DeepStruc: Towards structure solution from pair distribution function data using deep generative models

Andy S. Anker[*1], Emil T. S. Kjær[*1], Marcus N. Weng[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
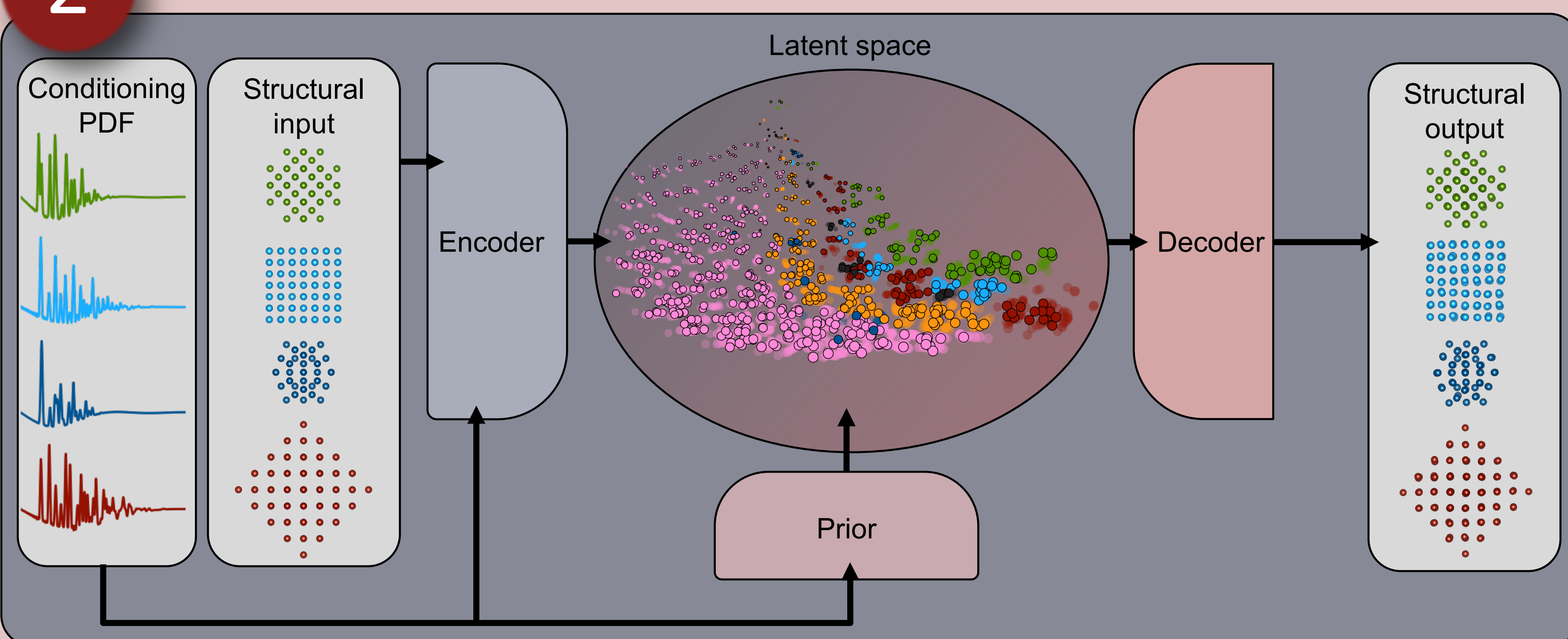
1 Introduction

Structure solution of nanostructured materials that have limited long-range remains a bottleneck in materials development. We present a deep learning algorithm, DeepStruc, that can solve a simple nanoparticle structure directly from a Pair Distribution Function (PDF) obtained from total scattering data by using a conditional variational autoencoder (CVAE). We first apply DeepStruc to PDFs from seven different structure types of monometallic nanoparticles, and show that structures can be solved from both simulated and experimental PDFs, including PDFs from nanoparticles that are not present in the training distribution. We also apply DeepStruc to a system of *hcp*, *fcc* and stacking faulted nanoparticles, where DeepStruc recognizes stacking faulted nanoparticles as an interpolation between *hcp* and *fcc* nanoparticles and is able to solve stacking faulted structures from PDFs. Our findings suggests that DeepStruc is a step towards a general approach for structure solution of nanomaterials.
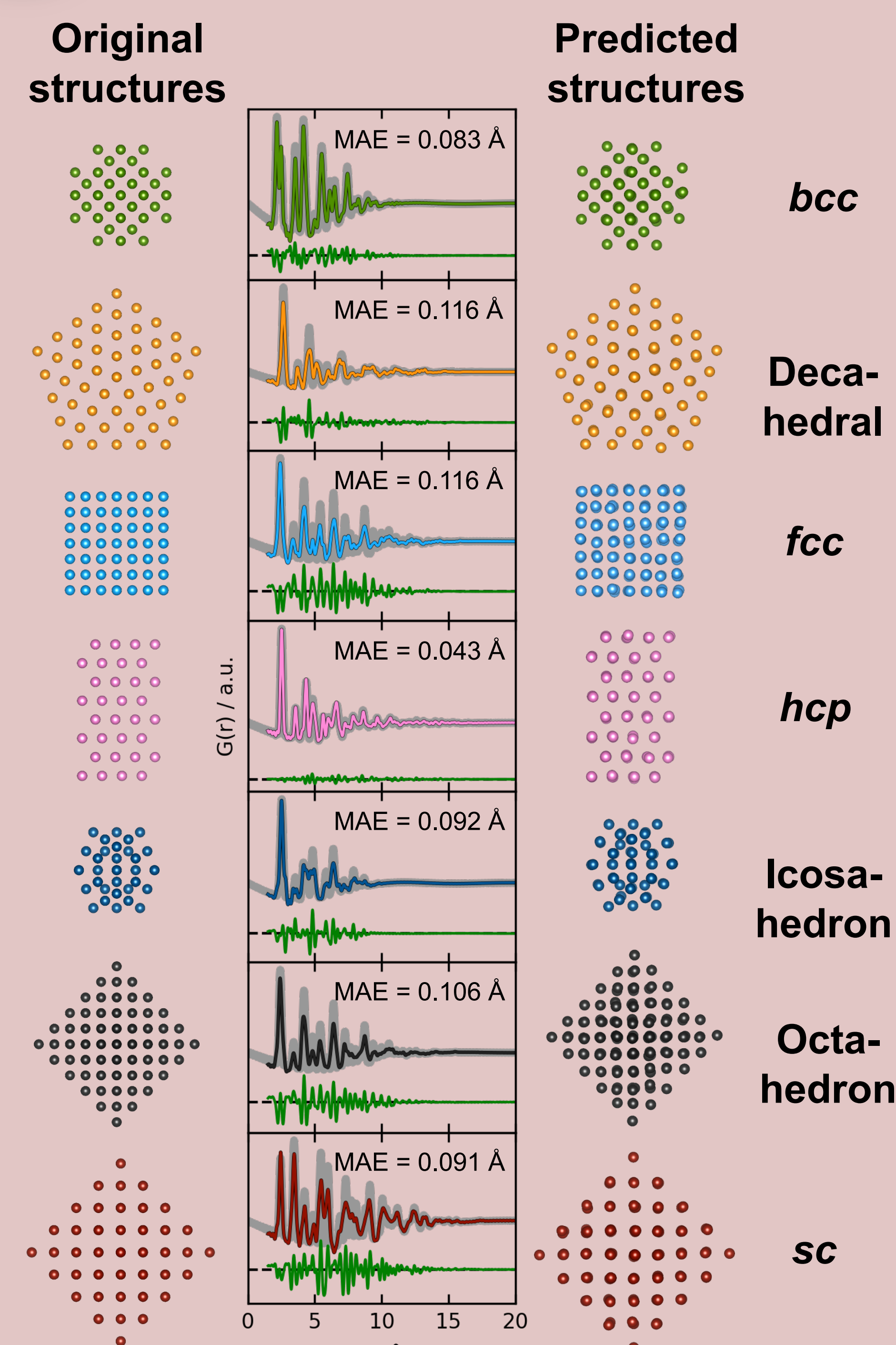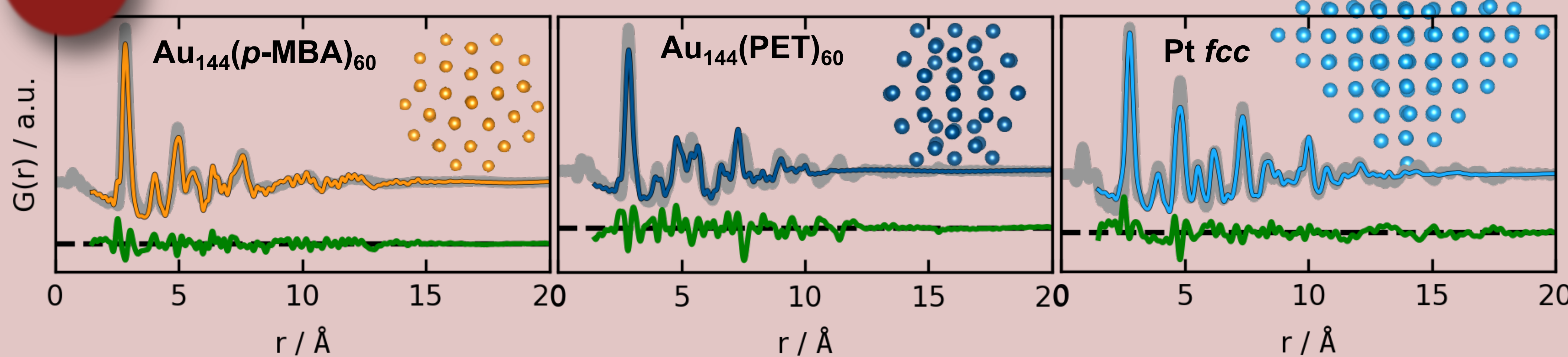
2 DeepStruc architecture

3 Solving simulated PDFs with DeepStruc

Original structures — Predicted structures: bcc (MAE = 0.083 Å), Deca-hedral (MAE = 0.116 Å), fcc (MAE = 0.116 Å), hcp (MAE = 0.043 Å), Icosa-hedron (MAE = 0.092 Å), Octa-hedron (MAE = 0.106 Å), sc (MAE = 0.091 Å)

MAE – mean absolute error of the atomic coordinates

4 Solving experimental PDFs with DeepStruc

$Au_{144}(p\text{-MBA})_{60}$ — $Au_{144}(PET)_{60}$ — Pt *fcc*

5 DeepStruc can extrapolate to data far from the training distribution

FCC with ~203 atoms
FCC with ~371 atoms
FCC with ~1368 atoms

Training area
200 atoms
Size + symmetry Extrapolation area
~400 atoms
Size Extrapolation area
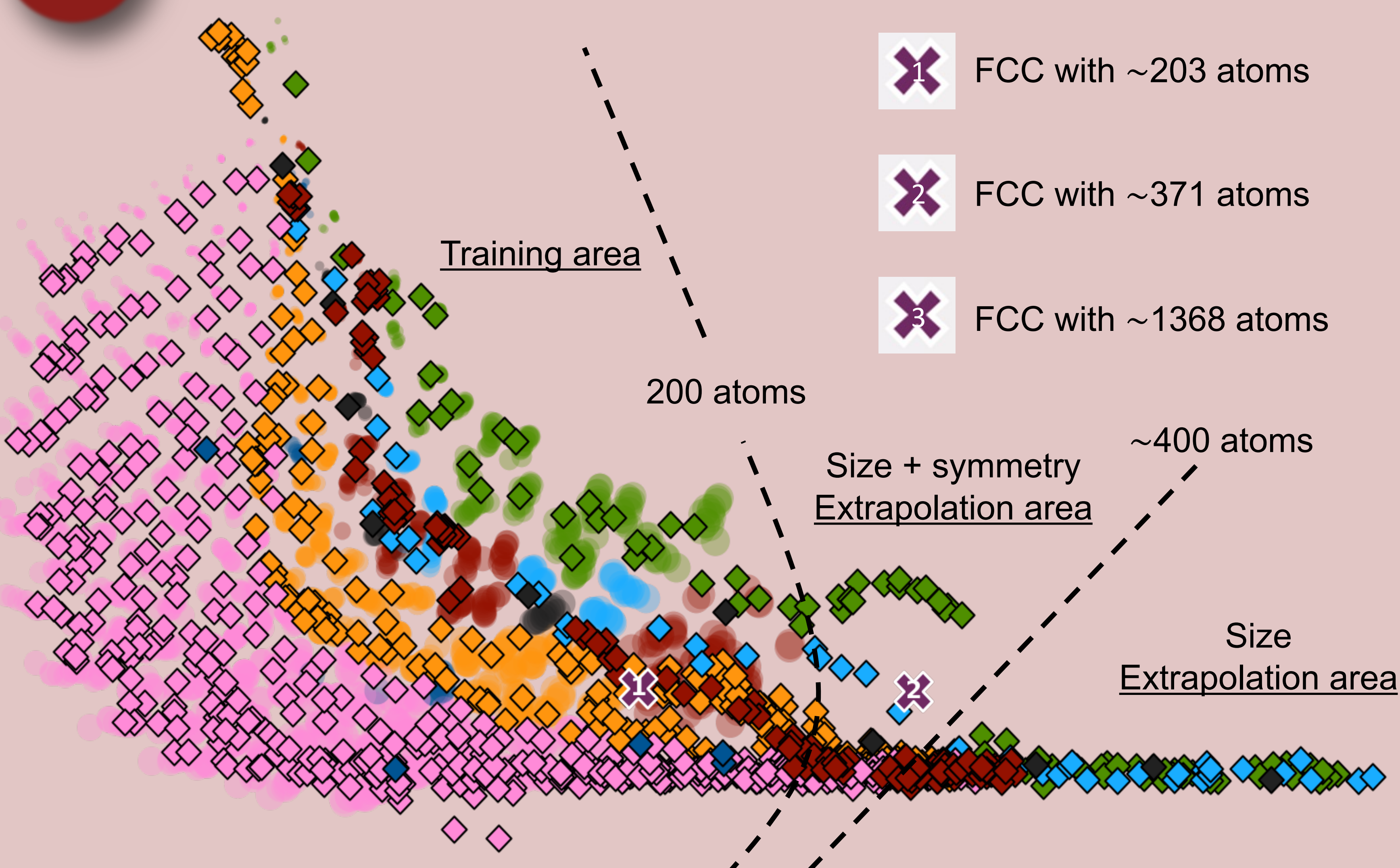
Legend:
- FCC
- BCC
- SC
- Octahedron
- Icosahedron
- Decahedron
- HCP
- Experimental data

6 DeepStruc identifies stacking faulted structures to be in between FCC and HCP in the latent space

fcc — hcp — SF

MAE = 0.038 Å, MAE = 0.023 Å, MAE = 0.031 Å, MAE = 0.029 Å

Acknowledgments