MS38 Computations with/for Pair Distribution Functions

MS38-01

Nanostructure characterization with pair distribution function analysis: tools for automated analysis

K.M.Ø. Jensen [1]

[1] *University of Copenhagen - Copenhagen Ø (Denmark)*

Abstract

The development of advanced, functional materials builds on an understanding of the intricate relationship between material structure and properties, and over the past century, crystallographic methods using scattering and diffraction have thus been essential for materials science. Crystallography allows ab initio determination of crystal structures from diffraction data, and has provided us with the vast knowledge of crystal chemistry that is now used in design of functional materials. However, in the case of nanomaterials with limited long-range order, crystallographic methods are challenged, and ab initio structure determination, or structure solution, is not currently possible. Over the past decades, total scattering with Pair Distribution Function (PDF) analysis has become an essential tool for characterisation of nanomaterial structure.[1] However, structure solution from the PDF is not possible except in a very few simple cases.[2-4] In the absence of broadly applicable ab initio nanostructure determination methods, it is therefore necessary to propose reasonable starting models and to then 'refine' the model parameters against the data using local minimization methods. The step of finding a starting model can be a major challenge and is thus a bottleneck in complex material characterization. Recently, new automated approaches have made it possible to test thousands of models against a dataset,[5-7] but these methods are computationally expensive, and analysing the output, i.e., extracting structural information from the resulting fits in a meaningful way is challenging. Here, we show how machine learning (ML) methods cain aid. Our Machine Learning based Motif Extractor[8] (ML-MotEx) trains an ML algorithm on thousands of automatically generated fits, which means that a larger strucural space can be invesigated. The use of explainale ML furthermore allows to use SHAP (SHapley Additive exPlanation) values to identify which model features are important for the fit quality. We have used the method on several different systems, and shown how it can be applied for PDF analysis of nanoclusters in solutions as well as disordered oxide materials.

As a second example of the use of ML for scattering data analysis, we show how autoencoders can be a step towards structure solution directly from PDF. Our algorithm DeepStruc[9] can solve a simple, monometallic nanoparticle structure from a Pair Distribution Function (PDF) obtained from total scattering data by using a conditional variational autoencoder. We show that this is possbile from both simulated and experimental PDFs, and illustrate how it learns to identify e.g. stacking faulted strcutures between fcc and hcp nanoparticles. These tools illustrate how new computational methods can aid in analysis of scattering data, and opens for new approaches to nanostucture characterization.

References

1: Billinge, S. J. L. & Kanatzidis, M. G. Beyond crystallography: the study of disorder, nanocrystallinity and crystallographically challenged materials with pair distribution functions. Chem. Commun., 749-760 (2004).

2: Cliffe, M. J., Dove, M. T., Drabold, D. & Goodwin, A. L. Structure determination of disordered materials from diffraction data. Phys. Rev. Lett. 104, 125501 (2010).

3: Juhás, P., Cherba, D. M., Duxbury, P. M., Punch, W. F. & Billinge, S. J. L. Ab initio determination of solid-state nanostructure. Nature 440, 655-658 (2006).

4: Juhás, P., Granlund, L., Duxbury, P. M., Punch, W. F. & Billinge, S. J. L. The Liga algorithm for ab initio determination of nanostructure. Acta Cryst. A 64, 631-640 (2008).

5: Banerjee, S., Liu, C.-H., Jensen, K. M. O., Juhas, P., Lee, J. D., Tofanelli, M., Ackerson, C. J., Murray, C. B. & Billinge, S. J. L. Cluster-mining: an approach for determining core structures of metallic nanoparticles from atomic pair distribution function data. Acta Crystallogr. A 76, 24-31 (2020).

6: Christiansen, T. L., Kjær, E. T. S., Kovyakh, A., Röderen, M. L., Høj, M., Vosch, T. & Jensen, K. M. Ø. Structure analysis of supported disordered molybdenum oxides using pair distribution function analysis and automated cluster modelling. J. Appl. Crystallogr. 53, 148-158 (2020).

7: Yang, L., Juhas, P., Terban, M. W., Tucker, M. G. & Billinge, S. J. L. Structure-mining: screening structure models by automated fitting to the atomic pair distribution function over large numbers of models. Acta Crystallogr. A 76, 395-409 (2020)

8: Anker, A. S. et al. Extracting Structural Motifs from Pair Distribution Function Data of Nanostructures using Explainable Machine Learning. ChemRxiv (2022). doi:10.26434/chemrxiv-2022-v1vfx (2022)

9: Kjær, E. S. T., Anker, A. S., Weng, M. N., Billinge, S. J. L., Selvan, R. & Jensen, K. M. Ø. DeepStruc: Towards structure solution from pair distribution function data using deep generative models. ChemRxiv doi:10.26434/chemrxiv-2022-0zrdl (2022).