

Extraction of Physical Parameters from X-ray Spectromicroscopy Data Using Machine Learning

Yuta Suzuki^{1,4}, Hideitsu Hino², Tetsuro Ueno³, Yasuo Takeichi⁴, Masato Kotsugi¹, and Kanta Ono^{4,*}

¹ Tokyo University of Science, Department of Materials Science and Technology, Tokyo, Japan.

² The Institute of Statistical Mathematics, Tokyo, Japan.

³ National Institutes for Quantum and Radiological Science and Technology, Hyogo, Japan

⁴ High Energy Accelerator Research Organization, Institute of Materials Structure Science, Ibaraki, Japan.

* Corresponding author, kanta.ono@kek.jp

Materials informatics has significantly accelerated the discovery and analysis of materials in the last decade. Spectroscopic data provide essential information about materials and hence are widely used for materials analysis. However, data analysis, that is, the extraction of physical parameters, of spectra is often conducted by manually comparing spectra and on-the-fly data analysis has not been realized yet. Considering that more than 100,000 X-ray absorption spectra (XAS) can be measured per day using the scanning transmission X-ray microscopy system at the Photon Factory [1], an automated analysis methodology is urgently required [2, 3]. If physical parameters are to be estimated from the spectra, the on-the-fly analysis can be realized by space mapping of these parameters using a high-throughput spectromicroscopy experiment capable of adaptive measurements [4]. XAS often shows complex spectral features with a few hundred or more “high-dimensional” data points, and the physical parameters, such as element, charge, and symmetry can be extracted from the XAS [5]. Feature extraction, a popular machine learning approach to treat high-dimensional datasets, involves the projection of data onto a few features (parameters), thus retaining the relevant information [6]. In this study, we propose a methodology to estimate the physical parameters from XAS using feature extraction with dimensionality reduction, as shown in Fig. 1.

We chose Mn XAS with different charges (2+, 3+, 4+) and crystal field parameters (10 Dq: 0–2.5 eV) as a simple example for the physical parameter estimation. The dimensionality reduction of Mn L-edge XAS was performed for the simulated XAS with CTM4XAS [7]. Although CT4XAS employs a very simplified model for XAS simulation, our methodology can also work with a more sophisticated first principles computational approach such as Bethe-Salpeter equation calculation [8-10]. We tested dimensionality reduction algorithms such as multi-dimensional scaling (MDS), ISOMAP, and t-distributed stochastic neighbor embedding [11-13].

Fig. 2 shows the results of the dimensionality reduction using MDS for the simulated XAS of the Mn compounds with different charges and crystal field parameters. The results indicated that the XAS data can be directly projected onto the physical parameter subspace, and automated estimation of charge and crystal field parameters from the XAS data was expected. The red circle in Fig. 2 corresponded to the experimentally measured XAS of MnO. It was shown that the physical parameters of Mn²⁺ with 10 Dq = 0.9 eV estimated from the figure agreed very well with the values estimated manually [5]. This methodology to estimate the physical parameters from the measured data should be applied to various X-ray microscopy analyses (i.e., for images, 3D data) and will significantly contribute to the analysis of large amounts of experimental data. This machine learning approach also contributes to achieve a spatially simplified description of the spectromicroscopy data [14].

References:

- [1] Y Takeichi *et al*, Rev. Sci. Instrum. **87** (2016), 013704.
 [2] M Lerotic *et al*, J. Electron Spectrosc. Relat. Phenom. **144-147** (2005), p. 1137–1143.
 [3] M Lerotic *et al*, J. Synchrotron Rad. **21** (2014), p. 1206–1212.
 [4] T Ueno *et al*, npj Comput. Mater. **4** (2018), p. 4.
 [5] F Groot, A Kotani in “Core Level Spectroscopy of Solids”, (CRC Press, Boca Raton, 2008) p. 225-285.
 [6] C M Bishop in “Pattern Recognition and Machine Learning”, (Springer-Verlag, New York, 2006) p. 559-603.
 [7] E Stavitski, F M F de Groot, Micron **41** (2010), p. 687–694.
 [8] E L Shirley, Phys. Rev. Lett. **80** (1998), p. 794-797.
 [9] J Vinson *et al*, Phys. Rev. B **83** (2011), p. 115106 1-7.
 [10] Y. Liang *et al*, Phys. Rev. Lett. **118** (2017), p. 096402 1-7.
 [11] I Borg, P Groenen in “Modern multidimensional scaling: theory and applications”, (Springer-Verlag New York, 1997) p. 1-18.
 [12] J B Tenenbaum, Science **290** (2000), p. 2319–2323.
 [13] L V D Maaten, G Hinton, J. Mach. Learn. Res. **9** (2008), p. 2579–2605.
 [14] This work was partly supported by JST CREST JPMJCR1761.

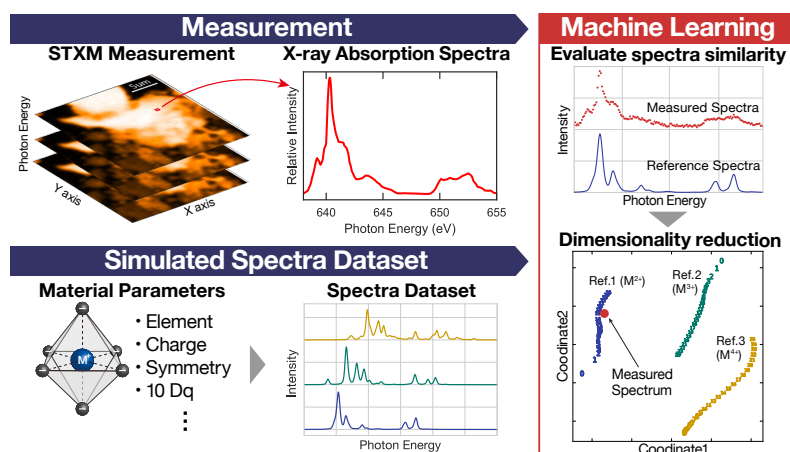


Figure 1. Machine learning methodology of efficient estimation of material parameters based on similarity of the data. The key to this method is a utilization of machine learning to feature extraction and comparison of spectra.

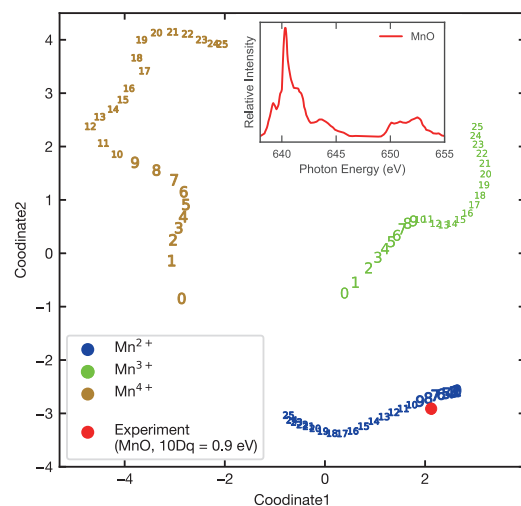


Figure 2. Dimensionality reduction and visualization of calculated Mn 2p XAS spectra and experimentally obtained XAS spectrum of MnO (inset)[5].

The experimentally obtained XAS spectrum of MnO with $10Dq = 0.9$ eV is located in the appropriate position. The numbers in this figure correspond to $10Dq$ (eV) multiplied by 10.