AI4X – Crystalline phase projection and identification from high temperature diffraction patterns using graph neural networks

Léonard Imbert [1] Christopher L. Hassam [1] Olivier Hernandez [1] David Berthebaud [1] David Portehault [2] Romain Gautier [1]

[1]*Nantes Université, CNRS, Institut des Matériaux de Nantes Jean Rouxel, IMN, Nantes F-44000, France* [2]*Sorbonne Université, CNRS, Laboratoire de Chimie de la Matière Condensée de Paris (CMCP), 4 place Jussieu, F-75005 Paris, France*. Correspondence to: Léonard Imbert leonard.imbert@cnrs-imn.fr.

1. Introduction

Powder X-Ray Diffraction (PXRD) is a common technique for the characterization of crystalline materials. It is based on observing and analysing the patterns produced by X-ray diffraction on a powder of crystalline material. The diffraction patterns originates from the reflection of X-Ray on the reticular plane of the lattice structure. The intensity of those reflections are characteristic of the atomic composition and arrangement of the structure. Because each crystal has its own structure – convolution of the atomic pattern with the lattice, the diffraction pattern is a fingerprint of the material.

This is the reason why PXRD became a common method for crystal identification and lattice parameters refinements. The phase identification consists of comparing the experimental signal to fingerprints registered in the database and selecting the chemically sounded closest one. This method still has issues. Most, notably in our case, the databases contain mostly phases from which the fingerprint has been measured at ambient temperature, a fingerprint that would be different if measured in our experimental conditions. Additionally, in the search of new materials, it is possible that the target crystal has not been measured yet, thus its absence in the database.

I present in this abstract a method to identify phases from the thermodiffraction patterns of multiphasic samples and to compare high temperature diagrams to structural databases using ambient condition projection of high temperature diagrams.

2. Effect of temperature on PXRD patterns

As a reminder, the reticular planes of a crystal – which define the peak position in diffraction pattern – are defined by the miller indices $[h, k, l]$ which are vector coordinates in the elemental lattice reciprocal space defined by the three basis vectors $\vec{a}^*, \vec{b}^*, \vec{c}^*$ and the three angles between them $\alpha^*, \beta^*, \gamma^*$. We can thus define the interplanar distance d_{hkl} between two equivalent reticular planes as:

$$\begin{aligned}\frac{1}{|d_{hkl}|^2} &= |a^*|^2h^2 + |b^*|^2k^2 + |c^*|^2l^2 \\ &+ 2|a^*||b^*|hk\cos(\gamma^*) + 2|a^*||c^*|hl\cos(\beta^*) \\ &+ 2|c^*||b^*|kl\cos(\alpha^*)\end{aligned} \tag{1}$$

Supplementary effects of temperatures can result in modifications of the peak shapes, due to evolution of the crystallisation level let it be degradation or growth of the crystallites.

2.1 Crystal thermal expansion approximation

The evolution of the lattice parameters with temperature can be approximated as an exponential for each translational symmetry direction [1]:

$$\vec{a}(T) \approx \vec{a}(T_0)e^{-\alpha_a(T-T_0)} \tag{2}$$

with $\vec{a}(T_0$ the base vector at an arbitrary temperature T_0. It has been observed experimentally that thermal expansion coefficients α_a are of the order 10^{-5} K^{-1}. Considering temperature difference $(T - T_0)$ of the order of 10^3 K, we approximate the interplanar distance depending on the temperature as:

$$d_{hkl}(T) \approx d_{hkl}(T_0)e^{\alpha_{hkl}(T-T_0)} \tag{3}$$

describing the peak drift over the temperature. An example of this drifting is shown in the Figure 1 where we can see the peak drifting with the temperature ramp between 0 and 35 minutes.

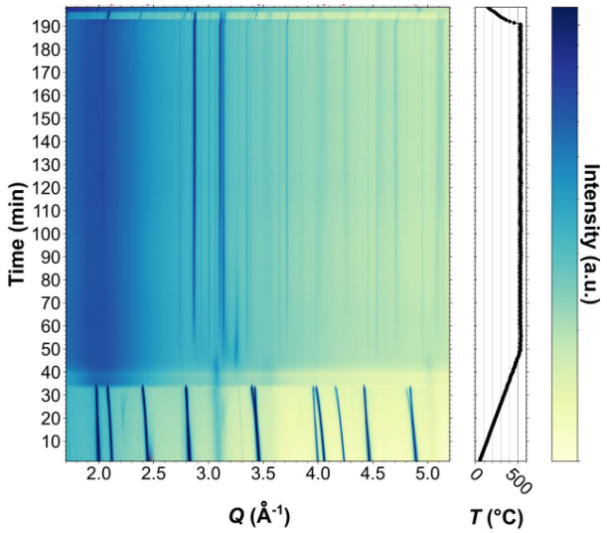

Fig. 1: Example of a thermodiffraction pattern. The Y axis is the time, X axis the is $Q = \frac{4\pi\sin(2\theta)}{\lambda}$ which represent the diffraction angle and the color the intensity of the diffraction. The time-temperature relationship is described by the curve on the right. We can see the peaks drifting during the temperature ramp between 0 and 35 minutes due to the lattice expansion. Figure taken from Song et. al [2]

3. Diffraction pattern processing

To process the thermodiffraction pattern we first analyse each diffraction pattern individually and independently. We extract each peak to transform the raw 1D diffraction patterns into a collection of peaks.

3.1 GNN for peak grouping

These peaks will form our graph, where each peak is a node and its features are the parameters presented before and the neighbouring relation is determined by the norm of $(t_1 - t_2, T_1 - T_2)$. The Graph Neural Network (GNN) is trained for edge classification, we ask the GNN for two classifications:

- Are the two peaks corresponding to the same reflection but different time and temperature ?
- Are they from the same crystalline phase ?

These two classifications allow us to create tracks of peaks from which we extract the thermal expansion coefficients and create diffraction pattern for each individual crystalline phases. Using that information, we can project the phases back to ambient temperature for comparison with the structural databases.

3.2 Synthetic training dataset

A large amount of training data is necessary for the GNN. Fortunately, because we are working with the aggregated features $(a, 2\theta, \sigma, \gamma)$ – the amplitude, position and shape parameters of the extracted peaks, the simulation of a synthetic training dataset is easier than if needed to simulate the entire thermodiffraction patterns. Using structures present in the Material Project database [3] we can simulate a large variety of aggregated thermodiffraction patterns.

Acknowledgments

This work is part of the MADNESS project of PEPR DIADEM and was supported by the French National Research Agency, as a part of the France 2030 program, under grant 23-PEXD-0011.

References

[1] G D Barrera, J A O Bruno, T H K Barron, and N L Allan. Negative thermal expansion. *Journal of Physics: Condensed Matter*, 17(4):R217, January 2005.

[2] Yang Song, Anissa Ghoridi, Fernando Igoa Saldaña, Isabel Gómez-Recio, Daniel Janisch, Edouard de Rolland Dalon, Dominique Thiaudière, Maria Luisa Ruiz-González, José M. González-Calbet, Benedikt Lassalle-Kaiser, Andrea Zitolo, Christel Laberty-Robert, and David Portehault. In and Out: Shuttling Atoms in Covalent Nanocrystals, from Synthesis in Molten Salts to Water-Splitting Electrocatalysis. *Journal of the American Chemical Society*, 147(24):20878–20887, June 2025.

[3] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, July 2013.