Automated Characterization of the Atomic Structure of Mono-Metallic Nanoparticles from X-ray Scattering Data using Generative Models

DanScatt · Columbia University · erc · University of Copenhagen · Villum Fonden

Andy S. Anker*[1], Emil T. S. Kjær*[1], Simon J. L. Billinge[3], Raghavendra Selvan[2] and Kirsten M. Ø. Jensen[1]

1: Nano-Science Center and Department of Chemistry, University of Copenhagen

2: Department of Computer Science and Department of Neuroscience, University of Copenhagen

3: Department of Applied Physics and Applied Mathematics, Columbia University and Condensed Matter Physics and Materials Science Department, Brookhaven National Laboratory, New York USA

*andy@chem.ku.dk and etsk@chem.ku.dk

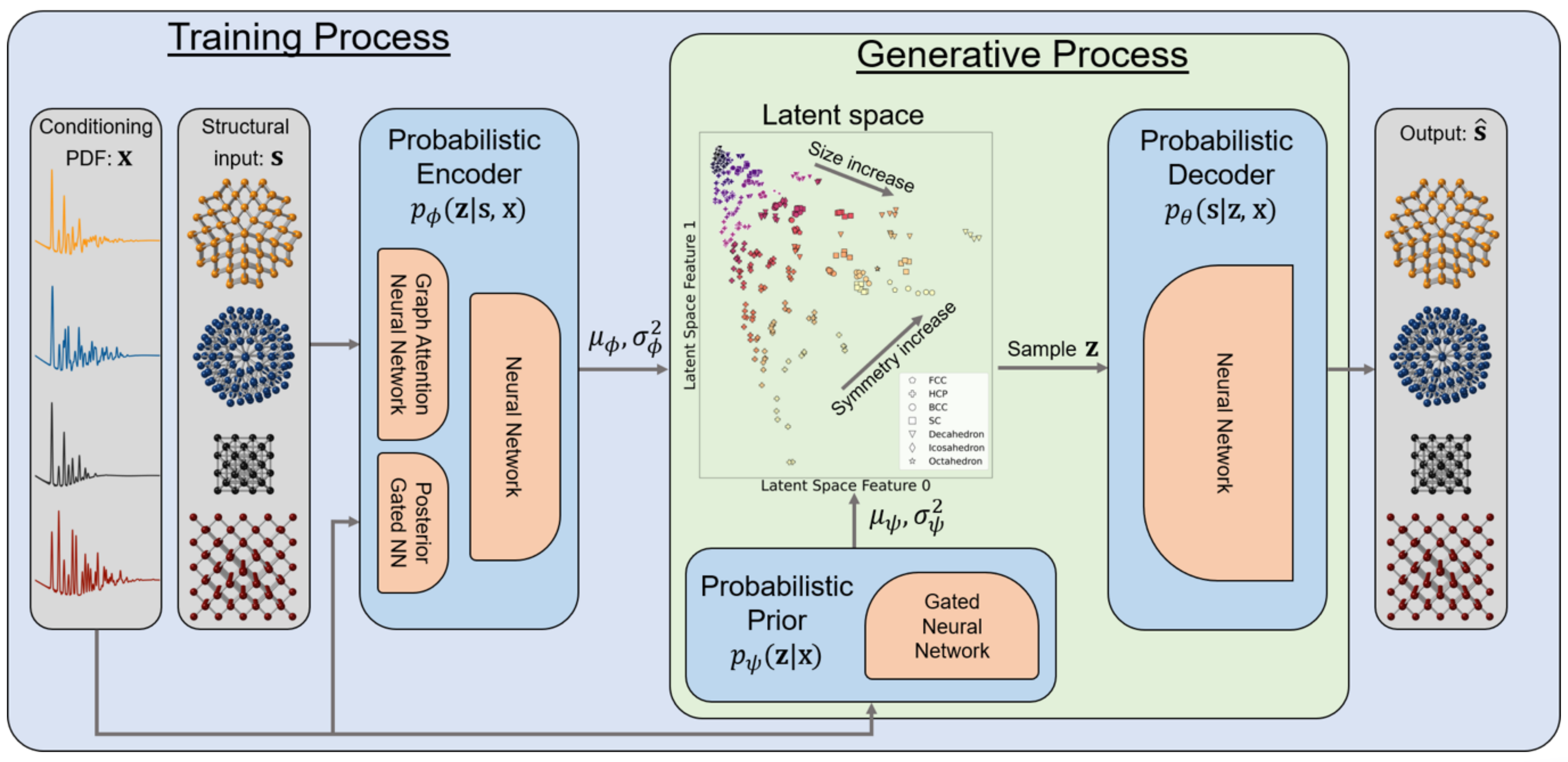
*Both authors contributed equally to this research

1: Introduction

During the past decades, research in materials science has been accelerated by the rapid development of synchrotron and neutron sources.[1] It is now common to measure terabytes of data with *in situ* and *in operando* experiments in order to study reactions in real time.[2, 3] However, conventional data analysis approaches using minimization techniques as least-squares fitting algorithms cannot keep up with the amount of measured data. The data analysis is, therefore, considered the bottleneck of materials science.[4, 5] With the continuing advancement of synchrotrons,[6] we need faster tools to analyse the data in order to fully utilize the full power of modern instruments. One way to overcome this bottleneck is by utilizing the computational efficiency of modern machine learning approaches.

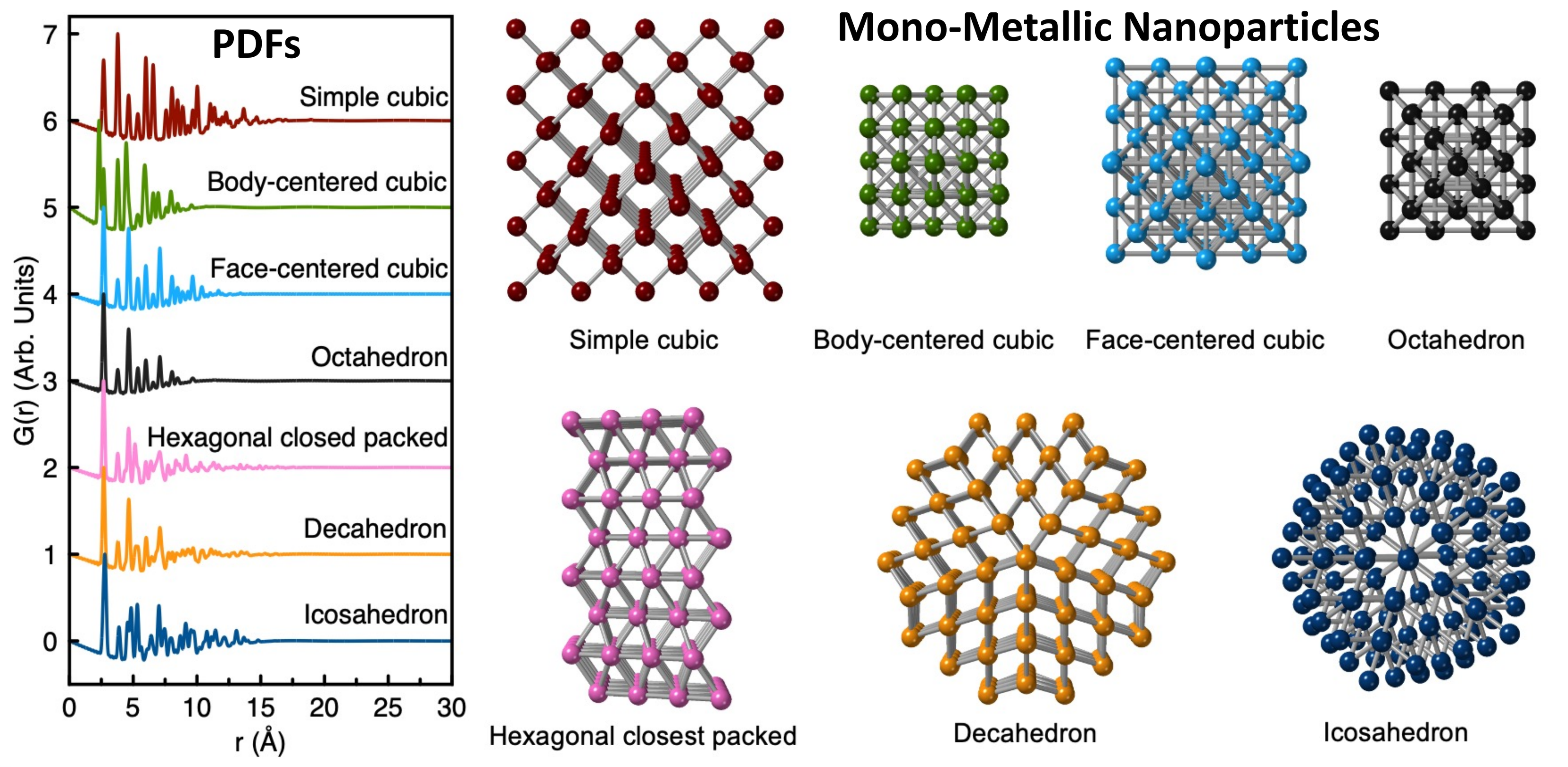
2: Conditioned Variational Autoencoder

We use a Conditioned Variational Autoencoder (CVAE) to automate data analysis. Graph based representation allows an elaborating input of the structure.



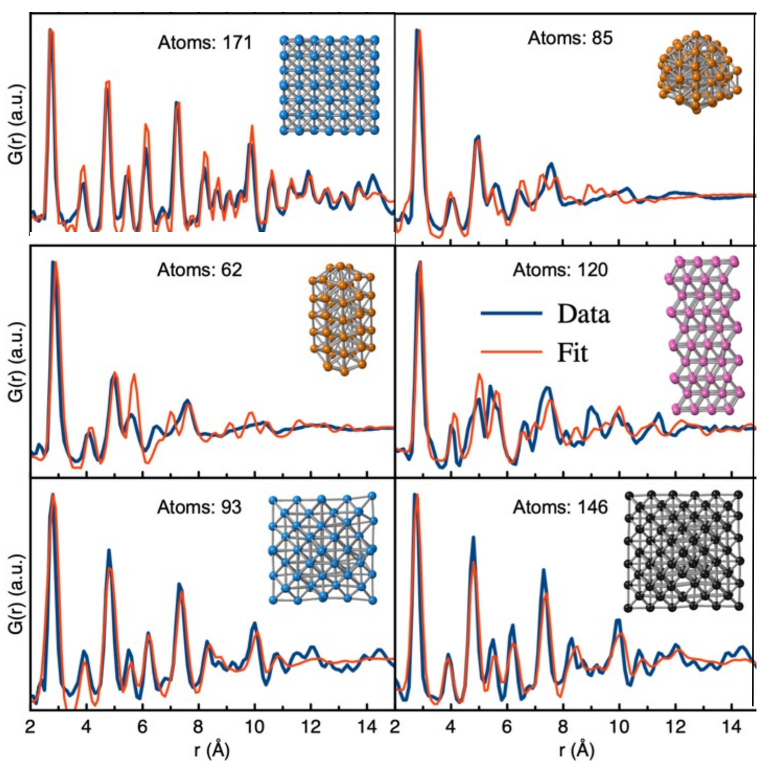
3: Mono-Metallic Nanoparticles

We train the CVAE on simulated Pair Distribution Function (PDF) data of mono-metallic nanoparticles.

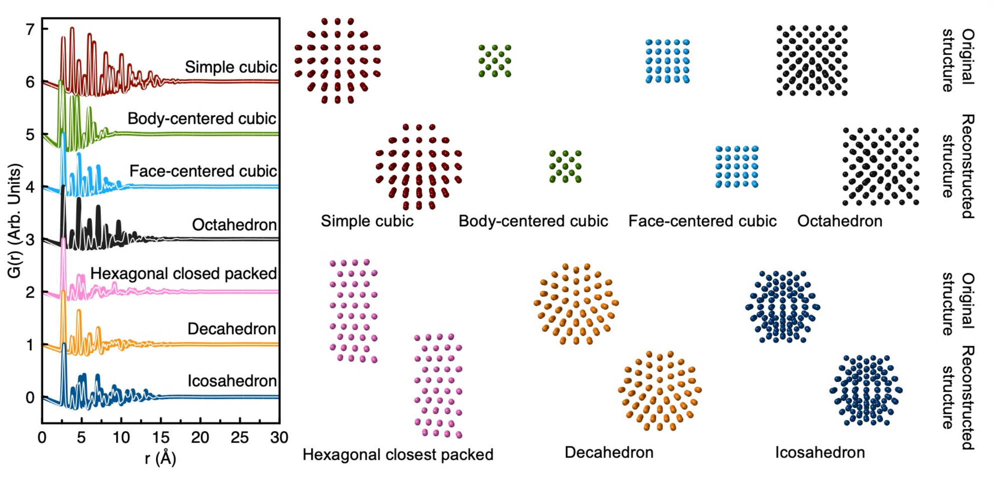


4: Instant Structure Classification

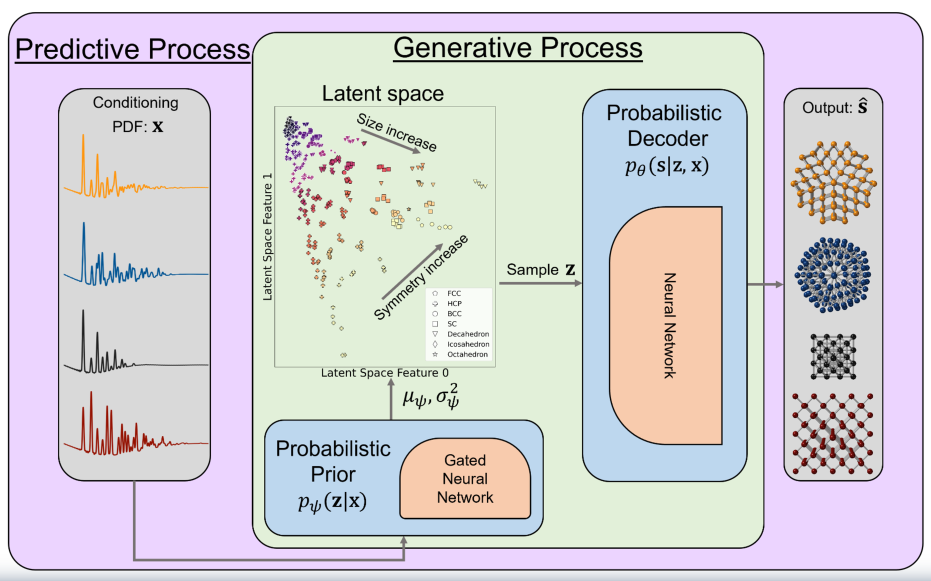
The reconstructed structures from 6 experimental datasets provide a good description the PDF.



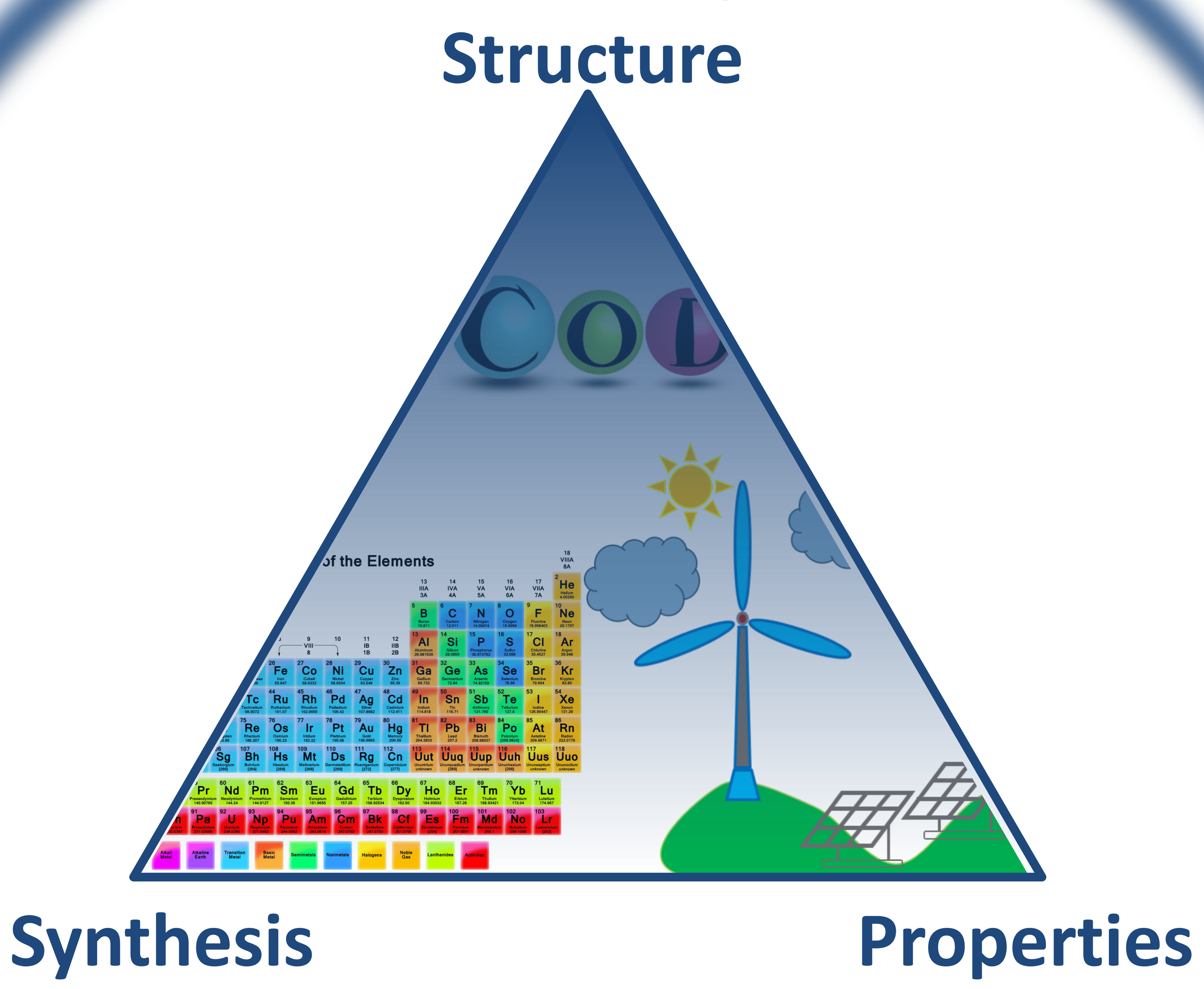
Reconstructions from simulated data.



The CVAE takes a PDF as input and predicts which Mono-metallic nanoparticle it is from.



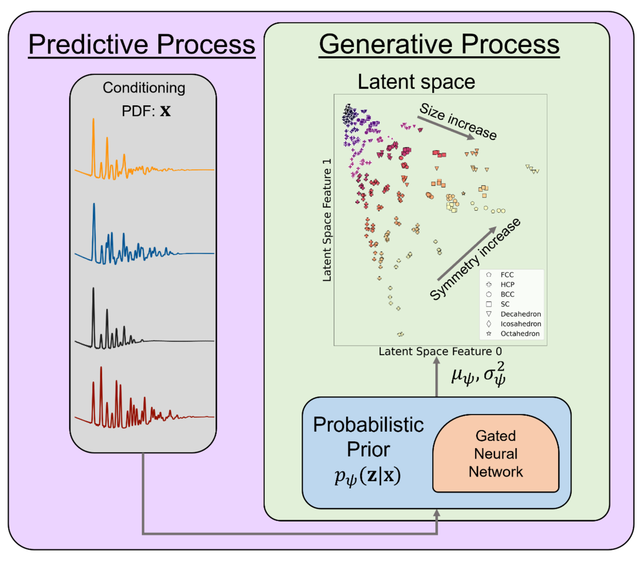
6: Future Perspectives



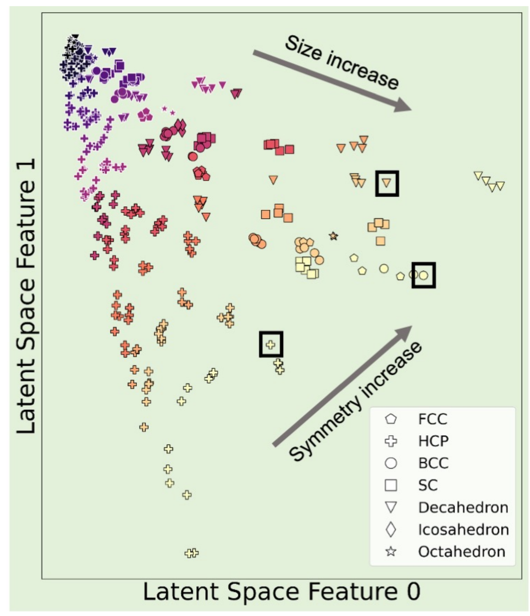
<u>The CVAE can be used to link synthesis, structure and properties by including all the parameters in the training phase of the algorithm.</u>

5: Analysing Similarities and Clustering Trends of the Mono-Metallic Nanoparticles

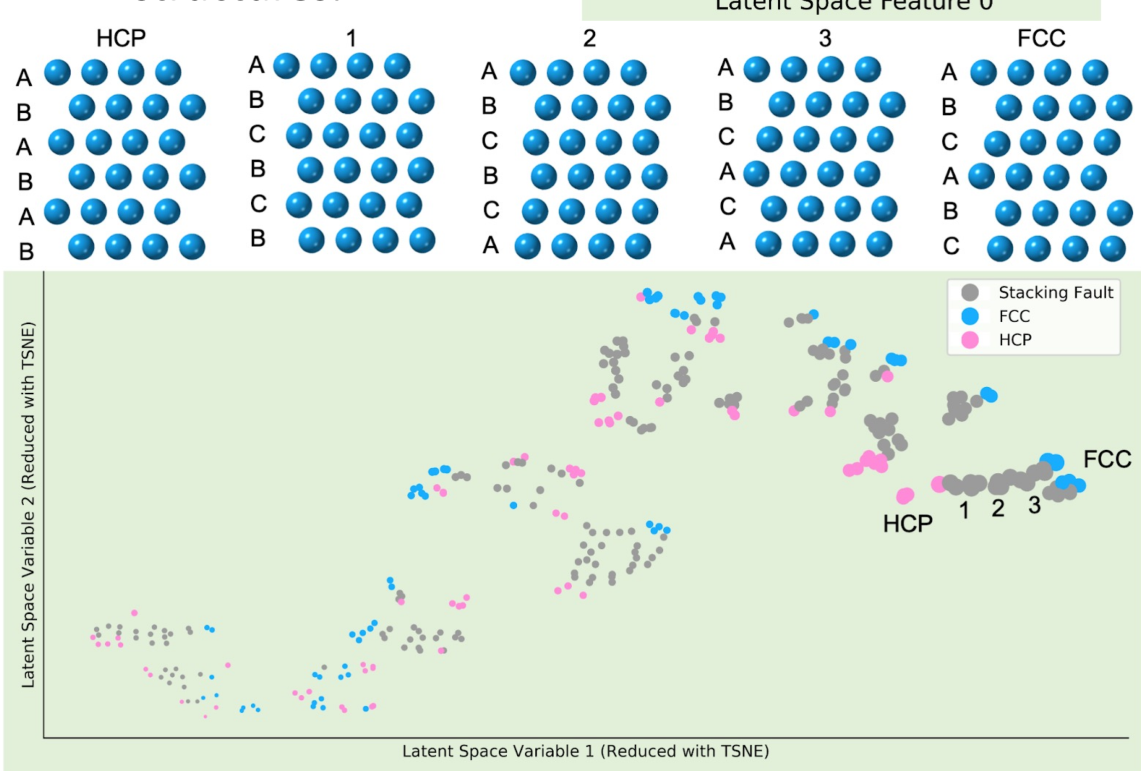
The CVAE can create a low-dimensional space where many PDFs are represented.



The low-dimensional space can be used to analyse similarities and clustering trends.



The CVAE can meaningfully place stacking faulted structures between *fcc* and *hcp* structures.



Acknowledgments

This work is part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, (Grant agreement no. 804066). Work in the Billinge group was supported by the US National Science Foundation through grant DMR-1922234.

References

1. Wang, C. et al., *Synchrotron Big Data Science,* Small **2018**
2. Jensen, K. M. Ø. et al., *X-Ray Diffraction Computed Tomography for Structural Analysis of Electrode Materials in Batteries,* Journal of The Electrochemical Society **2015**
3. Ravnsbæk, D. B. et al., *Extended Solid Solutions and Coherent Transformations in Nanoscale Olivine Cathodes,* Nano Lett. **2014**
4. Agrawal, A. & Choudhary, A., *Perspective: Materials informatics and big data: Realization of the "fourth paradigm" of science in materials science*, APL Materials **2016**
5. Butler, K. T. et al., *Machine learning for molecular and materials science,* Nature **2018**
6. Pacchioni, G., *An upgrade to a bright future*, Nature Reviews Physics **2019**