JATIS Journal of Astronomical Telescopes, Instruments, and Systems

RESEARCH PAPER

Recovering the O VII absorption distributions from X-ray data

Nichole Gray, Cameron T. Pratt, and Joel N. Bregman*
University of Michigan, Department of Astronomy, Ann Arbor, Michigan, United States

ABSTRACT. The absorption by gas toward background continuum sources informs us about the cosmic density of gas components as well as the hosts responsible for the absorption (galaxies, clusters, and cosmic filaments). Cosmic absorption line distributions are distorted near the detection threshold (S/N $\approx$ 3) due to true lines being scattered to a lower signal-to-noise (S/N) and false detections occurring at the same S/N. We simulate absorption line distributions in the presence of noise and consider two models for recovery: a parametric fitting of the noise plus a cut-off power law absorption line distribution and a non-parametric fit in which the negative absorption line distribution (emission lines) is subtracted from the positive S/N absorption line distribution (flip and subtract). We show that both approaches work equally well and can use data with S/N $\gtrsim$ 3 to constrain the fit. For an input of about 100 absorption line systems, the number of systems is recovered to $\approx$14%. This investigation examined the O VII X-ray absorption line distribution, but the approach should be broadly applicable for statistically well-behaved data.

Keywords: absorption lines; spectroscopy; X-rays; signal extraction

Paper 23049G received Apr. 21, 2023; revised Jul. 26, 2023; accepted Jul. 28, 2023; published Aug. 11, 2023.

1 Introduction

Approximately 90% of the baryons in the Universe remain in gaseous form,[1,2] with about 35% to 50% not yet detected with the usual absorption signatures of neutral and warm atomic gas. This remaining gas is predicted to lie in a hotter state ($T > 5 \times 10^5$ K) that is produced by the gravitational collapse of filaments, galaxies, and galaxy clusters (e.g., Refs. 3 and 4). For example, extended hot halos around galaxies are expected to have temperatures near the virial value, about 2×10^6 K for the Milky Way. Within these collapsed structures, star formation occurs, leading to the production of metals and their dispersal, typically as highly ionized outflows (e.g., Ref. 5).

Observational studies have focused on the high ionization ions of oxygen, in which lithium-like O VI has a peak ionization fraction at $10^{5.3}$ K and is accessible in the far ultraviolet (UV) (rest wavelengths of 1032, 1038 Å for the doublet). Many UV studies have been conducted with the Hubble Space Telescope and the Far Ultraviolet Explorer, showing that this ion is common in the Milky Way, halos of external galaxies, and the intergalactic medium (e.g., Refs. 6–9). The higher ionization ion of helium-like O VII has an ionization fraction near unity from $0.4 - 2 \times 10^6$ K, whereas for the hydrogenic form, O VIII is the dominant ion for $T > 2 \times 10^6$ K. Exploration of these ions is only possible in the X-ray regions as the strongest resonance lines fall at 0.574 keV (O VII) and 0.654 keV (O VIII).

*Address all correspondence to Joel N. Bregman, jbregman@umich.edu

Both O VII and O VIII have been detected in the Milky Way, using the grating spectrographs on Chandra (LETG) and XMM-Newton (RGS) and have been found to have column densities of more than an order of magnitude greater than O VI.[10–12] Extragalactic detection of O VII has been much more challenging, in that the equivalent widths (EWs) are expected to be several times smaller than those measured for the Milky Way. Currently, there is a singular detection, but with significant uncertainties,[13] and it will not be able to extend this effort to many sightlines, which is needed to assemble a statistically useful sample of O VII and O VIII absorption lines. This goal can be achieved if new X-ray observatories are built with an improved collecting area and spectral resolution.

Two observatories that would have the necessary capabilities are under development — Athena (ESA; Ref. 14) and Arcus (NASA[15]). Each would detect these oxygen lines against X-ray bright active galactic nucleus (AGNs) at redshifts suitable to sample a useful amount of physical space so that $\sim 10^2$ absorption systems are detected.[16] Such studies will yield the number of absorption line systems as a function of their column density (N) or EW, which provide fundamental information, such as the ratio of the mean ion densities to the critical density of the Universe, expressed as Ω (O VII) and Ω (O VIII). There are several steps to obtaining these quantities from the data, such as determining the shape of the EW distribution to the smallest feasible EWs.

Here we investigate the most effective approach to determining the shape of this distribution. Toward this purpose, we adopt the observational characteristics of an Arcus-type mission, in which a spectral resolution of 3000 to 4500 leads to a resolution element (65 to 100 km s^{-1}; 5 to 7 mÅ at $z = 0$ for the O VII Heα 21.60 Å resonance line) similar to the expected width of an absorption line (45 to 100 km s^{-1}). However, the result is rather general and can be applied to Athena or to UV absorption line studies, provided that the spectra are well-behaved and obey the assumed Gaussian statistical characterization.

An issue that will need to be considered is contamination by galactic absorption lines in the 10 to 40 Å range, which can be estimated from theoretical models or observations toward bright galactic X-ray sources (e.g., Refs. 17–21). These lines include the standard resonance lines, plus a number of inner shell lines, which are less well modeled due to the complexity of transitions from the inner shell. These inner shell lines are mostly due to oxygen, and in particular O I, O II, with contributions from O III - O VI, as well as similar lines from N, Ne, and Fe. The inner shell oxygen lines lie longward of the O VII Heα 21.60 Å line, so they can be confused with a redshifted O VII 21.60 Å feature. There are 15 inner shell lines of consequence in the 21.60 to 23.55 Å range, which if mistaken for redshifted O VII Heα, correspond to a redshift range of $0 < z < 0.09$. The median separation between the lines is about 1.6 resolution elements for a spectral resolution of 250 (typical of a good calorimeter in this energy range, such as found in Athena), so potential IGM lines are blended with Galactic lines, making line extraction difficult if not impossible.[21] A few other redshift regions would be problematic, such as at $0.30 < z < 0.37$ and $z \approx 0.44$.

Soft energy grating spectrometers can have much higher spectral resolution, which mitigates this issue, such as for Arcus. Arcus[22] is likely to have a resolution of ≈3500, leading to a median of 23 resolution elements between contaminating lines. This will lead to a modest decrease in the redshift search space in $0 < z < 0.09$ of about 5%. Therefore, we ignore this effect for the purposes of this work and concentrate on the recovery of the line distribution.

The remainder of the paper is structured in the following way: Sec. 2 gives some motivation for our work; Sec. 3 describes the methodology, specifically focusing on simulating a toy model and attempting to recover the input distribution; Sec. 4 presents our results and a discussion on the implications of this work and how it can be used in future studies; and Sec. 5 summarizes the main points of this paper.

2 Motivation

The scientific problem that motivated this study was the ability to detect the hot gas component of the Universe using highly ionized metals, such as O VII. The cosmological distribution of O VII absorbers as a function of EW has been calculated from numerical simulations by several authors.[23–27] There is remarkable agreement between studies, with a similar number of absorbers per unit redshift (at low z) for N(O VII) $>10^{15}$ cm^{-2}. Here, we used the work of Ref. 25, in which

the distribution was calculated in the column density range of 10^{13} cm^{-2} < N(O VII) < 2×10^{16} cm^{-2} (the summation that includes low temperature O VII). The distribution is convergent at both the high and low ends of the column density distribution, so to measure the total amount of N(O VII) present, one must be sensitive to sufficiently low absorption columns (we assume that the highest columns are easily measured). This can be investigated using the distribution of Ref. 25, in which the space density per unit redshift is above some N or EW. Here, we use EW instead of N, which for optically thin O VII is EW = 2.86 mÅ (N/10^{15} cm^{-2}). We approximate this distribution as

$$\mathrm{dn}(>\mathrm{EW})/\mathrm{dz} = 132 \times \exp(-\mathrm{EW}^{0.316}), \tag{1}$$

where the EW is in units of mÅ and the unit of dn(> EW)/dz is the number of absorption systems per unit redshift with an EW above some limiting value.

We prefer to work with the differential form of Eq. (1), which is the differential number of absorption systems per dEW, $\mathrm{d}^2\mathrm{n}/\mathrm{dzdEW}$. From this, we calculate the physically meaningful total integrated EW per unit redshift in the interval of a high and low EW, $\mathrm{dEW}(\mathrm{EW}_{\mathrm{min}}, \infty)/\mathrm{dz} = \int \mathrm{EW}(\mathrm{d}^2 n/\mathrm{dz}\,\mathrm{dEW})\mathrm{dEW}$. This integral is evaluated from $\mathrm{EW}_{\mathrm{min}}$ to infinity, where $\mathrm{EW}_{\mathrm{min}}$ corresponds to the limiting EW for which the underlying distribution can be extracted.

For a particular model, this quantity $\mathrm{dEW}(\mathrm{EW}_{\mathrm{min}}, \infty)/\mathrm{dz}$ can be corrected to the value $\mathrm{dEW}(0, \infty)/\mathrm{dz}$; we calculate and show this ratio, $\mathrm{dEW}(\mathrm{EW}_{\mathrm{min}}, \infty)/\mathrm{dz}/\mathrm{dEW}(0, \infty)/\mathrm{dz}$, in Fig. 1. The quantity $\mathrm{dEW}(0, \infty)/\mathrm{dz}$ for O VII is proportional to the cosmologically important Ω (O VII).

Our study was motivated by determining $\mathrm{EW}_{\mathrm{min}}$ and the ratio of $\mathrm{dEW}(\mathrm{EW}_{\mathrm{min}}, \infty)/\mathrm{dz}/\mathrm{dEW}(0, \infty)/\mathrm{dz}$ from the observational goals of particular future X-ray missions. Rather than using the derivative of Eq. (1) as the input, we use a simpler form as the precise form varies between investigators.

The Arcus mission has a plan to observe ~30 AGN to detect ~100 projected absorption line systems, probing a redshift space of $\Delta z \approx 15$;[15,22] Athena has similar goals.[14] There will be about 30,000 total spectral resolution elements examined in the O VII study for the above search space, with the Arcus resolution of 3000 to 4500 and a mean uncertainty per spectral resolution element of 0.75 mÅ. Nearly all of these resolution elements are expected to be null detections, with

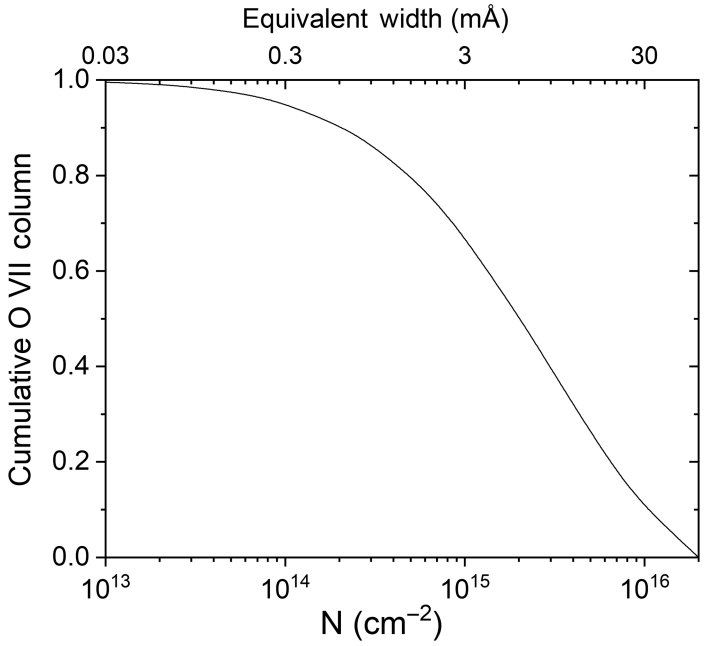

Fig. 1 Cumulative O VII column (or mass) as a function of the minimum limiting column of a study; the maximum column is 2×10^{16} cm^{-2} (60 mÅ). It shows that the cumulative fractions are 50% at 2×10^{15} cm^{-2} (6 mÅ), 75% at 6.4×10^{14} cm^{-2} (1.9 mÅ), and 90% at 2.1×10^{14} cm^{-2} (0.63 mÅ).

Gaussian noise causing each resolution to have a positive (absorption) or negative (spurious emission) value relative to the continuum. Thus, the challenge is to extract the underlying signal or distribution of absorption line systems from the ensemble of values.

There are a number of issues that will arise in the actual data from a mission, such as the uncertainty being wavelength dependent, which we ignore in this work. Also, the EW of an absorption line is modified by its redshift, an issue that we do not consider, for the sake of simplicity. These considerations can easily be dealt with in future studies.

3 Methods

3.1 Simulated Data

The main goal of our study is to determine the degree to which a simulated EW distribution can be recovered. We chose a distribution that captures the characteristics of the numerical simulation of Ref. 25, modified for simplicity. We work with the differential form of the EW distribution, $\mathrm{d}^2\mathrm{n}/\mathrm{dzdEW}$, where EW is the equivalent width, expressed in mÅ. It is given as

$$\mathrm{d}^2\mathrm{n}/\mathrm{dzdEW} = A_p\left(\frac{111}{1+\mathrm{EW}^m}\right), \tag{2}$$

where A_p is the normalization and m is the power law index. The flattening of $\mathrm{d}^2\mathrm{n}/\mathrm{dzdEW}$ at small EWs is somewhat arbitrary but does not significantly affect the determination of the two parameters, as we show below. Also, the distribution is cut off for EWs above 30 mÅ, which is motivated by the distribution of Ref. 25, as shown in Fig. 2. In this work, we set the normalization $A_p = 1$ and a power-law index of $m = 1.5$.

Added to this distribution is the random noise (N_r) in each resolution element, which forms a Gaussian ($\mathcal{N}$) centered at zero, parameterized by a normalization factor (A_g) along with a standard deviation (σ_g)

$$N_r = \mathcal{N}(A_g, \sigma_g). \tag{3}$$

Here we assume $A_g = 30{,}000$ and $\sigma_g = 0.75$ mÅ, based on the baseline expectation of Arcus.[15] The observed distribution, $\mathrm{d}^2\mathrm{n}/\mathrm{dzdEW}$, is the convolution between the noise and true distribution, as is illustrated in Fig. 3. For our analysis, we performed Monte Carlo simulations, drawing 10^3 random EW realizations; this number was chosen simply to be large enough to

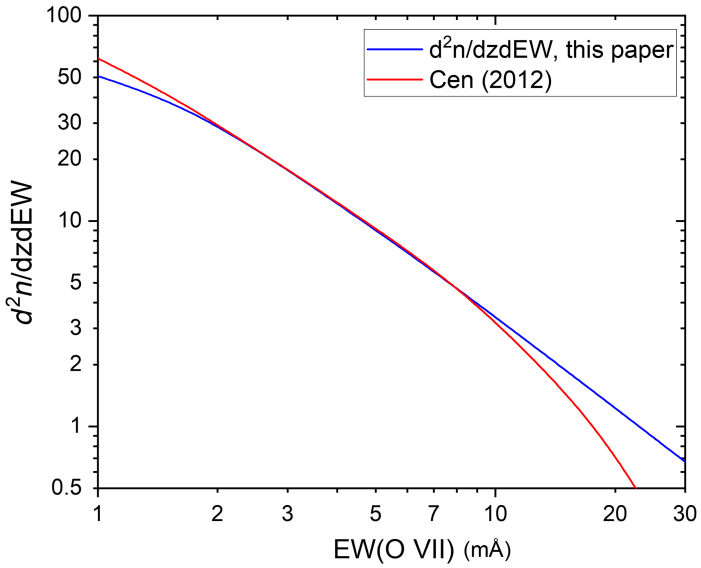

Fig. 2 Simplified differential column density distribution used here [Eq. (2), blue] compared with the distribution by Ref. 25 (red), in which the latter is normalized to be equal at 3 mÅ. The range shown here is chosen from below the EW values that provide significant constraints on the fit to the maximum EW value of the distribution.

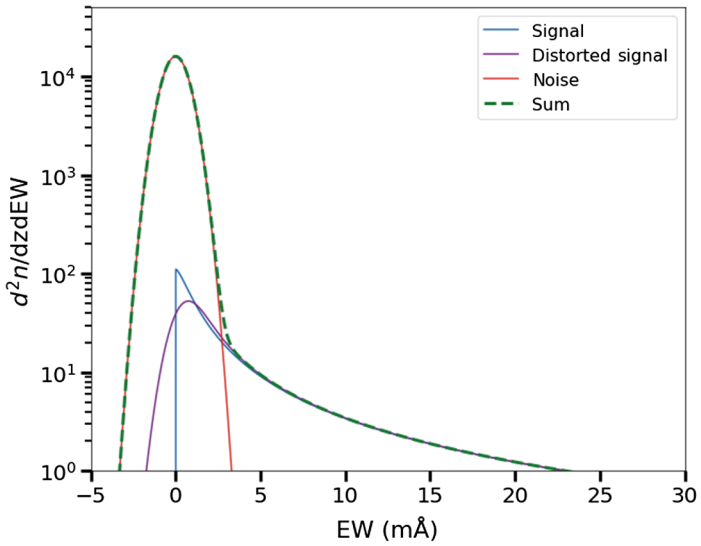

Fig. 3 Components of the model include the original distribution [Eq. (2), blue], the distribution convolved with the Gaussian error (black), the noise from the many spectral resolution elements without absorption lines (red), and the sum (green dashed).

gather robust statistics while remaining computationally inexpensive. From each realization, we obtained an EW distribution by binning the data into histograms, which we describe next.

3.2 Adaptive Binning

Recovering simulated distributions requires binning data to estimate the signal properties. Unfortunately, the choice of bins is arbitrary and must be determined *a priori*. In practice, one should develop a binning scheme that best represents the data, i.e., one that allows for the most accurate recovery of the input signal. For example, it would not be ideal to use linear spaced bins when trying to recover a rapidly changing distribution, which occurs here. The intrinsic EW distribution, $d^2n/dzdEW$, contains few counts at large EWs, so the bins are broader, whereas at lower EWs, the bins can be smaller. It is important to use consistent binning throughout the simulations even though each realization will produce a different distribution.

Here, we set a fiducial bin width of 0.25 mÅ for the low EW data. Moving toward larger EW values, we allowed the next bin to have the same width as long as the model predicted $\geq$15 counts for that bin. This ensured that the signal-to-noise (S/N) of each bin was $\gtrsim$3 using Poisson statistics. We acknowledge that one could use more elaborate binning schemes, such as Bayesian blocks,[28] which may be considered in future work.

3.3 Recovery Methods

We used two methods to recover our input parameters, one being a parametric fit for the four parameters of the model, such as A_p, m, A_g, and σ_g (the full-parameter fit, designated FPF). An advantage of this approach is that we can understand the accuracy to which each parameter is fit. A disadvantage is that we are requiring that four parameters be fit, compared with our other approach in which only two parameters were fit.

In the other model, we did not fit for the Gaussian, but rather subtracted the absolute values of the negative EW values (apparent emission lines) from those on the positive side, our flip and subtract (FAS) approach. This method was non-parametric in obtaining the $d^2n/dzdEW$ distribution, which was important considering it was convolved with the uncertainties. In reality, the continuum noise distribution may be more complicated than a single Gaussian defined by A_g and σ_g. The FAS approach did not assume a functional form for the noise; therefore, it would be the ideal choice for real observations. The only shortcoming to FAS is that the noise from the negative side of the Gaussian was added to the positive side, increasing the statistical uncertainty.

The uncertainties on the FAS result were calculated by adding the counts in the pairs of respective negative and positive bins in quadrature. Measuring counts in a bin obeys Poisson

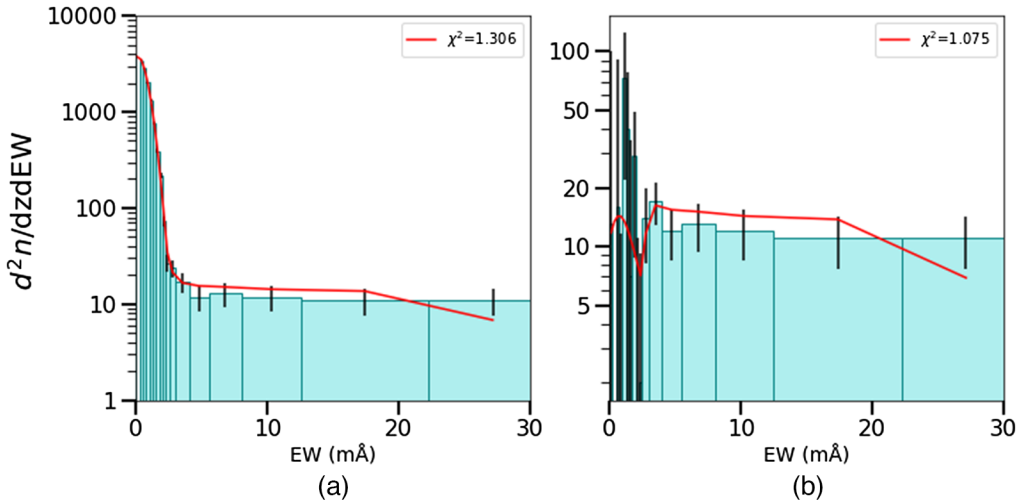

Fig. 4 Examples of the FPF (a) and FAS (b) recovery methods for a single realization. The best-fit model is shown in red for each, and the reduced χ^2 value is presented in the upper right of each panel. The fitting was done via least-squares regression.

characteristics, so the uncertainty takes the form of $1/\sqrt{T}$, with T being the number of counts per bin. Using this method, we fit for the underlying $d^2n/dzdEW$, convolved with the Gaussian error distribution. As an example, we show fits for each method given a single realization in Fig. 4.

4 Results and Discussion

4.1 Returned Parameter Spread

As mention in Sec. 3.3, we attempted to recover the input signal distribution using two different recovery methods. The best fit parameters for each method are presented in Table 1, and their distributions can be found in the corner plots in Figs. 6 and 7. Both the FPF and the FAS methods recovered the input signal parameters with similar precision. Also, both produced the same correlation between the power-law slope and normalization parameters as seen in Figs. 6 and 7.

We note that the FPF yielded a biased estimate on the normalization of the Gaussian, A_g. Given the estimated uncertainties in Table 1, the value of $A_g = 0.992$ differs from the true value

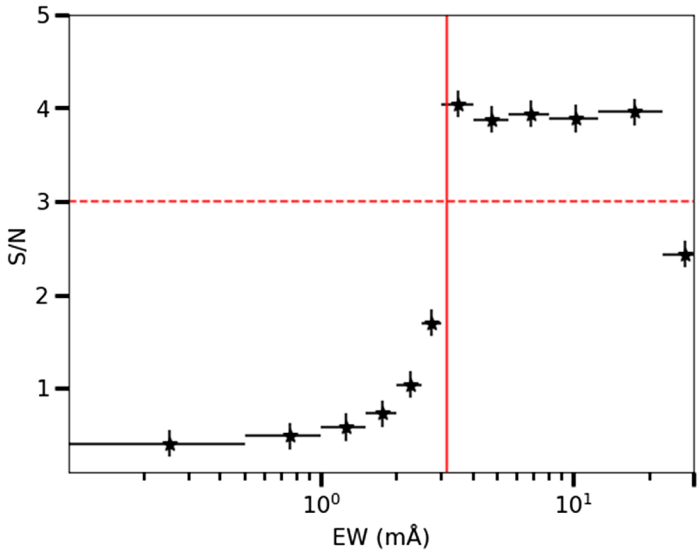

Fig. 5 Average S/N distribution from all realizations. The cutoff point for SN ≥ 3 is shown by the intersecting red lines on the left plot, with a value of 3.0 ± 0.3 mÅ. The horizontal error bars represent the width of each bin, and the vertical error bars denote the standard deviation using different realizations.

Table 1 Average results from the full simulation, reflecting the results in Figs. 6 and 7. The uncertainties on these values are obtained from the 16'th and 84'th percentiles.

FPF				FAS	
A_p	m	A_g	σ_g	A_p	m
$1.08^{+0.37}_{-0.28}$	$1.56^{+0.15}_{-0.15}$	$0.992^{+0.0013}_{-0.0017}$	$0.753^{+0.0031}_{-0.0031}$	$1.05^{+0.37}_{-0.27}$	$1.56^{+0.15}_{-0.15}$

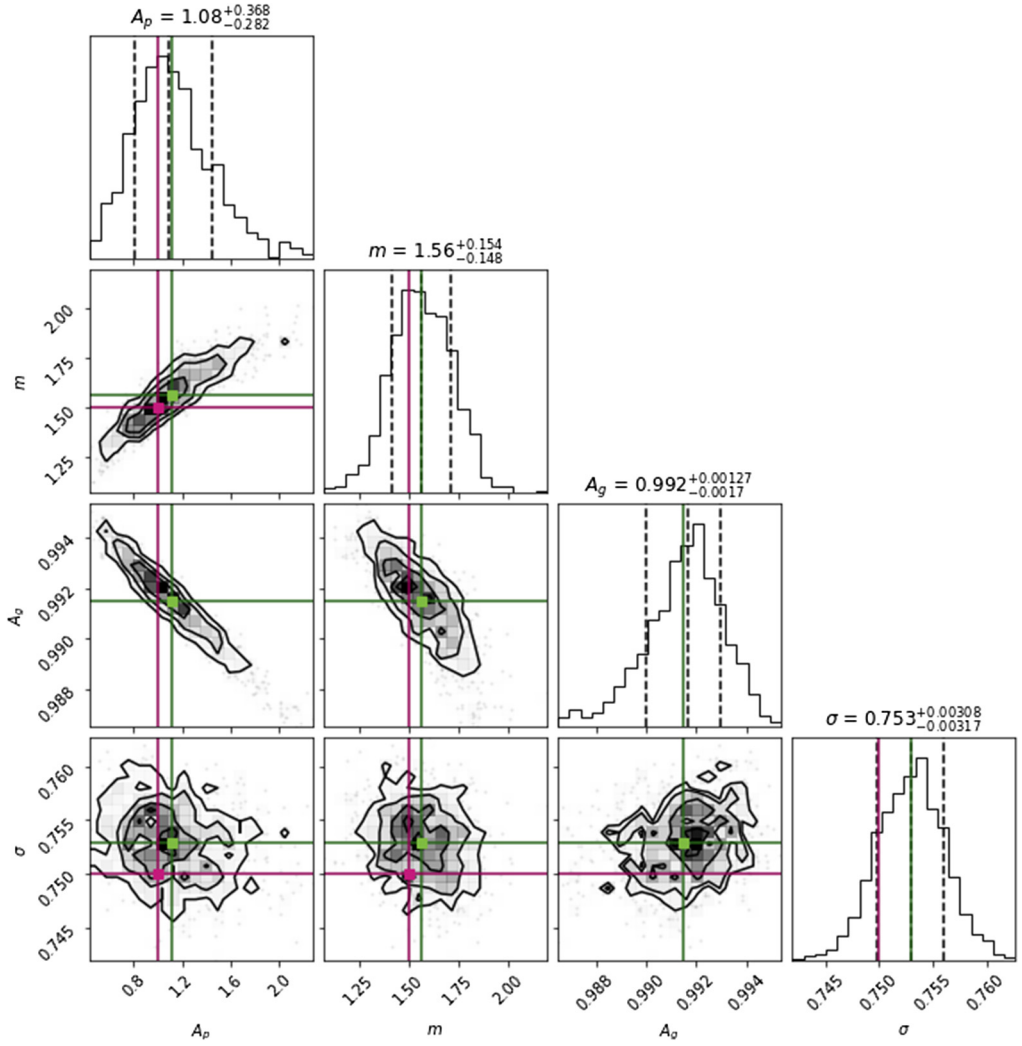

Fig. 6 Distribution of best-fit parameters obtained for each of the 1000 realizations from FPF. Red lines indicate the input values, and green lines give the mean. In the histograms, the dotted lines represent the 16'th and 84'th percentiles. The values above the histograms are equal to those in Table 1.

by $\approx 8\sigma$. This effect may be related to the correlations between the power-law slope and normalization parameters in Figs. 6 and 7. Regardless, this bias was near the 1% level and did not significantly affect our results.

4.2 Signal-to-Noise

An important empirical measurement is the S/N as a function of EW. In particular, one would like to know at what EW value the signal should separate from the noise in a significant way. In the following, we discuss how this was done using the FAS method.

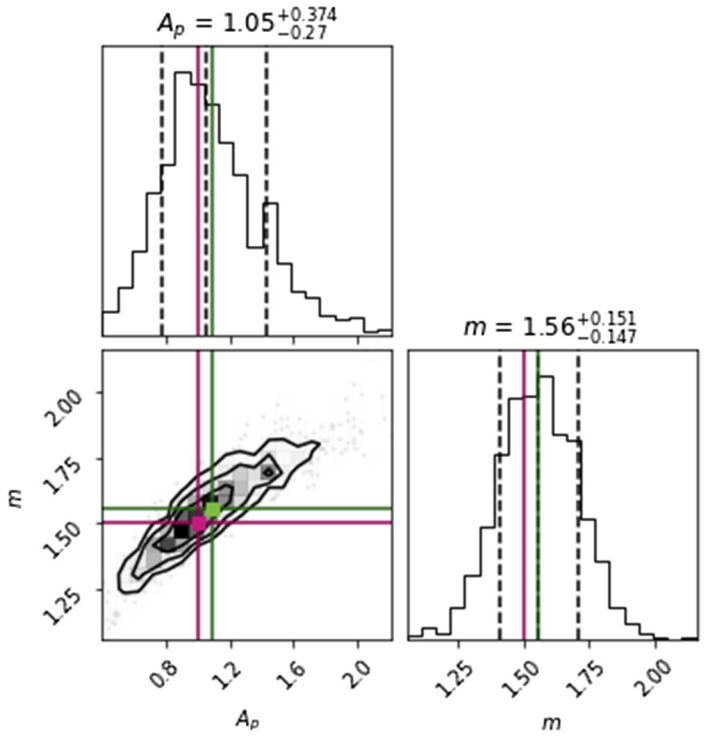

Fig. 7 Same as in Fig. 6 but for the FAS method.

The observed signal for the FAS method was obtained by subtracting the negative EW distribution from the positive side. The uncertainty for the remaining positive values were determined by simple error propagation (i.e., added in quadrature). An example of a single realization is shown in Fig. 4(b). One can notice that the uncertainties for each bin are larger in the right panel compared with the left due to the propagation of errors.

The FAS method was performed over 1000 realizations, and the average S/N for each bin is shown in Fig. 5. On this plot, we also show a horizontal line denoting where an $S/N = 3$ is achieved. As expected, the S/N values were small at low EWs due to the high presence of false detections. Eventually, the S/N rapidly improved and surpassed an S/N of 3 at an EW of 3.0 ± 0.3 mÅ for the FAS (similar results were obtained using FPF). In practice, this sets the theoretical limit for the minimum EW that can be used to reliably measure the EW distribution in the real Universe.

Another important measured quantity was the intrinsic uncertainty in the number of detected systems— that is a cosmological variance set by the fact that we can only observe one Universe. Counting the number of detections required an estimate of the signal using either FAS or FPF. The uncertainties were large at low EWs, so we measured the total number of systems above some EW threshold (here we use EW = 3.0 mÅ). We compared the total number of systems measured above the EW threshold with the model input. The FAS and FPF methods yielded a fractional uncertainty of ≈14%. These changed to 18% and 21% when we lowered the EW threshold to 2 and 1 mÅ, respectively. These uncertainties translate directly to the uncertainty in a cosmological density, such as Ω (O VII).

5 Summary and Conclusion

The motivation for this work was to understand the uncertainties involved in determining the cosmological O VII and O VIII metal contributions to the Universe, based on X-ray missions in the planning stages, such as Athena or Arcus. We adopted a functional form for $d^2n/dzdEW$ that is idealized (defined by two parameters) but similar to those predicted from large scale structure simulations. We created simulated datasets and used two approaches to recover them: an FPF, in which we fit for both the distorted absorption line distribution and the Gaussian noise profile, and a FAS method, in which the distribution of negative EW features (false apparent

emission features) is subtracted from the distribution of positive features (absorption lines). In the FAS method, the net spectrum was fitted with the model $d^2n/dzdEW$ that was convolved with the error distribution, which was assumed to be single Gaussian in this investigation.

We found that both models led to equivalently good fits and that, for ~100 O VII absorption systems with EW > 3 mÅ and 30,000 null detections (but smoothed by the intrinsic Gaussian), the power-law index m was recovered to 10% and the amplitude A_p to about 30% (1σ). The determination of the total EW or number of absorption systems was 14%, which sets the uncertainty in determining Ω (O VII). The fitting utilized all parts of the simulated data, but most of the power in fitting the input $d^2n/dzdEW$ came from absorption features with EW $\gtrsim$ 3 mÅ.

In an actual instrument, the sensitivity usually varies as a function of wavelength, so the noise distribution is a bit more complicated than the single Gaussian, but it can still be determined by the sum of Gaussians as a function of wavelength space. Another consideration is defining the location of the continuum, for which there is always some uncertainty. Incorrect locations of the continuum also broaden or distort the S/N distribution, particularly at low S/N values. An investigation of this effect is beyond the scope of this paper, but we suggest limiting the extraction of signals to $|S/N| > 3$ and using the FAS method. Values of $|S/N| < 3$ play no role in determining A_p and m.

This work was meant to address a focused issue relating to particular future missions. However, the application of this work is broader. The approach here can be applied to determining $d^2n/dzdEW$ in any wavelength region, such as for UV or optical absorption line systems, for which there are already large amounts of data.

6 Appendix A. Corner Plots

Here we present the distribution of best-fits based on 1000 simulations in Figs. 6 and 7, which show the covariance between parameters. In both FPF and FAS, there is a degeneracy between the power-law index m and the normalization of the power law A_p in the sense that a fit with a shallower power-law slope requires a larger normalization (and vice versa).

Data Availability

Data sharing is not applicable to this article as no new data were created or analyzed. All data in this work were generated randomly and can be easily replicated.

Acknowledgments

We would like to thank Zhijie Qu and Anne Blackwell for their comments and insights. We gratefully acknowledge support for this work through the University of Michigan and NASA (Grant Nos. 80NSSC19K1013 and 80NSSC22K0481).

References

1. M. Fukugita, C. J. Hogan, and P. J. E. Peebles, "The cosmic Baryon budget," *ApJ* **503**, 518–530 (1998).
2. J. M. Shull, B. D. Smith, and C. W. Danforth, "The Baryon census in a multiphase intergalactic medium: 30% of the Baryons may still be missing," *ApJ* **759**, 23 (2012).
3. J. Schaye et al., "The EAGLE project: simulating the evolution and assembly of galaxies and their environments," *Mon. Not. R. Astron. Soc.* **446**, 521–554 (2015).
4. P. F. Hopkins et al., "FIRE-2 simulations: physics versus numerics in galaxy formation," *Mon. Not. R. Astron. Soc.* **480**, 800–863 (2018).
5. B. D. Oppenheimer et al., "The multiphase circumgalactic medium traced by low metal ions in EAGLE zoom simulations," *Mon. Not. R. Astron. Soc.* **481**, 835–859 (2018).
6. B. P. Wakker et al., "The far ultraviolet spectroscopic explorer survey of O VI absorption in and near the galaxy," *Astrophys. J. Suppl. Ser.* **146**, 1–123 (2003).
7. J. K. Werk et al., "The COS-halos survey: physical conditions and Baryonic mass in the low-redshift circumgalactic medium," *ApJ* **792**, 8 (2014).
8. S. D. Johnson, H.-W. Chen, and J. S. Mulchaey, "On the possible environmental effect in distributing heavy elements beyond individual gaseous haloes," *Mon. Not. R. Astron. Soc.* **449**, 3263–3273 (2015).
9. Z. Qu and J. N. Bregman, "The mass and absorption column densities of galactic gaseous halos. II. The high ionization state ions," *ApJ* **862**, 23 (2018).

10. K. R. Sembach et al., "Highly ionized high velocity gas in the vicinity of the Milky Way," in *The IGM/Galaxy Connection. The Distribution of Baryons at z = 0*, J. L. Rosenberg and M. E. Putman, Eds., Astrophysics and Space Science Library, Vol. **281**, p. 155, Springer (2003).
11. T. Fang, J. Bullock, and M. Boylan-Kolchin, "On the hot gas content of the Milky Way halo," *ApJ* **762**, 20 (2013).
12. M. J. Miller and J. N. Bregman, "Constraining the Milky Way's hot gas halo with O VII and O VIII emission lines," *ApJ* **800**, 14 (2015).
13. F. Nicastro et al., "Observations of the missing baryons in the warm–hot intergalactic medium," *Nature* **558**, 406–409 (2018).
14. X. Barcons et al., "Athena: ESA's X-ray observatory for the late 2020s," *Astron. Nachr.* **338**, 153–158 (2017).
15. R. K. Smith et al., "Arcus: the x-ray grating spectrometer explorer," *Proc. SPIE* **9905**, 99054M (2016).
16. J. N. Bregman et al., "Strategies for detecting the missing hot Baryons in the universe," *J. Astron. Telesc. Instrum. Syst.* **1**, 045003 (2015).
17. J. M. Miller et al., "Chandra/high energy transmission grating spectrometer spectroscopy of the galactic black hole GX 339-4: a relativistic iron emission line and evidence for a seyfert-like warm absorber," *ApJ* **601**, 450–465 (2004).
18. M. F. Gu et al., "Laboratory measurement and theoretical modeling of k-shell x-ray lines from inner-shell excited and ionized ions of oxygen," *ApJ* **627**, 1066–1071 (2005).
19. E. Gatuzz et al., "Photoionization modeling of oxygen K absorption in the interstellar medium: the chandra grating spectra of XTE J1817-330," *ApJ* **768**, 60 (2013).
20. F. Nicastro et al., "Diffuse low-ionization gas in the galactic halo casts doubts on z $\simeq$ 0.03 WHIM detections," *Mon. Not. R. Astron. Soc.* **458**, L123–L127 (May 2016).
21. E. Gatuzz et al., "Measuring the hot ICM velocity structure function using XMM-Newton observations," arXiv:2307.02576 (2023).
22. R. SmithArcus Probe Team, "The Arcus Probe: exploring the formation and evolution of clusters, galaxies, and stars," in *American Astronomical Society Meeting Abstracts*, Vol. 55, p. 120.07 (2023).
23. R. Cen and J. P. Ostriker, "Where are the Baryons? II. Feedback effects," *ApJ* **650**, 560–572 (2006).
24. E. Branchini et al., "Studying the warm hot intergalactic medium with gamma-ray bursts," *ApJ* **697**, 328–344 (2009).
25. R. Cen, "Coincidences between O VI and O VII lines: Insights from high-resolution simulations of the warm-hot intergalactic medium," *Astrophys. J.* **753**(1), 17 (2012).
26. N. A. Wijers, J. Schaye, and B. D. Oppenheimer, "The warm-hot circumgalactic medium around EAGLE-simulation galaxies and its detection prospects with X-ray and UV line absorption," *Mon. Not. R. Astron. Soc.* **498**, 574–598 (2020).
27. T. Tuominen et al., "Cosmic metal invaders: intergalactic O VII as a tracer of the warm-hot intergalactic medium within cosmic filaments in the EAGLE simulation," *A&A* **671**, A103 (2023).
28. J. D. Scargle et al., "Studies in astronomical time series analysis. VI. Bayesian block representations," *ApJ* **764**, 167 (2013).

Nichole Gray received her BS degree in astronomy and astrophysics as well as in interdisciplinary physics from the University of Michigan in 2022.

Cameron T. Pratt is a graduate student in the Department of Astronomy at the University of Michigan. His research interests are extragalactic astronomy, signal processing, machine learning, and statistics.

Joel N. Bregman is a professor of astronomy at the University of Michigan with research interests in the gaseous content of the Universe as well as in mission planning, especially for NASA and ESA.