DARK ENERGY SPECTROSCOPIC INSTRUMENT

U.S. Department of Energy Office of Science

THE OHIO STATE UNIVERSITY

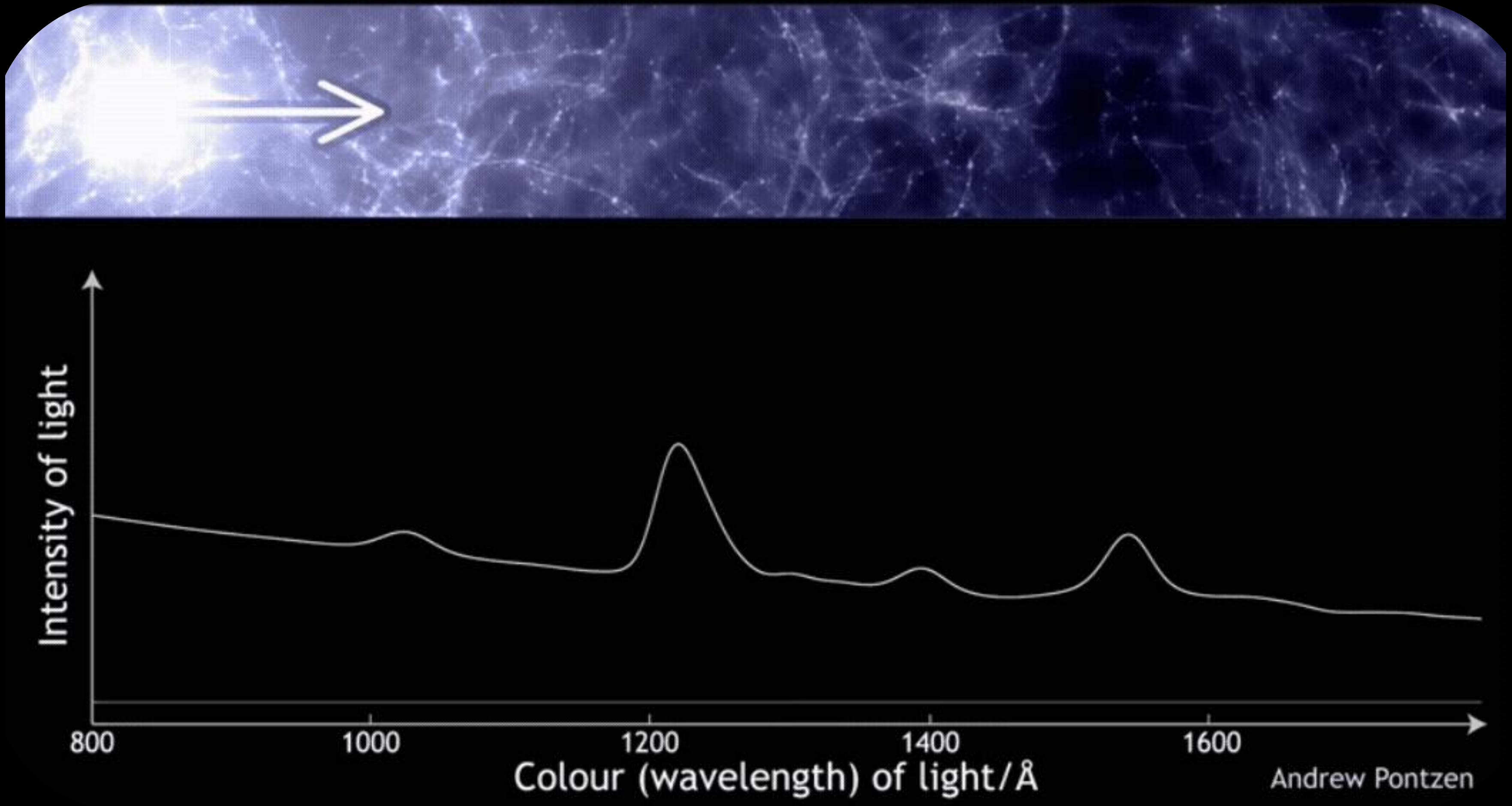
Novel Methods for Lyman-α Forest Cosmology

Wynne Turner

Ohio State University

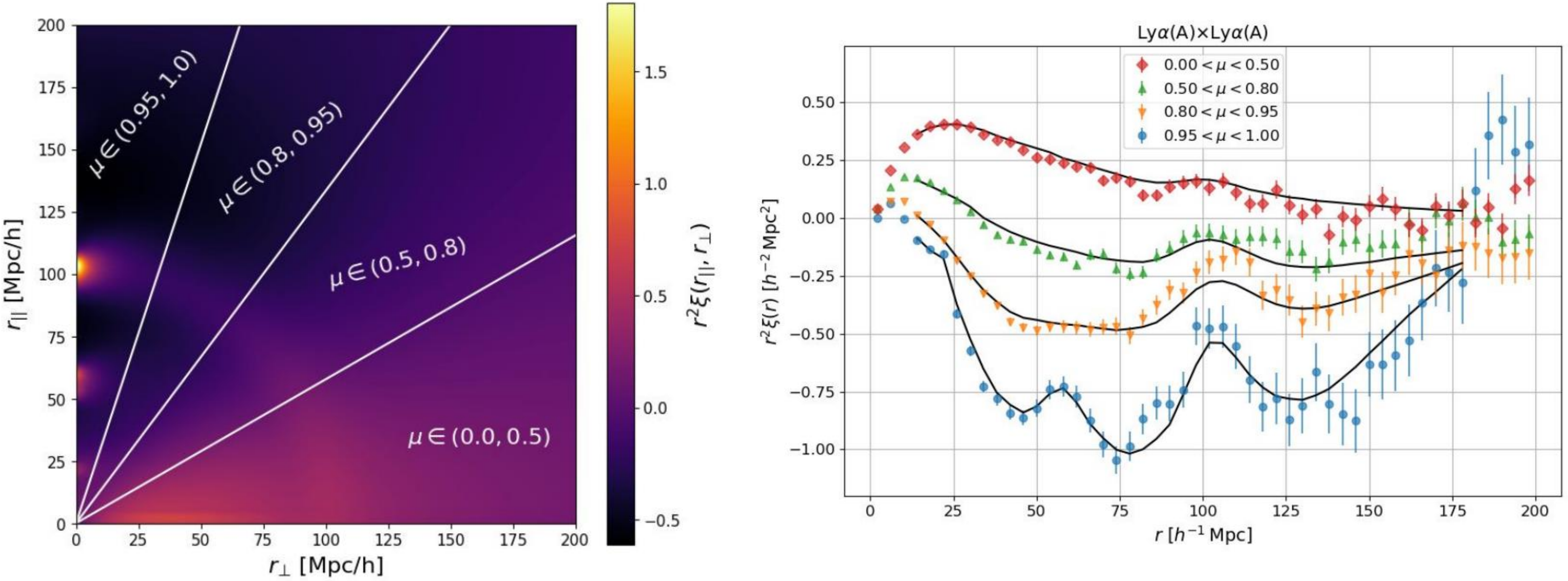
Cosmic Horizons Conference – July 2026

Why the Lyman-α forest?



- Probes ionization state, temperature, density of IGM
- Traces large-scale structure: great for $z > 2$ cosmology!
- DESI is measuring spectra of ~ one million Lyα forest quasars

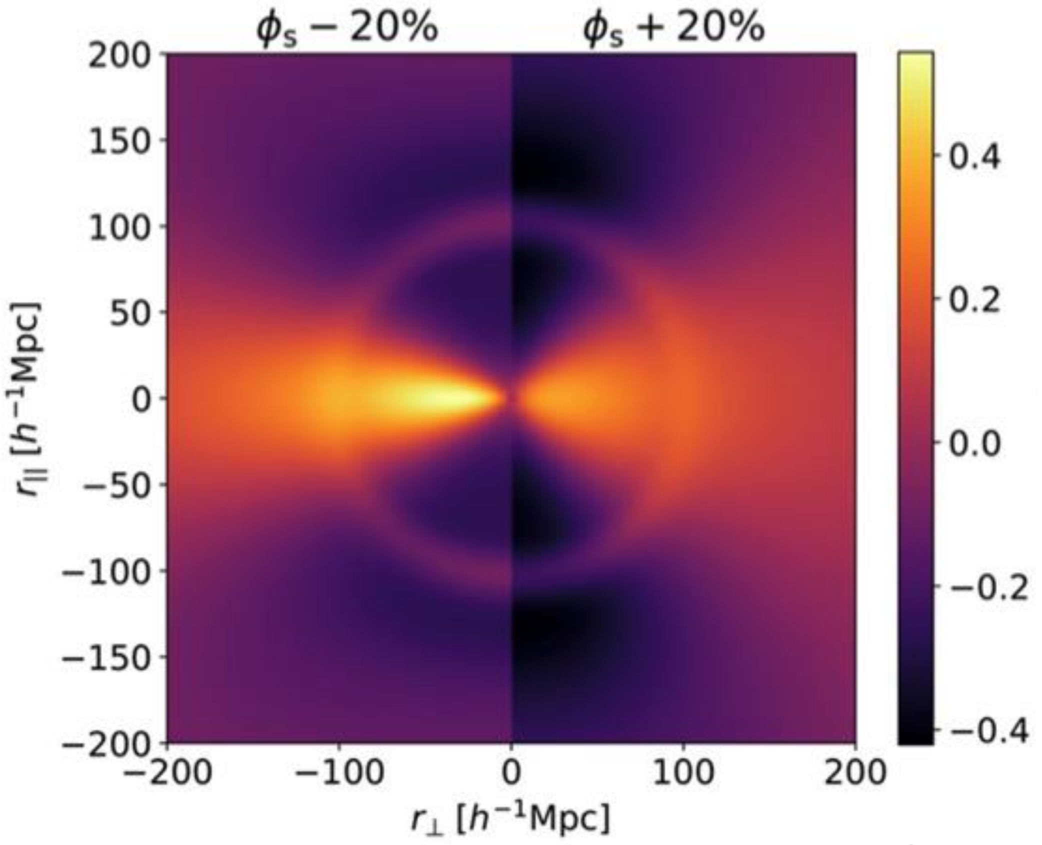
Lyα forest two-point correlations

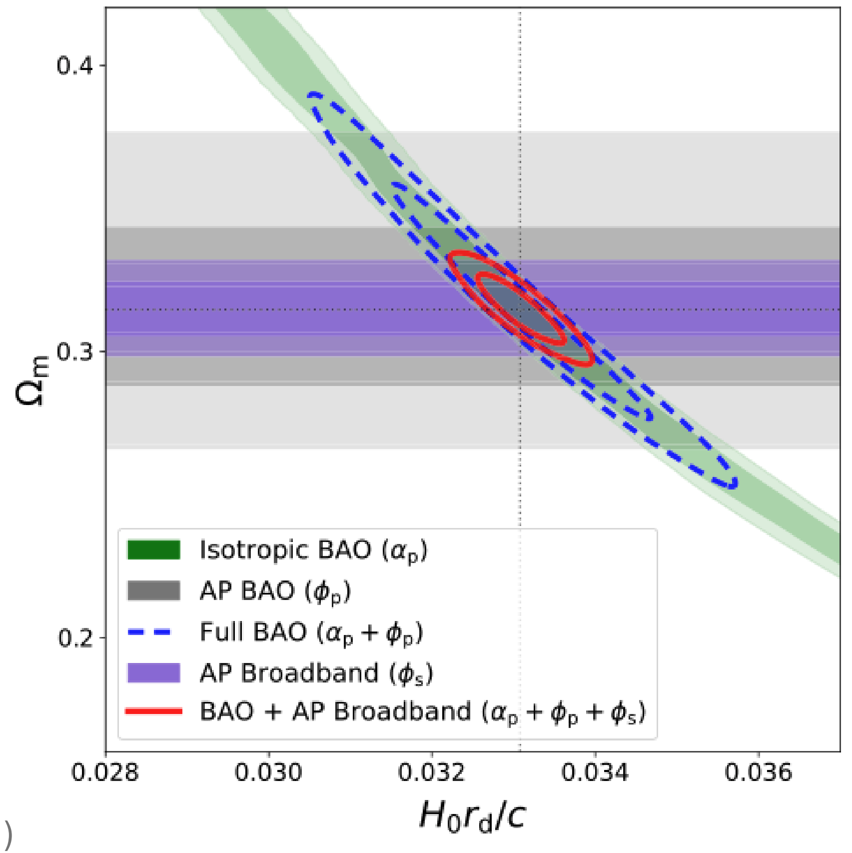


Credit: Andrei Cuceu and the DESI Collaboration

Cosmology beyond BAO

- More to be gained from the Alcock-Paczyński (AP) effect: factor of ~2 improvement
- Arises from anisotropy introduced by incorrect fiducial cosmology

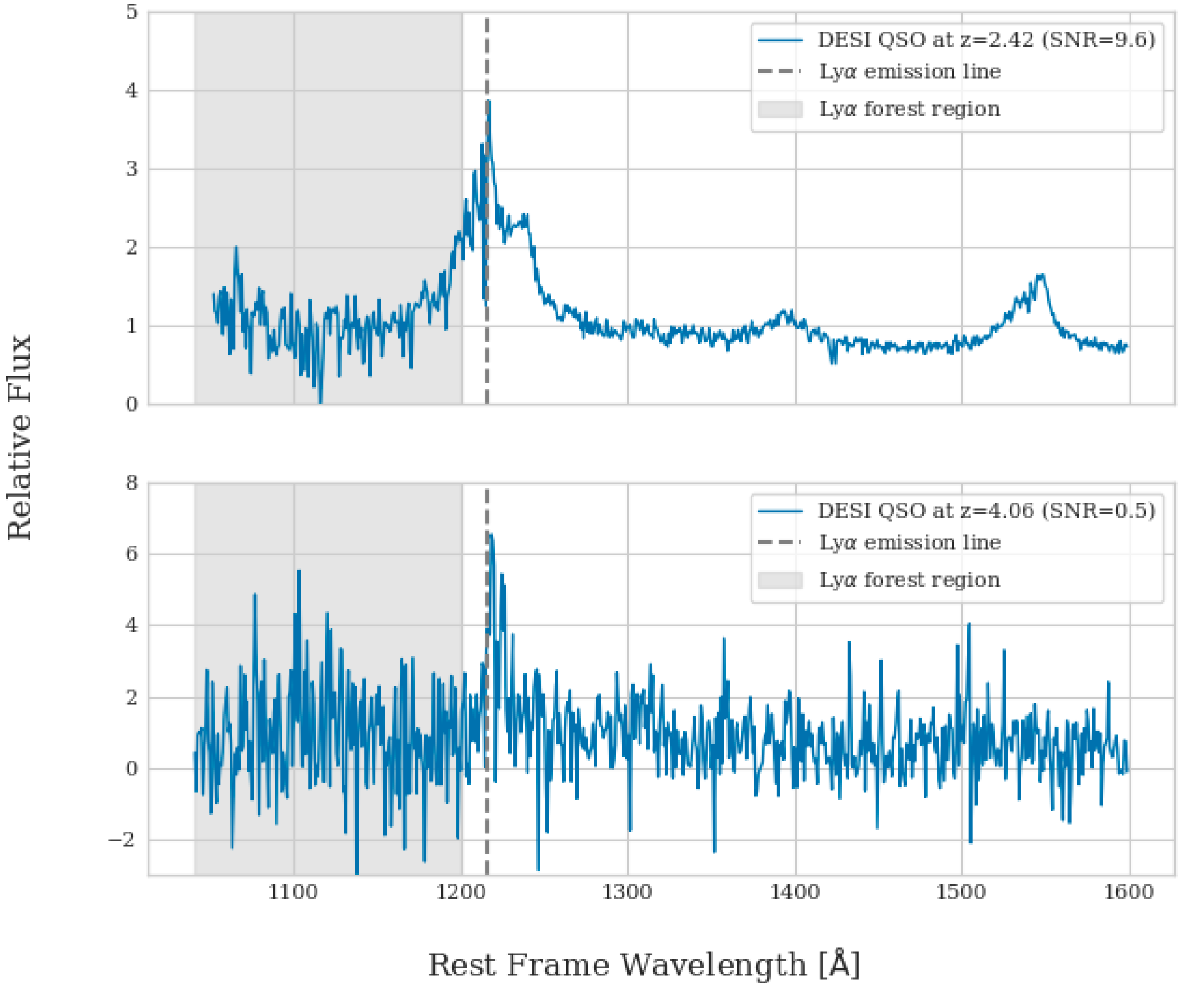




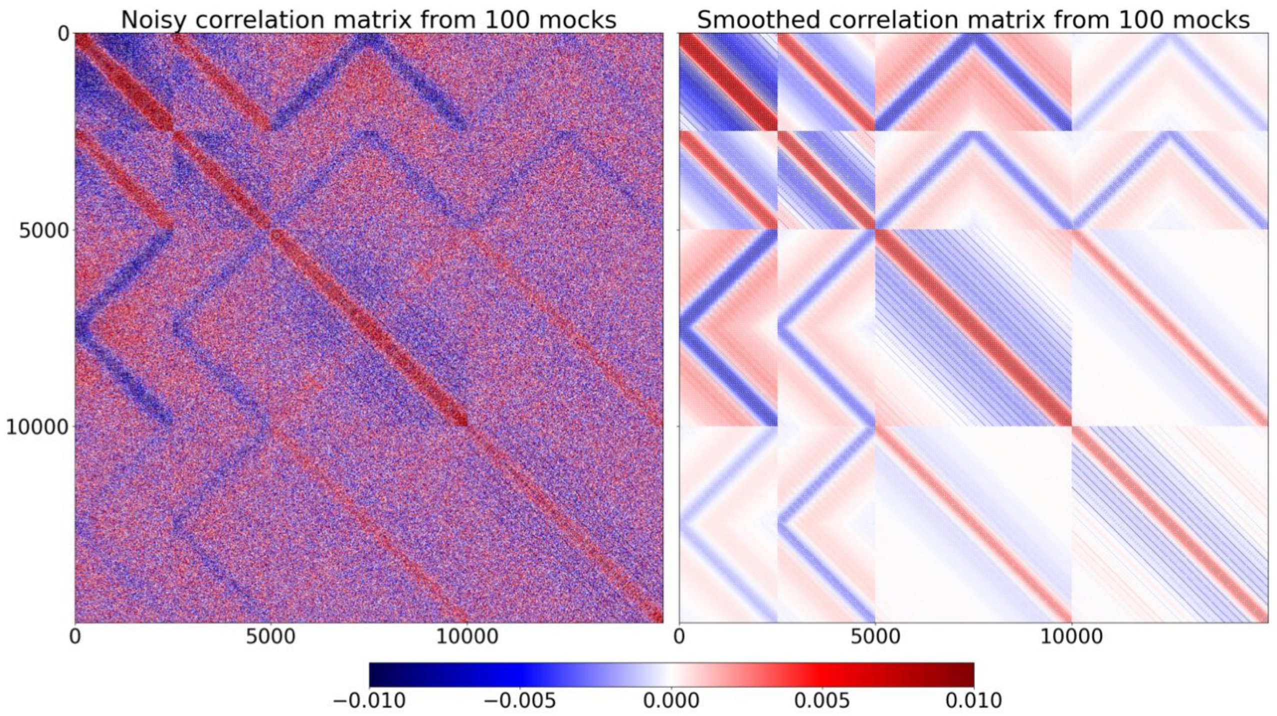
Cuceu et al. (2021)

Challenges in Lyα forest cosmology

Requires estimates of the quasar continuum



Covariance estimation: ad-hoc smoothing procedure

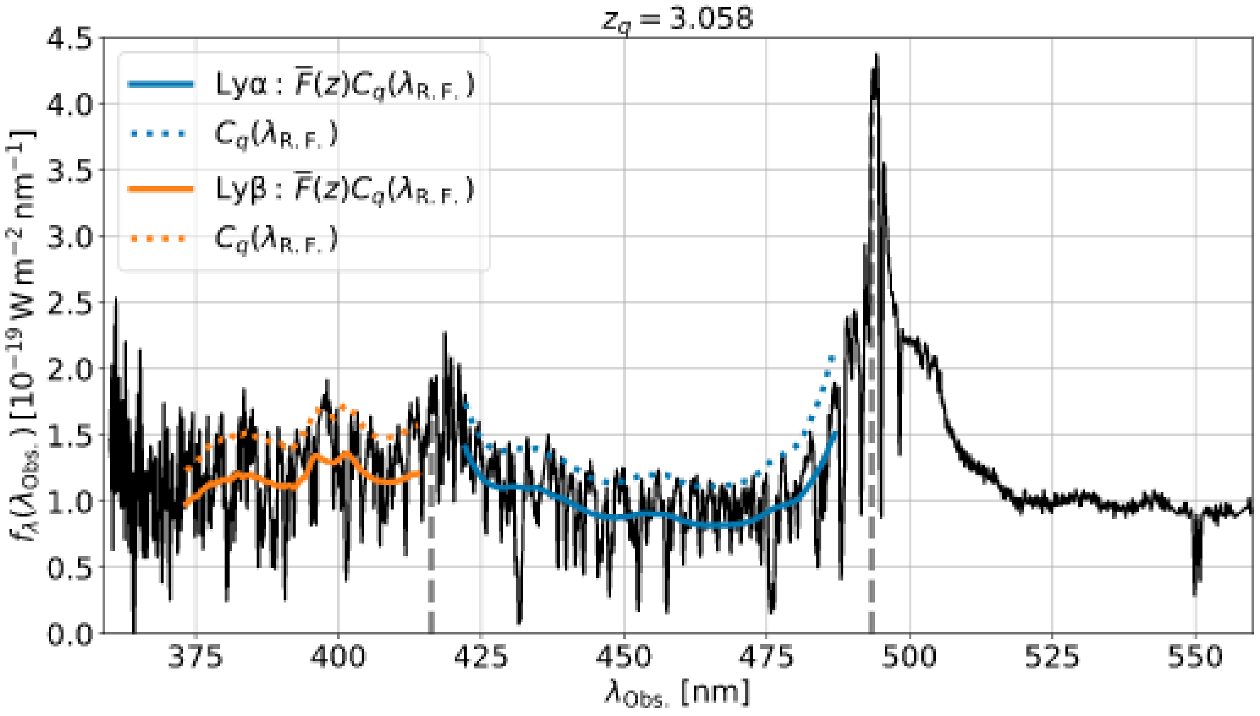


Cuceu et al. (2025)

Current continuum fitting method

- Forest continuum cannot be measured directly in DESI data
- Current method estimates the expected flux $\bar{F}(\lambda)C_q(\lambda)$ for each quasar
  - Mean continuum $\bar{C}(\lambda_{\mathrm{RF}})$ is measured from data
- Removes some information on large scales
  - Distorts correlation function on all scales
  - Not recovering all the information in the two-point correlation functions

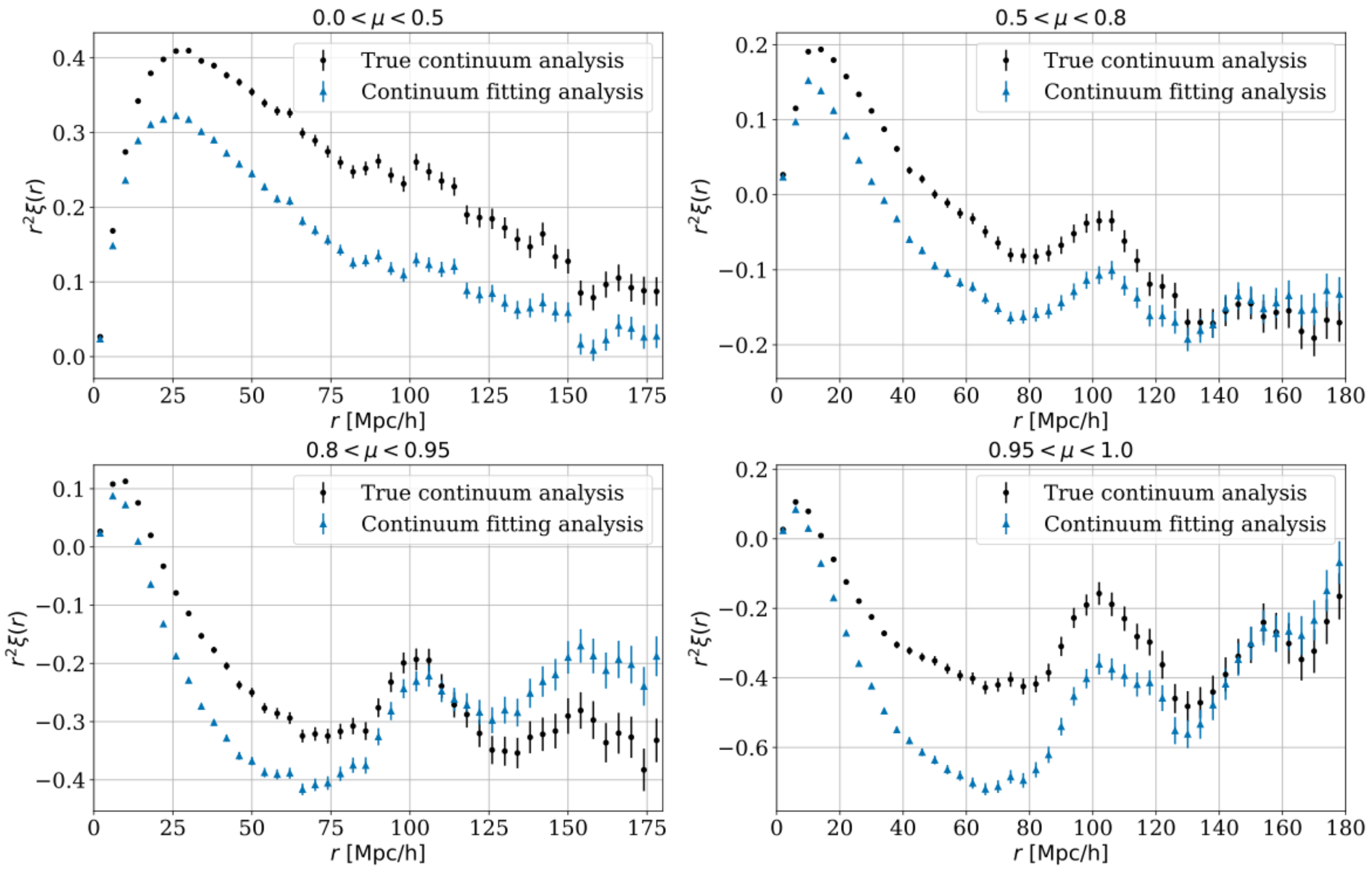
$$\bar{F}(\lambda)C_q(\lambda) = \bar{C}(\lambda_{\mathrm{RF}})\left(a_q + b_q \log\lambda\right)$$



du Mas des Bourboux et al. (2020)

Current continuum fitting method

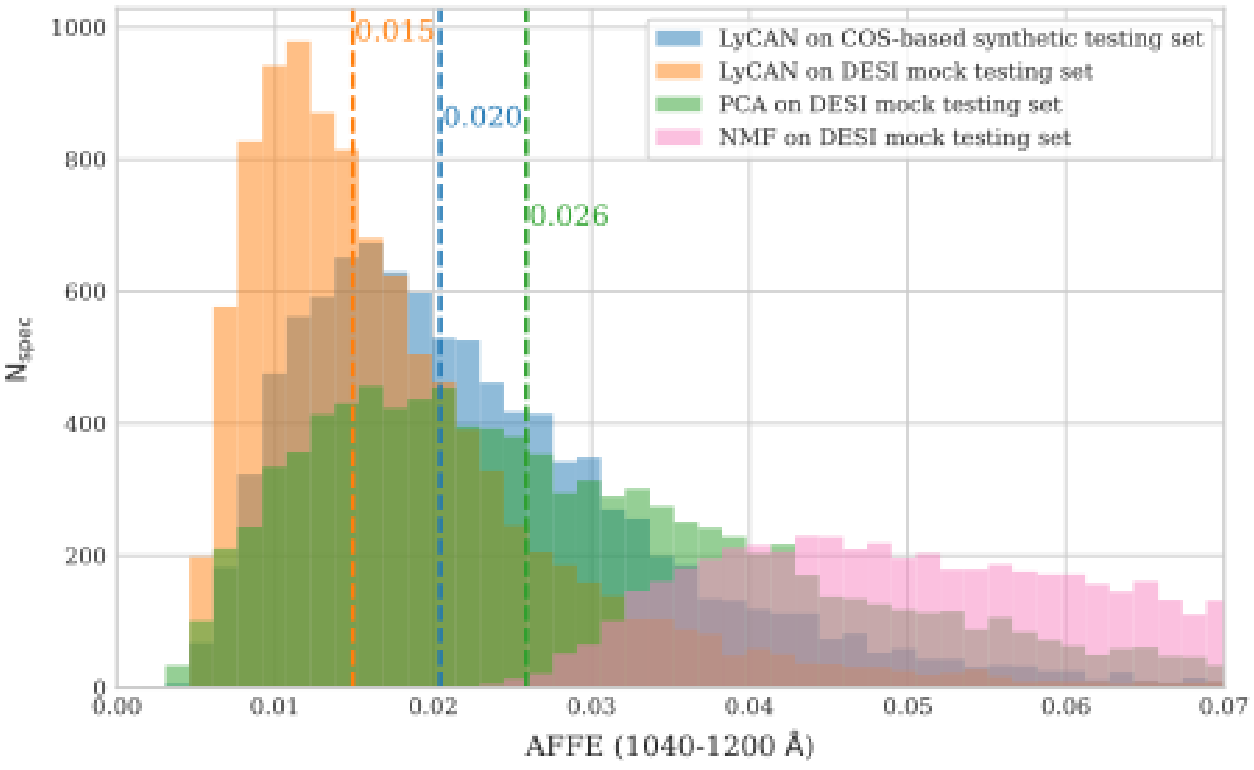
- Forest continuum cannot be measured directly in DESI data
- Current method estimates the expected flux $\bar{F}(\lambda)C_q(\lambda)$ for each quasar
    - Mean continuum $\bar{C}(\lambda_{\rm RF})$ is measured from data
- Removes some information on large scales
    - Distorts correlation function on all scales
    - Not recovering all the information in the two-point correlation functions

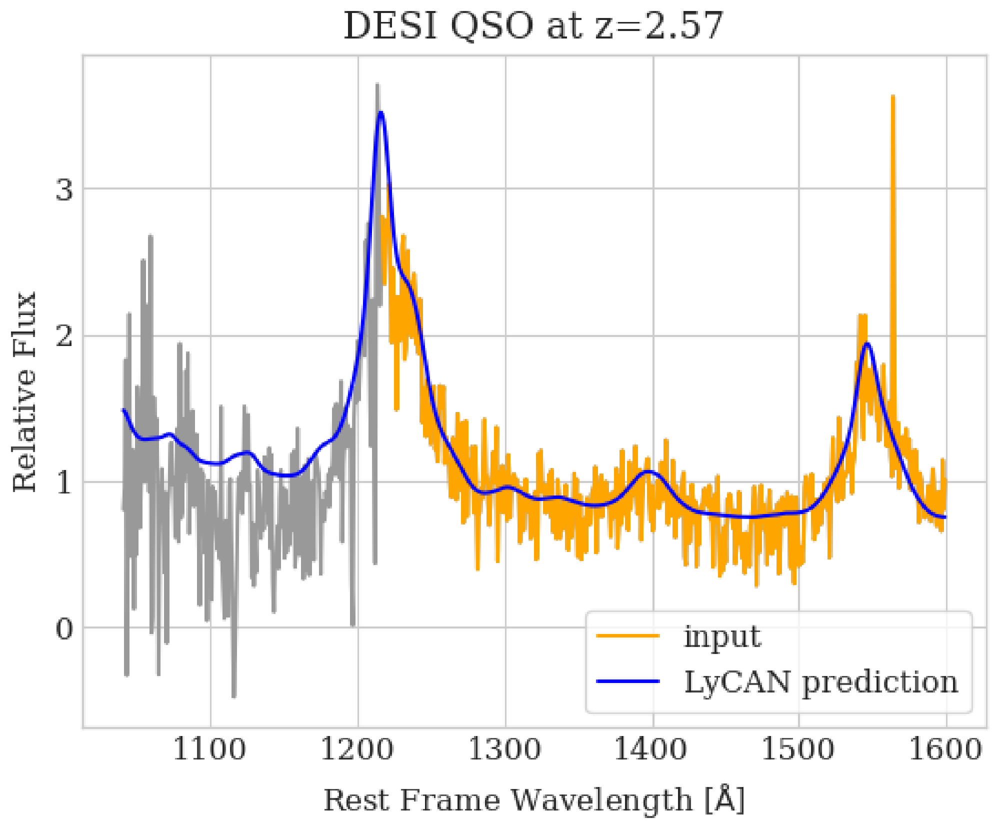


LyCAN

Lyα **C**ontinuum **A**nalysis **N**etwork

- CNN that predicts unabsorbed continuum from 1040-1600 Å based only on red side (1216-1600 Å)
- Trained on low-z COS spectra + DESI Y5 mock
- Measured optical depth evolution with DESI DR1



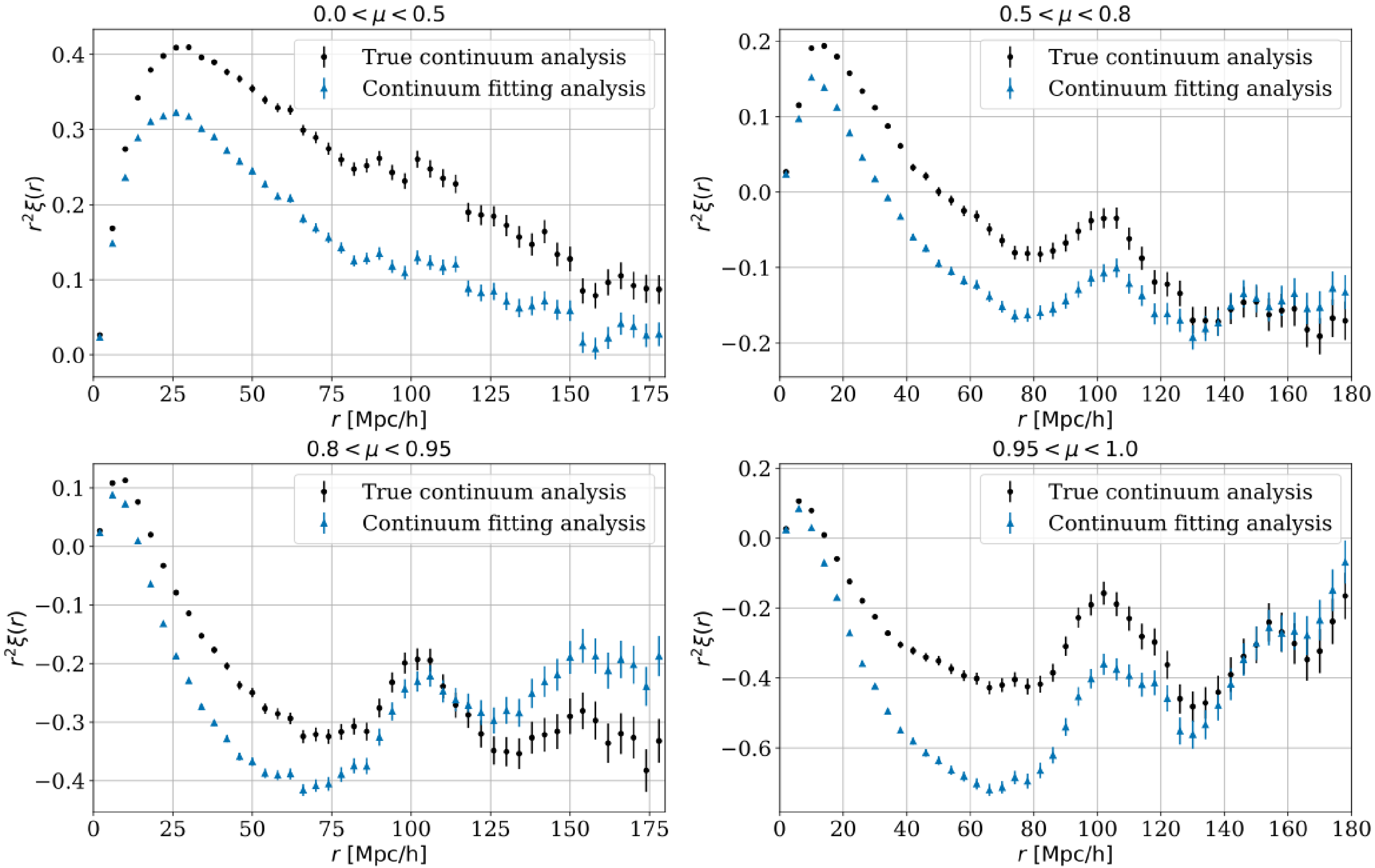


Achieves ~2% error in forest on testing set mocks

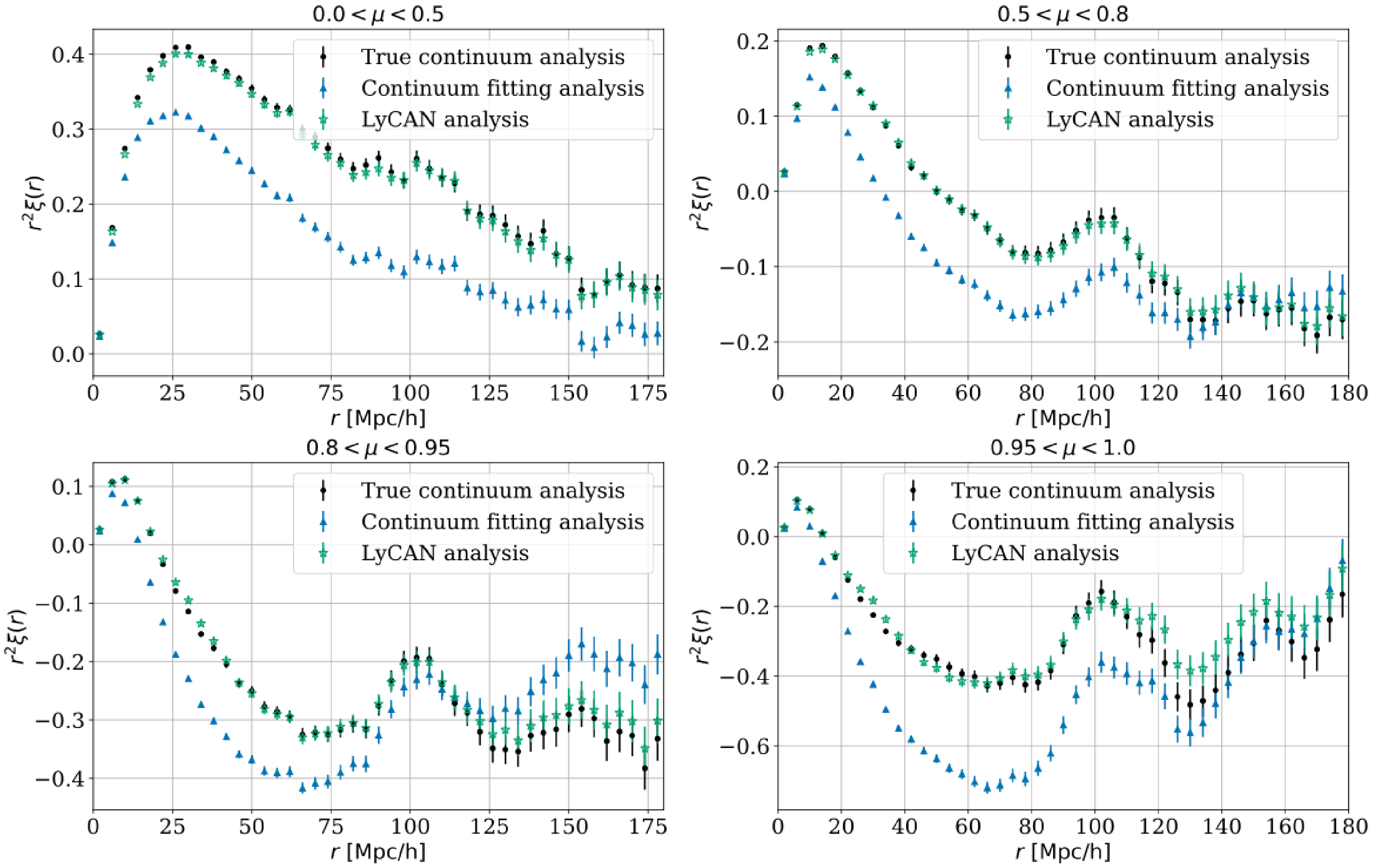
$$\mathrm{AFFE} = \frac{\int_{1040\text{Å}}^{1200\text{Å}} \left| \frac{F_{\mathrm{pred}}(\lambda) - F_{\mathrm{true}}(\lambda)}{F_{\mathrm{true}}(\lambda)} \right| d\lambda}{\int_{1040\text{Å}}^{1200\text{Å}} d\lambda}$$

Turner et al. (2024)

Can we measure undistorted correlation functions?



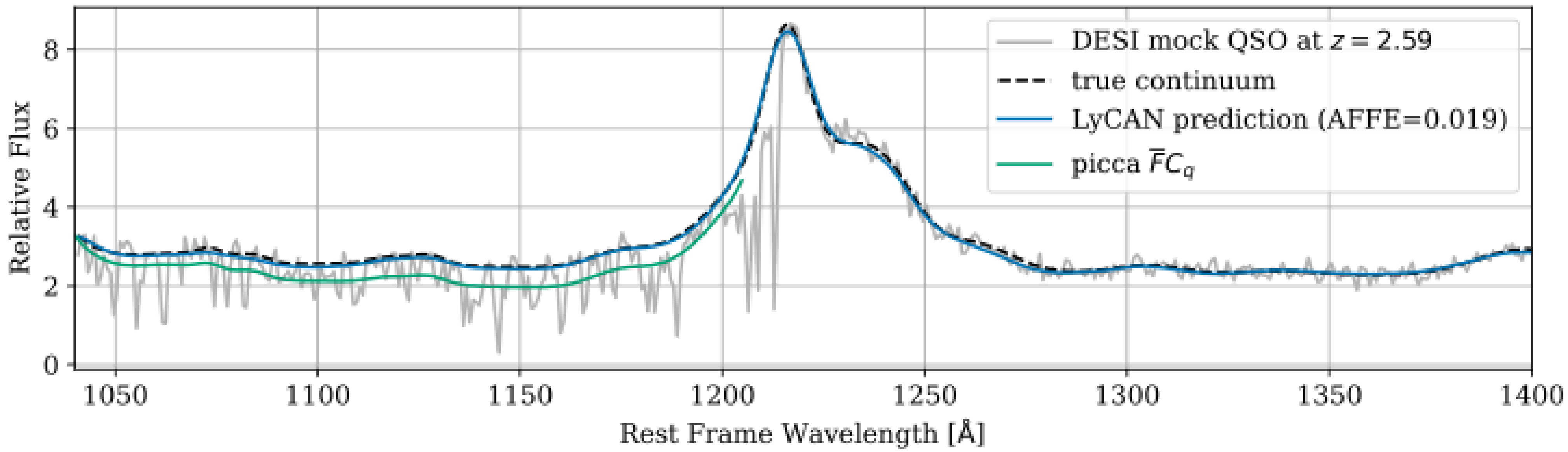
Can we measure undistorted correlation functions?



Turner et al. (2026)

Impact on cosmology

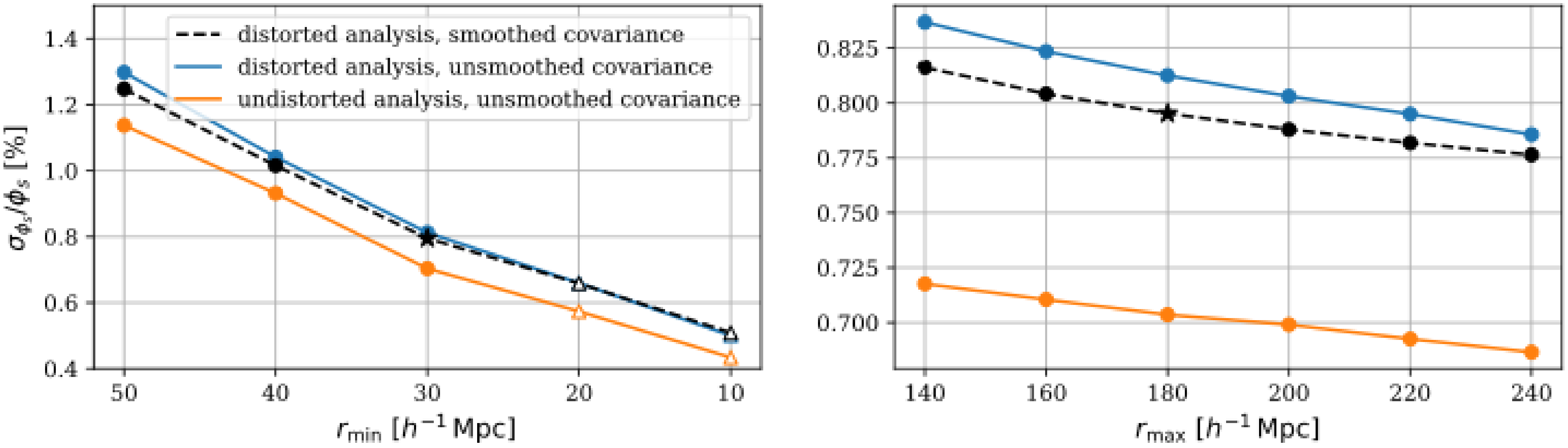
- **Question:** how much better can we measure cosmological parameters with knowledge of the true continuum?
- **Method:** full-shape forecasts using DESI DR2 mocks (idealized case without contaminants)
    - Covariance matrices from mocks with the true continuum vs continuum fitting
    - Also explored different scale cuts



Turner et al. (2026)

Forecast results

- ~10-15% improved AP constraints with the true continuum
  - → ~40% extension in area
- More information available at larger scales, but dominant effect is the true continuum



Turner et al. (2026)

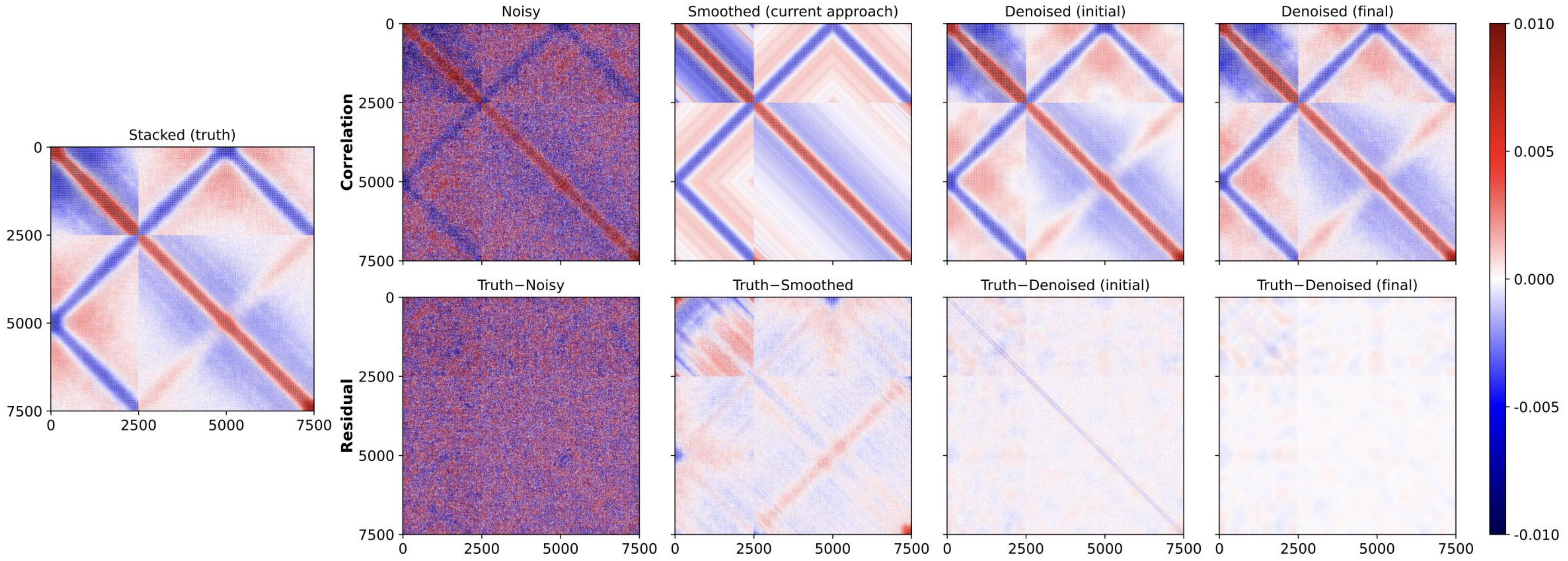
Covariance denoising

Step 1: project noisy matrix onto eigenbasis of mock-based reference

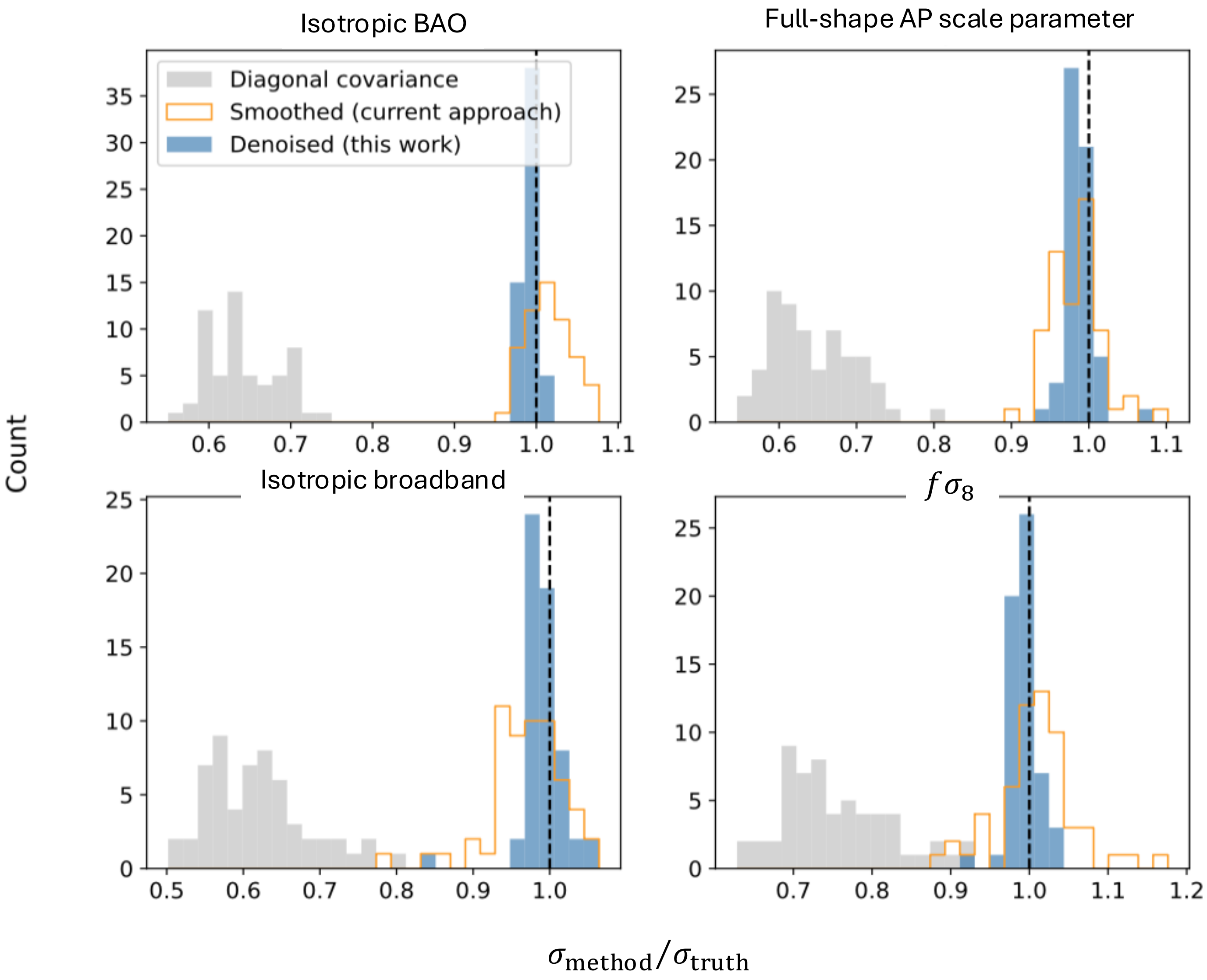
Construct initial denoised matrix

Classify relative to mock suites

Step 2: project noisy *residual* onto mock suite-weighted residual eigenbasis



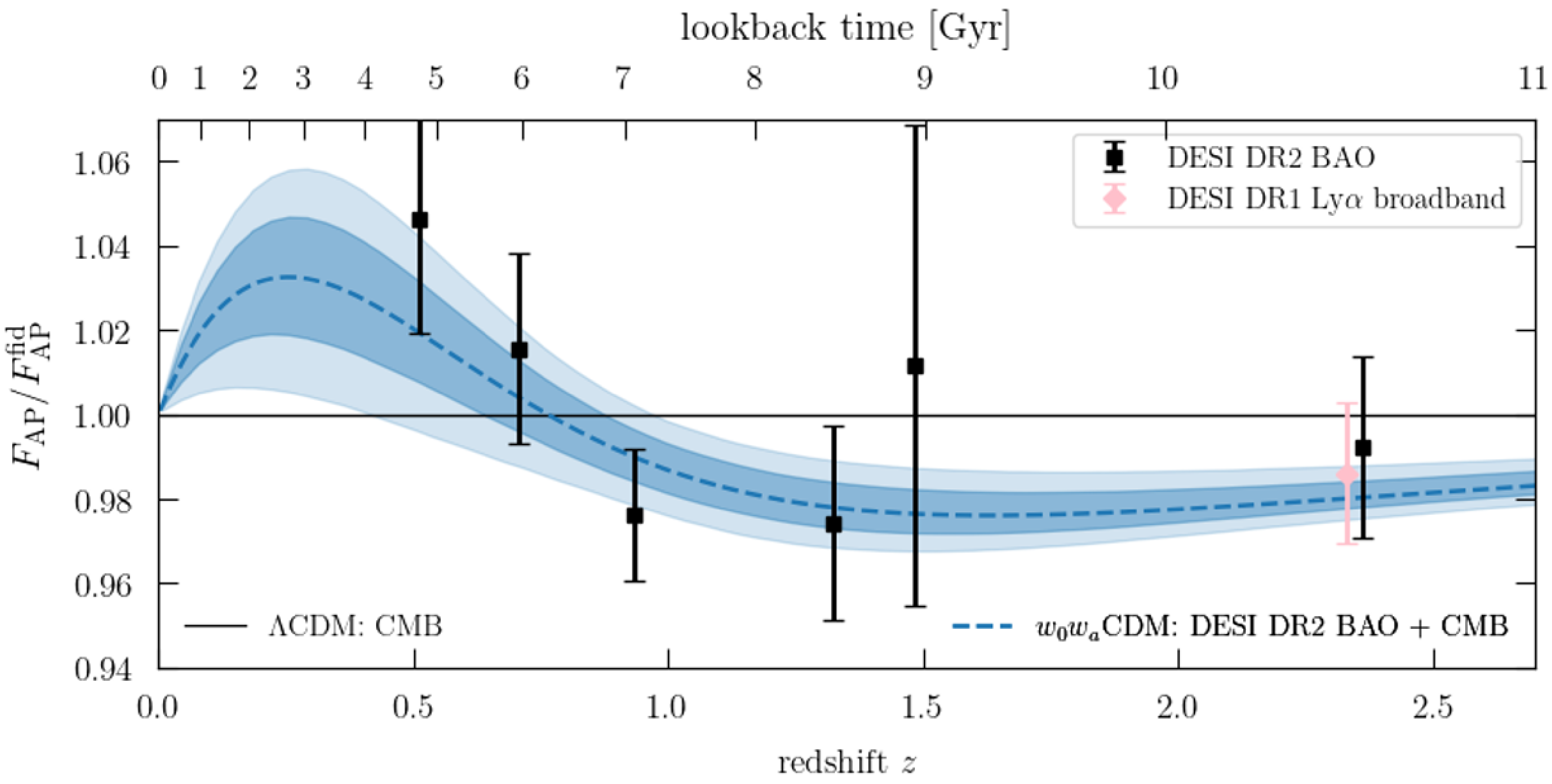
Effect on cosmological parameters



Denoised covariances result in most consistent uncertainty recovery across mocks

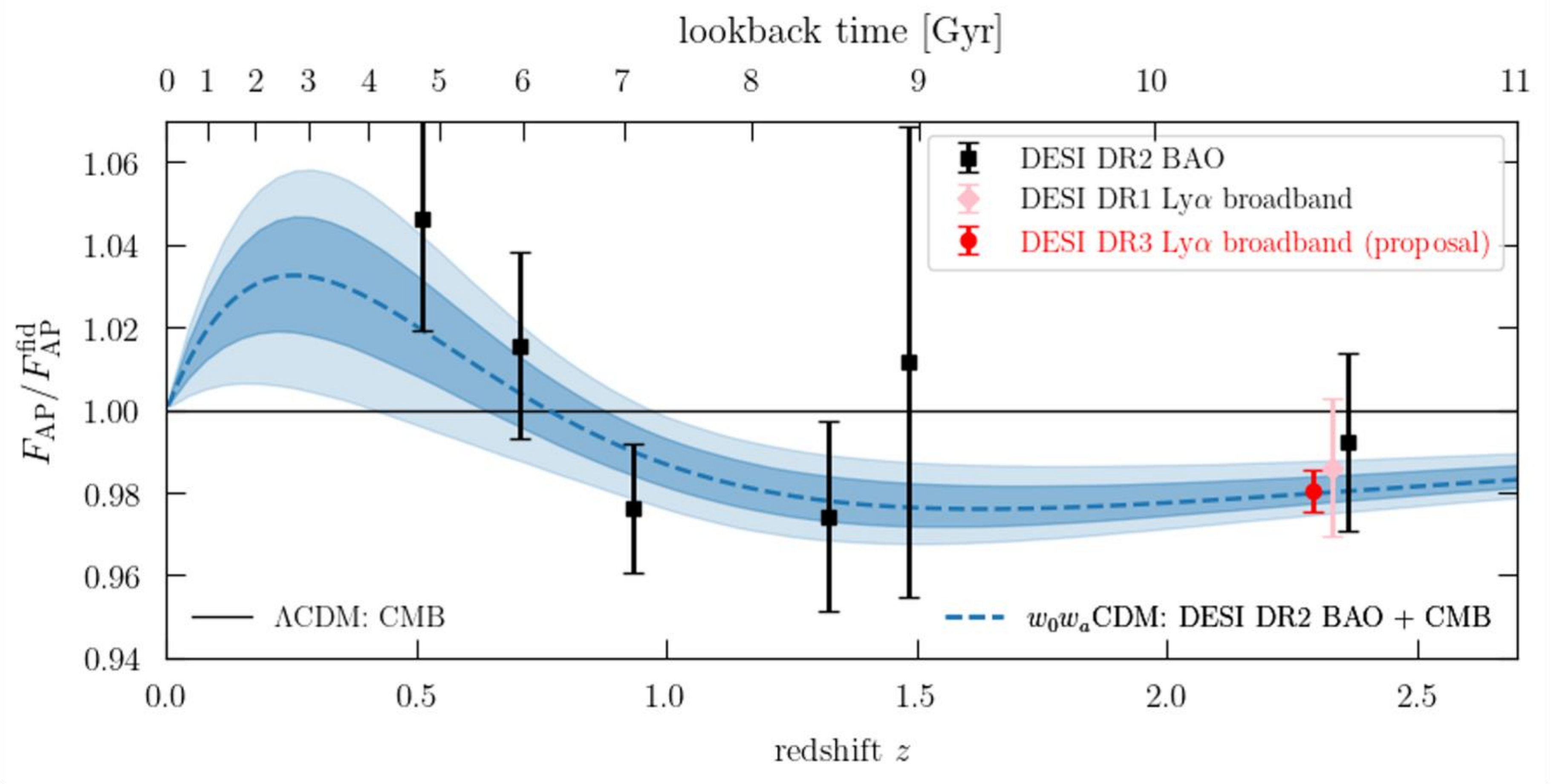
Turner (2026)

Future prospects



Adapted from Cuceu et al. (2025)

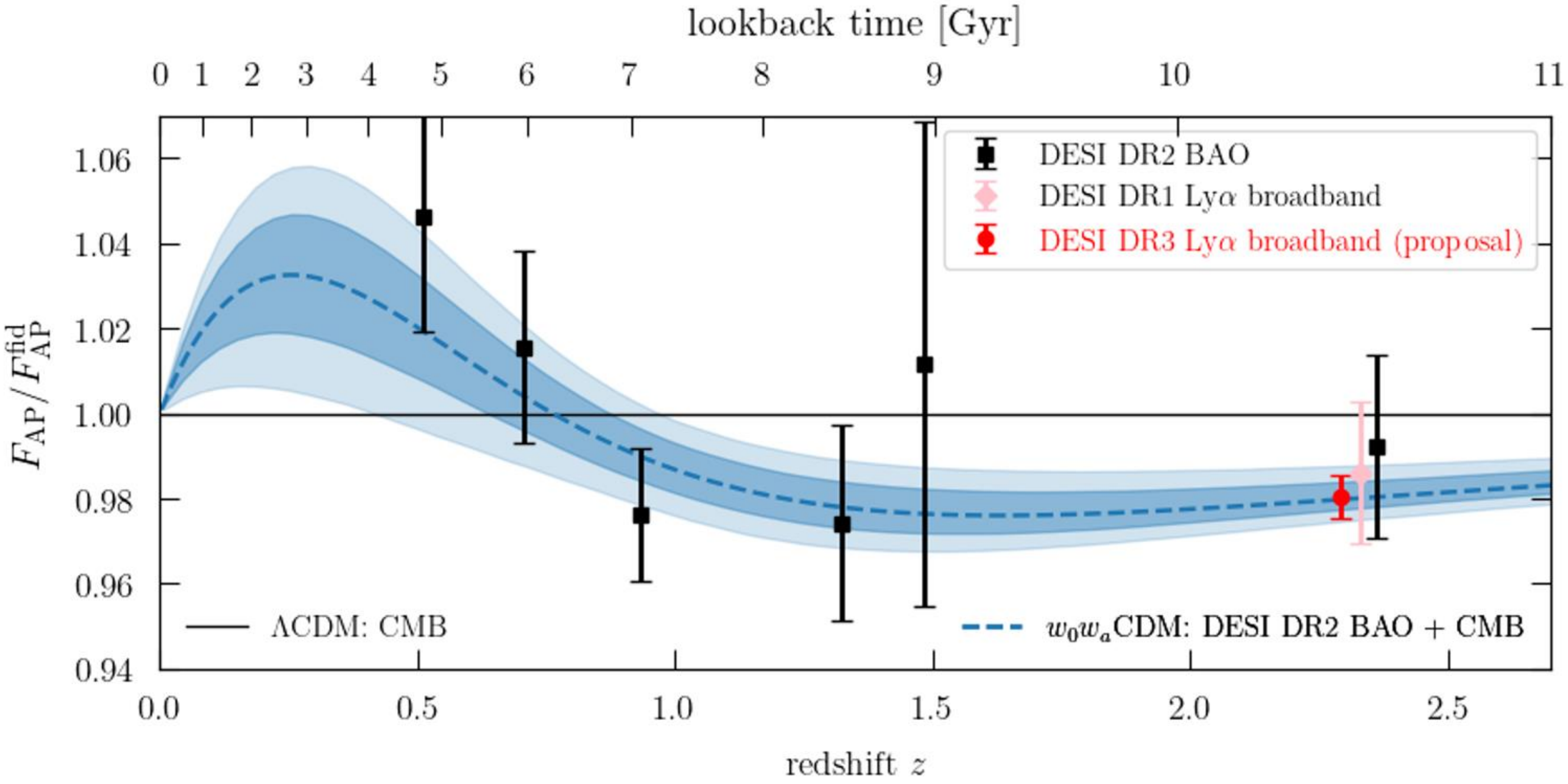
Future prospects



Adapted from Cuceu et al. (2025)

Summary

- Access to the unabsorbed continuum could enable ~10-15% tighter constraints on the AP parameter
  - May be feasible with a method like LyCAN and improved covariance estimation
- Will help the Lyα forest distinguish between different cosmological models at high redshift



Extra Slides

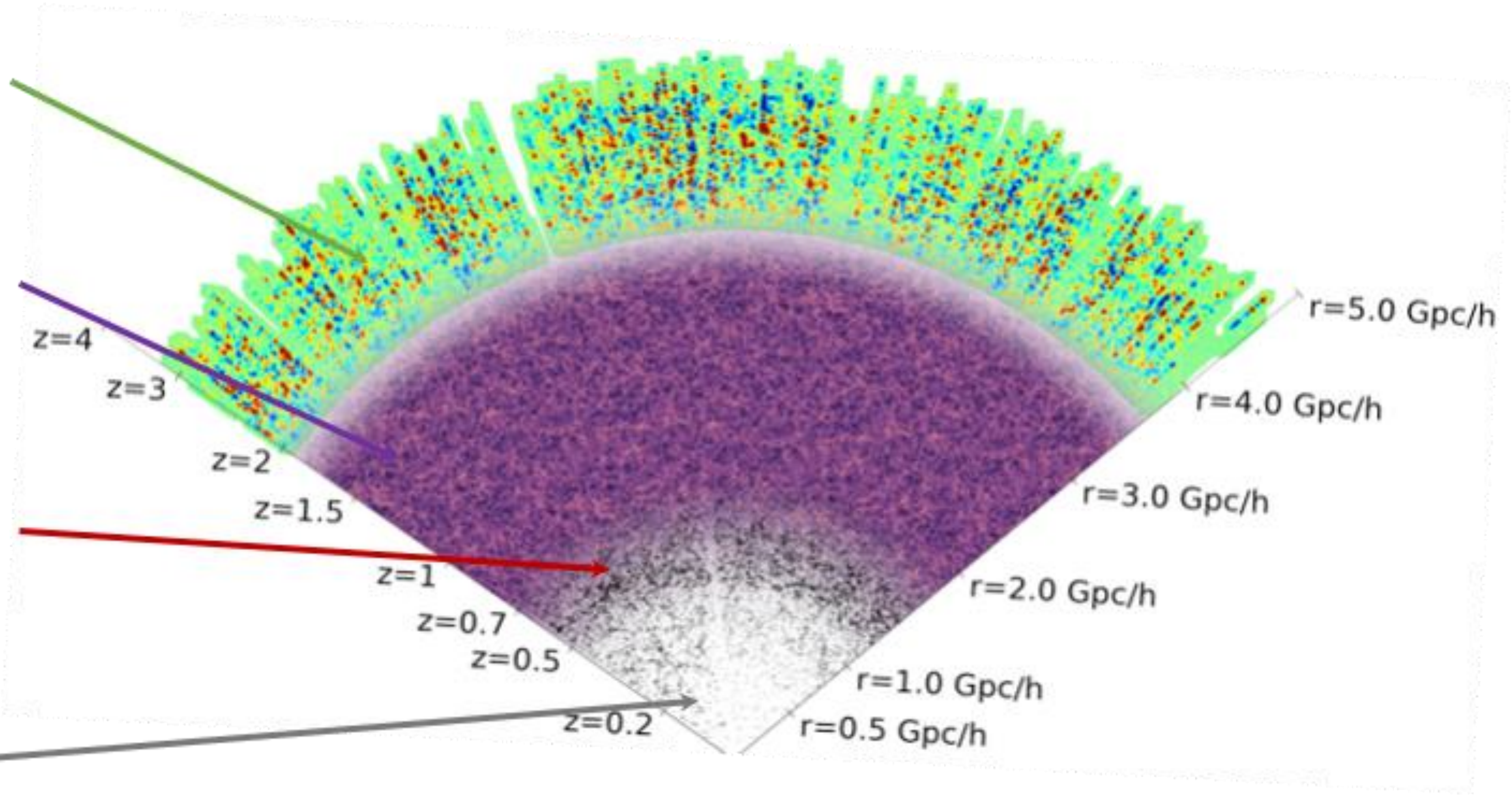
DESI tracers

3 million quasars + Ly-α forest ($1 < z < 3.5$)

16 million Emission Line Galaxies ($0.6 < z < 1.6$)

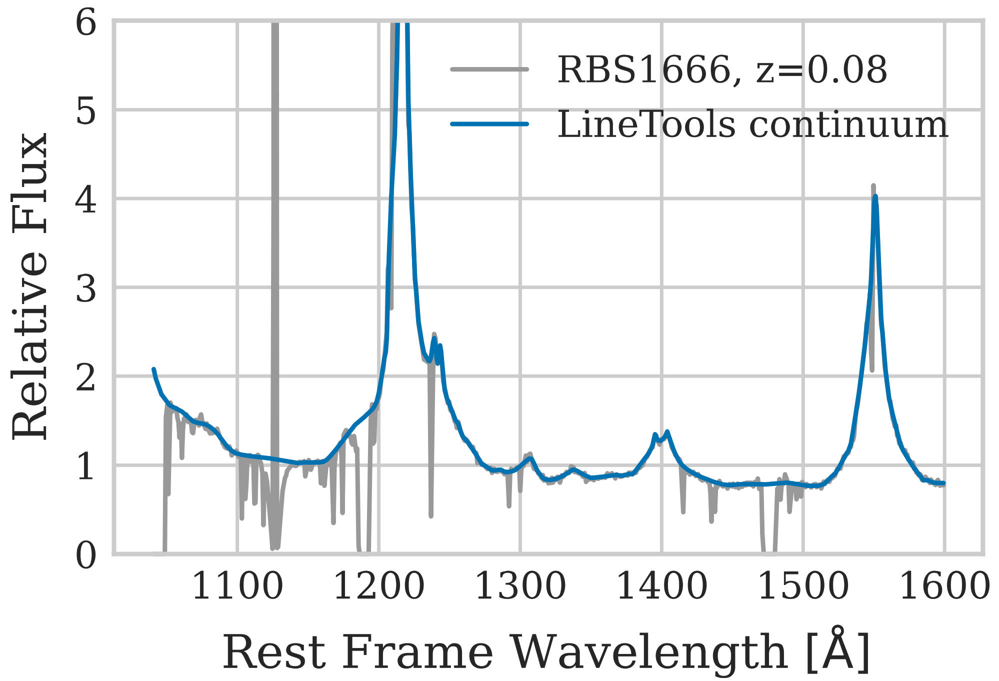
8 million Luminous Red Galaxies ($0.4 < z < 1.1$)

13 million Bright Galaxies ($0.0 < z < 0.4$)



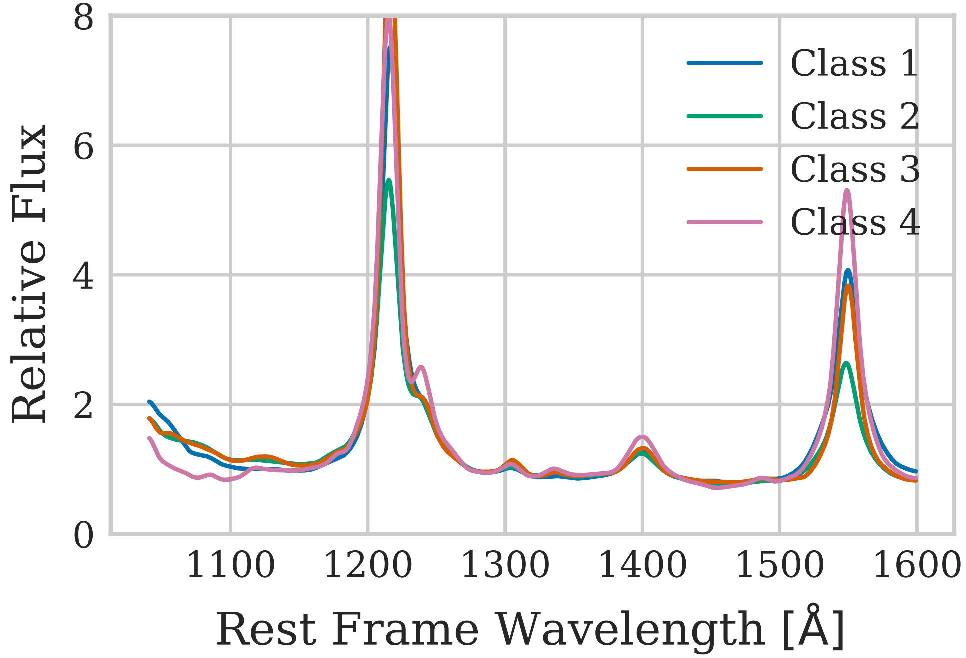
LyCAN training

Cosmic Origins Spectrograph (COS) spectrum



→

Gaussian Mixture Model representation of 38 COS spectra



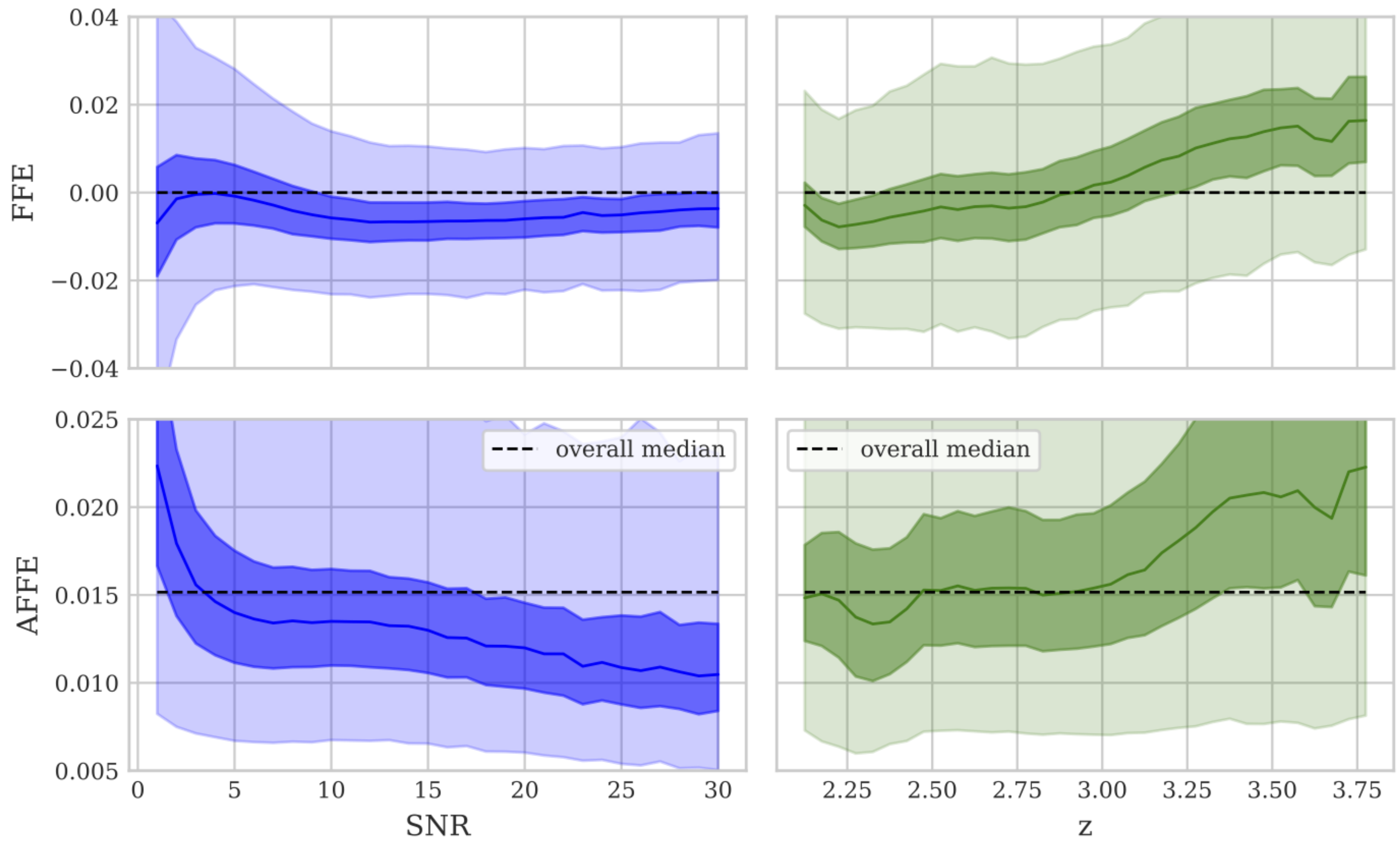
LyCAN architecture

- Convolutional Neural Network (CNN): layers convolve the input pixels with learned filters
- Hyperparameters optimized with pseudo-random search
- LyCAN: 6 hidden layers (3 convolutional, 3 fully-connected)

Layer type	Size	Activation function
Conv1D (input)	64	ReLU
MaxPooling1D	2	
Conv1D	32	ReLU
MaxPooling1D	2	
Conv1D	48	ReLU
MaxPooling1D	2	
Conv1D	16	ReLU
MaxPooling1D	2	
Flatten		
Dense	96	ReLU
Dense	64	ReLU
Dense	128	ReLU
Dense (output)	699	Linear

Bias

LyCAN tends to over-predict the continuum at high-z (~1% in FFE)

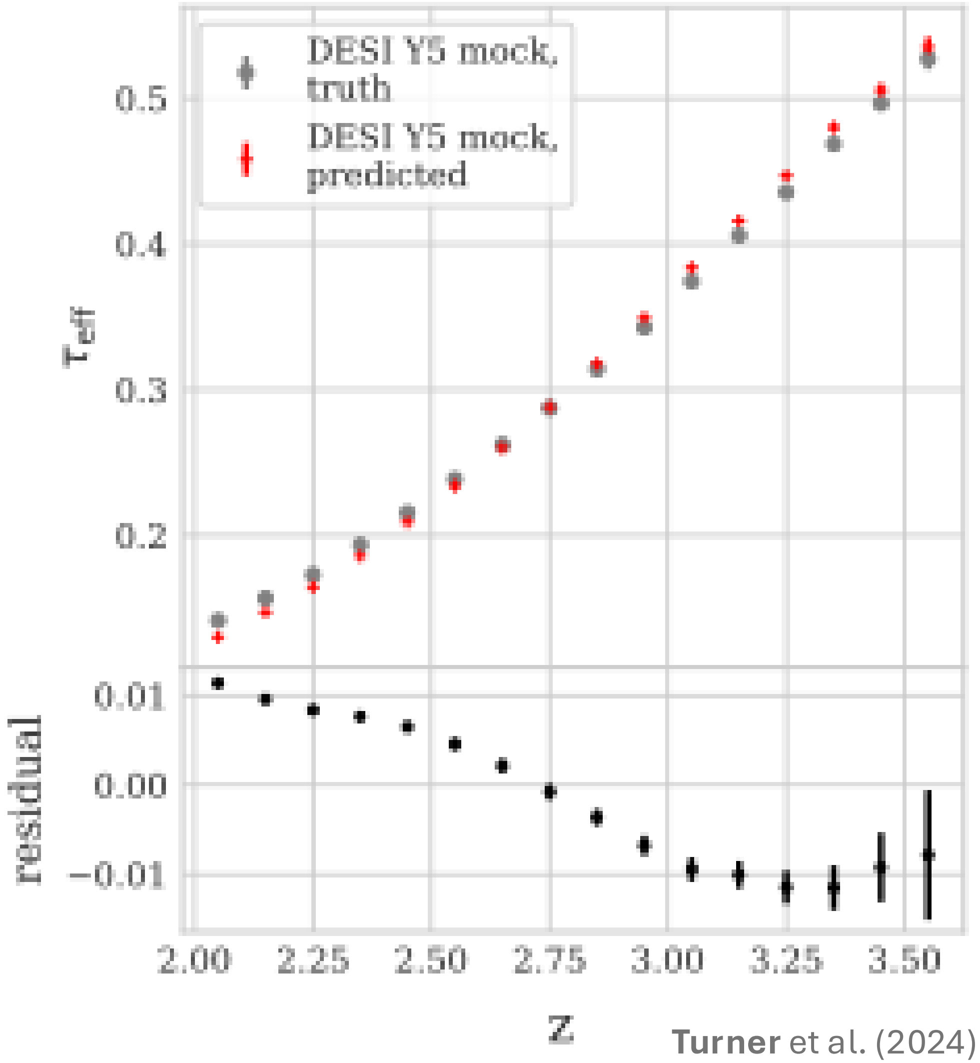


Turner et al. (2024)

Optical depth: measurement and corrections

- Measured transmitted flux fraction for each QSO in redshift bins of $\Delta z = 0.1$
- Covariance matrix computed with 10,000 bootstrap realizations
- Bias correction calibrated with results on mocks
- Corrected for absorption due to metals and optically thick absorbers

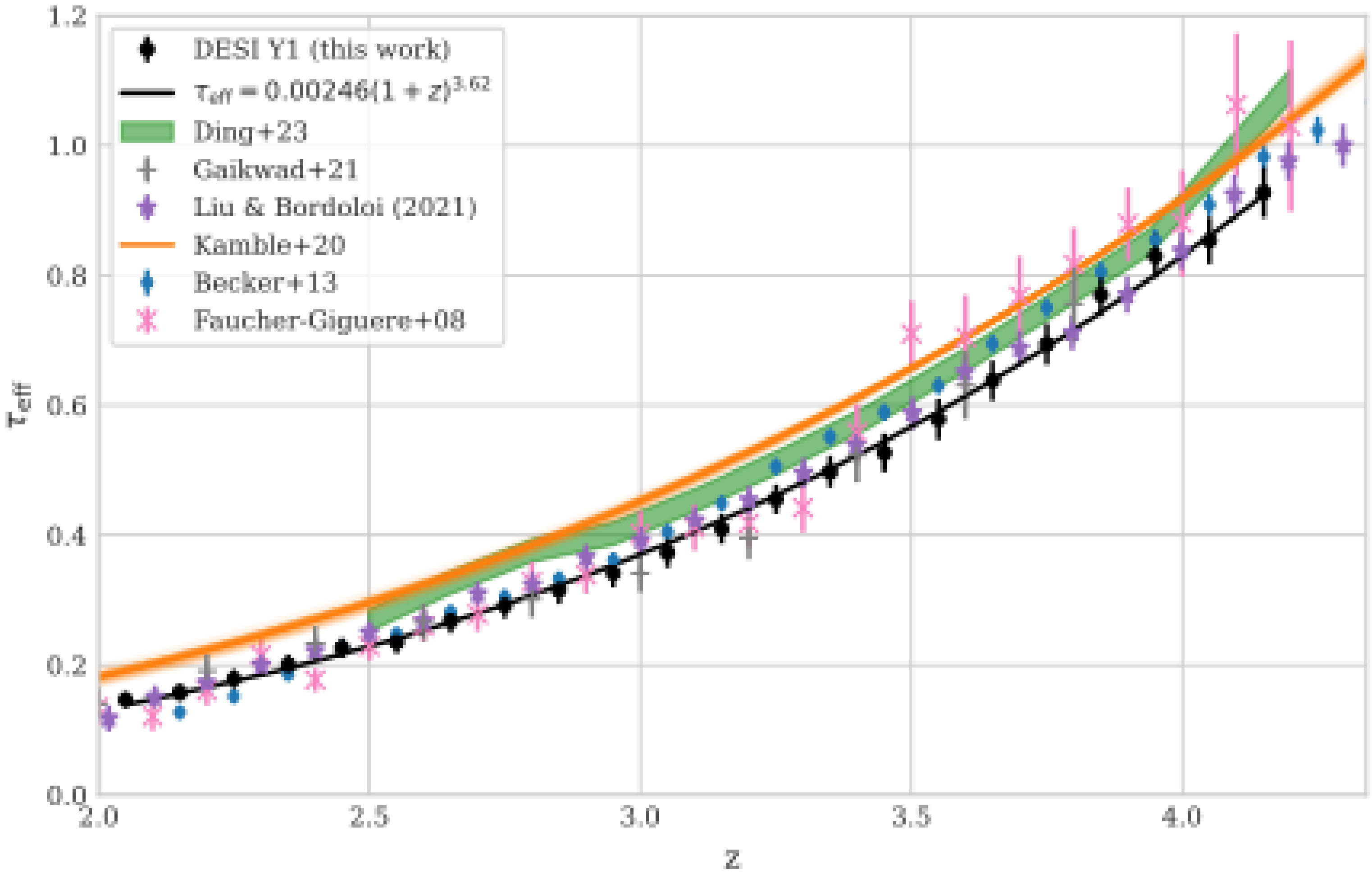
$$\tau = -\ln\langle F\rangle$$



Turner et al. (2024)

Optical depth evolution

- Probes the evolution of the IGM, including possible signatures of He II reionization
- Largest sample to date for measurement of τ_{eff} (~84,000 QSOs, SNR ≥ 5 only)
    - Error bars dominated by metal correction scheme (Schaye+03)
    - Best agreement with low-z measurements based on high-resolution spectra (e.g. Faucher-Giguere+08; Gaikwad+21)



Turner et al. (2024)

Full shape parameters

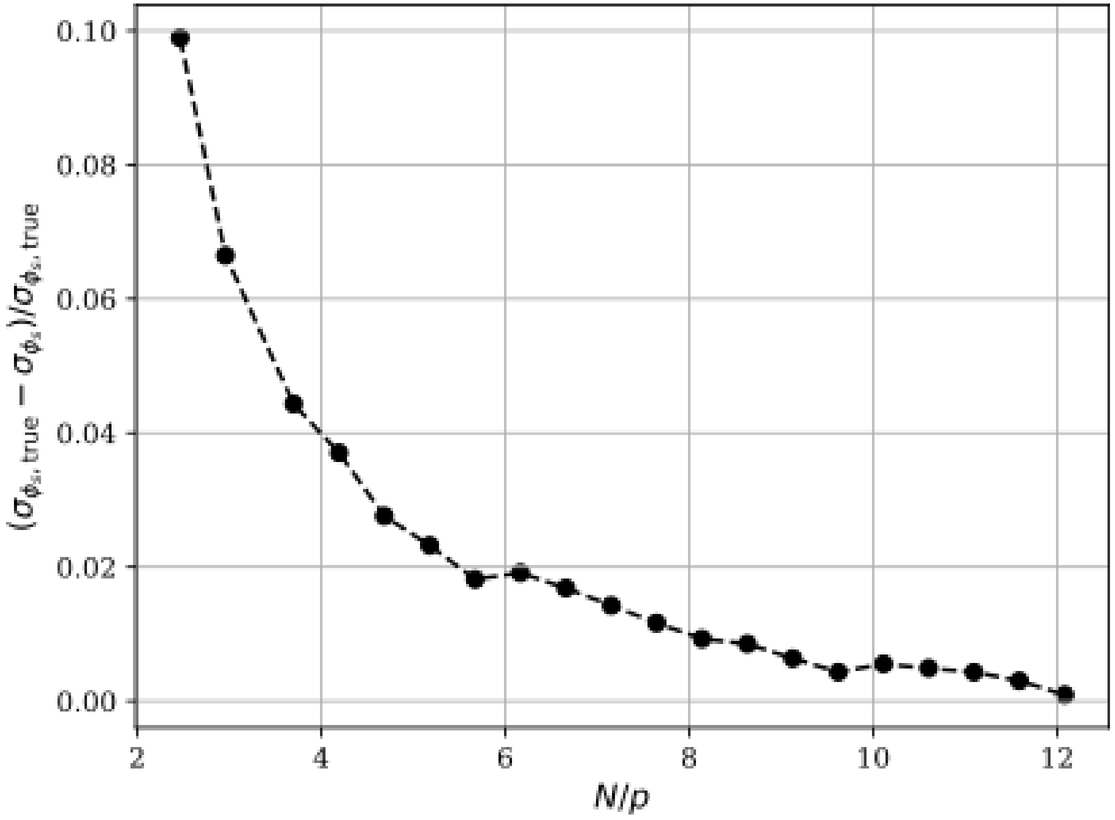
- Angles and redshifts $(\Delta\theta, \Delta z)$ $\xrightarrow{\text{Fiducial cosmology}}$ Comoving distances $(r_\parallel, r_\perp)$
- Model: Template P(k) at fiducial cosmology
- Rescale coordinates: $r'_\parallel = q_\parallel r_\parallel$ and $r'_\perp = q_\perp r_\perp$
- Parameterize anisotropy: $\phi = \frac{q_\perp}{q_\parallel}$ and $\alpha = \sqrt{q_\perp q_\parallel}$
- Measurements of ϕ and α correspond to:

$$\phi = \frac{D_M(z)H(z)}{[D_M(z)H(z)]_{\mathrm{fid}}} \qquad \alpha = \sqrt{\frac{D_M(z)D_H(z)/r_d^2}{[D_M(z)D_H(z)/r_d^2]_{\mathrm{fid}}}}$$

Covariance estimation

- Measured covariance matrices from mocks with the true continuum vs continuum fitting
  - Applied and validated the Hartlap correction to ensure robustness of the covariance matrix
- Hartlap correction: accounts for bias of inverse covariance when C is estimated from a limited number of samples
  - Multiplicative factor on C^{-1}
- Also investigated impact of extending analysis to larger scales
- Forecasts assuming Planck 2018 fiducial model

Error on the Hartlap correction



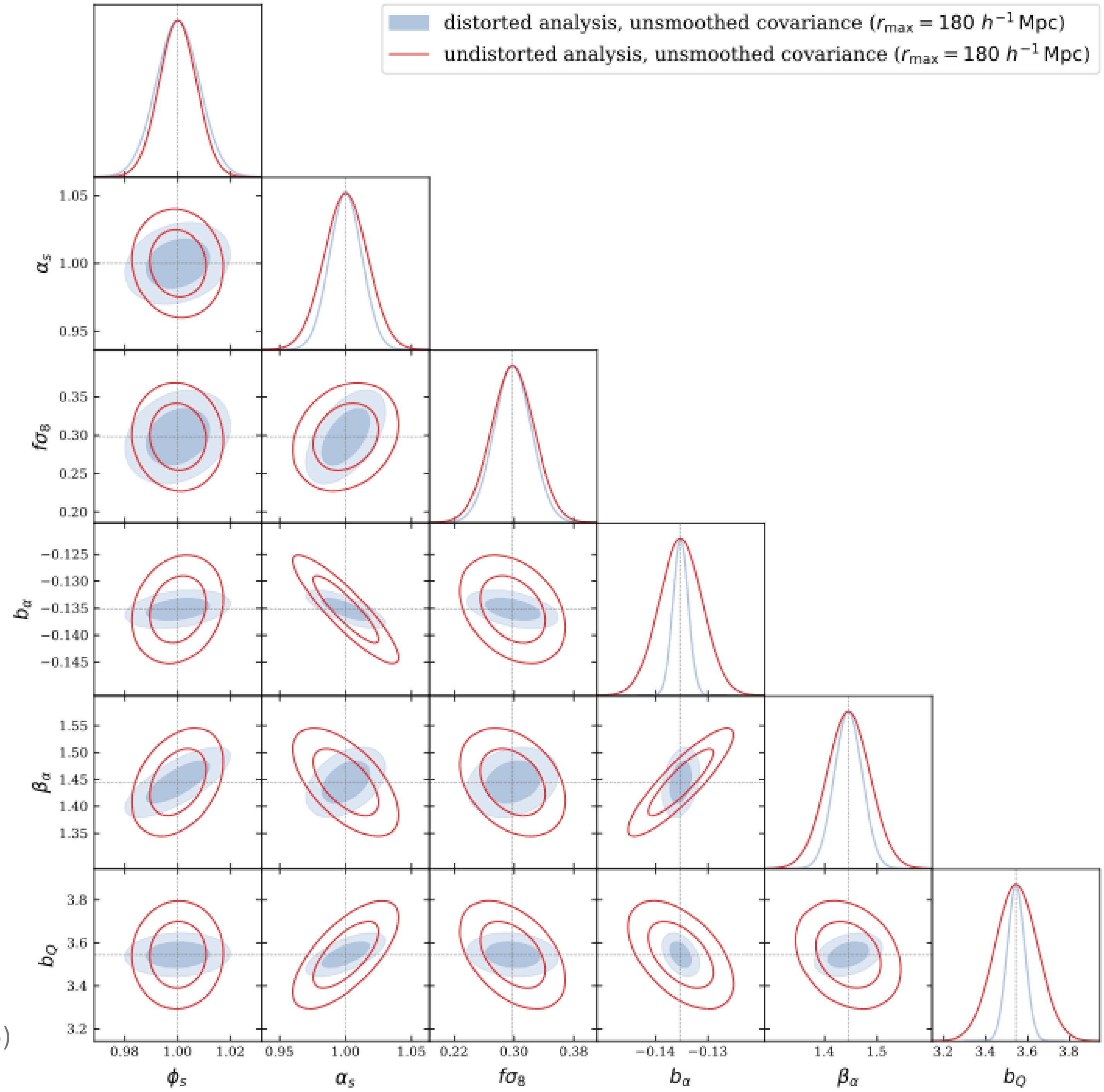
N: number of samples used for covariance measurement
p: length of data vector

Turner et al. (2026)

Other parameters

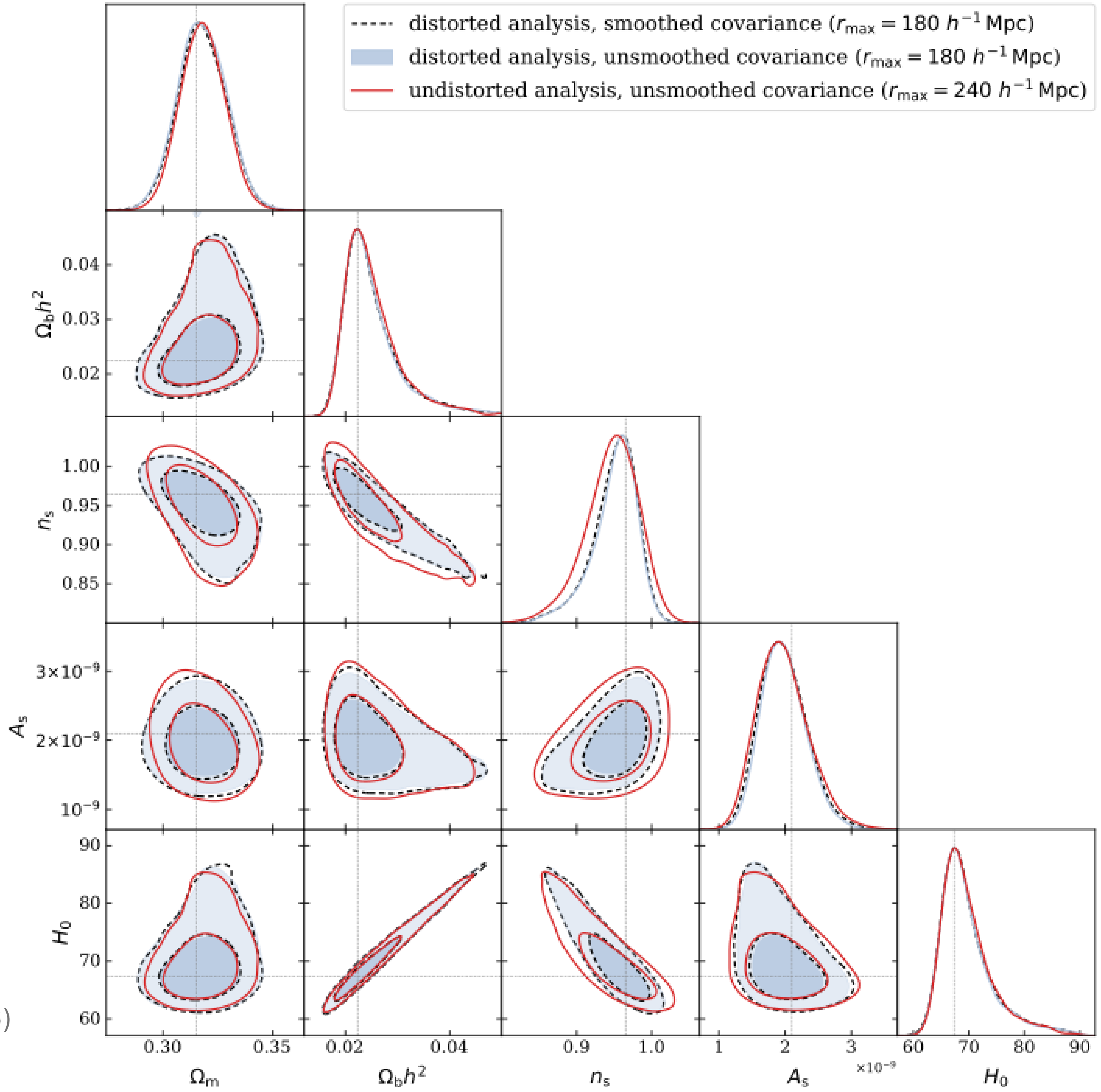
- Improvement seen in ϕ_s, ϕ_p (not shown)
- $f\sigma_8$ and bias parameters degrade
- Caveats:
  - Sub-optimal estimator!
  - $f\sigma_8$ is not well-constrained by the Lyα forest

Turner et al. (2026)



Direct cosmological inference

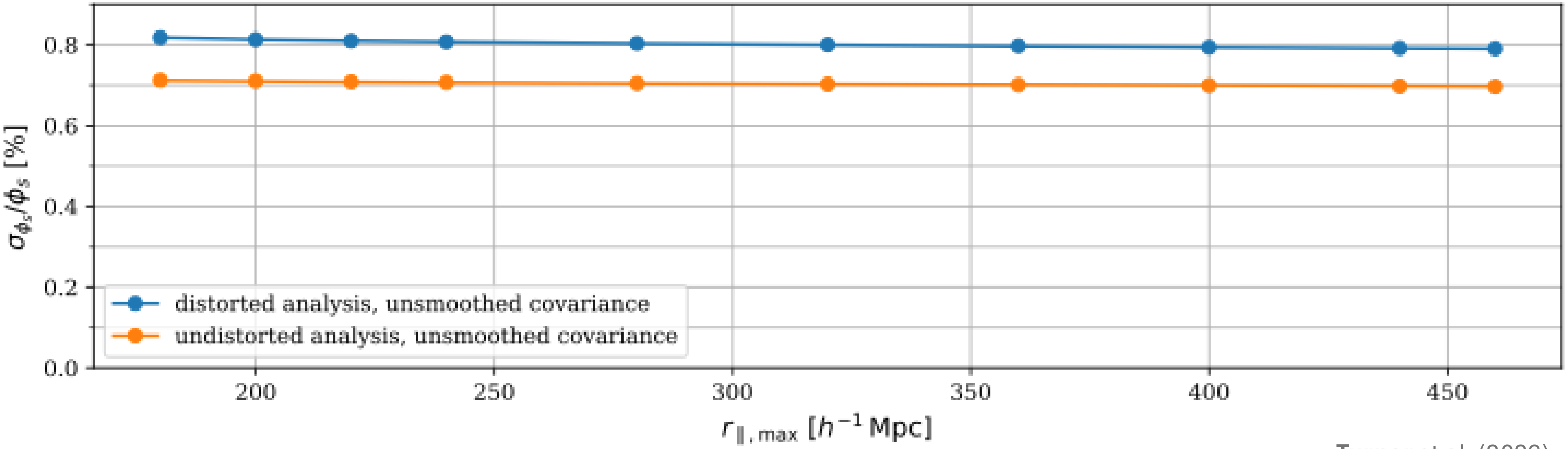
- Direct fit to full power spectrum instead of template fitting approach
- Improvement on AP translates to improvement on $\Omega_{\rm m}$
- Parameters with degradation are not well-constrained by the forest



Turner et al. (2026)

Impact of larger scales

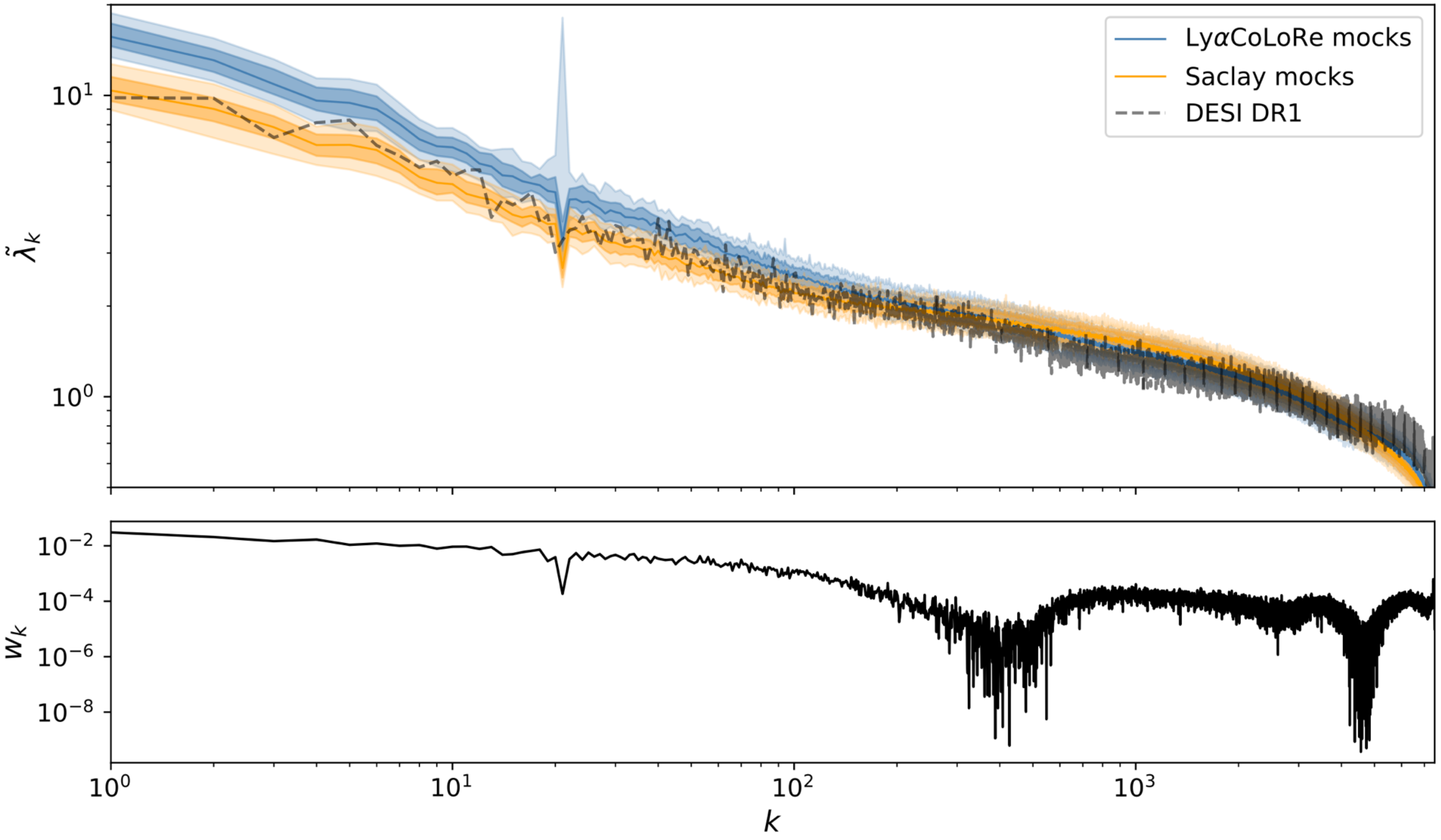
(Along the line of sight)



Turner et al. (2026)

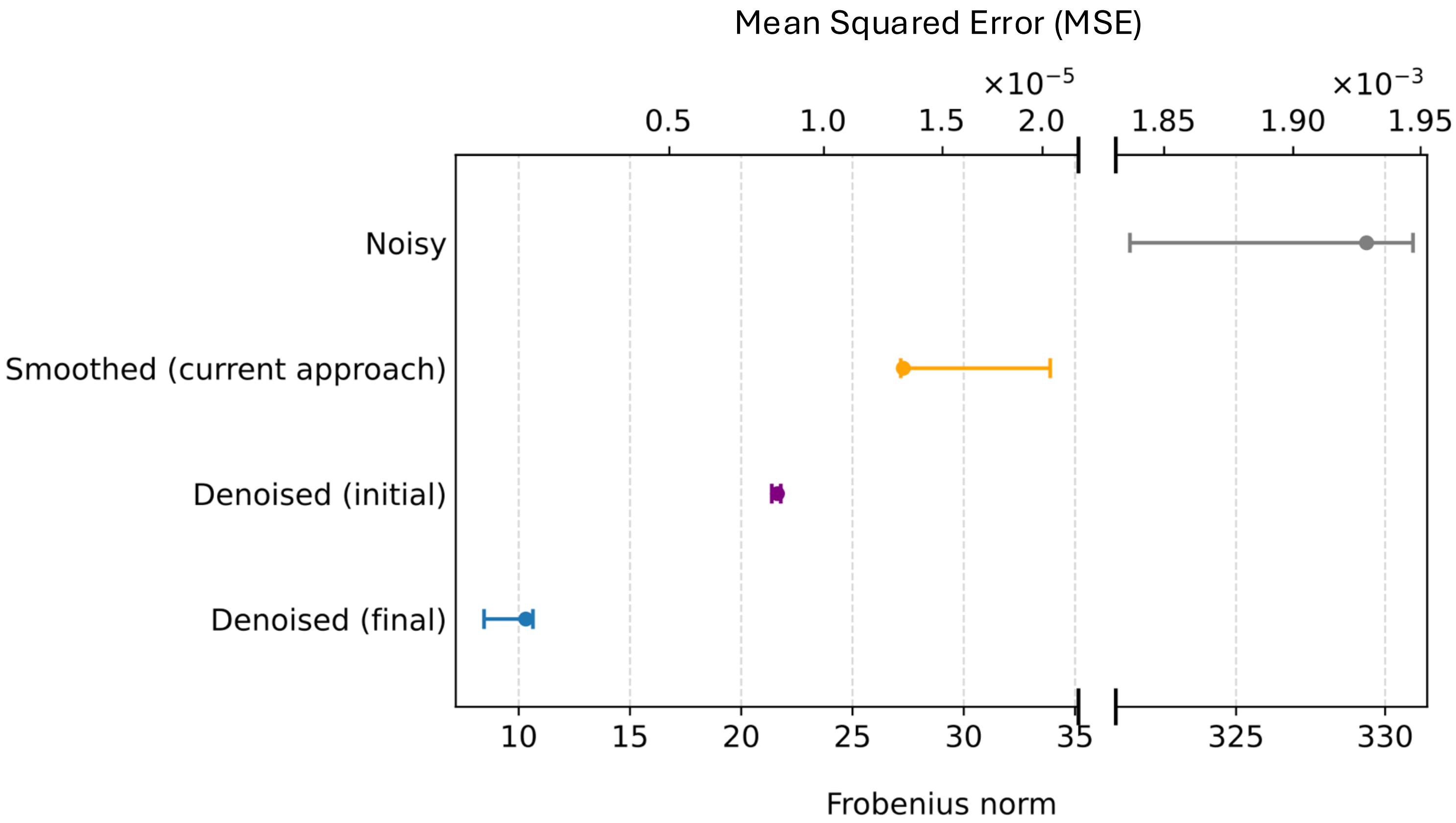
Some additional information available, but minimal

Covariance denoising: pseudo-eigenvalues



Turner (2026)

Matrix-level Reconstruction Metrics



Factor of ~30 (1000) improvement in median Frobenius norm (MSE)

Turner (2026)