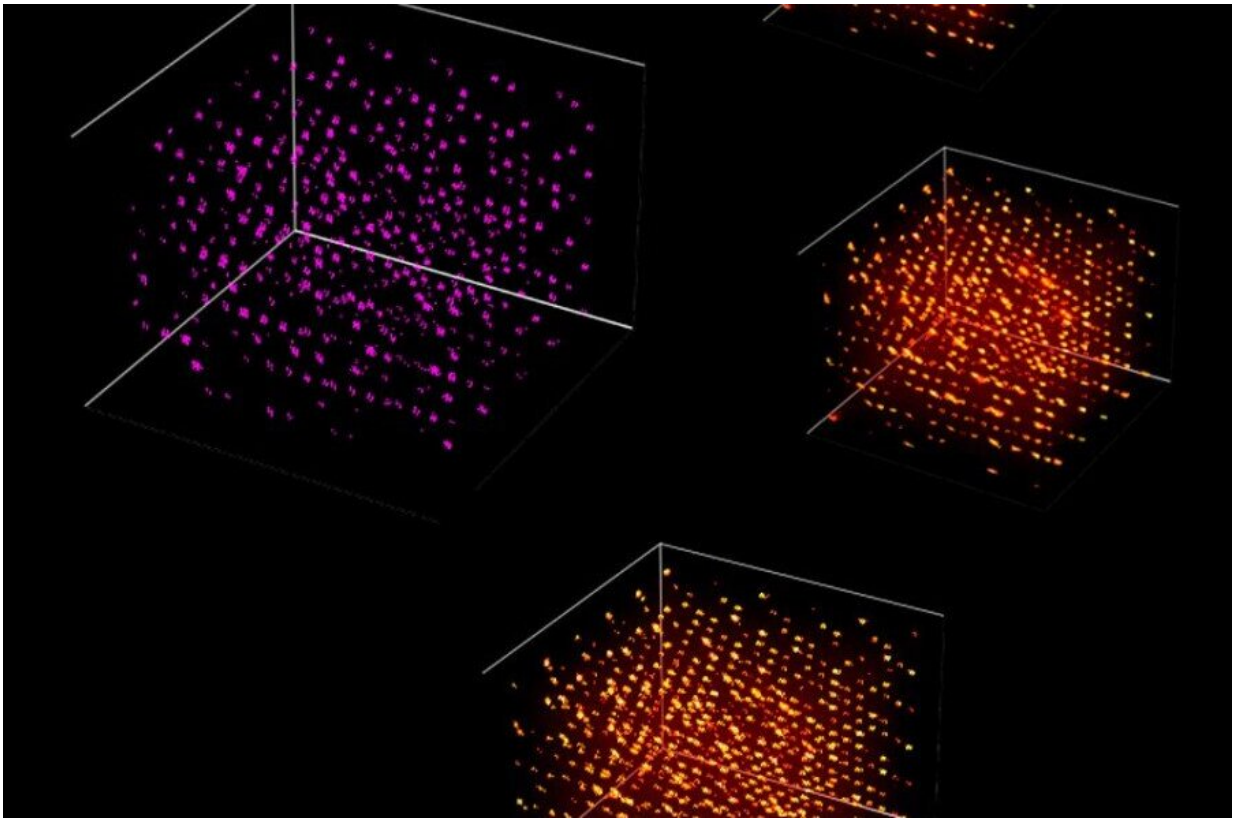# Harnessing machine learning to analyze quantum material

June 14 2022, by Kate Blackwood



An example of 3D X-ray diffraction data going through a phase transition upon cooling. The magenta plot shows special points associated with charge density wave formation as they were revealed by the machine learning algorithm X-TEC. Credit: Krishna Mallayya/Provided

Electrons and their behavior pose fascinating questions for quantum

physicists, and recent innovations in sources, instruments and facilities allow researchers to potentially access even more of the information encoded in quantum materials.

However, these research innovations are producing unprecedented—and until now, indecipherable—volumes of data.

"The information content in a piece of material can quickly exceed the total information content in the Library of Congress, which is about 20 terabytes," said Eun-Ah Kim, professor of physics in the College of Arts and Sciences, who is at the forefront of both quantum materials research and harnessing the power of machine learning to analyze data from quantum material experiments.

"The limited capacity of the traditional mode of analysis—largely manual—is quickly becoming the critical bottleneck," Kim said.

A group led by Kim has successfully used a machine learning technique developed with Cornell computer scientists to analyze massive amounts of data from the quantum metal $Cd_2Re_2O_7$, settling a debate about this particular material and setting the stage for future machine learning aided insight into new phases of mater.

The paper, "Harnessing Interpretable and Unsupervised Machine Learning to Address Big Data from Modern X-ray Diffraction," published June 9 in *Proceedings of the National Academy of Sciences*.

Cornell physicists and computer scientists collaborated to build an unsupervised and interpretable machine learning algorithm, XRD Temperature Clustering (X-TEC). The researchers then applied X-TEC to investigate key elements of the pyrochlore oxide metal, $Cd_2Re_2O_7$.

X-TEC analyzed eight terabytes of X-ray data, spanning 15,000 Brillouin

zones (uniquely defined cells), in minutes.

"We used unsupervised machine learning algorithms, which are a perfect fit to translate high dimensional data into clusters that make sense to humans," said Kilian Weinberger, professor of computer science in the Cornell Ann. S Bowers College of Computing and Information Science.

Thanks to this analysis, the researchers discovered important insights into electron behavior in the material, detecting what is known as the pseudo-Goldstone mode. They were trying to understand how atoms and electrons position themselves in an orderly fashion to optimize the interaction within the astronomically large "community" of electrons and atoms.

"In complex crystalline materials, a specific structure of multiple atoms, the unit cell, repeats itself in a regular arrangement like in a high-rise apartment complex," Kim said. "The repositioning we discovered happens at a scale of each apartment unit, across the entire complex."

Because the arrangement of the units stays the same, she said, it is difficult to detect this repositioning by watching from the outside. However, the repositioning almost spontaneously breaks a continuous symmetry, which results in a pseudo-Goldstone mode.

"The existence of pseudo-Goldstone mode can reveal the secret symmetries in the system that can be hard to see otherwise," Kim said. "Our discovery was enabled by X-TEC."

This discovery is significant for three reasons, Kim said. First, it shows that machine learning can be used to analyze voluminous X-ray powder diffraction (XRD) data, serving as a prototype for applications of X-TEC as it scales up. X-TEC, available to researchers as a software package, will be integrated into the synchrotron as an analysis tool at the

Advanced Photon Source and at the Cornell High Energy Synchrotron Source.

Second, the discovery settles a debate concerning the physics of $Cd_2Re_2O_7$.

"To the best of our knowledge, this is the first instance of the detection of a Goldstone mode using XRD," Kim said. "This atomic scale insight into fluctuations in a complex quantum material will be only the first example of answering key scientific questions accompanying any discovery of new phases of matter … using information-rich voluminous diffraction data."

Third, the discovery showcases what collaboration between physicists and computer scientists can accomplish.

"The mathematical inner workings of machine-learning algorithms are often not unlike models in physics but applied to high dimensional data," Weinberger said. "Working with physicists is a lot of fun, because they are so good at modeling the natural world. When it comes to data modeling, they truly hit the ground running."

Provided by Cornell University