Unlocking the Potential of EXAFS: Machine Learning Approaches for Spectroscopic Data

Javier Heras-Domingo[1], Andrea Ruiz-Ferrando[1], Stephan Pollitt[2], Sharon Mitchel[3] , Olga Safonova[2], Javier Pérez-Ramírez[3] and Núria López[1]
[1]Institut Català d'Investigació Química ICIQ, Av. dels Països Catalans 16, 43007 Tarragona, Spain
[2]Paul Scherrer Institute, CH-5232 Villigen, Switzerland
[3]Institute for Chemical and Bioingenieering, Department of Chemistry and Applied Biosciencies (ETH Zurich), Vladimir-Prelog-Weg 1, 8093 Zurich, Switzerland

jheras@iciq.es

Sustainability will only be achieved through the pursuit of precise control of materials, for instance utilizing minimal amounts of critical raw materials, which stands as the **Holy Grail** in heterogeneous catalysis. Heterogeneous catalysts featuring atomically dispersed active and stable metal atoms were coined in 2011 as single-atom catalysts (SACs) [1]. Since then, SACs exhibit remarkable and exciting performances in numerous industrially significant reactions. The host can be different materials, including oxides, and most commonly carbon-based ones. Due to the varied synthetic forms of carbon matrix SACs offer a very wide speciation (charge and coordination of the metal atom) imparting unique geometric and electronic properties needed in the catalysis field that requires to be identified via characterization tools like microscopy [2,3] and X-ray absorption spectroscopy.

Generally, in heterogeneous catalysis (and more critically in SACs) there is no direct way to inverse engineer active sites (i.e., targeting a catalytic property and identifying the synthesis route) and thus material synthesis and testing is followed by strict, long and expensive characterization protocols aimed to provide hints for structures that can be then evaluated via Density Functional Theory (DFT) simulations to trace structure-activity relationships. The ultimate consequences are that the development of materials is limited by the characterization step, which the demand of synchrotron characterization techniques is skyrocketing, even if in many cases the data treatment is done manually. In addition, bench solutions like hiXAS (HP Spectroscopy) and easyXAS are being developed and the cost of this equipment is getting reduced. Therefore, the amount of data and users is growing at a fast pace, highlighting the pressing need to leverage the newest advances in data science and artificial intelligence (AI) [4] to the field.

In this work, we have developed intelligent and automated EXAFS tools for the rational design of single-atom catalysts, while leveraging artificial intelligence (AI) to enhance data extraction, pivoting to quantitative information analysis like the relative propulations of different SAC environments.

The goal is to comprehensively understand the intricate structural changes occurring in series of single-atom catalysts (SACs) anchored on carbon-based host materials. We have leveraged **automated workflows** to generate high-quality theoretical data using state-of-the-art open-source software packages like Atomate [5], Fireworks [6], Pymatgen [7]. To ensure robustness and accuracy, we will not only utilize theoretical data but also incorporate experimental synchrotron EXAFS data, in collaboration with the SuperXAS beamline at PSI for both training and validation steps.

The collected data has been used to train and validate a new **Graph Neural Network** (GNN) model that is able to predict the EXAFS spectra of a given structure. Afterwards, the model only trained on simulated data has been fine-tuned with experimental data to include effects that are not well captured with simulated data, for instance the Debye-Waller effects also known as thermal disorders, among others. Our methodology is depicted in Figure 1.

The above-mentioned strategy has been used to follow the synthesis of Pt single-atom catalyst (SAC) hosted in PTI poly(triazine imide) support. Such synthetic route consisted in the following steps: i) Impregnation, ii) Wash, iii) low-temperature annealing and iv) high-temperature annealing. Furthermore, we could select from a pool of simulated structures to determine the most reasonable structures for each of the synthetic steps showcasing the generation and activating of the single-atom catalyst active site.

Finally, the significance of this work lays in the fact that **better characterization techniques are key to designing new catalysts with tailored properties**. We have leveraged artificial intelligence (AI) to enhance the widespread characterization technique EXAFS, ensuring the transition from qualitative to more quantitative information. Our new methodology has been showcased by following the synthesis of Pt-PTI catalyst where the potential structure on every step is elucidated.

References

[1] Qiao, B., Wang, A., Yang, X., Allard, L. F., Jiang, Z., *et. al.* Nat. Chem., **2011**, 3(8), 634-641
[2] Mitchell, S., Parés, F., Faust Akl, D., Collins, S. M., Kepaptsoglou, D. M., Ramasse, *et. al*, JACS, **2022**, 144(18), 8018-8029.
[3] Rossi, K., Ruiz-Ferrando, A., Akl, D. F., Abalos, V. G., Heras-Domingo, J., *et. al.*, Adv. Mat., **2023**, 2307991.
[4] Wang, H., Fu, T., Du, Y. *et. al*. Nature, **2023**, 620, 47–60
[5] Mathew, K., Montoya, J. H., Faghaninia, A., Dwarakanath, S., Aykol, M., Tang, H., Chu, L., Smidt, T., Bocklund, B., Horton, M., *et. al.*, Comput. Mater. Sci., **2017**, 139, 140-152.

[6] Jain, A., Ong, S. P., Chen, W., Medasani, B., Qu, X., Kocher, M., *et. al.*, Concurrency and Computation, **2015**, *27*(17), 5037-5059.
[7] Horton, M., *Nature Reviews Physics*, 2023, *5*(2), 76-76.

Figures

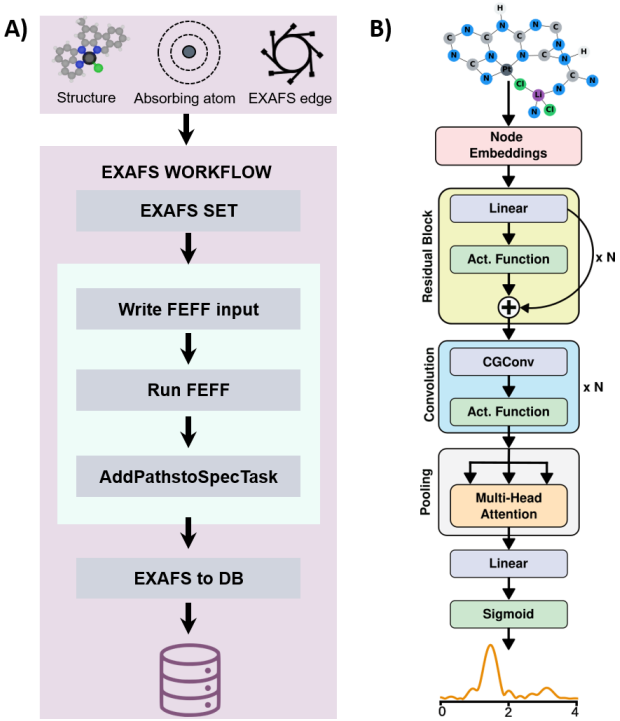

Figure 1. A) Automated workflow to generate simulated EXAFS spectra given a pool of structures, B) Graph Neural Network architecture consisting in a residual block, graph convolution block and a pooling layer with multi-head attention.