Automated Workflows and Machine Learning models for X-ray spectra simulations: applications to Li-ion battery materials

Michael A. Hernandez Bertran[1,2], Federico Grasselli[2], Davide Tisi[3], Elisa Molinari[1,2], Michele Ceriotti[3], Deborah Prezzi[1]
[1]Istituto Nanoscienze, Consiglio Nazionale delle Ricerche CNR, via Campi 213/A 41125, Modena, Italy.
[2]FIM, University of Modena and Reggio Emilia, via Campi 213/A 41125, Modena, Italy.
[3]Institut des Sciences et Génie Matériaux, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1015 Lausanne, Switzerland.

michaelalejandro.hernandezbertran@unimore.it

In the framework of material science, core-level spectroscopies are established strategies to probe the electronic structure and chemical environment of materials. While these techniques provide valuable information, their interpretation in complex systems is not straightforward, emphasizing the importance of theoretical insights from ab-initio approaches. On the other hand, the time required for computing an X-ray spectrum increases with the number of non-equivalent atoms, becoming prohibitively expensive for complex, disordered materials. This challenge can be tackled by employing a surrogate model that combines the precision of ab initio methods with computational efficiency [1]. We trained a machine-learning (ML) model for predicting X-ray Photoelectron Spectroscopies (XPS), based on Kernel Ridge Regression (KRR) [2] and atom-density descriptors. Core-electron Binding Energies (BE) are used as the target quantity. A comprehensive automated AiiDA workflow [3], integrating first-principles XPS simulation with sample sub-selection via Farthest Point Sampling (FPS), was employed to generate the critical amount of data needed for the training process. The ML models were trained on a representative dataset comprising about 250 lithiated Si-based structures, previously obtained through ab initio molecular dynamics (MD), DeePMD and grand canonical Monte Carlo simulations [4]. Validation on a dataset of around 50 structures demonstrated an accuracy of 0.1 eV, aligning with typical XPS experimental resolution. To accurately sample the relevant range of Li concentrations, we generated an additional dataset of structures, containing over 70,000 atoms each, which was prepared employing a melt-quench-anneal procedure and neuroevolution potentials (NEPs) [5]. We leveraged the ML model to build a stoichiometry map in order to identify the LixSi phases that form at various potentials in Si-based anodes.

References

[1] F. M. Paruzzo, A. Hofstetter, F. Musil, S. De, M. Ceriotti, and L. Emsley, Nat. Commun. 9 (2018) 4501 .
[2] V. L. Deringer, A. P. Bartók, N. Bernstein, D. M. Wilkins, M. Ceriotti, and G. Csányi, Chem. Rev. 121 (2021) 10073.
[3] P. N. O. Gillespie, M. A. Hernandez Bertran, X. Wang, G. Pizzi, E. Molinari, and D. Prezzi, Preprint (2024).
[4] F. Fu, X. Wang, L. Zhang, Y. Yang, J. Chen, B. Xu, C. Ouyang, S. Xu, F. Z. Dai, W. E, Advanced Functional Materials, 33 (2023) 37.
[5] P. Pegolo, and F. Grasselli, Frontiers in Materials 11 (2024) 1369034.