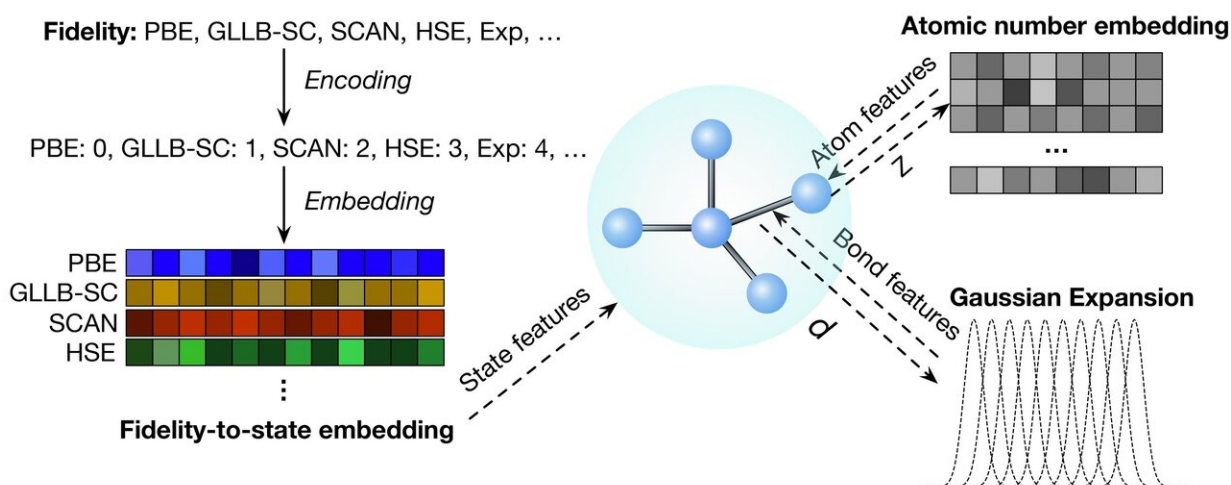


# New method makes better predictions of material properties using low quality data

January 14 2021



A schematic of the multi-fidelity graph network approach to predicting the properties of new materials. Credit: Materials Virtual Lab at UC San Diego

Advancements in energy technologies, healthcare, semiconductors and food production all have one thing in common: they rely on developing new materials—new combinations of atoms—that have specific properties enabling them to perform a needed function. In the not-too-distant past, the only way to know what properties a material had was by performing experimental measurements or using very expensive computations.

More recently, scientists have been using machine learning algorithms to rapidly predict the properties that certain arrangements of atoms would have. The challenge with this approach is it requires a lot of highly accurate data to train the model, which often does not exist.

By combining large amounts of low-fidelity data with the smaller quantities of high-fidelity data, nanoengineers from the Materials Virtual Lab at UC San Diego have developed a new machine learning method to predict the [properties of materials](#) with more accuracy than existing models. Crucially, their approach is also the first to predict the properties of disordered materials—those with atomic sites that can be occupied by more than one element, or can be vacant. They detailed their multi-fidelity graph networks approach on January 14 in *Nature Computational Science*.

"When you are designing a new material, one of the key things you want to know is if the material is likely to be stable, and what kind of properties it has," said Shyue Ping Ong, a professor of nanoengineering at the UC San Diego Jacobs School of Engineering and the paper's corresponding author. "The [fundamental problem](#) is that valuable [accurate data](#), such as experimental measurements, is difficult to come by, even though we have large databases of less accurate computed properties. Here, we try to get the best of both worlds—combine the large low-fidelity data and the smaller high-fidelity data to improve the models' accuracy in high value predictions."

While other multi-fidelity approaches exist, these methods do not scale well or are limited to only two fidelities of data. They are not as accurate or dynamic as this new multi-fidelity graph network approach, which can work with an unlimited number of data fidelities and can be scaled up very quickly.

In this paper, the nanoengineers looked specifically at materials' band

gaps—a property used to determine electrical conductivity, the color of the material, solar cell efficiency, and more—as a proof-of-concept. Their multi-fidelity graph networks led to a 22-45% decrease in the mean absolute errors of experimental band-gap predictions, compared to a traditional single-fidelity approach. The researchers also showed that their approach can predict high-fidelity molecular energies accurately as well.

"There is no fundamental limitation as to what properties this can be applied to," said Ong. "The question is which kind of properties we have data on."

In the near term, Ong's team plans to use this new method to develop better materials for energy storage, photovoltaic cells and semiconductor devices.

While predicting the properties of ordered materials, the team made another serendipitous discovery—in the graph deep learning model they use, atomic attributes are represented as a learned length-16 embedding vector. By interpolating these learned embedding vectors, the researchers found they were able to also create a predictive model for disordered materials, which have atomic sites that can be occupied by more than one element or can be vacant at times, making them harder to study using traditional methods.

"While the bulk of computational and machine learning works have focused on ordered materials, disordered compounds actually form the majority of known materials," said Chi Chen, an assistant project scientist in Ong's lab, and first author of the paper. "Using this approach, multi-fidelity graph network models can reproduce trends in the band gaps in disordered materials to good accuracy."

This opens the door to much faster and more accurate design of [new](#)

[materials](#) to meet key societal needs.

"What we show in this work is you can actually adapt a machine learning algorithm to predict the properties of disordered materials. In other words, now we are able to do materials discovery and prediction across the entire space of both ordered and disordered materials rather than just ordered materials," said Ong. "As far as we know, that is a first."

**More information:** *Nature Computational Science*, [DOI: 10.1038/s43588-020-00002-x](#)

Provided by University of California - San Diego

Citation: New method makes better predictions of material properties using low quality data (2021, January 14) retrieved 20 March 2024 from <https://techxplore.com/news/2021-01-method-material-properties-quality.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--