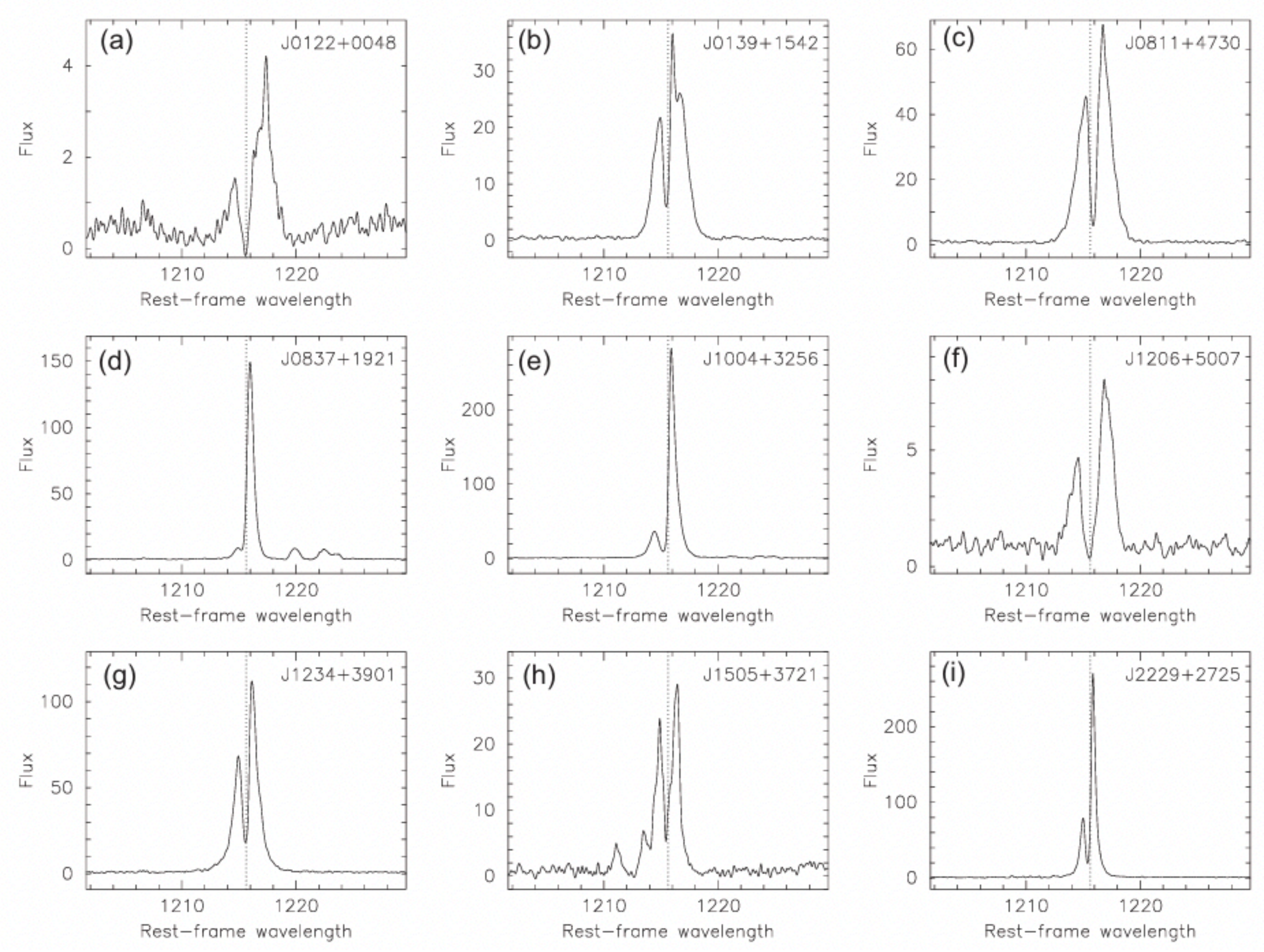
Expectations for HST/COS studies

- Our XTREM-20K sample of potential leakers add statistical significance to the local xLAEs found by Izotov+24
- We explore a similar range of faint luminosities, low-mass and metallicity, and ionization conditions but at higher redshifts
  - Our predictions are in line with the XTREM sample of high O32, high Te, are xLAEs with very strong LyC escape
  - The XTREM redshift range makes them better suit for LyC studies (e.g. LzLCS)









Izotov+24

Summary and Conclusions

- We present XTREM: an unprecedented sample of 20k EELGs selected from DESI EDR using a KMEANS methodology to filter outliers and minor spectral classes. We find EELGs (including XMPs, "Green Peas" and rare AGNs or "Little Red Dots" analogs) are outliers of any classification scheme and our method is robust to isolate these systems.

- XTREM sample is an ideal training set to understand efficient ways to exploit DESI DRs and other surveys e.g J-PAS for the selection and study of complete samples of EELGs (see Antonio's talk later)

- XTREM contains a golden subsample of ~1500 [OIII]4363 emitters, including the most metal-poor ones (XMPs, Z < 0.1 Z_{sun}), which is ideal for Lyman escape studies. New objects w.r.t previous works.

  - We discovered new EELGs with extreme ionization conditions (O32>10-70) and electron temperatures (Te>20kK). Selection effects due to S/N cuts may prevent to find these objects.

  - Cloudy models with standard stellar ionization (BPASS) show their O32-R23 and Te properties are only predicted by incomplete or optically-thin nebulae or highly porous ISM conditions consistent with "density-bounded" assumptions.

    - We establish a limit in the O32-R23 plane above which ionization-bounded models (fesc(LyC)=0) cannot explain their line ratios. Importantly, no need of exotic or very hard radiation fields is found (and *no sign* of very high density, or high ionization lines such as HeII or NeV is particularly found in this sample)

    - While strong leakage is still possible in not so extreme conditions, our XTREM subsample are great candidates for prodigious Lya and LyC emission, being one of the closest analogs to the most extreme LAEs at z>6 found by JWST

  - Future HST/COS follow-up will confirm their xLAE and LCE candidacy and test the low-mass LAE/LCE regime

Thank you for your attention!!

Grant PID2023-147386NB-I00 funded by:

MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

Cofinanciado por la Unión Europea

AGENCIA ESTATAL DE INVESTIGACIÓN

INSTITUTO DE ASTROFÍSICA DE ANDALUCÍA

EXCELENCIA SEVERO OCHOA

CSIC

Escape of Lyman radiation from galactic labyrinths
7-11 April 2025, OAC, Kolymbari, Crete

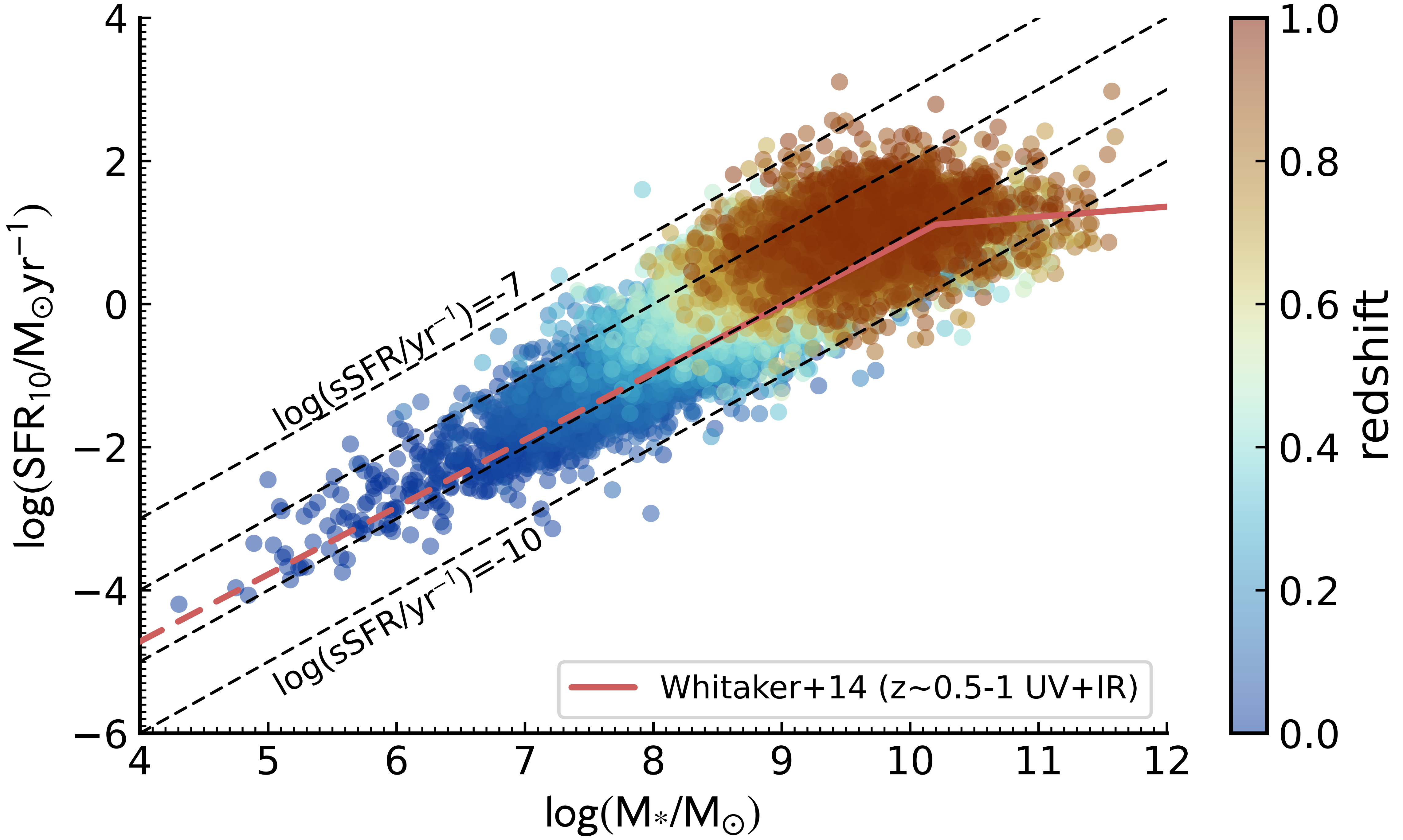
EELGs in DESI EDR

Scaling relations

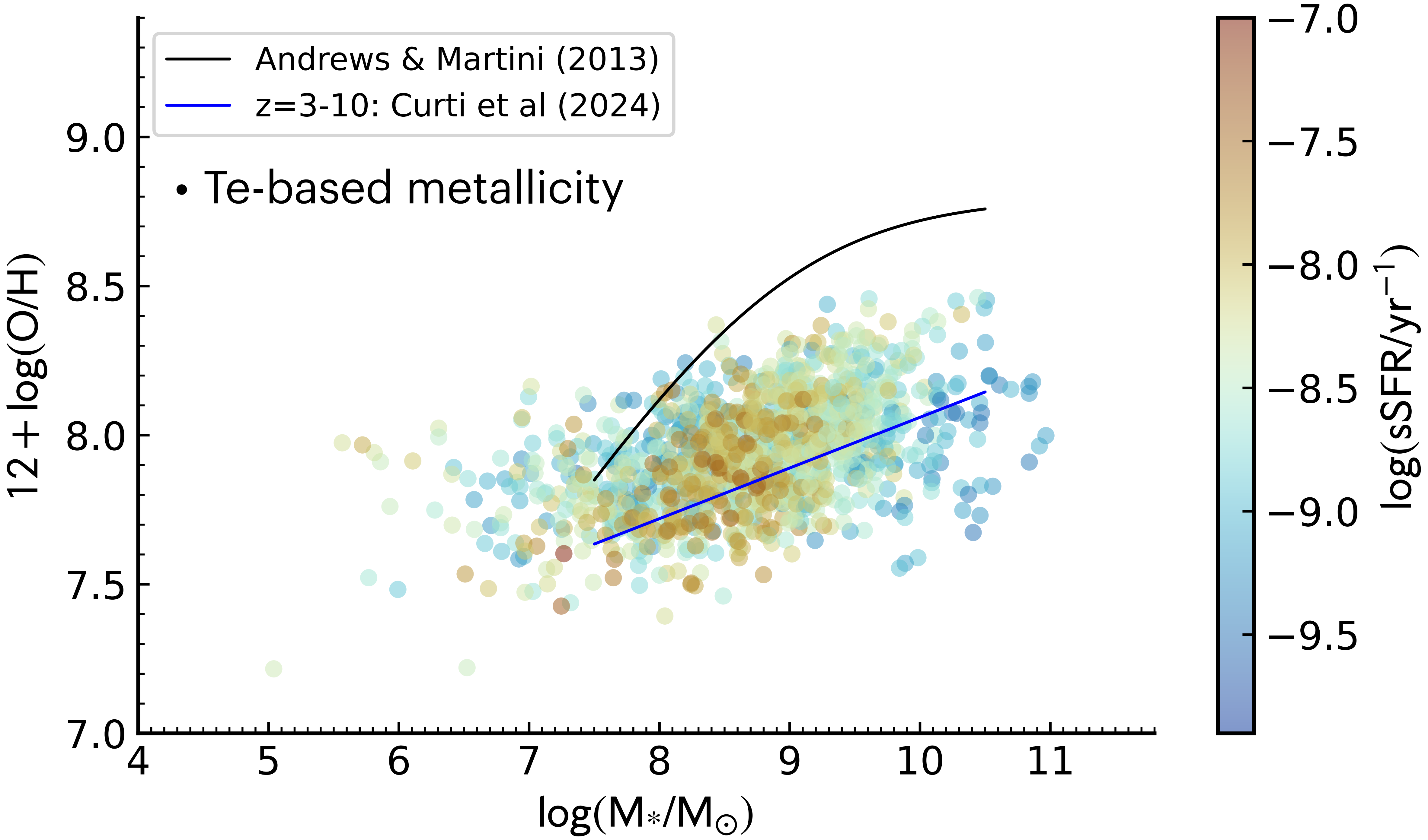
Preliminary results

- A very large sample allow us to probe mass-metallicity and other key scaling relations in excellent detail

Mass-SFR relation



Mass-metallicity relation



Amorín+25, in prep