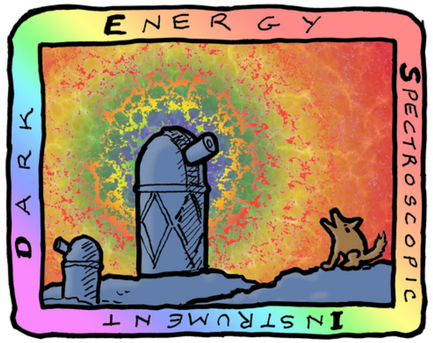


CENTER FOR
ASTROPHYSICS
HARVARD & SMITHSONIAN

Bayesian Component Separation for DESI LAE Automated Spectroscopic Redshifts and Photometric Targeting

arXiv:2504.06870

Ana Sofía M. Uzsoy

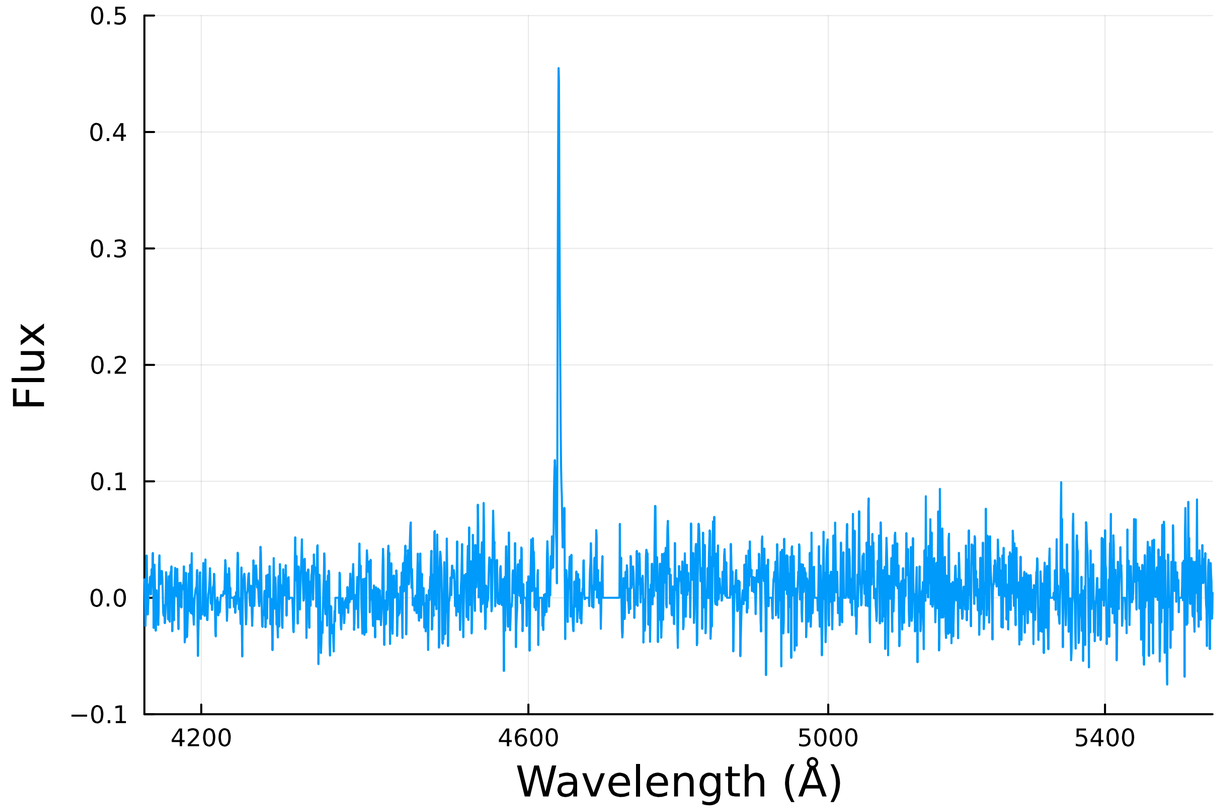
with Andrew Saydjari, Arjun Dey, Anand Raichoor, Douglas Finkbeiner, and others

Escape of Lyman Radiation from Galactic Labyrinths, OAC, Crete

April 11, 2025

Motivation

- DESI (Dark Energy Spectroscopic Instrument) collects millions of galaxy spectra
- DESI-2: focus on high-redshift targets (aims to observe ~1.5 million LAEs)
- We want to optimize LAE selection and analysis for future surveys like DESI-2
- DESI pipeline currently does not automatically determine redshifts for LAEs

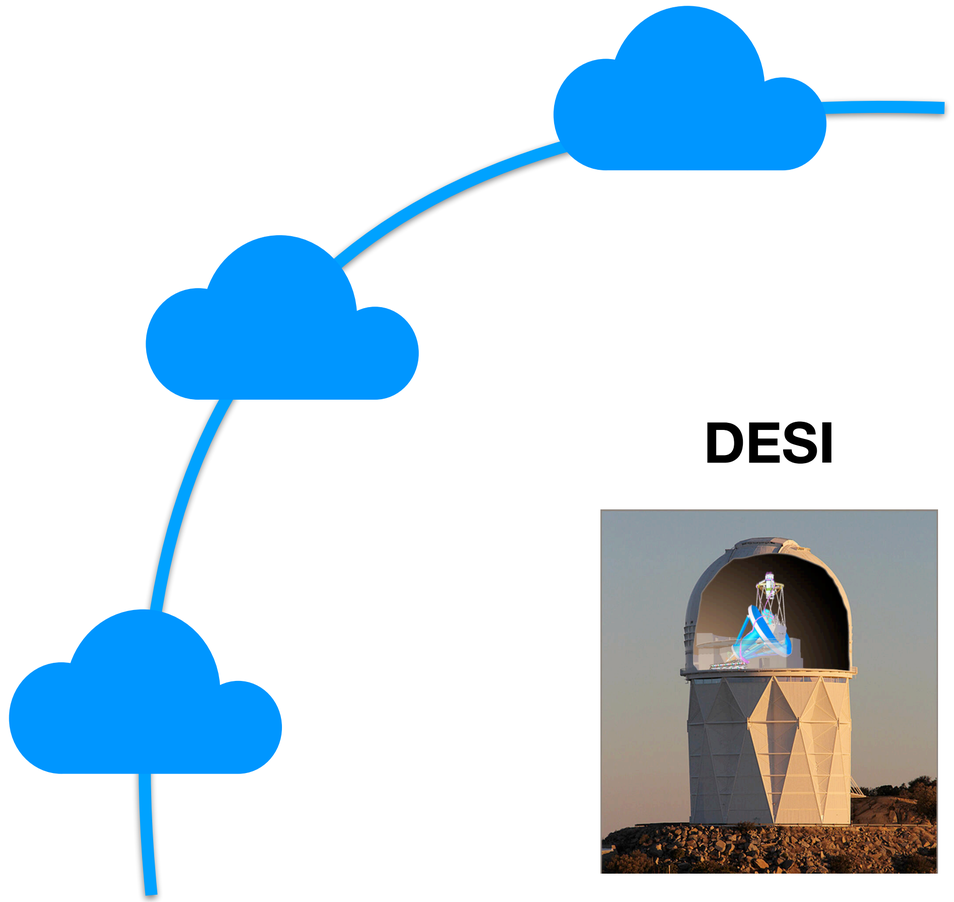


Schlegel et al. (2022)

DESI

Ana Sofia Uzsoy

Sky/atmosphere



DESI

Ana Sofia Uzsoy

Interstellar & Intergalactic medium

Sky/atmosphere

DESI

Ana Sofia Uzsoy

Target LAE galaxy at redshift z

Interstellar & Intergalactic medium

Sky/atmosphere

DESI

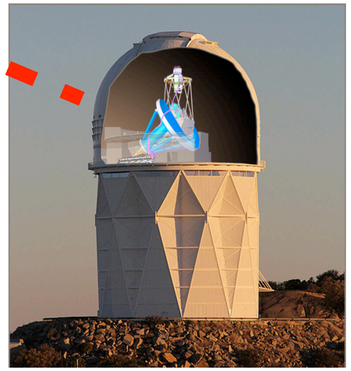
Ana Sofia Uzsoy

Target LAE galaxy at redshift z

Interstellar & Intergalactic medium

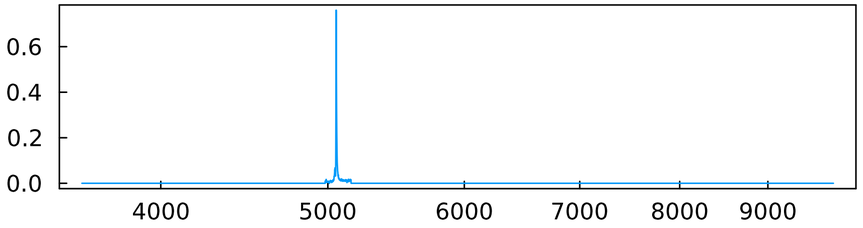
Sky/atmosphere

DESI

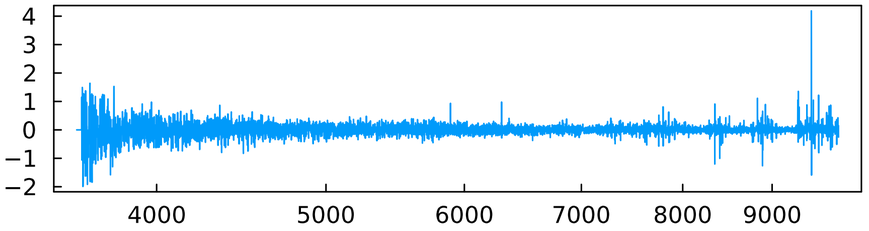


Ana Sofia Uzsoy

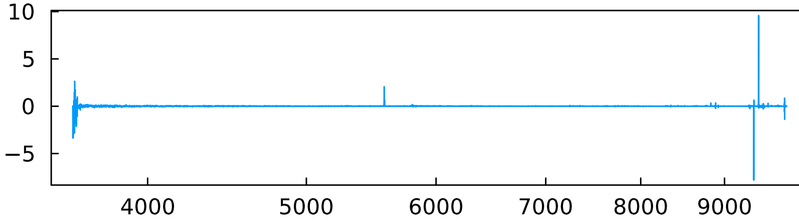
Target LAE galaxy at redshift z



Interstellar & Intergalactic medium



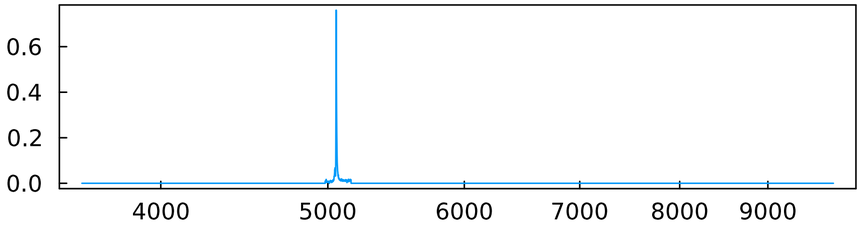
Sky/atmosphere



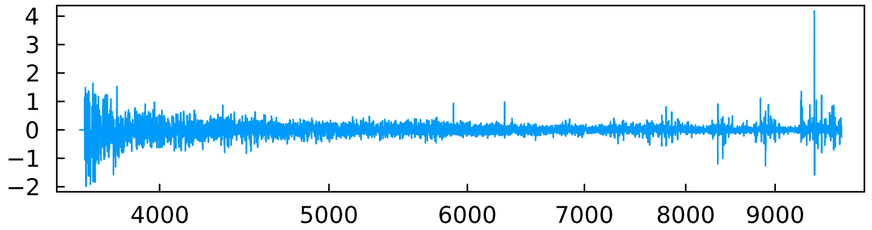
DESI

Ana Sofia Uzsoy

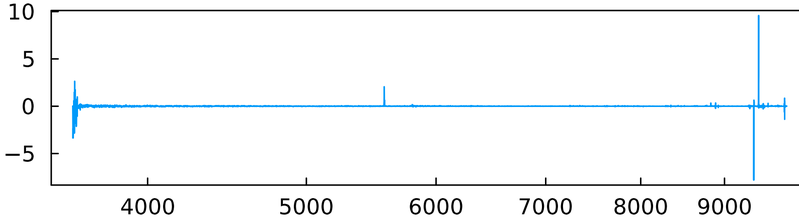
Target LAE galaxy at redshift z



Interstellar & Intergalactic medium



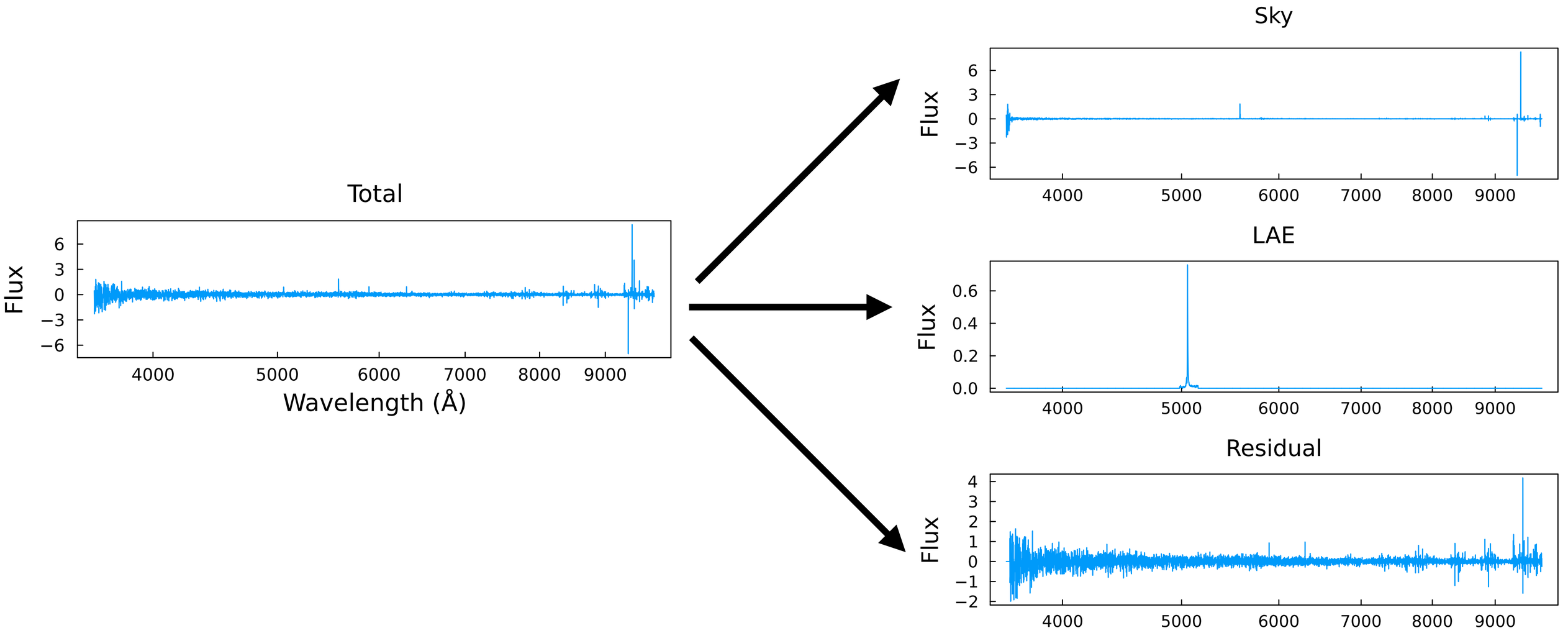
Sky/atmosphere



Total

DESI

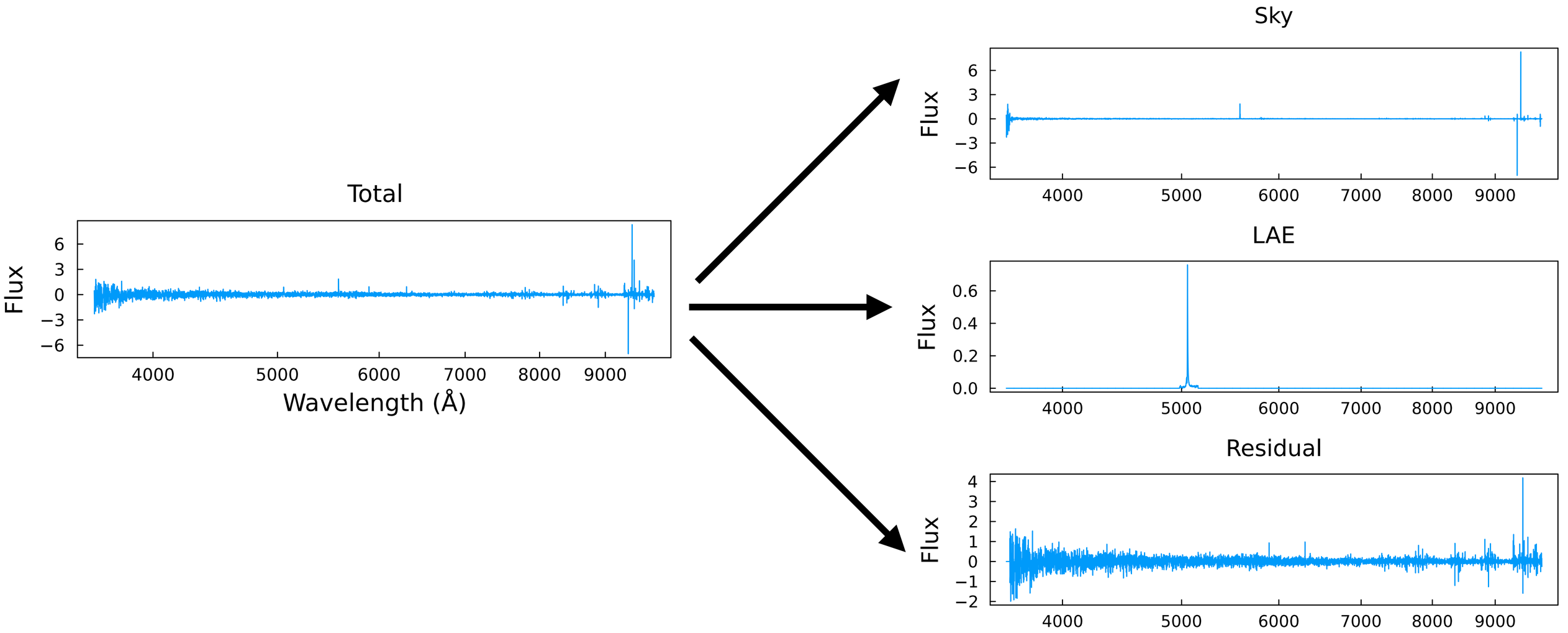
How can we separate a spectrum into its components?



Ana Sofia Uzsoy

How can we separate a spectrum into its components?

How can we automatically determine the redshift of LAEs?



Ana Sofia Uzsoy

MADGICS:

Marginalized **A**nalytic **D**ataspace **G**aussian **I**nference for **C**omponent **S**eparation

Bayesian component separation technique

Ana Sofia Uzsoy

MADGICS Formalism

Known # of components k

Prior for each component i as covariance matrix C_i

$$\sum_{i}^{k} C_i = C_{tot}$$

$$D_i = C_i C_{tot}^{-1} D_{tot}$$

Saydjari et al. (2023)

We use data-driven covariance matrices to put priors on our components.

- MADGICS assumes that you know what you're looking for
- We use known component data to create priors as covariance matrices
- Data-driven method allows us to effectively express systematics
- Components sum exactly to the data

D = matrix of n data vectors

$$C = \frac{DD^T}{n}$$

$$\sum_i^k C_i = C_{tot}$$

$$D_i = C_i C_{tot}^{-1} D_{tot}$$

Saydjari et al. (2023)

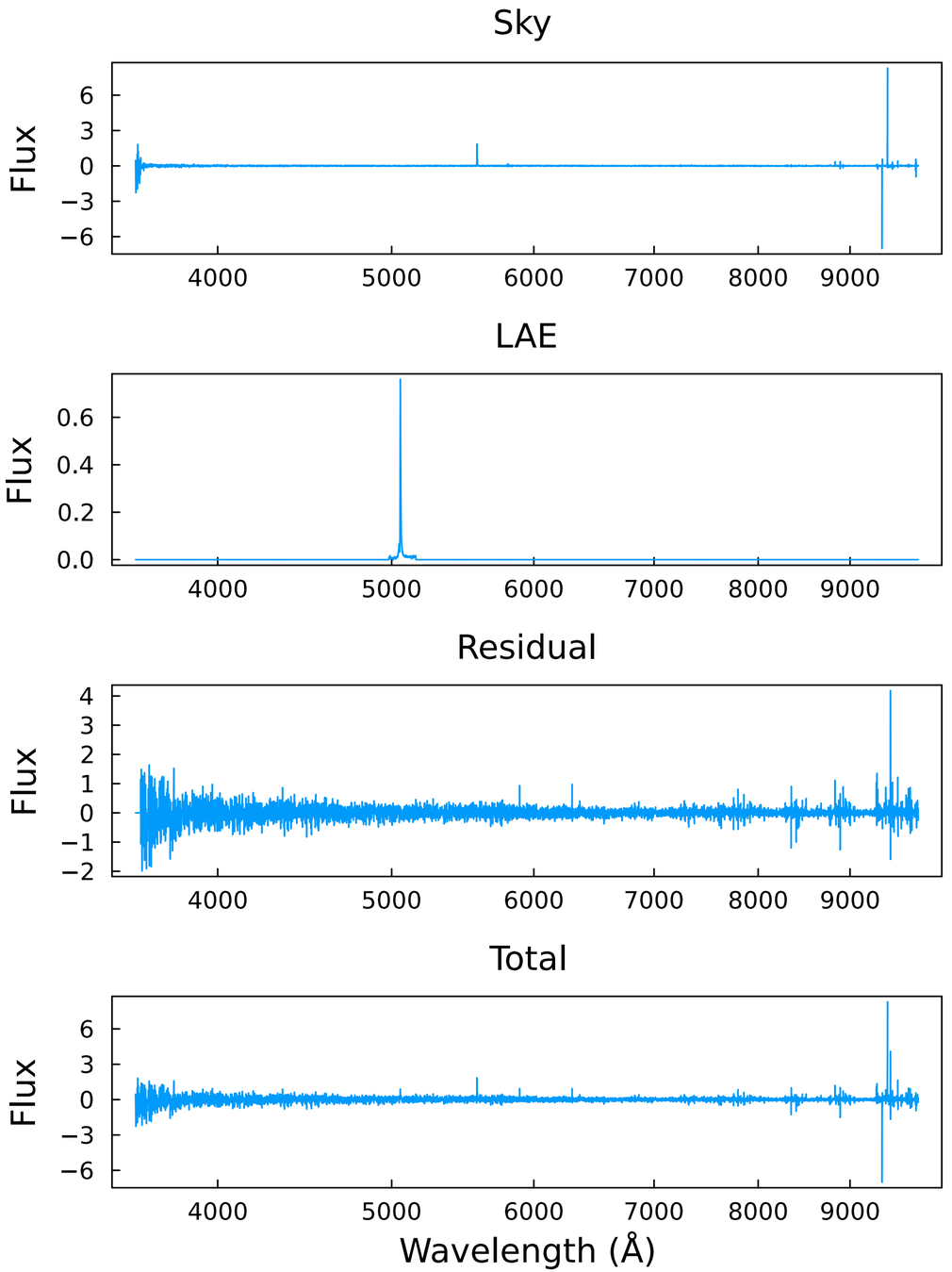
We can model LAE spectra as a sum of components.

spectrum = sky + LAE + residual

Prior covariance matrices:

$$C_{sky}, C_{LAE}, C_{res}$$

Wavelength-wavelength covariances!



Ana Sofia Uzsoy

We can model LAE spectra as a sum of components.

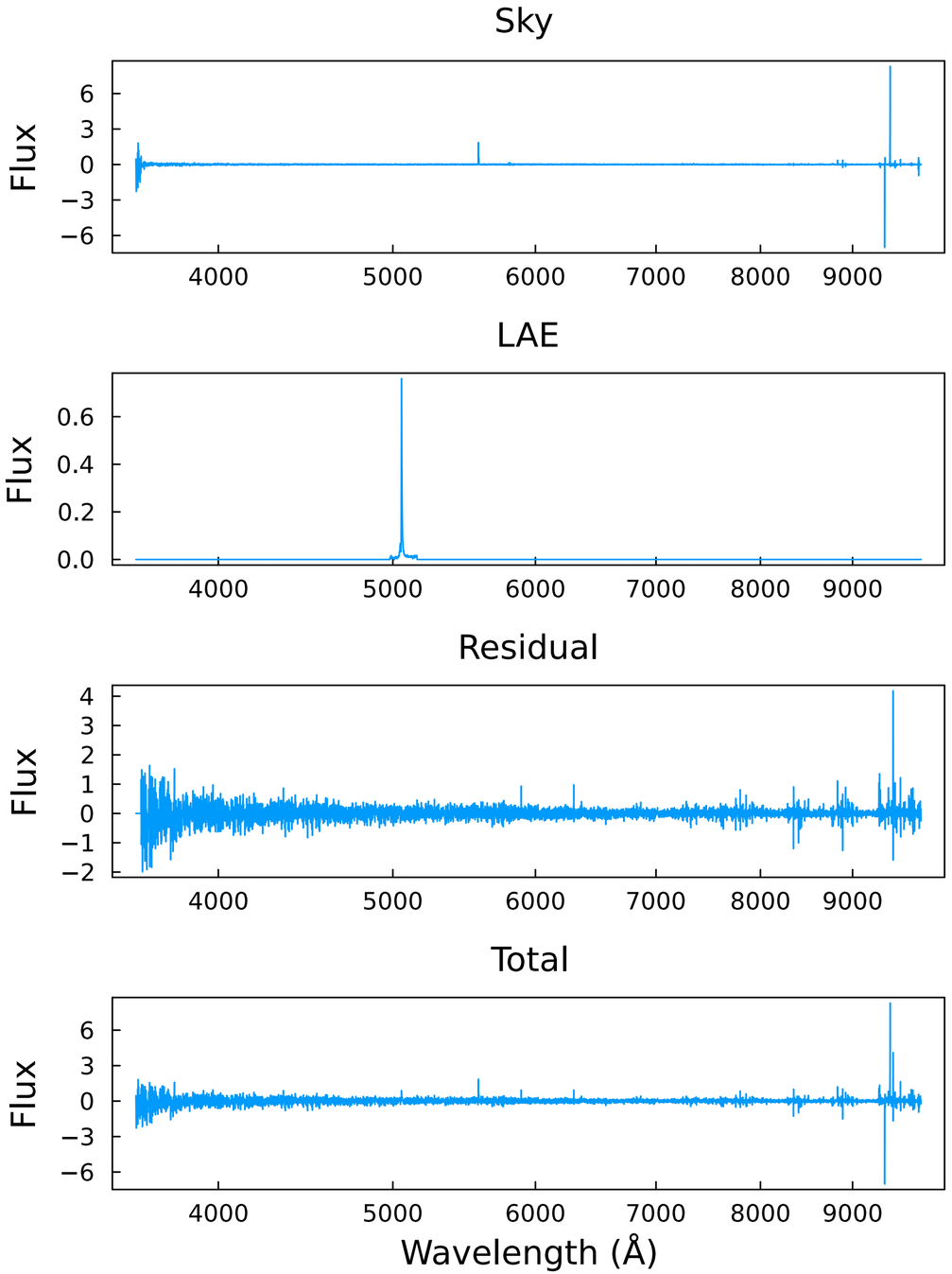
$$C_{tot} = C_{sky} + C_{LAE} + C_{res}$$

$$d_{sky} = C_{sky} C_{tot}^{-1} d_{tot}$$

$$d_{LAE} = C_{LAE} C_{tot}^{-1} d_{tot}$$

$$d_{res} = C_{res} C_{tot}^{-1} d_{tot}$$

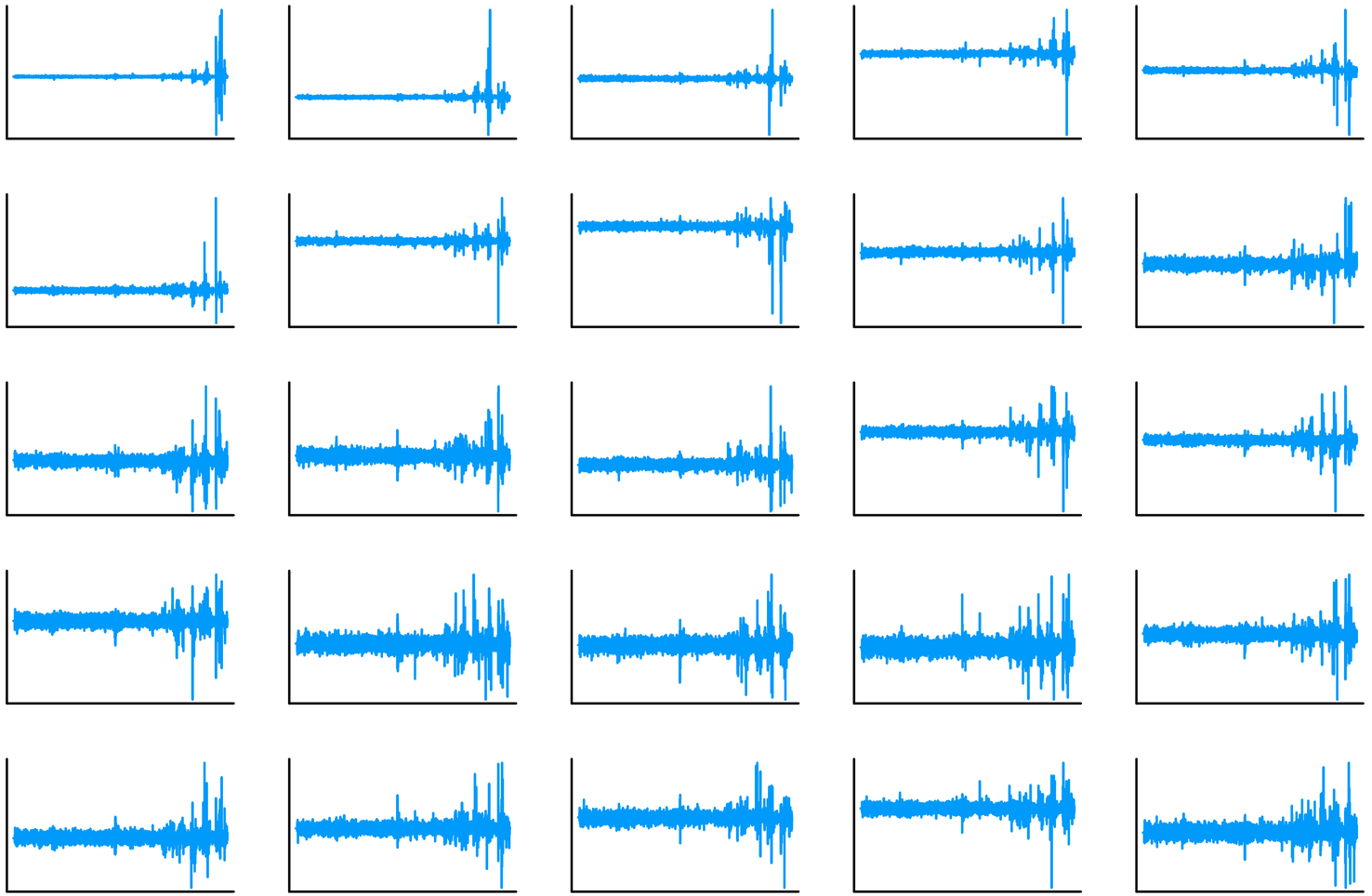
d_{tot} : total spectrum



Ana Sofia Uzsoy

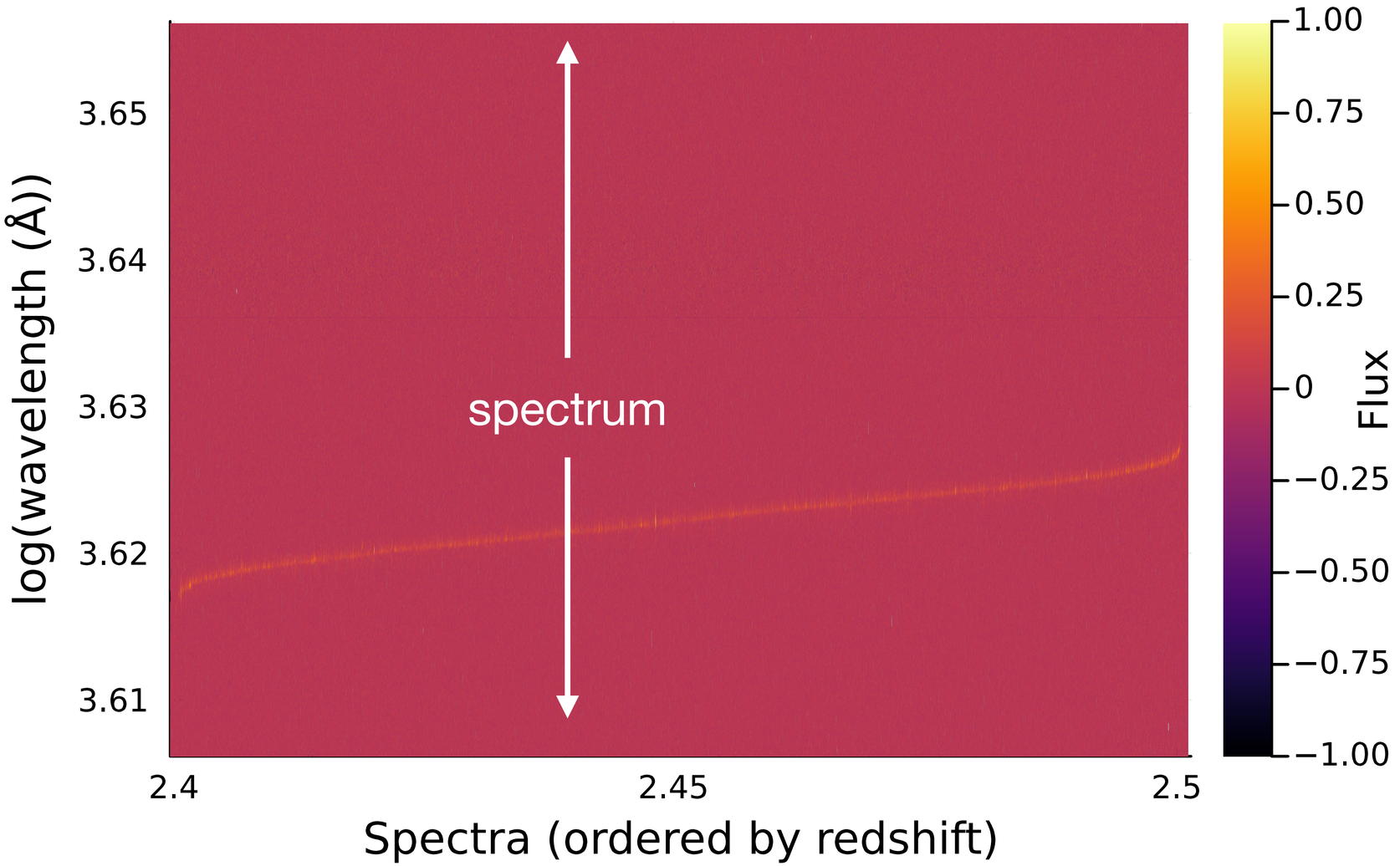
We create a prior covariance matrix of sky spectra.

C_{sky} : low-rank matrix approximation from top 50 sky eigenvectors



Our LAE prior is made of visually inspected ODIN LAE spectra.

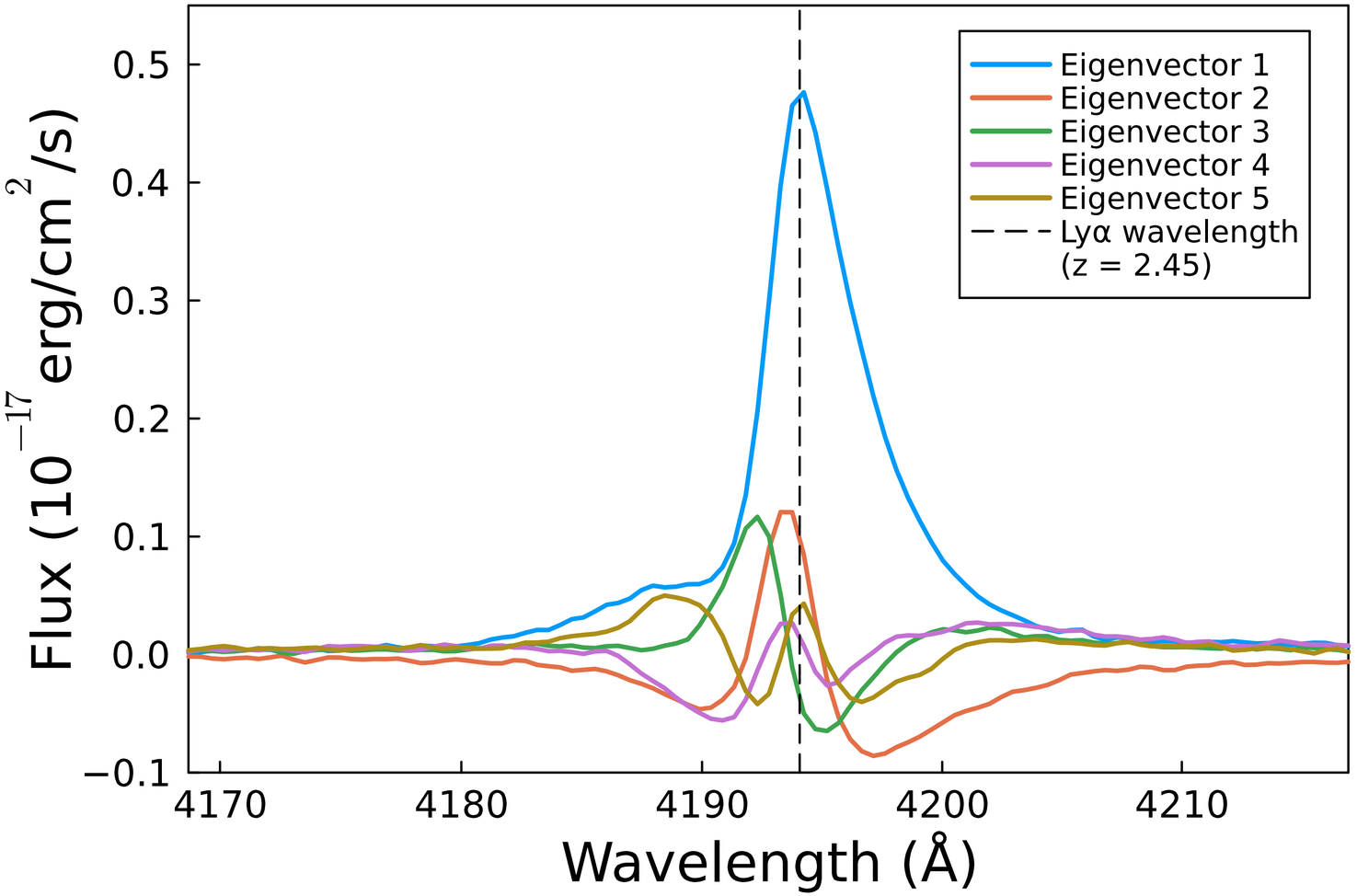
C_{LAE} : data-driven covariance matrix from ODIN LAE spectra shifted to z=2.45



Lee et al. (2023), Firestone et al. (2024)

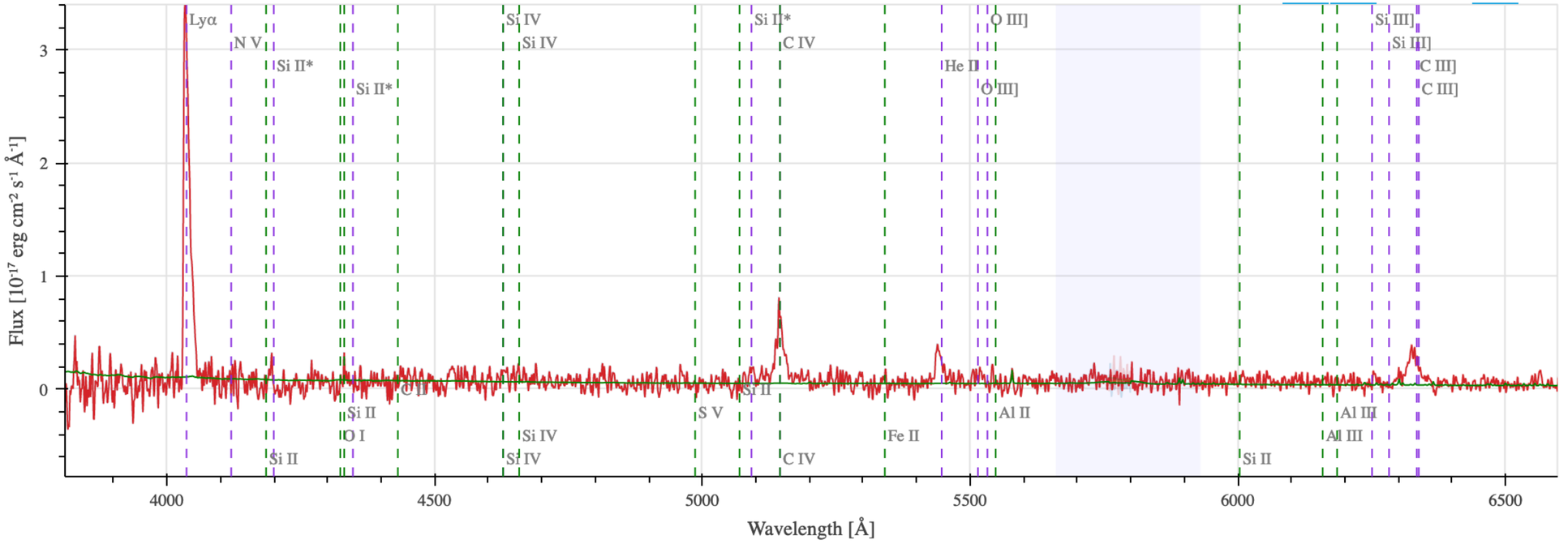
Our LAE prior is made of visually inspected ODIN LAE spectra.

C_{LAE} : data-driven covariance matrix from ODIN LAE spectra shifted to z=2.45



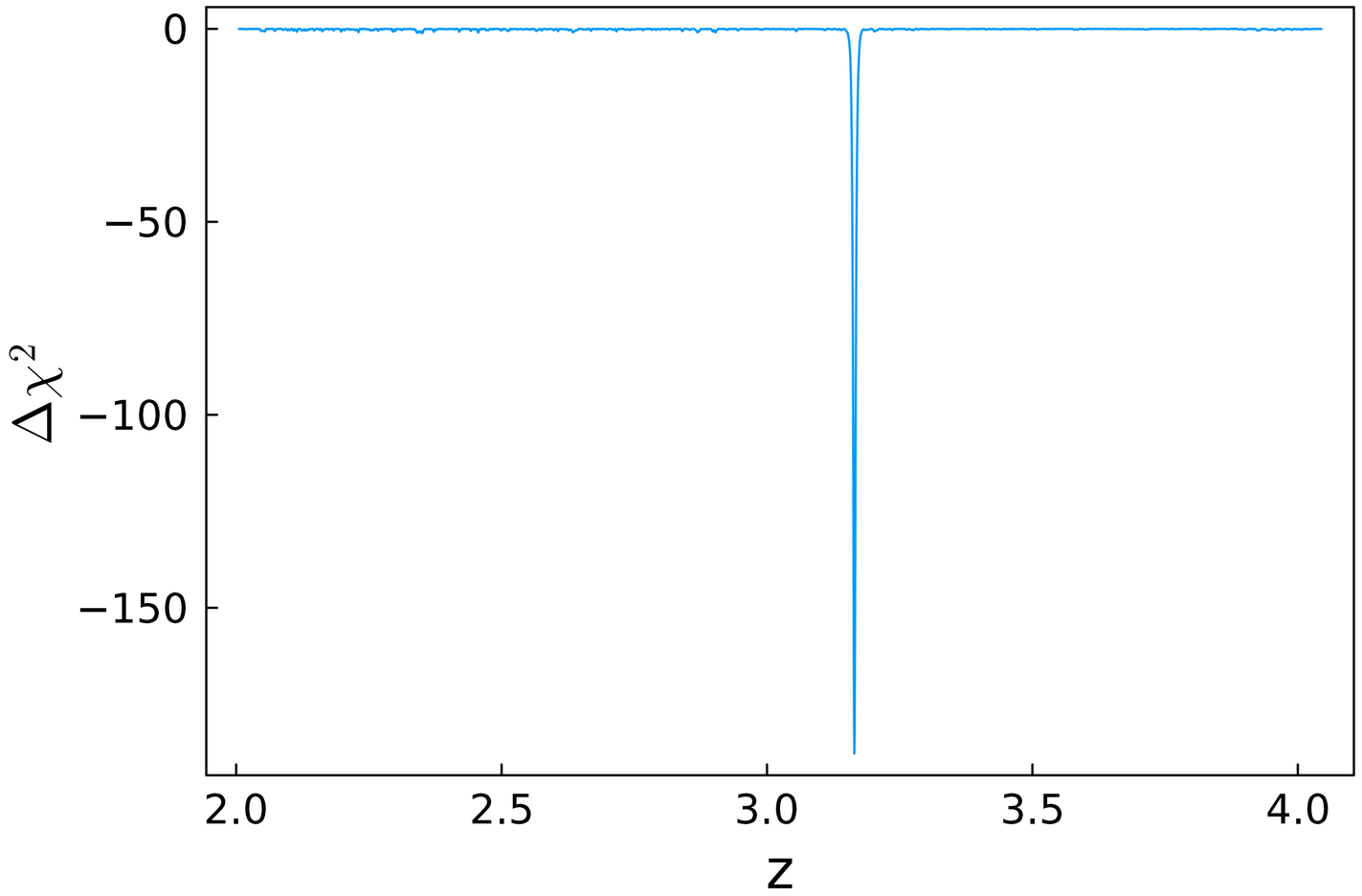
We want to automatically determine spectroscopic redshifts.

- DESI spectra for 910 LAE targets with medium-band photometry at $2 < z < 4$
- Visually inspected redshifts for comparison

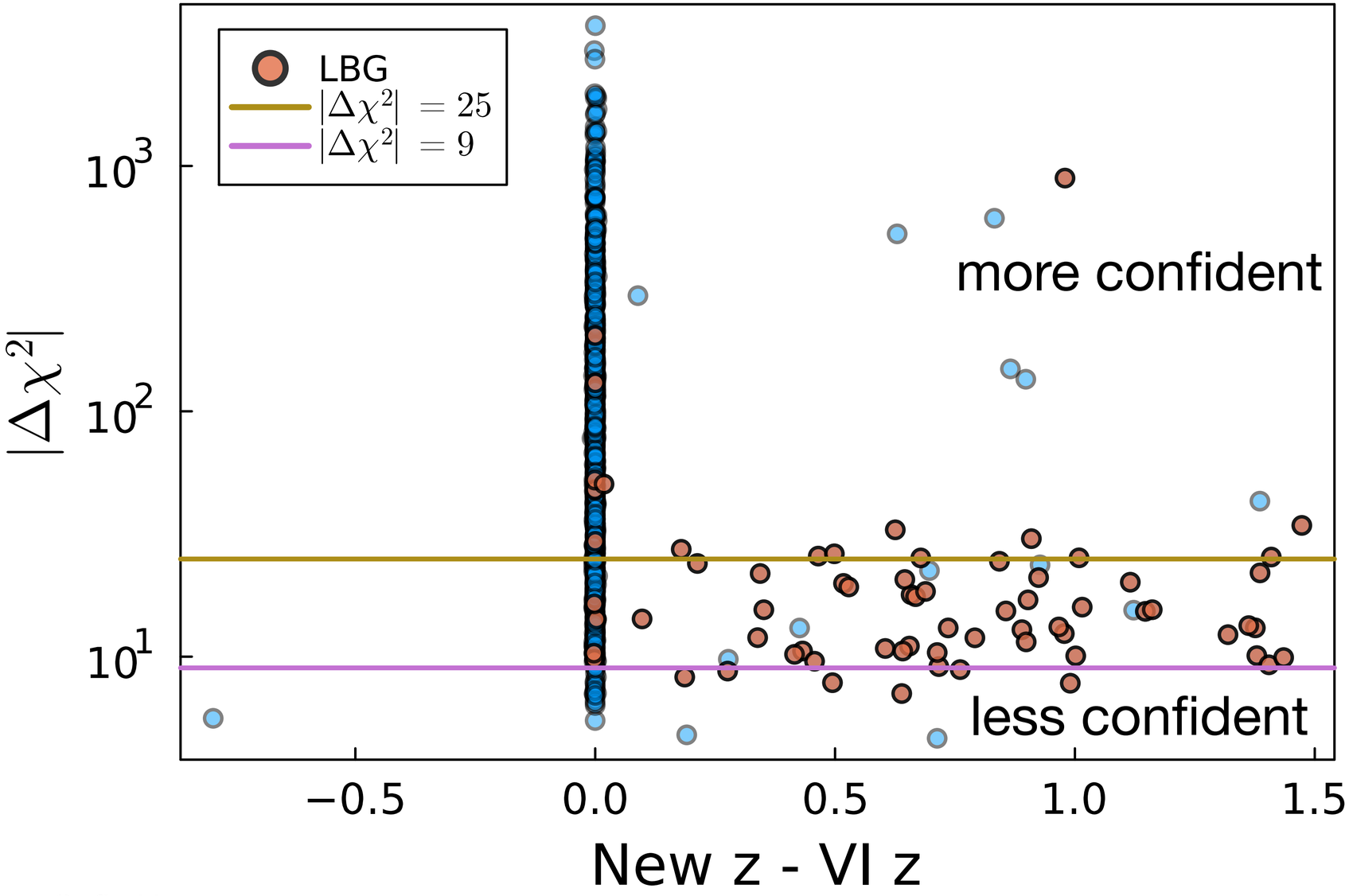


We determine redshift by minimizing $\Delta\chi^2$.

$$\Delta\chi^2(z) = \chi^2(\text{sky + residual}) - \chi^2(\text{sky + residual + LAE}(z))$$



Our method is able to accurately determine redshift.

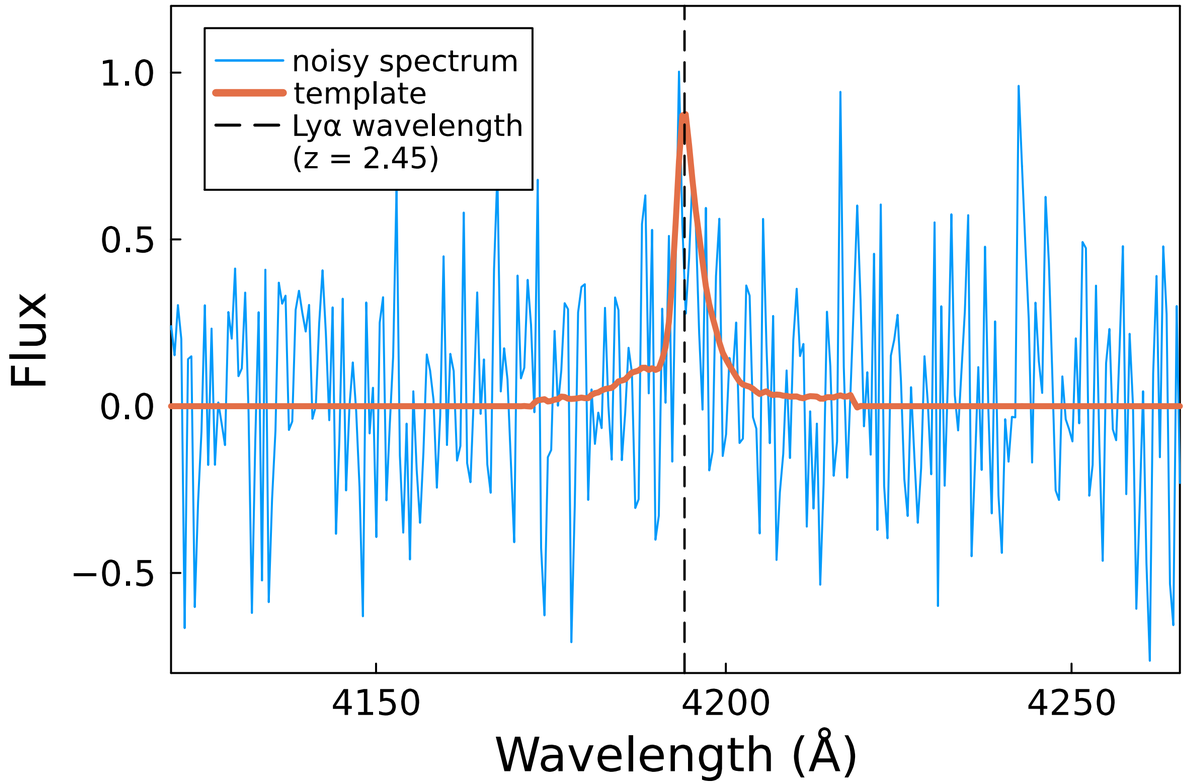


>90% within 0.005
of VI redshift

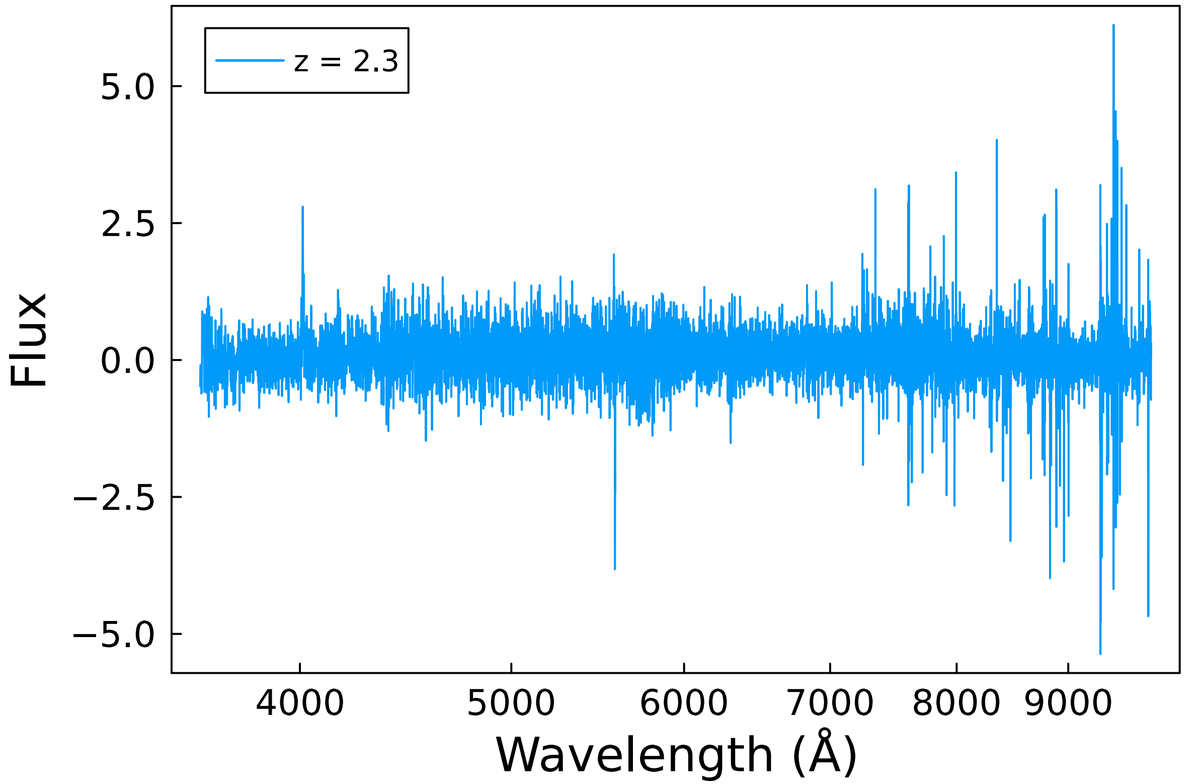
~0.4 s/spectrum

Ana Sofia Uzsoy

We validate our method with injection-recovery tests.

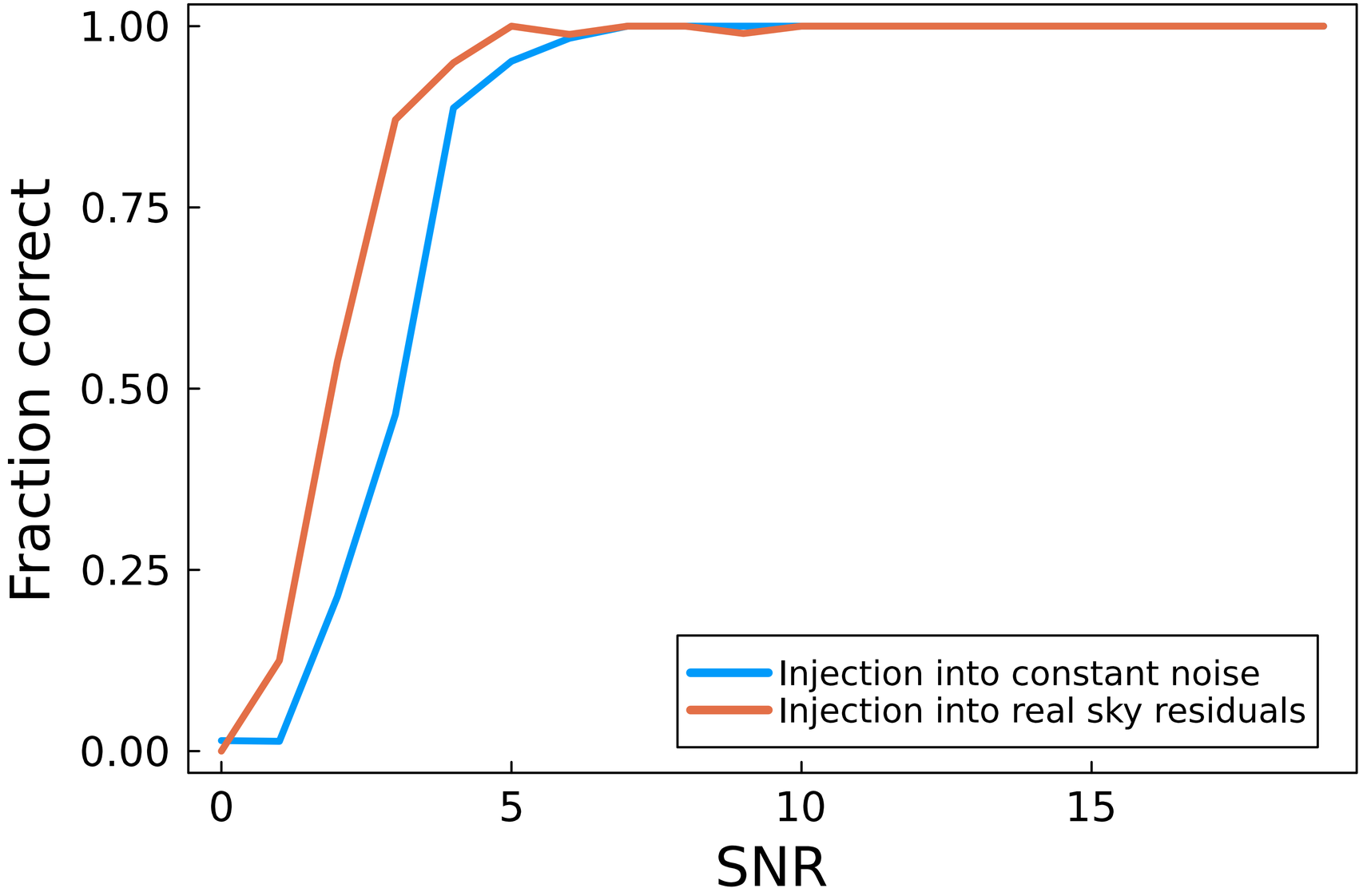


Injection into constant noise



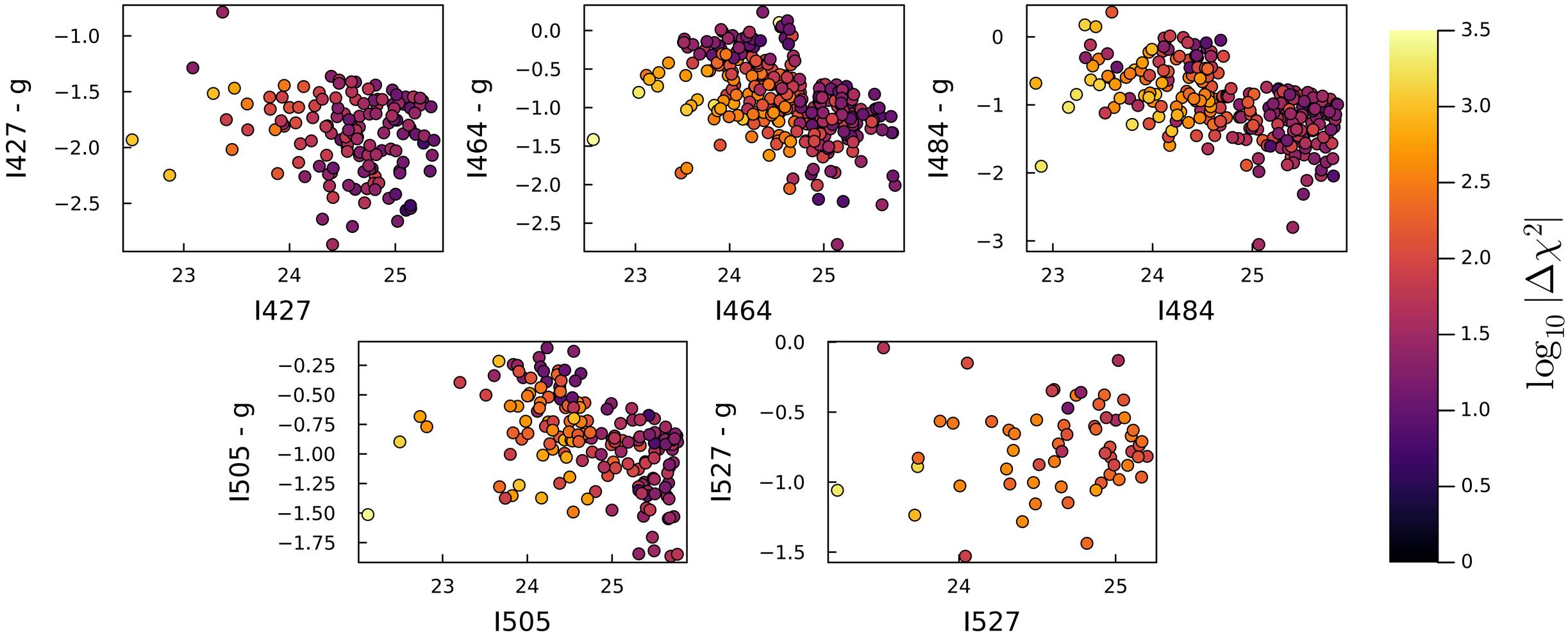
Injection into actual sky spectra

We validate our method with injection-recovery tests.



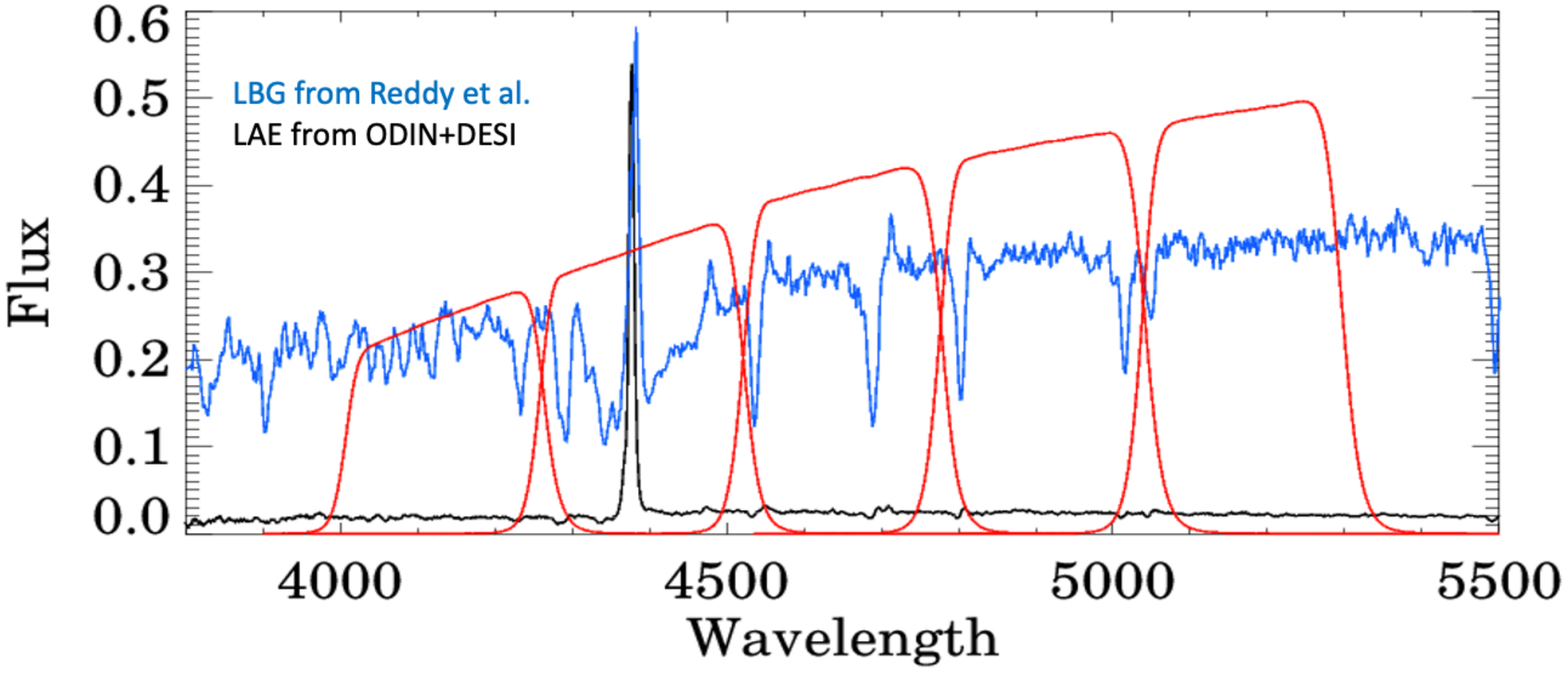
Ana Sofia Uzsoy

Implications for targeting with medium bands



Ana Sofia Uzsoy

Intermediate-Band Imaging Survey (IBIS)

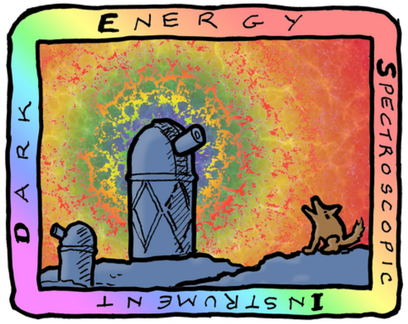


Ana Sofia Uzsoy

PIs: Arjun Dey, David Schlegel


ENERGY
DARK
SPECTROSCOPIC
INSTRUMENT
DARK ENERGY
SPECTROSCOPIC
INSTRUMENT
U.S. Department of Energy Office of Science
DEPARTMENT OF ENERGY · UNITED STATES OF AMERICA ·
NSF
NERSC
FROM RESEARCH TO INDUSTRY
cea
GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMIA Y COMPETITIVIDAD
CONAHCYT
HEISING-SIMONS FOUNDATION
GORDON AND BETTY MOORE FOUNDATION
Science & Technology Facilities Council
Thanks to our sponsors and
72 Participating Institutions!

Summary



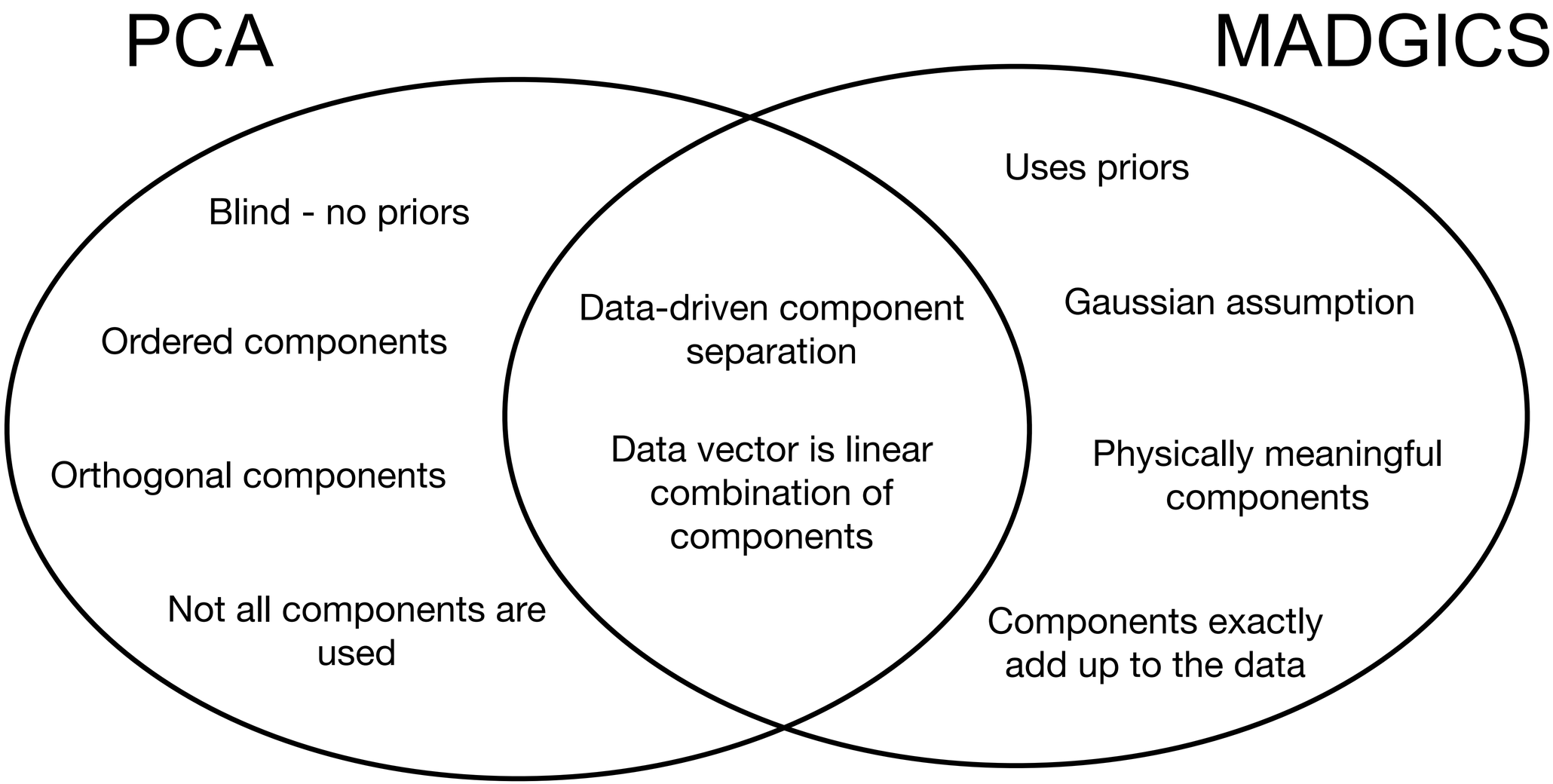
CENTER FOR ASTROPHYSICS
HARVARD & SMITHSONIAN

- MADGICS is a Bayesian component separation technique that separates data into any number of Gaussian components.
- We create data-driven priors and express an LAE spectrum as a sum of a sky, LAE, and residual component.
- We present an automated spectroscopic redshift determination pipeline for LAEs in DESI and explore implications for selecting LAEs with medium-band photometry.

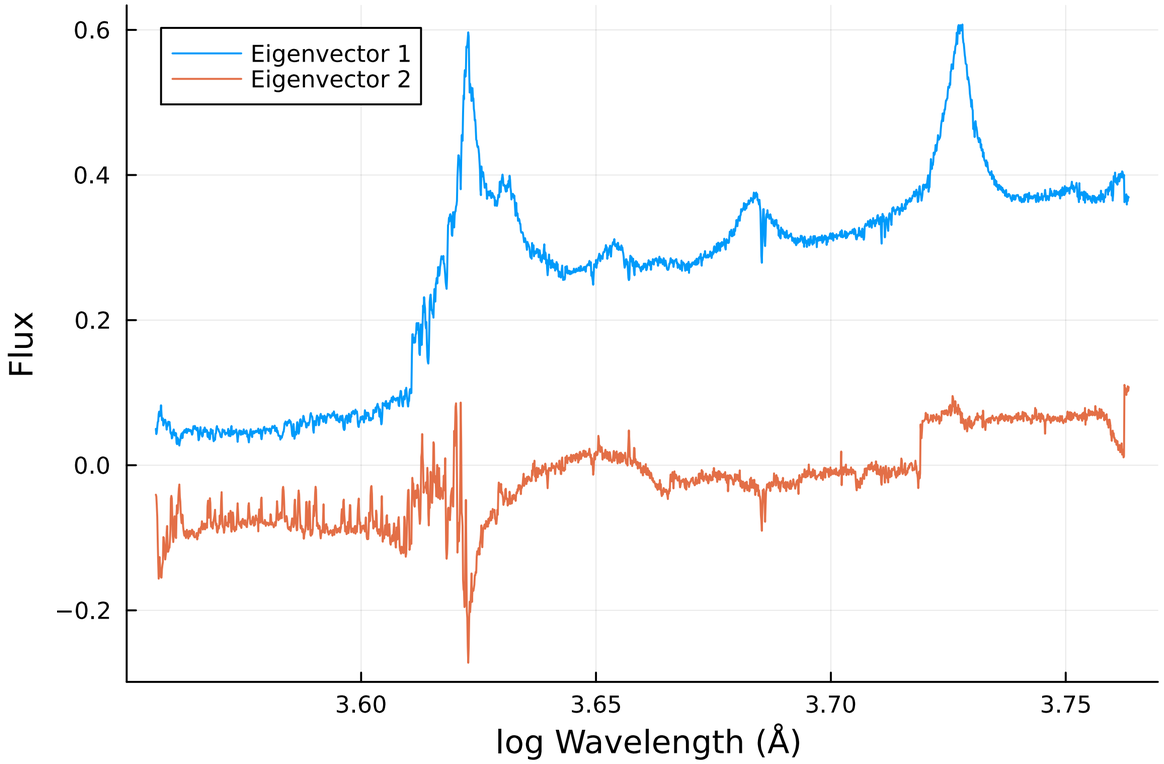
arXiv:2504.06870

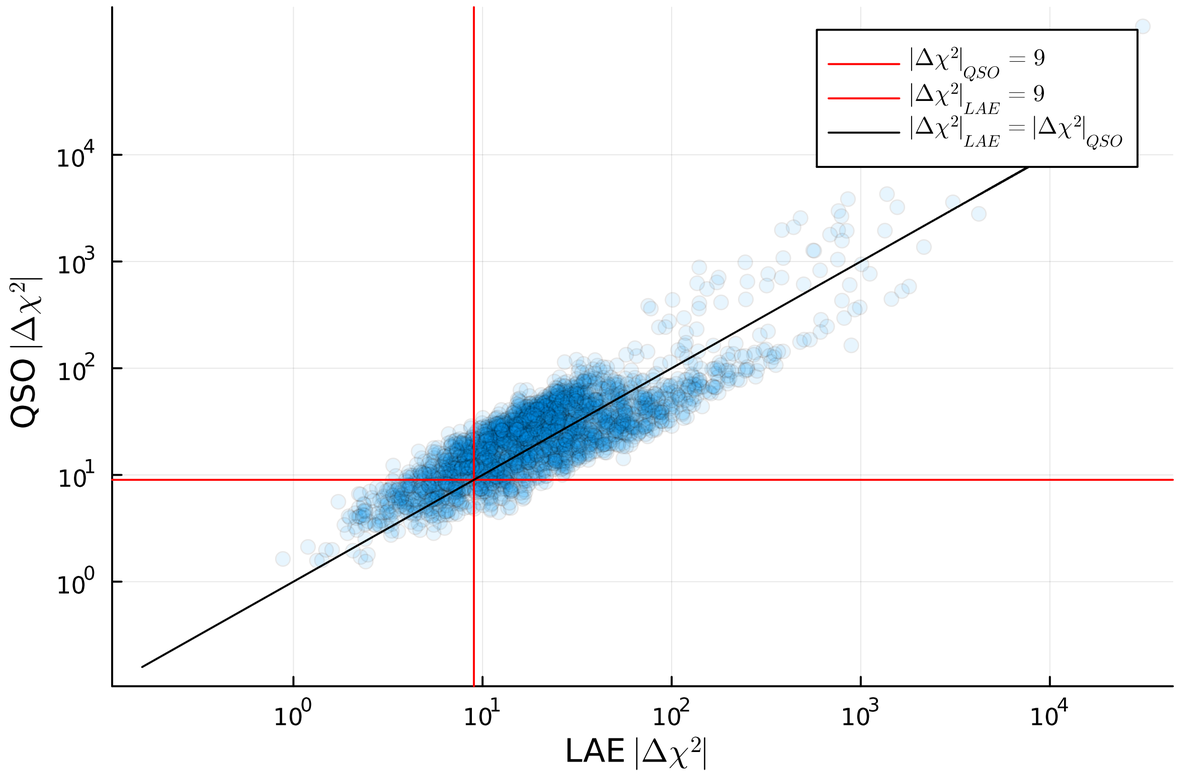
Ana Sofia Uzsoy ana_sofia.uzsoy@cfa.harvard.edu

How is this not just PCA?



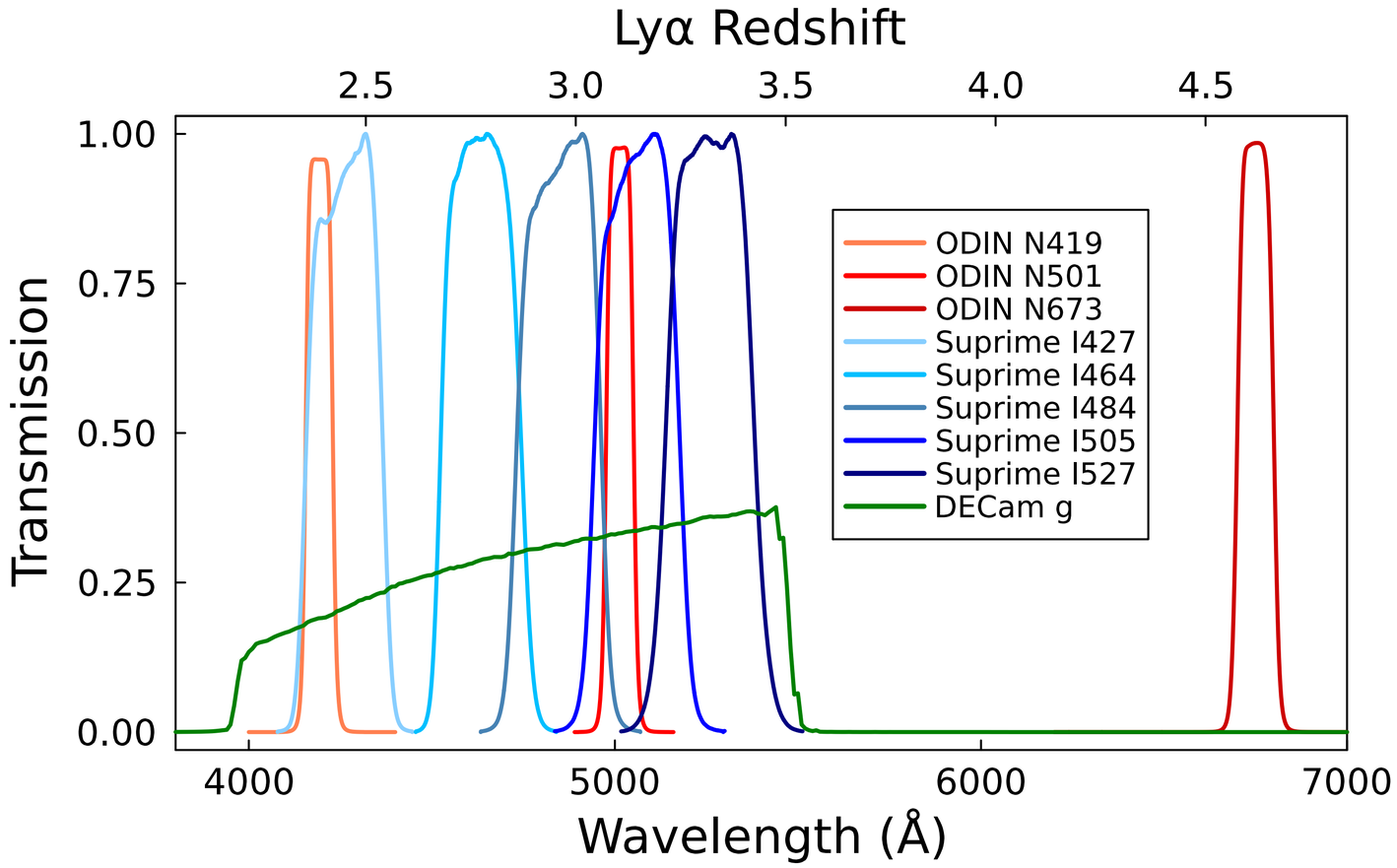
LAE vs. quasar classification





Ana Sofia Uzsoy

We use DESI spectra from LAEs targeted with two different photometric surveys.



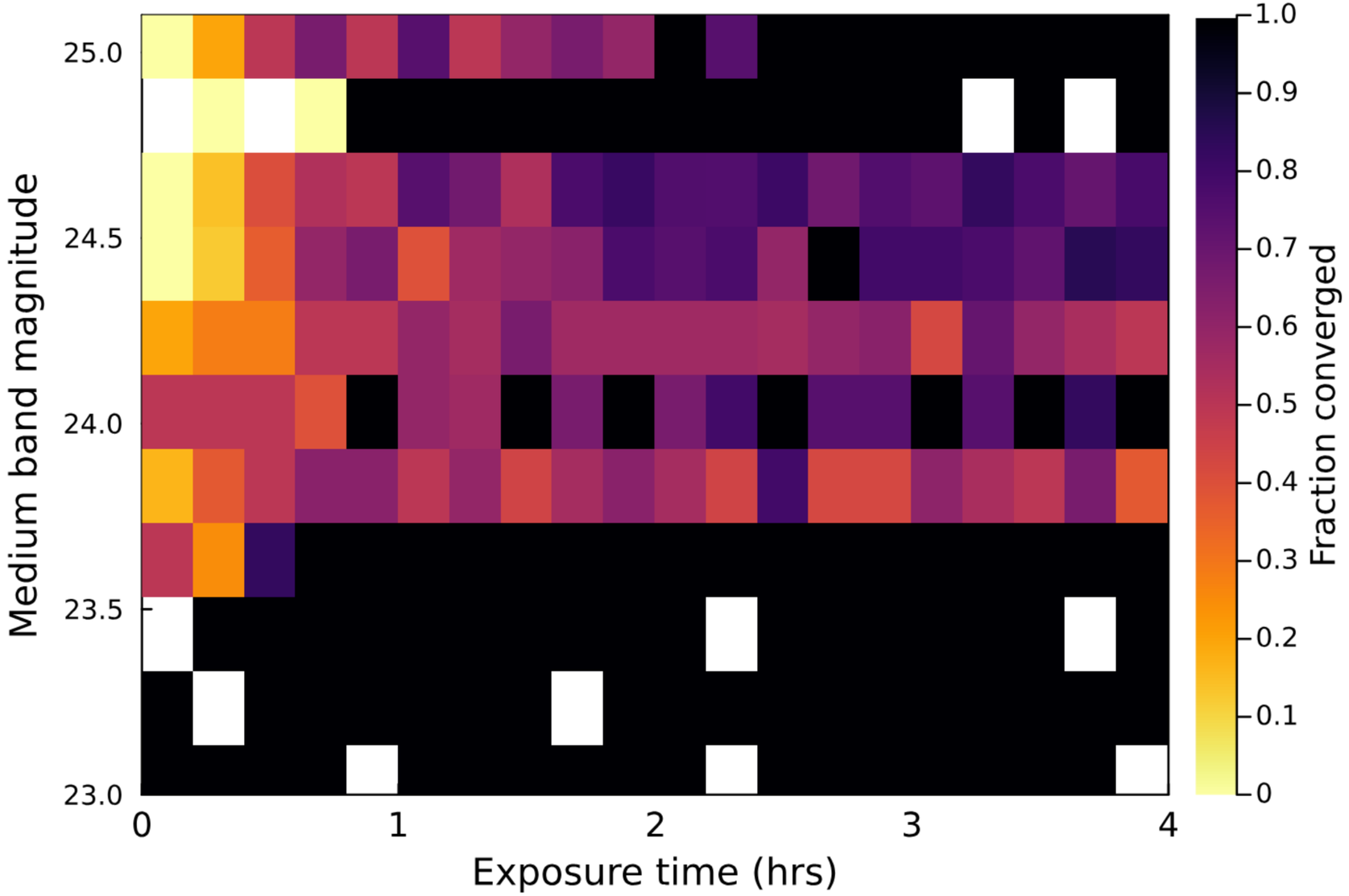
- ODIN
  - DECam
  - Narrow bands, ~100 Å FWHM
- COSMOS
  - Subaru Suprime-Cam
  - Medium bands, ~250 Å FWHM

Summary of priors

- C_{sky} : low-rank matrix approximation from top 50 sky eigenvectors
- C_{LAE} : data-driven covariance matrix from VI ODIN LAE targets all shifted to be at the same redshift
- C_{res} : 1/ivar on diagonal for each target

$$C_{tot} = C_{sky} + C_{LAE} + C_{res}$$

For how long should we observe LAEs?



Ana Sofia Uzsoy

How do we calculate $\Delta\chi^2$?

$$\chi^2 = X_{tot}^T C_{tot}^{-1} X_{tot}$$

We can then evaluate the change in χ^2, or $\Delta\chi^2$, from including vs. not including the LAE component:

$$\Delta\chi^2 = X_{tot}^T \left(C_{sky} + C_{LAE} + C_{res}\right)^{-1} X_{tot} - X_{tot}^T \left(C_{sky} + C_{res}\right)^{-1} X_{tot}$$

where C_{sky}, C_{LAE}, and C_{res} are the covariance matrices for the sky, LAE, and residual components