Development of an XRD Spectrum Generation Model for Crystal Structure Prediction Supporting Multi-site Substitution

Koya Watanabe
Nara Intitute of Science and Technology

Elemental substitution techniques are important in exploring new compositions of inorganic compounds. Recently, examples of elemental substitutions where the number of sites and oxidation states of elements differ before and after substitution have been reported that we define as multi-site substitution in this study. While multi-site substitution is expected to provide a new method for exploration in materials development, methods for directly predicting the crystal structures of obtained substitution candidate compounds remain unestablished. Existing machine learning models make predictions based on similarity to known structures, making accurate predictions difficult for multi-site substitutions that may deviate significantly from known structures. Therefore, this study proposes a crystal structure prediction method using two-stage machine learning that combines XRD spectrum generation and crystal structure identification. By generating XRD spectra that consider changes in crystal structure before and after substitution and predicting crystal structures from these outputs, we address the problem of dependence on known structures and enable prediction of after substitution structures. We constructed a Variational Autoencoder (VAE) conditioned on chemical composition and developed a Latent Diffusion Model to enable generation considering polymorphism. We evaluated the accuracy of generated XRD spectra and assessed the interpretability of the latent space.