

XAS Between the Stars

Zack Gainsforth^{1*}, Anna L. Butterworth¹, John A. Tomsick¹ and Andrew J. Westphal¹

¹. Space Sciences Laboratory, University of California, Berkeley, CA, USA.

* Corresponding author: zackg@ssl.berkeley.edu

X-ray absorption spectroscopy (XAS) played a crucial role in understanding several microscopic grains returned by NASA's Stardust mission from the interstellar dust stream and which have a high likelihood of originating in the present-day interstellar medium[1]. In addition, XAS finds common use within the planetary science community to study other extraterrestrial samples such as meteorites, and interplanetary dust particles. Meanwhile, in the last couple of decades, the astronomy community has embraced the use of XAS to study the chemical characteristics of circumstellar and interstellar dust[2]. Telescopes such as Chandra (NASA), and Newton/XMM (ESA), have enabled high resolution X-ray spectroscopy in space. Chandra has an energy resolution better than 1 eV at the Fe-L edge, which enables a determination of Fe oxidation state. Synchrotron X-ray absorption spectra of solar system materials cannot be compared directly against interstellar X-ray spectra from space telescopes, because scattering is an important component of astronomical extinction spectra. However, extinction spectra can be derived from absorption spectra using the Kramers-Kronig relation and an assumption of the size distribution of interstellar dust. Bright X-ray sources such as black hole or neutron star binaries are needed to back-illuminate interstellar dust, and careful consideration of the physics at play within these sources is required to ensure a clean interpretation of interstellar spectra[3].

A significant challenge remains in identifying and extracting spectra in the presence of noise and interferences from undesired X-ray processes such as ionized atom transitions and superimposed continuum sources. In order to identify spectra for study, we wrote a code in python to search through the Transmission Grating Catalog of X-ray spectra from Chandra (TGCat)[4], which has over 1000 observations with sufficient energy resolution for our use. Figure 1 shows the location of all these sources superimposed on a dust map generated by the Planck Collaboration[5].

Because of spectral contamination, source variability, and other caveats noted above, it is impractical to analyze spectra summed across the entire sky. Instead, we applied machine learning to select spectra with characteristics matched to our individual studies. For study at the Fe-L edge, we extracted spectral features from all spectra in the TGCat database. Features included a polynomial pre-edge, post-edge, Fe- L₃ absorption line as well as a characterization of noise properties. We were then able to separate signals using the HDBSCAN clustering algorithm[6] and eliminate spurious Fe-L detections and ignore sources which added more noise than signal. Figure 2 shows a spectrum in blue (top) generated by summing all spectra in the TGCat database which represents an "all sky" spectrum at the Fe-L edge. Many features are present which can be assigned to ionized species, and other phenomena not related to our study. Plotted in orange (bottom) is a spectrum generated by summing only sources selected to optimize the features of interest to us. The difference is clear, the latter spectrum has an improved signal to noise and will allow for a cleaner scientific analysis since it lacks contaminating signals.

References:

- [1] AL Butterworth et al., *Meteoritics and Planetary Science* **49** (9) (2014), p. 1562.
 [2] JC Lee, *Space Science Reviews* **157** (1-4) (2010), p. 93.
 [3] AJ Westphal et al., *The Astrophysical Journal* **872** (1) (2019), p. 66.
 [4] DP Huenemoerder et al., *The Astronomical Journal* **141** (4) (2011), p. 129.
 [5] Planck Collaboration et al., *Astronomy and Astrophysics* **571** (2014), p. A11.
 [6] L McInnes, J Healy and S Astels, *The Journal of Open Source Software* **2** (2017), p. 11.

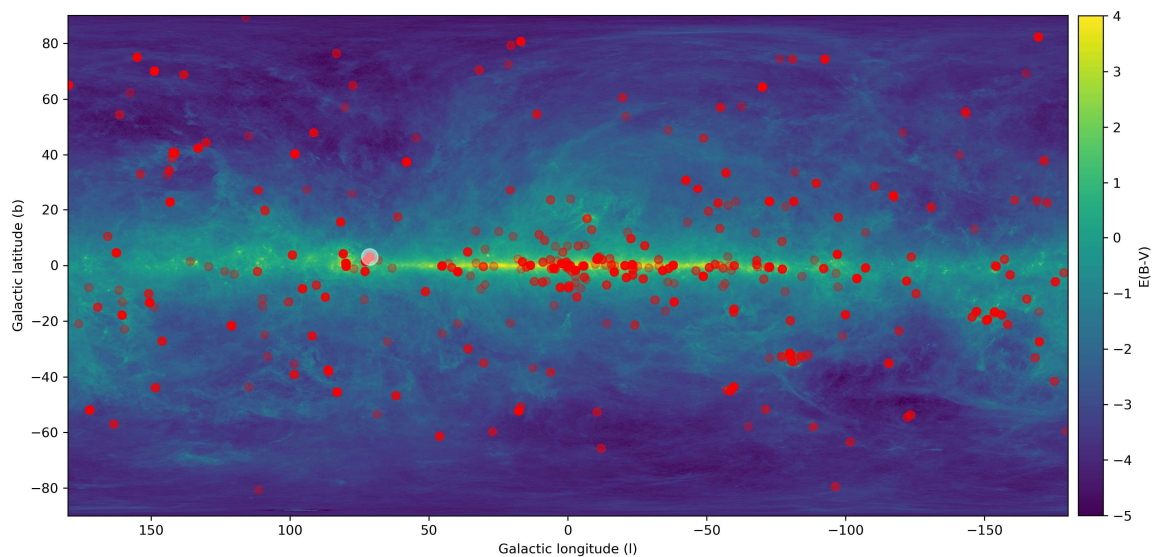


Figure 1. Location of X-ray sources with spectra containing sufficient resolution for XAS plotted over the Planck Collaboration Dust map. Cyg X-1 is highlighted white.

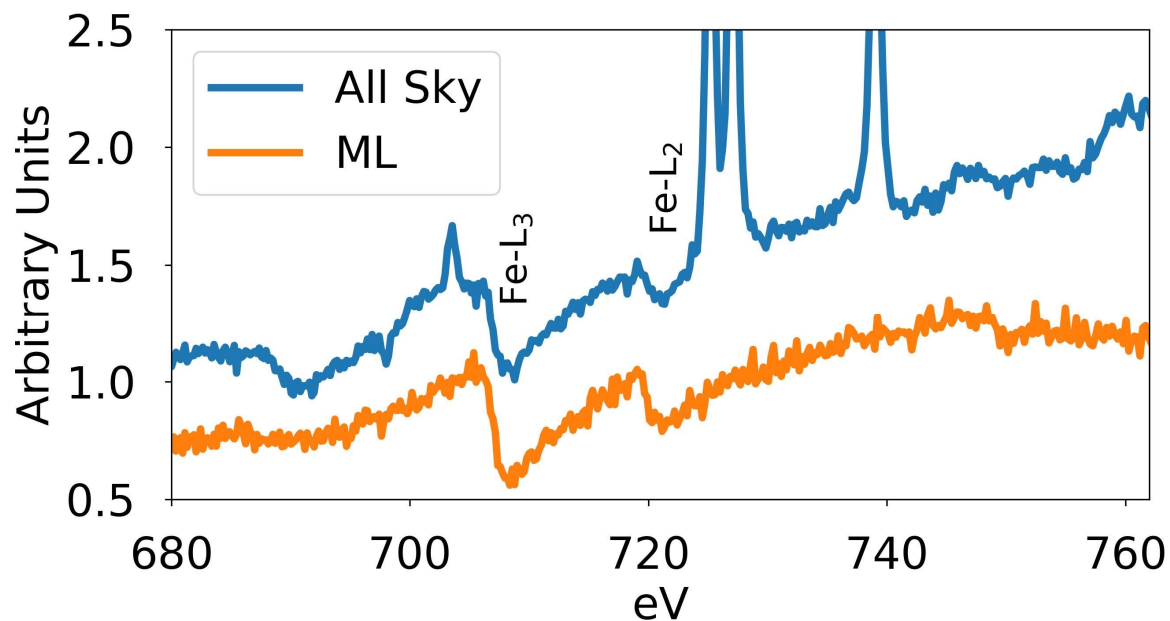


Figure 2. Comparison of an Fe-L spectrum generated by summing over the whole sky (> 1000 observations) with a spectrum generated by summing over specific sources selected using machine learning (19 observations).