

The Elucidation of Non-equilibrium States of Heterogenous Catalysis by Data-driven Multiscale Simulation: A Case Study of Methanol Synthesis

Harry H. Halim, Ryo Ueda, Yamada Yuki, and Yoshitada Morikawa

大阪大学 大学院工学研究科

1. Introduction

Global warming is world class problem that becomes more worrying over the years. A promising solution comes from the conversion of CO₂ (i.e., greenhouse gas) to higher-value chemical such as methanol which can be achieved by heterogenous catalysis. The advantage of this conversion is two-fold: reducing the amount of CO₂ in atmosphere and producing methanol which can be used to generate electricity through fuel cell [1]. Understanding the non-equilibrium states (i.e., states when the transformation occurs) of catalytic process of methanol synthesis is the key factor in designing the best catalyst. However, observing the non-equilibrium states is inaccessible in most experiments due to the atomistic size and time scale of the events.

Though computational approaches such as Density Functional Theory (DFT) and Molecular Dynamics (MD) can observe detail atomistic events, it suffers from huge tradeoff between speed and accuracy. For instance, DFT can be used to accurately describe the stability of the catalytic system by calculating the energy and the atomic forces. However, it is computationally expensive and typically applied only for models with small size and short observation time. On the other hand, MD method can be used to explicitly simulate the trajectory of thousands to ten thousand of atoms, enabling the direct observation to the non-equilibrium states with

observation time up to seconds. Despite the efficiency of MD for simulating large system, it has unreliable accuracy when dealing with chemical reactions, which are fundamental in the heterogenous catalysis. This originates from the inadequacy of the classical interatomic-potential (i.e., function governing atomic interactions) in MD to take into account the electronic structures of the system, which is eventually obtainable from DFT method.

In the emergence of data science, it is possible to accelerate the MD simulation while maintaining high accuracy at the DFT level. This can be achieved by constructing a machine-learning (ML) interatomic potential by learning the energy and forces of atomic structures defined in a set of DFT data. The ML then acts as a bridge for the low-scale but accurate DFT, and the high-scale but inaccurate MD, resulting in multiscale simulation. Armed with this framework, we conducted direct observation of non-equilibrium states of initial stage of methanol synthesis on Cu-based catalyst, including the three main reactants: CO₂, CO, and H₂. We note that in this project, these three reactants are still treated independently. Interacting all the reactants is much more complicated and becomes our project in the future.

2. Multiscale Simulation

The framework of the multiscale simulation is shown schematically in Fig. 1. There are four important building blocks, namely DFT, ML, MD, and the analysis tool that we called *elucidator*. Each of them is discussed in the following.

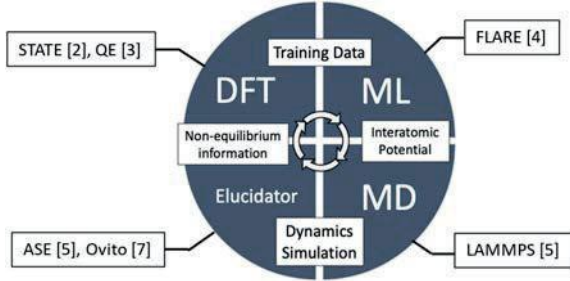


Fig. 1 : The framework of multiscale simulation

2.1 Density Functional Theory (DFT)

DFT is used to start the framework by generating the training data used as input in the ML algorithm. In this project, we tried two popular DFT software, namely STATE [2] and Quantum Espresso [3]. The STATE is used in the multiscale simulations of CO₂ and H₂ while Quantum Espresso is used for the simulations of CO. All of these reactants are interacted with Cu surface. The training data of Cu-H system, Cu-CO system, and Cu-CO₂-H₂ system consist of atomic environments within structures of the Cu surface interacting with various configurations molecules (both in gas and adsorbed states). The typical training structure of Cu-H, Cu-CO, and Cu-CO₂-H are depicted in Fig 2, Fig 3, and Fig 4, respectively. In practice, the DFT software can run using MPI and OMP parallelization and thus increase the efficiency of the calculations especially in large super computer. We note that the generation of database is the most time consuming considering the computational expense of DFT.

2.2 Machine-Learning (ML)

After the database was generated, we proceed to the construction of machine-learning interatomic potential. In this project we used ML algorithm called Gaussian Process Regression (GPR) which is implemented in FLARE software [4]. GPR is chosen since it can provide the uncertainty of prediction based on the database itself. This enables the learning scheme called *active learning* in which the high uncertainty value is used as the criterion for an atomic environment to be included in the database. This learning scheme thus ensures each data is sufficiently different from each other, resulting in a compact and low-correlated database. Further, *on-the-fly* learning scheme can be implemented by employing the force-field in dynamic simulation to generate more high-uncertainty data that can be used to improve the force-field continuously. The GPR that is implemented in FLARE can use OMP parallelization in a single node.

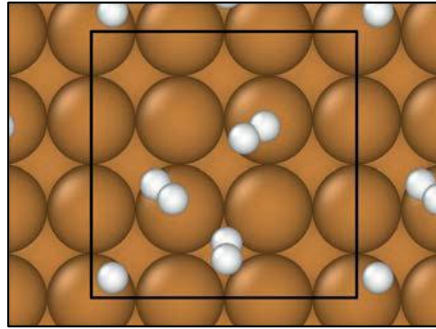


Fig. 2 : The Cu-H training structure

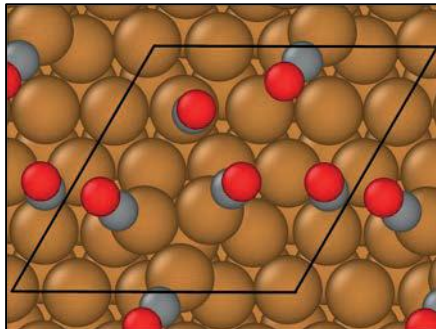


Fig. 3 : The Cu-CO training structure

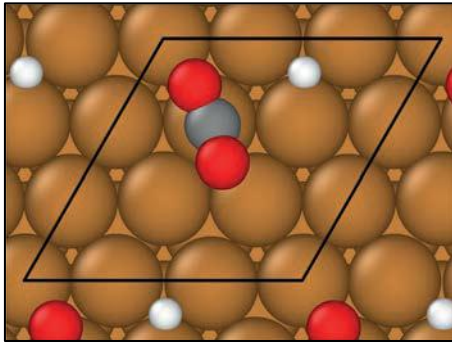


Fig. 4 : The Cu-CO₂-H training structure

2.3 Molecular Dynamics (MD)

The LAMMPS [5] software is used to perform large scale molecular dynamics in parallel both using MPI and OMP parallelization. The machine-learning interatomic potential is read by the software and the energy and forces are calculated at every timestep. The non-equilibrium states of the catalysis then can be obtained from the atomic trajectories of MD simulations.

2.4 Elucidator

Elucidator is a set of analysis tool that we used to analyze the trajectory of the MD simulation. This tool is Python-based program which relies on two libraries, namely ASE [6] and Ovito [7]. Tasks that are performed using this tool including: the analysis of splitting and recombination rate of hydrogen, the surface reconstruction of induced by CO, and the vibrational states of CO₂ hydrogenation.

3. Validation of Machine Learning Interatomic Potential

Before starting a production run of MD simulations, it is necessary to validate the accuracy of the interatomic potential produced by ML. The validations were done in terms of mean absolute error (MAE) of atomic forces (i.e., the difference between the forces calculated by DFT and the forces predicted by ML). All the results of

MAE for the Cu-H, Cu-C), and Cu-CO₂-H systems are shown in Fig 5, Fig 6, and Fig 7, respectively. We found that the MAEs are smaller than 50 meV/Å, which is reasonably good,

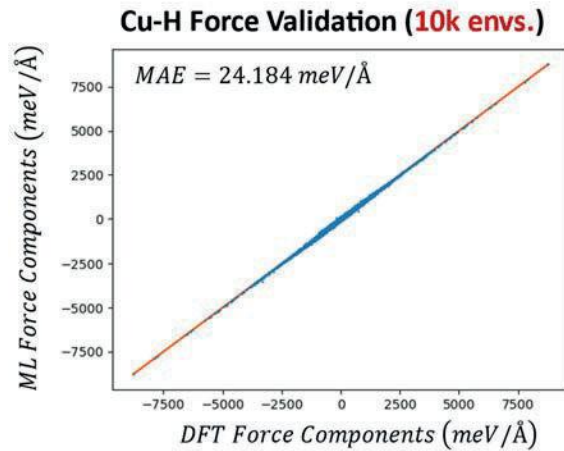


Fig. 5: Validation of atomic forces for Cu-H system

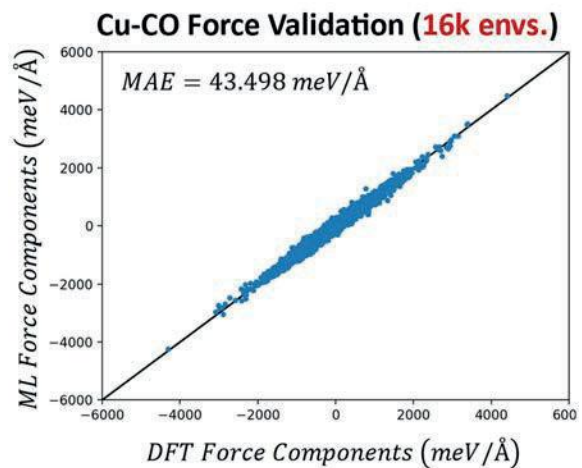


Fig. 6: Validation of atomic forces for Cu-CO system

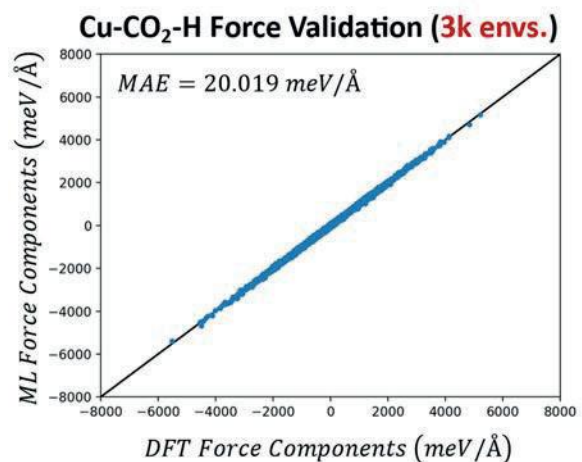


Fig. 7: Validation of atomic forces for Cu-CO₂-H system

4. Non-equilibrium states of H₂, CO, and

CO₂ interacting with Cu surface

After confirming the accuracy of the interatomic potential, the production run of MD simulations was performed. The resulting dynamics of H₂ interacting with Cu surface is shown in Fig.8. The figure shows the splitting of hydrogen molecules into two hydrogen atoms (shown in red color). The hydrogen atom is involved in almost every intermediate state of methanol synthesis. Fig 9 shows the snapshot of dynamics of Cu-CO system of which the CO co-adsorbed on the Cu surface. The analysis of the CO induced-surface reconstruction is ongoing. The snapshot depicting the hydrogenation process of CO₂ is shown in Fig 10. The current simulation has able to show successful and unsuccessful hydrogenation of CO₂.

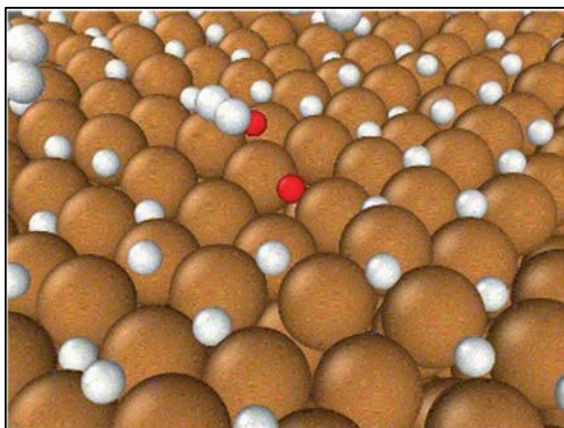


Fig. 8 : Snapshot of dynamics of Cu-H system

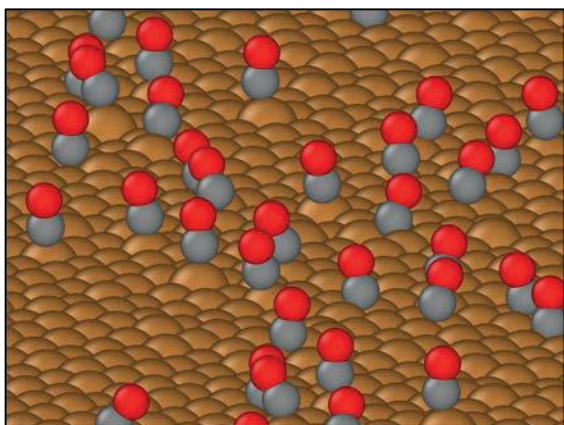


Fig. 9 : Snapshot of dynamics of Cu-CO system

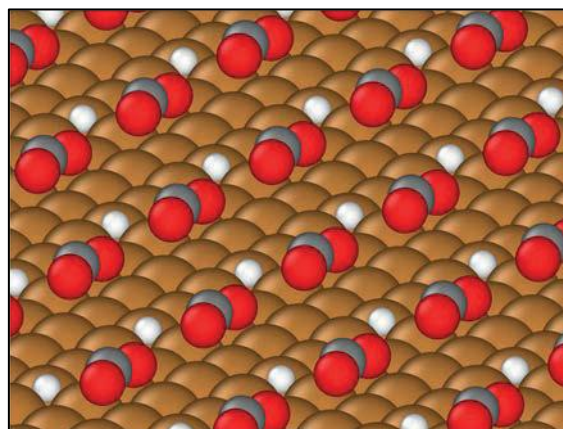


Fig. 10 : Snapshot of dynamics of Cu-CO₂-H system

5. Conclusion

Overall, we have been able to provide the direct observation to the dynamics of reactants in the methanol synthesis. The analysis of the resulting dynamics is being prepared and expected to be published in near future. When this research is successfully completed, detail knowledge of non-equilibrium states (the ‘missing link’ in experimental and computational studies) can be provided, which lead to more confident understanding of heterogenous catalysis.

Bibliography

- (1) M.H. de Sa, et al., *Energies.*, **15**, 17, (2022).
- (2) Y. Morikawa., et.al., *Appl. Sur. Sci.*, **11**. 169-170, (2001).
- (3) P. Giannozzi., et.al., *J. Phys. Condens. Matter.*, **21**, 39550, (2009).
- (4) J. Vandermause., et.al., *npj Comput Mater.*, **6**, 20, (2020).
- (5) A.P. Thompson, et.al., *Comp Phys Comm*, **271** 10817, (2022).
- (6) A. H. Larsen, et al., *J. Phys.: Condens. Matter.*, **29**, 273002, (2017).
- (7) A. Stukowski., *Modelling Simul. Mater. Sci. Eng.*, **18**, 015012 (2010).