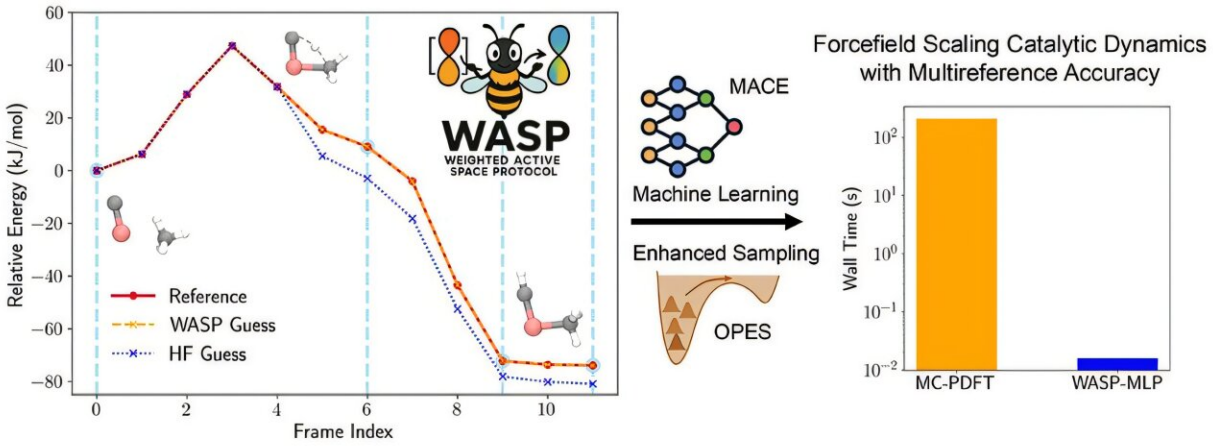
Machine learning and quantum chemistry unite to simulate catalyst dynamics

September 16 2025



Accuracy and speedup achieved with the Weighted Active Space Protocol (WASP) for methane activation on titanium carbide. Credit: Seal et al.

Catalysts play an indispensable role in modern manufacturing. More than 80% of all manufactured products, from pharmaceuticals to plastics, rely on catalytic processes at some stage of production. Transition metals, in particular, stand out as highly effective catalysts because their partially filled d-orbitals allow them to easily exchange electrons with other

molecules. This very property, however, makes them challenging to model accurately, requiring precise descriptions of their electronic structure.

Designing efficient transition-metal catalysts that can perform under realistic conditions requires more than a static snapshot of a reaction. Instead, we need to capture the dynamic picture—how molecules move and interact at different temperatures and pressures, where atomic motion fundamentally shapes catalytic performance.

To meet this challenge, the lab of Prof. Laura Gagliardi at the University of Chicago Pritzker School of Molecular Engineering (UChicago PME) and Chemistry Department has developed a powerful new tool that harnesses electronic structure theories and machine learning to simulate transition metal catalytic dynamics with both accuracy and speed.

"Over the past decade, machine-learned potentials have significantly advanced the way in which we simulate molecular dynamics, offering speed and scalability. Yet, accurately capturing the electronic structure of transition metal catalysts has remained an unsolved challenge. Our new method bridges this gap by integrating multireference quantum chemistry methods with machine-learned potentials, delivering both accuracy and efficiency," Gagliardi said.

The results are published in *Proceedings of the National Academy of Sciences*.

Enabling machine learning to speed up simulations

Over the past decade, the Gagliardi group has developed multiconfiguration pair-density functional theory (MC-PDFT), a quantum-chemistry method capable of describing the intricate electronic structures of transition metal reactions. While MC-PDFT provides high

accuracy, it is prohibitively slow for simulating the dynamics of catalytic systems—a critical step in predicting how catalysts truly behave under realistic conditions.

To address this challenge, the team turned to machine-learned interatomic potentials (ML-potentials), which can capture molecular dynamics with remarkable efficiency. ML-potentials have been applied widely in materials science, but until now, they had never been successfully combined with multireference methods like MC-PDFT.

The reason lies in a long-standing obstacle: labeling consistency. Machine learning models require unique and reliable property labels—such as energies and forces derived from wave functions—for every molecular geometry along a reaction pathway. For multireference quantum chemistry methods, assigning such labels uniquely had remained an unsolved problem.

To overcome this challenge, Ph.D. student Aniruddha Seal, jointly advised by Gagliardi and Prof. Andrew Ferguson, developed a novel algorithm that generates consistent wave functions for new geometries as a weighted combination of wave functions from previously sampled molecular structures. The closer a new geometry is to a known one, the more strongly its wave function resembles that of the known structure. This approach ensures that every point along a reaction pathway is assigned a unique, consistent wave function, making it possible to train ML-potentials accurately on multireference data.

"Think of it like mixing paints on a palette," Seal explained. "If I want to create a shade of green that's closer to blue, I'll use more blue paint and just a little yellow. If I want a shade leaning toward yellow, the balance flips. The closer my target color is to one of the base paints, the more heavily it influences the mix. WASP works the same way: it blends information from nearby molecular structures, giving more weight to

those that are most similar, to create an accurate prediction for the new geometry."

This innovation forms the basis of the Weighted Active Space Protocol (WASP), a framework that combines the accuracy of MC-PDFT with the efficiency of machine learning, developed through a close collaboration with the Parrinello Group at the Italian Institute of Technology, Genoa, bringing together expertise in electronic structure theory and machine-learned potentials. WASP delivers dramatic speedups: simulations with multireference accuracy that once took months can now be completed in just minutes.

Impact: Bridging accuracy and efficiency in catalyst design

By uniting accuracy and speed, WASP opens the door to designing catalysts that can withstand realistic conditions—high temperatures and high pressures. Transition metals are central to countless large-scale processes, but their complexity has made catalyst rational design challenging.

A prime example is the Haber–Bosch process, where iron serves as the catalyst to convert nitrogen and hydrogen into ammonia. Despite being developed more than a century ago, this iron catalyst still dominates ammonia production worldwide. With WASP, researchers now have the tools to explore alternatives that could increase efficiency, reduce byproducts, and lower the environmental cost.

So far, WASP has been successfully demonstrated for thermally activated catalysis—reactions driven by heat. The next frontier is adapting the method to light-activated reactions, which are essential for the design of new photocatalysts. Photocatalysts hold enormous promise

for technologies, from water treatment to energy production.

The new tool is available publicly at https://github.com/GagliardiGroup/wasp.

More information: Aniruddha Seal et al, Weighted Active Space Protocol for Multireference Machine-Learned Potentials, *Proceedings of the National Academy of Sciences* (2025). DOI: 10.1073/pnas.2513693122