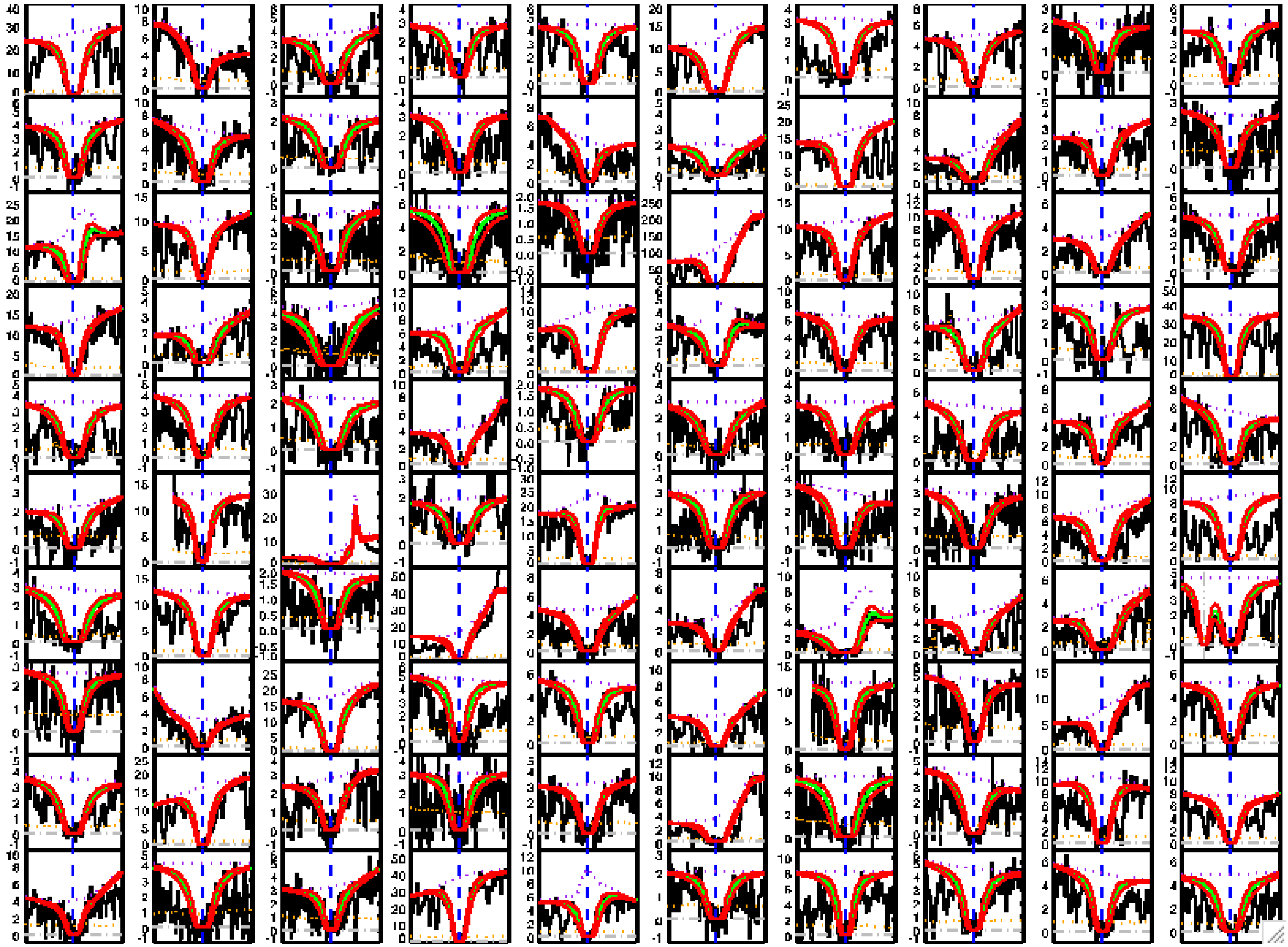
Deep Learning of Quasar Spectra

David Parks
J. Xavier Prochaska
S. Dong
Z. Cai

(UC Santa Cruz)



https://github.com/davidparks21/qso_lya_detection_pipeline

Arxiv astro-ph/1709.04962

Quasar Spectroscopy (Early Days)

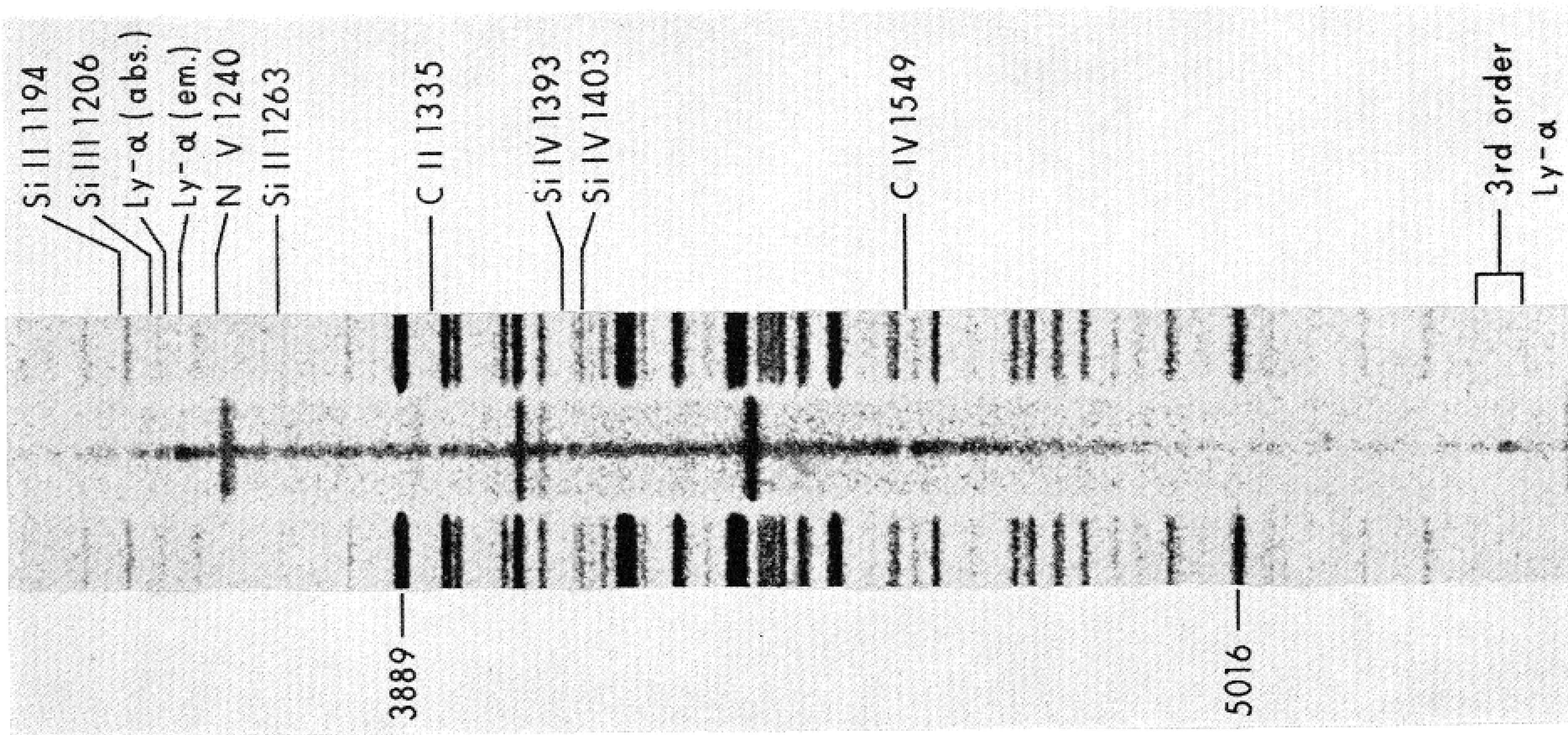


Fig. 2.—Lick spectrum of 3C 191 obtained in February, 1966, with the prime-focus spectrograph on the 120-inch telescope. The comparison spectrum shown is that of He + Ar.

Quasar Spectroscopy (Early Days)

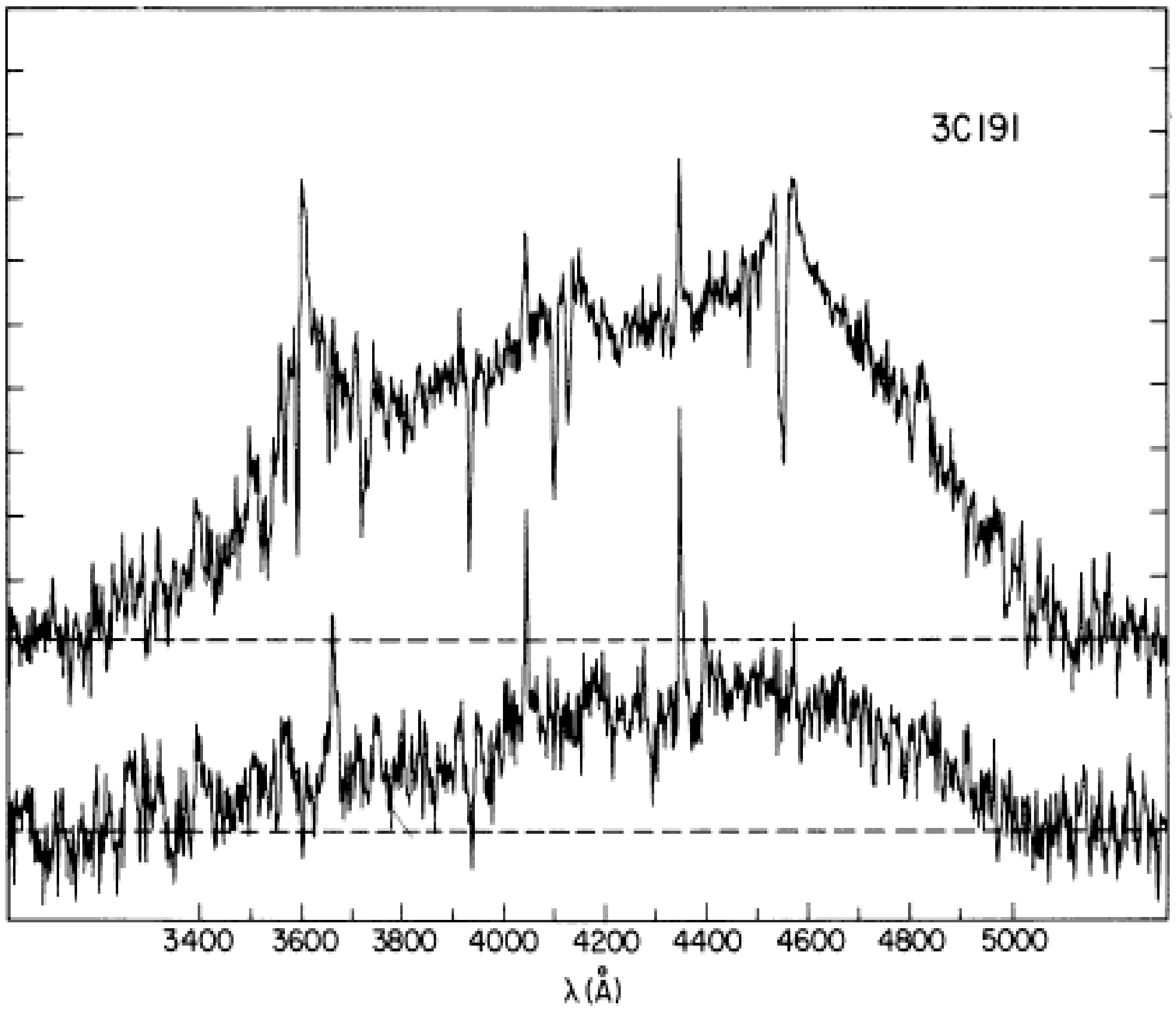
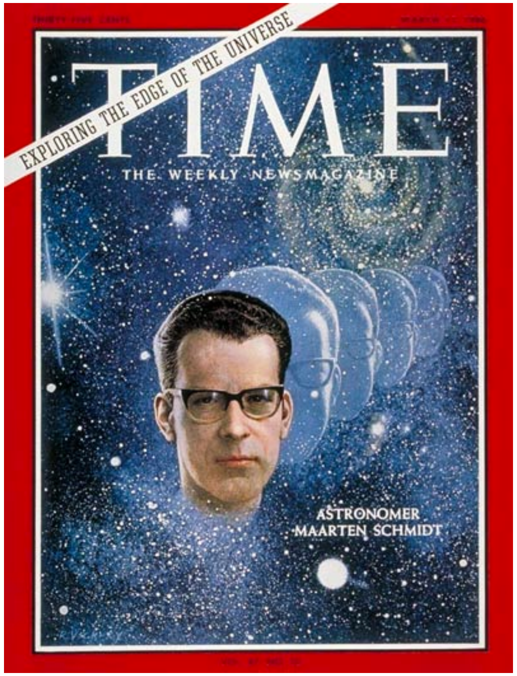
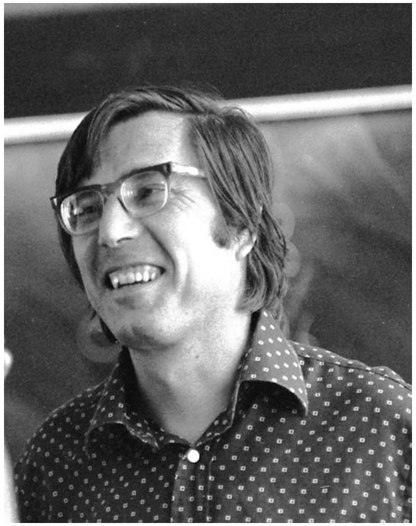


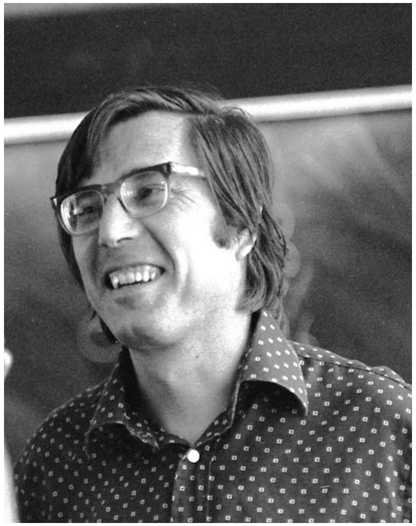
Fig. 1.—Density tracings of the spectrum of 3C 191 and the nearby night-sky spectrum on the same plate are shown. The two tracings have the same vertical magnification but have been shifted by an arbitrary amount. The strong emission lines in the night-sky spectrum are due to mercury city lights.



Quasar Spectroscopy (Disco Days)



Quasar Spectroscopy (Disco Days)



Quasar Spectroscopy (Disco Days)

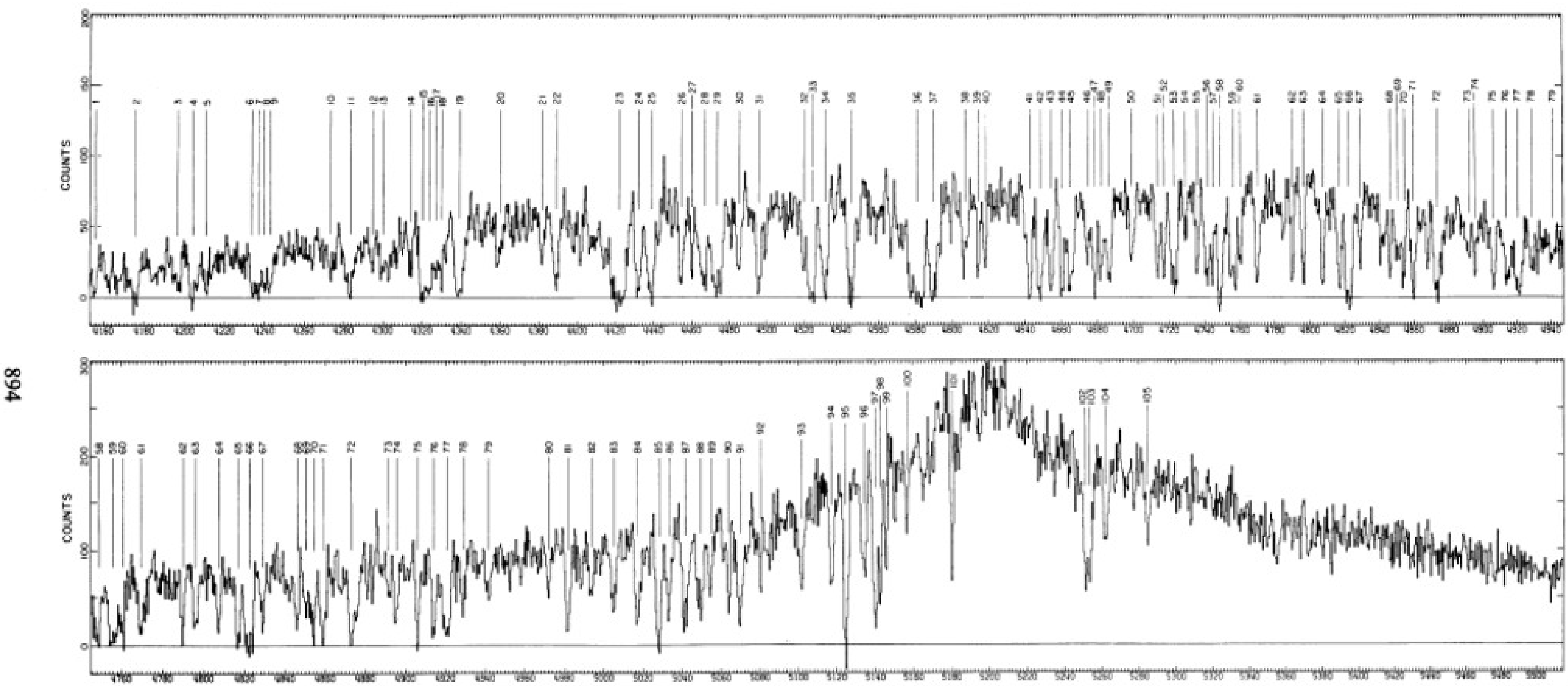
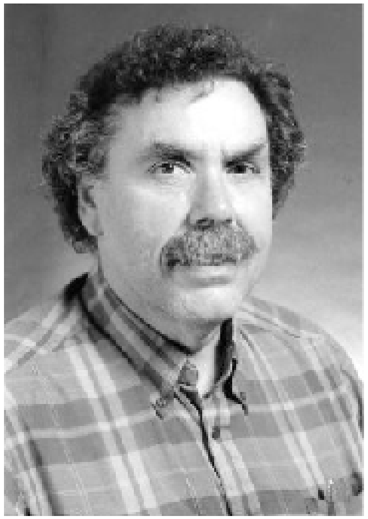
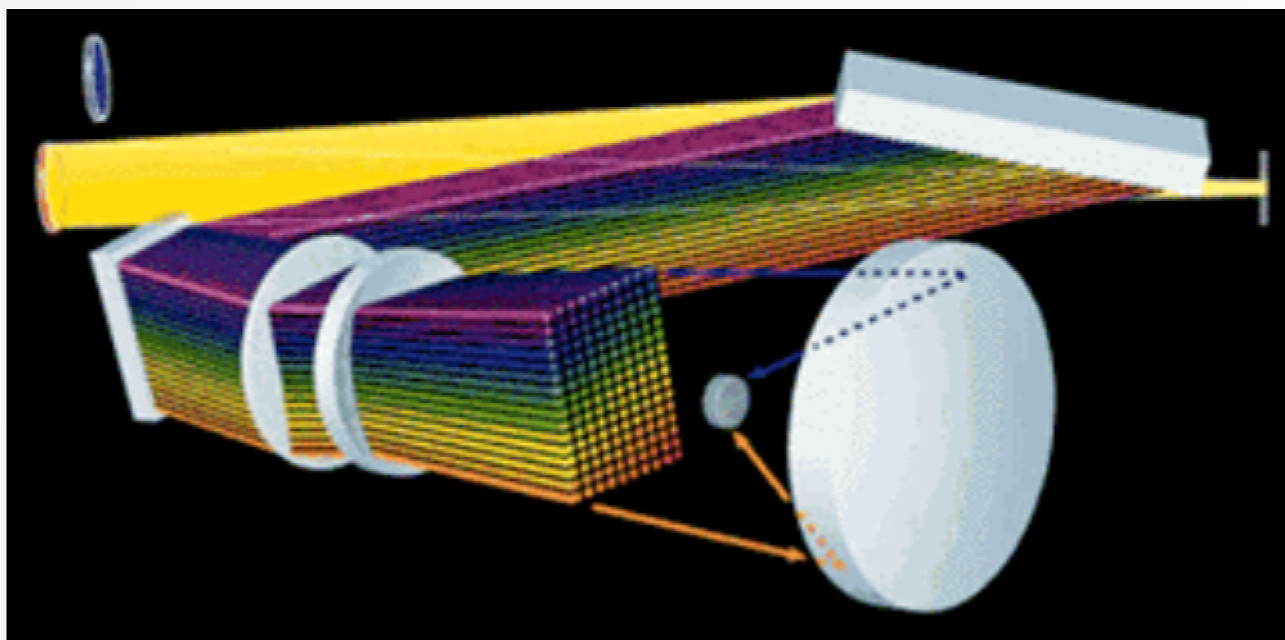


FIG. 1.—Spectrum of PKS 2126−158, showing four overlapping, independent observations covering the range 4153–6807 Å. Each bin is equivalent to 25.78 km s^{-1} (the wavelength axis is logarithmic). The zero intensity level in each observation is indicated by the horizontal lines. The 113 absorption lines listed in Table 3 are marked and labeled.

Quasar Spectroscopy (Keck/HIRES)

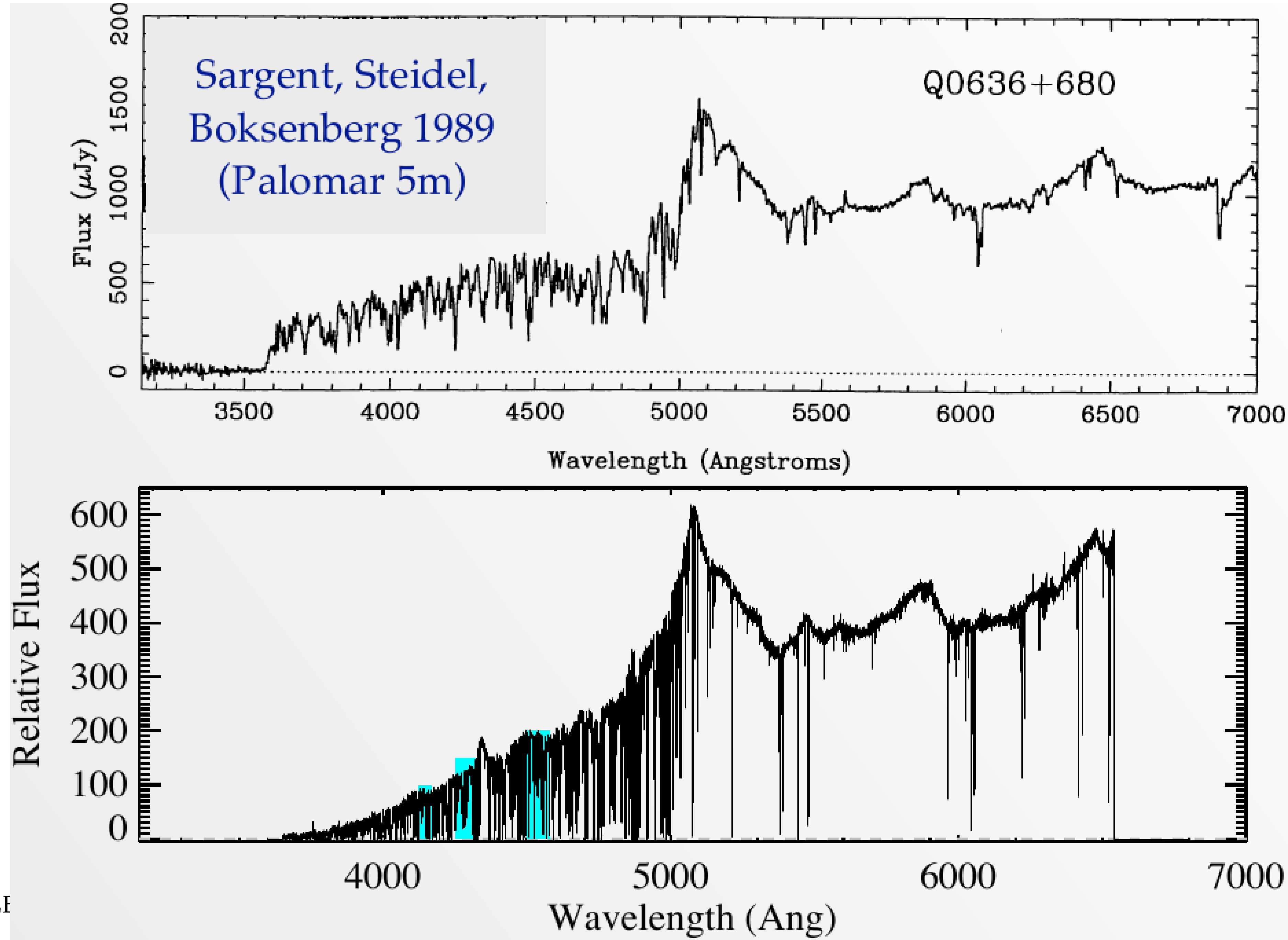




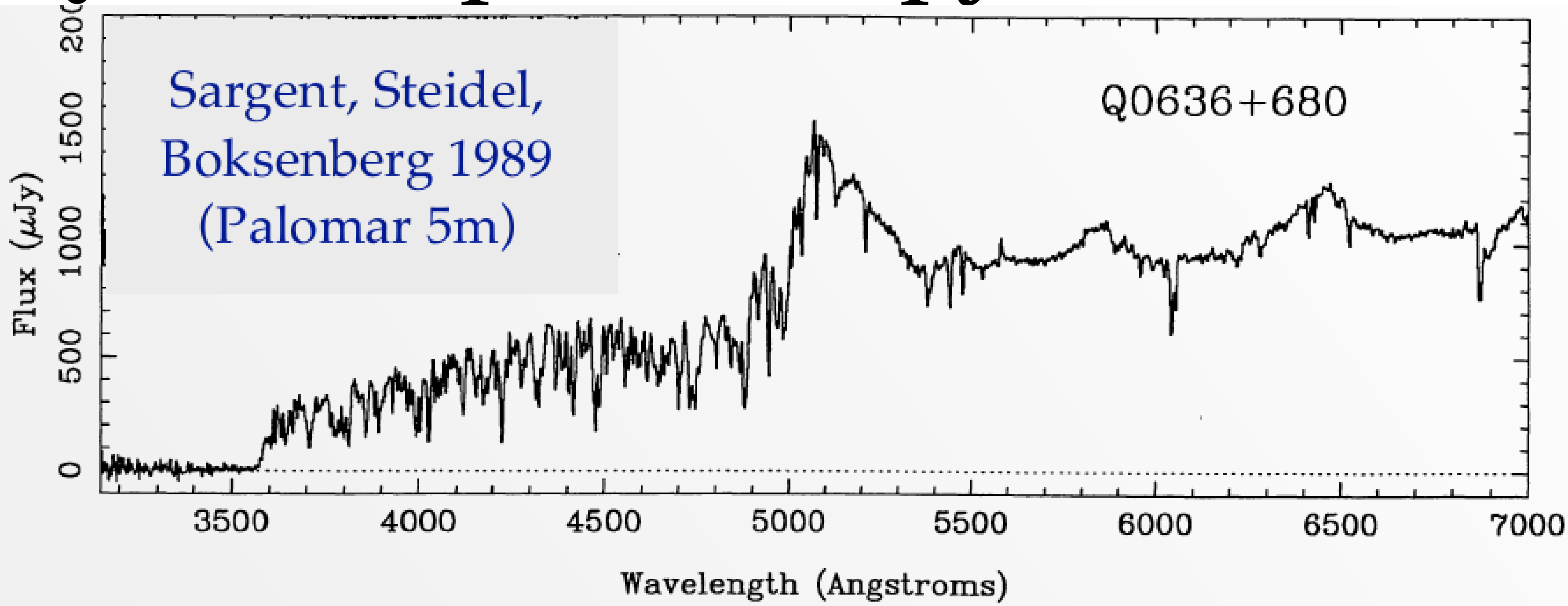
Steve Vogt (UCO)

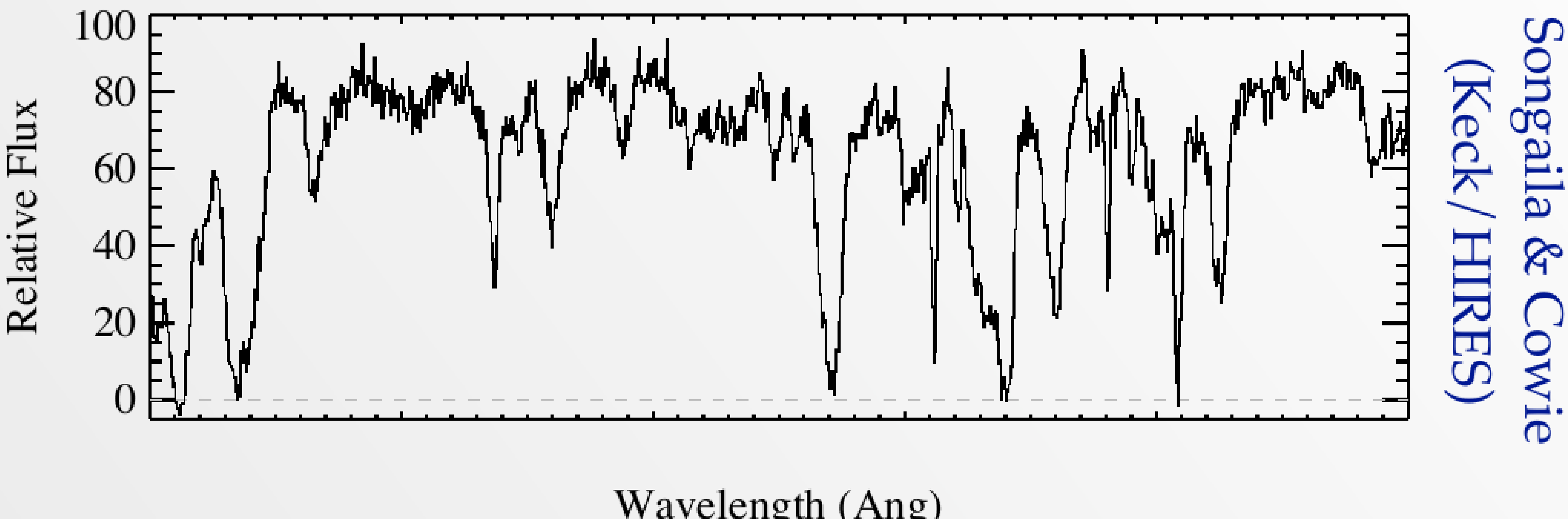
HIRES

Quasar Spectroscopy (Keck/HIRES)

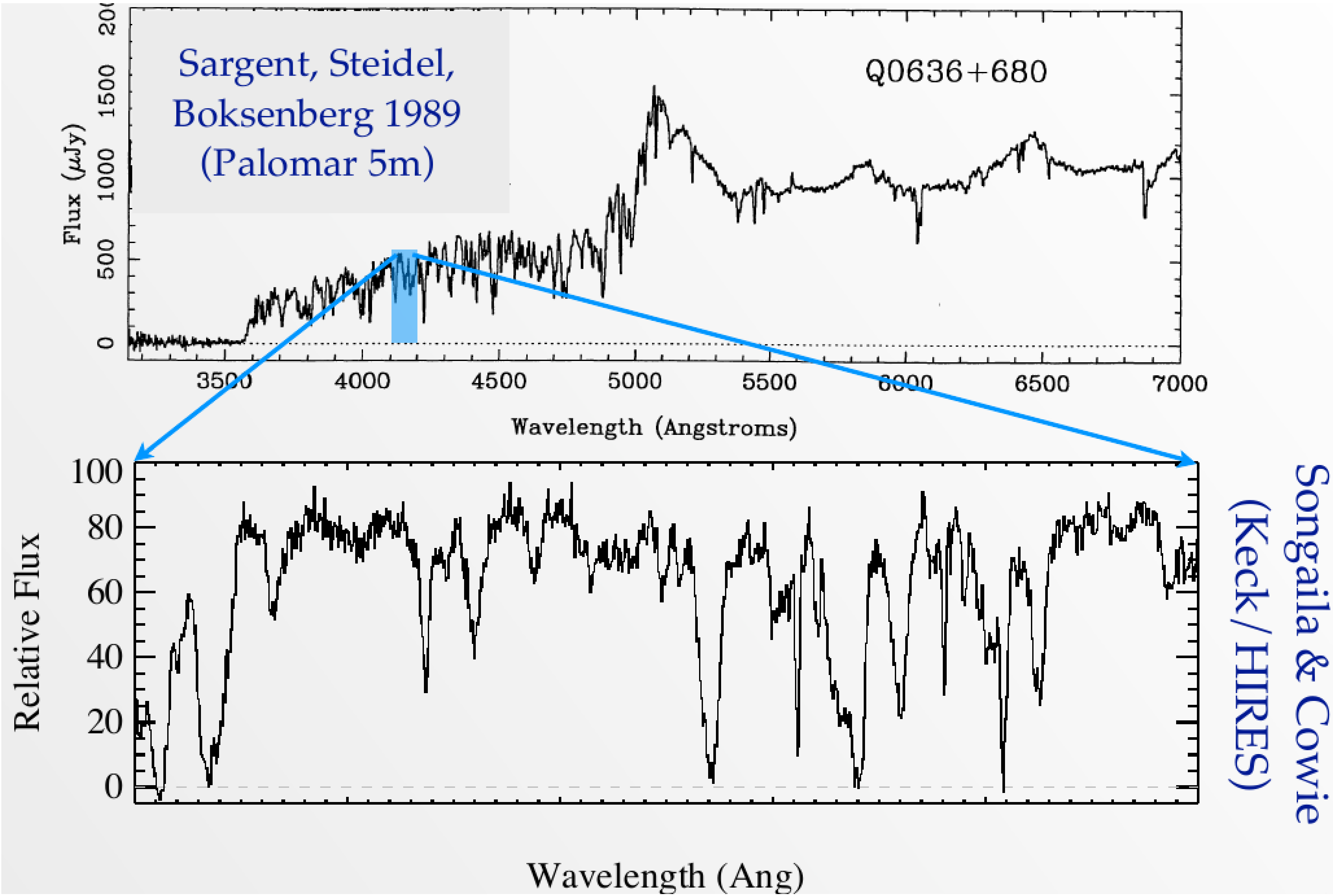


Quasar Spectroscopy (Keck/HIRES)

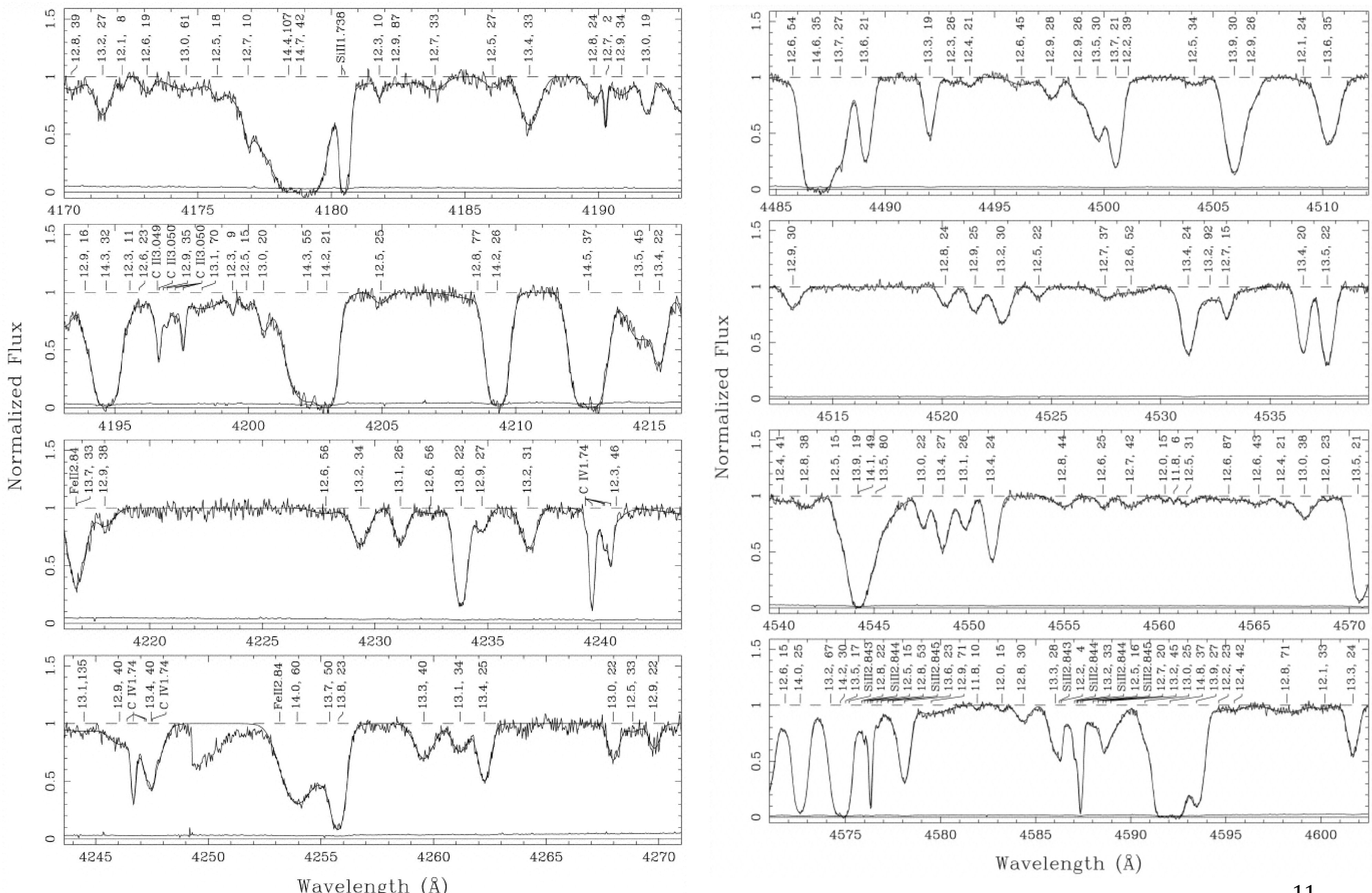




Quasar Spectroscopy (Keck/HIRES)

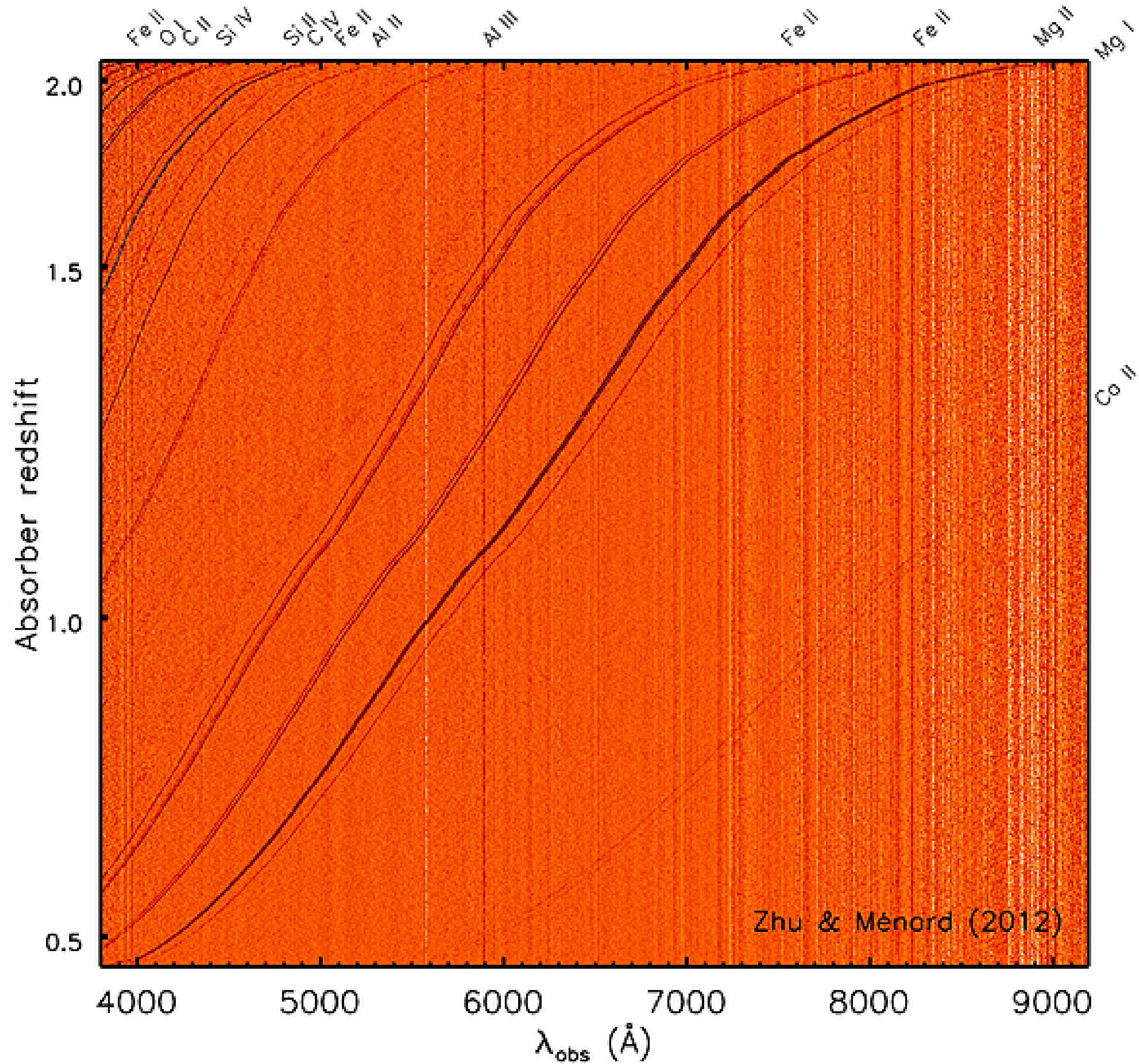


Quasar Spectroscopy (Keck/HIRES)



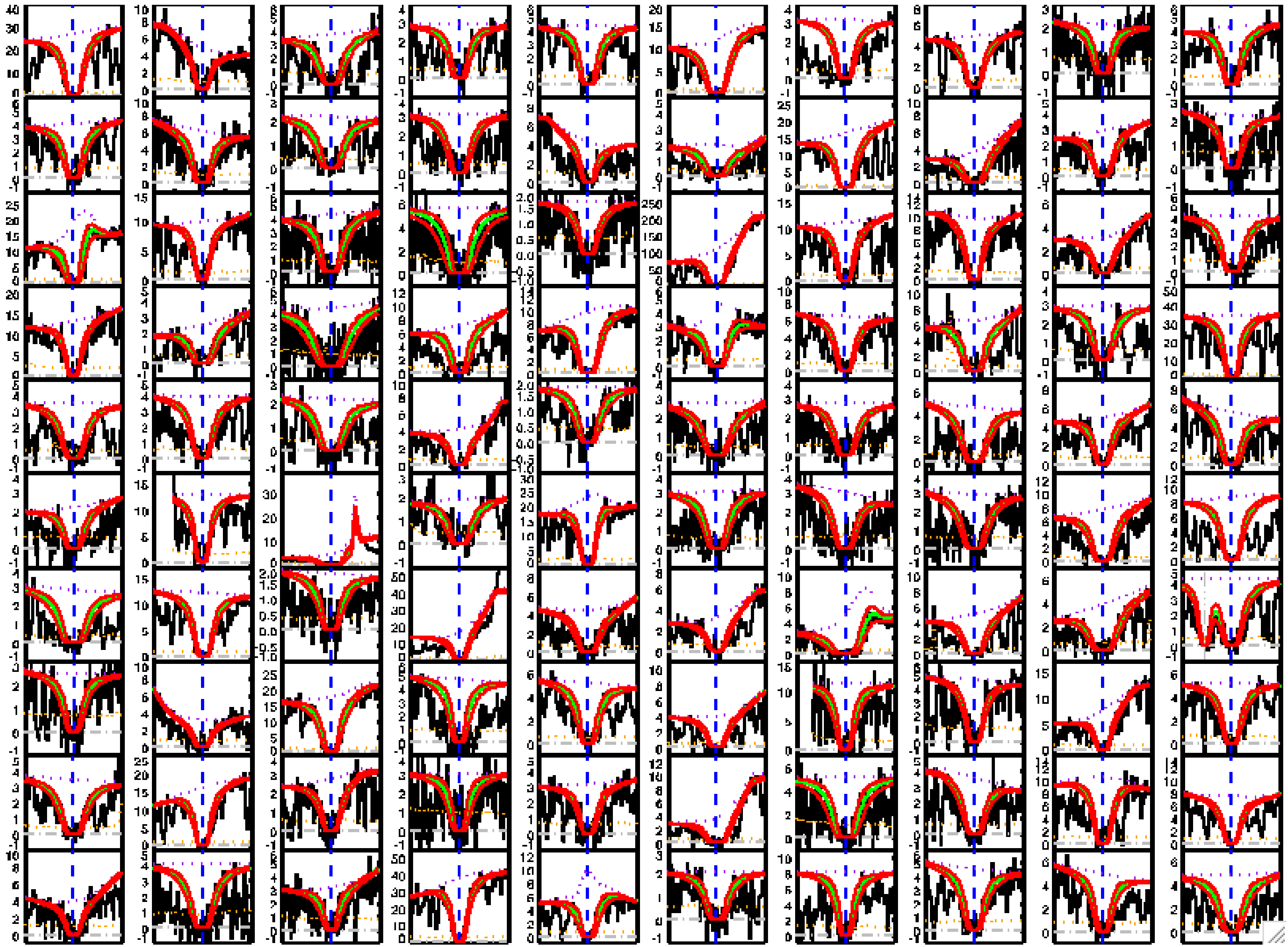
Kirkman&Tytler 97

Quasar Spectroscopy (SDSS)



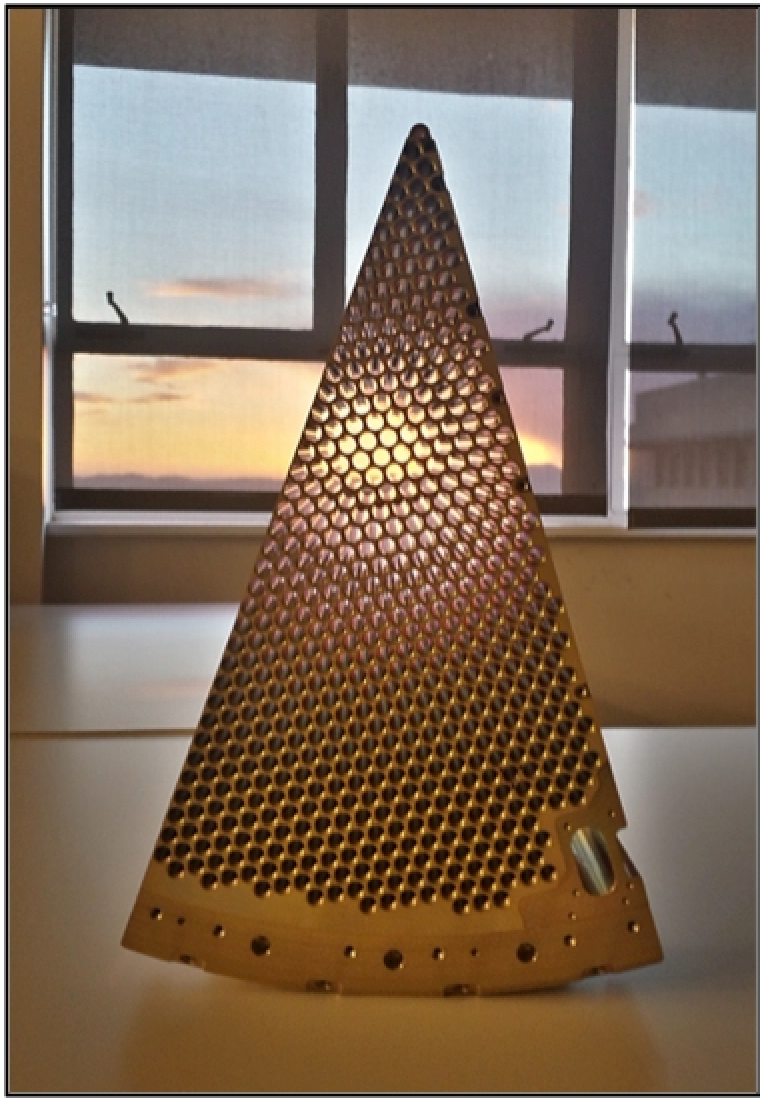
Zhu & Menard 2012

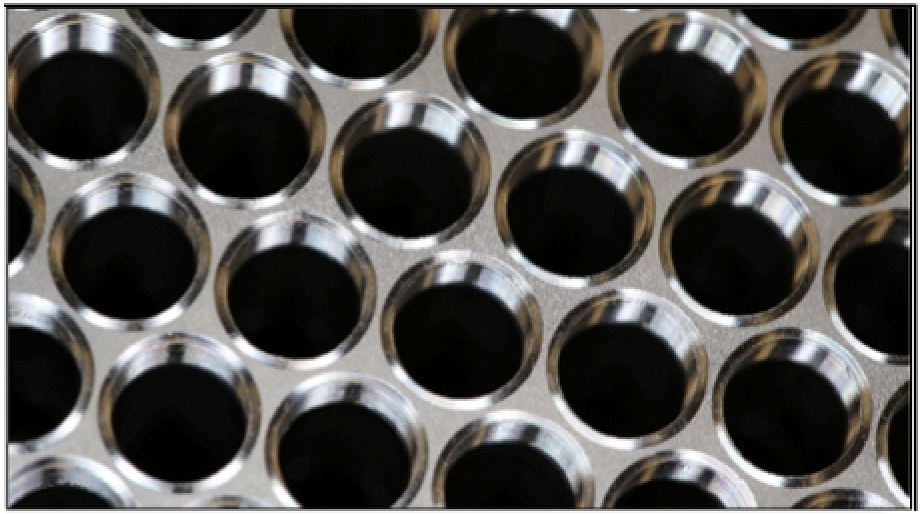
Quasar Spectroscopy (SDSS)

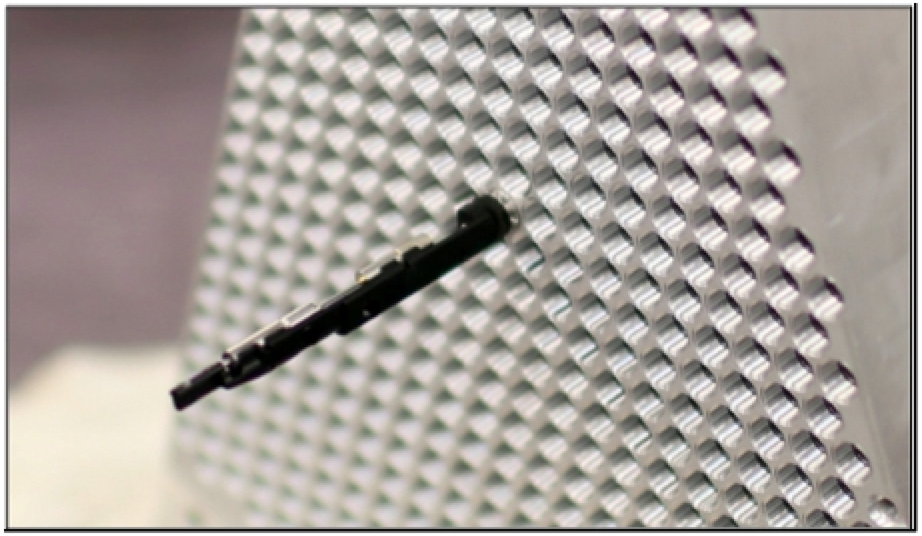


PWH 2005

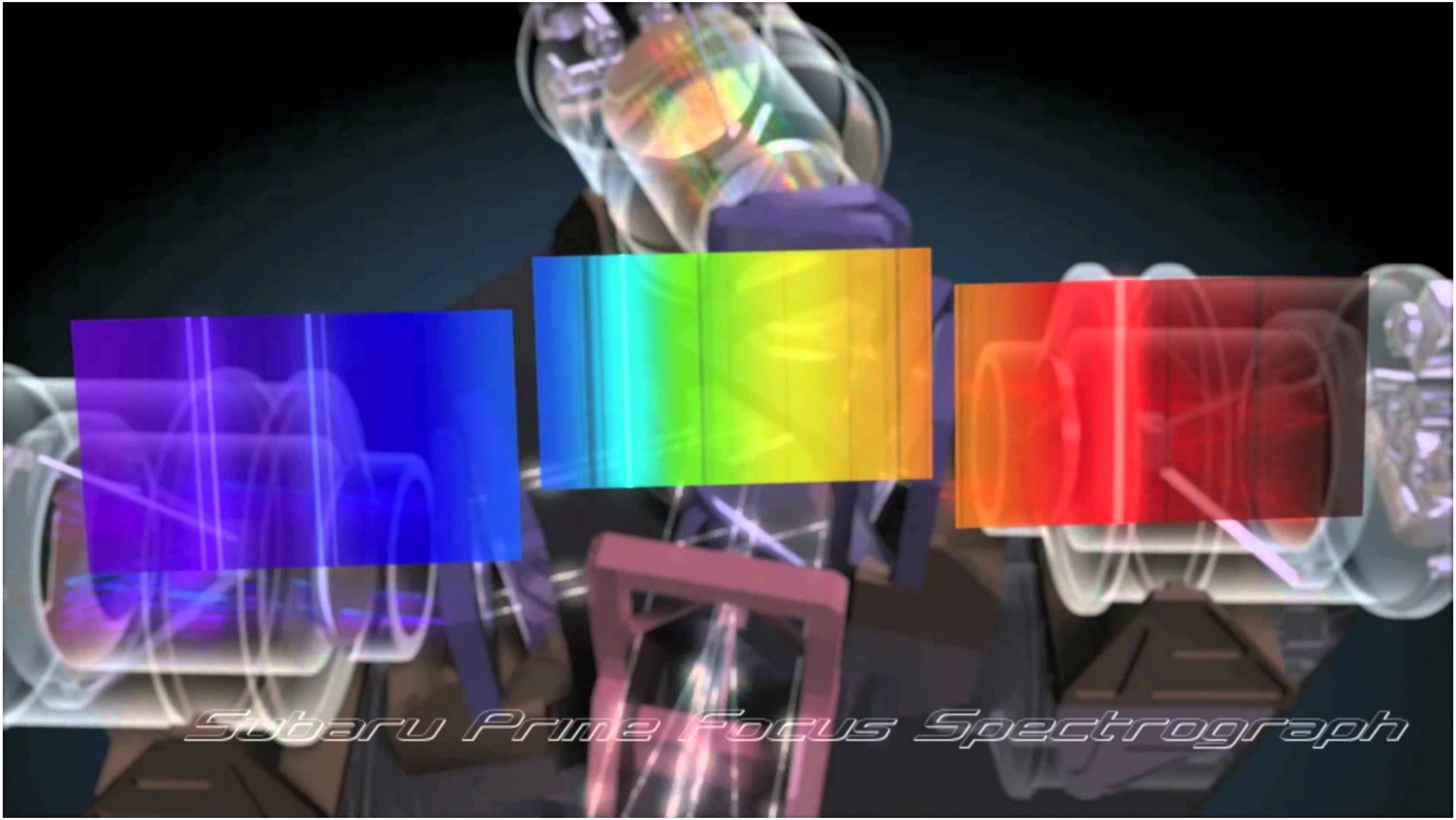
Quasar Spectroscopy (Future)





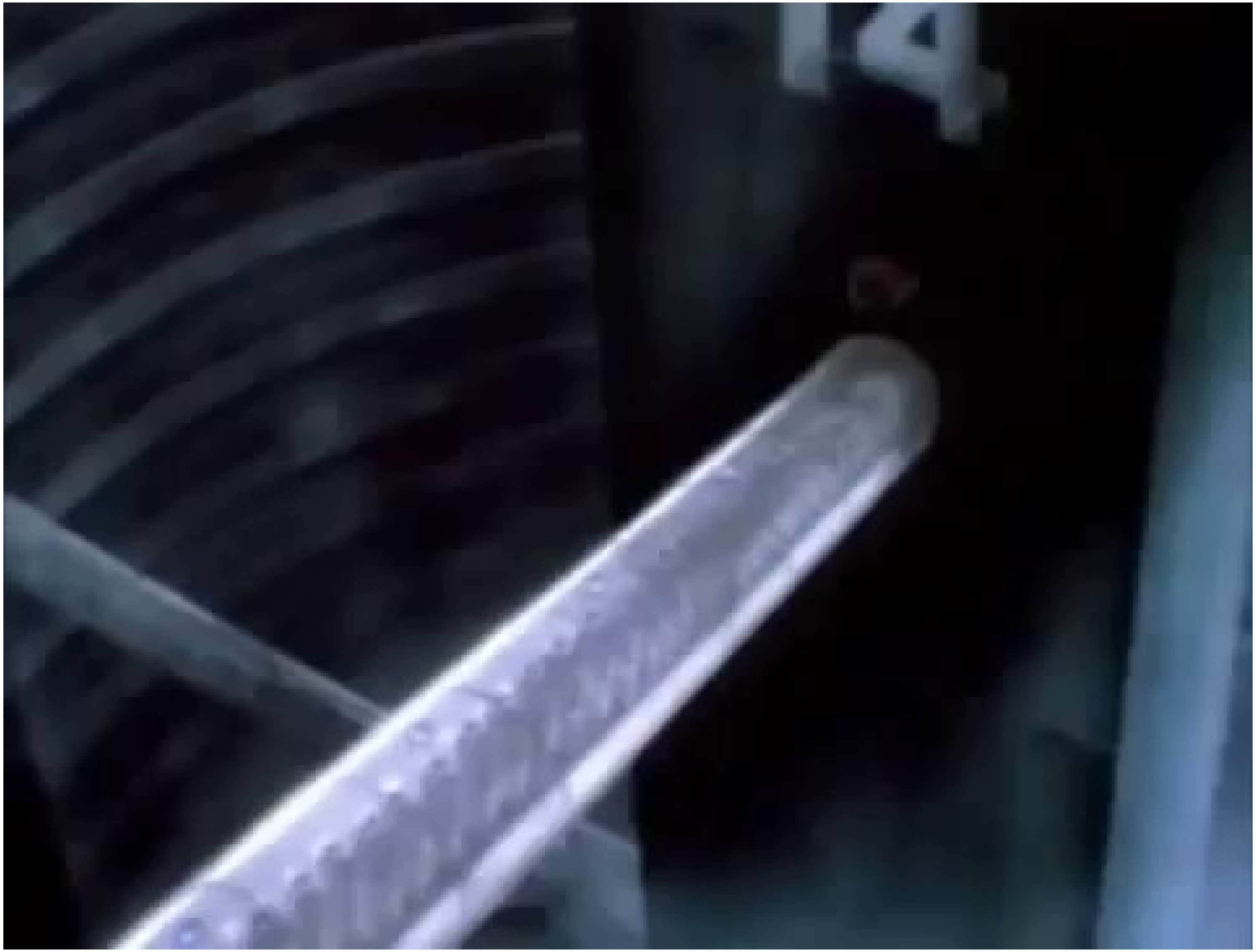


Quasar Spectroscopy (Future)

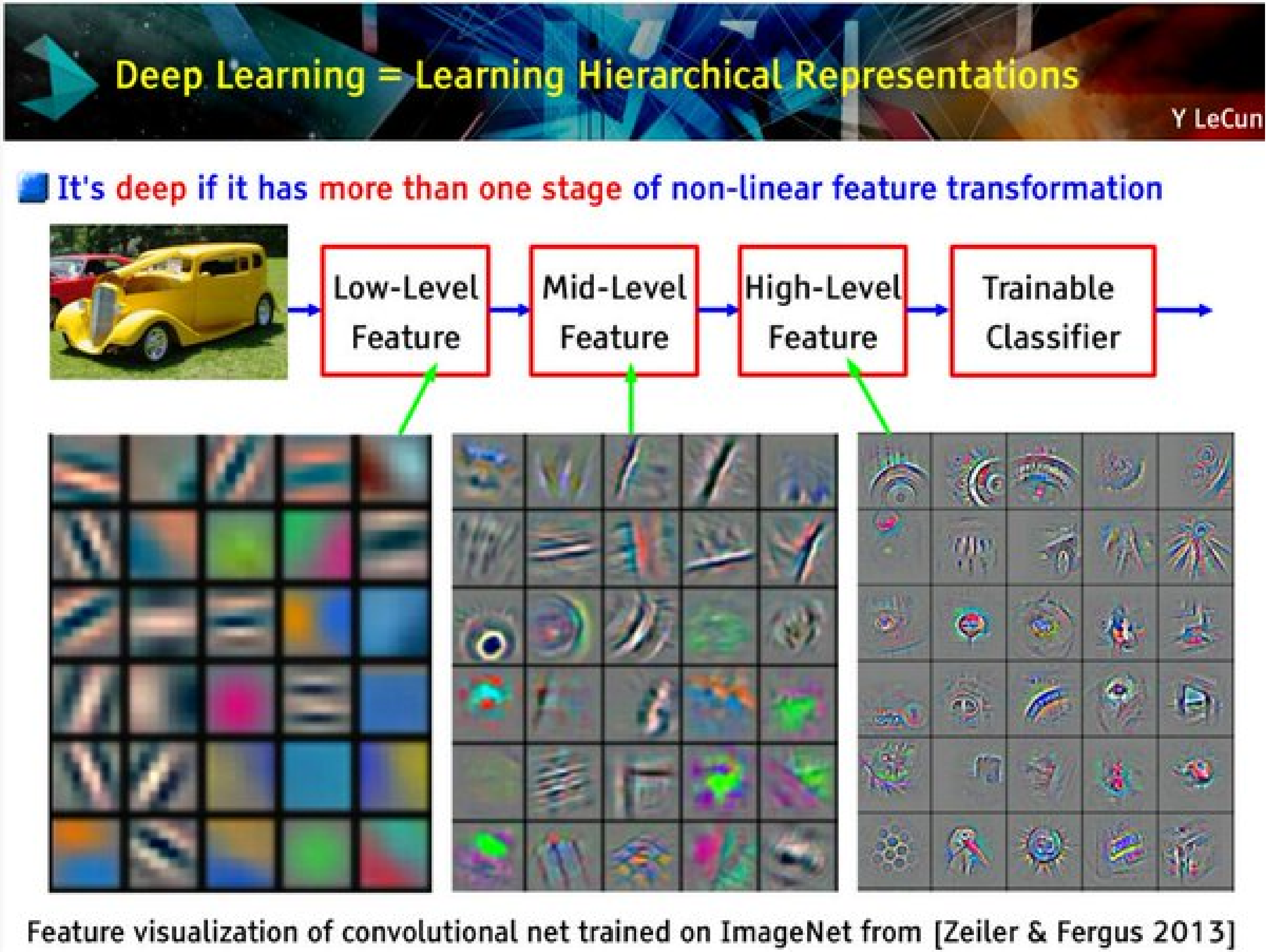


Machine Learning

Machine Learning

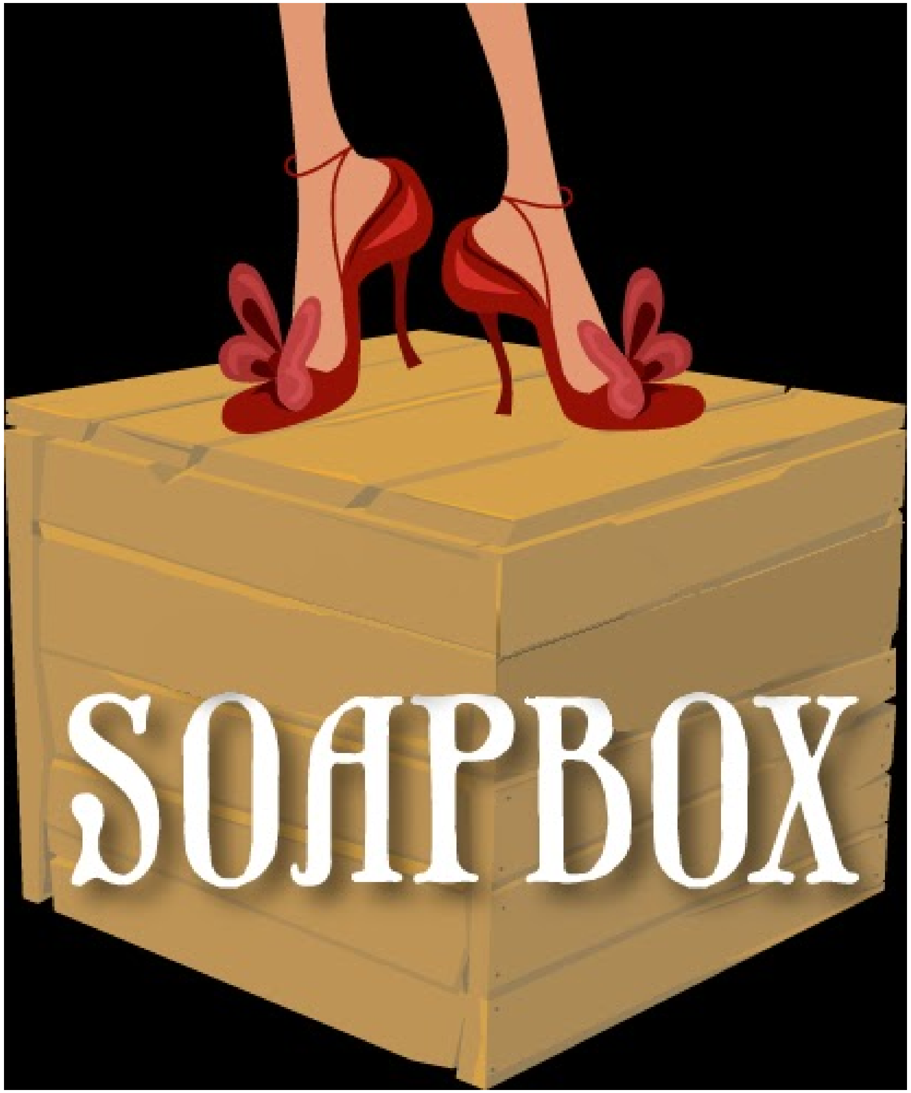


Deep Learning



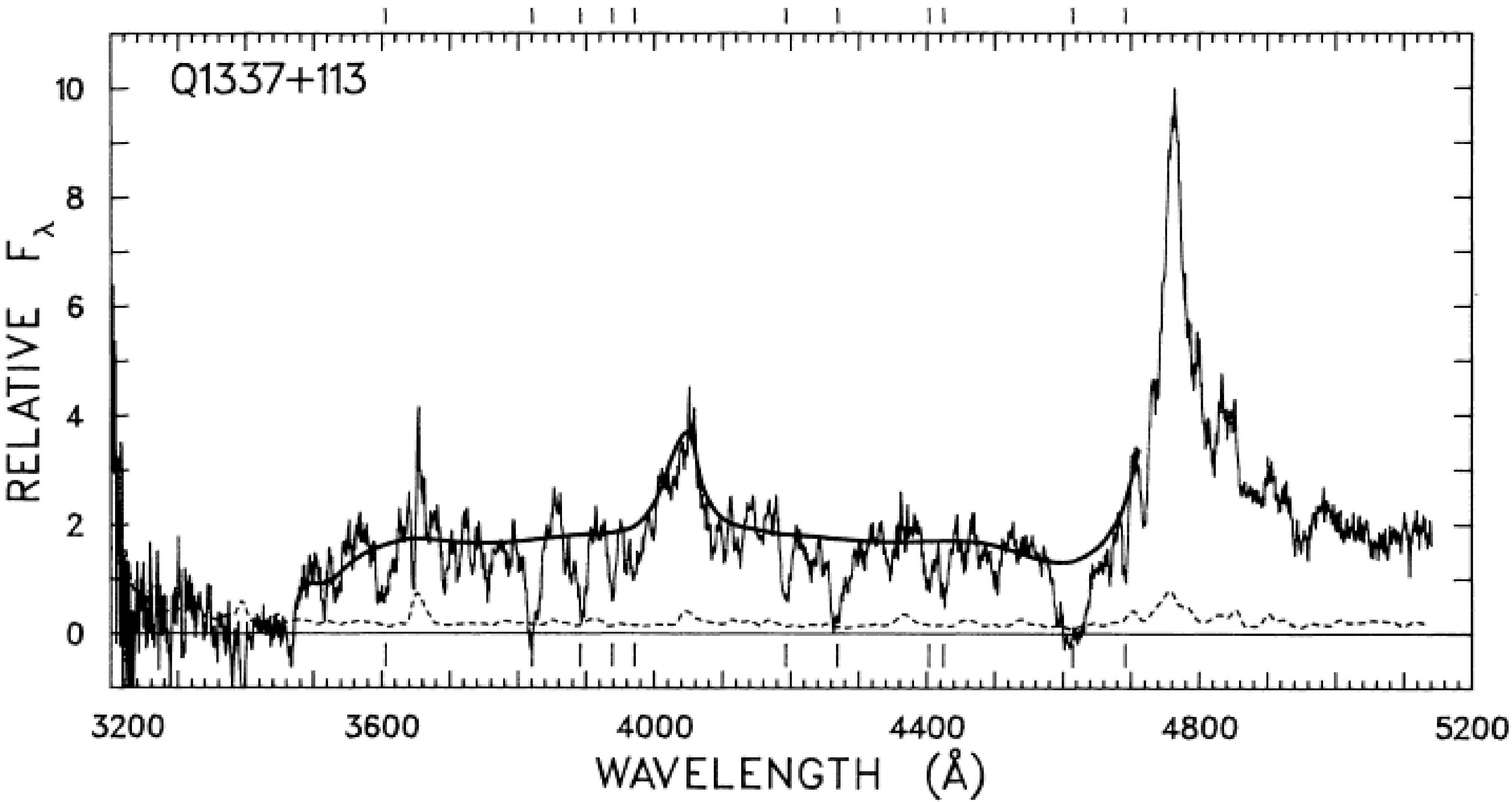
Machine Learning in Astronomy

- Astronomy is rife with tasks demanding human labor
  - Source identification
  - Continuum fitting
  - Line identification
- Machine Learning
  - Can perform many of these tasks
  - Auto-magically, repeatably, better!
- Astrophysics and ML
  - I harbor my doubts…



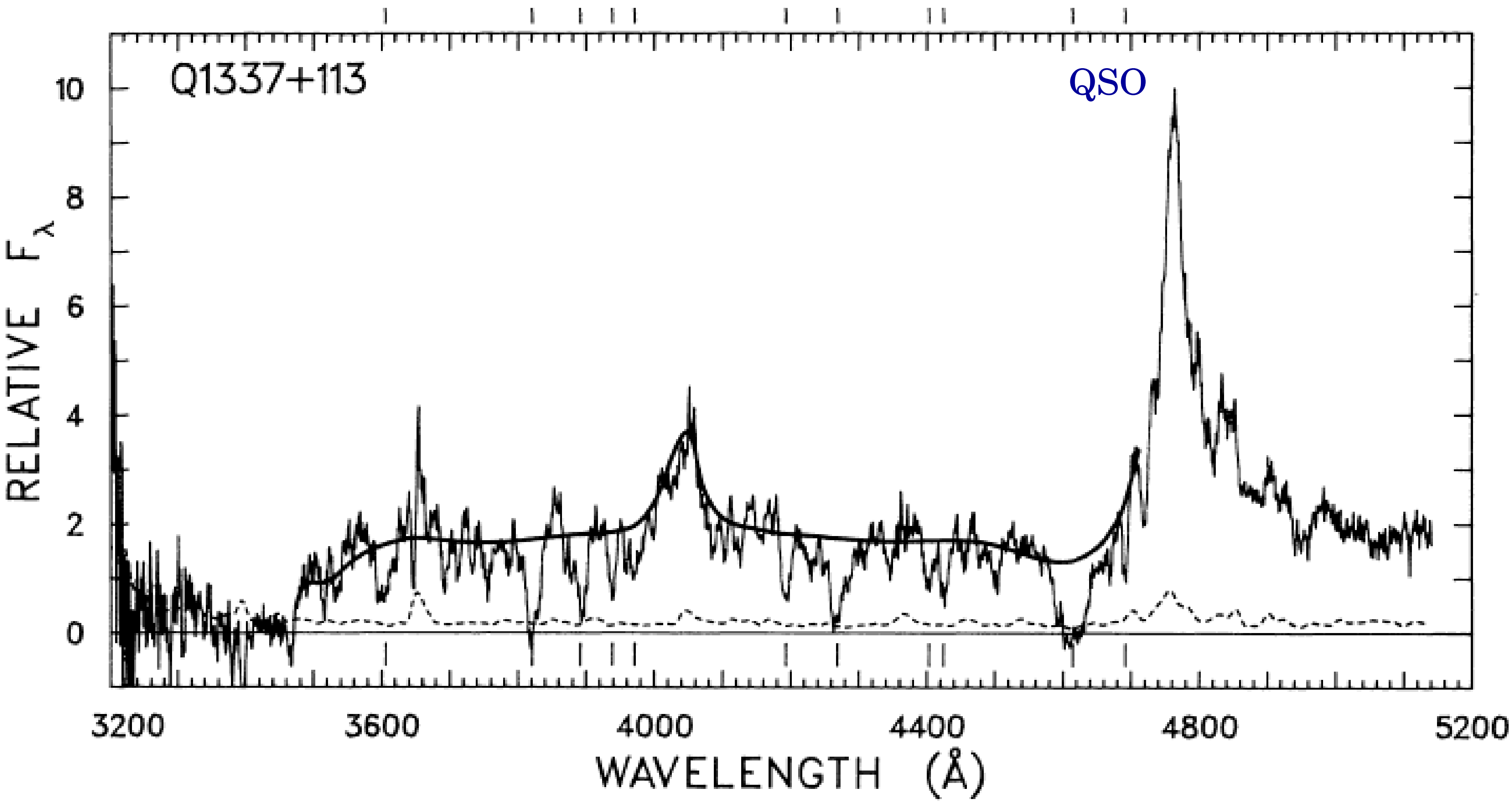
The Damped Lyα Systems (DLAs)

Wolfe+86



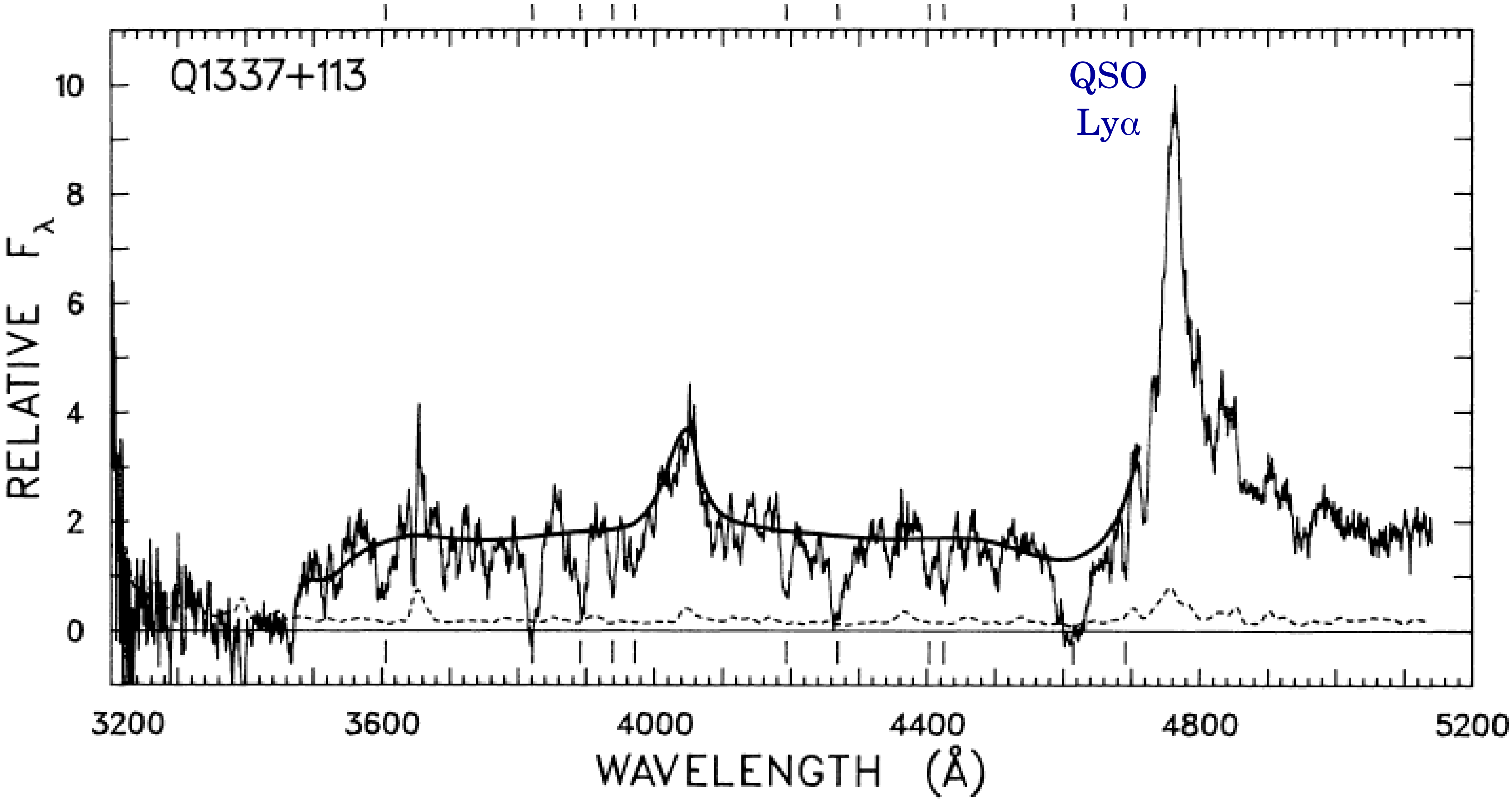
The Damped Lyα Systems (DLAs)

Wolfe+86



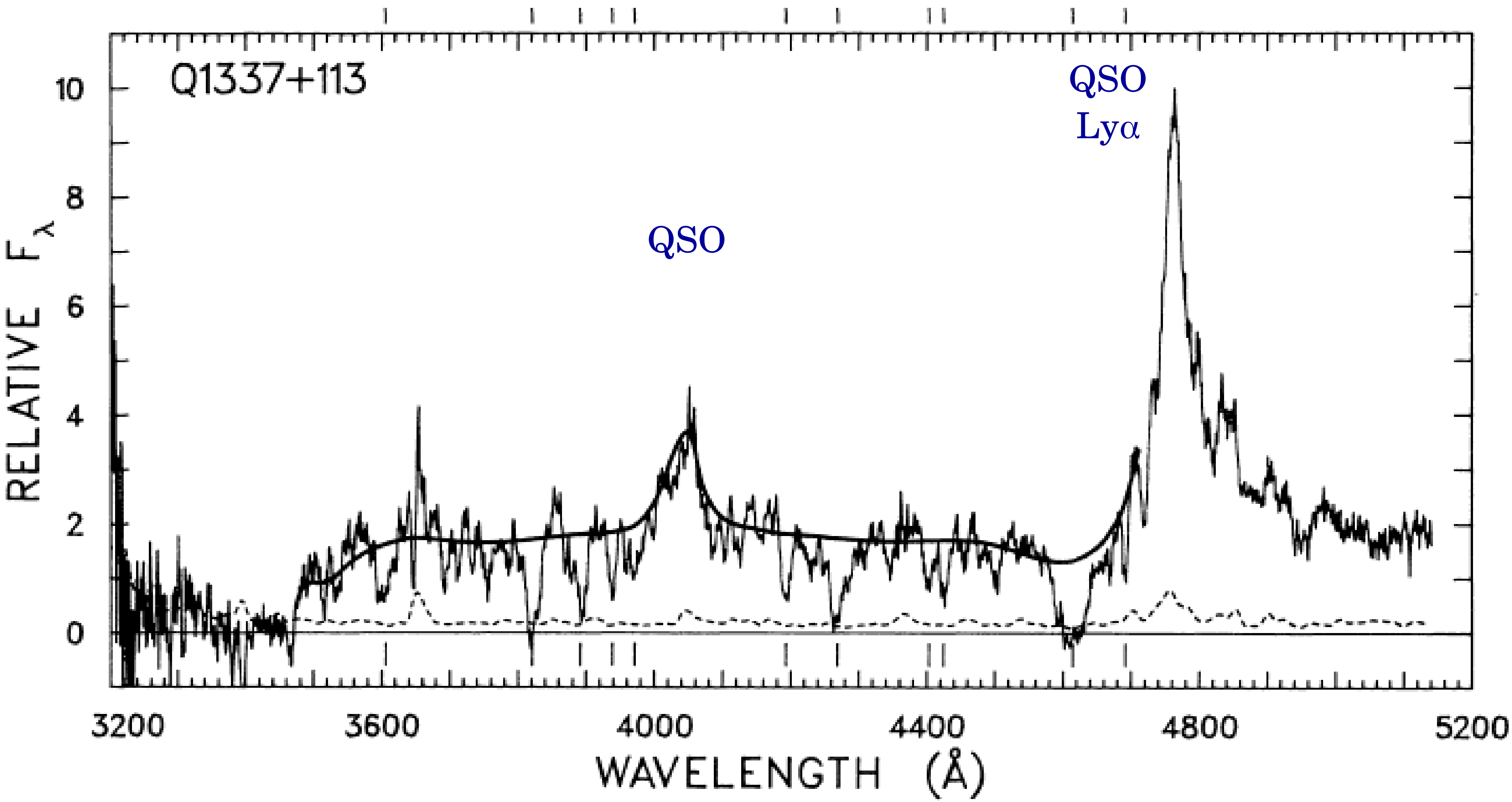
The Damped Lyα Systems (DLAs)

Wolfe+86



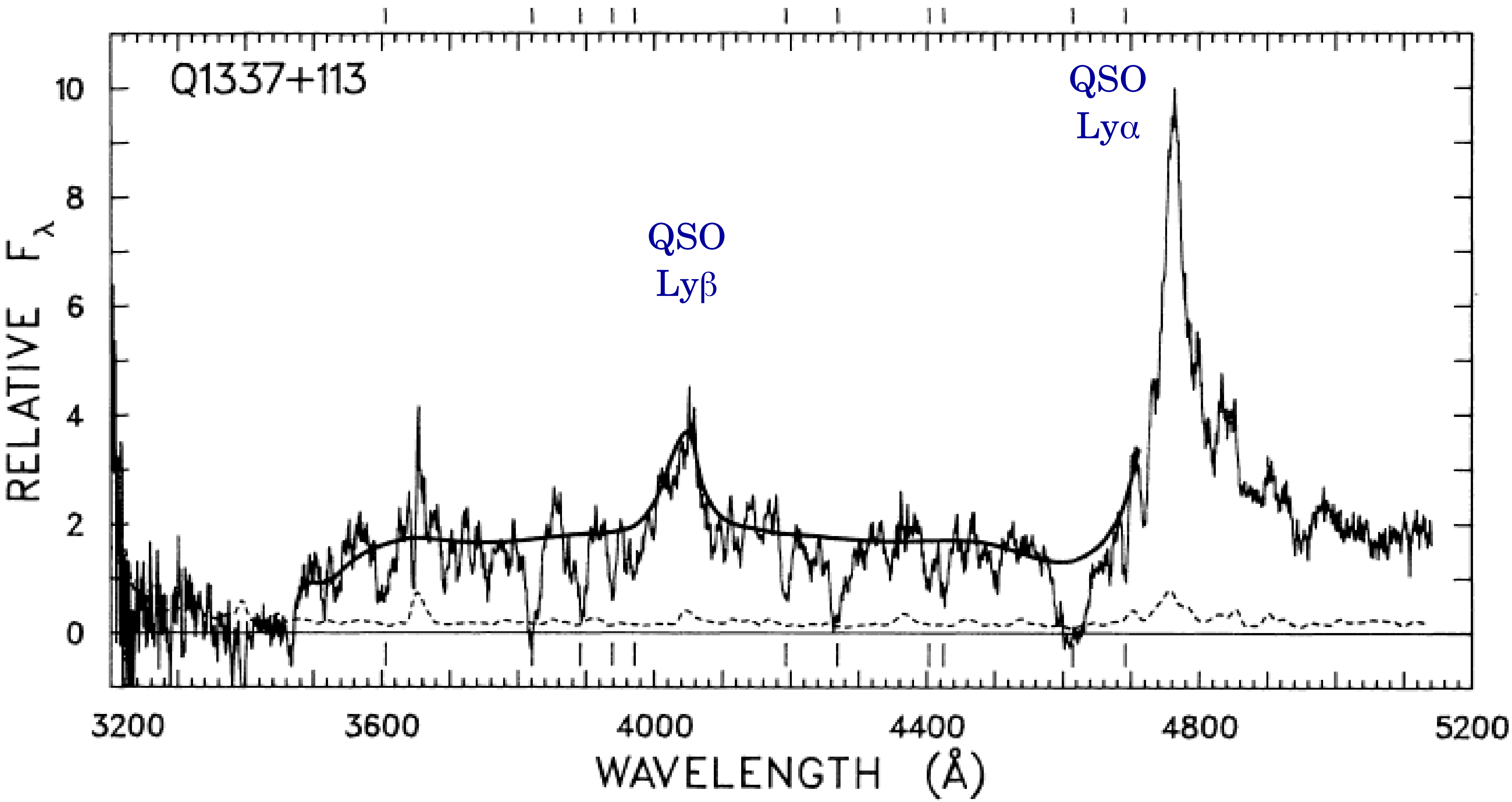
The Damped Lyα Systems (DLAs)

Wolfe+86



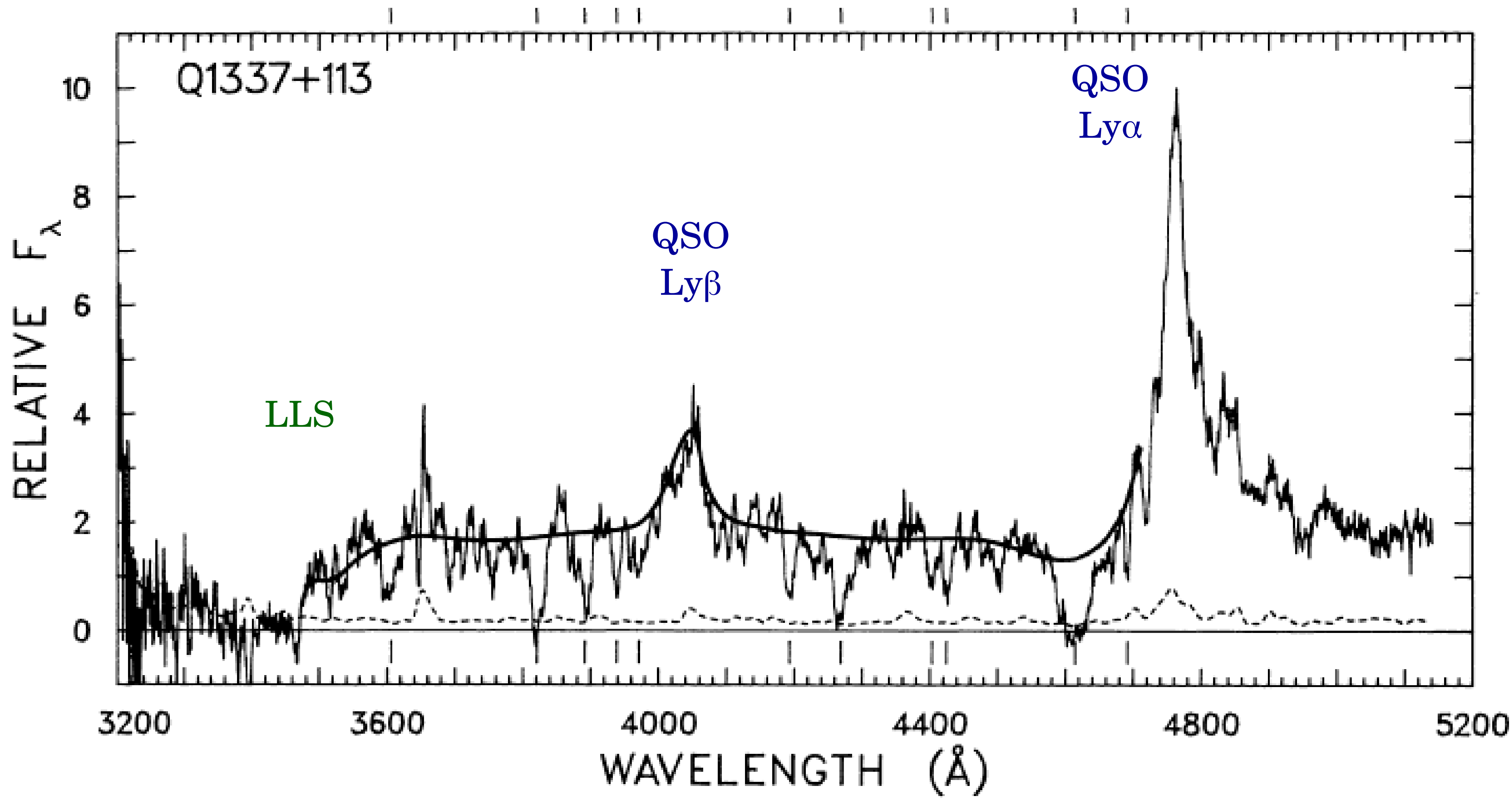
The Damped Lyα Systems (DLAs)

Wolfe+86



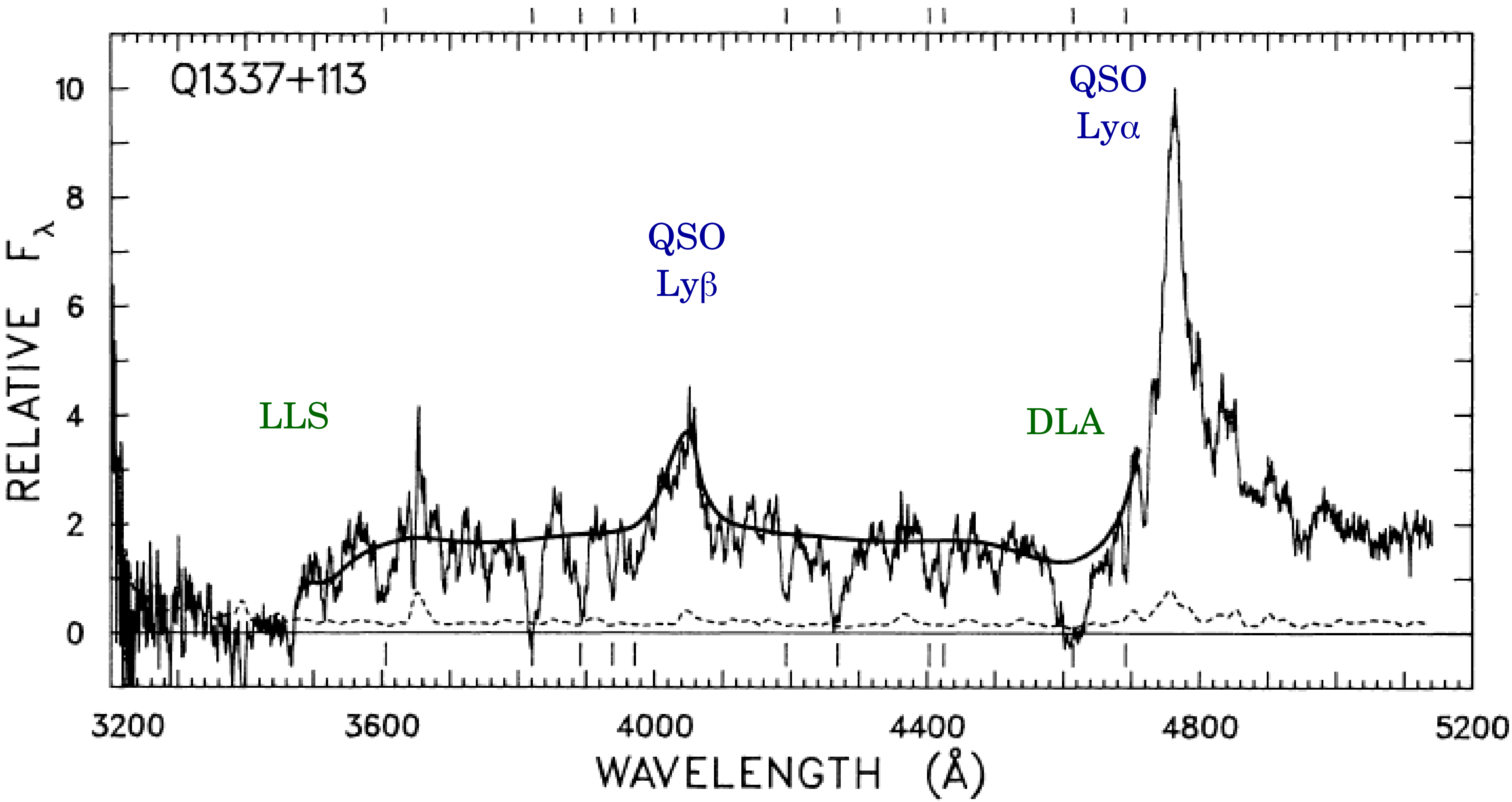
The Damped Lyα Systems (DLAs)

Wolfe+86



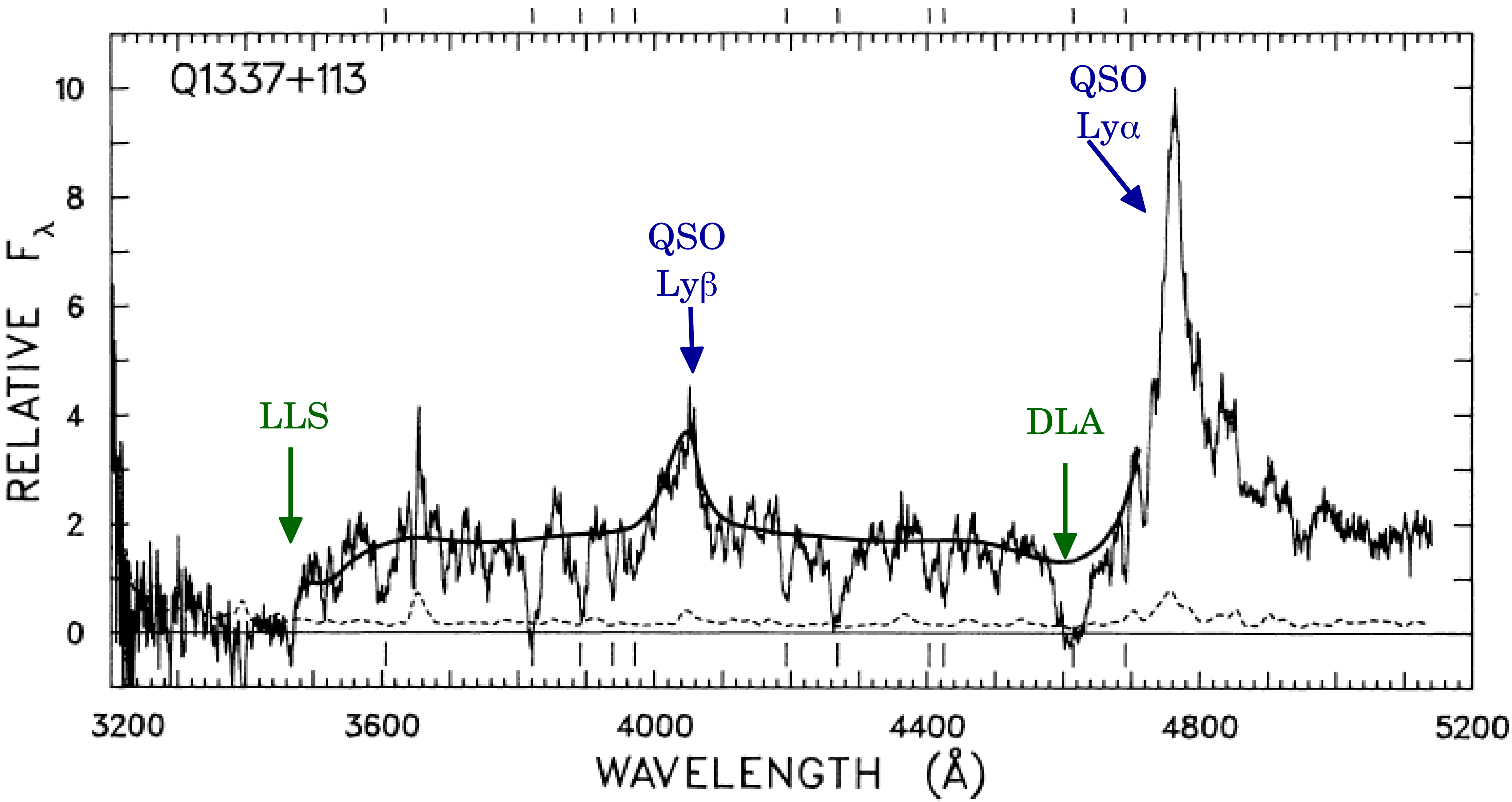
The Damped Lyα Systems (DLAs)

Wolfe+86

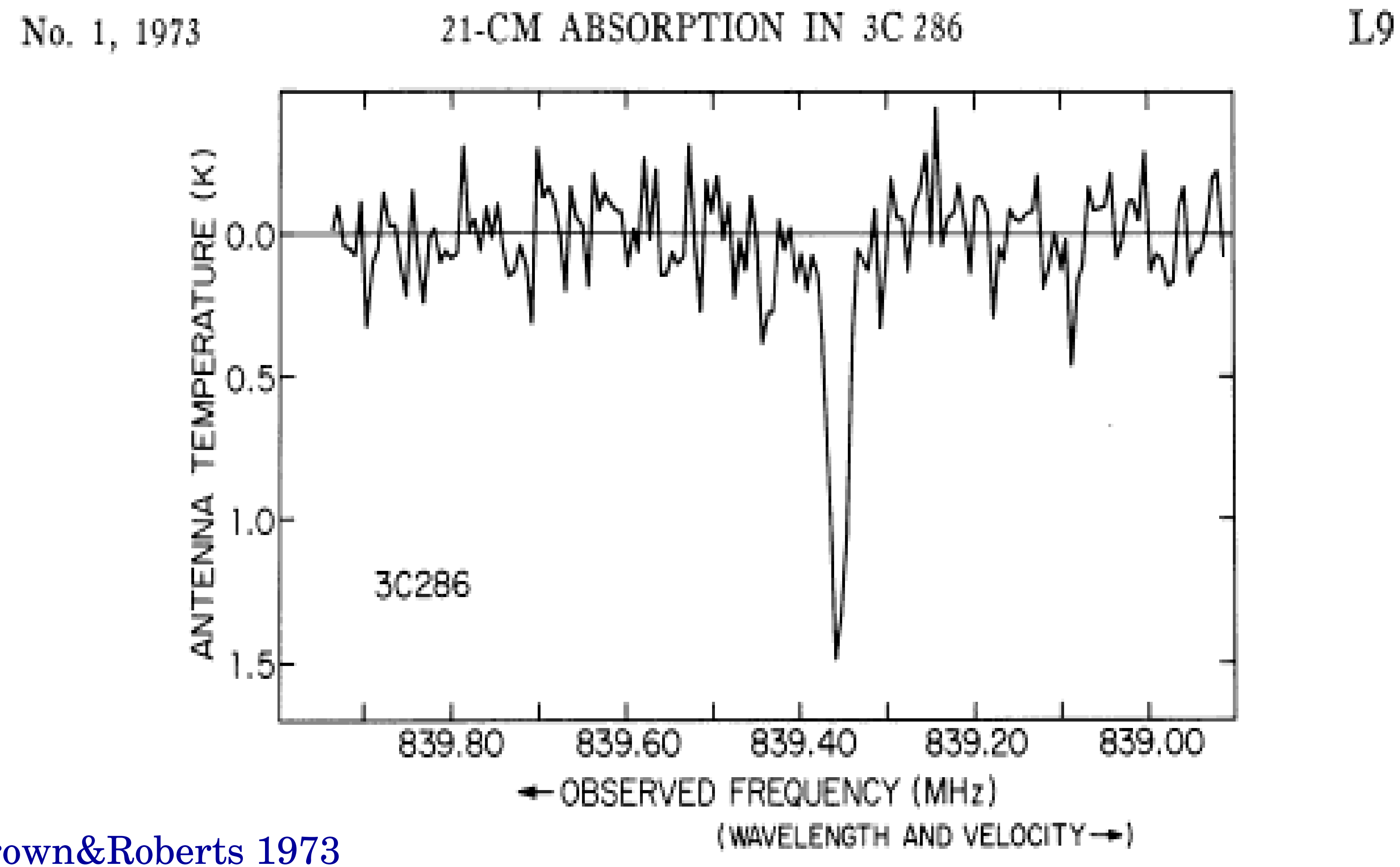


The Damped Lyα Systems (DLAs)

Wolfe+86



DLAs from 21cm (Wolfe)



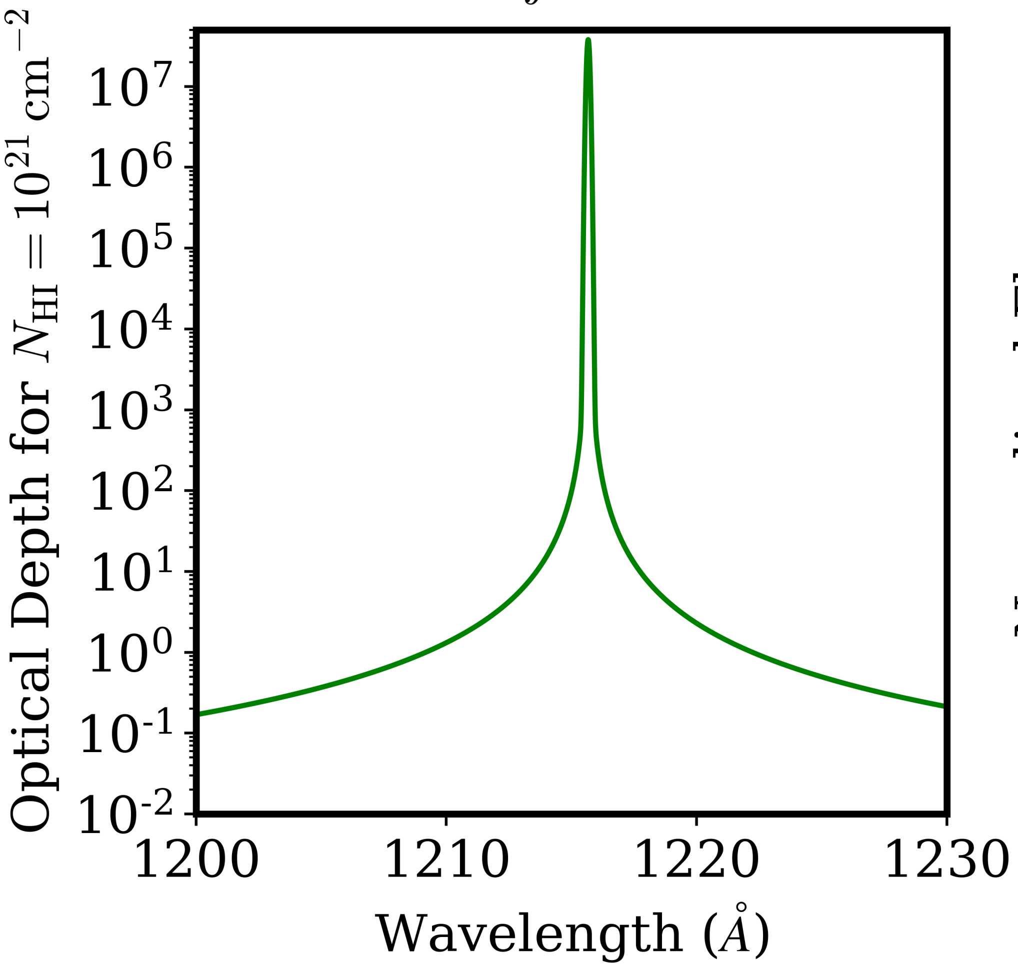
Brown&Roberts 1973

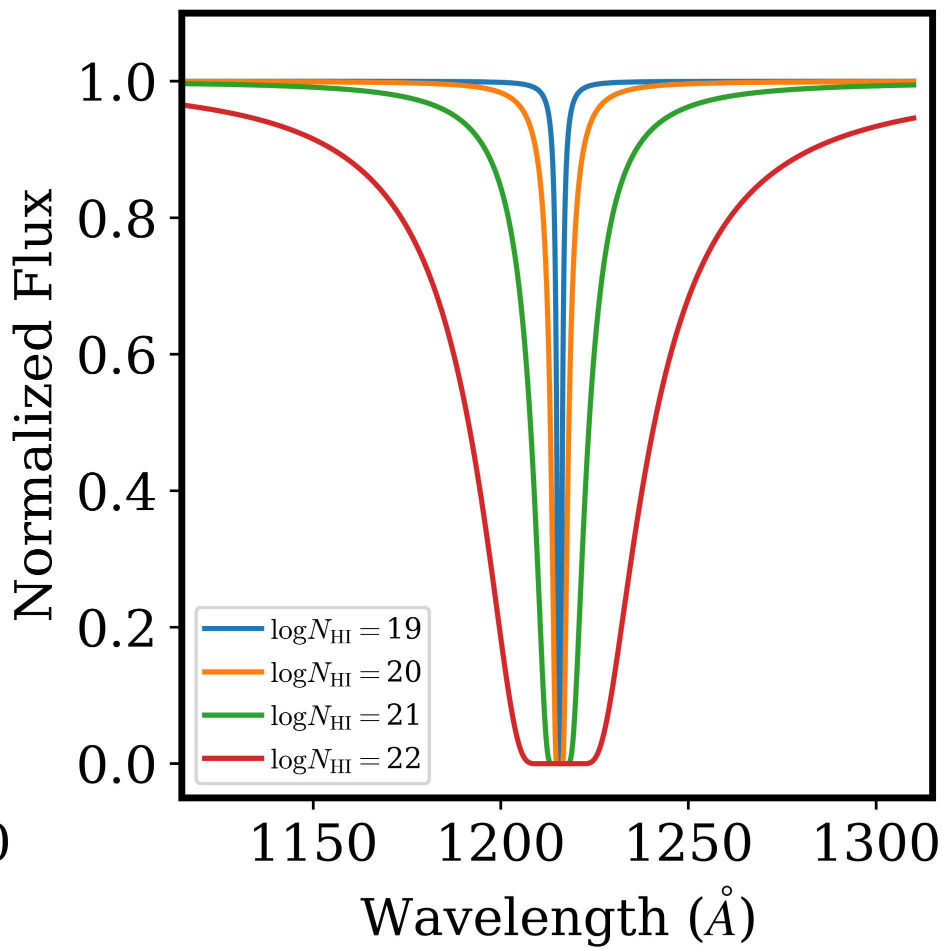
DLAs: Quantum Mechanics

$$\kappa_\nu = n_j \sigma_{jk}(\nu) \quad \tau_\nu = \sigma_\nu \int n_j ds$$

$$\sigma_\nu = \sigma_{jk}\phi_\nu \qquad N_j \equiv \int n_j ds$$

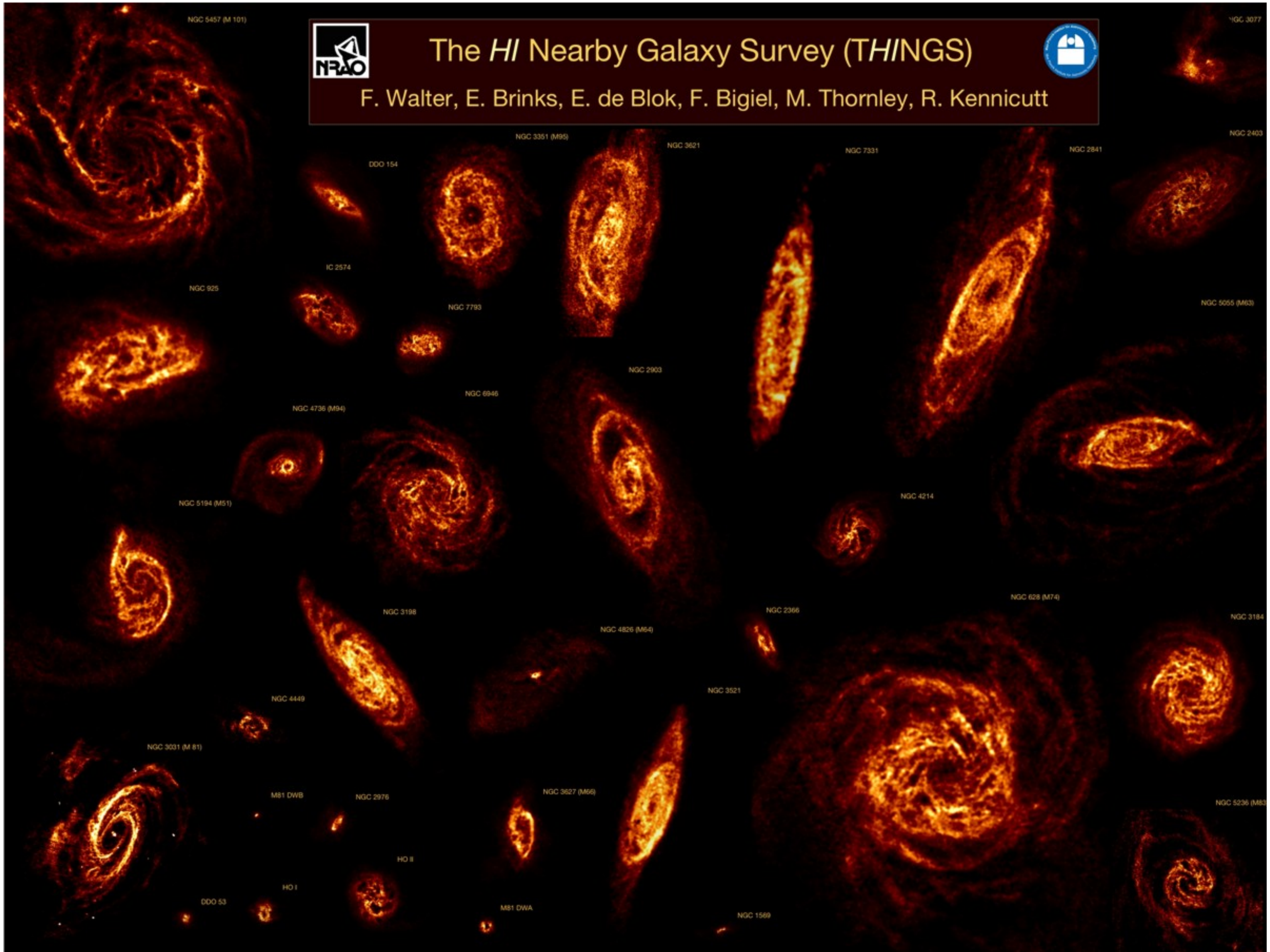
$$\phi_V(\nu) = \frac{\gamma}{4\pi} \int_{-\infty}^{\infty} \frac{\left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} \exp\left(-\frac{mv^2}{2kT}\right)}{(\nu - \nu_{jk} - \nu_{jk} v/c)^2 + (\gamma/4\pi)^2} dv$$





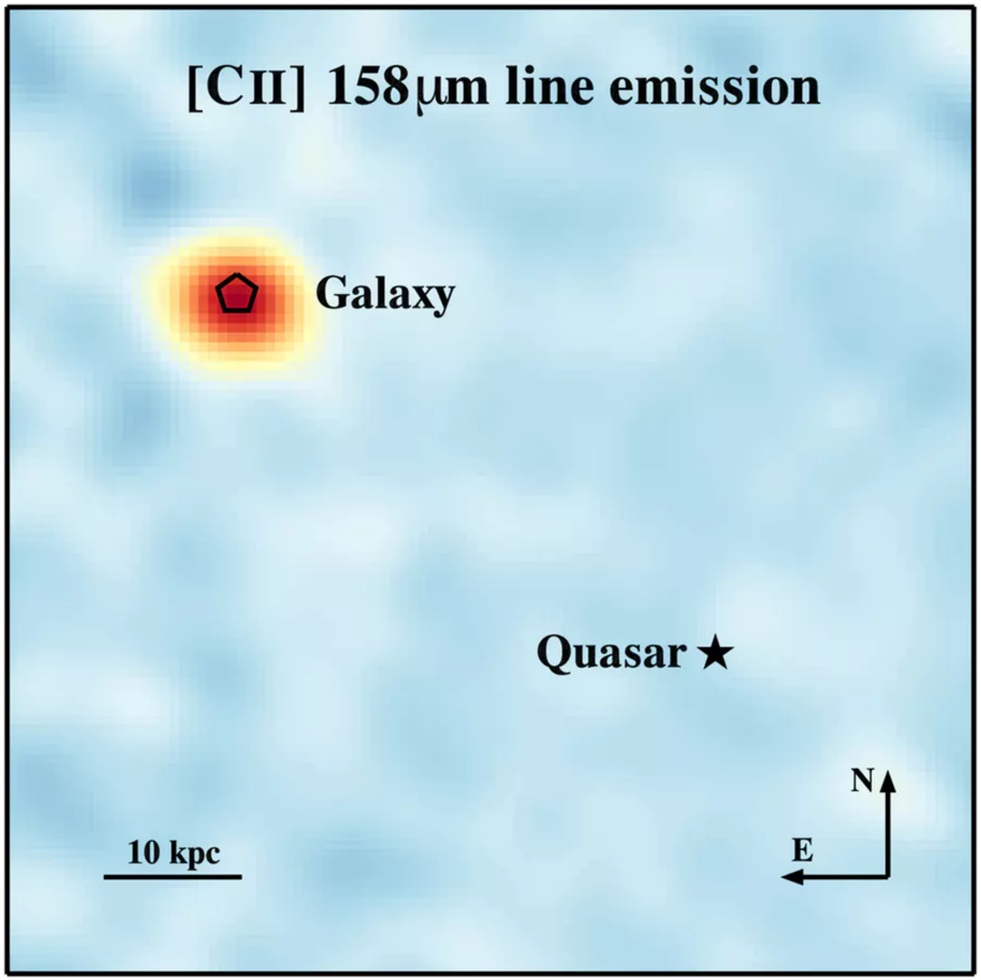
DLAs in the Modern Universe

Kim+
Zwaan+
Walter+

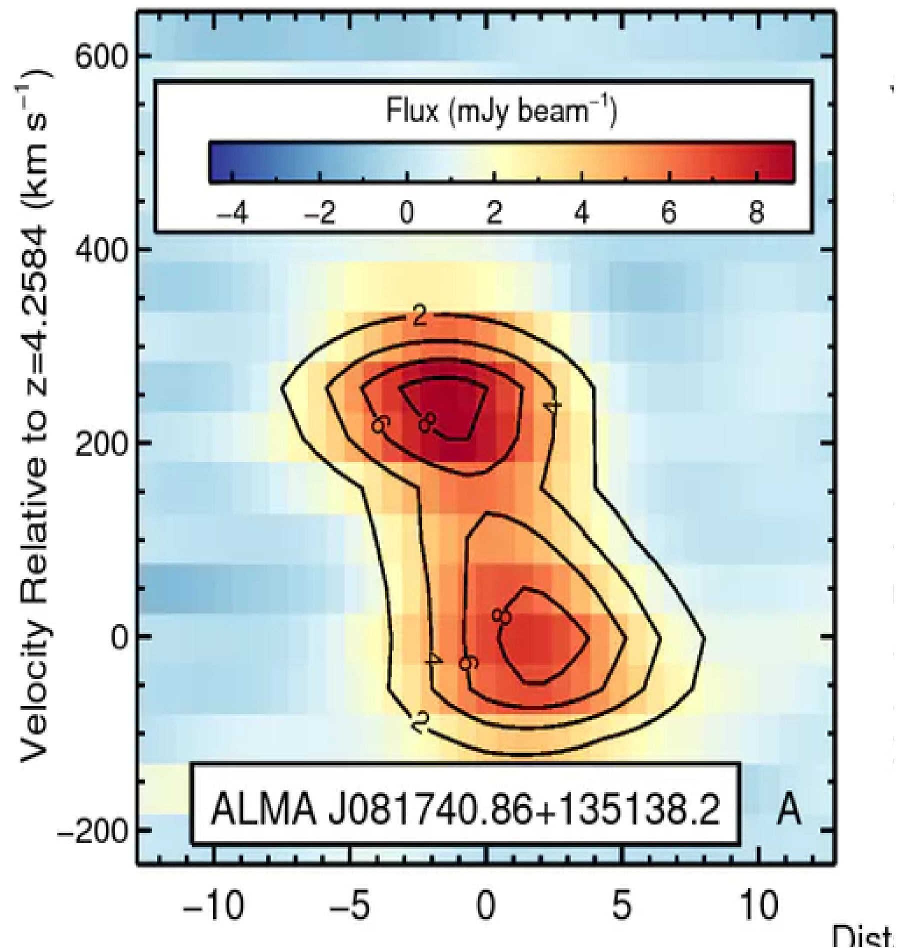


Cold, neutral gas in galaxies. Aka the Interstellar Medium (ISM)

DLAs in the Distant Universe

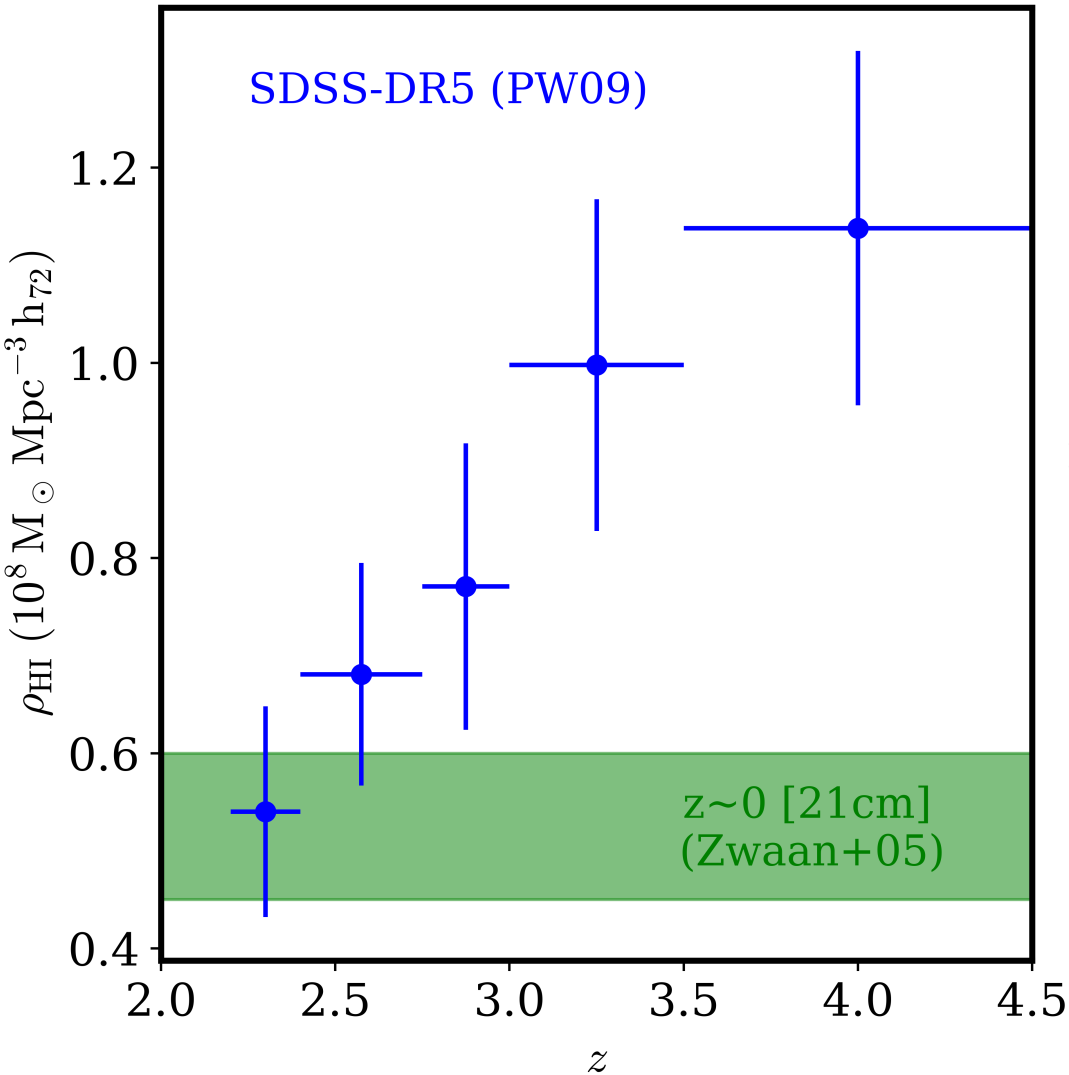


Neeleman+17

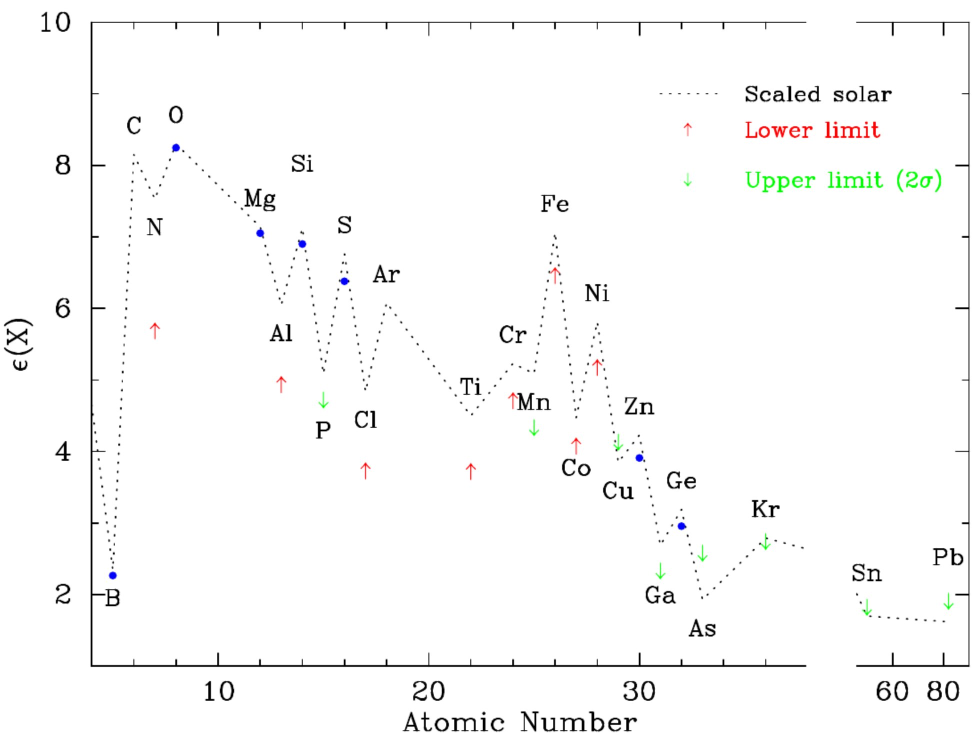


Cold, neutral gas in galaxies? Or around galaxies (aka CGM)??!
[see https://vimeo.com/209248385]

DLA Science

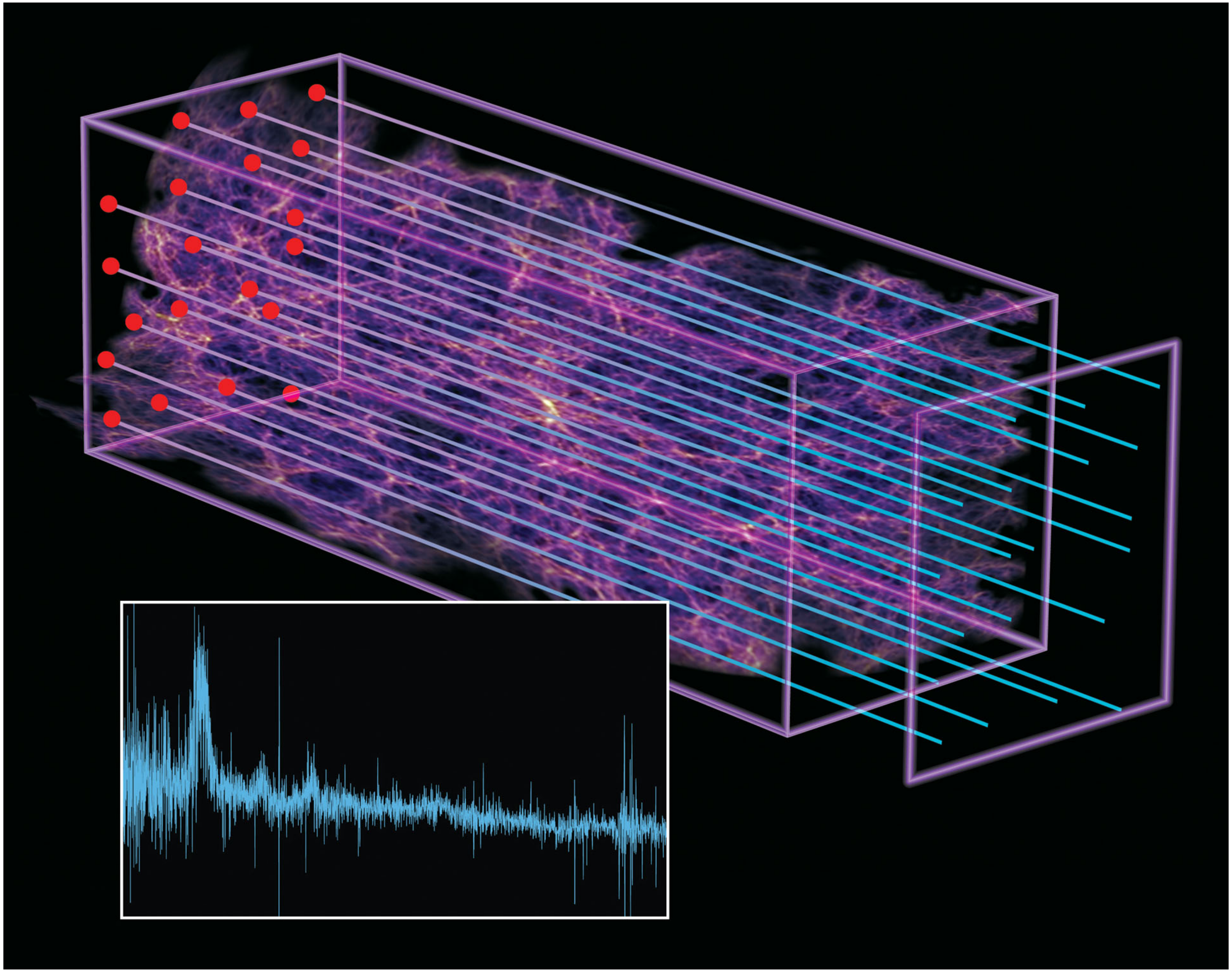


XHW03



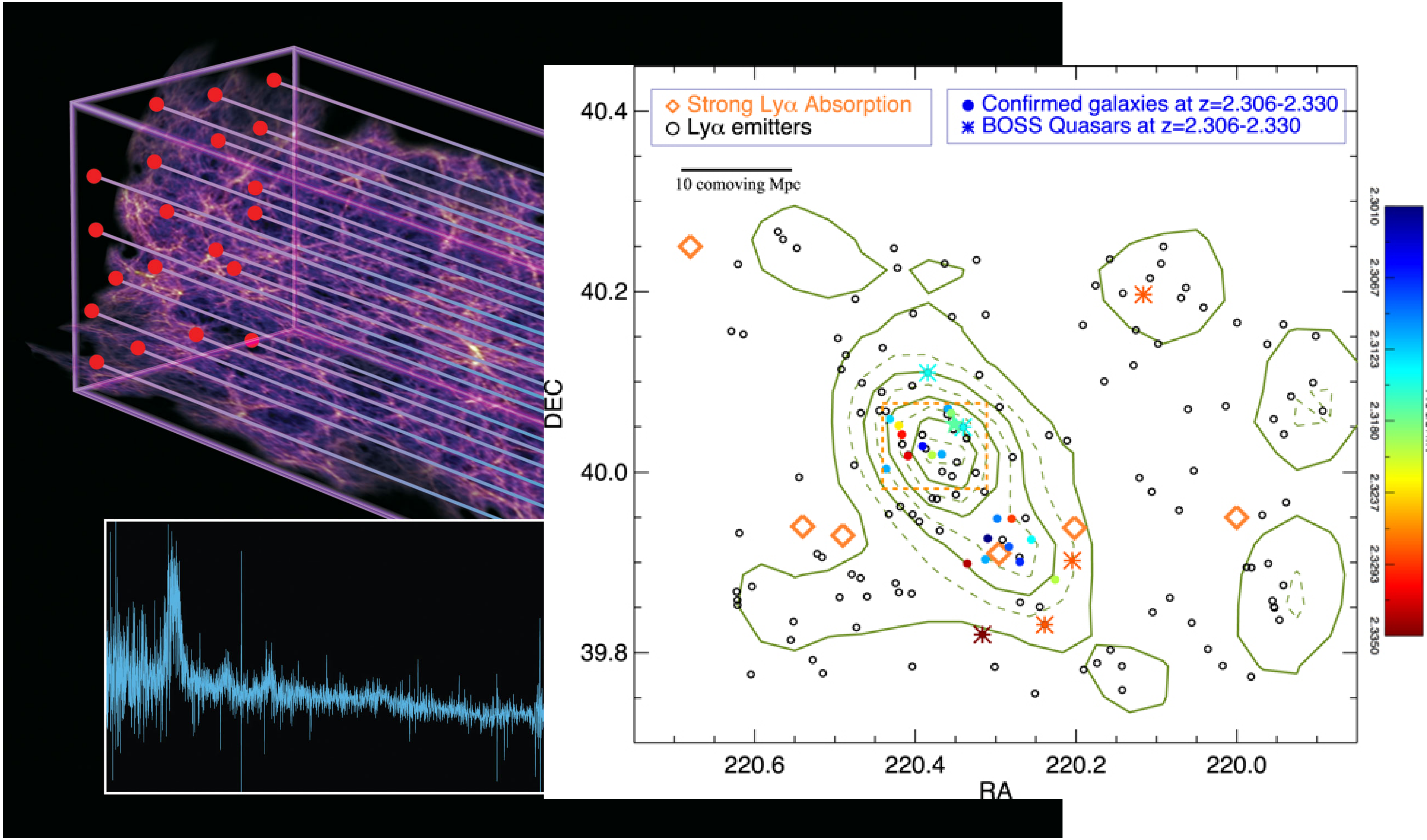
Cosmological measurements; nucleosynthesis; galaxy formation

DLA as Contaminants



Baryonic Acoustic Oscillations (Solzar+15)
Search for protoclusters (MAMMOTH; Cai+17)

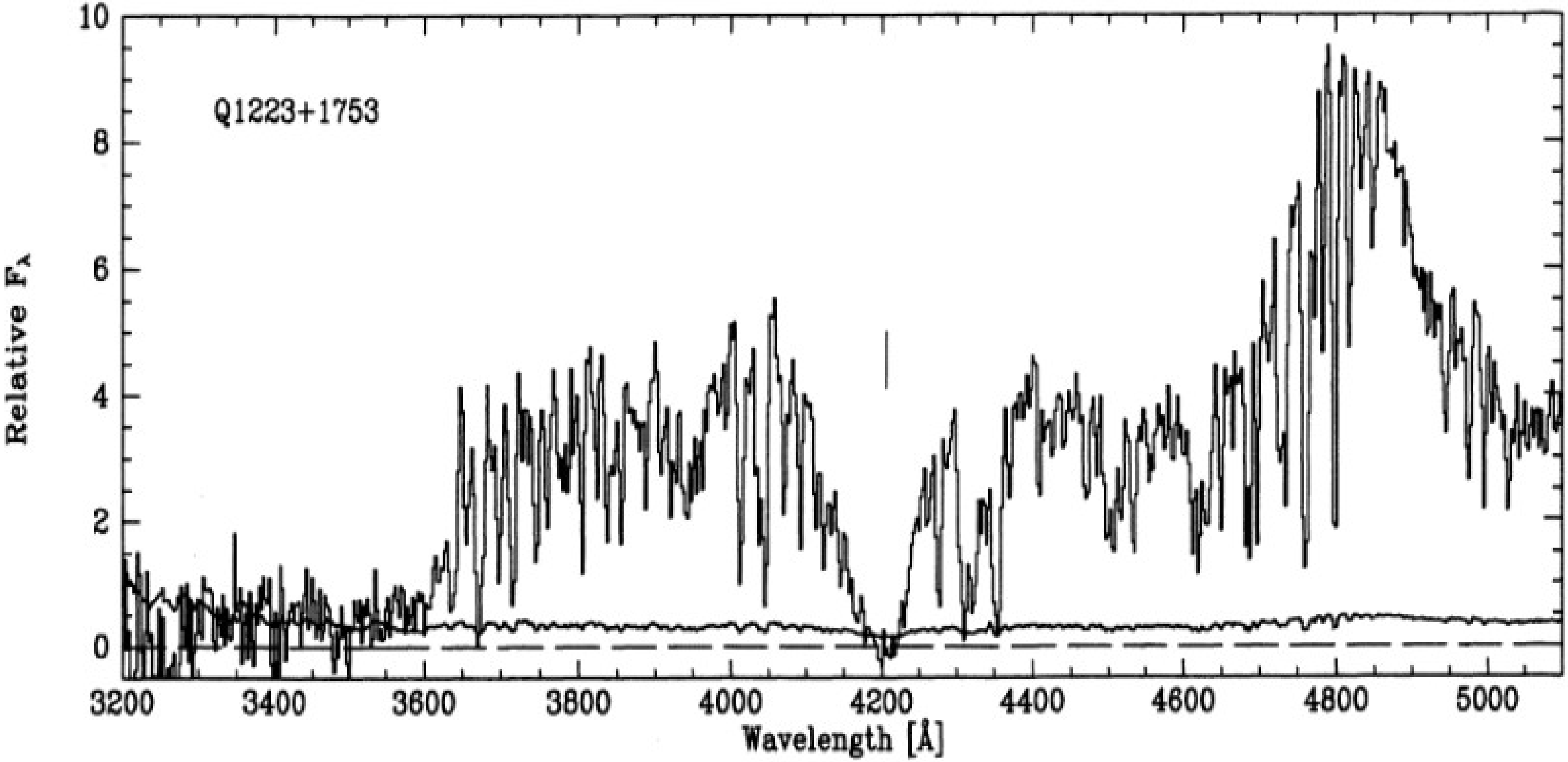
DLA as Contaminants



Baryonic Acoustic Oscillations (Solzar+15)
Search for protoclusters (MAMMOTH; Cai+17)

DLA Analysis (Old School)

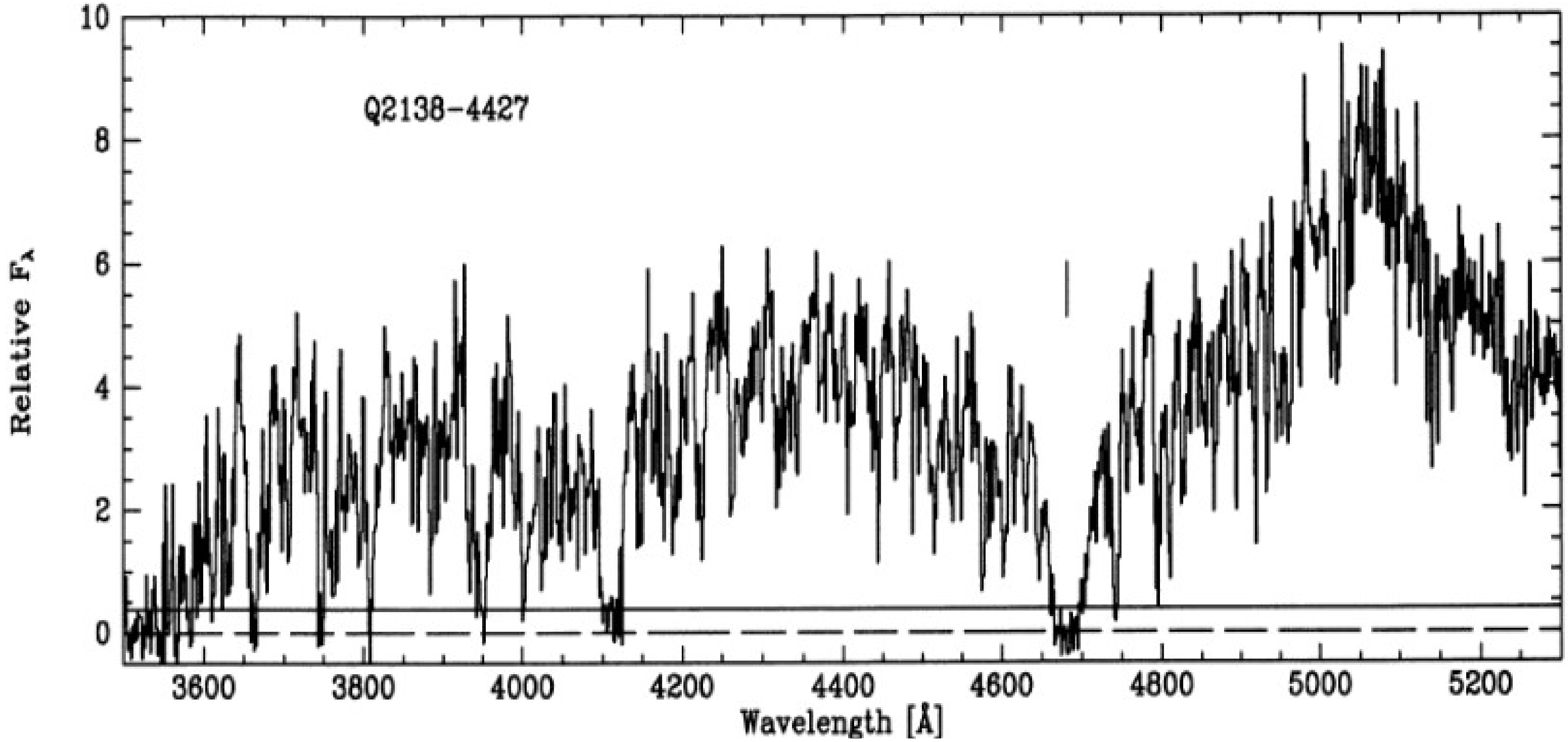
Wolfe+95



Visual inspection (first error array!); by-eye N_{HI} fitting

DLA Analysis (Old School)

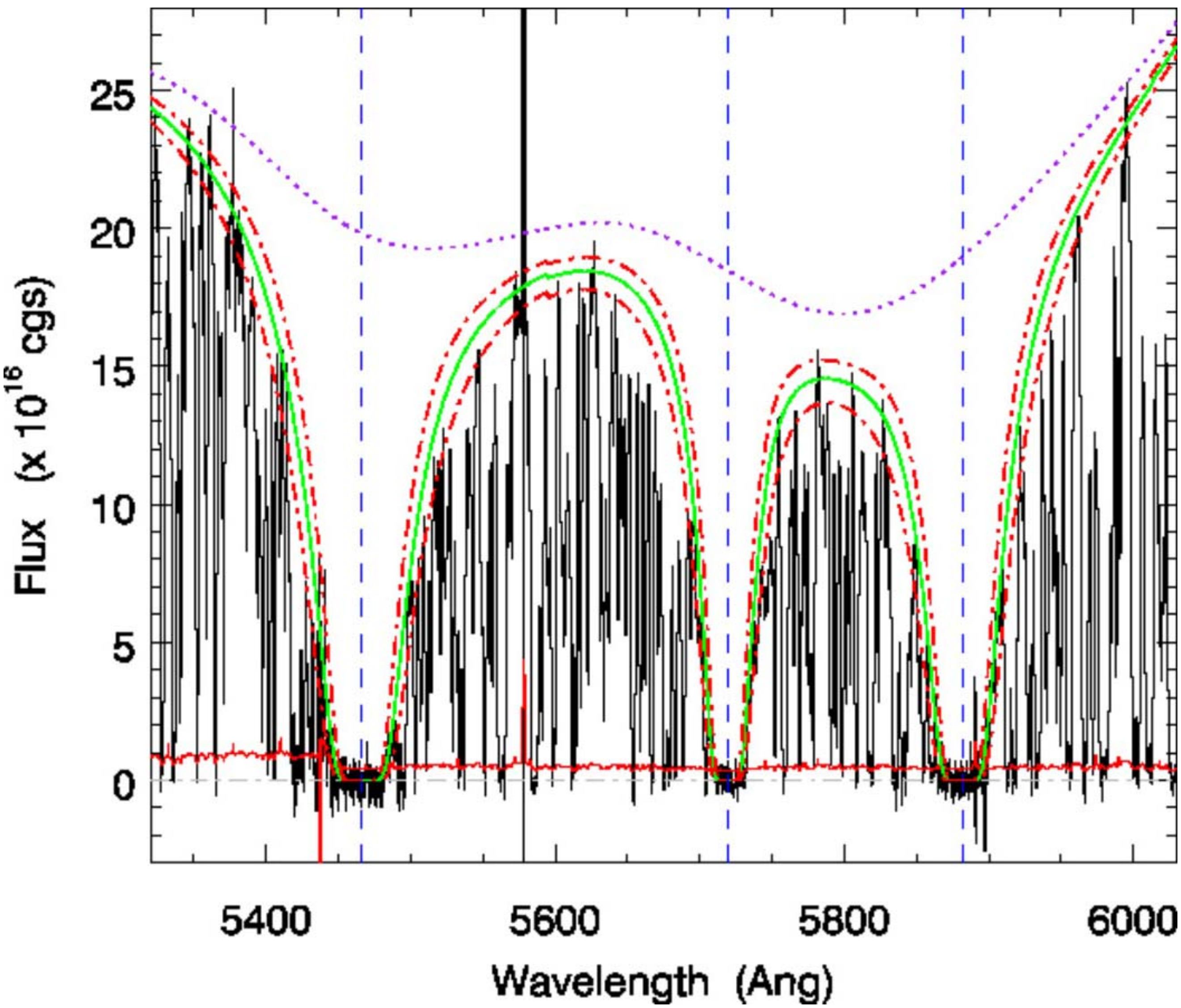
Wolfe+95



Visual inspection (first error array!); by-eye N_{HI} fitting

DLA Analysis (Old School)

X+03



Visual inspection (first error array!); by-eye N_{HI} fitting

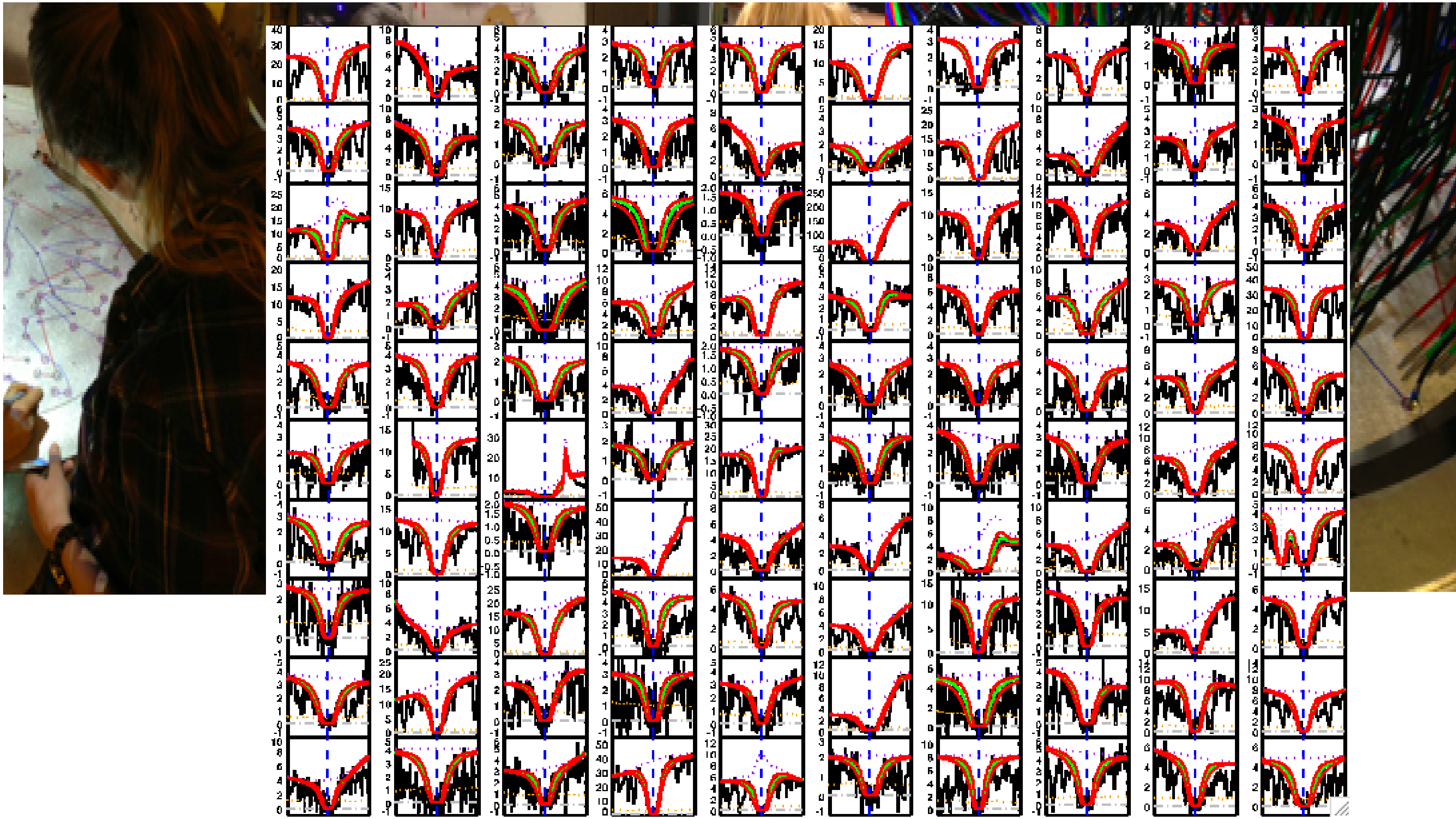
DLA Analysis (Early SDSS)

X+03
X+05

Visual inspection (first error array!); by-eye N_{HI} fitting

DLA Analysis (Early SDSS)

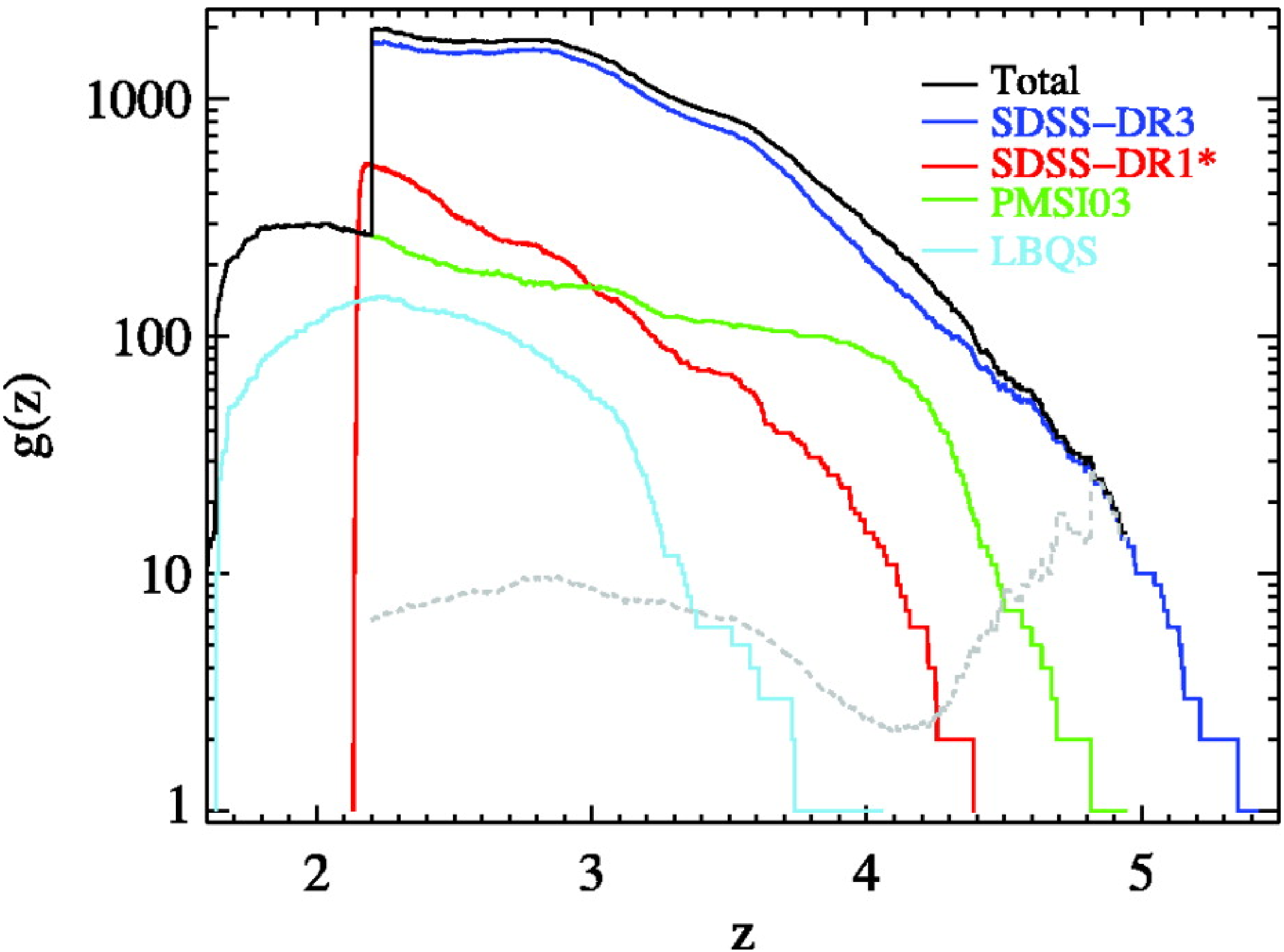
X+03
X+05



Visual inspection (first error array!); by-eye N_{HI} fitting

DLA Analysis (Early SDSS)

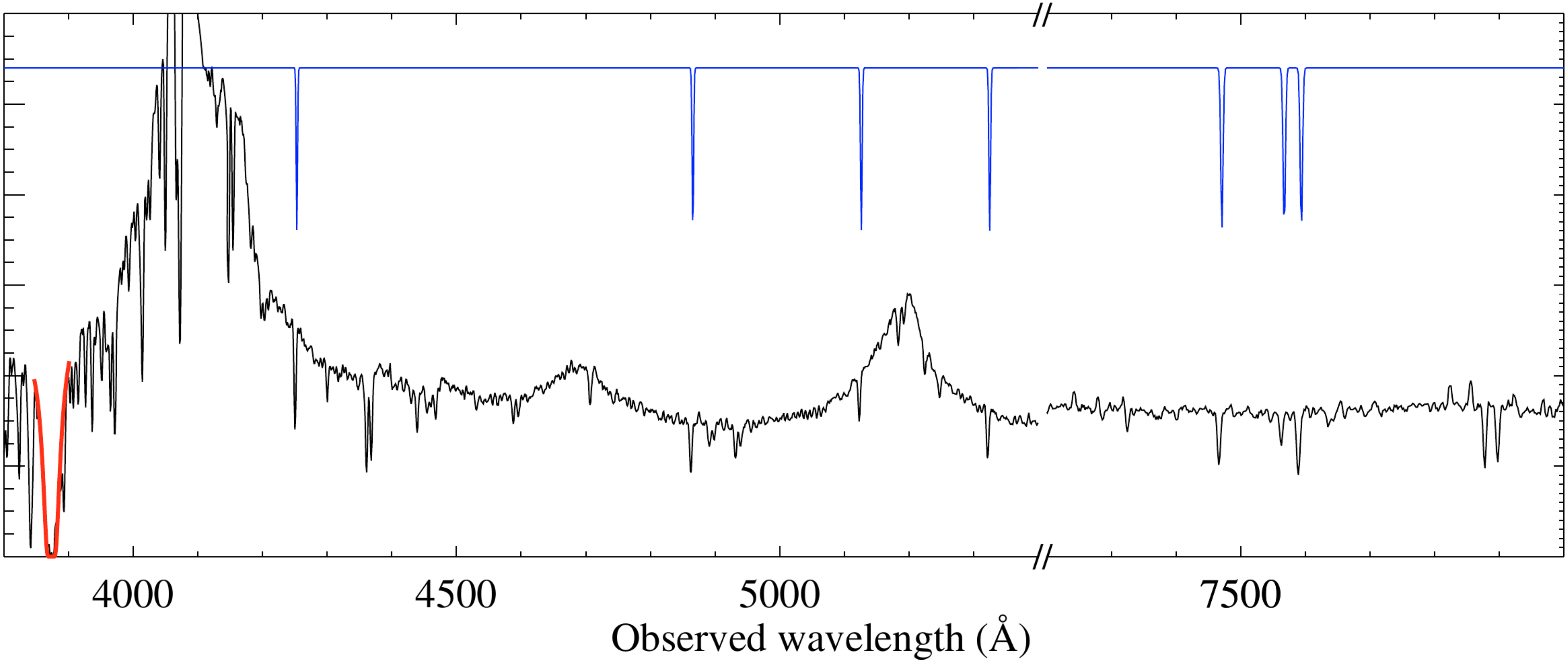
X+03
X+05



Survey path: Number of QSOs examined for DLAs

DLA Analysis (BOSS)

Noterdaeme+09,12



Spearman Correlation Analysis

DLA Analysis (BOSS)

$$\Pr(\mathcal{M} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid \mathcal{M}) \Pr(\mathcal{M})}{p(\mathcal{D})} = \frac{p(\mathcal{D} \mid \mathcal{M}) \Pr(\mathcal{M})}{\sum_i p(\mathcal{D} \mid \mathcal{M}_i) \Pr(\mathcal{M}_i)},$$

$$p(f) = \mathcal{GP}(f; \mu, K),$$

Garnett, Ho,
Bird 2016

Learned Gaussian Process model

DLA Analysis (BOSS)

$$\Pr(\mathcal{M} \mid \mathcal{D}) = \frac{p(\mathcal{D} \mid}{p}$$

$$p(f) = \mathcal{GP}(f; \mu, K),$$

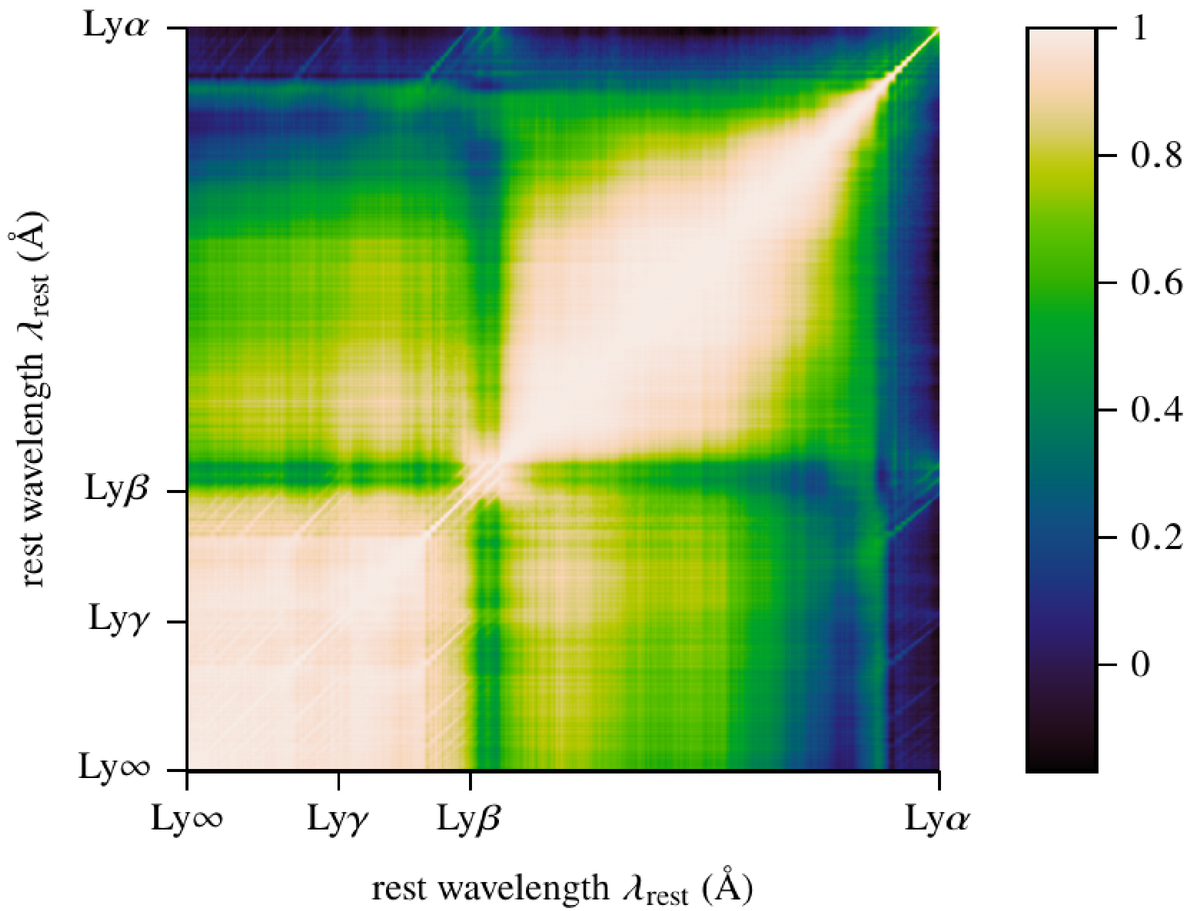
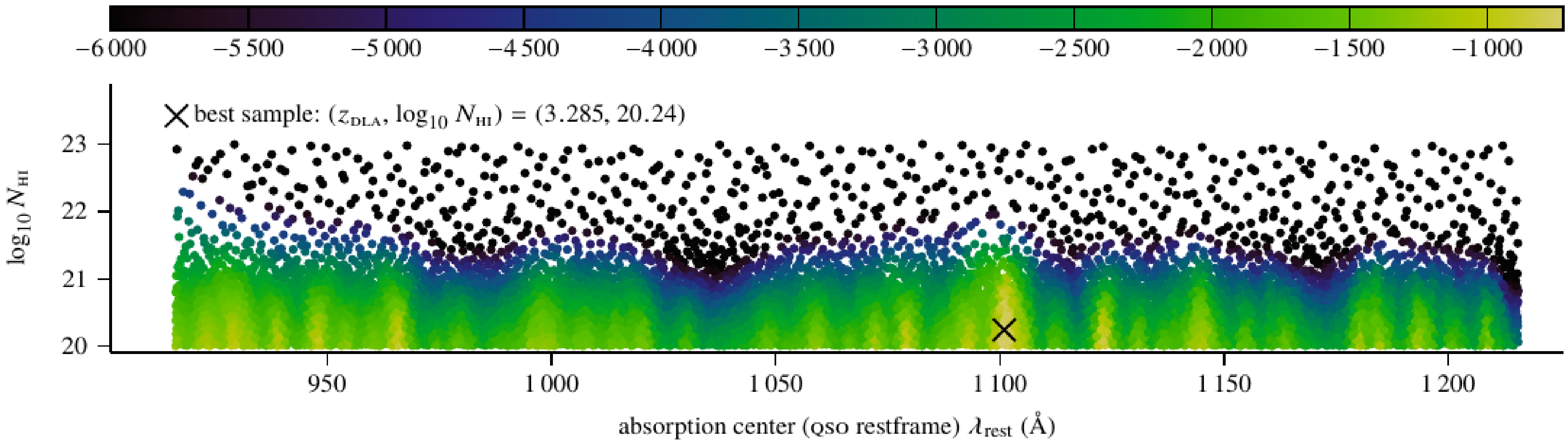


Figure 5. The observation covariance matrix **K** corresponding to the learned parameters shown in Figure 4. The entries have been normalized to give unit diagonal; the entries are therefore correlations rather than raw covariances.

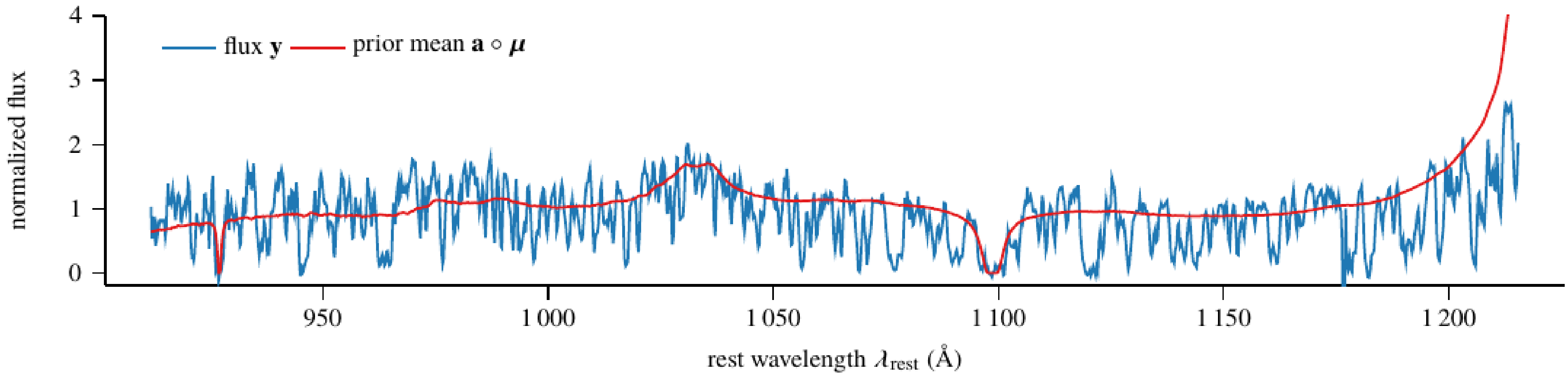
Garnett, Ho,
Bird 2016

Learned Gaussian Process model

DLA Analysis (BOSS)



(b) Sample log likelihoods for DLA model, $\{\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\text{QSO}}, \theta_i, \mathcal{M}_{\text{DLA}})\}$.

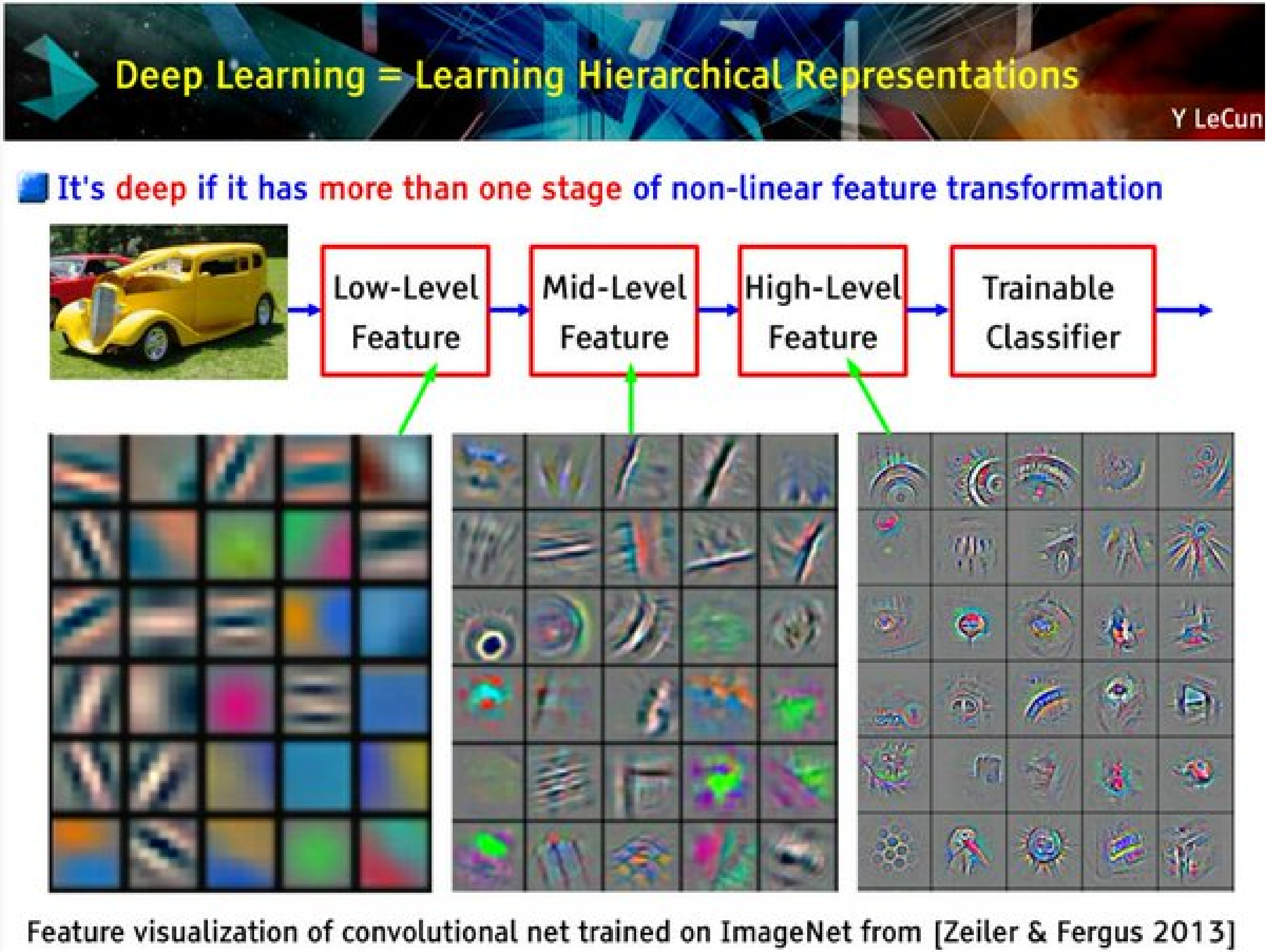


(c) DLA model: $\log p(\mathbf{y} \mid \boldsymbol{\lambda}, \mathbf{v}, z_{\text{QSO}}, \mathcal{M}_{\text{DLA}}) = -730.$

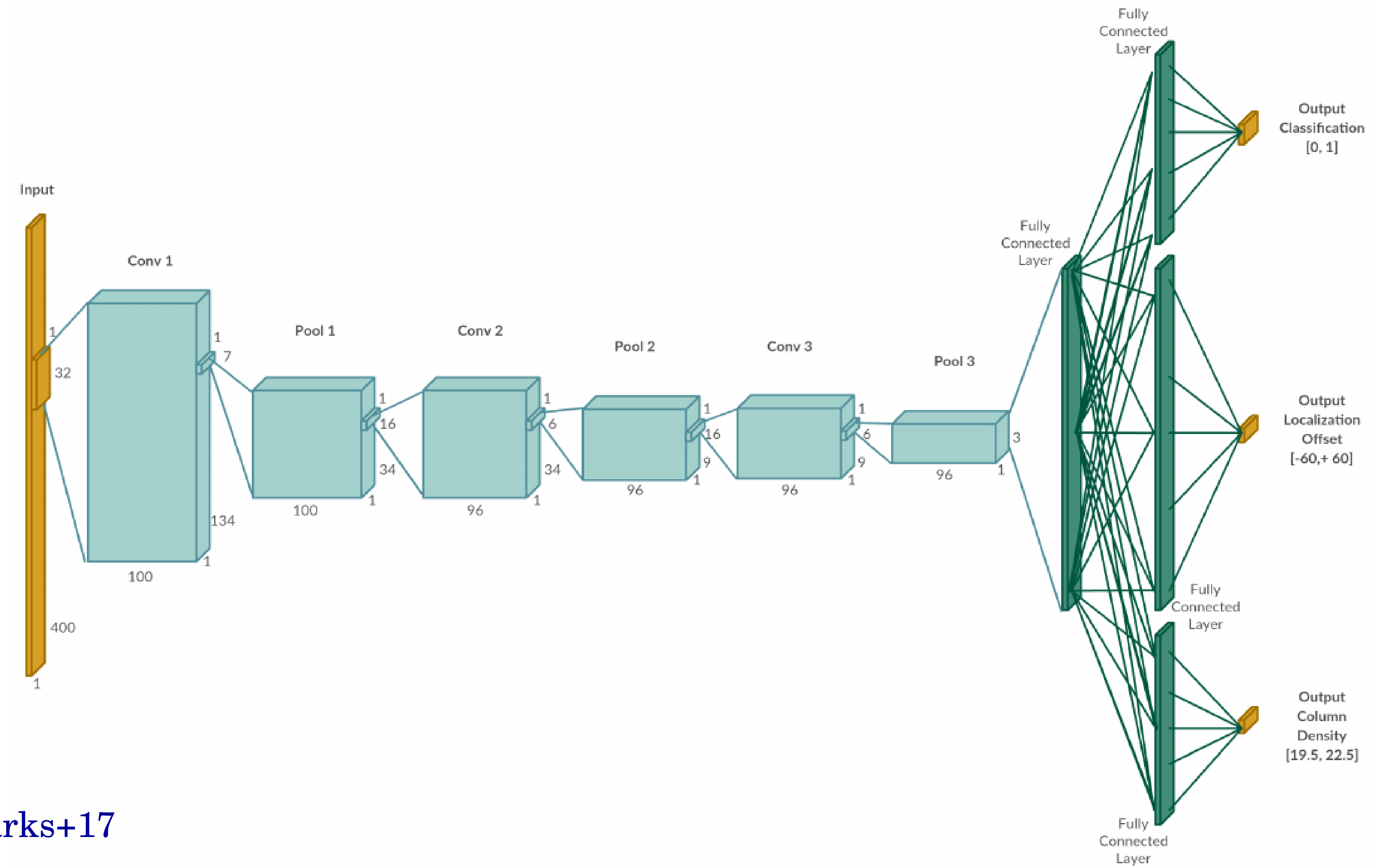
Garnett+16

Learned Gaussian Process model

Deep Learning



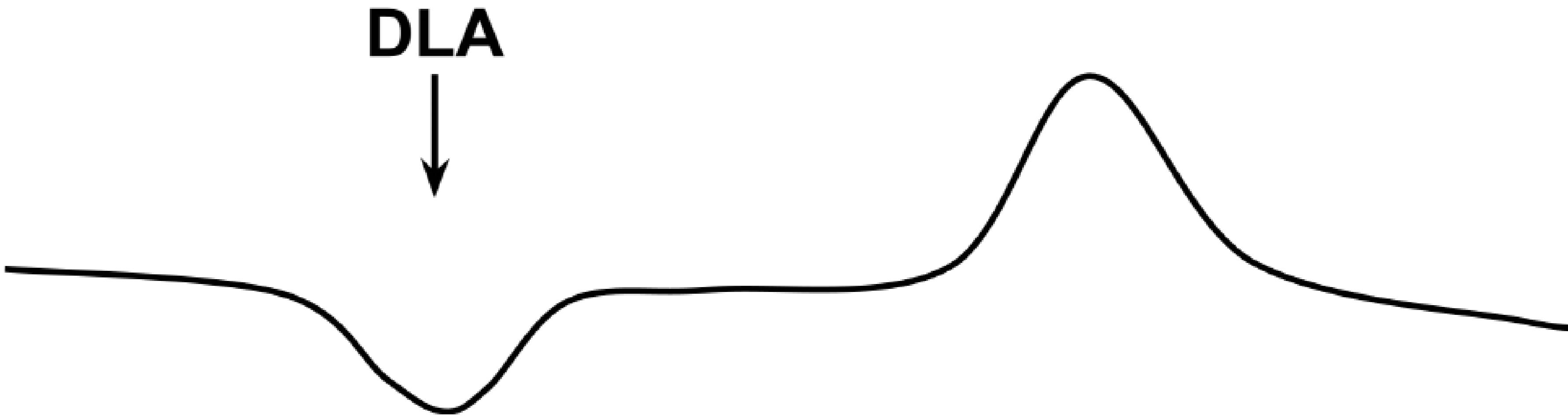
Deep Learning of DLAs



Parks+17

1709.04962 Neural Network architecture for our DLA Algorithm: Cast the spectrum snippet as a 1D image

CNN Labels and Learning



0	0	0	0	-60	...	-3	-2	-1	0	+1	+2	+3	...	+60	0	0	0	0	Localization
0	0	0	0	21.4	...	21.4	21.4	21.4	21.4	21.4	21.4	21.4	...	21.4	0	0	0	0	NHI (Column Density)
0	0	0	0	1	...	1	1	1	1	1	1	1	...	1	0	0	0	0	Classification

Three labels at each pixel: (1) Localization [i.e. redshift];
(2) HI column density; (3) Classification

Multi-task Learning: Combined loss function for all three labels

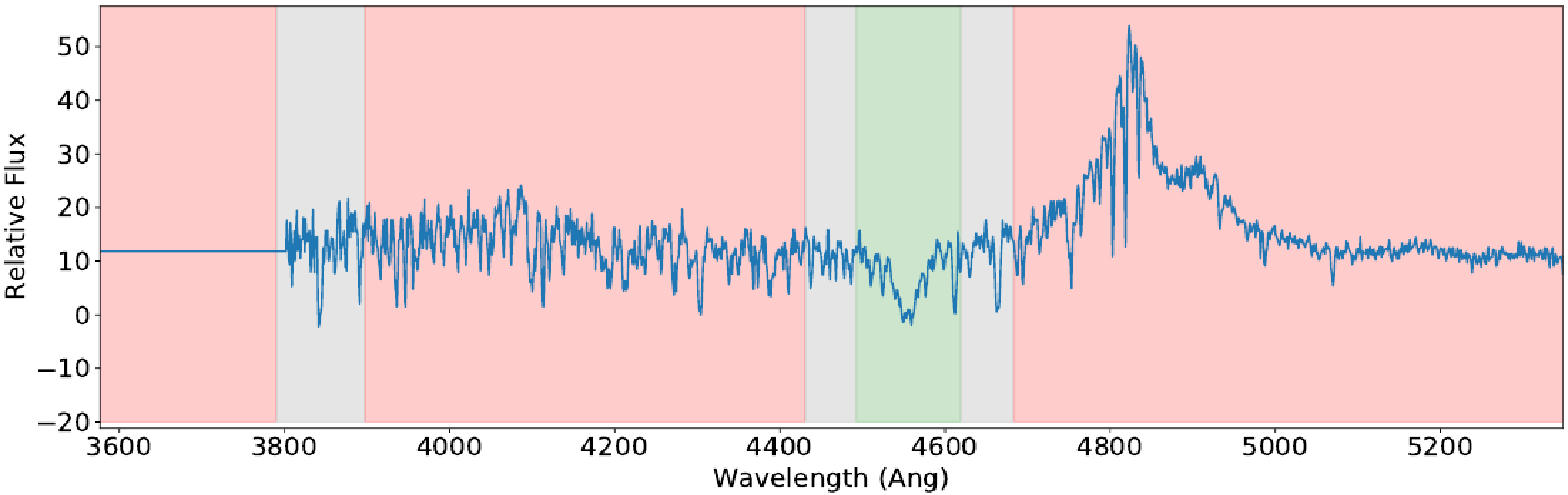
$$\mathcal{L}_c = -y_c \log(\hat{y}_c) - (1 - y_c) \log(1 - \hat{y}_c)$$

$$\mathcal{L}_o = (y_o - \hat{y}_o)^2$$

$$\mathcal{L}_h = \left(\frac{y_c}{y_c + \epsilon}\right) (y_h - \hat{y}_h)^2$$

Parks+17

CNN Training

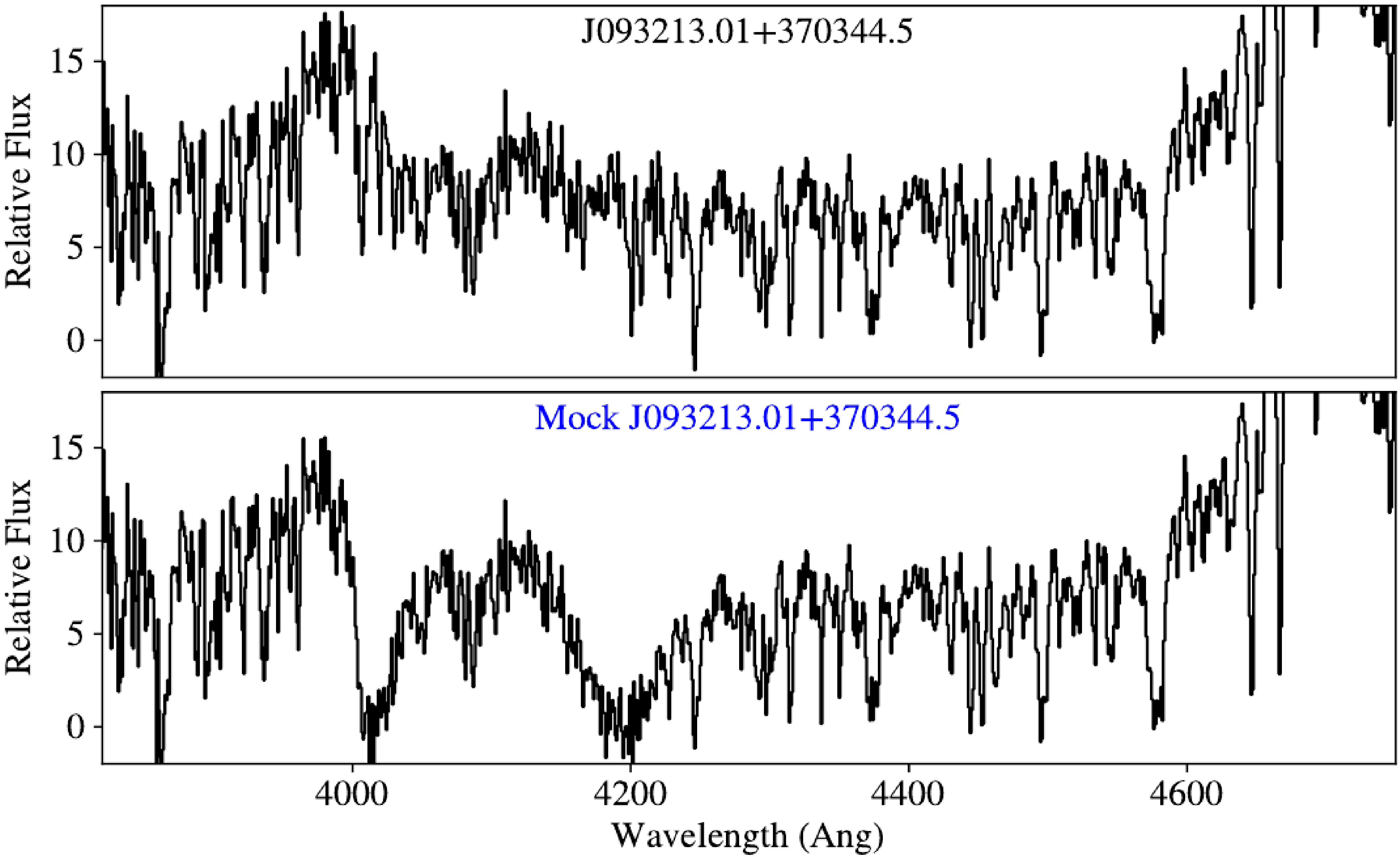


200,000 sightlines of DLAs injected into 'DLA-free' quasar spectra from the SDSS-DR5.

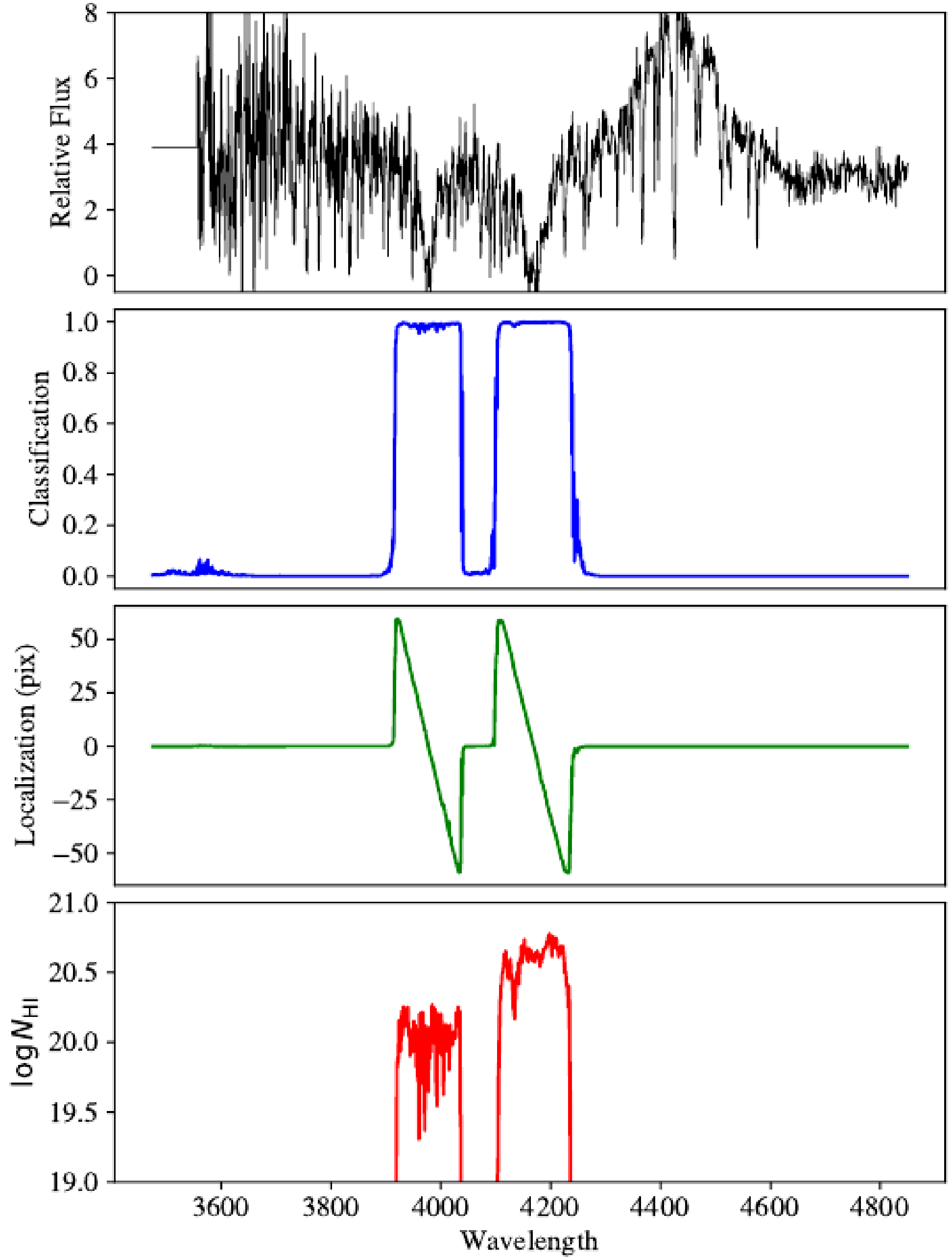
By-hand addition of additional training sets: high N_{HI}, SLLS

Parks+17

CNN Training

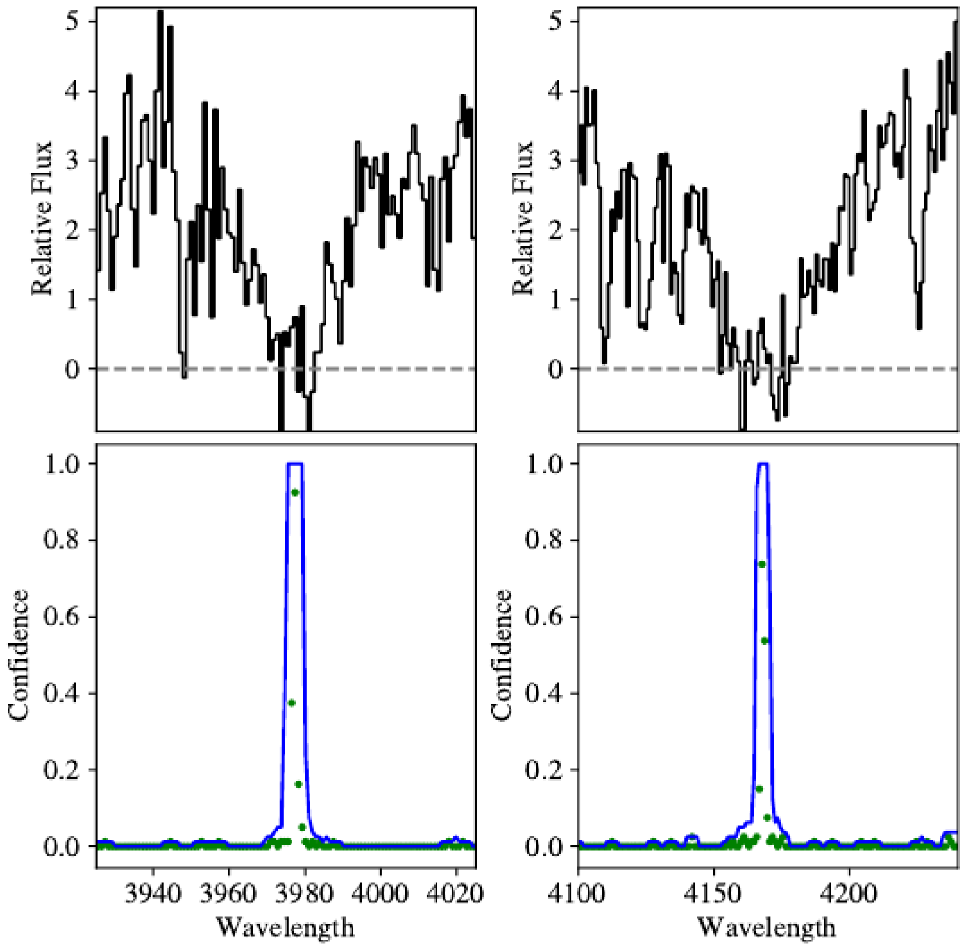


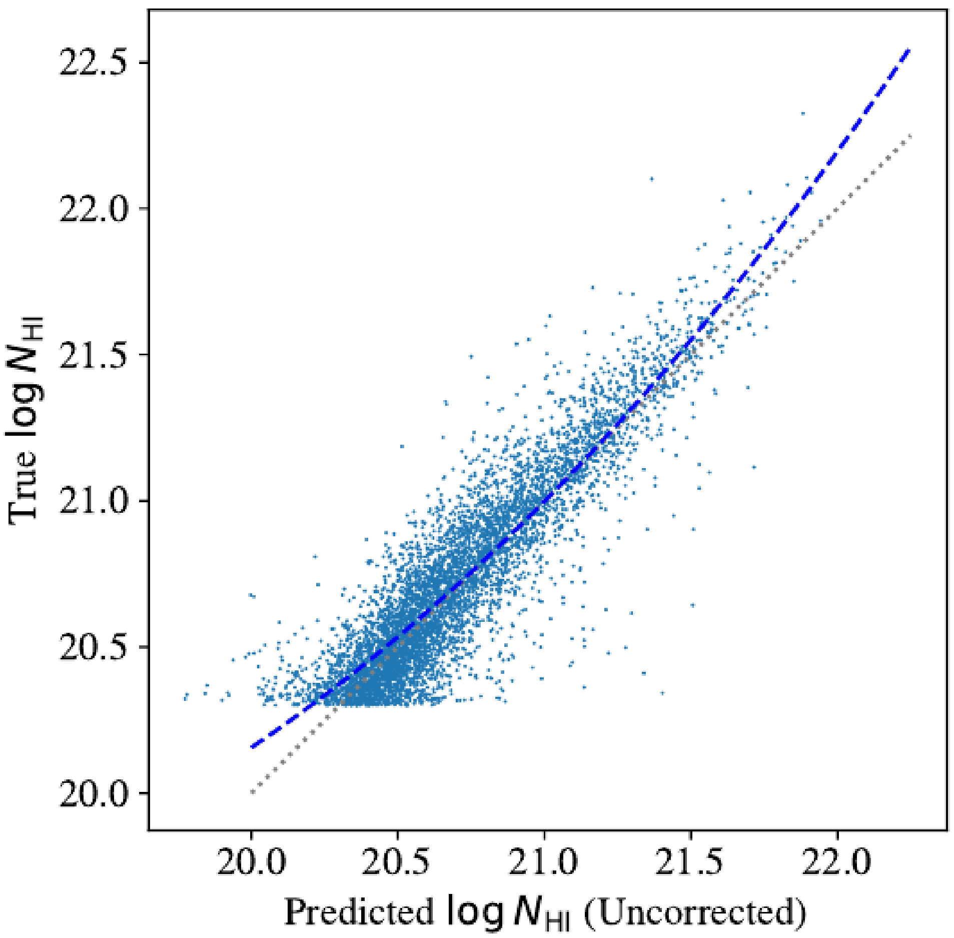
CNN Validation



Parks+17

CNN Validation

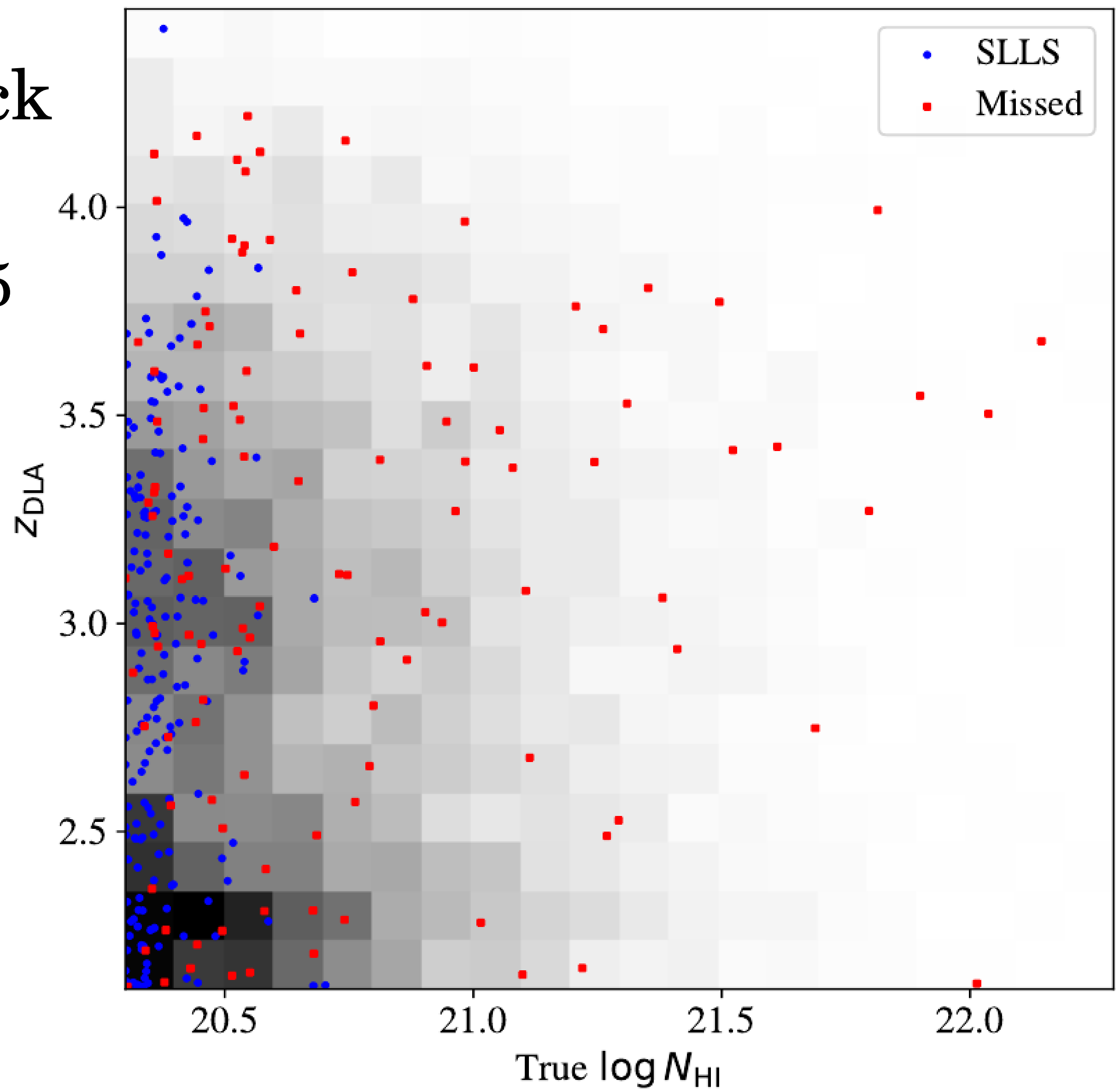




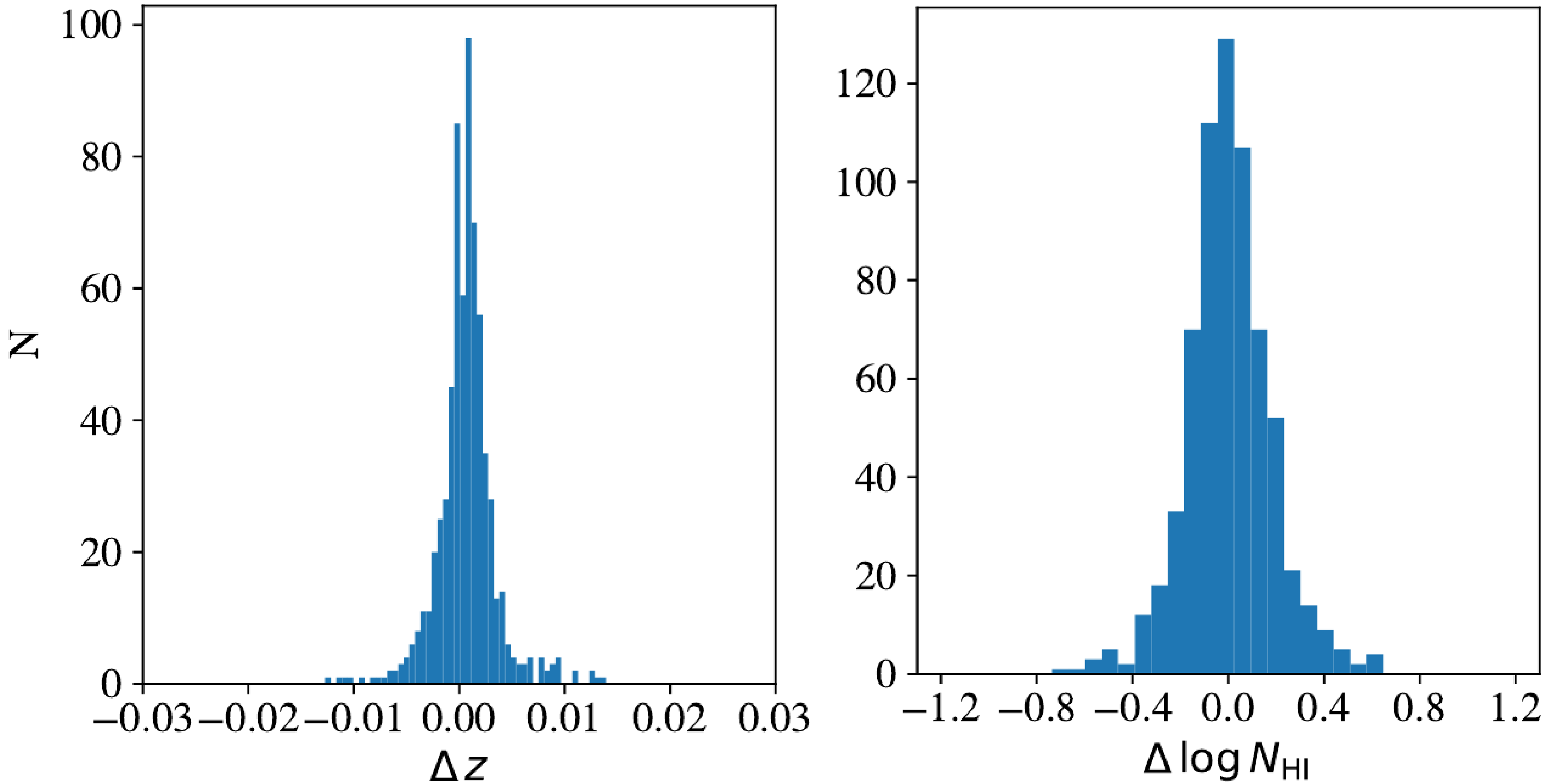
Parks+17

CNN Validation

- 11,110 DLAs held back
  - Matched 99% of these to within |dz| < 0.015
- ~300 false negatives
  - Primarily SLLS
  - Overlapping DLAs
- 74 ‘false positives’
  - Some may be real
  - Overlapping DLAs



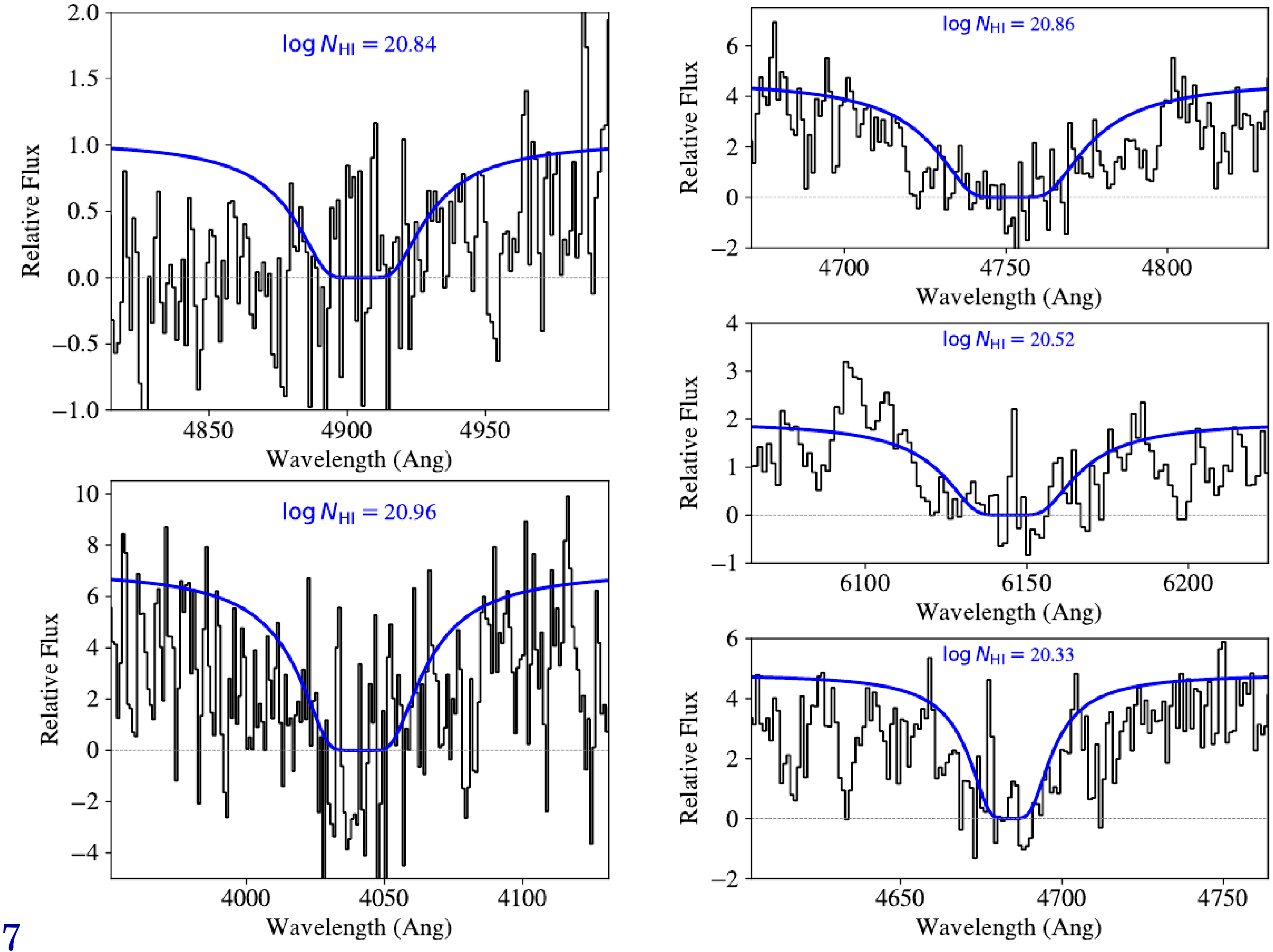
CNN Validation



Parks+17

Precise DLA measurements **without** Quantum Mechanics!!

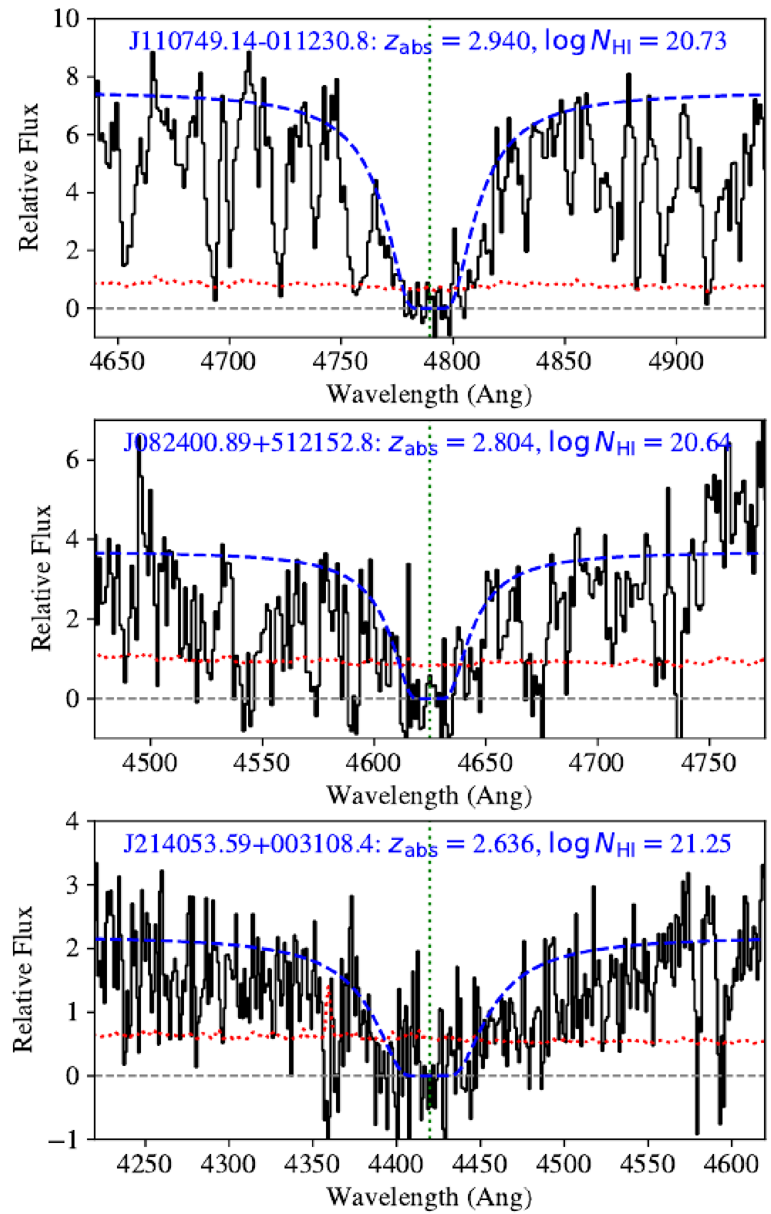
CNN Validation



Parks+17

'Fun' examples for DLA officiandos

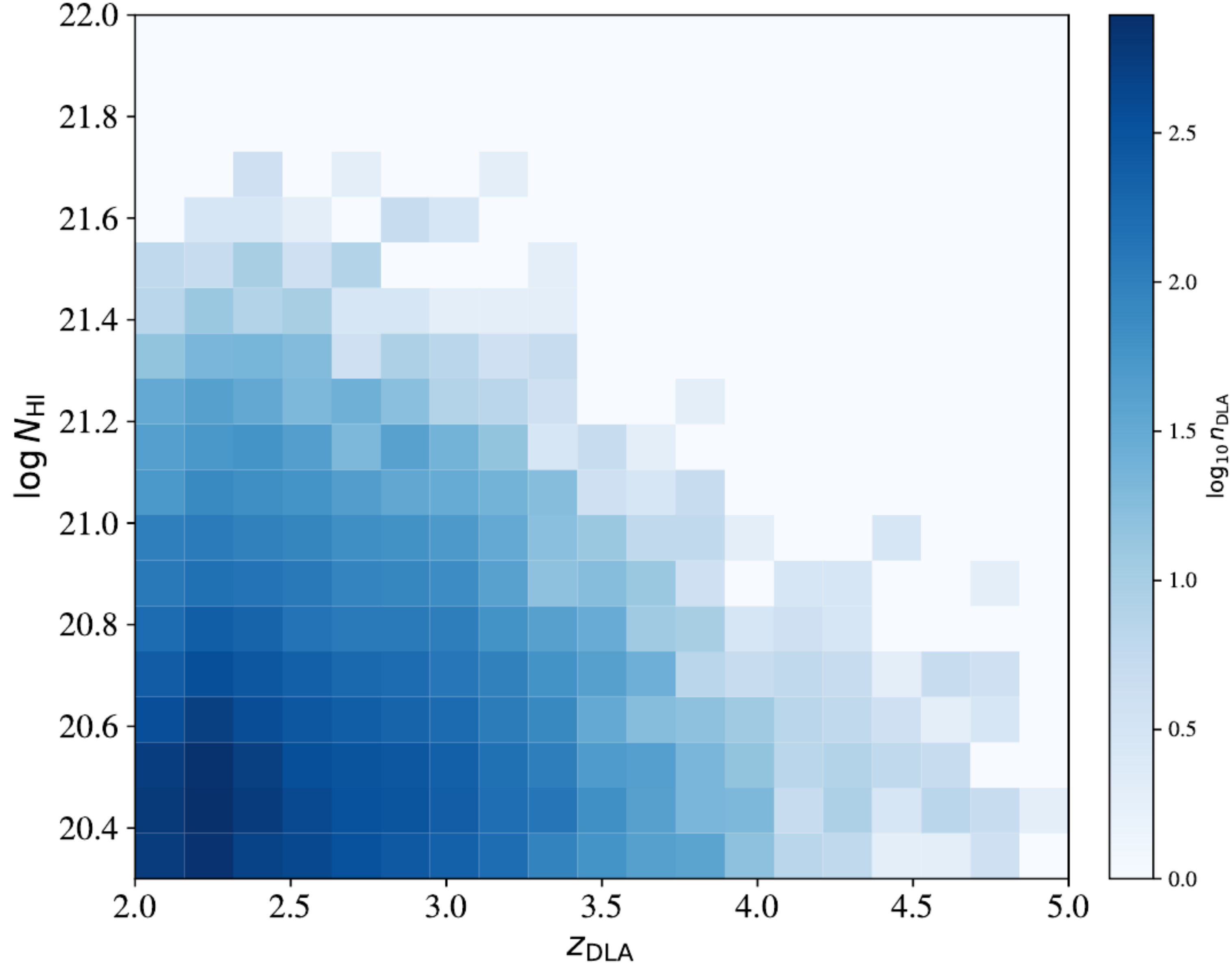
DLA Results (CNN + SDSS-DR7)



Parks+17

1,659 DLAs in DR7 not reported previously (~2x gain).

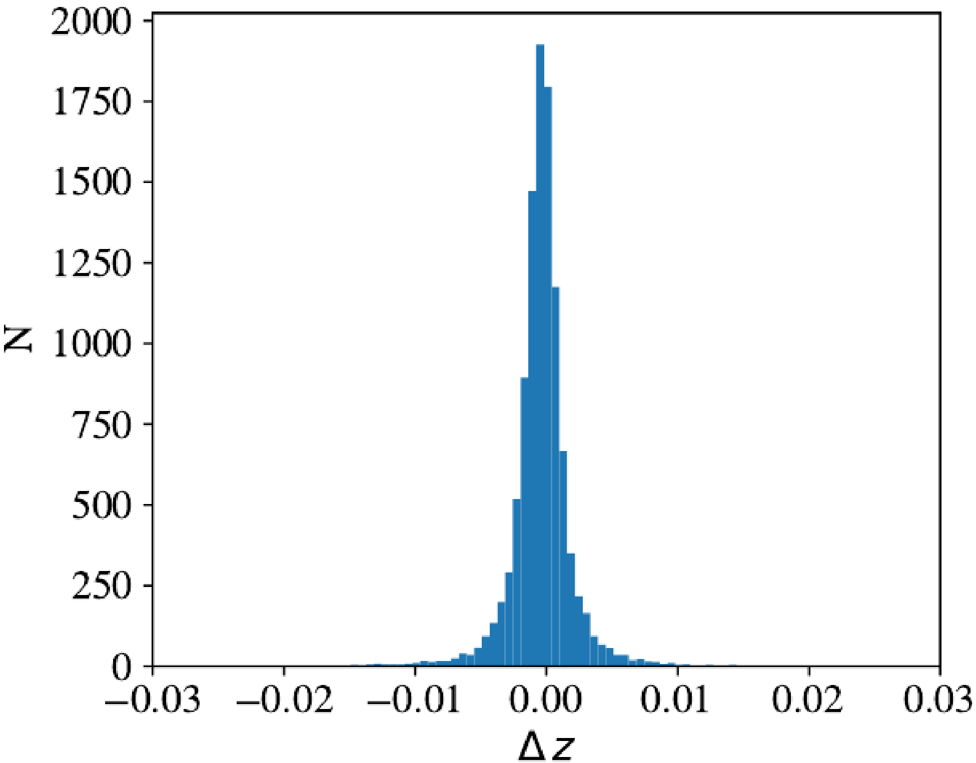
DLA Results (CNN + BOSS)

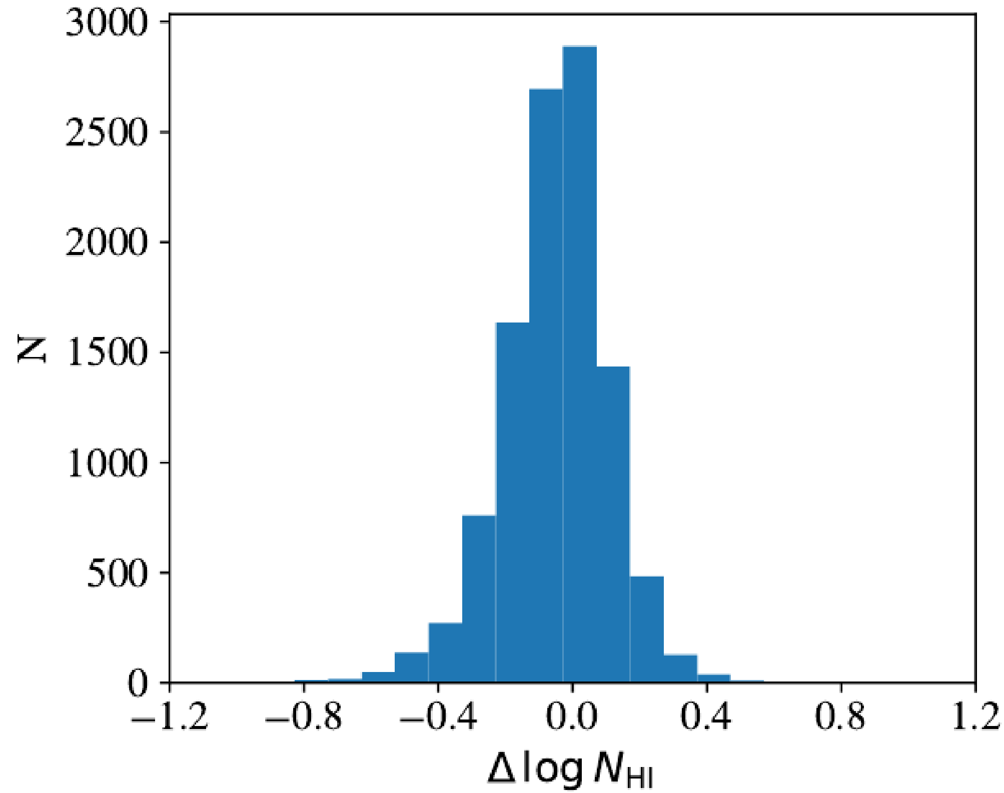


Parks+17

~19,000 DLAs with z>2

DLA Results (CNN/GP + BOSS)

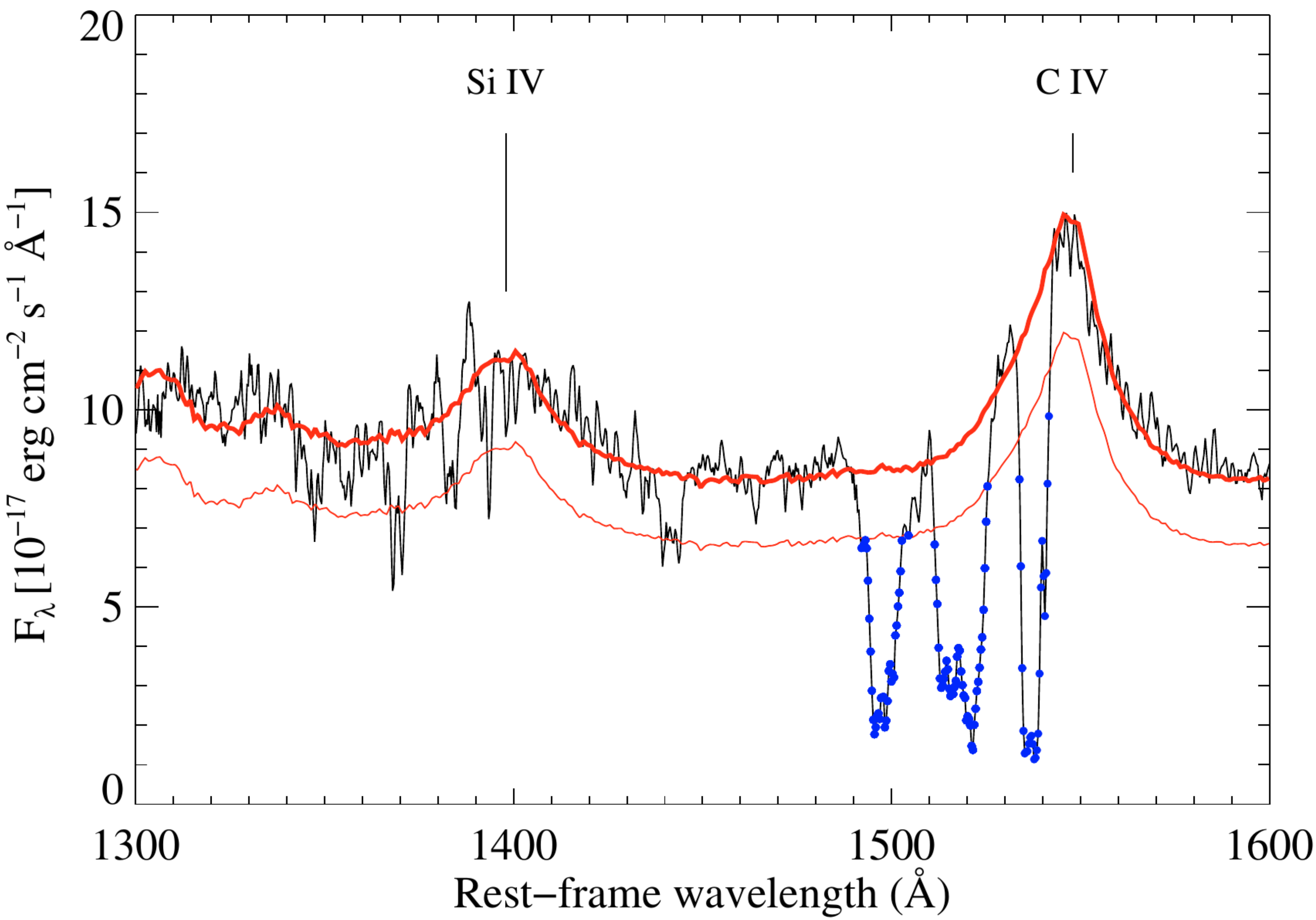




Excellent complementarity between GP and CNN

Parks+17

Future Ideas: BALs

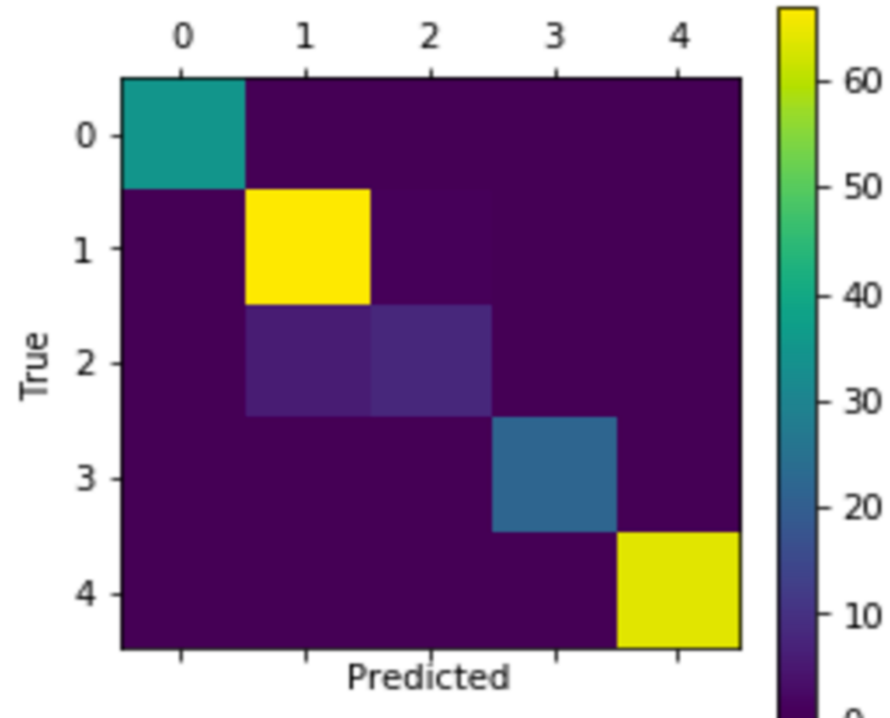


Jiang et al., in prep.

I am advised to use RNNs.

Future: Spectral Image Classification

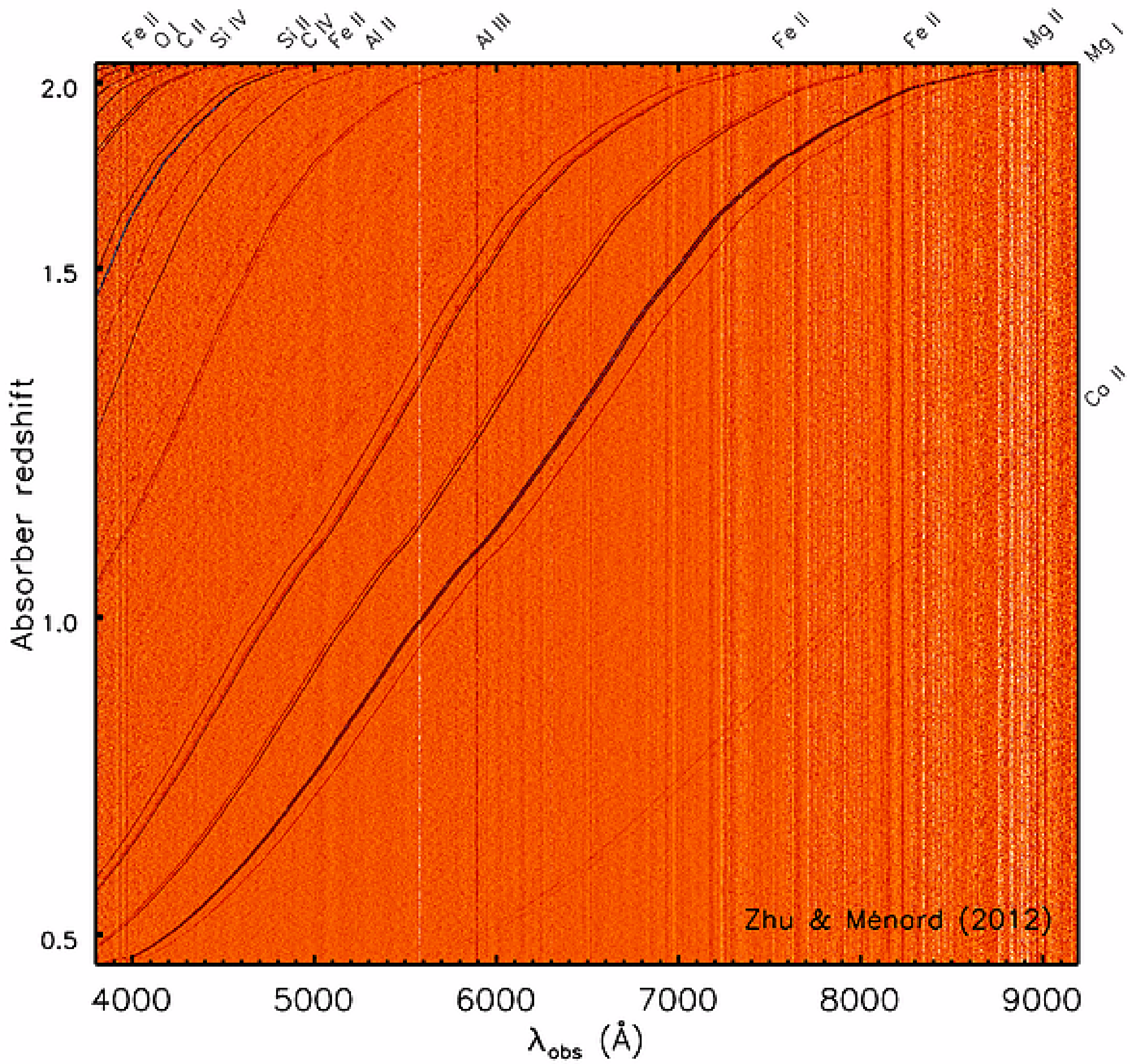
```
Confusion Matrix:
[[35  0  0  0  0]
 [ 0 67  1  0  0]
 [ 0  6  8  0  0]
 [ 0  0  0 22  0]
 [ 0  0  0  0 64]]
```



Yankoff+17

Done. Just need to write it up and implement

Future: Other ideas...



??+17

Lots of things brewing

Machine Learning in Astronomy

- Astronomy is rife with tasks demanding human labor
  - Source identification
  - Continuum fitting
  - Line identification
- Machine Learning
  - Can perform many of these tasks
  - Auto-magically, repeatably, better!
- Astrophysics and ML
  - I harbor my doubts…