Machine Learning Methods for Improved Interpretation of XES in the Analytical Determination of the Oxidation State Distribution Function

Gerald T. Seidler1 (\*), Lukas Nies1,2, William M. Holden1

1Physics Department, University of Washington, Seattle, United States

2II. Physics Institute, Justus Liebig University Giessen, Giessen, Germany

(\*) seidler@uw.edu

Many XANES studies of, e.g., transition metal compounds aim to estimate the fraction of different local electronic and structural environments for the metal species. It is common to measure the XANES of a system of interest along with a chemically-reasonable ensemble of reference standards, and then to use linear combination fits or principal component analysis to extract the nominal estimators of local properties. This method is contingent on some degree of orthogonality between reference spectra and, more critically, on the validity of reference spectra. This latter issue encompasses both the selection of reference samples but also systematic error because the disordered sample of interest will typically not have the same second- or farther-neighbors than the crystalline reference samples.

The scientific goal is to extract specific, well-defined estimators including oxidation state distributions, immediate bonding configurations, and coordination numbers. From the perspective of machine learning (ML), such quantities are the desired ‘classifications’ and the the prerequisite task is to generate a data set for directed training of a neural network. Timoshenko et al.1 have recently taken first steps to show how this approach can be performed via the considerable theoretical solution of the forward problem (i.e., accurate theoretical XANES), with impressive results for interpretation of XANES.

We explore a different but related problem: how to most reliably determine the distribution of oxidation states in K XES analysis of sulfur or phosphorus.2 This should be a far easier problem than XANES interpretation because the classification is simpler. Therefore, an ensemble of physically meaningful K spectra as a complete training ensemble includes the commonly observed ranges of energy shifts, K1/K2 ratios, peak splittings and individual peak widths. This is, then, essentially a deconvolution problem where one must allow for a degree of uncertainty in the underlying response function of the convolution.

We report the methods and results of directed ML machine learning to address this problem in benchtop analytical XES in the tender x-ray range.2,3 In particular, we compare and contrast the use of ML methods with the usual linear decomposition analysis based on reference compounds for the problems of sulfur speciation in biochars4 and for phosphorus speciation in InP core-shell nanoparticles.5 This comparison addresses the dependence of statistical inferences on assumed prior knowledge, a problem even more relevant for the harder XANES interpretation problem. From this work, we have developed a new, user-friendly analytical tool for estimating oxidation state distribution functions for S and P.

This work was supported by the US Department of Energy and The University of Washington.

1. J. Timoshenko, *et al.,* J. Phys. Chem. Lett. **8**, 5091 (2017).
2. L. Nies, W.M. Holden, G.T. Seidler, *in preparation* , J. Synch. Rad. (2018).
3. W.M. Holden, et al., Rev. Sci. Instrum **88**, 073904 (2017).
4. W.M. Holden, G.T. Seidler, S. Cheah, *submitted*, Analytical Chemistry (2018).
5. J. Stein, B. Cossairt, W.M. Holden, *et al*., *in preparation*, J. Amer. Chem. Soc. (2018).
6. E. Jahrman, G.T. Seidler, J.R. Sieber, *submitted*, Analytical Chemistry (2018).