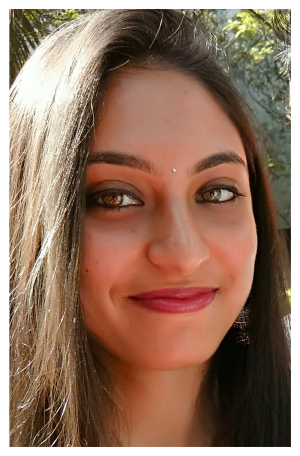
Automated Physics-based Detection and Identification of Intergalactic clouds using Probabilistic Programming

Rashmeet Kaur Nayyar[1], Mansi Padave[2], Sanchayeeta Borthakur[2], Siddharth Srivastava[1]
[1]School of Computing, Informatics, and Decision Systems Engineering, Arizona State University
[2]School of Earth and Space Exploration, Arizona State University



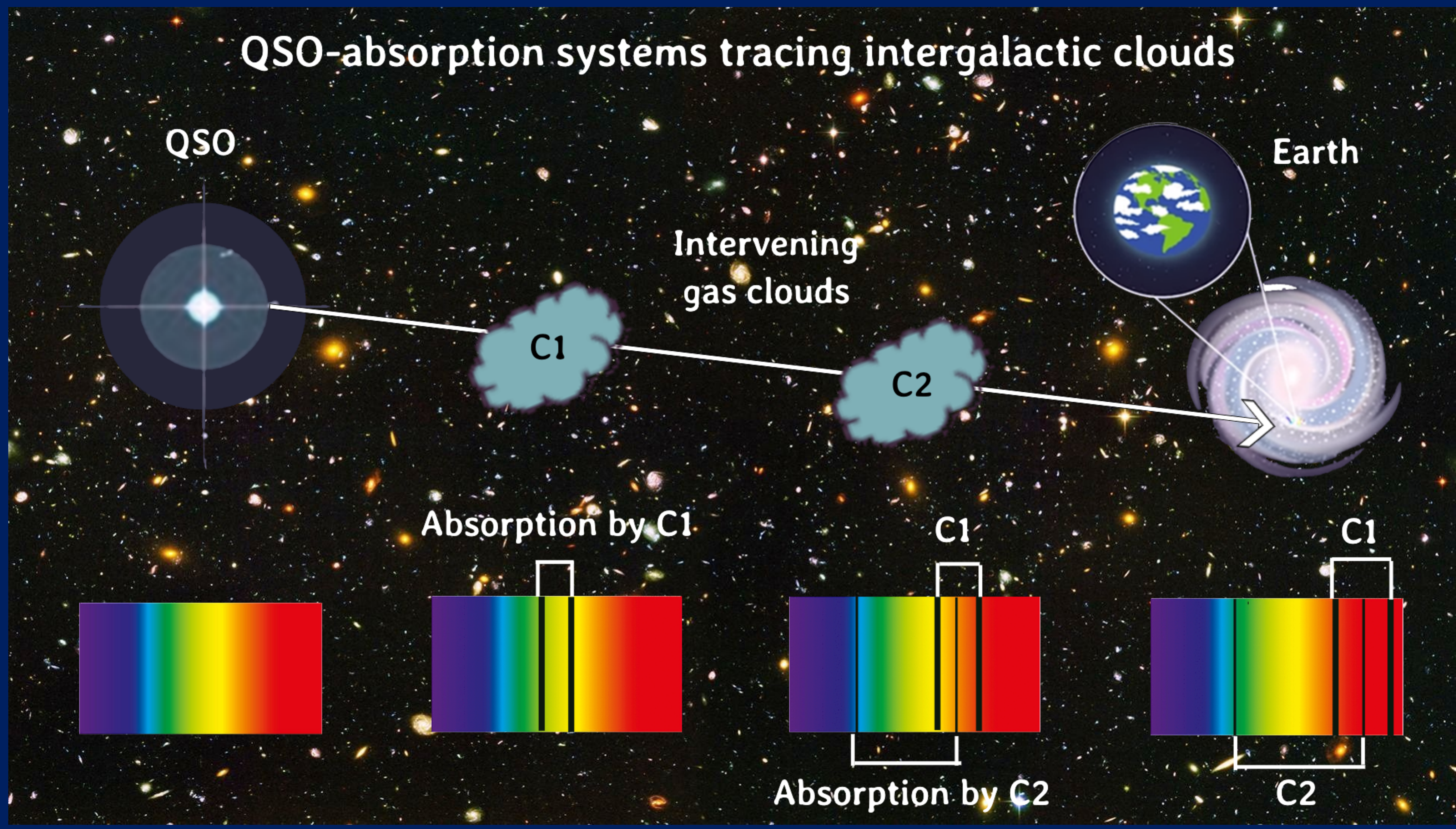
1. Objective

To develop an automated system that identifies the intergalactic clouds in QSO spectra.

We are testing our system on UV spectra obtained with the Cosmic Origins Spectrograph aboard the Hubble Space Telescope.

2. Challenges

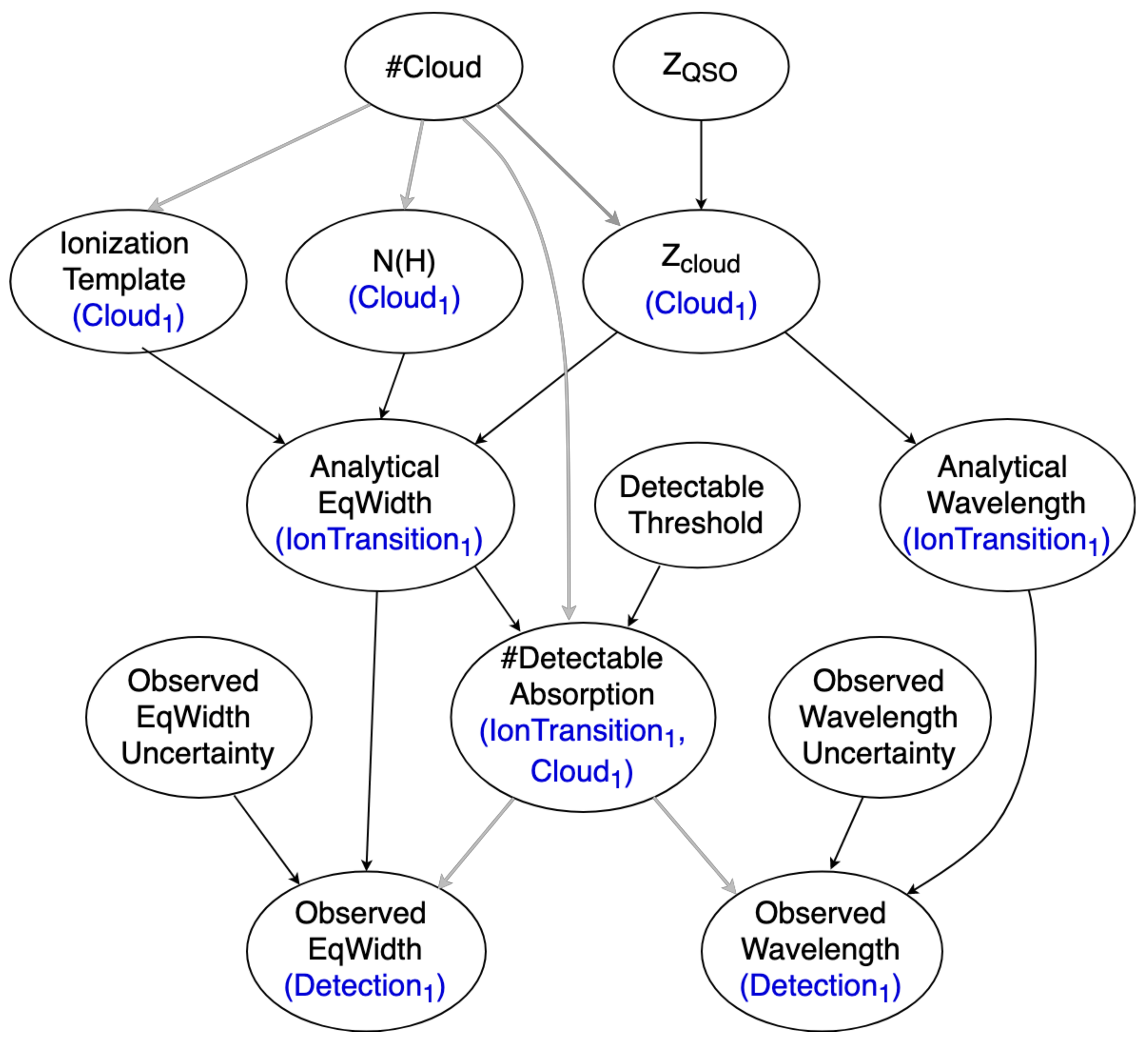
- ❖ Redshifts of absorption lines are unknown.
- ❖ Approaches based on probabilistic inference typically require prior knowledge of the number of clouds which are unknown.
- ❖ Observational uncertainties make it difficult to use combinatorial search techniques.



3. Advantages Of Our System

- ❖ Automated detection and characterization of matter in the intergalactic medium from absorption spectra
- ❖ Reliable and reproducible
- ❖ Scalable to large datasets that are expected in the next decade (e.g. observations with GMT, TMT, ELT etc.)

Contingent Bayesian Network for our Probabilistic BLOG Model



4. Our Approach

- ❖ Uses **Bayesian Logic (BLOG)** [4] [5] [6], a probabilistic programming language that extends first-order logic semantics with probability theory and allows efficient specifications for physics-based probabilistic models.
- ❖ Accounts for the **observational uncertainties** as well as the **unknown numbers and types of intervening clouds**.
- ❖ Utilizes a rich query language for investigating the properties and evaluates this representational approach using **techniques for approximate probabilistic inference.**
- ❖ Uses collisional ionization templates [3] and published values for line-transitions (e.g. wavelength and f-value) [1].

Such an automated system for species identification in QSO-absorption lines does not exist and our system complements the automated system by Danforth et al. [2].

5. Results

We tested our approach with simulated data from one intervening gas cloud. In this test, we provided the following observations:

- ❖ there are 3 detections in the spectra between 1220 Å and 1470 Å, at 1458.80 Å, 1238.31 Å, and 1245.14 Å
- ❖ redshift of the background QSO, $z_{QSO} = 0.36$

Our system correctly identified the *species associated with the detected absorption lines* and in the process predicted the correct values for the *number of clouds*, and their *redshifts and column densities.*

Property	Correct Value	Computed Estimate (MPE)	Predicted Probability
#Cloud	1	1	0.90
Z_{cloud}	0.20	0.19 - 0.21	1.00
Ionization Template	CIE at 10^6 K	CIE at 10^6 K	0.90
N(H)	1e20	1e19 - 1e21	0.82

Ion transitions	Correct Value	Computed Estimate (MPE)	Predicted Probability
H I - 972.00	False	False	0.83
H I - 1025.00	False	False	0.97
H I - 1215.67	True	True	1.00
C II - 1036.34	False	False	1.00
C II - 1334.53	False	False	1.00
C IV - 1548.20	False	False	0.58
C IV - 1550.78	False	False	0.79
N V - 1238.82	False	False	0.69
N V - 1242.80	False	False	0.82
O VI - 1031.93	True	True	1.00
O VI - 1037.62	True	True	1.00
SI II - 1190.42	False	False	1.00
SI II - 1193.29	False	False	0.99
SI IV - 1206.42	False	False	1.00

6. Conclusion And Future Work

- ❖ Preliminary results show that this approach has the potential to be able to identify the sources of absorptions and characterize their properties. We were able to get good estimates for all the properties for a single cloud.

Next Steps:

- ❖ Investigate and evaluate more efficient methods for probabilistic inference, e.g., the Metropolis-Hastings algorithm.
- ❖ Extend the system to simulated and real spectra with larger numbers of intervening clouds.

References

[1] Morton et al., 1988, ApJS, 68, 449 .
[2] Danforth et al., 2016, ApJ, 817, 111.
[3] Gnat & Sternberg, 2007, ApJS, 168, 213.
[4] Milch, B., Marthi, B., Russell, S. J., Sontag, D., Ong, D. L., & Kolobov, A. BLOG: Probabilistic models with unknown objects. IJCAI, pp. 1352–1359, 2005a.
[5] Milch, Brian, and Stuart Russell. "General-Purpose MCMC inference over relational structures." *Proceedings of the Twenty-Second Conference on Uncertainty in Artificial Intelligence*. AUAI Press, 2006.
[6] Wu, Y., Srivastava, S., Hay, N., Du, S. S., & Russell, S. "Discrete-Continuous Mixtures in Probabilistic Programming: Generalized Semantics and Inference Algorithms.", 2018.