Machine-learning analysis of X-ray data picks out key catalytic properties
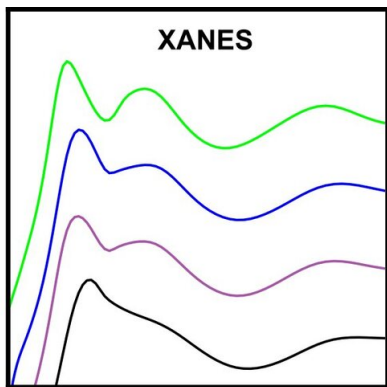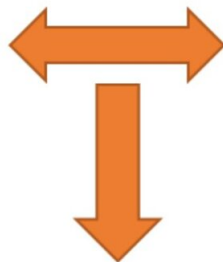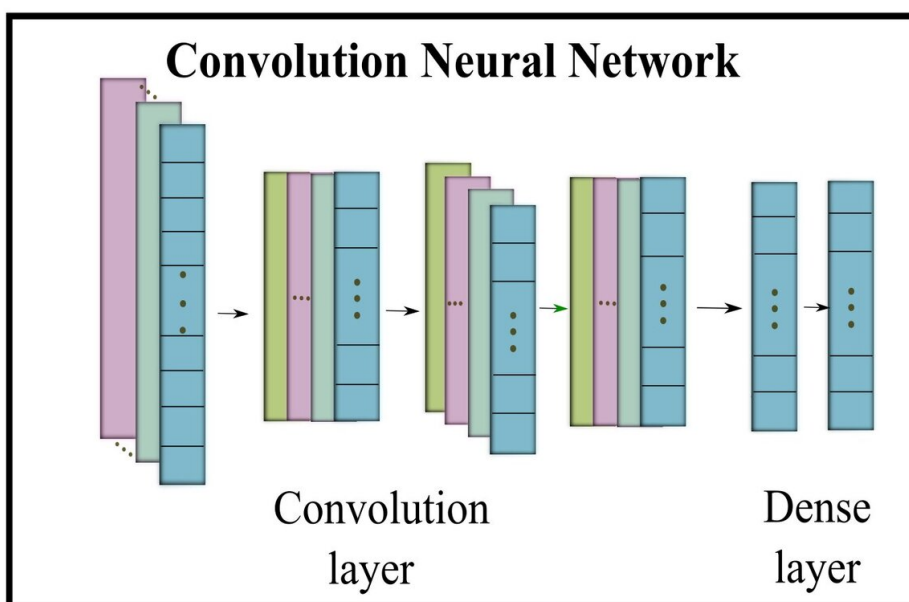
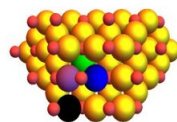October 23 2019, by Karen McNulty Walsh

XANES

Theory

TRAINING
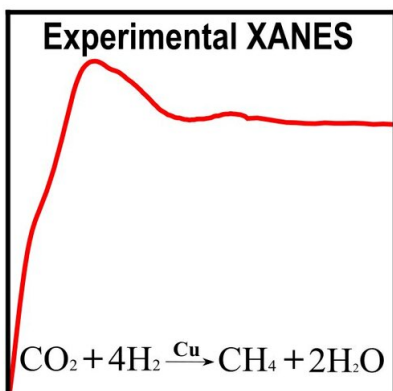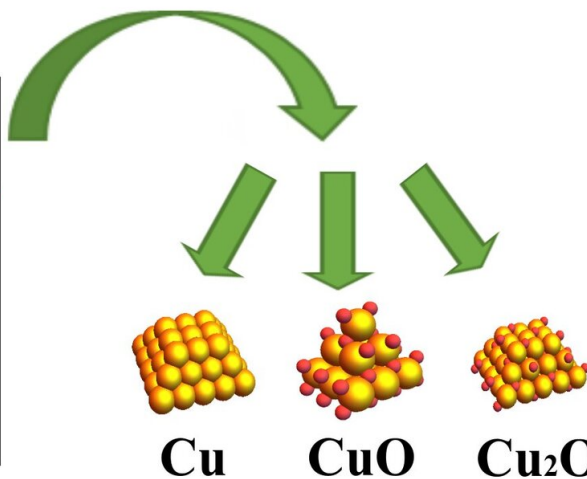
Convolution Neural Network

Convolution
layer

Dense
layer

APPLICATION

Experimental XANES

$$CO_2 + 4H_2 \xrightarrow{Cu} CH_4 + 2H_2O$$

Operando Experiment

Cu   **CuO**   **Cu₂O**

This schematic shows how spectra derived from theory calculations using known structures (top) can be used to train a neural network (center), which can then use its assembled "knowledge" to translate spectra measured in operando experiments into the corresponding structures (bottom). Credit: Brookhaven National Laboratory

Scientists seeking to design new catalysts to convert carbon dioxide (CO_2) to methane have used a novel artificial intelligence (AI) approach to identify key catalytic properties. By using this method to track the size, structure, and chemistry of catalytic particles under real reaction conditions, the scientists can identify which properties correspond to the best catalytic performance, and then use that information to guide the design of more efficient catalysts.

"Improving our ability to convert CO_2 to methane would 'kill two birds with one stone' by making a sustainable non-fossil-fuel energy source that can be easily stored and transported while reducing carbon emissions," said Anatoly Frenkel, a chemist with a joint appointment at the U.S. Department of Energy's Brookhaven National Laboratory and Stony Brook University.

Frenkel's group has been developing a machine-learning approach to extract catalytic properties from X-ray signatures of catalysts collected as chemicals are transformed in reactions. The current analysis is described in a paper just published in the *Journal of Chemical Physics*, based on X-ray data collected at DOE's Argonne National Laboratory.

The team of Argonne senior chemist Stefan Vajda, now at the J. Heyrovský Institute of Physical Chemistry in Prague, prepared size-

selective clusters of copper atoms. Then they used mass spectrometry and x-rays at Argonne's Advanced Photon Source (APS) to study how various size clusters performed in the reaction and how their [oxidation state](#) evolved during the reaction of carbon dioxide with hydrogen.

Copper has shown promise as a catalyst that can lower the temperature of the CO_2-to-methane reaction. Size-selective copper clusters may also help drive the reaction efficiently to the desired outcome—selectively producing just methane and water vapor—without channeling reactants down a variety of pathways toward other products.

"There are, broadly speaking, two major challenges towards implementing this idea," said Frenkel. "First is the lack of knowledge of the structure of the prepared clusters; the smaller they are, the more variations there may be in shapes and structures—even when the number of atoms in each [cluster](#) is the same.

"Second, even if we start the reaction with clusters of a certain size and shape, they may transform beyond recognition during the reaction to various forms of oxides."

Some of the oxides might enhance reactivity; others might impede the reaction. To understand how the catalyst works, the scientists need to know what types of oxides form during the reaction—and how they affect catalytic performance.
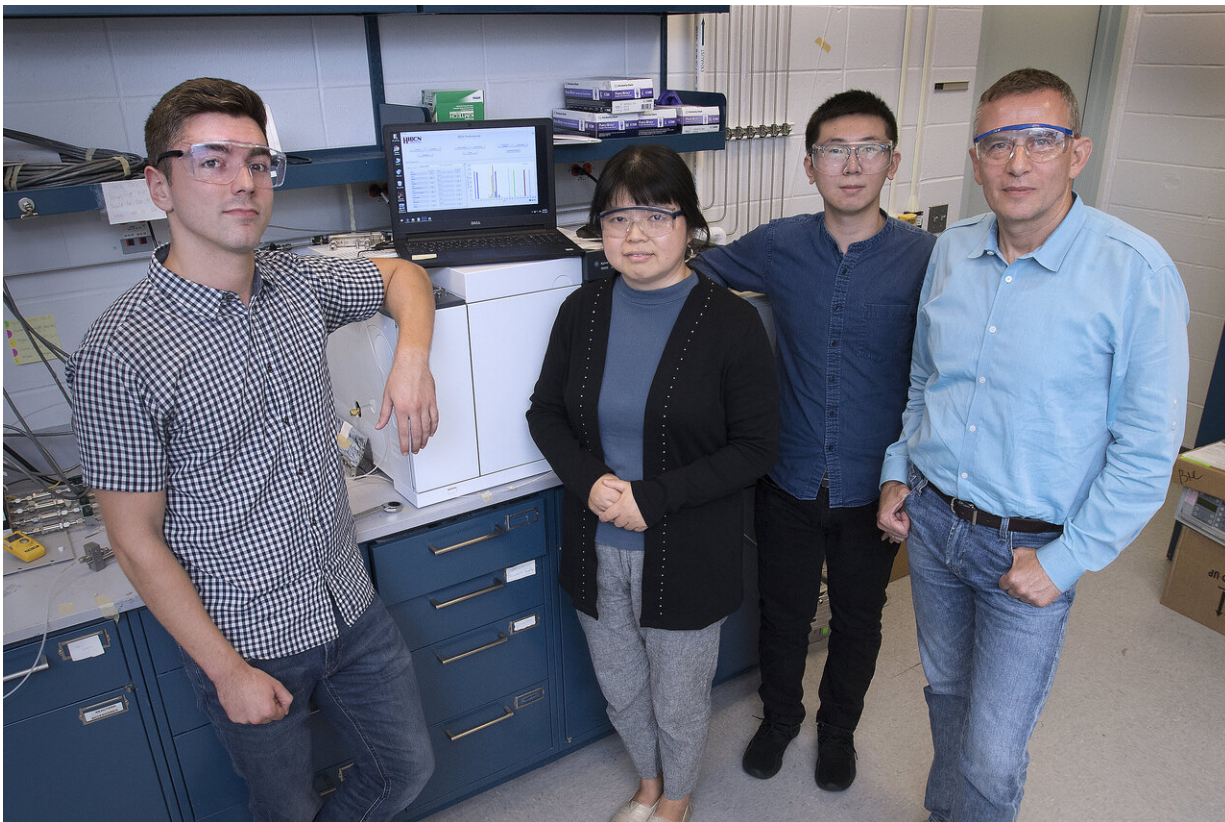
Research team members: Stony Brook University (SBU) graduate student Nicholas Marcella, Brookhaven Lab chemist Ping Liu, SBU graduate student Yang Liu, and SBU-Brookhaven Lab joint appointee Anatoly Frenkel. Credit: Brookhaven National Laboratory

Collecting spectral data

X-ray data collected while analyzing catalysts at the APS or other synchrotron light sources (including the National Synchrotron Light Source II at Brookhaven Lab) contains a wealth of information about chemical composition and structure because these properties determine how x-rays interact with the sample. But extracting that information from data collected from ultra-dilute samples made up of tiny clusters (containing as few as four atoms per cluster) presents a big challenge.

"These samples are too small for X-ray scattering or imaging methods commonly used to characterize materials on the nanoscale," Frenkel said.

Instead, the scientists analyzed how individual copper atoms absorb the synchrotron x-rays.

The amount of X-ray energy absorbed tells them how much energy it takes to "kick" an electron out of orbit from each copper atom, which depends on its oxidation state—how many electrons the atom has available to share in forming chemical bonds. The less oxidized the copper atom is (meaning it's holding on to its electrons), the less energy it takes for the x-rays to kick an electron out—because the electrons left behind help shield the escaping electron from the attractive positive charge of the copper nucleus. The more oxidized (with fewer electrons), the more energy it takes to kick a remaining electron out—because the positive pull of the unshielded nucleus is harder to overcome.

The X-ray absorption spectrum therefore contains information about the oxidation state and other details that reveal features of the atomic structure, including how many adjacent atoms each copper atom is bound to. But to extract this information the scientists needed a way to relate the measured spectra to known structural arrays of copper atoms with various oxidation states.

That's where the artificial intelligence comes in. The scientists developed an artificial neural network "trained" to recognize key features in spectra from known structures so it could then find the unknown structures just by analyzing the measured spectra.

Training the network

Developing a library of known structures that they could use to train the

network presented its own challenges. For help, Frenkel's group turned to Ping Liu in Brookhaven's Chemistry Division.

"The structure of the clusters depends very strongly on how the particles interact with the support substrate they are deposited on and the reactive environment," said Liu, a theorist with vast experience modeling catalytic activity. "We have constructed model systems for the supported metal and metal oxide clusters, complex enough to capture the structures and catalytic behaviors during the reactions as observed experimentally," she said. "These operating models provide a strong basis enabling the accuracy and efficiency of the machine learning."

Then the team used numerical methods to generate the spectra these samples would produce—a fairly straightforward approach—and used these theoretically generated spectra to train the neural network.

Once the computer running the neural network program had learned the relationships between the spectral features and the key characteristics of the known clusters—the oxidation states, number of neighbor atoms, and so on—the scientists could feed the spectra measured from their experimental clusters into the network and it would tell them the cluster characteristics for those samples.

Characteristics of the clusters

In the copper catalyst experiment, the scientists used this approach to analyze X-ray absorption spectra from clusters made up of four, twelve, or twenty copper atoms.

"During the reaction, these clusters go through many different oxidation states depending on the stage of the reaction. We collected the spectra at these different stages and used our machine learning approach to identify the different oxidation states of the clusters at different stages

of the reaction. We also correlated the oxidation states with the observed catalytic activity to determine which structures are the best catalysts," Frenkel said.

Data from other experimental methods already existed for the two smaller cluster sizes, so it could serve as a cross-check of the new technique. "This comparison showed that we were able to recognize the oxidation states corresponding to metallic copper or the different types of metal oxide using our neural network approach," Frenkel said.

This was the first time Frenkel had applied his machine learning approach to solving anything other than pure metallic clusters.

"It's the first time we were able to train the network to recognize different types of oxides," he said.

It is also the first time Frenkel's method has been used in a predictive capacity—to determine the oxidation states and other characteristics of the 20-copper-atom clusters, for which no other data exist.

It turns out that the most catalytically active state of the copper catalyst is a mixture of metallic clusters (where copper is bound only to other copper atoms) and two different copper-oxides (CuO and Cu2O).

"There are many reactions where the catalyst turns out to be most active when it is neither fully oxidized nor fully reduced," Frenkel said. "Those clusters that are able to form this mixture of the three different states in the right proportions will be the most active."

Frenkel's group is continuing their analysis to learn more about the catalytic mechanism and will publish their results in the future.

  More information: Yang Liu et al, Mapping XANES spectra on

structural descriptors of copper oxide clusters using supervised machine learning, *The Journal of Chemical Physics* (2019). DOI: 10.1063/1.5126597

Janis Timoshenko et al. "Inverting" X-ray Absorption Spectra of Catalysts by Machine Learning in Search for Activity Descriptors, *ACS Catalysis* (2019). DOI: 10.1021/acscatal.9b03599

Janis Timoshenko et al. Subnanometer Substructures in Nanoassemblies Formed from Clusters under a Reactive Atmosphere Revealed Using Machine Learning, *The Journal of Physical Chemistry C* (2018). DOI: 10.1021/acs.jpcc.8b07952

Provided by Brookhaven National Laboratory