WORKSHOP #9

Curation, Data Analysis and Computational Modeling of X-ray Absorption Spectroscopy

Organizers: Eli Stavitski (NSLS-II), Daniel Allan (NSLS-II), Deyu Lu (CFN), Xiaohui Qu (CFN) and Matthew R. Carbone (CSI)

Synchrotron X-ray absorption spectroscopy (XAS) is a premier materials characterization technique that has broad applications in condensed matter physics, chemistry, materials science, and biology. To accelerate the scientific discovery at NSLS-II and CFN, a set of robust and efficient tools for curation, computational modeling and data analysis are critical for uncovering the structural and electronic characteristics of the system of interest, which are convoluted in the XAS signal. The workshop will provide lectures and hands-on tutorials on the following state-of-the-art tools. All software and tutorials will be open source.

- AIMM DB: Traditionally, XAS experimental results, simulations, and sample metadata were stored in highly varied forms using disparate formats and conventions. To apply new analysis tools, especially high-throughput and AI-guided methods, the data should be curated to be accessible, searchable, and usable. We developed and deployed "AIMM DB", a database and web application to capture and curate the data generated by a multi-laboratory collaboration. This tool is an extension of the Bluesky Tiled project from NSLS-II, adding new search capabilities and uploading analysis results, as opposed to only raw data, and by enforcing metadata standards relevant to X-ray spectroscopy. Participants will use Jupyter to connect to AIMM DB, learning how to access remote data and use it in interactive data analysis.
- MCR [https://github.com/AI-multimodal/mcr_const]: a matrix factorization-based method to transform a spatial or temporal spectral sequence into a simpler view of a few pure chemical species. Methods to incorporate prior knowledge to obtain physically meaningful results using specialized mathematical constraints will be introduced. Participants will have the opportunity to explore the effects of different constraints on spectral deconvolution result with example notebooks provided by the workshop.
- Lightshow [github.com/AI-multimodal/Lightshow]: a Python package for generating computational X-ray absorption spectroscopy input files for five popular first-principles simulation codes: FEFF, XSPECTRA, VASP, OCEAN and exciting. The Lightshow code provides an easy-to-use workflow and default parameters for each simulation package. This significantly lowers the barrier of computational core-level spectroscopy for non-expert and improves the reproducibility and reliability of the simulation data. Participants will have the opportunity to run XAS simulations on the computational resources provided by the workshop.

Agenda

Start Time (ET)	Title	Speaker (Affiliation)
9:00 – 9:10 a.m.	Opening Remarks	
9:10 – 9:50 a.m.	Calculating ab initio XANES with *exciting*	Benedikt Maurer (Humboldt-Universität Berlin)
9:50 – 10:30 a.m.	Calculating X-ray Spectroscopy with the OCEAN Code	John Vinson (NIST)
10:30 – 10:45 a.m.	Break	
10:45 – 11:25 a.m.	X-ray spectroscopy data: computation, database, extraction, featurization	Maria Chan (Argonne National Laboratory)
11:25 – 12:05 p.m.	Improved Calculations of XAS with FEFF and Corvus	Joshua J. Kas (University of Washington)
12:05 – 12:30 pm	A brief introduction to the challenges of modeling XAS	John Rehr (University of Washington)
12:30 – 1:30 p.m.	Lunch Break	
1:30 – 2:00 p.m.	AIMMdb: a data access and search service for multimodal XAS data	Juan Marulanda (NSLS-II/BNL)
2:00 – 3:45 p.m.	Lightshow Demo and Hands-on	Matthew Carbone (CSI/BNL) and Fanchen Meng (CFN/BNL)
3:45 – 4:00 p.m.	Break	
4:00 – 5:00 p.m.	MCR Demo and Hands-on	Xiaohui Qu (CFN/BNL)
5:00	Adjourn	

Calculating ab initio XANES with *exciting*
Benedikt Maurer (Humboldt-Universität Berlin)

X-ray Absorption Near Edge Spectroscopy (XANES) is a powerful tool for studying the core states and local electronic structure of solids and molecules. With the ability to reveal fingerprints of inequivalent atomic sites, it provides valuable insights into many chemical properties. The state-of-the-art approach to XANES analysis involves solving the Bethe-Salpeter equation (BSE) within the realm of many-body perturbation theory to study bound states between a core hole and an excited electron. To achieve true ab-initio calculations, a solution to Dirac's equation for the core states is required. In this presentation, we will cover the fundamentals of the theory and demonstrate how we solve it with the full-potential, all-electron solver, Exciting. We will present a selection of results, including examples for carbon K edges, oxygen K edges, and Ca L2,3 edges.

Calculating X-ray Spectroscopy with the OCEAN Code
John Vinson (NIST)

In this talk I will be introducing the OCEAN code for calculating electronic excitations using the Bethe-Salpeter equation method. OCEAN can be used to calculate a variety of different spectroscopies, including valence/UV, x-ray absorption, non-resonant inelastic x-ray scattering (or x-ray Raman

scattering), and resonant inelastic x-ray scattering (RIXS), and brief examples will be shown for each. Since today's workshop is dedicated to computational x-ray spectroscopy, I will try and outline some of the particular implementation choices made in OCEAN that set it apart from other codes and what these choices mean for OCEAN users. I will show several examples that highlight how computational spectroscopy can be used to help interpret measured spectra and note some examples where common approximations break down.

X-ray spectroscopy data: computation, database, extraction, featurization
Maria Chan (Argonne National Laboratory)

There has been extensive interest in machine learning (ML) of x-ray spectroscopy data. In this talk, we will discuss several important questions surrounding these efforts – Where does the data come from? How can we make the data FAIR? What is the best way to represent the data for ML models? The goal of this talk is both to discuss current efforts (including in the AIMM project involving scientists from Brookhaven, Argonne, and Berkeley National Labs) and raise questions for discussion.

Improved Calculations of XAS with FEFF and Corvus
Joshua J. Kas (Univ. of Washington and SLAC)

X-ray and related spectroscopies are powerful probes of atomic, vibrational, and electronic structure. Advanced calculations can require multiple scientific methods and codes to obtain accurate results. Here we describe a variety of recent improvements in modeling XAS, giving a variety of illustrative examples. In addition, we describe the use of the FEFF real-space multiple-scattering code together with the Corvus workflow machinery to facilitate accurate calculations.