Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Multiple Scattering Theory Simulation for Single-Crystal Fe-Based Compounds and Automated Material Characterization Based on Machine Leaning



Project Manager
Xiankang Tang

Principal Investigator
Jun. Prof. Dr. Hongbin Zhang

Project Term
2024 - 2025

Clusters
Lichtenberg II Cluster Darmstadt

Additional Software
FEFF10, abTEM, AutoMC

Institute
Theory of Magnetic Materials

University
Technische Universität Darmstadt

Introduction

Spectroscopy plays a pivotal role in materials science by providing fundamental insights into the composition, structure, and properties of materials at the atomic and molecular level. It encompasses a suite of non-destructive analytical techniques that probe the interaction between matter and electromagnetic radiation (like light, X-rays, or infrared). By analyzing how materials absorb, emit, or scatter this radiation, scientists can identify specific elements and molecules present, determine chemical bonding states and electronic structure, characterize crystallinity and phase composition, measure layer thicknesses, detect defects, and monitor surface reactions or degradation processes. This study combines machine learning methods and multiple scattering theory to take a step forward in predicting the spectra of inorganic materials. In this study, forward mapping and inverse mapping for spectroscopy of inorganic materials are established. The data required to train the model is obtained by high-throughput computation, which is the reason why we apply so many CPU core times.

Methods

Training and test datasets were generated using the FEFF

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

software. Atomic environments were encoded using the Smoothed Overlap of Atomic Positions (SOAP) descriptor from the Dscribe library. Spectra were predicted by training a random forest regressor. Three models were trained on each of the three variants of the same spectral dataset: a min-max normalized model, an average normalized model, and an unnormalized model. Hyperparameter tuning was performed using the optuna framework. A random forest regressor can be trained to predict the EXAFS spectra. On the other hand, we have modeled the inverse mapping of AFM and HRTEM using Bayesian optimization, which has greatly accelerated the development of materials science.

Results

We trained a random forest regression model using EXAFS data for selected Fe-Co compounds. Three separate models were trained on three variants of the same spectra dataset: with minimum-maximum normalization, with mean normalization, and without normalization. Hyperparameter tuning is performed using the optuna framework.
It was possible to train the random forest regressor to predict the EXAFS spectra. The best results were obtained using minimum-maximum normalization with an R2 score of 8.96e-1 on the training dataset and an R2 score of 8.24e-1 on the test dataset. The
other models show greater tendencies toward overfitting. These results demonstrate the feasibility of predicting EXAFS spectra using machine learning models, although further work is needed to improve accuracy and generalization.

Discussion

This research seeks to capture the spectra of compounds with Fe atom as the absorbing atom using multiple scattering theory and machine learning methods, with the goal of establishing an inverse and forward mapping of the crystal structure and spectra of the materials. The results show that our model is very effective in realizing the intended conception. In the future, we will also try generative modeling to further investigate the inverse mapping. In addition, how to combine microscopy data into the model as well, so as to realize multi-modal machine learning models, is also something worth considering in our next step.

Last Update: 2025-09-26 18:34