Neural Network Reconstruction of Physical Quantities From the Lyman-α Forest

Student: Runxuan(Max) Wang

Instructor: Prof. Rupert Croft

Physics Carnegie Mellon University

Carnegie Mellon University

Abstract

The Lyman-α forest is a collection of absorption lines by neutral hydrogen in the spectrum from a distant light source. The wavelengths of these absorption can be mapped to the distance along the line of sight. We aim to reconstruct the optical depth, density, and temperature of intergalactic medium along a line of sight from the observed flux spectrum in the Lyman-α forest using Convolutional Neural Network. Neural Network models were trained on noised mock observational data derived from cosmological hydrodynamic simulations with various redshifts (z=3, 2.5, 2). Currently, the models reconstructing optical depth and density have not outperformed the naive numerical methods. Yet, the model reconstructing temperature is able to exploit the thermal broadening of the spectrum to detect spatial inhomogeneities without any general density-temperature relation that the naive method must assume. The ongoing work aims to further improve the accuracy and the generality of these Neural Networks.

Data

A hydrodynamic cosmological simulation of the ΛCDM model was used. The parameters are h = 0.702, Ω_Λ = 0.725, Ω_m = 0.275, Ω_b = 0.046, n_s = 0.968 and σ_8 = 0.82. The mass per particle is 1.19 × 10^7 $h^{-1}M_\odot$ (gas) and 5.92 × 10^7 $h^{-1}M_\odot$ (dark matter). A cube of side length 400 h^{-1}Mpc has 65536 = 256^2 evenly spaced sightlines. The simulation outputs optical depth in real space τ_{real}, peculiar velocity v, density ρ, temperature T, etc. The optical depth in redshift space τ_{red} is obtained by convolving τ_{real} and v. The observed fluxes are then calculated by

$$F = e^{-\tau_{red}}$$

- **Low Resolution Data**

Gaussian noise with standard deviation $\sigma_N = 1/5\,\sigma_F$ is added to F, where σ_F is the root mean square of the flux spectrum. 512 pixels are sampled out of the original 4096 per sightline, each spanning 83.5 km/s. For large structure information, the spectrum is further smoothed by a 1D gaussian filter of kernel 6 pixels.

- **High Resolution Data**

High resolution data contains all 4096 pixels, each spanning 10.45 km/s. Only 20,000 sightlines are trained on due to limited GPU memory. With this high resolution, the thermal broadening effect becomes significant enough to spread the flux at each pixel by a Gaussian with

$$\sigma \propto \sqrt{T}$$

The simulation gives a homogeneous power-law relation between density and temperature. As shown in Fig A,

$$T = B\,(\frac{\rho}{\bar{\rho}})^k$$

For the z = 3 data set, $B = 11582$ K, $k = -0.538$.

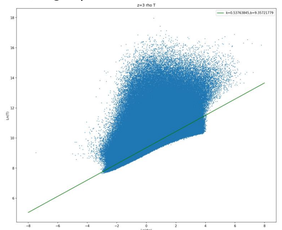
To account for more diverse and inhomogeneous ρ-T relations in real observations, we imposed the temperature for three additional cases:

Case 1. Impose a constant $T = 20{,}000\,K$ within two quasar dominated regions per sightline, each with width 50 Mpc/h and a random position. T at other pixels are the same as the simulation.

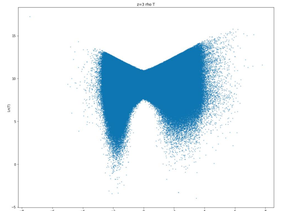
Case 2. Same as case 1, except that the T is randomly determined between $50{,}000\,K$ and $150{,}000\,K$.

Case 3. For every sightline, impose a random ρ-T relation with $B \sim Uniform(3000, 40000)$, and $k \sim Uniform(-1,1)$. Fluctuations are added as a Gaussian centered at 0 with $\sigma = B/8$, with overall ρ-T relation shown in Fig B.

↓ Fig A. ρ-T Relation from Simulation

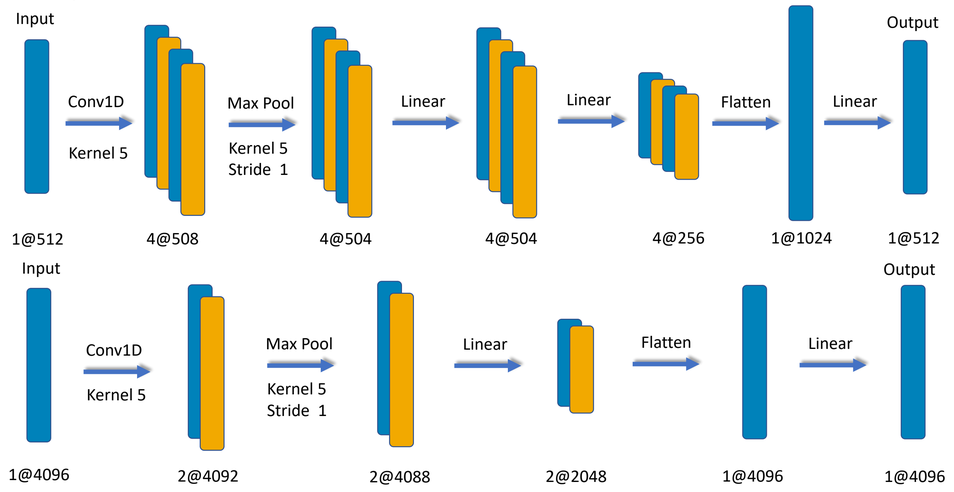


↓ Fig B. An Imposed ρ-T Relation



Models

Two architectures used are visualized in Fig C below. The top one is for reconstructing τ and ρ and the bottom one is for T. SeLU is used as the activation function for the convolution layer and the linear layers.



↑ Fig C. NN Architectures

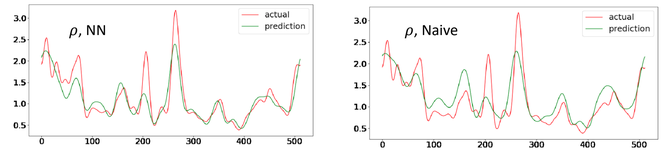
Current Results and Discussions

The naive method to reconstruct density, referred below as "naive", is to first reconstruct $\tau = -\log(F_s)$, where F_s is the flux smoothed by a Gaussian filter. Next, fit $\rho = a\tau^3 + b\tau^2 + c\tau + d$. Table A shows that, for optical depth and density, the Neural Network currently cannot extract more information than the naive method. The main challenge is that naturally the data size of low-value pixels is much larger than that of high-value ones. This unbalance in the training set makes the Neural Net always underpredict ρ and τ high-value regions. Example sightlines plotting predictions versus actual values for ρ and τ are shown in Fig D.

RMSE τ	Overall	$\tau < 2$	$\tau > 2$
NN	0.49	0.15	1.53
Naive	0.54	0.14	1.70

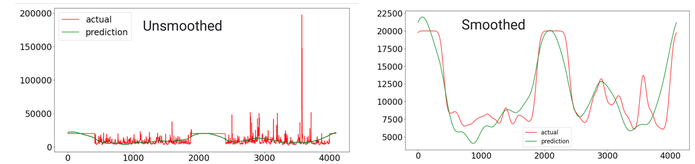
RMSE ρ	Overall	$\rho < 2$	$\rho > 2$
NN	0.34	0.18	1.55
Naive	0.34	0.19	1.52

↑ Table A. Root-Mean-Square Errors for τ and ρ



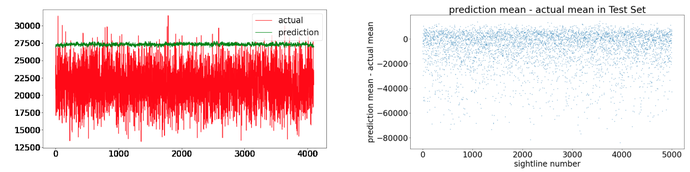
↑ Fig D. Predicted and Actual ρ for an Example Sightline

For high resolution data Case 1 and 2, the Neural Net successfully reconstruct the quasar-dominated high temperature regions from the smoothing effect due to thermal broadening. No simple numerical method can achieve this since the positions and/or T in these regions are randomly generated. Fig E shows the Neural Net prediction for a sightline in Case 1.



↑ Fig E. Predicted and Actual T for an Example Sightline; Case 1

For Case 3, currently the Neural Net can only provide a flat spectrum due to limited architecture complexity, an example sightline and the bias on the mean for 5000 sightlines in the test set are shown in Fig F below. The average bias on the mean is -7217 K.



↑ Fig F. T Prediction and Prediction Bias; Case 3

Future Work

Future work will focus on:

1. to further evaluate and improve the generality of the Neural Net's ability to recognize the smoothing effect on the flux.
2. to include pressure broadening effect.
3. to improve the accuracy of high-value regions for ρ and τ.

References

[1] Huang, Lawrence, Rupert A. C Croft, and Hitesh Arora. "Deep Forest: Neural Network Reconstruction of the Lyman-Alpha Forest." (2020): n. pag. Print.